Transactions of the ASME®

Journal of Engineering for Gas Turbines and Power

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 123 • NUMBER 1 • JANUARY 2001

Transactions of the ASME, Journal of Engineering for Gas Turbines and Power (ISSN 0742-4795) is published quarterly (Jan., April, July, Oct.) by The American Society of Mechanical Engineers, Three Park Avenue, New York, NY 10016. Periodicals postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Engineering for Gas Turbines and Power, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.
CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.
STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or ... printed in its publications (B7.1, par. 3).

INDEXED by Applied Mechanics Reviews and Engineering Information, Inc. Canadian Goods & Services Tax Registration #126148048

TECHNICAL PAPERS

Gas Turbines: Advanced Energy Systems

1 **Thermodynamic Optimization of the HAT Cycle Plant Structure—Part I: Optimization of the "Basic Plant Configuration"**
A. Lazzaretto and F. Segato

8 **Thermodynamic Optimization of the HAT Cycle Plant Structure—Part II: Structure of the Heat Exchanger Network**
A. Lazzaretto and F. Segato

17 **Entropy Production as a Predictive Performance Measure for Turbomachinery**
D. M. Paulus, Jr., R. A. Gaggioli, and W. R. Dunbar

Gas Turbines: Coal, Biomass, and Alternative Fuels

22 **A Study of Combustion Characteristics of Gasified Coal Fuel (99-GT-398)**
T. Hasegawa, M. Sato, and T. Nakata

Gas Turbines: Combustion and Fuel

33 **Effects of Fuel Nozzle Displacement on Pre-Filming Airblast Atomization (98-GT-360)**
Y. M. Han, W. S. Seol, D. S. Lee, V. I. Yagodkin, and I. S. Jeung

41 **Vortex Generators in Lean-Premix Combustion (98-GT-487)**
A. Eroglu, K. Döbbeling, F. Joos, and P. Brunner

Gas Turbines: Manufacturing, Materials, and Metallurgy

50 **Generation of Five-Axis Cutter Paths for Turbomachinery Components**
D. M. Tsay, W. F. Yan, and H. C. Ho

57 **Reconditioning of Gas Turbine Components by Heat Treatment (99-GT-444)**
J. A. Brown, R. Freer, and A. T. Rowley

Gas Turbines: Pipelines and Applications

62 **Measurement Uncertainties Encountered During Gas Turbine Driven Compressor Field Testing (98-GT-001)**
K. Brun and R. Kurz

70 **Degradation in Gas Turbine Systems (00-GT-345)**
R. Kurz and K. Brun

Gas Turbines: Structures and Dynamics

78 **Multiobjective Optimum Design of Rotor-Bearing Systems With Dynamic Constraints Using Immune-Genetic Algorithm**
B.-K. Choi and B.-S. Yang

82 **Reduced-Order Nonlinear Dynamic Model of Coupled Shaft-Torsional and Blade-Bending Vibrations in Rotors**
B. O. Al-Bedoor

89 **Component-Mode-Based Reduced Order Modeling Techniques for Mistuned Bladed Disks—Part I: Theoretical Models (00-GT-360)**
R. Bladh, M. P. Castanier, and C. Pierre

(Contents continued on inside back cover)

This journal is printed on acid-free paper, which exceeds the ANSI Z39.48-1992 specification for permanence of paper and library materials. ♾™
♻ 85% recycled content, including 10% post-consumer fibers.

(Contents continued)

Journal of Engineering for Gas Turbines and Power JANUARY 2001

Volume 123, Number 1

100 Component-Mode-Based Reduced Order Modeling Techniques for Mistuned Bladed Disks—Part II: Application (00-GT-361)
R. Bladh, M. P. Castanier, and C. Pierre

Special Section–Internal Combustion Engines

Cold start phenomena

109 Analysis of the Piston Ring/Liner Oil Film Development During Warm-Up for an SI-Engine
K. Froelund, J. Schramm, T. Tian, V. Wong, and S. Hochgreb

117 Simulation of a Single Cylinder Diesel Engine Under Cold Start Conditions Using Simulink
H.-Q. Liu, N. G. Chalhoub, and N. Henein

Emissions

125 Use of an Electrically Heated Catalyst to Reduce Cold-Start Emissions in a Bi-Fuel Spark Ignited Engine
G. N. Coppage and S. R. Bell

132 Fuel Composition Effects on Emissions From a Spark-Ignited Engine Operated on Simulated Biogases
K. C. Midkiff, S. R. Bell, S. Rathnam, and S. Bhargava

Flow, heat transfer, and combustion

139 In-Cylinder Tumble Flow Field Measurements and Predictions
C.-W. Hong and S.-D. Tarng

146 Forced Convective Heat Transfer of Parallel-Mode Reciprocating Tube Fitted With a Twisted Tape With Application to Piston Cooling
S. W. Chang

157 Study of Using Oxygen-Enriched Combustion Air for Locomotive Diesel Engines
D. N. Assanis, R. B. Poola, R. Sekar, and G. R. Cataldi

167 An Experimental Investigation of the Effects of Common-Rail Injection System Parameters on Emissions and Performance in a High-Speed Direct-Injection Diesel Engine
P. J. Tennison and R. Reitz

175 The Effect of Mean Turbulent Strain Rate on the Flame Speed of Premixed, Growing Flames
D. S.-K. Ting and M. D. Checkel

182 A Mechanism of Combustion Instability in Lean Premixed Gas Turbine Combustors
T. Lieuwen, H. Torres, C. Johnson, and B. T. Zinn

190 Laminar Burning Velocity of Methane—Air—Diluent Mixtures
M. Elia, M. Ulinski, and M. Metghalchi

197 Analysis and Modeling of the Fluid-Dynamic Effects in Branched Exhaust Junctions of ICE
F. Payri, E. Reyes, and J. Galindo

Hydrogen-fueled engines

204 An Investigation of a Cause of Backfire and Its Control Due to Crevice Volumes in a Hydrogen Fueled Engine
J. T. Lee, Y. Y. Kim, C. W. Lee, and J. A. Caton

211 Experimental Study of a Hydrogen-Fueled Engine
R. Sierens and S. Verhelst

Fuels and Combustion Technology

217 Scroll Compressor Simulation Model
C. Schein and R. Radermacher

226 Determination of Propane and Air Maximum Mixing Times
D. Brasoveanu and A. K. Gupta

231 Influence of Imperfections in Working Media on Diesel Engine Indicator Process
S. N. Danov and A. K. Gupta

Power

240 Design and Performance Verification of a 62-MWt CFB Boiler
K. Cen, P. Basu, L. Cheng, M.-X. Fang, and Z.-Y. Luo

246 Dynamic Optimization of Startup and Load-Increasing Processes in Power Plants—Part I: Method
J. Bausa and G. Tsatsaronis

251 Dynamic Optimization of Startup and Load-Increasing Processes in Power Plants—Part II: Application
J. Bausa and G. Tsatsaronis

Journal of Engineering for Gas Turbines and Power JANUARY 2001

Volume 123, Number 1

TECHNICAL BRIEF

255 Analytical Prediction Techniques for Axisymmetric Flow in Gas Labyrinth Seals
U. Yucel and J. Y. Kazakia

ANNOUNCEMENTS AND SPECIAL NOTES

258 Information for Authors

259 Preparing and Submitting a Manuscript

260 Preparation of Graphics

263 New Reference Format

Thermodynamic Optimization of the HAT Cycle Plant Structure—Part I: Optimization of the ''Basic Plant Configuration''

A. Lazzaretto
e-mail: lazza@dim.unipd.it
Mem. ASME

F. Segato

Department of Mechanical Engineering,
University of Padova,
via Venezia 1,
35131 Padova, Italy

A method for the thermodynamic optimization of the humid air turbine cycle plant structure is presented here. The method is based on the optimization of a ''basic configuration of the plant'' including ''basic components'' (compressor, turbine, combustion chamber, regenerator, and saturator), always present and connected in the same way in the plant structure and a heat exchange section which is viewed as a ''black-box'' where the heat transfer between hot and cold thermal flows occurs regardless of how many heat exchangers there are and of how they are interconnected. The optimal boundary conditions between the basic components and black-box are determined by calculating the conditions of maximum heat transfer in the black-box independently of the structure of the heat exchanger network. This is done by defining optimal composite curves in a Fortran routine at each step in the main optimization procedure. Once the structure of the heat exchanger networks that fulfill the optimal boundary conditions have been found, the optimal structure of the whole plant is obtained (see Section 2). The method is useful in a general sense as it can be applied to highly integrated energy systems in which it is difficult to define the optimal structure of the heat exchanger network in advance.
[DOI: 10.1115/1.1338999]

1. Introduction

The humid air turbine (HAT) cycle [1,2] achieves high efficiencies with relatively simple plant structures. However, the literature suggests there are plant structures of various complexity which nonetheless always include an intercooled compression, the regenerative heating of the combustion air, and to increase the turbine mass flow rate and power, the saturation of the compressed air. The main differences are usually associated with the number of heat exchangers (intercoolers, aftercoolers, and economizers) and the interconnections between them. Although the performances of these various structures were often found to be ''optimal,'' the assumptions used to determine them were different [3–10].

It has been suggested that a thermodynamic optimization of both the structure and the working parameters of the cycle can be performed by considering a ''superstructure'' which includes all the plant structures supplying good performances [8]. In this case the optimization indicates which components of the superstructure are to be included in the plant and what the optimal values of the decision variables are (pressures, temperatures, mass splitting ratios, etc.). However, this method requires a thorough preliminary study of all the configurations included in the superstructure and does not guarantee that good solutions are not neglected.

A different method for the thermodynamic optimization of the plant structure was first suggested in [11], and it is presented here in a more systematic and general way. The method is based on the optimization of a so called ''basic configuration'' of the plant, which derives from considering two sections in the total plant:

- The first section includes compressor, turbine, combustion chamber, regenerator, and saturator that are called ''basic components'' as they are always present and connected in the same way in the plant structure;
- The second section, called heat transfer section, includes intercoolers, aftercoolers, and economizers, the interconnections and numbers of which can vary within quite broad limits.

The latter is viewed as a ''black-box'' (see also [9]), in which the heat transfer between hot and cold thermal flows occurs without having defined the structure of the heat exchanger network (HEN), i.e., the number of heat exchangers and the interconnections between them.

To analyze the results in each step of the optimization procedure, rather than using an optimization algorithm, several thermodynamic simulations are carried out under variations of the decision variables within ranges that supply high values for the thermal efficiency of the total plant. Prior to carrying out the simulations, the decision variables are chosen through thermodynamic analyses on the behavior of single components or of zones of the plant that indicate the independent variables to be maximized, minimized, or assuming constant values.

The basic configuration is a configuration of the *whole* plant in which the heat transfer section (black-box) is not actually isolated from the rest of the plant, but simply considered without the physical devices that perform the heat transfer. It thus becomes crucial to obtain the optimal conditions at the boundaries of the black box. This is done by calculating the conditions of maximum heat transfer in the black-box independently of the structure of the heat exchanger network. To obtain these conditions, composite curves are implemented in a Fortran routine at each step in the main optimization procedure (see, e.g., [12]) (see Section 2.3.1)

After performing the optimization of the plant's basic configuration, the optimal values of all the variables at the boundaries of the black-box are known, in particular, the thermal capacities and the initial and final temperatures of the hot and cold thermal flows. Consequently, the structure of the HEN can be defined using a combination of the practical rules of Pinch Technology [12] and Second Law insights (see, for instance, [13,14]), as shown in the second part of the paper. There are HENs of various complexity that fulfill the optimal black-box boundary conditions and consequently guarantee the maximum total plant thermal efficiency.

Contributed by the Advanced Energy Systems Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the AES Division, June 29, 1999; final revision received by the ASME Headquarters, Jan. 23, 2000. Editor: H. D. Nelson.

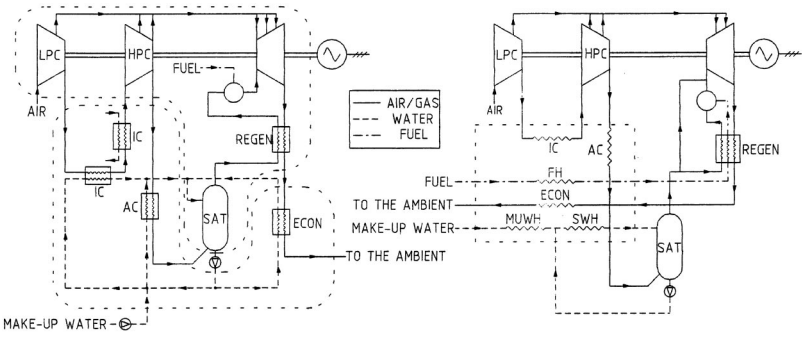


Fig. 1 One of the HAT cycle plant configuration proposed in the literature. The ''basic components'' and the heat transfer section are included within dotted lines. The ''basic configuration'' of the HAT cycle proposed here.

Among these HENs, the one having the minimum number of heat exchangers is selected. In this way, the thermodynamic optimization proposed here defines plant structures having maximum efficiencies with a minimum number of heat exchangers.

Even if the best possible HEN should be chosen by considering the costs of the components, in order to conduct a thermoeconomic optimization (which is beyond the scope of this paper), the indications regarding the plant structure obtained here can provide a good basis for subsequent cost evaluations.

The simulations were carried out on a PC P120 using the computer code Aspen Plus [15].

2. The Optimization Methodology

The main steps of the optimization methodology are:

2.1 Define a basic configuration of the plant in which the heat exchange section is considered to be a ''black-box'' where the heat transfer occurs independently of the structure of the heat exchanger network.

2.2 Perform preliminary thermodynamic analyses aimed at understanding the operating conditions of the basic configuration that lead to the maximum total plant efficiency.

2.3 Perform a thermodynamic optimization of the basic plant configuration aimed at maximizing the total efficiency by carrying out repeated simulations for different values of the decision variables. The boundary conditions between the black-box and the rest of the plant are optimized in each simulation (optimization step) by maximizing the heat transfer in the black-box, using a suitably developed Fortran routine.

Details for each of these steps are given in the following sections.

2.1 Definition of the ''Basic Configuration'' of the Plant. A simplified plant configuration of the HAT cycle is shown in Fig. 1 to highlight the basic components (see Sec. 1) and the heat transfer processes that occur in the heat exchange section, i.e.,

- the intercooling, aftercooling and saturation of the compressed air (the latter also involves a mass exchange);
- the humid air preheating in the regenerator;
- the water preheating at the saturator inlet in the intercooling and aftercooling processes of the compressed air and in the exhaust gas cooling.

These processes are explicitly represented as heat exchangers hot or cold sides in the ''basic configuration'' of the plant, shown in Fig. 1, in order to highlight the hot and cold thermal flows. In this way, the heat transfer section is viewed as a ''black-box'' where the heat transfer between hot and cold thermal flows occurs independently of the number of heat exchangers and of the interconnections between them. The black-box can only release heat to the environment since there are no external heat sources.

2.2 Preliminary Thermodynamic Analyses of the Basic Plant Conflguration. If the independent variables are suitably selected, the optimal value for some of them can be defined before running the simulations of the optimization procedure. This, in turn, can reduce the number of simulations in the optimization and provide preliminary insights on the thermodynamic behavior of the basic plant configuration.

The ''free'' structure of the heat exchange section complicates the process of choosing the independent variables. Thus, preliminary studies on single components or parts of the plant were performed to select them. These studies made it possible to choose a suitable set of independent variables of the basic plant configuration and to identify the ones with constant values or clear trends or the ones to be optimized (decision variables) in the total plant optimization procedure. These variables and the values they assume in the optimization (Table 1) were derived according to the following thermodynamic considerations:

• The heat recovery from the exhaust gases must be maximized. Thus, the humid air temperature at the regenerator outlet ($T_{\mathrm{ha},R,\mathrm{out}}$) must reach the maximum allowed value (depending on the turbine outlet temperature—the minimum value of Δt is fixed

Table 1 Independent variables of the basic plant configuration (the values of the optimization variables within parentheses are those considered in the optimization procedure—see Table 3)

VARIABLE		GOAL
Total pressure ratio	β	Optimize (10 - 20 - 30)
Intercooling final temperature	T_{ic}	Optimize (25 - 60 - 100 °C)
Intermediate pressure ratio	β_{HP}/β_{LP}	Optimize (range: 0,3-1,8)
Aftercooling final temperature	T_{ac}	Optimize (range: 30-180 °C)
Humid air temp. at the regen. outlet	$T_{ha,R,out}$	Maximize
Fuel temp. at the regenerator outlet	$T_{f,R,out}$	Maximize
Make-up water mass flow rate	m_{muw}	Maximize
Water temp. at the saturator inlet	T_w	Maximize
Final temp. of the exhaust gases	T_{exg}	Minimize (min. val.: 105 °C)
Relative air humidity at the sat. outlet	φ_{ha}	Fix at constant value (1)
Fuel temp. at the regenerator inlet	$T_{f,R,in}$	Fix at constant value (140 °C)
Inlet temp. of the make-up water	$T_{muw,in}$	Fix at constant value (15 °C)
Make-up water temperature at MUWH outlet (see Figure 1b)	$T_{muw,out}$	Fix at the value of the water temperature at the bottom of the saturator
Turbine inlet temperature	TIT	Fix at const. value (1250 °C)
Inlet air mass flow rate	m_a	Fix at const. value (100 Kg/s)
Inlet air temperature (ambient temp.)	T_a	Fix at constant value (15 °C)
Inlet air pressure (ambient pressure)	p_a	Fix at const. value (1,013 bar)

Table 2 Operating parameters of the HAT cycle components

Compressor	Maximum blade temperature: 780 °C
Inlet pressure losses: 1 Kpa	(1st nozzle: 800 °C)
Polytropic efficiency: 0,875	**Electric Generator**
Mechanical efficiency: 0,995	Efficiency: 0,98
Air Leakage (at the HP compressor	**Air/water heat exchangers**
outlet): 0,80%	Pressure losses (Hot side): 1%
Combustion Chamber	ΔT_{min}: 10 °C
Pressure losses: 3%	**Regenerator**
Heat losses: 1%	Pressure losses (Both sides): 2%
Fuel	ΔT_{min}: 25 °C
Composition: Methane (CH_4)	**Saturator**
Temperature: 15 °C	Pressure losses (Air): 0,7%
Pressure: 40 bar	**Pumps**
Turbine	Efficiency: 0,7
Isentropic eff.- Cooled stages: 0.875	Electrical efficiency: 0,96
Isentropic efficiency – Uncooled	**Ambient conditions**
stages: 0,886 - 0,894 - 0,899 (*)	Temperature: 15 °C
Mechanical efficiency: 0,995	Pressure: 1,01325 bar
Cooling flows factor (k): 0,04505	Relative humidity: 60%

(*) The three values, respectively associated with $\beta = 10, 20, 30$, take account of the different expansion ratios of the non-cooled stages.

at 25°C, see Table 2) and the exhaust gas temperature at the economizer outlet T_{exg} must be fixed at the minimum value (105°C).

• The make-up water flow rate ($\dot{m}_{muw}$) is equal to the water flow rate that evaporates in the saturator. Therefore $\dot{m}_{muw}$ must be maximized in order to obtain the maximum mass flow rate and power in the turbine.

• Either the water mass flow rate or the temperature (or both of them) at the saturator inlet must be increased in order to increase the amount of water that evaporates in the saturator. However, only the temperature (T_w) must be maximized since the mass flow rate is not an independent variable of the basic configuration (we have taken $\dot{m}_{muw}$ as independent variable).

• The relative air humidity (ϕ_{ha}) at the saturator outlet is set to its maximum value (equal to one) since, at a constant value of the outlet temperature, the complete saturation corresponds to the maximum water content of the humid air. On the other hand, when the water content has a constant value, a decrease in the relative humidity at the saturator outlet results in an increase in the humid air temperature. Thus, the thermal flow in the regenerator decreases and, consequently, the thermal flow in the economizer increases; this flow is then transferred from the air to the water in the saturator. This heat transfer from the regenerator to the economizer has a small influence on the total plant efficiency.

• The fuel temperature at the regenerator outlet ($T_{f,R,out}$) must be maximized to increase the heat recovery from the exhaust gases at high temperatures.

• The fuel preheating before the regenerator inlet generates a decrease in the regenerator thermal flow and, consequently, an increase in the exhaust gas temperature at the economizer inlet. This, in turn, leads to an increase in the water temperature at the saturator inlet (T_w).

• The temperature of the fuel at the regenerator inlet ($T_{f,R,in}$) is an optimization variable. However, to simplify the calculations, it is fixed at 140°C, which is approximately the value of the humid air temperature at the regenerator inlet. Since the preheating thermal flow is small in comparison to the other thermal flows in the black-box, there is little error.

2.3 The Optimization Procedure. The flow-chart of one of the steps of the optimization procedure of the basic plant configuration (Fig. 1) is shown in Fig. 2. First a set of values for the decision variables (see Table 1) is fixed. Then the initial value of the make-up mass flow rate ($\dot{m}_{muw}$) is fixed and the first simulation of the optimization procedure is performed. In this simulation

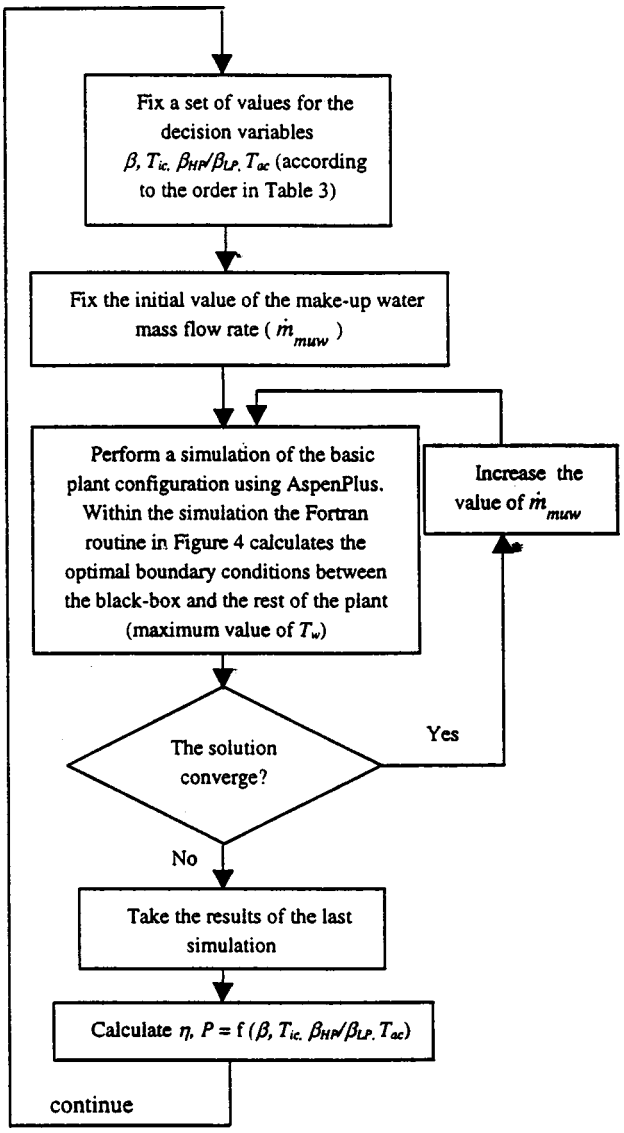


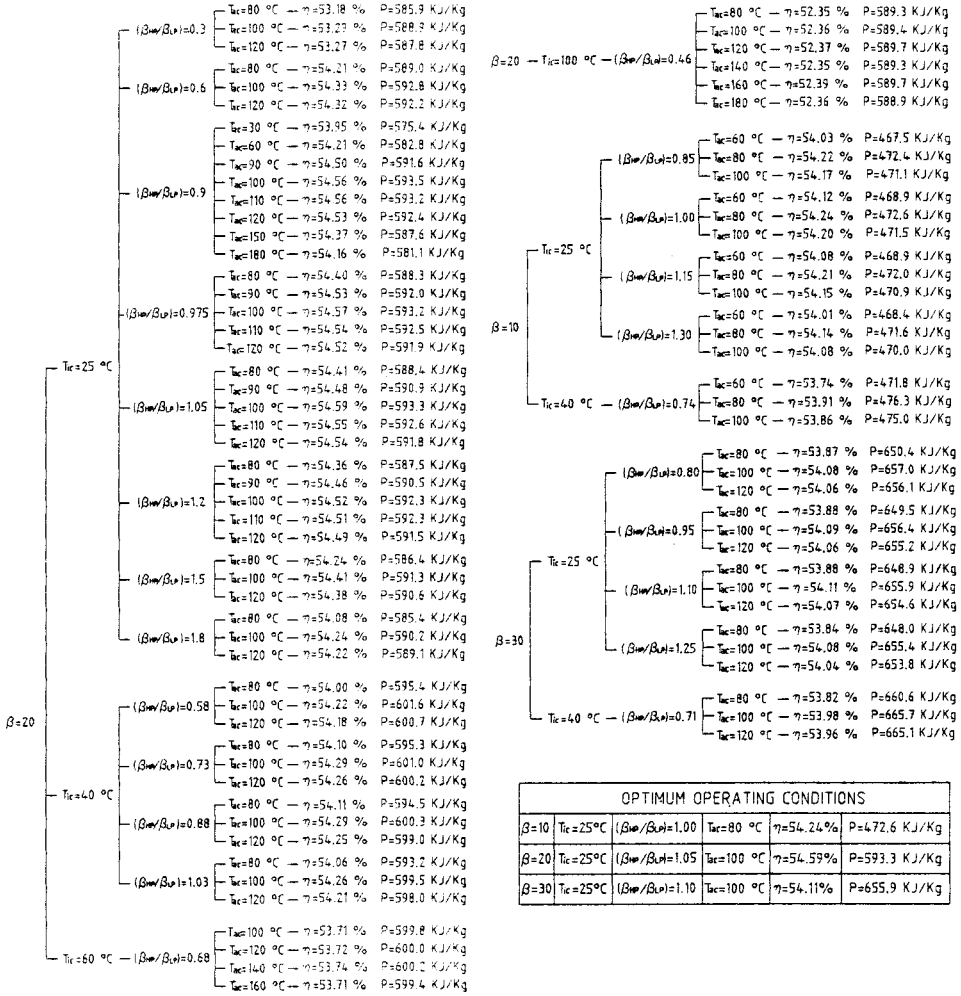
Fig. 2 Flow chart of one of the steps of the optimization procedure of the basic plant configuration (η and P represent the total plant efficiency and specific power, respectively)

the optimal boundary conditions between the black-box and the rest of the plant are determined using the Fortran routine described in Section 2.3.1. If the solution converges, the value of $\dot{m}_{muw}$ is increased and the simulation is repeated. This procedure is repeated until the solution no longer converges (i.e., until the selected value of $\dot{m}_{muw}$ is greater than its possible maximum value).

At this point, the maximum value of $\dot{m}_{muw}$ is the one used in the previous simulation, from which the total plant efficiency (η) and specific power (P) values associated with the selected decision variable values are obtained (see Table 3). The complete procedure is then repeated for a different set of decision variable values. The order used to vary these values is presented in Table 3.

2.3.1 Optimal Boundary Conditions between the ''Black-Box'' and the Rest of the Plant. The behavior of the heat exchange section in the basic plant configuration is not independent of the behavior of the rest of the plant. Thus, it is necessary to appropriately define the optimal boundary conditions between the two sections, i.e., the conditions maximizing the thermal exchange in the black-box.

Table 3 Results of the simulations of the basic plant configuration under variations of β, T_{ic}, β_{HP}/β_{LP}, and T_{ac} (η and P indicate the efficiency and specific power of the total plant, respectively)



OPTIMUM OPERATING CONDITIONS					
β=10	Tic=25°C	(βsw/βLP)=1.00	Tac=80 °C	η=54.24%	P=472.6 KJ/Kg
β=20	Tic=25°C	(βsw/βLP)=1.05	Tac=100 °C	η=54.59%	P=593.3 KJ/Kg
β=30	Tic=25°C	(βsw/βLP)=1.10	Tac=100 °C	η=54.11%	P=655.9 KJ/Kg

The hot streams (LPA, HPA, EXG) and cold streams (MUW, RW, SW, FUEL) of the black-box are represented in Fig. 3.

In the same figure the independent variables ($T_{\mathrm{muw,in}}$, T_{ic}, T_{ac}, T_{exg}, $\dot{m}_{\mathrm{muw}}$, $T_{f,R,\mathrm{in}}$, T_w) of the basic plant configuration at the black-box boundaries are shown (the dependent variables are not shown).

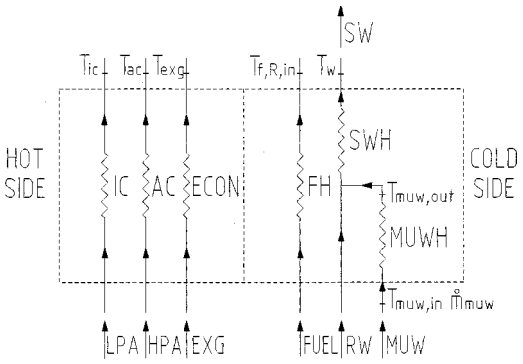


Fig. 3 Black-box with hot streams (LPA—low pressure air in the intercooler, HPA—high pressure air in the aftercooler EXG—exhaust gases in the economizer) and cold streams (MUW—make-up water, RW—recuperated water at the bottom of the saturator, SW—water sent to the saturator, FUEL)

As mentioned in Section 2.2, the outlet temperature of the stream EXG (T_{exg}) is fixed at 105°C (see Table 1) and the values of the outlet temperatures (T_{ic},T_{ac}) of the streams LPA and HPA are to be determined in the optimization procedure. They can range from the inlet values of the same streams to 25°C (equal to $T_{\mathrm{ambient}}+\Delta T_{\mathrm{min}}$, where $T_{\mathrm{ambient}}=15$°C and $\Delta T_{\mathrm{min}}=10$°C) because of the possibility of discharging heat to the environment.

If the structure of the heat transfer section were defined, the outlet temperatures (T_w,$T_{f,R,\mathrm{in}}$) of the cold strcams (SW, FUEL) would be dependent variables directly calculated by the simulation model. Since this structure is not defined in the basic plant configuration, the thermal flows in the heaters SWH and FH are unknown (see Fig. 3) and T_w, $T_{f,R,\mathrm{in}}$ become independent variables (their ranges of variation depend on the amount of heat available from the hot streams and on their temperature level). As explained at the end of Section 2.2, the temperature $T_{f,R,\mathrm{in}}$ is fixed at a constant value (140°C), whereas the temperature and mass flow rate at the inlet of the saturator (T_w,$\dot{m}_{\mathrm{muw}}$) are independent variables to be maximized in each optimization step (see Table 1). The optimal values of these variables are obtained here by defining the composite curves (CCs) (see, e.g., [12]) under conditions of maximum heat exchange in the black-box, using a Fortran routine, the flow chart of which is shown in Fig. 4.

In each optimization step the routine calculates the maximum value of T_w for fixed values of $\dot{m}_{\mathrm{muw}}$. As already mentioned in Section 2.3, the calculation is repeated for increasingly higher

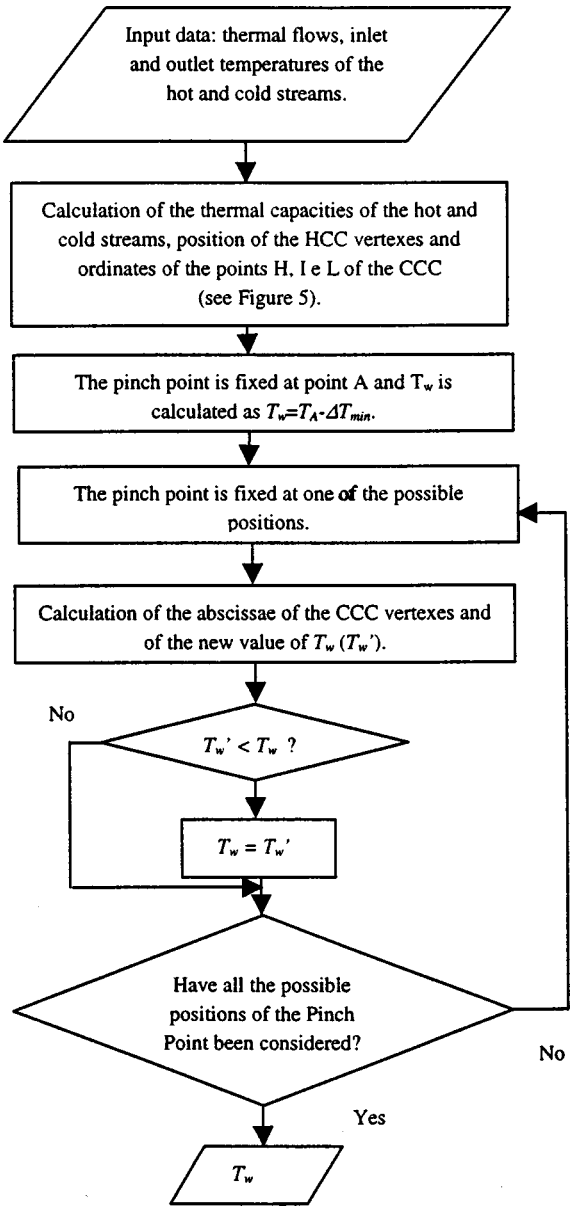


Fig. 4 Flow-chart of the Fortran routine used to determine the optimal value of the temperature at the saturator inlet (T_w)

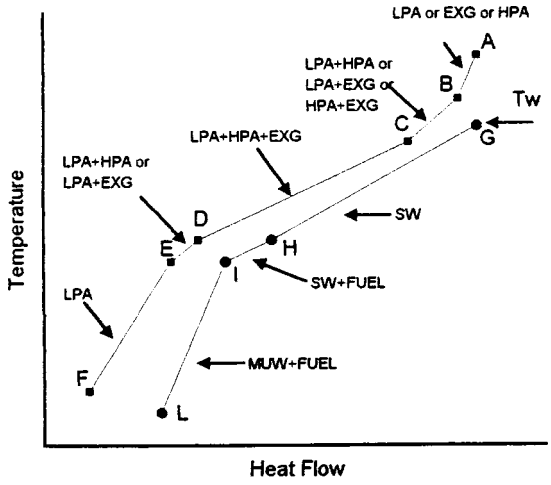


Fig. 5 Composite curves in the black-box. In each segment heat can be exchanged using various combinations of the hot and cold streams shown in Fig. 3

values of $\dot{m}_{muw}$ (see Fig. 2), until its maximum value, which corresponds to the maximum amount of water evaporated in the saturator, is obtained.

The shape of the CC (number of segments, slope, and ordinates of the vertexes) is not defined before running the simulations.

We only know:

- the values of $T_{muw,in}$, T_{exg}, $T_{f,R,in}$ (see Table 1);
- that the abscissae of the extreme right points of the two CCs must coincide since there are no available external heat sources; accordingly, the maximum value of the temperature T_w is obtained when the temperature difference at the Pinch Point is equal to ΔT_{min}.

In this context, the following assumptions regarding the cold and the hot sides are made in the Fortran routine to determine the CC:

Cold streams—The cold stream MUW (Fig. 3) is heated separately from RW up to the temperature of RW at the inlet of the black-box, in order to recover a part of the heat available from the hot streams at low temperature (from the ambient temperature to the temperature of the water at the bottom of the saturator).

Hot streams—The inlet temperatures of the hot streams are assumed to be different from the outlet ones. This hypothesis leads to a hot composite curve (HCC) made up of five segments (Fig. 5). When T_{ic} is fixed at the minimum value (25°C), there is one more segment since the thermal capacity of the stream LPA (Fig. 3) is different in the segments at higher and lower temperature where a partial condensation of the air humidity occurs.

Given these assumptions, the Fortran routine calculates T_w as follows, using the data of each simulation step.

The shape of the HCC (i.e., the position of vertexes and the slope of each segment), the slope of the CCC (cold composite curve) segments and the ordinates of the points *H*, *I*, and *L* are determined using the inlet and outlet temperatures of all the streams and the respective thermal flows. The length of the segment GH is still unknown (it is only known that the abscissae of *G* and *A* have to coincide) because the CCC curve can still move horizontally. This length is determined in order to maximize the temperature T_w. This is done by imposing the value of ΔT_{min} (10°C) in all the possible positions of the pinch point, i.e., where the HCC and CCC have cusps facing each other or in the extreme points. Each position corresponds to a horizontal movement of the CCC, and consequently to a value of T_w. The smallest value of T_w is chosen because it is the only one that assures the heat exchange feasibility. In other words, the CCC is moved horizontally until ΔT is greater than ΔT_{min} in all the vertexes except the pinch point.

Note that the methodology suggested here for optimizing the black-box boundary conditions can be applied to any kind of CC. However, the only CC of practical use besides the ones suggested are those in which the CCC vertex in *H* moves to *I* (these curves are not considered here).

2.4 Total Optimization Results. As mentioned in Sections 1 and 2, a step-by-step optimization of the basic plant configuration is performed using repeated simulations. The range of variation for each decision variable (see Table 1) is determined according to the technological constraints and to the achievement of good performances. The order of variation for the decision variables is shown in Table 3. The first variable is the overall pressure ratio β, since it affects the total plant operating conditions. The second and third variables are the final intercooling temperature T_{ic} and the intermediate pressure ratio β_{HP}/β_{LP} respectively (the optimal value of which depends on the value of T_{ic}). The last variable is the aftercooler outlet temperature T_{ac} which ranges

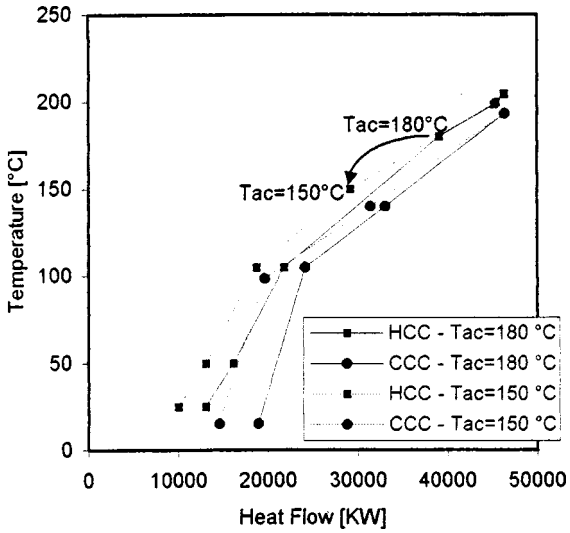


Fig. 6 Composite curves for T_{ac}=150°C and T_{ac}=180°C (above the optimum)

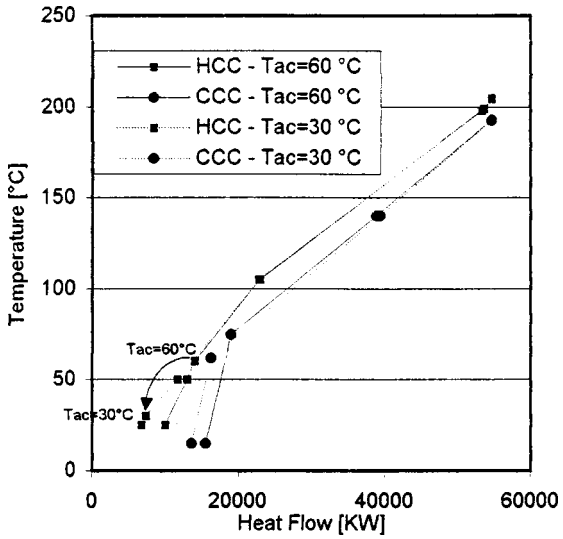


Fig. 7 Composite curves for T_{ac}=60°C and T_{ac}=30°C (below the optimum)

between 30°C (slightly higher than the ambient temperature) and 180°C (a little lower than the final temperature of the compressed air) depending on the overall pressure ratio.

The gas turbine model includes an intercooled compression with intermediate air bleeding for the cooling of the first three turbine stages (the first nozzle is cooled by humid air extracted at the saturator outlet). The cooling flows are calculated separately for each stator and rotor blade row using the mathematical relationships suggested by Stecco and Facchini [16]. The main operating parameters are shown in Table 2. Most of them are obtained by adjusting the model to the performances of an actual aeroderivative gas turbine (GE LM2500PE) operating in the northeast of Italy.

The simulation of the saturator is performed using the Radfrac module included in the Aspen Plus library. The adiabatic operating conditions and the value of the relative air humidity at the outlet are specified, whereas the mass flow rates and the thermodynamic conditions of the input streams are known from the calculations in the other components. Using an iterative process the code calculates the mass flow rates and the thermodynamic conditions of the outlet streams that fulfill the mass and enthalpy balance of the saturator.

The simulation results (Table 3) are explained in the following.

• The optimal aftercooler outlet temperature T_{ac} is approximately located in the middle of its variation range (30–180°C) and is a little higher than the temperature of the water at the bottom of the saturator. An example of CCs above the optimal value of T_{ac} is shown in Fig. 6. From this figure we note that a decrease in T_{ac} from 180°C to 150°C generates a decrease in the slope of the upper part of the CCC. This leads to an increase in the total water flow rate sent to the saturator (inversely proportional of the slope). An example of CCs below the optimal value of T_{ac} is shown in Fig. 7. For T_{ac}=60°C and T_{ac}=30°C the slope of the upper part of the CCC is almost constant since the thermal flow caused by the decrease of T_{ac} below its optimal value cannot be used to heat the water sent to the saturator because T_{ac} is lower than the temperature of the water at the bottom of the saturator. Below this temperature the stream LPA is enough to heat the cold streams (MUW and FUEL) and the aftercooling thermal flow is released to the environment. Moreover, the decrease in the air temperature at the saturator inlet (T_{ac}) generates a decrease in the water content of the humid air, resulting in a decrease of the total plant efficiency.

• The optimal value of the intercooling final temperature T_{ic} is the minimum allowed value (25°C), which also minimizes the compression work. The optimal value of the intermediate pressure ratio β_{HP}/β_{LP} is close to the ''local'' optimal value determined by deriving the expression of the polytropic compression work at constant values of the inlet and intercooling temperatures and overall pressure ratio (β). Higher values for T_{ic} and different values for β_{HP}/β_{LP} lead to higher temperatures of the water at the saturator inlet (T_w). However, the increase in the turbine power is lower than the increase in the compression work, resulting in a decrease in the net power and in the total plant efficiency.

• The optimal value of β is equal to 20. However, the total plant efficiency is not greatly affected by this variable, whereas the specific power increases as the pressure ratio increases (Table 3). In fact, in the regenerative cycles lower values of β generate higher exhaust gases temperatures and higher air temperatures at the combustion chamber inlet. In addition to this effect, when the values of β increase in the HAT cycle there is a positive effect due to the higher intercooling and aftercooling thermal flows.

3. Conclusions

The methodology for the thermodynamic optimization of a plant structure presented here mainly consists in optimizing a ''basic configuration'' of the total plant in which the physical structure of the heat exchanger network is left ''free'' and substituted by hot and cold thermal flows.

The optimization of the basic plant configuration is performed using repeated simulations and a suboptimization procedure (operating within the main optimization procedure) that determines the optimal boundary conditions between the heat exchange section and the rest of the plant independently of the structure of the heat exchanger network. Thus, the results of the optimization apply to all the heat exchanger networks that fulfill these boundary conditions. The second part of the paper discuss how to find these networks, and consequently the optimal structure of the total plant.

The main results of the optimization applied to the HAT cycle are:

(i) For each value of the total pressure ratio, the highest total plant efficiencies are obtained for values of the intermediate pressure ratio (β_{HP}/β_{LP}) and of the final intercooling temperature (T_{ic}) close to those that minimize the compression work.

(ii) The aftercooling has a positive effect on the total plant efficiency only when the aftercooling thermal flow can be used to heat part of the water at the bottom of the saturator.

(iii) The total pressure ratio does not have a significant influence on the total plant efficiency. The maximum value (54.59 percent) is obtained with a total pressure ratio $\beta=20$. Only slight decreases are obtained for $\beta=10$ (54,24 percent) and $\beta=30$ (54,11 percent).

This optimization methodology is of general validity since it can be used in all the systems that have a ''heat exchange section'' where the heat exchangers can be interconnected in different ways to get the best plant performances. As shown in the paper, it is crucial to correctly define and optimize the variables at the boundaries between the heat exchange section and the rest of the plant.

Acknowledgment

We would like to acknowledge Professor A. Mirandola for his helpful suggestions.

References

[1] Rao, A., Francuz, V., Mulato, F., Sng, B., West, E., J., K., Perkins, G., and Podolski, D., 1993, ''A Feasibility and Assessment Study for FT 4000 Humid Air Turbine (HAT).'' Final Report TR-102156, EPRl, Palo Alto, CA.

[2] Day, W., and Rao, A., 1993, ''Redefined Natural Gas HAT Cycle Produces Higher Output,'' Modern Power Systems, **13**, p. 6.

[3] Lindgren, G., Eriksson, J., Bredhe, K., and Annerwall, K., 1992, ''The HAT Cycle, a Possible Future for Power Generation,'' *Proc. of Florence World Energy Research Symposium (Flowers'92)*, S. Stecco and M. Moran, eds., Florence, Italy, June 7–12, Nova Science Publishers, Commack, NY.

[4] Stecco, S. S., Desideri, U., Facchini, B., and Bettagli, N., 1993, ''The Humid Air Cycle: Some Thermodynamic Considerations, ASME Paper 93-GT-77.

[5] Stecco, S. S., Desideri, U., and Bettagli, N., 1993, ''Humid Air Gas turbine Cycle: A Possible Optimization, ASME Paper 93-GT-178.

[6] Huang, F. F., and Naumowicz, T., 1994, ''Thermodynamic Study of the System Performance of an Externally-Fired Humid Air Turbine Cycle Power Plant,'' *Proc. of Florence World Energy Research Symposium (Flowers'94)*, E. Carnevale, G. Manfrida, and F. Martelli, eds., Florence, Italy, July 6–8, SGEditoriali, Padava, Italy.

[7] Rosen, P., Eidensten, L., Svedberg, G., Torisson, T., and Yan, J., 1994, ''Analysis of Small Scale Evaporative Gas Turbines. Part I: Directly Fired with Natural Gas,'' *Proc. of Florence World Energy Research Symposium (Flowers'94)*, E. Carnevale, G. Manfrida, and F. Martelli, eds., Florence, Italy, July 6–8, SGEditoriali, Padava, Italy.

[8] Chiesa, P., Lozza, G., Macchi, E., and Consonni, S., 1995, ''An Assessment of the Thermodynamic Performance of Mixed Gas Steam Cycles: Part B–Water-Injected and HAT Cycles,'' J. Eng. Gas Turbines Power, **117**, pp. 499–508.

[9] Bram, S., and De Ruyck, J., 1996, ''Exergy Analysis Tools for Aspen Applied to Evaporative Cycle Design,'' *Proc. of ECOS'96, Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems*, Stockholm, Sweden, P. Alvfors, L. Eidensten, G. Svedberg, and J. Yan, eds., pp. 217–224.

[10] Xiao, Y. H., Cai, R., and Lin, R., 1996, ''Modeling HAT Cycle and Thermodynamic Evaluation,'' *Proc. of ECOS '96 Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems*, Stockholm, Sweden, P. Alvfors, L. Eidensten, G. Svedberg, and J. Yan, eds., pp. 211–216.

[11] Lazzaretto, A., and Segato, F., 1999. ''A Systematic Approach to the Definition of the HAT Cycle Structure Using Pinch Technology,'' *Proceedings of the ECOS'99 International Conference on Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems*, M. Ishida, G. Tsatsaronis, M. J. Moran, and H. Kataoka, eds., Tokyo Institute of Technology, Tokyo, June 8–10, pp. 215–222.

[12] Linhoff, B., Townsend, D. W., Boland, D., Hewitt, G. F., Thomas, B. E. A., Guy, A. R., and Marsland, R. H., 1982–1994, *A User Guide on Process Integration for the Efficient Use of Energy*, The Institute of Chemical Engineering, Rugby, Warks, UK.

[13] Sama, D. A., 1995a, ''The Use of the Second Law of Thermodynamics in Process Design,'' J. Energy Resour. Technol., **117**, pp. 179–185.

[14] Sama, D. A., 1995b, ''Differences between Second Law Analysis and Pinch Technology,'' J. Energy Resour. Technol., **117**, pp. 186–191.

[15] Aspen Plus, 1996, ''Aspen Plus User Guide'', release 9.3, Aspen Technology Inc., Cambridge, MA.

[16] Stecco, S. S., and Facchini, B., 1989, ''A Computer Model for Cooled Expansion in Gas Turbine,'' *Proc. of 1989 ASME Cogen-Turbo Symposium*, Nice, France, pp. 201–209.

Thermodynamic Optimization of the HAT Cycle Plant Structure—Part II: Structure of the Heat Exchanger Network

A. Lazzaretto
e-mail: lazza@dim.unipd.it
Mem. ASME

F. Segato

Department of Mechanical Engineering,
University of Padova,
via Venezia 1,
35131 Padova, Italy

In Part I of the paper a thermodynamic optimization methodology was presented for the ''basic'' configuration of a humid air turbine cycle plant in which the heat exchange section is viewed as a black-box separated from the rest of the plant (basic components), having a fixed structure. The results of the optimization apply to all the heat exchanger networks that fulfill the optimal boundary conditions between the black-box and the rest of the plant. The aim of this part is to define these heat exchanger networks using a combination of Pinch Technology and Second Law insights. The possibility of a further reduction in the number of heat exchangers is then investigated in order to achieve the best compromise between high performances and structural simplicity.
[DOI: 10.1115/1.1339000]

1. Introduction

The optimization methodology presented in the first part of the paper considers the part of the humid air turbine (HAT) cycle plant having a fixed structure separate from the part associated with the internal thermal exchanges, i.e., the so-called black-box. In fact, the possible variations in plant structure are usually limited to the latter, where the heat exchangers can be interconnected in different ways to maximize the thermal exchanges. This methodology makes it possible to optimize more than one plant structure by optimizing only the one defined as the ''basic plant configuration.'' As explained in the first part, this is possible if the boundary conditions between the black-box and the rest of the plant are appropriately defined and optimized.

After optimizing the operating conditions of the basic plant configuration, the heat exchangers within the black-box can be interconnected in different ways, each one fulfilling the optimal black-box boundary conditions. This consists in defining heat exchanger networks (HENs) starting from known values of thermal capacities and initial and final temperatures of the hot and cold flows in the black-box. Practical rules of Pinch Technology [1] and Second Law insights [2,3] are used to define the HENs.

Thus, according to this methodology, the choice of the optimal plant structure is reduced to the choice of the optimal structure of the HEN, i.e., the one having the minimum number of heat exchangers among those that fulfill the optimal conditions at the black-box boundaries.

The effects of further reductions in the number of heat exchangers are then investigated in order to find the best compromise between high efficiency and structural simplicity. The investigation begins with the observation that even remarkable variations from the optimal values of some choice variables (found in the first part) do not significantly affect the total plant efficiency.

2. Composite Curves of the Heat Exchange Section in the Optimum Conditions

The composite curves (Section 2.3.1—Part I) of the heat exchange section in the conditions of maximum total plant efficiency for the total pressure ratios $\beta=10$, 20, 30 are shown in Figs. 1*a*, 1*b*, and 1*c* respectively. Segments *GI* and *CD* should be as parallel as possible in order to maximize the amount of water that evaporates in the saturator. This condition was obtained in the two different plant structures suggested in Chiesa et al. [4] and Xiao et al. [5] where the heat capacity of the water sent to the saturator was set equal to the sum of the heat capacities of the hot streams. If this condition is imposed, it is possible that the minimum allowed exhaust gas temperature (105°C) be violated (i.e., lower values can be obtained). For this reason the mass flow rate of the water sent to the saturator is considered here to be a dependent variable and the exhaust gas temperature is fixed at 105°C (the make-up mass flow rate is considered independent, see Section 2.2—Part I). According to these hypotheses, segments *GI* and *CD* in Fig. 1*c* are almost parallel in the optimum conditions for $\beta=30$. In this case, the temperatures T_{ac} and T_{exg} almost coincide (point *D*) and are 10°C higher than the temperature of the water at the bottom of the saturator (T_I in Figs. 1*a*, 1*b*, and 1*c*). Vice versa, segments *GI* and *CD* cannot be parallel when shifting from $\beta=30$ to $\beta=20$ and $\beta=10$ (Figs. 1*a* and 1*b*) because, while T_I decreases progressively, T_{exg} is fixed at the minimum value (105°C).

3. Heat Exchanger Networks That Achieve the Maximum Total Plant Efficiency

The HEN structures which perform the heat exchange within the black-box in the conditions of total plant maximum efficiency (found in Part I) are defined here using a combination of practical rules of Pinch Technology and Second Law insights. More than one HEN can satisfy the optimal values of the inlet and outlet temperatures and mass flow rates of the streams in the black-box (Fig. 2—Part I). Each of these HENs leads to the same overall plant performances (see also Bram and De Ruyck, [6]). The existence of at least one HEN that perform the heat exchange is assured, since the composite curves were used to define the optimal boundary conditions of the black-box (Section 2.3.1—Part I).

After defining the composite curves, the first step in the HEN design is the ''grid representation'' [1], which is drawn separately above and below the pinch point, according to the rules of Pinch Technology.

3.1 Total Pressure Ratio (β) Equal to 20. The design of the HENs for $\beta=20$ is presented in this section. For $\beta=20$ the

Contributed by the Advanced Energy Systems Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the AES Division, June 29, 1999; final revision received by the ASME Headquarters Jan. 23, 2000. Editor: H. D. Nelson.

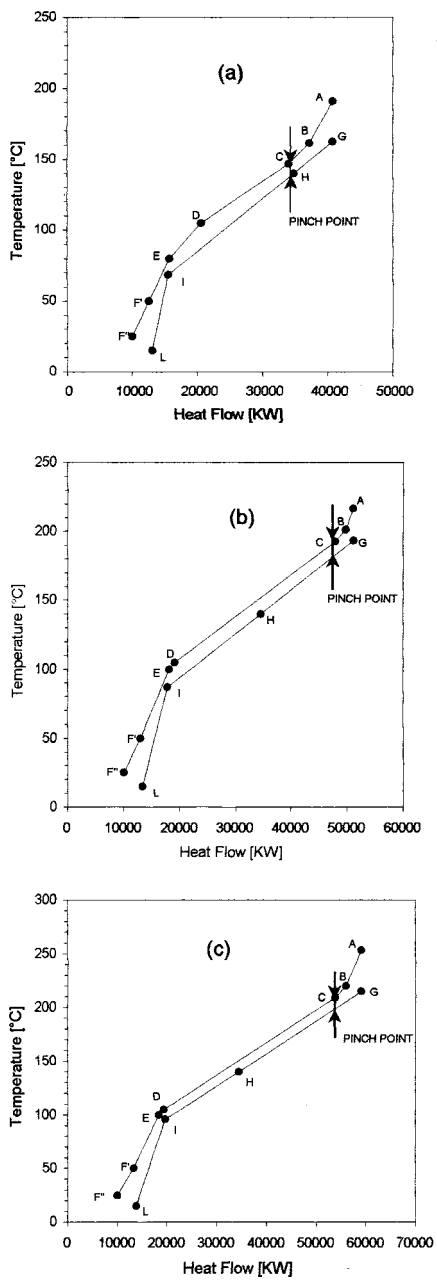


Fig. 1 Composite curves in the optimum operating conditions of the basic configuration for different values of the total pressure ratio. *a* β=10, *b* β=20, *c* β=30

pinch point in the optimum conditions is located at 192.82°C and 182.82°C for the hot and cold streams, respectively ($\Delta T_{\min}$ =10°C) (see Fig. 1*b*).

First HEN (Two Aftercoolers—Two Economizers). Above the pinch point the hot streams HPA and EXG (Fig. 3 in Part I) must be cooled without releasing heat to the environment (see Linhoff et al. [1]); thus, the cold stream SW is split into at least the two streams SW1 and SW2 (Fig. 2). According to the Second Law, to make a heat exchange feasible immediately below the pinch point, the thermal capacity of each cold stream must not exceed the thermal capacity of the hot stream matched to it. To assure this, the SW stream is split into three streams matched with the hot streams LPA, EXG, and HPA in the heat exchangers IC1, EC2, AC2.

The thermal capacities of the streams SW3 and SW4 are calculated in order to complete the thermal exchange of the hot streams HPA and EXG. In this way the cold stream SW is heated in AC1 and EC1 above the pinch and in AC2, EC2, and IC1 below the pinch. The make-up water (MUW) is heated in IC2. The heat required for fuel preheating is supplied in IC1 and IC2. The intercooler IC3 completes the cooling of the LPA stream to 25°C, releasing heat to the environment.

The number of heat exchangers in Fig. 2 can be decreased by eliminating IC2. This is done by mixing the stream MUW with the water flow (SW5) sent to IC1 (see Fig. 3*a*). In this case, the optimum conditions found in part I of this work are not precisely fulfilled (the fuel is preheated only up to 80°C), but the total plant efficiency decreases very little (54.50 percent). Indeed, the optimization procedure does not lead to ''the optimum'' but to a point ''close to the optimum.'' Thus, small deviations from this point do not lead to significant changes in the plant performances.

Second HEN (One Aftercooler—One Economizer). The economizers EC1 and EC2 and the aftercoolers AC1 and AC2 in Figs. 2 and 3*a* are grouped into the single units AC and EC (Fig. 3*b*). Accordingly, the streams SW1, SW3 and SW2, SW4 become single streams.

As a result, the SW stream is split into three streams above the pinch, the third of which (SW5) is not heated. According to the Second Law, the thermal capacities of a cold stream above the pinch must be greater than or equal to that of the hot stream matched to it, vice versa below the pinch. Thus, the thermal capacities of the cold streams SW1 and SW2 must be equal to those of the hot streams HPA and EXG, respectively. This results in a thermal exchange in EC and AC with a constant minimum temperature difference (10°C).

To get this value at the cold end of EC and AC as well, SW1 and SW2 must be heated up to 90°C and 95°C in the heat exchangers IC5 and IC4, respectively. Nevertheless, IC5 can be eliminated by imposing the HPA outlet temperature (T_{ac}) 10°C higher than the SW inlet temperature, without significant effects on the total plant efficiency (54.55 percent). It is worth noting that if IC4 is eliminated, the outlet temperature of the EXG stream would be lower than its minimum value (105°C). Furthermore, the intercooler IC2 cannot be eliminated because of the presence of IC4.

Third HEN (One Aftercooler—Two Economizers). A third option consists in grouping the aftercoolers AC1 and AC2 (in Figs. 3*a* and 3*b*), while keeping the economizers EC1 and EC2 separate (see Fig. 3*c*). In this case, IC5 is needed to precisely fulfill the optimal boundary conditions of the black box, whereas IC4 is not required. However, IC2 and IC5 can be eliminated as explained for the first and second HENs, respectively. This results in the plant structure in Fig. 4, having the minimum number of heat exchangers for β=20. Also in this case, the effect on the total plant efficiency (54.50 percent), due to the decrease in the final temperature of the FUEL stream (85°C), is negligible. Note that the values of some of the thermodynamic parameters obtained from the simulation of the complete plant structure in Fig. 4 are slightly different from those in Fig. 3*c*, which were obtained from the simulation of the basic plant configuration.

Fourth HEN (Two Aftercoolers—One Economizer). In the fourth option (Fig. 3*d*) the aftercoolers AC1 and AC2 are kept separate, whereas the economizers EC1 and EC2 are put together into the single unit EC. The intercoolers IC4 and IC2 cannot be eliminated (see the second HEN for details—Fig. 3*b*), whereas IC5 is not required. The HEN includes two more heat exchangers than the third one. The total plant efficiency is 54.55 percent.

3.2 Total Pressure Ratio (β) Equal to 10. For β=10, in the optimum conditions the pinch is located at a temperature of 147.02°C for the hot streams and 137.02°C for the cold streams (see Fig. 1*a*).

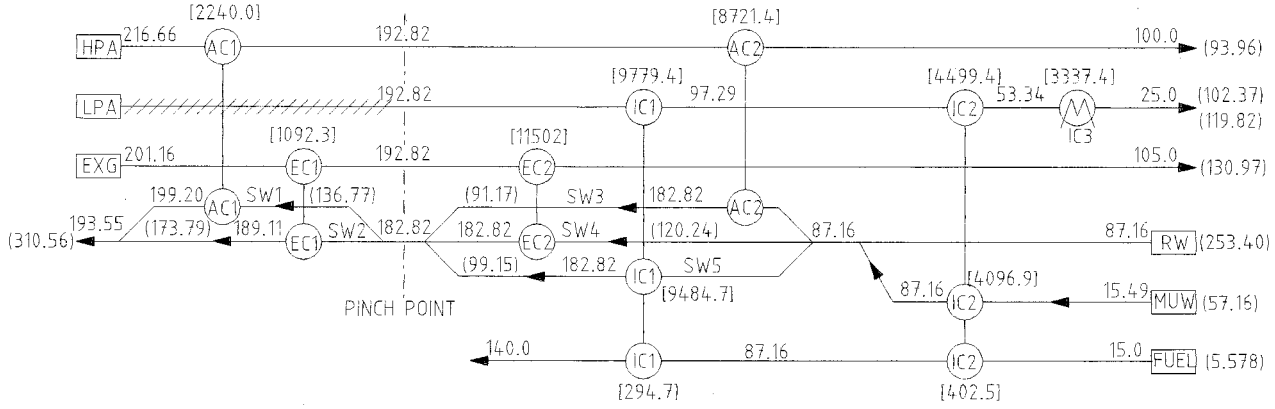


Fig. 2 Grid representation of the first HEN for β=20 (two aftercoolers—two economizers). Numbers represent temperatures (°C), the heat flows (kW) and the thermal capacities (kW/°C) are included within square and round parentheses, respectively.

First HEN (Two Aftercoolers—Two Economizers). In addition to the SW stream, the FUEL stream must be heated as well above the pinch point for β=10 (Fig. 5*a*). Nevertheless, it cannot be directly matched to a hot stream (HPA or EXG) because its thermal capacity is too low. Therefore, SW is split into the two streams SW1 and SW2 and the FUEL stream is heated in EC1 together with SW2.

Below the pinch it would be possible to directly match the FUEL stream with each one of the three hot streams but the sum of the heat capacities of the two remaining ones would not be sufficiently high to heat SW up to the pinch. Therefore, it is necessary to split SW into three streams matched with the three hot streams using the heat exchangers EC2, AC2, and IC1, and to heat the FUEL stream in one of these heat exchangers (we chose IC1).

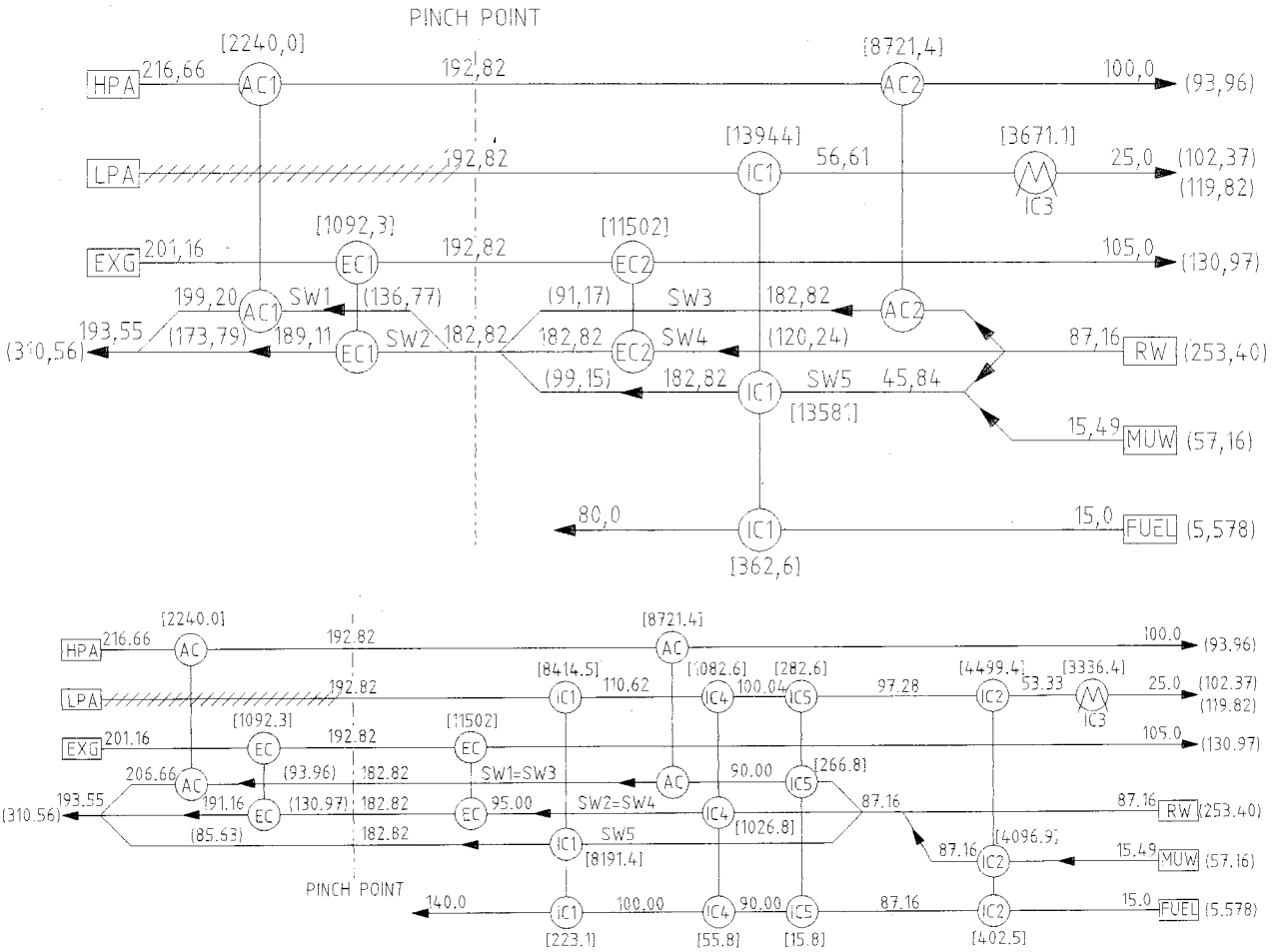


Fig. 3 Grid representations of the four HENs fulfilling the optimal boundary conditions at the black-box boundaries for β=20; *a* two aftercoolers, two economizers; *b* one aftercooler—one economizer; *c* one aftercooler, two economizers; *d* two aftercoolers, one economizer. Numbers represent temperatures (°C); the heat flows (kW) and the thermal capacities (kW/°C) are included within square and round parentheses, respectively.

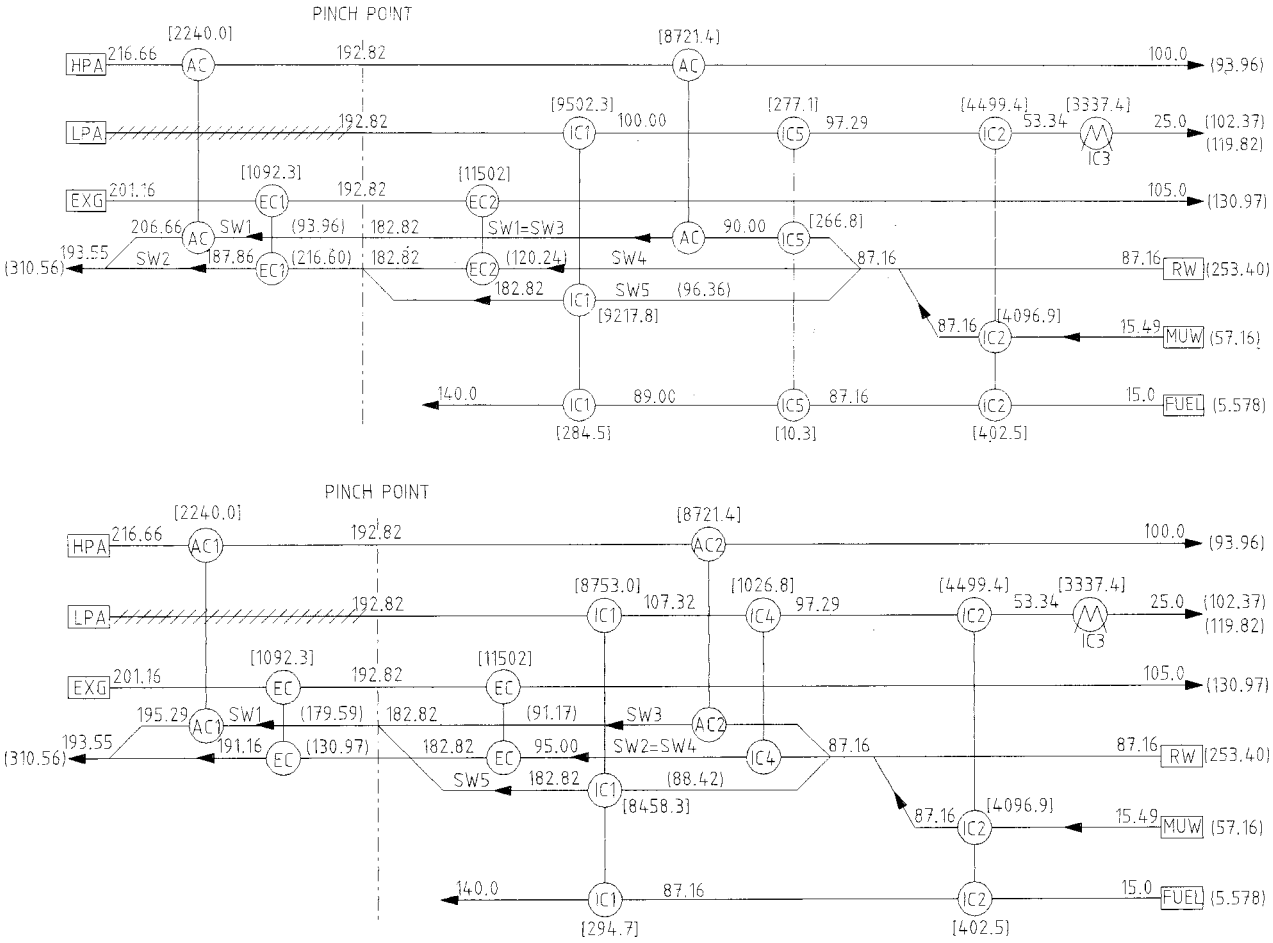


Fig. 3 (*Continued*)

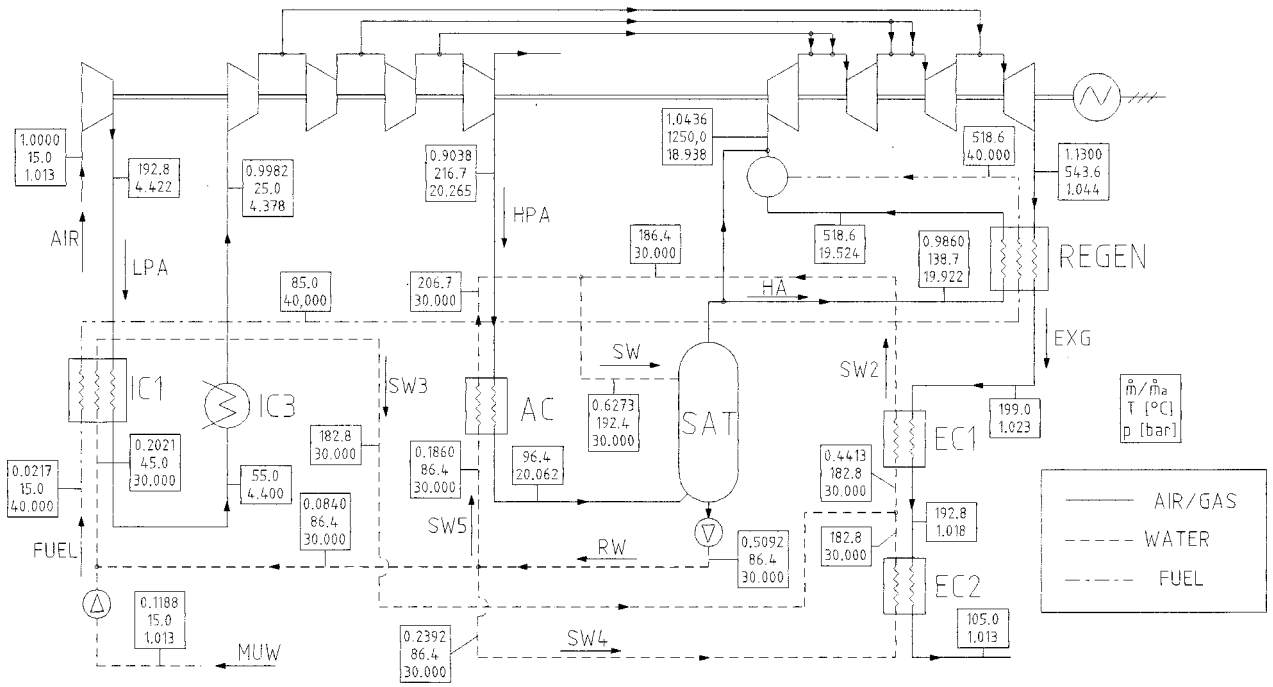


Fig. 4 Optimum plant structure for β=20 (third HEN: one aftercooler—two economizers). IC2 and IC5 in the grid representation in Fig. 3*c* have been eliminated by mixing the stream MUW with the water flow sent to IC1 (SW3) and by imposing the HPA outlet temperature 10°C higher than the SW5 inlet temperature.

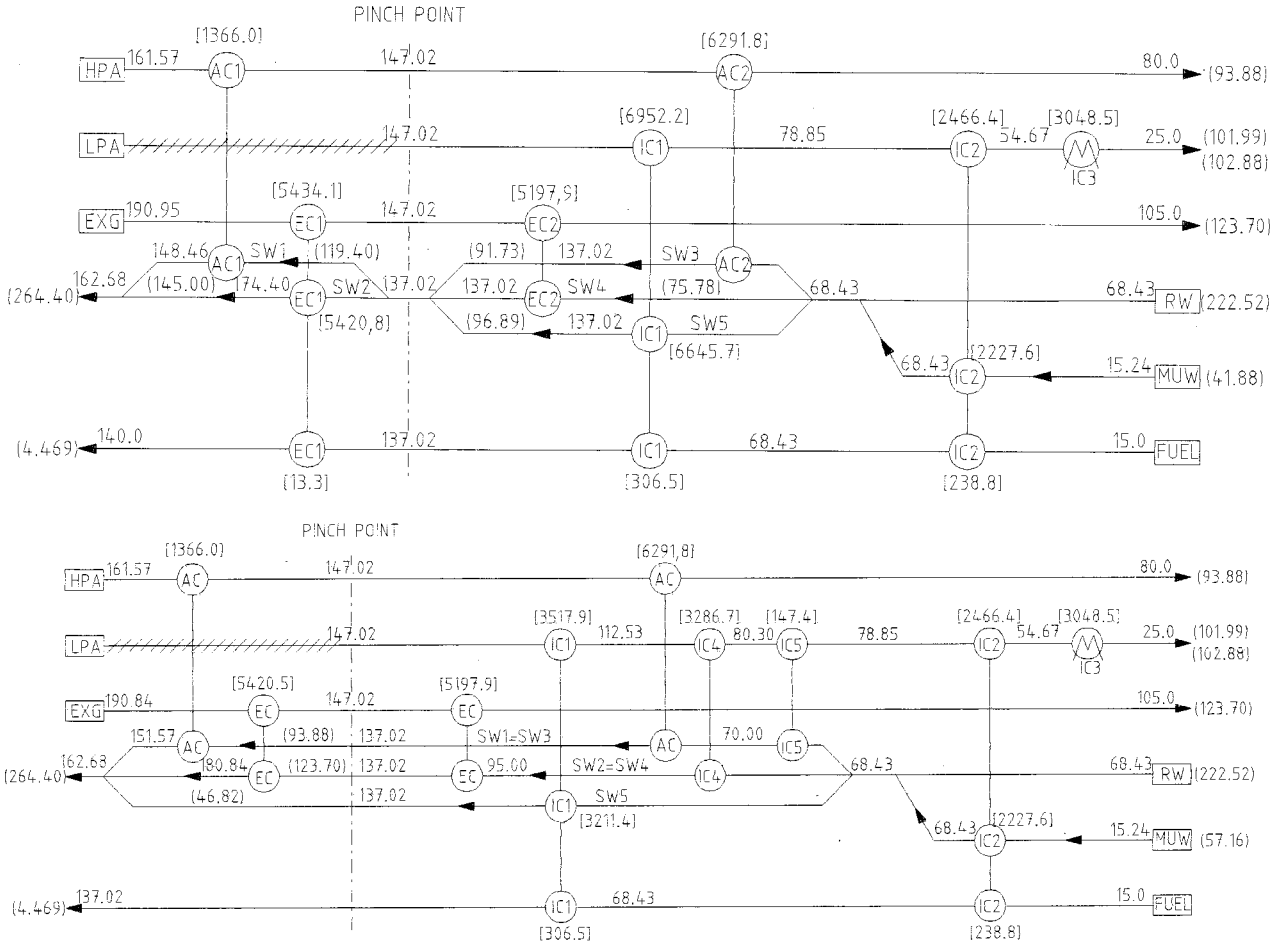


Fig. 5 Grid representations of the four HENs fulfilling the optimal boundary conditions at the black-box boundaries for β=10. *a* two aftercoolers, two economizers; *b* one aftercooler—one economizer; *c* one aftercooler, two economizers; *d* two aftercoolers, one economizer. Numbers represent temperatures (°C), the heat flows (kW) and the thermal capacities (kW/°C) are included within square and round parentheses, respectively.

The thermal capacities of the streams SW3 and SW4 are calculated in order to complete the thermal exchange in EC2 and AC2. The rest of the HEN is designed as for β=20. The intercooler IC2 can be eliminated for β=10 as well, as explained for β=20. In this case the aftercooling final temperature is fixed at 78°C, i.e., 10°C higher than the inlet temperature of SW3. The total plant efficiency is equal to 54.24 percent.

Second HEN (One Aftercooler—One Economizer). As was already done for β=20, it is possible to group the heat exchangers operating above and below the pinch into the units *AC* and *EC* (Fig. 5*b*) for β=10 as well. Nevertheless, the FUEL stream cannot be heated in *AC* and *EC* because the thermal capacities of the hot and cold streams in these heat exchangers must be equal to each other (see the second HEN for β=20). Therefore, the FUEL stream can be heated only up to the pinch point (137.02°C). The thermal flow required to reach 140°C is taken from the regenerator and this results in a small drop in the EXG inlet temperature (from 190.95 to 190.84°C). However, this stream makes the exact thermal flow required to heat SW up to its final temperature (162.68°C) available since the lower thermal flow available in the black-box from EXG is exactly compensated by the lower heat requirement of FUEL.

The heat exchangers IC4 and IC5 are needed to preheat the streams SW1 and SW2 up to 70 and 95°C, respectively. As has already been explained for β=20, IC5 can be eliminated by slightly modifying the final temperature of the HPA stream (from 80 to 78°C). In this case the total plant efficiency is 54.25 percent.

Third and Fourth HENs (One Aftercooler—Two Economizers, Two Aftercoolers—One Economizer). The third and fourth HENs are almost equal to the corresponding ones for β=20 except for the FUEL heating above the pinch, which must be performed in the heat exchangers EC1 and AC1 (see Figs. 5*c* and 5*d*). The total plant efficiencies are equal to 54.22 percent and 54.25 percent for the third and fourth HENs, respectively. As for β=20, the third HEN includes the minimum number of heat exchangers, and the related plant structure is shown in Fig. 6.

3.3 Total Pressure Ratio (β) Equal to 30. For β=30, in the optimum conditions the temperatures at the pinch point for the hot and cold streams are equal to 208.99°C and 198.99°C, respectively (see Fig. 1*c*). Consequently, the two streams to be cooled above the pinch are HPA and LPA (Fig. 7) (instead of HPA and EXG for β=10 and 20). For the same reasons explained for β=20, the stream SW must be split into at least two parts above the pinch and in three parts below it. The aftercoolers and intercoolers operating above and below the pinch can be grouped or left separate (only the option with grouped aftercoolers and intercoolers is presented here).

For β=30 the initial temperature of the stream SW is too high (95.7°C) to cool the hot streams HPA and EXG to their final

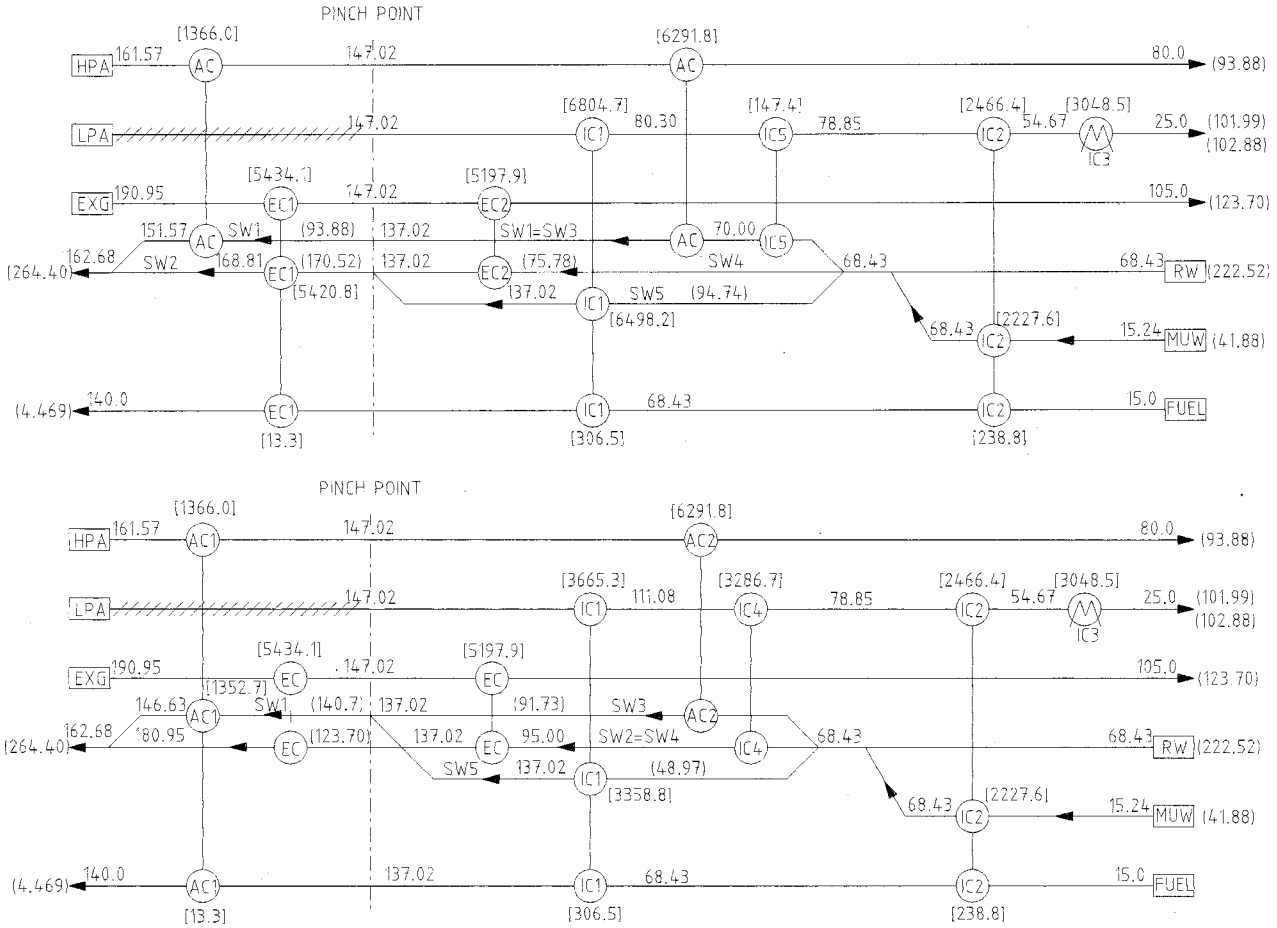


Fig. 5 (*Continued*)

temperatures (100 and 105°C). To strictly fulfill the black-box boundary conditions, it is therefore necessary to use the cold streams MUW and FUEL to cool the hot streams HPA and EXG in the heat exchangers AC2 and EC2. Nevertheless, AC2 and EC2 can be eliminated with negligible penalties on the total plant efficiency, by imposing the final temperatures of the streams HPA and EXG equal to the temperatures at the outlet of AC1 and EC1, respectively. Accordingly, the plant structure takes the form in Fig. 8 and the plant efficiency is 54.07 percent. In this structure the make-up water stream (MUW) is mixed with the water stream (SW2) sent to IC1, the initial temperature of which becomes 42.6°C.

3.4 Remarks. The structure of the various HENs in the heat exchange section depends on the black-box boundary conditions found in Part I of the paper, which are different for different values of the total pressure ratio.

From the discussion in Sections 3.1–3.3 we inferred that the heat can be recovered separately in the two sections above and below the pinch by considering the heat exchangers operating above and below the pinch to be (i) separate or (ii) joint, or the intermediate options to be (iii) joint aftercoolers—separate economizers and (iv) separate aftercoolers—joint economizers.

The four options lead to very similar HENs for $\beta=10$ and 20 (the only differences are related to the heating of the FUEL stream). For $\beta=30$ the different hot streams above the pinch (HPA and LPA instead of HPA and EXG) result in different HENs, particularly when using separate heat exchangers above and below the pinch.

(i) Using the first option (separate heat exchangers) for $\beta=10$ and 20 the water stream SW can be suitably split to complete the cooling of the hot streams (HPA and EXG) to their optimal final temperatures. However, this option implies the preheating of the water streams MUW and SW in the intercoolers IC2 and IC1. However, IC2 can be eliminated with small adjustments to the temperatures at the black-box boundaries, so that the total number of heat exchangers is six (AC1, AC2, EC1, EC2, IC1, IC3). For $\beta=30$, the option with separate heat exchangers above and below the pinch is not presented here. As for $\beta=10$ and 20, it would include one aftercooler, one economizer, and two intercoolers (including the dissipative one) below the pinch, whereas it would need one intercooler and one aftercooler (instead of one economizer and one aftercooler) above the pinch.

(ii) The second option consists of grouping aftercoolers and economizers (for $\beta=10$ and 20), or aftercoolers and intercoolers (for $\beta=30$). When using this option, the thermal capacities of the streams crossing the pinch point in the joint heat exchangers must be equal to achieve the maximum heat exchange, thus obtaining a thermal exchange with constant minimum temperature difference (10°C). To fulfill this condition for $\beta=30$ it is necessary to slightly increase the optimal final temperature of HPA by setting its value 10°C higher than the temperature of the water at the bottom of the saturator (initial temperature of RW). A thermal exchange with $\Delta T_{\min}=10$°C can be also obtained in the economizer by imposing a $\Delta T=10$°C at the cold end of this heat exchanger. In this way the final temperature of the exhaust gases (107.6°C) is higher than 105°C (minimum allowed value) since the temperature of the water at the bottom of the saturator is 97.6°C (see Fig. 8). In this case the HEN has only four heat exchangers (two intercoolers, one aftercooler, and one econo-

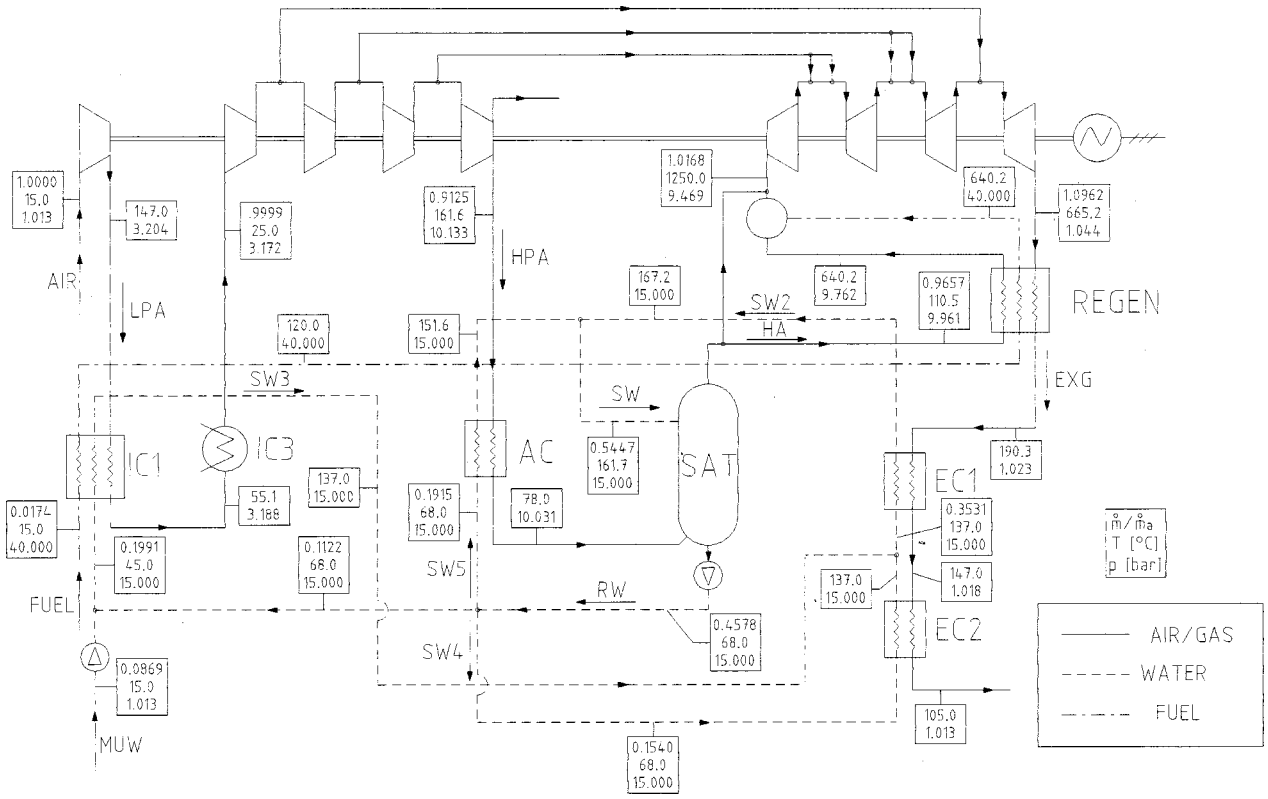


Fig. 6 Optimum plant structure for β=10 (third HEN: one aftercooler—two economizers). IC2 and IC5 in the grid representation in Fig. 5*c* have been eliminated by mixing the stream MUW with the water flow sent to IC1 (SW3) and by imposing the HPA outlet temperature 10°C higher than the SW5 inlet temperature.

mizer) and differs only for the fuel preheating from the HEN suggested by Xiao et al. [5]). For β=10 and 20 this solution cannot be applied to the EXG stream because its final temperature would be lower than the minimum allowed value (105°C). Thus, it is necessary to preheat the water flow rate sent to the economizer in a supplementary heat exchanger (IC4). In this way the total number of heat exchangers (six) is reduced of only one unit compared to the option using separate heat exchangers.

(iii) The minimum number of heat exchangers for β=10 and 20 is obtained by grouping the aftercoolers operating above and below the pinch while leaving the economizers separate (Figs. 4 and 6). These HENs include five heat exchangers, i.e., only one more economizer than the optimal HEN for β=30.

(iv) When aggregating the economizers for β=10 and 20, it is necessary to preheat the water entering the economizer in the intercooler IC4 (from 87.16 to 95°C) to fulfill the lower limit of the exhaust gases temperature (105°C) and consequently to heat the MUW stream in the intercooler. (Figs. 3*d* and 5*d*). Thus, the total number of heat exchangers (six) is increased by one unit with respect to option iii.

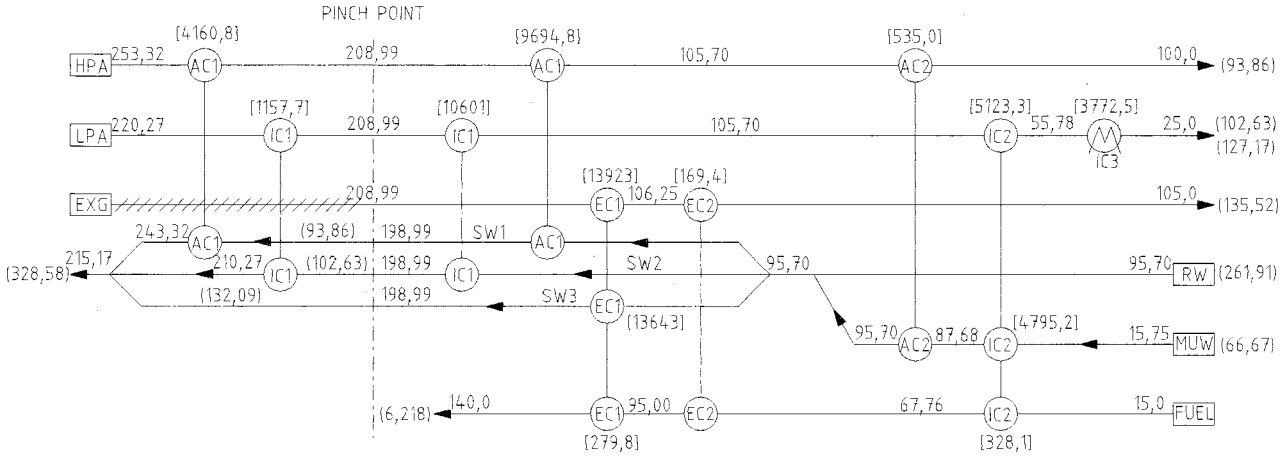


Fig. 7 Grid representation of the HEN for β=30 (one aftercooler—one economizer). Numbers represent temperatures (°C), the heat flows (kW), and the thermal capacities (kW/°C) are included within square and round parentheses, respectively.

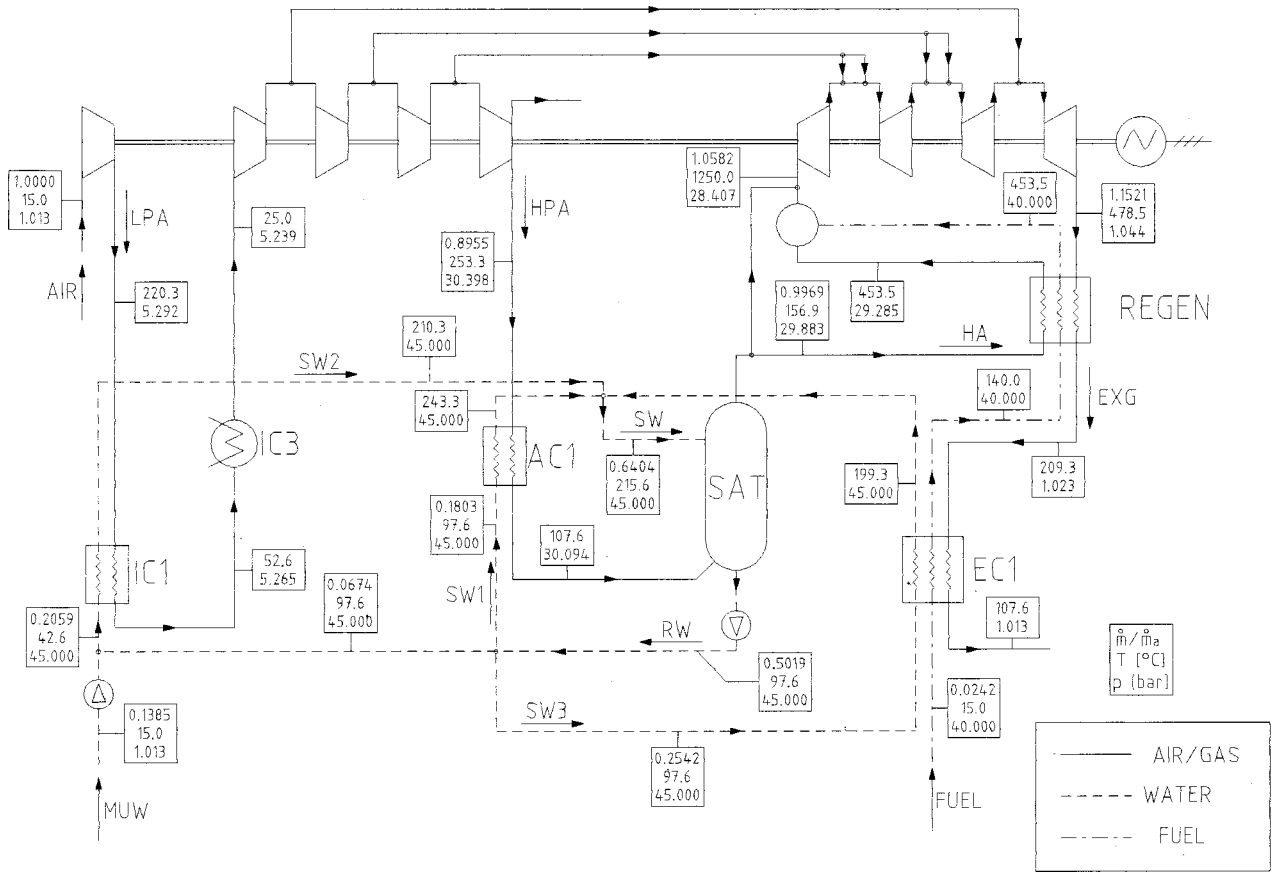


Fig. 8 Optimum plant structure for β=30 (one aftercooler—one economizer). AC2 and EC2 in the grid representation in Fig. 7 have been eliminated by imposing the final temperatures of the streams HPA and EXG equal to the temperatures at the outlet of AC1 and EC1, respectively.

The optimal HEN for β=30, having the minimum number of heat exchangers, can also be used for β=10 and 20. Two possible situations may occur in this case:

1 The thermal capacities of the hot and cold streams in the intercooler and in the aftercooler are set equal to each other. This implies that the thermal capacity of the water sent to the economizer is lower than the thermal capacity of the exhaust gases, or that the lower limit for the exhaust gases temperature (105°C) is violated.

2 The thermal capacities of the hot and cold streams in the economizer are set equal to each other by increasing the aftercooling final temperature (T_{ac}) until the temperature of the water at the bottom of the saturator is 95°C. In this case the ΔT at the cold end of the aftercooler is higher than 10°C, so that the thermal capacity of the water stream in the aftercooler is lower than the thermal capacity of the air stream.

In both cases the optimal conditions at the black-box boundaries are not exactly fulfilled. For β=20 the differences are negligible (the water temperature gain in the intercooler IC4 before the economizer inlet is only 9°C) and the total plant efficiency is 54.54 percent. For β=10 the water mass flow rate sent to the saturator ($\dot{m}/\dot{m}_a$=0.581) is much higher than the optimal one ($\dot{m}/\dot{m}_a$=0.5447—Fig. 6), whereas its temperature is lower (156.5°C instead of 161.7°C). In any case the effect on the total plant efficiency is not significant (η=54.21 percent).

4. Conclusions

In Part I of this paper the optimum operating conditions of the ''basic HAT cycle plant configuration'' were obtained. In this second part the various HENs that perform the heat exchange process in the black-box of the basic plant configuration were defined and analyzed, focusing attention on those having the minimum number of heat exchangers.

The main results of the analysis can be summarized as follows:

- The need to release heat to the environment explains why there is a dissipative heat exchanger in all the plant structures (the intercooler IC3 in Figs. 2–8 is chosen here—this heat exchanger is not mentioned again in the following for brevity).
- When the total pressure ratio β is equal to 30 the optimal conditions at the black-box boundaries can be approximately fulfilled using one intercooler, one aftercooler, and one economizer (only slight deviations from the optimal values of the final temperatures of the hot flows were observed).
- For lower pressure ratios (β=20, 10) the minimum allowed value for the temperature of the exhaust gases imposes the existence of at least one more heat exchanger (an economizer).
- Even quite important deviations from the optimal values of the choice variables found in the first part of the paper might not have significant effects on the total plant efficiency (the optimization process leads to points of relative maximum). This led us to consider the simplest plant structure for the lower pressure ratios (β=10 and 20) as well. In this case, for β=20, the value of the total plant efficiency is practically unchanged with respect to the one obtained using the optimal structure. For β=10, this efficiency decreases by only 0.4 percent.
- Thus, we can conclude that in the range of pressure ratios considered here (β=10–30) the maximum total plant efficiency can be attained either exactly or approximately by including one intercooler, one aftercooler, and one economizer in the plant

structure (besides the dissipative intercooler). Further increases in the number of heat exchangers are not useful for improving the total plant efficiency even with lower pressure ratios when they would be suggested by the strict agreement with the black-box optimal boundary conditions.

• The indications deriving from this thermodynamic optimization are believed to be a good basis for the choice of the actual plant structure. This choice is to be made in any case after introducing in the analysis the cost of the devices and performing a thermoeconomic optimization [7].

Acknowledgment

We would like to acknowledge Professor A. Mirandola for his helpful suggestions.

References

[1] Linhoff, B., Townsend, D. W., Boland, D., Hewitt, G. F., Thomas, B. E. A., Guy, A. R., and Marsland, R. H., 1982–1994, *A User Guide on Process Integration for the Efficient Use of Energy*, The Institute of Chemical Engineering, Rugby, Warks, UK.

[2] Sama, D. A., 1995a, ''The Use of the Second Law of Thermodynamics in Process Design'' J. Energy Resourc. Technol., **117**, pp. 179–185.

[3] Sama, D. A., 1995b, ''Differences between Second Law Analysis and Pinch Technology,'' J. Energy Resourc. Technol., **117**, pp. 186–191.

[4] Chiesa, P., Lozza, G., Macchi, E., and Consonni, S., 1995, ''An Assessment of the Thermodynamic Performance of Mixed Gas Steam Cycles: Part B—Water-Injected and HAT Cycles,'' J. Eng. Gas Turbines Power, **117**, pp. 499–508.

[5] Xiao, Y. H., Cai, R., and Lin, R., 1996, ''Modeling HAT Cycle and Thermodynamic Evaluation,'' *Proc. ECOS '96 Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems*, P. Alvfors, L. Eidensten, G. Svedberg, and J. Yan, eds., Royal Institute of Technology, Stockholm, Sweden, pp. 211–216.

[6] Bram, S., and De Ruyck, J., 1996, ''Exergy Analysis Tools for Aspen Applied to Evaporative Cycle Design,'' *Proc. of ECOS '96, Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems*, P. Alvfors, L. Eidensten, G. Svedberg, and J. Yan, eds., Royal Institute of Technology, Stockholm, Sweden, pp. 217–224.

[7] Bejan, A., Tsatsaronis, G., and Moran, M., 1996, *Thermal Design and Optimization*, John Wiley and Sons, New York.

D. M. Paulus, Jr.
e-mail: dmpaulus@execpc.com

R. A. Gaggioli

W. R. Dunbar

Department of Mechanical
and Industrial Engineering,
Marquette University,
Milwaukee, WI 53233

Entropy Production as a Predictive Performance Measure for Turbomachinery

It is proposed that consideration be given to an alternative, streamlined manner for mathematical modeling of the performance of energy conversion and transfer equipment. We make the case, here, by application to compressors. It is advocated that, instead of using an expression for efficiency as one of the governing equations, performance can be accounted for directly, with entropy production. It is shown that: (1) the modeling is more straightforward, using fewer relations, and (2) that compressor performance (e.g., maps) can be represented equally well. [DOI: 10.1115/1.1339983]

Introduction

Traditionally, various expressions for efficiency have been used to describe the performance of energy conversion hardware. These expressions, along with mass and energy balances, property relations, and appropriate boundary conditions are used to thermodynamically model a piece of hardware.

An equally valid set of governing relations can be obtained by deleting the efficiency and replacing it with an alternative, an expression for the entropy production in the hardware. This is desirable for two main reasons.

First, repetition that is present in the current modeling procedures is avoided. Consider compressors for example; today's modeling first solves for the performance of a hypothetical, ideal compressor. Then after correcting the calculated power input by using the efficiency, the energy balance and property relations are used a second time in order to determine the outlet conditions from the machine. Furthermore, if it is desired to evaluate the entropy production, the entropy balance needs to be used again. In any case, note that an entropy balance has been used, at least implicitly, in order to solve for the ideal performance.

Second, the use of dimensionless entropy production, instead of efficiency, presents a unified approach to equipment modeling, independent of the type of equipment. Today, each type is described with one of several efficiency definitions, such as isentropic efficiency, thermal efficiency, COP, heat exchanger effectiveness, and so on.

While these two drawbacks of the traditional procedures may appear to be only pedagogical, an entropy production approach could have advantages in simulation systems programming by reducing the amount of code necessary to model many pieces of equipment.

This paper will demonstrate the feasibility of replacing traditional isentropic efficiency with entropy production in the modeling of a compressor. The method outlined here should be equally applicable to other turbomachinery. To begin, the set of governing equations for both methods are shown.

Governing Equations for a Compressor[1]

The equations developed here are for a simple compressor with no extractions. Kinetic energy contributions are neglected for clarity. The compressor is considered adiabatic. The boundary conditions for both cases are taken as the inlet conditions (temperature, pressure, and composition), the rotational speed (N) of the compressor and the pressure ratio (Π). The desired output of the mathematical model is power and the outlet state (e.g., T_2).

With Isentropic Efficiency. A set of 11 governing equations is necessary for modeling a compressor given the above boundary conditions. These equations consist of balances, property relations, and performance relations.

Energy Balances

$$\dot{m}(h_1-h_2)+\dot{W}=0 \tag{1}$$

$$\dot{m}(h_1-h_{2,\text{ideal}})+\dot{W}_{\text{ideal}}=0 \tag{2}$$

Entropy Balance

$$s_1-s_{2,\text{ideal}}=0 \tag{3}$$

Property Relations

$$h_{2,\text{ideal}}=h(p_2,s_{2,\text{ideal}}) \tag{4}$$

$$h_1=h(p_1,T_1) \tag{5}$$

$$s_1=s(p_1,T_1) \tag{6}$$

$$T_2=T(p_2,h_2) \tag{7}$$

Performance Relations

$$\Pi=\frac{p_2}{p_1} \tag{8}$$

$$\eta_s=\frac{h_{2,\text{ideal}}-h_1}{h_2-h_1}=\frac{\dot{W}_{\text{ideal}}}{\dot{W}} \tag{9}$$

$$\eta_s=\eta(\Pi,N,R,T_1) \tag{10}$$

$$\dot{m}=\dot{m}(\Pi,N,R,T_1,p_1) \tag{11}$$

The foregoing system of equations is representative of industrial practice, where the right hand side of Eqs. (10) and (11) are given by a performance map. In the typical textbook problem, the right-hand side of these two equations would be replaced by *given* numbers, and rotational speed would not enter.

With Entropy Production. An equivalent set of governing equations, with the efficiency expressions replaced by an entropy production, can be formed with two fewer equations.

Energy Balance

$$\dot{m}(h_1-h_2)+\dot{W}=0 \tag{1}$$

[1]With sign changes, these sets of governing equations would be equally valid for a turbine.

Contributed by the Advanced Energy Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the AEDDivision, August 17, 2000; final revision received by the ASME Headquarters September 5, 2000. Editor: H. D. Nelson.

Entropy Balance

$$\dot{m}(s_1 - s_2) + \dot{S}_\pi = 0 \tag{12}$$

Property Relations

$$h_2 = h(p_2, s_2) \tag{13}$$

$$h_1 = h(p_1, T_1) \tag{14}$$

$$s_1 = s(p_1, T_1) \tag{6}$$

$$T_2 = T(p_2, s_2) \tag{15}$$

Performance Relations

$$\Pi = \frac{p_2}{p_1} \tag{8}$$

$$\dot{S}_\pi = \dot{S}_\pi(\Pi, N, R, T_1) \tag{16}$$

$$\dot{m} = \dot{m}(\Pi, N, R, T_1, p_1) \tag{11}$$

Not only are there two fewer equations here, but all of them relate, directly, to ''reality''—none are for hypothetical circumstances. (Also, the same system of equations could likewise be used to model other fluid machinery; e.g., a gas turbine.)

The question remains, ''Are the performance relationships represented by Eqs. (11) and (16) amenable to convenient, accurate representation by performance maps, dimensionless or quasidimensionless?''

Dimensional Analysis

In practice, it is impractical to map compressor performance for all possible operating points. So, efficiency and speed are mapped as parameters, with pressure ratio and dimensionless mass flow as coordinates. Therefore, a dimensional analysis will now be presented to find a dimensionless entropy production. Then it will be tested to compare its suitability as a replacement for efficiency, on a performance map.

Buckingham Pi Theorem. First, we allow

$$\dot{S}_\pi = \dot{S}_\pi(\dot{m}, P_1, P_2, T_1, N, D, R, \gamma, \nu) \tag{17}$$

This gives us a function of nine variables in four fundamental dimensions, which will allow a dimensionless expression for entropy production to be expressed as a function of five other dimensionless terms.

Four variables containing all four fundamental dimensions are selected, yielding the set $(\dot{m}, T_1, D, R)$. The variables have fundamental dimensions as follows:

$$\dot{S}_\pi[=]\frac{L^2 M}{t^3 \Theta} \tag{18a}$$

$$\dot{m}[=]\frac{M}{t} \tag{18b}$$

$$T_1[=]\Theta \tag{18c}$$

$$D[=]L \tag{18d}$$

$$R[=]\frac{L^2}{t^2 \Theta} \tag{18e}$$

We can multiply powers of these terms together in order to find the first pi term.

$$\dot{S}_\pi \dot{m}^\alpha T_1^\beta D^\chi R^\delta = \Pi_1 \tag{19}$$

Or, in terms of dimensions,

$$\left(\frac{L^2 M}{t^3 \Theta}\right)\left(\frac{M}{t}\right)^\alpha (\Theta)^\beta (L)^\chi \left(\frac{L^2}{t^2 \Theta}\right)^\delta = 1 \tag{20}$$

which gives four equations with four unknowns:

$$2 + \chi + 2\delta = 0 \tag{21a}$$

$$1 + \alpha = 0 \tag{21b}$$

$$-3 - \alpha - 2\delta = 0 \tag{21c}$$

$$-1 + \beta - \delta = 0 \tag{21d}$$

The solution of these equations yield $\alpha = -1$, $\beta = 0$, $\chi = 0$, $\delta = -1$, or $\dot{S}_\pi / \dot{m} R$, which will be referred to subsequently as σ, the dimensionless entropy production.

Similar manipulation of the other variables [1], along with multiplication by one another and the raising of them to powers, yield the groups

$$\frac{\dot{m}\sqrt{RT_1}}{pD^2}, \quad \frac{p_2}{p_1}, \quad \frac{ND}{\sqrt{RT_1}}, \quad \frac{ND^2}{\upsilon}$$

These are the traditional groups on which efficiency is dependent. Compressor maps typically will be based on the first three, or modifications of them. (For example, D might not be included.)

Equivalence of Dimensionless Entropy Production with Efficiency. Beginning with the expression for entropy production,

$$\dot{S}_\pi = \dot{m}(s_2 - s_1) \tag{22}$$

Assuming constant specific heat and dividing through by mass flow yields

$$\frac{\dot{S}_\pi}{\dot{m}} = c_p \ln\left(\frac{T_2}{T_1}\right) - R \ln\left(\frac{p_2}{p_1}\right) \tag{23}$$

It can readily be shown that

$$\frac{T_2}{T_1} = \frac{\left(\frac{P_2}{T_1}\right)^{R/c_p} + 1}{\eta} \tag{24}$$

Substituting (24) into (23), dividing by R, and using $\Pi = p_2/p_1$ yields

$$\frac{\dot{S}_\pi}{\dot{m}R} = \frac{c_p}{R}\ln\left(\frac{\Pi^{R/c_p} + 1}{\eta}\right) - \ln(\Pi) \tag{25}$$

As

$$\sigma = \frac{\dot{S}_\pi}{\dot{m}R}$$

and

$$\frac{\gamma - 1}{\gamma} = \frac{R}{c_p},$$

$$\sigma = \frac{\gamma}{\gamma - 1}\ln\left(\frac{\Pi^{\gamma - 1/\gamma} + 1}{\eta}\right) - \ln(\Pi) \tag{26}$$

Equation (26) shows that dimensionless entropy production is a function only of efficiency and variables upon which both it and efficiency depend. Therefore, dimensionless entropy production must work *at least as well as* efficiency in describing the performance of turbomachinery.

Testing of Dimensionless Relationships

Although the equivalence of dimensionless entropy production and efficiency were shown above, the authors desired to test this equivalence for a practical case. For this test, a hypothetical compressor off-design performance was computed from both maps of efficiency and dimensionless entropy production. Specific heats were not assumed constant in this testing.

Creation of Performance Map Based on Efficiency. The first step in testing was to create a performance map based on

efficiency, which could later be converted to dimensionless entropy production. A map for a hypothetical compressor was created, based on typical performance of an axial compressor. The map was made quasidimensionless, with an x axis of

$$M^* = \dot{m}\sqrt{\frac{RT}{R_0 T_0}}\frac{p_0}{p} \tag{27}$$

a y axis of pressure ratio, islands of constant efficiency, and lines of rotational speed,

$$N^* = N\sqrt{\frac{R_0 T_0}{RT}} \tag{28}$$

Here the subscript 0 represents design conditions, taken to be ISO conditions (288.15 K, 101.3 kPa, 60 percent relative humidity).

This graphical map was converted to computer use by programming interpolation functions, which worked from tabulated data. Mass flow and efficiency were made functions of speed and pressure ratio.

The efficiency map is shown in Fig. 1.

Creation of Performance Map Based on Dimensionless Entropy Production. Next, a map with dimensionless entropy production replacing efficiency was created. This was accomplished simply by using the efficiency map to calculate a dimensionless entropy production for numerous pressure ratios and speeds *at design conditions*. This map is shown in Fig. 2.

Comparison of Results for Off-Design Operation. Finally, the dimensionless relationship was tested by varying conditions *away from design*, calculating compressor performance with dimensionless entropy production and back calculating efficiency. Ideally, the efficiency calculated for a given M^* (or N^*) and pressure ratio should be the same as that determined directly from the efficiency performance map. The differences between original and back-calculated efficiencies are shown in Figs. 3–5. These figures show differences at various pressure ratios for several N^*'s. Maximum difference was 0.1 percent at 30°C for changes in temperature, 0.038 percent at 1.1 bar for changes in inlet pressure, and 2.0 percent at $R = 287.1$ kJ/kg K for changes in composition. This difference is small and certainly within the accuracy of a dimensionless map. Also, because dimensionless entropy production stems from an arguably more basic set of governing equations, off design performance might be actually better represented by it.

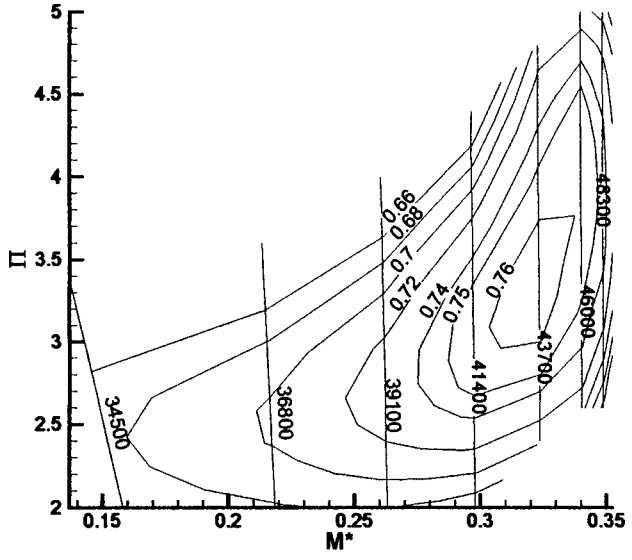


Fig. 1 Compressor map with efficiency

Example Calculations

An example calculation is shown here using both maps shown in Figs. 1 and 2. Specific heats are assumed constant.

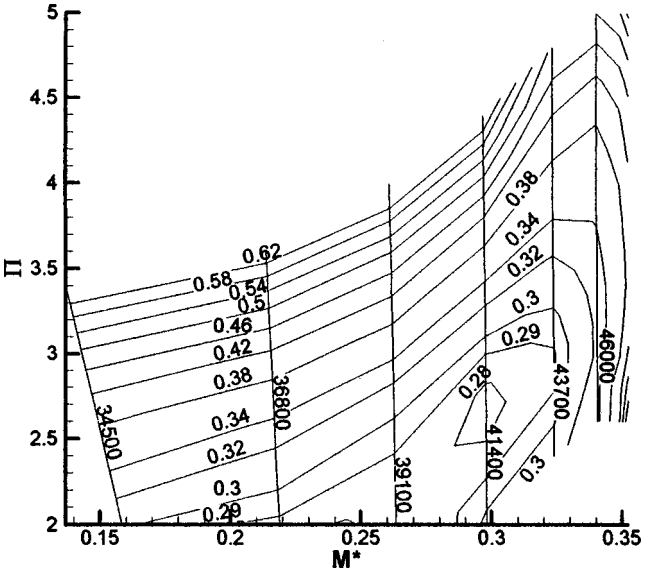


Fig. 2 Compressor map with dimensionless entropy production

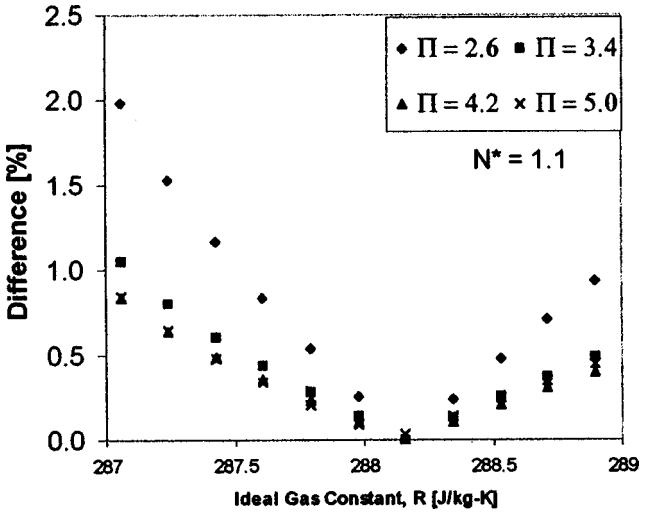


Fig. 3 Differences for varying gas constant

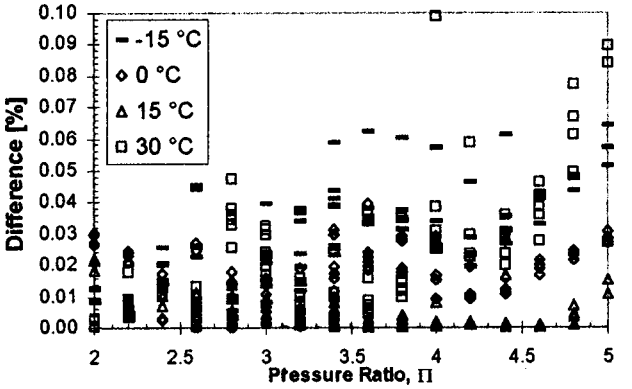


Fig. 4 Differences at various inlet temperatures

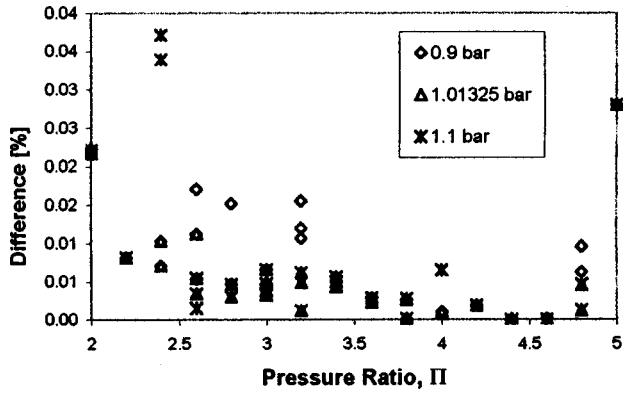


Fig. 5 Differences at various inlet pressures

With Isentropic Efficiency. For the conditions of $T_1 = 288.15$ K, $p_1 = 101.3$ kPa, relative humidity = 60 percent, and a pressure ratio of three, at an angular speed of 41,400 rpm, T_2 is calculated as follows:

The inlet conditions are ISO, so $N^* = N = 41{,}400$. From Fig. 1, $\eta = 0.755$. Using the set of governing equations that contain efficiency, T_{2s} is found to be 394.4 K, and T_2 is found to be 428.9 K.

If the composition of the air is assumed to be the same, the exit temperature can be found for the same conditions, with the exception of $T_1 = 303.15$ K. Using Eq. (28),

$$N^* = 41{,}400\sqrt{\frac{288.15}{303.15}} = 40{,}360$$

The efficiency corresponding to a pressure ratio of 3 and an angular speed of 40,360 rpm is 0.743. T_2 is found to be 453.7 K.[2]

With Dimensionless Entropy Production. The above calculations are now repeated with Fig. 2 and dimensionless entropy production. For the first case, with $T_1 = 288.15$ K ($N^* = 41{,}400$), σ is found from Fig. 2 to be 0.292.

With the appropriate governing equations, T_2 is found to be 428.7 K, the same as before within three significant figures. For the second case, with $T_1 = 303.15$ ($N^* = 40{,}360$), from Fig. 2 $\sigma = 0.313$. This results in $T_2 = 453.7$ K, the same as previously calculated.

Importantly, no intermediate calculation of a hypothetical T_{2s} is needed when using dimensionless entropy production.

Conclusions

Mathematical Modeling. It has been shown that the governing equations using a term for entropy production work at least as well as those using efficiency. Moreover, the governing equations using the terms for entropy production reduce the total number of equations necessary for modeling a compressor. We maintain that the explicit use of entropy balances, avoiding recourse to a two-step solution is more ''natural.'' In all real processes, entropy is generated, reflecting the kinetics of transport and conversion processes—the causes of dissipation. That dissipation must be accounted for in the governing equations. For *modeling* (i.e., simulation) purposes, rather than account for dissipation with ''efficiency''—in a different manner with a different definition of efficiency for each type of equipment—dimensionless entropy production provides a common technique for all types. Moreover, explicit rather than implicit use of the entropy balance avoids the need to solve for the ideal case before modeling the actual case, and hence reduces the number of governing equations. In summary, especially for students and newcomers, much confusion is avoided: (1) by providing a common technique, (2) by reducing the number of governing equations, and (3) by avoiding the circuitous route of solving for a hypothetical, ideal case on the way to the actual case.

In order to formulate a set of governing equations for modeling of equipment or processes, what are the guiding principles? Today the guiding principles generally are implicit, from experience—either personal experience or ''tradition'' (the experience of authorities, who have gone before us). We advocate the following ''guiding principles.''

1 Inquire, ''What are the additive properties that are independently being (a) transferred to or from the system, and/or (b) produced or consumed inside?'' For each of these write a balance equation.

2 To evaluate the transport terms and the net production terms in each of the balances, find appropriate kinetic relations (or, equivalently, for overall equipment, performance relations).

3 To evaluate the content of each additive property within the system, and the transport terms associated with flows, find the relevant thermostatic property relations.

4 Complement the foregoing balances, kinetic relations, and property relations with the pertinent initial and/or boundary conditions.

The resultant system of mathematical equations needs to include, of course, as many independent equations as there are unknowns. This formalism is certainly not foolproof.[3] Nevertheless, it is useful—capturing in explicit form what we generally do implicitly based upon experience. We find that it is helpful not only in practice but also pedagogically. The students see a general approach, applicable to all devices and processes, instead of system-specific canned—''Here's how you solve this type of problem,''—maneuvers. Hence, being less preoccupied with the different specific procedures, they get a better grasp of the fundamentals.

Performance Relations. Because compressor performance is normally presented in a dimensionless or quasidimensionless form, a dimensional analysis was performed on the governing equations including the entropy production. From this analysis, a dimensionless term for entropy production was found. Testing of this term revealed that it could replace efficiency on dimensionless maps with no perceptible difference in precision. Furthermore, because this term stems from a more natural set of governing equations there is a strong possibility that this method would lead to more accurate off-design predictions than efficiency.

One argument that could prevail against the use of entropy production versus efficiency is this: A quantitative value of the entropy production *apparently* gives the user little ''feel'' for how poorly (or how well) a device is performing. From an energy standpoint, this difficulty can be alleviated by multiplying the rate of entropy production by the appropriate temperature. (In the case

[2]In practice, the exit temperature is often calculated using the following two formulas, which are derivable from the set of governing equations that include efficiency:

$$\frac{T_{2s}}{T_1} = \left(\frac{p_2}{p_1}\right)^{(\gamma-1)/\gamma} \tag{29a}$$

$$T_2 - T_1 = \frac{T_{2s} - T_1}{\eta} \tag{29b}$$

The first represent ideal operation and the second corrects for the real operating efficiency.

When the entropy production approach is used, the latter pair of equations can be replaced by the single relationship

$$T_2 = T_1\left(e^{\sigma}\frac{p_2}{p_1}\right)^{(\gamma-1)/\gamma} \tag{30}$$

The latter equation (which is subject to the same approximations as the preceding pair) is readily derivable from the entropy balance and property relations:

$$\sigma = \frac{\dot{S}_\pi}{\dot{m}R} = \frac{\gamma}{\gamma-1}\ln\left(\frac{T_2}{T_1}\right) - \ln\left(\frac{p_2}{p_1}\right) \tag{31}$$

[3]For example, this formalism gives no guidance regarding the choice between a lumped or continuum model, a three-dimensional or one-dimensional model.

of compressors, this should be the inlet T; see Ref. [2].) That product is a measure of wasted power because of irreversibilities in the equipment.[4]

We emphasized ''apparently'' because, interestingly, the *dimensionless* entropy production is a direct measure of the (exergetic) inefficiency of the equipment. This can be observed by comparing the peak value of dimensionless entropy production in Fig. 2, with the peak value of efficiency in Fig. 1. The graphs are not exact inverses of one another, as the isentropic efficiency of a compressor approximates, but does not exactly equal the exergetic efficiency.

[4] $\eta_{II}=\dot{m}[h_2-h_1-T_0(s_2-s_1)]/\dot{m}(h_2-h_1)=(\dot{W}-\dot{m}RT_0\sigma)/\dot{W}$ (32)

With $T_0=T_1$ for the compressor,

$$\sigma=\dot{W}(1-\eta_{II})/\dot{m}RT_1 \tag{33}$$

Thus, σ is a dimensionless measure of the extent to which the actual exergy input to drive the compressor exceeds the minimum conceivable exergy input required for the exergy increase (of the compressed gas) that is actually accomplished by the compressor. The last expression may be rewritten as $(1-\eta_{II})/\dot{m}RT_1/\dot{W}$, which shows that σ is proportional to the *in*efficiency.

Future articles will extend these concepts to other types of energy conversion hardware. Similar work has already been accomplished for heat exchangers [3]. Extensions of this work could allow a unified equipment-independent approach to thermodynamic modeling [4].

Acknowledgment

This paper was written with partial support from a U.S. Air Force Office of Scientific Research grant.

References

[1] Lakshminarayana, B., 1996, *Fluid Dynamics and Heat Transfer of Turbomachinery*, John Wiley and Sons, New York.

[2] Gaggioli, R. A., Richardson, D. H., Bowman, A. J., and Paulus, D. M., Jr., 1999, ''Available Energy: Gibbs Revisited and Extended,'' AES, **39**, pp. 285–295.

[3] Sekulic, D. P., 1990, ''The Second Law Quality of Energy Transformation in a Heat Exchanger,'' ASME J. Heat Transfer, **112**, pp. 295–300.

[4] Paulus, D. M., Jr., Gaggioli, R. A., and Dunbar, W. R., 1999, ''Entropy Production as a Predictive Performance Measure: I. Turbomachinery,'' AES, **39**, pp. 531–535.

A Study of Combustion Characteristics of Gasified Coal Fuel

T. Hasegawa

M. Sato

Central Research Institute
of Electric Power Industry,
Yokosuka, Kanagawa, Japan

T. Nakata

Department of Aeronautics
and Space Engineering,
Tohoku University,
Sendai, Miyagi, Japan

The development of integrated, coal-gasification combined cycle (IGCC) systems provides cost-effective and environmentally sound options for meeting future coal-utilizing power generation needs in the world. The Japanese government and the Electric Power Industries in Japan promoted research and development of an IGCC system. We have being working on developing a low-NO_x combustion technology used in gas turbine combustors for IGCC. Each gaseous fuel produced from some raw materials contained CO and H_2 as the main combustible components, and a small amount of CH_4. Compositions and calorific values of gasified coal fuels varied widely depending on raw materials and gasifier types. Gaseous fuel, produced in various gasifiers, has a calorific value of 4–13 MJ/m^3, which is about one-tenth to one-third that of natural gas. The flame temperatures of fuels increase as the fuel calorific value rises. When the fuel calorific value rises 8 MJ/m^3 or higher, the flame temperature is higher than that of natural gas, and so NO_x production from nitrogen fixation is expected to increase significantly. Also, some gasified coal fuels contain fuel nitrogen, such as ammonia, if the hot/dry type gas cleaning system is employed. These factors affect the combustion characteristics of the gasified coal fuel. In this paper, we clarified the influence of gasified coal fuel properties on NO_x and CO emissions through experiments using a small diffusion burner and through numerical analysis based on reaction kinetics. The main results were as follows: 1 NH_3 conversion to NO_x increases with increasing CH_4 concentration in gaseous fuel. 2 If gaseous fuel contains CH_4, there will be some specific equivalence ratio in the primary combustion zone for the minimum NH_3 conversion to NO_x in the two-staged combustion. 3 Its specific equivalence ratio in the primary combustion zone increases with decreasing CH_4 concentration in gaseous fuel. 4 If the fuel contains a small percent of CH_4, there is no influence of the CO/H_2 molar ratio in the fuel on the conversion rate of NH_3 to NO_x, while there is an influence in the case where fuel contains no CH_4. The conversion rate increases with rises in the CO/H_2 molar ratio. 5 As the pressure increases, the conversion rate of NH_3 to NO_x slightly decreases and the CO emission declines significantly.
[DOI: 10.1115/1.1287586]

Introduction

Coal is positioned as a vital source of energy both at the present time and into the future because coal is widely distributed all over the world and reserves are abundant. However, compared to oil and natural gas, coal contains greater quantities of sulfur, nitrogen, and ash, etc., and produces more CO_2 for the same fuel calorific value. Therefore, the usage methods of coal must be highly efficient and have excellent environmental protection features. With respect to these conditions, research and development into integrated coal gasification combined cycle system (IGCC system) as the next method of thermal power generation using coal have been carried out inside and outside of Japan [1].

Japan has a national research and development project into air-blown entrained-flow IGCC system using a pilot plant with a capacity of 200 tons per day. An oxygen-blown IGCC system is under development overseas [2]. The calorific value of gasified coal fuel differs according to the type of gasification agent used in the gasifier. If the gasification agent is air, then gasified coal fuel forms a low calorific fuel of about 4 MJ/m^3, but if the agent is oxygen, then the fuel becomes a medium calorific fuel between approximately 9–13 MJ/m^3. To increase the thermal efficiency of IGCC, it is necessary to use a hot/dry type gas cleaning system, but this produces ammonia derived from nitrogenous compounds in coal in the gasifier. This NH_3 is then fed into the gas turbine where it forms fuel-NO_x in the combustion process. For this reason, technology to suppress fuel-NO_x is important.

The research into the basic characteristics of gasified coal fuel combustion includes research into the flammability limit of mixed gas consisting of CH_4 or H_2 diluted with N_2, Ar, or He [3]; the impact of N_2 on burning velocity [4]; the effect of N_2 and CO_2 on flammability limits [5,6]; the combustion characteristics of low calorific fuel [7,8]; basic research into the premixed combustion of medium calorific fuel [9]; investigations using model combustors [10,11]; and research into gas turbines for power generation and other industrial uses [12–15]. However, there is very little systematic research into the effect of fuel composition or calorific value on the formation characteristics of NO_x from gasified coal fuel, including NH_3, or into the low NO_x combustion method.

With regard to the gasified coal fuel, which contains CO and H_2 as the main combustible components, we researched the characteristics of combustion using a small diffusion burner [16], the production characteristics and suppression methods of thermal-NO_x and fuel-NO_x, by carrying out combustion tests and analyzing reaction kinetics [17–19]. Based on the knowledge acquired from this research, we carried out further studies into low NO_x gas turbine combustors for low-Btu gasified coal fuel at exhaust gas temperatures of 1573 K (1300 °C) and at 1773 K (1500 °C) [20–25], and research into gas turbine combustors designed for oxygen-blown medium-Btu fuel [26].

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-398. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

This paper reports the effects of fuel calorific value, the fuel CO/H_2 molar ratio, CH_4 and NH_3 concentrations in the fuel on the characteristics of NO_x production and CO emission in the combustion process, using small diffusion burners at atmospheric pressure, as part of the development of low NO_x combustion technology for gas turbines using gasified coal fuel. We carried out researches into the application and effectiveness of two-stage combustion, which is a method of combustion aiming to suppress the production of thermal-NO_x and fuel-NO_x, and into the effect of pressure in the combustor on combustion characteristics. We also clarified the fuel-NO_x formation mechanism through numerical analysis based on reaction kinetics.

Test Device and Methods

Test Device. Figure 1 is an outline of the test device. The air used for combustion is input by an air compressor. After the flow of the compressed air is controlled to the required amount by the orifice meter and flow control valve, the air is heated to the prescribed temperature by an air heater, then blown into the combustor. Primary air is injected into the combustor through a primary air swirler positioned around the fuel injection nozzle, and secondary air is input through air holes in the side walls of the combustor.

CO and H_2, which are the main combustible components of the fuel, are mixed in the prescribed proportions and fed into a cylinder to be used. The combustible components are diluted with N_2 to create the required calorific value, heated to the prescribed temperature by an electric heater, and fed through the fuel injection nozzle. The small amounts of NH_3 and CH_4 are controlled by a thermal mass flow controller, and premixed with the fuel before reaching the fuel injection nozzle.

Figure 2 is an outline of the combustor and the tested burner. The combustor consists of a cylinder-style combustion chamber with an inner diameter of 90 mm and a length of 1000 mm, and the primary air swirler and fuel injection nozzle. The combustion chamber is lined with heat insulating material and the casing is cooled with water. There are four sections for secondary air on the side walls of the combustion chamber along the direction of flow, to simulate two-stage combustion. We used the secondary air inlets at a distance from the edge of the fuel injection nozzles of $3\times D$, where D is the inner diameter of the combustor. The diameter of the secondary air inlets at the entry to the combustor is 13 mm; altogether, twelve inlets are positioned on the perimeter of one cross-section.

The tested burner consists of a fuel injection nozzle and a primary air swirler. There are twelve injection inlets with a diameter

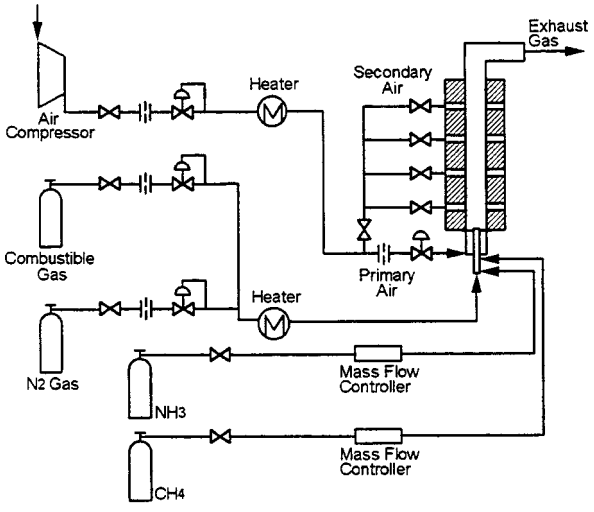


Fig. 1 Schematic diagram of experimental device

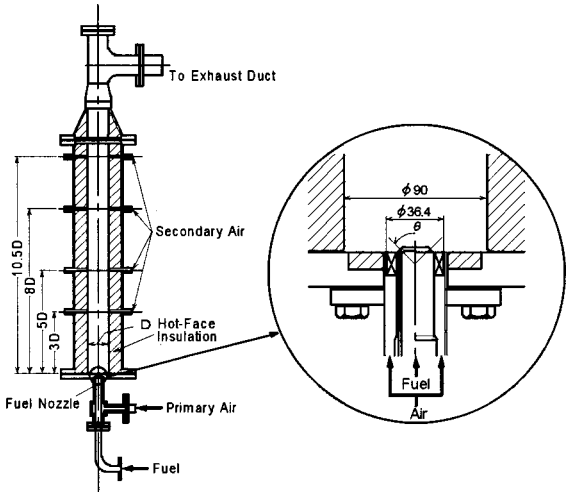


Fig. 2 Combustor and burner

of 1.5 mm on the fuel injection nozzle with an injection angle, θ, of 90 deg. The primary air swirler has an inner diameter of 24.0 mm, an outer diameter of 36.4 mm, and twelve vanes with a swirl angle, θa, of 45 deg. Swirl number, S, is 0.84. Sample gases were extracted from the exit of the combustor through a water-cooled stainless steel probe and continuously introduced into an emission console which measured CO and CO_2 by infrared analysis, NO_x by chemiluminescence analysis, O_2 by paramagnetic analysis, and hydrocarbons by flame ionization.

Tested Fuel. The typical compositions [15,27–34] of gaseous fuel produced in air or oxygen-blown gasifiers are shown in Table 1. Each type of gaseous fuel produced raw material with CO and H_2 as the main combustible components, and small amounts of CH_4. Fuel calorific values varied widely (4.2–13.0 MJ/m^3), from about one-tenth to one-third of that of natural gas, depending upon the raw material, the gasification agent, the gasifier type and IGCC system. For example, a gaseous fuel derived from biomass contained 30–40 percent steam in the gaseous fuel.

Table 2 shows typical conditions of tested fuel in the case of a 4.4 MJ/m^3 higher heating value. In the tests, fuel calorific values were varied between 3.0–11.4 MJ/m^3 with nitrogen dilution and the fuel CO/H_2 molar ratios were set to values of 0.43, 1.00, and 2.33, to reveal the combustion characteristics of gaseous fuel produced from different coal gasifiers and the effect of including small quantities of NH_3 and CH_4 in the fuel.

Higher fuel calorific values are generally considered to emit more thermal-NO_x because of the increased flame temperature. Therefore, in oxygen-blown IGCC, N_2 produced by the air separation unit is utilized for recovering power in order to increase the thermal efficiency of the plant and also to reduce the emission of NO_x by reducing the flame temperature.

Figure 3 shows the adiabatic flame temperature of fuels, which were (1) gaseous fuels with fuel calorific values (HHV) of 12.6, 8.4, 4.2 MJ/m^3, and (2) methane as the main component of natural gas. Flame temperatures were calculated using a CO/H_2 mixture (CO/H_2 molar ratio of 2.33:1) which contained no CH_4 under any condition, and the fuel calorific value was adjusted with nitrogen.

In the case of gasified coal fuel, as the fuel calorific value increases, the theoretical adiabatic flame temperature also increases. Fuel calorific values of 4.2 MJ/m^3 and 12.6 MJ/m^3 produce maximum flame temperatures of 1980 K and 2510 K respectively. At fuel calorific values of 8.4 MJ/m^3 or higher, the maximum flame temperature of the gaseous fuel exceeded that of high calorific gases such as methane, while the fuel calorific value was as low as one-fifth of methane. Furthermore, as the fuel calorific value increases, the equivalence ratio, which indicates maxi-

Table 1 Typical compositions delivered from the gasifiers

Fuel	Coal						Biomass		Heavy residue	Orimulsion™
Gasifier type	Fixed	Fluidized	Entrained				Entrained		Entrained	Entrained
Fuel feed	Dry	Dry	Dry			Slurry				
Developer	BGL [27]	BC [28]	IGC [29]	Shell [30]	Hycol [31]	Texaco [15]	Tampella [32]	[33]	Texaco	CRIEPI [34]
Oxydizer	O2	Air	Air	O2	O2	O2	Air	O2	O2	O2
Composition										
CO	56.4%	7.9–14.7%	25.9–27.6%	65.2–69.5%	55.2–59.4%	49.0%	8.0–15.0%	21.9–23.1%	51.7%	43.5%
H2	25.6%	13.2–15.0%	10.9– 9.4%	28.8–31.0%	31.1–33.7%	34.0%	8.0–12.0%	12.5–22.4%	43.1%	42.2%
CH4	6.6%	1.5– 2.8%	1.4– 0.5%	0.01–0.03%	1.0– 2.0%	0.2%	4.0– 8.0%	2.2%	0.2%	0.4%
CO2	2.8%	10.0–12.0%	6.7– 5.4%	1.0– 2.8%	7.6–10.4%	9.7%	13.0–18.0%	20.7–18.6%	3.2%	11.8%
H2O	–(a)	11.5–18.4%	–(a)	(Dry base)	–(a)	(Dry base)	7.0–15.0%	40.9–31.5%	(Dry base)	(Dry base)
NH3	–(a)	500–1000ppm	1000ppm(b)	100–600ppm	–(a)	–(a)	–(a)	0–200ppm	–(a)	–(a)
H2S+COS	20ppm	–(a)	404–714ppm	0.14– 1.1%	–(a)	–(a)	–(a)	0.285–1.132%	1.6%	1.35%
Others(N2 etc)	8.6%	45.9–47.3%	54.2–56.1%	–(a)	–(a)	7.1%	–(a)	1.800–1.048%	0.2%	0.75%
CO/H2 mole ratio	2.2	–(a)	2.4–3.0	2.1– 2.4	1.6–1.9(b)	1.4	–(a)	1.0–1.8	1.2	1.0
HHV[MJ/m^3]	13.0	3.6–4.1	4.9–5.2	12.2–12.5	12.0(b)	10.6	4.0–7.0	5.2–6.6	12.1	11.0

(a):Faulty description, (b):Estimated value

mum flame temperature, also increases. This is because the effect of thermal dissociation increases as the adiabatic flame temperature increases.

CO and H_2 are the main combustible components of gasified coal fuel, but the quantities differ according to the gasification method. Figure 4 shows the correlation between the equivalence ratio and the theoretical adiabatic flame temperature with the fuel CO/H_2 molar ratio as a parameter, when the fuel calorific value was set at 4.2 MJ/m^3 and CH_4 was not contained in the fuel [35]. Figure 4 shows that the theoretical adiabatic flame temperature is affected very little by the CO/H_2 molar ratio.

Numerical Analysis Based on an Elementary Reaction Model. We clarified the fuel-NO_x production characteristics, derived from ammonia in the gasified coal fuel, through a numerical analysis based on the reaction kinetics.

Table 2 Typical composition of the tested fuel

Composition	**CO**	**24.3 %**
	H2	**10.4 %**
	CH4	**0.0 %**
	CO2	**0.0 %**
	H2O	**0.0 %**
	NH3	**1000 ppm**
	N2	**65.2 %**
HHV	**4.4 MJ/m³ (1051 kcal/m³)**	
LHV	**4.2 MJ/m³ (1002 kcal/m³)**	

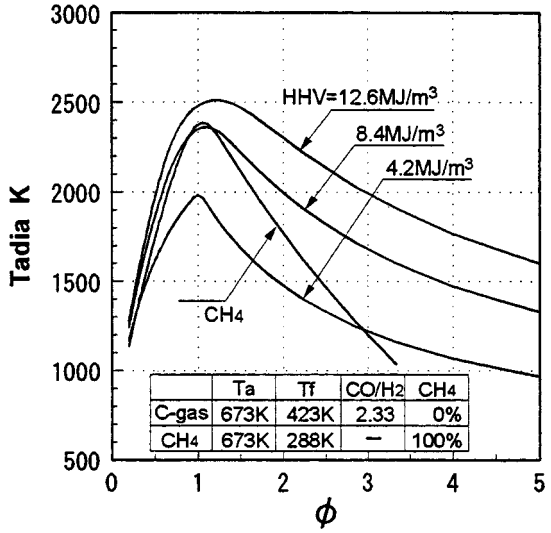


Fig. 3 Adiabatic flame temperature

The reaction model employed here was proposed by Miller and Bowman [36], and the appropriateness of the model for non-catalytic reduction of ammonia using NO [37] and the oxidation of ammonia by premixed flame of methane [36] has been confirmed by comparison with test results.

The reaction scheme is composed of 248 elementary reactions and 50 species are taken into account: CH_4, CH_3, CH_3O, CH_2OH, CH_2O, CH_2, CH, CO, CO_2, C, H_2, H, O, OH, O_2, H_2O, HO_2, H_2O_2, N_2, N, NO, NO_2, N_2O, HCN, CN, H_2CN, NH_3, NH_2, NH, NNH, HNO, HCNO, HOCN, HNCO, NCO, C_2N_2, HCO, CH_2CO, C_2H, C_2H_2, C_2H_3, C_2H_4, C_2H_5, C_2H_6, HCCO, HCCOH, C_3H_2, C_3H_3 and C_4H_2, C_4H_3. Values for thermodynamic data are taken from the JANAF thermodynamics tables [38]. For those species not listed in the tables, the values are calculated based on the relationship between the Gibbs' standard energy of formation ($\Delta G°$) and the chemical equilibrium constant (K).

$$\Delta G° = RT \ln(\mathrm{K}) \tag{1}$$

In the above formula, the value of $\Delta G°$ is obtained from the CHEMKIN database [39]. This study used the GEAR method [40] for the numerical analysis, as an implicit, multi-stage solution.

The model of the flow inside the combustor introduces the Pratt model [41] that linearly simplifies flows and each stage combustion zone is assumed to be a perfectly stirred reactor. Also, it is

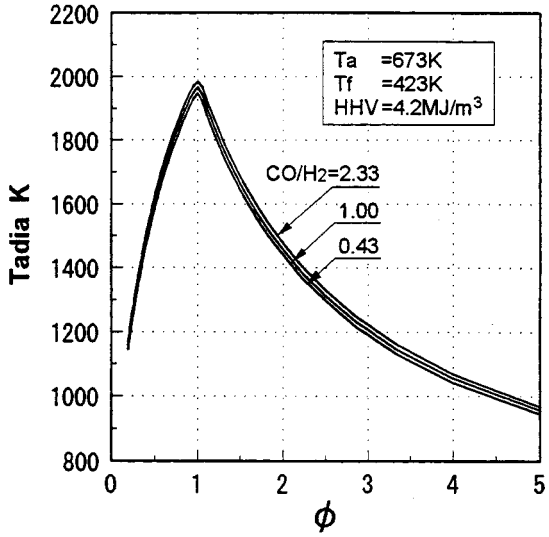


Fig. 4 Effects of CO/H_2 molar ratio in the fuel on adiabatic flame temperature

assumed that in the reaction process, the species are evenly mixed, and diffusion and stirring processes are not taken into consideration.

Fundamental Characteristics of Combustion

Nitrogenous compounds in fuel produces NO_x, called fuel-NO_x, in the combustion process. Fuel-NO_x has a different production mechanism from thermal-NO_x, which is derived from N_2 in the air. NH_3 in gasified coal fuel is easily converted into fuel-NO_x in the combustion process in the gas turbine and most of the NO_x emitted from the plant is fuel-NO_x. Therefore, technology for low NO_x combustion which suppresses fuel-NO_x is required for gas turbine combustors using gasified coal fuel.

NO_x Emission Characteristics in Non-Two-Stage Combustion. Two-stage combustion, which has a primary stage combustion under a fuel-rich condition, and a secondary stage which completely combusts the remaining unburned fuel, is widely accepted as being combustion technology which suppresses production of fuel-NO_x [42,43]. Non-two-stage combustion was researched before investigating the NO_x production characteristics in two-stage combustion of gasified coal fuel.

(1) Effect of Equivalence Ratio. Figure 5 shows the relationship between the equivalence ratio, ϕ ex, and both the conversion rate of NH_3 in fuel to NO_x, and the total concentration of NO_x emissions, at a fuel calorific value of 4.4 MJ/m^3 (at 273 K, 0.1 MPa) and when 1000 ppm of NH_3 was included in the fuel. Figure 5 also shows the concentration of thermal NO_x in the emission when NH_3 was not included in the fuel.

All combustion air was provided from the primary air swirler around the fuel injection nozzle. During the test, the temperature of the combustion air was preset and maintained at 673 K and the temperature of the fuel at 423 K. The equivalence ratio was adjusted by controlling the flow of combustion air and maintaining a constant flow of fuel. To obtain the conversion rate of NH_3 in the fuel to NO_x, the concentration of thermal-NO_x was first measured after stopping the supply of NH_3, then the concentration of total NO_x was measured while NH_3 was supplied, and finally fuel-NO_x was calculated by deducting the concentration of thermal-NO_x from that of total NO_x.

As shown in Fig. 5, the conversion rate of NH_3 in the fuel to NO_x was high, at approximately 84 percent, when the equivalence ratio was less than 1, and rapidly dropped when the equivalence ratio exceeded 1. Measured concentrations of emitted thermal-NO_x and total NO_x (dry base) reached a maximum when the equivalence ratio was around 1, at approximately 10 ppm and 520 ppm, respectively.

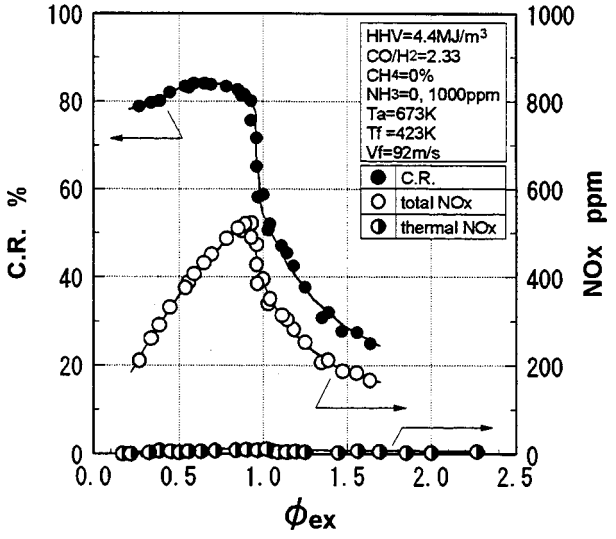


Fig. 5 NO_x emission characteristics in low-Btu gas flame

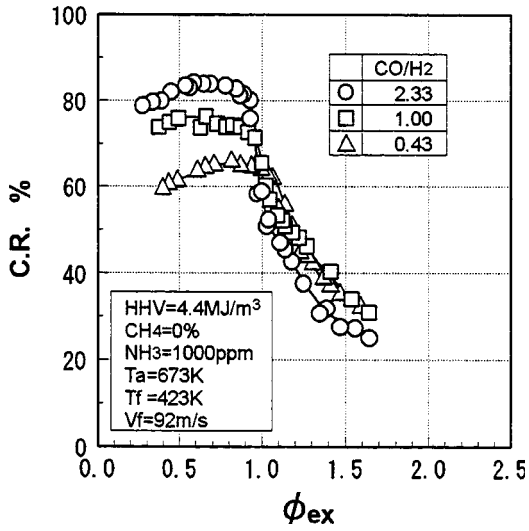


Fig. 6 Effects of CO/H_2 molar ratio on the conversion rate of NH_3 in the fuel to NO_x defining by the experiments

(2) Effect of the CO/H_2 Molar Ratio In the Fuel. The production characteristics of fuel-NO_x, derived from NH_3 in the fuel, are significantly affected by the characteristics of the flame zone in which most of the combustion reaction progresses, that is the characteristics of the flame zone, such as the temperature, the equivalence ratio, the oxidation of the fuel, and the concentration of radicals (O, OH, H, etc.) [44], and the production of fuel-NO_x is also affected by the type of fuel.

Figure 6 shows the correlation between the equivalence ratio, ϕex, and the conversion rate of NH_3 in the fuel to NO_x, with the fuel CO/H_2 molar ratio as a parameter, when the NH_3 concentration in the fuel and the calorific value were set and maintained at 1000 ppm and 4.4 MJ/m^3, respectively. The tested fuel included no CH_4 and had CO/H_2 molar ratios of 0.43, 1.00, and 2.33. The fuel calorific value was controlled using N_2. When the equivalence ratio was less than 1 (i.e., fuel-lean condition), the conversion rate of NH_3 in the fuel to NO_x was almost constant at any CO/H_2 molar ratio. However, when the equivalence ratio is more than 1 (i.e., fuel-rich condition), the conversion rate decreased rapidly. Concerning the effect of the CO/H_2 molar ratio, when the equivalence ratio is less than 1 (i.e., fuel-lean condition), higher CO/H_2 molar ratios increased the conversion rate, and when the equivalence ratio is more than 1 (i.e., fuel-rich condition), no CO/H_2 molar ratio was observed. This is because both H_2 is oxidized more rapidly than CO, and preferential diffusion of H_2 in the presence of excess air grows strongly as the CO/H_2 molar ratio becomes higher. That is to say, when the CO/H_2 molar ratio increases, the O_2 consumption rate decreases. Consequently, NH_3 is oxidized in the area with higher concentration of O_2, and as a result the NO_x production rate is increased. On the other hand, when the equivalence ratio is more than 1, O_2 is consumed by the oxidation of CO and H_2, and the production of NO_x by the oxidation of NH_3 is restrained.

Figure 7 shows the correlation between the equivalence ratio, ϕex, and the CO emission concentration, with the molar ratio of CO/H_2 in the fuel as a parameter. Combustion efficiency was calculated from the unburned fuel loss. The fuel calorific value, HHV, was set to 4.4 MJ/m^3, Ta to 673 K, Tf to 423 K, the fuel flow velocity, Vf, set and maintained at 92 m/s and no CH_4 contained in the fuel. When ϕex was less than 1 (i.e., fuel-lean condition), CO was not detected in the exhaust gas and combustion efficiency was nearly 100 percent at any CO/H_2 molar ratio.

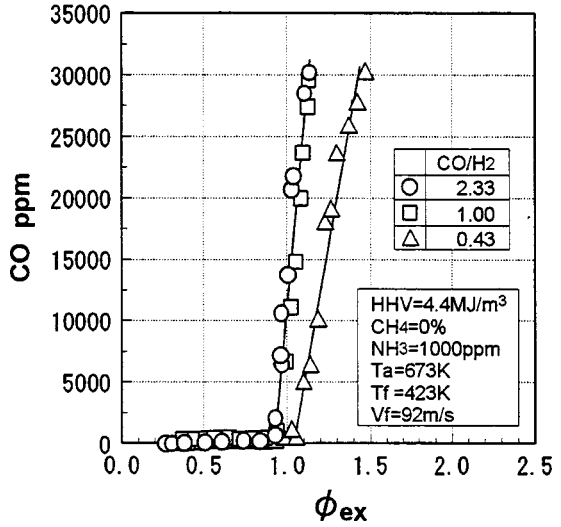


Fig. 7 Effects of CO/H_2 molar ratio in the fuel on CO emission characteristics defining by the experiments

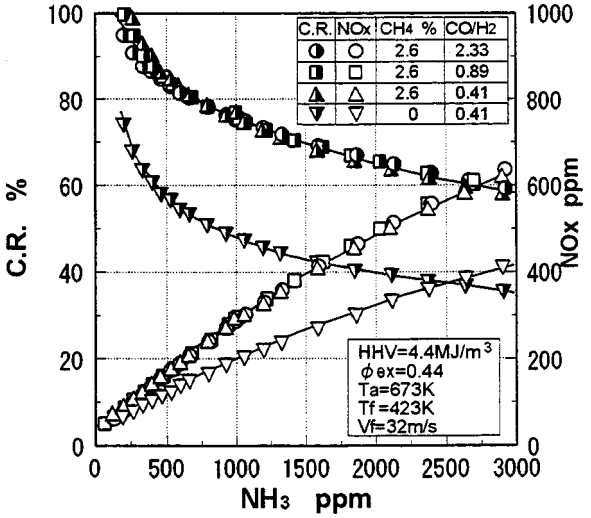


Fig. 9 Effect of NH_3 concentration on the conversion rate of NH_3 in the fuel to NO_x with CH_4 concentration and CO/H_2 molar ratio as parameters defining by the experiments

However, when ϕex exceeded 1 (i.e., fuel-rich condition), the CO emission concentration increased rapidly and the effect of the CO/H_2 molar ratio was detected. When the CO/H_2 molar ratio in the fuel was set to 2.33 and 1.00, the CO emission concentration was almost unchanged, but when the CO/H_2 molar ratio was reduced to 0.43, the combustion stability improved and the CO emission concentration declined.

(3) Effect of the Concentration of NH_3 in the Fuel. The concentration of NH_3 in gasified coal fuel is largely determined by the gasifier, the type of coal, and gasification conditions. Therefore, we carried out a study into the effect of NH_3 concentrations in the fuel on NO_x emission characteristics. Figure 8 shows the correlation between the concentration of NH_3 in the fuel and both the conversion rate of NH_3 to NO_x, and the NO_x emission concentration, when CH_4 was not contained in the fuel and the fuel calorific value was set and maintained at 4.4 MJ/m^3. As the concentration of NH_3 in the fuel increases, the concentration of emitted NO_x increases, but the conversion rate declines. It is generally accepted that increased concentration of nitrogenous compounds in the fuel suppresses the conversion rate of those nitrogenous compounds into NO_x [45–48]. Similar findings were observed in gasified coal fuel.

Figure 9 shows the correlation between the concentration of NH_3 in the fuel and the characteristics of NO_x emission, with the concentration of CH_4 and the CO/H_2 molar ratio as parameters. When the concentration of CH_4 in the fuel was set and maintained at 2.6 percent, no effect of the CO/H_2 molar ratio in the fuel on NO_x emission characteristics was detected. A comparison of conversion rates for fuel containing CH_4 (at a concentration of 2.6 percent) and fuel containing no CH_4, while maintaining the CO/H_2 molar ratio in the fuel at 0.41, shows that the NH_3 to NO_x conversion rate in the fuel containing CH_4 increases by 25 percent, regardless of the concentration of NH_3 in the fuel.

(4) Effect of Fuel Calorific Value. The calorific value of gasified coal fuel changes according to the type of gasification agent and conditions in the coal gasification process. Therefore, we studied the characteristics of NO_x emission using different fuel calorific values.

Figure 10 shows the correlation between the equivalence ratio,

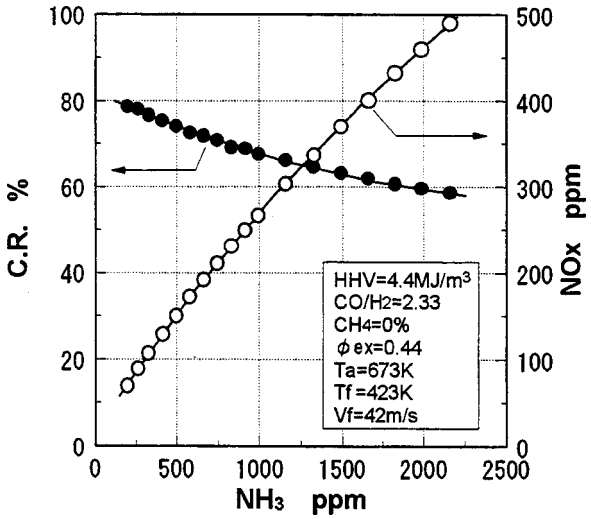


Fig. 8 Effects of NH_3 concentration in the fuel on the conversion rate of NH_3 in the fuel to NO_x defining by the experiments

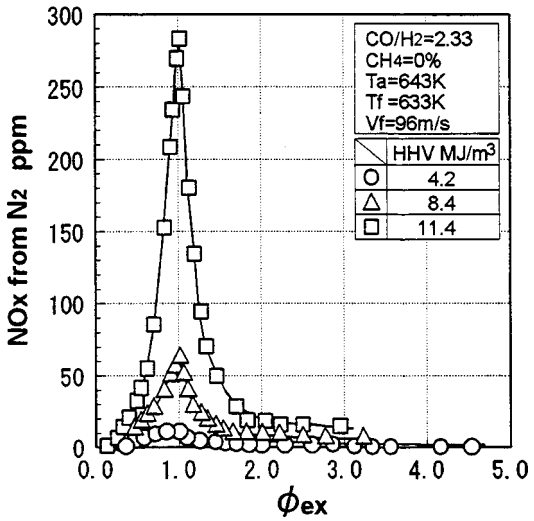


Fig. 10 Effects of fuel calorific value on NO_x emission characteristics from nitrogen fixation defining by the experiments

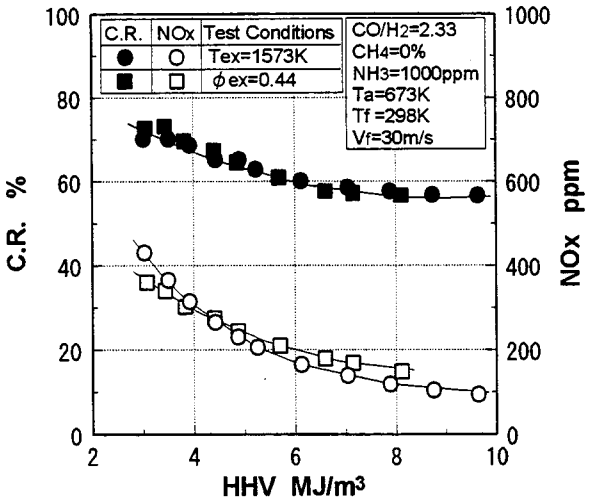


Fig. 11 Effects of fuel calorific value on the conversion rate of NH_3 in the fuel to NO_x and NO_x emission defining by the experiments

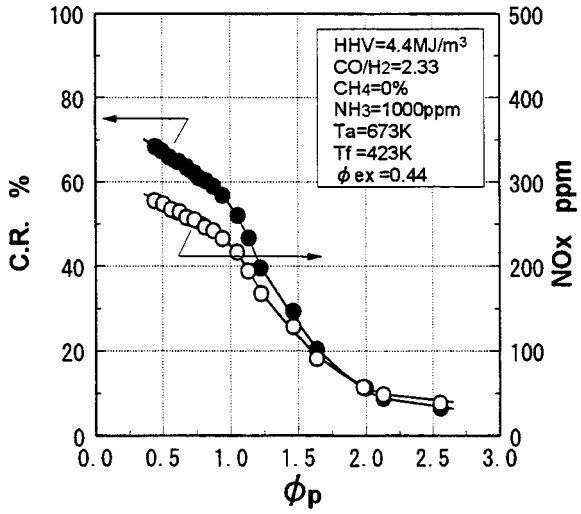


Fig. 12 NO_x emission characteristics in two-staged combustion defining by the experiments

ϕex, and thermal-NO_x emission, with fuel calorific value as a parameter, when NH_3 and CH_4 were not included. Figure 10 shows that the thermal-NO_x emission concentration increases as the fuel calorific value increases. When the fuel calorific value was set to 4.2 MJ/m^3, NO_x was produced at a concentration of 10 ppm or less, at equivalence ratios between 0.1–3.0. When the fuel calorific value was set to 11.4 MJ/m^3, the NO_x emission concentration reached a maximum of approximately 280 ppm.

Figure 11 shows the correlation between the fuel calorific value and both the conversion rate of NH_3 in the fuel to NO_x, and the NO_x emission concentration, when CH_4 was not included. The effect of varying the fuel calorific value was tested under conditions of (1) varying the equivalence ratio, ϕex, to maintain a constant theoretical flame temperature of 1573 K, and (2) maintaining a constant equivalence ratio, ϕex, of 0.44. Concerning the effect of changes in the fuel calorific value on the production characteristics of fuel-NO_x, Fenimore [49] reported that the conversion rate of nitrogenous compounds in the fuel into NO_x increases as the concentration of H_2 or CH_4 is diluted with inert gas. Our tests on gasified coal fuel also showed that reducing the fuel calorific value tends to increase the conversion rate. This effect is due to the decomposition process of NH_3 in the fuel proceeding into the fuel-rich region when the fuel calorific value is increased (by reducing the amount of N_2 dilution), and the NH_i produced reacts with NO to create N_2.

NO_x Emission Characteristics in Two-Stage Combustion. It is known that the fuel-NO_x production mechanisms of hydrocarbon fuels such as CH_4 and non-hydrocarbon fuels such as CO and H_2 due to reaction with nitrogenous compounds in the fuel, are different [50–53]. Gasified coal fuel consists of CO and H_2 as main combustible components, but also contains between thousands of ppm and a small percent of CH_4. Therefore, we investigated the effect of CH_4 on the production characteristics of fuel-NO_x caused by NH_3 in the fuel.

(1) Effect of Primary Equivalence Ratio. Figure 12 shows the correlation between the primary equivalence ratio, ϕp, and both the conversion rate of NH_3 in the fuel to NO_x, and the NO_x emission concentration, when the fuel did not contain CH_4, the equivalence ratio at the combustor exit was set to 0.44, and the exhaust gas temperature was 1573 K. The secondary air for the two-stage combustion was input from inlets positioned at a distance from the edge of the fuel injection nozzles of $3\times D$, where D is the inner diameter of the combustor. Even when the exit of the combustor has a condition of excess air and an equivalence ratio is 0.44, by setting a fuel-rich condition when the primary equivalence ratio is 1 or higher, the NH_3 to NO_x conversion rate was significantly decreased. This result shows that two-stage combustion is an effective combustion method to suppress the production of fuel-NO_x caused by NH_3 in fuel.

(2) Effect of the CO/H_2 Molar Ratio in the Fuel. Figure 13 shows the correlation between the primary equivalence ratio in two-stage combustion, ϕp, and the conversion rate of NH_3 in the fuel to NO_x, when the equivalence ratio at the combustor exit, ϕex, was set to 0.44, for fuel containing CH_4 and for fuel not containing CH_4, with the CO/H_2 molar ratio as a parameter. When the fuel did not contain CH_4, the CO/H_2 molar ratio exerted a significant effect on the conversion rate of NH_3 to NO_x when the primary equivalence ratio, ϕp, was less than 1 and under the condition of excess air. When the primary equivalence ratio, ϕp, was set to 0.44, and the CO/H_2 molar ratios were set to 0.41 and 2.33, the conversion rates were 47 percent and 68 percent respectively, showing that the conversion rate increases proportionate to the CO/H_2 molar ratio. However, when CH_4 is con-

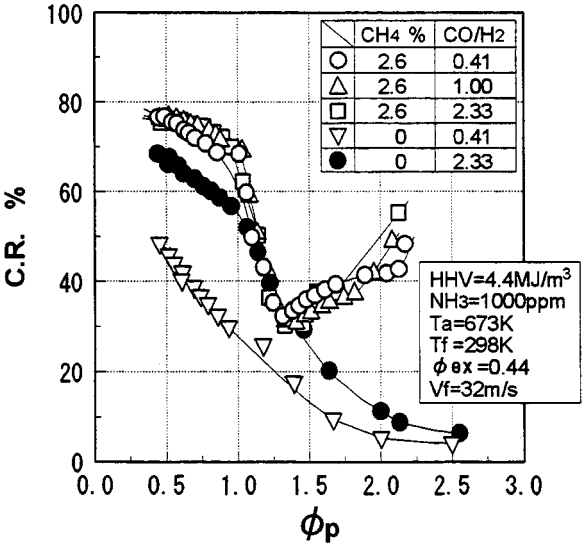


Fig. 13 Effect of CO/H_2 molar ratio on the conversion rate of NH_3 in the fuel to NO_x in two-staged combustion defining by the experiments

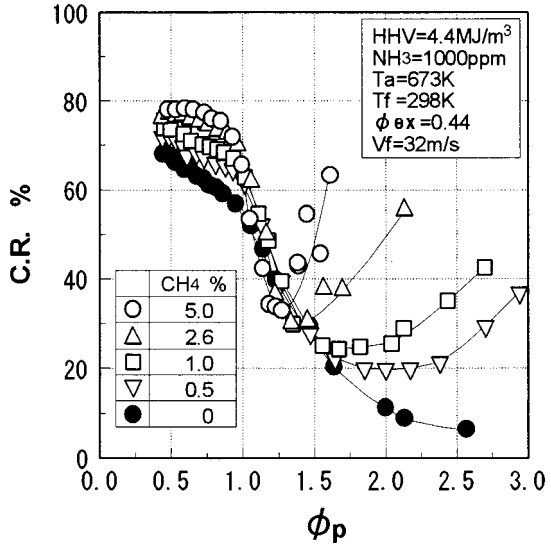


Fig. 14 Effect of CH_4 concentration on the conversion rate of NH_3 in the fuel to NO_x in two-staged combustion defining by the experiments

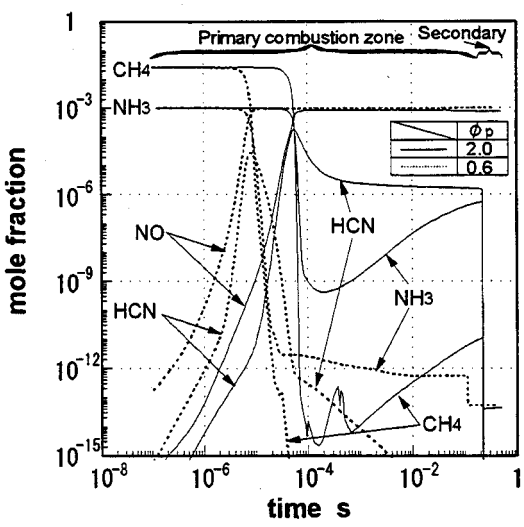


Fig. 15 Mole fraction-time products, comparing $\phi p = 0.6$, 2.0 in two-stage combustion as defined by the calculation

tained in the fuel at a concentration of 2.6 percent, no effect of the CO/H_2 molar ratio was observed on the NH_3 to NO_x conversion rate, when the primary equivalence ratio, ϕp, was set between 0.44 and 5.0.

When the fuel did not include CH_4, the effect of the CO/H_2 molar ratio was significant. In the same way with non-two-stage combustion described in Fig. 6, this is due to CO oxidizing more slowly than H_2, and so the higher the CO/H_2 molar ratio, the slower the O_2 consumption rate. Consequently, NH_3 is oxidized in the area with higher concentration of O_2, and as a result the NO_x production rate is increased. However, it is different from non-two-stage combustion that the conversion rate of NH_3 into NO_x is affected by the CO/H_2 molar ratio, even when the primary equivalence ratio, ϕp, increases more than 1. This is because the secondary air flows backward into the flame and decreases local equivalence ratio around the flame. On the other hand, the CO/H_2 molar ratio had no effect when the fuel contained CH_4, because CH_i produced by decomposition of CH_4 affects the decomposition of NH_3, rapidly producing in the flame zone HCN and NH_i, etc., which quickly oxidizes to NO_x in the secondary combustion zone.

(3) Effect of the Concentration of CH_4 in the Fuel. Figure 14 shows the correlation between the primary equivalence ratio in two-stage combustion, ϕp, and the conversion rate of NH_3 in the fuel to NO_x, with the concentration of CH_4 as a parameter at a fuel calorific value of 4.4 MJ/m^3, a NH_3 concentration of 1000 ppm in the fuel, and a CO/H_2 molar ratio set to 2.33. In the range of the examination of Fig. 14, the concentration of thermal-NO_x measured reached a maximum of 10 ppm, probably because the maximum theoretical adiabatic flame temperature of gasified coal fuel with a fuel calorific value of 4.4 MJ/m^3 is nearly 400 K lower than that of CH_4. When the primary equivalence ratio, ϕp, is less than 1, the NH_3 to NO_x conversion rate increases in direct proportion to the concentration of CH_4 in the fuel: and when ϕp is more than 1, under fuel-rich conditions in the primary combustion zone and CH_4-free fuel, the conversion rate decreases as the primary equivalence ratio increases. Fuel containing CH_4 has an appropriate primary equivalence ratio, (ϕp^*), at which the conversion rate drops to a minimum. ϕp^* is varied by adjusting the concentration of CH_4 (that is, ϕp^* increases as the concentration of CH_4 decreases). From these results, we conclude that when using two-stage combustion to suppress fuel-NO_x when burning gasified coal fuel containing CH_4, the primary equivalence ratio, ϕp^*, which minimizes the conversion rate of NH_3 in the fuel to NO_x, must be determined depending on the CH_4 concentration in the fuel.

The production characteristics of fuel-NO_x caused by NH_3 in gasified coal fuel were studied using reaction kinetics of elementary reactions. Figure 15 shows the behavior over time of chemical species NH_3, NO, HCN, and CH_4 in two-stage combustion, using primary equivalence ratio as a parameter. Under the condition of the equivalence ratio in the secondary combustion zone, ϕex was set to 0.45, the equivalent of the theoretical adiabatic flame temperature of 1573 K. We compared two cases in which the primary equivalence ratio, ϕp, was (a) 0.6 and (b) 2.0. For the analysis, the residence time in each combustion zone was calculated on the assumption that mixing gas temperature is regarded as the adiabatic flame temperature corresponding to each equivalence ratio. When the primary equivalence ratio was set at 0.6, the residence time in the primary combustion zone was about 0.11s, and the total residence time in the combustor was about 0.40s. The vertical axis represents the mole fraction of each species when the initial fraction of the fuel is set at 1 mole. The solid line in Fig. 15 shows the case where the primary equivalence ratio was set at 2.0, and the dotted line shows the case of the primary equivalence ratio of 0.6. Before the decomposition of NH_3, radical chemical species O, H, and OH are formed rapidly. The NH_3 in the fuel reacts with radicals, such as O, H, and OH, and decomposes to NH_i, then converted to NO. Following the decomposition of CH_4 and NH_3, the intermediate product HCN was formed in the primary combustion process in each case. When the primary equivalence ratio was 2.0 (in fuel-rich conditions), both the decomposition of CH_4 and NH_3, and the production of NO were delayed, and HCN was produced in high concentration and then oxidized slowly, while the HCN was rapidly decomposed to NO in the primary combustion zone in the case where the primary equivalence ratio was 0.6. Following the decomposition of HCN, NH_3 and CH_4 were reproduced, and more N_2 was produced in the primary combustion zone in fuel-rich conditions ($\phi p = 2.0$). Consequently a part of NH_3 was reduced to N_2 in the primary combustion zone when the primary equivalence ratio was set to 2.0 (in fuel-rich conditions). However, when the primary equivalence ratio rose higher, more NH_3 was reproduced by the decomposition of HCN in the primary combustion zone. This NH_3 was oxidized to NO_x in the secondary combustion zone and the conversion rate of NH_3 to NO_x increased. From above reason, there is

an appropriate primary equivalence ratio where the NO_x conversion rate reaches a minimum in fuel-rich conditions, as shown in Fig. 14.

It has been widely reported that HCN is not produced in non-hydrocarbon flame, even if the fuel contains nitrogenous compounds such as NH_3, but that if fuel includes hydrocarbons such as CH_4, then HCN is produced [52]. In the hydrocarbon flame, HCN is produced from nitrogen in the air by reaction (2).

$$CH_i(i=1,2)+N \rightleftharpoons HCN+NH_{i-1} \quad (2)$$

$$R\text{-}CH+NH_i \rightleftharpoons HCN+R\text{-}H_i \quad (3)$$

The HCN produced above is rapidly oxidized and produces prompt NO. If the fuel includes NH_3, then HCN is produced by reaction (3) in the fuel-rich region of the hydrocarbon flame and is oxidized to NO in the fuel-lean region. The nitrogen of NH_3 in the fuel has weaker bonding power than N_2. In the combustion process, NH_3 reacted with the OH, H, and O radicals and then easily decomposed into the intermediate NH_i by the following reactions.

$$NH_3+OH(O,H) \rightleftharpoons NH_2+H_2O(OH,H_2) \quad (4)$$

$$NH_i(i=1,2)+OH(H) \rightleftharpoons NH_{i-1}+H_2O(H_2) \quad (5)$$

When hydrocarbon is not contained in the fuel, NH_i is converted into N_2 by reacting with NO in the fuel-rich region.

$$NH_2+NO \rightleftharpoons N_2+H_2O \quad (6)$$

$$NH_2+NO \rightleftharpoons NNH+OH \quad (7)$$

$$NNH+NO \rightleftharpoons N_2+HNO \quad (8)$$

$$NH_2+H \rightleftharpoons NH+H_2 \quad (9)$$

$$NH+H \rightleftharpoons N+H_2 \quad (10)$$

$$N+NO \rightleftharpoons N_2+O \quad (11)$$

Also, part of the NH_2 produced as a product of the decomposition of NH_3 is oxidized into HNO by the O radical through reaction (12). Some HNO produced by reactions (8) and (12) is decomposed into N_2 and the remainder is oxidized into NO by reactions (13)–(16).

$$NH_2+O \rightleftharpoons HNO+H \quad (12)$$

$$HNO+OH \rightleftharpoons NO+H_2O \quad (13)$$

$$HNO+H \rightleftharpoons NO+H_2 \quad (14)$$

$$2HNO \rightleftharpoons N_2O+H_2O \quad (15)$$

$$N_2O+H \rightleftharpoons N_2+OH \quad (16)$$

Because of the high concentration of CO in the gasified coal fuel, NO is reduced to an N radical with oxidization of CO to CO_2 by reaction (17), and the N radical promotes the reduction of NO into N_2 by reaction (11).

$$NO+CO \rightleftharpoons N+CO_2 \quad (17)$$

$$N+NO \rightleftharpoons N_2+O \quad (11)$$

If fuel contains CH_4, HCN is produced by reaction (3) and the HCN is oxidized to NO in the fuel-lean zone. Therefore, increasing the concentration of CH_4 in gasified coal fuel increases the quantity of HCN, derived from NH_3, and of NO_x, produced from HCN in the secondary combustion zone under the condition of excess air. Because HCN is produced in the fuel-rich zone, as the equivalence ratio in the primary combustion zone increases, the production of HCN is increased and NO_x produced from HCN in the secondary combustion zone increases.

Figure 16 shows the relationship between the primary equivalence ratio, ϕp, for two-stage combustion, and both the NO_x conversion rate and the CO emission concentration, with the concentration of CH_4 as a parameter, and the equivalence ratio at the

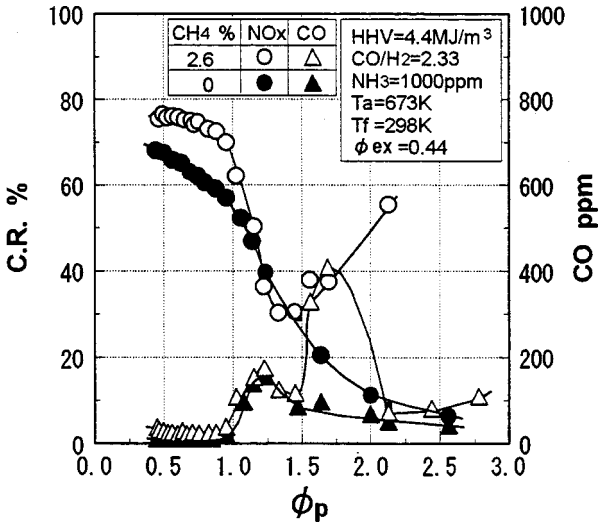


Fig. 16 Effect of CH_4 concentration on the conversion rate of NH_3 in the fuel to NO_x and CO emission characteristics in two-staged combustion defining by the experiments

combustor exit, ϕex, set to 0.44. As the concentration of CH_4 in the fuel was varied, the CO and H_2 content was adjusted to maintain the fuel calorific value at 4.4 MJ/m^3 and the CO/H_2 molar ratio at 2.33. When the fuel did not contain CH_4, the CO emission concentration peaked when the primary equivalence ratio, ϕp, was approximately 1.2. As ϕp increased further, the CO emission concentration tended to decline slowly. When the fuel contained 2.6 percent CH_4 and ϕp was 1.5 or less, the CO emission characteristics were almost the same as those for fuel which did not contain CH_4. However, for ϕp values between 1.5 and 2.1, the CO emission concentration increased rapidly to a maximum when ϕp was at approximately 1.7 and then decreased gradually, so that when ϕp reached 2.1 the CO emission concentration was almost the same as for fuel which did not contain CH_4. When ϕp was increased further, the CO emission concentration tended to increase gradually, because, if fuel includes CH_4, and ϕp is between 1.5 and 1.7, then HCN, produced in the fuel-rich primary combustion zone, is oxidized rapidly in the secondary combustion zone and produces NO and CO at the same time, leading to a rapid increase in the CO emission concentration. For ϕp values between 1.7 and 2.1, fuel which has not been combusted is burnt in the secondary combustion zone by the secondary air, a process which increases the production of NO_x, but at the same time tends to decrease the emission of CO. Values of ϕp above 2.1 destabilize combustion in the primary combustion zone, which suppresses decomposition of CH_4 in the primary combustion zone, leading to a gradual increase in the CO emission concentration at the exit of the combustor.

Figure 17 shows the correlation between ϕp and the CO emission concentration for two-stage combustion, with the concentration of CH_4 in the fuel as a parameter at a combustor exit equivalence ratio, ϕex, of 0.44. As the concentration of CH_4 in the fuel was varied, the CO and H_2 content was adjusted so as to maintain the fuel calorific value at 4.4 MJ/m^3 and the CO/H_2 molar ratio at 2.33. The CO emission concentration reached a maximum when ϕp exceeded 1.2 at all CH_4 concentration levels, increasing in direct proportion with the concentration of CH_4. The value of the primary equivalence ratio, ϕp, at which the CO emission concentration reached a maximum, was almost the same as the most appropriate primary equivalence ratio, ϕp^*, which minimizes the conversion rate of NH_3 to NO_x, as shown in Fig. 14, and there was a negative correlation between the concentration of CH_4 in the fuel and the appropriate primary equivalence ratio, ϕp^*. As the concentration of CH_4 in the fuel increases, the quantity of air for combustion at the same primary equivalence ratio increases

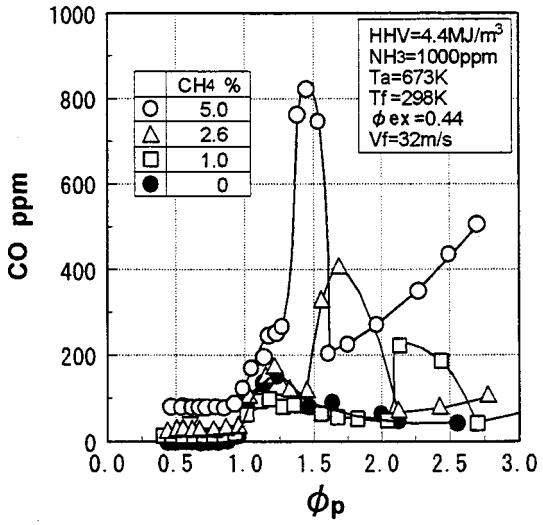


Fig. 17 Effect of CH_4 concentration on CO emission characteristics in two-staged combustion defining by the experiments

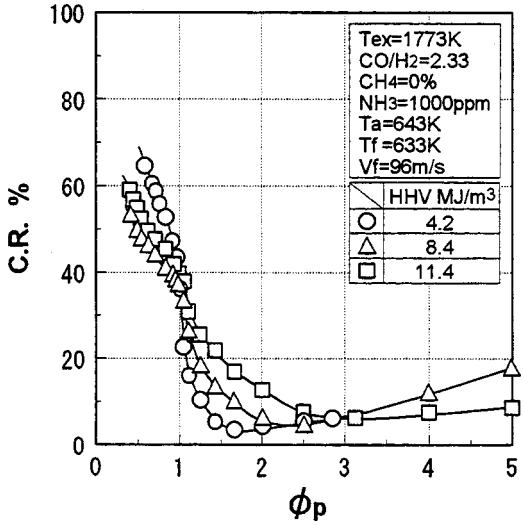


Fig. 19 Effect of fuel calorific value on the conversion rate of NH_3 in the fuel to NO_x in two-staged combustion defining by the experiments

and the flame temperature in the primary zone decreases. Therefore, increased CH_4 content in the fuel increased the rate of production of HCN at low primary equivalence ratios, and increased the quantity of NO_x produced in secondary combustion zone as well as primary zone.

(4) Effect of Fuel Calorific Value. Figure 18 shows the thermal-NO_x emission concentration when the equivalence ratio of the primary combustion zone, ϕp, is varied, for fuel containing no CH_4, with the fuel calorific value as a parameter. The equivalence ratio, ϕex, at the exit of the combustor was set for each fuel calorific value, so that the exhaust gas temperature of the combustor was set and maintained at 1773 K. At any fuel calorific value, the NO_x emission concentration indicates a maximum when the primary equivalence ratio, ϕp, was approximately 1. And increasing the fuel calorific value increased the NO_x emission concentration. These results lead the authors to conclude that it is possible to reduce thermal-NO_x emission at the exit of the combustor by using two-stage combustion in which ϕp is set to 1 or higher, with a fuel-rich condition in the primary combustion zone.

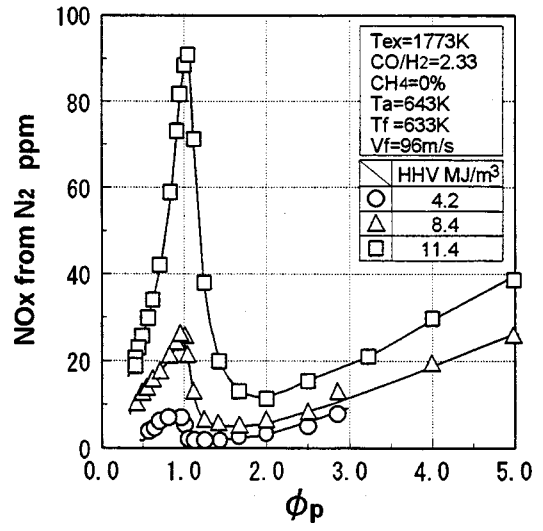


Fig. 18 Effect of fuel calorific value on NO_x emission characteristics from nitrogen fixation in two-staged combustion defining by the experiments

We studied the effectiveness of two-stage combustion in suppressing fuel-NO_x emission. Figure 19 shows the interrelationship between the primary equivalence ratio, ϕp, and the conversion rate of NH_3 in the fuel, containing no CH_4, to NO_x, when 1000 ppm of NH_3 was added to the fuel, in comparison with the conditions used for Fig. 18. Values of ϕp greater than 1, when the primary combustion zone is fuel-rich, lead to a rapid suppression of the conversion rate at all fuel calorific values. At values of ϕp between 1 and 2.5, increased fuel calorific values raised the conversion rate. This is because increasing the fuel calorific value increases the secondary air/fuel ratio, and the proportion of NO_x converted from NH_i and HCN, which are produced by decomposition of NH_3 in the primary combustion zone.

(5) Effect of Pressure in the Combustor. It is known that the effect of pressure in the combustor on the combustion property is significant. Based on the experimental data and numerical analysis, we designed a combustor for two-stage combustion using gasified coal, low calorific fuel, tested it and evaluated the combustion characteristics under pressurized conditions [25]. As a result of these tests, we found the following: (1) the influence of pressure on fuel-NO_x emission is very small in the case of gasified coal, low calorific fuel, (2) the influence of pressure on thermal-NO_x emission originating from the nitrogen fixation is smaller compared with the natural gas, and (3) the emission concentration of CO is deeply affected by pressure.

Figure 20 shows the correlation between the pressure in the combustor and both the conversion rates of NH_3 into NO_x and the thermal-NO_x emission characteristics under conditions where the combustor outlet gas temperature was 1773 K. The CH_4 concentration in the fuel was fixed at 1 percent, the NH_3 concentration at 1000 ppm, and the fuel calorific value set at 4.2 MJ/m^3. The conversion rate of NH_3 to NO_x is hardly affected by the pressure in the combustor, and slightly decreases with increases in the pressure. In the case where fuel contained 1.0 percent of CH_4, the conversion rate decreases to 40 percent or below when the pressure was 0.4 MPa and higher. One reason is that the concentration of each radical such as OH, O radical which promotes the NO_x formation decreases as the pressure increases.

On the other hand, the thermal-NO_x emission which is produced from the nitrogen fixation increased gradually from 4 ppm (with the correction of 16 percent O_2) under atmospheric pressure with the rise in pressure, reaching 6 ppm under operational conditions of 1.4 MPa. Thus, the pressure has a smaller effect on

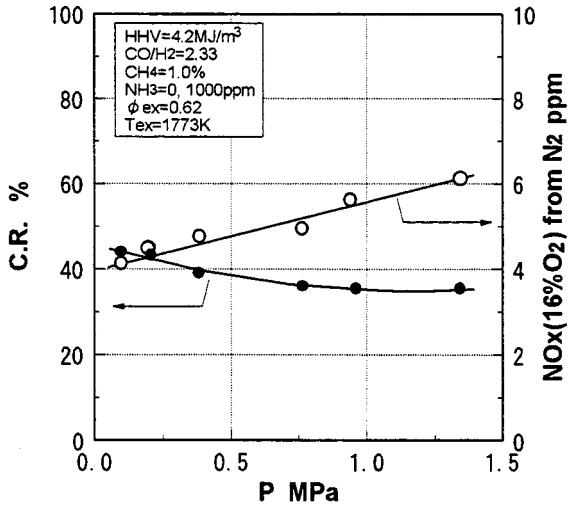


Fig. 20 Effect of pressure on the conversion rate of NH_3 in the fuel to NO_x and thermal-NO_x emission characteristics in a real combustor

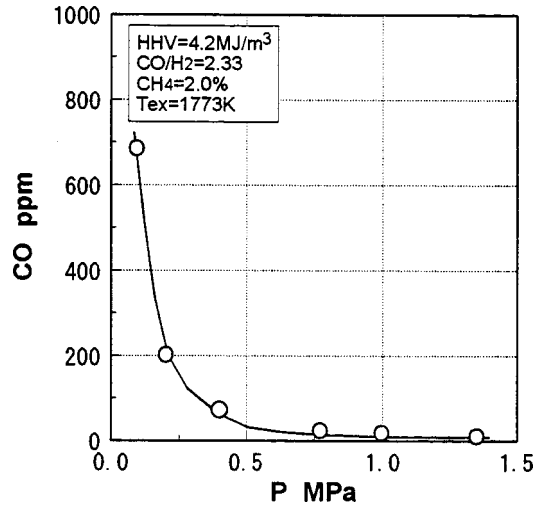


Fig. 21 Effect of pressure on CO emission characteristics in a real combustor

thermal-NO_x formation with the gasified coal fuel when compared to that of conventional fuels such as natural gas, and increases in accordance with approximately the power of 0.2 of the pressure. This is because the combustor is adapted to the two-stage combustion method. The flame temperature of gasified coal, low calorific fuel is lower than that of the hydrocarbon flame and hardly rises if the pressure increases.

Figure 21 shows the influence of pressure on CO emission characteristics, when the combustor outlet gas temperature was set at a constant value of 1773 K. CO emission was affected by the pressure and declined significantly as the pressure increased. The CO decreased to around 10 ppm at the operating pressure of the gas turbine, while the CO emitted about 700 ppm at the atmospheric pressure. That is to say that the CO emission decreased up to one-hundredth at the operating condition. This is because reaction rates increase and oxidation reactions of CO are enhanced when the pressure in the combustor increases.

Conclusions

We conducted our research using a small diffusion burner to clarify the effect of gaseous fuel properties on NO_x and CO emission characteristics and examined the results using numerical analysis based on reaction kinetics. Our principal conclusions are as follows:

1 NH_3 conversion to NO_x increases in direct proportion to the CH_4 concentration in gaseous fuel.

2 Gaseous fuels containing CH_4 have a specific equivalence ratio in the primary combustion zone for the minimum conversion of NH_3 to NO_x in two-stage combustion.

3 The specific equivalence ratio in the primary combustion zone increases with decreasing CH_4 concentration in gaseous fuel.

4 For fuels containing a small percent of CH_4, the effect of the CO/H_2 molar ratio in the fuel on the conversion rate of NH_3 to NO_x vanishes, while there is a positive correlation between the conversion rate and the CO/H_2 molar ratio in the case where fuel contains no CH_4.

5 As the pressure increases, the conversion rate of NH_3 to NO_x slightly decreases and the CO emission declines significantly.

Acknowledgment

The authors wish to express their appreciation to Mr. Koji Yamauchi of Shikoku Electric Power Corporation., Inc., who has contributed in this investigation.

Nomenclature

CH_i = one or more of CH_3, CH_2, CH, or C
CO/H_2 = molar ratio of carbonmonoxide and hydrogen in the fuel
C.R. = conversion rate from ammonia to NO_x (percent)
D = inner diameter of the combustor (mm)
HHV = higher heating value of the fuel (MJ/m^3)
K = chemical equilibrium constant
NH_i = one or more of NH_2, NH, or N
R = gas constant ($J/kg \cdot K$)
Tadia = adiabatic flame temperature (K)
Ta = air inlet temperature (K)
Tf = fuel inlet temperature (K)
Tex = combustor exit gas temperature (K)
Vf = velocity of fuel injection (m/s)
θa = angle of swirl vane
θ = angle of fuel injection nozzle
ϕ = equivalence ratio (inverse way of air/fuel ratio)
ϕex = equivalence ratio at combustor exit
ϕp = equivalence ratio in the primary combustion zone
ϕp^* = primary equivalence ratio which minimizes C.R.
ΔG° = Gibbs' standard energy of formation ($J/kg \cdot K$)
$NO_x(16\%\ O_2)$ = NO_x concentration corrected at 16 percent O_2 in the exhaust gas (ppm)
R- = hydrocarbon radical
S = swirl number $S = 2/3 \times (1 - B^3)/(1 - B^2) \times \tan \theta a$ B(boss ratio of swirl vane)$=0.66$

References

[1] Mizutani, H., 1995, ''Future View of Combined Cycle,'' Thermal and Nuclear Power Eng. Soc., **46**, No. 469, pp. 1198–1222.
[2] See, for example, Regenbogen, R. W., 1995, *Proc. the 12th Annual Int. Pittsburgh Coal Conf.*, p. 78.
[3] Ishizuka, S., and Tsuji, H., 1981, ''An Experimental Study of Effect of Inert Gases on Extinction of Laminar Diffusion Flames,'' *Proc. 18th Symp. (Int.) on Combust*. The Combust. Inst., Pittsburgh, PA, p. 695.
[4] Morgan, G. H., and Kane, W. R., 1962, *Proc. 9th Symp. (Int.) on Combust.*, The Combust. Inst., Pittsburgh, PA, p. 313.
[5] Coward, H. F., and Jones, G. E., 1971, ''Flammability Characteristics of Combustion Gases and Vapors,'' Bulletin 627, Bureau of Mines.
[6] Ishibasi, Y., and Kuroda, R., 1978, *Proc. the 6th Annual Conf. Gas Turbine Soc. Jpn.*, p. 7.
[7] Folsom, B. A., Courtney, C. W., and Heap, M. P., 1980, ''The Effects of LBG

Composition and Combustor Characteristics on Fuel NO_x Formation,'' ASME J. Eng. Power, **102**, p. 459.
[8] Drake, M. C., Pitz, R. W., Correa, S. M., and Lapp, M., 1984, ''Nitric Oxide Formation from Thermal and Fuel-Bound Nitrogen Sources in a Turbulent Nonpremixed Syngas Flame,'' *Proc. 20th Symp. (Int.) Combust.*, The Combust. Inst., Pittsburgh, PA, pp. 1983–1990.
[9] Döbbeling, K., Eroglu, A., Winkler, D., Sattelmayer, T., and Keppel, W., 1996, ''Low NO_x Premixed Combustion of MBtu Fuels in a Research Burner,'' ASME Paper 96-GT-126.
[10] Pillsbury, P. W., Cleary, E. N. G., Singh, P. P., and Chamberlin, R. M., 1976, ''Emission Results form Coal Gas Burning in Gas Turbine Combustors,'' ASME J. Eng. Power, **98**, p. 88.
[11] Clark, W. D., Folsom, B. A., Seeker, W. R., and Courtney, C. W., 1982, ''Bench Scale Testing of Low-NO_x LBG Combustors,'' ASME J. Eng. Power, **104**, p. 120.
[12] Battista, R. A., and Farrell, R. A., 1979, ''Development of an Industrial Gas Turbine Combustor Burning a Variety of Coal-Derived Low Btu Fuels and Distillate,'' ASME Paper 79-GT-172.
[13] Beebe, K. W., et al., 1982. DOE/NASA/13111-11, NASA TM, 82985.
[14] Döbbeling, K., Knöpfel, H. P., Polifke, W., Winkler, D., Steinbach, C., and Sattelmayer, T., 1994, ''Low NO_x Premixed Combustion of MBtu Fuels Using the ABB Double Cone Burner (EV Burner),'' ASME Paper 94-GT-394.
[15] Cook, C. S., Corman, J. C., and Todd, D. M., 1994, ''System Evaluation and LBtu Fuel Combustion Studies for IGCC Power Generation,'' ASME Paper 94-GT-366.
[16] Sato, M., Nakata, T., and Yamauchi, K., 1990, ''NO_x Emission Characteristics of Coal-Derived Low Btu Gas Fuel,'' J. Fuel Soc. Jpn., **69**, No. 10, pp. 952–959.
[17] Nakata, T., and Sato, M., 1991, ''Reaction Analysis of Coal Gaseous Fuel in a Gas Turbine Combustor,'' J. Jpn. Inst. Energy, **71**, No. 1, pp. 84–41.
[18] Yamauchi, K., Sato, M., and Nakata, T., 1991, ''The Effect of CH_4 Contained in Coal Gas Fuel on NO_x Formation,'' Trans. Jpn. Soc. Mech. Eng., Ser. B, **57**, No. 535, pp. 811–818.
[19] Nakata, T., Sato, M., and Hasegawa, T., 1998, ''Reaction of Fuel NO_x Formation for Gas Turbine Conditions,'' Trans. ASME: J. Eng. Gas Turbines Power, **120**, No. 3, pp. 474–480.
[20] Sato, M., Nakata, T., Yoshine, T., and Yamada, M., 1990, ''Development of a 1300 °C-Class Gas Turbine Combustor Burning Coal-Derived Low-Btu Gaseous Fuels,'' Trans. Jpn. Soc. Mech. Eng., Ser. B, **56**, No. 530, pp. 3147–3154.
[21] Sato, M., Ninomiya, T., Abe, T., Yoshine, T., and Hasegawa, H., 1990, ''Development of a 1300 °C-Class Gas Turbine Combustor Burning Coal-Derived Low-Btu Gaseous Fuels,'' Trans. Jpn. Soc. Mech. Eng., Ser. B, **56**, No. 532, pp. 3903–3909.
[22] Sato, M., Nakata, T., Yoshine, T., and Yamada, M., 1991, ''Development of a 1300 °C-Class Gas Turbine Combustor Burning Coal-Derived Low-Btu Gaseous Fuels,'' Trans. Jpn. Soc. Mech. Eng., Ser. B, **57**, No. 535, pp. 803–810.
[23] Nakata, T., Sato, M., Ninomiya, T., Yoshine, T., and Yamada, M., 1992, ''Development of a 1300 °C-Class Gas Turbine Combustor Burning Coal-Derived Low-Btu Gaseous Fuels,'' Trans. Jpn. Soc. Mech. Eng., Ser. B, **58**, No. 553, pp. 2890–2897.
[24] Nakata, T., Sato, M., and Hasegawa, T., 1994, ''A Study on Low NO_x Combustion in LBG-Fueled 1500 °C-Class Gas Turbine,'' ASME Paper 94-GT-218.
[25] Hasegawa, T., Sato, M., and Ninomiya, T., 1997, ''Effect of Pressure on Emission Characteristics in LBG-Fueled 1500 °C-Class Gas Turbine,'' ASME Paper 97-GT-277.
[26] Hasegawa, T., Hisarnatsu, T., Katsuki, Y., Sato, M., Yamada, M., Onoda, A., and Utsunomiya, M., 1998, ''A Study of Low NO_x Combustion Characteristics in Medium-Btu Fueled 1300 °C-Class Gas Turbine Combustor,'' ASME Paper 98-GT-331.
[27] *Modern Power Systems Review*, 1993, ''Clean Coal 5 Eyes BGL Gasification at Camden,'' August 1993, pp. 21–24.
[28] Kalsall, G. J., Smith, M. A., and Cannon, M. F., 1994, ''Low Emissions Combustor Development for an Industrial Gas Turbine to Utilize LCV Fuel Gas,'' Trans. ASME J. Eng. Gas Turbines Power, **116**, p. 559.
[29] Ichikawa, K., and Araki, S., 1996, ''Test Results of the IGCC System by the 200T/D Nakoso Pilot Plant,'' *Proc. The 9th Thermal Engineering Symp.*, pp. 11–12.
[30] Bush, W. V., Baker, D. C., and Tijm, P. J. A., 1991, ''Shell Coal Gasification Plant(SCGP-1) Environmental Performance Results,'' EPRI Interim Report No. GS-7397, Project 2695-1.
[31] Ueda, T., Kida, E., Nakaya, Z., Shikata, T., Koyama, S., and Takagi, M., 1995, ''Design of the HYCOL Gasifier,'' *Proc. Int. Conference Power Engineering-'95*, pp. 242–247.
[32] *Modern Power Systems*, 1994, ''Biomass Looks Good for Gasification Process,'' April 1994, pp. 61–65.
[33] Consonni, S., Larson, E. D., and Berglin, N., 1997, ''Black Liquor-Gasifier/Gas Turbine Cogeneration,'' ASME Paper 97-GT-273.
[34] Ashizawa, M., Takahashi, T., Taki, M., Mori, K., Kanehira, S., and Takeno, K., 1996, ''A Study, on Orimulsion Gasification Technology,'' *Power-Gen '96 Int.*, **8**, pp. 235–243.
[35] Hasegawa, T., and Sato, M., 1997, ''Study on NO_x Formation Characteristics of Medium Btu Coal Gasified Fuel,'' Trans. Jpn. Soc. Mech. Eng., Ser. B, **63**, No. 613, pp. 3123–3130.
[36] Miller, J. A., and Bowman, C. T., 1989, ''Mechanism and Modeling of Nitrogen Chemistry in Combustion,'' Prog. Energy Combust. Sci., **15**, pp. 287–338.
[37] Hasegawa, T., and Sato, M., 1998, ''Study of Ammonia Removal from Coal-Gasified Fuel,'' Combust. Flame, **114**, pp. 246–258.
[38] Chase, Jr., M. W., Davies, C. A., Downey, Jr., J. R., Frurip, D. J., McDonald, R. A., and Syverud, A. N., 1985, ''JANAF Thermodynamical tables 3rd Edition.,'' J. Phys. Chem. Ref. Data, **14**.
[39] Kee, R. J., Rupley, F. M., and Miller, J. A., 1990, Sandia Report, SAND 87-8215B.
[40] Hindmarsh, A. C., 1974, Lawrence Livermore Laboratory, Univ. California, Report No. UCID-30001, Rev. 3.
[41] Pratt, D. T., Bowman, B. R., and Crowe, C. T., 1971, *AIAA paper* 71-713.
[42] Martin, F. J., and Dederick, P. K., 1977, ''NO_x From Fuel Nitrogen in Two-Stage Combustion,'' *Proc. 16th Symp. (Int.) on Comb.*, The Comb. Institute., Pittsburg, PA, pp. 191–198.
[43] Yamagishi, K., Nozawa, M., Yoshie, T., Tokumoto, T., and Kanegawa, Y., 1974, ''A Study of NO_x Emission Characteristics in Two-Stage Combustion,'' *Proc. 15th Symp. (Int.) on Comb.*, The Comb. Institute., Pittsburgh, PA, pp. 1157–1165
[44] *JSME Data Book: Formation Mechanisms and Controls of Pollutants in Combustion System*, 1980, **47**, JSME, Japan.
[45] Sarofim, A. F., Williams, G. C., Modell, M., and Slater, S. M., 1975, ''Conversion of Fuel Nitrogen to Nitric Oxide in Premixed and Diffusion Flames,'' AIchE Symp. Ser., **71**, No. 148, pp. 51–61.
[46] Kato, K., Fujii, K., Kurata, T., and Mori, K., 1976, Trans. Jpn. Soc. Mech. Eng., **42**, No. 354, pp. 582–589.
[47] Fenimore, C. P., 1972, ''Formation of Nitric Oxide From Fuel Nitrogen in Ethylene Flames,'' Combust. Flame, **19**, No. 2, pp. 289–296.
[48] Takagi, T., Ogasawara, M., Daizo, M., and Tatsumi, T., 1977, Trans. Jpn. Soc. Mech. Eng., **43**, No. 368, p. 1426–1439.
[49] Fenimore, C. P., 1976, ''Effects of Diluents and Mixing on Nitric Oxide from Fuel-Nitrogen Species in Diffusion Flames,'' *Proc. 16th Symp. (Int.) on Comb.*, The Comb. Institute., Pittsburgh, PA, pp. 1065–1071.
[50] Heap, M. P., Tyson, T. J., Cichanowicz, E., Gershman, R., Kau, C. J., Martin, G. B., and Lanier, W. S., 1976, ''Environmental Aspects of Low BTU Gas Combustion,'' *Proc. 16th Symp. (Int.) on Comb.*, The Comb. Institute., Pittsburgh, PA, pp. 535–545.
[51] Takagi, T., Tatsumi, T., Ogasawara, M., and Tatsumi, K., 1978, Trans. Jpn. Soc. Mech. Eng., **44**, No. 388, pp. 4282–4291.
[52] Kato, K., Fujii, K., Kurata, T., and Mori, K., 1977, Trans. Jpn. Soc. Mech. Eng., **43**, No. 365, pp. 280–292.
[53] Takagi, T., Ogasawara, M., Daizo, M., and Tatsumi, T., 1976, ''NO_x Formation From Nitrogen in Fuel and Air During Turbulent Diffusion Combustion,'' *Proc. 16th Symp. (Int.) Combust.*, The Comb. Institute, Pittsburgh, PA, pp. 181–188.

Y. M. Han

W. S. Seol

D. S. Lee

Aero-propulsion Department,
Korea Aerospace Research Institute,
Taejon, Korea

V. I. Yagodkin
Central Institute of Aviation Motors,
Moscow, Russia

I. S. Jeung
Department of Aerospace Engineering,
Seoul National University,
Seoul, Korea

Effects of Fuel Nozzle Displacement on Pre-Filming Airblast Atomization

In gas turbine combustors, optimum arrangement between a fuel nozzle and a swirler/prefilmer module must be sought to achieve satisfactory ignition and stability characteristics in addition to reduced level of emissions. However, due to thermal expansion of the combustor or misalignment of the fuel nozzle, the location of a fuel nozzle may vary. Displacement of a fuel nozzle may change the amount of fuel injected to the pre-filming device (usually the inner swirler wall) and the location of attachment, which in turn affects the thickness of pre-filming liquid sheet on the wall. As a result, the spray structure formed by pre-filming airblast atomization may be significantly changed. An experimental investigation is carried out to study the effects of fuel nozzle displacement on the structure of a spray formed by a dual orifice pressure atomizer and a counter-rotating dual swirler. The inner wall of the swirler is designed to be used as a pre-filming device. The behavior of droplets, the flow characteristics of the swirling air flow, and the interaction between droplets and the flow are studied. Optical diagnostic methods including a flow visualization and an Adaptive Phase/Doppler technique are used. Distributions of droplet size, number density, and liquid phase volume flux are presented for various fuel nozzle displacements, in addition to gas phase velocity. [DOI: 10.1115/1.1335480]

Introduction

To guide the design and development of a gas turbine combustor, it is essential to acquire details of the fuel injector performance. Most of the systems now in service are of the prefilming type because of their potential for achieving significant reductions in soot formation and exhaust smoke in gas turbine engines of high pressure ratio.

The prefimling concept for airblast atomization evolved from the studies of Lefebvre and Miller [1]. They showed that the minimum drop sizes were obtained by using atomizers designed to provide maximum physical contact between the air and the liquid, and the increase in the thickness of liquid sheet would tend to produce drops of larger size. Rizkalla and Lefebvre [2] investigated the prefilming airblast atomization in detail. They found that the key factors governing drop size were the surface tension and the air momentum (density, velocity) for low viscosity liquids and the viscosity for relatively high viscosity liquids. Rizk et al. [3] observed that the thinner initial liquid sheet produced finer drops. Many studies were performed to find main factors in prefilming atomization such as the ratio of air and liquid density, the diameter of prefilming lip, the atomizer size, etc. [4–6]. Sattelmayer and Wittig [7] showed that in the prefilming airblast atomizer the wavy liquid films influence the velocity of gas phase. Wittig et al. [8] found that the pre-filming flow was almost laminar, and the turbulent air or the waviness of the film didn't induce significant turbulence into the film. Breña de la Rosa [9] inspected the effect of air swirl on the structure of a liquid spray. Wang et al. [10] conducted studies on the spray characteristics in the CFM56 swirl cup assembly which features two co-axial counter-rotating swirl air streams, a simplex atomizer, and the prefilming wall. An experimental study was also conducted on the influence of flare sleeve geometry on the continuous phase flow field and the droplet dispersion pattern [11]. An attempt to formulate a common approach that could provide the spray parameter of airblast atomizers was performed by using a semi-analytical calculation procedure [12]. In review of previous works, the important parameters influencing atomization in prefilming airblast atomizer are the initial thickness of liquid sheet and the air momentum. In addition, the interaction between swirling air and droplets is very complex.

When fuel injectors are separate from the frontal device, the nozzle shift with respect to the design position might occur due to thermal expansion of the combustor or misalignment of the fuel nozzle. It was observed that radial shift of a fuel nozzle deteriorates the symmetry of spray [13]. In the present work, effects of axial displacement of a fuel nozzle on the spray characteristics are studied. Displacement of a fuel nozzle may change the amount of fuel injected to the prefilming wall and the location of attachment due to changes in the interaction between the spray liquid and the swirling air. The above changes may affect the thickness of liquid film formed on the wall and the spray structure formed by prefilming airblast atomization.

An experimental investigation is carried out to measure the gas phase flow field without fuel injection as well as the spray characteristics such as drop size, velocity, number density, and volume flux. The main, annular nozzle of a dual orifice atomizer and a counter-rotating dual swirler are used. The convergent wall of the inner swirler is designed to be used as a prefilming device. This wall stabilizes swirling air flow and prevents flow separation within the passage, which encourage the liquid film to adhere to the wall. To asses the global structure of the spray, a visualization is performed using a laser sheet illumination technique. The adaptive phase/doppler velocimetry is used to quantify the gas phase flow field and the spray characteristics.

Experimental Apparatus and Procedure

Experimental Apparatus. The experimental setup employed in the present study is shown in Fig. 1. This includes a test chamber, a pressurized fuel supply system, an air supply system, an injector assembly, a spray catcher, a liquid collection and suction system, a visualization system and an adaptive phase/doppler velocimetry system. Air from the compressor passes through a

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Stockholm, Sweden, Jun. 2–5, 1998; ASME Paper 98-GT-360. Manuscript received by IGTI Mar. 19, 1998; final revision received by the ASME Headquarters Oct. 20, 1999. Associate Technical Editor: R. Kielb.

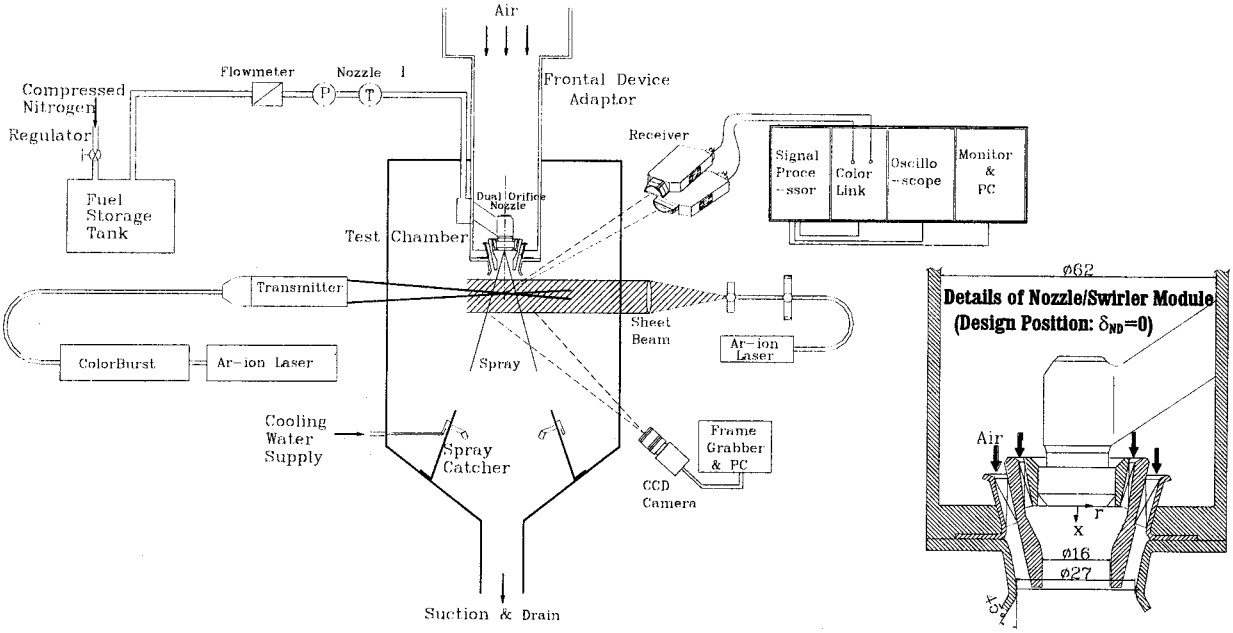


Fig. 1 Experimental setup and details of nozzle/swirler module

regulator/filter assembly, a flowmeter, an electric heater, a control valve, and enters to the frontal device adaptor. Temperature and pressure of the air are measured at this adaptor.

The counter rotating swirler has twelve flat vanes in each of the inner and the outer swirlers. The vane angle of the inner swirler is 40 deg, opposite to the fuel swirl direction, and that of the outer swirler is 45 deg, opposite to the inner swirl direction. The effective area of air passage is 0.875 cm^2 for the inner swirler and 1.17 cm^2 for the outer swirler. The fuel nozzle employed in this study is the main, annular nozzle of a dual orifice atomizer, for which $FN = 14.7 \times 10^{-7}$ m^2. The main fuel nozzle has six 60 deg helical slots in its swirl chamber and an exit orifice of 2.4 mm around the pilot orifice of 0.8 mm which is not working in the present study [14].

The fuel supply system is designed to accommodate a maximum fuel flow rate of 1.8 l/min for the main nozzle. The maximum pressure for the injection system is 1.5 Mpa. Continuous and steady fuel injection is obtained from the tank pressurized by nitrogen.

To assess global structures of the spray, visualizations with laser sheet illumination are performed. The Ar-ion laser light is transmitted via a single-mode fiber optic cable to the collimating micro-objective lens on the optical bench. After being collimated, the laser sheet is produced with a cylindrical lens and collimated by a convex lens. The beam thickness is about 0.3 mm at beam waist and the height is about 90 mm. Spray images are captured with a color CCD camera (Pulnix TMC-74) with 768(H)×493(V) pixels that is oriented perpendicular to the laser sheet. The captured image is processed with a color frame grabber (Data Translation DT2871). The CCD camera is equipped with a magnifying objective and the exposure time is 1/60 s.

The adaptive phase/doppler velocimetry (APV) system manufactured by TSI, Inc. is used to obtain the size and velocity information of droplets. A two-component system using green (514.5 nm) and blue (488 nm) beams from an Ar-ion laser operating at 0.8 W power output is used. The APV system is composed of a transmitting unit, a receiving system, a signal processing and data analysis unit, etc.

The transmitting unit is composed of a laser source (Coherent Innova 70-5), a color burst (TSI model 9201), and a fiber-optic probe (model 9832, 83 mm diameter). The color burst uses a Bragg cell operating at 40 MHz to shift laser beam frequency. The polarization of beams is 90 deg. The receiving system is composed of a dual adaptive receiver and a color-link. The signal processor (IFA755) is a digital burst autocorrelation signal processor with maximum frequency of 90 MHz. The details of APV setup are shown in Table 1.

In the present experiments, the receiver optics are located 30 deg off-axis from the forward scatter direction, the included angle between the receivers is 0 deg, and the clear aperture height is 35 mm (for droplet size range up to 121 μm) or 15 mm (for droplet size range up to 223 μm). The photomultiplier voltages are set to provide the maximum sensitivity for a given range of droplet sizes. The frequency shifts are set from 5 MHz to 10 MHz in order to prevent high frequency shift bias. The signal processor is set to have a single measurement per burst, the coincidence time of 100 ms, the number of minimum cycles per burst of 8, and the minimum threshold of 110 mV. The APV simulation software, SIMAP (for Mie scattering calculations) supplied by TSI, Inc. is used to find the phase-diameter relation for a given optic geometry. The accuracy of the APV system is checked against a monodisperse aerosol generator (Aerometrics DPG100). The error is found to be less than 3 percent for the droplets of about 90 μm. The transmitting probe and the receiving optics are mounted on a common rail which is fixed to a three-axis traverse system with 0.01 mm resolution. To measure the gas phase flow field without

Table 1 Details of APV setup

Transmitter focal length(mm)		512
Fringe number		24
Measurement Volume Diameter(μm)	514.5nm	128
	488nm	121
Fringe spacing(μm)	514.5nm	5.27
	488nm	5.00
Receiver focal length - front(mm)		500
Receiver focal length - back(mm)		300
Clear aperture height(mm)		15, 35
Spacing between receiving unit(degrees)		0

fuel injection, 4–6 μm propylene glycol particles produced by the six-jet atomizer (TSI, model 9306) are introduced upstream of the frontal device adaptor.

Experimental Procedure. The air pressure drop across the frontal device is 4.5 kPa and the ambient pressure is about 1 bar (pressure drop of about 4.5 percent). The corresponding air flow rate through the swirlers is about 24.4 g/sec. All tests are performed with kerosene at room conditions (temperatures of the air and the kerosene are about 22°C and 20°C, respectively). The pressure drop across the main fuel nozzle is 0.6 bar, the spray cone angle is about 80 deg, and the fuel flow rate is 10.2 g/sec. The corresponding air fuel ratio (AFR) is about 2.4, which is small compared to the normal combustor operating condition. This rather unrealistic AFR is due to limitations imposed by the atmospheric test conditions. The fuel injection pressure (i.e., fuel flow rate) may not be further reduced since the spray cone angle is not maintained.

The geometry permits to shift the fuel nozzle in the range from +2.5 mm to −3 mm with respect to the design position. To have maximum effects, nozzle displacements of +2.5 mm, 0.0 mm, and −3.0 mm, corresponding to about 15~18 percent of the nozzle outer diameter, are investigated. The positive sign (+) indicates downward displacement.

The height of visualized region is about 90 mm from the end of the frontal device. The APV measurements are conducted at four axial locations; 35, 45, 65, and 95 mm from the nozzle exit (design position). In gas phase measurements, several measurement locations are added to find the more detailed flow field such as recirculation zone, swirl components, etc.

Results and Discussions

Visualization Results. The spray liquid sheet injected from the fuel nozzle firstly meets the swirling air flow through the inner swirler. The swirling air will make the liquid sheet unstable and disintegrate the sheet into ligaments and droplets. The smaller drops follow the flow, but the larger drops and fragments of the sheet attach to the prefilming wall and then form liquid film on the wall. This liquid film flows interfacing the inner swirling air. The liquid film is subjected to the counter-swirling air on both sides at the prefilmer lip. The liquid sheet-air interaction produces waves that become unstable and disintegrate into fragments. These fragments become ligaments and, in turn, break down into drops. The attaching point of ligaments or larger drops and the thickness of the liquid film vary as the fuel nozzle shifts. This change may affect the spray characteristics significantly.

Images of the spray obtained at various nozzle displacements are shown in Fig. 2. The spray image (e.g., Fig. 2(*b*)) reveals two bright regions; one near the center and one near the spray edge. Near the center, the scattered light is mostly from relatively large droplets from the broken liquid sheet by the inner swirling flow. Near the edge, the scattered light is mostly from relatively fine droplets produced by airblast atomization on the prefilmer lip. As the fuel nozzle shifts downwards (Fig. 2(*a*)), the spray is less dispersive and the amount of fuel subjected to the prefilming atomization is decreased (the scattered light intensity is smaller near the edge while that near the center is larger). As the fuel nozzle shifts upwards (Fig. 2(*c*)), the amount of fuel subjected to the prefilming atomization is increased (the scattered light intensity near the edge is larger). The change in the spray structure is more significant when the fuel nozzle shifts downwards (positive displacement) than when the nozzle shifts upwards (negative displacements).

Gas Phase Flow Field. Figure 3 presents the gas phase flow field obtained without fuel injection. The figure shows a strong recirculation zone near the exit of the frontal device. The recirculation zone has a central toroidal structure. The width of the reverse flow zone is close to the outer swirler diameter (36 mm). The slight asymmetry is caused by the fuel feed arm of the fuel

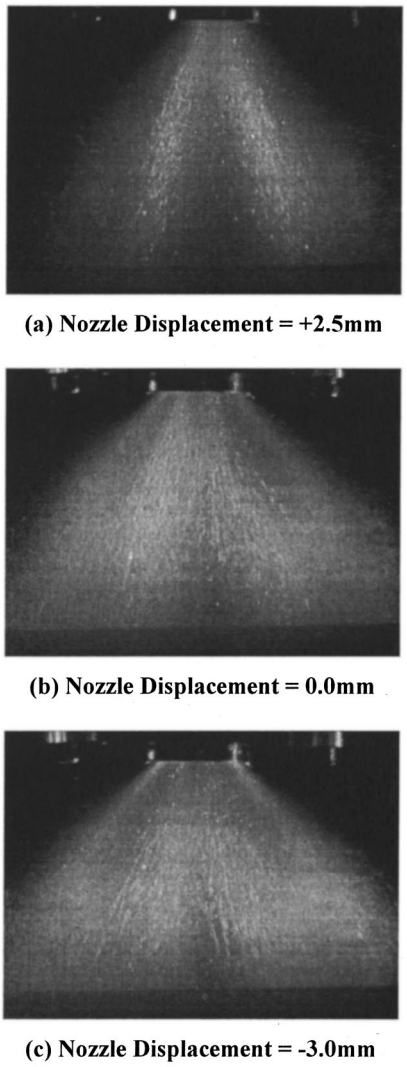


Fig. 2 Visualization of the effects of nozzle displacement on sprays (ΔP_a=4.5 kPa, ΔP_f=0.6 bar)

nozzle. In Fig. 4, radial profiles of swirl velocities are shown. Positive values of swirl velocity represent counterclockwise rotation as imparted by the outer swirler. Near the exit of flare (x = 23, 25 mm), the swirl velocity distribution has two peaks; one inside the reverse flow zone and one outside. The outer peak is

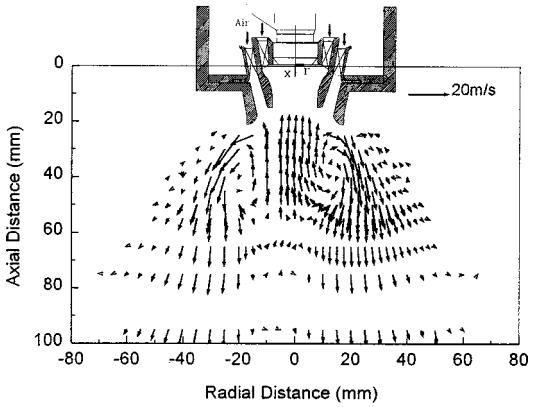


Fig. 3 Axial-radial velocity vector of gas phase without fuel injection (ΔP_a=4.5 kPa, δ_{ND}=0.0 mm)

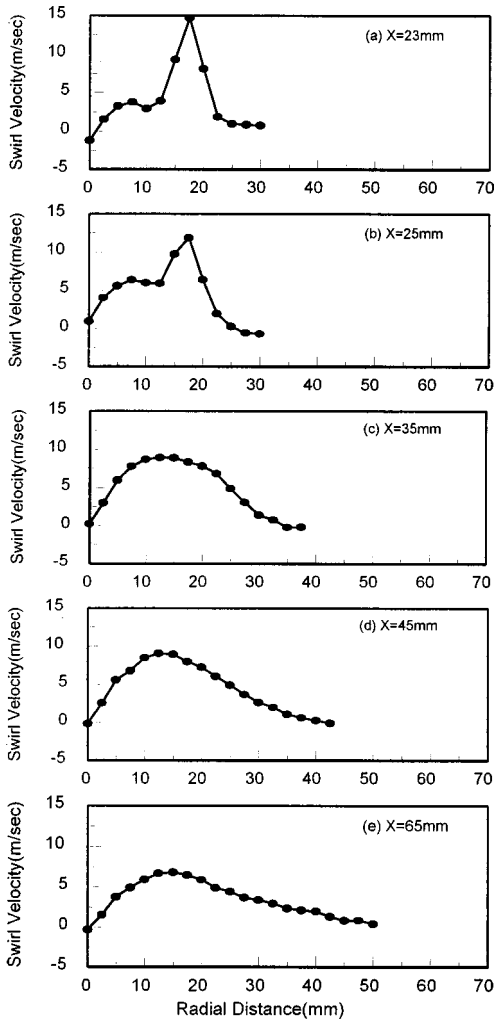


Fig. 4 Swirl velocity of gas phase without fuel injection ($\Delta P_a = 4.5$ kPa, $\delta_{ND} = 0.0$ bar)

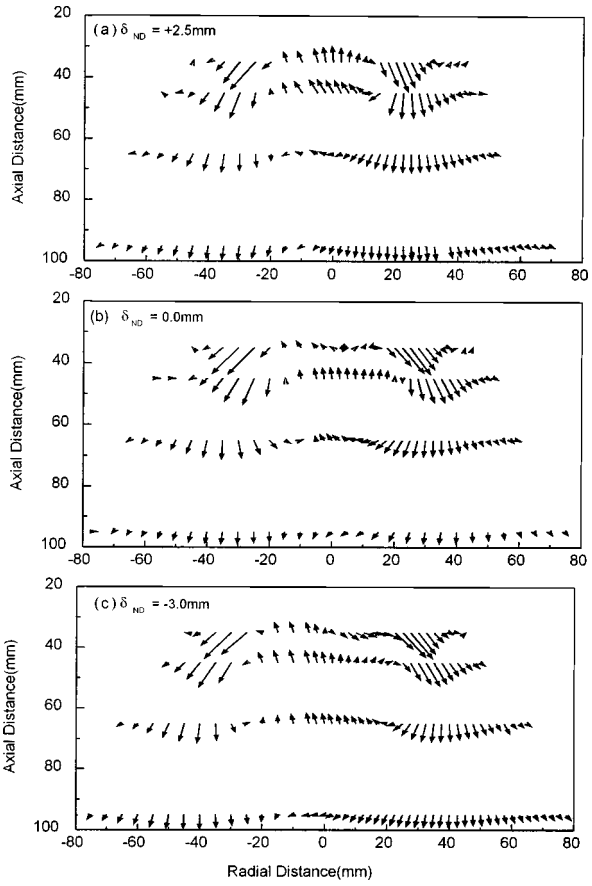


Fig. 5 Effect of nozzle displacement on velocity fields of 10–20 μm driplets ($\Delta P_a = 4.5$ kPa, $\Delta P_f = 0.6$ bar)

produced by the outer swirling flow, and the inner peak is due to reverse flow which has the same, counterclockwise rotation as the outer swirling flow. The inner swirling flow, which rotate clockwise, appears to be decayed due to the interaction between the inner and the outer swirl flow. The effect of the inner swirling flow on flow field is mainly to increase the shear stress in mixing layers due to opposite flow direction with respect to the outer swirling flow [15,16]. Dissipation of the swirl in the shear layer causes a strong pressure gradient in the axial direction, which in turn causes a reverse flow. Dual peaks are not apparent from $x = 35$ mm, and the swirl intensity decreases monotonously and disperses radially as the air flows downstream. Similar results have been reported by Wang et al. [17] and McDonell et al. [18,19] for CFM 56 swirl cup assembly.

Velocity Fields of Fuel Droplets. The velocity fields of 10–20 μm fuel droplets for the three displacements are shown in Fig. 5. The global flow features of 10–20 μm droplets for $\delta_{ND} = 0.0$ mm are similar to those of the continuous phase shown in Fig. 3. However, the fuel injection makes the main air flow more divergent, possibly due to momentum decrease in the inner swirling flow. When the injected fuel meets the inner swirling flow, a part of angular momentum of the inner swirling flow is transferred to the fuel and lost as the fuel attaches to the prefilming wall. As a result, the resultant angular momentum at the flare exit becomes larger, and the main air flow becomes more divergent. Some droplets having momentum transferred by the inner swirl flow enter into the recirculation zone.

Another significant change due to fuel injection is related to the location of recirculation zone. With fuel injection, the recirculation zone forms in further downstream zone, possibly because the major decay of resultant swirl, which causes the axial pressure gradient driving reverse flow, occurs further downstream due to the momentum decrease in the inner swirling flow.

The reverse velocity of the fuel droplets at $x = 45$ mm (about 4 m/s) is much smaller than the corresponding reverse velocity of the gas phase (about 10 m/s). The reason may be as follows. Firstly, the axial pressure gradient caused by decay of swirl diminishes as the fuel droplets are added. Secondly, the momentum exchange between fuel droplets and the reverse flow reduces the reverse velocity. Thirdly, fuel droplets may not follow the air flow exactly.

Swirl velocity profiles of 10–20 μm droplets for $\delta_{ND} = 0.0$ mm are shown in Fig. 6. At $x = 25$ mm, the negative swirl velocity near the center indicates that droplets having the same clockwise rotation as the inner swirling flow exist. But, the clockwise swirl disappears and the counter-clockwise swirl due to the recirculating flow appears near the center at $x = 35$ mm. At further downstream, the droplets have the counter-clockwise swirl only as the gas phase flow field does (cf. Fig. 4). However, the magnitude of swirl velocity of 10–20 μm droplets is smaller than that of gas phase.

As the nozzle displaces, the recirculation zone is significantly changed. This is because the nozzle displacement alters the resultant swirl and the axial pressure gradient at the exit of prefilmer lip. As the nozzle shifts downwards ($\delta_{ND} = +2.5$ mm), the momentum decrease of the inner swirl due to fuel injection is smaller

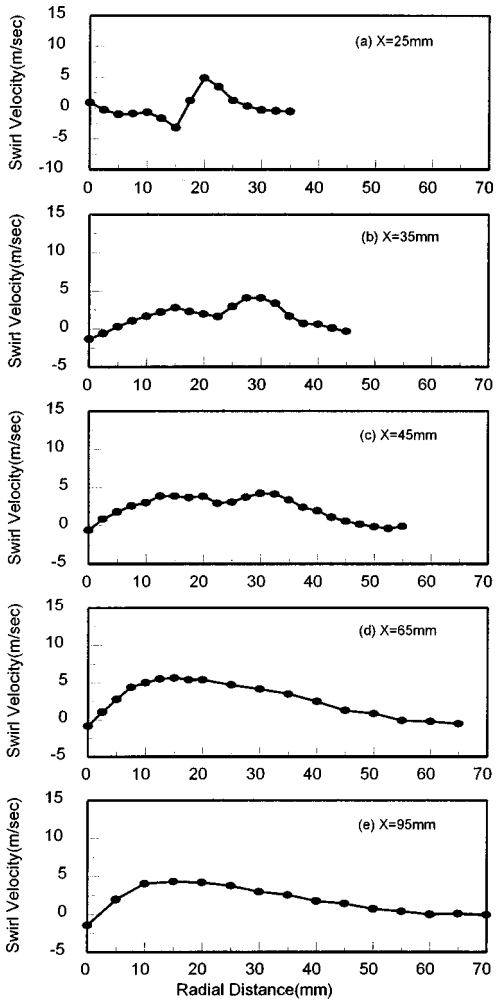


Fig. 6 Swirl velocity of 10–20 μm droplets for δ_{ND}=0.0 mm (ΔP_a=4.5 kPa, ΔP_f=0.6 bar)

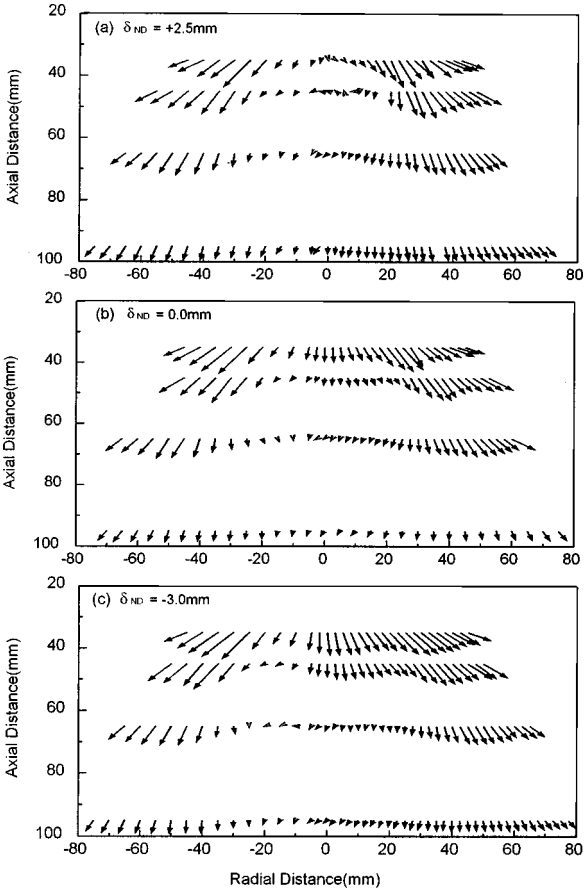


Fig. 7 Effects of nozzle displacement on velocity fields of 90–100 μm droplets (ΔP_a=4.5 kPa, ΔP_f=0.6 bar)

than that for $\delta_{ND}=0.0$ mm. This results in a smaller resultant swirl and a larger axial pressure gradient near the prefilmer exit, and the flow field becomes closer to that without fuel injection. The reverse axial velocity at $x=35$, 45 mm is larger than that for $\delta_{ND}=0.0$ mm but smaller than that without fuel injection. For upward nozzle displacement ($\delta_{ND}=-3.0$ mm), the momentum decrease of the inner swirl due to fuel injection is larger than that for $\delta_{ND}=0.0$ mm. This results in a larger resultant swirl and a smaller axial pressure gradient near the prefilmer exit, and the recirculation zone becomes wider and shifts downwards.

Figure 7 presents the velocity field of 90–100 μm droplets. Large droplets do not show any significant reverse velocity. At $x=35$, 45 mm, for $\delta_{ND}=0.0$ mm, 90–100 μm droplets have positive axial velocity while 10–20 μm droplets have negative axial velocity. The magnitude of outward radial velocity is larger than that for 10–20 μm droplets, especially near the edge of spray. This is believed to be due to the inertia of larger droplets. However, 90–100 μm droplets near the center are decelerated by reverse flow and lose the initial momentum quickly.

As the nozzle displaces, the velocity field of 90–100 μm droplets is significantly changed. The change corresponds to that for 10–20 μm in a sense described in the following. As described for 10–20 μm droplets, the downward nozzle displacement ($\delta_{ND}=+2.5$ mm) shifts the recirculation zone upwards, but the upward nozzle displacement ($\delta_{ND}=-3.0$ mm) shifts the recirculation zone downwards. Thus, for $\delta_{ND}=+2.5$ mm, deceleration of large droplets due to reverse flow occurs near the prefilmer exit, so that the velocity of 90–100 μm droplets near the center is close to zero even at $x\approx35$ mm. However, for $\delta_{ND}=-3.0$ mm, deceleration of large droplets occurs farther downstream, so that the velocity of 90–100 μm droplets near the center becomes close to zero at $x\approx65$ mm.

Figure 8 presents swirl velocity of 90–100 μm droplets for $\delta_{ND}=0.0$ mm. The distributions are in general similar to those of 10–20 μm droplets. At $x=25$ mm, negative swirl of larger droplets is stronger than that of small droplets because of larger inertia of droplets. At farther downstream locations, swirl velocity of 90–100 μm droplets is smaller than that of 10–20 μm droplets.

Spray Characteristics. In Fig. 9, the spray characteristics measured at $x=45$ mm for $\delta_{ND}=0.0$ mm are shown. There are two maxima in SMD of about 120 μm at $r\approx\pm10$ mm and three minima in SMD of about 100 μm at the center and of about 80 μm at $r\approx\pm40$ mm. The maxima in SMD at $r\approx\pm10$ mm are due to large droplets produced by atomization of the liquid sheet directly meeting the inner swirling flow before the fuel attaches to the prefilmer wall. This is evident from the visualization of spray; the bright region near the center. The large SMD near the edge appears because large droplets move outward due to their initial momentum and a part of fluid colliding with the outer flare wall produces a few large droplets. The minimum in SMD at the center of spray may be due to the characteristics of simplex atomization and the small droplets recirculated from downstream. The minima in SMD at $r\approx\pm40$ mm is due to small droplets produced by prefilming airblast atomization.

There are four maxima in number density. The maxima in number density at $r\approx\pm20$ mm are due to droplets produced by airblast atomization and droplets recirculated from downstream which are captured in the stagnant region. The location of maxima ($r\approx\pm20$ mm) are close to the boundary of reverse flow region.

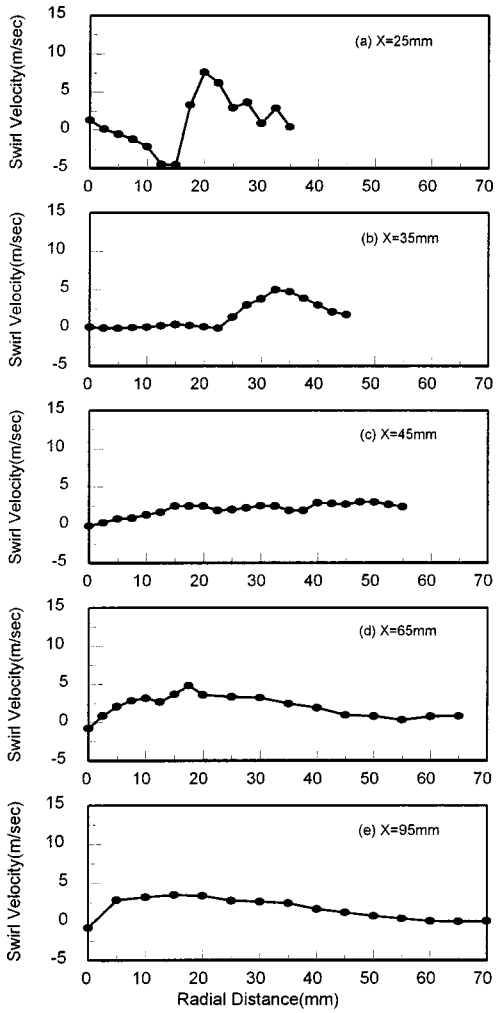


Fig. 8 Swirl velocity of 90–100 μm droplets for δ_{ND}=0.0 mm

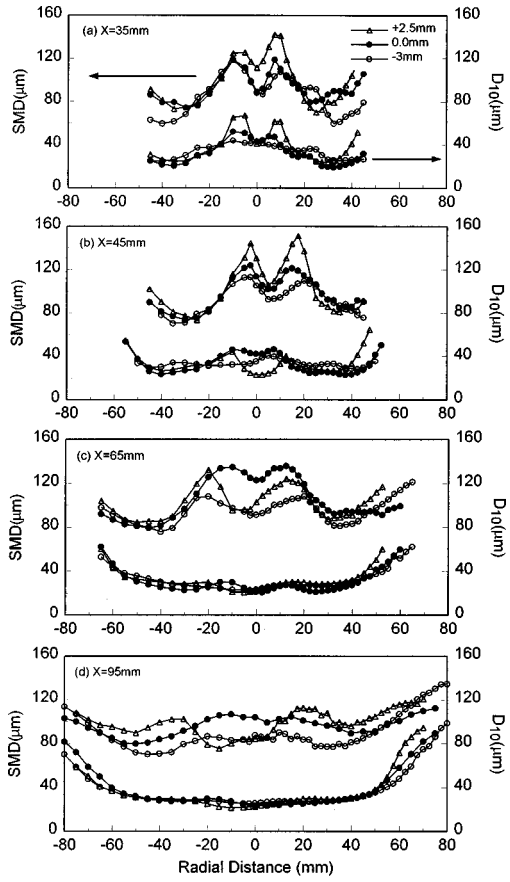


Fig. 10 Effects of nozzle displacement on SMD distribution

The maxima at $r \approx \pm 40$ mm are due to droplets produced by airblast atomization and captured in the outer, low speed region. The maxima in volume flux at $r \approx \pm 10$ mm are due to the droplets with large size. At $r \approx \pm 30$ mm, there are small peaks due to high velocity flow carrying large number of small droplets produced by airblast atomization.

Figure 10 shows SMD (D_{32}) and number diameter (D_{10}) distributions for the three δ_{ND}'s. When the nozzle displaces positively ($\delta_{ND}=+2.5$ mm), SMD near the center of spray ($r< \pm 10$ mm) increases (at $x=35$ mm). This may be because the portion of droplets which are not produced by the prefilming process but by the interaction with the inner swirling air increases. The droplets produced by the latter mechanism have larger size due to lower Weber number. On the other hand, when the nozzle displaces negatively. SMD near the edge of spray decreases (at $x=35$ mm). This may be because the prefilming length of the fuel increases so that the thickness of the film at the exit of the prefilmer decreases. As the spray travels downstream, D_{32} and D_{10} decreases near the center but increases near the edge of spray. This may be explained by outward dispersion of large droplets.

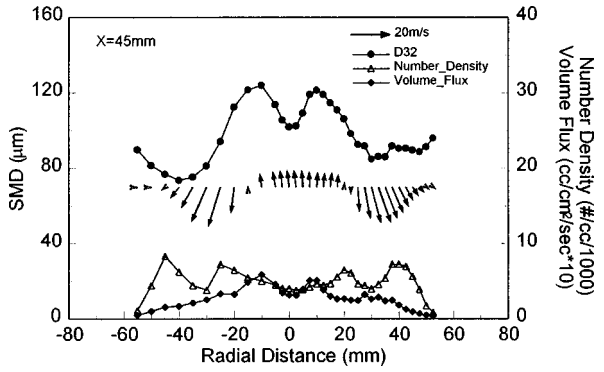


Fig. 9 Spray characteristics

Figure 11 presents the number density distributions for the three δ_{ND}'s. The number density have four maxima at $x=35$ mm as explained above. At this axial position, the number density increases and its distribution becomes wider as the fuel nozzle displaces upwards ($\delta_{ND}=-3.0$ mm). At $x=65$, 95 mm, the number density distribution for the three δ_{ND}'s are significantly different. At $x=65$ mm, the number density for $\delta_{ND}=+2.5$ mm has a large peak near the center. At $x=95$ mm, the number density for $\delta_{ND}=-3.0$ mm has a large peak near the center. The location of the peak corresponds to the boundary of the recirculation region. This indicates that recirculation region extends to further downstream as the nozzle displaces upwards ($\delta_{ND}=-3.0$ mm). Near the edge of spray, the number density for three δ_{ND}'s decreases as the spray travels downstream and disperses. The main parameter which influences the number density distribution at $x>65$ mm is the recirculation pattern.

Figure 12 presents the volume flux distributions for the three δ_{ND}'s. Large peak exist near the center at $x=35$ mm and $x=45$ mm, when the nozzle displaces downwards ($\delta_{ND}=+2.5$ mm). The distribution of volume flux becomes more uniform as the nozzle displaces upwards ($\delta_{ND}=-3.0$ mm). This indicates that the positive nozzle displacement results in more con-

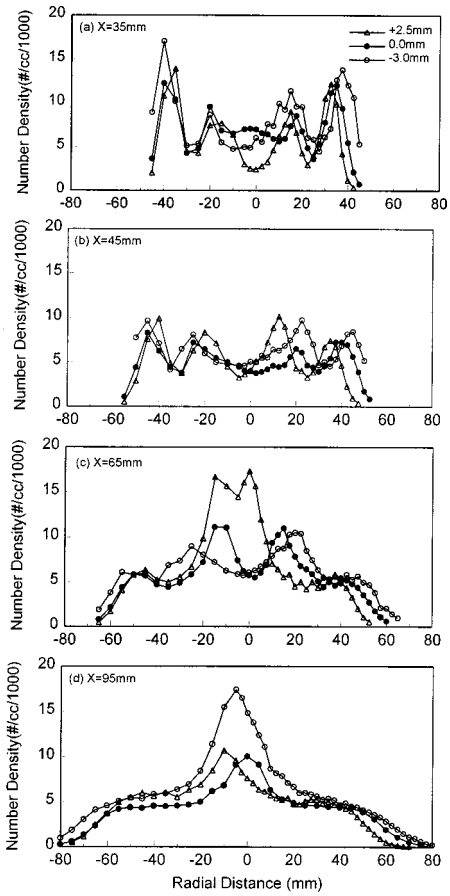


Fig. 11 Effects of nozzle displacement on number density distribution

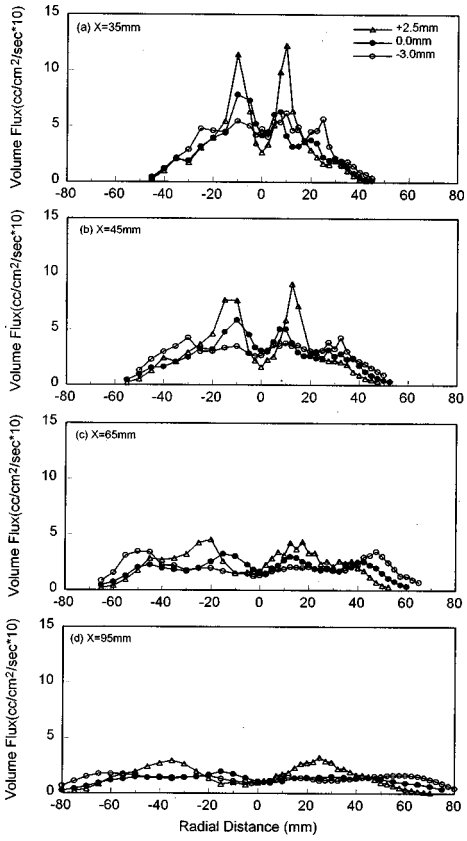


Fig. 12 Effects of nozzle displacement on volume flux distribution

centrated volume flux distribution near the center, while the negative nozzle displacements results in more uniform volume flux distribution.

In addition to the deterioration in volume flux distribution, nozzle shift in the downstream direction of about 15 percent of the nozzle diameter ($\delta_{ND}=+2.5$ mm) causes SMD to increase by about 20 percent in the center of spray near the prefilmer exit. On the other hand, nozzle shift in the upstream direction of about 18 percent of the nozzle diameter ($\delta_{ND}=-3.0$ mm) causes SMD to decrease by about 10 percent in the center of spray and by about 20 percent in the spray edge near the prefilmer exit.

Conclusions

An experimental study was carried out to investigate the effects of nozzle displacement on the structure of fuel spray. It was shown by flow visualization that significant changes occur in the spray structure as the fuel nozzle displaces with respect to the swirler/prefilmer module. Nozzle shift in the downstream direction (positive displacement) causes the amount of fuel subjected to the prefilming atomization to decrease and makes the spray less divergent. Nozzle shift in the upstream direction (negative displacement) causes the amount of fuel subjected to the prefilming atomization to increase and makes the spray more divergent.

The recirculation pattern of the flow is significantly changed as the nozzle displacement alters the resultant swirl and the axial pressure gradient at the exit of prefilmer lip. Nozzle shift in the downstream direction results in a smaller resultant swirl and a larger axial pressure gradient near the prefilmer exit and causes the recirculation zone to become narrower and to shift upstream. Nozzle shift in the upstream direction results in a larger resultant swirl and a smaller axial pressure gradient near the prefilmer exit and causes the recirculation zone to become wider and to shift downstream.

Effects of nozzle shift on the spray structure are as follows. As the nozzle shifts in the downstream direction, SMD near the center of spray increases since the portion of fuel subjected to prefilming atomization diminishes. On the other hand, as the nozzle shifts in the upstream direction, SMD near the edge of spray decreases since the prefilming length increases so that the thickness of the film at the exit of the prefilmer decreases. Near the prefilmer exit, the number density of fuel droplets increases and its distribution becomes wider as the fuel nozzle shifts in the upstream direction. In the region away from the prefilmer exit, the main factor which influences the number density distribution is the recirculation pattern of the flow, and the nozzle displacement affects the number density distribution by influencing the recirculation pattern. As indicated by the visualization and the measurement results described in the above, nozzle shift in the downstream direction results in more concentrated volume flux distribution near the center, while nozzle shift in the upstream direction results in more uniform volume flux distribution.

Acknowledgments

This work has been sponsored by the Korea Ministry of Science and Technology. A partial support by the Korea Federation of Science and Technology Societies is also gratefully acknowledged.

Nomenclature

D_{32} = Sauter mean diameter (SMD), μm
ΔP_a = air pressure drop across the swirler, kPa
ΔP_f = fuel pressure drop across the fuel nozzle, bar
δ = displacement

Subscripts

ND = nozzle displacement

References

[1] Lefebvre, A. H., and Miller, D., 1966, ''The Development of an Air Blast Atomizer for Gas Turbine Application,'' COA-Report-Aero-193, College of Aeronautics, Cranfield, England.

[2] Rizkalla, A., and Lefebvre, A. H., 1975, ''The Influence of Air and Liquid Properties on Airblast Atomization,'' ASME J. Eng. Gas Turbines Power, **97**, No. 3, pp. 316–320.

[3] Rizk, N. K., and Lefebvre, A. H., 1980, ''Influence of Liquid Film Thickness on Airblast Atomization,'' ASME J. Eng. Gas Turbines Power, **102**, pp. 706–710.

[4] Lorenzetto, G. E., and Lefebvre, A. H., 1977, ''Measurements of Drop Size on a Plain Jet Airblast Atomizer,'' AIAA J., **15**, No. 7, pp. 1006–1010.

[5] Jasuja, A. K., 1979, ''Atomization of Crude and Residual Fuel Oils,'' ASME J. Eng. Gas Turbines Power, **101**, No. 2, pp. 250–258.

[6] El-Shanawany, M. S. M. R., and Lefebvre, A. H., 1980, ''Airblast Atomization: The Effect of Linear Scale on Mean Drop Size,'' J. Energy, **4**, No. 4, pp. 184–189.

[7] Sattlemayer, T., and Wittig, S., 1986, ''Internal Flow Effects in Prefilming Airblast Atomizers: Mechanisms of Atomization and Droplet Spectra,'' ASME J. Eng. Gas Turbines Power, **108**, pp. 465–472.

[8] Wittig, S., Himmelsbach, J., Noll, B., Feld, H. J., and Samenfink, W., 1992, ''Motion and Evaporation of Shear-Driven Liquid Films in Turbulent Gases,'' ASME J. Eng. Gas Turbines Power, **114**, pp. 395–400.

[9] Breña de la Rosa, A., Wang, G., and Bachalo, W. D., 1992, ''The Effect of Swirl on the Velocity and Turbulence Fields of a Liquid Spray,'' ASME J. Eng. Gas Turbines Power, **114**, pp. 72–81.

[10] Wang, H. Y., McDonell, V. G., and Samuelsen, G. S., 1992, ''The Two Phase Flow Downstream of a Production Engine Combustor Swirl Cup,'' *Proceedings, Twenty-Fourth Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA, pp. 1457–1463.

[11] Wang, H. Y., McDonell, V. G., and Samuelsen, G. S., 1993, ''Influence of Hardware Design on the Flow Field Structures and the Patterns of Droplets Dispersion: Part I—Mean Quantities,'' ASME Paper 93-GT-199.

[12] Rizk, N. K., and Mongia, H. C., 1992, ''Calculation Approach Validation for Airblast Atomizers,'' Transactions of the ASME, **114**, April, pp. 386–394.

[13] Lefebvre, A. H., 1995, ''The Role of Fuel Preparation in Low Emissions Combustion,'' ASME Paper 95-GT-465.

[14] Han, Y. M., Seol, W. S., Yoon, M. S., Lee, D. S., Yagodkin, V. I., and Jeung, I. S., 1997, ''An Experimental Study on Modeling of Fuel Atomization for Simulating the Idle Regime of a Gas Turbine Combustor by Atmospheric Testing,'' ASME Paper 97-GT-152.

[15] Bach, T. V., and Gouldin, F. C., 1982, ''Flow Measurements in a Model Swirl Combustor,'' AIAA J., **20**, No. 5, pp. 642–651.

[16] Blümcke, E., Eickhoff, E., Eassa, C., and Koopman, J., 1987, ''Analysis of the Flow Through Double Airblast Atomizers,'' AGARD-CP-422, pp. 40_1-40_13.

[17] Wang, H. Y., McDonell, V. G., Sowa, W. A., and Samuelsen, G. S., 1993, ''Scaling of the Two-Phase Flow Downstream of a Gas Turbine Combustor Swirl Cup: Part I—Mean Quantities,'' ASME J. Eng. Gas Turbines Power, **115**, pp. 453–460.

[18] McDonell, V. G., Lee, S. W., and Samuelsen, G. S., 1994, ''Effect of Semi-Confinement on Spray Behavior in the Flow Field Downstream of an Aero-Engine Combustor Dome,'' ICLASS-94 Rouen, France, July, Paper VII-6, pp. 726–733.

[19] McDonell, V. G., Lee, S. W., and Samuelsen, G. S., 1995, ''Spray Behavior in Reacting and Non-Reacting Flow Fields Downstream of an Aero-Engine Combustor Dome,'' in *Mechanics and Combustion of Droplets and Sprays*, H. H. Chiu, and N. Chigier, eds., Begell, Housers, Inc., New York, pp. 281–294.

Vortex Generators in Lean-Premix Combustion

A. Eroglu

K. Döbbeling

ABB Corporate Research Ltd.,
Baden-Dättwill, Switzerland

F. Joos

P. Brunner

ABB Power Generation Ltd.,
Baden, Switzerland

A novel fuel-air mixing technique on the basis of vortex generators has been developed and successfully implemented in the worlds first lean-premix reheat combustor of ABB's GT24/GT26 series industrial gas turbines. This technique uses a special arrangement of delta-wing type vortex generators to achieve rapid mixing through longitudinal vortices, which produce low pressure drop and no recirculation zones along the mixing section. In this paper, after a short introduction to the topic, the motivation for utilizing vortex generators and the main considerations in their design are explained. A detailed analysis of the flow field, pressure drop and the strength of the vortices generated by a single vortex generator are presented as one of the three main geometrical parameters is varied. The results obtained through water model tests indicate that an optimum vortex generator geometry exists, which produces the maximum circulation at a relatively low pressure drop price. Moreover, the axial velocity distribution along the mixing section stays uniform enough to assure flash-back free operation despite the elevated inlet temperatures encountered in a reheat combustor. After selecting this optimized geometry, the process of the arrangement of multiple vortex generators in an annular combustor segment is described. The optimum arrangement presented here is suitable both for gaseous and liquid fuel injection, since it requires only one injection location per combustor segment.
[DOI: 10.1115/1.1335481]

Introduction

On the way to reach the ultra low emission targets of modern gas turbines, lean premixed combustion appears to be the most promising technique available today. This technique requires, on the one hand, passing the maximum amount of air through the combustor, and on the other, a complete mixing of the air and fuel injected into it. Only after complete and uniform mixing of air with fuel, locally lean conditions can be achieved within the combustion zone, which in turn guarantees low NO_X formation. The success of this technique depends primarily on the quality of mixing that can be achieved prior to combustion.

However, achieving sufficiently good mixing within limited space and residence time available in the mixing section of a gas turbine combustor is not a simple task. Difficulties arise due to conflicting requirements from different aspects of the combustor design, such as mixing, flashback safety, pressure drop, robustness and reliability of the design.

In order to obtain proper mixing of the fuel and the air streams both large-scale distribution and fine-scale mixing are necessary. However, given the fact that the mass flowrate of the fuel accounts only a few percent of the mass flowrate of the air, and the momentum of the fuel injection is limited with the available supply pressure, it is not possible to distribute the fuel uniformly within the air stream when a limited number of injection points are employed. Unfortunately, a potential multi-point injection solution runs into its limits too, when the size of injection orifices fall below an allowed limit. Additionally, suitability of such a solution for liquid fuel injection, which requires special atomizers, is questionable. Another problem which led to abandonment of this path in the past has been the combustion instabilities observed, especially under high pressure conditions.

A simple way of overcoming several problems associated with achieving proper mixing quality is utilizing the momentum of the air stream via vortex generators. By this way, large scale vortices can be created by the vortex generators which are employed first for bulk distribution of the fuel and subsequent fine scale mixing. A comprehensive research program has been undertaken at ABB Corporate Research Center to develop fast mixing techniques for low emission combustion, based on vortex generators. The requirements that no recirculation or low velocity zones can be tolerated and the pressure drop due to vortex generators has to be as low as possible led to vortex generators which generate exclusively streamwise vortices.

A design on the basis of these vortex generators have been developed and implemented in the second combustor (SEV) of GT24/GT26 as shown in Fig. 1. A more detailed view of the vortex generators, as pictured from the downstream end of the combustor, is presented in Fig. 2.

Previously, a detailed account of the development of the GT24/GT26 machines, EV burners and SEV combustors have been presented [1–5]. In this paper the fundamental findings acquired with various vortex generator geometries during the development of the SEV combustor are described. First, flow field measurements with laser Doppler anemometry (LDA) and mixing quality investigations with laser induced fluorescence (LIF) techniques from water model tests are presented. Then, conclusions are drawn on the optimum vortex generator geometry based on mixing speed, pressure drop and flashback danger. Additionally, the arrangement of multiple vortex generators with integrated fuel injection in a combustor segment and resulting vortex pattern and mixing quality are presented. Finally, main results and conclusions from these investigations are summarized.

Experimental Techniques

The mixing and aerodynamics investigations reported here have been carried out in a water channel with transparent models which have full optical access from all sides and from the downstream end for LIF and LDA measurements. Additionally, static pressure measurements have been conducted in these models with pressure tabs installed upstream and downstream of the mixing section model.

The water model has been operated in closed-cycle mode, with the exception of the LIF tests, during which it was switched to open-loop mode in order to prevent contamination of the main stream with injected dye. The mean axial velocity of the main flow in water model was 1 m/s, resulting in a Reynolds number on the basis of the hydraulic diameter of about 54,000. Upstream of

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Stockholm, Sweden, June 2–5, 1998; ASME Paper 98-GT-487. Manuscript received by IGTI March 25, 1998; final revision received by the ASME Headquarters October 20, 1999. Associate Technical Editor: R. Kielb.

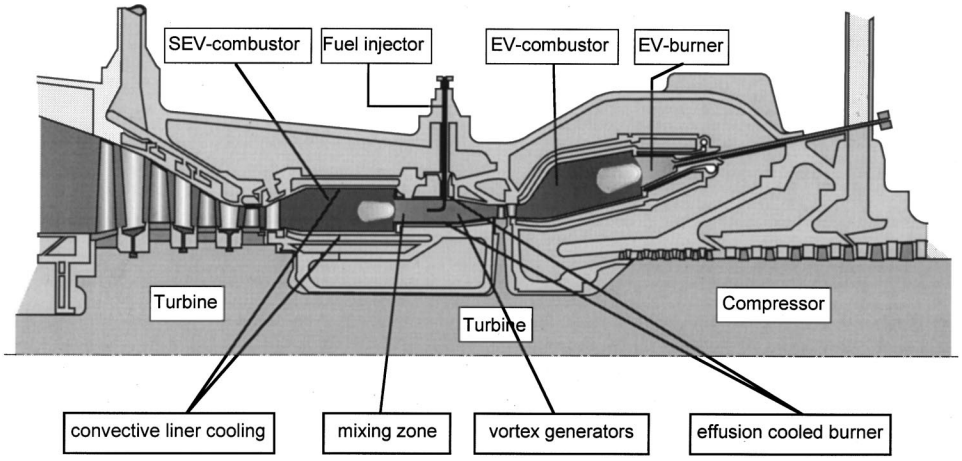


Fig. 1 Combustion system of GT24/GT26 gas turbines

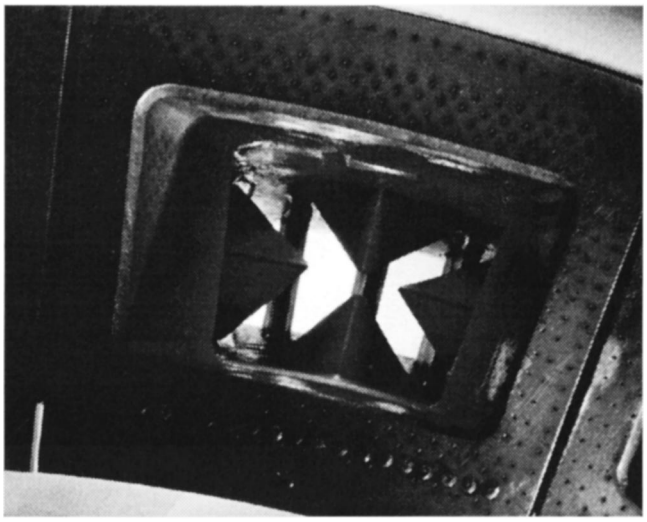


Fig. 2 SEV combustor as viewed from the downstream end

the test section, a settling chamber, equipped with honeycomb-screen type of flow straighteners, followed by a 9:1 area ratio matched-cubic contraction ensured nonuniformities in the mean axial velocity of less than 2 percent and a free-stream turbulence intensity of about 1.3 percent.

The LIF tests have been conducted with the blue line (488 nm) of a 5 W Argon-Ion laser beam, transmitted to the test rig via fiber-optic cable and expanded to a sheet of about 0.6 mm thickness via a cylindrical lens. The injected fluid is a weak solution of disodium fluorescein, as the main stream is free of dye. The injectant concentrations visualized with this technique have been recorded with a monochrome CCD camera and digitized with the help of an 8 bit frame grabber board, which was installed on a computer. A commercial image processing software was used for evaluation of statistical values such as the mean injectant concentration and the standard deviation by averaging 10 successive pictures. The accuracy of this method is estimated to be better than 5 percent.

The velocity measurements have been carried out with a two-component LDA system, operated in backscatter mode. Two components of the velocity are measured simultaneously and the third one is measured separately by rotating the probe 90 deg with respect to first measurement. The measurement probe was traversed with a computer controlled three-axis traverse system.

Reference Case

As a starting point, the mixing characteristics of a fuel injector without any vortex generators are of interest. A number of different injection geometries, including transverse jets injected from inner and outer walls, in-stream tube injectors and central single point injection have been tested. Due to practical considerations (e.g., number of parts, cooling, suitability for liquid fuel injection, thermoacoustic instabilities) an injector geometry as depicted in Fig. 3 has been selected. This injector consists of a single tube which is inserted into the flow through outer liner and bent 90 deg in the flow direction. Four injection holes are located at the tip, pointing slightly away from the horizontal plane in order to achieve the best possible distribution over the channel cross sec-

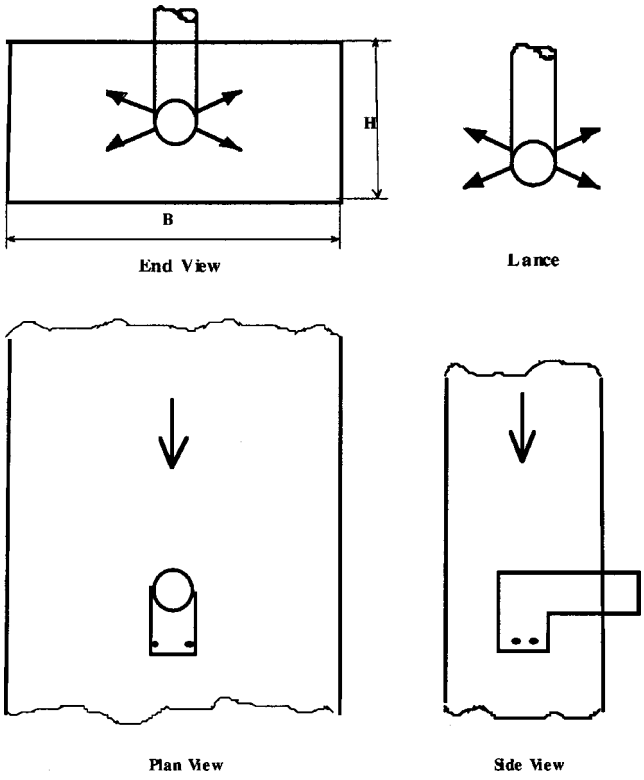


Fig. 3 Injection geometry for the reference case without vortex generators

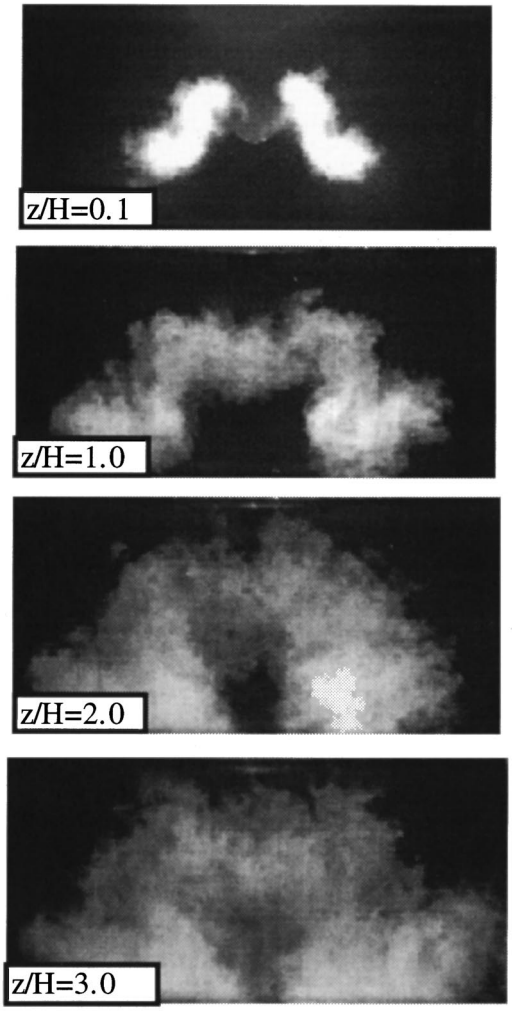


Fig. 4 The LIF pictures of fuel concentration at four successive stations downstream of a central lance without vortex generators

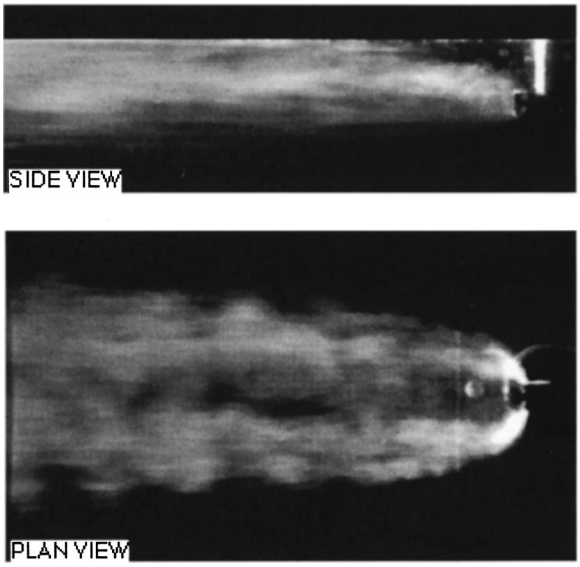


Fig. 5 The LIF pictures of fuel concentration in streamwise planes downstream of a central lance without vortex generators

tion without fuel engulfment in the wake of the injector and without impingement of the fuel jets on the liner walls.

The LIF pictures of injectant concentration over planes perpendicular to mainflow direction are given in Fig. 4, at streamwise distances of $z/H=0.1$, 1, 2, and 3. As can be clearly observed from these pictures, there exists zones with a wide variation in fuel concentration next to each other in addition to regions where no fuel exists, even at three channel-height streamwise distance from the injection. Additionally, with the light sheet positioned parallel to flow direction, radial and circumferential symmetry planes are given in Fig. 5. The flow direction is from right to left. The injector is partly visible at the right side of the picture. These concentration pictures indicate that the quality of both large and fine-scale mixing is far from being acceptable. Additional variants of the same injection geometry with increased injection momentum or other injection hole arrangements deliver similar results. Namely, only incremental modifications in the large-scale distribution pattern can be achieved when relied on the momentum of the injection alone. A major improvement in mixing requires utilizing the momentum of the main stream via vortex generators.

Single Vortex Generator

Upon recognition of the fact that the mass and momentum flow rate of the injection are not high enough to achieve the mixing quality needed within an acceptable axial length, methods of utilizing the momentum of the main stream have been investigated. The main requirements from a successful technique are as follows:

1 no recirculation zones or regions of low velocity along the mixing section where self-ignition may occur or the flame can be attached
2 low pressure drop
3 applicable both for gaseous and liquid fuel injection
4 suitable geometry for cooling in case it is necessary
5 simple and robust design

The requirements of low pressure drop and high safety against flame in the mixing zone led exclusively to delta-wing type vortex generators which can generate streamwise vortices without any recirculation zones. In order to satisfy other practical considerations such as mechanical integrity and cooling, a tetrahedral geometry, which consists of two half delta-wing side surfaces and a full-delta-wing upper surface has been chosen. A sketch of this device is given in Fig. 6, labeled with the principal parameters which define the geometry.

A series of water model tests with this type of vortex generators have been carried out in order to optimize the strength of the vortices produced, the flow velocities downstream, and the pressure drop. These tests have been conducted in a rectangular plexiglas channel. The ratio of channel height to VG height H/h was above two in order to minimize the influence of channel walls on the flow field. The measurement planes are perpendicular to the flow direction and located at $z/H=0.1$, 0.5, 1.0, 2.0, and 3.0. The height to width ratio of the vortex generator was fixed as the length to width ratio of the vortex generator is varied in four

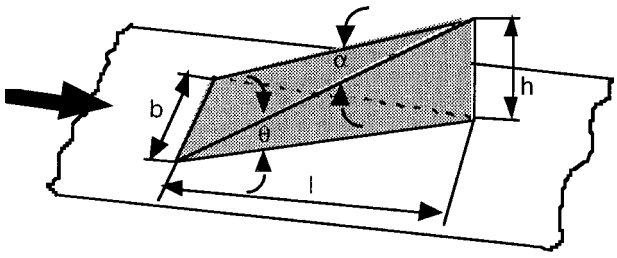


Fig. 6 The geometry of a single vortex generator element

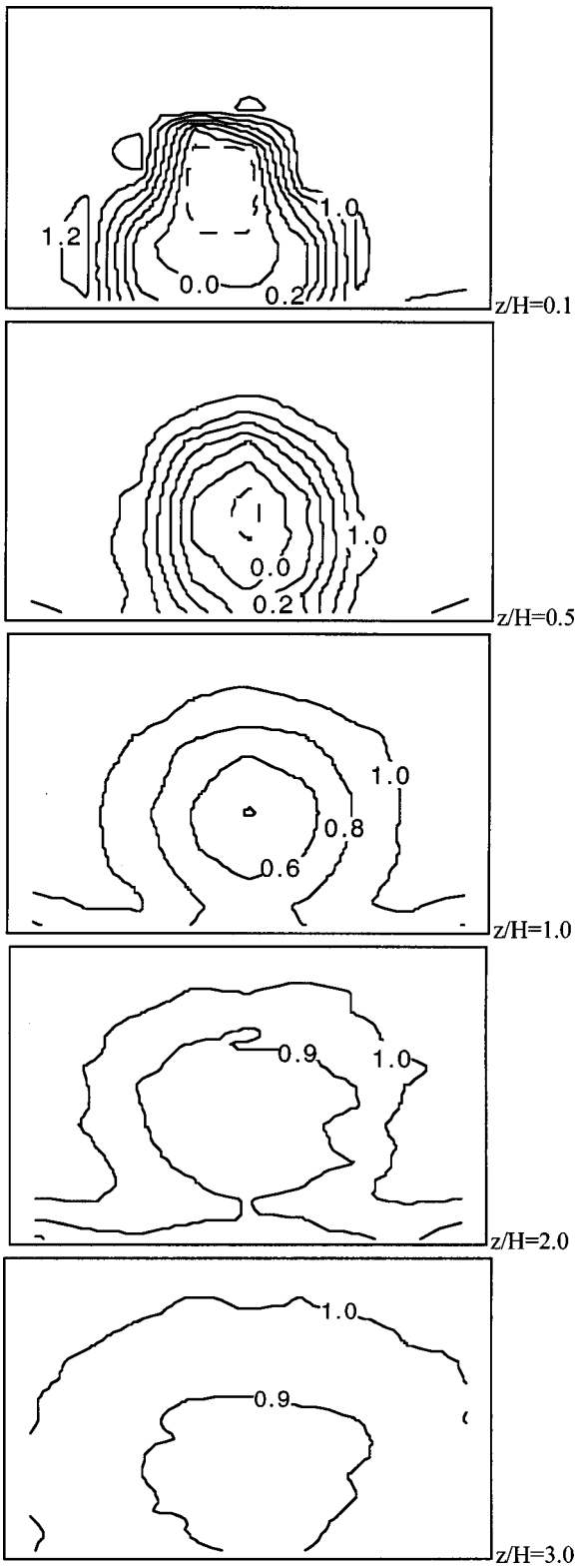


Fig. 7 Normalized axial velocity (U/U_∞) distributions at five successive planes downstream of vortex generator version A

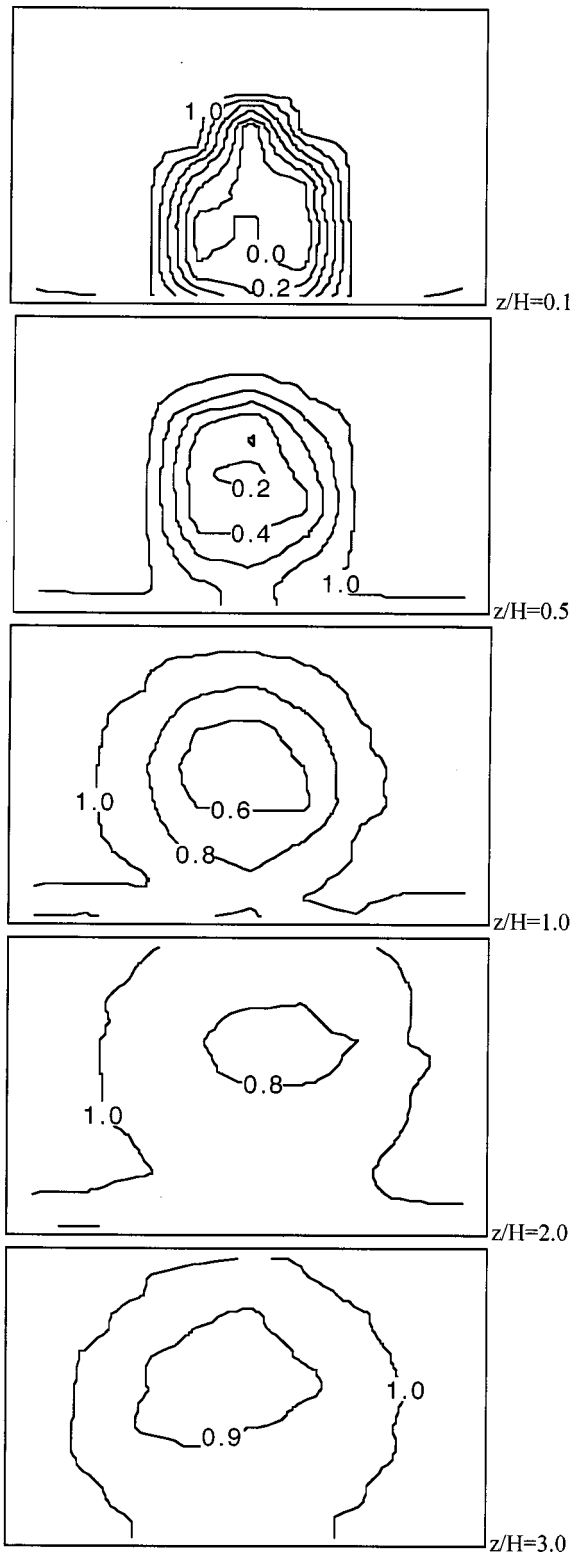


Fig. 8 Normalized axial velocity (U/U_∞) distributions at five successive planes downstream of vortex generator version B

discrete steps. These four steps are labeled as versions A: length less than width, B: length equal to width, C: length 50 percent greater than width, and D: length equal to twice the width.

Axial Velocity Distribution Along the Mixing Section. All three components of the flow velocity have been measured at successive stations downstream of a vortex generator for constant vortex generator height (h) and width (b), as the length of the vortex generator (l) has been varied. Among all velocity components, the mean axial velocity component is a good indicator of the potential flame stabilization danger in the mixing zone. In Figs. 7–10, a series of contour plots showing the distribution of

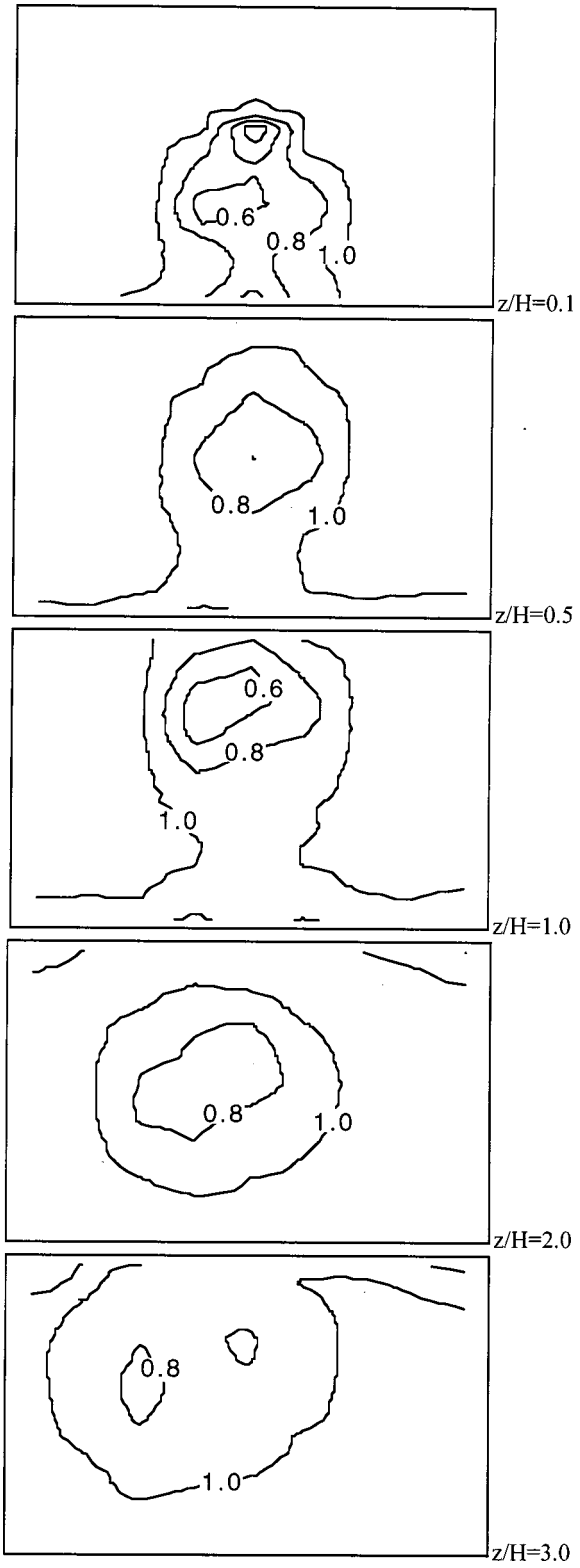


Fig. 9 Normalized axial velocity (U/U_∞) distributions at five successive planes downstream of vortex generator version C

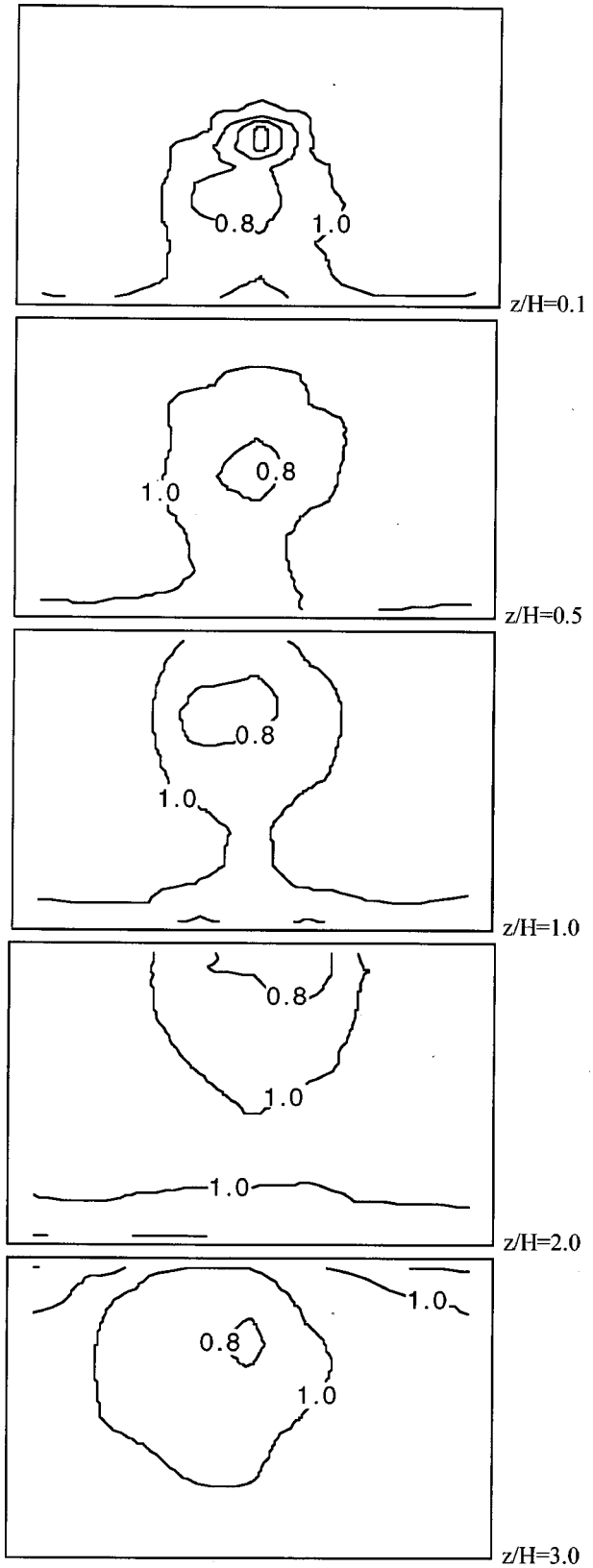


Fig. 10 Normalized axial velocity (U/U_∞) distributions at five successive planes downstream of vortex generator version D

the mean axial velocity over transverse planes downstream of the vortex generator are presented for all four geometries tested. The horizontal axis length of the plots correspond to the central 50 percent of the total channel width. The vertical axis of the plots covers the bottom 62.5 percent of the total channel height. The velocity values shown are nondimensional, normalized with the mean axial velocity.

The data from the first measurement plane at $z/H = 0.1$ reveal that the first two geometries, namely A and B give rise to a recirculation zone at this plane immediately downstream of the vortex generator. The other two geometries, namely C and D do not exhibit any negative or zero velocity zones.

Further downstream at $z/H=0.5$, the recirculation zone downstream of version B is already closed, as the negative velocities of version A are replaced by zero velocities. Evidently, all variants tested exhibit positive axial velocities starting from $z/H=1.0$. The growth rate of the region which is influenced by the vortex generator is the largest for the version A. At $z/H=3.0$, the influence of the vortex generator is visible over almost the entire channel cross-section for this shortest variant tested as the region influenced by the longest one is still relatively small at this station.

These results demonstrate that, for a given h/b ratio, a minimum l/b ratio is necessary in order to prevent the formation of a recirculation zone downstream of the vortex generator. The exact value of this critical parameter can be best determined after considering the remaining two aspects of the design. The first one is the circulation, which is a measure of the vortex strength or, indirectly, the quality of mixing that can be achieved by a vortex generator. The second one is the pressure drop caused by this device, which influences the efficiency and power output of the whole cycle.

Circulation Versus Pressure Drop. In addition to providing the basis for the assessment of the potential danger of flame stabilization in the mixing section, the velocity measurements mentioned above served to determine the strength of the vortices generated by each vortex generator. Based on the velocity distribution over transverse planes downstream of the vortex generator, it was possible to determine the vorticity distribution, which is defined as the curl of the velocity vector, namely

$$\vec{\omega}=\vec{\nabla}\times\vec{V}=\left(\frac{\delta}{\delta_x}\vec{e}_x+\frac{\delta}{\delta_y}\vec{e}_y+\frac{\delta}{\delta_z}\vec{e}_z\right)\times(U\vec{e}_x+V\vec{e}_y+W\vec{e}_z). \quad (1)$$

As far as the strength of the longitudinal vortices is concerned, the streamwise component of the vorticity vector is of main interest, which is defined as

$$\omega_z=\left(\frac{\partial V}{\partial x}-\frac{\partial U}{\partial y}\right). \quad (2)$$

The partial derivatives in Eq. (2) can be approximated by using a central differencing scheme

$$\frac{\partial V}{\partial x}=\frac{V_{i+1,j}-V_{i-1,j}}{2\Delta x} \quad (3)$$

and

$$\frac{\partial U}{\partial y}=\frac{U_{i,j+1}-U_{i,j-1}}{2\Delta y}. \quad (4)$$

Substituting (3) and (4) in (2) yields,

$$\omega_z=\left(\frac{V_{i+1,j}-V_{i-1,j}}{2\Delta x}-\frac{U_{i,j+1}-V_{i,j-1}}{2\Delta y}\right), \quad (6)$$

where i and j correspond to the coordinates of the grid point in x and y directions, respectively.

The circulation, Γ, as the value of the net vorticity over a region of the flow, is defined by

$$\Gamma=-\oint\vec{V}\cdot d\vec{s}. \quad (7)$$

This line integral can be converted to a surface integral by employing the theorem of Stokes [6], namely

$$\Gamma=-\int_S\int(\vec{\nabla}\times\vec{V})\cdot d\vec{S} \quad (8)$$

or

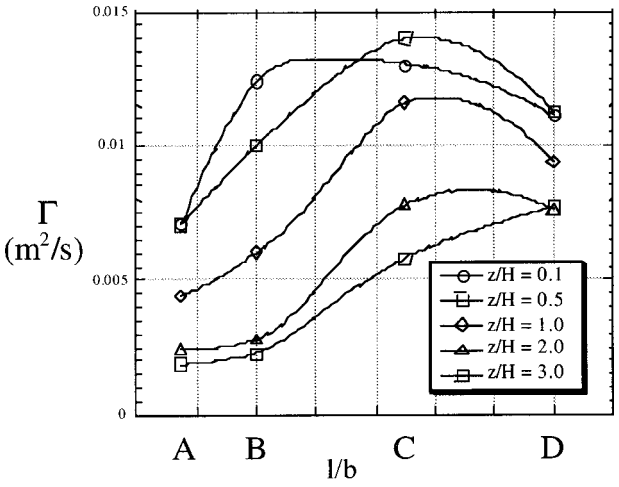


Fig. 11 Circulation values calculated from the velocity measurements at five successive planes as a function of length to width ratio of the vortex generator

$$\Gamma=-\int_S\int\vec{\omega}\cdot d\vec{S}. \quad (9)$$

This surface integral can now be calculated in discrete steps in order to calculate the circulation for each measurement plane as the sum of streamwise vorticity multiplied with the surface area of the element

$$\Gamma=-\sum_{j=1}^{M}\sum_{i=1}^{N}\omega_{i,j}\Delta x\Delta y. \quad (10)$$

This procedure has been repeated over all measurement planes for all four variants tested. The results are given in Fig. 11 for five successive planes, in the form of the circulation calculated with this method as a function of vortex generator length to width ratio.

Additionally, static pressure difference across the vortex generators, as measured between a plane located at one channel height upstream and a second plane located at eight channel height downstream of the vortex generator, are presented in Fig. 12. These measurements have been carried out both with a single element and four identical elements positioned side-by-side.

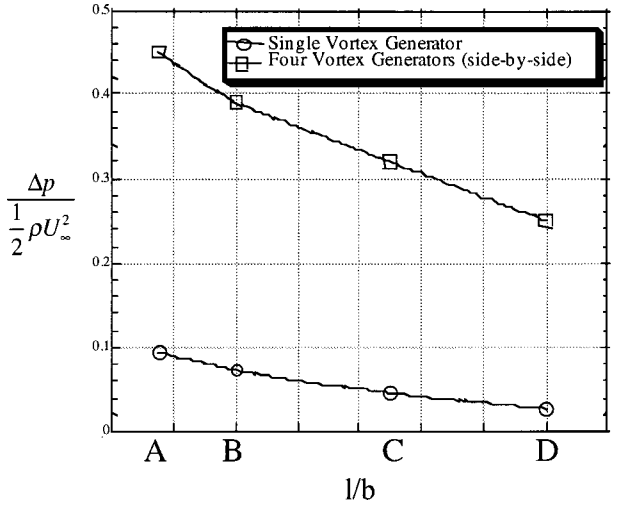


Fig. 12 The pressure drop coefficient of a single and four vortex generators as the length to width ratio is varied

The dependence of pressure drop coefficient on the vortex generator length is monotonic, as one would expect. In other words, the pressure drop caused by a vortex generator increases steadily with decreasing length. This can be attributed to the fact that with decreasing vortex generator length, the portion of the dynamic pressure which can be recovered downstream is reduced as the steeper angle of attack leads to higher transverse velocity components and eventually to a recirculation zone with zero and negative axial velocities.

On the other hand, the circulation distribution reaches a peak around version C as the vortex generator length is reduced. Further reduction in the length after this point causes the circulation to decrease. This also can be explained in the light of the velocity measurements which are in agreement with the previous studies on swirling flows and vortex breakdown. Studies on delta-wing type vortex generators have shown that for a fixed angle of sweep, increasing the angle of attack leads to breakdown position of the vortices to move upstream. In the extreme case of very high angle of attack, the breakdown position is located at the leading edge of the vortex generator. After breakdown, part of the streamwise vorticity is transformed into spanwise vorticity within the recirculation zone.

Based on this information, it can be concluded that for the tetrahedral vortex generator geometry selected, an optimum length exists which generates the strongest vortices for a given width and height. By selecting the geometry of the vortex generator at this optimum value, the most important requirement on the way to achieve the best mixing quality at the minimum pressure drop price is fulfilled.

Multiple Vortex Generators in a Duct Segment

Once the geometry of a single vortex generator element is defined for maximum vortex strength and for minimum pressure drop, the next step is the arrangement of these elements in a duct segment in a way which is compatible with the fuel injection. Ideally, the fuel is injected from a single location for each segment such that the same injection nozzle could be utilized for gaseous and liquid fuels.

The arrangement of the vortex generators has to be selected with this consideration in mind. Namely, the vortex pattern generated by a specific arrangement should be capable of distributing both gaseous and liquid fuel from a single injection location over the entire cross-section of the segment. A number of potential candidates for the arrangement have been considered and tested. A relatively quick and inexpensive method of determining whether a special arrangement is suitable for the large and fine-scale mixing purposes is the laser induced fluorescence (LIF) tests in water model. Here two example arrangements are presented.

A Pair of Opposing Vortex Generators with a Central Fuel Lance. In this arrangement, a pair of identical vortex generators are mounted on the opposing walls of the segment, as shown in Fig. 13. The fuel lance is located at the center, injecting through four holes which are pointing slightly away from the symmetry axis.

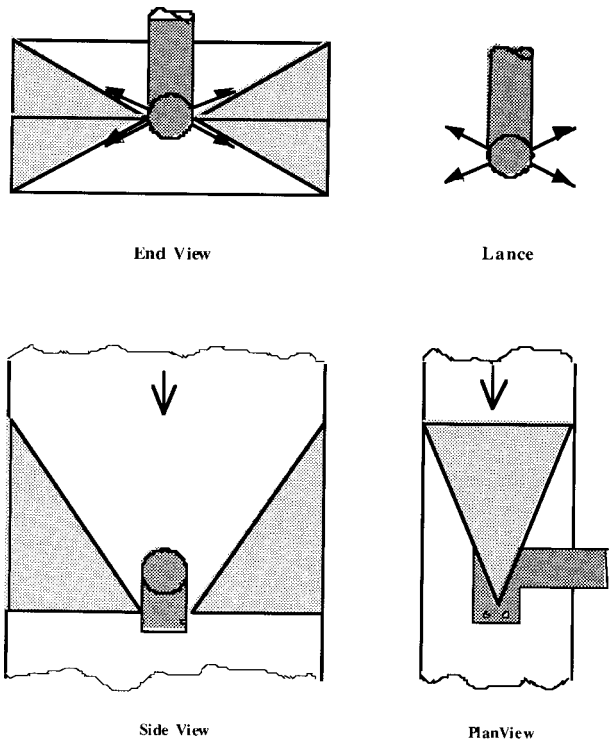


Fig. 13 The geometry of two opposing vortex generators with central injection

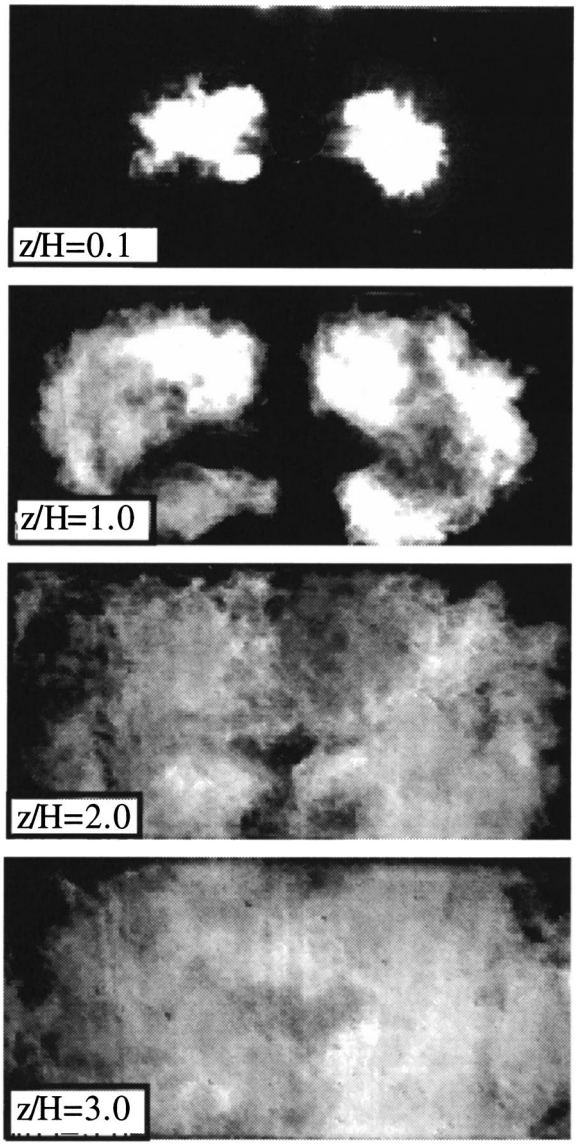


Fig. 14 The LIF pictures of fuel concentration over four successive planes downstream of a pair of opposing vortex generators with central injection

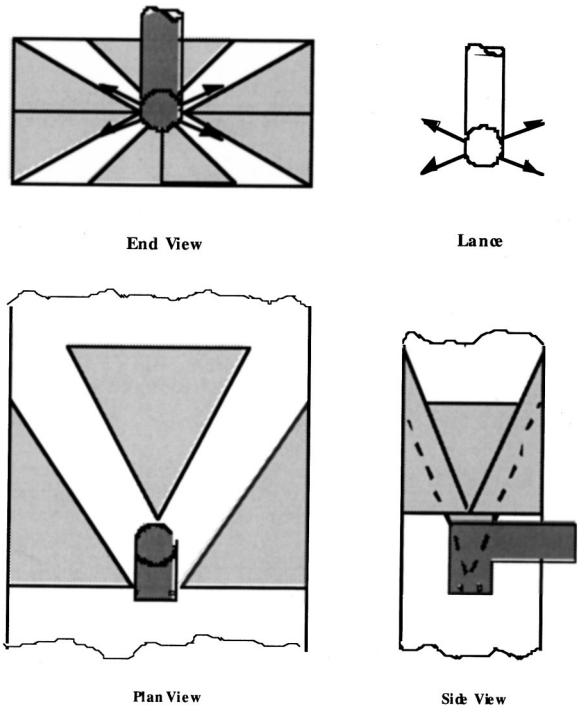


Fig. 15 The geometry of two pairs of opposing vortex generators with central injection

The LIF pictures associated with this case are presented in Fig. 14. It is evident from these pictures that the large scale distribution of the injected fuel is not complete even at three channel height downstream from the injection. Given the fact that the length of the mixing section in the machine is less than two channel heights, it is difficult to achieve the emission targets with this configuration. Additionally, due to almost completely fuel-free outer regions, insufficient flame stabilization is to be expected with such a configuration, since the recirculation zones after the sudden expansion into the combustor could not be efficiently used.

Further attempts in the direction of optimization of the mixing quality with this configuration, by modifying the injection angle and momentum of the fuel jets or the geometry of vortex generators, brought only marginal improvements at the cost of increased pressure drop and flashback danger.

Additional tests with the similar configuration, but with vortex generators mounted on the upper and lower walls have produced similar results, namely, a good portion of the channel cross-section not receiving enough fuel, resulting in large nonuniformities in mixing. As in the reference case mentioned earlier, when the momentum of the main stream is not used to the extent it is needed, problems with increased flashback danger arise due to overloading of the vortex generators by increasing the angles of attack and sweep in order to increase the mixing quality. This problem can be avoided only by increasing the number of the vortex generators, namely from a pair to two pairs as will be explained in the next section.

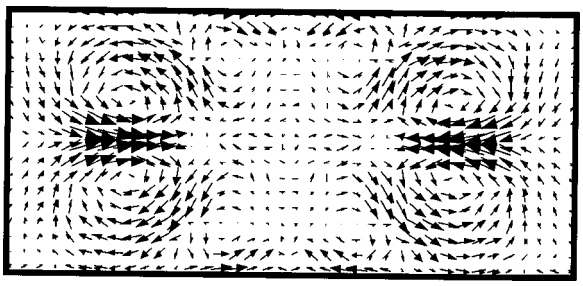


Fig. 16 The vortex pattern generated by two pairs of opposing vortex generators at the fuel injection plane

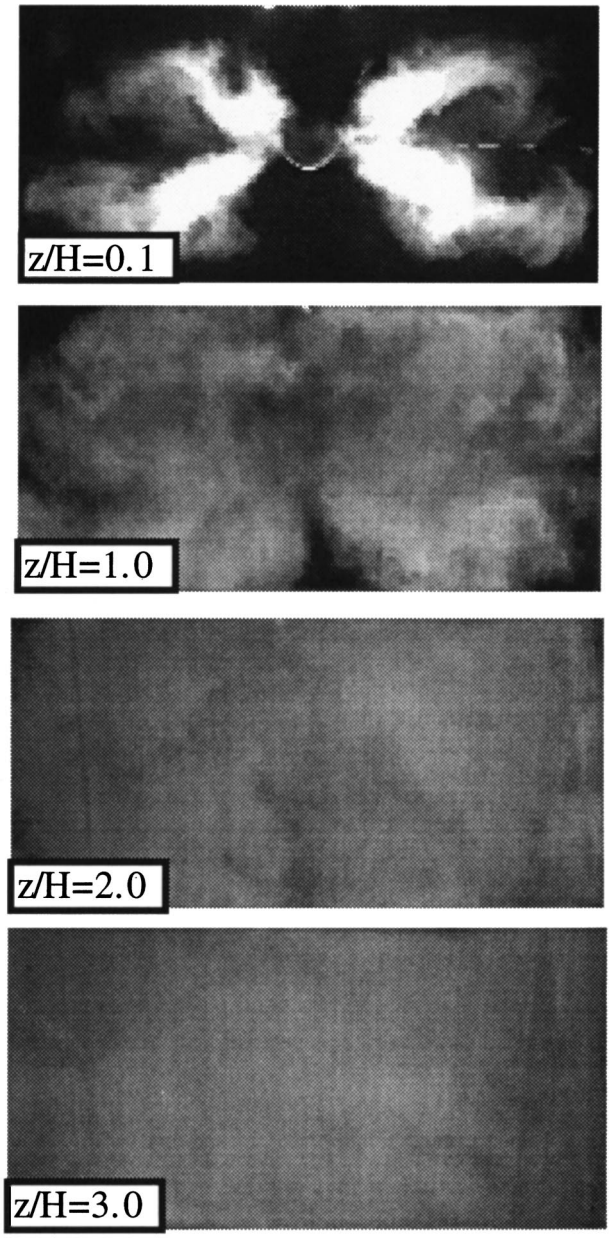


Fig. 17 The LIF pictures of fuel concentration over four successive planes downstream of two pairs of opposing vortex generators with central injection

Two Pairs of Opposing Vortex Generators with a Central Fuel Lance. This arrangement is similar to the previous one, with an additional pair of vortex generators mounted on the upper and lower walls, as shown in Fig. 15. The axial positions of the ends of the pairs are shifted with respect to each other, in order to provide the space needed for the central injector.

The vortex pattern generated by this arrangement is shown in Fig. 16, composed of velocity components perpendicular to the mean flow direction. This measurement has been carried out at the axial location of the tip of the fuel lance, in the absence of the lance itself. All four pairs of counter-rotating vortices are visible,

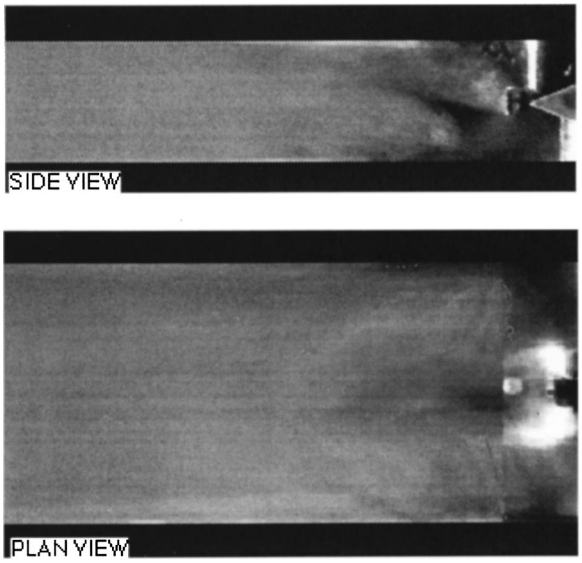


Fig. 18 The LIF pictures of fuel concentration over streamwise planes downstream of two pairs of opposing vortex generators with central injection

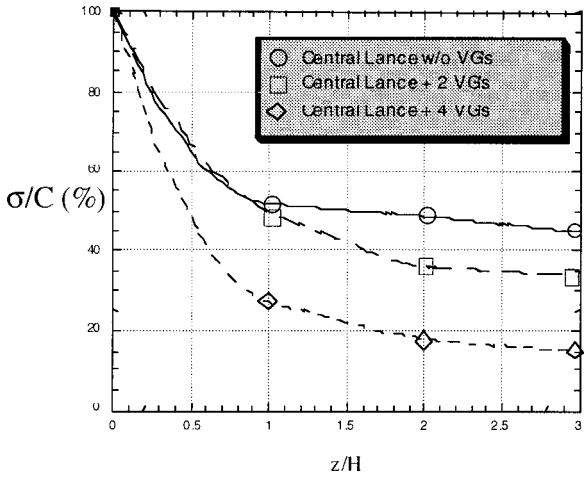


Fig. 19 Variation coefficient (standard deviation to mean concentration ratio) as a function of axial distance

each generated by one of the vortex generators. The pairs from the side vortex generators are clearly larger, in order to provide better penetration and distribution of the fuel to both sides, which are distanced further from the lance with respect to upper and lower regions.

The LIF pictures of fuel concentration for this configuration over four transverse planes downstream of the injection are given in Fig. 17. A dramatic increase in the quality of the mixing is observed from these pictures, when compared to the case with a single pair of vortex generators. Already at one channel height downstream from the injection plane, the large-scale distribution of the fuel is practically completed. Further downstream the fine-scale mixing progresses, yielding locally more uniform concentration distribution.

The LIF pictures from the same configuration in streamwise planes as observed from the side and plan views are given in Fig. 18. As it is evident from these pictures, the injected fluid is rapidly transported away from the injector and entrained into the vortices shown in Fig. 16. Moreover, no significant impingement of injectant to the side or top and bottom walls is visible, indicating that, in the case of liquid fuel injection, the coke formation danger due to droplets contacting hot surfaces prior to evaporation is practically nonexistent. An additional advantage of this configuration is that the fuel is transported away from the injector initially, to come back in the middle of the channel after the recirculation zone in the wake of the lance is closed completely, thus, ensuring no flashback danger due to fuel entrainment in the lance wake.

The fuel concentration recordings from LIF pictures have been evaluated with digital image processing techniques in order to quantify the mixing quality. The data from all three cases reported are presented in Fig. 19, where the standard deviation in local fuel concentration, σ, normalized with the mean concentration is plotted as a function of the axial distance from injection. Apparently, in the absence of any vortical structures, other than those generated by the injection itself, the mixing progresses very slow. After the first channel height axial length, the reduction in the normalized standard deviation is very gradual, reaching to a value of only about 45 percent at $z/H=3$. A single pair of vortex generators brings an improvement of about 12 percent in mixing quality at $z/H=3.0$, when compared to the reference case. The case with a pair of vortex generators shows the best value of 15 percent variation coefficient at $z/H=3.0$.

Summary

A comprehensive investigation has been carried out with delta-wing type vortex generators with the aim of fuel-air mixing in a reheat combustor. This investigation produced an optimum vortex generator geometry, which provides the maximum vortex strength at the minimum pressure drop price. Owing to the longitudinal vortices generated by this geometry, no zones with reverse flow or unacceptably low velocities are encountered along the mixing section, thus ensuring reliable operation without flame attachment or flashback prior to complete mixing.

In addition to optimization of a single element, an arrangement of two opposing pairs of vortex generators in a rectangular combustor segment is presented. This arrangement allows rapid mixing with a single, central injector, both for gaseous and liquid fuels. The experience gained with this arrangement through a number of high pressure tests [5] verified the feasibility of the concept presented here. More recently, starting with the first GT24 machine in Gilbert, NJ, and the first GT26 machine in Birr, Switzerland, a number of gas turbines with this technology have been commissioned into service, confirming the findings from the development tests.

Acknowledgments

The authors would like to acknowledge the cooperation of Dr. Rolf Althaus, Dr. Jaan Hellat, Dr. Yansong Liu, Dr. Yau-Pin Chou and Dr. Bettina Paikert. We are also grateful to Mr. Richard Straessle for invaluable assistance with the experiments and Mr. Barend Jenje for his careful measurements and evaluation of LDA data.

References

[1] Frutschi, H. U., 1994, ''Advanced Cycle System With New GT24/GT26 Turbines—Historical Background,'' ABB Rev., **1**, pp. 21–25.
[2] Sattelmayer, T., Felchlin, M., Haumann, J., Hellat, J., and Steyner, D., 1990, ''Second Generation Low Emission Combustors for ABB Gas Turbines: Burner Development and Tests at Atmospheric Pressure,'' ASME Paper 90-GT-162.
[3] Aigner, M., Mayer, A., Schiessel, P., and Strittmatter W., 1990, ''Second Generation Low Emission Combustors for ABB Gas Turbines: Tests Under Full Engine Conditions,'' ASME Paper 90-GT-308.
[4] Senior, P., Luturm, E., Polifke, W., and Sattelmayer, T., 1993, ''Combustion Technology of the ABB GT13E2 Annular Combustor,'' 20th CIMAC G22 1993, London.
[5] Joos, F., Brunner, P., Schulte-Werning, B., Syed, K., and Eroglu, A., 1996, ''Development of the Sequential Combustion System for the ABB GT24/GT26 Gas Turbine Family,'' ASME Paper 96-GT-315.
[6] Kuethe, A. M., and Chow, C.-Y., 1986, *Foundations of Aerodynamics*, John Wiley and Sons, New York.

D. M. Tsay[1]
Professor

W. F. Yan
Research Engineer

H. C. Ho
Graduate Student

Department of Mechanical Engineering,
National Sun Yat-Sen University,
Kaohsiung 80424, Taiwan

Generation of Five-Axis Cutter Paths for Turbomachinery Components

A simple, yet useful algorithm is developed to generate tool paths with global interference checking for five-axis point milling of turbomachinery components. Based on the projected distance between the surface data and the cutter-axis of a cylindrical ball-end mill, interference between the surface of a workpiece and the cutter can be detected. Given the cutter contact points of the surface and the cutter's size, it can produce the cutter location data without incurring interference through relatively rotating and tilting the workpiece. Applications of the developed procedure to five-axis machining of centrifugal compressor impellers with 13 and 15 blades are illustrated to demonstrate the usefulness and reliability of the approach. [DOI: 10.1115/1.1340639]

Introduction

To efficiently machine turbomachinery components such as centrifugal compressors, axial compressors, and fans having complex, overlapping surfaces, five-axis numerically controlled (NC) machine tools possessing many degrees of freedom are necessarily utilized to perform point milling. To produce such components with desired precision, correct cutter contact (CC) data must first be generated considering the scallop height [1,2], overcutting, and local gouging [3] of the part surfaces. More importantly, before machining, the possibility of tool interference must be avoided so that proper cutter location (CL) data can be decided and tool paths can be schemed. Otherwise, desired geometries of parts may be cut and/or the impact may result in the fracture of the cutter and the damage of the machine tool.

To machine parametric, sculptured surfaces within specified tolerances, a number of studies have been reported to characterize the scallop height and/or to improve it for local gouging. Loney and Ozsoy [4] proposed a numerical procedure to find the parametric value that will produce the largest chordal deviation. Vickers and Quan [5] determined the effective radius of a flat-end cutter as a function of cutter-tilting angles for machining low curvature surfaces. By establishing triangular polyhedron models [6,7], CL data without gouging can be yielded by checking the distances from the center of the ball-end mill to the CC data along tool paths. Oliver et al. [3] developed two techniques to analyze the chordal deviation and to detect gouging. By five-axis surface machining, Choi et al. [8] generated CL data from the given CC paths to give minimum cusp heights. Recently, Wu [9] as well as Elber and Fish [10] developed flank (side) milling techniques to machine turbomachinery parts with ruled surfaces to obtain smaller scallop heights using five-axis mill centers.

In order to remove the collision problem between the cutter and the workpiece, various approaches have been attempted to deal with global cutter interference. Typically, constructing solid models, Wang [11] proposed an algorithm to calculate interference-free zones utilizing a moving frame of reference. Using recursive subdivision of the control polygon of Bezier curves and surfaces, Tseng and Joshi [12] showed algorithms to determine tool-approach directions without interference. Based on the solid modeling technique, Takeuchi and Idemura [13] as well as Takeuchi and Watanabe [14] generated collision-free tool paths by checking the shape data of the tool that is never within the workpiece. Also, employing the control polygon of a sculptured surface, Lee [15] developed a procedure to find the feasible cutter orientation range without collision. Approximating surfaces into a set of polygons, Li and Jerard [16] recently used a bucketing scheme [17] to detect interference between a flat-end cutter and a polygon, and adjusted the tool-axis to avoid it.

Generally the complete geometry of a turbomachinery component is composed of various types of surfaces, i.e., blending surfaces, ruled surfaces, and sculptured surfaces. To machine such a part successfully, it becomes necessary to cut it by point milling. The procedure described here provides a relatively simple and effective tool to produce cutter paths with global tool interference checking. Regardless of the type of surfaces, in this article to detect global interference, it is more convenient to select CC points for checking. Furthermore, as mentioned earlier, the number of CC points are also determined according to the analysis of surface roughness for manufacturing tolerances. Based on the projected distance between the CC point and the cutter-axis as well as the size of a cylindrical ball-end mill, interference between the desired surface and the cutter can be detected. As a result, admissible cutter rotating and tilting ranges can then be determined. Therefore, the CL data required in the machining process can be generated through rotating and tilting the workpiece. The steps needed to fulfill the approach will be outlined based on a popular type of five-axis machining shown in Fig. 1. By the use of this method, practical application examples for machining compressor impellers by the type of a five-axis machining center shown in Fig. 1 are given to illustrate its effectiveness and reliability.

Background

As shown in Fig. 2, during the machining process, the cutter is guided to touch a prescribed CC point, $\mathbf{r}$, on the surface of a workpiece. With an offset equal to the radius, R, of the cylindrical ball-end cutter, the offset point of $\mathbf{r}$ can be denoted as $\mathbf{p}=\mathbf{r}+R\mathbf{n}$. $\mathbf{n}$ is the unit normal vector at point $\mathbf{r}$, and it can be easily evaluated for a general parametric surface [18]. Corresponding to $\mathbf{r}$, the CL data, $\mathbf{L}=(\mathbf{b},\mathbf{u})$, can be expressed by the position of the tip of the cutter, $\mathbf{b}$, and the unit vector of the cutter-axis. Note that $\mathbf{b}$ can be written as

$$\mathbf{b}=\mathbf{r}+R(\mathbf{n}-\mathbf{u}) \tag{1}$$

The global interference problem is defined as any portion of the desired geometry of a workpiece is within the flank of a cylindrical ball-end cutter. To overcome this difficulty, by rotating and

[1]Correspondent—Tel: 886-7-525-4229; Fax: 886-7-525-4299; Email: dermin@mail.nsysu.edu.tw

Contributed by the Gas Turbine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Gas Turbine Division January 1, 1999; final revision received by the ASME Headquarters September 18, 1999. Technical Editor: H. Nelson.

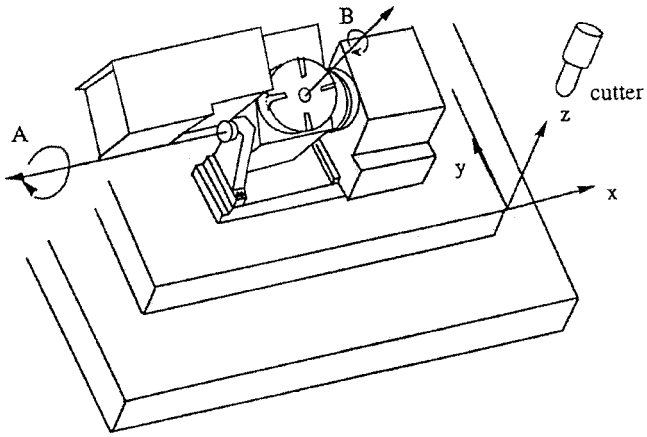


Fig. 1 Schematic illustration of a type of five-axis machining

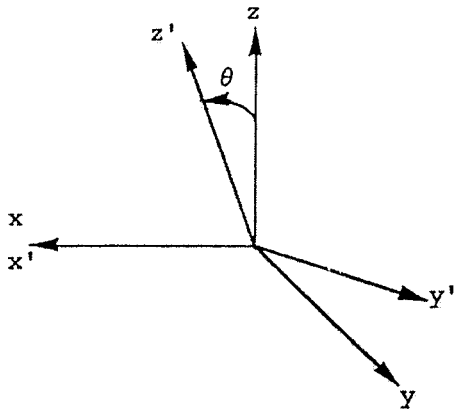


Fig. 3 Coordinate systems

tilting the rotary table of the machining center plotted in Fig. 1, the direction of the tool axis can be relatively adjusted so that a new vector of **u** needed in Eq. (1) can be evaluated. As a result, a CL data without interference may be found from Eq. (1).

In order to find an admissible direction of the cutter-axis without interference, based on the configuration of a five-axis machine tool illustrated in Fig. 1, the coordinates of a CC point on the workpiece is first rotated about the *z*-axis for an angle of ϕ and then tilted about the *x*-axis for an angle of θ in the *x*-*y*-*z* coordinate system shown in Fig. 3. In such a sequence, the involved process of calculations for finding an admissible direction of the cutter-axis will be more concise. This process of transformation can be written in a matrix form as

$$[R_{zx}]=[R_z\phi][R_x\theta] \tag{2}$$

However, the result above is exactly identical to that obtained by tilting the workpiece for an angle of θ about the *x*-axis and then rotating it about the z'-axis in the x'-y'-z' coordinate system illustrated in Fig. 3 for an angle of ϕ. In operation of the rotary table plotted in Fig. 1, note that the rotation (*B*-axis) is performed about the z'-axis whose direction is changed due to a tilt of θ about the *x*-axis. The transformation matrix of this step can be expressed as

$$[R_{xz'}]=[R_x\theta][R_{z'\phi}] \tag{3}$$

The identity between Eqs. (2) and (3) is established in the Appendix.

Determination of Admissible Cutter Orientation

As described earlier, the task of machining turbomachinery components with complex geometries can be conveniently accomplished by using five-axis machine tools especially for the global interference problem that must be overcome. For producing interference-free tool paths, the center of the ball-end of a cutter with a specified orientation is placed at the offset point of a CC point with an offset equal to the radius of the cutter. Then, as depicted in Fig. 4, the CC point is checked for interference between the workpiece and the cutter. To create interference-free cutter paths by adjusting tilting and rotating axes of a five-axis machine tool, all the schemed CC points must be outside of the flank of the used cutter.

Chordal Deviation and Scallop Height. Since the limitation of the shape of tool cutters and the characteristic of the machined surface, it is difficult to avoid incurring the state of surface roughness in point milling. The criterion to determine the number of CC points is based on the surface roughness. For this concern, as shown in Figs. 5 and 6, two kinds of surface roughness, the scallop height and the chordal deviation, are taken into account. Shown in Fig. 6, $\mathbf{r}(u)$ is the cutter contact point on an isoparametric tool path in the primary direction (u) of the machined sur-

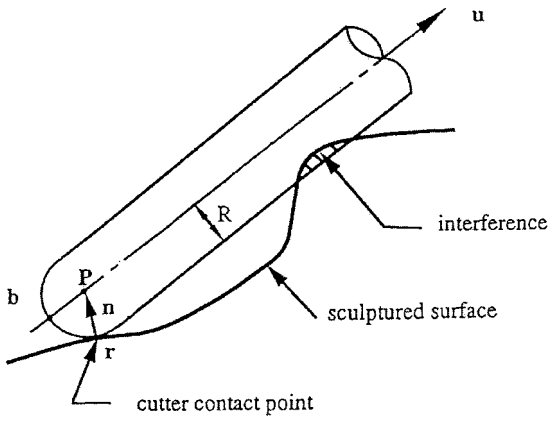


Fig. 2 Schematic illustration of interference and cutter location

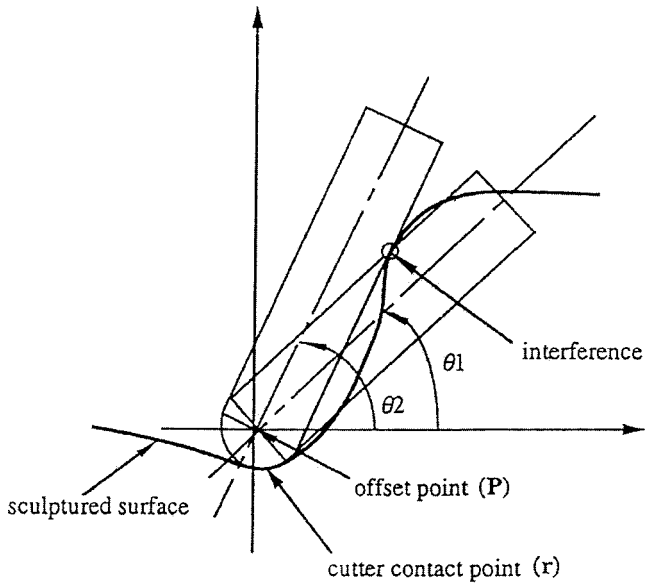


Fig. 4 Schematic illustration of rotational range of a cutter and interference

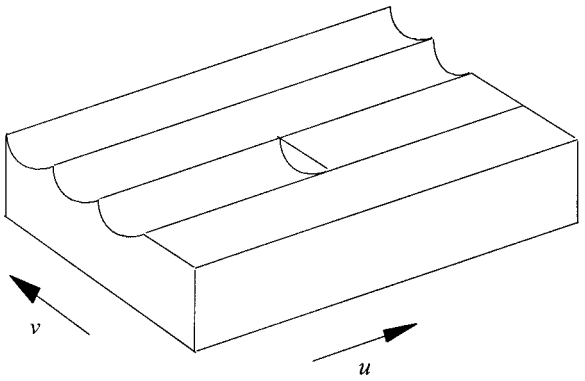


Fig. 5 Compositions of surface roughness

face. The chord bounded from $\mathbf{r}(0)$ to $\mathbf{r}(1)$ is denoted by vector $\mathbf{c}$. The perpendicular vector $\mathbf{d}$ from the chord to point $\mathbf{r}(u)$ of the curve can be written as

$$\mathbf{d}=\mathbf{r}(u)-\mathbf{r}(0)-\lambda\mathbf{c} \tag{4}$$

$$\text{where } \lambda=\frac{[\mathbf{r}(u)-\mathbf{r}(0)]\cdot\mathbf{c}}{|\mathbf{c}|^2}$$

As shown in Fig. 6, the maximum chordal deviation occurs when the tangent vector $\dot{\mathbf{r}}$ is perpendicular to $\mathbf{d}$, i.e.,

$$\dot{\mathbf{r}}\cdot\mathbf{d}=0, \tag{5}$$

and it is a true maximum when

$$\ddot{\mathbf{r}}\cdot\mathbf{d}<0 \tag{6}$$

Under the maximum tolerant chordal deviation, it is possible to determine the isoparametric step length in the primary direction (u) from Eqs. (4) to (6).

The other view of the surface roughness is along the cross direction (v) of the machined surface, i.e., the scallop height between two adjacent tool paths of a ball nose cutter. Both of the convex portion and the concave portion must be concerned. Figures 7 and 8 illustrate the geometric relationships for cases of the convex surface and the concave surface, respectively. Once the radius of the ball nose cutter R is given, the step length s in the convex portion can be found as

$$s=2\left\{R^2-\left[\frac{2R\rho-\delta^2-2\delta\rho}{2(\delta+\rho)}\right]^2\right\}^{1/2} \tag{7}$$

where δ is the scallop height and ρ is the radius of curvature of the curve. Since the radius of curvature does not vary too much in a small region, we can assume that the radius of curvature between two adjacent touches are approximately equal. In a similar way, the step length in the concave portion can also be obtained as

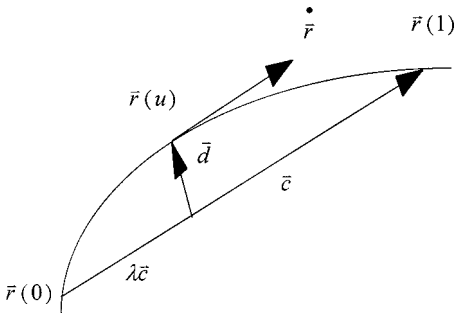


Fig. 6 Chordal deviation

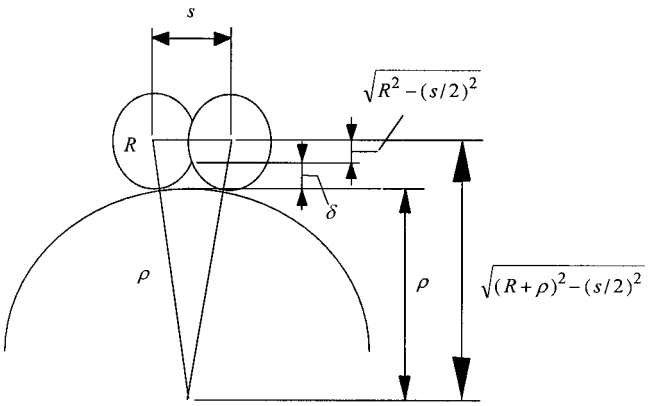


Fig. 7 Scallop height of the convex surface

$$s=2\left\{R^2-\left[\frac{2\delta\rho-2R\rho-\delta^2}{2(\delta-\rho)}\right]^2\right\}^{1/2} \tag{8}$$

Once the maximum tolerant scallop height has been determined, from Eqs. (7) and (8), the proper step length in the cross direction can be decided.

According to the analyses of the chordal deviation and the scallop height above, the CC points on the surface of a turbomachinery part are determined from the step lengths in both parametric directions. Therefore, CC points could be given under required accuracy. Moreover, the cutting speed which enormously influences the surface quality of a part is proportional to the active radius of the cutter. It is intended to avoid utilizing the center of the cutter for no effective milling [19] since it will usually leave tool marks on the part surface. Hence, we should also adjust the direction of the axis of the cutter that is not collinear to the direction of the normal of a CC point while determining rotating and tilting angles for global interference.

Algorithm. Given CC points and based on the projected distance between the CC point and the axis of a cylindrical cutter to detect interference, the algorithm to generate interference-free tool paths can be described in a step-wise fashion in the following outlines:

(1) Select an appropriate size of a cylindrical cutter with a ball-end, considering the geometric features of a workpiece;

(2) Calculate the offset point mentioned earlier for each CC point. And, then position the center of the ball-end of the cutter at the offset point by moving the three translation axes of the five-axis machine tool;

(3) Choose a proper angle of ϕ for turning the B-axis of the machine tool indicated in Fig. 1 by considering the selected cutter

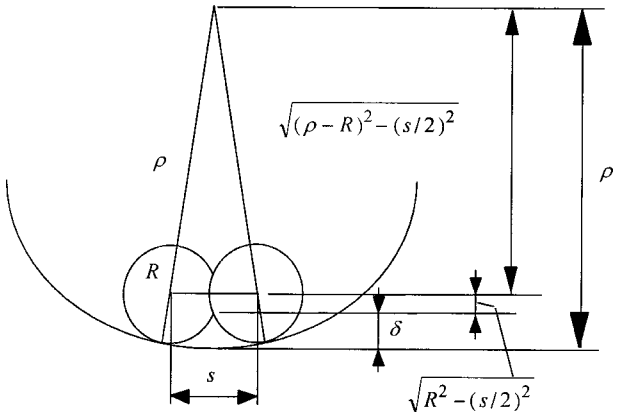


Fig. 8 Scallop height of the concave surface

size and the geometric features of a particular workpiece. For example, the spacing between two neighboring blades of a centrifugal impeller;

(4) Calculate the angular range of θ for the *A*-axis shown in Fig. 1, based on the angle of ϕ determined in the step above;

(4.1) Compute the possible angular range of a CC point by rotating the cutter (flank) about the offset point to touch all other CC points that are assumed to be interference points illustrated in Fig. 4;

(4.2) Find the intersection of angular ranges determined above. The common range will constitute the possible tilting range for the single CC point;

(4.3) Decide a tilting angle by avoiding the direction of the axis of the ball-end cutter being collinear to that of the CC point normal;

(4.4) Generate the CL data for the CC point based on the angles of ϕ and θ;

(5) Repeat steps of (4.1) through (4.4) for all CC points on the schemed tool path.

The CL data obtained above are applicable to a simultaneous five-axis machine tool without incurring collision. If the motion control of a five-axis machine tool is not simultaneous for its rotary table, one more step after step (5) above is needed to determine the intersection of range for tilting angles corresponding to the CC points along a schemed tool path. Then, a specific tilting angle decided from the intersection of the tilting ranges for a single tool path can be used with such a machine tool.

Two Types of Five-Axis Machine Tools. The procedure described above can be easily implemented to obtain the admissible direction of the cutter-axis without incurring interference by relatively rotating and tilting the workpiece. The calculations and solution required in step (4.1) above for a CC point will be illustrated in detail below for two types of five-axis machine tools.

Referring to Fig. 1, for a point, (x,y,z), in space, we select an angle of ϕ for rotation about the *z*-axis and this transformation can be expressed as

$$(x',y',z')=(x,y,z)\begin{bmatrix}\cos\phi & \sin\phi & 0\\ -\sin\phi & \cos\phi & 0\\ 0 & 0 & 1\end{bmatrix}$$

$$=(x\cos\phi-y\sin\phi, x\sin\phi+y\cos\phi, z) \tag{9}$$

Then, the transformation of the point in Eq. (9) tilted about the *x*-axis for an angle of θ can be written as

$$(x'',y'',z'')=(x',y',z')\begin{bmatrix}1 & 0 & 0\\ 0 & \cos\theta & \sin\theta\\ 0 & -\sin\theta & \cos\theta\end{bmatrix}$$

$$=(x',y'\cos\theta-z'\sin\theta, y'\sin\theta+z'\cos\theta) \tag{10}$$

From the result of the coordinate transformation given in Eq. (10), if a CC point just touches the flank of a cutter, then the projected distance from the CC point to the offset point, i.e., center of the cylindrical ball-end of the cutter, is equal to the radius of the mill and can be shown as

$$(x''_{cc}-x''_{of})^2+(y''_{cc}-y''_{of})^2=R^2 \tag{11}$$

where subscripts, *cc* and *of*, respectively denote coordinates of the CC point and the offset point. Equation (11) can be further described as

$$[(y'_{cc}-y'_{of})\cos\theta-(z'_{cc}-z'_{of})\sin\theta]^2=R^2-(x'_{cc}-x'_{of})^2 \tag{12}$$

Substituting $b=y'_{cc}-y'_{of}$, $c=z'_{cc}-z'_{of}$, and $a=R^2-(x'_{cc}-x'_{of})$ into Eq. (12) yields

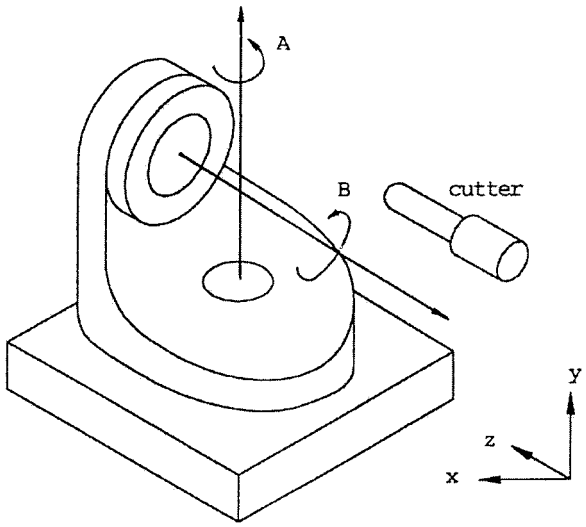


Fig. 9 Schematic illustration of a different type of five-axis machining

$$(b\cos\theta-c\sin\theta)^2=a \tag{13}$$

Using trigonometric relationships, we can also write Eq. (13) as

$$b^2\left(\frac{1+\cos 2\theta}{2}\right)+c^2\left(\frac{1-\cos 2\theta}{2}\right)-2bc\cos\theta\sin\theta=a \tag{14}$$

Moreover, Eq. (14) can be represented as

$$e\cos 2\theta-f\sin 2\theta=g \tag{15}$$

where $e=(b^2-c^2)/2$, $f=bc$, and $g=a-[(b^2+c^2)/2]$.

Letting $\tan\theta=t$ and substituting $\sin 2\theta=2t/(1+t^2)$ as well as $\cos 2\theta=(1-t^2)/(1+t^2)$ into Eq. (15), we can derive the solution of Eq. (15) as

$$t_{1,2}=\frac{-f\pm\sqrt{f^2-(g^2-e^2)}}{(g+e)} \tag{16}$$

or

$$\theta_{1,2}=\tan^{-1}(t_{1,2}) \tag{17}$$

Note that Eq. (17) has two roots for the admissible tilting angle of a CC point that is illustrated in Fig. 4. As described earlier, tilting angles for other CC points can be solved in a similar way.

Based on a different type of five-axis machining shown in Fig. 9, the derivation of the tilting angle for a CC point is given here. Comparing the configuration of this machine tool to that shown in Fig. 1, we can see that the *A*-axis of this machine is tilted about the *y*-axis and the *B*-axis is rotated about the *z*-axis. Therefore, one can modify Eq. (10) as

$$(x'',y',z'')=(x',y',z')\begin{bmatrix}\cos\theta & 0 & -\sin\theta\\ 0 & 1 & 0\\ \sin\theta & 0 & \cos\theta\end{bmatrix}$$

$$=(x'\cos\theta+z'\sin\theta, y', -x'\sin\theta+z'\cos\theta) \tag{18}$$

In a similar manner, if a CC point just touches the flank of the cutter, then the projected distance, Eq. (11), from the CC point to the offset point can be written as

$$[(x'_{cc}\cos\theta+z'_{cc}\sin\theta)-(x'_{of}\cos\theta+z'_{of}\sin\theta)]^2+(y'_{cc}-y'_{of})^2$$
$$=R^2 \tag{19}$$

Letting $b=x'_{cc}-x'_{of}$, $c=z'_{cc}-z'_{of}$, and $a=R^2-(y'_{cc}-y'_{of})^2$, one can express Eq. (19) as

$$(b\cos\theta - c\sin\theta)^2 = a \tag{20}$$

It is obvious that the formula above has the same form as that of Eq. (13). Hence, for a five-axis machine tool of this type, the same solution given in Eq. (20) can be again employed in the process of finding the admissible tilting range for a CC point.

Application Examples

To verify the usefulness and reliability of the approach described above, two application examples are illustrated to show five-axis machining of centrifugal compressor impellers. In these two cases, one has 13 blades and the second is with 15 blades. To machine the impellers with various numbers of blades, based on the developed algorithm introduced earlier, an interactive computer code in C language is written and run on a personal computer to implement the procedure to generate collision-free tool paths.

Before performing the algorithm here, the basic definitions about the geometry of impellers are first to be introduced. The geometry of an impeller is basically characterized by its blade and hub surfaces. Here represented by ruled surfaces [20], the blades are defined by camber surfaces [21]. A camber surface is used to form the pressure side surface and the suction side surface by its offset surfaces with an offset from both sides. As indicated in Fig. 10, a camber surface is a ruled surface constructed by linearly interpolating between two known boundary curves, the hub camberline (curve A-E) and the shroud camberline (curve M-Q). Therefore, the pressure side and suction side surfaces are also ruled surfaces. By referring to Fig. 11, the hub surface is generated by rotating the hub camberline with respect to the axis of the impeller. In addition, a constant-radius blending technique [22] is applied to construct the blends between the blade and the hub regions for pressure and suction sides.

In the first application example, the dimensions of the blank (aluminum alloy 7075) for the impeller having 15 blades are 118 mm and 36 mm for its diameter and height, respectively. The minimal gap between two adjacent blades of this 15 blade impeller is about 9 mm. The diameter of the chosen cylindrical ball-end cutter is 4 mm. Also, two constant-radius blending surfaces are constructed based on the cutter size. One blending surface is built between the hub surface and the pressure surface. And, the second blending surface is generated between the hub surface and the suction surface.

With the consideration of the gap between two neighboring blades of the impeller and the size of the mill, the angle of $\phi = 12$ deg is selected for calculations. For checking collision, the numbers of CC points in each block between two neighboring blades for the pressure side surface, the suction side surface, and

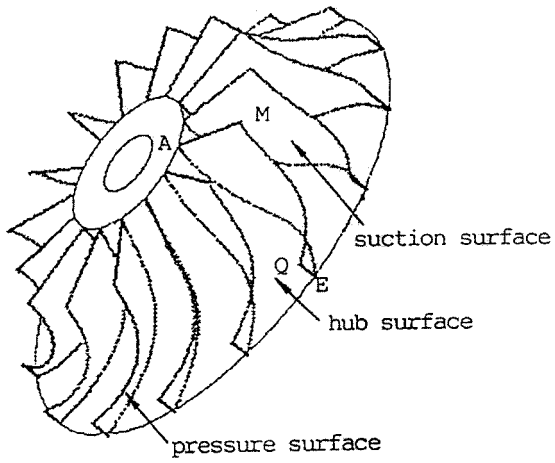


Fig. 11 Schematic illustration of a 15-blade impeller

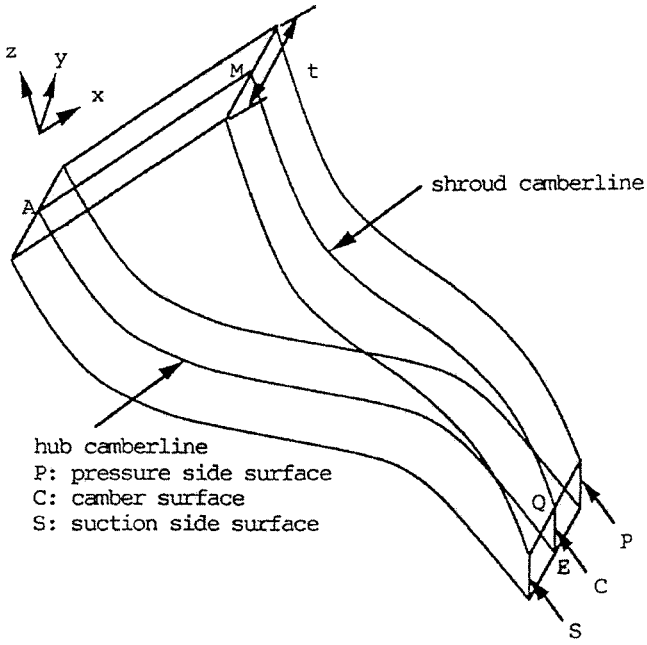


Fig. 10 Schematic illustration of camber, pressure side, and suction side surfaces of an impeller

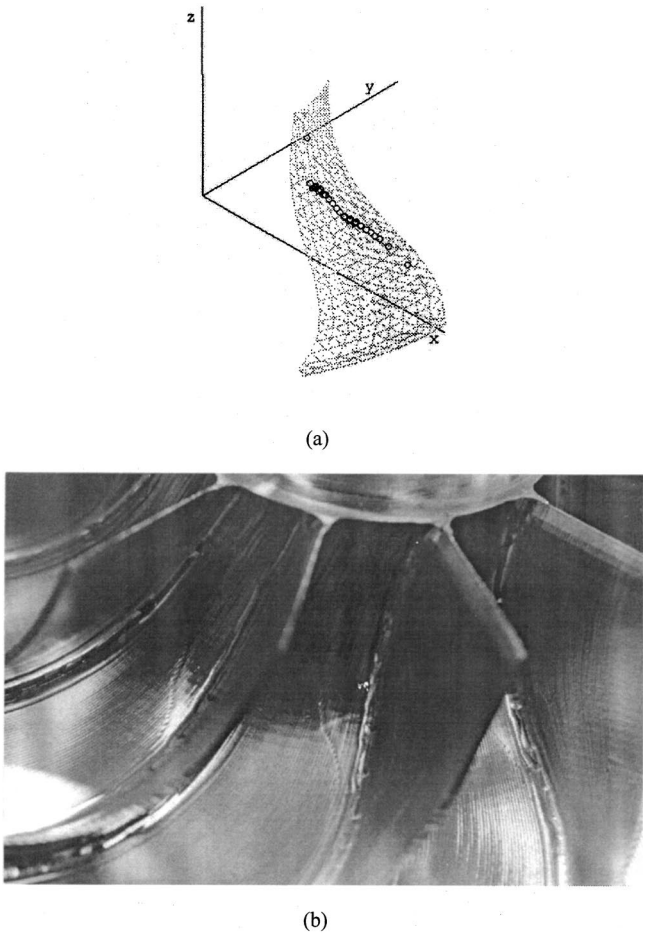
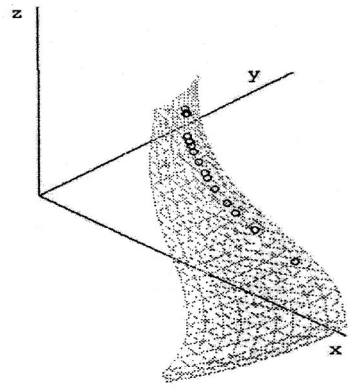
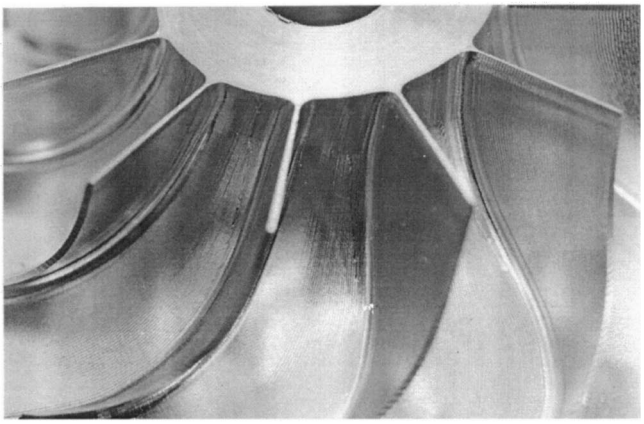


Fig. 12 (*a*) Simulated result for tool marks; and (*b*) tool marks left on the cut hub surface



(a)



(b)

Fig. 13 (*a*) **Simulated result for improved tool marks;** (*b*) **improved tool marks left on the cut hub surface**

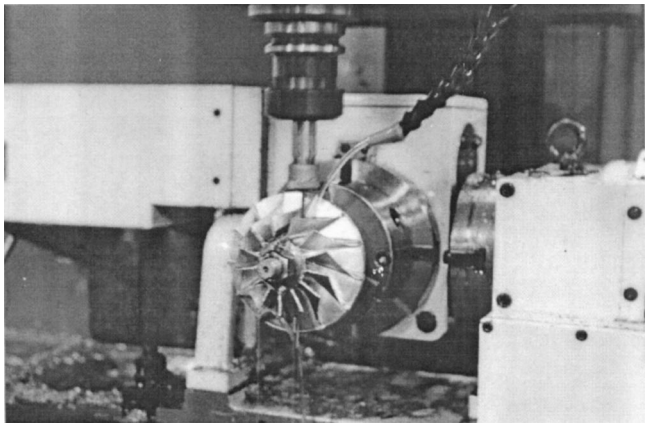


Fig. 15 Five-axis machining of an impeller with 13 blades

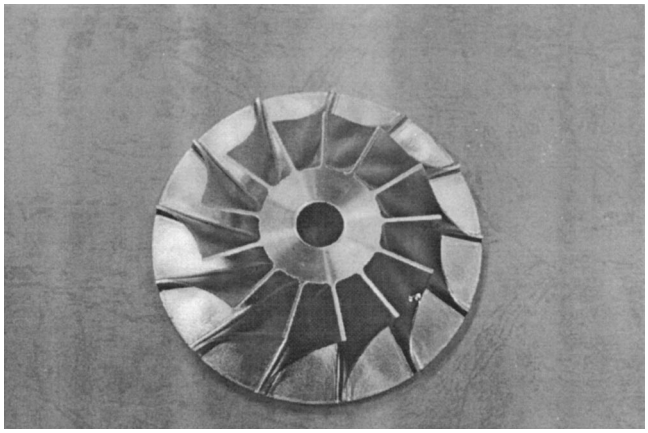


Fig. 16 A machined impeller with 13 blades

the hub surface are determined based on the analyses of the chordal deviation and the scallop height for manufacturing tolerances. The resulting CL data are further postprocessed for employing the type of the five-axis machining shown in Fig. 1 to machine the impeller.

As mentioned earlier, for a CC point, a range for the tilting angles without incurring interference may be identified even though it is limited within a small zone. In order to select a tilting angle for each CC point for cutting the impeller, the orientation of the cutter-axis can further be checked. Since the tip of a cylindrical ball-end cutter has the lowest efficiency in material-removal rates, if possible, the direction of the cutter-axis should not be collinear to that of the CC point normal. Otherwise, as presented in Fig. 12 with the simulated result and the machined impeller, the tip of the cutter will often leave tool marks on the hub surface of the impeller. Improvement for this situation is reflected in Fig. 13 with the simulated result and the cut impeller. The photograph given in Fig. 14 shows the complete geometry of the successfully cut 15-blade impeller.

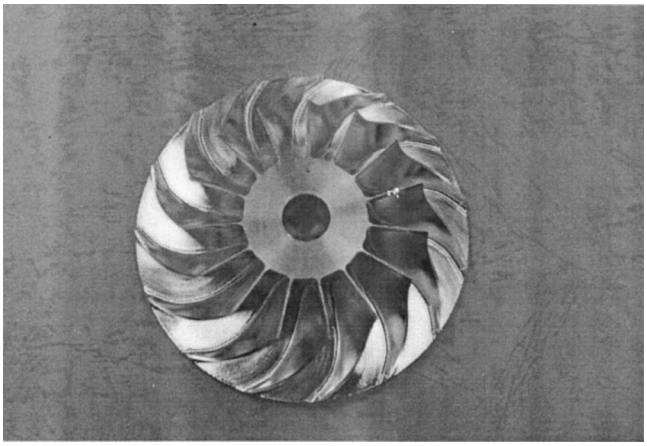


Fig. 14 A machined impeller with 15 blades

In a similar way, Fig. 15 illustrates the five-axis machining of an impeller with 13 blades for the second example. To generate interference-free tool paths, in this application, the used cutter size, material, and the size of the blank are the same as those of the first case. However, the shape of the blades is different. Furthermore, the constant-radius blending technique is again utilized to build blending surfaces. The machined impeller is photographed in Fig. 16.

Discussion and Conclusion

A simple and reliable algorithm that can be applied to generate collision-free tool paths for five-axis machining of turbomachinery parts is developed. Regardless of the type of surfaces on such a workpiece, only CC points are needed in the process of checking with the consideration of the chordal deviation and the scallop height. As has been demonstrated by practical application examples, the feasibility and versatility of the procedure is success-

fully verified. Though it is illustrated by a popular type of a five-axis machining center, the approach to its implementation does not depend on the particular choice of hardware.

In this article, only the type of cylindrical ball-end cutters is handled in the procedure. For machining general turbomachinery components, different shapes of cutters may be also used and it is well worth making the effort to include them in the algorithm. Also the choice of a tilting angle from the feasible range of cutter-axis for maximum material-removal rate is being observed.

Acknowledgments

This study is being supported by the National Science Council of Taiwan under grants of NSC 84-2212-E-110-005 and NSC 85-2622-E-110-007. The authors are grateful to Yi Tay Liu who developed the algorithm to produce geometric data from the preliminary design of impellers.

Appendix

The well-known complete transformation for a general case of rotation about an arbitrary axis in space can be described as (Ref. [20] p. 123)

$$[M]=[T][R_x][R_y][R_\phi][R_y]^{-1}[R_x]^{-1}[T]^{-1} \qquad (A1)$$

where $[T]$, $[R_x]$, and $[R_y]$ respectively denote the translation matrix, the transformation matrix for rotation about the x-axis, and the transformation matrix for rotation about the y-axis. And, $[R_\phi]$ is a z-axis rotation matrix for the rotation about the arbitrary axis. Others in the right-hand side of Eq. ($A1$) are inverse matrices. By referring to Fig. 3, since the z'-axis passes through the origin, the transformation for a rotation about the z'-axis is relatively simple and can be written as

$$[R_{z'}\phi]=[M_{z'}\phi]=[R_x][R_\phi][R_x]^{-1} \qquad (A2)$$

To make the z'-axis coincident with the z-axis, it is necessary to rotate an angle of $-\theta$ about the x-axis and the transformation matrix, $[R_x]$, is

$$[R_x]=[R_x(-\theta)]=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & \cos(-\theta) & \sin(-\theta) & 0\\ 0 & -\sin(-\theta) & \cos(-\theta) & 0\\ 0 & 0 & 0 & 1\end{bmatrix}=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & \cos\theta & \sin\theta & 0\\ 0 & -\sin\theta & \cos\theta & 0\\ 0 & 0 & 0 & 1\end{bmatrix}=[R_x\theta]^{-1} \qquad (A3)$$

The rotation matrix, $[R_\phi]$, is performed for an angle of ϕ about the z-axis, and it can be denoted as $[R_\phi]=[R_z\phi]$. Also, note that

$$[R_x]^{-1}=[R_x(-\theta)]^{-1}=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & \cos(-\theta) & \sin(-\theta) & 0\\ 0 & -\sin(-\theta) & \cos(-\theta) & 0\\ 0 & 0 & 0 & 1\end{bmatrix}^{-1}=\begin{bmatrix}1 & 0 & 0 & 0\\ 0 & \cos\theta & \sin\theta & 0\\ 0 & -\sin\theta & \cos\theta & 0\\ 0 & 0 & 0 & 1\end{bmatrix}=[R_x\theta] \qquad (A4)$$

Therefore, Eq. ($A2$) can be expressed as

$$[R_{z'}\phi]=[R_x(-\theta)][R_z\phi][R_x(-\theta)]^{-1}=[R_x\theta]^{-1}[R_z\phi][R_x\theta] \qquad (A5)$$

From Eq. ($A5$), one can see that the result of Eq. (2) coincides with that of Eq. (3) because

$$[R_{xz'}]=[R_x\theta][R_{z'}\phi]=[R_x\theta][R_x\theta]^{-1}[R_z\phi][R_x\theta]=[R_z\phi][R_x\theta]=[R_{zx}] \qquad (A6)$$

References

[1] Kim, B. H., and Chu, C. N., 1994, ''Effect of Cutter Mark on Surface Roughness and Scallop Height in Sculptured Surface Machining,'' Comput.-Aided Des., **26**, pp. 179–188.

[2] Suresh, K., and Yang, D. C. H., 1994, ''Constant Scallop-height Machining of Free-form Surfaces,'' Trans. ASME J. Eng. Ind., **116**, pp. 253–259.

[3] Oliver, J. H., Wysocki, D. A., and Goodman, E. D., 1993, ''Gouge Detection Algorithms for Sculptured Surface NC Generation,'' Trans. ASME J. Eng. Ind., **115**, pp. 139–144.

[4] Loney, G. C., and Ozsoy, T. M., 1987, ''NC Machining of Free Form Surfaces,'' Comput.-Aided Des., **19**, pp. 85–90.

[5] Vickers, G. W., and Quan, K. W., 1989, ''Ball-Mills Versus End-Mills for Curved Surface Machining,'' Trans. ASME J. Eng. Ind., **111**, pp. 22–26.

[6] Choi, B. K., and Jun, C. S., 1989, ''Ball-end Cutter Interference Avoidance in NC Machining of Sculptured Surfaces,'' Comput.-Aided Des., **21**, pp. 371–378.

[7] Hwang, J. S., 1992, ''Interference-free Tool-path Generation in the NC Machining of Parametric Compound Surfaces,'' Comput.-Aided Des., **24**, pp. 667–676.

[8] Choi, B. K., Park, J. W., and Jun, C. S., 1993, ''Cutter-location Data Optimization in 5-axis Surface Machining,'' Comput.-Aided Des., **25**, pp. 377–386.

[9] Wu, C. Y., 1995, ''Arbitrary Surface Flank Milling of Fan, Compressor, and Impeller Blades,'' Trans. ASME J. Eng. Gas Turbines Power, **117**, pp. 534–539.

[10] Elber, G., and Fish, R., 1997, ''5-Axis Freeform Surface Milling Using Piecewise Ruled Surface Approximation,'' Trans. ASME J. Manuf. Sci. Eng., **119**, pp. 383–387.

[11] Wang, W. P., 1990, ''Three-Dimensional Collision Avoidance in Production Automation,'' Comput. Ind., **15**, pp. 169–174.

[12] Tseng, Y., and Joshi, S., 1991, ''Determining Feasible Tool-approach Directions for Machining Bezier Curves and Surfaces,'' Comput.-Aided Des., **23**, pp. 36–379.

[13] Takeuchi, Y., and Idemura, T., 1991, ''5-Axis Control Machining and Grinding Based on Solid Model,'' Ann. CIRP, **40/1**, pp. 455–458.

[14] Takeuchi, Y., and Watanabe, T., 1992, ''Generation of 5-Axis Control Collision-Free Tool Path and Postprecessing for NC Data,'' Ann. CIRP, **41/1**, pp. 539–542.

[15] Lee, Y. S., 1993, Automatic Planning and Programming for Five-axis Sculptured Surface Machining, Ph.D. thesis, Purdue University, West Lafayette, IN.

[16] Li, S. X., and Jerard, R. B., 1994, ''5-axis Machining of Sculptured Surfaces with a Flat-end Cutter,'' Comput.-Aided Des., **26**, pp. 165–178.

[17] Jerard, B., Drysdale, R. L., and Magewick, J., 1989, ''Methods for Detecting Errors in Numerically Controlled Machining of Sculptured Surfaces,'' IEEE Comput. Graphics Appl., Jan., pp. 26–39.

[18] Faux, I. D., and Pratt, M. J., 1979, *Computational Geometry for Design and Manufacture*, Ellis Horwood, Chichester, UK.

[19] Schulz, H., 1995, ''High-Speed Milling of Dies and Moulds-Cutting Conditions and Technology,'' Ann. CIRP, **44**, pp. 35–38.

[20] Rogers, D. F., and Adams, J. A., 1990, *Mathematical Elements for Computer Graphics*, 2nd Ed., McGraw-Hill, New York.

[21] Smith, D. J. L., and Merryweather, H., 1973, ''The Use of Analytic Surfaces for the Design of Centrifugal Impellers by Computer Graphics,'' Int. J. Numer. Methods Eng., **7**, pp. 137–154.

[22] Choi, B. K., 1991, *Surface Modeling for CAD/CAM*, Elsevier, Amsterdam.

J. A. Brown
UMIST/E A Technology,
Postgraduate Training Partnership,
Capenhurst, Chester,
CH1 6ES, U.K.

R. Freer
Materials Science Centre,
UMIST, Manchester,
M1 7HS, U.K.

A. T. Rowley
E A Technology Limited,
Capenhurst, Chester,
CH1 6ES, U.K.

Reconditioning of Gas Turbine Components by Heat Treatment

At the high operating temperatures experienced within a gas turbine, creep damage is a major life-limiting factor. This is especially true for components which are highly stressed and closest to the hot gas inlet end of the machine, such as the rotating blades in this area. The ability to recover the creep properties of used gas turbine components might enable their service lives to be increased considerably. Thereby, maintenance costs could be reduced. Previous research into the feasibility of reconditioning crept superalloys by heat treatment will be reviewed and perceived limitations identified. The current work will add to knowledge on the effects of conventional recovery techniques and will explore potential new heat treatment regimes. The experimental and analytical methods to be used will be both described and supported by preliminary results.
[DOI: 10.1115/1.1287593]

Introduction

The hot gas path components of a gas turbine predominantly comprise nickel-based superalloys. They operate under temperature and stress conditions which will result in creep damage over time. Probably the most costly manifestation of the problem is the need for regular replacement of the stages of rotating blades which are located closest to the hot gas inlet end of the machine. In a large power generating gas turbine, first stage blades can require replacement after only 50,000 equivalent operating hours (EOH). These blades are expensive, primarily due to their complex fabrication route (they are usually of the directionally solidified or single crystal type). Hence, there is considerable incentive to increase their service lives.

The potential to recover the creep properties of pre-crept superalloys has been investigated by numerous authors during the last four decades [1–13]. Recovery of pre-crept specimens has been reported when utilizing heat treatment regimes such as

(a) simply holding unloaded Nimonic 80A specimens at the creep temperature for around 12 hours to effect cavity sintering [1]
(b) repetition of the original heat treatment regime in order to restore the initial microstructure [6]
(c) hot isostatic pressing (HIPing) of IN738 followed by reheat treatment to both sinter cavities and effect microstructural modification [11]

The observed recovery has been attributed to

- changes in dislocation structure
- movement of grain boundaries (via grain growth) such that cavities are removed from them and, hence, cannot continue to grow via grain boundary diffusion
- microstructural modification (via re-solution/re-precipitation)
- changes in cavity shape and their sintering (by vacancy diffusion under the action of surface tension and/or applied pressure)

In general, however, previous studies have been based upon a small sample batch size and have predominantly assessed recovery in terms of creep parameters; by subjecting pre-crept specimens to recovery heat treatments and repeating the creep test. This is of concern given the normal scatter of results expected on creep testing and may be illustrated by reference to Fig. 1 of the current work. Here eight specimens of Nimonic 80A have been tested to rupture at a temperature of 750°C under the action of a constant load which was equivalent to a tensile stress of 263 N/mm^2 at the start of each test. Specimen X exhibits over twice the time to failure of Specimen Y. Had the former been fortuitously selected for recovery heat treatment a quite significant, yet false, life extension would have been observed. The majority of studies have also failed to incorporate control specimens, where a pre-crept specimen would be subject to repetition of the creep test without seeing any recovery heat treatment. Therefore, any influence of the unloading cycle on apparent recovery has not been assessed. Changes in density are clearly a means of indicating void sintering on heat treatment and, therefore, could add weight to other evidence of recovery. Whilst the technique has been used in evaluating the ability of HIPing to sinter creep cavities [14], assessment of recovery by creep testing was not featured in the same study. It does not appear that density measurements accompanied any of the studies featuring recovery of superalloys by conventional heat treatment.

Clearly, in view of the potential benefits associated with the ability to recover the creep properties of gas turbine components, it is important to investigate the phenomena more thoroughly. By testing a number of specimens to rupture and by producing numerous control samples, in addition to multiple specimens recovered by each heat treatment regime, this work aims to add to the body knowledge on recovery by conventional heat treatment. Subsequently, recovery will also be assessed in terms of density change. Further, an alternative recovery technique which utilizes microwaves in addition to conventional heating will be investigated. It is considered that the electric fields associated with microwaves can enhance sintering rates by influencing the diffusion of vacancies during thermal/microwave heating [15]. Furthermore there is some evidence that the application of magnetic fields, which are a component of microwaves [16], can reduce defect density and affect stress relaxation in a range of materials [17,18]. If the ability of microwaves to sinter cavities is established, then the technique might be utilized to avoid the high pressures associated with HIPing. Nimonic 80A was selected for the study on the basis of its microstructural simplicity relative to other superalloys and previous work [1,2] indicating that its recovery could be affected by a single stage heat treatment at only 750°C for around 12 h.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-444. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters May 15, 2000. Associate Technical Editor: D. Wisler.

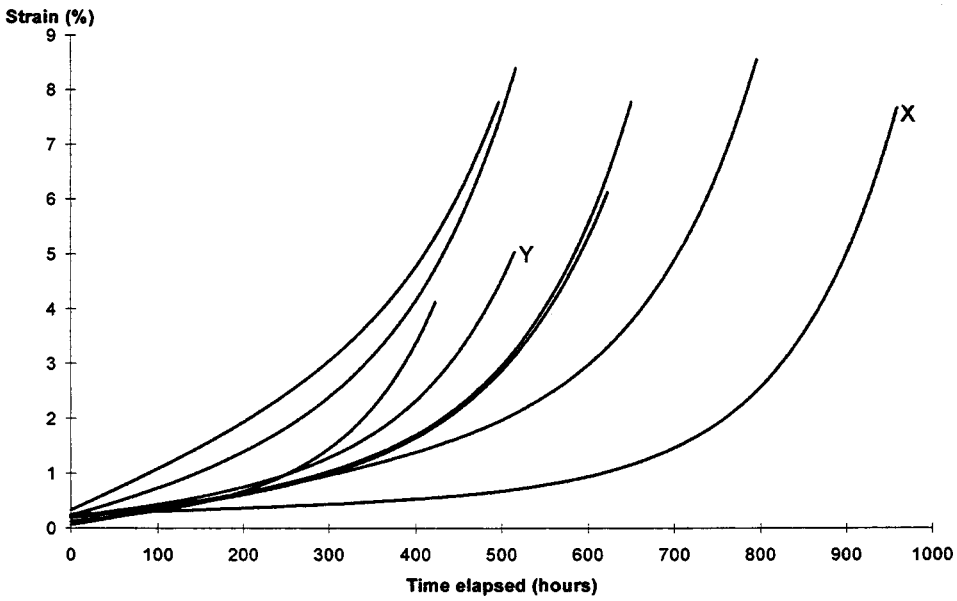


Fig. 1 Creep curves to rupture for Nimonic 80A (tested at 750°C, constant load and an initial stress of 263 N/mm^2)

Experimental

Sample Preparation. Solution treated Nimonic 80A bar was sectioned and machined to produce creep test specimens of 50 mm gauge length and 6.4 mm in diameter. The specimens were then subjected to a multi-stage heat treatment under argon to develop optimum creep resistant properties. The heat treatment regime comprised the following:

(i) heating to 1200°C and holding for 8 h
(ii) cooling to 1150°C and holding for 3 h
(iii) cooling to 1080°C and holding for 24 h and
(iv) cooling to ambient temperature
(v) heating to 820°C and holding for 18 h

Creep Testing. All creep tests were performed under constant load conditions and at a temperature of 750°C. The applied load was calculated to give an initial stress of 263 N/mm^2 in all cases (being based upon specimen dimensions prior to any creep testing). Gauge length extensions were monitored every 10 min and logged by a computer via extensometers clamped to the specimen and linear displacement transducers mounted in apertures within the extensometer legs. The temperature distribution along the specimen was monitored by a pair of thermocouples attached at either end of the gauge length.

Whilst testing to rupture and generating pre-crept specimens, the temperature was allowed to stabilize for a minimum of 2 h prior to loading and was controlled to within +/−5°C of the selected test temperature. Data logging commenced at zero load, such that the elastic extension on loading was recorded as well as creep strain with time. The tests were terminated at a pre-defined strain value by cooling under load. A pre-strain of 1.5 percent was initially selected as being representative of beyond the normal

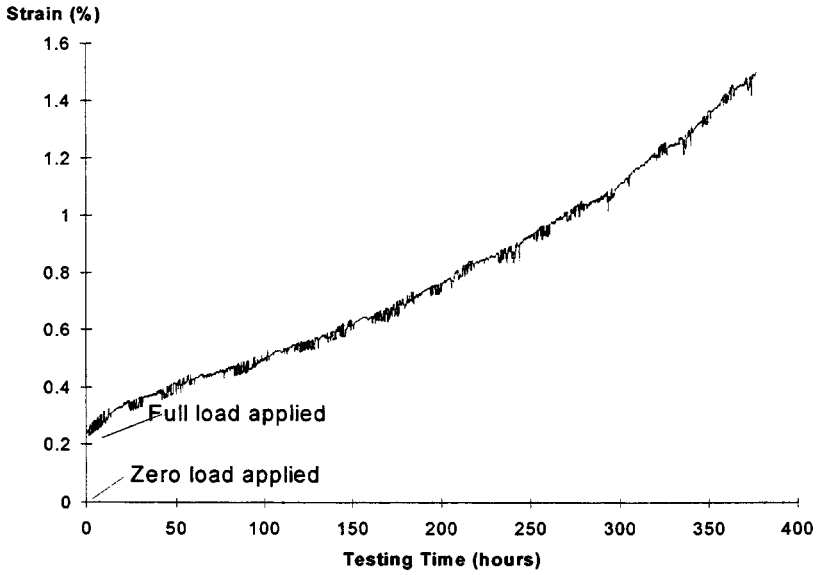


Fig. 2 Typical data logged on creep testing to 1.5 percent strain, where ε_o = strain at full load minus strain at zero load

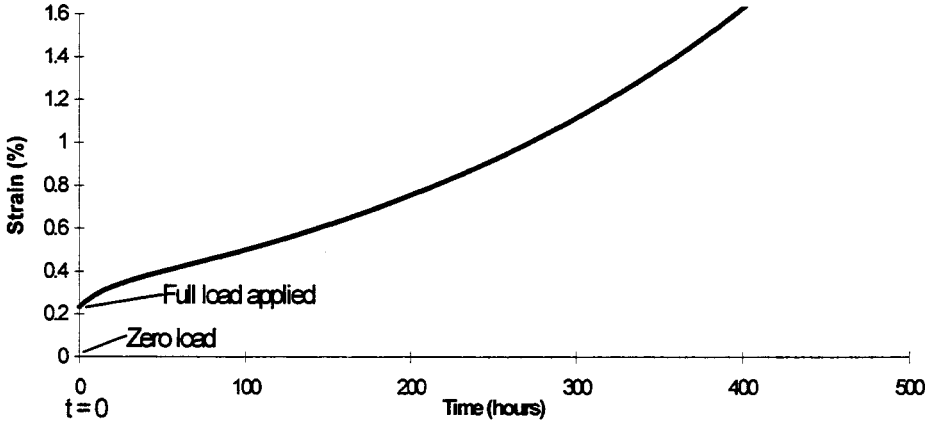


Fig. 3 Data manipulation in spreadsheet, where ε_o = strain at full load minus strain at zero load and occurs at $t = 0$

service life for a gas turbine blade—this comprised both elastic and creep strain components as shown from typical as-logged data in Fig. 2. Pre-crept specimens were then either returned to creep testing as controls or subject to a reconditioning heat treatment before re-testing. Thus far, reconditioning has been attempted by soaking at the creep temperature for 750 min, both with and without the application of a microwave field. In the former case, the thermal heat input was automatically adjusted to compensate for any microwave heating. The second stage of creep testing was affected by heating to the creep temperature as quickly as possible and loading the specimen when the temperature was within 5°C of the target. This stage of testing was terminated by specimen rupture.

Analysis. The logged data was entered into a spreadsheet such that plots of Strain (percent) versus Time (hours) were generated. Time $t = 0$ was defined as being at the end of the loading period such that the intercept with the y axis represented the elastic strain, as illustrated in Fig. 3. It was a straightforward operation to determine both the total strain and time to rupture for specimens taken to destruction in a single test. Both control and recovered specimens were initially analyzed by adding the time and creep strain during the second stage of the creep test to the respective figures from the pre-straining portion of the test. Hence, a measure of total strain and time to rupture was obtained for all interrupted creep tests. The minimum creep rate for each test was determined by fitting the θ-projection equation [13,19] (Eq. 1), modified to incorporate the elastic strain via Eq. 2, over the first 1.5 percent total strain of each curve and assessing the resulting gradients using Eq. 3. The θ-projection equation is claimed to be capable of predicting long term creep behavior from relatively few short term constant stress tests. In the present study it has merely been used for curve fitting, with the change in stress acting on the test piece due to its reduction in area being regarded as insignificant up to 1.5 percent strain. The values of minimum

Table 1 Summary of creep test results

Description of Test	Sample No	Minimum Creep Rate (1/ hrs)	Time to Min Creep Rate (hrs)	Total Strain at Rupture (%)	Total Time to Rupture (hrs)	Statistics
To rupture	3	0.00162333	60	7.68	650	Mean=676
	12	0.0002912	123	8.26	960	Std Dev=176
	14	0.00142899	68	6.5	633	
	15	0.00343544	19	9.32	517	
	17	0.00470513	16	8.28	497	
	25	0.00138785	46	9.45	796	
Control	52	0.00037607	49	4.44	1285	Mean=863
	53	0.00196009	63	7.99	759	Std Dev=259
	57	0.00088491	53	7.89	828	
	58	0.00118659	23	7.23	855	
	59	0.00165362	44	7.44	587	
Recovered at 750 °C for 750 minutes	21	0.00356094	53	7.65	600	Mean=707
	32	0.00226413	43	8.53	561	Std Dev=166
	42	0.0010612	19	6.79	738	
	45	0.00101508	121	7.28	929	
Recovered at 750 °C with of microwaves for 750 minutes	18	0.00262628	32	5.19	550	Mean=605
	20	0.00400177	29	9.7	641	Std Dev=47
	22	0.00298951	51	8.56	584	
	31	0.00155136	11	6.62	609	
	33	0.00227844	26	7.75	570	
	41	0.00157882	58	8.55	676	

creep rate and time to rupture for specimens tested directly to rupture were then utilized to determine a generalized Monkman–Grant relationship [20] via Eq. 4. The results of control tests were also analyzed using the same technique and data generated from creep tests on reconditioned specimens was assessed relative to the rupture and control data obtained.

Results and Discussion

The results of creep tests carried out to date are summarized in Table 1. In determining the minimum creep rate by fitting the θ-projection equation, it was noted that the creep curve could be extrapolated for an average of 50 h. In the case of specimens subject to recovery heat treatment, this is also evident and might be indicative of the absence of recovery. Consideration of the mean and standard deviations for each of the four sets of data, further reinforces this preliminary conclusion.

The data generated from both control specimens and testing to rupture is represented in Figure 4. Here curves based upon the Monkman–Grant relationship, given in Eq. 4, have been fitted to both sets of data. Similar values for both β and C, within the accuracy of the fit, are obtained. When both the control and rupture data are merged together, the fitted line (shown in Fig. 5) serves as a baseline for comparison with data generated from re-

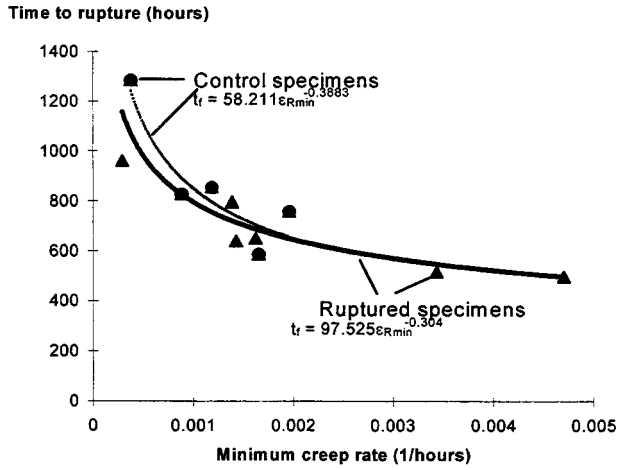


Fig. 4 Monkman-Grant relationships from rupture and control data

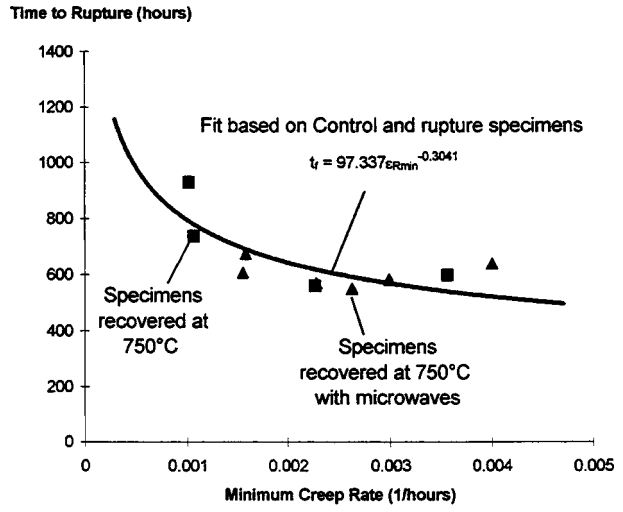


Fig. 5 Monkman-Grant relationship from combined rupture and control data, together with recovery data

covered specimens. Recovery upon heat treatment would be evident if associated points lay above the baseline. The inclusion of all data from specimens heat treated both thermally and with the application of microwaves in Fig. 5 clearly indicates that neither heat treatment regime has effected significant recovery of specimens pre-strained to 1.5 percent total strain. It is noted, however, that some results taken in isolation (as might occur when testing a smaller batch size) could be considered to indicate recovery.

Previous work has indicated that recovery becomes increasingly difficult as the amount of tertiary creep damage increases. Therefore, the present pre-straining of specimens to 1.5 percent total strain may represent too much creep damage for the particular heat treatment processes employed to recover. However, preliminary findings from tests conducted on specimens which were pre-strained to only 0.5 percent total strain prior to re-heat treatment without the application of microwaves are also indicative of the absence of recovery under the currently selected conditions. Hence, the present results may also indicate that the recovery observed by other authors was merely the fortuitous result of statistical variation and small sample batch size.

Conclusions

Work undertaken to date has demonstrated the limitations of assessing the recovery of creep properties on heat treatment by tests conducted on a small number of samples. It has also been found, within the accuracy of this technique, that neither unloading, nor the thermal (or thermal plus microwave heat treatment) regimes employed can affect the recovery of Nimonic 80A specimens pre-strained to 1.5 percent total strain. Hence, future work will focus on the recovery of creep properties in a number of specimens subject to a lesser pre-strain (say 0.5 percent), and will also involve re-solution treatment and re-precipitation both with and without the application of microwaves. It is further considered important that any density changes resulting from the various heat treatments be determined.

Equations

$$\varepsilon = \theta_1(1-\exp(-\theta_2 t)) + \theta_3(\exp(\theta_4 t)-1) \tag{1}$$

However, inserting into $\varepsilon_f = \varepsilon_o + \varepsilon_c$ gives

$$\varepsilon_f = \varepsilon_o + \theta_1(1-\exp(-\theta_2 t)) + \theta_3(\exp(\theta_4 t)-1) \tag{2}$$

$$\varepsilon_R = d\varepsilon_f/dt = \theta_1\theta_2\exp(-\theta_2 t) + \theta_3\theta_4\exp(\theta_4 t) \tag{3}$$

$$(\varepsilon_{R\,\min})^{\beta} t_f = C \tag{4}$$

Acknowledgments

This work was carried out under the EA Technology/UMIST Postgraduate Training Partnership (PTP). The authors gratefully acknowledge the financial support of both the Engineering and Physical Sciences Research Council (EPSRC), the DTI and EA Technology, and also the permission of EA Technology to publish these results. F. C. R. Wroe of EA Technology is also acknowledged for providing the idea behind the project and R. Pilkington of UMIST for provision of laboratory facilities.

Nomenclature

ε_f = total strain to failure
ε_o = elastic strain on loading
ε_c = creep strain
θ_{1-4} = constants
t = time
ε_R = creep rate
$\varepsilon_{R\,\min}$ = minimum creep rate
β = constant of <1
t_f = time to failure
C = constant

References

[1] Davies, P. W., Dennison, J. P., and Evans, H. E., 1966, ''Recovery of Properties of a Nickel-Based High-Temperature Alloy after Creep at 750°C,'' J. Inst. Met., **94**, pp. 270–275.

[2] Davies, P. W., Dennison, J. P., and Evans, H. E., 1967, ''The Kinetics of the Recovery of Creep Properties During Annealing of Nimonic 80A After Creep at 750°C,'' J. Inst. Met., **95**, pp. 231–234.

[3] Hart, R. V., and Gayter, H., 1968, ''Recovery of Mechanical Properties in Nickel Alloys by Re-Heat-Treatment,'' J. Inst. Met., **96**, pp. 338–344.

[4] Davies, P. W., Dennison, J. P., and Sidey, D., 1973, ''The Recovery of Mechanical Properties of Nimonic 105 by Heat-Treatment After High-Temperature Creep,'' J. Inst. Met., **101**, pp. 153–161.

[5] Cina, B., and Myron, S., 1976, ''Means for Extending the Secondary Creep Life of Nickel-Base Superalloys,'' *Second International Conference on Mechanical Behavior of Materials*, Federation of Materials Societies, Boston, pp. 2025–2030.

[6] Davies, P. W., and Dennison, J. P., 1975, ''The Use of Heat-Treatment to Recover the Creep Properties of Nimonic 115 After High-Temperature Creep,'' Met. Sci. J., **9**, pp. 319–323.

[7] Dennison, J. P., Holmes, P. D., and Wilshire, B., 1978, ''The Creep and Fracture of the Cast Nickel-Based Superalloy IN 100,'' Mater. Sci. Eng., **33**, pp. 35–47.

[8] Stevens, R. A., and Flewitt, P. E. J., 1979, ''Regenerative Heat Treatments for the Extension of the Creep Life of Superalloy IN 738,'' *Proceedings of the Fifth International Conference on Strength of Metals and Alloys*, Vol. 1, Aachen, Germany, Pergamon, New York, pp. 27–31.

[9] Ross, M. D., Bennett, G. T., and Stewart, D. C., 1979, ''Rejuvenation of Turbine Blade Material by Thermal Treatment,'' Pratt and Witney Aircraft Report AD-A077527.

[10] McLean, M., and Tipler, H. R., 1984, ''Assessment of Damage Accumulation and Property Regeneration by Hot Isostatic Pressing and Heat Treatment of Laboratory-Tested and Service-Exposed IN 738LC,'' *Proceedings of the Fifth International Symposium on Superalloys*, Gell, M., et al., eds. Metals Society of AIME, pp. 73–82.

[11] Koul, A. K., Immarigeon, J.-P., Castillo, R., Lowden, P., and Liburdi, J., 1988, Rejuvenation of Service-Exposed IN 738 Turbine Blades, ''*Superalloy's 1988*, S. Reichman et al., eds., The Metallurgical Society, Champion, PA, pp. 755–764.

[12] Baldan, A., 1992, ''Recovery of the Creep Resistance of a Conventionally Cast Nickel-Base Superalloy,'' J. Mater. Sci. Lett., **11**, pp. 1319–1321.

[13] Girdwood, R. B., and Evans, R. W., 1996, ''Recovery of Creep Properties of the Nickel-Base Superalloy Nimonic 105,'' Int. J. Pressure Vessels Piping, **66**, pp. 141–153.

[14] Stevens, R. A., and Flewitt, P. E. J., 1979, ''The Role of Hydrostatic Pressure on the Sintering of Creep Cavities in a Nickel-2 Percent Chromium Alloy,'' Acta Metall., **27**, pp. 67–77.

[15] Wroe, R., and Rowley, A. T., 1996, ''Evidence for a Non-Thermal Microwave Effect in the Sintering of Partially Stabilised Zirconia,'' J. Mater. Sci., **31**, pp. 2019–2926.

[16] Bruce, R. W., 1988, ''New Frontiers in the Use of Microwave Energy: Power and Metrology,'' Mater. Res. Soc. Symp. **124**, pp. 3–15.

[17] Zavalishin, Y. K., Masalov, Y. A., and Postnikov, S. N., 1995, ''Pulsed Magnetic Field Treatment (PMFT Method),'' Surf. Eng. Appl. Electrochem., **2**, pp. 65–74.

[18] Hochman, R. F., Tselesin, N., and Drits, V., 1988, ''Magnetic Fields: Fertile Ground for Metals Processing,'' Adv. Mater. Proc., **8**, pp. 36–41.

[19] Evans, R. W., and Wilshire, B., *Creep of Metals and Alloys*, The Institute of Metals, London.

[20] Nabarro, F. R. N., and de Villiers, H. L., 1995, *The Physics of Creep*, Taylor and Francis, London.

Measurement Uncertainties Encountered During Gas Turbine Driven Compressor Field Testing

K. Brun
Principal Project Applications Engineer

R. Kurz
Manager, Systems Analysis and Field Testing

Solar Turbines Incorporated,
San Diego, CA

Field testing of gas turbine compressor packages requires the accurate determination of efficiency, capacity, head, power and fuel flow in sometimes less than ideal working environments. Nonetheless, field test results have significant implication for the compressor and gas turbine manufacturers and their customers. Economic considerations demand that the performance and efficiency of an installation are verified to assure a project?s return on investment. Thus, for the compressor and gas turbine manufacturers, as well as for the end-user, an accurate determination of the field performance is of vital interest. This paper describes an analytic method to predict the measurement uncertainty and, thus, the accuracy, of field test results for gas turbine driven compressors. Namely, a method is presented which can be employed to verify the validity of field test performance results. The equations governing the compressor and gas turbine performance uncertainties are rigorously derived and results are numerically compared to actual field test data. Typical field test measurement uncertainties are presented for different sets of instrumentation. Test parameters that correlate to the most significant influence on the performance uncertainties are identified and suggestions are provided on how to minimize their measurement errors. The effect of different equations of state on the calculated performance is also discussed. Results show that compressor efficiency uncertainties can be unacceptably high when some basic rules for accurate testing are violated. However, by following some simple measurement rules and maintaining commonality of the gas equations of state, the overall compressor package performance measurement uncertainty can be limited and meaningful results can be achieved. [DOI: 10.1115/1.1340628]

Introduction

Gas turbine compressors are frequently tested after their installation in the field to verify the manufacturer's performance predictions. Failure to meet this performance may mean stiff financial penalties to the manufacturer as well as significant discontent on the end-user side. However, due to the non-ideal measurement conditions typically encountered during the field testing, uncertainties may be so significant that a simple single performance guarantee point cannot be defined. It is therefore imperative to define an allowable measurement uncertainty range, acceptable to both manufacturer and end-user, prior to the field test. To maintain impartiality, this uncertainty band should not be based on empirical ''experience'' values, but rather on uncertainty equations derived from the singular mathematical expressions for the gas turbine compressor performance and from the basic instrument and sensor device inaccuracies.

Limited information is available in the public domain describing actual gas turbine field test measurement uncertainties. The primary source of information employed by manufacturers and end-users are test codes such as provided by the American Society of Mechanical Engineers (ASME), the International Organization for Standardization (ISO) and the Verein Deutscher Ingenieure (VDI). These specifications cover thermodynamic calculation methods, instruments, site preparation and the reporting of turbomachinery test results in various degrees of detail. ASME Power Test Codes (PTC) 10 [1], 19.1 [2] and 22 [3], ISO 2314 [4], VDI 2045 [5] and 2048 [6] are typical examples of these standards.

A number of researchers from both academia and industry have to a very limited degree discussed field testing errors and uncertainties. For example, Fozi [7] discussed practical challenges of field performance tests. Steady state measurement conditions were shown to be critical for obtaining accurate field test results. Schmitt and Thomas [8] compared achievable field to factory test accuracies. Results showed that if field tests were properly conducted, comparable accuracies between field and factory tests were achievable. Meier and Rhea [9] performed compressor tests and discussed proper test practices and procedures. Instrumentation requirements for achievement of reasonable accuracy in field testing are defined and discussed. McRoberts [10] reported on compressor performance tests which were conducted in the challenging environment of an offshore platform. Finally, Klein and Draughton [11] undertook field performance tests of a gas reinjection compressor-turbine train. PTC 10 class I and PTC 10 class III factory test results were compared to field performance tests.

Thus, this paper aims to provide information for the public domain on gas turbine driven compressor field testing uncertainties. A simple analytic method to determine compressor and gas turbine measurement uncertainties is described. For this method a set of uncertainty equations is rigorously derived from the physical turbomachinery performance relations. Typical field test measurement experience errors are presented for different sets of instrumentation/sensors. By introducing these experience errors into the uncertainty equations, the total efficiency, power, and head uncertainty for a gas turbine compressor can be predicted prior to the field test for a given operating condition. Two actual field test examples are employed to demonstrate the viability of this method. Also, a number of parametric studies describing uncertainty trends are included. Parameters are identified that most significantly affect the total package uncertainty. This helps the user to optimize his field test instrumentation selection. Finally, the effect of different equations of state on the calculated performance is discussed.

Methodology

For the uncertainty analysis it is assumed that all measurement parameters can be considered to be independent and that param-

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Stockholm, Sweden, June 2–5, 1998; ASME Paper 98-GT-001. Manuscript received by the IGTI Division Mar. 1998; final revision received by the ASME Headquarters Nov. 2000. Associate Editor: R. Kielb.

eters have associated statistical bounds (such as a 95 percent confidence interval Δu) rather than absolute limits of errors. All parameters are also assumed to have Gaussian normal distributions around their respective mean values such that the uncertainties can be properly combined using the root-square sum method [12]. However, an uncertainty correction is added for parameters which have sample sizes smaller than 30; i.e., the uncertainty is widened for individual parameters to account for a Student-*t* type distribution [13]. The total uncertainty, ΔF, for a given function, $F=f(u_1,u_2,\ldots u_n)$ is, thus, determined from [14]:

$$\Delta F=\sqrt{\left(\Delta u_1\frac{\partial f}{\partial u_1}\right)^2+\left(\Delta u_2\frac{\partial f}{\partial u_2}\right)^2+\ldots\left(\Delta u_n\frac{\partial f}{\partial u_n}\right)^2} \quad (1)$$

For this method the overall uncertainty ΔF has the same statistical meaning as the individual uncertainties Δu. Namely, if Δu represents a 95 percent confidence than the result for the total uncertainty ΔF is also a 95 percent confidence interval. Note, that since it is not possible to distinguish between bias errors and data scatter in a field test environment, they will not be treated independently.

Compressor Performance Equations. The parameters that are measured during a centrifugal compressor field test are (i) the compressor inlet and outlet stagnation temperature, T_s and T_d, (ii) the inlet and outlet static pressure, p_s and p_d, and (iii) the actual inlet volume flow rate, Q (Fig. 1). Furthermore, knowledge of the gas composition is required to determine, based on a variety of available equations of state (SRK, LKP, BWRS, etc.), the gas molecular weight, MW, the specific heat ratio as a function of static temperature, $\gamma(t)$, and the gas compressibility factor as a function of static pressure and temperature $Z(p,t)$. Using this information the following simplified Eqs. (2) through (9) can be used sequentially to determine head, H, efficiency, η, and required driver power, P.

Gas constant:

$$R=\frac{R_{\text{Universal}}}{\text{MW}} \quad (2)$$

Density:

$$\rho=\frac{p}{R\cdot Z\cdot T} \quad (3)$$

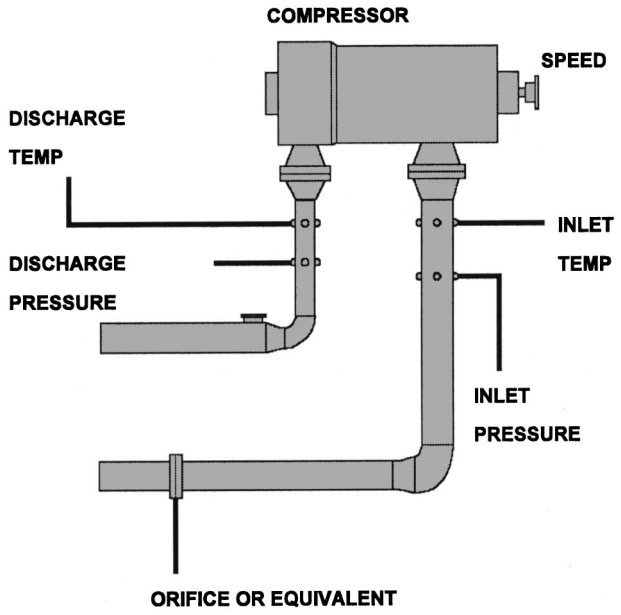


Fig. 1 Test instrumentation for a gas compressor

Specific heat:

$$c_p=Z\cdot R\cdot\left(\frac{\gamma}{\gamma-1}\right) \quad (4)$$

Head=actual enthalpy difference:

$$H=c_p(T_d)T_d-c_p(T_s)T_s \quad (5)$$

Isentropic outlet temperature:

$$T_d^*=T_s\cdot\left(\frac{p_d}{p_s}\right)^{\gamma 1/\gamma} \quad (6)$$

Isentropic enthalpy difference:

$$h^*=c_p(T_d^*)T_d^*-c_p(T_s)T_s \quad (7)$$

Isentropic efficiency:

$$\eta^*=\frac{h^*}{H} \quad (8)$$

Driver power:

$$P=\frac{H\cdot Q\cdot\rho}{\eta_M}=\frac{H\cdot Q\cdot p_s\cdot\text{MW}}{\eta_M\cdot R_{\text{Universal}}\cdot T_s} \quad (9)$$

In reality, the above simplified equations are not used in this form to determine compressor performance, but rather represent a close functional description of the actual physical equations. Nonetheless, since the total uncertainty is primarily based on the product of the individual uncertainty value and the gradient (df/du) of the governing equation, the simplified functional form is completely adequate to perform an accurate uncertainty analysis.

The simplified equations do not entirely represent the behavior of real gases: The effect of the pressure dependency of the heat capacity has been neglected. Also, in Eq. (6), the real gas behavior would require the isentropic coefficient rather than the ratio of specific heats. Both effects, while being important for the calculation of the absolute values, have very little influence on the functional form, and thus on the results of the uncertainty equations.

Note that for centrifugal compressor field tests one often assumes that stagnation and static values for pressure and temperature are very close so that their difference can be neglected for the performance calculation. Care should be taken with this assumption, though; this is only a valid assumption as long as compressor local inlet and outlet local flow Mach numbers are below approximately 0.1. For example, if in a pipeline compressor the local outlet flow Mach number reaches 0.1 it leads to a stagnation-static pressure difference of 0.6 percent and temperature difference of 0.15 percent. Based on the uncertainty equations, which are derived below, one can show that these deviations lead to a negligible compressor isentropic efficiency uncertainty of 0.12 percent. On the other hand, if the Mach number approaches 0.3 (which would coincide with about 125 m/s flow velocity in a natural gas pipeline), typical pressure and temperature errors are 7 percent and 1.6 percent respectively.

While in the calculation of the total pressure from the static pressure is easy, once the flow velocity is known, the difference between stagnation and static temperature imposes a more serious problem: The temperature measuring device in a thermowell will show a temperature that is somewhere between the stagnation and the static temperature. The potential temperature error is therefore defined by the difference between stagnation and total temperature. For the sake of brevity it is assumed herein that stagnation and static values are identical.

Compressor Uncertainty Equations. By substituting Eqs. (2), (3), and (4) into Eq. (1) and letting $L=(\gamma-1)/\gamma$, the following relation for the specific heat uncertainty Δc_p is obtained:

$$\sqrt{\left(\Delta Z\frac{R_{\text{Universal}}}{L\cdot \text{MW}}\right)^2+\left(\Delta L\frac{R_{\text{Universal}}\cdot Z}{L^2\cdot \text{MW}}\right)^2+\left(\Delta \text{MW}\frac{R_{\text{Universal}}\cdot Z}{L\cdot \text{MW}^2}\right)^2} \tag{10}$$

The above Eq. (10) is valid if the physical gas properties, specific heat ratio, compressibility factor, and molecular weight are directly determined from laboratory experiments. However, if these physical properties are calculated from a given gas composition via an equation of state, they cannot be considered to be independent.

For this case, Eq. (10) has to be slightly modified; namely, the absolute value sum of the individual uncertainty terms instead of the root-square sum should be taken to account for their functional dependency:

$$\Delta F=\left|\Delta u_1\frac{\partial f}{\partial u_1}\right|+\left|\Delta u_2\frac{\partial f}{\partial u_2}\right|+\ \ldots\ \left|\Delta u_n\frac{\partial f}{\partial u_n}\right| \tag{11}$$

This leads to a more conservative estimate of the uncertainty [12]. Furthermore, a physical property uncertainty, due to the effect of applying uncertainties in T and p to the nonideal gas state equation has to be included; i.e., since there is a measurement error in T and p, there will be an added error in determining c_p from the gas equation. This uncertainty is most conveniently obtained numerically by varying temperatures and pressures parametrically in the gas equation and, thus, determining the gradients $d\gamma/dT$, $d\gamma/dp$, dZ/dT, and dZ/dp indirectly. Recognizing that $d\gamma/dT=dL/dT$ and $d\gamma/dp=dL/dp$, one can easily determine corrections for ΔZ and ΔL:

$$\Delta Z=\sqrt{\left(\Delta T\cdot\frac{\partial \gamma}{\partial T}\right)^2+\left(\Delta p\cdot\frac{\partial \gamma}{\partial p}\right)^2} \tag{12a}$$

$$\Delta Z=\sqrt{\left(\Delta T\cdot\frac{\partial Z}{\partial T}\right)^2+\left(\Delta p\cdot\frac{\partial Z}{\partial p}\right)^2} \tag{12b}$$

The uncertainty in c_p is also affected by the variation of the gas properties during the duration of the test. This effect is again mathematically difficult to describe but can be easily handled numerically using a procedure similar to the one shown above for the variations in T and p. It is beyond the scope of this paper to list all possible gas composition variations; however, it is important to realize that they can strongly affect Z, L, and MW. For example, if a simple gas mixture of 90 percent Methane and 10 percent Ethane has a constituency uncertainty of 1 percent, then the resulting uncertainties in ΔMW, ΔZ, and ΔL are 0.9 percent, 0.05 percent, and 0.07 percent, respectively.

The authors would also like to point out that the consistent application of the gas state equations is imperative. For identical gas compositions one finds significant physical property output differences between the commonly employed gas equations of state (SRK, BWRS, and LKP). Beinecke and Luedtke [15] noted typical differences in Z values of 0.5 percent to 1.5 percent when comparing the output from SRK, BWRS, and LKP for a standard pipeline application. Thus, if one compressor manufacturer employs the SRK equation and another the BWRS gas equation, predictions for identical compressors may vary significantly. Also, if the compressor measurements are being employed to verify the driver (gas turbine) output performance, deviations of the gas equation output from actual physical values will lead to incorrect results for the required power.

By substituting Eq. (10) into Eq. (5), the uncertainty of the compressor head is determined. However, since the head uncertainty is not dependent on the absolute temperatures, but rather on the temperature difference (T_d-T_s), and since the specific heats (c_p) for the discharge and suction are functionally related, the temperature difference (T_d-T_s) should be employed for the derivation rather than the absolute temperature values (T_d,T_s).

$$\Delta H=\sqrt{(\Delta c_p\cdot(T_d-T_s))^2+(\Delta T_d\cdot c_pT_d)^2+(\Delta T_s\cdot c_pT_s)^2} \tag{13}$$

Uncertainties in the temperature T originate from the five following major sources of error: (i) location: incorrect position of the thermal sensor in the gas stream, (ii) installation: wall conduction heat transfer to and from the sensor due to inadequate insulation, (iii) calibration: instrument drift, nonlinearities, cold junction, and reference temperature errors, (iv) device: inherent accuracy limitations of the sensor device, (v) acquisition: amplifier, transmission, noise, read, and analog-digital conversion errors.

The first three of these errors can be minimized easily in production or laboratory test facilities. However, for field testing this is more difficult; namely, time and cost constraints can force the test engineer to accept field test arrangements with improperly located, installed, and calibrated instruments. For example, a minimum of four circumferentially spaced pressure taps should be employed to accurately measure gas pressures behind a pipe elbow. Due to the above mentioned constrains, typically only one tap is employed in field testing. While it is often impossible to correct these problems during the short field test duration, it is still imperative to recognize them and to advise the customer of consequential measurement accuracy limitations. The ASME PTC 10 [1] code specifies proper installation and location of the temperature sensors and, thus, should be used as a guideline and reference when defining field test procedures. Table 1 shows some typical values for the above five main sources of temperature measurement errors encountered during field tests. All units are in degrees Celsius.

To obtain the total temperature uncertainty, ΔT, the individual uncertainties have to be added using the root-square sum method. Table 1 shows that the location, installation, and calibration errors are the dominant factors while the device and acquisition errors are a smaller contribution to the total temperature error. Also note that field test device and acquisition errors are significantly larger than values quoted by instrument manufacturers ($>$0.005 percent full scale). Again, the circumstances and limitations encountered in field test may not always allow for ideal handling of the sensitive measurement instruments.

When substituting Eq. (6) into Eq. (1), the uncertainty for the isentropic (ideal) compressor outlet temperature is obtained.

Table 1 Typical magnitudes of temperature measurement errors [°C][1,2]

Sensor	Location	Installation	Calibration	Device	Acquisition
Hg Thermometer	0.1	0.1	0.0	0.03	0.1
Thermistor	0.1	0.1	0.2	0.02	0.05
Thermocouple	0.1	0.1	0.2	0.01	0.05
RTD	0.1	0.1	0.2	0.01	0.05
Infrared Sensor	0.3	0.0	0.3	0.05	0.05

[1]Aust [17] achieved under laboratory conditions ± 0.2 **K** with shielded NiCrNi-thermocouple probes. He mentioned deviations which depended on the flow velocity. Since the flow velocities in the nozzle of gas compressors are clearly subsonic, inaccuracies due to recovery factors should be negligible. The high accuracy will not easily be achieved in the field, since Aust used compensation elements kept in an oil bath of 0.0 ± 0.1°C. This procedure is not practical in the field. In any case, he found the calibration curves given in DIN43710 to describe the behavior of the thermocouples precisely. In the field, a total inaccuracy of ± 0.5 K seems achievable. Cleveland [16] shows instrument accuracies for thermometers to be 0.25–1.0 K, thermocouples 0.25–1.0 K and RTDs 0.0025–2.5 K. Schmitt and Thomas [8] report 0.5 K for RTDs. VDI 2045 [5] assumes a tolerance of ± 1.0 K for thermocouples and RTDs.

[2]Note, that the location error is based on the use of two probes. Lower errors can be achieved by using four probes, circumferentially distributed. Higher location errors occur if measurements are taken close to elbows and other flow obstructions. Also, on the discharge side of compressors, especially of compressors with volutes, the flow field can have significant temperature gradients. In all these cases, four probes are necessary to maintain the location error margins of Table 1. As a rule of thumb, the location error decreases by a multiplier of about $1/M$ with M being the number of probes evenly distributed in the measurement plane.

Isentropic temperature:

$$\Delta T_d^* = \sqrt{\left(\Delta T_s \cdot \left(\frac{p_d}{p_s}\right)^L\right)^2 + \left(\Delta p_d \cdot \frac{LT_s p_d^{L-1}}{p_s^L}\right)^2 + \left(\Delta p_s \cdot \frac{LT_s p_d^L}{p_s^{L+1}}\right)^2} \tag{14}$$

Uncertainties in pressure can also be categorized into the following five major groups of errors: (i) location: incorrect position or alignment of the pressure probe in the gas stream, (ii) installation: piping vibration transmission to the pressure pick-up due to inadequate damping, (iii) calibration: instrument drift, nonlinearities, hysteresis, and reference pressure errors, (iv) device: inherent accuracy limitations of the sensor device, (v) acquisition: amplifier, transmission, noise, and analog-digital conversion errors. Again, error sources (i), (ii), and (iii) are significantly larger in field testing than in production or laboratory tests. Although the ASME PTC 10 code [1] provides clear guidelines for correct installation and location of pressure probes one often finds these to be the main sources of pressure measurement errors. Table 2 shows some typical values for the five sources of pressure measurement errors encountered during field tests. All values are percent full scale. The uncertainty of the compressor efficiency, $\Delta\eta^*$, is determined by substituting Eqs. (7) and (8) into Eq. (1). Similar to Eq. (13), the temperature difference should be used rather than the absolute temperature values for the derivation of the isentropic enthalpy Eq. (15).

Isentropic enthalpy:

$$\Delta h^* = \sqrt{(\Delta c_p \cdot (T_d^* - T_s))^2 + (\Delta T_d^* \cdot c_{pT_d^*})^2 + (\Delta T_s \cdot c_{pT_s})^2} \tag{15}$$

Isentropic efficiency:

$$\Delta\eta^* = \sqrt{\left(\frac{\Delta h^*}{H}\right)^2 + \left(\Delta H \cdot \frac{h^*}{H^2}\right)^2} \tag{16}$$

Finally, the driver power uncertainty, ΔP, is obtained by combining Eqs. (9) and (1) to get the following:

Mass Flow:

$$\Delta W^2 = \left(\Delta p_s \cdot \frac{\mathrm{MW} \cdot Q}{R_{\mathrm{Universal}} Z T_s}\right)^2 + \left(\Delta \mathrm{MW} \cdot \frac{p_s Q}{R_{\mathrm{Universal}} Z T_s}\right)^2 + \left(\Delta Q \cdot \frac{p_s \cdot \mathrm{MW}}{R_{\mathrm{Universal}} Z T_s}\right)^2 + \left(\Delta Z \cdot \frac{p_s \cdot \mathrm{MW} \cdot Q}{R_{\mathrm{Universal}} Z^2 T_s}\right)^2 + \left(\Delta T_s \cdot \frac{p_s \cdot \mathrm{MW} \cdot Q}{R_{\mathrm{Universal}} Z T_s^2}\right)^2 \tag{17}$$

Power:

$$\Delta P = \sqrt{\left(\Delta H \cdot \frac{W}{\eta_M}\right)^2 + \left(\Delta W \cdot \frac{H}{\eta_M}\right)^2 + \left(\Delta \eta_M \cdot \frac{H \cdot W}{\eta_M^2}\right)^2} \tag{18}$$

The flow rate uncertainty, ΔQ, depends strongly on the device type employed for the measurements. A detailed discussion of flow measurement uncertainty is provided in ASME PTC 19.1 [3] and is, thus, not further discussed herein. Table 3 provides some typical values of ΔQ from field testing experience.

By evaluating Eqs. (10)–(18), estimates of the total measurement uncertainties for the compressor efficiency, head, and required driver power[7] can be obtained. However, one source of measurement uncertainty that is often overlooked is the uncertainty due to a finite sample size. The above uncertainty statistics, particularly Eq. (1) is valid only for mean parameters with an assumed Gaussian normal distribution. This is a good assumption for measurements where sample sizes are larger than 30. But for field tests it is sometimes difficult to maintain a steady-state system operating condition for a time period adequate to collect 30 or more samples. For small sample sizes it is more appropriate to assume a Student-*t* distribution instead of a Gaussian normal. Thus, an additional uncertainty due to a finite number of samples should be introduced into the total performance parameter uncertainty. This uncertainty is quantified by conservatively assuming

Table 2 Typical magnitudes of pressure measurement errors [percent full scale][3]

Sensor	Location[4]	Installation	Calibration	Device[5]	Acquisition
Simple Static	2.0	0.05	0.3	0.1	0.01
2 Static	1.0	0.05	0.3	--	--
4 Static	0.5	0.05	0.3	--	--
Pitot	1.0	0.1	0.3	0.1	0.01
Kiel	0.2	0.1	0.3	0.1	0.01

[3]Aust [17] achieved 0.1 percent of max. value for a calibrated system for wall static pressures, not including location. The location should not cause an error larger than 1 percent of dynamic pressure. Boelcs and Suter [18] show the dependency of the wall static error on wall shear stress. Wall taps need to be exactly perpendicular and flush to the surface—no burrs or slag are acceptable. On a compressor where the pressure taps were blown into the pipe with a cutting torch the resulting spread of the four pressure transducers was seen to be almost 2.0 percent of the static pressure. Preferably, the wall tap should be followed by a larger diameter hole [18]. The ratio of dynamic to static pressure for most applications will be below 1 percent, since the pipe diameters (and the compressor nozzle diameters) are selected to avoid high flow velocities. Also, the distortion that is seen by the impeller will be reduced due to fact that the flow normally gets accelerated considerably between the nozzle and the impeller eye.

VDI 2045 [5] assumes 0.2 percent of full scale for transducers and gauges. One should emphasize, that the instruments is selected such that most measured values are in the upper 25 percent of full scale. Liquid columns can be more precise, but the precision depends largely on the accuracy of reading the scales.

[4]This error will depend largely on the uniformity of the flow at the measuring location.

[5]Also reported by Schmitt and Thomas [8] and by various transmitter/ instrumentation manufacturers.

Table 3 Typical magnitudes of volume flow rate measurement errors [percent Full Scale][6] (values obtained with quality calibrated devices in laboratory conditions are shown in parentheses)

	Measurement Error	Sensitivity to Location	Pressure Loss
Orifice	1.5% (0.5%)	medium to low	high
Venturi	1.5% (0.5%)	medium	moderate
Ultrasound	0.5% (0.2%)	medium to high	low
Vortex	1.0% (0.3%)	--	--
Flow Meter	1.0-2.0% (0.5%)	medium	high
Annubar	1.0-1.5% (0.5%)	medium to low	moderate

[6]ISO2314 [4] assumes 0.5 percent accuracy for orifice flow measurements. This seems to be unrealistically low for field tests. Schmitt and Thomas [8] report a uncertainty for an orifice metering run per ISO5167 of 1.4 percent. Even for pipeline applications flow measurement error margins are 1 percent or larger [9]. Particularly cumbersome is the analysis if the gas samples are taken far away from the compressor or upstream of a separator or knock-out vessel [10]. A properly selected and calibrated fuel flow measuring device may be suitable to achieve measurement accuracies of ± 0.5 percent of the measured quantity [4]. However, the additional effort to get from 1 percent accuracy to 0.5 percent accuracy might not always be justified when the driven compressor is used for the power measurement. For example, with a 3 percent accuracy of the power measurement, the engine heat rate will have a tolerance of 3.16 percent and 3.04 percent respectively, i.e., the relative gain is only 0.1 percent in accuracy. According to a recent publication by Valenti [19] on the efforts at the Southwest Research Institute to evaluate gas flow measuring methods, measurement errors for orifices flow meters were 1 percent (mainly depending on piping configuration and diameter ratio of the orifice) and measurement errors for turbine flow meters were above 1 percent (mainly due to flow pulsations). This publication emphasizes the wide flow range and the very low pressure losses of ultrasonic flow meters. Ongoing research is trying to quantify the sensibility of this device to distortions of the flow profile.

[7]If a torque meter is used, the engine power can be measured with a total uncertainty as low as 1 percent. Even lower values were reported by Schmitt and Thomas [8].

Table 4 Parameter uncertainty due to limited sample size [percent]

Number of Samples	Head	Efficiency	Power
20	0.043	0.058	0.082
22	0.038	0.052	0.072
24	0.033	0.046	0.062
26	0.028	0.040	0.052
28	0.023	0.034	0.042
30	0.018	0.028	0.032

that all individual uncertainty bands are within their 95 percent confidence intervals and then determining the percent difference between Gaussian and Student-t distribution limits. Table 4 shows some typical values of the added uncertainty due to limited number of samples for the compressor head, efficiency, and power. Note that the values in Table 4 are dependent on inlet/outlet temperatures, inlet/outlet pressures, and gas composition. The limited sample size uncertainty should be added to the total parameter uncertainty using the root-square sum method.

Turbocompressor Package Uncertainty. To complete the above field test measurement uncertainty evaluation one also needs to look at the complete turbocompressor train (gas turbine and compressor efficiency) performance. The gas turbine shaft output power has to equal the compressor required power ($P_{GT}=P$). Thus, the following two equations can be used to define the gas turbine thermal efficiency, η_{TH}, and the total package efficiency, η_P:

Thermal efficiency:

$$\eta_{TH}=\frac{P}{W_{fuel}\cdot q} \tag{19}$$

Package efficiency:

$$\eta_P=\eta^*\cdot\eta_M\cdot\eta_{TH} \tag{20}$$

Here W_{fuel} is the fuel flow into the engine and q is the fuel heating value. The fuel flow is typically measured using an orifice plate in a metering run and the heating value is determined from the chemical composition of the fuel (often the centrifugal compressor discharge gas is used as the fuel gas). Based on the above equations the corresponding gas turbine uncertainty, $\Delta\eta_{TH}$, and package uncertainty, $\Delta\eta_P$, are given by:

Thermal efficiency:

$$\Delta\eta_{TH}=\sqrt{\left(\Delta P\frac{1}{w_{fuel}q}\right)^2+\left(\Delta W_{fuel}\frac{p}{w_{fuel}^2 q}\right)^2+\left(\Delta q\frac{P}{W_{fuel}q^2}\right)^2} \tag{21}$$

Package efficiency:

$$\Delta\eta_P=\sqrt{(\Delta\eta^*\eta_M\eta_{TH})^2+(\Delta\eta_M\eta^*\eta_{TH})^2+(\Delta\eta_{TH}\eta_M\eta^*)^2} \tag{22}$$

To complete the above Eqs. (21) and (22) the only additional information needed is the fuel flow uncertainty and the fuel heating value uncertainty. Since the fuel flow is measured in the same way as the flow through the gas compressor, uncertainty values in Table 3 can be used. Also, since the heating value is obtained directly from gas composition, the same percent uncertainty as was obtained for the specific heat [Eq. (10)] can be approximately used, namely:

$$\frac{\Delta q}{q}=\frac{\Delta c_p}{c_p} \tag{23}$$

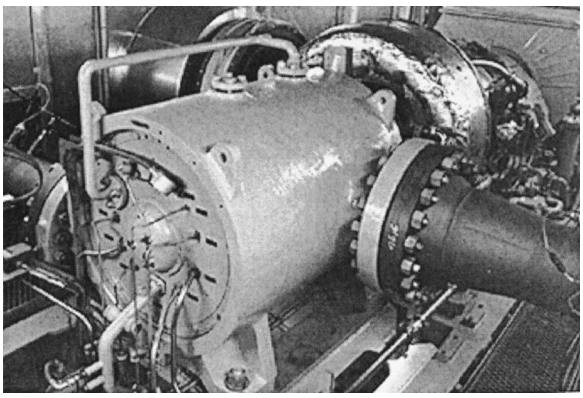


Fig. 2 Typical gas turbine driven compressor during a field performance test

By introducing the uncertainty experience values from Tables 1 through 4 into Eqs. (10)–(23), the measurement uncertainty for a field test can be predicted prior to the test. Consequently, the above method allows the gas turbine/compressor manufacturer and the end-user to determine reasonable test uncertainties as well as necessary requirements for the test instrumentation prior to the test. This method can also be employed to resolve observed variations of field test performance results from theoretically predicted and/or factory test results.

Uncertainty Calculations Examples and Comparisons to Field Test Results. Two examples are shown below to demonstrate how the total measurement uncertainty is obtained for some actual field test experiments. Figure 2 shows a typical gas turbine driven compressor during a field performance test. Also, a comparison to some factory test results is presented in example 1:

Example 1:

During field testing of a multistage turbocompressor the following conditions were encountered:

- Compressor application: Gas reinjection
- Instrumentation: single RTD and single pressure tap on each suction and discharge; orifice flow meter for compressor flow and turbine flow meter for engine fuel flow
- Test conditions: $p_s=65$ bar, $p_d=226$ bar, $T_s=303$ K, $T_d=448.5$ K ($\eta^*=62.0$ percent), $Q=18.4$ m^3/min
- Gas composition: fluctuating by 5 percent in constituency composition; during test 93.1 percent Methane, 5.8 percent Ethane, 0.4 percent Propane, 0.15 percent I-Butane, 0.25 percent *n*-Butane, 0.3 percent Nitrogen

The compressor field test uncertainty is predicted by evaluating Eqs. (10)–(23) with assumed measurement uncertainties as suggested in Tables 1 through 3 ($\Delta t_{s,d}=0.41$ K, $\Delta p_{1,2}=2.31$ percent, $\Delta Q=1.5$ percent, $\Delta\eta_M=0.1$ percent). The influence of the gas constituency fluctuation on the total uncertainty is accounted for by using the method as outlined in Eq. (12), which yields $\Delta Z=0.68$ percent, $\Delta\gamma=0.72$ percent, and ΔMW$=0.08$ percent for this case. From the resulting absolute uncertainty values for the specific heat, the actual head, the isentropic efficiency, and the required power (Δc_p, ΔH_A, $\Delta\eta_I$, ΔW) the percent total field test uncertainties are determined; namely: $\Delta c_p/c_p=2.46$ percent, $\Delta H/H=2.49$ percent, $\Delta\eta^*=3.24$ percent, $\Delta P/P=3.79$ percent.

Thus, total efficiency uncertainties of above 3 percent are seen for this set of field test measurements. The main cause for these rather large uncertainties is the inadequate number of pressure taps installed at the compressor suction and discharge. Increasing the number of pressure taps from one to four on each suction and discharge would reduce the predicted compressor and power un-

certainties to $\Delta\eta^*/\eta^* = 1.84$ percent, $\Delta P/P = 1.95$ percent, respectively; a significant measurement improvement. For the above example, a factory closed loop test had been performed on the compressor, in which an efficiency of 64.1 percent for a similar operating point was achieved. Taking the test uncertainty range into account, the efficiency measured during the field test (62 percent) therefore confirms the factory test results.

Example 2:

During a field test on a gas turbine driving a single stage pipeline compressor, the following compression conditions were encountered:

- Compressor application: Pipeline
- Instrumentation: four thermocouples and four pressure taps on each suction and discharge; annubar flow meters for the compressor flow; orifice flow meters for engine fuel flow
- Test conditions: $p_s = 45.36$ bar, $p_d = 55.1$ bar, $T_s = 25.8°C$, $T_d = 41.9°C$ ($\eta^* = 88.1$ percent), $Q = 6.64\ m^3/s$, $W_{\text{fuel}} = 0.47$ kg/s, $\eta_{\text{TH}} = 29.6$ percent, $\eta_P = 26.0$ percent
- Gas composition: 96.1 percent Methane, 1.8 percent Ethane, 0.3 percent Propane, 0.3 percent I-Butane, 0.1 percent n-Butane, 0.6 percent Carbon Dioxide, 0.8 percent Nitrogen

As a comparison, when applying the above test conditions to Eqs. (10)–(23) and assuming measurement uncertainties as shown in Tables 1–3 ($\Delta T_{s,d} = 0.3$ K, $\Delta p_{s,d} = 0.5$ percent, $\Delta Q = 1.5$ percent, $\Delta\eta_M = 0.1$ percent) the following percent uncertainties are obtained: $\Delta c_p/c_p = 2.40$ percent, $\Delta H/H = 3.6$ percent, $\Delta\eta^* = 5.5$ percent, $\Delta P/P = 4.0$ percent. Similarly, evaluating Eqs. (21) and (23) for the package and gas turbine total uncertainties with $\Delta Q_{\text{fuel}} = 1.0$ percent yields $\Delta\eta_{\text{TH}} = 4.5$ percent and $\Delta\eta^*_P = 4.3$ percent, respectively.

Note that despite a generally higher accuracy than in the first example, the uncertainties for example 2 are higher. This is typical for applications with low pressure ratios, which are very sensitive to uncertainties. Also note, that an improvement of the flow uncertainties would increase the overall accuracy only marginally.

In example 2, the accuracy for the power measurement is higher than the accuracy for the efficiency measurement. This is due to the fact, that the uncertainties for the actual head H are lower than for the isentropic head H^*. Because the calculation of the isentropic head requires the isentropic discharge temperature T_d^* [Eq. (14)], the associated uncertainties for T_d^* are usually higher than the uncertainties for T_d, which is directly measured. Thus, the power can be evaluated more accurately than the efficiency in many cases, as long as the accuracy of the flow measurement will be sufficiently high.

As in the previous example, one should remember that the theoretical uncertainties as determined from Eqs. (10)–(23) yield results for the total possible performance parameter uncertainty range (which is a very conservative estimate), while the deviation is just a comparison between two different measurements. One would thus anticipate the significantly larger theoretical uncertainty values.

Overall, the uncertainty equations are seen to predict measurement fluctuations well. However, the general experience with this method is that the uncertainty method conservatively over predicts errors by a small margin.

Parametric Studies and Trends. A number of parametric studies were undertaken to determine the effect of varying individual measurement variables on the total performance uncertainties. These studies also help to identify parameters which have the most significant effect on the total uncertainty. Hence, the effect of varying a number of field test measurement parameters on the efficiency and power uncertainty are evaluated. The compressor operating conditions as presented as example 1 above were used as a basis for the below studies.

Results of varying the compressor head and suction temperature (keeping all other conditions, including isentropic efficiency, constant) on the efficiency and power uncertainty are shown in Figs. 3 and 4. As the head is increased, the efficiency and power uncertainties are seen to decrease. The uncertainty rate of decrease is seen to be a direct function of the increasing compressor head magnitude. Also, the uncertainties are seen to be lower for increased compressor suction temperatures. Both of these results are expected and can be explained as follows: the larger the pressure or temperature suction-discharge differential, the smaller the relative influence of the individual measurement errors on the total performance uncertainties.

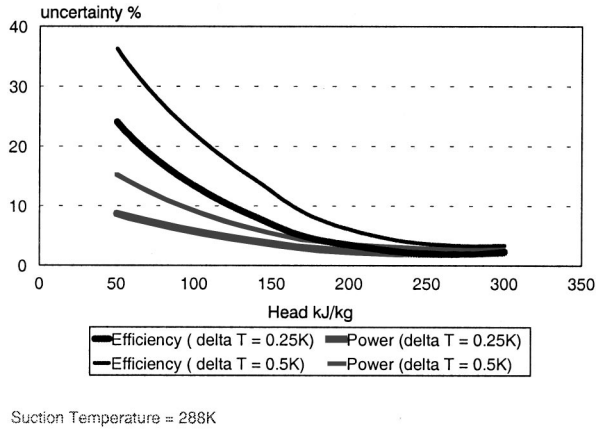


Fig. 3 Power and efficiency uncertainty versus compressor head at 288 K

Results in Figs. 3 and 4 are interesting, but do not provide any guideline as to which uncertainty most significantly affects the overall performance parameter uncertainties. Thus, Tables 5 and 6 show a quantitative comparison of the individual measurement uncertainties effect on the compressor isentropic efficiency uncertainty. For these studies only one measurement uncertainty parameter was varied while all others were left at the values described in example 1 above.

Results in Table 5 show that the influence of the suction and discharge pressure uncertainties on the total isentropic efficiency uncertainty are identical. This is consistent with Eq. (16); i.e., the influence on the isentropic enthalpy and efficiency uncertainty due to the gas composition uncertainty is negligible when compared to the influence due to the head uncertainty. Table 6 shows that the suction temperature uncertainty has a stronger influence on the total isentropic efficiency uncertainty than the discharge temperature uncertainty. Also, at higher temperature uncertainty levels this trend is seen to be more pronounced.

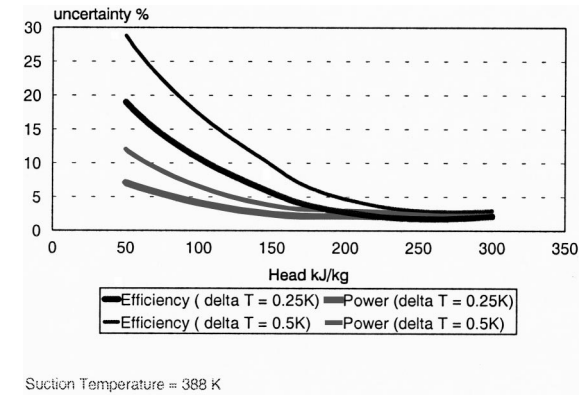


Fig. 4 Power and efficiency uncertainty versus compressor head at 388 K

Table 5 Compressor isentropic efficiency uncertainties [percent]

Error, %	Δp_s	Δp_d
1.0	2.95	2.95
2.0	3.15	3.15
3.0	3.46	3.46
4.0	3.86	3.86

Table 6 Compressor isentropic efficiency uncertainties [percent]

Error	ΔT_s	ΔT_d
0.25° K	3.21	3.23
0.50° K	3.27	3.25
1.00° K	3.46	3.30
2.00° K	4.12	3.46
4.00° K	6.04	4.00

Table 7 Compressor required power uncertainties [percent]

Error, %	Δp_s	Δp_d	ΔQ
1.0	3.17	3.79	3.62
2.0	3.61	3.79	4.02
3.0	4.25	3.79	4.60
4.0	5.00	3.79	5.30

Table 8 Compressor required power uncertainties [percent]

Error	ΔT_s	ΔT_d
0.25° K	3.77	3.77
0.50° K	3.81	3.81
1.00° K	3.90	3.90
2.00° K	4.22	4.17
4.00° K	5.20	5.02

Similarly, Tables 7 and 8 show a quantitative comparison of the individual measurement uncertainties affect on the compressor required power uncertainty. Again, for this study only one measurement uncertainty parameter was varied while all other were left at the values described in example 1 above.

Table 7 shows that the power uncertainty is strongly affected by the suction pressure uncertainty but not at all affected by the discharge pressure. Namely, the only pressure influence on the power uncertainty is through Eqs. (17) and (18) which only include the suction and not the discharge pressure. Also, Table 8 shows that the power uncertainty almost identically affected by the suction and discharge temperature uncertainty. Only at higher temperature uncertainty levels does the influence of the suction temperature uncertainty become more pronounced. The reason for the trend in Table 8 is that the required power uncertainty is primarily dependent on the temperature difference [equal influence of T_s and T_d as seen in Eq. (13)] and not on the absolute temperature values. The power uncertainty deviation at higher temperature uncertainty levels is due to the fact that the suction temperature also affects the power uncertainty through the mass flow Eq. (17).

Results in Tables 5–8 show that the temperature and flow rate measurement uncertainties have a stronger effect on the overall performance uncertainties than the pressure measurement uncertainties. However, due to the added complexity of pressure measurements, the field test pressure measurement uncertainties are often much larger than temperature measurement uncertainties and, thus, pressure measurement uncertainties should not be simply discounted.

Summary

The equations governing the compressor and gas turbine performance uncertainties were rigorously derived and results were numerically compared to two actual field tests. Typical field test measurement experience uncertainties were presented for different sets of instrumentation and sensors. By introducing the uncertainty experience values into the uncertainty equations the performance measurement uncertainty for a field test can be predicted prior to the test. This allows the gas turbine/compressor manufacturer and the end-user to agree on an allowable machinery performance range rather than having to set a singular guarantee point. The above method can also be employed to resolve observed variations of field test performance results from theoretically predicted and/or factory test results.

A number of parametric studies describing uncertainty trends are included. Test parameters that correlate to the most significant influence on the performance uncertainties were identified and suggestions are provided on how to minimize their measurement errors. Results showed that compressor efficiency uncertainties can be unacceptably high when some basic rules for accurate testing are violated. However, by following some simple measurement rules and maintaining commonality of the gas equations of state, the overall compressor package performance measurement uncertainty can be limited and meaningful results can be achieved.

Acknowledgments

The authors wish to thank Leon Sapiro, Daryl Legrand, Jerry Hammer, Tom Blattner, and Sebouh Ohanian for their various contributions to this paper.

Nomenclature

c_p = specific heat at constant pressure
γ = ratio of specific heats
h = enthalpy
k = isentropic exponent
Ma = mach number
MW = molecular weight
p = pressure
P = power
Q = volumetric flow
q = fuel heating value
R = gas constant
ρ = density
T = temperature
W = mass flow
Z = compressibility factor
η = efficiency
H = head
EOS = equation of state
BWRS = Benedict-Webb-Rubin-Starling
LKP = Lee-Kesler-Ploecker
PR = Peng-Robinson
RK = Redlich-Kwong
SRK = Soave-Redlich-Kwong

Subscripts

amb = ambient
d = discharge
f = fuel
M = mechanical
p = package
s = suction
TH = thermal

Superscripts

* = isentropic

References

[1] ASME PTC 10, 1979, ''Compressors and Exhausters.''
[2] ASME PTC 19.1, 1985, ''Measurement Uncertainties.''
[3] ASME PTC 22, 1985, ''Gas Turbine Power Plants.''
[4] ISO 2314, 1989, ''Gas Turbines-Acceptance Tests,'' Geneve.
[5] VDI 2045, 1993, ''Acceptance and Performance Tests on Turbo Compressors and Displacement Compressors,'' Duesseldorf.
[6] VDI 2048, 1978, ''Inaccuracies in Acceptance Test Measurements,'' Duesseldorf.
[7] Fozi, A.A., 1995, ''Field Testing of Compressor Packages,'' Solar Turbomachinery Technology Seminar, Paper No. 98.
[8] Schmitt, W., and Thomas, V., 1995, ''Comparison of Test Measurements Taken on a Pipeline Compressor/Gas Turbine Unit in the Workshop and at Site,'' ASME Paper No. 95-GT-125.
[9] Meier, R.H. and Rhea, C.S., 1982, ''Centrifugal Compressor Testing Experience and Practice,'' ASME Paper No. 82-GT-320.
[10] McRoberts, I.W., 1984, ''Performance Trials of Two Centrifugal Compressor Trains on an Offshore Platform,'' IMechE C48/84.
[11] Klein, J.M., and Draughton, P.A., 1991, ''Field Performance Testing of an Uprated Gas Reinjection Compressor/Turbine Train,'' ASME Paper No. 91-GT-48.
[12] Doeblin, E.O., 1966, *Measurement Systems*, McGraw-Hill, New York.
[13] Brun, K., 1996, ''Analysis of the Automotive Torque Converter Internal Flow Field,'' Ph.d. dissertation, University of Virginia.
[14] Scarborough, J.B., 1955, *Numerical Mathematical Analysis*, Johns Hopkins University, Baltimore, MD.
[15] Beinecke, D., and Luedtke, K., 1983, ''Die Auslegung von Turboverdichtern unter Berucksichtigung des Realen Gasverhaltens,'' VDI Berichte Number 487, pp. 271–279.
[16] Cleveland, A., 1982, ''Performance Testing of Gas Turbine Compressor Sets,'' ASME Paper No. 82-GT-199.
[17] Aust, N., 1988, ''Ein Verfahren zur Digitalen Simulation Instationarer Vorgange in Verdichteranlagen,'' Ph.D. Dissertation, University of the Federal Armed Forces, Hamburg, Germany.
[18] Boelcs, A., and Suter, P., 1986, ''Transsonische Turbomaschinen,'' G. Braun, Karlsruhe, Germany.
[19] Valenti, M., 1997, ''Determining Gas Meter Accuracy,'' Mech. Eng. (Am. Soc. Mech. Eng.), **3**, No. 119, pp. 102–106.

Degradation in Gas Turbine Systems

R. Kurz

K. Brun

Solar Turbines Incorporated,
San Diego, CA 92123

Any prime mover exhibits the effects of wear and tear over time. The problem of predicting the effects of wear and tear on the performance of any engine is still a matter of discussion. Because the function of a gas turbine is the result of the fine-tuned cooperation of many different components, the emphasis of this paper is on the gas turbine and its driven equipment (compressor or pump) as a system, rather than on isolated components. We will discuss the effect of degradation on the package as part of a complex system (e.g., a pipeline, a reinjection station, etc.). Treating the gas turbine package as a system reveals the effects of degradation on the match of the components as well as on the match with the driven equipment. This article will contribute insights into the problem of gas turbine system degradation. Based on some detailed studies on the mechanisms that cause engine degradation, namely, changes in blade surfaces due to erosion or fouling, and the effect on the blade aerodynamics; changes in seal geometries and clearances, and the effect on parasitic flows; and changes in the combustion system (e.g., which result in different pattern factors), the effects of degradation will be discussed. The study includes a methodology to simulate the effects of engine and driven equipment degradation. With a relatively simple set of equations that describe the engine behavior, and a number of linear deviation factors which can easily be obtained from engine maps or test data, the equipment behavior for various degrees of degradation will be studied. A second model, using a stage by stage model for the engine compressor, is used to model the compressor deterioration. The authors have avoided to present figures about the speed of degradation, because it is subject to a variety of operational and design factors that typically cannot be controlled entirely. [DOI: 10.1115/1.1340629]

Introduction

The function of a gas turbine (Fig. 1) and of a gas turbine package is the result of the fine-tuned cooperation of many different components. Any of these parts can show wear and tear over the lifetime of the package, and thus can adversely affect the operation of the system. In particular, the aerodynamic components, such as the engine compressor, the turbines, the driven pump, or compressor have to operate in an environment that will invariably degrade their performance. The understanding of the mechanisms that cause degradation, as well as the effects that the degradation of certain components can cause for the overall system are a matter of interest.

Degradation Mechanism. Several mechanisms cause the degradation of engines: Fouling is caused by the adherence of particles to airfoils and annulus surfaces. The adherence is caused by oil or water mists. The result is a build-up of material that causes increased surface roughness and to some degree changes the shape of the airfoil (if the material build up forms thicker layers of deposits). Many of the contaminants are smaller than 2 μm. Fouling can normally be eliminated by cleaning. Hot corrosion is the loss of material from flow path components caused by chemical reactions between the component and certain contaminants, such as salts, mineral acids, or reactive gases. The products of these chemical reactions may adhere to the aero components as scale. High-temperature oxidation, on the other hand, is the chemical reaction between the components metal atoms and oxygen from the surrounding hot gaseous environment. The protection through an oxide scale will in turn be reduced by any mechanical damage such as cracking or spalling, for example, during thermal cycles.

Erosion is the abrasive removal of material from the flow path by hard particles impinging on flow surfaces. These particles typically have to be larger than 20 μm in diameter to cause erosion by impact. Erosion is probably more a problem for aero engine applications, because state of the art filtration systems used for industrial applications will typically eliminate the bulk of the larger particles. Erosion can also become a problem for driven compressors or pumps where the process gas or fluid carries solid materials.

Damage is often caused by large foreign objects striking the flow path components. These objects may enter the engine with the inlet air, or the gas compressor with the gas stream, or are the result of broken off pieces of the engine itself. Pieces of ice breaking off the inlet, or carbon build up breaking off from fuel nozzles can also cause damage.

Abrasion is caused when a rotating surface rubs on a stationary

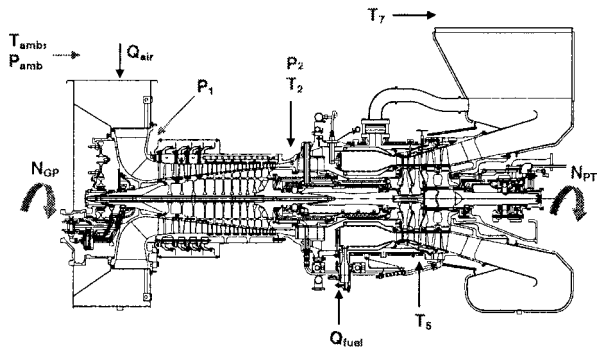


Fig. 1 Gas turbine cross section

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, 2000; Report No. 00-GT-345. Manuscript received by IGTI February 2000; final revision received by ASME Headquarters November 2000. Associate Editor: M. Magnolet.

surface. Many engines use abradable surfaces, where a certain amount of rubbing is allowed during the run-in of the engine, in order to establish proper clearances. The material removal will typically increase seal or tip gaps. While some of these effects can be reversed by cleaning or washing the engine, others require the adjustment, repair, or replacement of components [1].

It should be noted that the determination of the exact amount of performance degradation in the field is rather difficult. Test uncertainties are typically significant, especially if package instrumentation as opposed to a calibrated test facility is used [2]. Even trending involves significant uncertainties, because in all cases the engine performance has to be corrected from datum conditions to a reference condition [3].

Degradation of Components

Airfoils. In order to judge the degradation of aerodynamic components, we will first evaluate the effect of fouling, erosion, deposits, corrosion, and other damage on the individual airfoil. Fouling, and to some extent erosion generate a blade surface with increased roughness. Any increased roughness can increase the friction losses. It also may cause early transition from laminar to turbulent boundary layers, which increases loss production. Erosion, deposits, or damages to the airfoil change the geometric shape of the airfoil. On a well designed airfoil, optimized for the application, this will always reduce the performance of the airfoil. Bammert and Sonnenschein [4] reported early investigations on the change of turbine blade performance due to alterations of the blade geometry due to erosion, corrosion and fouling, which was modeled after results of the inspection of actual gas turbine blades [5]. The deterioration of the turbine blades is accompanied by changes in exit angles and increased losses. They report profile loss increases of 100% for 1% added or reduced thickness on well designed subsonic turbine blades. More recently, the same topic was reviewed by [6], with emphasis on the increased surface roughness due to degradation. It became clear that the main influence of degradation appears around the optimum incidence angles, while the far off-optimum performance hardly was influenced. It also became clear, that added roughness on the pressure side of the blades has a very small effect compared to added roughness on the suction side On the other hand, Boelcs and Sari [7], who tested transonic turbine blades with and without deposits on the pressure side, conclude that in general, ''a carefully designed blade will not significantly change its flow properties in a soiled condition.'' Kurz [8] found a 100% increase in profile losses on a transonic annular turbine cascade at the blades with scratches on the suction side. It must be noted that on transonic turbine cascades the transition to turbulent flow is often determined by the position of the shock, and is thus less dependent on the boundary layer condition. Losses will increase more pronounced, because the shock position on the suction side will move upstream, typically causing an earlier transition to turbulent flow [7,9].

If the blade operates at or near transonic velocities, deposits, or added roughness (with the associated growth in boundary layer thickness) will also reduce the possible flow through the blade row. Thicker boundary layers on the blades and sidewalls reduce the flow capacity, especially near choking conditions. On the other hand, if the trailing edge erodes, the throat width of the blade is increased, thus allowing more flow, but with less head reduction. Schmidt [10] describes the effect of a deliberate reduction of the chord length in a power turbine nozzle, that significantly reduced the work output of said turbine.

Milsch [11] reports increases in profile losses from 2% at $k_s/s=0.3\times10^{-3}$ to 10% at $k_s/s=5\times10^{-3}$ for NACA65-(12)06 compressor cascades, as well as significantly reduced turning. This reduction in performance is caused by significantly increased boundary layer growth, premature transition to turbulent flow, and premature flow separation.

In general, all these influences will create higher losses and less turning. This means that the following row of airfoils will see different incidence angles, higher temperatures, lower (for compressors) or higher (for turbines) pressures, and densities.

Erosion typically has the most significant effects on the blade leading edges. This can significantly affect the location and extent of the transition from laminar to turbulent boundary layers [12]. Because the heat transfer characteristics of a boundary layer depend, in addition to its thickness, on its state (i.e., aminar, turbulent, transitional, separated), leading edge erosion can influence the heat balance of the blade.

Clearances. Clearances between stationary and rotating parts (i.e., between stationary blades and the rotating hub or between rotating blades and the stationary casing) have a tendency to open up during the aging process of the equipment. This results in higher and unwelcome leakage flows. These leakage flows reduce the possible head capability and the efficiency of the components [13–15]. In compressors, an increase of the rotor tip clearance from 1 percent of blade chord to 3.5% of blade chord reduces the pressure ratio of the stage by up to 15%.

The loss production is due to intensive mixing of leakage flows with the main flow, thus producing losses and reducing the effective through flow area (blockage). A simulation by Singh et al. [16] on various compressor stages suggests that the effects of tip clearance and added roughness after ingestion of sand lead to about the same magnitude of performance deterioration. However, the degradation mechanism by sand ingestion is more relevant for aero applications than for industrial applications.

On driven equipment such as compressors and pumps, increased clearances in interstage labyrinth seals, balance piston seals, or shroud seals will increase the absorbed power. Especially for high pressure ratio machines with low flow (i.e., low specific speed impellers), increased balance piston recirculation due to worn balance piston seals can cause a significant increase in absorbed power.

Compressor. Three major effects determine the performance deterioration of the compressor: Increased tip clearances [14], changes in airfoil geometry [16], and changes in airfoil surface quality [17]. While the first two effects typically lead to nonrecoverable degradation, the latter effect can be at least partially reversed by washing the compressor [18].

The overall effect of degradation on an engine compressor can be described through added losses and lower capability of generating head. Typically, a degraded compressor also will have a reduced surge or stall margin [19]. This will not have any significant effect on the steady state operation.

For a given speed of a degraded compressor, each subsequent stage will see lower Mach numbers (because of the higher temperature) and an increased axial velocity component (because $\rho = p/\mathrm{RT}$, where p is reduced, T is increased, thus the density gets reduced).

The net effect will be that while in the new machine all stages were working at their optimum efficiency point at design surge margins, the degradation will force all stages after the first one to work at off optimum surge margins and lower than design efficiency. This will not only lower the overall efficiency and the pressure ratio than can be achieved, but also the operating range.

Saying this, it should be possible by careful readjusting variable geometry where available, to counteract some of the effects of degradation.

Singh et al. [16] presented a model and sample calculations for the deterioration of three different single stage compressors. While the model includes the effects of increased roughness and tip clearance due to erosion, it omits the effects blunting of the blade leading edges and shortening of the chord. The latter effects are typically not seen on compressors of industrial gas turbines due to more diligent methods of inlet air filtration. The results show, that while the shape of the φ-ψ-η characteristics of the

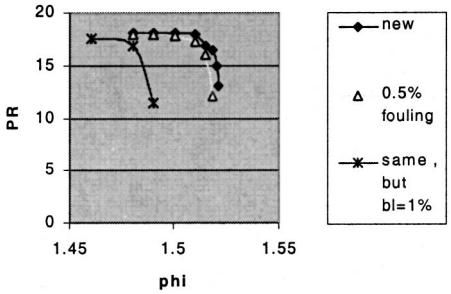


Fig. 2 Effect of fouling on an axial compressor, comparing a new compressor, a compressor with 0.5 percent fouling on head and efficiency, and a compressor with additionally 1 percent blockage: Pressure ratio

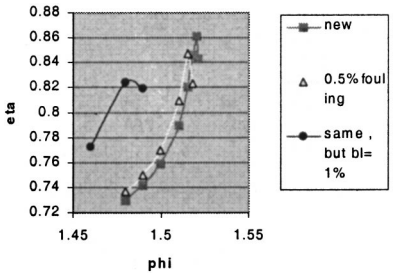


Fig. 3 Effect of fouling on an axial compressor comparing a new compressor, a compressor with 0.5 percent fouling on head and efficiency, and a compressor with additionally 1 percent blockage: Efficiency

stage are hardly affected, the achievable pressure ratio and the efficiency for the stage get reduced. The effects of blade roughness and tip clearance contribute about equally to the deterioration. The loss of aerodynamic performance was higher for the higher loaded test case. It also should be noted, that further increased roughness from initially $R_a = 4.0\ \mu$m to $R_a = 6.0\ \mu$m for the stator and 8.0 μm for the rotor, had created only a small additional deterioration. In addition to a reduced efficiency, added clearances will also increase the axial flow blockage, and thus will cause reduced through flow, and increased velocities in the main flow. Any reduction of turning will reduce the head rise of the stage.

With the above information, we are able to design a computational model that can simulate the effects of the degradation of single components on the performance of the compressor. The following assumptions were made.

At the design point, all stages operate at the same φ_{opt}. In the real compressor this could be accomplished by shaping the meridional plane of the compressor. In the model, the through flow areas were sized such that this was achieved for all stages.

Mach number dependency was implemented by linear variations of ψ. This can be justified by the fact, that for the changes in operating points for the model will be relatively small.

Inlet conditions for the next stage are calculated using updated pressure, temperature, and volume.

Stage characteristics (φ-ψ and φ-η) are modeled as third order polynomials with explicit limits for surge and choke. The stage characteristics for each stage are identical.

Fouling is simulated by multiplying the head and efficiency of the affected stage by two constant ''fouling factors'' (one each for head and efficiency).

The effects of increasing gaps are modeled by reducing the effective blade diameters, thus modeling the increasing blockage.

Applying the above model to a typical axial compressor reveals that the combined effects of airfoil fouling and increased clearances (Figs. 2 and 3) lead to loss of pressure ratio, loss of efficiency, and loss of range. In particular the increased clearances cause choke at lower flow. The model thus generates the same effects as mentioned in Refs. [19], [13], [16].

Turbine. Thicker boundary layers on the blades and sidewalls may reduce the flow capacity, especially near choking conditions, but one might think this effect is counteracted by the lower Mach numbers. However, lower Mach numbers are usually not seen at the first stage nozzle of the GP turbine (because TIT determines the temperature there) or at the first stage PT nozzle, because the necessary work, and thus temperature drop is dictated by the compressor power requirement. Boyle [20] found for a two stage turbine efficiency losses of 2.5% for a 10.2 μm surface roughness when compared to smooth blades. He also found that the most pronounced differences appear at the optimum operating point at the turbine, whereas the far off-optimum efficiency was almost the same for rough and smooth blades. It should be noted, that the losses due to clearances were in the same order of magnitude as the profile losses.

These clearance losses are approximately linear with the clearance δ [21]:

$$\Delta\eta = k_1 \cdot (\delta + k_2) \tag{1}$$

Traupel [21] also indicates that these losses increase with the increase of flow. In addition to a reduced efficiency, added clearances will also increase the axial flow blockage, and thus will cause reduced through flow, and increased velocities in the main flow. Radtke and Dibelius [22] report a reduction in efficiency of a multistage turbine by 0.6% when they increased the radial clearances from 0.5% of the blade height at the rotors/ 0.4 percent of the blade height at the stators to 0.8% of the blade height at rotors and stators.

However, if the degradation of the turbine section leads to material removal, especially in the nozzle area, we will see the opposite effect: The flow capacity increases for any given pressure ratio. The flow capacity of any nozzle is limited by the effective throat area. If the trailing edge erodes, the throat area increases. At the same time, the exit flow angle becomes more axial. Trailing edge erosion in the power turbine causes a shift in the distribution of pressure ratios $(p_3/p_5)/(p_5/p_7)$, i.e., a larger portion of the overall pressure ratio p_3/p_7 is consumed by the gas generator turbine. This causes a shift in the match temperature to higher ambient temperatures.

Looking into the polygons for turbines we can see that a reduction of turning in the stator and the rotor will lead to reduced work extraction for this stage.

Combustor. The combustion system is not likely to be the direct cause for performance deterioration. The combustion efficiency will usually not decrease. However, deterioration could potentially lead to a variation in the combustor exit temperature profile.

The problems with a distorted exit temperature distributions are threefold.

(1) Local temperature peaks can damage the turbine section.

(2) The altered temperature profile will increase secondary flow activity, thus reducing the turbine efficiency.

(3) Because the control temperature is measured at discrete circumferential points the average of these measured temperature is not the same as the true thermodynamic average temperature in this plane. Therefore, in the factory test, the correlation between the measured average and the true thermodynamic average is established. If the temperature field is altered due to (1) or (2), this correlation is no longer correct. Figure 4 shows two temperature fields in the T_5 measuring plane with probe locations. Both temperature fields have the same thermodynamic average temperature, but different probe averages. The engine could therefore be over fired (thus producing more power, but shortening the life) or under fired, thus losing power.

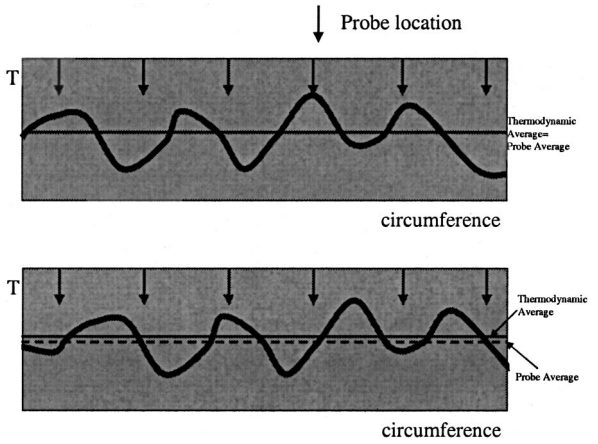


Fig. 4 Effect of probe location on the measured average temperature

Driven Equipment. Not only the driver will show performance degradation over time, but also the driven equipment. Solid and liquid contaminants in the process gas can damage impellers, or can form a dirt layer on the aerodynamic surfaces, which can increase friction losses. Cases are known, where pipeline compressors lost 5% in efficiency due to oily dust layers on the impeller and diffuser surfaces. One of the known sources of contamination are often reciprocating compressors which release significant amounts of lube oil into the process piping. It is also known that heavier hydrocarbons tend to form polymers at elevated temperatures, which can become the cause of fouling [23].

Gas Turbine Model

The overall behavior of the gas turbine was modeled using the following governing equations [24] for the flow compatibility equation for the gas generator

$$\frac{W\sqrt{T_3}}{p_3}=\frac{W\sqrt{T_1}}{p_1}\frac{p_1}{p_2}\frac{p_2}{p_3}\sqrt{\frac{T_3}{T_1}} \tag{2}$$

the work compatibility equations for the gas generator

$$\frac{\Delta T_{34}}{T_3}=\frac{\Delta T_{12}}{T_1}\frac{T_1}{T_3}\frac{c_{pa}}{c_{pg}\eta_m},$$

$$\eta_m c_{pg}\Delta T_{34}=c_{pa}\Delta T_{12}, \tag{3}$$

and the flow compatibility for the power turbine

$$\frac{W\sqrt{T_5}}{p_5}=\frac{W\sqrt{T_3}}{p_3}\frac{p_3}{p_5}\sqrt{\frac{T_5}{T_3}}. \tag{4}$$

Another condition is imposed by the flow capacity of the gas generator and power turbine.

From a pressure balance, we get the identity

$$\frac{p_3}{p_5}=\frac{p_3}{p_2}\frac{p_2}{p_1}\frac{p_1}{p_5}, \tag{6}$$

where p_3/p_2 is determined by the pressure drop in the combustor. This means that for any compressor pressure ratio, the gas generator pressure ratio is fixed by the power turbine pressure ratio. In particular, if the power turbine is in choke, then the gas generator pressure ratio is a unique fixed value.

Furthermore, the fuel flow is determined from

$$W_f=\frac{c_{pg}(T_3-T_2)W_1}{LHV}. \tag{7}$$

Any component of the engine can be represented by a performance map. In order to allow for small deviations from the design point, the dependency of operating parameters can be linearized, i.e., we do not necessarily need to model the entire performance map (Table 1). Even with these relatively simple relationships, the typical behavior of a two shaft engine can be simulated: The limitation by maximum speed at inlet temperatures below the match temperature, while being limited by maximum TIT at temperatures above the match temperature.

Table 1 Linearized operating parameters

Linearized parameter	Depending on	Equation
Compressor pressure ratio	Corrected gas generator speed	$k_{\mathrm{prc}}=\Delta \mathrm{TT}_c/\Delta N_{\mathrm{gpcorr}}$
Compressor flow function	Corrected gas generator speed	$K_{Wc}=\Delta W_c/\Delta N_{\mathrm{gpcorr}}$
Compressor efficiency	Corrected gas generator speed	$k_{\eta c}=\Delta \eta_c/\Delta N_{\mathrm{gpcorr}}$
Gas generator pressure ratio	Compressor pressure ratio	$k_{\mathrm{gpt}}=\Delta \mathrm{TT}_{\mathrm{gpt}}/\Delta \mathrm{TT}_c$
Gas generator turbine efficiency	Gas generator turbine speed	$k_{\eta\mathrm{gpt}}=\Delta \eta_{\mathrm{gpt}}/\Delta(\lvert Q_3/N_{\mathrm{gp}}\rvert)$

The optimum speed of the power turbine will depend on the gas generator operating point. With

$$\left(\frac{u}{c_{\mathrm{ax}}}\right)_{\mathrm{opt}}\propto\left(\frac{N_{pt}}{Q_5}\right)_{\mathrm{opt}}, \tag{8}$$

we derive the optimum power turbine speed from

$$N_{\mathrm{Pt,opt}}=\left(\frac{N_{pt}}{Q_5}\right)_{\mathrm{opt}}Q_5=\left(\frac{N_{\mathrm{Pt}}}{Q_5}\right)_{\mathrm{opt}}W\frac{RT_5}{p_5}. \tag{9}$$

The reduction in power turbine efficiency at off-optimum speeds can be approximated by

$$P=P_{\mathrm{opt}}\left[2.0\frac{N_{\mathrm{Pt}}}{N_{\mathrm{Pt,opt}}}-\left(\frac{N_{\mathrm{Pt}}}{N_{\mathrm{Pt,opt}}}\right)^2\right]. \tag{10}$$

These conditions are sufficient to explain the behavior of gas turbines at off-design conditions, for example, at ambient temperatures different from the ''match'' temperature: At higher ambient temperatures, the pressure ratio p_2/p_1 is reduced because the same head produces less pressure ratio, and the mass flow is reduced because the choked power turbine nozzle limits the flow (temperature topping). At lower ambient temperatures, gas generator speed limits override a limit in firing temperatures (speed topping).

Increasing the power turbine nozzle flow area reduces the power turbine resistance (i.e., increases the flow capacity), allowing a higher portion of the overall pressure ratio across the gas generator turbine. The TIT needed for the gas generator turbine to drive the compressor is reduced. When an engine matched to give rated TIT at rated gas generator speed at, for example, 15°C (59°F) is operated at 49°C (120°F), the rated TIT occurs before the rated gas generator speed can be obtained. To increase the gas generator speed, the power turbine area would have to be increased.

Effect of Degradation on the Gas Turbine

The following relationships are used in the model above to describe the amount of degradation. The loss of compressor efficiency $\Delta\eta_c$, compressor pressure ratio $\Delta\pi_c$, compressor air flow ΔW_c, gas generator turbine efficiency $\Delta\eta_{\mathrm{GPT}}$, and power turbine efficiency $\Delta\eta_{\mathrm{PT}}$.

For two shaft engines, the effects of degradation depend also on the control mode the engine is in. If the engine is operating at ambient temperatures below the match temperature, it will be in speed topping mode, i.e., the gas generator output is limited by the gas generator speed. If the engine is operating at ambient tempera-

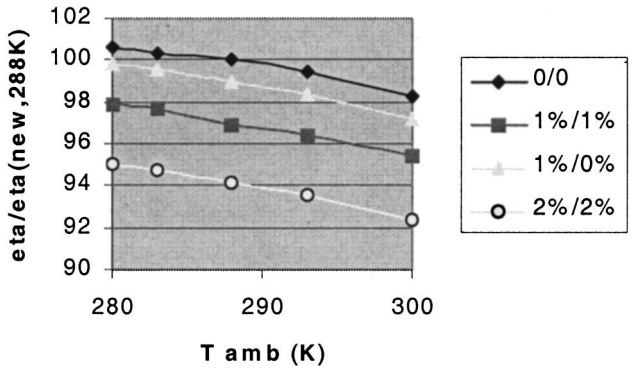


Fig. 5 Degradation of engine efficiency at full load

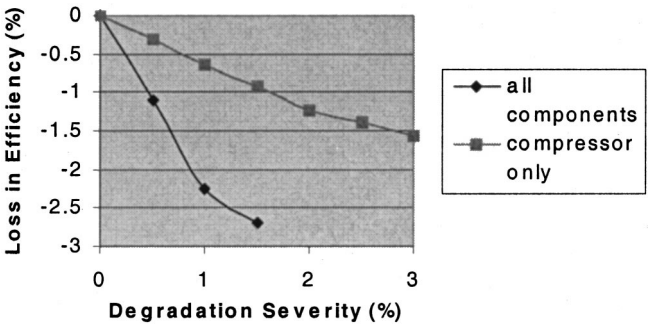


Fig. 7 Degradation of engine efficiency at constant load

tures above the match temperature, it will be in turbine inlet temperature topping mode, i.e., the gas generator output is limited by the firing temperature. This is similar to the situation seen in the comparison of Tarabrin et al. [25] between single and two shaft engines. The single shaft, generator drive will by nature have to run a fixed speed. Similar to that, the two shaft engine at speed topping will run at maximum speed during full load operation. The effects of degradation thus are less severe than for an engine at temperature topping, which is confirmed by the model. Effectively, degradation of the compressor and the gas generator turbine section move the match point to lower ambient temperatures. Figures 5 and 6 show the full load output and efficiency for an engine at different ambient temperatures assuming different levels of degradation. The different levels of degradation include case 1/0, where compressor flow, pressure ratio and efficiency are reduced by 1%. Case 1/1 is similar to the previous case, except that additionally, the gas generator turbine and power turbine efficiencies are reduced by 1%. Case 2/2 has all values of case 1/1 reduced by 2%. While the case for as new performance (0/0) shows the correct match temperature of 288 K (recognizable by the change in slope for the power and efficiency line), all other cases exhibit a match point at lower temperatures. Therefore, the effect of degradation at temperatures below 288 K is somewhat reduced.

Figure 7 shows the engine efficiency at a given load for various ambient temperatures assuming different levels of degradation. The degradation levels for ''all components'' are such, that 1% would correspond to case 1/1 in the previous example, while 1% degradation severity for ''compressor only'' corresponds to case 1/0. It shows that the decline in efficiency slows down at higher levels of degradation, both for the cases where all components are affected by degradation as well as cases where only the compressor is affected. This is due to the fact that the operating point in the example was at 95% load for the new engine, while for the respective end points of the curves, the engine is operating again at design firing temperature, thus at full load.

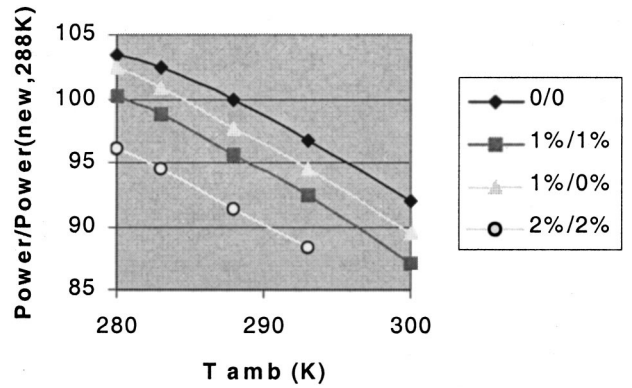


Fig. 6 Degradation of engine power output

Another point worth mentioning is the part load behavior. An operating point at part load (i.e., below maximum firing temperature and below max. NGP) can still be maintained with a degraded engine, albeit at a higher firing temperature and a higher gas generator speed than for the new condition. The relative loss in efficiency is significantly lower than for an engine at full load for the same amount of degradation.

Compressor deterioration by itself will usually cause higher power losses than losses in heat rate, because higher compressor exit temperature (due to lower efficiency) at a fixed firing temperature will reduce the possible fuel flow.

Tarabrin et al. [25] point out, the effect of degradation is typically more severe for two shaft engines than for single shaft engines. However, this is only true, if compressor degradation on the two shaft engine leads to a drop in gas generator speed. If variable geometry is available (in the example of Tarabrin et al. [25] adjustable PT nozzles, in our example adjustable compressor IGV's and stators), this drop in speed can be avoided, and the effects of compressor degradation are roughly the same for single and two-shaft engines.

Another form of (rather severe) degradation of the gas generator turbine nozzle occurs when material is removed due to erosion or corrosion. In this case, the pressure ratio over the turbine and its efficiency will drop. This will lead to a reduction in gas generator speed (i.e., again a shift in the match point to a lower temperature). The effect on the overall engine performance depends on the speed characteristic of the compressor, but if we assume that the compressor efficiency does not change with speed, then we see a significant reduction in engine output and efficiency. For some compressor designs, the design operating point is at slightly higher speeds than the optimum efficiency point of the compressor. In this case, the increase in compressor efficiency at reduced speeds can to some extent counteract the loss of engine efficiency.

Yet another effect is introduced by the fact that in most engines the turbine inlet temperature (TIT) is not measured directly. Rather, the power turbine inlet temperature (PTIT) or (on single shaft engines) the exhaust gas temperature (EGT) are measured, and TIT is calculated. The measured PTIT or EGT is not the thermodynamic average temperature, but the arithmetic average of several point measurements. Degradation of engines can lead to a shift in the true ratio between TIT and PTIT, thus causing the engine to over or under fire. This shift can be due to several reasons: change in flame patterns with resulting change in temperature patterns or change in gas generator turbine efficiency.

The second cause in the list has in fact a pronounced effect, mainly on the engine output, but only a small effect on engine efficiency. It should be noted that these losses can be recovered by

readjusting the control system. It can be shown that a reduction of gas generator turbine efficiency will lead to a underfiring of the engine under these conditions.

Degradation also affects the optimum power turbine speed. If a lower engine compressor ratio or deterioration at the gas generator turbine itself leads to a lower gas generator turbine pressure ratio, then the actual flow through the power turbine will increase slightly, thus increasing the optimum speed. However, the effect on the engine output at fixed PT speed is rather small (less then 0.1%) for modern power turbines with relatively wide operating range.

For comparison purposes, a second method to study the system degradation of gas turbines was developed using the simple nonideal gas turbine cycle analysis concept. Hence, a simple nonideal simple Brayton cycle analysis program was written using compressor, combustor, and turbine operating maps for a typical midsize two-shaft gas turbine. For this method it was assumed that the gas turbine degradation behavior can be accurately modeled using a nonideal Brayton cycle analysis if the individual component operating maps accurately reflect each components performance when the gas turbine operates at off-design points. While the previous method linearized the pressure ratio flow and efficiency relationships for each component, this method employs the actual nonlinear (higher order polynomials) operating maps for each components and is thus more accurate at higher gas turbine degradations and off-design. The following examples will prove that the linearization is accurate as long as the degradation remains small enough.

Engine Data

MacLeod et al. [26], who investigated the effect of component deterioration on overall engine performance on a single shaft turboprop engine, found that an increase in GP nozzle through flow area of 6% due to erosion, cracking, bowing and corrosion, caused an increase in heat rate by 3.55% but virtually no loss in power. Our model for a two shaft engine shows under these conditions a significant reduction in gas generator speed, with both significant (7% and 3%) losses in power and efficiency.

Frith [15] tested a helicopter gas turbine with compressor blades cropped to simulate increased clearances. A 3% crop on the axial compressor stages reduced airflow by 4.6% and pressure ratio by 3%. The compressor discharge temperature is reported to remain unchanged, which indicates a reduction in compressor efficiency by about 2.5%. Since the gas turbine is essentially the same as tested by Schmidt [10], the compressor, gas generator turbine and power turbine characteristics were available. The model predicts a loss in power of 7.8% and a loss in efficiency by 2.7%, which matches the engine test data (loss in power 8% and loss in efficiency 3.4%) very closely.

Figure 8 shows the result of a variety of measures to reverse the effects of degradation. In the example, the engine was returned to the factory after several thousand hours of operation. The initial run of the engine, without any adjustments whatsoever and at the IGV and T5 topping temperature from the original factory test, showed that the TIT was 40°F (22°C) below design T_3 and the gas generator speed was 3% below design speed. After adjusting the guide vanes to get back to the desired design T_3 and gas generator speed, the engine improved power and reduced heat rate. Then, the engine was detergent washed and continued to improve performance. Next the individual stages of compressor variable vanes were adjusted to the factory settings improving power. After all of the adjustments the engine, compared to the factory testing when the engine was new, had lost 2.5% in power and 1.2% in heat rate. These results show that a significant amount of apparent degradation can be reversed by washing and adjustments. Another engine was operated 3500 hours without a detergent wash. The environment was laden with jet engine exhaust and salt air. Borescope inspections had shown deposits on the compressor and turbine. A detergent wash was recommended. Control system data was recorded at full load before and after detergent washing and are displayed in Fig. 9. The data was taken with the ambient temperature below the design match temperature.

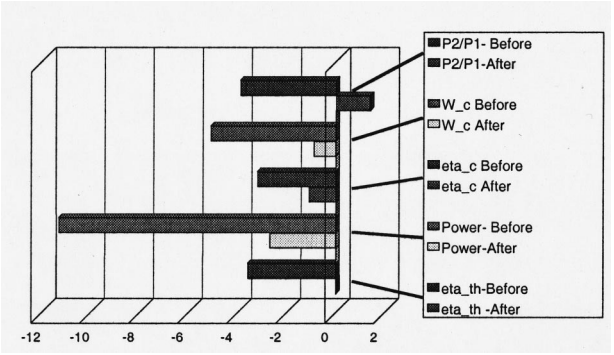


Fig. 9 Effect of detergent washing

The improvement in compressor pressure ratio, compressor efficiency, power, and heat rate are as expected, because the washing detergent was reported to be very black and dirty. The engine was also under fired, because one of the effects of degradation is a the reduction in the T_3/T_5 ratio. These improvements, explain the very substantial 9.7% increase in output power. After washing, and given the test uncertainties, this engine appears to be performing essentially the same as when new. The model predicts exactly this behavior: if we use a 2.1% loss in compressor efficiency, a 5% reduction in airflow and pressure ratio and a 0.5% reduction in gas generator turbine efficiency, we see a reduction in power of 8.6%, the efficiency drops by 3.5%, while the engine speed stays almost the same, and T_3 drops due to a reduction in T_3/T_5 ratio (Fig. 10).

Interestingly, while Tarabrin [25] indicate that smaller engines may be more susceptible to degradation, the study of Aker and Saravanamuttoo [27] reaches exactly the opposite conclusion.

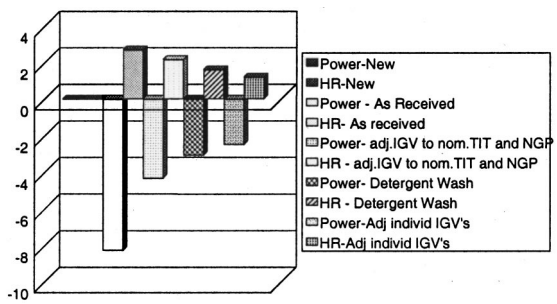


Fig. 8 Performance recovery

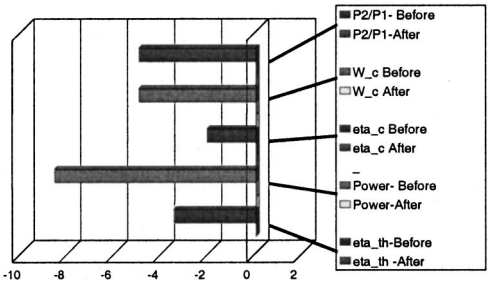


Fig. 10 Simulation of effects of detergent washing

Effects of Degradation on the Overall System

As we saw in the previous chapter, the net effect of engine degradation is a reduction of available power and an increase in heat rate, while the driven equipment will require more power for given operating points. The question is now, what does that mean with regards to the capability of the installation to fulfil its duty. We need to look into the different application scenarios, first for compression applications.

Compression Where Head Requirements Depend on the Flow. A typical example is a gas turbine driven compressor operating in a pipeline. The pressure losses in the pipeline increase roughly with the square of the actual flow. Because these losses have to be overcome by the compressor, the relationship between head and flow of the compressor is also roughly quadratic (Fig. 11). This means, that the net effect of degradation will be less flow, but (say) 4% less power from the driver will reduce the flow by less than 4%. Since most station designs have drivers sized for the highest ambient temperatures, strategies such as line packing will allow to maintain the station performance, albeit with a higher fuel consumption. As discussed earlier, the loss in power will be more significant than the loss in efficiency.

Compression Where the Head Requirements are Constant. In applications, where the suction and discharge pressures are fixed, such as in cases where the station feeds into a pipeline that operates at a certain pressure, or on processing applications, the compressor has to produce a certain head in order to generate any flow. If the available power is reduced, the flow is reduced. This will move the operating point of the compressor closer to the surge line. At one point in time the available power will no longer be sufficient to pump enough flow for the compressor to operate. It is applications similar to this, where equipment degradation may cause the shutdowns at the hotter parts of a day. Depending on the critical need for the station, it may be prudent to oversize the driver somewhat.

Power Generation. For power generation applications, the engine degradation will lead to higher fuel consumption, less power generation, but also more exhaust heat.

Pipeline Application. Figure 11 shows the move of an operating point for a pipeline compressor for the degradation levels from case 1/1. In this example, the pipeline compressor operation is represented by a map showing isentropic head versus actual flow for different operating speeds. It also shows lines of constant pipeline compressor efficiency, and lines of constant shaft power, from

$$\frac{P}{\rho} = \frac{QH}{\eta}. \tag{11}$$

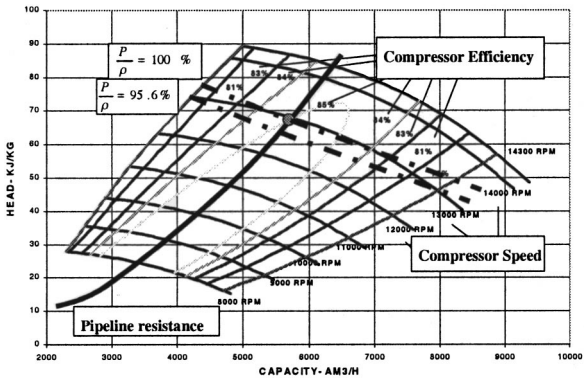


Fig. 11 Effect of system degradation on a gas turbine driven pipeline compressor

Table 2 Effect of various cases of degradation on power reduction and flow reduction

Case	Reduction in power (%)	Reduction in flow (%)
0/0	0	0
1/0	3.4	2.5
1/1	4.4	3.3
2/2	8.6	6.5

Because the resistance of any pipeline is approximately proportional to the square of the volumetric flow, any operating point of the pipeline compressor will reside on a single line. In particular, the maximum flow for any available power is at the intersection of the shaft power line and the pipeline resistance line. We also need to take into consideration that the available shaft power is a function of the speed [Eq. (10)], too. Considering the above, we can see that the flow is reduced, once the engine suffers from degradation. But, as Table 2 indicates, the loss in flow capability is smaller than the actual loss in power, because any reduction in flow also leads to a reduction in head.

Recoverable and Nonrecoverable Degradation

The distinction between recoverable and nonrecoverable degradation is somewhat misleading. The majority of degradation is recoverable, however, the effort is very different depending on the type of degradation. The recovery effort maybe as small as water or detergent online washing, or detergent on-crank washing. The degradation recovery by any means of washing is usually referred to as recoverable degradation. However, a significant amount of degradation can be recovered by engine adjustments (such as resetting variable geometry). Last, but not least, various degrees of component replacement in overhaul can bring the system performance back to as-new conditions.

Protection Against Degradation

While engine degradation cannot entirely be avoided, certain precautions can clearly slow the effects down. These precautions include the careful selection and maintenance of the air filtration equipment, and the careful treatment of fuel, steam or water that are injected into the combustion process. It also includes to obey manufacturers recommendations regarding shut-down and restarting procedures. For the driven equipment surge avoidance, process gas free of solids and liquids, and operation within the design limits need to be mentioned. With regards to steam injection, it must be noted, that the requirements for contaminant limits for a gas turbine are, due to the higher process temperatures, more stringent than for a steam turbine.

The site location and environment conditions, which dictate airborne contaminants, their size, concentration and composition, need to be considered in the selection of air filtration. Atmospheric conditions, such as humidity, smog, precipitation, mist, fog, dust, oil fumes, industrial exhausts, will primarily effect the engine compressor. Fuel quality will impact the hot section. The cleanliness of the process gas, entrained particles or liquids, will affect the driven equipment performance. Given all these variables, the rate of degradation is impossible to predict with reasonable accuracy.

While the rate of deterioration is slowed by frequent online washing, thorough on-crank washing can yield a more significant recovery. However, if the wrong detergents are used, some of the solvent may stick to the blades, or the washing process simply transports contaminant from the front stages of the compressor to rear stages or the turbine section. No matter how good the washing, the rear stages of the compressor will not get cleaned. If the compressor blades can be accessed with moderate effort (for example, when the compressor casing is horizontally split), hand cleaning of the blades can be very effective.

Conclusions

The study presents a methodology to simulate the effects of engine and driven equipment degradation. With a relatively simple set of equations that describe the engine behavior, and a number of linear deviation factors, which can easily be obtained from engine maps or test data, the equipment behavior for various degrees of degradation can be studied. The authors have avoided presenting figures about the rate of degradation, because it is subject to a variety of operational and design factors that typically cannot be controlled entirely.

Nomenclature

A = area
a = speed of sound
c_{ax} = axial velocity
cp = specific heat capacity
H = head
k_i = constants
k_a = equivalent roughness
LHV = lower heating value
N = speed
P = power
PTIT = power turbine inlet temperature
p = pressure
Q = volumetric flow
R_a = roughness
s = chord length
T = temperature
TIT = gas generator turbine inlet temperature
u = tip speed
W = mass flow
η = isentropic efficiency
$\psi = H/(u^2/2)$ = nondimensional head
$\varphi = Q/(u/A)$ = nondimensional flow
π = pressure ratio
ρ = density

Subscripts:

a = air
c = compressor
corr = corrected
f = fuel
g = exhaust gas
gp = gas generator
gpt = gas generator turbine
m = mechanical
opt = optimum
pt = power turbine
1 = inlet plane
2 = compressor diffuser exit plane
3 = gas generator turbine inlet plane
5 = power turbine inlet plane
7 = turbine diffuser exit plane.

References

[1] Diakunchak, I. S., 1991, ''Performance Degradation in Industrial Gas Turbines,'' ASME Report No. 91-GT-228.

[2] Brun, K., and Kurz, R., 1998, ''Measurement Uncertainties Encountered During Gas Turbine Driven Compressor Field Testing,'' ASME Report No. 98-GT-1.

[3] Bakken, L. E., and Skorping, R., 1996, ''Optimum Operation and Maintenance of Gas Turbines Offshore,'' ASME Report No. 96-GT-273.

[4] Bammert, K., and Sonnenschein, H., 1967, ''Der Einfluss verdickter und verduennter Turbinenschaufeln auf die Gittereigenschaften,'' Arch. Eisenhuettenwes., **38**, Heft 4, pp. 287–299.

[5] Bammert, K. et al., 1965, ''Ueber Korrosion, Erosion and Verschmutzung bei Gichtgasturbinen,'' VIK Bericht No. 164, Essen.

[6] Kind, R. J., Serjak, P. J., and Abbott, M. W. P., 1996, ''Measurements and Predictions of the Effects of Surface Roughness on Profile Losses and Deviation in a Turbine Cascade,'' ASME Report No. 96-GT-203.

[7] Boelcs, A., and Sari, O., 1988, ''Influence of Deposit on the Flow in a Turbine Cascade,'' ASME J. Turbomach., **110**, pp. 512–519.

[8] Kurz, R., 1995, ''Effects of Nonuniform Blade Pitch on the Flow Through an Annular Turbine Nozzle,'' Int. Rotating Machinery, **2**, No. 1, pp. 59–65.

[9] Kurz, R., 1991, ''Transonic Flow Through Turbine Cascades With 3 Different Pitch-to-Chord Ratios,'' Proc. 10. ISABE, Nottingham.

[10] Schmidt, K. J., 1992, ''Experimentelle und theoretische Untersuchungen zum instattionaeren Betriebsverhalten von Gasturbinentriebwerken,'' VDI, Duesseldorf.

[11] Milsch, R., 1971, ''Systematische Untersuchung ueber den Einfluss der Rauhigkeit von Verdichterschaufeln auf den Gitterwirkungsgrad,'' Ph.D. thesis, University of Hannover.

[12] Pinson, M., and Wang, T., 1994, ''Effects of Leading-Edge Roughness on Fluid Flow and Heat Transfer in the Transitional Boundary Layer of a Flat Plate,'' ASME Report No. 94-GT-326.

[13] Shabbir, A., Celestina, M. L., Adamczyk, J. J., and Strazisar, A. J., 1997, ''The Effect of Hub Leakage on Two High Speed Axial Flow Compressor Rotors,'' ASME Report No. 97-GT-346.

[14] Khalid, S. A., Khalsa, A. S., Waitz, I. A., Tan, S. C., Greitzer, E. M., Cumpsty, N. A., Adamczyk, J. J., and Marble, F. E., 1998, ''Endwall Blockage in Axial Compressors,'' ASME Report No. 98-GT-188.

[15] Frith, P. C., 1992, ''The Effect of Compressor Rotor Tip Crops on Turboshaft Engine Performance,'' ASME Report No. 92-GT-83.

[16] Singh, D., Hamed, A., and Tabakoff, W., 1996, ''Simulation of Performance Deterioration in Eroded Compressors,'' ASME Report No. 96-GT-422.

[17] Elrod, W. C., King, P. I., and Poniatowsky, E. M., 1990, ''Effects of Roughness, Freestream Turbulence, and Incidence Angle on the Performance of a 2-D Compressor Cascade,'' ASME Report No. 90-GT-208.

[18] Stalder, J. P., 1998, ''Gas Turbine Compressor Washing State of the Art Field Experiences,'' ASME Report No. 98-GT-420.

[19] Spakovszky, Z. S., Gertz, J. B., Sharma, O. P., Paduano, J. D., Epstein, A. H., and Greitzer, E. M., 1999, ''Influence of Compressor Deterioration on Engine Dynamic Behavior and Transient Stall Margin,'' ASME Report No. 99-GT-439.

[20] Boyle, R. J., 1994, ''Prediction of Surface Roughness and Incidence Effects on Turbine Performance,'' ASME J. Turbomach., **116**, pp. 745–751.

[21] Traupel, W., 1988, *Thermische Turbomaschinen*, Springer, Berlin.

[22] Radtke, F., and Dibelius, G., 1980, ''Reynoldszahleinfluss bei hochbelasteten axialen Turbinenbeschaufelungen,'' VDI-Berichte No. 361.

[23] Haq, I. U., Bu-Hazza, A. L., and Al-Baz, K., 1998, ''Multistage Centrifugal Compressor Fouling Evaluation at High Power Settings,'' ASME Report No. 98-GT-53.

[24] Cohen, H., Rogers, G. F. C., and Saravanamuttoo, H. I. H., 1996, *Gas Turbine Theory*, Longman, Harlow.

[25] Tarabrin, A. P., Schurovsky, V. A., Bodrov, A. I., and Stalder, J. P., 1998, ''Influence of Axial Compressor Fouling on Gas Turbine Unit Performance Based on Different Schemes and With Different Initial Parameters,'' ASME Report No. 98-GT-416.

[26] MacLeod, J. D., Taylor, V., and Laflamme, J. C. G., 1991, ''Implanted Component Faults and Their Effects on Gas Turbine Engine Performance,'' ASME Report No. 91-GT-41.

[27] Aker, G. F., and Saravanamuttoo, H. I. H., 1988, ''Predicting the Gas Turbine Performance Degradation due to Compressor Fouling Using Computer Simulation Techniques,'' ASME Report No. 88-GT-206.

Multiobjective Optimum Design of Rotor-Bearing Systems With Dynamic Constraints Using Immune-Genetic Algorithm

B.-K. Choi
Department of Mechanical
and Aerospace Engineering,
Arizona State University,
Tempe, AZ 85287-6106
e-mail: bgchoi@imap1.asu.edu

B.-S. Yang
School of Mechanical Engineering,
Pukyong National University,
San 100, Yongdang-dong, Nam-Ku,
Pusan 608-739, South Korea

In this paper, the combined optimization algorithm (immune-genetic algorithm) is proposed for multioptimization problems by introducing the capability of the immune system to the genetic algorithm. The optimizing ability of the proposed combined algorithm is identified by comparing the result of optimization with sharing genetic algorithm for the two-dimensional multipeak function. Also the combined algorithm is applied to minimize the total weight of the shaft and the transmitted forces at the bearings. The results show that the combined algorithm can reduce both the weights of the shaft and the transmitted forces at the bearing with dynamic constraints. [DOI: 10.1115/1.1338952]

Introduction

Many engineering problems can be formulated as optimization problems, which optimize some objective function. Meanwhile, most cases show the multiobjective function. It is not easy to solve such multiobjective functions by conventional optimization methods. Therefore a method for solving this is needed in the real world.

In the past year, the immune system, which has powerful abilities such as memory, recognition, and learning how to respond to invading antigens, has been applied to many engineering algorithms [1]. Some papers have shown the attempt that the immune system has made combined with genetic algorithm [2]. This paper proposes the combined optimization algorithm for multioptimization problems by introducing only the capability of the immune system that controls the proliferation of clones and suppresses the production of the antibody to the genetic algorithm [3,4].

In this paper, the optimizing ability of the proposed combined algorithm is identified by comparing the results of optimization with sharing genetic algorithm (SGA) [5] for the two-dimensional multipeak function, which have many local optimums. The combined algorithm is also applied to minimize the total weight of the shaft and the transmitted forces at the bearings. The inner diameter of the shaft and the bearing stiffness are chosen as the design variables. The results show that the combined algorithm can reduce both the weights of the shaft and the transmitted forces at the bearing with dynamic constraints.

Immune Algorithm

Immune System. The immune system is a complex network of organs containing cells that recognize foreign substances in the body and destroy them [6]. Figure 1 shows the production mechanism of antibodies in the immune system. Several types of cells compose the immune system such as stem cell, precursor cell, lymphocytes, antibodies, and antigen. The immune system produces a variety of antibodies and controls the production of the antibody combining the antigen and the antibody or two antigens. The immune system is well documented in many papers [1,2]. Therefore the details will be omitted here.

Contributed by the Structures & Dynamics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the S&D Division, Jan. 1999; final revision received by the ASME Headquarters, Sept. 28, 1999. Editor: H. D. Nelson.

Proposition of Combined Algorithm (IGA)

This section proposes a combined algorithm [immune-genetic algorithm (IGA)] by introducing only the capability of the immune system, which controls the proliferation of clones and suppresses the production of the antibody, to the genetic algorithm in order to prevent the initial convergence of the genetic algorithms and to search for the multioptimum solution. The abilities of self control and suppression added to the genetic algorithm are as follows [1]:

1 The differentiation mechanism of memory and suppressor cells
2 The suppressor mechanism of the production of antibodies

The antibody in the immune algorithm is the individual in the IGA. Figure 2 shows the flow chart of the IGA. The dotted line indicated in Fig. 2 is the additional part of the immune algorithm to simple genetic algorithm. The procedure of IGA is as follows:

Step 1. Production of the initial chromosome: This is the first process of IGA and it is executed only one time for the IGA procedure. In order to prevent the convergence of the local optimum according to the inclined distribution of the initial population, the uniform distribution method of genes is used. First, half of the population is initialized randomly and then the others are initialized with the complement of the produced binary code.

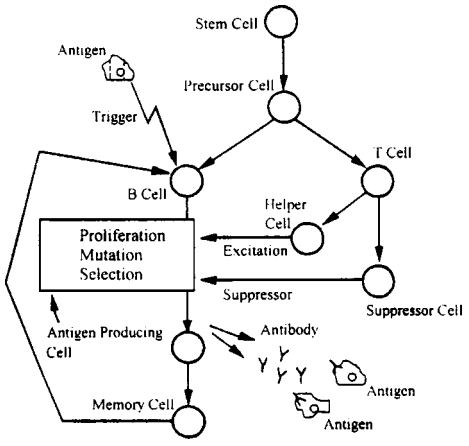


Fig. 1 The production mechanism of antibodies in the immune system

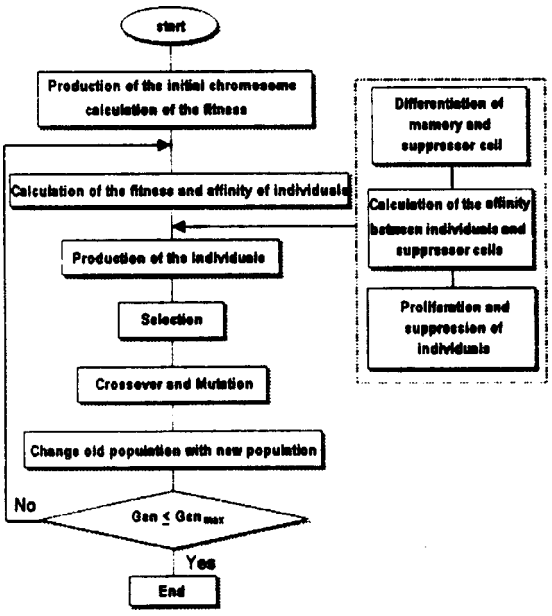


Fig. 2 The flow chart of IGA

Step 2. Calculation of affinity and fitness: The affinity $ay_{v,w}$ and the fitness are calculated by Eq. (1) and the objective function, respectively.

$$ay_{v,w}=1/(1+H_{v,w}) \tag{1}$$

where $H_{v,w}$ is the hamming distance between antibodies v and w. When $H_{v,w}$ equals zero, antibodies v and w are completely the same. $ay_{v,w}$ has a maximum value of 1 at this time.

Step 3. Differentiation of the memory cell and suppressor cell: The concentration of each individual is calculated. The individuals are dispersed to the memory cells and suppressor cells when the concentration c_v of the individual v exceeds over a threshold level T_c. The individuals are destroyed when the affinity with the suppressor cell exceeds T_{ac1}.

$$c_v=\frac{1}{N}\sum_{w=1}^{N} ac_{v,w}, \quad ac_{v,w}=\begin{cases}1: & ay_{v,w}\geqslant T_{ac1}\\ 0: & \text{otherwise}\end{cases} \tag{2}$$

where T_{ac1} is the limit of the similarity, $ac_{v,w}$ is the affinity when the affinity between the antibody v and w exceed T_{ac1} and N is the total number of the individuals. In this paper, 0.45 and 0.5 are used as T_c and T_{ac1}, respectively [1].

Step 4. Proliferation and suppression of the individual: The expected value e_v of the proliferation of the individual is defined by Eq. (3). In the individuals of the present generation, half of N are destroyed according to their level of fitness

$$e_v=f_v\prod_{s=1}^{S}(1-as_{v,s}^{k})\Big/\left(c_v\sum_{i=1}^{N} f_i\right),$$

$$as_{v,s}=\begin{cases}ay_{v,s}: & ay_{v,s}\geqslant T_{ac2}\\ 0: & \text{otherwise}\end{cases} \tag{3}$$

where f_v is the fitness of the individual v, S is the total number of the suppressor cell, k is the suppressor power, T_{ac2} is the limit of the similarity, and $as_{v,s}$ is the affinity when the affinity between the antibody v and the suppressor cell s exceeds T_{ac2}. 0.5 is used as T_{ac2} in this paper [1].

Equation (3) indicates that the individuals of higher fitness have a higher probability of remaining in the next generation than the individuals of higher concentration and the affinity with the suppressor cell. Equation (3) regulates the concentration and the variety of individuals in the population.

Step 5. Production of individual: The new individuals are randomly generated as much as the destroyed number in step 3.

Step 6. Crossover and mutation: This procedure can generate the individuals by a genetic reproduction operator such as mutation or crossover using the individuals that remained in step 4 and generated in step 5. Since the genetic operator is well documented in many papers [7], the details will be omitted here. In returning to step 2, the searching process is iterated until the final generation.

Numerical Result and Discussion

Optimization of Multipeak Function. First the optimization result of IGA is compared with that of SGA [5] for the two-dimensional multipeak functions which have many local optimums in order to identify the optimizing ability of the IGA. Equation (4) shows the objective function, called the Shubert function,

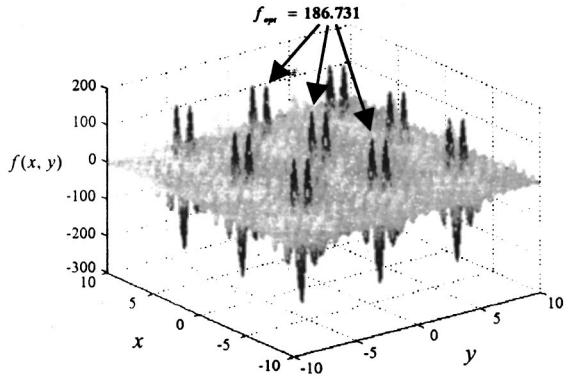


Fig. 3 Shubert function

Table 1 Parameters of SGA and IGA

	SGA [5]	IGA
Population size, N	50	35
Length of chromosome	15	20
Crossover probability, p_c	0.6	1.0
Mutation probability, p_m	0.005	0.2
Maximum generation, G	1000	1000

Table 2 Comparison of optimization results

Individual No.	SGA [5]			IGA		
	X	y	f	x	y	f
1	-1.4225	-7.0849	186.710	4.8583	5.4839	186.729
2	-1.4225	-7.0849	186.710	-7.087	4.8583	186.704
3	-1.4225	-7.0873	186.682	4.8583	-7.087	186.704
4	-1.4225	-7.0873	186.682	5.4839	-7.7713	186.686
5	-1.4249	-7.0898	186.644	-7.7126	5.4839	186.686
6	-1.4225	-7.0769	186.619	-7.7126	-7.087	186.661
7	-1.4420	-7.0849	186.067	5.4839	-1.4174	186.589
8	4.8253	-7.0837	184.258	-1.4174	5.4839	186.589
9	4.8253	-7.0837	184.258	4.8583	-0.7918	186.570
10	-7.0672	-1.4573	183.774	-0.7918	4.8583	186.570
11	-7.0672	-1.4573	183.774	-7.087	-1.4174	186.564
12	-1.4426	-0.8341	183.510	-1.4174	-7.087	186.564
13	-0.8176	-1.4591	183.427	-7.7126	-0.7918	186.521
14	-0.8176	-1.4591	183.427	-0.7918	-1.4174	186.431
15	-0.8176	-1.4591	183.427	-1.4174	-0.7918	186.431
16	-0.8176	-1.4591	183.427	-7.0674	-7.7126	186.120
17	-7.0672	-1.4597	183.403	5.4643	4.8583	185.970
18	-7.0672	-1.4597	183.403	-0.8113	-7.7322	185.150

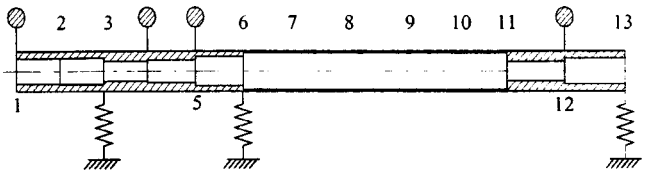


Fig. 4 Schematic of the rotor-bearing system

which is frequently used for the identification of multipeak function optimization. The Shubert function has the local optimum in the number of 760 and the global optimum in the number of 18, $f_{opt}=186.731$, among the local optimum. Figure 3 shows the Shubert function. Table 1 shows the parameters of SGA and IGA

$$f(x,y)=-\left[\sum_{i=1}^{5} i\cos\{(i+1)x+i\}\right]\times\left[\sum_{i=1}^{5} i\cos\{(i+1)y+i\}\right]$$
$$(-10\leqslant x,y\leqslant 10) \tag{4}$$

A comparison of the optimization result of sharing genetic algorithm and IGA is shown in Table 2. While sharing method does not find all of the global optimum, IGA can find all of the global optimum with good accuracy. Therefore, the optimization ability of IGA for the multipeak function is identified and demonstrated through the result of a numerical example.

Table 3 Initial and multiobjective optimization design (*a*) Inner radius of shaft element and bearing stiffness; (*b*) Critical speed, shaft weight, transmitted forces, maximum amplitude and bending stress

Multiobjective : W and F_3^b			
Shaft Element	Initial	WM [9]	IGA
	Inner Radius(*cm*)		
1	1.882	2.296	2.6890
2	1.940	2.369	2.6895
3	1.466	2.690	2.6770
34	1.660	2.618	2.6900
5	2.151	2.690	2.5745
6	2.690	2.690	2.6900
7	2.690	2.690	2.6900
8	2.690	2.690	2.6900
9	2.690	2.690	2.6900
10	2.690	2.690	2.6900
11	1.420	2.441	2.2165
12	1.880	2.287	2.6900
Bearing	Bearing Stiffness($N/m\times10^8$)		
1	0.035	0.035	0.035
2	1.27	1.28	1.74832
3	0.12	0.035	0.035

(a)

	Initial	WM	IGA
	Critical Speed (*rad/s*)		
2 nd	685.84	514.44	519.22
3 rd	2646.14	2365.15	2366.27
	Shaft Weight (*kg*)		
W	10.235	6.228	5.84656
	Transmitted Force (*N*)		
B1	13.562	9.303	8.86
B2	475.599	338.591	327.866
B3	461.366	146.158	141.018
	Maximum Bending Stress (*MPa*)		
MaxBS	45.29	22.688	23.54
	Maximum Amplitude (μm)		
MaxAmp	78.9	65.3	64.9

(b)

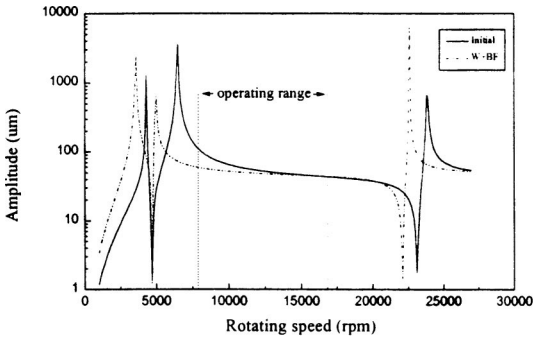


Fig. 5 Unbalance response of initial and optimum design

Optimization of Rotor-Bearing System. The rotor-bearing system [8,9] used in this paper consists of a single spool and three bearings located at stations 3, 6, and 13. Figure 4 shows a schematic of the rotor-bearing system. The fourth disk has unbalance $U=0.00108$ kg m. The details of the rotor configuration are well documented in Shiau's paper. Also, the same constraints and design variables are used in order to compare the optimization result of Shiau's paper [9]. The constraints on the critical speeds are taken as follows:

$$g_1(\mathbf{d})=\omega_2^c-691.67\ \text{rad/s}\leqslant 0,\quad g_2(\mathbf{d})=2301\ \text{rad/s}-\omega_3^c\leqslant 0 \tag{5}$$

The limitations on bending stress and unbalance response are also considered. The behavior constraints are as follows:

$$g_3(\mathbf{d})=150\ \text{MPa}-\sigma^*\leqslant 0,\quad g_4(\mathbf{d})=100\ \mu\text{m}-\delta^*\leqslant 0 \tag{6}$$

where σ^* and δ^* represent the allowable stress and allowable steady-state response, respectively. The inner radius of the shaft elements (r_{Ii}) and the bearing stiffness (k_i^b) are primarily taken as the design variables. The side constraints of the design variables are given by

$$0.0142\ \text{m}\leqslant r_{Ii}\leqslant 0.0269\ \text{m},\quad i=1-12 \tag{7}$$

$$3.50\times10^6\ \text{N/m}\leqslant k_i^b\leqslant 1.75\times10^8\ \text{N/m},\quad i=1-3 \tag{8}$$

The total shaft weight and transmitted force at the third bearing are considered as the objectives to be minimized, simultaneously. In order to identify the merit of the IGA, the optimization result is compare to Shiau and Chang's result [9] using the weighting method (WM). Table 3 shows a comparison of the optimization result. It is indicated that the optimum weight and transmitted force of the IGA are smaller than that of WM, satisfying all constraints. This implies that the IGA finds more exact solutions than WM using the characteristic of IGA: the diversity of the indi-

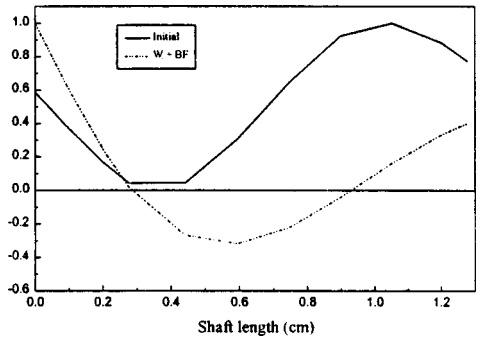


Fig. 6 Mode shape of initial and optimum design (2nd mode)

vidual and the production of efficient individuals. Figures 5 and 6 show the unbalanced response and mode shape corresponding to the initial and optimum design. In Fig. 5, the second critical speed of the optimum rotor is decreased in order to obtain separation margin and the unbalanced response of the optimum rotor also decreased in the operating range. Comparing the mode shape in Fig. 6, the second mode changed more because the response of the third bearing must decrease in order to reduce the transmitted force of the third bearing.

Conclusions

This paper proposes the combined optimization algorithm (IGA) for multoptimization problems by introducing only the capability of the immune system, which controls the proliferation of clones and suppresses the production of the antibody to the genetic algorithm. In this paper, the optimization result of IGA is compared with that of the sharing genetic algorithm for the two-dimensional multipeak function, which have many local optimums in order to identify the optimizing ability of the proposed IGA. For a comparison of the optimization result, IGA can find all of the global optimum with much accuracy, while the sharing method does not find all of the global optimum. Also, the combined algorithm is applied the multiobjective optimization of the rotor-bearing system to demonstrate the merit of the combined algorithm. For optimization of the rotor-bearing system, the results show that the IGA can reduce both the weights of the shaft and the transmitted forces at the bearing with dynamic behavior constraints. A comparison of the results between the IGA and WM shows that the IGA finds more exact solutions than the WM using the characteristic of IGA.

References

[1] Mori, K., Tsukiyama, M., and Fukuda, T., 1996, ''Multi-Optimization By Immune Algorithm with Diversity and Learning,'' *ICMAS'96*, Second International Conference on Multiagent Systems, Kyoto, Japan, pp. 118–123.

[2] Isao, T., Seiichi, K., and Hironori, H., 1997, ''An Evolutionary Optimization based on the Immune System and its Application to the VLSI Floorplan Design Problem,'' Trans. Inst. Electr. Eng. Jpn., Part C, **117**, pp. 821–827.

[3] Davis, L. ed., 1991, *Handbook of Genetic Algorithms*, Van Nostrand Reinhold, New York.

[4] Goldberg, D. E., 1989, *Genetic Algorithms in Search, Optimization & Machine Leaning*, Addision Wesley, New York.

[5] Shima, T., 1995, ''Global Optimization by a Niche Method for Evolutionary Algorithm,'' J. Systems Control Information, **8**, pp. 94–96 (in Japanese).

[6] Ivan, M. R., Jonathan, B., and David, M., 1989, *Immunology*, Grower Medical Publishing, New York.

[7] Choi, B. G., and Yang, B. S., 2000, ''Optimum Shape Design of Rotor Shafts Using Genetic Algorithm,'' J. Vibration Control, **6**, pp. 207–222.

[8] Shiau, T. N., and Hwang, J. L., 1990, ''Optimum Weight Design of a Rotor Bearing System with Dynamic Behavior Constraints,'' Trans. ASME: J. Eng. Gas Turbines Power, **112**, pp. 454–462.

[9] Shiau, T. N., and Chang, J. R., 1993, ''Multiobjective Optimization of Rotor-Bearing System with Critical Speed Constraints,'' Trans. ASME: J. Eng. Gas Turbines Power, **115**, pp. 246–255.

Reduced-Order Nonlinear Dynamic Model of Coupled Shaft-Torsional and Blade-Bending Vibrations in Rotors

B. O. Al-Bedoor
Assistant Professor,
Mechanical Engineering Department,
King Fahd University of Petroleum & Minerals,
KFUPM Box 841, Dhahran 31261, Saudi Arabia

In this study, a reduced-order nonlinear dynamic model for shaft-disk-blade unit is developed. The multibody dynamic approach with the small deformation theory for both blade-bending and shaft-torsional deformations is adopted. The equations of motion are developed using Lagrange's equation in conjunction with the assumed modes method (AMM) for approximating the blade transverse deflection. The model showed strong coupling between the blade bending and shaft torsional vibrations in the form of inertial nonlinearity, modal coupling, stiffening, softening, and parametric excitations. The model is suitable for extensive parametric studies for predesign stage purposes as well as for diagnostics of rotor malfunctions, when blade and shaft torsional vibration interaction is suspected. [DOI: 10.1115/1.1341203]

1 Introduction

The problem of modeling blade vibrations in turbo-machinery has received extensive research efforts due to the difficulties encountered in measuring these vibrations. The objective of a mathematical model is to aid in designing a blade that exhibits reduced vibration problems. Many blade and rotor designers used to adopt the assumption of no coupling between the blade vibrations and the rotor dynamics. Blades are driven via shafts and couplings that have inherent torsional flexibility and there exists a direct coupling between the driven blades and the driving shaft and coupling. Furthermore, there is an increasing need for operating rotor systems at higher speeds with more flexible components. Due to the aforementioned facts, in-depth studies on the vibration of a rotating blade that account for the shaft/coupling torsional flexibility are highly in demand.

The effects of rotation parameters on the natural frequencies and mode shapes of a rotating beam were first reported by Shilhansi [1] and Pruelli [2]. These studies showed that increasing the rotation speed strengthens the beam and results in higher natural frequencies. Likins et al. [3] reported a study on the mathematical modeling of spinning elastic bodies, wherein they highlighted the existence of beam stiffening due to the axial shortening. Kaza and Kvaternik [4] reported results of a study on the nonlinear flap-lag-axial equations of a rotating beam. Yokoyama [5] investigated the effect of shear deformations on the free vibration characteristics of rotating beams, using the finite-element method and Timoshenko beam theory. In the aforementioned studies, the effect of rotation was taken care of as a kinematic variable in the form of angular velocity and angular acceleration employed in the elastic equations, which are solved for the natural frequencies and mode shapes. Al-Bedoor and Khulief [6] reported a general case, when the beam is sliding through a rotating hub. This study considered both the elastic degrees of freedom and the reference rigid-body motions in their coupled format.

Srinivasan [7] reported a survey on the vibrations of bladed disk assemblies. In this survey, he highlighted the importance of modeling the disk-blade vibrations and classified these vibrations into two main categories; namely, structure-induced vibrations and aeroelastic-induced vibrations. The survey was focused on the structural-induced vibrations and their modeling. The effect of stager angle on the inertial and elastic coupling in bladed disks was reported by Crawley and Mokadam [8]. Tang and Wang [9] reported a model for rotor fuselage dynamic analysis. Loewy and Khader [10] studied the effect of flexural shaft flexibility on the dynamics of bladed-disk assemblies. Crawley et al. [11] reported analytical and experimental results on the coupled blade-disk-shaft whirl of a cantilevered turbofan. The effect of shaft torsional flexibility on the vibrations of rotating blades was not considered in the previously cited studies. Okabe et al. [12] highlighted the necessity for modeling both blade bending and shaft torsional deformations in turbomachinery. They introduced an equivalent reduced-order model that coupled shaft torsional and blade tangential vibrations. Their model adopted the modal synthesis procedure, wherein the blade was modeled as a simple mass-spring subsystem and the shaft as the other discrete subsystem. The two subsystems were coupled and the natural frequencies were obtained. Okabe et al. [12] compared their model predictions to actual measurements and reported they are in close agreement. Huang and Ho [13] reported results of a study on the coupled shaft torsion and blade-bending vibrations of a rotating shaft-disk-blade unit. The shaft torsional and blade-bending deformations were modeled, separately, using the assumed modes method. They used the weighted residual method and the receptances at the connection between the disk and the blade to couple the shaft-torsional and blade-bending deformations. Al-Bedoor [14] based on the multibody dynamic approach, developed a coupled model for shaft-torsional and blade-bending vibrations in rotors. The model employed the finite-element method to discretize the blade deformations. The study identified the nonlinear interaction and the destabilizing effect that the blade and shaft could introduce to excite each other. Due to the difficulty encountered in quantifying the nature of nonlinear coupling when the finite-element method is used, a reduced-order nonlinear dynamic model for shaft-torsional and blade-bending vibrations that adopts the assumed

Contributed by the XXX Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the XXX Division, April 28, 2000; final revision received by the ASME Headquarters May 16, 2000. Editor: H. D. Nelson.

modes method (AMM) for approximating blade deformations is needed. The AMM produces uncoupled blade equations and enables monitoring the vibration of each blade mode separately.

The present work is devoted to developing a coupled and reduced-order dynamic model for blade-bending and shaft-torsional vibrations in shaft-disk-blade units. The multibody dynamic approach is followed in developing the dynamic model through using body axes attached to different system components. The angular transformations using the total angle of system rotation, θ, and the small angle of torsional deformations, ψ, are utilized. The small deformation theory for blade-bending and shaft-torsional deformations is adopted. The model is developed using the Lagrangian dynamics in conjunction with the AMM for approximating blade deformations. The equations of motion are represented in a compact matrix form, wherein all the nonlinear terms are interpreted based on their associated stiffening, softening, damping, and inertial effects.

2 The Dynamic Model

A schematic diagram of a shaft-disk-blade system driven by an electrical motor is shown in Fig. 1. The disk is assumed to be rigid and the flexible blades are attached radially to the disk as shown. The beam is assumed to be inextensible and the Euler-Bernoulli beam theory is adopted. The beam is discretized using the AMM and the mode shapes of a cantilever beam are used. The model adopts the small deformation theory for both blade-bending and shaft-torsional deformations. The coordinate systems used in developing the model are shown in Fig. 2, wherein *XY* is the inertial reference frame, $x^m y^m$ is a body coordinate system of the motor shaft, $x^d y^d$ is a body coordinate system of the disk, and $x^b y^b$ is the blade coordinate system that is attached to the root of the blade such that x^b is always directed along the blade centerline.

Kinetic Energy Expressions. The kinetic energy of the shaft-disk-blade unit is made up of the motor kinetic energy T_M, the disk kinetic energy T_D, and the blade kinetic energy T_B. The shaft inertia is lumped into the disk and motor inertia.

The motor and the disk kinetic energy expressions can be written, respectively, as

$$T_M = \tfrac{1}{2} J_M \dot{\theta}^2$$
$$T_D = \tfrac{1}{2} J_D (\dot{\theta} + \dot{\psi})^2 \tag{1}$$

To develop the kinetic energy expression for the blade, the deformed configuration of the shaft-disk-blade system, shown in Fig. 2, is used. The global position vector of a material point, *P*, on the blade can be written as

$$R_P = [A(\theta)][B(\psi)] r_P^d \tag{2}$$

where r_P^d is the position vector of point *P* in the disk coordinate system $x^d y^d$, $[A(\theta)]$ is the rotational transformation matrix from the motor coordinate system to the inertial reference frame, *XY*, and $[B(\psi)]$ is the rotational transformation matrix from the disk coordinate system, $x^d y^d$, to the motor coordinate system. The position vector of the material point *P* in the $x^d y^d$ coordinate system can be written in the form

$$r_P^d = (x + R_D) i + u(x,t) j \tag{3}$$

where R_D is the disk radius, at which the root of the blade is attached, x and $u(x,t)$ are the axial position of the material point along the blade and its deflection measured with respect to the blade coordinate system, $x^b y^b$. The rotational transformation matrices $[A(\theta)]$ and $[B(\psi)]$ can be represented, respectively, as

$$[A(\theta)] = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \tag{4}$$

$$[B(\psi)] = \begin{bmatrix} 1 & -\psi \\ \psi & 1 \end{bmatrix} \tag{5}$$

where θ represents the motor rigid-body rotation and ψ represents the torsional deformation angle measured with respect to the motor coordinate system, $x^m y^m$. The transformation matrix $[B(\psi)]$ is linearized based on the assumption of small torsional deformations. It should be noted that no small motion linearization is used for the system rigid-body rotation transformation matrix, $[A(\theta)]$, and its derivatives.

The velocity vector of the material point in the inertial reference frame can be obtained by differentiating Eq. (2) as follows:

$$\dot{R}_P = [A(\theta)][B(\psi)]\dot{r}_P^d + \dot{\theta}[A_\theta(\theta)][B(\psi)] r_P^d + \dot{\psi}[A(\theta)][B_\psi(\psi)] r_P^d \tag{6}$$

where $[A_\theta]$ and $[B_\psi]$ represent the derivatives $[dA/d\theta]$ and $[dB/d\psi]$, respectively.

Upon substituting $[A_\theta]$, $[B_\psi]$, r_P^d, and $\dot{r}_P^d$ into Eq. (6), the velocity vector of the material point *P* in the inertial reference frame can be represented in the form

$$\dot{R}_P = \begin{Bmatrix} -\alpha\cos\theta - \beta\sin\theta \\ \beta\cos\theta - \alpha\sin\theta \end{Bmatrix} \tag{7}$$

where

$$\alpha = \dot{\theta}[\psi(x + R_D) + u(x,t)] + \dot{\psi}u(x,t) + \psi\dot{u}(x,t)$$
$$\beta = \dot{\theta}[(x + R_D) - \psi u(x,t)] + \dot{\psi}(x + R_D) + \dot{u}(x,t) \tag{8}$$

The kinetic energy of the blade can be found from

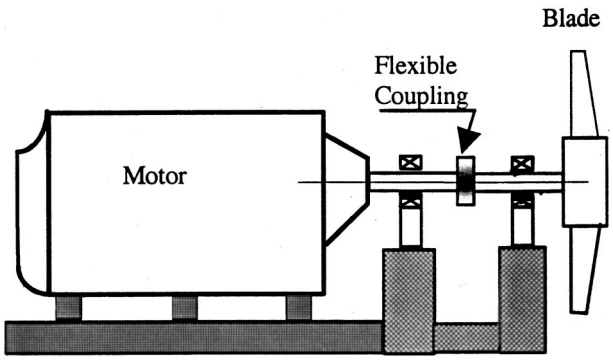


Fig. 1 Schematic diagram of blade-disk-shaft system

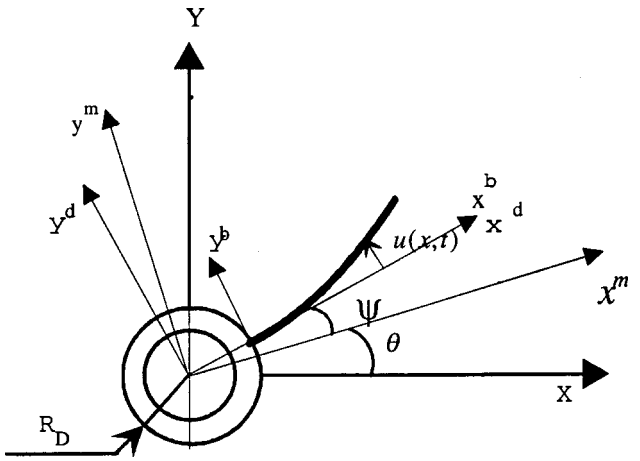


Fig. 2 System deformed configuration and coordinate system

$$T_B = \tfrac{1}{2}\int_0^L \rho \dot{R}_P^T \dot{R}_P dx \tag{9}$$

where ρ is the blade mass per unit length and L is the blade length.

Substituting Eq. (7) into Eq. (9), the blade kinetic energy expression is found in the form

$$\begin{aligned} T_B = {} & \tfrac{1}{2}\dot{\theta}^2\psi^2 \int_0^L \rho(x+R_D)^2 dx + \tfrac{1}{2}(\dot{\theta}+\dot{\psi})^2 \int_0^L \rho(x+R_D)^2 dx \\ & + \tfrac{1}{2}[(\dot{\theta}+\dot{\psi})^2 + \psi^2\dot{\theta}^2] \int_0^L \rho u^2 dx \\ & + \tfrac{1}{2}(1+\psi^2) \int_0^L \rho \dot{u}^2 dx + \psi\dot{\psi} \int_0^L \rho u \dot{u} dx \\ & + [\dot{\theta}(1+\psi^2) + \dot{\psi}] \int_0^L \rho(x+R_D)\dot{u} dx \end{aligned} \tag{10}$$

Now the total kinetic energy expression of the system can be written as follows:

$$T = T_M + T_D + T_B \tag{11}$$

Potential Energy Expressions. The system potential energy is made up of the blade-bending strain energy, V_B, the torsional strain energy, V_T, and the potential energy of the axial shortening due to transverse deformations and the motion-generated inertial forces, V_A. The elastic blade-bending strain energy for a blade with flexural rigidity $EI(x)$ is given by

$$V_B = \tfrac{1}{2}\int_0^L EI(x) \frac{\partial^2 u}{\partial x^2} dx \tag{12}$$

The torsional elastic potential energy stored in the flexible coupling is given by

$$V_T = \tfrac{1}{2} k_T \psi^2 \tag{13}$$

The axial shortening due to transverse deformations in conjunction with the radial inertial forces contributes to the system elastic potential energy by the known axial shortening potential energy. The axial shortening due to blade-bending deformations can be approximated [6] by

$$d\delta \cong -\frac{1}{2}\left(\frac{\partial u}{\partial x}\right)^2 dx \tag{14}$$

The inertial force on the material point P of the blade that results from the rotational motion of the blade can be expressed in the form

$$F_P = \int_x^L \rho(R_D + x)(\dot{\theta}+\dot{\psi})^2 dx \tag{15}$$

Now, the work that results from the axial shortening under the effect of the inertial forces can be called the axial shortening potential energy and can be written in the form

$$V_A = \int_0^L F_P d\delta \tag{16}$$

Evaluating the integral of Eq. (15) and substituting for F_P and $d\delta$ into Eq. (16), the axial shortening potential energy becomes

$$\begin{aligned} V_A = {} & \tfrac{1}{2}(\dot{\theta}+\dot{\psi})^2 \left\{ \tfrac{1}{2}\int_0^L \rho(L^2 - x^2)\left(\frac{\partial u}{\partial x}\right)^2 dx \right. \\ & \left. + R_D \int_0^L \rho(L-x)\left(\frac{\partial u}{\partial x}\right)^2 dx \right\} \end{aligned} \tag{17}$$

The system potential energy is

$$V = V_B + V_T + V_A \tag{18}$$

The Assumed Mode Method (AMM). The assumed mode method is used in approximating the blade elastic deformation, $u(x,t)$, shown in Fig. 2, relative to the blade coordinate system, as follows:

$$u(x,t) = \sum_{i=1}^{N} \phi_i(x) q_i(t) \tag{19}$$

where N is the number of modes, q_i is the vector of modal coordinates, which is time-dependent, and ϕ_i is the vector of the assumed mode shapes, which are those of a cantilever beam in this study. The mode shapes of a cantilever beam, which are normalized with respect to the mass per unit length of the blade, can be written in the form

$$\phi_i(x) = \frac{1}{\sqrt{\rho L}}[\cosh \lambda_i x - \cos \lambda_i x - \alpha_i(\sinh \lambda_i x - \sin \lambda_i x)] \tag{20}$$

where ρ and L are the blade mass per unit length and the blade length, respectively. The parameter λ_i is found from the solution of the transcendental equation,

$$\cos \lambda_i L \cosh \lambda_i L + 1 = 0 \tag{21}$$

and the parameter α_i is given by

$$\alpha_i = \frac{\sinh \lambda_i L - \sin \lambda_i L}{\cosh \lambda_i L + \cos \lambda_i L}$$

Upon substituting the AMM approximation for the blade deformations in the Lagrangian expression, a number of coefficient vectors and matrices have developed. The coefficient vectors and matrices can be expressed as follows:

$$[a] = \int_0^L \rho x[\phi(x)] dx \tag{22}$$

$$[b] = \int_0^L \rho[\phi(x)] dx \tag{23}$$

$$[I] = \int_0^L \rho[\phi(x)]^T[\phi(x)] dx \tag{24}$$

$$[k] = \int_0^L \rho[\phi''(x)]^T[\phi''(x)] dx \tag{25}$$

$$[k_{s1}] = \int_0^L \rho(L-x)[\phi'(x)]^T[\phi'(x)] dx \tag{26}$$

$$[k_{s2}] = \int_0^L \rho(L^2 - x^2)[\phi'(x)]^T[\phi'(x)] dx \tag{27}$$

$$[k_s] = R_D[k_{s1}] + \tfrac{1}{2}[k_{s2}] \tag{28}$$

$$[h] = [a] + R_D[b] \tag{29}$$

where $\phi' = d\phi/dx$ and $[I]$ is the identity modal mass matrix that has developed in this form as a result of the mode shape orthogonality condition. $[k]$ is the diagonal elastic stiffness matrix. The effect of stiffening due to blade rotation appeared in the matrices $[k_{s1}]$, $[k_{s2}]$, and $[k_s]$. R_D is the disk radius. The vectors $[a]$, $[b]$ represent the inertia coupling vectors between the system reference rotational motion, the system torsional flexibility degree of freedom, and the blade-bending degrees of freedom.

The Equations of Motion. By substituting the discretized forms of the kinetic and potential energy expressions, the Lagrangian of the system can be written in the following form:

$$\Gamma = \tfrac{1}{2}J_M\dot{\theta}^2 + \tfrac{1}{2}J_D(\dot{\theta}+\dot{\psi})^2 + \tfrac{1}{2}J_B[\dot{\theta}^2\psi^2 + (\dot{\theta}+\dot{\psi})^2] + \tfrac{1}{2}[\dot{\theta}^2\psi^2 + (\dot{\theta}+\dot{\psi})^2]\{q\}^T[I]\{q\} + \tfrac{1}{2}(1+\psi^2)\{\dot{q}\}^T[I]\{\dot{q}\} + \psi\dot{\psi}\{q\}[I]\{\dot{q}\} + [\dot{\theta}(1+\psi^2)+\dot{\psi}][a]\{\dot{q}\} + [\dot{\theta}(1+\psi^2) + \dot{\psi}]R_D[b]\{\dot{q}\} - \frac{1}{2}\frac{EI}{\rho L^4}\{q\}^T[k]\{q\} - \tfrac{1}{2}k_T\psi^2 - \tfrac{1}{2}(\dot{\theta}+\dot{\psi})^2R_D\{q\}^T[k_{s1}]\{q\} - \tfrac{1}{4}(\dot{\theta}+\dot{\psi})^2\{q\}^T[k_{s2}]\{q\} \tag{30}$$

Now using the discretized form of the system Lagrangian, performing the required differentiation and mathematical manipulation for the system degrees of freedom θ, ψ, and $\{q\}$, the system equations of motion can be represented in the following compact matrix form:

$$\begin{bmatrix} m_{\theta\theta} & m_{\theta\psi} & [m_{\theta q}] \\ m_{\psi\theta} & m_{\psi\psi} & [m_{\psi q}] \\ [m_{q\theta}] & [m_{q\psi}] & [m_{qq}] \end{bmatrix} \begin{Bmatrix} \ddot{\theta} \\ \ddot{\psi} \\ \{\ddot{q}\} \end{Bmatrix} + \begin{bmatrix} c_{\theta\theta} & c_{\theta\psi} & [c_{\theta q}] \\ c_{\psi\theta} & c_{\psi\psi} & [0] \\ [c_{q\theta}] & [0] & [c_{qq}] \end{bmatrix} \times \begin{Bmatrix} \dot{\theta} \\ \dot{\psi} \\ \{\dot{q}\} \end{Bmatrix} + \begin{bmatrix} 0 & 0 & [0] \\ 0 & k_{\psi\psi} & [0] \\ [0] & [0] & [k_{qq}] \end{bmatrix} \begin{Bmatrix} \theta \\ \psi \\ \{q\} \end{Bmatrix} = \begin{Bmatrix} F_\theta \\ F_\psi \\ \{F_q\} \end{Bmatrix} \tag{31}$$

where

$$m_{\theta\theta} = J_M + J_D + J_B(1+\psi^2) + \{q\}^T[(1+\psi^2)[I] - [k_s]]\{q\} \tag{32}$$

$$m_{\psi\psi} = m_{\theta\psi} = m_{\psi\theta} = J_D + J_B + \{q\}^T[[I] - [k_s]]\{q\} \tag{33}$$

$$[m_{\theta q}] = [m_{q\theta}]^T = (1+\psi^2)[h] \tag{34}$$

$$[m_{\psi q}] = [m_{q\psi}]^T = [h] + \psi\{q\}^T[I] \tag{35}$$

$$[m_{qq}] = (1+\psi^2)[I] \tag{36}$$

$$c_{\theta\theta} = 2\psi\dot{\psi}[J_B + \{q\}^T\{q\}] + 2\{q\}^T((1+\psi^2)[I] - [k_s])\{\dot{q}\} \tag{37}$$

$$c_{\psi\psi} = c_{\theta\psi} = c_{\psi\theta} = 2\{q\}^T([I] - [k_s])\{\dot{q}\} \tag{38}$$

$$[c_{\theta q}] = [c_{q\theta}]^T = 2\psi\dot{\psi}[h] \tag{39}$$

$$[c_{qq}] = 2\psi\dot{\psi}[I] \tag{40}$$

$$k_{\psi\psi} = k_T - \dot{\theta}^2(J_B + \{q\}^T\{q\}) - 2\dot{\theta}[h]\{\dot{q}\} \tag{41}$$

$$[k_{qq}] = \frac{EI}{\rho L^4}[k] + (\dot{\theta}+\dot{\psi})^2[[k_s] - [I]] + (\dot{\psi}^2 - \psi^2\dot{\theta}^2)[I] \tag{42}$$

$$J_B = \rho\left(\frac{L^3}{3} + R_D L^2 + R_D^2 L\right) \tag{43}$$

The equations of motion (31) are arranged in such a way that all nonlinear terms are in a position to contribute to the system inertia, damping, and stiffness. It is noteworthy to mention that structural damping can be easily accommodated in the damping matrix of the model. The first coefficient matrix in Eq. (31) is the system inertia matrix that contains nonlinear functions of the system motion parameters. The entry $m_{\theta\theta}$ described by Eq. (32) is the system reference rotational inertia, which is a nonlinear function of the shaft torsional and the blade-bending deformations. The entry $m_{\psi\psi}$ is the torsional degree of freedom rotational inertia that is a nonlinear function of the blade deflection. In the same equation, the terms $m_{\theta\psi}, m_{\psi\theta}$ represent the inertial coupling between the rigid-body rotation and the torsional degree of freedom. The inertia coupling between the system rigid-body rotation and the blade deformation is given by the vector $[m_{\theta q}]$, Eq. (34), which is a function also of the torsional deformation angle. The inertia coupling between the torsional degree of freedom and the blade deflection coordinates is shown to be a function of both ψ and $\{q\}$ in Eq. (35). The modal inertia matrix, Eq. (36), of the blade is affected by the square of the system torsional deformation.

The second coefficient matrix in Eq. (31) is the system-generalized damping matrix, which is developed as a result of coupling between the system degrees of freedom. All nonzero entries of the system-damping matrix are nonlinear functions of the shaft and blade deformation parameters; namely, ψ, $\dot{\psi}$, $\{q\}$, and $\{\dot{q}\}$. Structural damping can be added to this generalized damping matrix.

The third coefficient matrix in Eq. (31) is the system-generalized stiffness matrix where the entries corresponding to the system-generalized torsional stiffness, $k_{\psi\psi}$, and the generalized blade-bending stiffness, $\lfloor k_{qq} \rfloor$, are the only nonzero ones. The generalized stiffness $k_{\psi\psi}$ is given in Eq. (41), wherein the torsional system is softened by two terms. The first term combines the effects of the system rotational speed, the blade inertia, and the blade deformations. The second softening term is the Coriolis-like term, which combines the system rotational speed and the blade-bending modal velocity. The developed generalized torsional stiffness provides an interesting expression that relates the torsional stiffness, blade inertia, and system rotational velocity for stability purposes. The blade-bending generalized stiffness $\lfloor k_{qq} \rfloor$ is shown, in Eq. (42), to be under the combined stiffening and softening effects. Finally, the right-hand-side vector is the forcing vector corresponding to each degree of freedom of the system.

3 Numerical Simulation and Discussion

The computational process started with evaluating the coefficient vectors and matrices, Eqs. (22)–(27), for the first five cantilever beam mode shapes. The Gauss-quadrature 16-point integration scheme was used.

The system of nonlinear second-order equations, Eq. (31), is simulated using a multistep variable order predictor-corrector algorithm. The dimension and material properties of the blade-shaft-disk system are given in Table 1. For the considered system data, the uncoupled natural frequencies are 78.2 Hz for the shaft torsional degree of freedom and 26 Hz, 163.3 Hz, 457.43 Hz, 896.38 Hz, and 1481.73 Hz for the first five bending modes of the blade, respectively. In the simulation of the system, structural damping in the form of Rayleigh proportional damping is used. Damping ratios $\zeta_T = 0.0005$ and $\zeta_B = 0.001$ are taken for the torsional and for each of the blade-bending modes, respectively.

The motor torque that is designed using the inverse dynamic procedure to rotate the system to reach a constant speed of 1000 RPM in 1.0 sec is shown in Fig. 3 and the resulting system speed-time relation is shown in Fig. 4. The resulting blade-tip deflection is shown in Fig. 5(*a*) with the corresponding frequency spectrum shown in Fig. 5(*b*). Figure 5(*a*) shows that the blade tip deflection exhibits sustained oscillations, which are expected to damp out as a result of the used structural damping. The blade tip deflection

Table 1 Blade-disk-shaft data

Property	Value
Blade length L	0.4 m
Blade mass per unit length, ρ	1.35 kg/m
Blade flexural rigidity, EI	$75N.m^2$
Disk Radius, R_D	0.05 m
Disk moment of inertia, J_D	$8.8\times10^{-3}kg.m^2$
Motor moment of inertia, J_M	$8.8\times10^{-2}kg.m^2$
Torsional stiffness k_T	9880 N.m/Rad

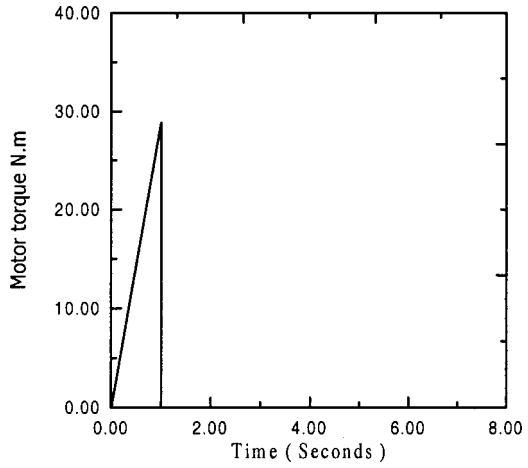


Fig. 3 Motor torque to rotate the system to a speed of 1000 RPM in 1 sec

frequency spectrum, Fig. 5(*b*), shows two frequency components in the blade vibration signal. The first frequency component is around 48 Hz and the second is around 290 Hz. These two components do not represent any of the blade natural frequencies or the coupling torsional natural frequency. This result and the result of sustained oscillation lead to a conclusion that there is a continuous parametric excitation of the blade. The coupling torsional vibration is shown in Fig. 6, wherein damped behavior can be seen as a result of the structural damping in the coupling.

In order to gain insight into the system behavior, the blade individual mode vibrations, corresponding to the blade tip deflection of Fig. 5, are shown in Figs. 7–11. Figures 7–11 show that all blade vibration modes are excited, almost to the same order of amplitude. This result stresses that there is a continuous parametric forcing excitation into the blade that comes from the coupling with the system torsional vibration. Moreover, when the system is simulated without any structural damping, the torsional vibration shows sustained oscillations, as expected. On the other hand, the blade vibration shows unstable behavior (growing vibration) when no damping is used for the blade vibration modes.

The sustained and unstable behavior can be explained by a greater understanding of the developed model. The model has inertial nonlinearities, as can be seen in Eq. (30), softening and stiffening effects depending on the system rotational velocity, and Coriolis effects. The modes are coupled through the system motion and the torsional vibration. Furthermore, the torsional vibration acts on the blade as base excitation, which under certain

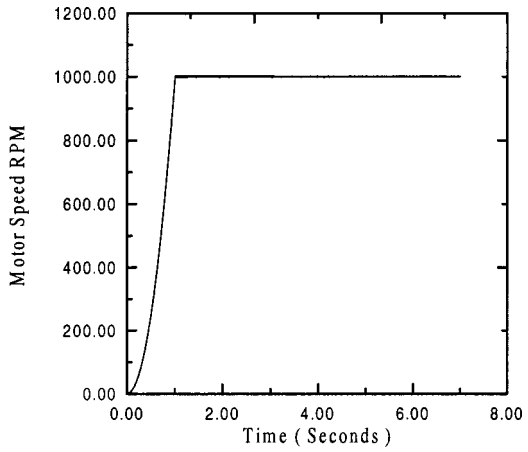


Fig. 4 Motor speed $\dot{\theta}$

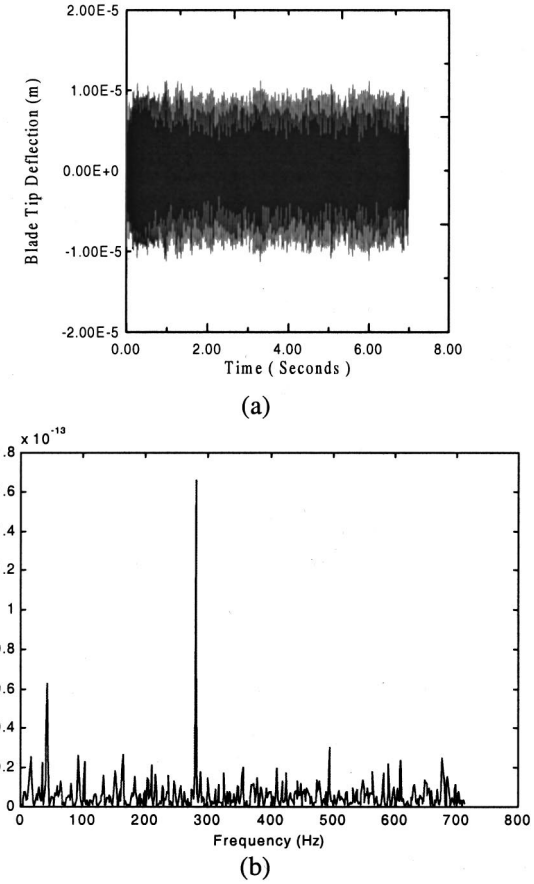


Fig. 5 (*a*) Blade deflection and (*b*) frequency spectrum

combination of system and motion parameters can act as a continuous excitation to the blade. Extracting the equation of motion of the first blade mode from Eq. (31), for constant rotational speed $\ddot{\theta}=0$, gives the equation

$$(h(1)+\psi q_1)\ddot{\psi}+\dot{q}_1+2\psi\dot{\psi}\dot{q}_1 + \left(\left[\frac{EI\lambda_1^4}{\rho L^4}+(\dot{\theta}+\dot{\psi})^2\right](k_s(1,1)-1)+(\dot{\psi}^2-\psi^2\dot{\theta}^2)\right)q_1=0 \tag{44}$$

Equation (44) is affected by the torsional degree of freedom ψ and its time derivatives through inertial input to the system in the form of base excitation as shown in the first term. Stiffening and

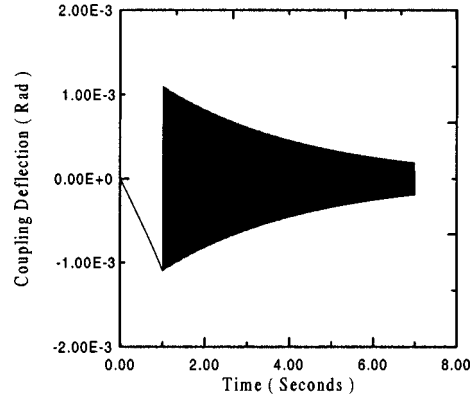


Fig. 6 Coupling torsional deflection

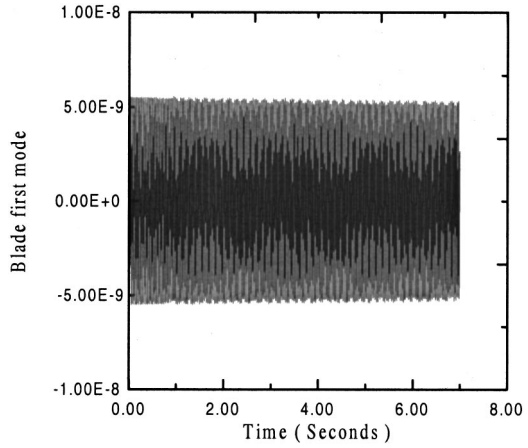


Fig. 7 Blade first bending mode vibrations

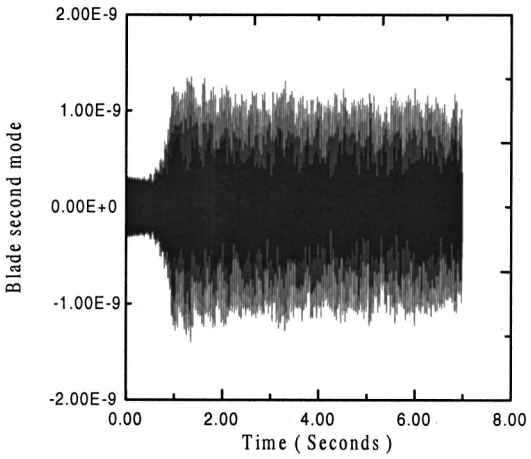


Fig. 8 Blade second bending mode vibrations

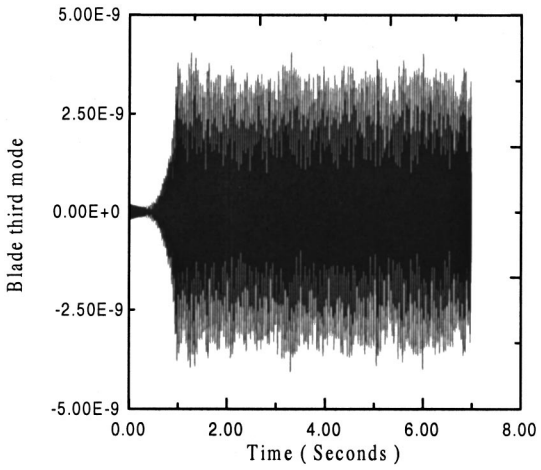


Fig. 9 Blade third bending mode vibrations

softening are shown by the terms associated with the stiffness. Moreover, when the system is excited by a torque for a short duration of time as shown in Fig. 3, the torsional degree of freedom will be responding sinusoidal with a frequency equal to the system torsional natural frequency. When this response, $\psi\dot{\psi}$, is substituted for in Eq. (44), the blade first bending mode will be exposed to sinusoidal damping that leads to parametric-like excitation [15]. This excitation together with the coupling between

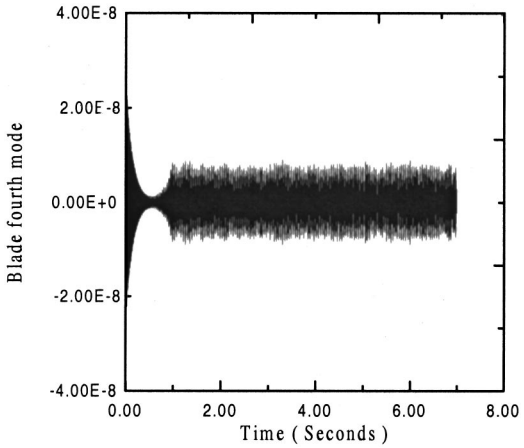


Fig. 10 Blade fourth bending mode vibrations

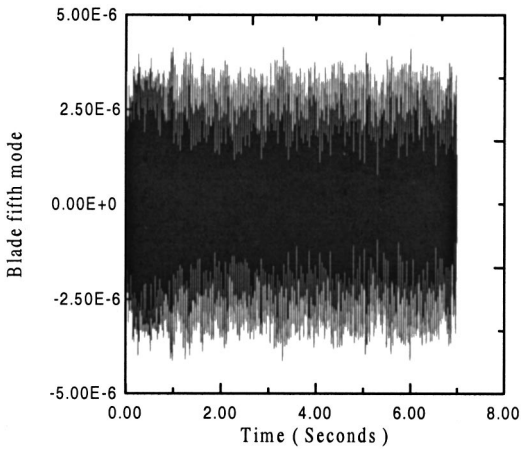


Fig. 11 Blade fifth bending mode vibrations

blade modes and the effect of base excitation are, most likely, the key factors for the shown unstable blade vibrations.

4 Conclusions

A reduced-order nonlinear dynamic model for the shaft-disk-blade unit is developed in this study. The multibody dynamic approach, in which coordinate systems are attached to different system components, is utilized. The motion angle θ and the small torsional deformation angle ψ transformations are used in developing the system energy expressions. The small deformation theory for both blade-bending and shaft-torsional deformations is adopted. The equations of motion are developed using Lagrange's equation in conjunction with the assumed modes method (AMM) for approximating the blade transverse deflection. The model showed strong coupling between the blade-bending and the shaft-torsional vibrations in the form of inertial nonlinearity, modal coupling, stiffening, softening, and parametric excitations. The model is suitable for extensive parametric studies for predesign stage purposes as well as for diagnostics of rotor malfunctions. Furthermore, more studies on the stability of shaft-blade-disk units using the developed model, as related to system design and operation parameters, are recommended.

Acknowledgments

The author acknowledges the support of King Fahd University of Petroleum and Minerals, Dhahran, Saudi Arabia, for this work.

Nomenclature

$A(\theta)$ = rotational transformation matrix
$B(\psi)$ = torsional transformation matrix
A_θ = derivative of rotational transformation matrix with respect to θ
B_ψ = derivative of torsional transformation matrix with respect to ψ
a, b = coefficient vectors
$c_{\theta\theta}$ = reference D.O.F. (degree of freedom) damping coefficient
$c_{\psi\psi}$ = torsional D.O.F. damping coefficient
$c_{\psi\theta}$ = damping coupling coefficient
c_{qq} = blade modal damping matrix
EI = blade flexural rigidity
$\{F\}$ = external moment/force vector
h = coefficient vector
I = identity matrix
J_B = blade moment of inertia
J_D = disk moment of inertia
J_M = motor moment of inertia
k_{qq} = blade-bending generalized stiffness matrix
k_T = torsional stiffness
k_{s1}, k_{s2}, k_s = axial shortening stiffness matrices
$k_{\psi\psi}$ = shaft-torsional generalized stiffness
L = blade length
$m_{\theta\theta}$ = reference D.O.F. generalized inertia
$m_{\psi\psi}$ = torsional D.O.F. generalized inertia
$m_{\psi\theta}$ = inertia coupling coefficient
m_{qq} = blade generalized inertia matrix
N = number of assumed modes
p = material point along the blade
q_i = modal coordinates
R_D = disk radius
T = system kinetic energy
T_B = blade kinetic energy
T_D = disk kinetic energy
T_M = motor kinetic energy
$u(x,t)$ = blade transverse deflection
V = system potential energy
V_A = blade axial shortening potential energy
V_B = blade-bending potential energy
V_T = torsional elastic potential energy
XY = inertial reference frame
$xy, x^b y^b$ = blade coordinate system
$x^d y^d$ = disk coordinate system
$x^m y^m$ = motor coordinate system
α = kinematic expression
α_i = modal factor
β = kinematic expression
δ = axial shortening
ϕ_i = mode shape
λ_i = modal frequency parameter
Γ = the system Lagrangian
ρ = mass per unit length
θ = reference rotation
ψ = torsional deformation angle

References

[1] Shilahansi, L., 1958, ''Bending Frequencies of a Rotating Cantilever Beam,'' ASME J. Appl. Mech., **25**, pp. 28–30.
[2] Pruelli, D., 1972, ''Natural Bending Frequency Comparable to Rotational Frequency in Rotating Cantilever Beam,'' ASME J. Appl. Mech., **39**, pp. 602–604.
[3] Likins, P. W., Barbera, F. J., and Baddeley, V., 1973, ''Mathematical Modeling of Spinning Elastic Bodies,'' AIAA J., **11**, No. 9, pp. 1251–1258.
[4] Kaza, K., and Kvaternik, R., 1977, ''Nonlinear Flap-Lag-Axial Equations of a Rotating Beam,'' AIAA J., **15**, No. 6, pp. 871–874.
[5] Yokoyama, T., 1988, ''Free Vibration Characteristics of Rotating Timoshenko Beams,'' Int. J. Mech. Sci., **30**, No. 10, pp. 743–755.
[6] Al-Bedoor, B. O., and Khulief, Y., 1997, ''General Planar Dynamics of a Sliding Flexible Link,'' J. Sound Vib., **206**, No. 5, pp. 641–661.
[7] Srinivasan, A. V., 1984, ''Vibrations of Bladed-Disk Assemblies: A Selected Survey,'' ASME J. Vibr. Acoust., **106**, pp. 165–168.
[8] Crawley, E., and Mokadam, D., 1984, ''Stager Angle Dependence of the Inertial and the Elastic Coupling in Bladed Disks,'' ASME J. Vibr. Acoust., **106**, pp. 181–188.
[9] Tang, D., and Wang, M., 1984, ''Coupling Technique of Rotor-Fuselage Dynamic Analysis,'' ASME J. Vibr. Acoust., **106**, pp. 235–238.
[10] Loewy, R. G., and Khader, N., 1984, ''Structural Dynamics of Rotating Bladed-Disk Assemblies Coupled With Flexible Shaft Motions,'' AIAA J., **22**, No. 9, pp. 1319–1327.
[11] Crawley, E. F., Ducharme, E., and Mokadam, D., 1986, ''Analytical and Experimental Investigation of the Coupled Bladed Disk Shaft Whirl of a Cantilever Turbofan,'' ASME J. Eng. Gas Turbines Power, **108**, pp. 567–576.
[12] Okabe, A., Otawara, Y., Kaneko, R., Matsushita, O., and Namura, K., 1991, ''An Equivalent Reduced Modeling Method and Its Application to Shaft-Blade Coupled Torsional Vibration Analysis of a Turbine-Generator Set,'' Proc. Inst. Mech. Eng., **205**, pp. 173–181.
[13] Huang, S. C., and Ho, K. B., 1996, ''Coupled Shaft-Torsion and Blade-Bending Vibrations of a Rotating Shaft-Blade Unit,'' ASME J. Eng. Gas Turbines Power, **118**, pp. 100–106.
[14] Al-Bedoor, B. O., 1998, ''Vibrations of a Rotating Blade With Flexible Coupling in the Drive System,'' ASME Pressure Vessels Piping, PVP, **368**, pp. 69–76.
[15] Szabelski, K., and Warminski, J., 1995, ''Parametric Self-Excited Nonlinear System Vibration Analysis With Inertial Excitation,'' Int. J. Nonlinear Mech., **30**, No. 2, pp. 179–189.

Component-Mode-Based Reduced Order Modeling Techniques for Mistuned Bladed Disks—Part I: Theoretical Models

R. Bladh
Graduate Student

M. P. Castanier
Assistant Research Scientist

C. Pierre
Professor

Department of Mechanical Engineering,
The University of Michigan,
Ann Arbor, MI 48109-2125

Component mode synthesis (CMS) techniques are widely used for dynamic analyses of complex structures. Significant computational savings can be achieved by using CMS, since a modal analysis is performed on each component structure (substructure). Mistuned bladed disks are a class of structures for which CMS is well suited. In the context of blade mistuning, it is convenient to view the blades as individual components, while the entire disk may be treated as a single component. Individual blade mistuning may then be incorporated into the CMS model in a straightforward manner. In this paper, the Craig–Bampton (CB) method of CMS is formulated specifically for mistuned bladed disks, using a cyclic disk description. Then a novel secondary modal analysis reduction technique (SMART) is presented: a secondary modal analysis is performed on a CB model, yielding significant further reduction in model size. In addition, a straightforward non-CMS method is developed in which the blade mistuning is projected onto the tuned system modes. Though similar approaches have been reported previously, here it is generalized to a form that is more useful in practical applications. The theoretical models are discussed and compared from both computational and practical perspectives. It is concluded that using SMART, based on a CB model, has tremendous potential for highly efficient, accurate modeling of the vibration of mistuned bladed disks. [DOI: 10.1115/1.1338947]

1 Introduction

The adverse effects of structural irregularities, or mistuning, among the blades of turbomachinery rotors is a persisting concern in the gas turbine community. Mistuning is caused by manufacturing tolerances, deviations in material properties, or nonuniform operational wear; therefore, mistuning is unavoidable. Furthermore, even small mistuning can have a dramatic effect on the vibratory behavior of a rotor, because it can lead to spatial localization of the vibration energy. As a result, certain blades may experience forced response amplitudes and stresses that are substantially larger than those predicted by an analysis of the nominal (tuned) design. Unfortunately, these random uncertainties in blade properties, and the immense computational effort involved in obtaining statistically reliable design data, combine to make this aspect of rotor design cumbersome.

Since the 1960s, several researchers have documented the effects of mistuning on blade vibrations by analyses of representative lumped parameter models, using numerical, statistical, and perturbation methods (Wagner [1]; Dye and Henry [2]; Ewins [3,4]; El-Bayoumy and Srinivasan [5]; Griffin and Hoosac [6]; Wei and Pierre [7,8]; Lin and Mignolet [9]). See Srinivasan [10] for a comprehensive survey of the literature. Unfortunately, in order to accurately represent an actual bladed disk design with a lumped parameter model, one must perform a difficult parameter identification which becomes infeasible as the number of model degrees of freedom (DOF) increases. Hence, to gain practical usefulness, there is a pressing need to employ accurate finite element models of rotor designs in mistuning studies.

To address this issue, there have been several efforts to generate reduced order models systematically from finite element models using component mode synthesis (CMS) methods (Irretier [11]; Kruse and Pierre [12,13]; Castanier et al. [14]; Bladh et al. [15]), receptance techniques (Yang and Griffin [16]), and classical modal analysis with mistuning projection (Yang and Griffin [17]). In CMS, the original structure is subdivided into smaller substructures, or components, for which normal modes are computed independently. The global structure is then represented by a truncated set of component modes that are assembled in a systematic fashion through compatibility constraints. This process yields highly reduced order models for bladed disks that are based on finite element models of arbitrary complexity.

The focus of this study is on the development of reduced order models of mistuned bladed disks, based primarily on CMS techniques. In this first part of the two-part paper, component-mode-based reduced order modeling techniques are derived and presented. In particular, the Craig–Bampton CMS method [18] is reformulated specifically for the analysis of mistuned bladed disks. This tailored formulation uses a cyclic symmetry description of the disk.

In addition, a straightforward non-CMS technique is formulated. This method consists of a modal analysis of the nominal (tuned) system, with a subsequent projection of the blade mistuning data onto the nominal system modes of vibration. This mistuning projection approach is a generalization of the mistuning formulation for shrouded blade assemblies developed in Bladh et al. [15]. Also, Yang and Griffin [17] presented an analogous technique for the case in which each blade is mistuned by a small deviation in its Young's modulus.

Finally, the general idea of a secondary modal analysis reduction technique (SMART) is introduced. In the SMART approach, a secondary modal analysis is performed on a model already reduced by CMS. This two-step reduction decreases dramatically the order of the original model. Also, in contrast to the non-CMS mistuning projection method, the SMART mistuning is introduced in the modal domain rather than the physical domain. This signifi-

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-360. Manuscript received by IGTI January 14, 2000; final revision received by ASME Headquarters April 2, 2000. Associate Editor: M. Mignolet.

cantly reduces the associated computational effort, and it allows a straightforward implementation of individual blade natural frequency mistuning.

This paper is organized as follows. General assumptions and model issues are summarized in Sec. 2. The modeling theory begins in Sec. 3 with a derivation of the Craig–Bampton (CB) technique tailored to a mistuned rotor with a cyclic symmetry description of the disk. In Sec. 4, the method of Castanier et al. [14], which uses a prescribed interface motion approach, is reviewed. In Sec. 5, the mistuning projection method is outlined. In Sec. 6, the secondary modal analysis reduction technique is formulated for mistuned rotors using the CB method for the intermediate model. In Sec. 7, a comparison is presented of the number of floating point operations required for the various methods. The conclusions are summarized in Sec. 8.

2 Computational Issues and Assumptions

An *N*-bladed disk assembly can be divided into one disk component (*d*) and *N* individual blades (*b*). It is assumed that the disk features cyclic symmetry, meaning that it is composed of *N* identical sectors. A disk sector and a blade component are depicted in Fig. 1, which also outlines the index notation used throughout this paper for the components and the interfaces with neighboring components.

Initially, it is assumed that each disk sector or blade is an identical (tuned) and physically isolated substructure. At this point, the stiffness matrix of all disk sectors and blades, $\mathcal{K}$, has a block-diagonal structure:

$$\mathcal{K}=\begin{bmatrix}\mathcal{K}^{\mathbf{d}} & \mathbf{0}\\ \mathbf{0} & \mathcal{K}^{\mathbf{b}}\end{bmatrix}=\begin{bmatrix}\mathbf{I}\otimes\mathbf{K}^{\mathbf{d}} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\otimes\mathbf{K}^{\mathbf{b}}\end{bmatrix}, \tag{1}$$

where $\mathbf{I}$ is an identity matrix of dimension *N*, the symbol $\otimes$ denotes the Kronecker product (see Appendix A), and $\mathbf{K}^{\mathbf{d}}$ (disk) and $\mathbf{K}^{\mathbf{b}}$ (blade) are the stiffness matrices of the two fundamental, stand-alone substructures. Using the notation of Fig. 1, the displacement vectors and the corresponding stiffness matrices may be partitioned in detail as

$$\mathbf{x}^{\mathbf{d}}=\begin{Bmatrix}\mathbf{x}_{\mathbf{D}}^{\mathbf{d}}\\ \mathbf{x}_{\Gamma}^{\mathbf{d}}\\ \mathbf{x}_{\alpha}^{\mathbf{d}}\\ \mathbf{x}_{\beta}^{\mathbf{d}}\end{Bmatrix}\quad \mathbf{K}^{\mathbf{d}}=\begin{bmatrix}\mathbf{K}_{\mathbf{DD}}^{\mathbf{d}} & \mathbf{K}_{\mathbf{D}\Gamma}^{\mathbf{d}} & \mathbf{K}_{\mathbf{D}\alpha}^{\mathbf{d}} & \mathbf{K}_{\mathbf{D}\beta}^{\mathbf{d}}\\ \mathbf{K}_{\mathbf{D}\Gamma}^{\mathbf{d}^{\mathbf{T}}} & \mathbf{K}_{\Gamma\Gamma}^{\mathbf{d}} & \mathbf{K}_{\Gamma\alpha}^{\mathbf{d}} & \mathbf{K}_{\Gamma\beta}^{\mathbf{d}}\\ \mathbf{K}_{\mathbf{D}\alpha}^{\mathbf{d}^{\mathbf{T}}} & \mathbf{K}_{\Gamma\alpha}^{\mathbf{d}^{\mathbf{T}}} & \mathbf{K}_{\alpha\alpha}^{\mathbf{d}} & \mathbf{K}_{\alpha\beta}^{\mathbf{d}}\\ \mathbf{K}_{\mathbf{D}\beta}^{\mathbf{d}^{\mathbf{T}}} & \mathbf{K}_{\Gamma\beta}^{\mathbf{d}^{\mathbf{T}}} & \mathbf{K}_{\alpha\beta}^{\mathbf{d}^{\mathbf{T}}} & \mathbf{K}_{\beta\beta}^{\mathbf{d}}\end{bmatrix} \tag{2}$$

$$\mathbf{x}^{\mathbf{b}}=\begin{Bmatrix}\mathbf{x}_{\mathbf{B}}^{\mathbf{b}}\\ \mathbf{x}_{\Gamma}^{\mathbf{b}}\end{Bmatrix}\quad \mathbf{K}^{\mathbf{b}}=\begin{bmatrix}\mathbf{K}_{\mathbf{BB}}^{\mathbf{b}} & \mathbf{K}_{\mathbf{B}\Gamma}^{\mathbf{b}}\\ \mathbf{K}_{\mathbf{B}\Gamma}^{\mathbf{b}^{\mathbf{T}}} & \mathbf{K}_{\Gamma\Gamma}^{\mathbf{b}}\end{bmatrix}. \tag{3}$$

The mass matrices are partitioned in exactly the same fashion.

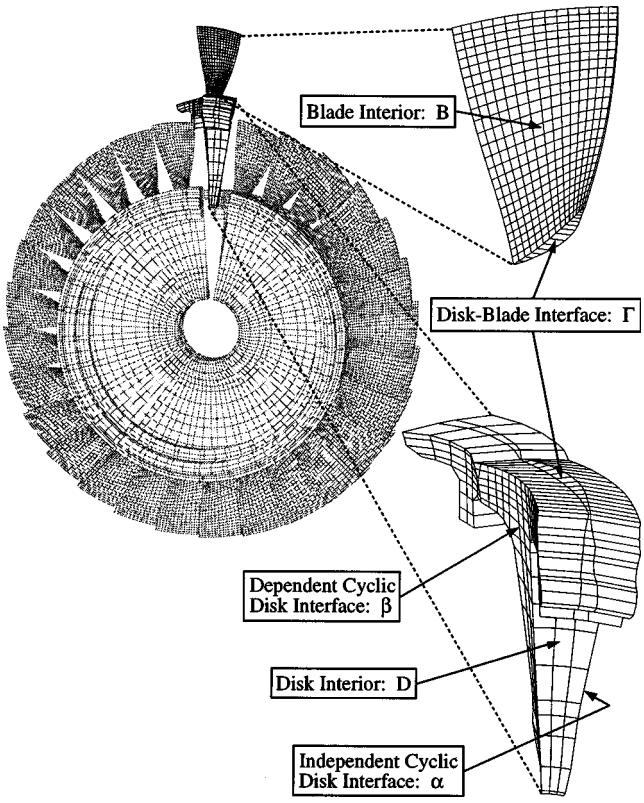


Fig. 1 Substructuring approach and Index notation

2.1 Cyclic Symmetry Description of the Disk Component.

The treatment of the disk component is greatly simplified by using a cyclic symmetry analysis. There are two principal benefits from treating the disk as a cyclic assembly rather than an assembly of *N* arbitrary components. First, the DOF at each interface between adjacent disk sectors are eliminated by cyclic constraints, which yields a smaller CMS model. Second, the disk component mode shapes bear a greater resemblance to the system modes, which improves modal convergence.

The cyclic symmetry analysis employed here is identical to the implementation in the commercial finite element software package MSC/NASTRAN™. This approach is a real-valued formulation, and it was outlined by Joseph [19]. From the theory of symmetrical components (Fortescue [20]), some quantity $\mathbf{x}$ (displacements, forces, etc.) in physical coordinates for all *N* disk sectors may be expressed as a linear combination of the corresponding quantity $\tilde{\mathbf{u}}$ in cyclic coordinates for the fundamental disk sector as

$$\mathbf{x}=\begin{Bmatrix}\mathbf{x}_1\\ \mathbf{x}_2\\ \mathbf{x}_3\\ \vdots\\ \mathbf{x}_N\end{Bmatrix}=(\mathbf{F}\otimes\mathbf{I})\begin{Bmatrix}\tilde{\mathbf{u}}^0\\ \tilde{\mathbf{u}}^{1,c}\\ \tilde{\mathbf{u}}^{1,s}\\ \tilde{\mathbf{u}}^{2,c}\\ \vdots\\ \tilde{\mathbf{u}}^{P}\end{Bmatrix}=\hat{\mathbf{F}}\tilde{\mathbf{u}}, \tag{4}$$

where $\mathbf{F}$ is the real-valued Fourier matrix defined in Appendix B. Moreover, the dimension of $\mathbf{I}$ is equal to the number of elements in $\tilde{\mathbf{u}}^{\mathbf{h}}$, *h* represents the harmonic order, and *P* is the highest possible harmonic for an *N*-bladed disk, which is defined as $P=\text{int}[N/2]$. The ''tilde'' notation is used throughout this paper to indicate when a quantity is represented in cyclic coordinates. The cyclic coordinate representation is essentially a Fourier series expansion of the disk motion. However, it is *not* an approximation, as the series contains the complete set of admissible circumferential shapes of the cyclic assembly.

The cyclic coordinate transformation of Eq. (4) yields a set of trigonometric relations that describe the motion of the disk sector's dependent cyclic boundary (β) relative to the independent cyclic boundary (α) as follows:

$$\begin{cases}\tilde{\mathbf{u}}_{\beta}^{\mathbf{h,c}}=\tilde{\mathbf{u}}_{\alpha}^{\mathbf{b,c}}\cos\phi_h+\tilde{\mathbf{u}}_{\alpha}^{\mathbf{h,s}}\sin\phi_h\\ \tilde{\mathbf{u}}_{\beta}^{\mathbf{h,s}}=-\tilde{\mathbf{u}}_{\alpha}^{\mathbf{h,c}}\sin\phi_h+\tilde{\mathbf{u}}_{\alpha}^{\mathbf{h,s}}\cos\phi_h\end{cases}\quad h=0,\ldots,P, \tag{5}$$

where $\phi_h=2\pi h/N$ is the interblade phase angle for the *h*th harmonic. Note that the two equations in Eq. (5) collapse to one for harmonics 0 and, if it exists, *N*/2. These are ''single'' harmonics and require only a single sector description, as in Eq. (2). All other harmonics are ''double'' and require a two-sector description:

$$\mathbf{K}_{\mathbf{2x}}^{\mathbf{d}}=\begin{bmatrix}\mathbf{K}^{\mathbf{d}} & \mathbf{0}\\ \mathbf{0} & \mathbf{K}^{\mathbf{d}}\end{bmatrix}. \tag{6}$$

For generality, a ''double'' harmonic is considered in the following. By enforcing Eq. (5) on the double disk sector's stiffness

matrix in Eq. (6), the dependent β DOF are eliminated. Moreover, for the disk's subsequent CMS implementation, it is necessary to partition with respect to interior (D and α) DOF—which are collectively denoted sector (S) DOF—and disk–blade interface (Γ) DOF. With this in mind, the cyclic disk stiffness matrix and displacement vector of harmonic h ($h \neq 0$, $h \neq N/2$) may be written as

$$\tilde{\mathbf{K}}_{\mathbf{d}}^{\mathbf{h}} = \begin{bmatrix} \tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h}} & \tilde{\mathbf{K}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}} \\ \tilde{\mathbf{K}}_{\mathbf{d,S\Gamma}}^{\mathbf{h^T}} & \tilde{\mathbf{K}}_{\mathbf{d,\Gamma\Gamma}}^{\mathbf{h}} \end{bmatrix} \quad \tilde{\mathbf{u}}_{\mathbf{d}}^{\mathbf{h}} = \begin{Bmatrix} \tilde{\mathbf{u}}_{\mathbf{S}}^{\mathbf{h,c}} \\ \tilde{\mathbf{u}}_{\mathbf{S}}^{\mathbf{h,s}} \\ \cdots \\ \tilde{\mathbf{u}}_{\Gamma}^{\mathbf{h,c}} \\ \tilde{\mathbf{u}}_{\Gamma}^{\mathbf{h,s}} \end{Bmatrix}. \tag{7}$$

The interior (SS) partition is subpartitioned as

$$\tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h}} = \begin{bmatrix} \tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h,0}} & \tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h,1}} \\ \tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h,1^T}} & \tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h,0}} \end{bmatrix}, \tag{8}$$

where

$$\tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h,0}} = \begin{bmatrix} \mathbf{K}_{\alpha\alpha}^{\mathbf{d}} + (\mathbf{K}_{\alpha\beta}^{\mathbf{d}} + \mathbf{K}_{\alpha\beta}^{\mathbf{d^T}})\cos\phi_h + \mathbf{K}_{\beta\beta}^{\mathbf{d}} & \mathbf{K}_{\mathbf{D}\alpha}^{\mathbf{d^T}} + \mathbf{K}_{\mathbf{D}\beta}^{\mathbf{d^T}}\cos\phi_h \\ \mathbf{K}_{\mathbf{D}\alpha}^{\mathbf{d}} + \mathbf{K}_{\mathbf{D}\beta}^{\mathbf{d}}\cos\phi_h & \mathbf{K}_{\mathbf{DD}}^{\mathbf{d}} \end{bmatrix}$$

$$\tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h,1}} = \begin{bmatrix} (\mathbf{K}_{\alpha\beta}^{\mathbf{d}} - \mathbf{K}_{\alpha\beta}^{\mathbf{d^T}})\sin\phi_h & -\mathbf{K}_{\mathbf{D}\beta}^{\mathbf{d^T}}\sin\phi_h \\ \mathbf{K}_{\mathbf{D}\beta}^{\mathbf{d}}\sin\phi_h & \mathbf{0} \end{bmatrix}.$$

The remaining partitions (SΓ and $\Gamma\Gamma$) take on the following forms:

$$\tilde{\mathbf{K}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}} = \begin{bmatrix} \mathbf{K}_{\Gamma\alpha}^{\mathbf{d^T}} + \mathbf{K}_{\Gamma\beta}^{\mathbf{d^T}}\cos\phi_h & -\mathbf{K}_{\Gamma\beta}^{\mathbf{d^T}}\sin\phi_h \\ \mathbf{K}_{\mathbf{D}\Gamma}^{\mathbf{d}} & \mathbf{0} \\ \mathbf{K}_{\Gamma\beta}^{\mathbf{d^T}}\sin\phi_h & \mathbf{K}_{\Gamma\alpha}^{\mathbf{d^T}} + \mathbf{K}_{\Gamma\beta}^{\mathbf{d^T}}\cos\phi_h \\ \mathbf{0} & \mathbf{K}_{\mathbf{D}\Gamma}^{\mathbf{d}} \end{bmatrix} \tag{9}$$

$$\tilde{\mathbf{K}}_{\mathbf{d,\Gamma\Gamma}}^{\mathbf{h}} = \begin{bmatrix} \mathbf{K}_{\Gamma\Gamma}^{\mathbf{d}} & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_{\Gamma\Gamma}^{\mathbf{d}} \end{bmatrix}. \tag{10}$$

Note that for $h=0$ and, if it exists, for $h=N/2$, the ''sine blocks'' are zero. By rearranging the DOF order in this case, two identical blocks form on the diagonal of $\tilde{\mathbf{K}}_{\mathbf{d}}^{\mathbf{h}}$, while the off-diagonal block is zero. Thus, as indicated earlier, one such block on the diagonal alone provides the complete disk description in these two special cases. The corresponding cyclic disk partitions for ''single'' harmonics are given by

$$\tilde{\mathbf{K}}_{\mathbf{d,SS}}^{h} = \begin{bmatrix} \mathbf{K}_{\alpha\alpha}^{\mathbf{d}} + (\mathbf{K}_{\alpha\beta}^{\mathbf{d}} + \mathbf{K}_{\alpha\beta}^{\mathbf{d^T}})\cos\phi_h + \mathbf{K}_{\beta\beta}^{\mathbf{d}} & \mathbf{K}_{\mathbf{D}\alpha}^{\mathbf{d^T}} + \mathbf{K}_{\mathbf{D}\beta}^{\mathbf{d^T}}\cos\phi_h \\ \mathbf{K}_{\mathbf{D}\alpha}^{\mathbf{d}} + \mathbf{K}_{\mathbf{D}\beta}^{\mathbf{d}}\cos\phi_h & \mathbf{K}_{\mathbf{DD}}^{\mathbf{d}} \end{bmatrix} \tag{11}$$

$$\tilde{\mathbf{K}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}} = \begin{bmatrix} \mathbf{K}_{\Gamma\alpha}^{\mathbf{d^T}} + \mathbf{K}_{\Gamma\beta}^{\mathbf{d^T}}\cos\phi_h \\ \mathbf{K}_{\mathbf{D}\Gamma}^{\mathbf{d}} \end{bmatrix} \quad \tilde{\mathbf{K}}_{\mathbf{d,\Gamma\Gamma}}^{\mathbf{k}} = \mathbf{K}_{\Gamma\Gamma}^{\mathbf{d}}. \tag{12}$$

Again, the manipulations of the mass matrix are completely analogous.

2.2 Engine Order Excitation Force. In this section, an external excitation force vector, $\mathbf{Q}$, is constructed for all the blade DOF of the assembly. The restriction to blade DOF is not an absolute requirement, but it leads to a more compact formulation, and it should be sufficient from a practical perspective. An engine order excitation is assumed, which is harmonic in time and differs only in phase from blade to blade. The phase at the nth blade, φ_n, is given by

$$\varphi_n = \frac{2\pi C(n-1)}{N}, \tag{13}$$

where C is the engine order of the excitation. The external force vector can then be expressed as

$$\mathbf{Q} = \begin{Bmatrix} \mathbf{Q_B} \\ \mathbf{Q}_\Gamma \end{Bmatrix} = \begin{Bmatrix} \sqrt{N}\mathbf{e_{C+1}} \otimes \mathbf{f_B} \\ \sqrt{N}\mathbf{e_{C+1}} \otimes \mathbf{f}_\Gamma \end{Bmatrix}, \tag{14}$$

where $\mathbf{e_{C+1}}$ is the $(C+1)$th column of the complex Fourier matrix defined in Appendix B. The column vectors $\mathbf{f_B}$ and $\mathbf{f}_\Gamma$ contain the forces on, respectively, the interior and interface DOF of a fundamental blade.

2.3 Mistuning Implementation. In this work, the blade mistuning is modeled by offsets in modal stiffnesses Λ—or, equivalently, offsets in natural frequencies ω—of the blades while fixed at the base (cantilevered). The mistuned modal stiffness of the kth cantilevered blade mode for the nth blade may be expressed as

$$\overline{\Lambda}_{b,n}^{k} = (\overline{\omega}_{b,n}^{k})^2 = (1+\delta_n^k)\Lambda_b^k, \tag{15}$$

where Λ_b^k is the modal stiffness of the kth tuned cantilevered blade mode, and δ_n^k is the corresponding mistuning parameter for the nth blade. Note that this implies that the mistuned modes of a blade may be realized by a linear combination of the tuned modes (i.e., it is assumed that they span the same space). Also, note that it is assumed throughout this paper that mode shapes are normalized to yield unit modal masses.

Due to simple implementation and validation, most of the published studies on mistuned bladed disks have considered variations in Young's modulus as the only source of blade mistuning. This implies a uniform rescaling of the blade stiffness matrix, which translates into a uniform rescaling of the modal stiffnesses as well. The mistuning parameter δ_n^k in Eq. (15) is then replaced by δ_n, which represents the offset of Young's modulus from its nominal value for the nth blade.

3 Fixed Interface Method (Craig–Bampton)

The Craig–Bampton (CB) method (Craig and Bampton [18]; Craig [21,22]) employs two sets of modes to represent the motion of each component:

- Φ, a truncated set of normal elastic modes of vibration with the DOF at component interfaces held fixed (see Fig. 2(*a*))
- Ψ, a complete set of static elastic constraint modes induced by successive unit deflections of each interface DOF while all other interface DOF are held fixed (see Fig. 2(*b*)).

The modes in Φ are linearly independent by definition, and the manner in which the interface DOF are successively displaced ensures linear independence among the constraint modes in Ψ, as well as linear independence between the two mode sets. Furthermore, if all modes are retained, the number of modes in Φ will equal the total number of interior DOF in the component. By construction, the number of modes in Ψ always equals the number of interface DOF. Hence, in the limit, linear independence and completeness combine to yield the exact solution for the CB method relative to the parent finite element model (i.e., it spans the complete deformation space of the finite element model).

The following subsections outline the assembly of the CB model as applied to mistuned bladed disks with a cyclic disk description.

3.1 Blade Component. For now, all N blades are assumed to be identical (tuned). Furthermore, from a component perspective, the blades are not directly coupled, since only unshrouded blades are considered in this study. Hence, for computing the component quantities, it is sufficient to look at a single blade and then expand to N blades.

First, the normal modes for the cantilevered blade are obtained

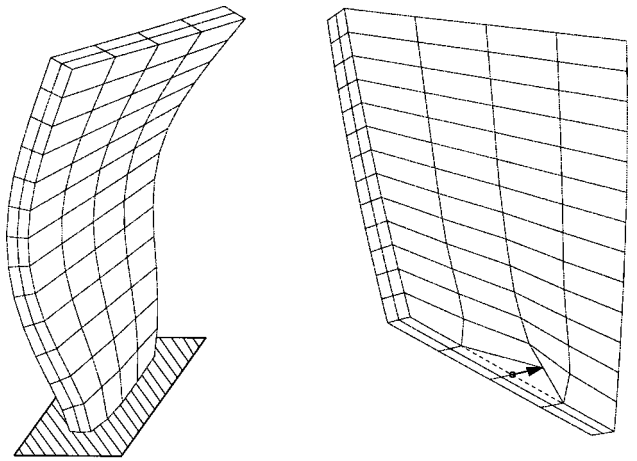


Fig. 2 Craig–Bampton component modes: (*a*) fixed-interface normal modes of vibration; (*b*) static constraint modes due to successive unit deflections of interface DOF

from the conventional eigenvalue problem:

$$[\mathbf{K}_{\mathbf{BB}}^{\mathbf{b}} - \omega^2 \mathbf{M}_{\mathbf{BB}}^{\mathbf{b}}]\phi_{\mathbf{b}} = \mathbf{0}. \tag{16}$$

The mode shapes $\phi_{\mathbf{b}}$ of interest are collected into a matrix $\Phi_{\mathbf{b}}$, and the corresponding eigenvalues form the elements in a diagonal matrix $\Lambda_{\mathbf{b}}$. Second, the constraint modes, $\Psi_{\mathbf{b}}$, are computed from the static problem

$$\begin{bmatrix} \mathbf{K}_{\mathbf{BB}}^{\mathbf{b}} & \mathbf{K}_{\mathbf{B\Gamma}}^{\mathbf{b}} \\ \mathbf{K}_{\mathbf{B\Gamma}}^{\mathbf{b}^{\mathbf{T}}} & \mathbf{K}_{\mathbf{\Gamma\Gamma}}^{\mathbf{b}} \end{bmatrix} \begin{bmatrix} \Psi_{\mathbf{b}} \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{R}_{\Gamma} \end{bmatrix}, \tag{17}$$

where $\mathbf{R}_{\Gamma}$ contains the reaction forces due to the imposed unit displacements, **I**. Solving the first block of equations in Eq. (17) yields

$$\Psi_{\mathbf{b}} = -\mathbf{K}_{\mathbf{BB}}^{\mathbf{b}^{-1}} \mathbf{K}_{\mathbf{B\Gamma}}^{\mathbf{b}}. \tag{18}$$

Note that the matrix inverse need not be computed, since the columns of $\Psi_{\mathbf{b}}$ are the solution vectors **x** of $\mathbf{K}_{\mathbf{BB}}^{\mathbf{b}}\mathbf{x} = -\mathbf{K}_{\mathbf{B\Gamma}}^{\mathbf{b}}$.

The physical blade displacements can now be expressed in terms of the two sets of component modes, which form the traditional Craig–Bampton modal matrix, $\mathbf{U}_{\mathbf{cb}}^{\mathbf{b}}$, as

$$\mathbf{x}^{\mathbf{b}} = \begin{Bmatrix} \mathbf{x}_{\mathbf{B}}^{\mathbf{b}} \\ \mathbf{x}_{\Gamma}^{\mathbf{b}} \end{Bmatrix} = \begin{bmatrix} \Phi_{\mathbf{b}} & \Psi_{\mathbf{b}} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{Bmatrix} \mathbf{p}_{\mathbf{b}}^{\mathbf{b}} \\ \mathbf{p}_{\mathbf{c}}^{\mathbf{b}} \end{Bmatrix} = \mathbf{U}_{\mathbf{cb}}^{\mathbf{b}} \mathbf{p}^{\mathbf{b}}. \tag{19}$$

The transformation from physical blade coordinates $\mathbf{x}^{\mathbf{b}}$ to CB blade modal coordinates $\mathbf{p}^{\mathbf{b}}$ via $\mathbf{U}_{\mathbf{cb}}^{\mathbf{b}}$ yields the reduced CB mass and stiffness matrices for the blade component:

$$\begin{aligned} \mu^{\mathbf{b}} &= \mathbf{U}_{\mathbf{cb}}^{\mathbf{b}^{\mathbf{T}}} \mathbf{M}^{\mathbf{b}} \mathbf{U}_{\mathbf{cb}}^{\mathbf{b}} = \begin{bmatrix} \mathbf{I} & \mu_{\mathbf{bc}} \\ \mu_{\mathbf{bc}}^{\mathbf{T}} & \mu_{\mathbf{cc,b}} \end{bmatrix} \\ \kappa^{\mathbf{b}} &= \mathbf{U}_{\mathbf{cb}}^{\mathbf{b}^{\mathbf{T}}} \mathbf{K}^{\mathbf{b}} \mathbf{U}_{\mathbf{cb}}^{\mathbf{b}} = \begin{bmatrix} \Lambda_{\mathbf{b}} & \mathbf{0} \\ \mathbf{0} & \kappa_{\mathbf{cc,b}} \end{bmatrix}, \end{aligned} \tag{20}$$

where

$$\begin{aligned} \mu_{\mathbf{bc}} &= \Phi_{\mathbf{b}}^{\mathbf{T}} [\mathbf{M}_{\mathbf{BB}}^{\mathbf{b}} \Psi_{\mathbf{b}} + \mathbf{M}_{\mathbf{B\Gamma}}^{\mathbf{b}}] \\ \mu_{\mathbf{cc,b}} &= \Psi_{\mathbf{b}}^{\mathbf{T}} [\mathbf{M}_{\mathbf{BB}}^{\mathbf{b}} \Psi_{\mathbf{b}} + \mathbf{M}_{\mathbf{B\Gamma}}^{\mathbf{b}}] + \mathbf{M}_{\mathbf{B\Gamma}}^{\mathbf{b}^{\mathbf{T}}} \Psi_{\mathbf{b}} + \mathbf{M}_{\mathbf{\Gamma\Gamma}}^{\mathbf{b}} \\ \kappa_{\mathbf{cc,b}} &= \mathbf{K}_{\mathbf{\Gamma\Gamma}}^{\mathbf{b}} + \mathbf{K}_{\mathbf{B\Gamma}}^{\mathbf{b}^{\mathbf{T}}} \Psi_{\mathbf{b}}. \end{aligned}$$

In the absence of direct blade-to-blade structural coupling, Eq. (20) is expanded for all *N* blades as

$$\begin{aligned} \overline{\mu}^{\mathbf{b}} &= \begin{bmatrix} \mathbf{I} & \mathbf{I} \otimes \mu_{\mathbf{bc}} \\ \mathbf{I} \otimes \mu_{\mathbf{bc}}^{\mathbf{T}} & \mathbf{I} \otimes \mu_{\mathbf{cc,b}} \end{bmatrix} \\ \overline{\kappa}^{\mathbf{b}} &= \begin{bmatrix} \mathbf{I} \otimes \Lambda_{\mathbf{b}} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \otimes \kappa_{\mathbf{cc,b}} \end{bmatrix}. \end{aligned} \tag{21}$$

3.2 Cyclic Disk Component. The construction of the required quantities for the disk component is more computationally intensive than for the blade component, although the steps are the same. For the setup of the cyclic structural matrices involved, the reader is referred to Sec. 2.1, Eqs. (7)–(12).

To begin, the cyclic normal modes for the disk component are obtained for each harmonic *h* from the cyclic eigenvalue problem:

$$[\tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h}} - \omega^2 \tilde{\mathbf{M}}_{\mathbf{d,SS}}^{\mathbf{h}}]\tilde{\phi}_{\mathbf{d}}^{\mathbf{h}} = \mathbf{0}, \quad h = 0, \ldots, P. \tag{22}$$

Moreover, the cyclic constraint modes, $\tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}}$, are computed for each harmonic *h* from the cyclic static problem

$$\begin{bmatrix} \tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h}} & \tilde{\mathbf{K}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}} \\ \tilde{\mathbf{K}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}^{\mathbf{T}}} & \tilde{\mathbf{K}}_{\mathbf{d,\Gamma\Gamma}}^{\mathbf{h}} \end{bmatrix} \begin{bmatrix} \tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}} \\ \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{R}_{\Gamma}^{\mathbf{h}} \end{bmatrix}. \tag{23}$$

Again, solving the first block of equations in Eq. (23) for the cyclic disk constraint modes yields

$$\tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}} = -\tilde{\mathbf{K}}_{\mathbf{d,SS}}^{\mathbf{h}^{-1}} \tilde{\mathbf{K}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}}. \tag{24}$$

While keeping normal and constraint modes separated, the retained cyclic normal modes and the cyclic constraint modes of the disk are merged into a cyclic Craig–Bampton modal matrix, $\tilde{\mathbf{U}}_{\mathbf{cb}}^{\mathbf{d}}$, as

$$\tilde{\mathbf{U}}_{\mathbf{cb}}^{\mathbf{d}} = \begin{bmatrix} \underset{h=0,\ldots,P}{\tilde{\mathbf{B}}\mathbf{diag}} [\tilde{\Phi}_{\mathbf{d}}^{\mathbf{h}}] & \underset{h=0,\ldots,P}{\tilde{\mathbf{B}}\mathbf{diag}} [\tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}}] \\ \mathbf{0} & \mathbf{I} \end{bmatrix}, \tag{25}$$

where $\tilde{\mathbf{B}}\mathbf{diag}[\cdot]$ denotes a pseudo-block-diagonal matrix (see Appendix B), with the argument being the *h*th ''block,'' and the range of *h* is shown. In this context, the *h*th ''block'' pertains to the *h*th harmonic. Similarly, the eigenvalues corresponding to the retained cyclic normal modes form the elements in a pseudo-block-diagonal generalized stiffness matrix $\tilde{\Lambda}_{\mathbf{d}}$, where each block is diagonal in itself.

Using Eq. (25), the physical disk displacements can now be expressed in terms of the two sets of cyclic component modes by virtue of Eq. (4):

$$\overline{\mathbf{x}}^{\mathbf{d}} = \begin{Bmatrix} \overline{\mathbf{x}}_{\mathbf{S}}^{\mathbf{d}} \\ \overline{\mathbf{x}}_{\Gamma}^{\mathbf{d}} \end{Bmatrix} = \mathbf{U}_{\mathbf{cb}}^{\mathbf{d}} \begin{Bmatrix} \tilde{\mathbf{p}}_{\mathbf{d}}^{\mathbf{d}} \\ \tilde{\mathbf{p}}_{\mathbf{c}}^{\mathbf{d}} \end{Bmatrix} = \mathbf{U}_{\mathbf{cb}}^{\mathbf{d}} \tilde{\mathbf{p}}^{\mathbf{d}}, \tag{26}$$

where

$$\mathbf{U}_{\mathbf{cb}}^{\mathbf{d}} = \begin{bmatrix} \underset{h=0,\ldots,P}{\hat{\mathbf{F}}\tilde{\mathbf{B}}\mathbf{diag}} [\tilde{\Phi}_{\mathbf{d}}^{\mathbf{h}}] & \underset{h=0,\ldots,P}{\hat{\mathbf{F}}\tilde{\mathbf{B}}\mathbf{diag}} [\overline{\Psi}_{\mathbf{d}}^{\mathbf{h}}] \\ \mathbf{0} & \hat{\mathbf{F}} \end{bmatrix}. \tag{27}$$

The transformation from physical disk coordinates $\overline{\mathbf{x}}^{\mathbf{d}}$ to CB cyclic disk coordinates $\tilde{\mathbf{p}}^{\mathbf{d}}$ via $\mathbf{U}_{\mathbf{cb}}^{\mathbf{d}}$ yields the reduced CB mass and stiffness matrices for the disk component:

$$\tilde{\mu}^{\mathbf{d}} = \begin{bmatrix} \mathbf{I} & \tilde{\mu}_{\mathbf{dc}} \\ \tilde{\mu}_{\mathbf{dc}}^{\mathbf{T}} & \tilde{\mu}_{\mathbf{cc,d}} \end{bmatrix} \quad \tilde{\kappa}^{\mathbf{d}} = \begin{bmatrix} \tilde{\Lambda}_{\mathbf{d}} & \mathbf{0} \\ \mathbf{0} & \tilde{\kappa}_{\mathbf{cc,d}} \end{bmatrix}, \tag{28}$$

where

$$\tilde{\mu}_{\mathbf{dc}} = \underset{h=0,\ldots,P}{\tilde{\mathbf{B}}\mathbf{diag}}\ [\tilde{\Phi}_{\mathbf{d}}^{\mathbf{h}^T}[\tilde{\mathbf{M}}_{\mathbf{d,SS}}^{\mathbf{h}}\tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}} + \tilde{\mathbf{M}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}}]]$$

$$\tilde{\mu}_{\mathbf{cc,d}} = \underset{h=0,\ldots,P}{\tilde{\mathbf{B}}\mathbf{diag}}\ [\tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}^T}[\tilde{\mathbf{M}}_{\mathbf{d,SS}}^{\mathbf{h}}\tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}} + \tilde{\mathbf{M}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}}] + \tilde{\mathbf{M}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}^T}\tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}} + \tilde{\mathbf{M}}_{\mathbf{d,\Gamma\Gamma}}^{\mathbf{h}}]$$

$$\tilde{\kappa}_{\mathbf{cc,d}} = \underset{h=0,\ldots,P}{\tilde{\mathbf{B}}\mathbf{diag}}\ [\tilde{\mathbf{K}}_{\mathbf{d,\Gamma\Gamma}}^{\mathbf{h}} + \tilde{\mathbf{K}}_{\mathbf{d,S\Gamma}}^{\mathbf{h}^T}\tilde{\Psi}_{\mathbf{d}}^{\mathbf{h}}].$$

Note that there is no coupling between the harmonics due to the orthogonality of the cyclic modes.

3.3 CMS Model Assembly. In the Craig–Bampton method, the CMS model assembly is achieved by satisfying displacement compatibility over the component interfaces, i.e., $\overline{\mathbf{x}}_{\Gamma}^{\mathbf{b}} = \overline{\mathbf{x}}_{\Gamma}^{\mathbf{d}}$. The physical interface displacements for the blades and the disk are found in Eqs. (19) and (26), respectively, and result in the following necessary condition:

$$\overline{\mathbf{x}}_{\Gamma}^{\mathbf{b}} = \begin{Bmatrix} \mathbf{x}_{\Gamma,1}^{\mathbf{b}} \\ \mathbf{x}_{\Gamma,2}^{\mathbf{b}} \\ \vdots \\ \mathbf{x}_{\Gamma,\mathbf{N}}^{\mathbf{b}} \end{Bmatrix} = \begin{Bmatrix} \mathbf{p}_{\mathbf{c},1}^{\mathbf{b}} \\ \mathbf{p}_{\mathbf{c},2}^{\mathbf{b}} \\ \vdots \\ \mathbf{p}_{\mathbf{c},\mathbf{N}}^{\mathbf{b}} \end{Bmatrix} = \overline{\mathbf{p}}_{\mathbf{c}}^{\mathbf{b}} = \hat{\mathbf{F}}\tilde{\mathbf{p}}_{\mathbf{c}}^{\mathbf{d}} = \overline{\mathbf{x}}_{\Gamma}^{\mathbf{d}}. \tag{29}$$

Hence, keeping $\tilde{\mathbf{p}}_{\mathbf{c}}^{\mathbf{d}}$ as active DOF, the substructure coupling is represented by the constraint transformation

$$\begin{Bmatrix} \tilde{\mathbf{p}}_{\mathbf{d}}^{\mathbf{d}} \\ \tilde{\mathbf{p}}_{\mathbf{c}}^{\mathbf{d}} \\ \overline{\mathbf{p}}_{\mathbf{b}}^{\mathbf{b}} \\ \overline{\mathbf{p}}_{\mathbf{c}}^{\mathbf{b}} \end{Bmatrix} = \begin{bmatrix} \mathbf{I} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I} \\ \mathbf{0} & \hat{\mathbf{F}} & \mathbf{0} \end{bmatrix} \begin{Bmatrix} \tilde{\mathbf{p}}_{\mathbf{d}} \\ \tilde{\mathbf{p}}_{\mathbf{c}} \\ \overline{\mathbf{p}}_{\mathbf{b}} \end{Bmatrix} = \mathbf{T}_{\mathbf{cb}}\mathbf{p}_{\mathbf{cb}}. \tag{30}$$

After this final transformation, the synthesized system mass and stiffness matrices for the CB method take on the following forms:

$$\mathbf{M}^{\mathbf{cb}} = \mathbf{T}_{\mathbf{cb}}^{\mathbf{T}}\begin{bmatrix} \tilde{\mu}^{\mathbf{d}} & \mathbf{0} \\ \mathbf{0} & \overline{\mu}^{\mathbf{b}} \end{bmatrix}\mathbf{T}_{\mathbf{cb}} = \begin{bmatrix} \mathbf{I} & \tilde{\mu}_{\mathbf{dc}} & \mathbf{0} \\ \tilde{\mu}_{\mathbf{dc}}^{\mathbf{T}} & \tilde{\mu}_{\mathbf{cc,d}} + \mathbf{I}\otimes\mu_{\mathbf{cc,b}} & \hat{\mathbf{F}}^{\mathbf{T}}(\mathbf{I}\otimes\mu_{\mathbf{bc}}^{\mathbf{T}}) \\ \mathbf{0} & (\mathbf{I}\otimes\mu_{\mathbf{bc}})\hat{\mathbf{F}} & \mathbf{I} \end{bmatrix} \tag{31}$$

$$\mathbf{K}^{\mathbf{cb}} = \begin{bmatrix} \tilde{\Lambda}_{\mathbf{d}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\kappa}_{\mathbf{cc,d}} + \mathbf{I}\otimes\kappa_{\mathbf{cc,b}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\otimes\Lambda_{\mathbf{b}} \end{bmatrix}, \tag{32}$$

where it is recognized that similarity transformations of expanded matrices follow $\hat{\mathbf{F}}^{\mathbf{T}}(\mathbf{I}\otimes\mathbf{A})\hat{\mathbf{F}} = \mathbf{I}\otimes\mathbf{A}$.

Note that the derived model is still for the tuned assembly. However, since the CB modal coordinates pertaining to the normal blade modes are for a cantilevered blade, the introduction of modal stiffness mistuning is exquisitely simple. The cantilevered blade modal stiffnesses on the diagonal of $\mathbf{I}\otimes\Lambda_{\mathbf{b}}$ (extreme lower-right partition in Eq. (32)) may be perturbed directly and individually to give the following $\mathbf{K}_{\mathbf{bb}}^{\mathbf{cb}}$ partition:

$$\mathbf{K}_{\mathbf{bb}}^{\mathbf{cb}} = \underset{n=1,\ldots,N}{\mathbf{Bdiag}}\left[\underset{k=1,\ldots,m_b}{\mathbf{diag}}\ [1+\delta_n^k]\Lambda_{\mathbf{b}}\right], \tag{33}$$

where m_b is the number of retained cantilevered blade modes. Introducing mistuning in this way implies that any mistuning effects on the constraint modes are neglected. This is not a severe approximation, but as shown in the second part of this study, it does affect the performance of the method. However, except in the simplistic case of varying Young's modulus, quantifying mistuning for the constraint modes is a rather ambiguous task.

Finally, the CB modal force is constructed. Projecting the component modal matrices of Eqs. (19) and (27) onto the physical blade force vector described in Eq. (14), while enforcing the constraints of Eq. (30), the modal force is obtained as

$$\mathbf{F}^{\mathbf{cb}} = \begin{Bmatrix} \mathbf{0} \\ \mathbf{F}_{\mathbf{c}}^{\mathbf{cb}} \\ \mathbf{F}_{\mathbf{b}}^{\mathbf{cb}} \end{Bmatrix}, \tag{34}$$

where

$$\mathbf{F}_{\mathbf{c}}^{\mathbf{cb}} = \begin{Bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \sqrt{N}\mathbf{f}_{\mathbf{C,c}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C+1}}\otimes\{\Psi_{\mathbf{b}}^{\mathbf{T}}\mathbf{f}_{\mathbf{B}} + \mathbf{f}_{\Gamma}\} \\ \sqrt{N}\mathbf{f}_{\mathbf{C,s}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C+1}}\otimes\{\Psi_{\mathbf{b}}^{\mathbf{T}}\mathbf{f}_{\mathbf{B}} + \mathbf{f}_{\Gamma}\} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{Bmatrix}$$

$$\mathbf{F}_{\mathbf{b}}^{\mathbf{cb}} = \sqrt{N}\mathbf{e}_{\mathbf{C+1}}\otimes\Phi_{\mathbf{b}}^{\mathbf{T}}\mathbf{f}_{\mathbf{B}}.$$

Note how orthogonality between columns of **E** and **F** of different interblade phase angles simplifies the modal force partition pertaining to the constraint modes.

Using Eqs. (30)–(34), the complete CB model of a mistuned bladed disk (in the absence of aerodynamic coupling) can now be set up as

$$\mathbf{M}^{\mathbf{cb}}\ddot{\mathbf{p}}_{\mathbf{cb}} + \mathbf{C}^{\mathbf{cb}}\dot{\mathbf{p}}_{\mathbf{cb}} + (1+Gj)\mathbf{K}^{\mathbf{cb}}\mathbf{p}_{\mathbf{cb}} = \mathbf{F}^{\mathbf{cb}}, \tag{35}$$

where j denotes the imaginary unit, $\sqrt{-1}$. To facilitate more realistic modeling of the structure's dynamic response, Eq. (35) includes structural damping with coefficient G, as well as viscous modal damping of the cantilevered blade modes, which is implemented as

$$\mathbf{C}^{\mathbf{cb}} = \begin{bmatrix} \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{I}\otimes\underset{k=1,\ldots,m_b}{\mathbf{diag}}\ [2\zeta^k]\sqrt{\Lambda_{\mathbf{b}}} \end{bmatrix}, \tag{36}$$

where ζ^k is the viscous damping coefficient associated with the kth cantilevered blade mode.

4 Disk-induced Blade Constraint Modes (REDUCE)

The method presented in this section is referred to by the name of the associated computer code, REDUCE. It was thoroughly outlined for free vibrations by Castanier et al. [14]. It was extended to the forced response by Kruse and Pierre [12] and subsequently revised and further extended to cover shrouded assemblies by Bladh et al. [15]. In terms of generality and versatility, REDUCE represents the current state of the art for modeling the structural dynamics of mistuned bladed disk assemblies. Moreover, REDUCE is actively used by several gas turbine engine companies. However, as will be shown in this study, alternative methods may be formulated that yield superior performance.

The REDUCE method employs two sets of modes:

- A truncated set of cantilevered blade modes—blade component modes with the disk–blade interface DOF held fixed (see Fig. 3(*a*)). The mode shapes and associated eigenvalues are collected into a modal matrix $\mathbf{u}^{\mathbf{b}}$ and a diagonal generalized stiffness matrix $\hat{\mathbf{K}}_{\mathbf{b}}$, respectively. The corresponding generalized coordinates are denoted **b**. Note that $\hat{\mathbf{K}}_{\mathbf{b}}$ is equivalent to $\mathbf{I}\otimes\Lambda_{\mathbf{b}}$ of the CB method.
- A truncated set of cyclic modes for the fundamental disk–blade sector with a massless blade (see Fig. 3(*b*)). The blade portions of the cyclic mode shapes are collected into modal matrices $\tilde{\mathbf{u}}_{\mathbf{h}}^{\mathbf{d}}$ for each harmonic h. These shapes may be described as disk-induced blade constraint modes. The modal matrices $\tilde{\mathbf{u}}_{\mathbf{h}}^{\mathbf{d}}$ and the

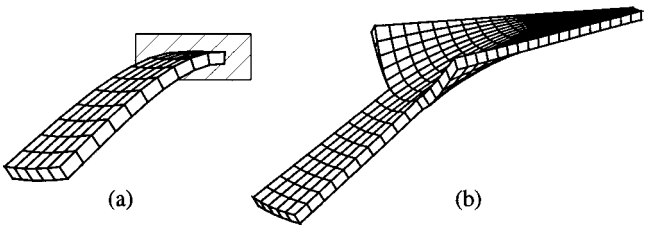


Fig. 3 REDUCE component modes: (*a*) fixed-interface (cantilevered) normal blade modes; (*b*) cyclic modes for the fundamental disk–blade sector with a massless blade

associated eigenvalues are assembled into $\tilde{\mathbf{U}}^{\mathbf{d}}$ and a diagonal generalized stiffness matrix $\hat{\mathbf{K}}_{\mathbf{d}}$, respectively. The corresponding generalized coordinates are denoted **a**.

By definition, each set of modes, $\mathbf{u}^{\mathbf{b}}$ and $\tilde{\mathbf{U}}^{\mathbf{d}}$, is linearly independent. Moreover, since $\mathbf{u}^{\mathbf{b}}$ does not contain any disk motion, it is clear that the two mode sets are linearly independent with respect to each other as well. Furthermore, in the limit, the number of modes in $\mathbf{u}^{\mathbf{b}}$ will equal the total number of interior (B) DOF in the blade, and the number of modes in $\tilde{\mathbf{U}}^{\mathbf{d}}$ will equal the number of independent disk interior $(\mathrm{D}+\alpha)$ plus interface (Γ) DOF. Hence, just as for the CB method, linear independence and DOF completeness in the limit combine to yield the exact solution for the REDUCE method relative to the parent finite element model. Finally, note that this is not a true CMS method, since the assembled disk–blade sector is needed to obtain the disk-induced blade constraint modes.

The resulting equations of motion are restated here for convenience:

$$\mathcal{M}\ddot{\mathbf{z}}+\mathcal{C}\dot{\mathbf{z}}+(1+Gj)\mathcal{K}\mathbf{z}=\mathcal{Q}, \tag{37}$$

where

$$\mathbf{z}=\begin{Bmatrix}\mathbf{a}\\ \mathbf{b}\end{Bmatrix} \quad \mathcal{C}=\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & (\mathbf{I}\otimes \underset{k=1,\ldots,m_b}{\mathbf{diag}}\,[2\zeta^k])\sqrt{\hat{\mathbf{K}}_{\mathbf{b}}}\end{bmatrix}$$

$$\mathcal{M}=\begin{bmatrix}\mathbf{I}+\underset{h=0,\ldots,P}{\tilde{\mathbf{B}}\mathbf{diag}}\,[\tilde{\mathbf{u}}_{\mathbf{h}}^{\mathbf{d}^{\mathrm{T}}}\mathbf{M}^{\mathbf{b}}\tilde{\mathbf{u}}_{\mathbf{h}}^{\mathbf{d}}] & \tilde{\mathbf{U}}^{\mathbf{d}^{\mathrm{T}}}(\mathbf{I}\otimes\mathbf{M}^{\mathbf{b}}\mathbf{u}^{\mathbf{b}})\\ (\mathbf{I}\otimes\mathbf{u}^{\mathbf{b}^{\mathrm{T}}}\mathbf{M}^{\mathbf{b}})\tilde{\mathbf{U}}^{\mathbf{d}} & \mathbf{I}\end{bmatrix}$$

$$\mathcal{K}=\begin{bmatrix}\hat{\mathbf{K}}_{\mathbf{d}} & \tilde{\mathbf{U}}^{\mathbf{d}^{\mathrm{T}}}(\mathbf{I}\otimes\mathbf{K}^{\mathbf{b}}\mathbf{u}^{\mathbf{b}})\\ (\mathbf{I}\otimes\mathbf{u}^{\mathbf{b}^{\mathrm{T}}}\mathbf{K}^{\mathbf{b}})\tilde{\mathbf{U}}^{\mathbf{d}} & \underset{k=1,\ldots,N}{\mathbf{Bdiag}}\,[\underset{k=1,\ldots,m_b}{\mathbf{diag}}\,[1+\delta_n^k]]\hat{\mathbf{K}}_{\mathbf{b}}\end{bmatrix}$$

$$\mathcal{Q}=\{\mathcal{Q}_{\mathbf{d}}^{\mathrm{T}}\ \vdots\ \mathcal{Q}_{\mathbf{b}}^{\mathrm{T}}\}^{\mathrm{T}}=\{\mathbf{0}\cdots\mathbf{0}\ \mathcal{Q}_{\mathbf{d}}^{\mathbf{C}^{\mathrm{T}}}\mathbf{0}\cdots\mathbf{0}\ \vdots\ \mathcal{Q}_{\mathbf{b}}^{\mathrm{T}}\}^{\mathbf{T}}$$

$$\mathcal{Q}_{\mathbf{d}}^{\mathbf{C}}=\left\{\begin{array}{l}\sqrt{N}\{\mathbf{f}_{\mathbf{C},\mathbf{c}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C},\mathbf{1}}^{\mathbf{d},\mathbf{c}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}}\\ \quad -\mathbf{f}_{\mathbf{C},\mathbf{s}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C},\mathbf{1}}^{\mathbf{d},\mathbf{s}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}}\}\\ \sqrt{N}\{\mathbf{f}_{\mathbf{C},\mathbf{s}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C},\mathbf{1}}^{\mathbf{d},\mathbf{c}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}}\\ \quad +\mathbf{f}_{\mathbf{C},\mathbf{c}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C},\mathbf{1}}^{\mathbf{d},\mathbf{s}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}}\}\\ \quad\vdots\\ \sqrt{N}\{\mathbf{f}_{\mathbf{C},\mathbf{c}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C},\mathbf{m_d}}^{\mathbf{d},\mathbf{c}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}}\\ \quad -\mathbf{f}_{\mathbf{C},\mathbf{s}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C},\mathbf{m_d}}^{\mathbf{d},\mathbf{s}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}}\}\\ \sqrt{N}\{\mathbf{f}_{\mathbf{C},\mathbf{s}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C},\mathbf{m_d}}^{\mathbf{d},\mathbf{c}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}}\\ \quad +\mathbf{f}_{\mathbf{C},\mathbf{c}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C},\mathbf{m_d}}^{\mathbf{d},\mathbf{s}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}}\}\end{array}\right\}$$

$$\mathcal{Q}_{\mathbf{b}}=\sqrt{N}\mathbf{e}_{\mathbf{C}+\mathbf{1}}\otimes\mathbf{u}^{\mathbf{b}^{\mathrm{T}}}\mathbf{f}_{\mathbf{b}},$$

and $\mathbf{f}_{\mathbf{b}}$ is the force vector on the fundamental blade (i.e., a composite of $\mathbf{f}_{\mathbf{B}}$ and $\mathbf{f}_{\Gamma}$ from Sec. 2.2). In this formulation, quantities in physical coordinates (structural matrices, mode shapes) pertain to the blade part only.

Since the disk motion is described by the disk portion from the second mode set alone (Fig. 3(*b*)), no separate set of constraint modes for the disk is employed. This causes the disk to be too stiff at the interface, which degrades the performance (modal convergence) of the method. However, it has been found that artificial softening of the cantilevered blade modes yields significant accuracy improvements for both free and forced responses. This is achieved by adjusting the eigenvalues of the cantilevered blade modes in an iterative fashion, based on the finite element eigenvalues for blade-dominated assembly modes. Though it is heuristic, this technique has proved very efficient (Bladh et al. [15]).

5 Non-CMS Mistuning Projection Method

This method is based on the assumption that the mistuned modes of a bladed disk assembly may be realized by a linear combination of its tuned modes. This assumption is justified by two observations: (a) the local motion of a blade in a mistuned assembly is, to a large extent, merely an amplification of its tuned motion; and (b) any admissible disk shape, no matter how spatially localized, may be realized by a linear combination of its harmonic shapes in cyclic coordinates if all harmonics 0 through *P* are included in the model.

By separating disk and blade parts by partitioning as shown in Eq. (1), the stiffness matrix of the mistuned bladed disk may be represented as

$$\bar{\mathcal{K}}=\begin{bmatrix}\mathbf{I}\otimes\mathbf{K}^{\mathbf{d}} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\otimes\mathbf{K}^{\mathbf{b}}+\underset{n=1,\ldots,N}{\mathbf{Bdiag}}\,[\Delta\mathbf{K}_{\mathbf{n}}^{\mathbf{b}}]\end{bmatrix}, \tag{38}$$

where $\Delta\mathbf{K}_{\mathbf{n}}^{\mathbf{b}}$ is a matrix containing the stiffness deviations from the nominal stiffness matrix for the *n*th blade. Recall that only blade stiffness mistuning is considered here; thus, all other partitions remain unaffected.

First, a set of nominal, cyclic eigenvalues and eigenvectors are computed for each harmonic *h* from

$$[\tilde{\mathcal{K}}^{\mathbf{h}}-\omega^2\tilde{\mathcal{M}}^{\mathbf{h}}]\tilde{\mathbf{u}}^{\mathbf{h}}=\mathbf{0},\quad h=0,\ldots,P, \tag{39}$$

where

$$\tilde{\mathcal{K}}^{\mathbf{h}}=\begin{bmatrix}\tilde{\mathbf{K}}_{\mathbf{d},\mathbf{SS}}^{\mathbf{h}} & \tilde{\mathbf{K}}_{\mathbf{d},\mathbf{S}\Gamma}^{\mathbf{h}} & \mathbf{0}\\ \tilde{\mathbf{K}}_{\mathbf{d},\mathbf{S}\Gamma}^{\mathbf{h}^{\mathrm{T}}} & \tilde{\mathbf{K}}_{\mathbf{d},\Gamma\Gamma}^{\mathbf{h}}+\tilde{\mathbf{K}}_{\mathbf{b},\Gamma\Gamma}^{\mathbf{h}} & \tilde{\mathbf{K}}_{\mathbf{b},\mathbf{B}\Gamma}^{\mathbf{h}^{\mathrm{T}}}\\ \mathbf{0} & \tilde{\mathbf{K}}_{\mathbf{b},\mathbf{B}\Gamma}^{\mathbf{h}} & \tilde{\mathbf{K}}_{\mathbf{b},\mathbf{BB}}^{\mathbf{h}}\end{bmatrix}.$$

The cyclic disk quantities are defined in Sec. 2.1. Referring to Eq. (3), the cyclic blade quantities have the common form

$$\tilde{\mathbf{K}}_{\mathbf{b},\mathbf{xy}}^{\mathbf{h}}=\begin{cases}\mathbf{K}_{\mathbf{xy}}^{\mathbf{b}} & h=0,\quad h=\dfrac{N}{2}\ (\text{if it exists})\\ \begin{bmatrix}\mathbf{K}_{\mathbf{xy}}^{\mathbf{b}} & \mathbf{0}\\ \mathbf{0} & \mathbf{K}_{\mathbf{xy}}^{\mathbf{b}}\end{bmatrix} & h\neq 0,\quad h\neq\dfrac{N}{2},\end{cases} \tag{40}$$

where ''**xy**'' represents **BB**, **B**Γ, or $\mathbf{\Gamma\Gamma}$. The structure of $\tilde{\mathcal{M}}^{\mathbf{h}}$ is identical to that of $\tilde{\mathcal{K}}^{\mathbf{h}}$.

A fundamental step in this method is to use a small subset of the obtained cyclic modes in order to form a reduced order model by classical modal analysis. The disk–blade interface plus interior blade portions of the selected mode shapes are collected into a cyclic modal matrix $\tilde{\mathbf{U}}_{\mathbf{s}}$, while the associated eigenvalues are collected into a cyclic generalized stiffness matrix $\tilde{\Lambda}_{\mathbf{s}}$. In the typical case, the analyst concentrates on a particular family of blade

modes and/or eigenfrequency veerings that are deemed critical due to engine operating conditions. As shown by Yang and Griffin [17], mistuned bladed disks are particularly well adapted to system mode selections in relatively narrow frequency bands. Hence, the mode selection is typically based on a frequency range encompassing the blade mode family (or families) of interest. Thus, in most cases, the number of modes needed, m_s, is on the order of the number of blades, $\mathcal{O}(N)$. This is also the size of the resulting reduced order model.

In the tuned case, classical modal analysis simply results in a small, fully decoupled system, where the modal mass matrix is an identity matrix, and the modal stiffness matrix is a diagonal matrix with the selected eigenvalues as elements. The mistuned case, however, requires some additional work. By representing the mistuned blades by stiffness matrix deviations as indicated in Eq. (38), the mistuning is entered into the reduced order model by projecting the stiffness deviations onto the selected tuned modes of the assembly. Note, however, that the stiffness deviations are in physical coordinates while the nominal modes are cyclic. Hence, the mistuning projection takes the form:

$$\Delta\tilde{\mathbf{K}}^{\mathbf{b}} = \tilde{\mathbf{U}}_{\mathbf{s}}^{\mathbf{T}}\hat{\mathbf{F}}^{\mathbf{T}} \underset{n=1,\ldots,N}{\mathbf{Bdiag}} [\Delta\mathbf{K}_{\mathbf{n}}^{\mathbf{b}}]\hat{\mathbf{F}}\tilde{\mathbf{U}}_{\mathbf{s}}. \tag{41}$$

The reduced order model may then be formulated as

$$\ddot{\mathbf{q}} + \tilde{\mathbf{C}}^{\mathbf{mp}}\dot{\mathbf{q}} + (1+Gj)\ [\tilde{\Lambda}_{\mathbf{s}} + \Delta\tilde{\mathbf{K}}^{\mathbf{b}}]\mathbf{q} = \tilde{\mathbf{F}}^{\mathbf{mp}}, \tag{42}$$

where

$$\tilde{\mathbf{C}}^{\mathbf{mp}} = \underset{k=1,\ldots,m_s}{\mathbf{diag}} [2\zeta^k]\sqrt{\tilde{\Lambda}_{\mathbf{s}}}$$

$$\tilde{\mathbf{F}}^{\mathbf{mp}} = \begin{Bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \sqrt{N}\mathbf{f}_{\mathbf{C,c}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C+1}} \otimes \tilde{\mathbf{u}}_{\mathbf{s,b}}^{\mathbf{C}^{\mathbf{T}}}\mathbf{f}_{\mathbf{b}} \\ \sqrt{N}\mathbf{f}_{\mathbf{C,s}}^{\mathbf{T}}\mathbf{e}_{\mathbf{C+1}} \otimes \tilde{\mathbf{u}}_{\mathbf{s,b}}^{\mathbf{C}^{\mathbf{T}}}\mathbf{f}_{\mathbf{b}} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{Bmatrix},$$

$\mathbf{f_b}$ is the fundamental blade force vector (same as in Sec. 4), and $\tilde{\mathbf{u}}_{\mathbf{s,b}}^{\mathbf{C}}$ contains the blade portions of the selected mode shapes of the Cth harmonic. Note that the stiffness deviation projection matrix, $\Delta\tilde{\mathbf{K}}^{\mathbf{b}}$, does not possess any particular matrix structure other than symmetry. Thus, in general, the reduced order model stiffness matrix becomes fully populated when mistuning is introduced. Also, note that the viscous modal damping introduced here refers to the nominal assembly modes, not the cantilevered blade modes as in previous methods. As seen in Eq. (42), once the blade stiffness deviations have been established, this method is conceptually very straightforward. However, the process for obtaining the blade stiffness deviations requires some further consideration.

In the simplistic case of offsets in blade Young's modulus, this method allows for simple input of mistuning in the following manner:

$$\Delta\mathbf{K}_{\mathbf{n}}^{\mathbf{b}} = \delta_n\mathbf{K}^{\mathbf{b}}, \quad n=1,\ldots,N. \tag{43}$$

However, for input of mistuning individually for each mode of each cantilevered blade, as in the previous methods, the task becomes more cumbersome. In this case, an approach analogous to the mistuning of shrouded assemblies in the REDUCE method has to be adopted. This approach is outlined in Bladh et al. [15], but here it is reviewed briefly and adapted to current notation. To begin, a diagonal matrix containing the mistuned modal stiffnesses (measured or generated) of the nth blade may be written as

$$\underset{k=1,\ldots,m_b}{\mathbf{diag}} [1+\delta_n^k]\Lambda_{\mathbf{b}} = \mathbf{u}^{\mathbf{b}^{\mathbf{T}}}[\mathbf{K}^{\mathbf{b}} + \Delta\mathbf{K}_{\mathbf{n}}^{\mathbf{b}}]\mathbf{u}^{\mathbf{b}}, \tag{44}$$

where $\mathbf{u^b}$ contains the nominal mode shapes for a cantilevered blade, and $\Lambda_{\mathbf{b}}$ is a diagonal matrix containing the corresponding modal stiffnesses for a tuned cantilevered blade. Note that there is already an approximation made at this point, namely that the mode shapes of a mistuned blade are the same as those of a tuned blade. Rearranging Eq. (44) and identifying $\Lambda_{\mathbf{b}} = \mathbf{u}^{\mathbf{b}^{\mathbf{T}}}\mathbf{K}^{\mathbf{b}}\mathbf{u}^{\mathbf{b}}$, yields the following expression for the stiffness deviation:

$$\Delta\mathbf{K}_{\mathbf{n}}^{\mathbf{b}} = \mathbf{u}^{\mathbf{b}^{\mathbf{T}^{-1}}} \underset{k=1,\ldots,m_b}{\mathbf{diag}} [\delta_n^k]\Lambda_{\mathbf{b}}\mathbf{u}^{\mathbf{b}^{-1}}. \tag{45}$$

Now, making use of the eigenvector normalization assumption, it is realized that

$$\mathbf{I} = \mathbf{u}^{\mathbf{b}^{\mathbf{T}}}\mathbf{M}^{\mathbf{b}}\mathbf{u}^{\mathbf{b}} \Rightarrow \mathbf{u}^{\mathbf{b}^{\mathbf{T}^{-1}}} \approx \mathbf{M}^{\mathbf{b}}\mathbf{u}^{\mathbf{b}}$$

$$\Lambda_{\mathbf{b}} = \mathbf{u}^{\mathbf{b}^{\mathbf{T}}}\mathbf{K}^{\mathbf{b}}\mathbf{u}^{\mathbf{b}} \Rightarrow \Lambda_{\mathbf{b}}\mathbf{u}^{\mathbf{b}^{-1}} \approx \mathbf{u}^{\mathbf{b}^{\mathbf{T}}}\mathbf{K}^{\mathbf{b}}. \tag{46}$$

It is assumed here that $\mathbf{u}^{\mathbf{b}^{-1}}$ exists, i.e., that it is complete (square). This is seldom the case in practice, as it would be impractically large, and thus, as indicated, Eq. (46) is an approximation. The nonexistent inverse does not pose a problem, however, since it need not be computed. Moreover, the implied approximation is of the same order as the modal analysis itself, and it does not cause a noticeable decrease in accuracy.

By substituting Eq. (46) into Eq. (45), the stiffness deviation matrix may be expressed in its final form as

$$\Delta\mathbf{K}_{\mathbf{n}}^{\mathbf{b}} = \mathbf{M}^{\mathbf{b}}\mathbf{u}^{\mathbf{b}} \underset{k=1,\ldots,m_b}{\mathbf{diag}} [\delta_n^k]\mathbf{u}^{\mathbf{b}^{\mathbf{T}}}\mathbf{K}^{\mathbf{b}}. \tag{47}$$

In this manner, stiffness deviations of individual blades can be incorporated into the reduced order model by projection onto the selected cyclic modes of the assembly.

6 Secondary Modal Analysis Reduction Technique (SMART)

The CB implementation presented in Sec. 3 yields robust and highly reliable reduced order models. However, the retained physical interface DOF may lead to impractically large CMS models when using highly detailed parent finite element models. There is thus a need for further model order reduction. One option is to improve the CB representation by partial secondary modal analysis. In particular, consider a partial secondary modal analysis on: (a) the partitions of the matrices that pertain to the constraint modes; or (b) the partitions of the matrices that pertain to the disk normal modes *plus* the constraint modes. A truncated set of modes is then selected to form the new reduced order model. (It should be noted that the approach (a) was originally developed by Tan et al. [23] for power flow analysis.) Both approaches yield significant further model order reduction, while retaining the high accuracy level of the original CB model. They may also eliminate matrix ill-conditioning due to the mix of modal- and physical-coordinate matrix partitions present in the classical CB setup. Furthermore, these extensions to the CB technique are by no means restricted to mistuned bladed disks only, but may be applied to any CB synthesized system for additional model reduction. However, preliminary results indicate that the improvements gained by the partial secondary modal analyses may not be sufficient to justify the additional modeling efforts in this particular application. Hence, these two methods will not be considered in any further detail in this paper.

As a far more appealing alternative, a technique is introduced that whittles the size of the reduced order model to an absolute

minimum, without incurring severe truncation errors or sacrificing versatility. This is achieved by performing a full-scale secondary modal analysis on an intermediate model that has already been reduced through modal analysis in some fashion (e.g., by the CB method). The secondary modal analysis is based on those modes of the intermediate model that fall within a frequency range encompassing the blade mode family (or families) of interest during subsequent mistuned free and/or forced response analyses. In this respect, the approach is very similar to the mistuning projection method of the previous section. However, a distinct advantage of the SMART model is the fact that the required projection of mistuning data is carried out in the low-order modal domain, while the mistuning projection method must deal with the projections in the physical domain. This SMART idea may be applied to any intermediate model that is constructed from a CMS (or other) method. In this work, the intermediate models are constructed with the CB method, due to its robustness and excellent accuracy. Furthermore, the CB method is a natural choice in view of the mistuning implementation employed in this study, since it gives direct access to the blade modal stiffnesses.

As in the mistuning projection method, the first step is to obtain the tuned modes from which to form the new selected basis. This is done for each harmonic h using the tuned CB model entirely in cyclic coordinates, which results in the following set of eigenvalue problems:

$$[\tilde{\mathbf{K}}^{\mathbf{h}}_{\mathbf{cb}} - \omega^2 \tilde{\mathbf{M}}^{\mathbf{h}}_{\mathbf{cb}}]\tilde{\mathbf{u}}^{\mathbf{h}}_{\mathbf{cb}} = \mathbf{0}, \quad h = 0, \ldots, P, \tag{48}$$

where

$$\tilde{\mathbf{K}}^{\mathbf{h}}_{\mathbf{cb}} = \begin{bmatrix} \tilde{\mathbf{\Lambda}}^{\mathbf{h}}_{\mathbf{d}} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \tilde{\mathbf{K}}^{\mathbf{cb,h}}_{\mathbf{cc,d}} + \tilde{\mathbf{K}}^{\mathbf{cb,h}}_{\mathbf{cc,b}} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \tilde{\mathbf{\Lambda}}^{\mathbf{h}}_{\mathbf{b}} \end{bmatrix}$$

$$\tilde{\mathbf{M}}^{\mathbf{h}}_{\mathbf{cb}} = \begin{bmatrix} \mathbf{I} & \tilde{\mathbf{M}}^{\mathbf{cb,h}}_{\mathbf{dc}} & \mathbf{0} \\ \tilde{\mathbf{M}}^{\mathbf{cb,h^T}}_{\mathbf{dc}} & \tilde{\mathbf{M}}^{\mathbf{cb,h}}_{\mathbf{cc,d}} + \tilde{\mathbf{M}}^{\mathbf{cb,h}}_{\mathbf{cc,b}} & \tilde{\mathbf{M}}^{\mathbf{cb,h^T}}_{\mathbf{bc}} \\ \mathbf{0} & \tilde{\mathbf{M}}^{\mathbf{cb,h}}_{\mathbf{bc}} & \mathbf{I} \end{bmatrix}.$$

The disk partitions as well as the blade cc partitions are defined in Sec. 3. For the remaining cyclic blade partitions in the CB model, the definitions in Sec. 3.1 give

$$\tilde{\mathbf{M}}^{\mathbf{cb,h}}_{\mathbf{bc}} = \begin{cases} \mu_{\mathbf{bc}} & h = 0, \ h = \dfrac{N}{2} \ \text{(if it exists)} \\ \begin{bmatrix} \mu_{\mathbf{bc}} & \mathbf{0} \\ \mathbf{0} & \mu_{\mathbf{bc}} \end{bmatrix} & h \neq 0, \ h \neq \dfrac{N}{2} \end{cases}, \tag{49}$$

$$\tilde{\mathbf{\Lambda}}^{\mathbf{h}}_{\mathbf{b}} = \begin{cases} \Lambda_{\mathbf{b}} & h = 0, \ h = \dfrac{N}{2} \ \text{(if it exists)} \\ \begin{bmatrix} \Lambda_{\mathbf{b}} & \mathbf{0} \\ \mathbf{0} & \Lambda_{\mathbf{b}} \end{bmatrix} & h \neq 0, \ h \neq \dfrac{N}{2} \end{cases}. \tag{50}$$

Similar to the mistuning projection method, the next step is to select a small subset of the obtained CB cyclic modes. This subset is then used to form a further reduced order model by classical modal analysis, which is the secondary modal analysis indicated by the name of the method. The constraint-mode portions of the selected mode shapes, $\tilde{\mathbf{u}}^{\mathbf{h,c}}_{\mathbf{cb}}$, are collected into a cyclic modal matrix $\tilde{\mathbf{U}}^{\mathbf{s,c}}_{\mathbf{cb}}$, while the normal-blade-mode portions, $\tilde{\mathbf{u}}^{\mathbf{h,b}}_{\mathbf{cb}}$, are collected into $\tilde{\mathbf{U}}^{\mathbf{s,b}}_{\mathbf{cb}}$. The associated eigenvalues are collected into a cyclic generalized stiffness matrix $\tilde{\Lambda}^{\mathbf{s}}_{\mathbf{cb}}$. Both the mode selection process and the resulting reduced tuned system follow the discussion in Sec. 5 for the mistuning projection method. The difference, however, is that the selected eigenvalues are exact (with respect to the parent finite element model) in the mistuning projection method, while they are only as accurate as the intermediate model in the SMART case.

As shown in Eq. (33), mistuning is represented in the CB model by perturbing the diagonal elements of the normal-blade-mode (bb) partition, which represent the individual modal stiffnesses for each cantilevered blade mode of each blade in the assembly. Hence, using the selected tuned, cyclic modes as the basis for the secondary modal expansion, mistuning enters into the SMART model by projecting these modal stiffness perturbations onto the selected tuned modes of the assembly. Note that the perturbations are in normal coordinates while the nominal modes are cyclic. Hence, the mistuning projection takes the form:

$$\Delta\tilde{\mathbf{K}}^{\mathbf{s}}_{\mathbf{cb}} = \tilde{\mathbf{U}}^{\mathbf{s,b^T}}_{\mathbf{cb}} \hat{\mathbf{F}}^{\mathbf{T}} \ \mathbf{B} \ \underset{n=1,\ldots,N}{\mathbf{diag}} \left[\underset{k=1,\ldots,m_b}{\mathbf{diag}} \ [\delta^k_n] \Lambda_{\mathbf{b}} \right] \hat{\mathbf{F}} \tilde{\mathbf{U}}^{\mathbf{s,b}}_{\mathbf{cb}}. \tag{51}$$

Note that only the mode shape portions pertaining to the blade normal modes are involved in the projection, which normally are of modest size. Equation (51) represents the key to this method's versatility and suitability for statistical studies: versatility by enabling input of a practical measure of mistuning obtained for cantilevered blades; and suitability for statistical studies due to its computational efficiency, since this mistuning projection is made in the low-order modal domain. In Sec. 7, it is shown how the latter makes the task of obtaining comprehensive forced response statistics nearly effortless compared to other methods of comparable accuracy.

With the mistuning projection in place, the SMART CB model may be formulated as

$$\ddot{\mathbf{q}} + \tilde{\mathbf{C}}^{\mathbf{s}}_{\mathbf{cb}} \dot{\mathbf{q}} + (1 + Gj)[\tilde{\Lambda}^{\mathbf{s}}_{\mathbf{cb}} + \Delta\tilde{\mathbf{K}}^{\mathbf{s}}_{\mathbf{cb}}]\mathbf{q} = \tilde{\mathbf{F}}^{\mathbf{s}}_{\mathbf{cb}}, \tag{52}$$

where

$$\tilde{\mathbf{C}}^{\mathbf{s}}_{\mathbf{cb}} = \tilde{\mathbf{U}}^{\mathbf{s,b^T}}_{\mathbf{cb}} \left[\mathbf{I} \otimes \underset{k=1,\ldots,m_b}{\mathbf{diag}} \ [2\zeta^k] \sqrt{\Lambda_{\mathbf{b}}} \right] \tilde{\mathbf{U}}^{\mathbf{s,b}}_{\mathbf{cb}}$$

$$\tilde{\mathbf{F}}^{\mathbf{s}}_{\mathbf{cb}} = \begin{Bmatrix} \mathbf{0} \\ \vdots \\ \mathbf{0} \\ \tilde{\mathbf{u}}^{\mathbf{C,c^T}}_{\mathbf{cb}} \tilde{\mathbf{F}}^{\mathbf{C,c}}_{\mathbf{cb}} + \tilde{\mathbf{u}}^{\mathbf{C,b^T}}_{\mathbf{cb}} \tilde{\mathbf{F}}^{\mathbf{C,b}}_{\mathbf{cb}} \\ \mathbf{0} \\ \vdots \\ \mathbf{0} \end{Bmatrix}$$

$$\tilde{\mathbf{F}}^{\mathbf{C,c}}_{\mathbf{cb}} = \begin{Bmatrix} \sqrt{N}\mathbf{f}^{\mathbf{T}}_{\mathbf{C,c}} \mathbf{e}_{\mathbf{C}+1} \otimes \{\Psi^{\mathbf{T}}_{\mathbf{b}} \mathbf{f}_{\mathbf{B}} + \mathbf{f}_{\Gamma}\} \\ \sqrt{N}\mathbf{f}^{\mathbf{T}}_{\mathbf{C,s}} \mathbf{e}_{\mathbf{C}+1} \otimes \{\Psi^{\mathbf{T}}_{\mathbf{b}} \mathbf{f}_{\mathbf{B}} + \mathbf{f}_{\Gamma}\} \end{Bmatrix}$$

$$\tilde{\mathbf{F}}^{\mathbf{C,b}}_{\mathbf{cb}} = \begin{Bmatrix} \sqrt{N}\mathbf{f}^{\mathbf{T}}_{\mathbf{C,c}} \mathbf{e}_{\mathbf{C}+1} \otimes \Phi^{\mathbf{T}}_{\mathbf{b}} \mathbf{f}_{\mathbf{B}} \\ \sqrt{N}\mathbf{f}^{\mathbf{T}}_{\mathbf{C,s}} \mathbf{e}_{\mathbf{C}+1} \otimes \Phi^{\mathbf{T}}_{\mathbf{b}} \mathbf{f}_{\mathbf{B}} \end{Bmatrix}.$$

Note that this generally leads to a fully populated, symmetric stiffness matrix, similar to that of the mistuning projection method.

7 Comparison of Methods

In this section, the number of floating point operations (flops) required to set up and use a reduced order model is estimated for the previously introduced techniques—Craig–Bampton, REDUCE, mistuning projection, and SMART Craig–Bampton. From the outset it is realized that including REDUCE in this comparison is not quite fair, as REDUCE does not possess nearly the same accuracy as the other methods. However, since REDUCE has gained popularity with several industrial users, it is included as a benchmark for comparison.

The standard algorithms for matrix inversion (for solving $\mathbf{Ax} = \mathbf{b}$) and multiplication of square matrices are $\mathcal{O}(n^3)$ processes, where n is the order of the matrix (Strang [24]). Moreover, for this comparison it is essential to have an estimate of the com-

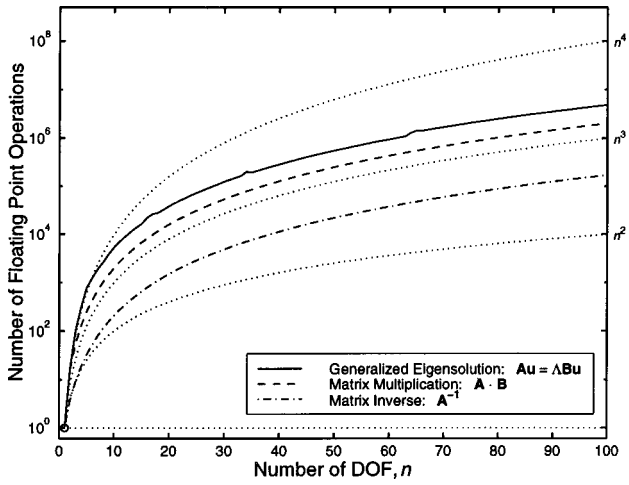


Fig. 4 Required number of floating point operations (flops) as a function of matrix size for generalized eigensolution, matrix multiplication, and matrix inversion. Note the slight ''bumps'' due to the iterative nature of the eigensolution.

putational effort involved in solving an eigenvalue problem, since it is a central component in all methods. However, this is not possible to obtain explicitly, because it is an iterative process for $n>2$. Since some estimate of the eigensolver cost is necessary, a brief numerical investigation was conducted using MATLAB™, which has a built-in flops counter. Two matrices, **A** and **B**, were constructed with increasing size n as follows (upper triangle only):

$$A_{ij}=\begin{cases} \dfrac{i^4}{j^2} & i=j \\ -\dfrac{i^2}{j} & i\neq j \end{cases} \qquad B_{ij}=\begin{cases} \dfrac{(i+j)^2}{i^2} & i=j \\ -\dfrac{(i+j)}{i} & i\neq j \end{cases}$$

$$j=1,\ldots,n \quad i=1,\ldots,j. \tag{53}$$

The generated matrices, which are fully populated, real-valued, symmetric, positive definite, and diagonally dominant, were used to get estimates of the required number of flops for the three basic matrix operations required here: generalized real, symmetric eigensolutions, $\mathbf{Au}=\Lambda\mathbf{Bu}$; multiplication of two square matrices, $\mathbf{A}\cdot\mathbf{B}$; and matrix inversion, $\mathbf{A}^{-1}$. The results are shown in Fig. 4.

The results in Fig. 4 confirm the n^3 proportionality of matrix inversion and multiplication. The eigensolution closely follows the n^3 curve, too. Hence, it is assumed that n^3 proportionality may be used for all three matrix operations to get fair comparisons. Note that the iterative nature of the eigenvalue problem shows up in Fig. 4 as a few slight ''bumps.''

To compare the methods pseudo-quantitatively, estimates of the required flops during statistical studies were obtained. The term ''pseudo'' is used because the flops count only considers the three basic types of operations mentioned above, not the various administrative tasks, variable initialization, passing of variables, etc. Furthermore, the flops count does not take efficient coding into consideration, such as using sparse matrix routines. Hence, the results presented here should not be taken as the true ''costs'' required by actual analyses. However, they should provide a fair basis for a rough comparison of the methods.

Table 1 Assumed model dimensions for flops count

Number of...	Variable	Value
Blades	N	50
Blade Interior (B) DOF	n_b	2000
Disk Independent (D+α) DOF	n_d	1000
Disk-Blade Interface (Γ) DOF	n_c	100
Blade Modes	m_b	3
Disk Modes (per harmonic)	m_d	3

Table 2 Essential steps during model construction and use in a forced response statistical study (M=Modal domain, MP=Mistuning Projection method, RED=REDUCE)

Calculation	C-B	SMART	M. P.	RED.
Sector Normal Modes		M	X	X
Blade Normal Modes	X	X		X
Blade Constraint Modes	X	X		
Disk Normal Modes	X	X		
Disk Constraint Modes	X	X		
Model Assembly	X	X		X
Modal Force	X	X	X	X
Mistuning Projection		M	X	
Response Amplitudes	M	M	M	M

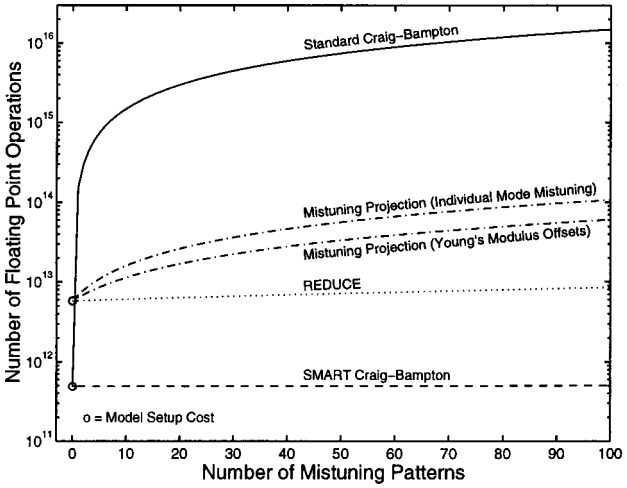


Fig. 5 Comparison of estimated cumulative numbers of floating point operations during statistical analyses, including the model setup cost

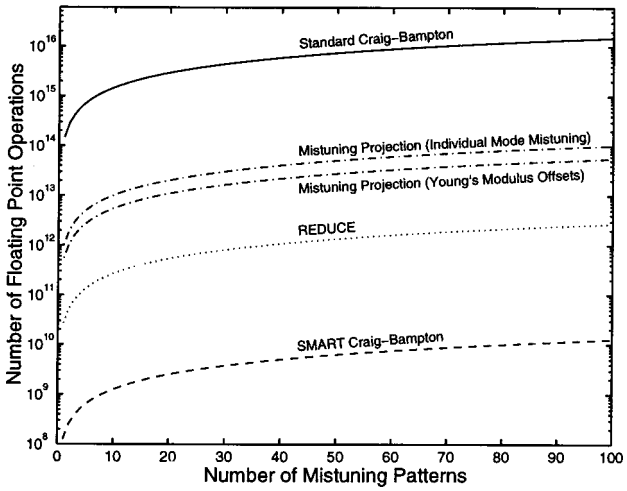


Fig. 6 Comparison of estimated cumulative numbers of floating point operations during statistical analyses, disregarding the model setup cost

The ''initial conditions'' for each method are the required finite element structural matrices for the components. Table 1 outlines the model dimensions used in this comparison. Table 2 outlines the essential steps considered for each method. The upper portion of Table 2 represents the model setup. The lower portion contains the required steps for the mistuned forced response, which must be repeated for each new mistuning pattern. The results are shown in Figs. 5 (model setup cost included) and 6 (model setup cost excluded).

As Fig. 5 indicates, the mistuning projection method and REDUCE carry a similar setup cost, due largely to the set of cyclic eigenvalue problems involving the fundamental sector they have in common. Recall, the blades are massless in the REDUCE method, but the problem size is unaffected by this fact. Note how insignificant is the effort required to go from a CB model to a SMART model. Of course, the set of cyclic eigenvalue problems using the CB model does result in an increased setup cost for the SMART model, but that cost is trivial on this scale.

A mistuned forced response statistics simulation was carried out for Figs. 5 and 6 by calculating the mistuned response at 1000 sampled excitation frequencies for 100 mistuning patterns. From Fig. 5, it is clear that the CB model suffers from carrying all the interface DOF in the reduced order model. It is equally clear that the mistuning projection in the physical domain severely degrades the performance of the mistuning projection method. Note that even in the simplistic case of mistuning via Young's modulus offsets there is a significant number of operations required to project the mistuning. In contrast, studying a sequence of mistuning patterns is relatively effortless with the SMART model.

Figure 6 shows the flops after the setup costs, to highlight the ''long-term'' performance of each method. REDUCE demonstrates good speed, and requires approximately one and four orders of magnitude fewer flops per mistuning pattern than mistuning projection and CB, respectively. However, it must again be emphasized that REDUCE lacks the high accuracy of those two methods. As shown, SMART cuts the flops even further: it is close to four orders of magnitude less computationally intensive per mistuning pattern than mistuning projection. Thus, it is deduced that SMART is clearly the most appealing method for performing vibration analyses of mistuned bladed disks.

8 Conclusions

Component mode synthesis (CMS) is an efficient technique for dynamic analyses of complex structures. However, the applications of CMS to mistuned bladed disk assemblies are remarkably scarce. In this first part of the two-part paper, it was shown how the Craig–Bampton (CB) method of CMS can be applied to mistuned bladed disks in a systematic manner. In addition, two new approaches for efficient and realistic modeling of mistuned bladed disks were presented:

- Mistuning projection method:

Classical modal analysis of the tuned finite element model, followed by a projection of individual blade mistuning onto the retained system modes.

- Secondary modal analysis reduction technique (SMART):

Formulation of a cyclic Craig–Bampton model, followed by classical modal analysis of the full CB model. The mistuning is input in the CB modal space and then projected to the SMART model via the blade portions of the retained secondary modes.

Both methods employ the assumption of a common deformation space for tuned and mistuned mode families. This results in minimized reduced order models, while requiring projection of blade mistuning data onto a selection of tuned modes.

The developed methods were compared by means of a theoretical count of floating point operations required for model setup and use in statistical forced response predictions. It was demonstrated that the most straightforward technique—the mistuning projection method—suffers a high computational cost due to carrying out the mistuning projections in the physical domain. In contrast, the SMART mistuning projections are performed in the low-order modal domain. Furthermore, it was shown that SMART analyses are exceptionally fast. Therefore, it is clear that the SMART approach is well-suited for performing comprehensive studies of mistuned forced response statistics.

Acknowledgments

This work is supported by the GUIde Consortium on blade durability at Carnegie Mellon University.

Appendix A: The Kronecker Product

The Kronecker product of two matrices is defined as

$$\mathbf{A}\otimes\mathbf{B}=\begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \cdots & a_{1N}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \cdots & a_{2N}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}\mathbf{B} & a_{N2}\mathbf{B} & \cdots & a_{NN}\mathbf{B} \end{bmatrix}. \tag{A1}$$

Selected useful properties of the Kronecker product:

$$(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D})=(\mathbf{AC})\otimes(\mathbf{BD}) \tag{A2}$$

$$(\mathbf{A}\otimes\mathbf{B})^{-1}=\mathbf{A}^{-1}\otimes\mathbf{B}^{-1} \tag{A3}$$

$$(\mathbf{A}\otimes\mathbf{B})^{\mathbf{T}}=\mathbf{A}^{\mathbf{T}}\otimes\mathbf{B}^{\mathbf{T}}. \tag{A4}$$

Appendix B: Circulant Matrices

The mass and stiffness matrices of any linear cyclic system may be cast in circulant or block-circulant form. Hence, the application of cyclic symmetry in this paper makes frequent use of the properties of circulant matrices and their eigenvectors. The properties of circulant matrices are thoroughly examined in Davis [25]. The general form of a square circulant matrix is

$$\mathbf{C}=\mathbf{circ}(c_1,c_2,\ldots,c_N)=\begin{bmatrix} c_1 & c_2 & \cdots & c_N \\ c_N & c_1 & \cdots & c_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_2 & c_3 & \cdots & c_1 \end{bmatrix}. \tag{B1}$$

All circulant matrices of order N possess N independent eigenvectors. In particular, they share the same set of eigenvectors that make up the complex Fourier matrix, **E**:

$$\mathbf{E}=[e_{ki}];\ \ e_{ki}=\frac{1}{\sqrt{N}}e^{j\alpha(i-1)(k-1)},\ \ k,i=1,\ldots,N, \tag{B2}$$

where $j=\sqrt{-1}$ and $\alpha=2\pi/N$. In addition, there exists an ''almost-equivalent'' real-valued form of Eq. (*B*2):

$$\mathbf{F}=[\mathbf{f_0}\ \mathbf{f_{1,c}}\ \mathbf{f_{1,s}}\cdots\mathbf{f_{n,c}}\ \mathbf{f_{n,s}}\cdots\mathbf{f_{N/2}}]=\begin{bmatrix} \frac{1}{\sqrt{N}} & \sqrt{\frac{2}{N}} & 0 & \cdots & \frac{1}{\sqrt{N}} \\ \frac{1}{\sqrt{N}} & \sqrt{\frac{2}{N}}\cos\alpha & \sqrt{\frac{2}{N}}\sin\alpha & \cdots & -\frac{1}{\sqrt{N}} \\ \frac{1}{\sqrt{N}} & \sqrt{\frac{2}{N}}\cos 2\alpha & \sqrt{\frac{2}{N}}\sin 2\alpha & \cdots & \frac{1}{\sqrt{N}} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{\sqrt{N}} & \sqrt{\frac{2}{N}}\cos(N-1)\alpha & \sqrt{\frac{2}{N}}\sin(N-1)\alpha & \cdots & \frac{(-1)^{N-1}}{\sqrt{N}} \end{bmatrix}, \tag{B3}$$

where the last column only exists if N is even.

Note that both **E** and **F** are orthonormal, or unitary, such that $\mathbf{E}^*\mathbf{E}=\mathbf{F}^\mathbf{T}\mathbf{F}=\mathbf{I}$, where **I** is an identity matrix of size N, and $*$ denotes the Hermitian adjoint (complex conjugate transpose). In addition, this implies that $\mathbf{E}^{-1}=\mathbf{E}^*$ and $\mathbf{F}^{-1}=\mathbf{F}^\mathbf{T}$, such that the typical transformation products $\mathbf{E}^*\mathbf{C}\mathbf{E}$ and $\mathbf{F}^\mathbf{T}\mathbf{C}\mathbf{F}$ are similarity transformations (Strang [24]).

The reason behind calling **F** ''almost-equivalent'' to **E** is that the columns of **F** are not true eigenvectors of **C**, and hence, the similarity transformation $\mathbf{F}^\mathbf{T}\mathbf{C}\mathbf{F}$ will not yield a diagonalized matrix. However, it will result in a matrix where all nonzero elements will be grouped into 2×2 blocks (''double'' harmonics) on the diagonal, except for the (1, 1) and, for N even, the (N, N) elements (''single'' harmonics). This matrix type is referred to as pseudo-block-diagonal.

These properties are readily extended to the case of block-circulant matrices by expanding **E** and **F** as $\mathbf{E}\otimes\mathbf{I}$ and $\mathbf{F}\otimes\mathbf{I}$, respectively. The scalar c_i then represents a matrix block $\mathbf{C_i}$, where $\mathbf{C_i}$ and **I** are of the same size.

References

[1] Wagner, J. T., 1967, ''Coupling of Turbomachine Blade Vibrations Through the Rotor,'' ASME J. Eng. Power, **89**, pp. 502–512.

[2] Dye, R. C. F., and Henry, T. A., 1969, ''Vibration Amplitudes of Compressor Blades Resulting From Scatter in Blade Natural Frequencies,'' ASME J. Eng. Power, **91**, pp. 182–188.

[3] Ewins, D. J., 1969, ''The Effects of Detuning Upon the Forced Vibrations of Bladed Disks,'' J. Sound Vib., **9**, pp. 65–79.

[4] Ewins, D. J., 1973, ''Vibration Characteristics of Bladed Disc Assemblies,'' J. Mech. Eng. Sci., **15**, pp. 165–186.

[5] El-Bayoumy, L. E., and Srinivasan, A. V., 1975, ''Influence of Mistuning on Rotor-blade Vibrations,'' AIAA J., **13**, pp. 460–464.

[6] Griffin, J. H., and Hoosac, T. M., 1984, ''Model Development and Statistical Investigation of Turbine Blade Mistuning,'' ASME J. Vib., Acoust., Stress, Reliab. Des., **106**, pp. 204–210.

[7] Wei, S. T., and Pierre, C., 1988, ''Localization Phenomena in Mistuned Assemblies with Cyclic Symmetry. I. Free Vibrations,'' ASME J. Vib., Acoust., Stress, Reliab. Des., **110**, pp. 429–438.

[8] Wei, S. T., and Pierre, C., 1988, ''Localization Phenomena in Mistuned Assemblies with Cyclic Symmetry. II. Forced Vibrations,'' ASME J. Vib., Acoust., Stress, Reliab. Des., **110**, pp. 439–449.

[9] Lin, C.-C., and Mignolet, M. P., 1997, ''An Adaptive Perturbation Scheme for the Analysis of Mistuned Bladed Disks,'' ASME J. Eng. Gas Turbines Power, **119**, pp. 153–160.

[10] Srinivasan, A. V., 1997, ''Flutter and Resonant Vibration Characteristics of Engine Blades,'' ASME J. Eng. Gas Turbines Power, **119**, pp. 742–775.

[11] Irretier, H., 1983, ''Spectral Analysis of Mistuned Bladed Disk Assemblies by Component Mode Synthesis,'' *Vibrations of Bladed Disk Assemblies*, ASME, New York, pp. 115–125.

[12] Kruse, M. J., and Pierre, C., 1996, ''Forced Response of Mistuned Bladed Disks Using Reduced-Order Modeling,'' *Proc. 37th AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference*, Vol. 4, AIAA, New York, pp. 1938–1950.

[13] Kruse, M. J., and Pierre, C., 1996, ''Dynamic Response of an Industrial Turbomachinery Rotor,'' *Proc. 32nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, AIAA, New York.

[14] Castanier, M. P., Óttarsson, G., and Pierre, C., 1997, ''A Reduced-Order Modeling Technique for Mistuned Bladed Disks,'' ASME J. Vibr. Acoust., **119**, pp. 439–447.

[15] Bladh, R., Castanier, M. P., and Pierre, C., 1999, ''Reduced Order Modeling and Vibration Analysis of Mistuned Bladed Disk Assemblies with Shrouds,'' ASME J. Eng. Gas Turbines Power, **121**, pp. 515–522.

[16] Yang, M.-T., and Griffin, J. H., 1997, ''A Reduced Order Approach for the Vibration of Mistuned Bladed Disk Assemblies,'' ASME J. Eng. Gas Turbines Power, **119**, pp. 161–167.

[17] Yang, M.-T., and Griffin, J. H., 1999, ''A Reduced Order Model of Mistuning Using a Subset of Nominal System Modes,'' *Proc. 44th ASME Gas Turbine and Aeroengine Technical Congress, Exposition and Users Symposium*, ASME, New York.

[18] Craig, R. R., and Bampton, M. C. C., 1968, ''Coupling of Substructures for Dynamics Analyses,'' AIAA J., **6**, pp. 1313–1319.

[19] Joseph, J. A., 1981, ''Cyclic Symmetry in MSC/NASTRAN,'' *MSC/NASTRAN Application Manual*, The MacNeal-Schwendler Corporation, Los Angeles, CA, Chapter 3.2, pp. 10–24.

[20] Fortescue, C. L., 1918, ''Method of Symmetrical Co-ordinates Applied to the Solution of Polyphase Networks,'' Trans. Am. Inst. Electr. Eng., **37**, pp. 1027–1115.

[21] Craig, R. R., 1981, *Structural Dynamics, An Introduction to Computer Methods*, John Wiley and Sons, New York.

[22] Craig, R. R., 1995, ''Substructure Methods in Vibration,'' ASME J. Mech. Des., **117**, pp. 207–213.

[23] Tan, Y.-C., Castanier, M. P., and Pierre, C., 1999, ''Modal Approximations of Power Flow Between Coupled Component Structures,'' *Proc. Sixth International Congress on Sound and Vibration*, Copenhagen, Denmark, Vol. 5, pp. 2315–2322.

[24] Strang, G., 1988, *Linear Algebra and Its Applications*, 3rd Ed., Saunders, Philadelphia, PA.

[25] Davis, P. J., 1979, *Circulant Matrices*, John Wiley and Sons, New York.

Component-Mode-Based Reduced Order Modeling Techniques for Mistuned Bladed Disks—Part II: Application

R. Bladh
Graduate Student

M. P. Castanier
Assistant Research Scientist

C. Pierre
Professor

Department of Mechanical Engineering,
The University of Michigan,
Ann Arbor, MI 48109-2125

In this paper, the component-mode-based methods formulated in the companion paper (Part I: Theoretical Models) are applied to the dynamic analysis of two example finite element models of bladed disks. Free and forced responses for both tuned and mistuned rotors are considered. Comprehensive comparisons are made among the techniques using full system finite element solutions as a benchmark. The accurate capture of eigenfrequency veering regions is of critical importance for obtaining high-fidelity predictions of the rotor's sensitivity to mistuning. Therefore, particular attention is devoted to this subject. It is shown that the Craig–Bampton component mode synthesis (CMS) technique is robust and yields highly reliable results. However, this is achieved at considerable computational cost due to the retained component interface degrees of freedom. It is demonstrated that this problem is alleviated by a secondary modal analysis reduction technique (SMART). In addition, a non-CMS mistuning projection method is considered. Although this method is elegant and accurate, it is seen that it lacks the versatility and efficiency of the CMS-based SMART. Overall, this work shows that significant improvements on the accuracy and efficiency of current reduced order modeling methods are possible. [DOI: 10.1115/1.1338948]

1 Introduction

The blades of a bladed disk are intended to be identical, but in fact there are always small, random differences among the blades, called mistuning. Mistuning can result in blade forced response amplitudes and stresses that are much larger than those predicted for a perfectly tuned rotor. (See Srinivasan [1] for a comprehensive survey of the literature.) Thus, mistuning has a critical impact on blade fatigue life in turbine engines, and it is of great importance to be able to predict the mistuned forced response in an accurate and efficient manner. Several recent studies have presented reduced order modeling techniques that are capable of generating low order models of bladed disks from parent finite element models (Irretier [2]; Zheng and Wang [3]; Kruse and Pierre [4,5]; Castanier et al. [6]; Yang and Griffin [7,8]; Bladh et al. [9]).

In the companion paper [10], some new reduced order modeling techniques were introduced that are well-suited for the efficient and accurate vibration analysis of mistuned bladed disks. These techniques are summarized as follows:

- **Craig–Bampton (CB) Method:** The CB method (Craig and Bampton [11]) of component mode synthesis (CMS), formulated specifically for bladed disks, is employed. A cyclic symmetry description is used for the disk component. Each blade is treated as a separate component.
- **Mistuning Projection Method:** A classical modal analysis is performed on the full finite element model, and then the individual blade mistuning is projected onto the cyclic system modes.
- **Secondary Modal Analysis Reduction Technique (SMART):** A primary reduced order model is generated using CMS (e.g., the CB method), and then a secondary modal analysis is performed on the full CMS matrices. Only the system modes of interest are retained. Blade mistuning is introduced in the CMS matrices (blade component degrees of freedom (DOF)) and then projected from the primary modal coordinates to the secondary modal coordinates.

In addition, the method of Castanier et al. [6] was revisited, which is here denoted REDUCE. Since the code is being actively used in industry, REDUCE represents the current state of the art.

In this paper, the above-mentioned techniques are applied to simple, yet representative finite element models of bladed disks. First, tuned free vibrations are considered, and the methods are compared by studying their modal convergence trends. The eigenfrequency veerings that occur when plotting natural frequencies versus number of nodal diameters have been shown to play a crucial role in determining a bladed disk's sensitivity to mistuning [4,12]. Therefore, the modal convergence study focuses on how well the derived methods capture the eigenfrequencies in a particular veering region.

Next, it is demonstrated how well the methods represent mistuned, localized modes of vibration. A modal convergence study is presented with respect to the capture of both mistuned natural frequencies and mistuned mode shapes using the well-known modal assurance criterion, or MAC [13]. This study is relevant, since the introduction of mistuning data differs among the methods and results in an approximation for some of them. Hence, a good tuned representation does not necessarily translate to a good mistuned representation.

Predictions of forced response amplitudes are then considered, for both tuned and mistuned configurations. This is the primary application of these methods. The efficiency and accuracy of the methods are examined via forced response frequency sweeps. The SMART approach is also applied to a large-size model, using a CB model as the intermediate CMS basis. The excellent accuracy and tremendous computational savings that result from using this approach are demonstrated.

This paper is organized as follows. The finite element models used in this study are presented in Sec. 2. The validation of the

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-361. Manuscript received by IGTI January 14, 2000; final revision received by ASME Headquarters April 2, 2000. Associate Editor: M. Mignolet.

developed methods begins with free vibration results in Sec. 3, where particular emphasis is placed on the capture of tuned natural frequency veerings and the mistuned mode shapes. Forced response results are examined in Sec. 4. In Sec. 5, the SMART approach is validated for a large-size model to further demonstrate the potential of this method. Finally, concluding remarks are given in Sec. 6.

2 Description of Example Finite Element Models

Two different finite element models of mistuned bladed disks are used for the validation of the developed methods:

- A simple, ''small'' model (see Fig. 1) that allows for extensive studies of modal convergence.
- A more realistic, ''large'' model (see Fig. 2) that is used to demonstrate the potential of the SMART approach.

The finite elements are all eight-noded brick (linear solid) elements. It should be noted that both meshes are relatively coarse compared to that of a typical industrial finite element model, and they may not represent precisely the behavior of actual bladed disks with the depicted geometries. However, this is not a concern for this study, since all of the methods are applied to the same parent finite element models.

The material properties for both models are those of steel (see Table 1). However, the models differ in the structural damping coefficient, G, used during the forced response. Note that viscous damping is not considered in this study.

For the mistuned results, a single mistuning pattern is used for each model. The mistuning patterns were sampled from a uniform distribution of mean zero, and one mistuning value, δ_n, is assigned to each blade. The mistuning is introduced to the full finite element model by varying Young's modulus in the blade elements:

$$E_n = (1+\delta_n)E_o, \quad n = 1, \ldots, N, \tag{1}$$

where n is the blade number and N is the total number of blades. Note that the methods developed in this study are not restricted to mistuning by variations in Young's modulus. However, individual mode mistuning is not considered in this paper.

2.1 Small Example Finite Element Model. The smaller of the two finite element models that are analyzed in this study is shown in Fig. 1. Due to its relatively low number of DOF, this model is used extensively in this study. This model features 12 blades, each with length 60 mm and base width 7.5 deg (~13 mm). The blades are slightly tapered along the radial direction, from 5 mm thickness at the base to 2 mm thickness at the tip. The disk is 5 mm thick. It has an outer radius of 100 mm, and an inner radius of 20 mm. The disk is clamped at the inner radius.

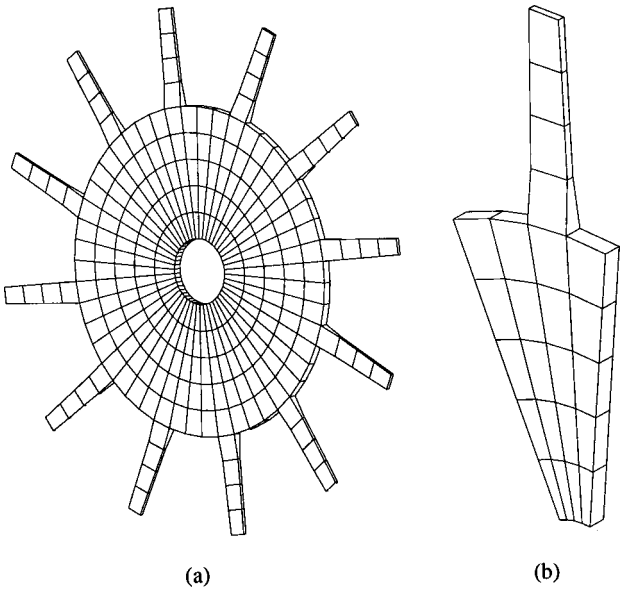


Fig. 1 Finite element meshes for the ''small'' example blisk: (*a*) the full model; (*b*) the fundamental sector

For the Craig–Bampton method, the fundamental sector in Fig. 1(*b*) is further substructured into a disk sector component (for a cyclic symmetry analysis) and a blade component. Model data for the components are listed in Table 2. For the forced response, the structural damping coefficient is taken to be $G = 0.01$.

The single mistuning pattern used for this model is taken from a uniform distribution of mean zero and standard deviation 3.0 percent. The mistuning parameters, δ_n, used for each blade are listed in Table 3.

2.2 Large Example Finite Element Model. The finite element model of the ''large'' test case rotor is depicted in Fig. 2. This model is used in Sec. 5 as a more realistic case. This rotor has 24 blades. Each blade has a base pitch of 30 deg (measured from the axial direction), and a uniform twist of an additional 30 deg over its length. The base radius is 212 mm, and the blade

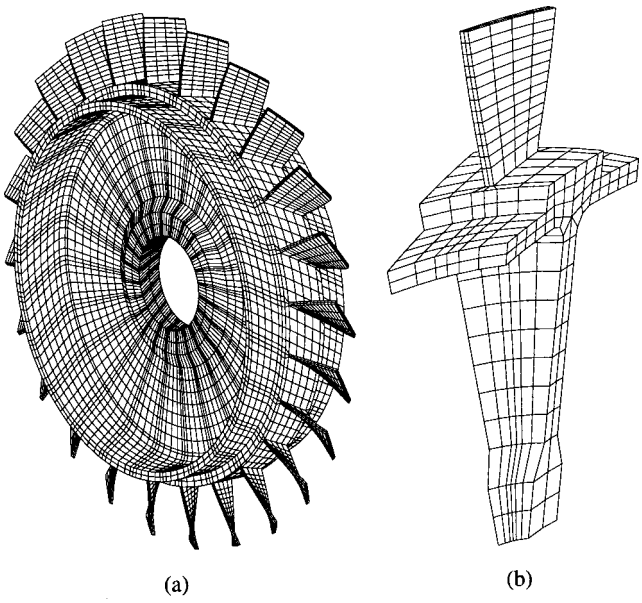


Fig. 2 Finite element meshes for the ''large'' example blisk: (*a*) the full model; (*b*) the fundamental sector

Table 1 Material properties (generic steel) for both models

Property	Notation	Value	Unit
Nominal Young's Modulus	E_o	$2.00 \cdot 10^5$	MPa
Poisson's Ratio	ν	0.25	—
Mass Density	ϱ	$7.86 \cdot 10^3$	kg/m^3

Table 2 Basic model data for the small model

Component	Elements	Nodes	DOF
Cantilevered Blade	4	20	48
Disk Sector	20	60	108
Disk-Blade interface	—	4	12
Fundamental Sector	24	76	168
Full Assembly	288	768	2016

Table 3 Single mistuning pattern for the small model

Component	Elements	Nodes	DOF
Cantilevered Blade	104	210	585
Disk Sector	372	651	1521
Disk-Blade interface	—	15	45
Fundamental sector	476	846	2151
Full assembly	11,424	18,072	51,624

Table 4 Basic model data for the large model

Blade	δ_n	Blade	δ_n
1	-0.0276	7	0.0132
2	0.0050	8	0.0207
3	0.0449	9	-0.0107
4	-0.0171	10	-0.0090
5	0.0161	11	0.0161
6	-0.0112	12	0.0351

Table 5 Single mistuning pattern for the large model

Blade	δ_n	Blade	δ_n
1	-0.00515	13	-0.08075
2	-0.01875	14	-0.04905
3	-0.01820	15	0.05935
4	-0.00390	16	-0.06925
5	-0.05005	17	-0.00415
6	-0.00850	18	0.00430
7	0.01415	19	0.01840
8	0.07620	20	-0.05475
9	0.02930	21	0.02395
10	0.02720	22	0.03810
11	0.02770	23	0.04110
12	-0.04925	24	0.03930

length is 68 mm. The rotor is fixed at the interfaces toward adjacent stages. This is believed to provide a reasonable description of the dynamics of a bladed disk assembly.

As was done with the smaller model, the fundamental sector is substructured into a disk sector and a blade component. The model data are found in Table 4. This model is assigned a structural damping coefficient of $G=0.0025$.

The single mistuning pattern is taken from a uniform distribution of mean zero and standard deviation 5.0 percent. The corresponding mistuning parameters, δ_n, are listed in Table 5.

3 Free Vibration

3.1 Tuned Assembly. In Fig. 3, an assessment is made of the small model's tuned characteristics by plotting the natural frequencies of the tuned system versus the number of nodal diameters. The nearly horizontal connecting lines correspond to assembly modes that are dominated by blade motion, while the slanted connecting lines correspond to disk-dominated modes. The rapid increase of the eigenfrequencies of the disk-dominated modes is due to stiffening of the disk as the circumferential wavelength decreases with increasing number of nodal diameters.

Two significant eigenfrequency veering regions are highlighted in Fig. 3: one located at one nodal diameter, around 3300 Hz (veering 1); and a second located at three nodal diameters, around 6500 Hz (veering 2). Earlier studies (Kruse and Pierre [4]; Bladh et al. [9]) have shown that the characteristics of eigenfrequency veerings are crucial in determining a design's sensitivity to mistuning. Furthermore, it has been found that maximum mistuned forced response amplitudes are likely to occur in veering regions (that is, when the frequency and engine order of excitation correspond to the frequency and nodal diameters of a veering). Therefore, the ability to capture veering regions is an important consideration when assessing the performance of a modeling method.

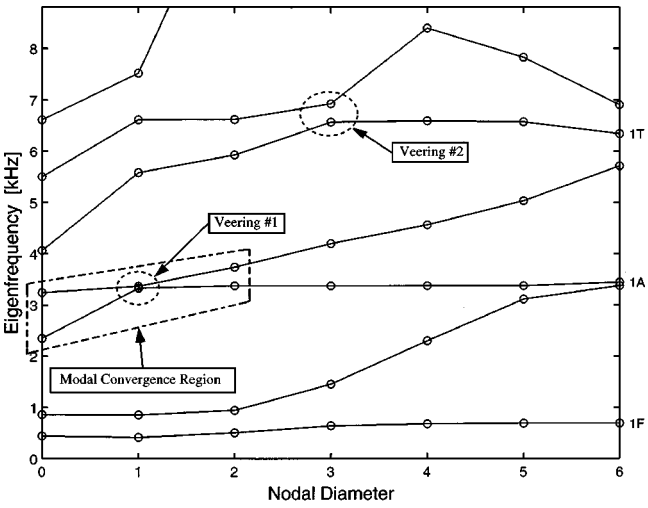


Fig. 3 Natural frequencies versus nodal diameters for the small example finite element model. The circles show the natural frequency values, while the connecting lines are drawn to aid in visualization of the mode families and the frequency veerings. The character of each family of blade-dominated modes is indicated on the right, where F=Flex, T=Torsion, and A=Axial (edgewise) bending.

As pointed out in Part I, all methods considered in this study are complete, in the limit, for the tuned case: They yield the finite element solution when all the component modes are included in the model. Hence, for the tuned case, one may compare the performance of the methods by examining their rates of modal convergence. In combination with the above-mentioned veering discussion, the methods are evaluated by comparing modal convergence trends based on the eigenfrequency errors in the ''modal convergence region'' of Fig. 3.

The modal convergence for each method (except for mistuning projection, which is irrelevant for the tuned case) is presented in Fig. 4. This figure shows the average percent error among the six distinct eigenfrequencies in the modal convergence region versus the number of retained modes in the reduced order model (ROM). Note that there are actually ten modes in this region: four double modes, and two single modes. However, to avoid counting the double-mode errors twice, this region is evaluated at only the six distinct eigenfrequencies. The results in Fig. 4 clearly illustrate the remarkable difference in both accuracy and modal convergence rate between the CB model and REDUCE. However, it must be emphasized that the superiority of the CB model is achieved at a considerable expense in terms of minimum model size relative to REDUCE. As shown, CB does not capture this veering at all with fewer than around 200 DOF, while REDUCE yields a fair veering representation with much fewer than 100 DOF.

The vertical line in Fig. 4 represents the SMART approach with a CB model as intermediate CMS basis. The implication of the vertical line is that the tuned SMART model is always as accurate as the intermediate CMS model from which it derives, while its size stays constant at the number of modes selected by the analyst. Hence, it may be viewed as collapsing the CB model onto an arbitrarily smaller, *n*-dimensional subspace. Alternatively, a graphical interpretation using Fig. 4 would be to collapse a CB model of certain accuracy horizontally to the left onto a point on an arbitrarily positioned vertical line. In Fig. 4, where only six

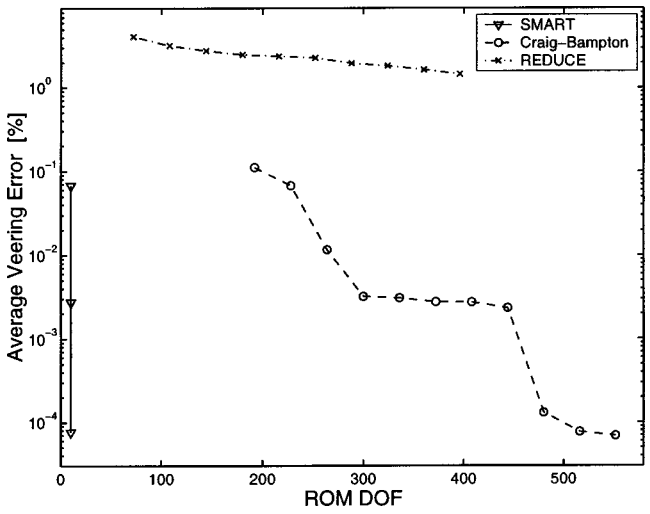


Fig. 4 Modal convergence trends in the region surrounding veering 1 for the small example model

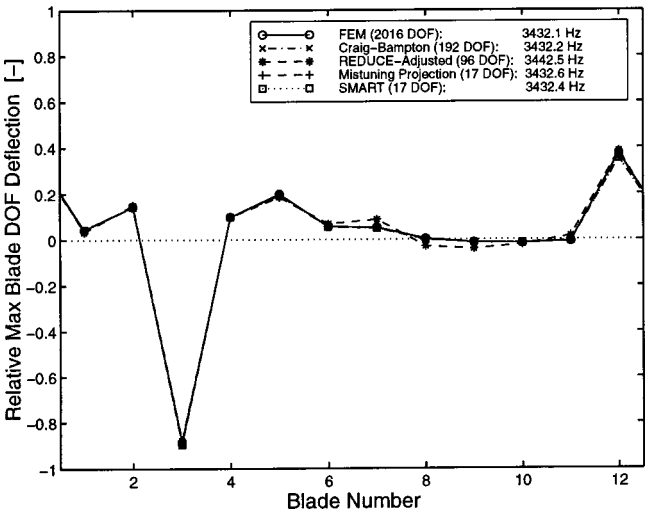


Fig. 6 Reduced order model representations of mistuned mode shape number 38 for the small example model

modes are considered, the SMART model has only the minimum six DOF. Note that using SMART for tuned free vibration analyses makes no sense; it is included here only to aid in demonstrating the approach.

3.2 Mistuned Assembly. The small model is now mistuned by offsets in blade Young's modulus, in accordance with Eq. (1) and Table 3. First, the selected methods' representations of two mistuned mode shapes are considered. The finite element mode shapes and the approximations obtained through the various methods are depicted in Figs. 5 and 6. The obtained mode shape representations come from first finding the maximum physical displacement of any DOF in any blade. The physical displacements of the DOF where the maximum was found are then plotted for each blade, normalized by the Euclidean norm of the N displacements as

$$A^n = \frac{z_m^n}{\sqrt{\sum_{k=1}^{N} z_m^{k^2}}}, \quad n = 1, \ldots, N, \tag{2}$$

where z_m^n is the displacement in physical coordinates for the nth blade's mth DOF (the DOF at which the maximum was found).

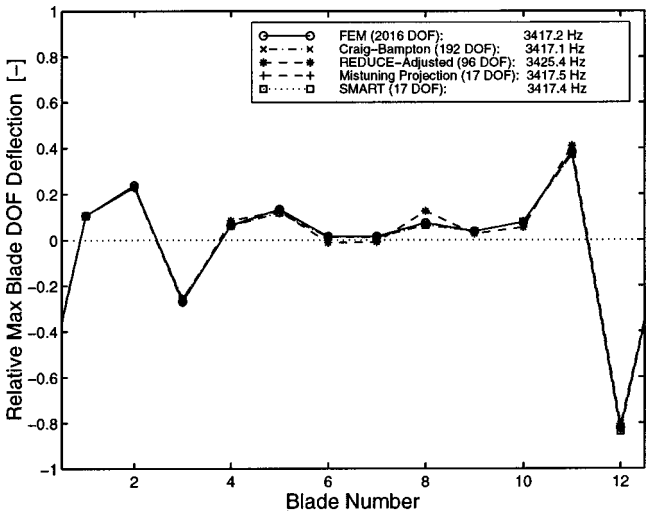


Fig. 5 Reduced order model representations of mistuned mode shape number 37 for the small example model

The two modes depicted in Figs. 5 and 6 belong to the second family of blade-dominated system modes (1A, see Fig. 3) located around 3300 Hz. The two mode shapes can be characterized as intermediately localized, which means that several blades participate in the motion, but there is still a clear dominance by one or two blades. This type of intermediate localization has been shown to cause the most severe increases in resonant forced response amplitudes [9,14].

The SMART model used here is based on an intermediate CB model that incorporates the first seven normal disk modes and the first six normal blade modes (plus the complete set of constraint modes). Furthermore, the 17 tuned modes that fall inside the frequency range 3000–3500 Hz are selected for both the mistuning projection method and the SMART approach. This means that the two models include all harmonics of the tuned modes associated with the second family of blade-dominated modes. This completeness in terms of harmonic content is a basic requirement for these methods.

Except for REDUCE, note the excellent agreement among eigenfrequencies in Figs. 5 and 6. The relative error is less than 0.02 percent for both modes. REDUCE, however, cannot quite match these excellent results with errors around 0.3 percent. These results translate nicely to the mode shapes as well, where the approximated shapes are virtually indistinguishable from those obtained by full finite element analysis. Although REDUCE's mode shape representations are not as accurate, the method does an excellent job in capturing the relative peaks as well as their locations.

Next, an aspect of method efficiency is investigated. All the obtained approximations of natural frequencies and mode shapes are compared to the finite element results, in order to establish how many system modes each method captures versus the number of possible modes, m_r (i.e., the total number of DOF in the ROM). A system mode is regarded as ''captured'' if the error is below a pre-defined tolerance—δ_f for natural frequencies, or δ_{mac} for mode shapes. The obtained natural frequencies, f_i^r, are compared with the finite element frequencies, f_j^{fe}, through straightforward fractions

$$\frac{f_i^r}{f_j^{\text{fe}}} - 1 \leqslant \delta_f, \quad \begin{matrix} i = 1, \ldots, m_r \\ j = 1, \ldots, m_{\text{fe}} \end{matrix}, \tag{3}$$

where m_{fe} is the total number of DOF in the full finite element model. If the above inequality is satisfied for any f_j^{fe}, then the

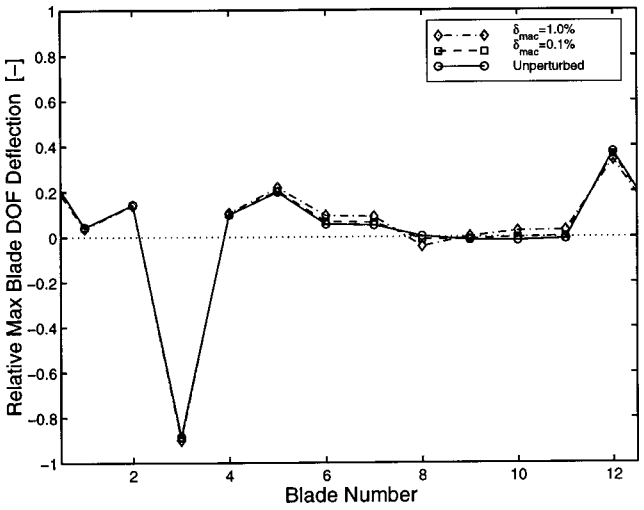


Fig. 7 Sensitivity of modal assurance criterion (MAC) values (mistuned mode shape number 38)

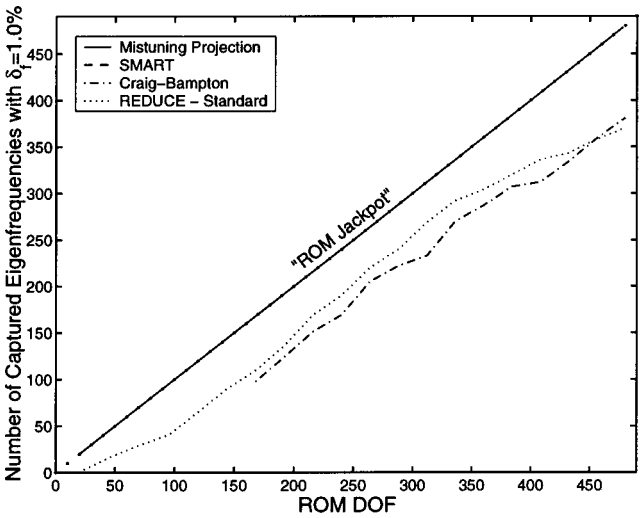


Fig. 9 Method efficiency with respect to mistuned natural frequencies for $\delta_f \leqslant 1.0$ percent for the small example model

natural frequency f_i^r is regarded as captured. In a similar fashion, the approximated mode shapes, $\psi_{\mathbf{i}}^{\mathbf{r}}$, are compared to the exact ones via MAC conformity:

$$1 - \frac{(\psi_{\mathbf{i}}^{\mathbf{r}^T} \psi_{\mathbf{j}}^{\mathbf{fe}})^2}{(\psi_{\mathbf{i}}^{\mathbf{r}^T} \psi_{\mathbf{i}}^{\mathbf{r}})(\psi_{\mathbf{j}}^{\mathbf{fe}^T} \psi_{\mathbf{j}}^{\mathbf{fe}})} \leqslant \delta_{\mathrm{mac}}, \quad \begin{matrix} i = 1, \ldots, m_r \\ j = 1, \ldots, m_{\mathrm{fe}} \end{matrix} \tag{4}$$

where $\psi_{\mathbf{j}}^{\mathbf{fe}}$ is the finite element mode shape of the jth mode.

It is clear that when an exact match between two mode shapes is obtained, the MAC ratio becomes one; it is between zero and one in all other cases. However, it is not clear just how good a match a MAC value of, for instance, 0.99 represents. To assess this, the 38th mistuned mode shape in Fig. 6 was perturbed by applying randomly generated scale factors to each blade's relative amplitude. The MAC conformity with respect to the unperturbed shape was then computed. The resulting perturbed mode shapes are shown in Fig. 7 for MAC values of 0.99 and 0.999. Note that a 1.0 percent deviation from a perfect match might be considered a very good representation of the mode shape.

The two different tolerance levels (1.0 percent and 0.1 percent) were employed in order to get an indication of how large the errors were among the unacceptable modes. The separate results for natural frequencies and mode shapes are shown in Figs. 8–11. As indicated, the straight line, $y=x$, represents what could be called the ''ROM Jackpot,'' which means that one system mode of desired accuracy is captured for each DOF in the ROM.

The CB formulation suffers considerably from having a full, physical set of interface DOF in the ROM. These DOF, the constraint modes, are of course necessary in order to form the flexible motion of the otherwise fully constrained interfaces. However, they also result in many purely computational modes, which have little or no physical meaning, leading to the relatively poor efficiency displayed by the CB method. Moreover, by relaxing the tolerance level to 1.0 percent, the results show that CB gives only fair (0.1–1.0 percent) representations of several modes.

Note that the REDUCE model quickly yields ''recognizable'' mode shapes. Moreover, observing the dramatic difference between Figs. 8 and 9, it is clear that most of the natural frequencies fall between the two tolerance levels. However, these results show that very few modes are predicted with high accuracy, which is

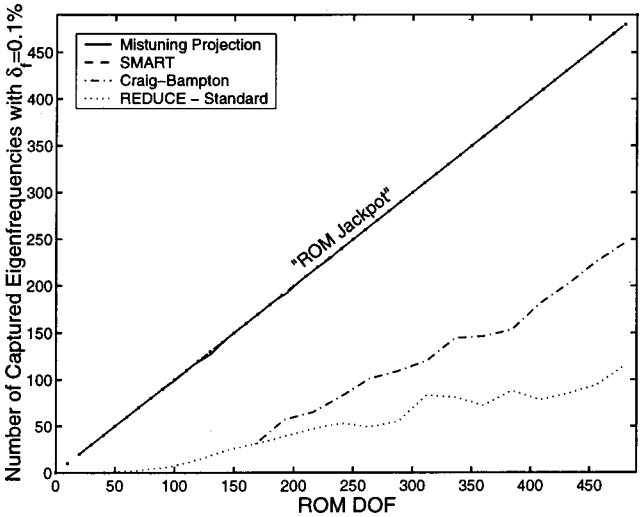


Fig. 8 Method efficiency with respect to mistuned natural frequencies for $\delta_f \leqslant 0.1$ percent for the small example model

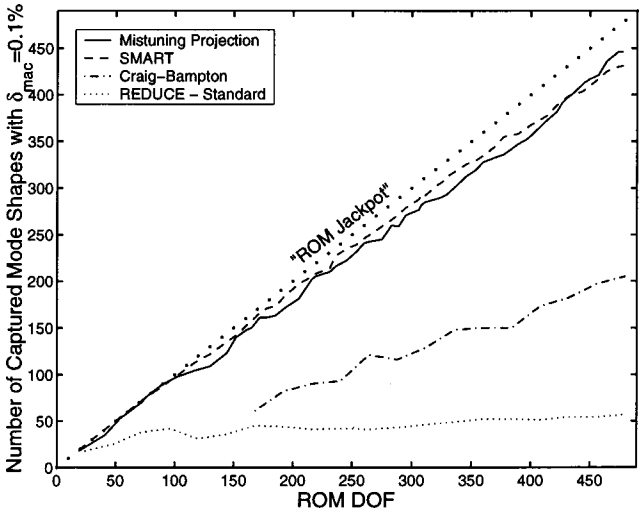


Fig. 10 Method efficiency with respect to mistuned mode shapes for $\delta_{\mathrm{mac}} \leqslant 0.1$ percent for the small example model

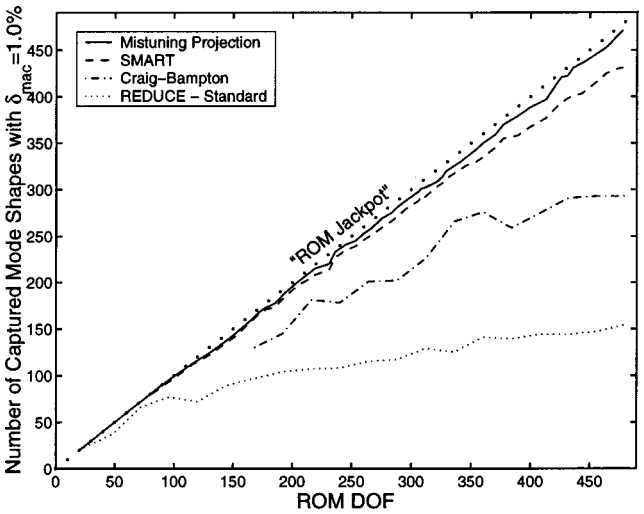


Fig. 11 Method efficiency with respect to mistuned mode shapes for $\delta_{mac} \leq 1.0$ percent for the small example model

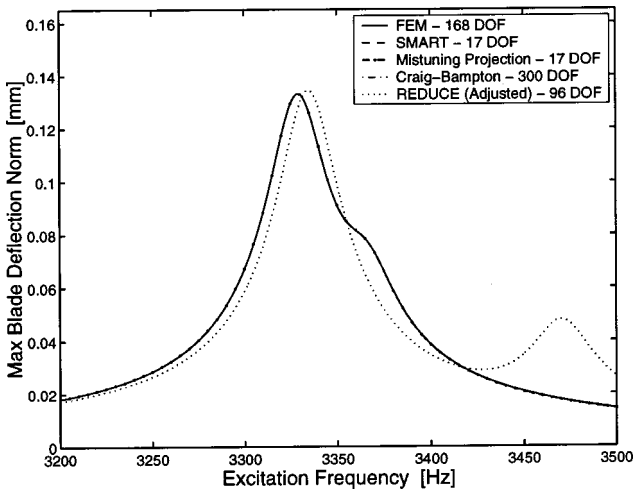


Fig. 12 Forced response frequency sweep through veering 1 for engine order one (1E) excitation of the tuned small example model

evidenced by the very poor progress as the number of retained modes increases for the smaller tolerance (Fig. 8). This is consistent with the tuned results in Fig. 4.

In contrast, both the mistuning projection method and the SMART approach come very close to the ideal ''jackpot'' state for the lower order modes. The capture of mistuned natural frequencies (Figs. 8 and 9) is outstanding. Both methods also display excellent efficiency in terms of mode shapes (Figs. 10 and 11), although they appear to be less impressive in this respect. However, the mode match requirements imposed by the MAC tolerances are very strict (see Fig. 7).

There is a notable deterioration in mode shape representation of the SMART model as the mode numbers increase. Note that the missed mode shapes in Fig. 10 are not accepted even with the relaxed tolerance in Fig. 11. This may be explained, in part, by the approximation implied by neglecting mistuning effects among the constraint modes in the intermediate CB model. First, for higher modes there is more and more local waviness in the structure. Second, since the constraint modes are obtained by successive unit interface DOF deflections, they are also very local in nature. Thus, it is hypothesized that the neglected mistuning among the constraint modes has more impact as the order of the mode increases. This is not a problem for the mistuning projection method, where the entire physical blade stiffness matrices are rescaled properly. (The mistuning projection method does degrade somewhat for higher order modes, but to a much lesser extent.) Furthermore, note that for this method (as opposed to the SMART model) several of the misrepresented mode shapes are still within the relaxed tolerance level. This degradation is likely due to the approximation incurred by representing the mistuned modes with a truncated set of tuned modes. Hence, while yielding accurate representations of the relatively smooth blade shapes at lower modes, the effects of this approximation become more pronounced for higher modes in which increasing local waviness is observed. However, in the limit, the mistuning projection method will yield the exact solution in the particular case of mistuning by Young's modulus offsets.

4 Forced Response

In this section, engine order excitation is considered in the two veering regions indicated in Fig. 3. For veering 1, this implies an engine order one excitation (1E), which for a 12-bladed assembly has a blade-to-blade forcing phase shift of 30 deg. Veering 2 requires an engine order three excitation (3E) with a blade-to-blade forcing phase shift of 90 deg. In both cases, the force is a unit nodal load, $(1, 1, 1)/\sqrt{3}$, applied to one of the nodes on the blade tip. Hence, this force is able to excite all the fundamental modes of the blade. This applied force was chosen arbitrarily, but it serves to verify the accuracy of the reduced order models.

The amplitude metric employed here, the maximum blade deflection norm, is a scalar value based on the Euclidean norm of the physical displacement vector of each blade. The plotted norms (i.e., the maximum norms) are thus obtained at each driving frequency as

$$A = \max\{A^n\} = \max\left\{ \sqrt{\sum_{j=1}^{n_c+n_b} |z_j^n|^2} \right\}, \quad n = 1, \ldots, N, \quad (5)$$

where $|z_j^n|$ is the magnitude of the complex displacement in physical coordinates for the jth DOF of the nth blade, and n_c and n_b are, respectively, the numbers of disk–blade interface and blade interior DOF for one blade. As an alternative, it is realized that from any reduced order model, the physical displacement vector for the nth blade, $\mathbf{z^n}$, is recovered from a modal expansion via some blade modal matrix, $\mathbf{U^n}$, as $\mathbf{z^n} = \mathbf{U^n q}$. Thus, the vector norms may be represented in an equivalent matrix notation as

$$A = \max\{A^n\} = \max\{\sqrt{\mathbf{q}^* \mathbf{U^{n*}} \mathbf{U^n q}}\}, \quad n = 1, \ldots, N, \quad (6)$$

where $*$ denotes the complex conjugate transpose. Note that the inner matrix products, $\mathbf{U^{n*}U^n}$, may be computed and saved prior to any forced response frequency sweeps or extensive statistical simulations. This allows for fast computations of the blade norms through low-order modal domain matrix projections, instead of tedious computations in physical coordinates.

The resulting tuned and mistuned forced response amplitudes from frequency sweeps over veering 1 are illustrated in Figs. 12 and 13, respectively. Clearly, the excellent accuracy displayed by CB, SMART, and the mistuning projection method in free vibrations translates nicely to the forced response. As demonstrated in Fig. 12, the tuned response approximations are practically indistinguishable from the exact solution for these methods. Also, the capture of the maximum amplitude and the general resonance behavior over the frequency range is outstanding for these three methods. Again, the REDUCE method lacks accuracy in comparison. In particular, in Fig. 12 REDUCE predicts a wide separation of the two tuned resonant frequencies, while, in reality, they are very close. Nevertheless, REDUCE gives a fair prediction of the mistuned peak amplitudes in Fig. 13, which is the key quantity in studies of forced response statistics.

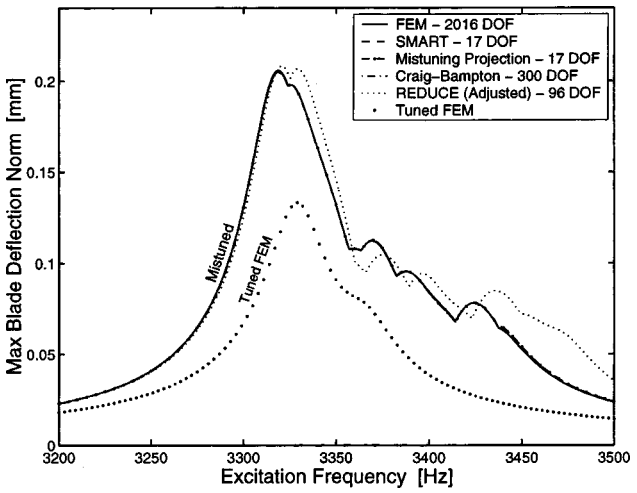


Fig. 13 Forced response frequency sweep through veering 1 for engine order one (1E) excitation of the mistuned small example model (tuned FEM solution included for reference)

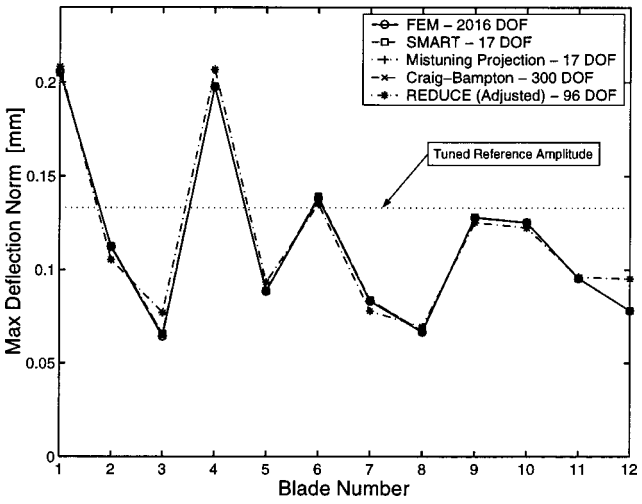


Fig. 15 Maximum response amplitudes for all blades from forced response frequency sweep through veering 1 for engine order one (1E) excitation of the mistuned small example model

A second sample of the methods' performances for the mistuned forced response is shown in Fig. 14, which illustrates a frequency sweep through veering 2 for an engine order three (3E) excitation. Qualitatively, the results are similar to those for veering 1, despite very different circumstances in terms of principal blade motion (different blade mode family) and engine order.

Another point of interest for the turbomachinery industry is blade-to-blade dispersion: blade-to-blade variations in maximum mistuned response amplitudes for a given rotor. This is used to assess mistuning levels from experimental data, and it is also an important aspect when matching computational models with experiments. Figure 15 displays the maximum response amplitudes obtained for each blade from the forced response frequency sweep of Fig. 13. Again, all methods but REDUCE produce results that are virtually indistinguishable from the finite element values.

Based on the presented results, it is concluded that the CB, SMART, and the mistuning projection methods are all consistent, high-performance reduced order modeling techniques suitable for mistuned bladed disks. Taking into consideration the differences in efficiency and the final ROM sizes, it is clear that SMART must be considered to be the premier method examined in this study.

5 Application of SMART to a Large-Size Model

The feasibility and performance of the SMART approach is further highlighted in this section, where it is applied to the large-size model introduced in Sec. 2.2. The REDUCE method is also included in this comparison to represent the current state of the art.

The tuned natural frequencies of the large model are plotted versus the number of nodal diameters in Fig. 16. Note the excellent global representation of the tuned characteristics displayed by the CB model. For its subsequent use in the SMART approach, it was desirable to get a very accurate CB model that also spanned a fairly wide frequency band. Therefore, as many as 12 normal disk modes and 16 normal blade modes were employed in its construction.

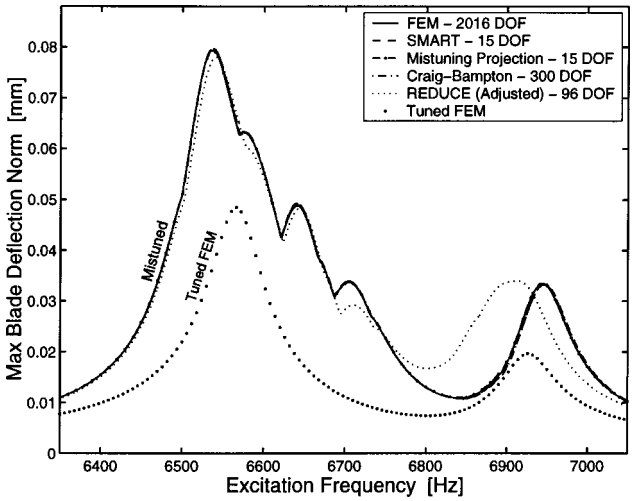


Fig. 14 Forced response frequency sweep through veering 2 for engine order three (3E) excltation of the mistuned small example model (tuned FEM solution included for reference)

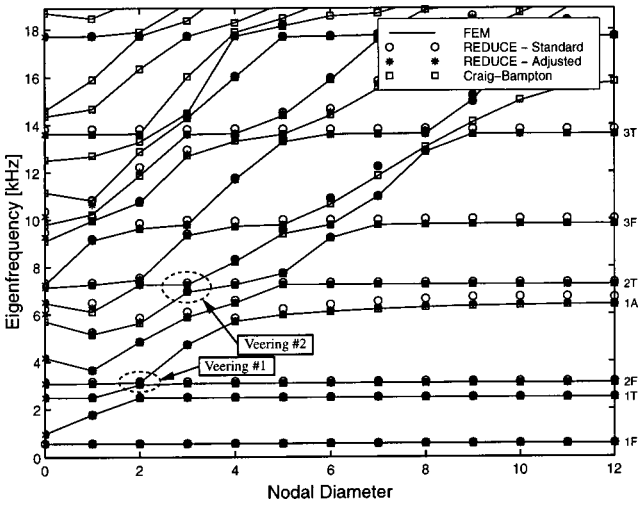


Fig. 16 Natural frequencies versus nodal diameters for the large example model. The character of each family of blade-dominated modes is indicated on the right, where F=Flex, T=Torsion, and A=Axial (edgewise) bending.

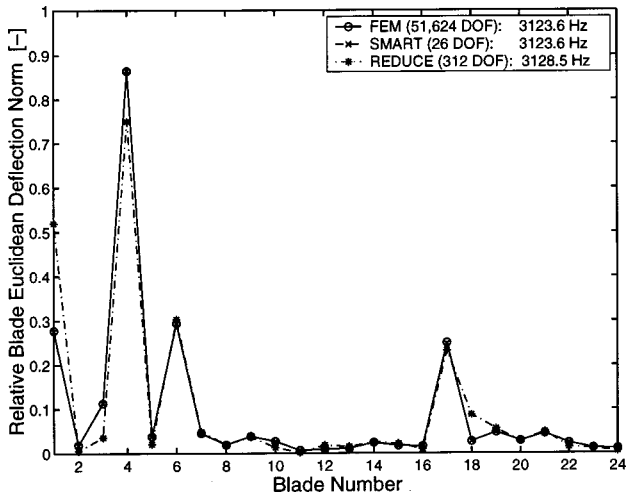


Fig. 17 Reduced order model representations of mistuned mode shape number 64 for the large example model

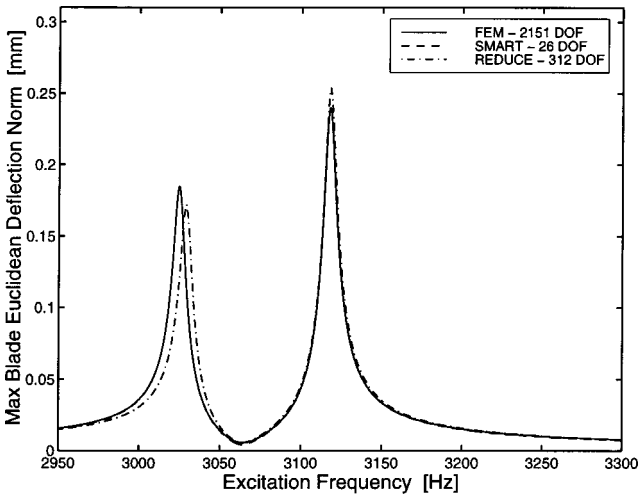


Fig. 19 Forced response frequency sweep through veering 1 for engine order two (2E) excltation of the tuned large example model

In contrast, the REDUCE model uses only 5 disk modes and 8 blade modes (312 DOF in total), which is why the REDUCE model is missing certain system modes. Considering this inequality in number of retained modes, it may seem unfair to compare the CB and REDUCE models. However, the improvement gained by including the same number of modes in the REDUCE model as in the CB model is fairly marginal (recall the poor modal convergence displayed by the REDUCE method in Fig. 4).

First, the SMART representation of mistuned mode shapes is investigated. The mode shapes are represented by the vector norms in accordance with Eq. (6), although scaled to represent the relative blade deflection. Using the notation of Eq. (6), the relative blade displacement norm for each blade is computed as

$$\overline{A^n} = \frac{\sqrt{\mathbf{q}^* \mathbf{U}^{n*} \mathbf{U}^n \mathbf{q}}}{\sqrt{\sum_{k=1}^{N} \mathbf{q}^* \mathbf{U}^{k*} \mathbf{U}^k \mathbf{q}}}, \quad n = 1, \ldots, N. \tag{7}$$

Figures 17 and 18 illustrate the 64th and 135th mistuned mode shapes obtained with the single mistuning pattern listed in Table 5. Note from the mistuned natural frequencies that the selected mode shapes are taken from the two families of blade modes that subsequently will be explored in terms of forced response. As indicated, the SMART models in the respective frequency bands contain 26 and 31 modes (DOF), compared to 51,624 DOF for the full mistuned finite element model. Hence, the model size is reduced by more than three orders of magnitude, and yet the SMART models continue to exhibit excellent accuracy in the mistuned case, both in terms of mistuned natural frequencies and mistuned mode shapes. Even though the REDUCE model is approximately one order of magnitude larger than the SMART model, it does not possess the same accuracy.

As shown in Fig. 16, the large model exhibits several significant veering regions. This section focuses on the two indicated veering regions, which correspond to a 2E and 3E excitation for the forced response (30 and 45 deg blade-to-blade forcing phase shift, respectively). The external excitation force used here consists of a unit nodal load applied in the axial direction on the tip of the blade's leading edge. Again, this applied force was chosen arbitrarily for the purpose of verifying the accuracy of the reduced order models.

Figure 19 depicts the tuned response in veering 1, which exhibits well-separated lower and upper resonances. As expected, the SMART model yields a close-to-perfect match with the finite el-

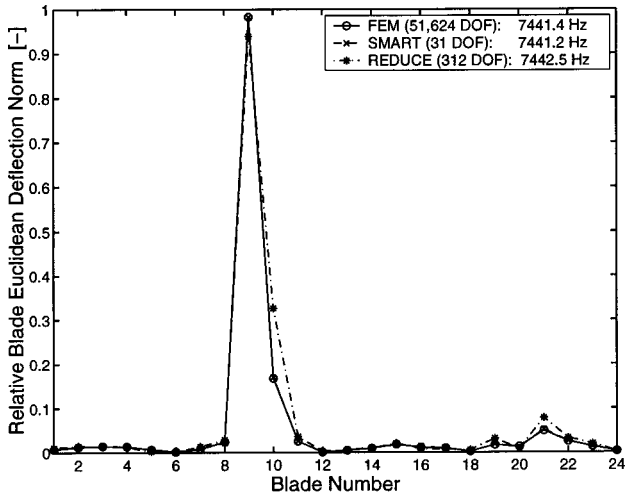


Fig. 18 Reduced order model representations of mistuned mode shape number 135 for the large example model

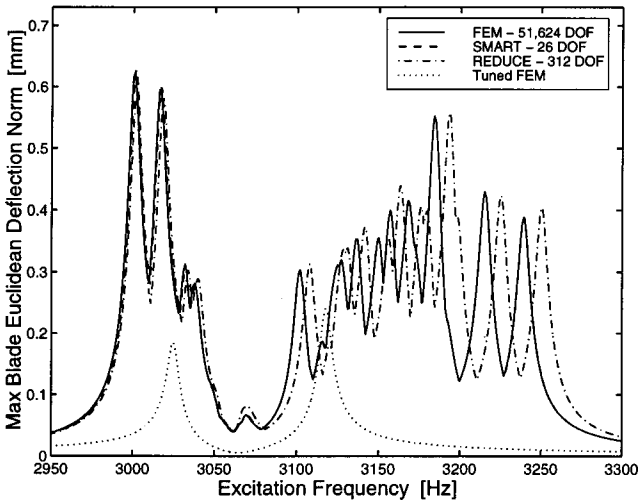


Fig. 20 Forced response frequency sweep through veering 1 for engine order two (2E) excitation of the mistuned large example model (tuned FEM solution included for reference)

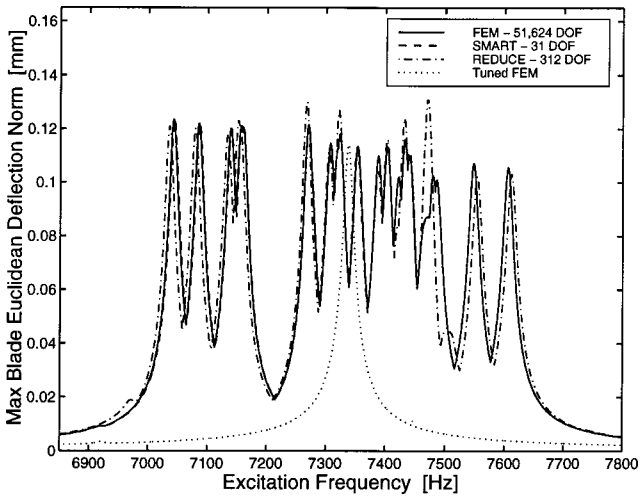


Fig. 21 Forced response frequency sweep through veering 2 for engine order three (3E) excitation of the mistuned large example model (tuned FEM solution included for reference)

ement solution, both in terms of resonant frequencies and peak amplitudes. It is further noted that REDUCE has slight offsets in the resonant frequencies, despite adjustments. What is more disturbing, however, is the fact that the maximum tuned amplitude is overestimated. Using this maximum tuned amplitude as reference in a statistical analysis might result in significantly *underestimated* amplifications due to mistuning, which is not acceptable.

The mistuned response in veering 1 is depicted in Fig. 20. Note the severe amplification of the maximum resonant amplitude due to mistuning—well over 100 percent in this case. A slight, insignificant frequency shift is observed for the SMART model, particularly in the group of resonances at the lower frequencies. The conformity is otherwise excellent even for this highly complex resonance pattern. REDUCE yields a fair approximation, although its frequency shift is more pronounced.

To consider another excitation case, the mistuned response in veering 2 is depicted in Fig. 21. Clearly, the disk–blade modal interaction is less critical in this veering, as there is only a very modest amplification of the maximum resonant amplitude. This is explained by the more shallow veering, which suggests strong disk–blade coupling beyond critical levels from a mistuning sensitivity standpoint. As expected, a scenario similar to Fig. 20 is seen here in terms of accuracy. Note that even though REDUCE generally matches the natural frequencies better than in the previous case, the predicted resonant amplitudes are significantly worse. From the results presented in this section, it is concluded that the SMART approach can yield excellent accuracy and efficiency even with a large parent finite element model.

6 Conclusions

The primary contribution of this paper is the application of two novel approaches for the reduced order modeling of mistuned bladed disks. These two methods—the mistuning projection method and the secondary modal analysis reduction technique (SMART)—utilize the assumption that tuned and mistuned mode families span the same deformation space. This results in very small reduced order models (ROMs), with matrix dimensions on the order of the number of blades.

In the mistuning projection method, a classical modal analysis is performed on the finite element model of a bladed disk. A coordinate transformation is performed to project the blade mistuning from the finite element domain onto the tuned system modes. In the SMART approach, a primary ROM is generated via component mode synthesis, and then a secondary modal analysis is performed to generate a smaller, secondary ROM. The mistuning data are implemented directly in the blade-component modal coordinates of the primary ROM, and the mistuning is then projected onto the secondary ROM. Using a low-order, modal-domain projection makes the SMART approach highly efficient.

The methods were compared in terms of modal convergence, mistuned mode shape representation, and tuned and mistuned forced response amplitude predictions. It was demonstrated that the mistuning projection and SMART methods outperform current techniques: the new methods exhibit comparable or improved accuracy, while being far superior in computational efficiency. The SMART approach was shown to be exceptionally fast for running simulations of mistuned rotor forced response. Thus, SMART appears to be the most appealing method to date for comprehensive studies of forced response statistics for mistuned bladed disks.

Acknowledgment

This work is supported by the GUIde Consortium on blade durability at Carnegie Mellon University.

References

[1] Srinivasan, A. V., 1997, ''Flutter and Resonant Vibration Characteristics of Engine Blades,'' ASME J. Eng. Gas Turbines Power, **119**, pp. 742–775.
[2] Irretier, H., 1983, ''Spectral Analysis of Mistuned Bladed Disk Assemblies by Component Mode Synthesis,'' *Vibrations of Bladed Disk Assemblies*, ASME, New York, pp. 115–125.
[3] Zheng, Z.-C., and Wang, F.-R., 1985, ''Dynamic Analysis of Blade Groups Using Component Mode Synthesis,'' *Vibrations of Blades and Bladed Disk Assemblies*, ASME, New York, pp. 97–103.
[4] Kruse, M. J., and Pierre, C., 1996, ''Forced Response of Mistuned Bladed Disks Using Reduced-Order Modeling,'' *Proc. 37th AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference*, Vol. 4, AIAA, New York, pp. 1938–1950.
[5] Kruse, M. J., and Pierre, C., 1996, ''Dynamic Response of an Industrial Turbomachinery Rotor,'' *Proc. 32nd AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, AIAA, New York.
[6] Castanier, M. P., Ottarsson, G., and Pierre, C., 1997, ''A Reduced-Order Modeling Technique for Mistuned Bladed Disks,'' ASME J. Vibr. Acoust., **119**, pp. 439–447.
[7] Yang, M.-T., and Griffin, J. H., 1997, ''A Reduced Order Approach for the Vibration of Mistuned Bladed Disk Assemblies,'' ASME J. Eng. Gas Turbines Power, **119**, pp. 161–167.
[8] Yang, M.-T., and Griffin, J. H., 1999, ''A Reduced Order Model of Mistuning Using a Subset of Nominal System Modes,'' *Proc. 44th ASME Gas Turbine and Aeroengine Technical Congress, Exposition and Users Symposium*, ASME, New York.
[9] Bladh, R., Castanier, M. P., and Pierre, C., 1999, ''Reduced Order Modeling and Vibration Analysis of Mistuned Bladed Disk Assemblies with Shrouds,'' ASME J. Eng. Gas Turbines Power, **121**, pp. 515–522.
[10] Bladh, R., Castanier, M. P., and Pierre, C., 2001, ''Component-Mode-Based Reduced Order Modeling Techniques for Mistuned Bladed Disks—Part I: Theoretical Models,'' ASME J. Eng. Gas Turbines Power, **123**, pp. 89–99.
[11] Craig, R. R., and Bampton, M. C. C., 1968, ''Coupling of Substructures for Dynamics Analyses,'' AIAA J., **6**, pp. 1313–1319.
[12] Bladh, R., Castanier, M. P., and Pierre, C., 1998, ''Reduced Order Modeling and Efficient Forced Response Statistics Prediction for Mistuned Bladed Disks,'' *Proc. 3rd National Turbine Engine High Cycle Fatigue Conference*, San Antonio, TX.
[13] Allemang, R. J., and Brown, D. L., 1982, ''A Correlation Coefficient for Modal Vector Analysis,'' *Proc. 1st International Modal Analysis Conference*, Union College, Schenectady, NY, pp. 110–116.
[14] Ottarsson, G., and Pierre, C., 1995, ''On the Effects of Interblade Coupling on the Statistics of Maximum Forced Response Amplitudes in Mistuned Bladed Disks,'' *Proc. 36th AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference*, AIAA, New York, Vol. 5, pp. 3070–3078. Also, J. Sound Vib. (in print).

K. Froelund

J. Schramm

Department of Energy Engineering,
The Technical University of Denmark,
Lyngby, Denmark

T. Tian

V. Wong

S. Hochgreb

Sloan Automotive Laboratory,
Massachusetts Institute of Technology,
Cambridge, MA

Analysis of the Piston Ring/Liner Oil Film Development During Warm-Up for an SI-Engine

A one-dimensional ring-pack lubrication model developed at MIT is applied to simulate the oil film behavior during the warm-up period of a Kohler spark ignition engine. This is done by making assumptions for the evolution of the oil temperatures during warm-up and that the oil control ring during downstrokes is fully flooded. The ring-pack lubrication model includes features such as three different lubrication regimes, i.e., pure hydrodynamic lubrication, boundary lubrication and pure asperity contact, nonsteady wetting of both inlet and outlet of the piston ring, capability to use all ring face profiles that can be approximated by piece-wise polynomials, and, finally, the ability to model the rheology of multigrade oils. Not surprisingly, the simulations show that by far the most important parameter is the temperature dependence of the oil viscosity. [DOI: 10.1115/1.1341206]

Introduction

Much effort has been made in the past to describe the lubrication of the interface between the piston rings and the cylinder liner. Early work in this field focused on improving the friction and wear characteristics of these parts, thereby giving better fuel economy and better durability [1]. More recently, new interest arose amid strong concerns over the influence of the piston ring pack on oil consumption and various HC-emissions related mechanisms. A review of the current status can be found in Ref. [2].

The description of the lubricating process is very complex and therefore the attempt to model piston ring-pack lubrication has been hampered by the need to make assumptions to enable the problem to be mathematically formulated. This is because real phenomena exist beyond current modeling capabilities—such as circumferential variations, secondary piston motions, and particles and degradation of the oil—just to mention some of the real world complications.

Previous ring pack lubrication simulations have been exercised under ''cyclic steady-state'' conditions. This paper introduces a ''cyclic quasi steady-state'' model for the ring pack lubrication during warm-up, where variations in parameters such as cylinder liner temperature are slowly varied, with the characteristic time much larger than the duration of an engine cycle.

The Ring-Pack Lubrication Model

The MIT ring pack model is described here briefly. A full presentation can be found in Ref. [3]. The model is one dimensional and delivers, therefore, no information for the circumferential variation in the film thickness caused by, for example, ring nonconformability.

Three lubrication regimes are considered; pure hydrodynamic lubrication (when the actual oil film thickness is bigger than four times the combined surface roughness), mixed lubrication (when the actual oil film thickness is between the combined roughness and four times the combined roughness), and pure boundary friction if there is not sufficient oil under the ring to provide hydrodynamic lubrication. A discussion of the combined roughness can be found in Appendix 1.

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, 23 April 1997; final revision received by the ASME Headquarters, 20 October 2000. Associate Editor: D. Assanis.

In the pure hydrodynamic regime, there is no asperity contact and the hydrodynamic pressure force under the ring face is in equilibrium with the gas pressure force and the ring elastic force. For the mixed lubrication, the contact force of the asperities is added to the radial force balance on the ring. Finally, for the pure boundary lubrication, ring load is solely supported by asperity contact.

For each ring there are four governing equations:

- the integrated average flow Reynolds equation;
- the nonsteady continuity applied to the leading wetted edge of the ring;
- the nonsteady continuity applied to the trailing wetted edge of the ring; and
- the radial force balance.

For each time step, the system equations are solved simultaneously for the four unknowns: the minimum average clearance between the ring and the liner, the leading wetted edge position on the ring face profile, the trailing wetted edge position on the ring face profile, and the oil flow rate at the outlet wetted position on the ring face profile.

The system equations can be used for solving the lubrication of a whole ring pack consisting of a compression ring, a scraper ring, and an oil control ring by applying an additional conservation equation for the oil flow between the rings; i.e., the amount of oil that leaves one ring is the same as that entering the adjacent ring.

It is assumed that during the downstroke the oil control ring is always fully flooded. With this knowledge, the minimum oil film thickness for the oil control ring can be found for the down stroke. Thereafter, solutions for the scraper ring and for the compression ring can be derived.

Assumptions for Using the Model for the Warm-up Period

The following assumptions are made for calculating the development of the oil film during the warm-up period:

- the model developed for the ''cyclic steady state'' can be used for predicting ''cyclic quasi steady state;''
- the speed and load of the engine are constant;
- the oil supply to the ring/liner begins without considerable time lag after start of the engine;
- the oil supply is sufficient such that the oil control ring is fully flooded on downstrokes [3];
- the oil temperature changes linearly in time [4–6];

- the oil temperature dependence on liner position follows the square root expression developed by Woschni at all times [7]; and
- oil viscosity is only a function of temperature and shear rate.

Quasisteadiness. Quasisteady conditions should prevail because the time scale for heating up the engine is much larger than the time scale for the lubrication process to adjust to changes. The cycle-to-cycle variations because of rising temperature are very low.

Constant Speed and Load. Although not true during the first few cranking cycles, the assumption is reasonable through most of the warm-up period. With this assumption, usage of the model for strong cycle to cycle variations of parameters is prevented.

Oil Supply. The fulfillment of the two assumptions about the oil supply depends on three factors: the choice of oil, the starting temperature, and the clearances in the bearings of the oil supply system. For a very low start temperature, the oil pump can have problems in pumping the oil, which clearly inhibits the oil supply. Also, with an improper choice of oil, considerable time lag can occur for the oil supply to reach engine components, such as the camshaft, during warm-up. Schwarze showed that the supply of oil to the various engine components can be simulated by simplifying the system as a network system of flow resistances, connected in series and parallel [8]. His model shows that excessive bearing clearance in one or more crankshaft bearings, either due to normal wear or bearing damage, can change the oil supply to the various engine components considerably.

The recommended oil for the Kohler engine is the SAE 10W30 multi-grade oil. The oil properties have been measured by Shell Research, Ltd. and are valid down to a start temperature of 20°C [9]. The warm-up simulations are therefore performed for a start temperature of 20°C.

Because of the choices of the start temperature and oil viscosity grade, the first oil supply assumption is expected to be valid. The evaluation of the second oil supply assumption is very difficult. From measurements or from the above mentioned network analysis for the oil supply system, the oil supply from the connecting rod bearings to the cylinder liner could be known. However, the processes of oil transport from the connecting rod to the piston skirt and from there to the cylinder liner have never been well understood. Until now, the underlying design philosophy has been to supply the ring/liner with excess oil to prevent oil starvation under all driving conditions. If this is true, the second oil supply assumption will also hold.

Oil Temperature. An important variable in the warm-up model is the local instantaneous temperature on the liner, since oil viscosity is a very strong function of temperature.

Very often the temperature distribution along the cylinder liner is assumed linear. However, thermocouple measurements on running engines suggest that it would be more appropriate to choose a spatial temperature profile that has a higher slope near the top dead center and a lower slope near the bottom dead center [7]. Here the relation developed by Woschni is used, where it is assumed that if the temperature in the dead centers is known as a function of time, then the spatial distribution at all times is given by

$$T(x) \approx T_{\mathrm{tdc}} - (T_{\mathrm{tdc}} - T_{\mathrm{bdc}}) \sqrt{\frac{x}{s}}, \tag{1}$$

where x is the distance from the top dead center of the compression ring, s is the stroke, and T_{tdc} and T_{bdc} are the liner temperatures for the top dead center and bottom dead center of the compression ring, respectively. The above equation is very practical, since experimental engines very often have thermocouples mounted in the liner at the dead-center positions of the compression ring. It should be noted here that x can be greater than s for the oil control ring and the scraper ring near bottom dead center, and the temperature will consequently be lower than the temperature at the bottom dead center for the compression ring.

The liner temperature as a function of time can be measured or simulated. Experiments have been performed on test engines with thermocouples mounted in the liner at the compression ring dead centers, but very often the cooling system deviated substantially from the one under practical driving conditions. Andrews and Trapy et al. measured the heating of the engine block under realistic circumstances and found that the temperature dependence on time is linear up to a point [4,5]. This statement is further supported by heat transfer network simulations made by Kaplan, where the engine components are assumed to be modeled by lumped heat capacitors connected by heat resistors [6]. The characteristic warm-up times are different for different engine components and vary with speed and load. It is assumed that the time constant t_{ss} for warming up the Kohler cylinder liner from 20°C to 100°C (at mid stroke position) is 3 min. This is in accordance with Kaplan, where a typical four cylinder production engine was simulated at constant speed and load (1500 rpm and 260 kPa bmep) during the warm-up period [6]. The temperatures of the dead centers are thus approximated as

$$T_{\mathrm{tdc}}(t) = (T_{\mathrm{tdc,ss}} - T_{\mathrm{tdc,0}}) \frac{t}{t_{ss}} + T_{\mathrm{tdc,0}}, \tag{2a}$$

$$T_{\mathrm{bdc}}(t) = (T_{\mathrm{bdc,ss}} - T_{\mathrm{bdc,0}}) \frac{t}{t_{\mathrm{ss}}} + T_{\mathrm{bdc,0}} \forall t \in [0; t_{\mathrm{ss}}], \tag{2b}$$

where $T_{\mathrm{tdc,ss}}$ and $T_{\mathrm{bdc,ss}}$ are the steady-state temperatures and $T_{\mathrm{tdc,0}} = T_{\mathrm{bdc,0}} = T_0$, i.e., the start temperature.

Implementing (2*a*) and (2*b*) into (1), one obtains the expression for the temperature of the cylinder liner as a function of time and liner position

$$T(x,t) = \left[\left(1 - \sqrt{\frac{x}{s}} \right) T_{\mathrm{tdc,ss}} + \sqrt{\frac{x}{s}} T_{\mathrm{bdc,ss}} - T_0 \right] \frac{t}{t_{\mathrm{ss}}} + T_0. \tag{3}$$

This is the liner temperature profile used in this work. It is further assumed that this cylinder liner temperature is equal to the temperature of the lubricating oil between the piston rings and the cylinder liner. The time duration for reaching steady-state, i.e., $t_{\mathrm{ss}} = 3$ min, are divided into seven time steps of 30 s each.

A plot of the applied cylinder liner temperature distribution for the simulation is shown in Fig. 1, where $T_0 = 20°\mathrm{C}$, $T_{\mathrm{tdc,ss}} = 120°\mathrm{C}$ and $T_{\mathrm{bdc,ss}} = 90°\mathrm{C}$. It can be seen in this figure that the relative position on the cylinder liner varies from 0, the TDC of the compression ring, to approximately 1.14, the BDC of the oil control ring. The group of curves shown on the plot are for different time duration after the start of the engine. The benefit of using the square root expression for the spatial temperature distribution instead of the commonly used linear approximation is limited here, but for larger diesel engines, the temperature difference between the dead centers can be substantial, and the deviation from the linear profile is more pronounced.

Oil Viscosity. The oil viscosity is generally dependent on the following factors in decreasing order: oil temperature, oil shear rate, oil pressure, oil degradation, and concentration of dissolved fuel in oil. The influence of the oil pressure tends to increase the viscosity by maximum 30 percent, whereas typical values of dissolved fuel concentrations in the oil on 1 percent, during the warm-up, tend to reduce the viscosity by approximately 20 percent [10]. The oil degradation tends to increase the viscosity and can be very significant [11]. Only the first two factors are taken into account in this work.

Engine details. The ring-pack model is applied to a Kohler single cylinder spark ignited CH-14 engine. The engine details are shown in Table 1. The engine has been used at MIT for extensive oil film thickness measurements using the laser induced fluores-

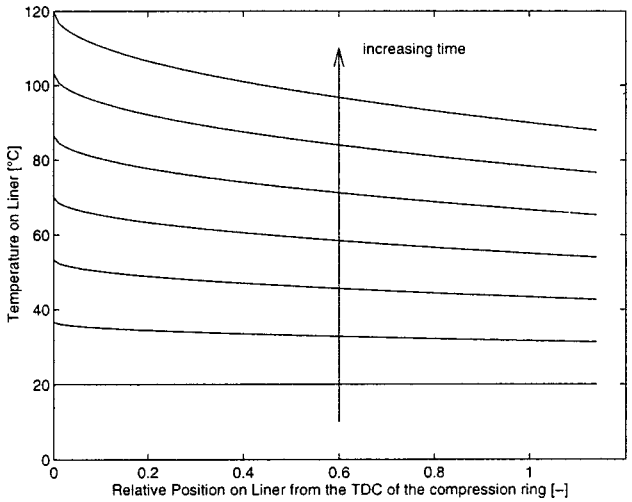


Fig. 1 Liner temperature as a function of liner relative position and time. The time steps are 0, 30, 60, 90, 120, 150, and 180 s.

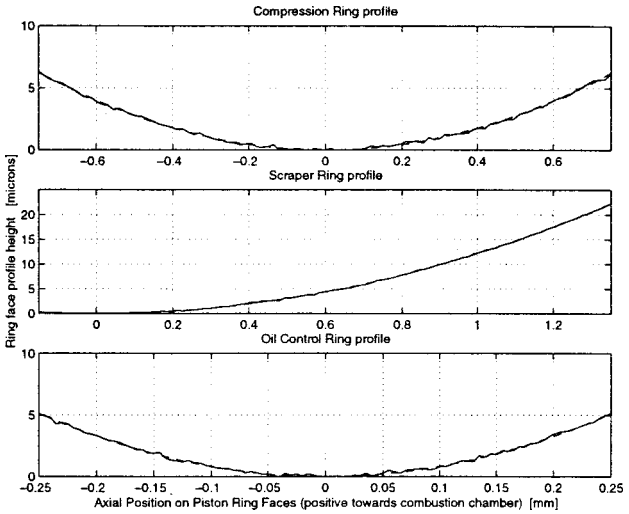


Fig. 2 Ring face profiles with a 0.10 μm RMS Gaussian roughness

cence technique [12]. For this engine, good agreement have been obtained between model predictions for various oils and operating conditions [3].

Simulation is done for a constant test condition, which is assumed to be set after the engine is started. The cylinder pressure is measured and the pressure between the compression and scraper ring is calculated using a new gas flow model at MIT [13].

The dimensional profiles of the piston rings are shown in Fig. 2. Measured worn ring profiles have been fitted to appropriate polynomial approximations. It has been found that for the compression ring a single parabolic fit without offset and for the scraper ring a parabolic fit with strong offset towards the oil sump direction were appropriate. The oil control ring was approximated by a single parabola without offset. The RMS-roughness of the rings is measured to be 0.10 μm after breakin. Therefore, the dashed curve shows the mean profile and the solid line the simulated actual profile. The selection of the rings is representative of SI engines, in general.

Oil Rheology. The oil rheology of oils of several SAE grades were compared. Three oils were selected for comparison: SAE 10W30, SAE 10W50, and SAE 30. An SAE 10W30 oil is the recommended oil for this engine. An SAE 10W50 oil is chosen because this oil has very high shear thinning properties. The SAE 30 is a monograde oil and is chosen because of its strong increase in viscosity with decreasing temperature.

The temperature dependence of the viscosity at low shear rate has been measured and fitted to the Vogel equation given by

$$\mu_0(T) = \kappa \cdot \exp\left(\frac{\theta_1}{\theta_2 + T}\right). \tag{4}$$

Figure 3 shows a plot of the low shear kinematic viscosity of the three selected oils. It can be seen that the oils show different low shear kinematic viscosity properties.

The shear rate dependence of the viscosity can be described by the Cross equation:

$$\mu(\dot{\gamma}) = \frac{\mu_0 + \mu_\infty(\dot{\gamma}/\beta)^m}{1 + (\dot{\gamma}/\beta)^m}, \tag{5}$$

where μ_0 is the low shear rate viscosity, μ_∞ is the high shear rate viscosity, $\dot{\gamma}$ the average shear velocity, β the critical shear velocity, and m is a parameter, specifying the width of the non-Newtonian region. It is assumed that $m=1$ for all oils [9]. The critical shear rate β is given by $\beta(T)=10^{a+bT}$. For all oils the same parameters for a and b are used, i.e., $a=2.43$ and $b=0.0218°C^{-1}$ [9]. In Table 2 the necessary temperature and shear velocity properties of the three chosen oils are shown.

Figure 4 shows the calculated average shear rate and critical shear rate at the minimum oil film thickness location for the baseline SAE 10W30 oil at 2500 rpm and 2/3 load. For the warm and cold engines, the critical shear rates are about $5\times10^4\ s^{-1}$ and 500 s^{-1}, respectively. The calculated average shear rate is, for most of the engine cycle, about one order of magnitude above these values, i.e., ranging from $10^6\ s^{-1}$ to $3\times10^7\ s^{-1}$. Therefore the oil behaves for all rings as if it has no dependence on the shear velocity for the major part of the engine cycle, and the viscosity characterizing the piston ring and linear lubrication is the one for

Table 1 Engine details

Engine	Kohler single cylinder spark ignited
Engine block	Cast aluminum with cast iron cylinder liner
Bore × Stroke	87.0 mm × 67.0 mm
Distance between: compression ring/scraper ring scraper ring/oil control ring	 4.1 mm 5.2 mm
Displacement	398 cm^3
Maximum power	10.5 kW @ 3600 rpm
Maximum torque	28.9 Nm @ 2600 rpm
Test condition	2/3 Load @ 2500 rpm

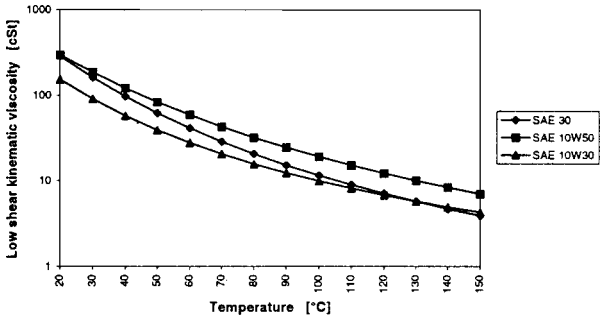


Fig. 3 Low shear rate viscosity of the three lubricants

Table 2 Properties of the three chosen oils

Oil type	SAE 10W30	SAE 10W50	SAE 30
ρ [kg/m^3]	880	880	888
κ [cSt]	0.1403	0.0352	0.0246
θ_1 [°C]	869.7	1658.9	1432.3
θ_2 [°C]	104.4	163.5	132.9
μ_∞ / μ_0 [-]	0.76	0.49	1.00

high shear rates. Only in the vicinity of the dead centers, the calculated average shear rate is less than the critical shear rate. At these positions the oil have the higher low shear dynamic viscosity, which enhances the squeezing motion of the piston ring.

Figure 5 shows the comparison of the high shear viscosity dependence on the temperature given in Table 2. All three oils have all approximately the same high shear rate viscosity at 150°C. From an engineering point of view, the three oils can therefore be chosen equally. It can further be seen that the order of the oil viscosities has changed so that the SAE 10W30 is the oil with lowest viscosity, the SAE 10W50 the intermediate, and SAE 30 the oil with highest viscosity.

Figure 6 shows the viscosity as a function of liner position and time for the baseline SAE 10W30 oil. A high viscosity has to be overcome during the warm-up, with high shear rate kinematic viscosity for the SAE 10W30 oil up to 200 cSt, which is approximately 40 times the warmed up value. At steady-state, the ratio of the viscosity at the dead centers is a factor of 2.

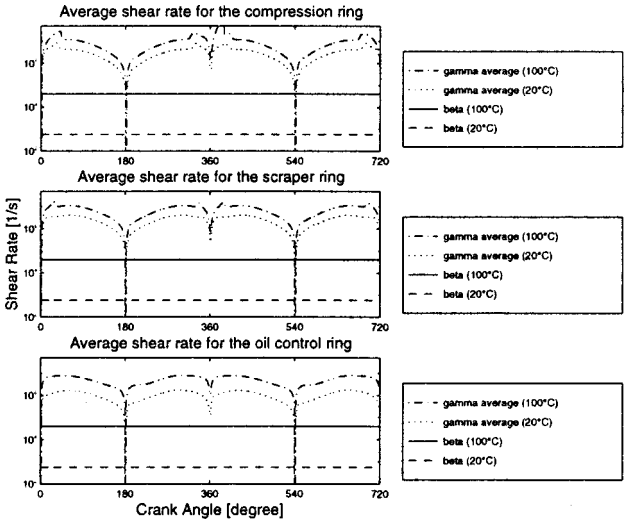


Fig. 4 Calculated average shear rate ($\dot{\gamma}$) and critical shear rate (β) for the three rings for cold (20°C) and warm (100°) conditions, respectively

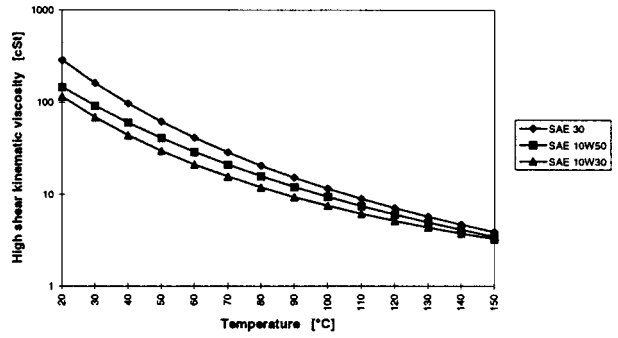


Fig. 5 High shear rate viscosity of the three lubricants

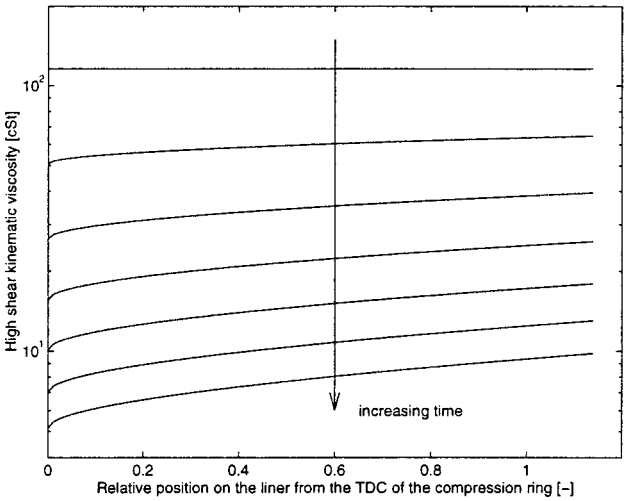


Fig. 6 High shear kinematic viscosity for the ring/liner interface during warm-up

Simulation Results

Figure 7 shows the results of the simulations for the SAE 10W30 oil. Similar calculations have been generated for the other oils, although these will not be shown here. Every figure is divided into subfigures denoted by (a)–(k). For Figs. 7(a)–7(j), the horizontal-axis is in crank angle degrees, where the origin is chosen to be the TDC of the expansion stroke, whereas for Fig. 7(k) the horizontal-axis is the liner position from the TDC of the compression ring.

In every sub-figure, seven curves are shown, one for every time step, i.e., for 0, 30, 60, 90, 120, 150, and 180 s after the start of the engine.

In the following section, the results for the different oils will be discussed.

Results for the SAE 10W30 Oil

Figures 7(a)–7(c) show the simulated oil film thickness for the whole ring pack. Two additional horizontal lines are drawn. The dashed line shows the place where the oil film thickness is equal to the combined roughness and the dotted line the place where the oil film thickness equals four times the combined roughness. The figures of the minimum oil film thickness, Figs. 7(a)–7(c), are thus used to provide information on how critical the lubrication is at the dead centers, and therefore must be compared with the contact pressures shown in Figs. 7(d)–7(f).

The behavior of the film thickness during the warm-up can be seen on Fig. 7(a) for the compression ring. As the engine heats up and the viscosity of the oil decreases, so does the minimum oil film thickness under the compression ring. This trend is clearly seen at midstroke of the ring, where the minimum oil film thickness typically falls by a factor of 2.5. At TDC the trend is not so clear. One reason is that even for the later time steps, the asperity contact pressure rises to levels comparable to the gas pressure, thereby preventing further ring penetration.

A comparison between Figs. 7(a)–7(f) shows that the compression ring for this ring-pack experiences the most contact with the liner at positions shortly after TDC. The reasons for this are the following: this ring is the most starved one, being the last one receiving oil supply; it also has the highest gas pressure load and it operates in the uppermost position of the ring pack, and is thereby lubricated by the thinnest oil.

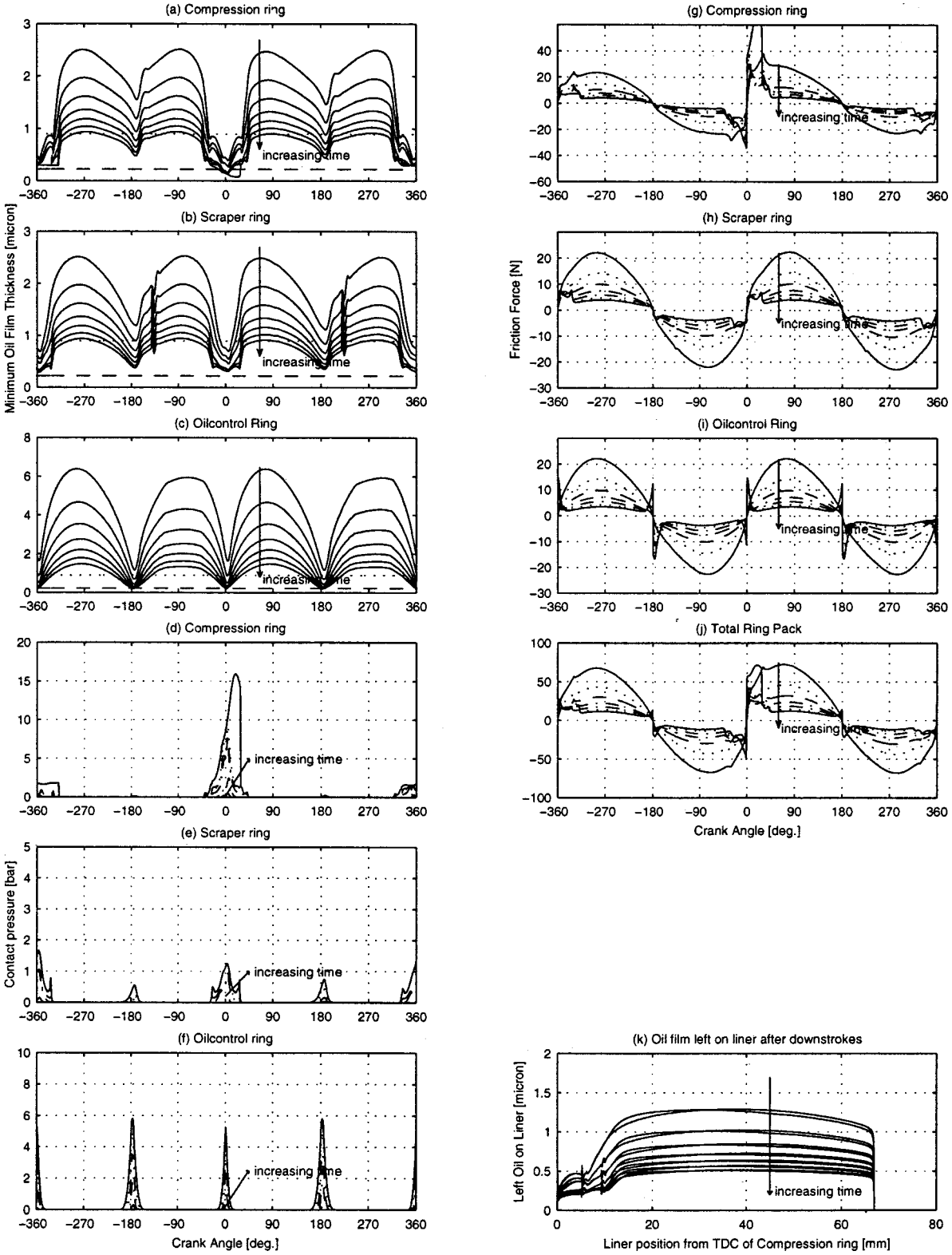


Fig. 7 Simulation results for the SAE 10W30 oil

The position where the cyclic minimum oil film thickness occurs seems to be about 5° ATDC, irrespective of the temperature, as long as the ring does not penetrate the oil film totally, as is the case in Fig. 7(*a*) for $t = 180$ s. This is important, because it can be related to the position on the liner for a given ring, where the maximum wear occurs. Here it shall be mentioned that for predicting the wear caused by the ring pack, the liner position corresponding to each ring in TDC is different (Table 1), so it has been suggested that one calculates the total wear on a liner position by summing up the wear contributions from each ring [14].

The plots of the oil film thickness for the scraper ring Fig. 7(*b*) show a transient in the solution for the minimum oil film thickness

shortly after 140° BTDC. The transition in the scraper ring profile can be calculated to occur at the position on the cylinder liner where the compression ring is at BDC.

The contact force Figs. 7(*d*)–7(*f*) show that the asperity contact starts to show nonzero values exactly when the corresponding transition between the pure hydrodynamic lubrication regime and the mixed lubrication regime is crossed in Figs. 7(*a*)–7(*c*). These contact pressures grow very fast with the penetration depth into the roughness profiles, because the used Greenwood and Tripp model for simulating the asperity force rises with the power of 6.8 of the penetration depth into the roughness profiles, which soon prevent further ring penetration [15,16]. Finally, the results for the contact force of the scraper ring Fig. 7(*e*), and for the oil control ring Fig. 7(*f*), show that these rings for the warmed-up conditions always exhibits some asperity contact at dead centers.

A comparison of Figs. 7(*a*)–7(*c*) shows that the shape of the oil film thickness over a stroke has a certain degree of nonsymmetry to the dead center positions for the compression and the scraper ring, whereas the oil control ring is close to symmetric to the dead center position. Several reasons contribute to this:

- the kinematic viscosity varies with liner position (Fig. 6);
- the piston speed is nonsymmetrical and is consequently not strictly harmonic but skew to the TDC side [17];
- the gas pressure loads varies (compression and scraper ring);
- the ring face may be nonsymmetric caused by parabola offsets (scraper ring, Fig. 2); and
- the different wetting conditions [3].

The oil control ring has no gas pressure load, the ring face is symmetric, and the wetting is usually fully flooded. The viscosity effect should tend to make the oil film thickness smaller near TDC and the piston velocity effect have the opposite effect.

The scaling of the oil film thickness with viscosity is in lubrication theory given by $\delta_{\text{oil}} \propto \sqrt{\mu N}$ for constant sliding speed and for a static loaded bearing. As the oil control ring is statically loaded, this is the only ring where the scaling can be expected to hold true. A comparison between the curves for $t=0$ sec. and for $t=180$ sec. in Fig. 7(*c*) at $+90°$ is therefore expected to follow the relation

$$\frac{\delta_{\text{oil},t=0}}{\delta_{\text{oil},t=180}} \approx \sqrt{\frac{\mu_{t=0}}{\mu_{t=180}}}.$$

The left-hand side gives 3.95 and the right-hand side approx. 4.4, where the viscosity is found in Fig. 6 for a relative position of 0.64 (half the stroke plus the distance between the compression ring and the oil control ring). This verifies that the lubrication theory holds and the small deviation is caused by varying sliding speed and different wetting conditions. The scaling of the compression and scraper rings does not follow this simple expression.

The friction plots of the single rings on Figs. 7(*g*)–7(*i*) show a twofold trend with temperature: when the viscosity level is high, i.e., at the engine start, the viscous force is high at midstroke; however, at the same time the boundary friction force near the dead centers is very low. The reverse is true for low viscosity levels at higher temperatures.

The friction of the whole ring pack is shown in Fig. 7(*j*). By averaging the numerical value of this friction force, the friction mean effective pressure of the ring pack can be found. This is discussed later in the paper. The friction profiles can easily be converted into power loss by multiplying the friction by the piston speed. Here it shall be mentioned that friction force at midstroke contributes most to the power loss.

Finally, the last plot, Fig. 7(*k*), shows the oil left on the cylinder liner after each of the two down strokes in one engine cycle. This oil is left after the compression ring passage in the downstrokes. For the parts of the stroke where the squeezing motion of the ring is negligible, the thickness of the oil layer left on the liner is approximately half of the minimum oil film thickness under the compression ring. The shape of the oil layer left on the cylinder liner is poorly examined in the literature, but the reasons for examining this layer are obvious:

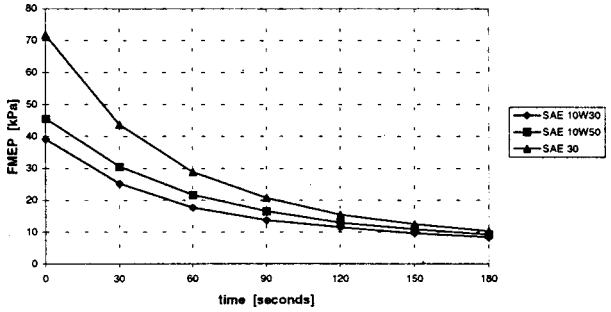


Fig. 8 Average friction mean effective pressure for the ring-pack during the warm-up

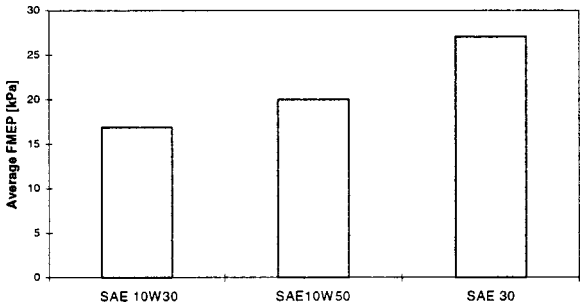


Fig. 9 Cycle averaged ring-pack friction mean effective pressure during the warm-up

- this is the layer in which the absorption and desorption process of gaseous hydrocarbons takes place;
- it is in this layer that diffusion of sulfur takes place in diesel engines for marine applications, thereby causing corrosion of the cylinder liner [18];
- it is this layer causing a significant part of the evaporative oil consumption [19].

Here the influence of the liner oil layer on the absorption/desorption processes is subsequently discussed in Appendix 2.

Average Friction Calculations

The total friction of the ring pack is averaged over a whole cycle in Fig. 8. Therefore, the friction of the ring pack can be studied as a function of time. The ring-pack FMEP for the warm condition is 9 kPa for selection of the standard SAE 10W30 oil. The friction of the ring-pack is shown to increase about 4 to 5 times, when the engine is at 20°C. The SAE 10W50 oil shows a bigger increase, whereas the ratio for the SAE 30 oil is about a factor of 8. This is the reason for the ongoing trend of choosing the low friction oils, such as SAE 10W30. In Fig. 9, the FMEP, calculated in Fig. 8, is averaged over the whole warm-up phase. Here a relative estimate can be seen for the total amount of work and thereby the total fuel consumption, assuming constant efficiency of the engine, for overcoming the whole warm-up phase. The SAE 10W50 and SAE 30 oils result in an increase of the average FMEP by 20 percent and 60 percent, respectively.

Conclusions

This is the first attempt to describe the oil film development during the warm-up period for SI engines. A new temperature profile of the cylinder liner has been proposed to make the calculations possible. Interesting findings include:

• For the piston ring lubrication, only the high shear viscosity is of importance. Multigrade oils effectively exhibit full shear thinning behavior for this application.

• For all temperatures between 20°C and the warm condition, all rings change lubrication regime near TDC to the mixed lubrication. Thereby, asperity contact occurs during the whole warm-up phase.

• The minimum oil film thickness between the oil control ring and cylinder liner scales fairly well with the square root of the viscosity. The thickness between other rings in the ring-pack does not scale with viscosity in any simple manner. While the thickness of the oil left on liner is coupled with the minimum oil film thickness of the compression ring, no simple scaling is found.

• The cycle averaged ring-pack FMEP increases four to five times at cold conditions (20°C) compared to the warm condition (100°C) for the baseline SAE 10W30 oil. By averaging the above-mentioned FMEP over the whole warm-up phase, the average warm-up FMEP (20°C→100°C) is twice the warm FMEP (100°).

• The oil film left on the liner is important for modeling the absorption/desorption mechanism of fuel hydrocarbons in the oil film. The thickness of this oil layer is predicted to be smaller than previously calculated and to be in the order of a half to two microns, and vary dependent on the oil temperature. Furthermore, the oil layer left on the liner is found to have a small, but still significant, thickness in the region not overrun by the oil control ring, i.e., the distance between TDC of the compression ring and TDC of the oil control ring. The thickness is on the order of 0.2–0.5 μm, depending on the oil temperature. This finding is important for a further study of the warm-up absorption/desorption process to estimate the contribution from this source to the engine-out unburned hydrocarbons.

The model of the oil film behavior during the warm-up has yet to be verified by experiments, and will be the topic of subsequent work.

Acknowledgments

Kent Froelund's exchange study at MIT has been sponsored by the Technical University of Denmark. The authors would like to thank the MIT Lubrication Consortium for providing the ring-pack lubrication simulation program and other resources in this study. In particular, the authors would like to thank Dr. Ian Taylor at Shell for providing the oil data needed for this paper. The authors would also like to thank Professor John Heywood at MIT and Jan-Roger Linna at Volvo Car Corp. for the helpful discussions.

Appendix 1: Combined Roughness

Measurements using the traditional Talysurf profile of the ring and the liner shows strong non-Gaussian roughness profiles. This is because it contains some deep honing grooves that acts as a reservoir for oil for the cold start or for other critical lubricating situations. The honing grooves have typical widths on the order of 0.1 mm and typical depths on the order of 5 μm. There are about one hundred of these in the circumferential direction, which can be seen on microscopic pictures. A simple calculation shows that the grooves stand for approximately 10 mm in the circumferential direction and the rest of the surface (approximately 300 mm) contains no grooves. The argument can be made that because of the sparse distribution of the honing grooves, they can simply be subtracted from the traditional Talysurf profile [20]. Thereby two goals are reached:

• The lubrication model is designed for Gaussian roughness profiles. The rest of the roughness profile after subtracting the deep honing grooves shows a much more Gaussian roughness profile.

• The RMS roughness becomes far lower, while the real liner shows a top Gaussian distributed flat surface (therefore the name ''plateau honing'') with a very smooth surface containing some deep grooves that tend to increase RMS.

The surface roughness of the liner is then reduced to 0.20 μm. With a ring surface roughness on 0.10 μm the combined surface roughness is calculated to be

$$\sigma = \sqrt{\sigma_{\text{liner}}^2 + \sigma_{\text{ring}}^2} = 0.22\ \mu\text{m}.$$

Appendix 2: Absorption/Desorption of HC in the Liner Oil Film

Early work on the influence of the liner film thickness on the absorption/desorption process has been done by using a constant oil film thickness on the liner [21–23]. Later studies have considered the effects of cold liner temperatures on the absorption/desorption process, still using a spatially constant liner oil film thickness [24,25]. A recent study by Min [26] takes into account the spatial distribution of the oil film thickness.

These studies suggest that the absorption/desorption process can generally be divided into two limiting regimes, dependent on the wall temperature [24]:

• A thick layer regime: for low wall temperatures, the absorption is diffusion controlled because the penetration depth of the fuel hydrocarbons into the oil film is smaller than the actual oil film thickness.

• A thin layer regime: for high wall temperatures, the absorption is controlled by the oil film thickness, i.e., scales with viscosity, because the penetration depth now is larger than the actual oil film thickness. Here the total amount of oil left on the liner is important for the absorption.

For all temperatures it is found that during one engine revolution there is enough time for the full desorption of all fuel hydrocarbons stored in the oil film.

The work presented here suggests that for the top liner region on figure (*k*), i.e., the region from the top dead center of the compression ring to the top dead center of the oil control ring (the first 10 mm on the liner), the oil film thickness is very small, ranging from 0.5 to 0.20 μm, depending on the liner temperature. This is clearly smaller than previous studies found by using the average oil film thickness over the whole engine stroke [21–25]. The importance of this finding is significant because calculations made by Min show that the upper 15 percent of the cylinder liner should be responsible for almost all the desorbed HC emissions. This study shows that the upper part of the liner oil layer can be assumed to be in the ''thin film'' regime at all times. Hence, the significance of the absorption/desorption process might be less pronounced than previous estimated.

Nomenclature

$T(x)$ = liner temperature at liner position x from the TDC of the compression ring

$T_{\text{tdc}}, T_{\text{bdc}}$ = liner temperature at TDC and BDC of the compression ring

s = engine stroke

ρ = density of oil

$\kappa, \theta_1, \theta_2$ = parameters in the Vogel equation

μ_0 = low shear dynamic viscosity

μ_∞ = high shear dynamic viscosity

β = critical shear velocity

$\dot{\bar{\gamma}}$ = average shear velocity

δ_{oil} = oil film thickness

N = engine rpm

σ_{liner} = surface roughness of the liner

σ_{ring} = surface roughness of the ring

D = binary diffusion coefficient of fuel in the oil

τ = characteristic diffusion time

References

[1] Ruddy, B. L., Dowson, D., and Economou, P. N., 1982, ''A Review of Studies of Piston Ring Lubrication,'' *Proceedings of the 9th Leeds-Lyon Symposium on Tribology*, Elsevier, New York.

[2] Mierbach, A., et al., 1995, ''Piston Ring Performance Modeling,'' *T&N Symposium*, Paper 15, Wuertzburg–Indianapolis.

[3] Tian, T., 1996, ''A Piston Ring-Pack Film Thickness and Friction Model for Multigrade Oils and Rough Surfaces,'' submitted to the 1996 Fall SAE Fuels and Lubricants Meeting.

[4] Andrews, G. E., Harris, J. R., and Ounzain, A., ''SI Engine Warm-Up: Water and Lubricating Oil Temperatures Influences,'' SAE Paper 892103.

[5] Trapy, J. D., and Damiral, P., ''An Investigation of Lubricating System Warm-Up for the Improvement of Cold Start Efficiency and Emissions of S.I. Automotive Engines,'' SAE Paper 902089.

[6] Kaplan, J. A., 1990, ''Modeling The Spark Ignition Engine Warm-Up Process to Predict Component Temperatures,'' M.Sc. thesis, Sloan Automotive Laboratory, MIT.

[7] Woschni, G., and Zeilinger, K., ''Vorausberechnung des Kolbenringverhaltens,'' FVV Workshop, ''Tribosystem Kolben-Kolbenring-Zylinderlaufflaeche,'' 10. Okt. 1989, VDMA-Haus Frankfurt.

[8] Schwarze, H., 1993, ''Die Auslegung von Oelversorgungssystemen in Verbrennungsmotoren,'' Tribologie+Schmiertechnik 40, Jahrgang 5/1993.

[9] Taylor, R. I., 1995, ''Piston Assembly Friction and Wear: The Influence of Lubricant Viscometry,'' *Proceedings of the International Tribology Conference*, Yokohama.

[10] Department of Energy Engineering, Technical University of Denmark, 1990, unpublished results.

[11] Taylor, R. I., and Bell, J. C., 1994, ''The Influence of Lubricant Degradation on Friction in the Piston Ring Pack,'' *Proceedings of the 20th Leeds-Lyon Symposium on Tribology*, Elsevier, New York.

[12] Tamai, G., 1995, ''Experimental Study of Engine Oil Film Thickness Dependence on Liner Location, Oil Properties and Operating Conditions,'' M.Sc. thesis, Sloan Automotive Laboratory, MIT, Cambridge, MA.

[13] Tian, T., 1996, ''Modeling Piston-Ring Dynamics, Blowby and Ring-Twist Effects,'' submitted to the Fall ASME ICE Conference, October.

[14] Ting, L. L., and Mayer, J. E., Jr., 1974, ''Piston Ring Lubrication and Cylinder Bore Wear Analysis, Part 1-Theory and Part 2-Theory Verification,'' Lubrication Technology, Transaction of the ASME, April.

[15] Greenwood, J. A., and Tripp, J. H., 1971, ''The Contact of Two Nominally Flat Surfaces,'' *Proceedings of the Institute of Mechanical Engineers*, Vol. 185, p. 625.

[16] Hu, Y., Cheng, H. S., Arai, T., and Kobayashi, Y., 1993, ''Numerical Simulation of Piston Ring in Mixed Lubrication-A Nonaxisymmetrical Analysis,'' ASME J. Tribol., **115**, pp. 1–9.

[17] Heywood, J. B., 1988, *Internal Combustion Engines*, McGraw-Hill, New York.

[18] Schramm, J., Henningsen, S., and Sorenson, S. C., ''Modelling of Corrosion of Cylinder Liner in Diesel Engines by Sulfur in the Diesel Fuel,'' SAE Paper 940818.

[19] Wahiduzzanan, S., Keribar, R., Dursunkaya, Z., and Kelly, F. A., ''A Model for Evaporative Consumption of Lubricating Oil in Reciprocating Engines,'' SAE Paper 922202.

[20] Radcliffe, C. D., and Dowson, D., 1995, ''Analysis of Friction in a Modern Automotive Piston Ring Pack,'' *30th Lubrication and Tribology Meeting*, Elsevier, New York.

[21] Korematsu, K., 1990, ''Effects of Fuel Absorbed in Oil Film on Unburnt Hydrocarbon Emissions from Spark Ignition Engines,'' JSME Int. J., Series 2, **33**, No. 3, pp. 362–368.

[22] Schramm, J., ''Smoereoliens indflydelse paa paa kulbrinteemissioner fra benzinmotorer,'' Ph.D. thesis, Laboratoriet for Energiteknik, Technical University of Denmark.

[23] Shyy, W., and Adamson, T. C., Jr., 1983, ''Analysis of Hydrocarbon Emissions From Conventional Spark-Ignition Engines,'' Combust. Sci. Technol., **33**, pp. 245–260.

[24] Norris, M., and Hochgreb, S., ''Novel Experiment on In-Cylinder Desorption of Fuel From the Oil Layer,'' SAE Paper 941963.

[25] Linna, J. R., and Hochgreb, S., 1995, ''Analytical Scaling Model for Hydrocarbon Emissions From Fuel Absorption in Oil Layers in Spark Ignition Engines,'' Combust. Sci. Technol., **109**, pp. 205–226.

[26] Min, K., and Cheng, W. K., 1996, ''Oil Layer as a Source of Hydrocarbon Emissions in SI Engine,'' submitted to the Fall ASME ICE Conference, October.

Simulation of a Single Cylinder Diesel Engine Under Cold Start Conditions Using Simulink

H.-Q. Liu
Graduate Research Assistant

N. G. Chalhoub
Associate Professor

N. Henein
Professor

Mechanical Engineering Department,
Wayne State University,
Detroit, MI 48202

A nonlinear dynamic model is developed in this study to simulate the overall performance of a naturally aspirated, single cylinder, four-stroke, direct injection diesel engine under cold start and fully warmed-up conditions. The model considers the filling and emptying processes of the cylinder, blowby, intake, and exhaust manifolds. A single zone combustion model is implemented and the heat transfer in the cylinder, intake, and exhaust manifolds are accounted for. Moreover, the derivations include the dynamics of the crank-slider mechanism and employ an empirical model to estimate the instantaneous frictional losses in different engine components. The formulation is coded in modular form whereby each module, which represents a single process in the engine, is introduced as a single block in an overall Simulink engine model. The numerical accuracy of the Simulink model is verified by comparing its results to those generated by integrating the engine formulation using IMSL stiff integration routines. The engine model is validated by the close match between the predicted and measured cylinder gas pressure and engine instantaneous speed under motoring, steady-state, and transient cold start operating conditions. [DOI: 10.1115/1.1290148]

1 Introduction

The design of high power-density diesel engines is influenced by many factors that affect the engine performance, fuel economy, and exhaust emissions. The optimization of the engine design is a very elaborate process which involves extensive hardware modifications, manufacturing and testing. Such a process is costly and time consuming. Mathematical models are being increasingly relied on to simplify and speed up the design optimization procedure. Most existing engine models are nonlinear and have been developed to predict the engine performance over a wide range of operating conditions. They can be classified into two groups. The focus of the first group is to predict the engine behavior at steady state. Whereas, the second group extends the simulation of the engine to include transient conditions. Depending on their levels of sophistication, these models can simulate to various degrees of accuracy the thermodynamics, gas flows, combustion process, heat release rate, heat transfer, dynamic characteristics, and frictional losses in engines [1–8]. These models can be used to predict the engine performance over a wide range of operating conditions. The simulation results will serve to guide the development of various engine components prior to their fabrication; thus, reducing the time and cost of development.

Most of the work done in this area have concentrated on steady-state models for the purpose of modifying engine design parameters in order to minimize emissions and maximize power and fuel economy of the engine. However, recent regulations have imposed stringent emissions and fuel economy standards that can no longer be addressed by a steady-state analysis of the engine. To contribute towards solving this problem, the current research work aims at developing a nonlinear dynamic model for a naturally aspirated, single cylinder, four-stroke, direct injection diesel engine, which can simulate the engine performance under transient and steady-state operating conditions. Moreover, the capability of the model in predicting the engine response under cold starting conditions is examined in this study.

To develop a highly interactive engine simulation package, the mathematical formulation of the model has been coded in modular form that is transparent to the user. The platform used for interconnecting the various engine submodels is selected to be the Simulink module of the general purpose software MATLAB. The main advantage of the Simulink module is its capability of representing the entire engine model by an assemblage of interconnected blocks. Each engine process, which is represented by a single block, can be written in either FORTRAN or C language or defined as an M-file irrespective of the language used in coding other blocks. Such a format would enable the user to test any new engine component model by simply inserting its code in the appropriate block. No modification of the remaining blocks will be required as long as the input and output vectors of the altered block are kept the same.

The mathematical formulations, describing the various engine processes, are presented in the next section. The Simulink engine model along with the discussion of its numerical accuracy are included in the third section. The qualitative validation of the model is carried out in Section 4 by comparing the digital simulation results with the experimental data that was obtained from a single cylinder, direct-injection, research Deutz diesel engine. Subsequently, the work is summarized and the main conclusions are highlighted.

2 Thermodynamic and Dynamic Modeling

A nonlinear model for a naturally aspirated, single cylinder, four-stroke, direct injection diesel engine is developed by taking into consideration the engine thermodynamics, the dynamics of the moving parts and the relationships with which these two engine subsystems would interact. The first part of the model deals with the flows through the intake and exhaust manifolds, intake and exhaust valves, and blowby. In addition, it accounts for the ignition delay, the combustion process and heat transfer. Whereas, the second part of the model deals with the dynamics of the piston-connecting rod-crankshaft mechanism along with the starter dynamics and the frictional losses associated with the pis-

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division February 19, 1997; final revision received by the ASME Headquarters February 23, 2000. Technical Editor: D. Assanis.

ton assembly, engine bearings, auxiliaries, and valve trains. The formulations pertaining to each of these processes are briefly discussed in this section.

Flow Through a Restriction. The formulation, derived for the computation of the mass flow rate through a restriction [6], has been used herein to model the flow rates across the valves and other restrictions. The flow is assumed to be quasi-steady, one-dimensional, and isentropic. The fluids are treated as ideal gases with constant specific heats. Based on these assumptions, the mass flow rate for a subsonic flow can be computed from

$$\dot{m}=c_d A_v P_o\left(\frac{k}{RT_o}\right)^{1/2}\sqrt{\frac{2}{k-1}\left[\left(\frac{P_s}{P_o}\right)^{2/k}-\left(\frac{P_s}{P_o}\right)^{(k+1)/k}\right]}. \quad (1)$$

However, when $P_s/P_o \leqslant [2/(k+1)]^{k/(k-1)}$, the flow becomes sonic and its mass flow rate can be determined from

$$\dot{m}=c_d A_v P_o\left(\frac{k}{RT_o}\right)^{1/2}\sqrt{\left(\frac{2}{k+1}\right)^{(k+1)/(k-1)}}, \quad (2)$$

where c_d and A_v for the intake and exhaust valves are adopted from previous work done on the same engine [3]. The discharge coefficient is expressed as the ratio of the measured air flow to the theoretical air flow through the intake and exhaust valves. The measured data are obtained from static bench tests for a number of fixed valve lifts. Whereas, the theoretical values are computed based on an isentropic flow through an orifice with a diameter equal to the inner seat diameter of the valve. Note that the theoretical values are computed under the same range of pressure drop across the port as to the one experienced when collecting the measured data. Moreover, the effective valve areas are determined from the valve train geometry and the valve lift.

The effective flow areas of the blow-by model are determined by computing the openings of the compliant piston rings. Their corresponding discharge coefficients are assumed to be constant and equal to 0.6.

Thermodynamic Processes: Governing Equations and Assumptions. Four control volumes are used in the model to represent the intake and exhaust manifolds, the combustion chamber and the space between the piston, piston rings, and cylinder walls (see Fig. 1). The formulation ignores the effect of pressure waves inside the control volumes and treats the gases as a homogeneous mixture of ideal gases which is assumed to be at a uniform temperature and pressure at each instant in time. Therefore, the instantaneous state of the mixture is dependent on the temperature, T, pressure, P, and equivalence ratio, ϕ. Moreover, the mass flows across the boundary of the combustion chamber, during the period when the intake and exhaust valves are closed, are assumed to be limited to the fuel injection and the blowby flow. In addition, the injected fuel is assumed to instantaneously evaporate.

From the continuity equation, one can write

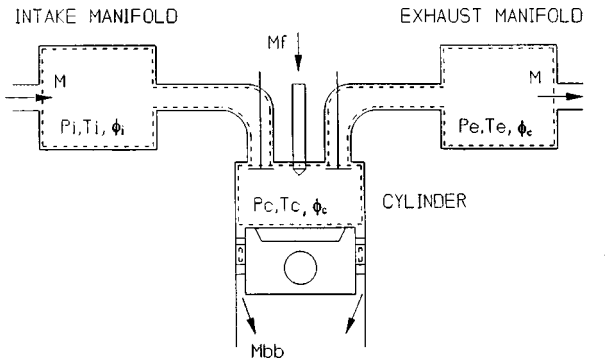


Fig. 1 Schematic of the engine control volumes

$$\frac{dm}{dt}=\dot{m}_{in}+\dot{m}_{ex}+\dot{m}_{bb}+\dot{m}_f, \quad (3)$$

where $\dot{m}_{in}$, $\dot{m}_{ex}$, and $\dot{m}_{bb}$ are determined from Eqs. (1) and (2).

Moreover, the implementation of the first law of thermodynamics yields

$$\frac{d(mu)}{dt}=-P\frac{dV}{dt}+\dot{Q}_{ht}+\sum_j h_j\dot{m}_j, \quad (4)$$

where $P(dV/dt)$ is the rate of work done by the piston displacement. $h_j\dot{m}_j$ is the energy carried out or brought in by the flowing gas and injected liquid fuel in addition to the work required to push the gases in or out at the j^{th} location of the system boundary.

Applying the state equation one obtains

$$PV=mRT. \quad (5)$$

In addition, u and R can be written as [9]

$$u=u(P,T,\phi)$$

$$R=R(P,T,\phi). \quad (6)$$

Moreover, the time derivative of the equivalence ratio is given by

$$\dot{\phi}=\frac{1+\phi\,\mathrm{FAR}_s}{m}\left(\frac{1+\phi\,\mathrm{FAR}_s}{\mathrm{FAR}_s}\dot{m}_f-\phi\frac{dm}{dt}\right). \quad (7)$$

Using Eqs. (5)–(7) into Eq. (4) and rearranging the terms, one obtains

$$\frac{dT}{dt}=\left(\frac{\partial u}{\partial T}+\frac{P}{T}\frac{\partial u}{\partial P}\frac{AA}{BB}\right)^{-1}\left\{-\frac{P}{m}\frac{dV}{dt}+\frac{1}{m}\left(\dot{Q}_{ht}+\sum_j h_j\dot{m}_j-u\frac{dm}{dt}\right)-\frac{\partial u}{\partial\phi}\frac{d\phi}{dt}-CC\right\}, \quad (8)$$

where

$$AA=1+\frac{T}{R}\frac{\partial R}{\partial T}\quad BB=1-\frac{P}{R}\frac{\partial R}{\partial P}$$

$$CC=\frac{P}{BB}\left[\frac{1}{V}\frac{dV}{dt}-\frac{1}{m}\frac{dm}{dt}-\frac{1}{R}\frac{\partial R}{\partial\phi}\frac{d\phi}{dt}\right]\frac{\partial u}{\partial P}.$$

Equations (3) and (8) can be numerically integrated to determine the mixture mass and temperature. The state equation can then be used to compute the mixture pressure.

Ignition Delay. The ignition delay, ID, is dependent on the fuel properties along with the mixture pressure, temperature and equivalence ratio. It is estimated once per engine cycle according to the following empirical formulation which was developed by Hardenberg and Hase [10]:

$$ID=(0.36+0.22v_p)\exp\left[E_a\left(\frac{1}{RT_a}-\frac{1}{17190}\right)+\left(\frac{21.2}{P_a-12.4}\right)^{0.63}\right], \quad (9)$$

where the activation energy, E_a, is defined as

$$E_a=\frac{618840}{CN+25}.$$

The mean temperature, T_a, and pressure, P_a, are computed as follows:

$$T_a=T_{amb}CR^{0.37}\quad P_a=P_{amb}CR^{1.37}.$$

Combustion Process. A single zone combustion model is used to describe the combustion process in the single cylinder, direct-injection, diesel engine. The combustion process is assumed to be complete and its overall effect is represented by the apparent fuel burning rate (AFBR). Moreover, the fuel is considered to burn in part as ''pre-mixed combustion'' and in part as

''diffusion controlled combustion'' [11]. These two phenomena are assumed to start at the ignition point and proceed simultaneously during the combustion period. Therefore, the total fuel burning rate can be expressed as

$$\dot{m}_f = \dot{m}_p + \dot{m}_d . \tag{10}$$

Initially, the ''pre-mixed combustion'' is considered to consume most of the evaporated fuel present in the combustion chamber at the end of the ignition delay period. The combustion process is then assumed to continue in the ''diffusion controlled mode'' only. A β term is introduced by Watson and co-workers [11] to quantify the portion of the fuel consumed in the ''pre-mixed'' burning mode. The burning factor depends on the length of the ignition delay period and the overall equivalent ratio prior to ignition, ϕ_{ig}. However, it should be mentioned that, in this study, the computation of ϕ_{ig} was performed based on the total amount of the injected fuel. β is defined as [11]

$$\beta = \frac{m_p}{m_f} = 1 - 0.926 \phi_{ig}^{0.37} ID^{-0.26}. \tag{11}$$

Thus, the total burning rate can now be written in the following nondimensional form

$$\dot{M}_f(\tau) = \frac{\dot{m}_f(\tau)}{m_f/(\theta_{\text{end}} - \theta_{ig})} = \beta \dot{M}_p(\tau) + (1-\beta)\dot{M}_d(\tau) \tag{12}$$

$$\tau = (\theta - \theta_{ig})/(\theta_{\text{end}} - \theta_{ig}),$$

where the component burning rates are written as [11]

$$\begin{aligned} \dot{M}_p(\tau) &= c_1 c_2 \tau^{c_1 - 1}(1 - \tau^{c_1})^{c_2 - 1} \\ \dot{M}_d(\tau) &= c_3 c_4 \tau^{c_4 - 1} \exp(-c_3 \tau^{c_4}). \end{aligned} \tag{13}$$

Note that $\dot{M}_d$ is a Wiebe function and the c_i's shape factors were determined in a previous investigation on the same engine [3].

Heat Transfer. The heat transfer rate from the cylinder mixture to the piston top, cylinder walls and cylinder head, $\dot{Q}_{ht}$, can be obtained from

$$\dot{Q}_{ht} = \alpha A_w (T_m - T_w), \tag{14}$$

where α, expressed in KW/m^2 °K, is defined as [12]

$$\alpha = 130 D^{\bar{m}-1} P_m^{\bar{m}} T_m^{0.75 - 1.62\bar{m}} \left[c_1 v_p + c_2 \frac{V_i T_1}{P_1 V_1}(P_m - P_{mt}) \right]^{\bar{m}}. \tag{15}$$

The exponent $\bar{m}$ is selected to be 0.8 which corresponds to turbulent fluid flow in pipes. The first term in the bracket, $c_1 v_p$, reflects the effect of the piston motion on the average gas velocity in the cylinder. c_1 is 6.18 during the scavenging period and 2.28 during the compression and expansion strokes [12]. The second term in the bracket accounts for the additional gas velocity induced by the combustion process. c_2 is constant and equal to 3.24×10^{-3} m/sec °C [12].

Engine Dynamics. Only the rigid body motion of the piston/connecting-rod/crankshaft mechanism is considered in this work. The connecting rod is treated as a two-lumped mass system. The first mass is assumed to be connected to the crank-pin while the second mass is attached to the piston assembly.

The equation of motion, derived based on the Lagrange principle, can be expressed as [13]

$$\begin{aligned} (I_e + m_r r^2 \psi^2) \frac{d^2\theta}{dt^2} + \left(m_r r^2 \Psi \frac{d\Psi}{d\theta} \right) \left(\frac{d\theta}{dt} \right)^2 \\ = T_{\text{gas}} + T_{\text{start}} - T_{\text{fric}} - T_{\text{load}}, \end{aligned} \tag{16}$$

where the reciprocating mass, m_r, consists of the piston assembly mass and the mass of the reciprocating part of the connecting rod. Ψ can be related to θ as follows:

$$\Psi = \sin(\theta) + \frac{(r/l)\sin\theta\cos\theta}{[1 - ((r/l)\sin\theta)^2]^{0.5}}. \tag{17}$$

The numerical integration of the above equation yields the instantaneous angular velocity of the engine. Furthermore, the gas pressure, used in determining T_{gas}, along with the angular velocity of the crankshaft serve to couple the thermodynamics and dynamics portions of the engine model.

Engine Starter. The general mathematical formulation for determining the starter torque is adopted from the work done by Poublon and his co-workers [14]

$$T_{\text{starter}} = 305 \exp\left(-0.0051 \frac{d\theta}{dt} \right). \tag{18}$$

It should be pointed out that the numerical values appearing in the above equation have been determined experimentally in our laboratory for the Deutz diesel engine.

Engine Frictional Losses. The engine friction torque, T_{fric}, is required by the rigid body equation of motion of the piston assembly/connecting-rod/crankshaft mechanism. The frictional losses formulation, used herein, follows the work done by Rezeka and Henein [15]. It accounts for the losses stemming from the piston rings viscous and mixed lubrication regimes ($T_{\text{fric}}^{(1)}$ and $T_{\text{fric}}^{(2)}$), piston skirt hydrodynamic lubrication regime ($T_{\text{fric}}^{(3)}$), valve train friction ($T_{\text{fric}}^{(4)}$), auxiliaries and unloaded bearing friction ($T_{\text{fric}}^{(5)}$), loaded bearing friction ($T_{\text{fric}}^{(6)}$). The total friction torque, T_{fric}, is then defined as

$$T_{\text{fric}} = \sum_{i=1}^{6} \alpha_i T_{\text{fric}}^{(i)}. \tag{19}$$

The α_i terms are determined by minimizing the magnitude of the error between the predicted and the measured friction torques [15]. The reader is referred to Appendix A for the expressions of the $T_{\text{fric}}^{(i)}$ terms along with the numerical values of their corresponding α_i terms.

The formulations, obtained by modeling the various engine components, have led to a set of nonlinear stiff differential equations. These equations are then numerically integrated to obtain the simulated engine performance.

3 Simulink Engine Model

The goal for developing a diesel engine model is to come up with a tool for assessing the overall engine performance under transient and steady-state conditions. Such a tool has to be user-friendly and should only require basic technical background from its users. Furthermore, the code has to be written in modular form that is transparent to the user. The rationale is to provide the user with the capability of testing his/her own model of any engine component within a framework of an overall engine simulation package. This can be easily done by replacing the specific module with the user's own module, without modifying the remaining modules as long as the input and output vectors of the new module match those of the old one.

The platform used for interconnecting the various submodels of the engine is selected to be the Simulink module of the general purpose software MATLAB. Each submodel can be introduced in the overall Simulink model of the engine as a subsystem block or a user-definable block called S-function block. The subsystem block representation describes the functionality of the engine subsystem by using blocks that are readily available in the Simulink libraries. Whereas, the S-function block representation enables one to define the functionality of the engine subsystem in three

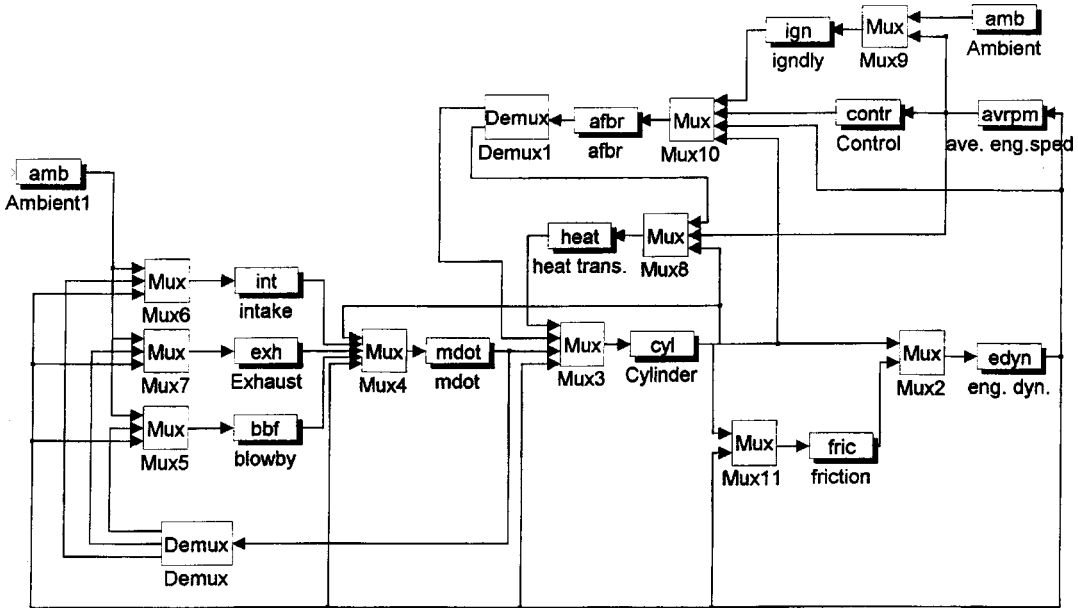


Fig. 2 Simulink model of a single cylinder engine

different ways, namely, a FORTRAN code, a C code or an M-script file. The latter consists of a code written based on MATLAB built-in functions. The greatest advantage of the MATLAB/Simulink modeling environment is its flexibility to accept submodels that are defined as an M-script file or written in either FORTRAN or C language irrespective of the languages that are used in coding the remaining blocks of the Simulink model. This is because each submodel is individually compiled and linked to a gateway routine in order to yield a MATLAB-callable C or FORTRAN program referred to as a MEX-file. The submodel is then introduced in the overall Simulink model by entering the name of the MEX-file in the name field of the dialog box of the S-function block. For further details on how to use MATLAB/Simulink module, the reader is referred to the work of Dabney and Harman [16].

Once the Simulink engine model is fully constructed, the stiff integrators that are readily available through MATLAB can be used to generate the numerical results.

To assess the feasibility of using MATLAB as the main platform for a highly interactive engine simulation package, all the equations derived for the engine model are first coded in FORTRAN and simultaneously integrated by using the International Mathematical and Statistical Library (IMSL) stiff integrator. This step serves a dual purpose. First, it aids in checking the integrity of the model after it has been imported into MATLAB. Second, the numerical accuracy and the speed of execution of the MATLAB stiff integrator can be compared to those of a commonly used and widely accepted IMSL gear stiff integrator.

Next, this study has investigated the effect of increased level of inter-connectivity between the blocks of the Simulink model on the execution time of the digital simulations. This was done by developing three Simulink models for the same engine. The first Simulink model consisted of a single S-function block. The entire engine formulation, which was originally coded in FORTRAN language, was complied and linked to a gateway routine in order to generate the corresponding MATLAB-callable FORTRAN program. The resulting MEX-file was then imported into the Simulink model by defining its name in the dialog box of the S-function block.

The same reasoning was followed in generating the second Simulink model, which consisted of two S-function blocks. The original FORTRAN code for the entire engine was partitioned into two parts. The MEX-files corresponding to the partitioned FORTRAN code were then inserted in their respective S-function blocks of the Simulink model.

Similarly, the entire FORTRAN code was rewritten in modular form in order to generate the third Simulink model that consisted of a collection of S-function blocks (see Fig. 2). The digital simulation results were then generated by using the MATLAB Adams/Gear integration algorithm.

The numerical accuracy of the Simulink engine model is verified by the strong agreement between the IMSL results and those generated by the single and multi-block Simulink models. Figure 3 shows the comparison between the IMSL and Simulink results for the crankshaft speed over the first 200 cycles of an arbitrary acceleration mode with no-load. Similarly, Fig. 4 illustrates the pressure traces, as generated by the two numerical integrators, for engine cycle number 200. Table 1 lists the ratios of the computation time required by a Simulink model to that of an IMSL model for the simulation of a single engine cycle. It should be mentioned that the data was generated based on the computational capability of a Sun 20 workstation. Table 1 reveals that the IMSL code is faster in generating the results than the other three engine Sim-

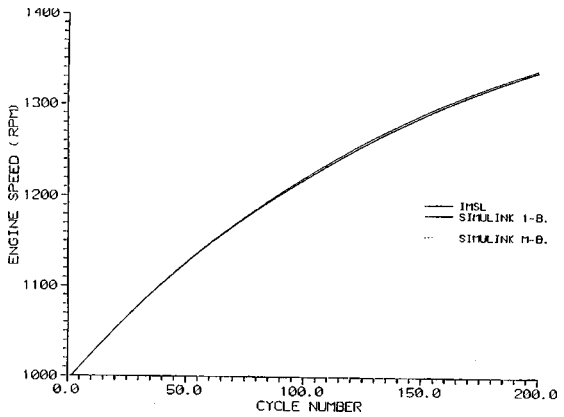


Fig. 3 Average engine speed as predicted by IMSL, Simulink single and multi-block models

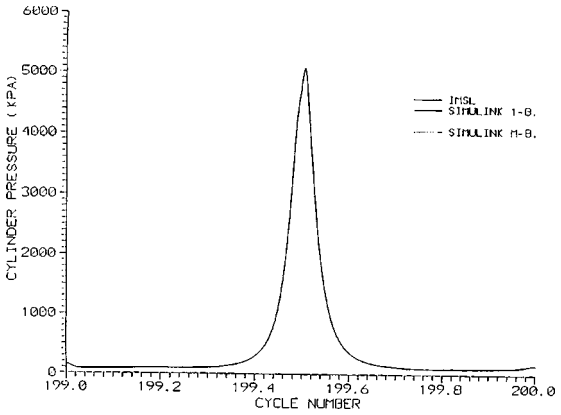


Fig. 4 Cylinder gas pressure as predicted by IMSL, Simulink single and multi-block models

Table 1 Ratios of the computation time required by a Simulink integrator, for three different forms of the engine model, to that of an IMSL integrator

Integration Algorithm and Modular Form	Ratio of Computational Time for One Engine Cycle
IMSL integration routine	1.0
Simulink one-block model	1.175
Simulink two-block model	1.675
Simulink multi-block model	4.025

ulink models. Moreover, it demonstrates that the computation time increases with the number of blocks in the Simulink model. This may be linked to the increased data transfer among the blocks.

It should be emphasized that this exercise serves to demonstrate the preservation of the numerical accuracy of the results as one moves from the IMSL environment to a more interactive and user-friendly environment. However, no conclusion can be made at this stage concerning the accuracy of the model in predicting the actual engine performance. This is made by comparing the theoretical results with experimental data in the following section.

4 Theoretical and Experimental Results

The numerical values of the parameters in the engine model are calibrated based on experimental data. They are listed in Table 2. As a first step towards qualitatively validating the engine model, the predicted and measured cylinder pressure and instantaneous engine speed are compared in Figs. 5 and 6 under motoring condition. Figure 5 shows a good agreement between the theoretical and experimental cylinder pressure curves. Whereas, the slight overestimation of the cylinder gas pressure during the compression stroke has resulted in a larger deceleration in the predicted engine speed before TDC (see Fig. 6). Moreover, the discrepancies between the theoretical and measured instantaneous engine speed curves can be attributed to inaccuracies in the friction formulations.

Figure 7 shows a comparison between the predicted and experimental cylinder gas pressure traces under firing conditions with no load while the average engine speed was 1795 rpm. In this case, the slight underestimation of the gas pressure during the compression stroke has resulted in an overestimation of the instantaneous engine speed near TDC (see Fig. 8).

Finally, the initial conditions of the simulation have been set to mimic the cold start of the engine. This is done by setting initial pressures and temperatures of all engine components to the atmospheric pressure and to room temperature of 30 °C. Figure 9 exhibits a close match between the theoretical and experimental cylinder gas pressure curves. Figure 10 demonstrates the capability of the model in qualitatively predicting the instantaneous engine speed. The discrepancy between the theoretical and experimental results may be due to the same reasons stated earlier as well as to the differences in the fueling rate. Furthermore, it should be men-

Table 2 Engine specifications

Deutz 210 Diesel Engine	
Cylinder bore	0.095 m
Stroke	0.095 m
Number of cylinders	1
Displaced volume	0.673 l
Length of the connecting rod	0.16 m
Compression ratio	17:1
Engine moment of inertia	$0.36 Kg.m^2$
Intake valve diameter	0.0415 m
Exhaust valve diameter	0.0365 m
Valve Timing:	
Opening of intake valve	23 BTDC
Closing of intake valve	63 ABDC
Opening of exhaust valve	63 BBDC
Closing of exhaust valve	23 ATDC

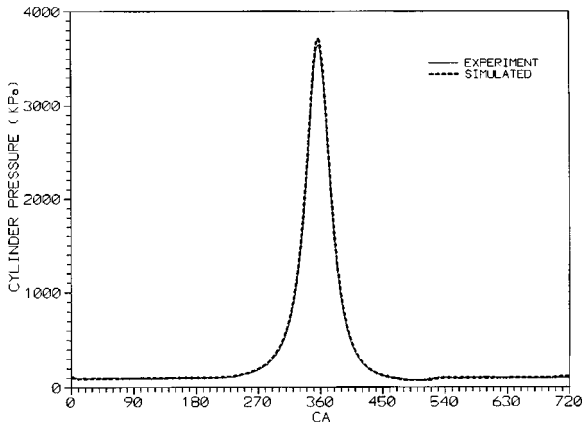


Fig. 5 Comparison between experimental and predicted cylinder gas pressure under motoring

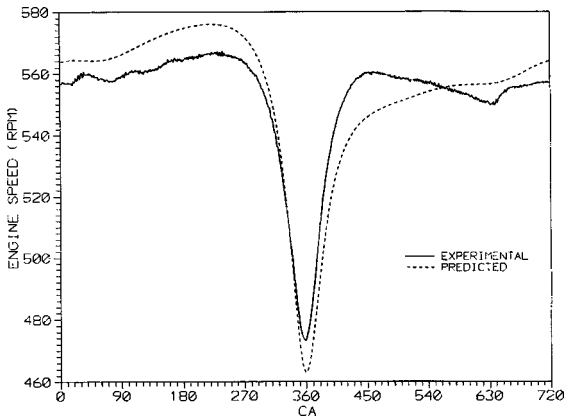


Fig. 6 Comparison between experimental and predicted instantaneous engine speed under motoring

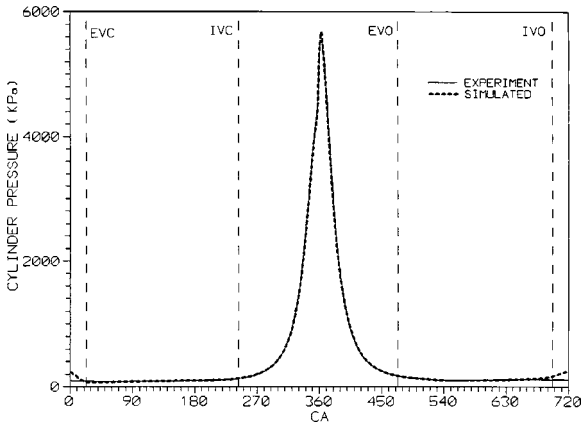


Fig. 7 Comparison between experimental and predicted cylinder gas pressure under firing

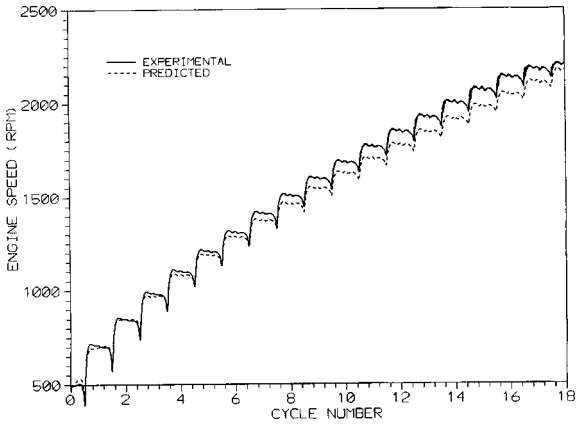


Fig. 10 Comparison between experimental and predicted transient engine speed

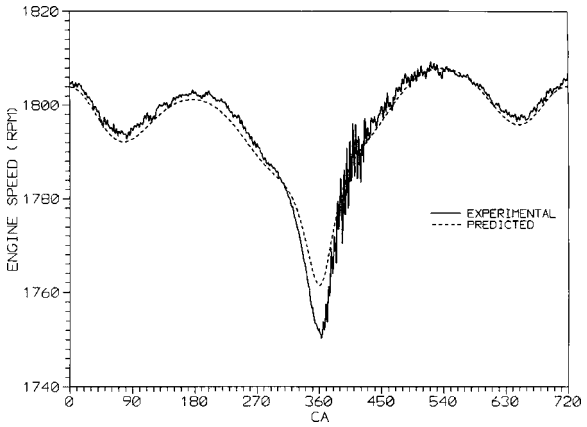


Fig. 8 Comparison between experimental and predicted instantaneous engine speed under firing

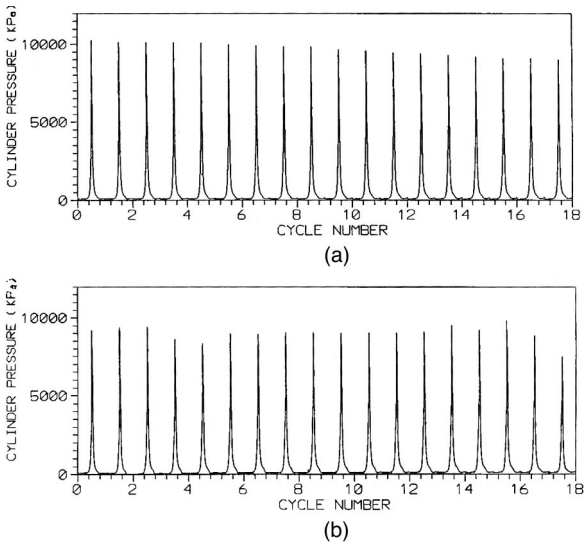


Fig. 9 Predicted and experimental cylinder gas pressure during cold starting: (*a*) simulation results and (*b*) experimental results

tioned that the effect of the engine governor has been excluded from the mathematical formulation of the model.

5 Summary and Conclusions

Stringent regulations, recently imposed on emissions and fuel economy, cannot be addressed by a steady-state analysis of the engine. The current research work has focused on developing a dynamic model for a naturally aspirated, single cylinder, four-stroke, direct injection diesel engine, which can simulate the engine performance under transient and steady-state operating conditions. The model accounts for the filling and emptying processes of the cylinder, blow by, intake and exhaust manifolds. A single zone combustion model is implemented and Woschni's formulation is used to determine the heat transfer between the cylinder charge and its boundaries. The derivations include the dynamics of the crank-slider mechanism and employ an empirical model to estimate the instantaneous frictional losses in different engine components.

In this study, a user friendly and interactive engine simulation package has been developed. This is done by coding the formulation in modular form whereby each module represents a single process in the engine. The entire engine model is defined by an assemblage of interconnected blocks within the Simulink module of the general purpose software MATLAB. Such a format would enable the user to test any new engine component model by simply inserting its code in the appropriate block. No modification of the remaining blocks will be required as long as the input and output vectors of the altered block are kept the same. The numerical accuracy of the Simulink model is verified by comparing its results to those generated by commonly used IMSL stiff integration routines. Furthermore, this study demonstrates that the IMSL code is faster than the Simulink engine models in generating the numerical results. Moreover, the computation time is proven to increase with the number of blocks used in the Simulink model. This may be linked to the increased data transfer among the blocks.

Next, the model is tested under motoring, steady-state and transient operating conditions. The validation of the model is demonstrated by the agreement between the predicted and experimental results of the cylinder pressure and the instantaneous engine speed. The discrepancies in the results are linked to unmodeled system dynamics and inaccuracies in the adopted values of the engine model parameters.

Future research work will aim at including in the model some of the engine processes that are excluded from the current formulation, taking into consideration the effect of the governor, improving the accuracy of the numerical values of the model param-

eters, extending the model to multi-cylinder engines, and simulating the engine performance under different ambient temperatures.

Acknowledgment

The authors would like to acknowledge the technical and financial support of the Automotive Research Center (consortium of five universities directed by the University of Michigan) by the National Automotive Center, located within the US Army Tank-Automotive Research, Development, and Engineering Center (TARDEC), Warren Michigan.

Nomenclature

A_v = cross sectional area of the restriction (m^2)
A_w = surface area of the cylinder walls, piston top and cylinder head (m^2)
c_d = discharge coefficient
CN = fuel cetane number
CR = compression ratio
D = cylinder bore (m)
FAR_s = stoichiometric fuel-air ratio
G = number of intake or exhaust valves per cylinder
h = clearance between the piston skirt and the cylinder liner (m)
h_i = enthalpy of the flux at the i^{th} location of the system boundary (kJ/kg)
I_e = mass polar moment of inertia of the crankshaft, flywheel, main gear and the rotating part of the connecting rod ($kg.m^2$)
k = ratio of specific heats
K = dummy number indicating the angle of the crank
l = length of the connecting rod (m)
L_s = spring load (N)
M = length of the piston skirt (m)
m = total mass of the cylinder contents (kg)
$\dot{m}_{bb}$ = blowby mass flow rate (kg/s)
$\dot{m}_d$ = fuel burning rate in the diffusion controlled mode (kg/s)
$\dot{m}_{ex}$ = mass flow rate through the exhaust valve (kg/s)
$\dot{m}_f$ = mass rate of the injected fuel (kg/s)
$\dot{m}_i$ = mass flow rate crossing the system boundary at the i^{th} location (kg/s)
$\dot{m}_{in}$ = mass flow rate through the intake valve (kg/s)
$\dot{m}_p$ = fuel burning rate in the pre-mixed mode (kg/s)
n_c = number of compression rings
n_o = number of oil rings
P = pressure (kPa)
P_e = elastic pressure of the ring (N/m)
P_{mt} = instantaneous cylinder pressure during a motoring cycle (kPa)
$\dot{Q}_{ht}$ = rate of heat transfer to the walls of the combustion chamber (KW)
R = gas constant (kJ/kg.°K)
r = crank radius (m)
r_c = journal bearing radius (m)
T = temperature, (°K)
T_{fric} = friction torque (N.m)
T_{gas} = torque generated by the gas pressure in the cylinder (N.m)
T_{load} = load torque (N.m)
T_{start} = Starter torque (N.m)
u = internal energy of the material contained inside the combustion chamber (kJ/kg)
V_i = instantaneous cylinder volume (m^3)
v_p = average piston speed (m/s)
w = ring width (m)
μ = oil film viscosity (Pa.s)
ϕ = equivalence ratio
θ = crank angle (radian)
ω = instantaneous angular velocity of the crankshaft (rad/sec)

Subscripts

c = indicates that the variable is evaluated under the cylinder conditions
e = denotes that the variable is evaluated under the exhaust manifold conditions
end = indicates that the variable is evaluated at the end of the combustion process
i = indicates that the variable is evaluated under the intake manifold conditions
ig = denotes that the variable is evaluated at the moment prior to ignition
m = denotes the cylinder mixture
o = indicates that the variable is evaluated at stagnation conditions
s = indicates that the variable is evaluated under static conditions
w = denotes a variable associated with the bounding surfaces of the cylinder contents
1 = indicates that the variable is evaluated under cylinder conditions at the moment when the intake valve closes
$(\cdot)$ = a dot above a variable denotes a mass flow rate

Appendix A

The total friction torque is treated as a linear combination of the friction torques of the various engine components. This appendix provides the expressions for the $T_{fric}^{(i)}$ terms along with the numerical values of the α_i weighting factors as defined by Rezeka and Henein [15]. The friction torque induced by the hydrodynamic lubrication regime of the piston rings is given by

$$T_{fric}^{(1)}=[\mu v_p(P_e+P_{gas})w]^{0.5}D(n_0+0.4n_c)r|K|. \quad (A1)$$

The friction torque due to the mixed lubrication regime at the piston rings is

$$T_{fric}^{(2)}=\pi Dn_c w(P_e+P_{gas})[1-|\sin\theta|\}r|K|. \quad (A2)$$

The friction torque resulting from the hydrodynamic lubrication regime of the piston skirt is

$$T_{fric}^{(3)}=\mu\frac{v_p}{h}DMrK. \quad (A3)$$

The valve train friction torque is

$$T_{fric}^{(4)}=GL_s r|K|(\omega)^{-1/2}. \quad (A4)$$

The auxiliaries and unloaded bearing friction torque is defined as

$$T_{fric}^{(5)}=\mu\omega. \quad (A5)$$

The loaded bearing friction is obtained from

$$T_{fric}^{(6)}=\frac{\pi}{4}D^2 r_c P_{gas}|\cos\theta|(\omega)^{-1/2}. \quad (A6)$$

The numerical values of the weighting coefficients are

$$\alpha_1=23 \quad \alpha_2=0.252 \quad \alpha_3=1.0$$
$$\alpha_4=0.26 \quad \alpha_5=9.6 \quad \alpha_6=0.5. \quad (A7)$$

References

[1] Dai, Y. M., and Fowler, A., 1993, ''Closed Cycle Diesel Engine Performance Evaluation by Computer Simulation,'' ASME Paper 93-ICE-13.
[2] Watson, N., and Marzouk, M., 1977, ''A Non-Linear Digital Simulation of Turbocharged Diesel Engines Under Transient Conditions,'' SAE Paper 770123.
[3] Gardner, T. P., and Henein, N. A., 1988, ''Diesel Starting: A Mathematical Model,'' SAE Paper 880426.

[4] Medica, V., and Giadrossi, A., 1994, ''Numerical Simulation of Turbocharged Diesel Engine Operation in Transient Load Conditions,'' ASME Trans. Eng. Syst. Des. Anal. PD-Vol. 64-8.3, ASME, New York, pp. 589–596.
[5] Assanis, D. N., and Heywood, J. B., 1986, ''Development and Use of a Computer Simulation of the Turbocompounded Diesel System for Engine Performance and Component Heat Transfer Studies,'' SAE Paper 860329.
[6] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.
[7] Benson, R. S., 1971, ''A Comprehensive Digital Computer Program to Simulate a Compression Ignition Engine Including Intake and Exhaust Systems,'' SAE Paper 710773.
[8] Ramos, J. I., 1989, *Internal Combustion Engine Modeling*, Hemisphere, New York.
[9] Krieger, R. B., and Borman, G. L., 1966, ''The Computation of Apparent Heat Release From Internal Combustion Engines,'' 66-WA/DGP-4, ASME, New York.
[10] Hardenberg, H. O., and Hase, F. W., 1979, ''An Empirical Formula for Computing the Pressure Rise Delay of a Fuel From its Cetane Number and From the Relevant Parameters of Direct-Injection Diesel Engines,'' SAE Paper 790493.
[11] Watson, N., Pilley, A. D., and Marzouk, M., 1980, ''A Combustion Correlation for Diesel Engine Simulation,'' SAE Paper 800029.
[12] Woschni, G., 1967, ''A Universally Applicable Equation for the Instantaneous Heat Transfer Coefficient in the Internal Combustion Engine,'' SAE Paper 670931.
[13] Rezeka, S. F., 1984, ''A Mathematical Model of Reciprocating Combustion Engine Dynamics for the Diagnosis of Deficient Energy Conversion,'' Ph.D. dissertation, Wayne State University, Detroit, MI.
[14] Poublon, M., Patterson, D. J., and Boerma, M., 1985, ''Instantaneous Crank Speed Variations as Related to Engine Starting,'' SAE Paper 850482.
[15] Rezeka, S. F., and Henein, N. A., 1984, ''A New Approach to Evaluate Instantaneous Friction and Its Components in Internal Combustion Engines,'' SAE Paper 840179.
[16] Dabney, J. B., and Harman, T. L., 1998, *Mastering Simulink 2*, Prentice-Hall, Englewood Cliffs, NJ.

Use of an Electrically Heated Catalyst to Reduce Cold-Start Emissions in a Bi-Fuel Spark Ignited Engine

G. N. Coppage

S. R. Bell
e-mail: sbell@coe.eng.ua.edu

Department of Mechanical Engineering,
The University of Alabama,
Box 870276,
Tuscaloosa, AL 35487-0276

Reduction of cold-start emissions using electrically-heated catalyst (EHC) technology was the focus of this work. Comprehensive emission measurements of CO, CO_2, NO_x, and total hydrocarbons (THC) are reported for a spark-ignited engine operated on baseline gasoline and compressed natural gas (CNG). Electric heating times of 0, 20, and 40 s with and without secondary air injection were investigated. The 40-second electric catalyst heating with secondary air injection scenario yielded the greatest catalyst system (EHC+OEM three-way catalyst) conversion efficiencies for THC, CO, and NO_x for gasoline and natural gas fueling. Electric catalyst heating coupled with secondary air injection significantly improved THC and CO emissions for gasoline fueling. THC oxidation was difficult for CNG fueling due to the high content of nonreactive methane in the fuel. The independence of NO_x emissions on heating time was demonstrated for all fueling cases. [DOI: 10.1115/1.1340640]

Introduction

In addition to strategic and economic concerns, the increase in global petroleum consumption has generated serious environmental concerns regarding urban air pollution in the United States. The three major pollutants generated by internal combustion engines (CO, NO_x, and THC) all play major roles in ozone and photochemical smog formation plus global warming.

The contribution of emissions from on-road vehicles to national pollutant levels of CO, NO_x, and VOC is substantial. (VOC, or volatile organic compounds, are associated with the EPA's ambient air standards.) On-road vehicles are divided into four major categories: (1) light-duty gasoline vehicles and motorcycles, (2) light-duty gasoline trucks, (3) heavy-duty gasoline trucks, and (4) diesels [1]. Overall levels of on-road vehicle emissions are dominated by light-duty gasoline vehicle (LDGV) emissions. A summary of the national emissions reported for 1994 for CO, NO_x, and VOC is given in Table 1. The contribution of LDGV emissions to the on-road vehicle totals for CO, NO_x, and VOC is also included in the following table [1]. Examination of the data reveals the importance of on-road vehicles, in particular light-duty gasoline vehicles, to overall emission levels in the United States. Although CO, NO_x, and VOC pollutant levels have recently stabilized, it is predicted that continual increases in vehicle miles traveled (VMT) will continue to offset pollutant conversion improvements.

Research Motivations

Several forms of legislation by federal and state governments (primarily California) have placed emission limits on light-duty vehicles based on model year. In Table 2, the emission limits defined by the EPA for light-duty vehicles dating from the 1980 model year are presented in units of grams per mile. The lowered emissions standards in 1994 were a result of the Clean Air Act Amendments in 1990 which called for significant reductions in vehicle tailpipe emissions. The emission limits are applicable to vehicles at 50,000 and extended to vehicles at 100,000 miles.

The State of California has adopted a stricter ''tier'' legislation in order to combat the severe environmental problems caused by elevated vehicle emissions in urban areas such as Los Angeles. The emission standards, or tiers, are designated by transitional, low, ultra-low, and zero emission vehicle labels (TLEV, LEV, ULEV, and ZEV, respectively). Automobile manufacturers will be required to meet each of these regulation levels in increasing percentages with the passage of time. Tables 3 and 4 present the tier legislation adopted by the California Air Resources Board, or CARB, for light-duty vehicles for 50,000 and 100,000 miles, respectively. Note that CARB expresses its regulated hydrocarbon standard in terms of nonmethane organic gases (NMOG). (NMOG totals are adjusted for the reactivity of individual hydrocarbon species, whereas NMHC totals are not.)

In order to meet these stricter emission regulation laws and avoid possible penalties for noncompliance, substantial improvements will be required by automobile manufacturers in the areas of engine management and catalyst system performance. One particular area of interest is the reduction of tailpipe emissions generated at engine cold-start conditions. At cold start, the standard three-way catalytic converter in current vehicles (i.e., 1992 and later models), requires approximately one minute of hot exhaust flow in order to reach operating or ''light-off'' temperature. Once light-off conditions are reached, catalyst conversion efficiencies attain very high values. However, during the first minute of cold engine operation, the standard catalyst system accomplishes essentially no pollutant conversion and CO, NO_x, and THC emissions are released to the atmosphere at near engine-out levels. It has been estimated that upwards of 75 percent of harmful CO and unburned hydrocarbon vehicle emissions occur during this initial period of cold behavior [2]. Therefore, identifying, testing, and implementing a viable solution to cold-start emissions reduction could assist future emissions regulation compliance.

Table 1 1994 National emissions summary (million kilograms)

CATEGORY	CO	NO_x	VOC
TOTAL	88,900	21,400	21,000
On-Road Vehicles	55,400	6,830	5,710
LDGV	35,500	3,390	3,520

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division June 1999; final revision received by the ASME Headquarters August 1999. Associate Editor: D. Assanis.

Table 2 EPA Emission limit summary for light-duty vehicles (grams per mile)

MODEL YEAR	CO	NO_x	THC
1980-1993	3.4	1.0	0.41
1994 (a)	3.4	0.4	0.25 (c)
1994 (b)	4.2	0.6	0.31 (c)

Notes: (a) These standards apply to vehicles at 50,000 miles. (b) For vehicles at 100,00 miles. (c) These limits are for nonmethane hydrocarbons, or NMHC.

Table 3 Carb emission limit summary for light-duty vehicles with 50,000 miles (grams per mile)

TIER	CO	NO_x	NMOG
TLEV	3.4	0.4	0.125
LEV	3.4	0.2	0.075
ULEV	1.7	0.2	0.040
ZEV	0	0	0

Table 4 Carb emission limit summary for light-duty vehicles with 100,000 miles (grams per mile)

TIER	CO	NO_x	NMOG
TLEV	4.2	0.6	0.156
LEV	4.2	0.3	0.090
ULEV	2.1	0.3	0.055
ZEV	0	0	0

Research Objectives

The overall objective of this research work was to evaluate the cold-start emission characteristics of a CNG/gasoline bi-fuel engine equipped with an electrically-heated catalytic converter (EHC), a stock three-way catalytic converter (TWC), and secondary air injection. The specific research goals were as follows:

- Develop an experimental setup capable of testing an EHC and supplying secondary air injection.
- Evaluate the emission levels for various cold-start electric heating and secondary air injection strategies for gasoline and CNG fueling.
- Determine the best cold-start emissions reduction strategy for bi-fuel operation.

EHC Operational Principles

Although numerous different designs exist for the EHC, the basic aim common to all design types is the reduction of the time required for catalyst light-off. In general, an exhaust system containing an EHC possesses three means for heating the catalyst: (1) electrical energy delivered to the EHC, (2) sensible energy of the exhaust feed stream which depends on exhaust temperature and mass flow rate, and (3) chemical energy liberated in the exothermic (energy-releasing) catalytic reactions once light-off temperature and stoichiometric composition are reached [3]. The relatively cold temperature of the exhaust gas passing through the catalyst system during the initial time of cold-start engine operation can limit the contribution of sensible energy to catalyst heating. Therefore, utilizing the chemical energy available from catalytic reactions is a potential option for achieving faster catalyst light-off and reduced cold-start emission levels.

Injecting a small amount of air just upstream of the catalyst system is referred to as secondary air addition. During the initial cold start period, the engine may operate slightly rich during this period and the addition of the secondary air promotes oxidation reactions in the catalyst which leads to earlier light-off and thus lower levels of cold-start emissions. Simplistic in theory, the EHC dissipates electrical power through resistance heating in a short amount of time by drawing electrical current through its metallic substrate. However, implementation of these concepts on an actual vehicle has raised serious questions regarding power control issues and economic viability.

Whittenberger and Kubsh [4] designed an on-board power controller that delivered 750 amps of current at 12 V to the monolith-shaped metallic substrate of an EHC. By combining precrank and postcrank electric heating, the substrate temperature reached light-off condition [350°C (662°F)] in less than 30 s. Hydrocarbon and carbon monoxide Bag 1 FTP (cold transient phase) emissions were reduced more than 50 percent, compared to results obtained using a standard ceramic-supported TWC. Testing conditions included normal ambient (75°F) and cold ambient (21°F) temperatures, gasoline fueling, plus the addition of secondary air just upstream of the EHC.

Instead of a precrank/postcrank combination heating strategy, Gottberg et al. [5] reported best light-off performance and hydrocarbon emissions reduction using only postcrank EHC heating. With secondary air added, light-off temperature was achieved in less than 20 s. The experimental setup for their work used a 12 V, 600 CCA (cold-cranking amps) battery for the energy supply, a high power electronic switch capable of carrying 400 A, required power cables, and a software controller unit. This control unit performed several duties, including temperature- or time-controlled electric heating and operation of the secondary air pump. In addition, the control unit monitored the integrity of the EHC and its temperature sensor. To optimize HC and CO oxidation at cold-start, secondary air injection was halted as soon as light-off temperature was achieved in order to avoid over-leaning the exhaust gas. This point in time was termed the Air Cut Point (ACP).

The EHC system configuration employed by Kaiser Maus, Swars, and Bruck [6] consisted of a heated element, a second light-off catalyst, and a third main catalyst. Substrate matrix diameter, substrate cell density, and catalyst volume increased in the direction of flow with each passing catalyst portion. An alternative to this cascade design was investigated by Brunson, Kubsh, and Whittenberger [7]. Instead of separated EHC and light-off catalysts, a single-core cascade arrangement was designed which incorporated both components into one metallic core. Subsequently, the EHC mass and energy demand were reduced, thereby adding flexibility to the heated zone design. The close coupling between the two catalyst zones allowed for improved heat transfer characteristics and faster achievement of light-off temperature. Furthermore, the EHC electrical system design was simplified. Although the single-core approach yielded compatible cold-start HC emissions reduction to a two-core cascade at lower energy levels, the authors identified the long-range mechanical durability of this design as an area for future research.

The EPA sought to test the viability of the EHC in a battery of FTP emissions tests conducted on a CNG-fueled vehicle [8]. The EHC was formulated especially for methane oxidation, as were two other main catalysts that were tested. However, the EHC was not coupled with a larger main catalyst as is the standard practice in EHC system design. A current of 500 A was drawn for 20 s of precrank and 40 s of postcrank resistive heating, and approximately 142 l per minute of secondary air was injected for 60 s. For the cold transient phase of the FTP, electrically heating the EHC reduced CH_4, NMHC, and CO emissions while slightly raising NO_x emissions compared to no heating results. Addition of secondary air adversely affected CH_4, NMHC, and CO emissions, probably due to the air's cooling effect and subsequent catalyst light-off delay; NO_x was virtually unchanged from secondary air injection.

Experimental Facilities

Experimental facilities required for this work included (1) the engine and dynamometer test stand, (2) the catalytic converters, (3) the secondary air injection system, (4) the emissions bag sampler, (5) the emissions analyzers, (6) the natural gas conversion

Table 5 Catalyst technical information

DESCRIPTION	DELPHI TWC	EMITEC EHC (HEATER)	EMITEC EHC (LIGHT-OFF)
Substrate Type	ceramic monolith	S-wound metallic	S-wound metallic
Substrate Volume	1.8 liters (0.9 liters ea. brick)	0.045 liters	0.954 liters
Substrate Dimensions	8.1 x 14.5 x 9.1 (length) cm (each brick)	ϕ90 x 7 mm	ϕ90 x 150 mm
Precious Metal Loading	25 grams/ft^3	300 grams/ft^3	300 grams/ft^3
Precious Metal Ratio (Pt:Pd:Rh)	10:0:1	0:1:0	0:1:0

system, and (7) the data acquisition systems. The overall facility development was to be capable of transient emissions testing under a controlled environment which would allow broad comparisons to be made using the electrically-heated catalyst system with either gasoline or compressed natural gas (CNG) fueling.

A four cylinder 2.2 l spark-ignited 1994 General Motors engine was used for this work. This engine possessed multiport fuel injection, feedback closed-loop fuel control, electronic spark timing, and exhaust gas recirculation. Proper engine operation was maintained by the electronic control module (ECM). A Model SF-901 programmable water brake dynamometer manufactured by the SuperFlow Corporation was coupled to the engine. In addition to its transient test capabilities, the SF-901 provided several support systems for the engine, including engine cooling via a shell-in-tube heat exchanger, engine instrumentation, and an operator interface via a remote control console.

Major modifications to the exhaust system were required. The stock three-way catalytic converter (TWC) manufactured by AC Rochester (now Delphi Automotive Systems) was located approximately 1.5 meters downstream of the exhaust manifold. An electrically-heated catalyst manufactured by Emitec was used. Comprised of two distinct portions, the Emitec EHC is actually a converter system. The first small-volume portion contains the EHC heater while the second larger volume portion contains the EHC light-off section. Table 5 summarizes the important physical characteristics of the TWC and EHC.

In order to optimize heat transfer from the EHC to the exhaust gas, the EHC unit was installed just upstream of the TWC. A photograph of the EHC/light-off unit incorporated in the catalyst system is shown in Fig. 1. To provide a consistent power supply to the EHC heater matrix, two heavy duty commercial batteries connected in parallel (separate from the engine electrical system) were utilized. A simple switching mechanism was then designed to carry the high current (approx. 150 A) to the EHC. Also, secondary air was delivered upstream of the EHC at a flow rate of approximately 125 l per minute for one minute after engine crank to aid CO and THC oxidation.

Exhaust emissions were collected in a mobile bag sampler and their molar fractions measured (in parts per million or percent volume) using an NO_x analyzer (chemiluminescence detector), CO_2 and CO analyzers (infrared detectors), an O_2 analyzer (paramagnetic oxygen measurement), plus a hydrocarbon analyzer (flame ionization detector). A sample gas conditioner was used to filter and dry exhaust samples (except THC) prior to measurement. To study instantaneous trends in emission concentrations and exhaust temperatures, existing data acquisition systems were modified and used for this work.

Another key facility component was the Gaseous Fuel Injection (GFI) natural gas conversion kit. It allowed for either gasoline or compressed natural gas (CNG) fueling based on the position of a selector switch mounted on the engine test stand. Replacement of the stock oxygen sensor with a heated one was the lone modification performed on the test engine. Two fiberglass-wrapped aluminum tanks with displacement volumes of 2364 cubic in. each provided CNG storage at 3000 psig. In order to decrease the fuel pressure from tank storage level to proper injection level, a high pressure single stage regulator was used. The GFI compuvalve (computer/metering valve combination) performed speed density calculations and controlled the amount of fuel injected into the engine's intake air stream. A PC interface allowed system calibration and monitoring.

Testing Procedures

Since the EHC is aimed at reducing cold-start engine emissions, only cold-start engine tests were performed for this work. The following criteria were minimum requirements for the achievement of cold-start conditions in the I.C. Engines lab:

- Six hours from completion of the previous engine test.
- Manifold, engine oil, fuel, and cooling water temperatures equal to lab ambient air temperature (which was approximately

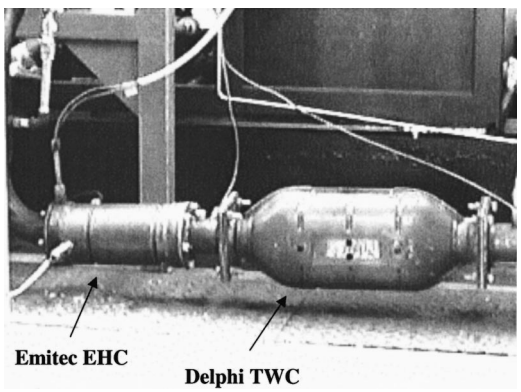


Fig. 1 Catalyst system

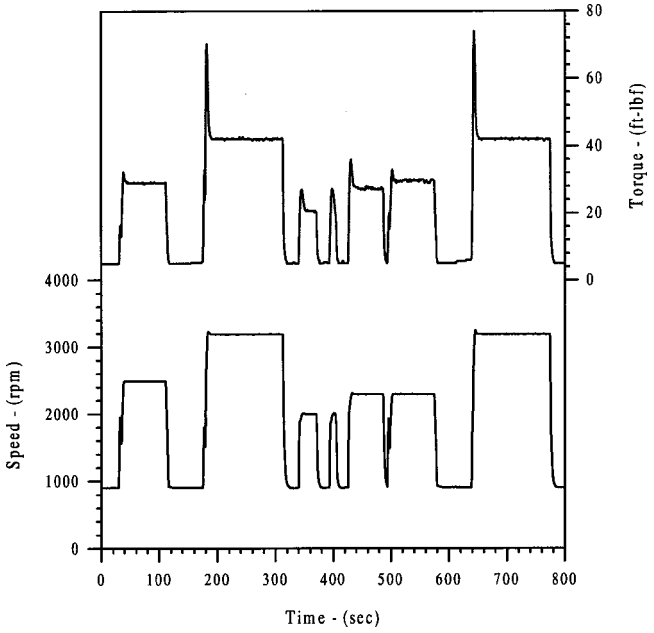


Fig. 2 Transient test cycle speed and load specifications

24°C during testing). Engine oil temperatures were measured in the oil pan and engine coolant temperatures were measured in the engine block and heat exchanger.

The engine speed and brake torque schedule used in testing is shown in Fig. 2. The frequency of the transient speed and load changes were taken from the first 505 s of the FTP-75 (Bag 1), the standard test cycle for light-duty vehicles in the United States developed by the EPA. FTP-75 tests are performed on actual test vehicles by utilizing a chassis dynamometer, whereas for this work a test engine was coupled to a water brake dynamometer. Therefore, emission results generated from this work should not be interpreted as FTP results. Also, direct comparisons between the results of this work and current emission regulations and standards should not be made.

Engine stock fueling rates were used for gasoline tests. In general, an engine with a TWC system operates very near stoichiometric equivalence ration utilizing closed-loop control for this purpose. In the case of natural gas fueling, the conversion kit used appeared to operate slightly lean of stoichiometric fueling. Therefore, a few additional tests were conducted with slightly enriched fueling (near stoichiometric), and those results are also reported.

Data Analysis

The dependent variables for the cold-start engine tests were the concentrations of NO_x, CO_2, CO, O_2, and total hydrocarbons (THC) in the exhaust gas. Sampling location (before EHC or after TWC), fuel type (gasoline or CNG), EHC heating time (0, 20, or 40 s), and secondary air injection (yes or no) were chosen as the independent variables. For each testing condition, multiple engine tests were conducted to establish repeatability. Statistical methods, including Chauvenet's criterion and the Student *t* statistic, were used to construct confidence intervals for the results.

In order to convert dry molar fractions (raw data) into units of grams per mile, several calculations were made. The details of these calculations are omitted here. A cycle distance of 5.60 miles was chosen for all emissions calculations. Consistent use of this number allows comparative results discussions involving gasoline and CNG fueling. The methodology used in determining the SF-901 cycle distance and procedures for calculating mass emissions from the measured mole fractions can be found in Ref. [9].

Any dilution effects on the bag samples from the secondary air injection would be slight as the amount of injected air was on the order of one percent of the engine intake air. The CO_2 data confirmed no significant secondary air dilution effects for the cases with and without the secondary air. The OEM engine spark timings were used in all testing reported here. Catalyst system conversion efficiencies were calculated for each emission constituent using the following equation:

$$\eta_{\text{catalyst}} = \left(1 - \frac{\text{after catalyst concentration}}{\text{before catalyst concentration}}\right) \cdot 100 \text{ percent.} \quad (1)$$

Results and Discussion

All cold-start engine tests followed the transient test cycle described earlier (see Fig. 2). Emissions sampling was taken before the EHC (engine-out emissions) and after the TWC using a bag sampler. Data acquisition systems were used in order to generate emission concentrations (in units of parts per million, or ppm) as a function of test cycle time (ie. continuous or ''real time'' sampling. For certain testing scenarios, secondary air was supplied upstream of the EHC at a volumetric flow rate of approximately 125 l per minute for 60 s. Table 6 provides a comprehensive summary of bag emission results reported in units of grams per mile for each testing scenario.

The following discussion is organized in two main sections. The first section presents and discusses emission results obtained from gasoline fueling, while the second section explores emission results from CNG fueling (stock and enriched). The effects of electric catalyst heating and secondary air injection are investigated for all fueling cases. Real time plots for THC, CO, and NO_x emissions are presented in Figs. 3–5 for gasoline fueling and in Figs. 6–8 for CNG fueling (with respect to gasoline fueling).

Gasoline Case Results. Electrically heating the catalyst without the aid of secondary air injection had very little effect on THC, CO, and NO_x emissions. As can be seen in Table 6, the bag emission results for CO and NO_x emissions are statistically equivalent for the three heating scenarios (0, 20, and 40 s). For THC emissions, a modest 23 percent reduction was realized for 40 s electric heating compared to no heating. At cold-start, the engine operates in open-loop mode for approximately one minute due to the cold oxygen sensor. Consequently, the intake mixture is fuel rich, which translates into an oxygen deficiency for the exhaust gas. As a result, THC and CO oxidation are inhibited. Two factors which directly increase NO_x formation in the cylinders, (1) high combustion temperatures and (2) slightly lean conditions, are not present at cold-start. Therefore, the relatively low concentration of NO_x at cold-start, compared to CO and THC concentrations, essentially eliminates the direct effect of electric heating on overall NO_x levels. This trend can be seen in Fig. 3, in which after cata-

Table 6 Comprehensive bag emission results with 95 percent confidence levels

Fuel	Sample Location	Time (sec)	Air Inj.?	THC (g/mi)	CO (g/mi)	NO_x (g/mi)	CO_2 (g/mi)
Gasoline	b. EHC	--	--	**2.66** +/- *0.09*	**17.15** +/- *0.79*	**2.97** +/- *0.16*	**541** +/- *6*
Gasoline	a. TWC	0	no	**0.44** +/- *0.01*	**5.92** +/- *0.33*	**0.73** +/- *0.14*	**579** +/- *11*
Gasoline	a. TWC	20	no	**0.40** +/- *0.05*	**4.50** +/- *1.54*	**0.69** +/- *0.05*	**529** +/- *18*
Gasoline	a. TWC	40	no	**0.34** +/- *0.03*	**5.13** +/- *1.32*	**0.70** +/- *0.13*	**528** +/- *8*
Gasoline	a. TWC	0	yes	**0.35** +/- *0.03*	**4.61** +/- *0.54*	**0.60** +/- *0.03*	**524** +/- *5*
Gasoline	a. TWC	20	yes	**0.26** +/- *0.03*	**3.07** +/- *0.22*	**0.59** +/- *0.02*	**530** +/- *6*
Gasoline	a. TWC	40	yes	**0.22** +/- *0.05*	**2.70** +/- *0.27*	**0.56** +/- *0.03*	**529** +/- *8*
CNG (stock)	b. EHC	--	--	**1.80** +/- *0.25*	**9.71** +/- *0.01*	**1.43** +/- *0.15*	**370** +/- *1*
CNG (stock)	a. TWC	0	no	**0.50** +/- *0.10*	**0.58** +/- *0.25*	**0.79** +/- *0.17*	**387** +/- *1*
CNG (stock)	a. TWC	20	no	**0.47** +/- *0.05*	**0.36** +/- *0.10*	**0.82** +/- *0.02*	**387** +/- *5*
CNG (stock)	a. TWC	40	no	**0.44** +/- *0.02*	**0.21** +/- *0.02*	**0.76** +/- *0.05*	**387** +/- *7*
CNG (stock)	a. TWC	0	yes	**0.41** +/- *0.05*	**0.34** +/- *0.02*	**0.80** +/- *0.12*	**381** +/- *5*
CNG (stock)	a. TWC	20	yes	**0.41** +/- *0.09*	**0.26** +/- *0.05*	**0.79** +/- *0.07*	**378** +/- *10*
CNG (stock)	a. TWC	40	yes	**0.39** +/- *0.01*	**0.17** +/- *0.02*	**0.76** +/- *0.05*	**381** +/- *5*
CNG (enriched)	b. EHC	--	--	**2.86**	**13.81**	**1.38**	**370**
CNG (enriched)	a. TWC	0	no	**0.52**	**1.62**	**0.08**	**389**
CNG (enriched)	a. TWC	40	yes	**0.43**	**0.98**	**0.08**	**387**

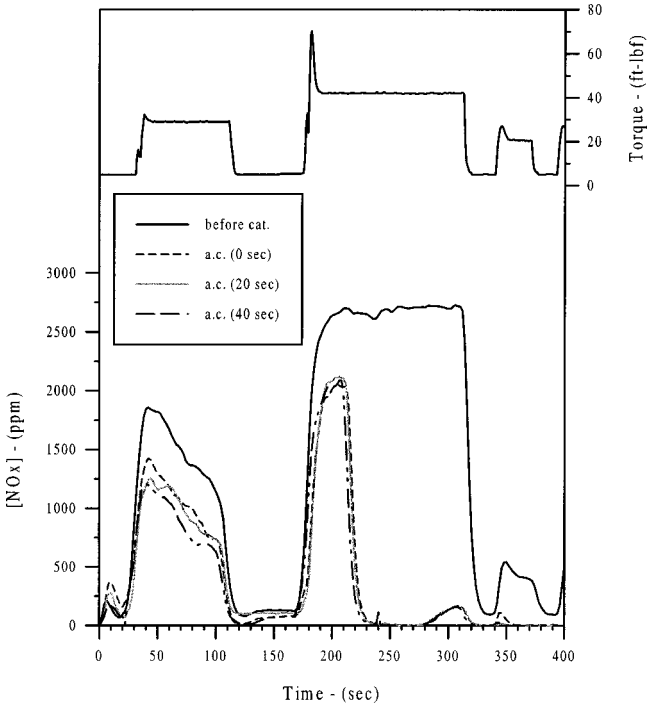


Fig. 3 NO_x concentration in ppm and brake torque as functions of time for gasoline fueling, before and after catalyst sampling, 0, 20, and 40 s electric catalyst heating, without secondary air injection

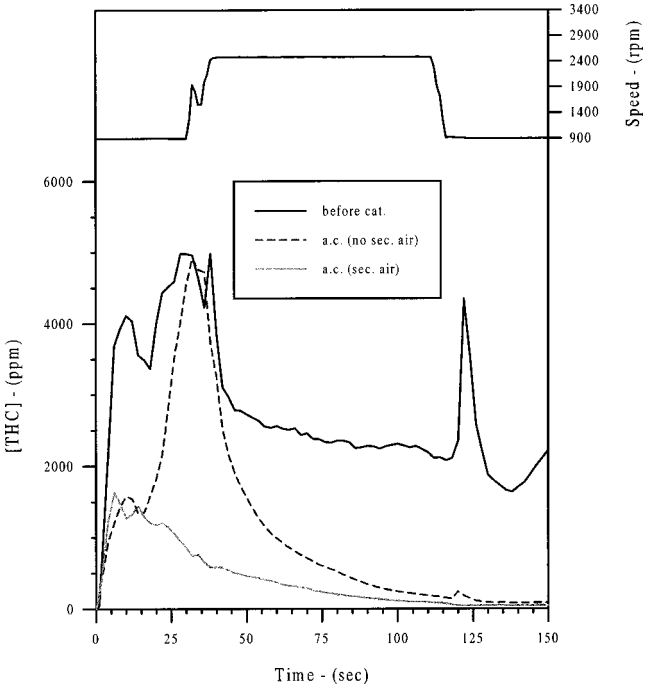


Fig. 5 THC concentration in ppmC and engine speed as functions of time for gasoline fueling, before and after catalyst sampling, 40, s electric catalyst heating, with and without secondary air injection at approximately 125 lpm for 60 s

lyst NO_x levels are basically indiscernible for the three heating scenarios. The same behavior was exhibited with the addition of secondary air injection. As the air-to-fuel ratio in the exhaust is shifted lean due to air injection (which aids CO and THC oxidation, two vital components for NO_x reduction), the NO_x reduction efficiency decreases.

The exhaust gas composition delivered to the catalyst system during the first one or two minutes of engine operation was still fuel-rich, independent of the quantity of electric energy that was added to the stream. Injecting secondary air upstream of the EHC to the rich exhaust would lean the exhaust and quicken the achievement of light-off conditions for THC and CO. Indeed, the

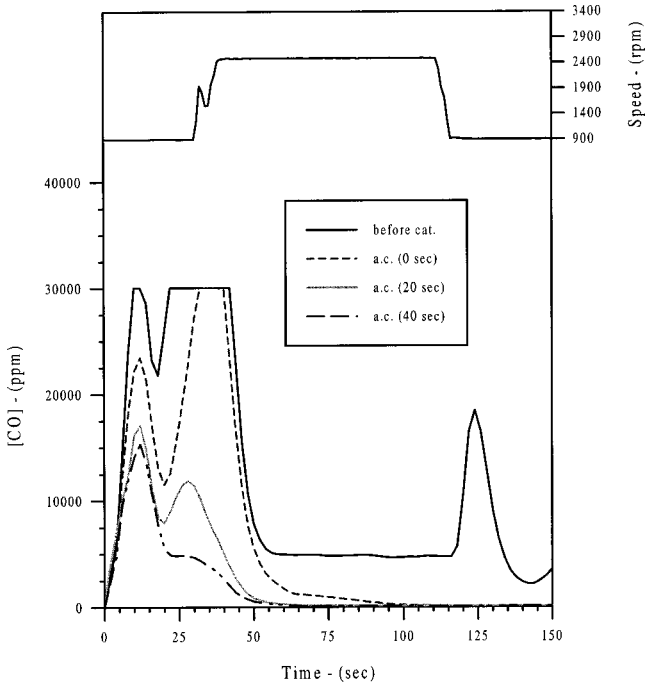


Fig. 4 CO concentration in ppm and engine speed as functions of time for gasoline fueling, before and after catalyst sampling, 0, 20, and 40 s electric catalyst heating, with secondary air injection at approximately 125 lpm for 60 s

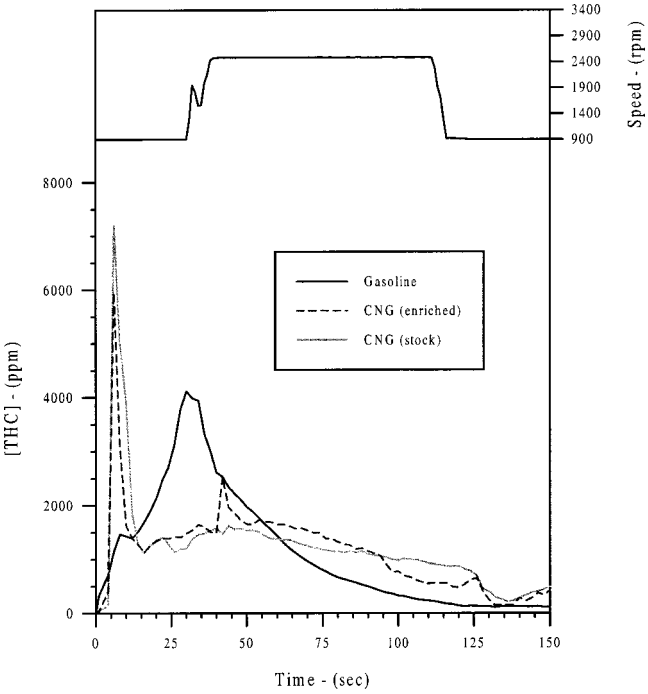


Fig. 6 THC concentration in ppmC and engine speed as functions of time for gasoline and CNG fueling, after catalyst sampling, 0 s electric catalyst heating, without secondary air injection, stock, and enriched CNG fueling

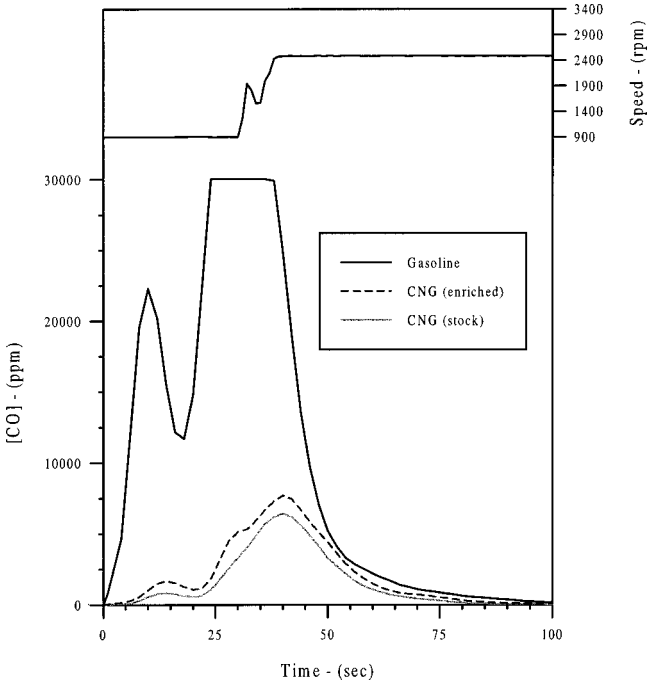


Fig. 7 CO concentration in ppm and engine speed as functions of time for gasoline and CNG fueling, after catalyst sampling, 0 s electric catalyst heating, without secondary air injection, stock, and enriched CNG fueling

effectiveness of electric heating was significantly improved when used in conjunction with secondary air. This trend is evident in Fig. 4, in which CO emissions are plotted for the first 150 s of the test cycle (30,000 ppm was the upper limit of the CO analyzer). For the 20 and 40 s heating cases, the energy available at the EHC and excess oxygen supplied from secondary air injection were sufficient to initiate and sustain CO oxidation during open-loop behavior. Bag emission results support this fact, with 20 and 40 s heating reducing CO emissions by 33 and 41 percent, respectively, compared to no heating. Similar results were obtained for THC emissions, as 20 and 40 s heating yielded emission reductions of 26 and 37 percent, respectively, compared to no heating. Furthermore, the 40 s heat with secondary air case reduced THC and CO emissions by 50 and 54 percent, respectively, compared to the no heat, no secondary air case.

A direct comparison between heating with and without secondary air injection for THC emissions is found in Fig. 5 for the 40 s heating scenario. The presence of exothermic THC oxidation reactions during open-loop behavior was evident for the test run involving secondary air injection, as THC concentration consistently declined during the first minute of engine operation and continued decreasing even after the termination of electric heating and air injection (recall that secondary air injection was terminated after one minute of engine operation). The energy released from oxidation then served to heat the catalyst system and sustain light-off conditions. Similar trends were observed for CO emissions.

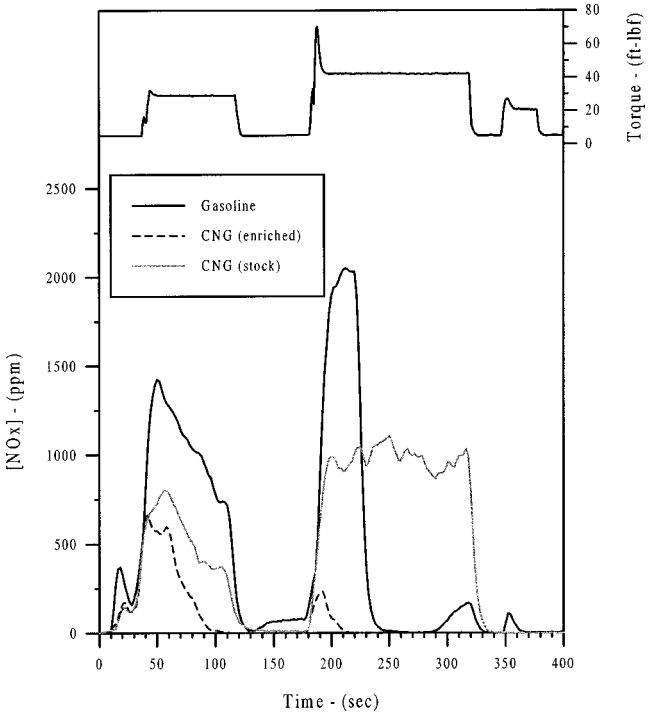


Fig. 8 NO_x concentration in ppm and brake torque as functions of time for gasoline and CNG fueling, after catalyst sampling, 0 s electric catalyst heating, without secondary air injection, stock, and enriched CNG fueling

Natural Gas Case Results. Two fueling cases were investigated for CNG. First, the engine was operated in the stock fueling mode of the GFI kit, in which no changes were made to the factory calibration. As will be discussed, this fueling scenario resulted in lean behavior (slightly leaner than stoichiometric equivalence ratio) and elevated NO_x emissions. In an attempt to improve NO_x emissions throughout the test cycle, the GFI stock calibration data was modified to deliver a slightly more rich fuel/air mixture. Although THC and CO emissions would increase from this strategy, the catalyst system was expected to handle this increase and help NO_x reduction. The following CNG results discussion is presented in two sections corresponding with the two conversion kit fueling modes. Comparative emission plots between gasoline and both CNG fueling modes are found in Figs. 6–8.

Stock Fueling. Due to the low reactivity and high content of methane in the CNG used for this work (approximately 98.5 percent by volume), THC emissions were extremely difficult to oxidize for natural gas fueling. As a result, electric catalyst heating with and without secondary air injection proved ineffective in reducing THC emissions from the no heat, no secondary air test case (see Table 6).

On the other hand, after catalyst CO emissions were reduced to extremely low levels for the stock CNG fueling case. Catalyst system conversion efficiencies of 94, 96, and 98 percent (see Eq. (1)) were realized for 0, 20, and 40 s electric catalyst heating, respectively, without secondary air injection. These efficiencies translated into emission values of 0.58, 0.36, and 0.21 g per mile for the engine test cycle used in this work. Compared to their gasoline counterparts of 4.61, 3.07, and 2.70 g per mile with secondary air injection, the lean nature of the GFI kit in stock fueling mode was evident. Although this lean nature contributed to decreasing CO emissions relative to gasoline levels, the EHC system still proved an effective tool in facilitating CO oxidation for this fueling mode. Specifically, electrically heating the catalyst for 40 s and adding secondary air for 60 s reduced CO emissions by 71 percent compared to the no heat, no secondary air test case.

NO_x reduction for CNG fueling relies mainly on the concentration of CO in the exhaust gas. Substantially lowered CO emissions, therefore, are detrimental to NO_x reduction and low conversion efficiencies are the result. Indeed, bag emission results (Table 6) proved this fact as NO_x emissions were statistically equivalent for all heating cases with and without secondary air injection. Catalyst conversion efficiencies no greater than 47 percent were realized for NO_x emissions in the GFI stock fueling mode.

Enriched Fueling. The GFI stock fueling calibration file was modified to deliver a two percent richer (compared to stock fuel-

ing) fuel-air mixture to the engine. Confidence intervals for the enriched results were not calculated due to the lesser number of test runs (1–2) performed, compared to stock fueling results (3–4 test runs). Therefore, care should be exercised when drawing conclusions concerning these results. Two test cases, (1) no heat without secondary air and (2) 40 s heat with secondary air, were analyzed for enriched fueling. Figures 6–8 present direct comparisons for THC, CO, and NO_x emissions, respectively, for all three fueling modes (gasoline, stock CNG, and two percent enriched CNG).

By increasing the amount of CNG fuel delivered to the cylinders for combustion, the concentration of THC in the exhaust was expected to increase. A modest after catalyst emission increase of 10 percent was realized for test case (2) compared to stock fueling. For test case (1), the THC emission values were statistically equivalent for CNG stock and enriched fueling. Compared to gasoline fueling, THC emissions for test case (1) were similar, with only an 18 percent increase realized for enriched CNG fueling. This trend is shown in Fig. 6. The initial spikes experienced for both CNG fueling modes are quickly counteracted by the GFI kit as closed-loop fueling is attained approximately 20 s after engine crank (from experimental observation). The reason for the faster transition to closed-loop operation was not investigated, but was perhaps in part due to the use of an electrically heat oxygen sensor in the GFI conversion system. For test case (2), THC emissions were increased 77 and 95 percent for stock and enriched CNG fueling, respectively, compared to gasoline fueling. However, the high concentration of nonreactive methane in the CNG exhaust lessens the severity of these increases. The EHC system was moderately successful in improving THC oxidation during enriched CNG fueling as test case (2) reduced THC emissions by only 17 percent compared to test case (1).

For test cases (1) and (2), after catalyst CO emission increases of 179 and 477 percent were realized, respectively, compared to stock CNG fueling results. These large increases should not be viewed as severely detrimental to CO emissions due to the extremely low CO levels associated with CNG stock fueling. The ''relative'' increase is shown in Fig. 7 along with gasoline CO emissions. Furthermore, the enriched catalyst conversion efficiency for CO was comparable to the efficiencies generated from stock fueling. Specifically, CO conversion efficiencies of 88 and 93 percent resulted for test cases (1) and (2), respectively, compared to 94 and 98 percent for the CNG stock fueling counterparts. The EHC system again aided CO oxidation, although not to the extent as seen with stock CNG fueling. Compared to test case (1), CO emissions were reduced 40 percent for test case (2).

By enriching the intake CNG fuel-air mixture and thus increasing the CO content of the exhaust, the reduction of NO_x was vastly improved compared to CNG stock fueling. After catalyst NO_x emission decreases of 90 percent were realized for both test cases compared to the stock fueling runs, thus approximately doubling the catalyst conversion efficiencies. The importance of the fuel management system for the engine is clearly evident from these results. The difference between NO_x emissions for gasoline and CNG fueling (both modes) is shown in Fig. 8. As was seen with stock CNG and gasoline fueling, cold-start NO_x emissions reduction was independent of EHC heating and secondary air injection for the two percent fueling enrichment test runs. A composite NO_x total of 0.08 g per mile was obtained for both test cases (see Table 6).

Summary and Conclusions

Cold-start emissions reduction in a gasoline/CNG bi-fuel spark ignited engine using an electrically-heated catalyst with secondary air injection was the focus of this work. The following summary statements and conclusions are offered.

1 Forty-second electric catalyst heating with secondary air injection yielded the maximum average catalyst conversion efficiencies for gasoline and CNG testing for this test configuration. Average catalyst conversion efficiency is defined as the mean between THC, CO, and NO_x conversion efficiencies for a particular test scenario (e.g., stock CNG fueling, 20 s electric heat, without secondary air injection);

- gasoline: 86 percent catalyst efficiency,
- stock CNG: 75 percent catalyst efficiency;
- enriched CNG: 91 percent catalyst efficiency.

2 For THC and CO oxidation, secondary air injection increased the effectiveness of EHC heating (gasoline fueling) by delivering a lean mixture to the catalyst system.

3 Electric heating was ineffective in reducing NO_x emissions (with and without secondary air injection) for all fueling cases. NO_x is not a major concern at cold-start due to lower combustion temperatures and rich fuel-air mixtures (open-loop operation). Therefore, electric heating has little impact on overall NO_x levels.

4 The stock CNG fueling for the GFI kit appeared to be slightly lean of stochiometric fueling, thus resulting in very low CO emissions and ineffective NO_x reduction.

5 By slightly enriching the CNG fueling (to near stoichiometric) substantially reduced NO_x emissions were achieved while slightly raising after catalyst CO emissions compared to stock fueling.

6 After catalyst THC emissions were comparable for all fueling cases. THC oxidation was more difficult for CNG fueling due to the high content of nonreactive methane in the exhaust.

7 The EHC/TWC system was adaptable to either gasoline or natural gas fueling.

8 The dominating influence from on-road vehicles on overall emission statistics drives the need to reduce vehicle cold-start emissions. The environmental benefits of a viable cold-start emissions reduction strategy for cars and trucks are paramount.

References

[1] U.S. Environmental Protection Agency Report, 1995, ''National Air Pollutant Trends, 1900–1994,'' EPA-454/R-95-011, Research Triangle Park, NC.

[2] Summers, J. C., and Silver, R. G., 1992, ''Catalyst Technologies to Meet Future Emission Requirements for Light-Duty Vehicles,'' *Catalytic Control of Air Pollution—Mobile and Stationary Sources*, R. G. Silver, J. E. Sawyer, and J. C. Summers, eds., American Chemical Society, Washington, D.C., pp. 3–10.

[3] Laing, P. M., 1994, ''Development of an Alternator-Powered Electrically-Heated Catalyst System,'' SAE Paper No. 941042, *LEV/ULEV Emission Technologies*, Society of Automotive Engineers, Warrendale, PA, pp. 137–144.

[4] Whittenberger, W. A., and Kubsh, J. E., 1990, ''Recent Developments in Electrically Heated Metal Monoliths,'' *Recent Trends in Automotive Emissions Control*, Society of Automotive Engineers, Inc., Warrendale, PA, SAE Paper No. 900503, pp. 61–70.

[5] Gottberg, I., Rydquist, J. E., Backlund, O., Wallman, S., Maus, W., Bruck, R., and Swars, H., 1991, ''New Potential Exhaust Gas Aftertreatment Technologies for 'Clean Car' Legislation,'' SAE Paper No. 910840.

[6] Kaiser, F. W., Maus, W., Swars, H., and Bruck, R., 1993, ''Optimization of an Electrically-Heated Catalytic Converter System Calculations and Application,'' SAE Paper No. 932722.

[7] Brunson, G., Kubsh, J. E., and Whittenberger, W. A., 1993, ''Combining Heated and Unheated Core Functions for Improved Cold Start Emissions Performance,'' SAE Paper No. 932722.

[8] Hellman, K. H., Piotrowski, G. K., and Schaefer, R. M., 1994, ''Evaluation of Specialized Methane Catalytic Converters on a CNG-Fueled Vehicle,'' *LEV/ULEV Emission Technologies*, Society of Automotive Engineers, Inc., Warrendale, PA, February, SAE Paper No. 940473, pp. 67–86.

[9] Coppage, G. N., 1996, ''Cold-Start Emissions Reduction in a Bi-Fuel Spark Ignited Engine Using an Electrically-Heated Catalyst,'' Master's thesis, University of Alabama, Tuscaloosa, AL.

Fuel Composition Effects on Emissions From a Spark-Ignited Engine Operated on Simulated Biogases

K. C. Midkiff
e-mail: cmidkiff@coe.eng.ua.edu

S. R. Bell

S. Rathnam

S. Bhargava

Department of Mechanical Engineering,
The University of Alabama,
Box 870276,
Tuscaloosa, AL 35487-0276

Measurements are reported for a spark-ignited (SI) engine burning natural gas and three simulated biogas fuels (natural gas, CO_2, and N_2 mixtures). Exhaust concentrations of CO, CO_2, O_2, NO_x, and unburned hydrocarbons, as well as brake power and brake specific energy consumption, were measured. Leaner mixtures, retarded spark timing and diluent addition (CO_2,N_2) yielded reduced NO_x emissions. NO_x reductions up to 50 percent were achieved at MBT timing through diluent addition. Reduced peak temperatures caused by diluent addition, lean conditions, and retarded spark timing reduced combustion quality slightly, as evidenced by small increases in CO and unburned hydrocarbons emissions. [DOI: 10.1115/1.1338951]

Introduction and Background

Landfill disposal of solid wastes and processing of wastes in sewage digester plants are increasing worldwide. McBean et al. [1] estimate that solid waste disposal is approaching 1 ton per person per year, and that annual landfilling will reach 190 million tons in the United States by the year 2000. About 93 percent of the overall municipal solid waste production (about 10^9 tons per year) is landfilled [2]. The decomposition of solid wastes in landfills by a combination of chemical, physical, and biological processes produces solid, liquid, and gaseous byproducts. The gas formed is often vented and flared to reduce local pollution and prevent a buildup of the potentially explosive gas. Landfill gas (LFG) is a flammable and potentially harmful mixture of methane, carbon dioxide, and numerous trace constituents; therefore, future disposal requirements are expected to be more stringent. In addition to landfill waste deposits, many organic wastes such as animal manure, crop residues, and sewage are rapidly decomposed in ''digesters,'' resulting in large quantities of byproduct digester gas. This gas has a composition and relatively low heating value similar to landfill gas, and must be disposed of in an environmentally acceptable manner.

The growing availability of landfill gases suggests that they be used for energy production rather than simply flared for disposal. It is estimated that for England and Wales the energy available from LFG is equivalent to an electrical generating capacity of about 600 MW/yr [3]. Power generation from landfill gas using spark ignition engines, dual fuel engines, and gas turbines has been successfully demonstrated in the UK for engine sizes ranging from about 0.5 to over 4 MW. In the U.S. as of 1990, 117 landfill methane power plants were in operation, each a few megawatts in size.

Landfill gas and digester gas are referred to as ''biogas'' because they are primarily produced through anaerobic, microbial decomposition of organic matter. Landfill gases typically have a medium-Btu content and consist of 40–60 mole percent methane with most of the balance made up by CO_2. Nitrogen may be present in quantities up to 10 percent depending upon the extent of air penetration into the interior of the landfill. Many trace components, such as hydrogen, higher alkanes, and sulfur compounds are also present, depending on the type of organic matter buried, the time history of the decomposition, and the ambient conditions at the burial site.

Gas composition can have a profound impact on the performance and emission characteristics of the engine. Although biogas mixtures vary considerably in composition, for convenience these fuels can be adequately represented by mixtures of methane containing substantial concentrations of the diluents CO_2, N_2, and water vapor. Flame speed, ignition delay, and knock characteristics are affected by fuel gas composition, which affect the optimization of spark timing and air/fuel ratio. Understanding the impact of varying gas composition on engine performance and emissions is, therefore, important.

Research Objectives

The overall goal of this research was to investigate, experimentally, the composition effects of three different simulated biogas fuels on engine performance and emissions compared to operation on natural gas baseline fuel. A parallel study to the work reported here focused on the combustion characteristics of biogas fuels [4]. The tasks undertaken to meet the performance and emissions research objectives were to

- develop data acquisition facilities to measure performance and emissions from the test engine and modify the test engine to increase its compression ratio from 8.3 to 11, which is closer to the compression ratios of large, stationary-power engines that operate at biogas production sites
- develop a fuel mixing setup to simulate dry landfill gas and digester gas composition by mixing carbon dioxide and nitrogen with pipeline natural gas
- conduct engine experiments with the baseline natural gas and the simulated biogas fuels and measure the effect on performance and emissions

Experimental Facilities and Procedure

The experimental facilities include the engine and accessories, the emissions monitoring equipment, and the data acquisition systems. A 1987 General Motors four-cylinder, four-stroke, 2.5 l, spark-ignited engine was used for this research. Although the stock compression ratio of this engine is 8.3:1, the compression ratio was increased to 11:1 by replacing the stock pistons with custom-made pistons. The detailed engine specifications are given

Contributed by the Internal Combustion Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, June 1999; final revision received by the ASME Headquarters, August 31, 1999. Editor: H. D. Nelson.

Table 1 Engine specifications

Engine Displacement	2.5 L
Number of Cylinders	4
Compression Ratio	11:1
Cylinder Bore	10.16 cm
Stroke	7.62 cm
Connecting Rod Length	15.37 cm
Displaced Volume Per Cylinder	625.0 cm^3
Clearance Volume Per Cylinder	85.6 cm^3
Intake Valve Head Diameter	4.37 cm
Exhaust Valve Head Diameter	3.81 cm

in Table 1. The original equipment manufacturer (OEM) ''rated power'' has not been included in Table 1 because the manufacturers rating for the stock 8.3 compression ratio is not relevant to the 11:1 ratio actually used. The OEM engine utilized exhaust gas recirculation, but this feature was defeated for this study. The accessories used included an electronic control module, a heads-up unit, a Himmelstein meter for measuring speed and torque, and a fuel mixing setup for mixing natural gas, N_2, and CO_2. The engine is equipped with an electronic control module (ECM) for controlling gasoline operation. The heads-up unit that interfaces with the ECM is used to adjust the spark timing manually. The heads-up unit is also used for changing the throttle position. Torque is measured by a noncontact strain gage torquemeter.

To simulate dried biogas, natural gas from the town gas line was mixed with nitrogen and carbon dioxide, which were stored in bottles. The mixture was introduced into the engine ahead of the throttle valve with adequate flow length for complete mixing [4]. To dispense refrigerated CO_2 from its vacuum-insulated bottle at the maximum rate of 230 l/min, the outside surface of the bottle was heated using two 1 kW strip heaters, and a heated pipeline (500 W) was installed between the bottle outlet and gas regulator. Brooks Instruments mass flow meters were used to measure flow rates of natural gas, N_2, and CO_2. These meters provided a 0–5 V analog signal to the data acquisition system.

Two different data acquisition systems were used: one for collecting the steady-state data (performance and emissions) and the other for collecting high-speed, in-cylinder pressure data [4]. Output voltages from measurement devices such as thermocouples, pressure sensors, gas analyzers, etc., were amplified to the 0–5 V range required by the analog-to-digital (A/D) board of the steady-state data acquisition system. Engine operating parameters, exhaust gas concentrations, and combustion stoichiometry were averaged and displayed during testing.

The engine was coupled to a General Electric cradle-type dc electric dynamometer capable of absorbing up to 75 kW (100 bhp). The dynamometer was used to motor and load the engine. Testing was essentially steady state. The following engine parameters were measured using the data acquisition system: engine speed and torque, intake manifold pressure, intake air temperature, coolant inlet and outlet temperatures, exhaust gas temperatures, natural gas flow rate, CO_2 flow rate, and N_2 flow rate. Air flow rate was determined by measuring pressure drop across a laminar flow element. Intake manifold pressure and the pressure drop across the throttle were also measured.

The emissions monitoring equipment used were carbon monoxide, carbon dioxide, oxides of nitrogen (NO_x), oxygen, and hydrocarbon analyzers. The instruments are mounted on a common rack that also holds zero and span gases and has air and sample drying capabilities. Exhaust hydrocarbons were measured as wet volume fractions using a Beckman Industrial Model 402 hydrocarbon analyzer equipped with a flame ionization detector. The hydrocarbon measurement instrumentation is incapable of distinguishing methane from other hydrocarbon species, and a nonmethane hydrocarbon analyzer was not available for these tests. A dried exhaust sample was passed through the remaining gas analyzers. Nondispersive infrared analyzers were used for measuring CO_2 (Rosemount Analytical Model 880) and CO (Infrared Industries Model IR-703). Oxides of nitrogen were measured as NO_x using a chemiluminescent analyzer (Rosemount Analytical Model 951A). Exhaust oxygen concentration was measured by a paramagnetic oxygen analyzer (Beckman Industrial Model 755A). Voltage outputs from all gas analyzers were fed to the emissions data acquisition system. A ''data point'' stored by the emissions data acquisition program consisted of computed averages of 30 rapidly acquired measurements.

Table 2 Fuel mixtures, stoichiometric air-to-fuel ratios and lower heating values

FUEL	COMPOSITION	Stoich A/F	LHV (Btu/lbm)
Baseline	Pipeline Natural Gas (NG)	17.02	20,800
Fuel 1	60% NG/40% CO_2	6.09	7,500
Fuel 2	75% NG/25% CO_2	9.00	11,100
Fuel 3	55% NG/35% CO_2/10% N_2	5.62	6,950

During this study the engine was operated using pipeline natural gas as a baseline fuel and the three mixtures of natural gas and diluents shown, along with lower heating value (LHV) and stoichiometric air-to-fuel ratio (A/F) information, in Table 2. A survey of many landfill and sewage digester gas compositions showed that the three mixtures used reasonably span the range of expected CO_2 (25–40 percent) and N_2 variation. Because it mostly consists of coal-seam methane produced from local wells, the town natural gas supply has a very high methane content, measured in excess of 98 mole percent in tests over the past several years. The balance is typically about 1.5 percent N_2, 0.2 percent CO_2, 0.1 percent O_2, 0.1 percent ethane, and the remaining trace constituents of less than 0.01 mole percent. Commercial bottled liquid CO_2 and gaseous N_2 were used.

The A/F was varied from stoichiometric to the lean misfire limit for two engine speeds (2350 and 2850 rpm), and two throttling positions (50 percent wide open throttle (WOT) and 100 percent WOT). The coefficient of variation of the indicated mean effective pressure, reported in Bell et al. [4], was used to quantify the degree of misfire. For each operating condition, the maximum brake torque (MBT) spark advance was determined. The engine was tested at three spark timings: MBT, retarded by 5 Deg, and retarded by 10 Deg from MBT timing. Performance, emissions, and in-cylinder pressure measurements were made and recorded. The fuel equivalence ratio Φ was displayed in near real time using measured exhaust emissions concentrations from the data acquisition system as input to the data analysis software. Variations in Φ were accomplished by manual adjustment of the fuel flow rate.

Reported fuel equivalence ratios were calculated from the measured concentrations of the exhaust constituents. A program developed by Thiagarajan et al. [5] was updated to calculate Φ for an arbitrary fuel represented by $C_nH_mO_rN_s$. The wet hydrocarbon (HC) and dry inorganic gas analysis method described by Heywood [6] was used to obtain equivalence ratio. Use of this program with gas analyzer inputs allowed continuous display and update of Φ. The validity of the stoichiometric equivalence ratio as calculated from exhaust emissions data was substantiated by the output voltage of the exhaust gas oxygen (EGO) sensor, which provides a clear indication of the transition from lean to fuel rich operation. The location of $\Phi=1$ as determined from emissions measurements agreed to within ± 1 percent of that indicated by the EGO sensor.

Results and Discussion

Experiments were conducted at 50 and 100 percent WOT at engine speeds of 2350 and 2850 rpm for stoichiometries ranging from $\Phi=1$ to near the lean misfire limit ($\Phi\cong 0.6$) and for three spark timings using baseline natural gas and the three simulated biogas mixtures. However, due to space limitations, the bulk of

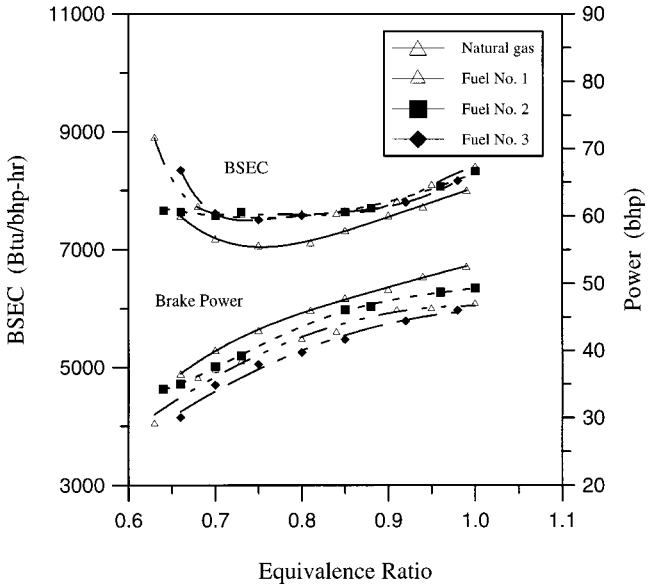


Fig. 1 Brake power and BSEC as a function of Φ for natural gas and fuels 1, 2, and 3 at 2350 rpm, 100 percent WOT, and MBT timing

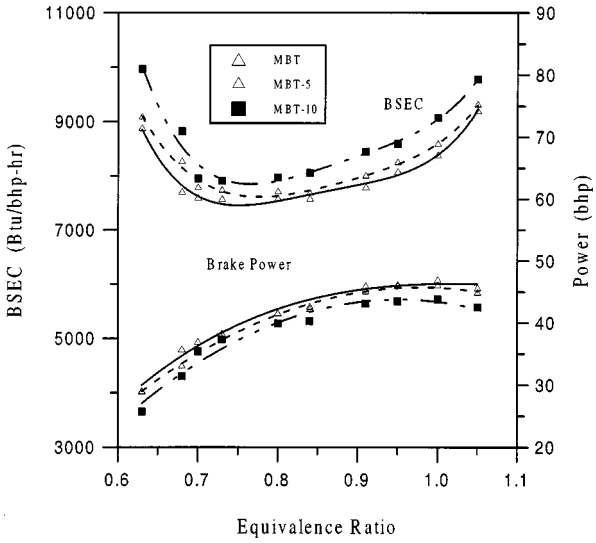


Fig. 2 Brake power and BSEC as a function of Φ for fuel 1 at 2350 rpm, 100 percent WOT, and MBT, MBT-5, and MBT-10 timing

the results presented here will be from the tests at 2350 rpm and 100 percent WOT. Significant variations in performance and emissions resulting from changes in throttle position and engine speed will be discussed where appropriate. The performance results that are reported include brake horsepower and brake specific energy consumption (BSEC). The emissions results reported include specific emissions of NO_x, unburned hydrocarbons, and CO.

Performance Results. Figure 1 shows the performance parameters, brake power, and BSEC as functions of fuel equivalence ratio for natural gas and fuels 1, 2, and 3 at MBT timing and 2350 rpm at 100 percent WOT. Brake power peaks at about 53 hp (about 40 kW) at $\Phi=1$ for baseline natural gas fueling. A peak power output of about 64 hp (about 48 kW) was measured for the 2850 rpm, 100 percent WOT case (not shown; see [4]). Brake power gradually declines as the air-to-fuel ratio is increased and a lower proportion of fuel is inducted into the cylinder. The loss of power as a function of decreasing Φ is clearly not due to declining thermal efficiency, because BSEC declines as Φ decreases (lower BSEC implies higher thermal efficiency or better fuel economy). Similarly, as the extent of diluent addition to the fuel increases at a fixed value of Φ, the brake power decreases. The presence of diluent, CO_2, or N_2 in a gaseous fuel reduces the volumetric heating value of the fuel and reduces the brake power output. As the heating value per unit volume of the fuel decreases, there is a corresponding decrease in the stoichiometric A/F on a volume (or mass) basis. As the stoichiometric A/F decreases at constant Φ, both the air charge and the fuel energy charge to the cylinder also decrease. Assuming that BSEC does not decrease dramatically, then the work output per cycle, or power, decreases as the diluent fraction increases. For fuel 1, the brake power output was reduced by 14 percent at $\Phi=1$ compared to natural gas.

An overall minimum BSEC of about 7000 Btu/bhp h (equivalent to a thermal efficiency of 36.4 percent based on LHV) is achieved for the baseline natural gas case at $\Phi\cong0.75$. BSEC is higher for the three simulated biogas fuels, where a flattened minimum BSEC of about 7500 Btu/bhp h is achieved over the $0.7<\Phi<0.8$ range. It was shown by Germane et al. [7] that thermal efficiency increases due to increased specific heat ratios, reduced dissociation losses, and reduced heat losses associated with leaner stoichiometries. Consequently, the BSEC is reduced until near the misfire limit, where incomplete combustion begins to cause an increase in BSEC. A significant difference in BSEC between the three simulated biogas fuels was not observed for the 2350 rpm, 100 percent WOT case shown here. BSEC values are higher for the 2850 rpm, 100 percent WOT case (not shown), as increased engine speed results in less complete combustion. For this higher speed case, BSEC shows greater dependence on the extent of diluent addition, with BSEC results for fuel 2 (~75 percent CH_4) lying above the natural gas results and below the fuel 1 (~60 percent CH_4) and fuel 3 (~55 percent CH_4) cases.

Figure 2 shows the measured performance parameters as a function of Φ for fuel 1 at 100 percent WOT and 2350 rpm at MBT, MBT-5, and MBT-10 spark timing. The brake power output varied between 46 bhp at $\Phi=1.03$ and 30 bhp at $\Phi=0.64$ for MBT timing. Karim and Wierzba [8] reported that the laminar burning velocity of a fuel mixture consisting of 60 percent methane and 40 percent CO_2 was only 60 percent that of pure methane. The reduced flame propagation rates increase the combustion period and ignition lag time, and reduce the burning rate. The authors concluded that advanced spark timing is effective in compensating for these detrimental effects. Spark timing, shown in Table 3, had to be advanced more for fuel 1 than for natural gas to achieve maximum brake torque, e.g., 10 Deg of additional spark advance were required to achieve MBT for fuel 1 at $\Phi=1$. Furthermore, as Φ decreased, increasingly larger degrees of spark advance were required. Due to limitations of the heads-up unit, a maximum spark advance of 60 Deg could be applied. It is observed from Fig. 2 that retarding the spark timing from MBT timing resulted in modest power and fuel economy reductions. At MBT-5, e.g., power declined by 2 percent and BSEC increased by 1 percent at $\Phi=1$. Compared to MBT timing, power output decreased by 3 percent and BSEC increased by 4 percent at $\Phi=1$ for

Table 3 MBT spark timing at 100 percent WOT and 2350 rpm

Fuel	Degree of Spark Advance			
	Φ = 1	Φ = 0.9	Φ = 0.8	Φ = 0.65
Natural Gas	30	34	39	55
Fuel 1	40	41	43	58
Fuel 2	38	39	42	60
Fuel 3	41	45	50	60

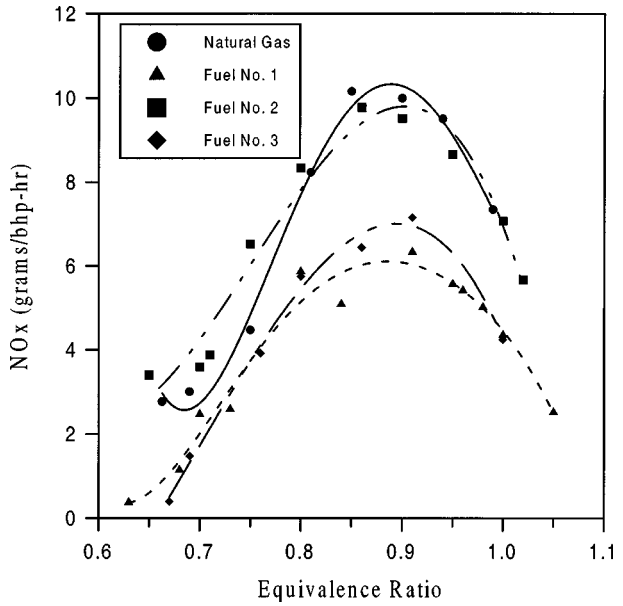


Fig. 3 Brake specific NO_x emissions as a function of Φ for natural gas and fuels 1, 2, and 3 at 2350 rpm, 100 percent WOT, and MBT timing

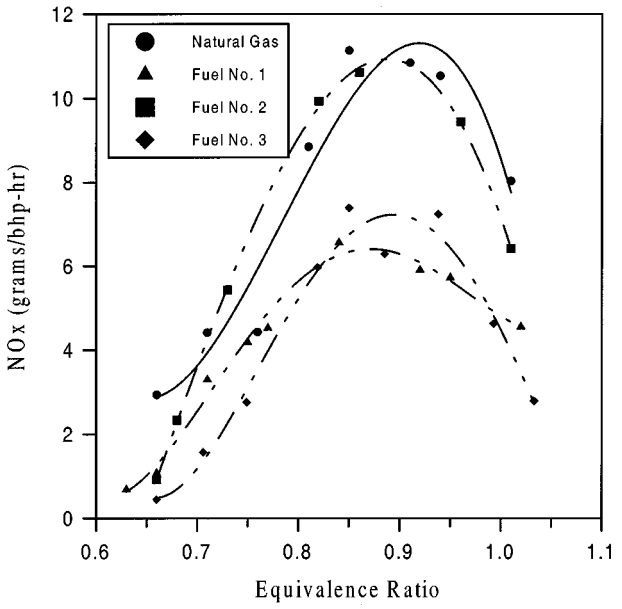


Fig. 4 Brake specific NO_x emissions as a function of Φ for natural gas and fuels 1, 2, and 3 at 2850 rpm, 100 percent WOT, and MBT timing

MBT-10 timing. The spark timing advance required to achieve MBT was even greater for the 2850 rpm cases (not shown) than for the 2350 rpm cases.

NO_x Emissions Results. Figure 3 shows brake specific NO_x emissions as a function of Φ for natural gas and fuels 1, 2, and 3 at MBT timing and 2350 rpm at 100 percent WOT. The polynomial curves fit to the experimental data are to aid the eye in following data trends; these curves are not based on the physical phenomena that determine species concentrations. For all four fuels NO_x emissions peak near $\Phi=0.88$. NO_x formation in engine combustion is governed principally by the Zeldovich mechanism (thermal NO_x). Average combustion temperatures have been estimated by a single-zone thermodynamic combustion model that uses in-cylinder pressures measured as a function of crank angle [4]. Representative results of the heat release model using measured in-cylinder pressures as input are presented in Table 4 for the MBT timing, 2350 rpm, 100 percent WOT case. Phase I is the period of ignition and flame development. Phase II is the principal flame propagation/combustion period, wherein peak temperatures are achieved. Estimates of average phase II temperatures closely parallel the dependence of NO_x levels on Φ. NO_x emissions decline in both directions away from the peak, but the decline is more pronounced in the direction of lower Φ. This same trend is seen in detailed model results for average phase II temperature.

Table 4 Selected heat release model results for MBT timing, 2350 rpm, and 100 percent WOT

Fuel	Φ	Phase I duration (degrees)	Phase I Avg. T (K)	Phase II duration (degrees)	Phase II Avg. T (K)
NG	1.0	30	830	32	1813
NG	0.85	36	833	34	1830
NG	0.7	49	794	39	1725
1	1.0	32	760	43	1610
1	0.8	38	730	61	1625
2	1.0	27	798	42	1690
2	0.8	38	832	43	1716
3	1.0	31	711	39	1507
3	0.8	40	746	55	1570

The effect of fuel composition on NO_x emissions is substantial. NO_x emissions peak at about 10 g/bhp h (1 g/bhp h=1.34 g/kW) for the baseline natural gas case and appear slightly lower for fuel 2 (although the difference is probably not statistically significant). Fuel 2 has the least diluent addition and has the highest estimated average phase II temperature (Table 4) of the three simulated biogases. Large reductions in NO_x emissions for all Φ are observed for fuels 1 and 3, with peak emissions of about 6 g/bhp h. Although the effect is small and possibly insignificant, it appears that NO_x emissions are lower for fuel 1 than for fuel 3. This result is similar in Fig. 4, which shows brake specific NO_x emissions as a function of fuel equivalence ratio for natural gas and fuels 1, 2, and 3 at MBT timing and 2850 rpm at 100 percent WOT. For this higher engine speed there is clearer separation in the NO_x emissions near $\Phi=0.9$, with natural gas showing the highest NO_x emissions, followed in reducing order by fuel 2, fuel 3, and then fuel 1.

Although the total diluent proportion of fuel 3 (55 percent CH_4, 35 percent CO_2, and 10 percent N_2) exceeds that of fuel 1 (60 percent CH_4, 40 percent CO_2), the fact that CO_2 has a substantially higher molar specific heat may result in a greater temperature depression effect for fuel 1. A simple estimation of the maximum temperature elevation can be calculated by equating the sensible heating of the fuel/air mixture (ideal gas) to the heat release of the lower heating value of the fuel. Rathnam [9] reports such calculations using specific heats (evaluated at combustion temperatures) of the various fuel/air mixtures tested. These results show that at the same Φ, combustion temperature rises are estimated to be reduced by 5–10 percent for fuels 1–3 compared to straight natural gas as a result of specific heat and diluent fraction variation. Contrary to the results of these simple estimates of the temperature effects of varying diluent levels and specific heats, the phase II temperatures predicted by the heat release model are somewhat higher for fuel 1 than for fuel 3. In summary, for the same Φ, NO_x emissions appear to be higher for fuel 3 than for fuel 1, a trend that agrees with temperature estimates based on mixture specific heats but disagrees with heat release model estimates of average phase II temperatures based on measured P–V data. A heat release model more accurate than the single-zone heat release model used here is needed to predict temperature with

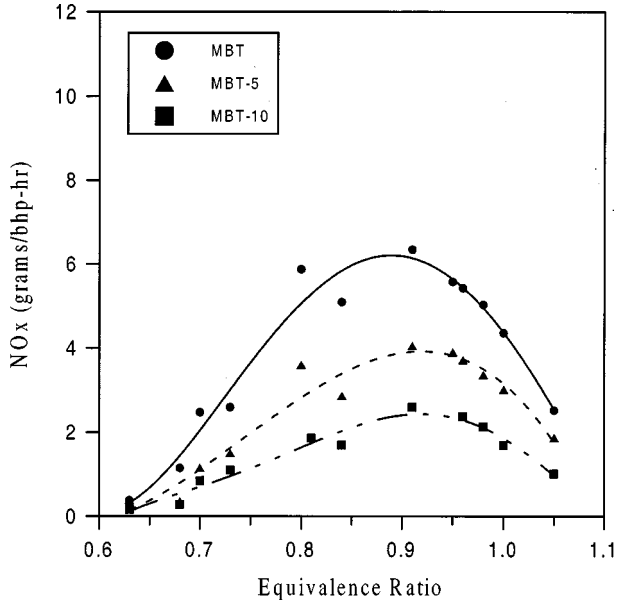


Fig. 5 Brake specific NO_x emissions as a function of Φ for fuel 1 at 2350 rpm, 100 percent WOT, and MBT, MBT-5, and MBT-10 timing

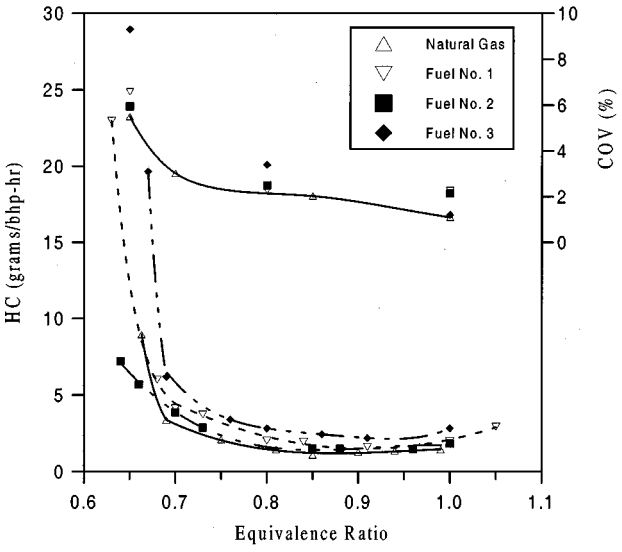


Fig. 6 Brake specific hydrocarbon emissions and coefficients of variation (COV) as a function of Φ for natural gas and fuels 1, 2, and 3 at 2350 rpm, 100 percent WOT, and MBT timing

sufficient accuracy to resolve NO_x emissions differences for relatively minor variations in fuel mixtures or operating parameters.

Comparing Fig. 4 to Fig. 3, the difference in NO_x emission levels attributable to engine speed is insignificant. Comparison of the 100 percent WOT cases to the 50 percent WOT cases (not shown) also reveals negligible differences in NO_x emissions.

Figure 5 presents measured specific NO_x emissions as a function of fuel equivalence ratio using fuel 1 at 2350 rpm and 100 percent WOT with MBT, MBT-5, and MBT-10 spark timing. Even these modest levels of spark timing delay result in substantial NO_x reductions. A 5 Deg reduction in spark advance yields a NO_x peak of about 4 g/bhp h compared to 6 g/bhp h for MBT timing, and a 10 Deg reduction results in a NO_x peak of about 2 g/bhp h. A 5 Deg spark advance reduction combined with lean operation of $\Phi<0.75$ also yields NO_x emissions of less than 2 g/bhp h. Retarding spark timing decreases in-cylinder gas pressures [6], resulting in lower peak burned gas temperatures, thus, lower NO_x emissions at retarded spark timing.

Unburned Hydrocarbon Emissions Results. Figure 6 shows the specific hydrocarbon emissions as a function of Φ for natural gas and fuels 1–3 at 2350 rpm and 100 percent WOT at MBT. Hydrocarbon emissions for baseline natural gas fueling ranged from 1.5 g/bhp h at $\Phi=0.8$ to 2 g/bhp h at $\Phi=1$. The decrease in HC emissions for $\Phi<1$ is due to the presence of excess oxygen that allows more complete combustion of the fuel. As the fuel equivalence ratio was leaned below $\Phi=0.8$, a gradual increase in hydrocarbon emissions was observed. This is in part due to lower in-cylinder temperatures, which lead to larger quench volumes near the cylinder wall. In addition, the slower flame speed associated with the lean condition results in increasingly incomplete combustion. At about $\Phi=0.65$, near the lean misfire limit, hydrocarbon emissions exceed 10 g/bhp h.

Emissions of unburned hydrocarbons for fuel 1 ranged from 2.3 g/bhp h at $\Phi=0.8$ to 2.5 g/bhp h at $\Phi=1$ for MBT timing. Hydrocarbon emissions for fuel 1 are higher than for natural gas throughout the fuel equivalence ratio range tested ($0.65<\Phi<1.05$). In addition to causing increased hydrocarbons, the addition of CO_2 also shifts the lean misfire limit toward higher Φ. Fuel 2 hydrocarbon emissions were closest to the baseline natural gas case of the three simulated biogas fuels, and fuel 3 emissions were highest. The difference in hydrocarbons emissions shown in Fig. 6 compared to those measured for the other engine speed/throttle position cases was insignificant. Retarded spark timing had a negligible effect on hydrocarbon emissions for all fuels tested at all engine speed/throttle position cases (effect of timing data not shown). Slightly higher lean limits (Φ values) were observed as the spark timing was retarded from MBT to MBT-10.

The relevance of exhaust temperatures to the extent of hydrocarbon burnup as the combustion gases exit the engine is well known (e.g., [6]), with higher exhaust temperatures resulting in reduced emissions. Figure 7 shows exhaust temperatures as a function of fuel equivalence ratio for natural gas and fuels 1, 2, and 3 at 100 percent WOT and 2350 rpm at MBT timing. Although the temperature data are bunched together at the two extremes of Φ examined, in the midrange of Φ, where combustion is

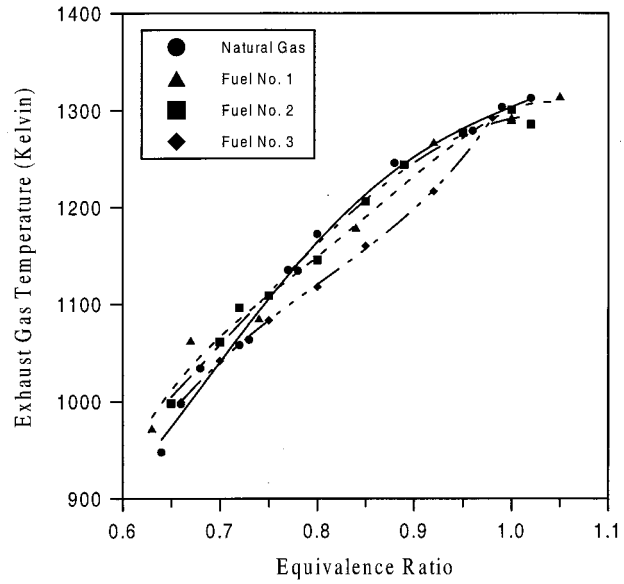


Fig. 7 Exhaust temperatures as a function of Φ for natural gas and fuels 1, 2, and 3 at 2350 rpm, 100 percent WOT, and MBT timing

generally complete for all four fuels studied, the exhaust temperatures, as expected, are inversely related to the level of unburned hydrocarbon emissions (shown in Fig. 6). At extremely low values of Φ incomplete combustion in the cylinder increases, particularly for fuels with larger diluent content. Because in-cylinder combustion is less complete for the high diluent cases, combustion in the exhaust is more prevalent than for the straight natural gas case. Consequently, both higher exhaust temperatures and higher unburned hydrocarbons by comparison to natural gas were measured for the high diluent content mixtures (fuels 1 and 3) at very low Φ.

Figure 6 also shows the coefficient of variation (COV) as a function of Φ for natural gas, and fuels 1, 2, and 3. The COV is the ratio of the standard deviation of indicated mean effective pressure (IMEP) to the IMEP itself (expressed here as a percentage). Values of IMEP were computed using measured in-cylinder pressures [4]. The COV remained below 2 percent for natural gas and below 4 percent for fuels 1, 2, and 3 for $0.7<\Phi<1$, but rose sharply as the lean misfire limits were approached. It is observed from the COV results that as diluent concentration increases, the flammability range is reduced and the lean misfire limit is shifted towards higher fuel equivalence ratio. COV parallels HC emissions for each fuel.

CO Emissions Results. Figure 8 shows measured specific CO emissions as a function of Φ for natural gas and fuels 1, 2, and 3 at 100 percent WOT, 2350 rpm, and MBT timing. CO emissions for natural gas fueling were level at about 2 g/bhp h between $\Phi=0.65$ and $\Phi=0.95$ and then rose sharply to exceed 20 g/bhp h at $\Phi=1$ as oxygen became scarce. The increase in CO emissions occurs before the stoichiometric condition. The specific emissions of CO (mass of emissions per unit power) increase slowly moving toward leaner stoichiometries. This occurs even though CO production is essentially constant because BSEC increases slightly at reduced brake power, so specific CO emissions increase slightly.

CO emissions are greater by approximately 10 percent between $\Phi=0.65$ and $\Phi=0.9$ for fuel 1 compared to natural gas. On the other hand, because of its high CO_2 content, CO_2 emissions for fuel 1 were almost 80 percent greater than CO_2 emissions for natural gas. This implies that most of the fuel CO_2 remains stable or quickly reforms during combustion, so that high exhaust emissions of CO do not result from substantial increases in fuel CO_2. Further evidence that CO_2 dissociation is not the cause of increased CO emissions for the high diluent fuels comes from estimations of average temperatures during the primary combustion period (phase II) reported by Bell et al. [4] and shown in Table 4 for the experiments described here. The average phase II temperatures, calculated from a single-zone thermodynamic model using in-cylinder pressure measurements, are highest for straight natural gas and decline substantially for increasing diluent fraction. If CO_2 dissociation was the predominant mode of CO formation, then CO emissions should decrease with increasing diluent addition. In fact, CO emissions increase with increasing diluent fraction, primarily because of decreased exhaust temperatures [6]. This contention is supported by the fact that CO emissions were highest for fuel 3, which experienced the lowest exhaust temperatures (Fig. 7). CO emissions are increased by an average of 16 percent for fuel 3 compared to natural gas. For all four fuels tested, the effect of spark timing on CO emissions was minimal. Similarly, the effects of throttle position and engine speed were insignificant on CO emissions for stoichiometries ranging from $0.7<\Phi<0.9$.

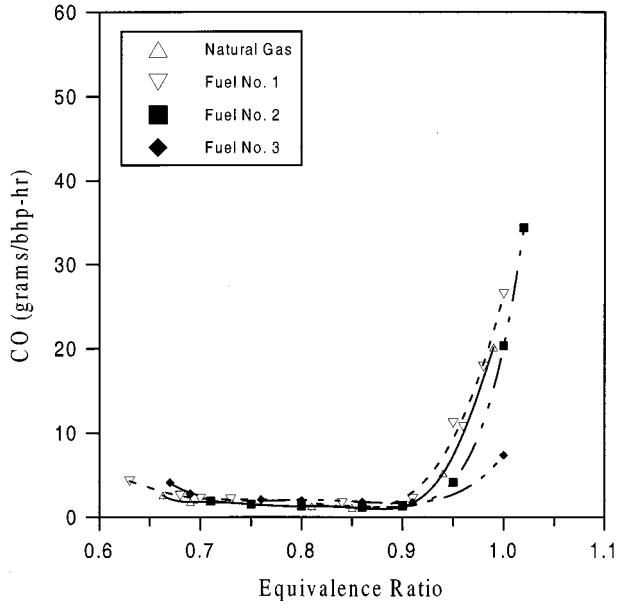


Fig. 8 Brake specific CO emissions as a function of Φ for natural gas and fuels 1, 2, and 3 at 2350 rpm, 100 percent WOT, and MBT timing

Summary and Conclusions

The primary objective of this study was to determine the effect of biogas fuel composition on emissions. Three simulated biogas fuels were tested in a conventional, automotive spark-ignited engine (with 11:1 compression ratio) and the results compared to baseline natural gas fueling. A secondary objective was to evaluate fuel effects on engine performance. The results of the experimental studies are summarized here.

A power decrease of 30–35 percent was observed as the equivalence ratio was reduced from stoichiometric to near the lean misfire limit. The decrease in power due to CO_2 addition at stoichiometric operation was found to be nearly linear. A 13–14 percent power loss was observed when the CO_2 content was increased to 40 percent by volume. The decrease in power due to CO_2 addition decreased as Φ was decreased. A slight power loss was observed when N_2 content was increased to 10 percent. BSEC was lowest between $\Phi=0.8$ and $\Phi=0.7$ for all operating cases. BSEC increased as the CO_2 content increased in the fuel, but the increase was nonlinear. The BSEC did not change significantly with the addition of N_2 between $\Phi=1.0$ and $\Phi=0.75$, but below $\Phi=0.75$ the BSEC increased with added N_2.

NO_x emissions were greatly reduced for all biogas fuels compared to natural gas because of lower peak temperatures resulting from diluent addition. Peak NO_x emissions, which occurred in the $0.85<\Phi<0.9$ range, were reduced by approximately 50 percent for fuel 1 compared to natural gas. Significant reductions in NO_x emissions also resulted from retarded spark timing throughout the range of Φ tested for all fuels (10–30 percent peak reductions for fuel 1 at MBT-5 timing). The measured exhaust temperature for fuel 3 was lowest of all fuel cases for all Φ, which resulted in the highest hydrocarbon and CO emissions of all of the fueling cases. Increased levels of CO appeared to be associated more with low exhaust temperatures (incomplete combustion) than with dissociation of fuel CO_2.

These research findings are useful in understanding the effect on both performance and emissions of the diluent content of a biogas. Compared to natural gas, there is a maximum reduction in brake power output of 15 percent at $\Phi=0.7$ for the maximum diluent addition of 45 percent (fuel 3). This power loss is accompanied by a significant (40 percent) reduction in NO_x, and smaller increases of 10 percent more CO and 30 percent more hydrocarbons. The hydrocarbons emissions consist primarily of methane, a nonreactive hydrocarbon that does not participate in reactions leading to photochemical smog. Further reductions in NO_x can be achieved by a slight delay in spark timing. For example, at $\Phi=0.75$, it was observed that retarding the spark timing by 5 Deg resulted in a 20–40 percent reduction in NO_x, but only about a 5 percent reduction in power output. Increased power output and combustion stability at lean fuel equivalence ratios can be ob-

tained by redesigning the fuel intake system, ignition system, etc., to accommodate biogas, and by using turbocharging.

Acknowledgments

The efforts of engineers Frank Newman, Ray Burkhalter, and Tim Patterson are gratefully acknowledged.

References

[1] McBean, A. E., Rovers, A. F., and Farquhar, J. G., 1995, *Solid Waste Landfill Engineering and Design*, Prentice-Hall, Englewood Cliffs, NJ.
[2] Christensen et al., 1989.
[3] Brown, A. K., and Maunder, H. D., 1994, ''Using Landfill Gas: A UK Perspective,'' Renewable Energy, **5**, pp. 774–781.
[4] Bell, S. R., Midkiff, K. C., Bhargava, S., and Rathnam, S., 1997, ''An Investigation of Low-Btu Gas Combustion in a Spark-Ignited Engine,'' *Natural Gas Engines*, ICE-Vol. 28-2, ASME, New York, pp. 57–64.
[5] Thiagarajan, S., Midkiff, K. C., Bell, S. R., and Green, M. N., 1995, ''Investigation of Fuel Composition Effects on a Natural Gas Fueled Spark-Ignited Engine,'' *Natural Gas and Alternative Fuels for Engines*, ICE-Vol. 24, ASME, New York, pp. 41–51.
[6] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.
[7] Germane, G. J., Wood, C. G., and Hess, C. C., 1983, ''Lean Combustion in Spark-Ignited Internal Combustion Engines—A Review,'' SAE Paper 831694.
[8] Karim, G. A., and Wierzba, I., 1992, ''Methane-Carbon Dioxide Mixtures as a Fuel,'' SAE Paper 921557.
[9] Rathnam, S., 1996, ''Emissions Characteristics of Biogas Combustion in a Spark-Ignited Engine,'' M.S. thesis, University of Alabama, Tuscaloosa, AL.

In-Cylinder Tumble Flow Field Measurements and Predictions

C.-W. Hong

S.-D. Tarng

Department of Power Mechanical Engineering,
National Tsing Hua University,
Hsinchu 30013, Taiwan

This paper presents the comparison between measured and predicted results of the in-cylinder tumble flow generated by a port-valve-liner assembly on a steady-flow test bench. The purpose was to advance the understanding of the stationary turbulence process via experimental and computational techniques in the same time. A baseline single-cylinder 4-stroke motorcycle engine was chosen. Its liner was replaced by a transparent acrylic-plastic tube and the piston was removed. This was to focus the research on the tumble flow generated by the geometry of its port, the passage of the canted inlet valve, and a dome-shaped combustion chamber. The in-cylinder turbulent flow field was measured via a 3-component laser Doppler velocimeter (LDV) point by point. A simultaneous computer simulation was carried out to predict the in-cylinder flow field of the same engine under the same operating condition, using KIVA3V—the most recent version of the KIVA code. The mean speed, turbulence intensity, tumble ratio, swirl ratio, and vortex circulation from both skills were all compared. A reasonably good level of agreement has been achieved. Both modern techniques are also validated. [DOI: 10.1115/1.1335479]

Introduction

It is an un-ended target for an engine researcher to further promote the performance, as well as to further reduce the emission, of existing engines. Controlling the in-cylinder flow pattern is a modern and effective technique to accomplish the target, especially in direct injection spark ignition (DISI) engines [1,2]. Evaluation of successful implementation of this technique depends on the improvement of the modern design tools, such as computational fluid dynamics (CFD) simulation and laser diagnostic experiments. Computer animation and/or experimental flow visualization of the invisible flow field inside the engine play the key role during the concept design stage. In these two methods, CFD is more convenient and more versatile to study the flow physics. However, its accuracy in terms of quantity and quality is always a question to the engine designer. To answer this suspicion, comparisons between the experimental and simulation result under the same operating condition were conducted. This was able to evaluate both skills and hope that the numerical test bench could replace the expensive experiments sometime in the future.

Extensive investigations have been reported on the subject of validation of the CFD codes with the flow field measurements of internal combustion engines [3–9]. Among the existing codes, the KIVA series are most favored by the academics because the available physical sub-models and numerical schemes. Users are able to modify the source code for their own special purpose. The research reported in this paper focuses on the in-cylinder tumble flow generated from the engine intake geometry and the shape of the cylinder head design. Simultaneous flow field measurements were also carried out using the laser Doppler velocimetry (LDV). Although the LDV technique is laborious, it is still the most accurate instrument in turbulent flow measurements.

This paper also intends to provide an evaluation of the updated KIVA3V code [10] by cross exam between the simulation and measurement results under the same operating conditions. Evaluation of the flow field parameters, including the mean speed, turbulence kinetic energy, vortex circulation, and overall tumble and swirl ratios, are carried out. The comparison work and validation results are described in the following sections.

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division April 26, 1999; final revision received by the ASME Headquarters July 27, 1999. Technical Editor: D. Assanis.

Experimental Apparatus and Test Procedure

A steady flow test rig was built up to measure the in-cylinder flow of a dismantled cylinder head from a 4-stroke motorcycle engine. Specifications of the engine are listed in Table 1. The schematic diagram of the experimental apparatus is illustrated in Fig. 1. The test rig basically consists of three major parts. The first one is an orifice meter to measure the overall air flow rate. The dimension of the pipe line system was designed according to the

Table 1 Engine specifications

Model	YAMAHA Fuzzy 125	
Number of cylinders	1	
Bore	51.5	mm
Stroke	60.0	mm
Displacement	125	c.c.
Compression ratio	10.3	
Inlet Port Diameter	20	mm
Inlet Valve Geometry		
Valve Diameter	25.5	mm
Stem Diameter	5	mm
Maximum Lift	7	mm
Seat Angle	30	degree
Inclination	30	degree

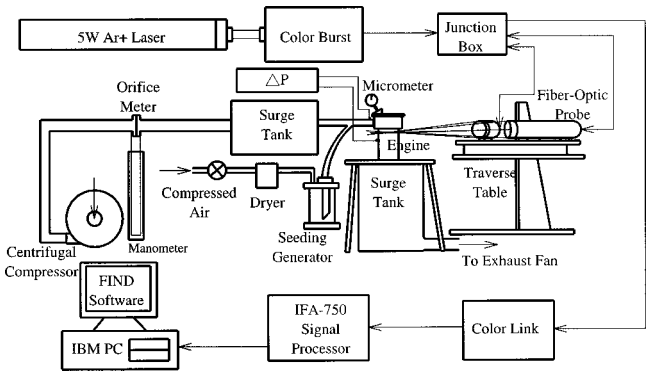


Fig. 1 Experimental apparatus

ISO 5167 standard. The combination of a centrifugal compressor upstream and an extraction fan downstream of the test rig was used to adjust the pressure difference across the inlet valve. The second part is the engine cylinder head with a dome combustion chamber and two canted valves. The liner has been replaced by a transparent acrylic-plastic tube for optical access. The piston is removed, hence, the bulk flow is kept steady but with stationary turbulence. The air flow rate is mainly controlled by the intake valve lift and the pressure difference across the valve. A micrometer shown in the figure was used to measure the valve position. The third major part is an LDV system which consists of a 5W Argon laser and two back-scattered fiber-optic probes. One probe creates four beams in green and blue, and the other probe transmits two beams in purple. Two probes are placed perpendicularly to each other, hence, the instantaneous particle velocity in three spatial components can be detected simultaneously. Powdered MgO particles were planted into the flow passage by a seeding generator. The scattered Doppler signals were processed by a digital burst correlator, TSI IFA 750, and then transferred into velocity data in an IBM PC.

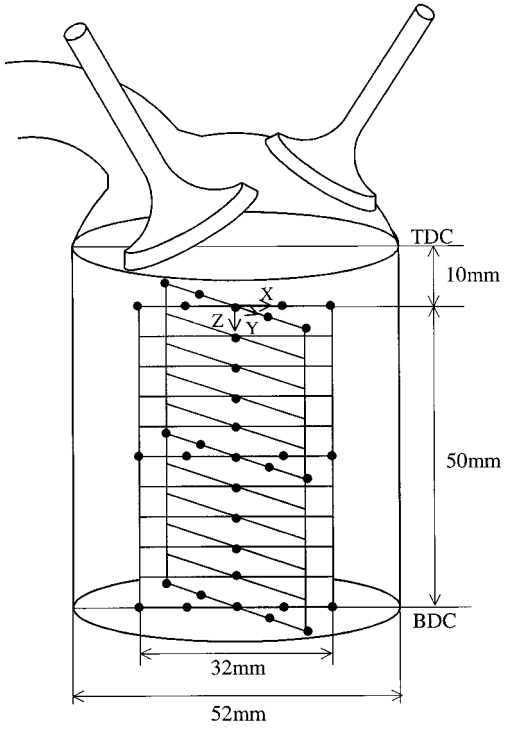


Fig. 2 LDV measurement locations (11 horizontal planes, 9 points per plane, 99 in total)

Figure 2 shows the measurement locations chosen to pin-point the in-cylinder flow field. There are eleven horizontal planes separated from each other by 5 mm distance. The top plane is 10 mm beneath the top-dead-center (TDC) surface, and the bottom plane is located at the bottom-dead-center (BDC). In each horizontal plane, namely XY plane in the figure, nine radial positions, including the center, are chosen. They are distributed in two orthogonal vertical planes, XZ and YZ, which are the tumble planes we are concerned. The total is 55 measurement points in each vertical plane, hence, the number of grid points in the whole cylinder is 99. Note that the distance from the nearest measurement location to the wall boundary is 10 mm. This was constrained by the distortion lights due to the optical path through the cylindrical wall. The steady flow test was preset under the condition of pressure difference at 3 cm water column and the valve position at the maximum (7 mm). The measured inlet pressure was 0.9867 atm, the outlet pressure was 0.9846 atm, so the actual pressure difference was 0.0021 atm. Note that the upstream pressure was measured at the location of the inlet port 50 mm upstream of the inlet valve. The downstream pressure was measured at the location of the BDC of the cylinder. These two pressures and positions were later input to the simulation code in the section of computer simulation.

Experimental Analysis and Results

For stationary turbulence measurements, the instantaneous velocity, e.g., in X component, at a specific position and at an instant, t, is represented by

$$U(t)=\overline{U}+u(t), \tag{1}$$

where $\overline{U}$ is the time average (or ensemble average) mean velocity, calculated by

$$\overline{U}=\lim_{\tau\to\infty}\left[\frac{1}{\tau}\int_{t_0}^{t_0+\tau}U(t)dt\right]. \tag{2}$$

The turbulence velocity, $u(t)$, is normally transferred to a root mean square (RMS) value, named turbulence intensity, u', to represent its random fluctuation. This is given mathematically by

$$u'=\lim_{\tau\to\infty}\left[\frac{1}{\tau}\int_{t_0}^{t_0+\tau}u(t)^2\right]^{1/2}. \tag{3}$$

It can also be represented by turbulence kinetic energy, κ, as

$$\kappa=\frac{1}{2}(u'^2+v'^2+w'^2), \tag{4}$$

where u',v',w' are turbulence intensities evaluated from X, Y, Z components, respectively.

Figure 3(*a*) shows the measured result of the flow pattern by means of the velocity vector diagram in XZ plane. This is the principal tumble motion we are concerned. Note that there exists a secondary tumble flow in the region of the left hand side of the figure along the liner wall boundary. The velocity vectors are distorted at the interface of these two tumble jets. Figure 3(*b*) shows that in the YZ plane, a pair of counter vortices are generated. Apparently, the tumble motion creates two re-circulation zones beneath the inlet valve. Figure 4 displays the top down view of the XY planes in four locations. The swirl flow doesn't appear in this engine because there is no inclined angle of the swirl axis in the intake system design.

The tumble motion is significant in both XZ and YZ planes from the above flow field diagrams. To quantify the tumble motion, tumble ratio as well as swirl ratio are general engineering terms to represent the overall flow characteristics. They are defined as organized rotations of the charge motion around specific cylinder axes, such as X, Y, and Z. The mathematical definitions of the tumble ratio, TR, and the swirl ratio, SR, of the flow field are evaluated by

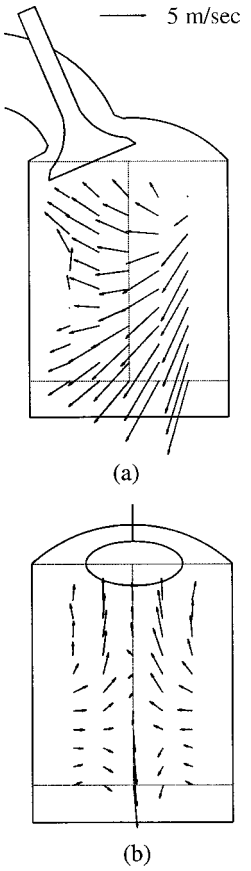


Fig. 3 LDV measurements of the tumble flow in (*a*) the XZ plane and (*b*) the YZ plane

$$TR_x=\frac{\sum_{i=1}^{n}(H_x)_i}{\sum_{i=1}^{n}(I_{xx})_i\omega},\quad TR_y=\frac{\sum_{i=1}^{n}(H_y)_i}{\sum_{i=1}^{n}(I_{yy})_i\omega},\quad SR_z=\frac{\sum_{i=1}^{n}(H_z)_i}{\sum_{i=1}^{n}(I_{zz})_i\omega}, \tag{5}$$

where H is the angular momentum per unit mass, I is the moment inertia per unit mass, and ω is the crank speed in radian per second. The angular momentum can be represented by a vector form as

$$\vec{H}=H_x\vec{i}+H_y\vec{j}+H_z\vec{k}=\vec{r}\times(\overline{U}\vec{i}+\overline{V}\vec{j}+\overline{W}\vec{k}). \tag{6}$$

The position vector $\vec{r}$ is a vector from the origin to the measurement location

$$\vec{r}=x\vec{i}+y\vec{j}+z\vec{k} \tag{7}$$

and $\overline{U},\overline{V},\overline{W}$ are mean velocities along the X, Y, Z axes respectively. The moment inertia per unit mass in three components are given by

$$I_{xx}=y^2+z^2,\quad I_{yy}=x^2+z^2,\quad I_{zz}=x^2+y^2. \tag{8}$$

The tumble and swirl ratios are normalized with respect to the crank speed, ω, which is calculated by

$$\omega=\frac{2\pi N}{60}, \tag{9}$$

where N is the equivalent engine speed in rpm.

The mathematical definition of the tumble ratio may mislead the quantitative characteristics of the rotational flow if there exists a pair of counter rotating vortices with the similar strength. To provide a means of evaluating both large and small cell motions of the vortex motion, fluid circulation is a better mathematical concept to represent the physical phenomenon. It is defined by fluid flowing about an enclosed contour or related to the flux of vorticity via Stokes theorem [11]. In a single rectangular mesh, the circulation with respect to the X-axis within the cell is calculated by

$$\Gamma_x=\sum_{i=1}^{4}\left(\left(\left(\frac{\partial\overline{W}}{\partial y}\right)_i-\left(\frac{\partial\overline{V}}{\partial z}\right)_i\right)\frac{A_i}{4}\right)_x, \tag{10}$$

where A_i is the rectangular area of the mesh. The similar definitions can be applied to the circulations w.r.t. Y and Z-axes. Normally, a root mean square of the vortex circulation is used to represent the overall quantity of the rotation motion for the whole plane, i.e.,

$$\Gamma_{rms}=\sqrt{\frac{\sum_{m=1}^{n}(\Gamma_m)^2}{n}}, \tag{11}$$

and n is the total number of small cells within the whole area. The resulting circulation is always a positive number, it an be used to quantify the bulk and/or net motion of the fluid both in large and small cells. It also removes the cancellation effect generated by counter rotating vortex structure.

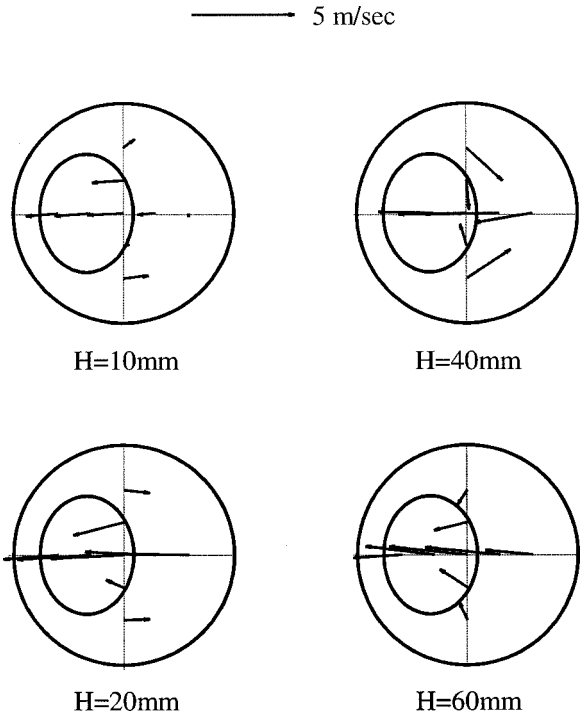


Fig. 4 LDV measurements of the non-swirl flow in XY planes (H stands for height beneath the TDC)

Computer Simulation

Computer simulation of the steady flow test for the same engine was conducted on an IBM SP2 workstation using a updated package, KIVA3V [10]. This simulation package was developed by Los Alomos Lab, aiming at flow study of the reciprocating engine with vertical or canted valves. The code uses a block-structure mesh with connectivity defined through indirect addressing. Hence, complex geometry of cylinder head, intake and exhaust ports can be adapted together to form a complete set of grids which emulates the flow passage of a realistic engine. Compared

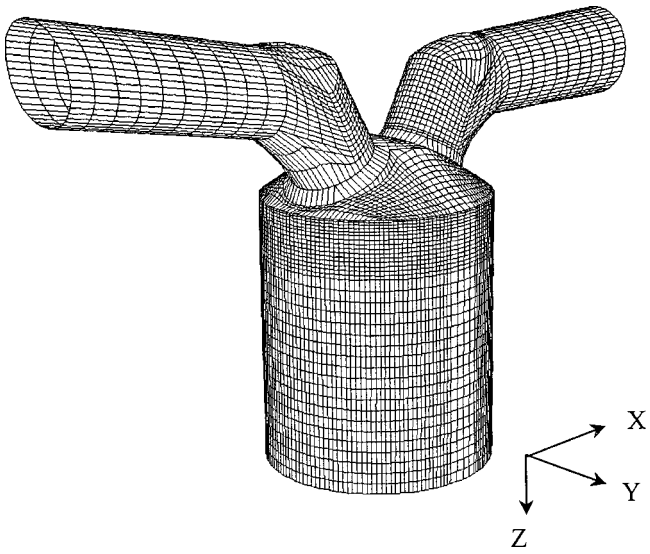


Fig. 5 Computational grids at BDC of the port-valve-liner assembly with dome-shaped combustion chamber and two canted valve seats

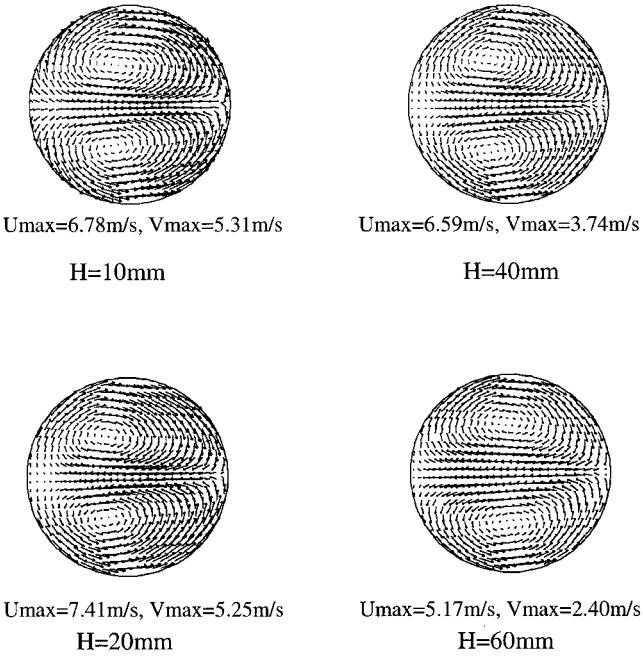


Fig. 7 Computed velocity vector diagrams in XY planes, which are located at 10 mm, 20 mm, 40 mm, and 60 mm beneath TDC, respectively

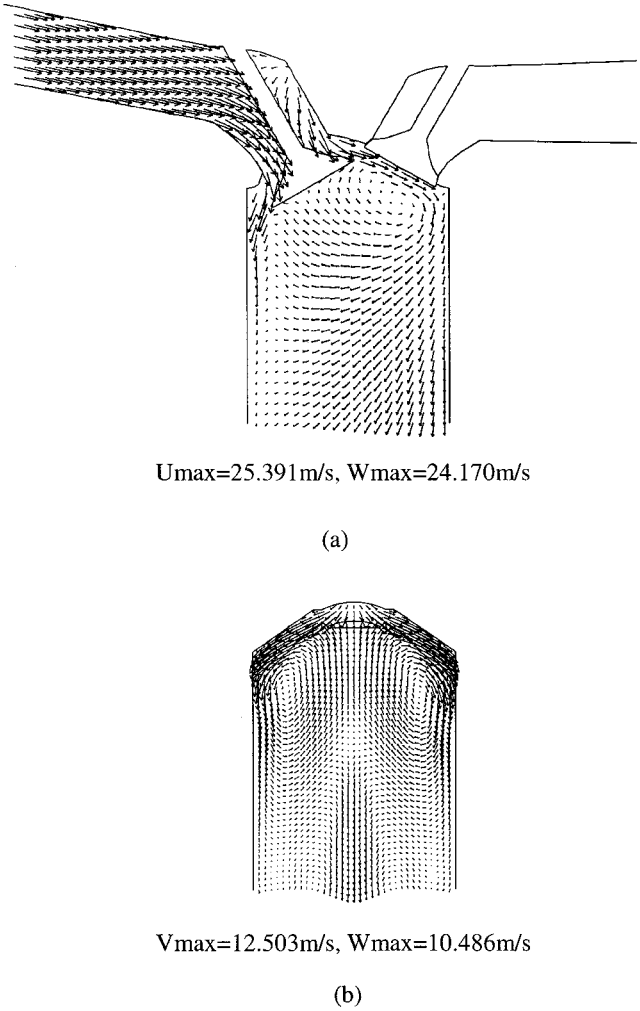


Fig. 6 Computed velocity vector diagrams in (*a*) the XZ plane and (*b*) the YZ plane

with its predecessor, the new version has the advantage of inputting a wedge type combustion chamber instead of pancake type in KIVA-3 [12]. However, it is still not readily set up to simulate the dome type combustion chamber, such as the baseline engine in this paper. By modifying the source code in grid generation, we implement a spherical equation to fit the dome chamber. This is given, in mm, by

$$x^2+y^2+(z-37.2)^2=134.4^2. \tag{12}$$

Steady-state computation was performed. The piston was removed and replaced by an outflow boundary. The valve position and the virtual piston position were kept constant in the simulation. Reshaping of the grid structure near the interface between the valve seat and the dome cylinder head was performed by the program itself. Figure 5 shows the final product of grid generation

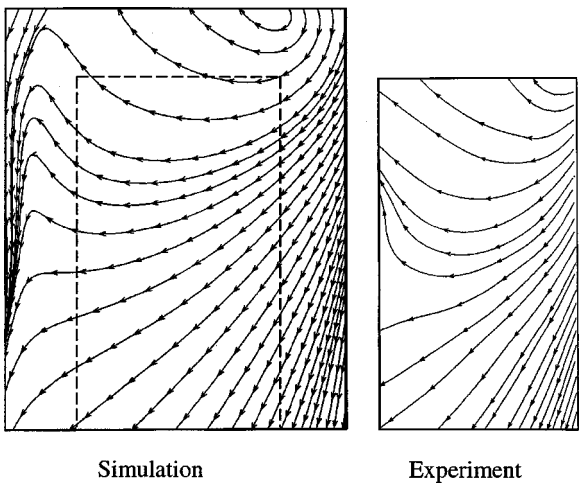


Fig. 8 Comparison of simulation and experimental results represented by streamlines on the XZ plane (the rectangle enclosed by dashed lines is the measurement area)

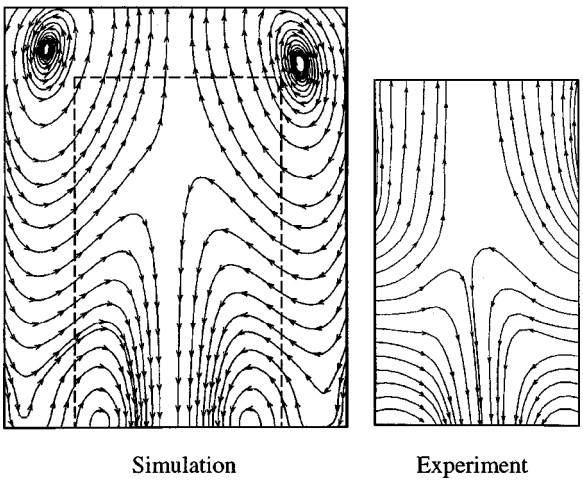


Fig. 9 Comparison of simulation and experimental results represented by streamlines on the YZ plane (the rectangle enclosed by dashed lines is the measurement area)

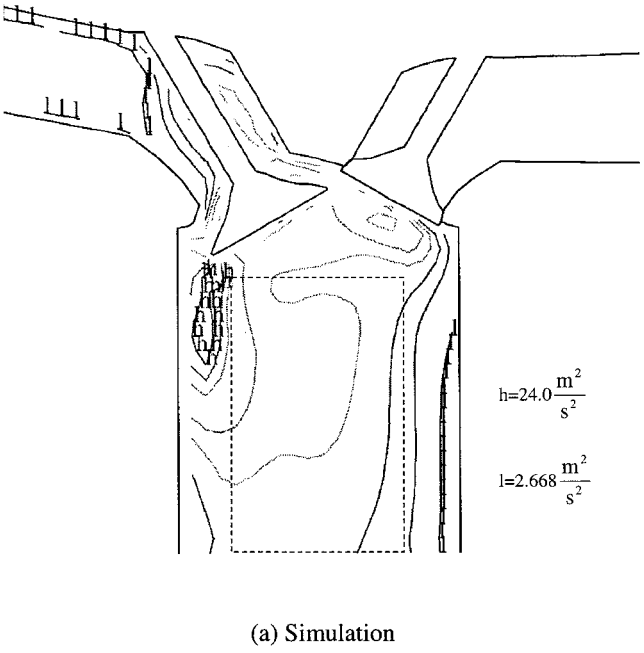


Fig. 11 Comparison of simulation and experimental results of turbulence kinetic energy represented by contour diagrams

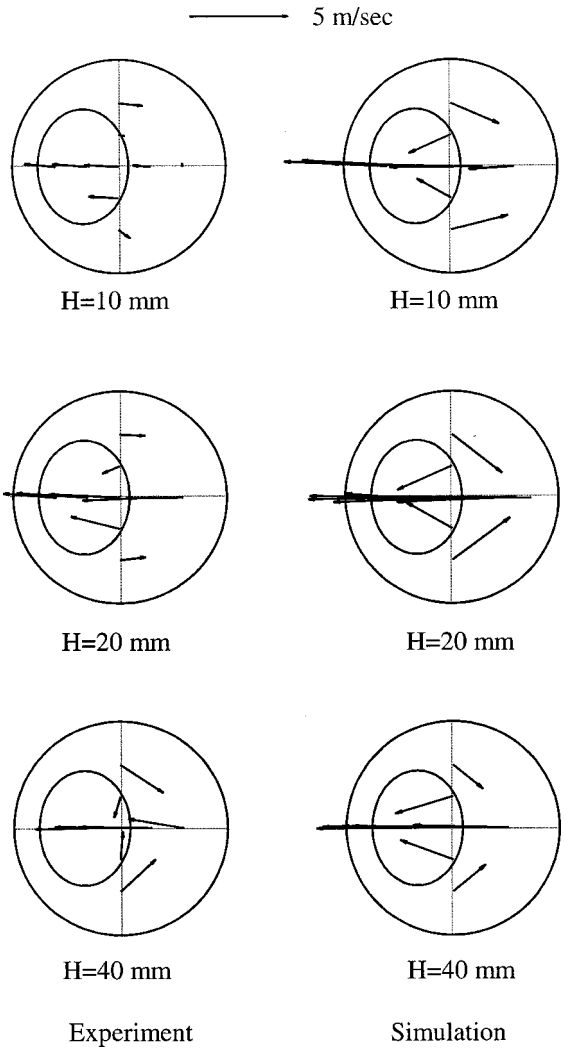


Fig. 10 Comparison of simulation and experimental velocities on the XY planes

of the port-valve-liner assembly with the dome cylinder head and two canted valves. The computational domain includes parts of the intake and exhaust pipes together with the valves, ports, combustion chamber, and the liner just as the model engine in the experiment. Full cylinder (360 deg) was taken into account considering reshaping of the complex geometry of the combustion chamber which makes the grid structure slightly in-symmetric. The grid is more dense in the top region, because the flow is expected to be more complicated in this area.

By specifying pressures at inlet and outflow boundaries (originally the moving piston), the package was converted to solve the compressible, stationary turbulent flows with fixed boundaries. The standard $\kappa-\varepsilon$ turbulence model was used with wall functions. Standard coefficients in the two-equation model were employed without modification. The steady inflow and outflow rates need at least 1500 computational cycles to converge. Grid sensitivity study has shown that the maximum difference in the mean velocity and turbulence intensity is only about 5 percent for increasing mesh number by 100 percent. Moreover, the computed volume flow rates are almost identical.

Simulation Results and Comparisons

The simulation results are illustrated by velocity vector diagrams, as displayed in Fig. 6 and Fig. 7. They are the cross sections on the XZ, YZ, and XY planes, respectively. Figure 6(*a*) reveals that there are two tumble jets inside the cylinder, one is

the main jet with a large vortex guided by the dome chamber; the other is a counter rotating jet with much smaller scale. Figure 6(*b*) exhibits two symmetric tumble vortices in the YZ plane. The maximum velocity is located around the inlet valve curtain. Note that the vectors point downwards at the BDC, because the piston was replaced by the outflow boundary. Figure 7 shows the velocity vector diagram in four horizontal planes. They located at 10 mm, 20 mm, 40 mm, and 60 mm beneath the TDC. With different scales, the flow patterns show similarities in different locations. The flow field tends to be more homogeneous as the flow further downwards.

If we compare the simulation results (Fig. 6 and Fig. 7) with their counterparts in experiments (Fig. 3 and Fig. 4) by streamline plots, Fig. 8 and Fig. 9 show the flow patterns are in good agreement. Although there are still some errors in both magnitude and direction of the mean velocity, the simulation clearly indicates the apparent feature of the whole flow field. Moreover, the simulation is able to show the details of the flow physics at some place where experiment is difficult to detect, e.g., flow field near the wall boundary or at corners. Figure 10 shows the comparison of the vector diagrams on the XY planes, 10 mm, 20 mm, and 40 mm beneath the TDC, respectively. The comparison of computed and measured flow velocities at these three cross sections show reasonably good level of agreement in direction prediction. However, simulation tends to over predict the magnitude. Figure 11 compares the turbulence kinetic energy in the XZ plane shown by a contour diagram. The distributions of the largest and the smallest

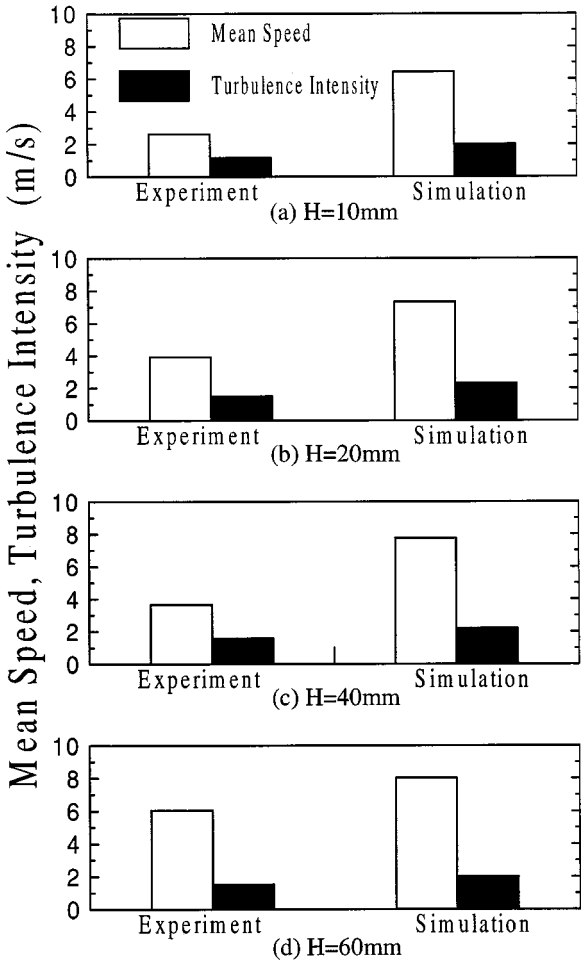


Fig. 12 Comparison of simulation and experimental results of mean speed and turbulence intensity at different locations in Z axis

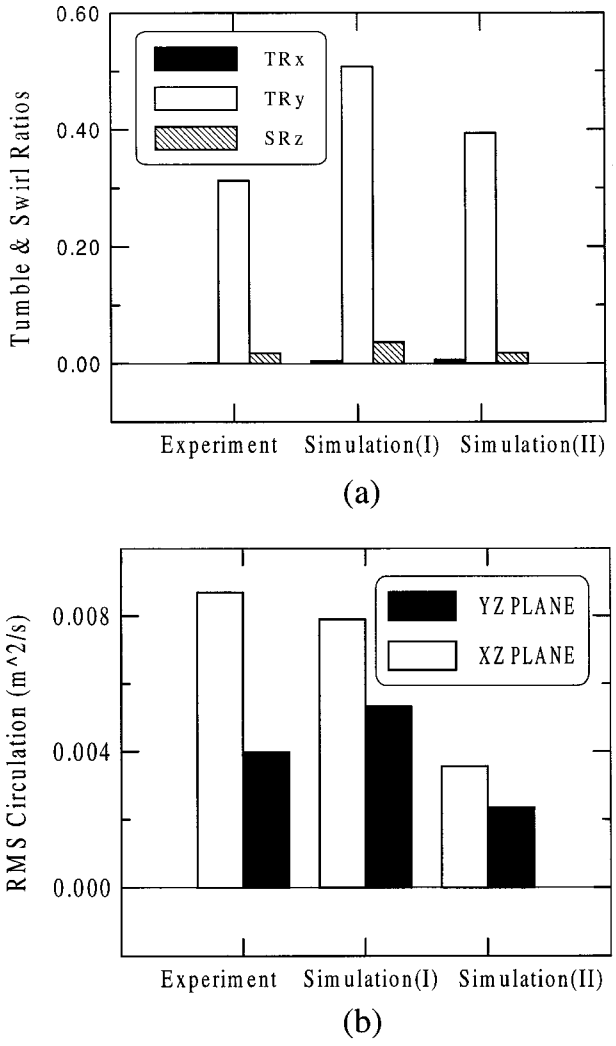


Fig. 13 Comparison of simulation and experimental tumble ratio, swirl ratio and RMS of circulations. Simulation (I) means calculation from the measurement area, while simulation (II) stands for the calculation from the whole area of the liner cross section.

turbulence intensity agree with each other. The largest turbulence intensity takes place at the interface of two rotating jets. Figure 12 shows the comparison of the mean speed, turbulence intensity in both simulation and experiments at several positions which locate at the *Z*-axis distanced 10 mm, 20 mm, 40 mm, and 60 mm beneath the TDC. Both the mean speed and the turbulence intensity in simulation tend to over predict the result. The tumble ratio and swirl ratio about *X*, *Y*, *Z*-axes are shown in Fig. 13(*a*). In the figure, simulation (I) means calculation from the measurement area, while simulation (II) stands for the calculation from the whole area in the linear cross section. Both simulation results over predict the ratios. The characteristics of the tumble motion are better represented by the RMS of the vortex circulation, as shown in Fig. 13(*b*). The comparison shows that it is in reasonably good agreement between the experimental and simulation results.

Conclusions

In-cylinder turbulent flow field measurements and predictions have been proved to be two useful tools in designing the engine geometry. Both techniques have their own advantages, but with inherent drawbacks. CFD provides useful qualitatively information of the flow field in details, but the accuracy is limited by the uncertainties of the turbulence modeling and some complex

boundary conditions. The LDV measurement adopted in this research is able to provide accurate quantitative answers; nowadays, however, the experiment is laborious and normally expensive as well. Both methods still have a long way to be improved from the academic viewpoint. Nevertheless, for industrial applications, they are useful design tools, provided that the errors are within acceptable limits.

This paper has validated the modern engine simulation code, KIVA3V, in the case of an engine with dome shape combustion chamber and canted valves. Simulation results tend to over predict the magnitude of the mean speed, the turbulence intensity and the tumble ratio. However, the flow patterns in the whole flow field are correctly predicted. Summarily, the simulation and experimental results are in reasonable level of agreement. The discrepancy is acceptable to the engine designer. Hence, based on the above research, further improvements to the baseline engine are being conducted by the motorcycle industry in Taiwan to design higher performance powertrain systems.

Acknowledgments

Support from both NSC 86-2212-E007-063 and ITRI 0986-053J3 contracts are greatly appreciated. Special thanks are due to Dr. S. C. Kong for his guidance in grid generation of the KIVA3V code.

Nomenclature

A = rectangular area of the mesh
H = angular momentum per unit mass
I = moment inertia per unit mass
N = engine speed
$\vec{r}$ = position vector
SR = swirl ratio
TR = tumble ratio
U = instantaneous velocity
$\overline{U},\overline{V},\overline{W}$ = mean speeds
u = turbulence velocity
u',v',w' = turbulence intensities
x,y,z = distance from origin in three components
ε = turbulence dissipation rate
κ = turbulence kinetic energy
ω = crank speed
Γ = circulation

References

[1] Zhao, F. Q., Lai, M. C., and Harrington, D. L., 1997, ''A Review of Mixture Preparation and Combustion Control Strategies for Spark-Ignition Direct-Injection Gasoline Engines,'' SAE Paper 970627.
[2] Hill, P. G., and Zhang, D., 1994, ''The Effects of Swirl and Tumble on Combustion in Spark-Ignition Engines,'' Prog. Energy Combust. Sci., **20**, pp. 373–429.
[3] Taghavi, R., and Dupont, A., 1989, ''Investigation of the Effect of Inlet Port on the Flow in a Combustion Chamber Using Multidimensional Modeling,'' ASME J. Eng. Gas Turbines Power, **111**, No. 7, pp. 479–484.
[4] Amsden, A. A., O'Rouke, P. J., Bulter, T. D., Meintjes, K., and Fansler, T. D., 1992, ''Comparisons of Computed and Measured Three-Dimensional Velocity Fields in a Motored Two-Stroke Engine,'' SAE Paper 920418.
[5] Hessel, R. P., and Rutland, C. J., 1995, ''Intake Flow Modeling in a Four-Stroke Diesel Using KIVA-3,'' J. Propul. Power, **11**, No. 2, pp. 378–384.
[6] Kong, S. C., and Hong, C. W., 1997, ''Comparisons of Computed and Measured Flow Processes in a Four-Stroke Engines,'' Paper No. 97-ICE-49, ICE-Vol. 29-2, 1997 ASME Fall Technical Conference, ASME, New York.
[7] Hampson, G. J., and Reitz, R. D., 1998, ''Two-Color Imaging of In-Cyliner Soot Concentration and Temperature in a Heavy-Duty DI Diesel Engine With Comparison to Multidimensional Modeling for Single and Split Injections,'' SAE Paper 980524.
[8] O'Connor, J. F., and McKinley, N. R., 1998, ''Comparisons of Intake Port CFD Flow Predictions Using Automatic Brick Meshing With Swirl Laser Sheet and LDA Flow Measurements,'' SAE Paper 980129.
[9] Le Coz, J. F., Henriot, S., Pinchon, P., 1990, ''An Experimental and Computational Analysis of the Flow Field in a Four-Valve Spark Ignition Engine-Focus on Cycle-Resolved Turbulence,'' SAE Paper 900056.
[10] Amsden, A. A., 1997, ''KIVA3V: A Block-Structured KIVA Program for Engines With Vertical or Canted Valves,'' Los Alamos National Lab, LA-13313-MS.
[11] Pao, R. H. F., 1967, *Fluid Dynamics*, Charles E. Merrill Books, Columbus, OH.
[12] Amsden, A. A., 1993, ''KIVA-3: A KIVA Program With Block-Structured Mesh for Complex Geometries,'' Los Alamos National Lab, LA-12503-MS.

Forced Convective Heat Transfer of Parallel-Mode Reciprocating Tube Fitted With a Twisted Tape With Application to Piston Cooling

S. W. Chang
Associate Professor,
Department of Marine Engineering,
National Kaohsiung Institute of
Marine Technology,
Post Code 811
Kaohsiung, Taiwan, R.O.C.
e-mail: swchang@mail.nkimt.edu.tw

This experimental study, motivated by the need to improve the cooling performance of a piston in a marine propulsive diesel engine, investigated turbulent flow heat transfer in a reciprocating tube fitted with a twisted-tape insert. The nonreciprocating experimental data, obtained from the tube fitted with twisted tape, confirmed that heat-transfer augmentation from plain-tube level occurs. When the test tube reciprocated, buoyancy effects became appreciable, and interacted with the reciprocating and inertial forces to provide considerable heat transfer modifications from nonreciprocating situation. When the reciprocating forces were relatively weak, a range of heat-transfer impediments, that could reduce local Nusselt numbers to levels of about 53% of nonreciprocating values, was observed. A further increase of the relative strength of the reciprocating force resulted in a subsequent heat-transfer recovery, and eventually led to heat-transfer enhancements relative to the nonreciprocating situation. For design considerations, heat-transfer enhancement due to the twisted-tape insert was confirmed, but the impediments from nonreciprocating levels at lower values of pulsating numbers needs particular attention in order to avoid overheating situations. [DOI: 10.1115/1.1340627]

Keywords: Reciprocating Heat Convection, Twisted-Tape Insert

In order to match the power/speed characteristics of marine propellers to those of the propulsive system, there has been a recent tendency to design marine diesel engines to achieve the required power at a relatively low engine speed. This obviates the need for heavy-duty reduction gearboxes with an attendant reduction in weight. Typically this involves the need for a continuous power output from a diesel engine of about 15,000 HP, with a shaft speed of about 80 rev/min. The diameter of the piston, under these circumstances, is usually in the range 60–90 cm. In order to achieve a relatively low engine speed, the use of a larger than normal stroke to bore ratios is being examined actively. In addition, the well-known thermodynamic advantages obtained by designing the engine to run with a high compression ratio and maximum combustion temperature form an active; and continuing, quest for improved performance and economy.

Unfortunately, the use of a high stroke to bore ratio, combined with a high compression ratio and high combustion temperature, greatly increases the thermal and mechanical loads imposed on the piston. In turn, this necessitates the use of cooled pistons if the mechanical integrity and life are to be maintained at a commercially realistic level. The internal piston cooling may use either water or oil as the cooling medium. Piston cooling flow circuits are usually comprised of interconnected passages that are aligned either in a direction parallel or orthogonal to the direction of piston reciprocation. The individual coolant passages involve interconnecting bends, and the internal surfaces of these passages may be smooth or deliberately rib roughened. The use of helical-twisted tapes fitted inside circular coolant channels is also a potentially useful method for enhancing the convective mechanism that controls the transfer of heat from the piston into the coolant; see Refs. [1–5]. The use of internal convection-enhancing devices, such as ribs and twisted tapes, has been the subject of much fundamental research over the years for stationary duct flow situations. If the coolant passage is reciprocating, as with the piston-cooling application, additional inertial forces act on the coolant. These forces can dramatically change the nature of the flow pattern that is established in the channel, and hence the heat-transfer performance. These effects must be fully understood if reliable cooling performance is to be predicted at the design stage.

The present paper describes the results of a series of experiments aimed at studying the effect of reciprocation on heat transfer in a channel of circular cross section, and fitted with a continuously twisted helical tape. The direction of reciprocation is aligned in the axial direction of the channel. The study forms part of an ongoing program of research into the performance of piston-cooling systems. When the channel is stationary, the presence of the tape creates secondary flows and turbulence enhancement that, in turn, improves the heat transfer. When the channel reciprocates, time-varied inertial forces further modify the flow and turbulence levels; see Ref. [6]. It is with the combined effect of the helical insert and the oscillation that the present investigation is concerned.

Physical Considerations

With a reciprocating tube, the fluid velocity field inside the tube will be influenced by the periodic motion in some way, and consequently by the internal heat transfer. The pertinent nondimensional groups associated with this modified forced convection situation may be derived from the basic conservation equations *of mass*, *momentum*, and *energy*. It is convenient to consider the fluid motion relative to the reciprocating tube. This means that the usual form of the momentum conservation equation must be modified to include the fact that the fluid motion is referred to a noninertial reference frame. A detailed explanation of the way by which these modifications are made may be found in Ref. [6]. Thus the form of the momentum equation for this oscillating geometry becomes

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division Apr. 12, 1999; final revision received by the ASME Headquarters Oct. 9, 2000. Associate Editor: D. Assanis.

$$\frac{D\nu}{Dt} = -\frac{1}{\rho}\nabla p^* + \upsilon\nabla^2\nu - \omega r \sin\omega t(\vec{k}\cdot\nabla)\nu + \beta(T-T_0)\omega^2 r\cos\omega t\vec{k} \quad (1)$$

All symbols are defined in the Nomenclature section. It may be worth noting that, the coordinate system used in Eq. (1) is a reciprocating reference frame. The flow pressure field p^* can be resolved as

$$p^* = p + \rho\omega^2 r \sin\omega t z, \quad (2)$$

in which the pressure field p of Eq. (2) is the externally applied pressure source that drives forced convective flow. The last component of Eq. (2) periodically varies in the time domain, and accounts for the effect of system reciprocation on the pressure field. Thus the nature of the pulsating flow in a stationary duct may be fundamentally different from the reciprocating duct flow. For pulsating flows in stationary ducts, the flow pulsation normally originated from the externally applied oscillatory pressure wave superimposed onto a pressure head. This oscillatory pressure or velocity potential is usually defined as the boundary condition either at the flow entrance or flow exit. As an illustrative example, the effect of flow pulsation on mass transfer could be linear, and the Sherwood number associated with a pulsating flow in a stationary duct could sometimes be expressed as the summation of Sherwood numbers of steady and oscillatory flows [7]. For flow in a reciprocating duct, the reciprocating force is induced by system reciprocation, and the relative strength of reciprocating force, in terms of $\omega r \sin\omega t$ in Eq. (1), is coupled with $(\vec{k}\cdot\nabla)\nu$ as a nonlinear term to affect the flow motion. Thus we expect the effect that the relative strength of the reciprocating force on heat transfer might be nonlinear. In view of the turbulent and vortical behavior, the differences in the laminar form of the momentum conservation equations between pulsating flow in the stationary duct and reciprocating duct flow would yield different turbulence and vortical flow motions. This fundamental difference may provide the essential differences in the physical phenomena between the present flow and pulsating flow in a stationary system.

For flow in a tube containing a twisted tape, the centrifugal and torsional forces arise naturally when the fluid traverses the twisted semicircular passages, and these forces can distort the axial velocity profile and modify the distributions of turbulent kinetic energy [1]. As spanwise oscillations have been shown to affect the longitudinal vortex initiated by the centrifugal or Rayleigh-Benard instability [8], we have the possibility that the buoyancy effect on heat transfer might not be negligible when a tube containing twisted tape is reciprocated. As a convenient reference for buoyancy level, the buoyancy parameter was defined as $\beta(T_w - T_f)$ to quantify the relative strength of the buoyancy effect at a particular set of flow rate and reciprocating frequency. As the temperature variation of the coolant did not make a significant change to a coolant Prandtl number for this class of experiment, its effect on heat transfer was not investigated. Having completed the nondimensionalized process for Eq. (1), and based on the aforementioned geometric and thermal boundary conditions described in Fig. 2, the nondimensional heat transfer relationship for forced convection in the reciprocating test section has the functional form

$$\mathrm{Nu} = \{\mathrm{Re}, \mathrm{Pu}, \beta(T_w - T_f), Z\}, \quad (3)$$

in which

$$\mathrm{Nu} = q_f d/k_f\{T_w - T_f\} \quad \text{(Nusselt number)} \quad (4)$$

$$\mathrm{Re} = W_m d/\upsilon \quad \text{(Reynolds number)} \quad (5)$$

$$\mathrm{Pu} = \omega r/W_m \quad \text{(Pulsating number)} \quad (6)$$

All symbols in Eqs. (3)–(6) are defined in the Nomenclature section. When the test section was reciprocated, the pulsating number, Pu, defined in Eq. (6), provided an indication of the time-averaged force ratio between the reciprocating and convective inertial forces.

Experimental Apparatus

Reciprocating Facility. Figure 1 shows the schematics of the reciprocating facility used for the experiments. An air receiver (1) was maintained at a pressure in the range 6–8 bar by a compressor (2). Air from the receiver was allowed to flow through the heated test section (3), which simulates the oscillating piston-cooling channel, via an air dryer (4), a combined pressure regulator and filter (5), a Tokyo Keiso TF-1120 mass flow meter (6), and a needle valve (7). A flow calming section (8), of a length equivalent to 20 tube diameters, was fitted immediately upstream of the entry plane of the test section. The needle valve (7) was used to set the airflow through the test section at a level appropriate for the Reynolds number at which a particular experiment was to be conducted.

A crank-wheel mechanism (9), driven by a 1500-W dc electric motor (10), via a reduction gearbox (11), was used to create the oscillatory motion of the test section (3). Speed fluctuations were minimized by means of a flywheel (11). A counter balancing weight (12) was fitted to the flywheel to maintain dynamic balance and a pair of guide ways was used to ensure a smooth reciprocating motion of the test section. An optical proximity transducer (13) was used to measure the rotation speed of the crank mechanism, and hence the reciprocating frequency of the test section.

Test Section Module. Figure 2 shows the constructional details of the heated test section. As shown, a heated stainless-steel tube (1) of circular cross section was used to simulate the piston-cooling channel. This tube had a bore diameter of 15.0 mm and a wall thickness of 1.5 mm. This tube was held between insulating bushes (2) and (3), so that the length of the actively heated test section was 225.0 mm. The inlet and exit bushes (2) and (3) were made from insulating material to minimize axial heat loss from the test section. The outer surface of the tubular test section was electrically heated by means of spirally wound nickel/chrome resistance wire having a diameter of 0.3 mm. The wire was wound onto the outer surface of the test section with a pitch of 2.0 mm.

The heated test section and supporting bushes were held together by means of inlet and exit flanges (4) and (5) together with four draw bolts (6). The air entered the test section via an upstream plenum chamber, of diameter 120 mm to create a ''sharp'' entry condition. The pressure and temperature of the entry air in this plenum chamber was measured with a pressure transducer linked to digital encoder and a type-*K* thermocouple respectively. The inlet Reynolds number, for a particular experiment, could be calculated and monitored using the measured air flow, plenum chamber pressure, and temperature measurements. This inlet Reynolds number was based on the bore diameter of the test section. A convergent nozzle, fitted into the exit flange (5), permitted the pressure in the entry plenum to be increased, for a given nominal Reynolds numbers Re. This meant that the axial air velocity through the test section could be reduced in response to a plenum chamber pressure rise. In turn, this meant that a range of pulsating numbers Pu could be examined for a specified nominal Reynolds number simply by adjusting the plenum chamber pressure.

The temperature of the test section wall was measured at nine of axial locations, along two opposite edges of the test section, using type-*K* thermocouples. The locations of these thermocouples, expressed in equivalent or hydraulic diameters from the entry plane, was 0.00, 1.33, 2.66, 4.00, 6.67, 8.00, 9.33, and 10.67, respectively. These wall thermocouples were actually embedded in the tube wall so that the junctions were 0.50 mm from the inner bore surface of the tube. In order to avoid an ''end heat loss'' effect and an ''upstream'' effect of the choked nozzle, the

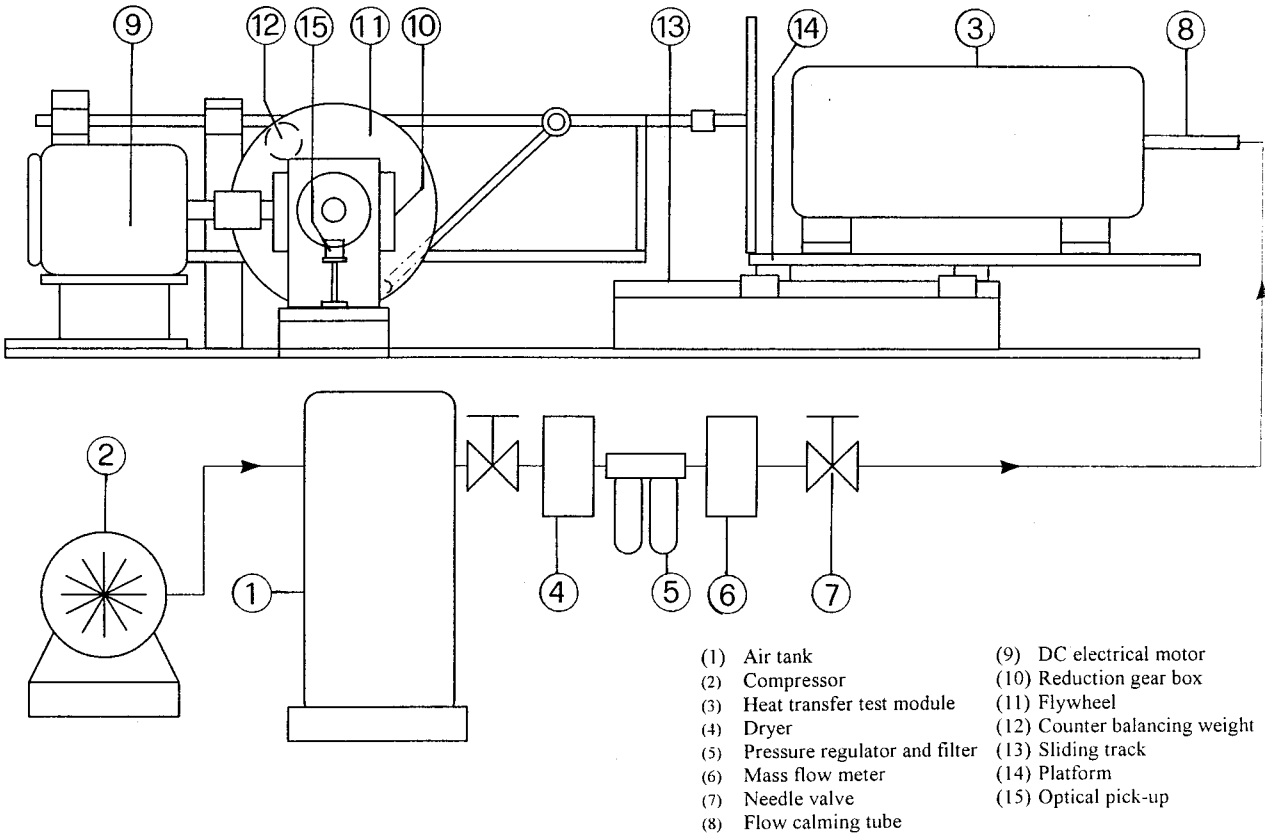


Fig. 1 Schematic layout of the reciprocating test facility

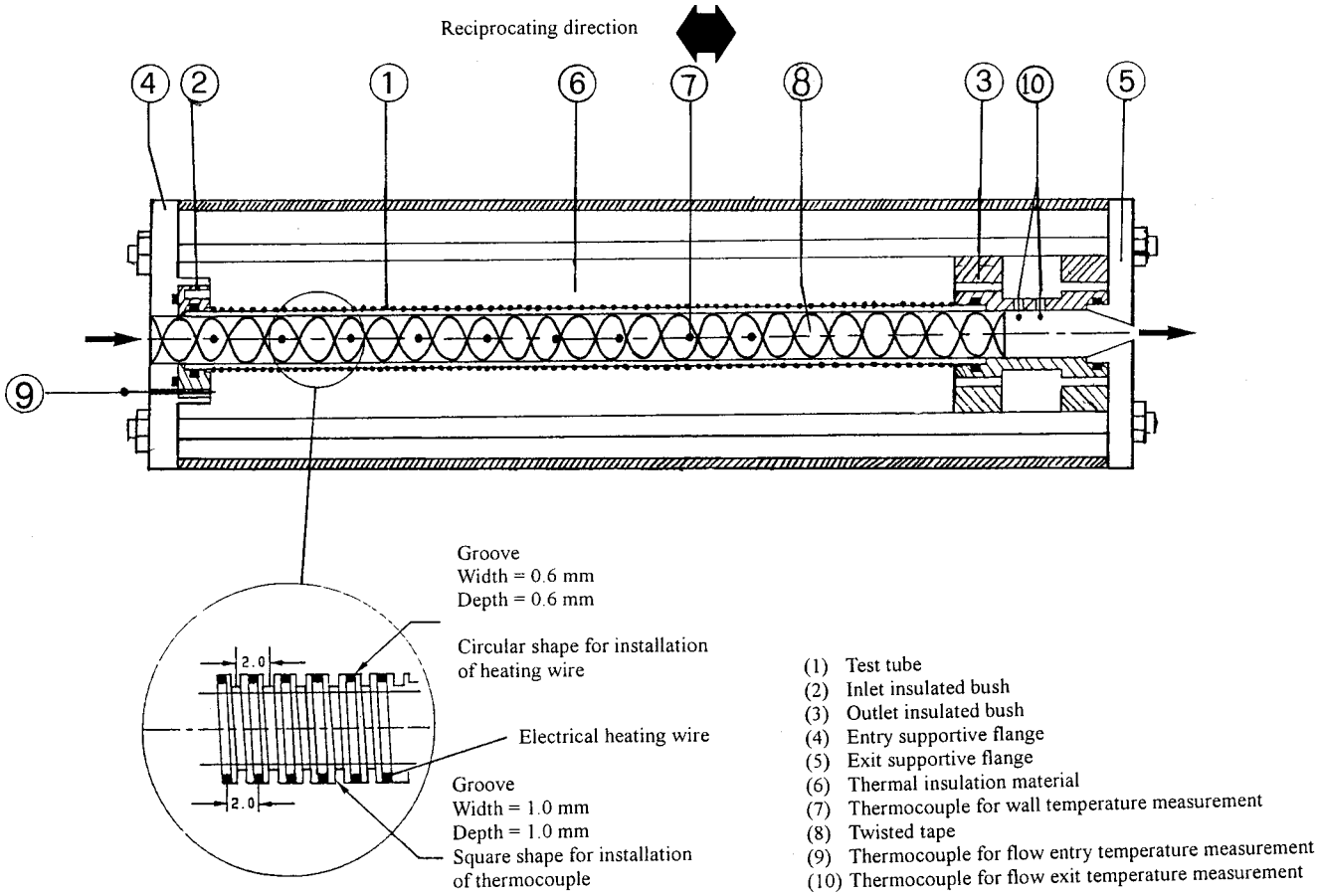


Fig. 2 Constructional details of a test section

final wall thermocouple was approximately 7.5 hydraulic diameters upstream of the exit plane of the test section.

The helical twisted strip (7) inserted in the test section was made from a strip of stainless steel 1.50 mm thick. The width of this tape corresponded to the inner bore of the test section and the helical pitch was 1.33 times the bore diameter of the test section. This tape traversed the entire length of the actively heated tube and the inlet-supporting bush. The fin parameter C_{fin}, defined as $(k_m/k_f)(\delta/d)$, was 57.7 for present study. Near the flow entrance and in the outlet insulated bush (3), three thermocouples penetrated into the flow. Thermocouple (9) measured the bulk flow entry temperature and two additional thermocouples (10) detected the exit flow bulk temperature. In order to avoid the upstream and downstream disturbances induced by these thermocouples for flow-temperature measurements, these thermocouples were not allocated within the heated flow path, as indicated in Fig. 2. The penetrating distances of these centerline thermocouples were far above the boundary layers in order to probe the temperature of the turbulent core. By using an enthalpy balance method, the flow bulk temperatures along the tube were sequentially calculated starting from the measured flow temperature at the entrance to the end of the test tube. In general, the difference between the estimated and measured flow bulk temperatures at the exit plane was less than 8%, and this set of experimental data was collected for subsequent data reduction.

Experimental Program and Data Reduction

The experimental program adopted involved three phases. Initially the test section and instrumentation was checked out with a series of heat-transfer experiments with a stationary test section, over a range of Reynolds number levels, without the twisted tape fitted. This was followed by a similar series of stationary base line experiments with the twisted tape inserted into the test section. Finally, heat-transfer experiments were undertaken with reciprocation of the test section, with the twisted tape inserted.

With the reciprocating experiments, for each Reynolds number/pulsating number pair, five different levels of heater power were selected to permit an assessment of the effect of buoyancy on heat transfer. Actual heater powers were selected to create a test section nominal wall temperature of 90 deg, 100 deg, 110 deg, 120 deg, and 130 deg at a downstream location of 6.67 effective diameters. The test section was reciprocated at three different frequencies, namely 0.42, 0.83, and 1.25 Hz. For each individual test, the apparatus was allowed approximately 30 min to achieve thermal equilibrium. The on-line data acquisition and storage system was then activated, and captured all relevant measurements for a total scan period of 10 s. In this respect, the time response for the measurement system was adjusted to 0.005 s/scan. This measurement rate was capable of resolving changes on the time scale for reciprocating tests at a maximum frequency of 1.25 Hz. However, this measurement rate might be insufficient to detect the full time-varied temperature variations for a reciprocating cycle. Therefore, the temporally maximum and minimum temperature levels over the 10 s ''equilibrium state'' were automatically logged. These measurements were used to define the oscillating amplitude of a temporal Nusselt number variations for a particular experiment. The raw experimental data were subsequently processed to generate the appropriate Reynolds number, pulsating number, and buoyancy number [see Eq. (2)], and the individual effect of these parameters on the Nusselt number distribution examined. The range of these non-dimensional parameters is given in Table 1.

Having captured the raw experimental data, the following procedure was adopted for data reduction, including the calculation of all relevant nondimensional groups. To evaluate the convective heat carried by the flow, q_f, for evaluation of the Nusselt number using Eq. (4), the net axial conductive heat along the test tube and the external heat loss were subtracted from the total heating power supply. To estimate the net axial conductive heat at any axial location, the measured axial wall temperature distribution was numerically curve fitted based on the wall temperature measurements and a simple finite-difference representation for Fourier conduction law. The characteristics of external heat losses at different reciprocating frequencies were revealed through a number of heat loss calibration runs at various reciprocating frequencies. For each set of calibration runs, the flow was blocked off, and the test section was filled with thermal insulation. Under such circumstance, the heat supplied to the test tube was entirely lost into the atmosphere and balanced with the amount of heat loss at correspondingly steady-state temperature distribution. It was found that the amount of heat loss was proportional to the locally prevailing temperature difference between the wall and the ambient, with the proportionality constant increasing with the reciprocating frequency. By incorporating these proportionality constants at different reciprocating frequencies into the data reduction program, and based on the measured wall and ambient temperatures, the amount of external heat loss for each heat transfer test were able to be estimated. The representative wall temperature T_w in Eqs. (3) and (4) was estimated by an one-dimensional radial conduction correction of the measured value to a value corresponding to the inner surface of test tube. For a 225-mm-long straight tube fitted with a twisted tape, the quasisteady pressure level in the tube was assumed as a constant for each reciprocating test. In general, based on the pressure measurements in front of the test section, less than 1.2% of pressure oscillation due to system reciprocation was found. Subsequently, all the local dimensionless groups in Eq. (3) along test tube were calculated with the fluid properties determined using the local bulk fluid temperature at the measured pressure level, for convenience.

Table 1 Range of experimental parameters investigated

Reynolds number range	3500–18,000
Pulsating number range	0–0.2
Buoyancy parameter range	0.05–0.28

An uncertainty approximation for the data reduction was conducted [9]. During data reduction process, the local properties of the coolant, needed in Eqs. (5) and (6) were estimated from the fluid bulk temperatures and thus one of the major sources of uncertainty was this temperature measurement. As the steady state of a nonreciprocating flow system was approximated when the local temperature variations were in the range of ± 0.3°C, the maximum uncertainty in the temperature measurement was estimated to be ± 0.3°C. The maximum percentages of error for the coolant's specific heat, mass flow rate and thermal conductivity, the fluid density, the hydraulic diameter of test duct, and the heating area were estimated as $\pm 0.05\%$, $\pm 4\%$, $\pm 0.25\%$, $\pm 0.3\%$, $\pm 0.05\%$, and $\pm 0.6\%$, respectively. With the temperature difference between wall and fluid varied from 20 to 60°C, the maximum uncertainty for Nusselt, Reynolds, and pulsating numbers and the buoyancy parameter were about 12%, 5%, 1%, and 0.5%, respectively. Since the Reynolds number, pulsating number, and buoyancy parameter appeared in the dimensionless fluid momentum equations as the ''coefficients'' in front of the forced convective, pulsating, and buoyancy terms, respectively, changes of the values of these dimensionless groups would yield the solution patterns of flow and therefore the heat transfer. This study is aimed at experimentally examining the effects of system reciprocation on heat transfer when the Reynolds, pulsating numbers, and buoyancy parameters were systematically varied. In this way the interactive influences induced by these dynamic forces present in a heated reciprocated duct could be examined.

Results and Discussion

Stationary Results. The axial variation of the Nusselt number was examined for the smooth-walled tube case over the entire

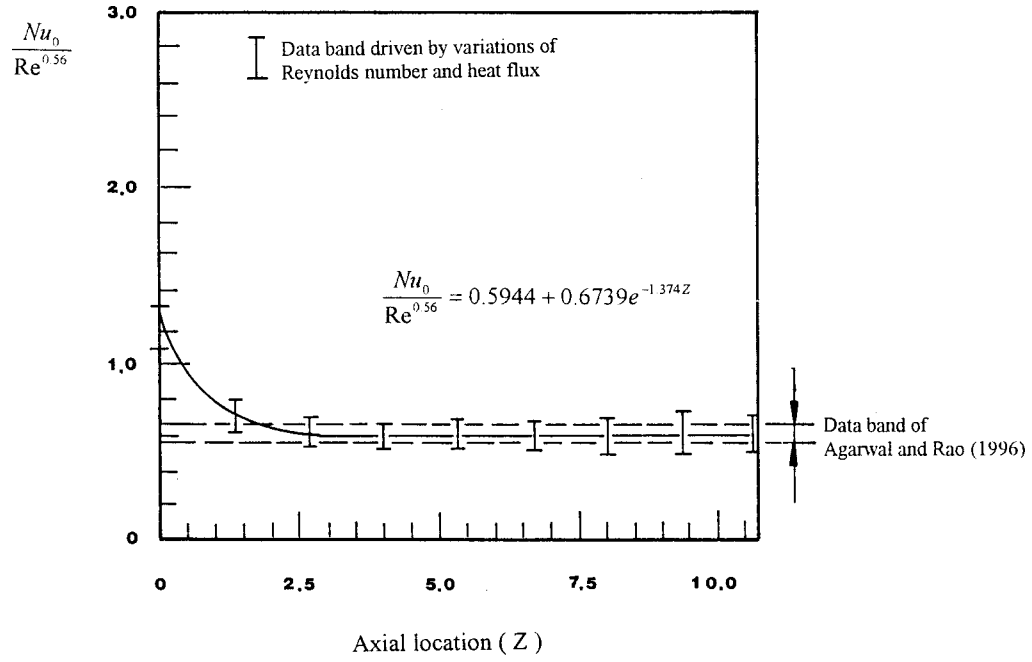


Fig. 3 Typical axial heat-transfer variations in a tube fitted with twisted tape

range of Reynolds numbers and heater power settings. In the immediate entry region of the test section, the local Nusselt number is relatively high. On moving downstream from the entry plane, the Nusselt number tended to decay and approach a fully developed level. This is consistent with the development of the wall boundary layer. It was found that all data acquired in the entry region could be represented with an equation having a simple exponential decay structure with reference to the axial location relative to the entry plane. This correlation is given in Eq. (7) as

$$Nu_\infty = [0.0203 + 0.05\exp(-1.389Z)]Re^{0.8} \quad (7)$$

Note that the Prandtl number variation, over the range of experiments conducted, is small. Thus Prandtl number effects are im-

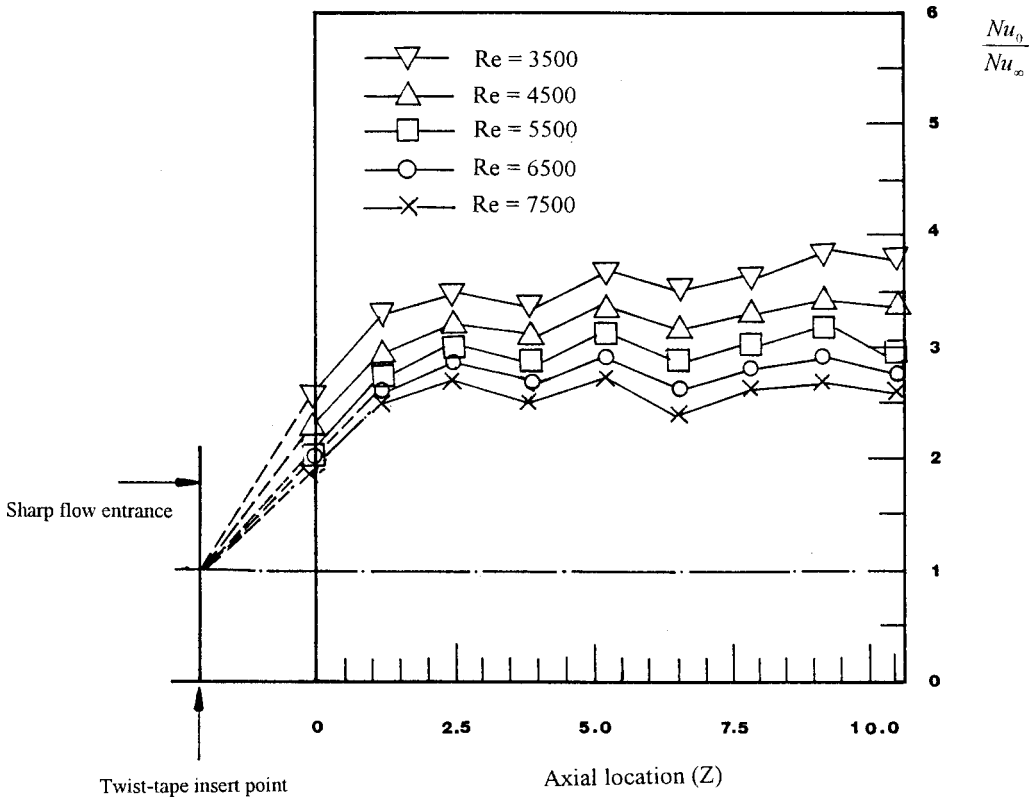


Fig. 4 Axial distributions of the heat-transfer performance ratio with various Reynolds numbers

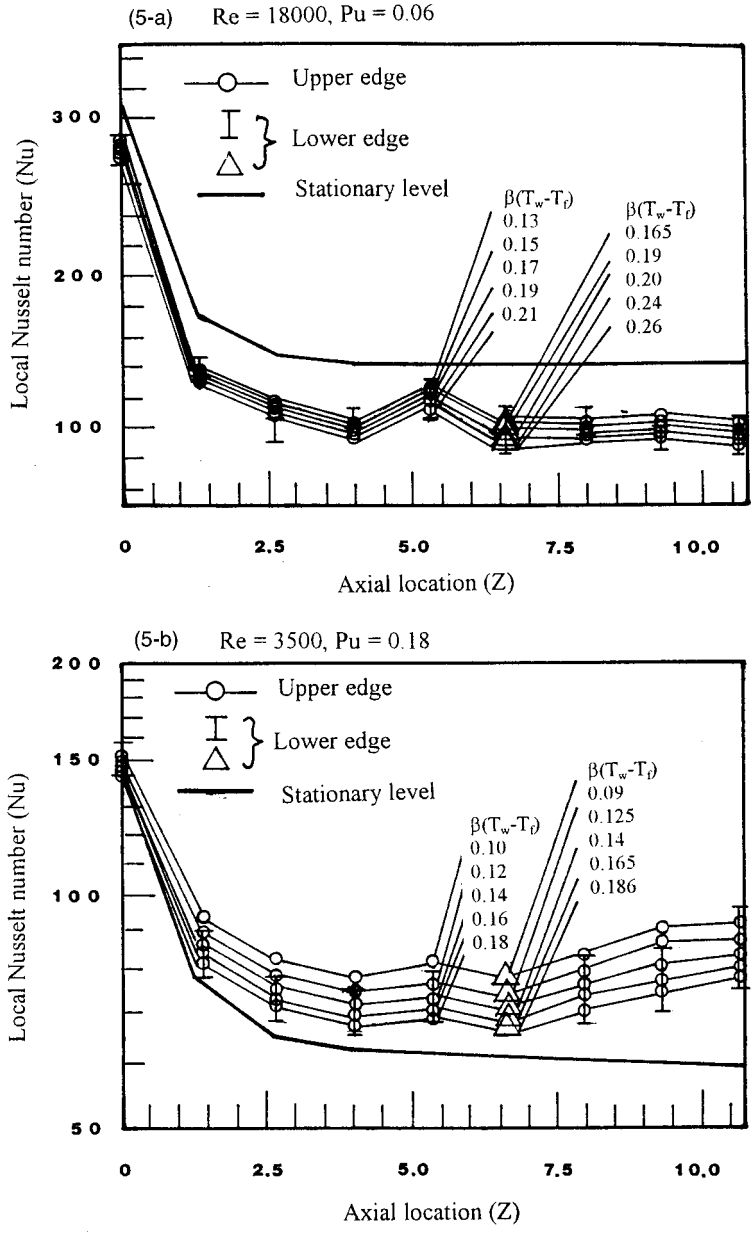


Fig. 5 Axial distributions of a reciprocating Nusselt number with various buoyancy levels at pulsating numbers of 0.06 and 0.18

plicitly included in the numerical constants shown in this equation. The fully developed Nusselt number is in good agreement with the well-known Dittus-Boelter [10] value.

A similar analysis was undertaken with the data obtained when the twisted tape was fitted into the test section. In this case the variation in local Nusselt number was correlated using Eq. (8):

$$Nu_0 = [0.5944 + 0.6739 \exp(-1.374Z)]Re^{0.56} \quad (8)$$

Figure 3 compares all stationary twisted tape data with Eq. (8). Also shown, for comparative purposes, in Fig. 3 is the experimental correlation of Agarwal and Rao [5]. Good agreement between the present data and that of Agarwal and Rao [5] is evident. As shown in Fig. 3, no evident heat-transfer differences were observed between two opposite edges at each axial location. Therefore, the swirls in the two semicircular twisted flow passages were symmetrical. Results obtained with different heater powers were collapsed into a narrow data scatter band in Fig. 3. The buoyancy effect in the stationary tube was thus negligible. With the twisted tape present the Nusselt number achieves a more or less fully developed level after about three effective diameters downstream of entry. This was also the case with the plain smooth-walled test section.

The overall increase in heat transfer resulting from the use of the twisted tape was assessed by standardizing the measured Nusselt number Nu_0 distributions with respect to the fully developed plain tube Nusselt number Nu_∞. Figure 4 shows the relative Nusselt number Nu_0/Nu_∞ distributions for Reynolds numbers of 3500, 4500, 5500, 6500, and 7500. Considerable enhancement in local heat transfer is produced by the insertion of the twisted tape. The swirl induced flow created by the twisted tape produces an augmentation of heat transfer in the range 206–385% compared to the plain tube situation. The enhancement produced increases in the direction of flow from the inlet station, and tends to reach a constant value in the developed flow regions at each Reynolds number. The enhancement in heat transfer produced decreased with increases in the Reynolds number at all axial locations. This is felt

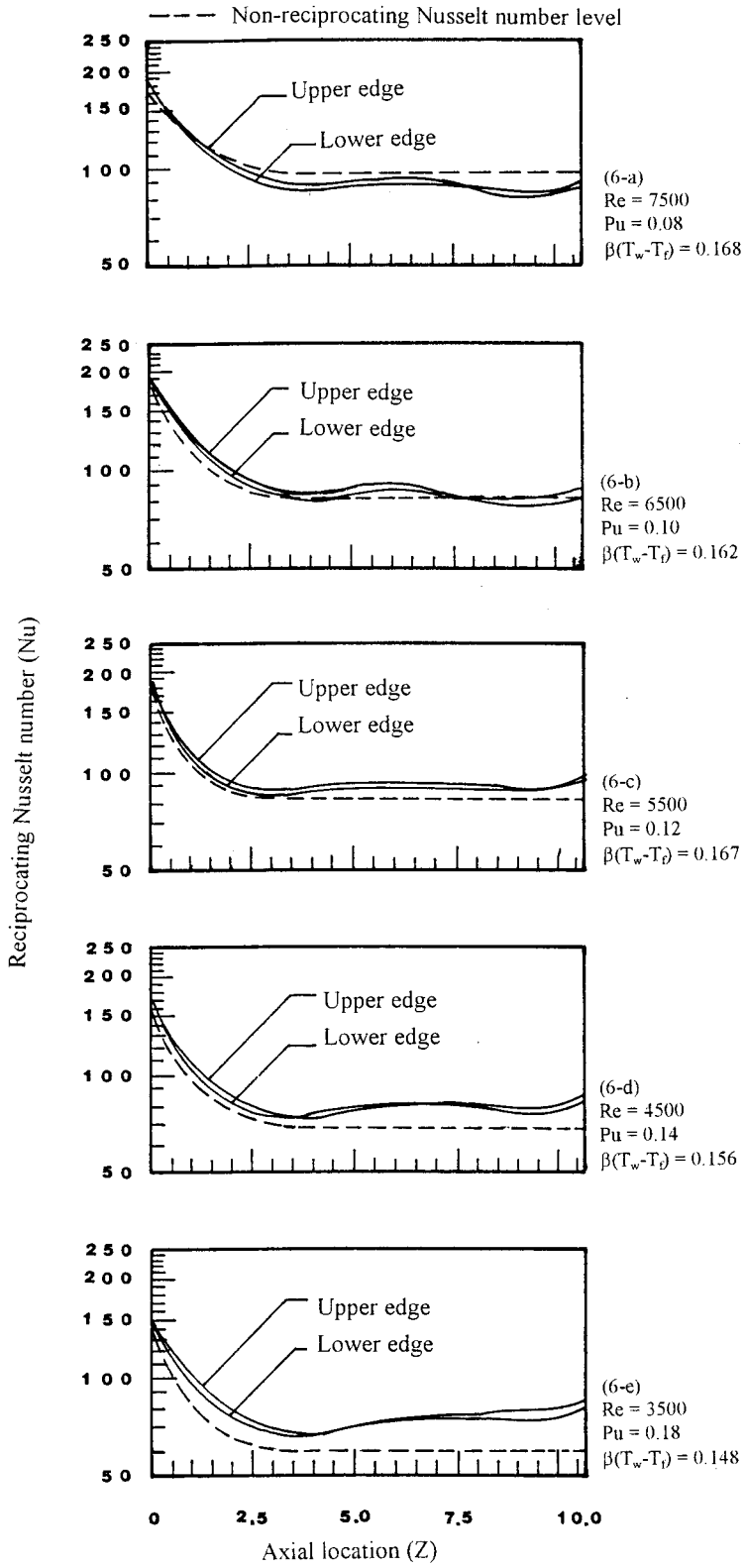


Fig. 6 Axial distributions of reciprocating Nusselt number with various pulsating and Reynolds numbers at reciprocating frequency of 0.83 Hz

to reflect the different Reynolds numbers exponents given in Eqs. (7) and (8). In the developed flow region, as shown in Fig. 4, the values of the relative Nusselt number Nu_0/Nu_∞ were about 3.85, 3.46, 3, 2.75, and 2.06 when the Reynolds numbers were 3500, 4500, 5500, 6500, and 7500, respectively. At the axial location $Z=0$, where the heating was initiated, the most thin data spread driven by the variation of the Reynolds number was found. However, by extrapolating the relative Nusselt number Nu_0/Nu_∞ upstream, Fig. 4 shows that all the relative Nusselt numbers collapse to unity at the location about 1.8 effective diameter upstream Z

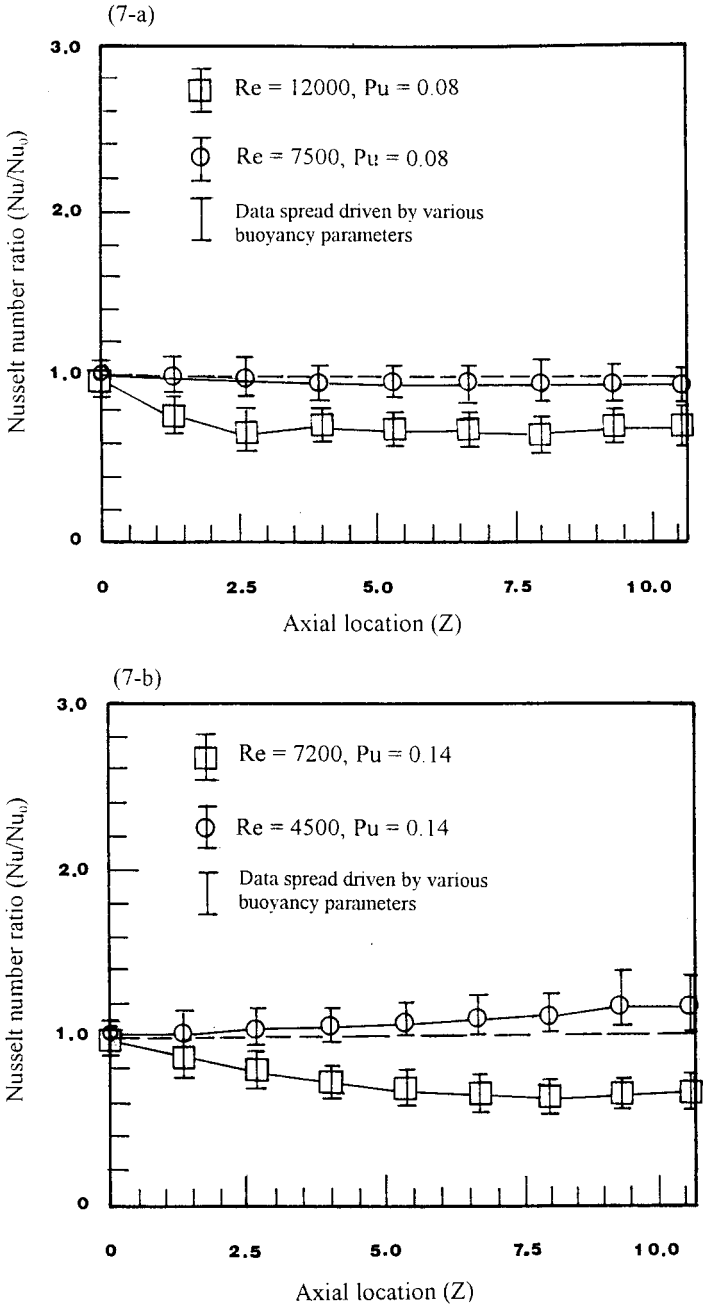


Fig. 7 Normalized axial Nusselt number distributions at a pulsating number of 0.08

$=0$ plane. Referring to the constructional details of the test section, the axial spot of $Z=-1.8$ was close to the location where the flow started to encounter the twisted tape. Therefore, the modification of flow structure by twisted tape actually commenced in front of the $Z=0$ plane. Along the test tube, the influences of twisted tape on heat transfer gradually emerged along with the merged boundary layers.

Reciprocating Results. When the tube is reciprocated, the pulsating flow convects cooler or warmer fluid within the pair of twisted semicircular passages in which swirling flows develop. Accompanying this possibility of pulsation of the bulk flow is a possible modification of the vortical flow cells, the turbulence structure, and the vorticity itself. As a result, under the basically uniform heat-flux heating condition, the wall temperature experiences a timewise variation. This produces an attendant temporal Nusselt number variation at each measuring station. Nevertheless, after taking the mean value of the timewise Nusselt number data for a period of 10 s, the time-averaged heater transfer results along the tube were determined. Figures 5(*a*) and 5(*b*) show the typical variation of the time-averaged Nusselt number along the measurement edges for nominal Reynolds of 3500 and 18,000, respectively, and pulsating numbers of 0.06 and 0.18. There was no discernible difference in the heat-transfer measurements on the two measuring edges which suggests that the flow conditions in both semicircular passage were the same even with test section oscillation. For comparison Figs. 5(*a*) and 5(*b*) also show the nonreciprocating heat-transfer distributions obtained with the same Reynolds number values. The heat-transfer characteristics shown in this figure were typical of those obtained with different running conditions. In general the heat-transfer dis-

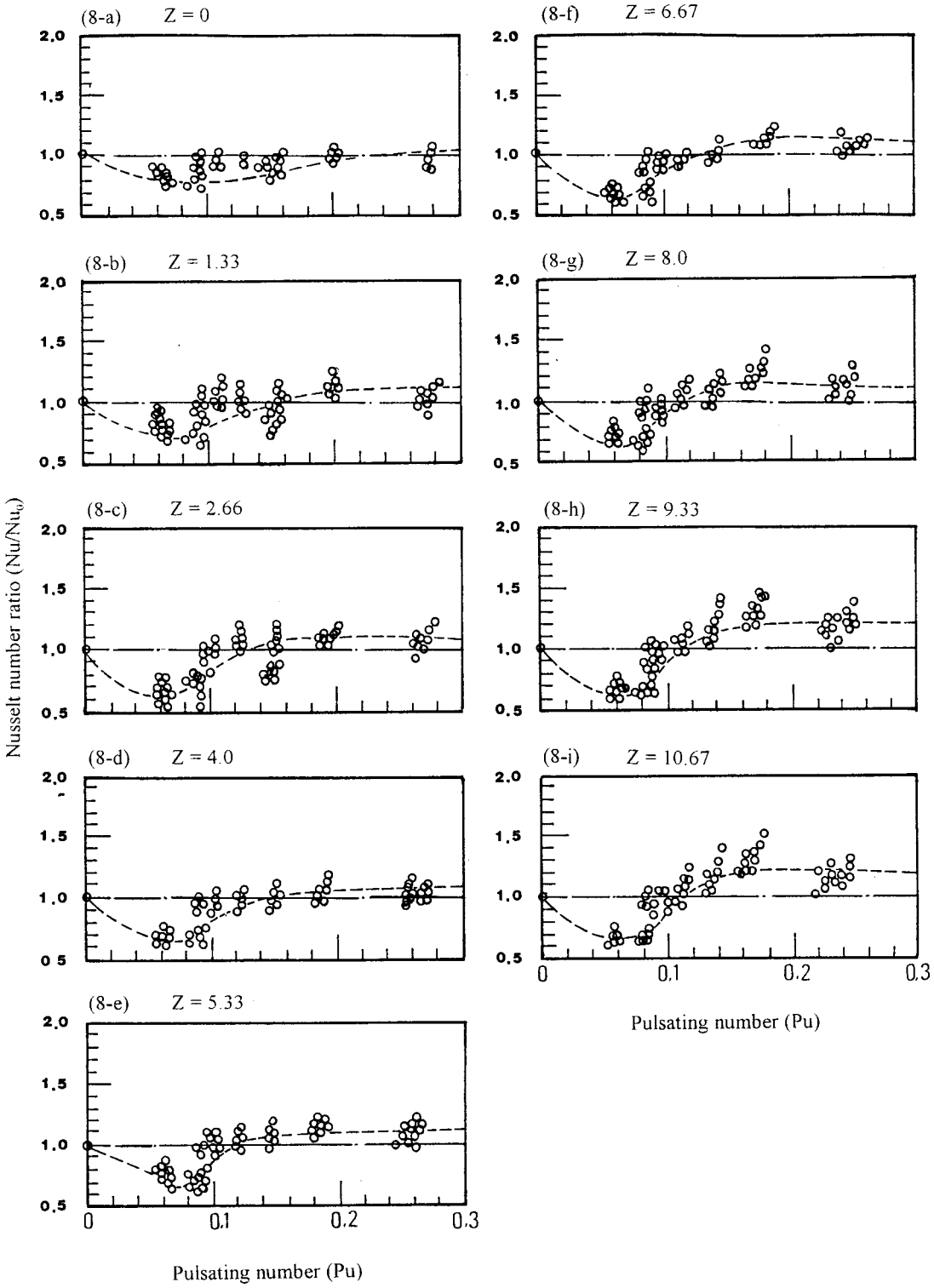


Fig. 8 Parametric presentation of overall reciprocating effects on heat transfer

tributions along the test section were not significantly different from the stationary results when the test section was reciprocated. However the reciprocating force could impede or enhance relative heat transfer when the pulsating numbers were low and high, respectively.

For each Reynolds number/pulsating number pair, the heat transfer was found to be sensitive to the heater powers used. In other words a buoyancy interaction was apparent when the test section was oscillated. As seen in both figures an increase in the buoyancy parameter, $\beta(T_w - T_f)$, creates a systematic decrease in the local Nusselt number. This sensitivity to buoyancy tends to increase with increases in the pulsating number for a specified axial location. Clearly, therefore, there is an interactive coupling between the pulsating force and the buoyancy. At the immediate entry section, the least data spread was noted because of the buoyancy effect. As the flow proceeds downstream, the sensitivity to buoyancy increases. The effect of buoyancy may be quantified by noting that, at a downstream location of approximately ten hy-

draulic diameters from the entry plane, heat transfer may be reduced by approximately 25% over the range of buoyancy parameters covered by the experiments.

The results described above demonstrated the combined effect of pulsating flow and the buoyancy effect. Because we are dealing with a real fluid it is not possible to entirely eliminate buoyancy, and examine the effect of flow pulsation in isolation. However, in order to attempt this, comparisons of the effect of the pulsating number were undertaken for tests where the average value of the buoyancy parameter, $\beta(T_w - T_f)$, are the same. Such a comparison is given in Fig. 6. In this figure the average buoyancy parameter over the length of the test section is 0.16. Figures 6(*a*) and 6(*b*) clearly show the impediment to relative heat transfer, which occurs for pulsating numbers of 0.08 and 0.10, respectively. When the pulsating number is greater than 0.14 an enhancement in relative heat transfer occurs, and this may be seen in Figs. 6(*d*) and 6(*e*). Thus it is thought that, as the strength of the pulsating force is increased by gradually reducing the Reynolds number, the flow becomes destabilized, and this causes improved relative heat transfer.

Variations of the pulsating number can also be achieved by adjusting the frequency rather than by varying the Reynolds number at a fixed frequency. Tests were undertaken at pulsating numbers of 0.08 and 0.14 under these conditions, and Figs. 7(*a*) and 7(*b*) show typical variations of the relative Nusselt number ratio obtained with two different Reynolds number values. Because the pulsating number was fixed for each of these figures, the relative strength of the reciprocating flow to flow inertia was also fixed for the two different Reynolds numbers concerned. The data were found to collapse into separate lines of fixed Reynolds numbers. This suggests that the relative Nusselt number ratio is not independent of Reynolds number. Higher levels of relative Nusselt number ratio were found to consistently occur with lower Reynolds number values. The influence of flow pulsation depends, therefore, on the absolute strength of the inertial force for a specified value of the pulsating number.

It is clear from this investigation that flow pulsation has a significant effect on forced convection in this particular tube geometry. The effects of flow pulsation and buoyancy are interactive, and the relative changes in heat transfer, compared to the stationary case, are dependent on the axial location. To illustrate this overall effect, Figs. 8(*a*)–8(*i*) shows the variation of the relative Nusselt number ratio with pulsating number for all tests undertaken. All data points collected at a specific axial location were examined in a single graph. Note that, because the heat transfer features in Eq. (6), which is used for the normalization process, the relative Nusselt number ratio becomes unity when the test section is static. Figure 8 shows that the patterns of departure of the heat-transfer performance from the static case were similar in all cases. Additionally, significant data spreading is indicated, caused by Reynolds number and buoyancy effects. At each value of the pulsating number shown in Fig. 8, increasing the Reynolds number or the buoyancy number brought about reductions in the relative Nusselt number ratio.

Generally, it may be noted that at relatively low pulsation numbers (e.g., 0.08) heat transfer is reduced, particularly at high values of the buoyancy parameter. For example, the local Nusselt number could fall to about 53% of the corresponding static value. However, as the pulsating number is increased (e.g., 0.18) the heat transfer tends to recover, particularly at downstream locations. These particular reciprocation-driven heat transfer results could be due to variations in the vortical flow structures present in the channel. There is evidence to show that flow pulsation in curved and helical tubes may induce strong secondary flow cells which have temporal fluctuations linked to the frequency and amplitude of the pulsation; see Refs. [11,12]. The presence of the twisted tape also involved centrifugal instabilities in the flow, and these may have similarities to the temporal fluctuations found with the curved and helical tubes. Based on the data shown in Fig. 8, it appears that low-frequency reciprocation tends to stabilize the flow, whereas increasing the frequency at fixed amplitude tends to create stronger secondary flow levels because the nonlinear pulsating term in Eq. (1) is dominant. Secondary flows may also occur in reciprocating ducts due to the presence of extended surfaces [13,14], and these may further interact with the nonlinear flow pulsation and thus influence the heat transfer in comparison with smooth-walled channels. The nonlinear interaction of swirling flows and reciprocating forces creates synergistic effects in the presence of the twisted tape, and these effects are not necessarily additive.

Additionally, it has been shown that the effect of reciprocation is dependent on the axial location in the tube. As the development of hydrodynamic and thermal boundary layers occurs, the influence of the reciprocation gradually increases in the downstream direction. The most significant influence of reciprocation was found to occur in the latter part of the channel, where the flow is expected to be reasonably well developed.

The motivation for this work was the desire to improve the propulsive efficiency of marine Diesel engines by operating at lower engine speeds. Under these circumstances, cooling channels located in the pistons may well operate with fairly low pulsation numbers and impediments to the internal heat transfer may result from the piston oscillation. Therefore, careful attention to the design of these cooling systems is necessary to ensure reliable operation.

Conclusions

An experimental study has been performed to investigate turbulent flow heat transfer inside a reciprocating tube fitted with an internal twisted tape. The effect of flow pulsation and buoyancy has been particularly studied. As a result of the study the following salient points emerge.

1 With a static test section, the spatial variations of relative Nusselt number ratio (i.e., heat transfer compared to fully developed smooth-walled data) gradually increased in the downstream direction, measured from the entry plane, and tended to an asymptotic level as the flow became developed. With fully developed flow, the heat transfer in the test section was found to increase, with respect to smooth-walled channels, by a factor in the range 2.1–3.9. The relative Nusselt number ratio, for this static case, was Reynolds number dependent with the ratio decreasing with increasing Reynolds number.

2 The spatial distribution of heat transfer, produced with a reciprocating test section, was similar to the forms found with the stationary case. However, reciprocation may create an impediment or enhancement in heat transfer depending on the value of the pulsating number. Generally heat-transfer impediment occurred at low values of the pulsating number with the enhanced heat transfer occurring at high values of the pulsating number. The symmetry of the twisted flows caused by the tape did not appear to be changed by the reciprocation.

3 A strong buoyancy effect was detected. Increasing the buoyancy parameter tended to reduce the heat transfer. Also increasing the pulsating number enhanced the influence of buoyancy.

4 The relative Nusselt number ratio Nu/Nu_0 is not independent of Reynolds number. Higher levels of the relative Nusselt number ratio were found to occur consistently with lower Reynolds number values. Therefore, the influence of flow pulsation depends on the absolute strength of the inertial force for a specified value of the pulsating number.

5 At a specified axial location, low values of the pulsating number produced significant reductions in heat transfer when compared to the static case. This reduction of local heat transfer could be reversed as the pulsating number was increased. The most significant influence of the pulsation was experienced in the developed region of flow.

6 From the viewpoint of design piston cooling systems, it is essential to ensure that the range of pulsating numbers and buoyancy parameters produced does not create an environment of reduced heat transfer. Otherwise serious overheating of the pistons may occur.

Acknowledgments

This work was financially supported by the National Science Council, Taiwan, Republic of China, under Grant No. NSC-87-2611-E-006-034. My special appreciation to associate Professor, L. M. Su at NKIMT and Professor W. D. Morris at the University of Wales, Swansea, for their long-term support and assistance to my research work.

Nomenclature

English Symbols

C_{fin} = fin parameter
d = hydraulic diameter of test duct
H = pitch of twisted tape
k_f = thermal conductivity of fluid
k_m = thermal conductivity of twisted tape
$\bar{k}$ = unit vector in the downstream direction of main flow
Nu = reciprocating Nusselt number
Nu_0 = nonreciprocating Nusselt number
Nu_∞ = Nusselt number value for plain tube
Pu = pulsating number
Pr = Prandtl number
p^* = pressure field of reciprocating flow
q_f = convective heat flux
r = radius of crank arm
Re = Reynolds number
T_f = flow bulk temperature
T_w = wall temperature
W_m = mean through flow velocity
z = axial location
Z = dimensionless axial location (z/d)

Greek Symbols

β = thermal expansion coefficient of coolant
ρ = density of fluid
ω = angular velocity of rotating crank
υ = fluid kinetic viscosity
δ = thickness of twisted tape

References

[1] Date, A. W., 1974, ''Prediction of Fully-Developed Flow in a Tube Containing a Twisted-Tape,'' Int. J. Heat Mass Transf., **17**, pp. 845–859.
[2] Hong, S. W., and Bergles, A. E., 1976, ''Augmentation of Laminar Flow Heat Transfer in Tubes by Means of Twisted-Tape Inserts,'' ASME J. Heat Transfer., **98**, pp. 251–256.
[3] Marner, W. J., Bergles, A. E., and Chenoweth, J. M., 1983, ''On the Presentation of Performance Data for Enhanced Tubes Used in Shell-and-Tube Heat Exchangers,'' ASME J. Heat Transfer, **105**, pp. 358–365.
[4] Manglik, R. M., and Bergles, A. E., 1988, ''Laminar Flow Heat Transfer in a Semi-Circular Tube With Uniform Wall Temperature,'' Int. J. Heat Mass Transf., **31**, No. 3, pp. 625–636.
[5] Agarwal, S. K., and Rao, Raja M., 1996, ''Heat Transfer Augmentation for the Flow of a Viscous Liquid in Circular Tubes Using Twisted Tape Inserts,'' Int. J. Heat Mass Transf., **39**, No. 17, pp. 3547–3557.
[6] Chang, S. W., and Su, L. M., 1997, ''Influence of Reciprocating Motion on Heat Transfer Inside Ribbed Duct With Application to Piston Cooling in Marine Diesel Engine,'' J. Ship Res., **41**, No. 4, pp. 332–339.
[7] Nishimura, T., and Kojima, N., 1995, ''Mass Transfer Enhancement in a Symmetric Sinusoidal Wavy-Walled Channel for Pulsatile Flow,'' Int. J. Heat Mass Transf., **38**, No. 9, pp. 1719–1731.
[8] Hu, H. C., and Kelly, R. E., 1997, ''Stabilization of Longitudinal Vortex Instabilities by Means of Transverse Flow Oscillations,'' J. Phys. Fluids, **9**, No. 3, pp. 648–654.
[9] Editorial Board of ASME Heat Transfer, 1993, ''Journal of Heat Transfer Policy on Reporting Uncertainties in Experimental Measurements and Results,'' ASME J. Heat Transfer, **115**, pp. 5–6.
[10] Dittus, F. W., and Boelter, L. M. K., 1930, Univ. California Publs. J. Engn., **2**, pp. 443–461.
[11] Zabielski, L., and Mestel, A. J., 1998, ''Unsteady Blood Flow in a Helically Symmetric Pipe,'' J. Fluid Mech., **370**, pp. 321–345.
[12] Lyne, W. H., 1971, ''Unsteady Viscous Flow in Curved Pipe,'' J. Fluid Mech., **45**, pp. 13–31.
[13] Chang, S. W., Su, L. M., Yang, T. L., and Hwang, C. C., 1999, ''An Experimental Study of Heat Transfer in Reciprocating Square Duct Fitted With Ribs Skewed to the Flow,'' ASME J. Heat Transfer, **121**, pp. 232–236.
[14] Chang, S. W., Su, L. M., Hwang, C. C., and Yang, T. L., 1999, ''Heat Transfer in a Reciprocating Duct Fitted With Transverse Ribs,'' J. Exp. Heat Transfer, **12**, pp. 95–115.

Study of Using Oxygen-Enriched Combustion Air for Locomotive Diesel Engines

D. N. Assanis
The University of Michigan,
Ann Arbor, MI 48109

R. B. Poola

R. Sekar

Argonne National Laboratory,
Argonne, IL 60439

G. R. Cataldi
Association of American Railroads,
Washington, D.C. 20019

A thermodynamic simulation is used to study the effects of oxygen-enriched intake air on the performance and nitrogen oxide (NO) emissions of a locomotive diesel engine. The parasitic power of the air separation membrane required to supply the oxygen-enriched air is also estimated. For a given constraint on peak cylinder pressure, the gross and net power output of an engine operating under different levels of oxygen enrichment are compared with those obtained when a high-boost turbocharged engine is used. A 4 percent increase in peak cylinder pressure can result in an increase in net engine power of approximately 10 percent when intake air with an oxygen content of 28 percent by volume is used and fuel injection timing is retarded by 4 degrees. When the engine is turbocharged to a higher inlet boost, the same increase in peak cylinder pressure can improve power by only 4 percent. If part of the significantly higher exhaust enthalpies available as a result of oxygen enrichment is recovered, the power requirements of the air separator membrane can be met, resulting in substantial net power improvements. Oxygen enrichment with its attendant higher combustion temperatures, reduces emissions of particulates and visible smoke but increases NO emissions (by up to three times at 26 percent oxygen content). Therefore, exhaust gas after-treatment and heat recovery would be required if the full potential of oxygen enrichment for improving the performance of locomotive diesel engines is to be realized. [DOI: 10.1115/1.1290590]

Introduction

Both railroad operators and engine manufacturers face a major challenge if they are to meet the standards for emissions of smoke, particulates, unburned hydrocarbons, carbon monoxide, and oxides of nitrogen (NO_x) set forth in the U.S. Environmental Protection Agency's regulations for locomotive diesel engines. With recent advances in AC traction and electronics, railroads also face an increasing demand for higher motive power so payloads in freight locomotives can be increased. Argonne National Laboratory and the Association of American Railroads initiated a cooperative research and development agreement to study the application of oxygen-enriched intake air as one method to meet the above challenges.

A number of analytical and experimental studies (e.g., [1–4]) have demonstrated the benefits of using oxygen-enriched intake air in diesel engines. Increasing the oxygen content of a reacting fuel-oxidizer mixture leads to faster burn rates and the ability to burn more fuel at the same stoichiometry (oxygen-to-fuel ratio). These effects have the potential to increase the thermal efficiency and specific power output of a diesel engine. The power increase from a given displacement engine can be suitably exploited to increase the number of locomotive cars or freight. In addition, oxygen enrichment can also be considered as a way to reduce the loss in power output when locomotives operate in underground tunnels and at high altitudes.

When oxygen is added to the combustion air, emissions of visible smoke, particulates, and unburned hydrocarbons decrease significantly over a wide-load range as a result of the more complete combustion (e.g., [5]). Added oxygen also leads to shorter ignition delays and offers the potential for burning lower grade and nonpetroleum fuels [1,5]. However, the increased oxygen content in the combustion air is also expected to increase NO emission levels because of the higher combustion temperatures. The anticipated increase in NO emissions could be controlled by retarding fuel injection timing and by using other state-of-the-art concepts such as exhaust post-treatment with monatomic nitrogen and lean-NO_x catalysts. The objective of the present study is to address, in a systematic way, the key technical issues associated with applying oxygen-enrichment technology to locomotive diesel engines. These issues include (1) optimizing the level of oxygen enrichment in the intake air, (2) designing an air separation membrane and assessing its parasitic power requirements, and (3) exploring the merits of injection retard for controlling peak pressure and NO emissions.

An appropriately modified computer simulation of a turbocharged diesel engine is used to study the effects of oxygen-enriched intake air on the performance and NO emissions of a locomotive diesel engine. The physical and operating characteristics of an air separation membrane are briefly reviewed, and a simplified model is proposed to estimate the module size and parasitic power required to supply the desired oxygen-enrichment level in the intake air. On the basis of engine and membrane models, the net increase in brake power achieved by an engine equipped with an air separation membrane to supply the oxygen-enriched intake air are computed. Use of oxygen-enriched air is compared with turbocharging to a higher inlet boost as a means to increase power output. The effects of retarded fuel injection timing are also studied in conjunction with oxygen-enriched intake air. Finally, the potential for enhanced exhaust heat recovery when using oxygen enrichment is assessed. Other aspects of oxygen enrichment, such as a reduction in particulates, smoke, and unburned hydrocarbons and the ability to burn lower grade fuels, are not the subject of this paper but have been studied earlier (e.g., [1–3]).

Summary of Diesel Engine Simulation

The model used for this study is based on a zero-dimensional, thermodynamic simulation of a turbocharged diesel engine [6]. The parent code has been validated against test results from several turbocharged diesel engines, and was previously modified to

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division December 21, 1999; final revision received by the ASME Headquarters March 16, 2000. Technical Editor: H. Nelson.

allow for various levels of oxygen enrichment in the intake air and for operation with water-emulsified fuels [7]. This section briefly summarizes the main assumptions of the simulation. Additional details on the parent code can be found in Assanis and Heywood [6].

The diesel four-stroke cycle is treated as a sequence of continuous processes: intake, compression, combustion (including expansion), and exhaust. Quasi-steady, adiabatic, one-dimensional flow equations are used to predict mass flows past the intake and exhaust valves. The compression process is defined to include the ignition delay period (i.e., the time interval between the start of the injection process and the ignition point). The total length of the ignition delay can be either specified or predicted by using an Arrhenius expression that is based on the mean cylinder gas temperature and pressure during the delay period. Combustion is modeled as a uniformly distributed heat release process. The rate of heat release is specified by using an algebraic burning rate correlation. Heat transfer is included in all the engine processes. Convective heat transfer is modeled by using available engine correlations that are based on turbulent flow in pipes. The characteristic velocity and length scales required to evaluate these correlations are obtained from a mean and turbulent kinetic energy model. Radiative heat transfer, based on the predicted flame temperature, is added during combustion.

For this work, a heat release correlation that was developed earlier [8] based on experimental data for oxygen-enriched diesel combustion was used, i.e.,

$$\frac{\dot{m}_{f,b}(\theta)}{m_{f,o}} = \theta^{B_p} \cdot \exp(A_p - C_p \cdot \sqrt{\theta}) + \theta^{B_d} \cdot \exp(A_d - C_d \cdot \sqrt{\theta}), \tag{1}$$

where $\dot{m}_{f,b}$ is the instantaneous rate of fuel burning; $m_{f,o}$ is the total mass of fuel injected; θ is the crank angle after ignition; A, B, and C are adjustable parameters; and subscripts p and d correspond to the premixed and diffusion combustion phases, respectively. The adjustable parameters are determined by matching the experimental heat release rates against curve-fitted profiles generated by using the above heat release correlation (Eq. (1)). On the basis of measured heat release data for a Caterpillar, single-cylinder diesel engine [8] operating with an oxygen content in the range of 21 percent to 35 percent by volume and an overall equivalence ratio of about 0.5, the following values are recommended for the constants:

$$\begin{aligned} A_p &= 16.810 - 0.61539\,[\mathrm{O_2}] \\ B_p &= 26.625 - 0.70921\,[\mathrm{O_2}] \\ C_p &= 28.152 - 0.76602\,[\mathrm{O_2}] \\ A_d &= -6.0295 - 0.16920\,[\mathrm{O_2}] \\ B_d &= 3.4698 + 0.10564\,[\mathrm{O_2}] \\ C_d &= 1.8202 + 0.03195\,[\mathrm{O_2}]. \end{aligned} \tag{2}$$

In Eq. (2), the oxygen concentration, $[O_2]$, is to be taken as percentage.

The formation and destruction of thermal NO from atmospheric nitrogen have been studied widely on the basis of the extended Zel'dovich mechanism [9], i.e.,

$$\begin{aligned} \mathrm{O + N_2 = NO + N} \\ \mathrm{N + O_2 = NO + O} \\ \mathrm{N + OH = NO + H}. \end{aligned} \tag{3}$$

For the purpose of implementing the NO kinetics model in the diesel simulation, the combustion chamber is divided in two zones: a burned zone containing products of combustion and an unburned zone containing air and residual gas. At any instant throughout combustion, an incremental flux of fuel given by Eq. (1) that is accompanied by a stoichiometric flux of air is assumed to cross from the unburned zone to the burned zone. The products of combustion are assumed to be at the adiabatic flame temperature, resulting from combustion of a stoichiometric fuel-air mixture of specified oxygen content at the instantaneous unburned gas temperature and pressure (uniform within the combustion chamber). The kinetic rate of change of the concentration of NO in the postflame gases, along with appropriate expressions for its temperature-dependent rate constants, are obtained from Lavoie et al. [9]. The instantaneous concentrations of NO, O, O_2, OH, H, and N_2 are approximated by their equilibrium values at the specified burned zone temperature and pressure. On the basis of a knowledge of the instantaneous concentration of NO in the burned zone, its overall concentration in the cylinder can be computed by multiplying the ratio of moles in the stoichiometric burned core times the total number of moles in the cylinder.

Air Separation Membrane

Operating Principle and Modes. Air delivered to engines can be oxygen-enriched by selective permeation through nonporous, polymer membranes via a well-known ''solution-diffusion'' mechanism. In this mechanism, air molecules dissolve into the polymer and then diffuse across it. Figure 1 illustrates a typical prototype module built with hollow fibers, which resembles a shell-and-tube geometry. The ambient air is fed from the tube side at a higher pressure ($P_1 >$ ambient) and holds the shell side (P_2) at a lower pressure. The pressure differential ($P_1 - P_3$), pressure ratio (P_1/P_2), and concentration gradient across the membrane provide the driving force for the dissolution and diffusion of oxygen and nitrogen molecules across the membrane. Because for a given polymeric material, the intrinsic rate of dissolution-diffusion of oxygen is greater than that of nitrogen, oxygen diffuses more rapidly and becomes enriched in the permeate side at pressure P_2. The portion of the air that is swept out without crossing the membrane called retentate is filled with more nitrogen and comes out on the other side of the tube at pressure P_3 and is necessary to effect separation.

Air separation membrane units can be operated in one of three modes as illustrated in Fig. 2: vacuum, pressure, or mixed. In the vacuum mode, the feed air P_1 is pressurized to only slightly above atmospheric pressure (about 5 to 20 kPag, overcoming the tube flow losses), and a vacuum is maintained on the permeate P_2 of the membrane. The retentate P_3 is vented at atmospheric pressure. Oxygen concentration in permeate is a function of pressure ratio (P_1/P_2) between the feed and permeate. The vacuum mode is efficient because it gives a high oxygen concentration due to high-pressure ratio at a low differential pressure ($P_1 - P_3$). However, because of the limited differential pressure, the vacuum mode requires a larger membrane area for given flow rate than does the pressure mode. In the pressure mode, the feed air is typically pressurized (by an air compressor to about 2 to 3 atmospheres), while permeate is maintained at about atmospheric pressure. Higher driving forces are obtained in this mode because the differential pressures are higher than those of the vacuum mode, resulting in reduced membrane area requirements. However, the

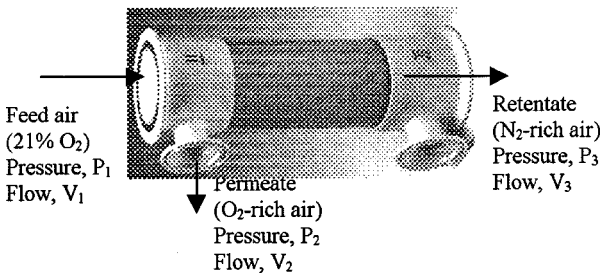


Fig. 1 Functional representation of a prototype membrane

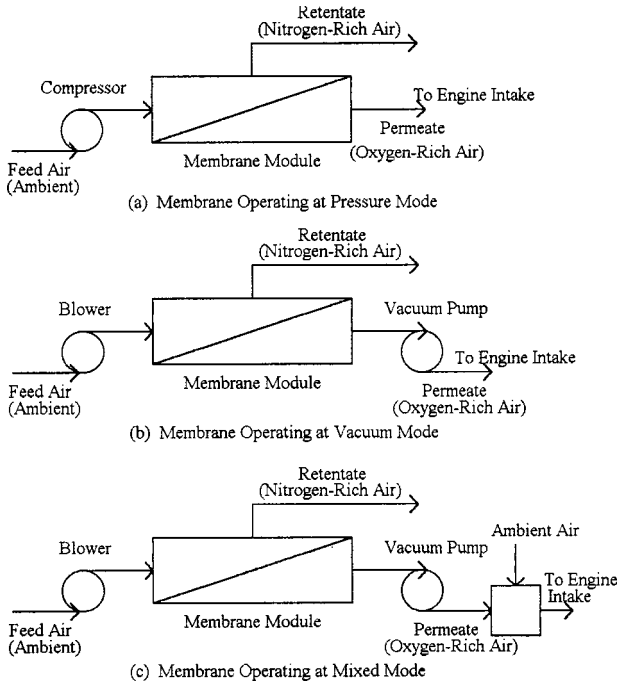


Fig. 2 Air separation membrane modes of operation

pressure mode is more energy intensive for oxygen-enriched product, because both permeate and retentate have to be compressed to higher pressures. Finally, in the mixed mode, the feed air is pressurized, and a partial vacuum is maintained on the permeate side to increase both the pressure ratio and differential pressure and thus the oxygen concentration.

Figure 3 illustrates one possible scheme of membrane operating in the mixed mode to supply oxygen-enriched air for a turbocharged, multi-cylinder, diesel engine.

Membrane Performance and Design. The performance of a membrane is characterized by its permeability and selectivity (intrinsic properties) and stage cut (or recovery). The overall permeability is a function of flow as well as of membrane composition. It can be described by Fick's law as follows [10,11]:

$$N_i = \frac{P_{ri} A \Delta P}{\delta}, \tag{4}$$

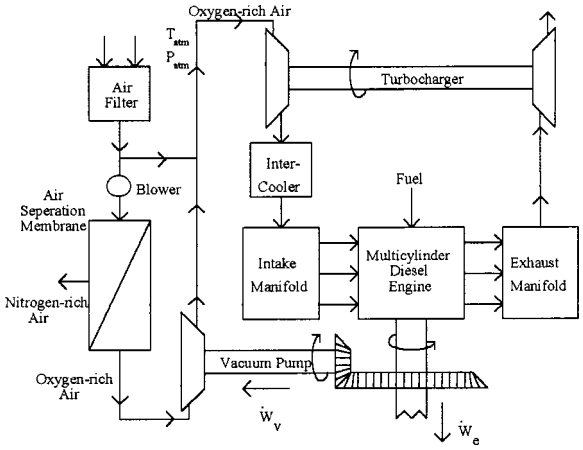


Fig. 3 Turbocharged diesel system configuration using oxygen-rich air supplied by air separation membrane

Table 1 Membrane properties and input parameters for two CMS materials

Properties	CMS-3	CMS-7
Permeability (cm^3 of STP-cm/cm^2-s-cm of Hg x 10^{10})		
Oxygen	340	990
Nitrogen	130	480
Selectivity (O_2/N_2)	2.71	2.05
Membrane skin thickness (μm)	0.1	0.1
Hollow fiber dimensions (OD/ID), μm	660/500	660/500
Packing density	0.5	0.5

where N_i is the flow rate of gas i (cm^3 of STP/s); P_r is the intrinsic permeability of gas i in the membrane (cm^3 of STP-cm/cm^2-s-cm of Hg); A is membrane area (cm^2); δ is the membrane separating barrier thickness (cm); and ΔP is the transmembrane partial pressure difference of gas i (cm of Hg).

Equation (4) indicates that the degree of separation between the gases clearly depends on the relative permeabilities of the gases to be separated. This ratio of gas permeabilities is known as selectivity or separation factor. The separation factor (α) between oxygen and nitrogen can be calculated as follows:

$$\alpha = \frac{P_{rO}}{P_{rN}}. \tag{5}$$

The larger the value of the separation factor is, the better is the separation. The stage cut (φ or recovery) of an air separator is another important measure of performance. It is simply the permeate flow rate divided by the feed flow rate.

The performance of an air separator depends on several other parameters such as membrane polymer structure, skin thickness, geometry of the fibers, fiber dimensions, flow pattern, feed direction, feed conditions, cartridge type, packaging density, and arrangement of separators. Details are provided in Winston Ho and Sirkar [12] and Koros and Chern [13]. The selection of a membrane to achieve the desired oxygen-enriched airflow is evaluated in terms of the power required to maintain the pressure ratio and differential pressure across the membrane and the amount of space it occupies.

Various mathematical models and calculation methods for designing hollow-fiber membranes have been reported in the literature [14,15]. The development of a model to study all the performance variables and their tradeoffs would be quite exhaustive and beyond the scope of this paper. In the present work, a Du Pont proprietary code was made available to establish a relationship between the permeate oxygen concentration and stage cut under the specified input conditions listed in Table 1 and at several pressure ratios between the feed and permeate (ranging from 1.5 to 4) in the following form:

$$[O_2] = A\varphi^3 + B\varphi^2 + C\varphi + D, \tag{6}$$

where $[O_2]$ is permeate oxygen concentration expressed as percent volume; A, B, C, and D are third-order polynomial coefficients as a function of the pressure ratio between the feed and permeate; and φ is the stage cut. Equation (6) was used to compute the stage cut at several desired permeate oxygen concentrations. Perfluoro-2-2-dimethyl-1-3 dioxole copolymerized with tetrafluoroethylene, being developed at Compact Membrane Systems, Inc., (CMS) is being considered as a membrane material [16] in the present model. From the stage cut, feed flow rates corresponding to permeate flow rate and oxygen concentration can be computed. The pumping work can be computed from the pressure ratio between the feed and permeate and from the feed and permeate flow rates.

In the absence of certain key design data, including the mean partial pressure of oxygen in the module and the fiber geometry, it is difficult to estimate membrane size. However, an attempt was

made to demonstrate the influence of membrane properties and other design variables on the membrane's size and power requirements. For comparison purposes, a pressure ratio of 3.0 between the feed and permeate across the membrane was assumed, and the mixed mode was considered as a mode of operation (the mixing stream's concentration was 28 percent by volume and its flow was about 72 percent of the total engine airflow). The module size also varies with the active length of the fibers, which is, in turn, a function of pressure drop across the fiber tubes. For purposes of this study, the active length was arbitrarily selected to be 0.9 or 1.2 m. Finally, to calculate the membrane surface area by using Eq. (4), the end pressures of oxygen across the feed and permeate were employed instead of the mean partial pressure of oxygen. Membrane volume and module size was computed on the basis of fiber geometry and packing density.

Results and Discussion

Engine Configuration. To explore the effect of various oxygen-enrichment levels in the intake air on engine performance and emissions, a representative 12-cylinder, GE 12-7FDL locomotive diesel engine was selected, and the diesel engine simulation was applied to model its performance. The specifications of the engine geometry are given in Table 2. Locomotive diesel engines are operated at a series of fixed settings (throttle notches 1 through 8), with each one providing a constant amount of power at a governed engine speed. Since turbomachinery maps were not available, the intake and exhaust manifolds were treated as constant pressure and temperature plenums. The intake plenum conditions were specified; the exhaust plenum conditions were predicted by means of ideal thermodynamic models of the compressor and turbine with specified efficiencies. The actual intake and exhaust valve events and the manifold control volumes and surface areas were obtained from a previous study conducted at the Southwest Research Institute [17].

Model Calibration and Validation. The diesel engine model was calibrated through tuning of its adjustable constants related to the computation of the discharge and heat transfer coefficients, and overall friction for operation at notch 8 (full load) with standard air. The frictional losses were estimated on the basis of the Millington and Hartles correlation [18]. The heat release correlation of Eq. (1) was used; the values assumed for the adjustable constants were those predicted by Eq. (2) for 21 percent oxygen content. After airflow and other available brake quantities were matched within two percent at notch 8, the ability of the simulation to predict performance at six other notch positions (without any further tuning) was explored. Measured data for the GE engine at the various notch positions included input parameters such as engine speed, mass of fuel injected, intake and exhaust temperature and pressure, and volumetric efficiency. Output parameters included brake power, brake-specific fuel consumption (bsfc), brake mean effective pressure (bmep), air mass flow, air/fuel ratio, and brake thermal efficiency.

Figure 4 compares predicted brake power, bsfc, bmep, and air mass flow rate with corresponding measurements at all six notch positions. Over the range of notch positions examined, the diesel engine simulation underpredicted brake power, bmep, and air flow by about two percent to seven percent, while it correspondingly overpredicted bsfc by about two percent to seven percent. Although agreement between baseline engine predictions and data could be improved by further tuning the model constants, it was felt that such an exercise was not warranted. First, a complete and reliable experimental database at all notch positions (specifically pressure and heat release rate) was not available when the diesel engine simulation was being matched to the engine. Further, the present study is intended to assess relative performance changes when the intake air for the engine is supplied with various oxygen levels, and not to focus on baseline engine performance predictions.

Table 2 Specifications for GE locomotive diesel engine

Bore (mm)	228.6
Stroke (mm)	266.7
Connecting rod length (mm)	589.7
Compression ratio	12.7:1
Displacement volume (L)	10.9
Number of cylinders	12
Number of valves	48
Rated speed (rpm)	1,050
Rated power (kW)	1,864
Standard injection timing (bTDC)	24

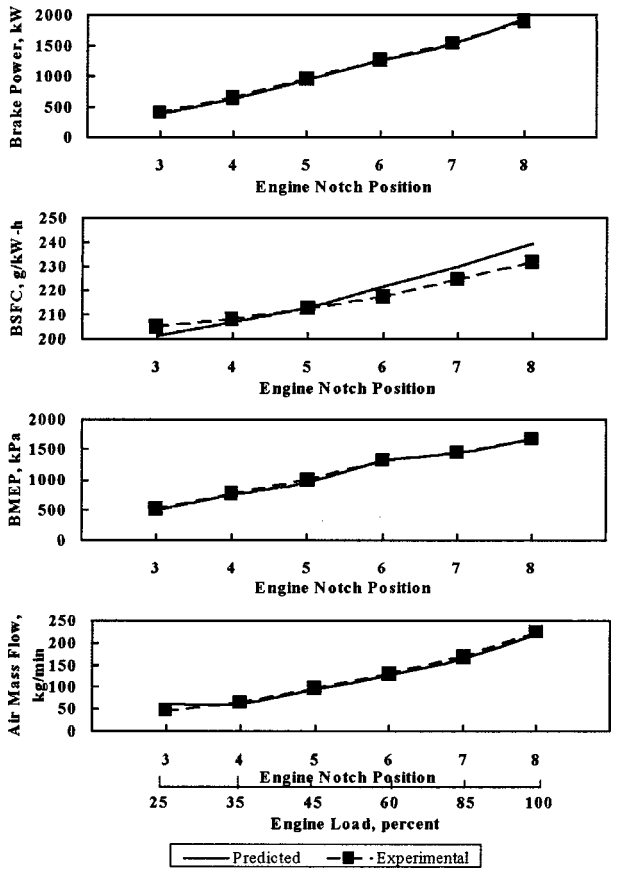


Fig. 4 Validation of diesel simulation with experimental results

Engine Gross Performance Under Various Oxygen Enrichment levels. A constant oxygen-to-fuel ratio was used as the basis to compare engine performance under different levels of oxygen enrichment in the intake air. The mass of fuel injected per cylinder per cycle was increased proportionally to the oxygen level in the intake air to maintain a constant oxygen-to-fuel ratio. The gross brake power, bsfc, bmep, and peak cylinder pressure obtained when oxygen levels ranged from 21 percent to 35 percent (by volume) are shown in Fig. 5 at all six notch positions. The amounts of power that would be required by the membrane to supply the desired intake air oxygen concentrations are not captured in the gross performance estimates but will be accounted for in a subsequent section to arrive at net performance improvements. The model predictions indicate that cylinder (gross) brake power significantly increases when the intake air oxygen concentration increases from 21 percent to 35 percent. A substantial (10 percent) output increase can be achieved as the result of a relatively small increase in oxygen content to 23 percent by volume,

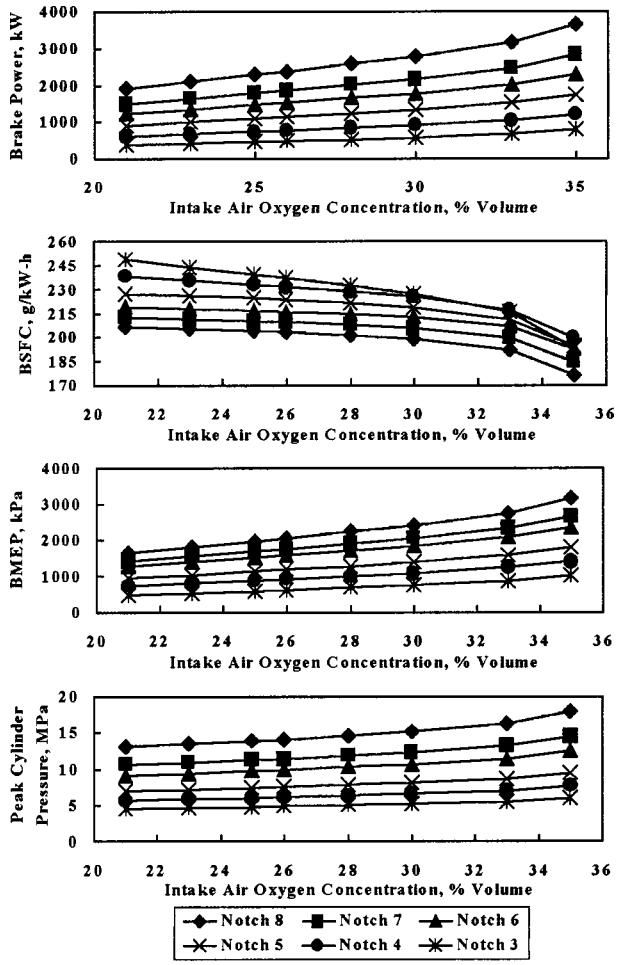


Fig. 5 Effects of intake air oxygen-enrichment on engine performance at various notch positions

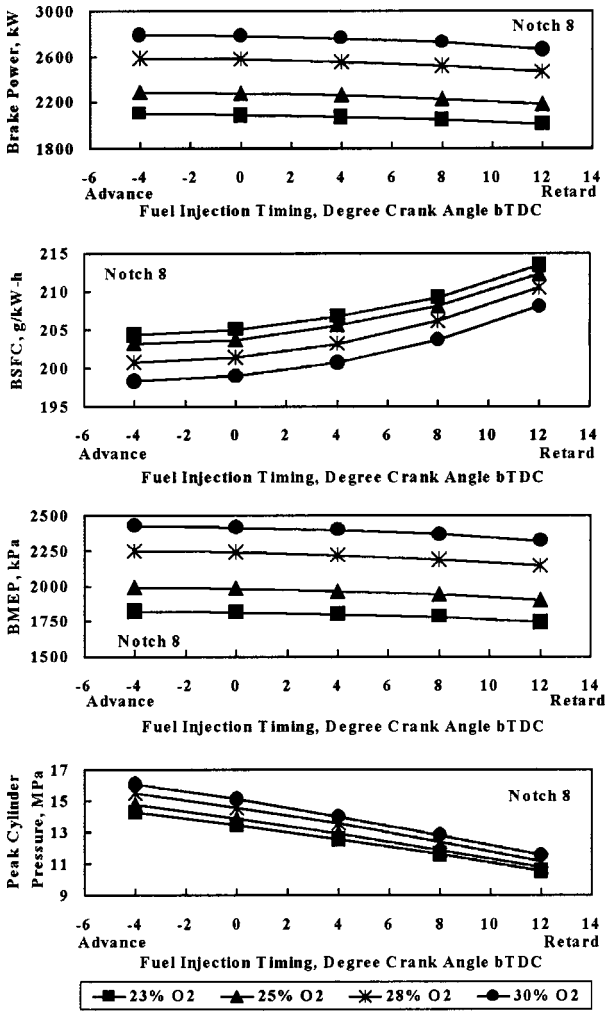


Fig. 6 Effects of fuel injection timing on performance characteristics at various oxygen-enrichment levels

while up to 90 percent power increase can be achieved when the oxygen content is increased to 35 percent. When the intake oxygen content was increased from 23 percent to 35 percent, the cylinder brake output increased from about 10 percent to 90 percent at notch positions 8 through 5. At lower notch positions (4 and 3), the cylinder output increased even more, by approximately 12 percent and 110 percent at oxygen-enrichment levels of 23 percent and 35 percent, respectively. The cylinder output is increased when oxygen-enriched air is used because of burning additional fuel, which is proportional to increase in the intake air oxygen concentration. The enhanced power output resulting from oxygen enrichment was accompanied by higher bmep and lower bsfc. At notch 8, the bsfc decreased from 206 to 176 g/kW-h (about 14 percent reduction) when the intake air oxygen level was increased from 21 percent to 35 percent. A similar reduction in bsfc was observed at other notch positions, more predominantly at lower notch positions (4 and 3). The implied thermal efficiency improvements are attributed to faster burn rates, particularly during the diffusion phase of combustion [2,7]. Despite the advantages of lower bsfc, higher cylinder output, and higher bmep, the peak cylinder pressures were higher by about three percent to 35 percent when the intake air oxygen level was increased from 23 percent to 35 percent, respectively, over ambient air. However, the increase in peak cylinder pressure was smaller than the increase in cylinder output. This feature of the oxygen-enriched engine is attractive, particularly because some of the other techniques for increasing power output (increased compression ratio, high boost turbocharging) typically yield power improvements proportional to peak pressure increases.

Effects of Fuel Injection Timing with Oxygen-Enrichment. Oxygen-enriched combustion leads to shorter ignition delays, higher diffusion burn rates, and faster completion of the combustion process [2]. These effects should permit the fuel injection timing to be retarded to contain the peak cylinder pressure within a specified ceiling while still allowing some of the performance benefits resulting from oxygen enrichment to be realized. Retarded fuel injection timing should also help to lower combustion temperatures and NO formation levels associated with oxygen enrichment.

Figure 6 shows the effects of fuel injection timing on the performance characteristics of an engine operating at full load (notch 8) with intake air at various oxygen-enrichment levels (23 percent to 30 percent by volume). As a result of retarding the fuel injection timing by four crank angle (CA) degrees, the deterioration in brake power, bsfc, and bmep was marginal (<1 percent) compared with the decrease in peak cylinder pressure (about 7 percent) at any given oxygen-enrichment level. Retarding the injection timing further (up to 12 CA degrees) combined with oxygen enrichment resulted in a reduction in the engine power output and bmep of about 4 percent and an increase in bsfc of about 4 percent when compared to standard injection timing of the respective oxygen-enrichment cases. However, the corresponding peak cylinder pressure decreased by approximately 22 percent, almost uniformly within the range of oxygen contents examined. When fuel injection timing is retarded to 12 CA degrees, oxygen-enriched combustion air provided higher gross power and bmep, and lower

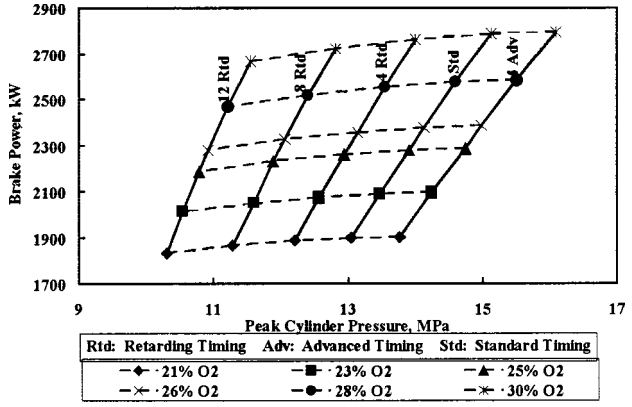


Fig. 7 Fuel injection timing map: brake power versus peak cylinder pressure at various intake air oxygen-enrichment levels

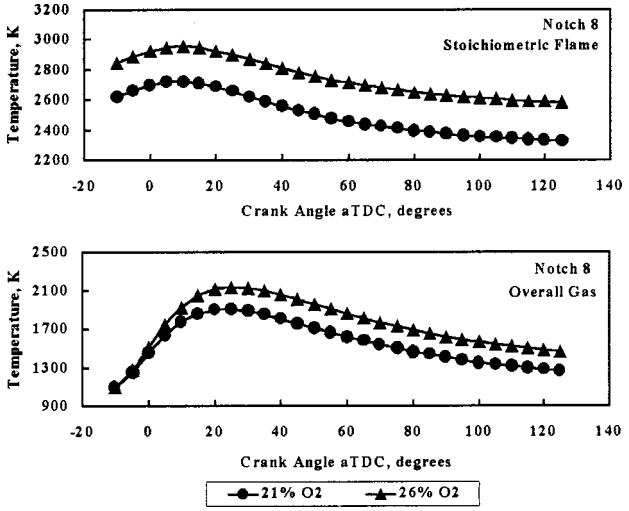


Fig. 8 Overall combustion and stoichiometric core-flame temperatures with intake air oxygen-enrichment

peak cylinder pressures when compared to those obtained using ambient air operating at standard injection timing. For example, when 25 percent oxygen-enriched air is used at 12 degrees CA retarded injection timing, the gross power is increased by about 15 percent, bmep is increased by about six percent, bsfc is increased by about three percent and peak cylinder pressure is decreased by about 17 percent when compared to those obtained using ambient air operating at standard injection timing.

The additional oxygen available in the combustion air shortened the ignition delay of the fuel-air mixture and thus the magnitude of the premixed heat release as captured by the heat release correlation. In addition, the faster diffusion burn rates still allowed combustion to proceed in an optimal manner, so it reached its peak within 10 to 16 CA degrees after TDC for the optimal retarding of timing. Figure 7 illustrates an injection-timing map showing the relationship between brake power and peak cylinder pressure for various oxygen levels in the intake air. By retarding fuel injection timing between 4 and 8 CA degrees for the various oxygen-enrichment levels, the peak cylinder pressure can be brought down to that achieved when standard air is used. The corresponding decrease in brake power and increase in bsfc at any particular oxygen-enrichment level was only about three percent. For example, at a constant peak cylinder pressure (like that of the base engine), the cylinder brake output increased by about 45 percent when the intake air oxygen level was increased from 21 percent to 30 percent and the fuel injection timing was retarded by about 8 CA degrees. The higher cylinder brake output at the same or a marginally higher peak cylinder pressure is clearly one advantage of oxygen-enriched combustion over other methods of increasing the power output from a given engine displacement.

NO Emissions. Higher post-flame temperatures and oxygen concentrations during the combustion process result in higher NO formation rates [19]. For a given combustion chamber, NO emissions correlate with variations in the stoichiometric flame temperature, which is dependent on inlet pressure and temperature and the chemical composition of the fuel and oxidant [20]. In the present work, variations in the stoichiometric, adiabatic flame temperature with changes in the oxygen concentration in the combustion air were studied by keeping all other influencing parameters (inlet pressure, temperature, and fuel composition) constant. The instantaneous traces of adiabatic flame temperature and overall, bulk gas temperature when 21 percent and 26 percent oxygen concentrations were used are depicted in Fig. 8. When the intake oxygen concentration was increased from 21 percent to 26 percent (by volume), the peak stoichiometric flame temperatures and the overall gas temperatures were higher by about 240 K. This result can be readily explained, since the extra number of oxygen molecules in oxygen-enriched combustion displace nonreacting nitrogen molecules, which normally act as a diluent and cool down the flame. This result corroborates the correlation of adiabatic flame temperature with intake air effects, as established by several other researchers [20,21].

The NO formation histories during the combustion period are shown in Fig. 9 for both ambient air and 26 percent oxygen-enriched intake air. Instantaneous NO concentrations are shown as a fraction of the contents in the burned core zone and the overall cylinder. The instantaneous NO concentrations were considerably higher with 26 percent oxygen-enriched intake air than with standard air. As combustion temperatures dropped, the NO concentrations froze at concentrations higher than equilibrium at the given temperature and pressure. The cycle-integrated, NO concentrations (as a fraction of the adiabatic burned gas and the overall gas) are shown in Fig. 10 for various oxygen-enrichment levels. The NO emissions increased significantly when the intake air oxygen concentrations were increased. For instance, when the intake oxygen was increased from 21 percent to 26 percent, the cylinder-out NO emissions increased by about 2100 parts per million (ppm). During operation at 30 percent oxygen content, the NO emissions increased by a factor of three, from 1600 to 5000 ppm. This increase in NO concentration corroborates previously published results on the effects of oxygen enrichment on NO emissions [5,22].

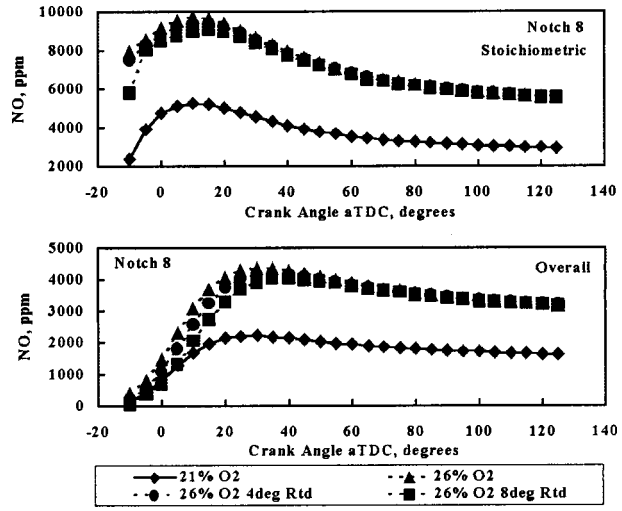


Fig. 9 Effects of oxygen and injection timing on NO emissions in stoichiometric and overall combustion regions

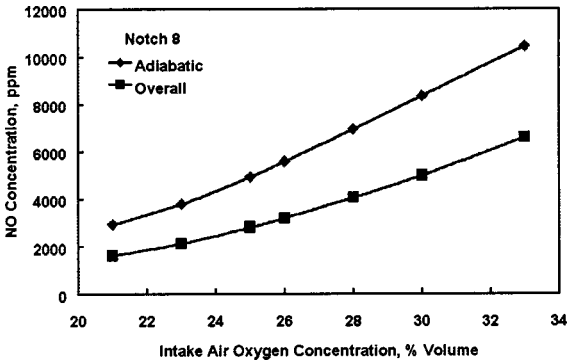


Fig. 10 Nitric oxide concentrations at different intake air oxygen-enrichment levels

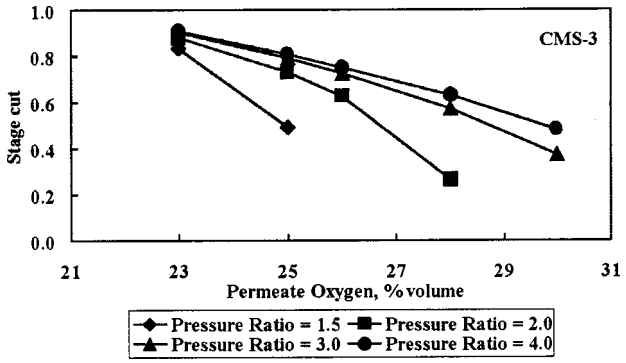


Fig. 11 Variation of permeate oxygen concentration with stage cut at various pressure ratios for CMS-3 membrane material

Higher NO emissions, which are associated with higher combustion temperatures, are one of the major drawbacks of oxygen-enrichment technology. As a possible remedy, fuel injection timing was varied to examine whether it could influence the NO formation associated with oxygen enrichment. When the injection timing was retarded, the peak cylinder pressure and maximum overall in-cylinder gas temperatures decreased. However, computations showed that engine-out NO emissions decreased by only 50 to 100 ppm when fuel injection timing was retarded 4 to 12 CA degrees and oxygen-enriched intake air was used. Clearly, the NO emission reductions obtained as a result of retarding fuel injection timing were marginal compared with the NO emission increases obtained when the oxygen concentration was increased from 21 percent to 26 percent. This result can be attributed to the fact that retarding injection timing does little to affect flame temperature, which is dominantly impacted by an increased oxygen concentration in the air. It appears that other control technologies are needed to overcome the NO penalty associated with oxygen enrichment. Some methods that have been explored to keep NO levels reasonably low when oxygen-enriched intake air is used include adding water in either the intake air or the fuel [7] and using monatomic nitrogen induced by a pulsed arc, as a post-treatment device [23].

Membrane Size and Power Requirements. An estimation of the power required by the membrane to supply the desired level of oxygen-enrichment is necessary to assess the net power improvement obtained from an engine operating with oxygen-enriched intake air. The simplified membrane model was used to compute the membrane power and size required for an engine operating at full load and different intake air oxygenenrichment levels. The predicted variation in permeate oxygen concentration with stage cut at various feed-to-permeate pressure ratios is shown in Fig. 11 for CMS-3 membrane material. To yield high stage cut and high purity, both selectivity and permeability should be high; these properties are difficult to achieve because of the inherent tradeoffs between membrane permeability and selectivity. Data on the feed flow rate, permeate flow rate, and pressure ratios across the module were used to obtain the membrane power data.

The total isentropic work required to drive the membrane module to produce the desired oxygen concentration and permeate flow rate at engine full load condition (supplying 3.31 m^3/s airflow) under the three alternative operating modes are shown in Fig. 12. The pressure mode is clearly the most power intensive because the pumping work is done on the feed air. On the other hand, the mixed mode requires the least power for moderate levels (up to 26 percent) of permeate oxygen. At higher permeate oxygen levels, the vacuum mode requires less power than the other two modes; mainly because the purity of the oxygen concentration of the mixing stream is limited to 30 percent or less (by volume) when either CMS-3 or CMS-7 membrane material is used. When vacuum mode is used, the isentropic work required producing oxygen concentration from 23 percent to 25 percent and then from 26 percent to 28 percent is small (18 and 47 kW, respectively). The minimum amount of isentropic work required to yield a prescribed level of permeate oxygen under full-load engine operation (notch 8) is shown in Fig. 13 for both CMS materials. CMS-3 requires less power because its selectivity and stage cut are relatively higher than those of CMS-7 membrane. In practice, the pumping losses should be accounted to realize the net power requirements of the membrane to supply desired oxygen enriched airflow.

Table 3 compares some membrane performance parameters, including surface area, module volume, and isentropic work, required supplying 26 percent oxygen-enriched air to a 12-cylinder locomotive engine operating at full load. The CMS-3 module occupied more space but required less power than the CMS-7 module. The tradeoff between membrane size and power is apparent. This tradeoff is inherent in membranes because of their intrinsic

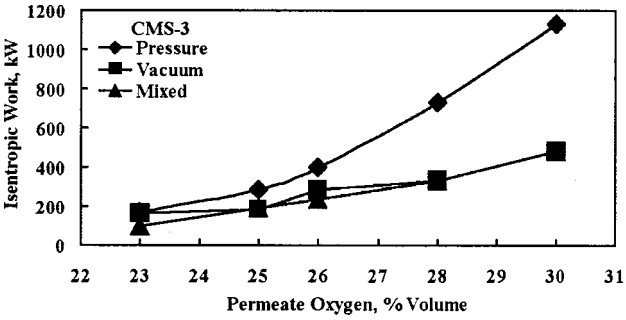


Fig. 12 Total isentropic work required to drive membrane under three operating modes

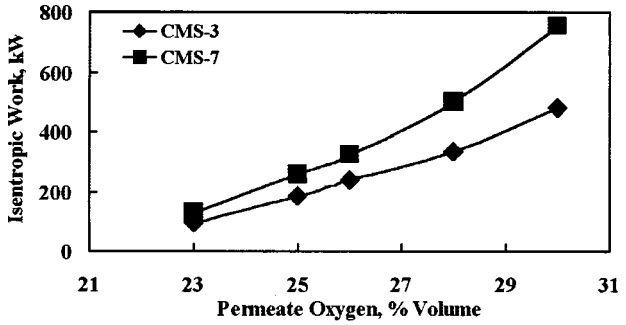


Fig. 13 Minimum isentropic work required by membrane module at various oxygen-enrichment levels

Table 3 Design estimates of membrane size and power requirements for two different materials operating under mixed mode to supply 26% oxygen in the intake air of an engine at full load

Material	Membrane surface area (m^2)	Active fiber length (m)	No. of fibers (10^3)	Module volume (m^3)	Approximate cylindrical module dimensions (L x D) (m)	Membrane isentropic work (kW)
CMS-3	2,200	0.914 1.219	1,532 1,150	1.05 0.79	0.914 x 1.219 1.219 x 0.914	237
CMS-7	755	0.914 1.219	527 395	0.36 0.27	0.914 x 0.7 1.219 x 0.52	321

properties (permeability versus selectivity). For example, CMS-3 has a relatively low permeability but high selectivity, and thus requires less power but a larger module size. Conversely, CMS-7 has a relatively high permeability but low selectivity, and thus required more power but a smaller module size. The tradeoffs among membrane properties and operating modes suggest that a detailed optimization study should be conducted to arrive at the smallest possible size and lowest amount of power required for any given application.

Performance of the Engine with Membrane Supplying Oxygen-Enriched Air. The net work required by an air separator membrane to supply the desired oxygen concentration in the engine intake air have to be considered to evaluate the net brake output obtained from the engine. Computation of the parasitic power requirements of an air separator module involves optimizing several physical and operating characteristics of the membrane. Because design information is limited and there is no detailed membrane model, the accuracy of estimating membrane power requirements is compromised. However, for the preliminary membrane design arrived at in the previous section, membrane power requirements were estimated as a function of the permeate oxygen level for two candidate materials (Fig. 13). Assuming that an air separator module made of CMS-3 membrane material and operating under the mixed mode supplies the desired oxygen concentration in the intake air, the net brake power output (cylinder brake power minus net membrane work) was computed at full load, as shown in Fig. 14. To obtain the net membrane work, 70 percent mechanical efficiency was assumed for both the vacuum pump and blower. On the basis of the present membrane model, the net engine power improvements were found to be about three percent to ten percent when the intake air oxygen was increased to 23 percent and 30 percent, respectively. The net brake power improvements obtained from the engine that used oxygen-enriched intake air were considerably smaller than the gross cylinder output (without allowing for membrane work). Any reduction in membrane net power requirements would directly enhance the net power improvements. However, for a detailed design and ultimate assessment of the membrane potential, more accurate computations of membrane power requirements should be carried out. They should consider all tradeoffs among membrane intrinsic properties and operating conditions.

Oxygen Enrichment Versus Turbocharging. Turbocharging is a proven, practical method to obtain a higher power output than the natural aspiration from a given engine displacement. However, there are practical limits to boosting intake manifold pressure before the complexity and losses associated with compressing the air are increased considerably (e.g., need for two-stage turbocharging and intercooling). Use of oxygenenriched intake air can be considered as an alternative or a complementary method to boost the power output. In the present work, an attempt was made to quantify the difference between highly boosted turbocharging versus oxygen enrichment plus regular boost, in terms of the brake power obtainable at any specified peak cylinder pressure.

To admit the same mass of oxygen per cycle as that which occurs with oxygen enrichment, turbocharging requires a considerably higher intake manifold pressure or lower temperatures. At full load, turbocharging requires the intake manifold pressure to be increased from 241 to 393 kPa as the oxygen content increased from 21 percent to 35 percent by volume in the intake air, respectively. The same mass of fuel is injected in both cases, so that they operate with the same oxygen-to-fuel ratio. The considerably increased compression work resulting from the highly boosted air and its extra mass of nitrogen are reduced in the case of oxygen enrichment. As a result, the gross cylinder power obtained is considerably higher with oxygen enrichment than with high-boost turbocharging at any constant peak cylinder pressure (see Fig. 15). When the net membrane power is subtracted from the gross cylinder output, the difference between oxygen enrichment and turbocharging appears to be minimal for operation at the base injection timing (Fig. 15). However, an oxygen-enriched engine, with its shorter delay and faster diffusion burn period, can operate with an injection timing that is retarded so that peak cylinder pressure is reduced, as discussed previously. When this strategy is followed, the net brake power obtainable at any constant peak cyl-

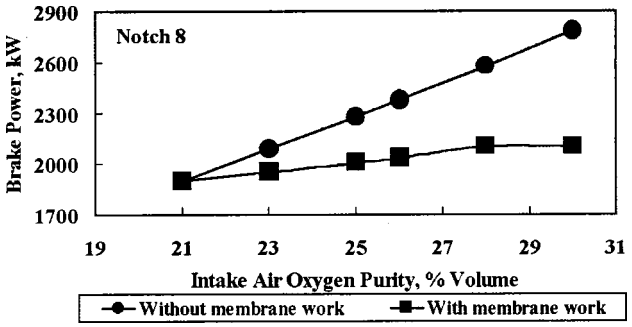


Fig. 14 Potential of brake power enhancement with oxygen-enrichment at full load

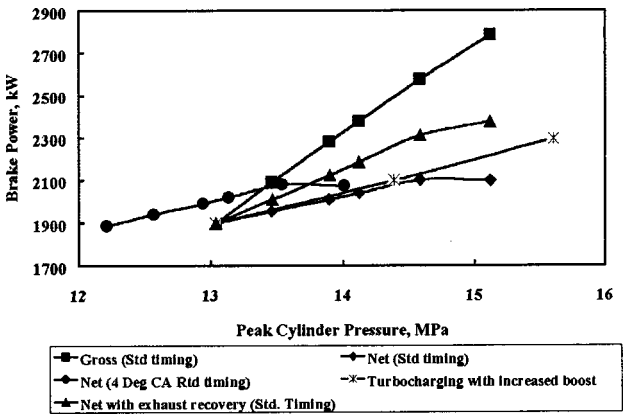


Fig. 15 Potential of oxygen-enrichment over turbocharging with increased boost to enhance the power output

inder pressure is higher with oxygen enrichment than with high-boost turbocharging. For instance, by using 28 percent oxygen concentration and fuel injection timing retarded by 4 CA degrees, a four percent increase in peak cylinder pressure could result in an increase in engine net power of approximately 10 percent, compared with an increase of only 4 percent by using turbocharging. In general, net power enhancement with increased turbocharging is fairly proportional to the increase in peak cylinder pressure, whereas the benefit is higher with oxygen enrichment.

At a constant oxygen-to-fuel ratio, exhaust gas temperatures at the turbine exit were considerably higher (100 to 500 K) with oxygen enrichment than with turbocharging because of the higher flame temperatures of the former. If part of the additional exhaust energy what is available with oxygen enrichment were recovered, the net power improvement of the oxygenenriched engine would be more attractive. To quantify the expected improvement, the difference in exhaust enthalpy past the turbine under operation with oxygen enrichment versus increased inlet boost turbocharging was assumed to be recovered by using an organic rankine cycle with 20 percent overall thermal efficiency. The additional power obtained from the recovered exhaust energy was then added to the net brake power of the oxygen-enriched engine, as shown in Fig. 15. For a 12 percent increase in peak cylinder pressure, the net power obtained by using 28 percent oxygen-enriched air with exhaust recovery increased from 10 percent to 22 percent, whereas the power improvement that resulted from using an increased inlet boost was only 13 percent. Hence, it is essential to recover part of the incremental exhaust energy with oxygen enrichment so that the net power improvements from a given engine are competitive with other conventional methods.

Summary and Conclusions

A thermodynamic diesel engine simulation and a simplified model of the air separator membrane were used to study the effects of different intake air oxygen-enrichment levels on the performance and NO emissions of a locomotive diesel engine. The following conclusions can be drawn from this investigation:

1 At notch 8 (full load), when the intake air oxygen concentration was increased from 21 percent to 35 percent (by volume) while a constant oxygen-to-fuel ratio was maintained in the combustion mixture, the cylinder (gross) brake output and brake mean effective pressure increased by up to 90 percent, and gross brake-specific fuel consumption decreased by down to 15 percent. However, the corresponding peak cylinder pressure increased by up to 35 percent.

2 Retarding the fuel injection timing when oxygen-enriched intake air is used has a beneficial effect with respect to peak cylinder pressure without imposing a severe penalty on power and fuel consumption. For intake air oxygen levels varied between 23 percent and 30 percent (by volume) and with retarded fuel injection timing (up to 12 deg crank angle), the peak cylinder pressure was decreased by up to 22 percent, with only a marginal penalty (on the order of 4 percent) in power and fuel consumption.

3 NO emissions computed on the basis of adiabatic flame temperature increased by up to 4 times when the intake air oxygen concentration was increased from 21 percent to 35 percent. Although retarding fuel injection timing can somewhat reduce NO emissions associated with oxygen-enriched combustion, another post-treatment device is needed to comply with the NO emission standards.

4 To provide intake air with 23 percent to 30 percent oxygen content (by volume) for a 12-cylinder locomotive diesel engine operating at full load requires isentropic work between 112 to 485 kW to drive the auxiliary equipment; the corresponding membrane area is estimated to occupy between 0.28 and 0.85 m^3, depending on the desired level of oxygen enrichment. However, an extensive optimization of several membrane parameters is necessary to refine the estimates of membrane power and size for a desired enrichment level and airflow.

5 On the basis of a preliminary membrane design and its power requirements, the net brake power improvement obtained from an engine operating with oxygen-enriched air (from 23 percent to 30 percent by volume) supplied by an air separation membrane was only between 3 percent and 10 percent. For the same inlet conditions, the cylinder brake output (without the parasitic membrane work) increased substantially, i.e., between 10 percent and 46 percent. Any reduction in net membrane power requirements has an obvious direct impact on net power improvement.

6 When peak cylinder pressure is controlled by retarding fuel injection timing, oxygen enrichment has the potential to result in greater improvements in brake power than those obtained by using other methods such as high-boost turbocharging. For instance, a 4 percent increase in peak cylinder pressure can result in a 10 percent increase in net engine power with oxygen enrichment but only a 4 percent increase with high-boost turbocharging. In addition, recovering part of the increased exhaust energy through a bottoming cycle can make the oxygen-enriched concept far more attractive than other methods of boosting power at a constant peak cylinder pressure.

7 Since the required modifications to the intake system are not complex, a membrane device can be retrofitted to in-use locomotives. Oxygen-enrichment technology promises to provide a significant power boost on sudden demand or continuously, alleviate concerns about certain exhaust emissions (visible smoke, particulates, and unburned hydrocarbons), and enable the use of lower grade fuels.

Acknowledgments

The authors acknowledge the support of the U.S. Department of Energy, Office of Energy Research, under contract W-31-109-Eng-38. The contributions of V.O. Markworth, Southwest Research Institute, in supplying experimental data are greatly appreciated. The authors thank Stuart Nemser, Compact Membrane Systems, Inc., for providing details on membrane properties and design. The contributions of Dr. Zoran Filipi in simulating the baseline locomotive engine are gratefully acknowledged. Assistance from Anthony Markel and Kevin Stork while the manuscript was being prepared is very much appreciated.

References

[1] Sekar, R. R., Marr, W. W., Cole, R. L., Marciniak, T. J., and Schaus, J. E., 1990, ''Diesel Engine Experiments With Oxygen Enrichment, Water Addition, and Lower-Grade Fuel,'' Twenty-Fifth Intersociety Energy Conversion Engineering Conference, Reno, Nevada.

[2] Sekar, R. R., Marr, W. W., Assanis, D. N., Cole, R. L., Marciniak, T. J., and Schaus, J. E., 1991, ''Oxygen-Enriched Diesel Engine Performance: A Comparison of Analytical and Experimental Results,'' ASME J. Eng. Gas Turbines Power, **113**, pp. 365–369.

[3] Sekar, R. R., Marr, W. W., Cole, R. L., and Marciniak, T. J., 1991, ''Effects of Oxygen Enrichment and Fuel Emulsification on Diesel Engine Performance and Emissions,'' ASME Meeting on Fuels, Controls and After-treatment for Low-Emissions Engines, ICE, **15**, pp. 21–28.

[4] Marr, W. W., Sekar, R. R., Cole, R. L., Marciniak, T. J., and Longman, D. E., 1993, ''Oxygen-Enriched Diesel Engine Experiments With a Low-Grade Fuel,'' SAE Paper 932805.

[5] Iida, N., and Sato, G. T., 1988, ''Temperature and Mixing Effects on NO_x and Particulates,'' SAE Paper 880424.

[6] Assanis, D. N., and Heywood, J. B., 1986, ''Development and Use of a Computer Simulation of the Turbocompound Diesel Engine System for Engine Performance and Component Heat Transfer Studies,'' SAE Paper 860329.

[7] Assanis, D. N., Sekar, R. R., Baker, D., Ciambekos, C. T., Cole, R. L., and Marciniak, T. J., 1990, ''Simulation Studies of Diesel Engine Performance With Oxygen Enriched Air and Water Emulsified Fuels,'' ASME Paper 90-ICE-17.

[8] Assanis, D. N., Karvounis, E., Sekar, R. R., and Marr, W. W., 1993, ''Heat Release Analysis of Oxygen-Enriched Diesel Combustion,'' ASME J. Eng. Gas Turbines Power, **115**, pp. 761–768.

[9] Lavoie, G. A., Heywood, J. B., and Keck, J. C., 1970, ''Experimental and Theoretical Investigation of Nitric Oxide Formation in Internal Combustion Engines,'' Combust. Sci. Technol., **1**, pp. 313–326.

[10] Gollan, A. Z., and Kleper, M. H., 1985, ''Research Into an Asymmetric Membrane Hollow Fiber Device for Oxygen-Enriched Air Production,'' DOE/ID-12429-1, U.S. Department of Energy.
[11] Ragland, K. W., and Whipple, J. G., 1989, ''Test and Evaluation of Polymeric Membranes for Oxygen-Enrichment of Air,'' DOE/ID-12710-1, U.S. Department of Energy.
[12] Winston Ho, W. S., and Sirkar, K. K., 1992, *Membrane Handbook*, Chapman & Hall, New York.
[13] Koros, W. J., and Chern, R. T., 1987, *Handbook of Separation Process Technology*, John Wiley and Sons, New York.
[14] Pan, C. Y., 1986, ''Gas Separation by High-Flux, Asymmetric Hollow-Fiber Membrane,'' AIChE J., **32**, No. 12, pp. 2020–2027.
[15] Chern, R. T., Koros, W. J., and Fedklw, P. S., 1985, ''Simulation of a Hollow-Fiber Gas Separator: The Effects of Process and Design Variables,'' Ind. Eng. Chem. Process Des. Dev., **24**, pp. 1015–1022.
[16] Nemser, S. M., and Roman, I. C., 1991, ''Perfluorodioxole Membranes,'' U.S. Patent 5,051,114.
[17] Markworth, V. O., Widener, S. K., Matheaus, A. C., and Mason, R. L., 1993, ''Locomotive Improvement Program — Twelfth Research Phase Final Report,'' Southwest Research Institute Report No. 034171/AAR Report No. R-841.
[18] Millington, B. W., and Hartles, E. R., 1968, ''Frictional Losses in Diesel Engines,'' SAE Paper 680590, SAE Trans., **77**, pp. 2390–2410.
[19] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.
[20] Plee, S. L., Ahmad, T., and Myers, J. P., 1982, ''Diesel NO_x Emissions—A Simple Correlation Technique for Intake Air Effects,'' Nineteenth Symposium (International) on Combustion, The Combustion Institute, pp. 1495–1502.
[21] Sawyer, R. F., Cernansky, N. P., and Oppenheim, A. K., 1973, ''Factors Controlling Pollutant Emissions from Gas Turbine Engines,'' Atmospheric Pollution by Aircraft Engines, AGARD CP-125, Paper No. 22.
[22] Ghojel, J., Hillard, J. C., and Levendis, J. A., 1983, ''Effect of Oxygen-Enrichment on the Performance and Emissions of IDI Diesel Engines,'' SAE Paper 830245.
[23] Ng, H. K., Novick, V. J., and Sekar, R. R., 1995, ''Using Monatomic Nitrogen Induced by a Pulsed Arc to Remove Nitrogen Oxides from a Gas Stream,'' ASME Fall Technical Conference, **1**, pp. 73–80.

An Experimental Investigation of the Effects of Common-Rail Injection System Parameters on Emissions and Performance in a High-Speed Direct-Injection Diesel Engine

P. J. Tennison

R. Reitz[1]

Engine Research Center,
University of Wisconsin-Madison,
Madison, Wisconsin 53706

An investigation of the effect of injection parameters on emissions and performance in an automotive diesel engine was conducted. A high-pressure common-rail injection system was used with a dual-guided valve covered orifice nozzle tip. The engine was a four-valve single cylinder high-speed direct-injection diesel engine with a displacement of approximately $\frac{1}{2}$ liter and simulated turbocharging. The engine experiments were conducted at full load and 1004 and 1757 rev/min, and the effects of injection pressure, multiple injections (single vs pilot with main), and pilot injection timing on emissions and performance were studied. Increasing the injection pressure from 600 to 800 bar reduced the smoke emissions by over 50 percent at retarded injection timings with no penalty in oxides of nitrogen (NO_x) or brake specific fuel consumption (BSFC). Pilot injection cases exhibited slightly higher smoke levels than single injection cases but had similar NO_x levels, while the single injection cases exhibited slightly better BSFC. The start-of-injection (SOI) of the pilot was varied while holding the main SOI constant and the effect on emissions was found to be small compared to changes resulting from varying the main injection timing. Interestingly, the point of autoignition of the pilot was found to occur at a nearly constant crank angle regardless of pilot injection timing (for early injection timings) indicating that the ignition delay of the pilot is a chemical delay and not a physical (mixing) one. As the pilot timing was advanced the mixture became overmixed, and an increase of over 50 percent in the unburned hydrocarbon emissions was observed at the most advanced pilot injection timing. [DOI: 10.1115/1.1340638]

Introduction

The automotive industry has long favored the spark-ignition (SI) engine with its relative simplicity, smooth operation, and good performance characteristics. The spark-ignition engine has offered the industry an engine with good overall performance over a wide range of engine speeds. With increasing fuel prices and ever increasing environmental [corporate average fuel economy and proposed CO_2 regulations] and emissions regulations, the automotive industry is looking towards other alternatives to the traditional SI engine. The compression ignition engine, or diesel engine, with compression ratios of 12–24, as compared to 8–12 in SI engines [1], offers a higher thermodynamic efficiency than the spark-ignition engine. Additionally, the diesel engine with its lack of pumping losses due to throttling the intake charge as in a SI engine, and the increased thermal efficiency, offers increased fuel economy when compared to a SI engine. The diesel engine also has a reputation for being a very reliable engine over a long service life, in part because of the sturdy construction needed to withstand the high compression ratios used.

Traditionally, relatively high combustion noise, oxides of nitrogen (NO_x) emissions, and particulate emissions have hindered diesel engine use in the automotive industry. These shortcomings are due to the heterogeneous nature of the spray combustion process occurring in a direct-injection compression-ignition engine. Common-rail injection systems for direct-injection diesel engines offer automotive diesel manufacturers an effective tool to help resolve these historical shortcomings of the diesel engine. Modern common-rail injection systems offer high injection pressures with flexible electronic control of fuel delivery, injection pressure, injection timing, and rate-of-injection (ROI) through the use of multiple injections. By effectively controlling the spray characteristics the resulting combustion process can be controlled.

Heavy-duty diesel engine manufacturers commonly use high injection pressures with smaller nozzle hole diameters to reduce particulate emissions. Reductions in particulate are attributed to better atomization of the fuel and increased mixing due to wall impingement, enhanced spray penetration, and additional air entrainment [2,3]. Additionally, research has shown that pilot and split injections can reduce both NO_x emissions and combustion noise significantly [2,4]. High-pressure split injections have been shown to simultaneously reduce both NO_x and particulate emissions [5,6].

Research has also been performed on the effect of exhaust gas recirculation (EGR) on exhaust emissions [7]. Additionally the combination of split injections with EGR has been shown to reduce both NO_x and particulate emissions simultaneously in heavy-duty engines [8]. Common-rail fuel injection systems should allow high-speed direct-injection (HSDI) diesel engine manufacturers to significantly reduce the shortcomings of the tra-

[1]Corresponding Author: Prof. Rolf D. Reitz, Engine Research Center, University of Wisconsin-Madison, Madison, WI 53706 E-mail: reitz@engr.wisc.edu

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division Nov. 16, 1999; final revision received by the ASME Headquarters, June 6, 1999. Associate Editor: D. Assanis.

ditional diesel engine mentioned earlier by effectively controlling the diesel spray characteristics and thus the combustion process.

Experimental Setup

The injection system used was an electrohydraulically controlled high-pressure common-rail injection system capable of up to two injections per combustion event. The injector had a dual-guided valve covered orifice nozzle tip. An electronic motor drove the high-pressure injection pump. The additional parasitic power of the electronic motor was accounted for during the injection pressure study. Instantaneous mass rate-of-injection profiles were obtained using a Bosch-type rate-of-injection meter. The details of the test bench used for the present research can be found in Bower [9]. A detailed study of the spray characteristics produced using this fuel injection system can be found in Tennison et al. [10]. Fuel system specifications are summarized in Table 1.

Table 1 Fuel system specifications

Injector type	Electro-hydraulically controlled injector
Injection pressure	Variable, max rail pressure = 1350 bar
Nozzle type	Dual guided VCO nozzle
Number of nozzle holes	6
Hole Diameter	160 μm
Hole L/D ratio	6.25
Included spray angle	145 °

The engine was a four-valve single cylinder HSDI diesel engine with a displacement of approximately $\frac{1}{2}$ liter. The test cell was designed to simulate turbocharging by controlling intake boost and exhaust back pressure. The engine was fully instrumented for temperature and pressure measurements, including in-cylinder pressure. The test cell had full emissions measurement capability including NO_x, CO, CO_2, total unburned hydrocarbons and smoke measurements. Gaseous engine emissions were measured using a Nicolet Fourier Transform Infrared (FTIR) emissions analyzer. Smoke measurements were performed using a Bosch RTT100 Smoke Opacimeter. Note that the RTT100 measures

Table 2 Engine specifications

Engine Type	Single cylinder diesel • Direct injection • Four valves • Ricardo Hydra block
Bore x Stroke	82.0 x 90.4 mm.
Compression Ratio	18.79:1
Displacement	477 cm^3.
Squish Clearance Volume	26.84 cm^3
Piston Type	Reentrant bowl
Intake Ports	1 high swirl, 1 tumble
Intake Valve Opening	10°BTDC
Intake Valve Closing	38°ABDC
Exhaust Valve Opening	38°BBDC
Exhaust Valve Closing	8.5°ATDC

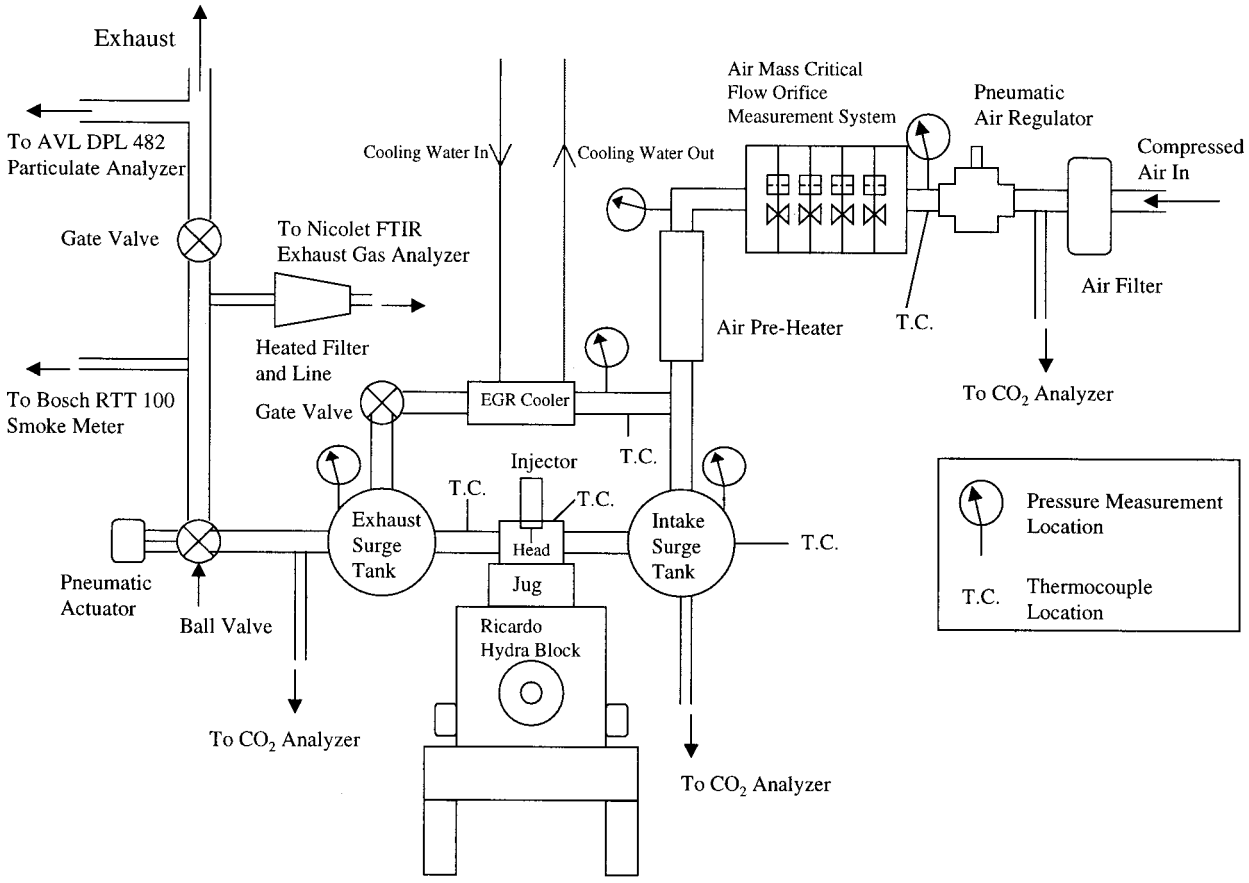


Fig. 1 Test engine cell layout

Table 3 Engine test parameters (baseline)

Rev/min	1004	1004	1757	1757	1757
Injection Pressure [bar]	600	800	1100	1100	1100
Injection timing					
Pilot duration [μs]	200	160	N/A	160	160
Dwell [μs]	4930	4930	N/A	3302	Varied
Main duration [μs]	1062	890	1186	1024	1024
SOI pilot °ATDC	-52	-52	N/A	-42	Varied
SOI main °ATDC	-4	-4	-5.5	-5.5	-5.5
Intake pressure [kPa]	117	117	196	196	196
Intake Temperature [°C]	37	37	50	50	50
Exhaust Pressure [kPa]	116	116	198	198	198
Equivalence ratio	0.91	0.91	0.76	0.76	0.76
Lambda	1.099	1.099	1.32	1.32	1.32

opacity and, through the use of internal correlation equations, the unit converts the opacity measurement to a particulate value with units of mg/m^3. For the present research, this data was converted to units of g/kW hr and is presented as Bosch C [g/kW hr] in the engine results. Engine specifications are given in Table 2 and Fig. 1 is a schematic of the test cell layout.

The fuel used was a commercially available #2 diesel fuel with an API gravity=32.7 at 60°F, viscosity of 2.43 cSt at 40°C, cetane rating of 44.3, C/H ratio of 6.8689, and a net heat of combustion of 18 266 BTU/lb.

Test Conditions

The present research focused on the effects of injection pressure and multiple injections. The test conditions were carried out at full load conditions. Injection pressure studies were performed at 1004 rev/min and the multiple injection study was performed at 1757 rev/min. Injection system timings are characterized by three energizing times (time during which the solenoid is excited) expressed in μs, related to the first injection, the dwell period, and the main injection. For example, a pilot injection case with a notation of 160/3302/1024 means that the first pulse was injected with a duration of 160 μs, a dwell time of 3302 μs, and a second pulse with a duration of 1024 μs.

The start-of-injection for both the pilot and the main injection are expressed in the number of crank angle (CA) degrees after top dead center (ATDC). There was a slight mechanical delay between the electronic signal sent by the electronic control unit (ECU) to the injector and the actual start of injection. This delay was determined to be typically $\frac{1}{4}$ ms. Data in this paper is presented in reference to the ECU timing signals sent to the injector. For each engine condition tested an injection timing sweep was performed. The timing was varied in increments of four crank angle degrees, typically up to twelve degrees both advanced and retarded from the baseline timing condition. Details of the engine test conditions (baseline) are shown in Table 3.

Results and Discussion

Injection Pressure Effects. The effect of injection pressure on emissions and performance was investigated for an engine speed of 1004 rev/min and full load. This case was chosen because it exhibited relatively high smoke emissions. The injection pressure was raised by 200 bar from the baseline condition while the overall fuel rate was kept constant. The solenoid activation time for the pilot was adjusted to try to keep the total mass of the pilot injection constant between the two injection pressures. The 600-bar case had a pilot consisting of 7.7 percent of the total mass injected (33.9 mg/inj), while the 800-bar case had a pilot consisting of 8.3 percent of the total injected mass (34.2 mg/inj). The difference in fuel rate for the two injection pressures was less than 1 percent.

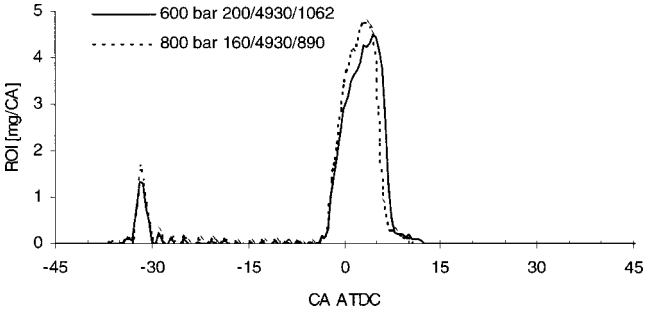


Fig. 2 ROI for 600 and 800-bar pressure, 1004 rev/min, SOI pilot=−35 ATDC, SOI Main=−4 ATDC

Figure 2 shows the Bosch rate-of-injection profiles for the two injection pressures at the baseline injection timing. A full timing sweep was performed at the same intervals as those performed for the baseline case. The dwell between the pilot and main injection was kept constant at 30 CA degrees for both injection pressures. The increased pumping power required for the higher injection pressure was accounted for in the results since the pump is driven externally by an electric motor. The increased pump power would be approximately 0.11 kW, and this was subtracted from the measured engine out power.

Figure 3 shows Bosch C vs NO_x for the two injection pressures. As can be seen, there was a significant reduction in smoke at the higher injection pressure, without a large increase in NO_x production. Indeed, the same NO_x production levels could be achieved with a slight retardation of injection timing for the higher injection pressure case. At the higher injection pressure, a main injection timing of 0° ATDC achieved the same NO_x level as the base injection timing of −4° ATDC without any penalty in brake specific fuel consumption (BSFC), as shown by the circled points in Fig. 3. However, at the same timing a smoke reduction of over 50 percent was achieved.

Figure 4 shows the effects of injection pressure on total hydrocarbon emissions (THC) vs start-of-main injection timing. The case of increased injection pressure appears to have a significant effect on total hydrocarbon emissions at retarded injection timings. This result is consistent with the findings of Oblander et al. [4]. The reduced THC emissions could be due to the enhanced atomization and better mixing created by the increased injection pressure.

Figures 5 and 6 show apparent-heat-release-rate (AHRR) and

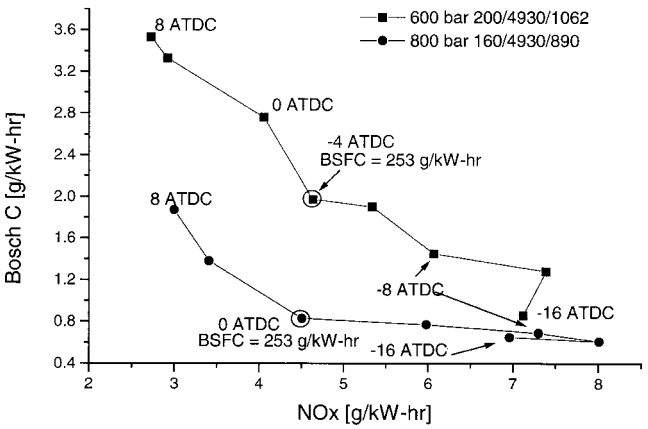


Fig. 3 Effects of injection pressure on Bosch C vs NO_x

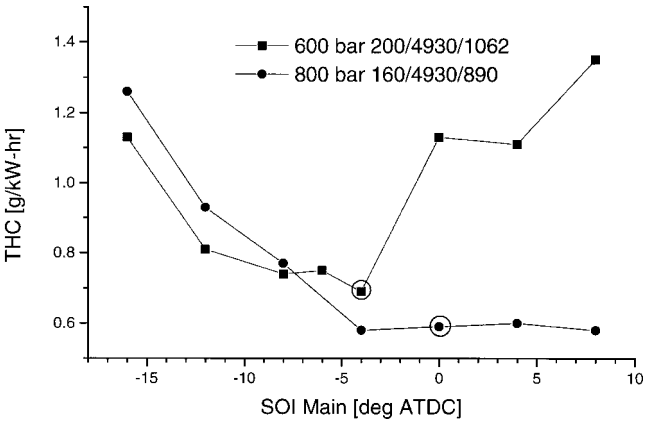


Fig. 4 Effects of injection pressure on THC vs SOI main injection

ROI at two different SOI timings for the injection pressures of 600 and 800 bar, respectively. As one can see, the rate of heat release for both injection pressures is quite similar. At the higher 800-bar injection pressure the rate of heat release for the main injection peaks earlier and drops off earlier than for the lower injection pressure for both injection timings shown. This is due to the greater injection velocities and shorter duration of the higher-pressure injection.

Multiple Injection Effects—Single vs Pilot-Main Injection. The effect of multiple injections on emissions and performance was investigated with the common-rail injection system. A single injection case and a split injection case with a pilot and main

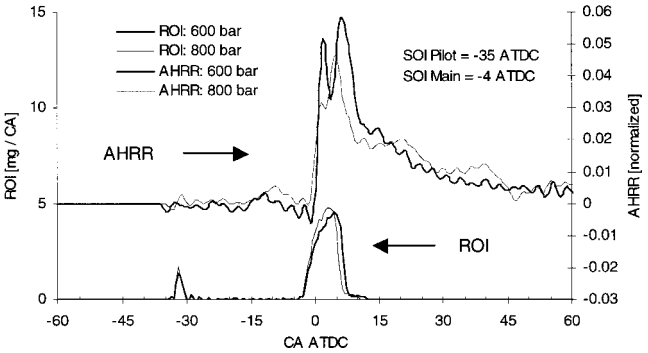


Fig. 5 AHRR and ROI for 600 and 800-bar injection pressures for base injection timing at 1004 rev/min

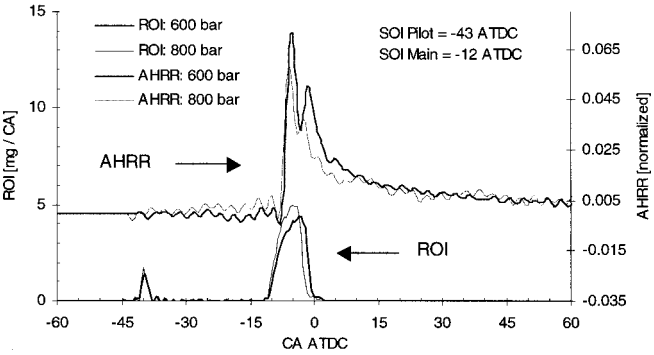


Fig. 6 AHRR and ROI for 600 and 800-bar injection pressures for an advanced injection timing at 1004 rev/min

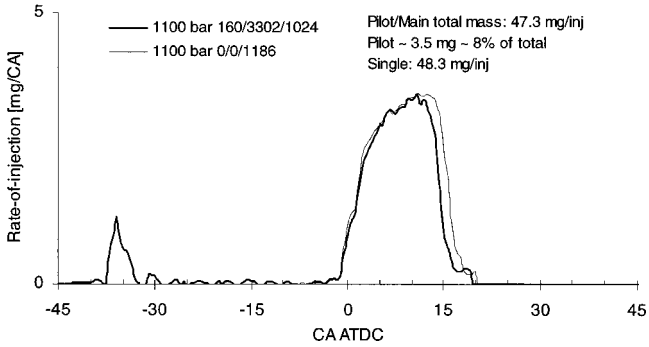


Fig. 7 ROI profiles for single and pilot injection cases at 1757 rev/min

injection were conducted at the engine speed of 1757 rev/min. The fuel rate for the two cases was set to be within 4 percent of each other and the effect of the start of pilot injection timing was investigated. The pilot timing was varied in 5 CA° increments while keeping the main injection at the baseline timing.

Figure 7 shows the ROI profiles for both the single and pilot injection cases at a main injection timing of −5.5 CA° ATDC. To match the fuel quantity of the pilot-main case the single injection had a slightly longer injection duration than the main in the pilot case. Figure 8 compares the effects of the single injection vs the split injection with a pilot on the smoke vs NO_x tradeoff curve. The timing sweep for the single injection was created using the same injection timing as the baseline sweep used for the main injection. The sweep was performed in four CA° increments. As can be seen from the figure the effect of a pilot injection is minimal when compared to a single injection. However, the pilot injection case does appear to produce a greater amount of smoke than the single, especially at some retarded injection timings. The NO_x levels are nearly identical at central injection timings.

Figure 9 shows BSFC vs main injection timing for the single and pilot injection cases. The single injection case has a lower BSFC over a wide range of injection timings. This is consistent with the findings of Oblander et al. [4]. Figure 10 shows cylinder pressure, AHRR, and ROI for the baseline main injection timing of −5.5° ATDC for both the single and pilot injection cases. The pilot injection burn is seen to cause a noticeable increase in cylinder pressure by the end of the compression stroke. With the pilot injection there does not seem to be a significant decrease in the ignition delay of the main injection, as other researchers have also found [2,4]. Under these conditions the rate of heat release for the main injected quantity of fuel does not appear to be significantly

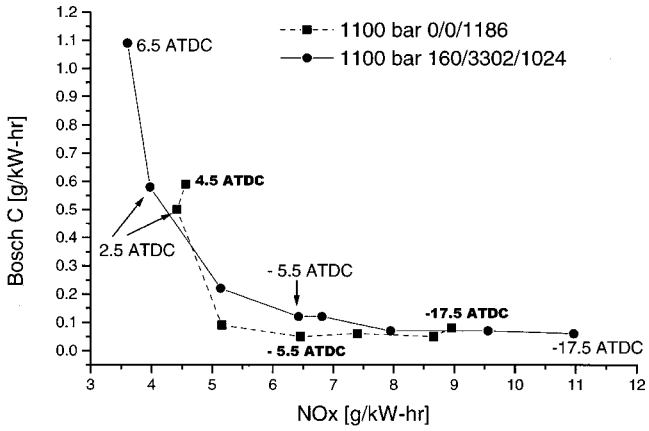


Fig. 8 Effects of single injection and pilot on smoke vs NO_x tradeoff curve

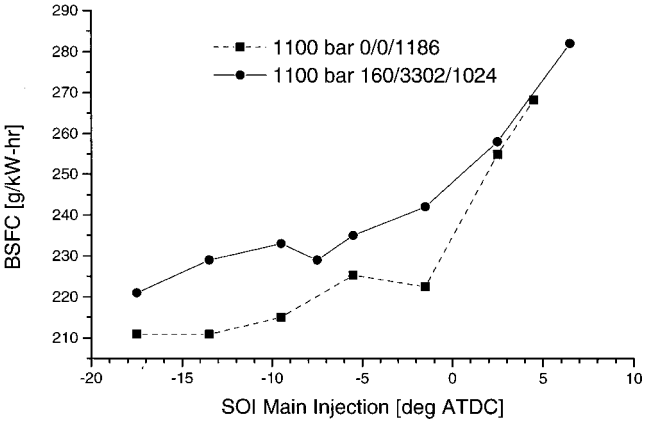


Fig. 9 Effects of single injection vs pilot on BSFC vs injection timing

different between the two cases. The peak for the single injection case is shifted slightly later than the pilot injection case. This is not consistent with results of Shimada et al. [2] who found that the peak rate of heat release was reduced by nearly a half with pilot injection. This difference is likely due to the fact that Shimada et al. used larger pilot injection quantities, whereas in this present research the pilot was only 8 percent of the total injected mass. The pilot injection does not seem to have much of an effect on the apparent heat release rate at these speed, load, and injection parameters.

Figure 11 shows THC vs main injection timing for the pilot and

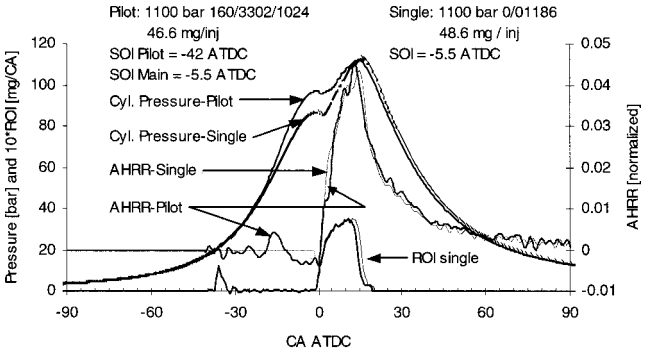


Fig. 10 Cylinder pressure, AHRR, ROI for single and pilot injection cases at 1757 rev/min

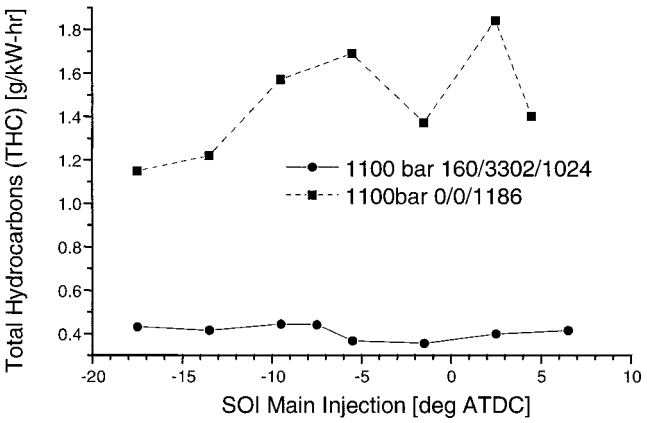


Fig. 11 THC vs SOI for single and pilot injection cases at 1757 rev/min

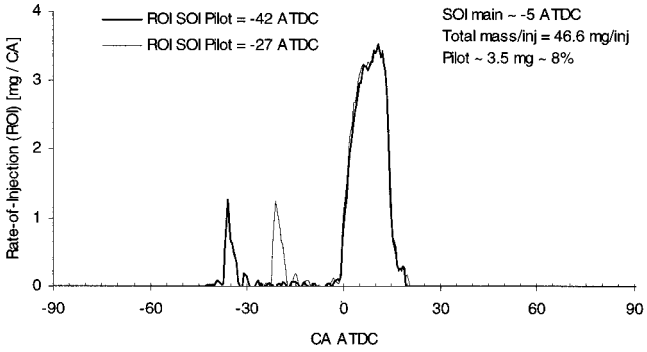


Fig. 12 ROI vs CA for two pilot injection timings 1757 rev/min

single injection cases. As can be seen from the figure there is a significant reduction, over an 80 percent reduction, in total hydrocarbon emissions when a pilot injection is used. This is consistent with the results of Oblander et al. [4] who showed similar benefits of using a pilot injection on unburned hydrocarbons.

Pilot Injection Timing. The effect of the pilot injection timing was studied for the condition at 1757 rev/min. In this case the pilot timing was varied in 5 CA° increments while holding the main injection timing constant. Figure 12 shows example ROI profiles at two different pilot timings.

In the following figures, the notation 1757AD (D indicates that the dwell was changed) refers to the pilot timing study and 1757A refers to the baseline case. Figure 13 shows the smoke vs NO_x tradeoff curve for the pilot timing study and Fig. 14 shows an expanded view of the tradeoff curve for both the pilot study and the baseline 1757 rev/min timing sweep cases. Note that, for the baseline case, the pilot and main injection timings were varied with a constant dwell of 35 CA° between them. For the pilot timing study only the SOI of the pilot was varied. Figure 13 shows a complicated soot-NO_x tradeoff exists as the pilot dwell is changed. In particular, at long dwells high NO_x and low soot is seen. As the dwell is decreased, soot increases and NO_x at first decreases and then hooks back at relatively short pilot-main dwells.

Figure 14 shows however that the pilot timing does not have as a significant effect on the smoke NO_x tradeoff as does the effect of main injection timing. The data of Fig. 13 are seen to be clustered around the −5.5° ATDC data point in Fig. 14. In Fig. 13 it is seen that the best improvement in NO_x occurs at a pilot timing of about −32° ATDC. The NO_x is reduced by 15 percent at this pilot timing when compared to the baseline timing (i.e., 35°

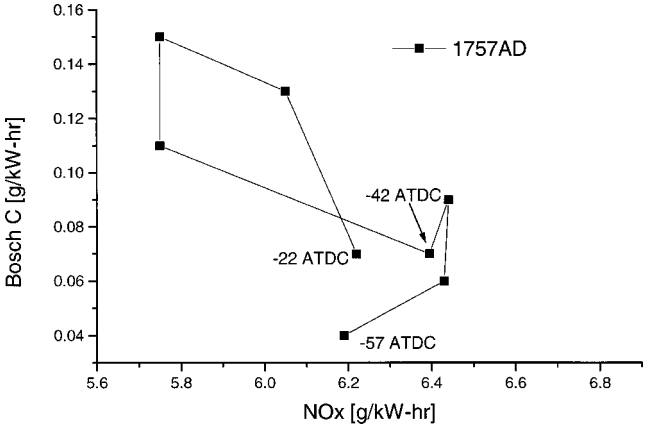


Fig. 13 Bosch C vs NO_x for pilot timing study. SOI main injection is constant at −5.5° ATDC

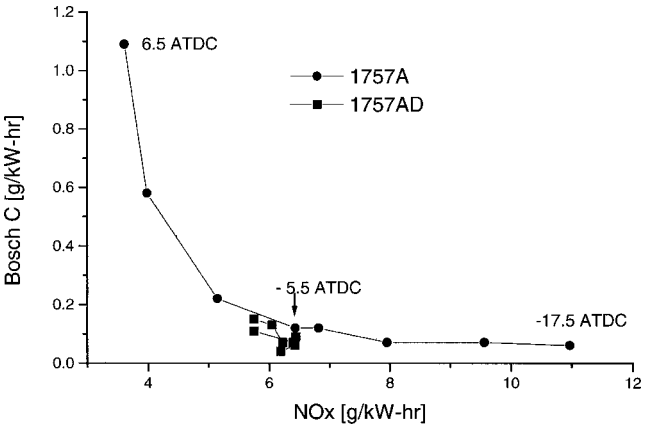


Fig. 14 Bosch C vs NO_x for pilot timing study and baseline timing sweep

dwell). However, this reduction in NO_x is accompanied by an increase of smoke of 85 percent when compared to the baseline pilot-main timing. The best timing point appears at the most advanced timing of −57° ATDC. This timing exhibits a reduction of NO_x by 8.6 percent and a simultaneous reduction of smoke by 43 percent when compared to the base injection timing. This reduction in emissions is accompanied by a reduction in BSFC by 2.5 percent. The overall effect of the pilot injection timing on BSFC was small in comparison to the effect of main injection timing. The best BSFC points appeared at the more advanced pilot injection timings. Overall, it is clear that the smoke-NO_x tradeoff is complex in shape as the dwell timing is changed and a meaningful conclusion on the effect of pilot injection timing on emissions cannot be made without further study.

Figure 15 shows cylinder pressure, AHRR and ROI (in order from top to bottom) for the base pilot timing (AD) and a retarded pilot injection timing of 10 CA° from the base (AD-10). The retarded pilot injection timing produces a significantly larger apparent rate of heat release for the pilot injection when compared to the more advanced timing. This larger heat release rate does not apparently have a significant effect on the heat release due to the main injection.

Figure 16 shows the AHRR for several different pilot injection timings. As the pilot timing is advanced, the heat release rate of the pilot burn is diminished, and as the pilot timing is retarded, the AHRR of the pilot is larger. Figures 15 and 16 present an interesting fact in that the pilot reaches the autoignition point at approximately the same crank angle regardless of pilot SOI timing. Autoignition of the pilot will not occur until the temperature and pressure in the cylinder have reached appropriate levels, and until sufficient mixing time has elapsed. The amount of fuel injected in the pilot is very small, therefore, vaporization of the fuel and

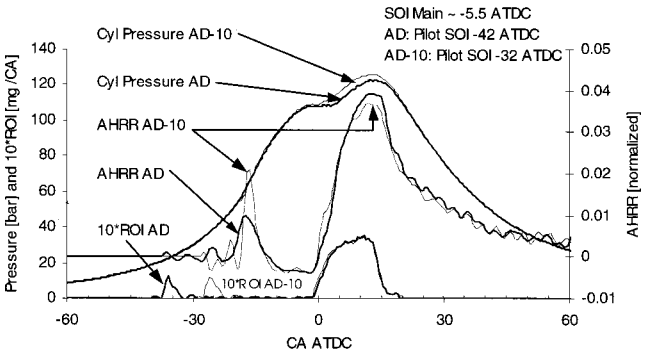


Fig. 15 Pressure, AHRR, and ROI for two cases from the pilot study

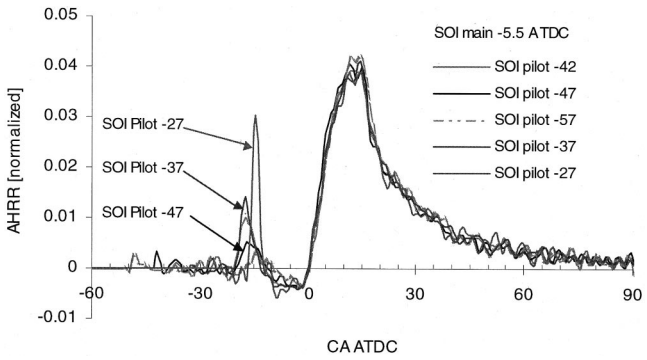


Fig. 16 AHRR for various cases from the pilot study

complete fuel-air mixing should occur within a short period of time. Therefore, it appears that the associated ignition delay with the pilot injection, is dominated by a chemical delay and not a physical one (mixing).

Figure 17 graphically demonstrates the fact that the autoignition point of the pilot is relatively independent of pilot injection timing. This figure depicts cases at 1757 rev/min at various injection timings with the pilot-main dwell held constant at the base value. Notice that the start of heat release for the main injection changes with injection timing, but the start of heat release for the pilot does not. The peak AHRR value of the pilot burn for the most advanced injection timing case is much lower than the other cases. It appears that if the injection timing were to be advanced further that the pilot burn would disappear. In fact, at 1004 rev/min one can see from Fig. 18 that the pilot burn does effectively disappear.

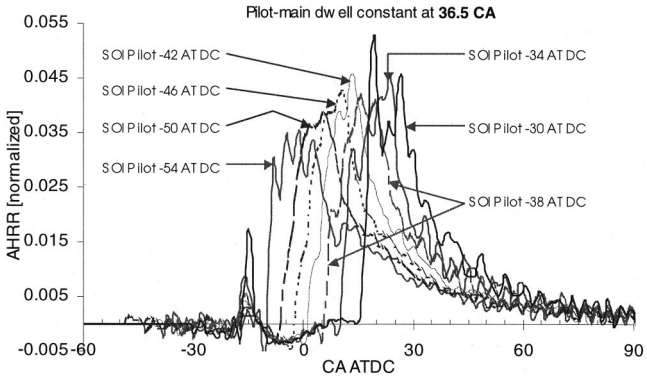


Fig. 17 AHRR at 1757 rev/min, 1100 bar injection pressure, pilot-main dwell constant, pilot SOI varied

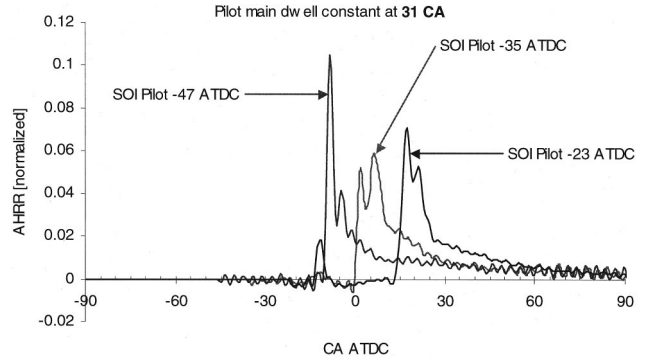


Fig. 18 AHRR at 1004 rev/min, 600-bar injection pressure, pilot-main dwell constant, pilot SOI varied

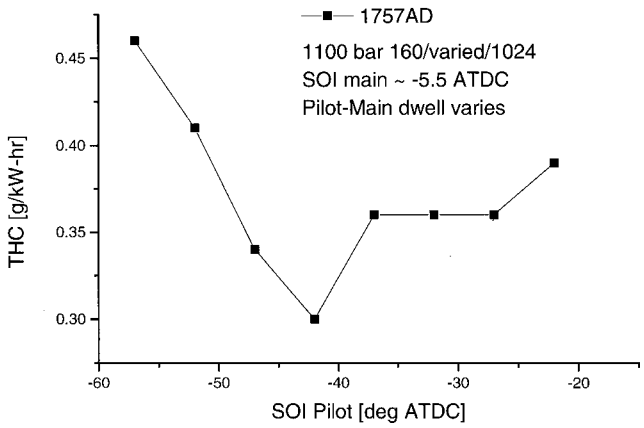


Fig. 19 THC vs pilot injection timing at 1757 rev/min

As discussed previously, the pilot ignition delay appears to be a chemical delay (temperature and pressure related) and not a physical mixing delay. Increased temperature and pressure of the gases in the cylinder prior to autoignition are a result of the compression process and should have roughly constant values at a given crank angle from cycle-to-cycle. The autoignition point of the pilot is independent of injection timing because the gases in the cylinder do not reach a high enough temperature and pressure for autoignition until a certain crank angle. For the 1757 rev/min case, from the pilot study shown in Fig. 16, the crank angle is about 15.5° before THC. No matter how early the pilot is injected, autoignition will not occur until that crank angle is reached. This behavior occurs for all test cases with pilot injection.

At highly advanced injection timings the peak apparent-heat-release-rate of the pilot injection diminishes. The fuel-air mixture could in fact be overmixed and too lean to burn for the very advanced pilot injection timings. The in-cylinder fuel-air mixture is most likely homogeneous in nature by the time the appropriate crank angle is reached. Figure 19 shows THC vs pilot injection timings. At very early pilot injection timings the amount of unburned hydrocarbon emissions increases rather dramatically. This supports the hypothesis that the fuel-air mixture may in fact be ''overmixed'' and too lean to burn in some regions. A diesel engine is typically run with fairly large amounts of excess air. The spray combustion process does not necessarily utilize all the air (O_2) in the combustion chamber. The well-mixed pilot fuel-air mixture may not be completely utilized during the combustion of the main injection.

Conclusions

The effects of injection pressure were investigated for a full load 1004 rev/min case. Injection pressure was increased from 600 to 800 bar and enhanced atomization and improved air utilization for the higher-pressure injection case resulted in a smoke reduction of as much as 50 percent without any appreciable penalty in BSFC. This was achieved at a slightly retarded injection timing to achieve the same NO_x level as with the baseline lower injection pressure case. The resulting decrease in soot was likely due to enhanced mixing from the higher injection pressure.

The effects of multiple injections were also studied. First, the effects of a pilot injection were studied with respect to using a single injection case. The pilot injection case was found to produce higher levels of smoke than the single injection case through most of the SOI timings considered in the tests. The two types of injections also had very similar NO_x emissions. However there was a significant reduction (over 80 percent) in total hydrocarbon emissions when a pilot injection was used. The best injection strategy appeared to be a pilot injection that is retarded slightly from the baseline timing point. This retarded timing condition had slightly lower THC emissions than the baseline timing case.

The pilot timing was also varied while holding the main injection timing constant to determine the effects of pilot injection timing and dwell. Overall, the effect of the pilot timing was found to be small compared to the effect of the main injection timing. The best pilot timing was found to be −57° ATDC. At this point a simultaneous reduction of NO_x of 8.6 percent, and a reduction in smoke of 43 percent, was found. This was accompanied by a reduction in BSFC of 2.5 percent. Pilot injection timing dwell was seen to have a minor effect on performance and emissions but, the effects did not follow a simple trend.

The point of ignition for the pilot injection was at the same CA regardless of pilot injection timing. This phenomenon can be seen in the pilot dwell timing study, where the SOI of the main injection was held constant, and at all baseline operating speeds. The ignition delay of the small pilot quantity of fuel was thus determined to be dependent on a chemical delay and not a physical mixing delay. There is an associated increase of THC at the very advanced injection timings. However there is no significant AHRR associated with the pilot quantity of fuel at these advanced pilot injection timings. The pilot injection appears to be overmixed and too lean to burn.

Acknowledgments

The authors would like to acknowledge in particular R. Thout, T. Lee, and D. Corgard for their assistance in the data collection and analysis for this research. Additionally, the faculty and staff of the ERC are thanked for their input and assistance during this research. Funding for this research was provided by Centro Ricerche Fiat and Sandia National Laboratories.

Nomenclature

ABDC = After Bottom Dead Center
AHRR = Apparent heat-release-rate
ATDC = After Top Dead Center
BBDC = Before Bottom Dead Center
BSFC = Brake Specific Fuel Consumption
BTDC = Before Top Dead Center
CA = Crank angle
CAFE = Corporate average fuel economy
CO = Carbon Monoxide
CO_2 = Carbon Dioxide
ECU = Electronic Control Unit
EGR = Exhaust Gas Recirculation
FTIR = Fourier Transform Infrared
HSDI = High Speed Direct Injection
NO_x = Oxides of Nitrogen
ROI = Rate-of-Injection
SI = Spark Ignition
SOI = Start-of-Injection
THC = Total Hydrocarbons
VCO = Valve Covered Orifice

References

[1] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.
[2] Shimada, T., Shoji, T., and Takeda, Y., 1989, ''The Effect of Fuel Injection Pressure on Diesel Engine Performance,'' SAE Paper 891919.
[3] Shundoh, S., Kakegawa, T., and Tsujimura, K., 1991, ''The Effect of Injection Parameters and Swirl on Diesel Combustion with High Pressure Fuel Injection,'' SAE Paper 910489.
[4] Oblander, K., Kollmann, K., Kramer, M., and Kutschera, L., 1989, ''The Influence of High Pressure Fuel Injection on Performance and Exhaust Emissions of a High Speed Direct Injection Diesel Engine,'' SAE Paper 890438.
[5] Tow, T. C., Pierpont, D. A., and Reitz, R. D., 1994, ''Reducing Particulate and NO_x Emissions by Using Multiple Injections in a Heavy Duty D.I. Diesel Engine,'' SAE Paper 940897.

[6] Pierpont, D. A., and Reitz, R. D., 1995, ''Effect of Injection Pressure and Nozzle Geometry on D.I. Diesel Emissions and Performance,'' SAE Paper 950604.
[7] Ladommatos, N. B., Horrocks, R. R., and Cooper, L., 1996, ''The Effect of Exhaust Gas Recirculation on NO_x Emissions in a High-Speed Direct-Injection Diesel Engine,'' SAE Paper 960840.
[8] Pierpont, D. A., Montgomery, D. T., and Reitz, R. D., 1995, ''Reducing Particulate and NO_x Using Multiple Injections and EGR in a D.I. Diesel,'' SAE Paper 950217.
[9] Bower, G. R., and Foster, D. E., 1991, ''A Comparison of the Bosch and Zuech Rate of Injection Meters,'' SAE Paper 910724.
[10] Tennison P. J., Georjon T. L., Farrell P. V., and Reitz R. D., 1998, ''An Experimental and Numerical Study of Sprays from a Common Rail Injection System for Use in an HSDI Diesel Engine,'' SAE 980810.

The Effect of Mean Turbulent Strain Rate on the Flame Speed of Premixed, Growing Flames

D. S.-K. Ting
Mechanical, Automotive & Material Engineering,
University of Windsor,
Windsor, Ontario, Canada N9B 3P4

M. D. Checkel
Mechanical Engineering Department,
University of Alberta,
Edmonton, Alberta, Canada T6G 2G8

This paper presents a flame growth model based on experimental measurements of flame speed and mean turbulent strain rate. Methane/air mixtures of 0.7 and 0.9 equivalence ratios were centrally spark-ignited in a 125 mm cubical chamber. Based on schlieren images and combustion pressure traces, a linear correlation was found between the turbulent flame speed and the turbulent strain rate. For these unity-Lewis-number and near-zero-Markstein-number flames, the effectiveness of turbulent strain in enhancing the flame speed was found to increase linearly with the mean flame radius over the range of conditions tested. [DOI: 10.1115/1.1339990]

Introduction

Accurate turbulent flame speed models are required for improved design of combustion equipment and for hazard evaluation of accidental explosions. Ideally, such models would be based on fundamental physics and be fully inclusive of all phenomena. However, given the number of physical and chemical phenomena involved and the complexity of their interaction, such models are not likely to become available in the near future and will be so complex as to limit their utility as design tools. This paper presents an elementary turbulent flame growth model which is based on experimental results, and it includes mixture chemistry in the form of laminar flame speed, mixture motion effects in the form of turbulent shear rate, and an additional scale effect depending explicitly on the mean flame size.

Flame speed as used in this paper is defined as

$$S=(1/\rho A)dm/dt, \tag{1}$$

where ρ is the density of the unburnt gas, A is the surface area of the sphere comprising the burnt volume, and dm/dt is the mass burning rate. The laminar and turbulent flame speeds, S_l and S_t, correspond to flame propagation in a quiescent and turbulent mixture, respectively.

The limitations of current premixed turbulent flame models can be inferred from the study by Trouve and Poinsot [1]. Their numerical simulations in a decaying turbulent flow showed that even for a ''planar'' turbulent flame front, the combustion phenomena of interest would vary significantly with time in a manner depending on Lewis number, Le (=thermal diffusivity/mass diffusivity). For example, the reaction rate, turbulent flame speed, and flame surface area were predicted to increase during an initial growth period of about two eddy turnover times before they become constants for unity Le mixtures. Since these trends were predicted for planar flame fronts in a decaying turbulent flow, the model would likely lead to increasing combustion rate for unity Le mixtures as well in situations where the turbulence remained constant or decayed more slowly. This is rather perplexing as a planar flame front of unity Lewis number and zero Markstein number in a constant turbulence flow field is expected to propagate quasi-steady.

To develop generally applicable models, turbulent flame speed is commonly correlated with the turbulence intensity, the flame surface fractal dimension, the turbulence structures, or the turbulent strain rate. The flame speed–turbulence intensity correlation is probably the simplest and most frequently used. One limitation of flame speed–turbulence intensity correlations is that turbulent length scale effects are not taken into account explicitly, knowing that the relative size of the turbulent eddies with respect to the flame front thickness and the flame size can be very important [2,3].

Flame front fractal dimension is difficult to estimate or measure consistently. A slight change in the inner or outer cutoff scale can lead to serious errors in measuring fractal dimensions [4]. Also, the fractal aspect of a flame is actually the outcome of the flame–fluid motion interaction and the underlying turbulence parameters which result in fractal behavior of the flame front are more fundamental, and usually easier to measure, than the fractal behavior.

An ability to directly model the turbulent mixture motion structures may ultimately solve the mystery of mixture turbulence action on the flame front. With the advancement in numerical computation power, the vortical structures in a turbulent flow field with low Reynolds (Re) can now be simulated from Navier–Stokes equations [5,6]. Unfortunately, these simulations require a large amount of computation power even at very low Re. Great advances in computation technology and numerical models are needed before large Re turbulent flows can be simulated accurately. The results from these simulations can be used to better estimate the effects of turbulence on a passive flame front. Further work will then be required to thoroughly model the interaction of the turbulence and a reactive flame front. This approach may ultimately generate detailed knowledge of turbulence–flame interaction.

Mean turbulent shear rate can be used to characterize turbulent mixture motion in a manner which incorporates both intensity and scale effects. Batchelor [7] deduced that the interface area of a nonreacting material surface in a homogeneous isotropic turbulent field grows in proportion to the rate of strain. Building on this idea, Thomas [8] indicated the importance of the rate of strain just ahead of a flame front in the flame surface development and mixture consumption rate. Studies such as [9–12] illustrate the promise in correlating turbulent flame speed with the mean turbulent strain rate. This paper takes a close look at the turbulent flame speed–mean strain rate correlation using experimental data collected in a constant-volume, cubical chamber.

The Rate of Strain

Even in a quiescent mixture, the flame usually exists with a certain amount of laminar strain. The laminar strain rate is also called the chemical strain rate [13]. For a laminar flame, the strain

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, June 6, 2000; final revision received by the ASME Headquarters, November 15, 2000. Associate Editor: D. Assanis.

rate is simply the flame or velocity gradient, S_l/δ_l [3], where S_l is the laminar flame speed and δ_l is the laminar flame front thickness.

In isotropic and isothermal turbulence, the mean turbulent strain rate can be expressed as u'/λ, where u' is the rms turbulent fluctuation intensity and λ is the Taylor microscale [14,15]. For isotropic turbulence, the turbulence dissipation is given by, $\epsilon = Cu'^3/\Lambda$, where Λ is the integral scale and C is a constant of the order of unity [16]. Based on the turbulence properties of flow on the centerline of a pipe, Abdel-Gayed et al. [9] recommended $C = 0.37$ which leads to

$$\lambda^2/\Lambda = C_1 v/u' \tag{2}$$

where v is the kinematic viscosity and C_1 is about 40.4 for turbulent Reynolds number based on integral length scale $\mathrm{Re}_\Lambda > 60$ [9]. Kido et al. [17] obtained a different expression based on turbulence properties measured in their constant volume combustion chamber. Specifically, the correlation is

$$\lambda/\Lambda = \sqrt{15}/C_2^{0.4}\,\mathrm{Re}\wedge^{0.2}, \tag{3}$$

where C_2 is a constant about 11. In a later section, the use of Eq. (3) will be shown to lead to less experimental scatter than Eq. (2). This is presumably because it gives a better representation of turbulence relations in a combustion chamber than the traditionally used pipe flow relationship. The point is that as the turbulent strain rate is not measured directly, its accuracy depends on the validity of the turbulence correlations used.

Turbulent Flame Development

If turbulence is modeled as a set of eddies with a range of scales, the small flame kernel right after spark ignition will be primarily wrinkled and sheared by eddies smaller than itself [9,18,19]. Larger eddies mostly convect the flame around. As the flame grows, progressively larger eddies become effective in wrinkling the flame front. Therefore, flame growth beyond a certain size is required to make the turbulence fully effective in affecting the reacting flame surface. Knowing that most of the turbulence energy is contained in the larger scale motions represented by the integral length scale, we would expect the flame ball to grow, as a minimum, somewhat larger than the integral scale for the turbulence to be fully effective. While the integral scale is defined as the mean energy-containing scale of turbulence, the much smaller Taylor microscale is defined in terms of mean rate of strain u'/λ [10,11]. It is, therefore, expected that turbulent straining could become significant well before the length scale of the reacting flame surface approaches the integral scale, and that shear effects could approach a maximum level while the flame is still relatively small.

It is a common assumption that once the flame grows larger than the integral scale, the flow field turbulence is at its maximum ability in affecting the turbulent flame speed. The flame would then be ''fully developed'' and the flame speed/turbulence intensity ratio would remain constant with further flame growth [10,20]. In other words, the fully developed turbulent flame is assumed to propagate with a quasisteady turbulent flame speed in a steady turbulence flow field. Studies such as [21–25], however, give the contradicting evidence, showing that turbulent flames continue to accelerate with increasing flame size even when the flame is much larger than the integral scale. There are a few possible reasons behind this continuous flame speed acceleration. First, it is shown by Ashurst [26] that the volume expansion of a wrinkled flame ball can lead to progressively more wrinkled flame, and hence, to a progressive stronger flame speed–turbulence correlation. Also, the instantaneous flame speed–turbulence correlation can be influenced by both current turbulence levels and by a ''memory'' of previous flame–turbulence interactions embodied in the developed flame surface [27–29].

Nevertheless, it is also reasonable to assume that there will be some limit to flame size effects. As the flame grows very large, the flame front farthest from the center is only infinitesimally affected by the expansion of the flame front close to the center [26]. Likewise, the development of flame surface complexity will probably reach some limit where it is balanced by flame surface destruction. Therefore, the developing turbulent flame may become fully developed eventually.

Lean methane–air mixtures were chosen so as to minimize effects aside from turbulence on the flame development. Stoichiometric and 72 percent stoichiometric methane–air mixtures have 1.01 and 0.97 Le, respectively [30,3]. The 0.7 and 0.9 equivalence ratio methane–air mixtures used in the experiments correspond to 1.8 and −0.2 Ma, respectively [31].

Experiment and Analysis

Turbulent flame growth experiments were conducted in a 125 mm cubical combustion chamber as sketched in Fig. 1. Turbulence was generated by pulling a 60 percent solid perforated plate across the chamber prior to ignition. Turbulence intensity and scale were set by choosing appropriate values for the plate velocity, plate hole diameter, and turbulence decay period. The integral scale generated was a strong function of the plate hole diameter, and it increased, but only very slowly, as the turbulence decayed. For the same plate, the intensity was increased by pulling the plate across the chamber at a faster speed and/or by reducing the time delay after the plate passage and ignition. The preignition turbulence was up to 2 m/s, while the integral scale varied from 2 to 8 mm, covering typical turbulence conditions in engines. Premixed methane–air mixtures of 0.7 and 0.9 equivalence ratios were ignited at initial conditions of 300 K and 101 kPa. These temperature and pressure are lower than that encountered in practical engines; future work aims at raising these values to that of an operating engine. A capacitive discharge/coil system with 312.5 mJ of stored energy supply was used to force a spark across a 5 mm spark gap at the center of the chamber. This amount of stored energy was chosen to ensure proper ignition of fuel-lean mixtures under high turbulence straining. The effects of the spark energy and of the spark gap on the burning rate had not been studied. A more detailed description of the experiment and analysis can be found in Ref. [32].

The flame growth rate and mass burning rate were measured in two ways. Schlieren flame growth images were recorded at 2000 frames/s using a high speed video camera. Concurrently, the combustion chamber pressure rise was traced and recorded using a pressure transducer and computer data acquisition system.

A multizone thermodynamic equilibrium model was used to deduce the flame speeds from pressure trace and the schlieren flame growth images. During the early flame growth period before the flame contacted the chamber walls, the 125 mm cubical com-

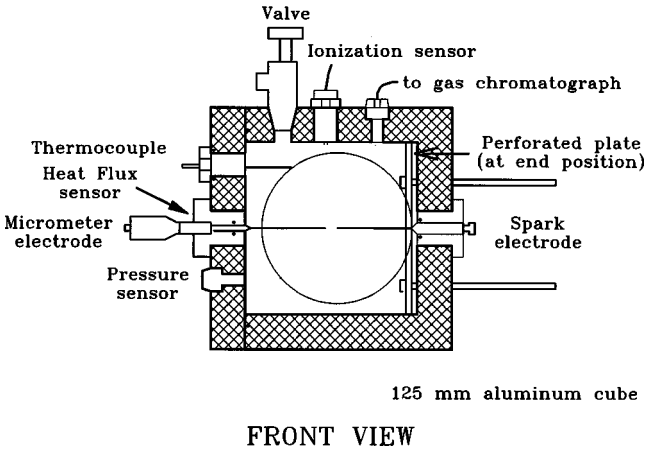


Fig. 1 The 125 mm cubical combustion chamber

bustion chamber was treated as a sphere of equivalent volume, with an equivalent combustion chamber radius of 76.6 mm. The multizone model divides the initial unburnt mixture into 1500 spherical shell elements of equal cross-sectional area. The flame was assumed to propagate isotopically with a flame front of negligible thickness. As each element burns, the thermodynamic equilibrium composition and volume of that element, (and all previously burnt elements), are recalculated, using a first-law energy balance to establish the correct total energy, and hence, the correct chamber pressure. Using this model, the mass and volume of burnt mixture can be established from the experimental pressure trace. To calculate the flame speed, the flame surface area was assumed to be that of a smooth sphere embracing 100 percent burnt mixture.

In analyzing experimental results, the unburnt mixture turbulence is *not* assumed to remain constant after ignition. Combustion chamber turbulence intensity and scale change significantly during the combustion period due to viscous decay and compression effects. The viscous decay is particularly affected by the turbulence scale so any experimental analysis attempting to determine scale effects must carefully consider decay. In this case, the decaying turbulence measured in the ''cold'' runs without combustion, is corrected for the increases in chamber pressure and the unburnt mixture temperature. The details of these adjustments and discussion of the validity of the turbulence model can be found in Ref. [32].

An explicit error analysis was not possible as the mean turbulent strain rate could not be measured directly, in addition to the adjustments in the flame front turbulence. Therefore, the error in the results presented in the following section depends heavily on the accuracy in estimating the mean turbulent strain rate; the values for the flame speed are estimated to be within ± 20 percent otherwise. Future studies will undeniably lead to refinement in the values of the constants of the flame–turbulence correlations given below. The qualitative conclusions deduced from this study, however, are expected to remain valid.

Results and Discussion

Even in the idealized, constant-volume, combustion chamber used in this study, turbulent flame growth is a highly transient phenomenon. Flame size, flame propagation speed, turbulence parameters, along with pressure and temperature, all vary significantly throughout the combustion period, so the only sensible correlations are between instantaneous burning rates and current turbulence conditions. Unless otherwise stated, Eq. (3) has been used in calculating the mean strain rate from directly measured and modeled values. The kinematic viscosity of air is calculated from Ref. [33] assuming ideal gas. The mixture viscosity is assumed to be equal to that of air, as methane is a minor component of the mixture and the kinematic viscosity of methane is rather close to that of air.

A typical plot of turbulent flame speed as a function of mean strain rate is shown in Fig. 2. The equivalence ratio, the coefficient of determination, and the mean flame radius are designated as ϕ, R^2, and r, respectively. Each data point represents the mean of three to five experiments conducted at almost identical conditions with flames propagating into turbulent mixtures of 0.9 equivalence ratio methane–air with 8 mm integral scale. All parameters were calculated for the point where the flames were at 55 mm radius. Over this range of strain rates, the data points fall along a straight line with a positive slope and an extrapolated y intercept corresponding closely to the laminar flame speed. Note that the corresponding laminar flame speed had been affected by changes in parameters such as pressure, temperature, ignition-induced turbulence, hydrodynamic instability, etc. All these effects were collected into the instantaneous laminar flame speed, and the turbulent flame speed only gave the additional effects due to turbulence parameters supplied via the plate movement. Consequently, any curve fitting used is expected to pass through the

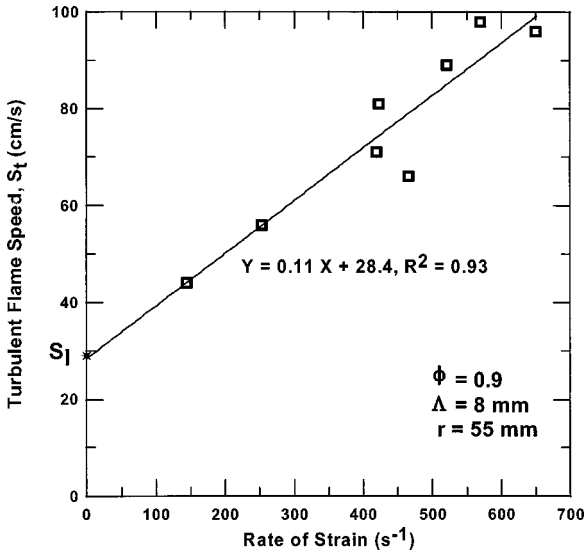


Fig. 2 Turbulent flame speed as a function of mean strain rate (ϕ=0.9, Λ=8 mm, r=55 mm)

corresponding laminar flame speed. For the wide range of shear rates portrayed in Fig. 2, there is a strong linear correlation between turbulent flame speed and mean turbulent shear rate. Specifically,

$$S_t = C_3 u'/\lambda + S_l, \tag{4}$$

where C_3 is a constant about 0.11 cm for this particular case. We chose to use linear fit as the scatter in the experimental data precluded the use of a higher order fit.

Figures 3*a* and 3*b* show more extensive plots of turbulent flame speed against rate of strain for 2 and 8 mm integral scale cases, respectively. The results for the $\Lambda = 4$ mm case, which fall between those of the $\Lambda = 2$ and 8 mm cases, are not shown here due to space limit. The values for 55 mm radius flames are from pressure trace analysis, while those for 19 and 27 mm radius flames are estimated from the schlieren images. It should be pointed out that measurement of turbulent flame speed based on schlieren images gives a slightly exaggerated value compared with the pressure trace analysis. This is because the schlieren image measures the leading edge of the three-dimensional projection of the flame front rather than the average position [32]. For lucidity of depiction, linear fits are used and the data points for 27 mm radius flames have not been shown.

The quality of the linear fits in Fig. 3 is illustrated by the coefficient of determination as well as the y intercept. A line with R^2 near unity and a y intercept near the laminar flame speed would be considered an ideal fit. The experimental scatter in values is relatively large in most cases aside from the $\Lambda = 8$ mm, $\phi = 0.9$ case. For specific Λ, ϕ, and r, all three figures show that there is a roughly linear relation between turbulent flame speed and mean strain rate. The values of both flame speed and turbulent strain rate are higher for the faster burning, $\phi = 0.9$ mixtures. This is because the turbulence in the faster burning mixture has less time to decay and hence, a higher level of straining is still present when the flame reaches the flame size at which measurements are taken. Similarly, turbulence with larger integral scale decays more slowly and hence, a higher rate of strain is obtained for the larger integral scale mixtures. For both mixtures, the slope of the turbulent flame speed–mean strain rate correlation appears to increase as the flame grows.

From Figs. 3*a* and 3*b*, the strain rate seems to become progressively more effective in enhancing the flame speed as the flame grows. This is illustrated by an increase in the slope C_3 with increasing flame radius r. From 19 to 55 mm r, the increase in C_3

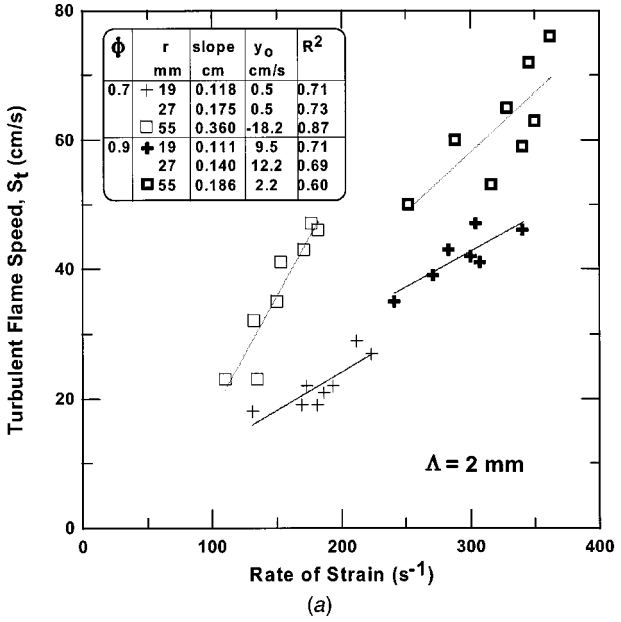


Fig. 3 Turbulent flame speed as functions of mean strain rate, equivalence ratio, and mean flame radius for: (*a*) Λ=2 mm turbulence, (*b*) Λ=8 mm turbulence

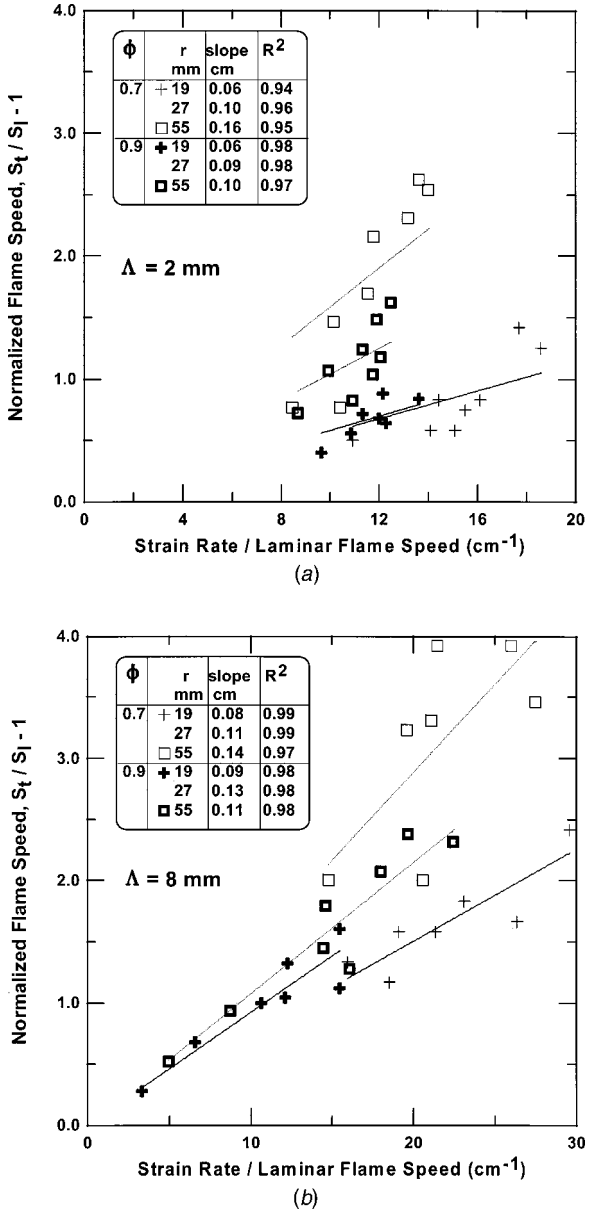


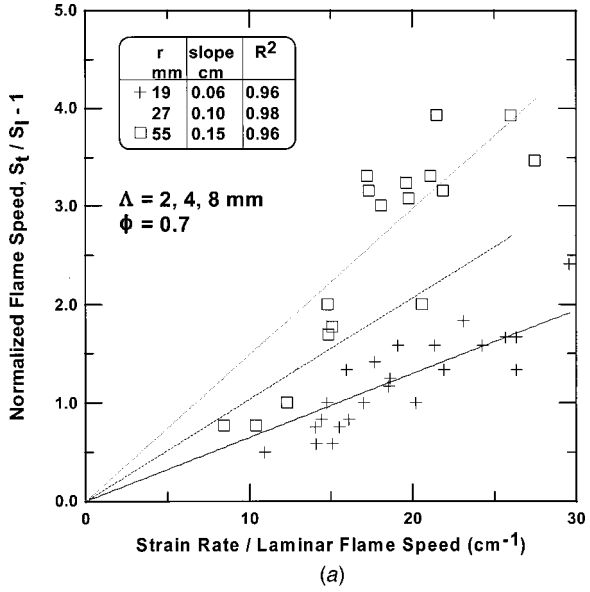
Fig. 4 Normalized turbulent flame speed as functions of mean strain rate, equivalence ratio, and mean flame radius for (*a*) Λ=2 mm turbulence, (*b*) Λ=8 mm turbulence

appears to be larger for the leaner, 0.7 equivalence ratio flames than the 0.9 equivalence ratio flames. This trend may be attributed to the small differences in Ma and turbulence-flame interaction period. The somewhat positive Ma, ϕ=0.9 flames tend to attenuate the increase in flame speed as the flame grows. In other words, the ϕ=0.9 flames are more stable compared to the ϕ=0.7 flames. The small difference in stability may lead to the apparent less effective straining for the richer flames. The longer flame–turbulence interaction period for the slower burning, leaner flames may also contribute to their more effective turbulent straining. The schlieren images seem to confirm this argument as the degree of flame front wrinkling appeared to be higher for the leaner flames.

For Le≈1 and Ma≈0 mixtures, the effect of equivalence ratio is generally accounted for by nondimensionalizing flame speed and turbulence intensity results with laminar flame speed. If both sides of Eq. (4) are so treated, the flame speed–mean strain rate correlation can be expressed as:

$$S_t S_l - 1 = C_s(1/S_l)(u'/\lambda), \tag{5}$$

where C_s is a dimensional constant with units in cm^{-1}. When treated in this form, the data set from Fig. 3 gives Figs. 4*a* and 4*b*. Linear fits passing through the origin are used to fit the data. This procedure retains the physics behind the results as zero turbulent strain rate corresponds to a quiescent mixture. Hence, the line is expected to pass through zero at zero turbulent strain; recall an earlier discussion that the instantaneous laminar flame speed includes all nonturbulence effects. Again, the data points for the intermediate flame size, 27 mm radius, are not shown to reduce the clutter of overlapping data sets, and the Λ=4 mm case is not shown due to space limit. Figure 3 showed that the ϕ=0.9 cases had a wider range (and higher) of absolute shear rates than the ϕ=0.7 cases, particularly for the larger flame sizes. This is be-



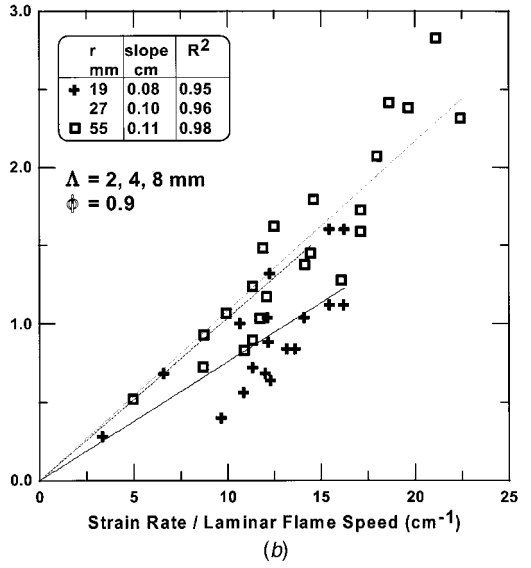
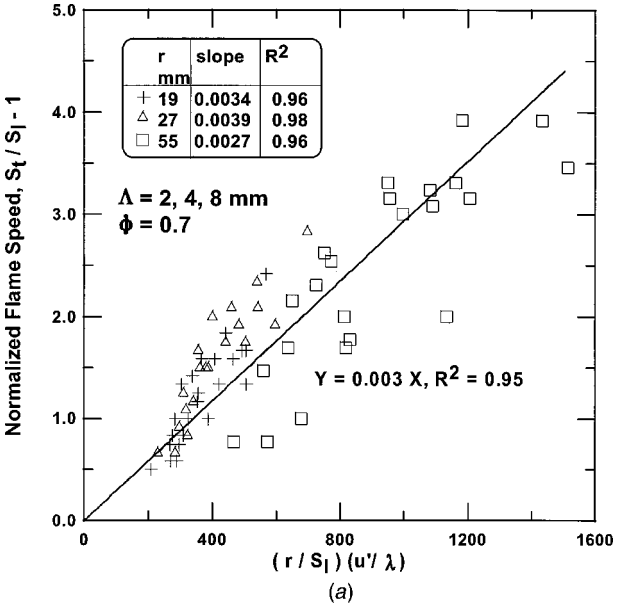


Fig. 5 Normalized turbulent flame speed S_t/S_l-1 as a function of mean strain rate/laminar flame speed for (*a*) ϕ=0.7 mixtures, (*b*) for ϕ=0.9 mixtures



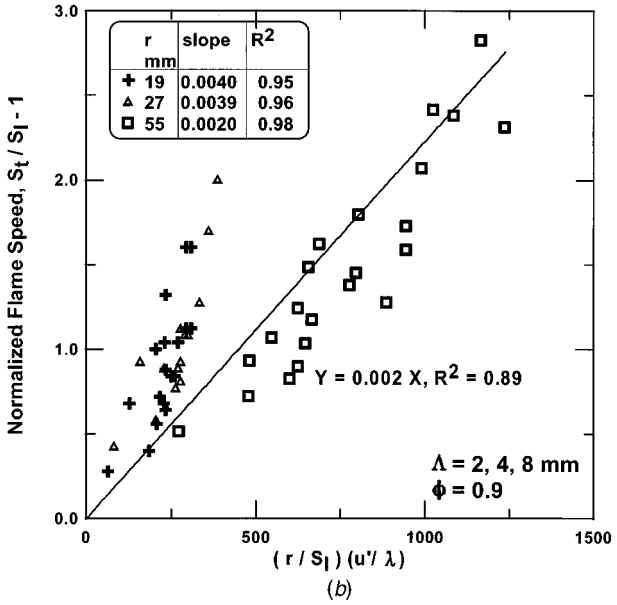


Fig. 6 Normalized turbulent flame speed S_t/S_l-1 versus $(r/S_l)(u'/\lambda)$ for (*a*) ϕ=0.7 mixtures, (*b*) ϕ=0.9 mixtures

lieved to be a consequence of the rapid turbulence decay, which substantially reduced the turbulent shear rate during combustion, particularly for the leaner, slower-burning mixtures.

It is notable in Figs. 4*a* and 4*b* that the normalized curves do not fall onto a single line. The slopes are similar for the smallest flame size (19 mm) but the difference in slopes is substantial for the largest flame size (55 mm). This indicates a significant size effect with the larger flame sizes having their relative flame speeds increased substantially more at the same level of relative shear. This trend is more significant for the leaner, 0.7 equivalence ratio flames due, probably, to differences in Ma and flame–turbulence interaction duration as discussed earlier.

By comparing the corresponding slopes for a single flame size and equivalence ratio across Figs. 4*a* and 4*b*, it becomes clear that the normalization procedure has virtually eliminated the turbulence scale effect. The slope is essentially the same for an integral scale of 2 mm, (4 mm), or 8 mm. Hence, it is possible to combine the data in Figs. 4*a* and 4*b* into groups of specific r and ϕ. The recombined data sets are plotted in Figs. 5*a* and 5*b* for the ϕ=0.7 and 0.9 cases, respectively. Again, the data sets are fitted with straight lines passing through the origin and the data points for the 27 mm radius flames are not shown to avoid clutter. From Figs. 4 to 5, the decrease in the coefficient of determination is rather small. This indirectly validates the assumption that turbulent length scale has been properly incorporated in mean shear rate.

Figures 5*a* and 5*b* depict progressively more effective enhancement of turbulent flame speed as the flame grows for both ϕ=0.7 and ϕ=0.9 mixtures. This progressive enhancement is most evident for the leaner ϕ=0.7 mixtures, probably because the ϕ=0.7 mixture has Ma of -0.2 while the ϕ=0.9 mixture has Ma of 1.8. As explained earlier, the somewhat positive Ma for ϕ=0.9 mixtures implies that the turbulent straining enhancement on the turbulent flame speed would be attenuated slightly. This attenuation, along with the shorter history of flame–turbulence interaction for faster burning mixtures may explain the lower increase in turbulent flame speed–mean strain rate correlation, as these faster-burning flames grow. Note that the progressive turbulent straining

effect on flame speed as the flame grows is much larger than the small discrepancy introduced by the two different techniques (schlieren versus pressure).

It appears from Fig. 5, that the slope C_s increases in proportion to the mean flame radius r. Thus, Eq. (4) may be rewritten as:

$$S_t/S_l - 1 = C_r(r/S_l)(u'/\lambda), \tag{6}$$

where C_r is a nondimensional constant with a value about 0.003. This expression is analogous to that in Ref. [34] which was derived from a flame growth model based on eddy structures in the turbulent mixture.

Figures 6*a* and 6*b* plot the very same data set turbulent flame speed–mean strain rate correlation in the form of Eq. (6) for $\phi=0.7$ and $\phi=0.9$ mixtures, respectively. Straight lines passing through the origin are used to fit the data points from each flame radius and also the whole data set from all flame sizes. For the $\phi=0.7$ mixtures, as shown in Fig. 6*a*, the data points follow Eq. (6) fairly closely. The fits for the schlieren results, the 19 and 27 mm cases, appear to give somewhat higher slopes. These higher slopes could be attributed to the exaggerated turbulent flame speeds determined from the schlieren images, as discussed earlier. In general, Eq. (6) appears to describe the overall result well. Hence, for unity Le and zero Ma flames, the turbulent flame speed–mean strain rate correlation can be expressed by Eq. (6) with a C_r of 0.003.

Figure 6*b* shows that the $\phi=0.9$ results do not fit Eq. (6) as well. The decrease in C_r with increasing r is much larger than the discrepancy introduced by the difference between schlieren and pressure trace analyses ($\approx$5 percent C_r); i.e., the progressive increase in the turbulent straining effectiveness as the flame grows is less linear for a $\phi=0.9$ mixture than for the $\phi=0.7$ case. The faster-burning $\phi=0.9$ flame may have approached a plateau of flame area enhancement by the time the flame reaches 55 mm radius, and/or, the slightly positive Markstein number might have attenuated the turbulent flame speed enhancement by straining.

Notes on Turbulence Modeling

In this study, the mean turbulent strain rate at the flame front has been adjusted for the effects of viscous decay and compression. While the turbulence decay rate has been confirmed by measurements in ''cold'' runs [35], the compression effects were difficult to confirm experimentally. Since this compression adjustment could affect the relationships evident in the final results, the relative importance of the compression effect is examined here. The ''rapid distortion'' compression enhancement of turbulence assumes negligible extra decay and thus maximizes the compression effects. Thus, the true turbulence level is likely to fall somewhere between the value used and the value obtained by simple decay. The influence of that uncertainty is examined by comparing experimental results processed with the compression effect included and neglected completely.

Figure 7 illustrates the effects of the compression turbulence adjustment on the turbulent flame speed–mean strain rate correlation for the largest flame sizes measured; where the effect of compression is expected to be maximum. The figure depicts that the inclusion of compression effects on top of the normal turbulence decay only lowers the slope of the S_t-u'/λ correlation ($\approx$30 percent); that is, it does not affect the coefficients of determination R^2 nor the extrapolated laminar flame speeds (y intercepts) significantly. In other words, inclusion or exclusion of compression effects only shifts the slopes of the relationships presented, it does not alter the conclusions.

As mentioned earlier, the mean strain rate correlation results depend heavily on the accuracy of the model used for estimating the strain rate from measured turbulence quantities. Figure 8 shows the effects of two strain rate models on the turbulent flame speed–mean strain rate correlation for $\phi=0.7$ mixtures. Linear fits are used to fit the set of data points. The figure shows that the strain rate relationships from combustion chambers [17] lead to much more consistent fits with lesser data scatters than expressions based on pipe turbulence [9]. The same results were obtained with the $\phi=0.9$ mixtures (the graph for this case is not shown due to space limitation). Using strain rates based on pipe turbulence also gave lower slopes which could be extrapolated to unreasonably high laminar flame speeds. Therefore, it appears that turbulence correlations obtained from pipe flow are not suitable for estimating the rate of strain in a closed chamber.

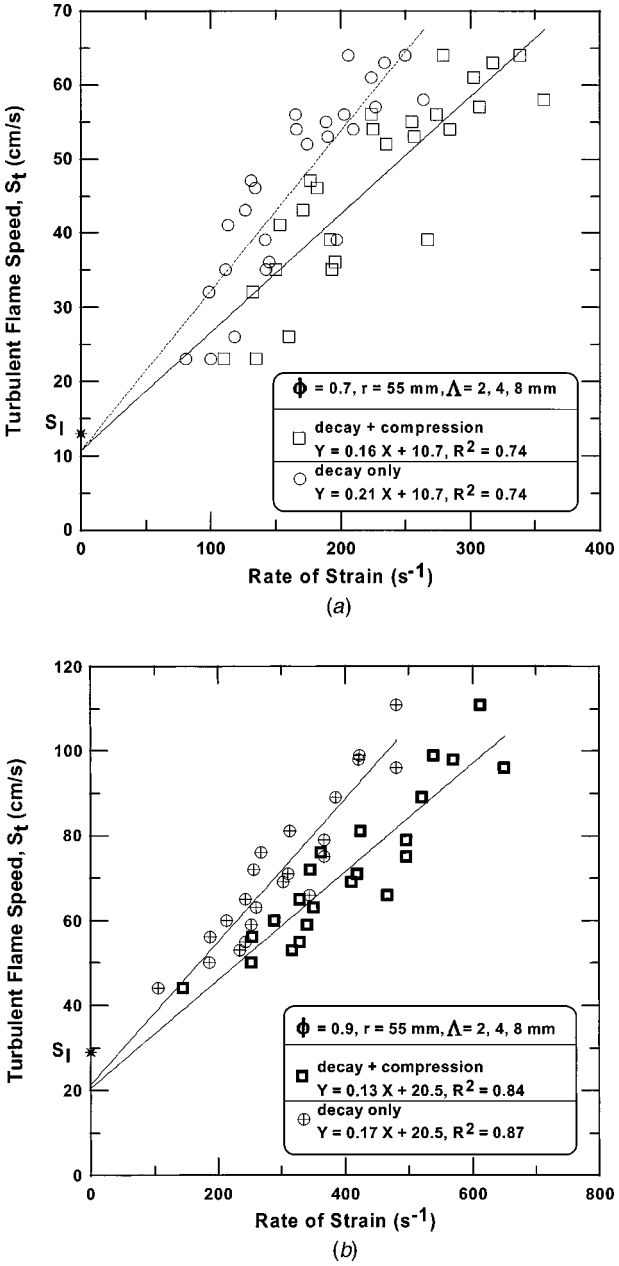


Fig. 7 The effects of flame front turbulence adjustment on the turbulent flame speed–mean strain rate correlation

Concluding Remarks

The turbulent flame speed–mean strain rate relation was studied in a 125 mm cubical combustion chamber with decaying turbulence. Using two different flame speed measurement techniques, and a range of flame sizes, it was found that there is an approximately linear correlation between the turbulent flame speed and the mean strain rate u'/λ. The turbulent straining becomes more effective in enhancing the turbulent flame speed as

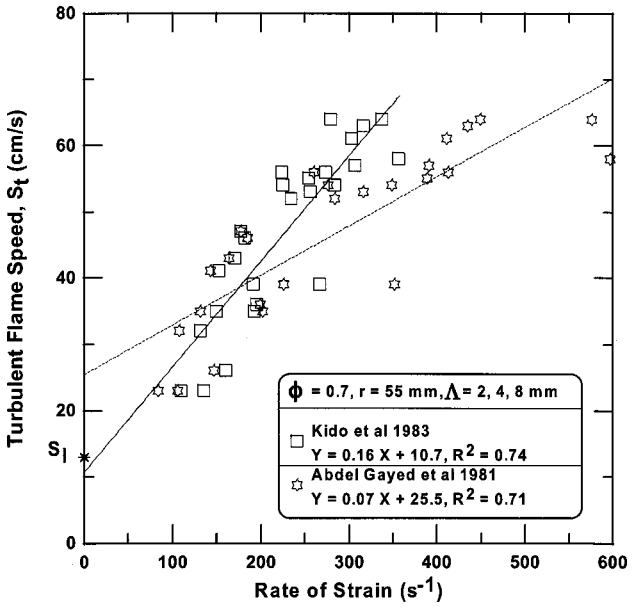


Fig. 8 The effects of strain rate model on the turbulent flame speed–mean strain rate correlation for ϕ=0.7 mixtures

the flame grows. For these methane–air flames with unity Lewis number and near zero Markstein number, the effectiveness of turbulent strain in enhancing flame speed appears to increase linearly with the mean flame radius. It is not possible to determine when the increase with flame size will end and the flame becomes ''fully developed.'' However, there is a suggestion that the increase is becoming more gradual for 0.9 equivalence ratio flames at 55 mm flame radius. Note that this flame size is large relative to typical flame travel distances in spark ignition engines. Hence, most or all of the flame travel in spark ignition engines will probably be within the developing flame regime.

The smaller increase in the effectiveness of turbulent strain in enhancing flame speed of the 0.9 equivalence ratio flames may be due to the slightly positive Ma, which tends to attenuate the turbulent straining, hence, stabilizing the flame as it grows. In addition, the shorter flame–turbulence interaction duration for the faster burning, 0.9 equivalence ratio flames may also lead to the lesser increase in the effectiveness of turbulent straining.

The effects due to flame radius, flame–turbulence interaction history, or time duration need to be examined in detail. The ideal experiment would involve maintaining a spherical, premixed, turbulent flame ball at specific sizes in a steady turbulent field over a period of time. The ability to hold the flame size and the turbulence level constant over the required time period could single out the relative flame size/turbulent eddy size effects from time period effects. If these experiments could be achieved, the effects of Lewis number and Markstein could be easily studied by adjusting mixture properties. The more realistic alternative, however, is to develop a similar experiment with significantly larger combustion chambers to observe more of the processes of a developing and maturing turbulent flame. The oscillating-grid and parallel-plate designs used by Srdic et al. [36] and by Shy et al. [37] would help in devising cleaner (nearer to isotropy) and higher intensity turbulence generation, respectively.

Acknowledgments

The authors gratefully appreciate financial support from NSERC, and several excellent comments from two particular reviewers.

References

[1] Trouve, A., and Poinsot, T. J., 1994, J. Fluid Mech., **278**, pp. 1–31.
[2] Williams, F. A., 1976, Combust. Flame, **26**, pp. 269–270.
[3] Abdel-Gayed, R. G., Bradley, D., Hamid, M. N., and Lawes, M., 1984, 20th Symposium (International) on Combustion, pp. 505–512.
[4] Mantzaras, J., Felton, P. G., and Bracco, F. V., 1988, SAE paper 881635 or SP-759.
[5] She, Z.-S., Jackson, E., and Orszag, S. A., 1990, Nature (London), **344**, pp. 226–228.
[6] Vincent, A., and Meneguzzi, M., 1994, J. Fluid Mech., **258**, pp. 245–254.
[7] Batchelor, G. K., 1952, R. Soc. London, A, **213**, pp. 349–366.
[8] Thomas, A., 1986, Combust. Flame, **65**, pp. 291–312.
[9] Abdel-Gayed, R. G., and Bradley, D., 1981, Philos. Trans. R. Soc. London, Ser. A, **301**, pp. 1–25.
[10] Abdel-Gayed, R. G., Bradley, D., and Lawes, M., 1987, Proc. R. Soc. London, Ser. A, **414**, pp. 389–413.
[11] Bradley, D., 1992, 24th Symposium (International) on Combustion, pp. 247–262.
[12] Checkel, M. D., and Thomas, A., 1994, Combust. Flame, **96**, pp. 351–370.
[13] Abdel-Gayed, R. G., Bradley, D., and Lau, A. K. C., 1998, 22nd Symposium (International) on Combustion, pp. 731–738.
[14] Andrews, G. E., Bradley, D., and Lwakabamba, S. B., 1975, Combust. Flame, **24**, pp. 285–304.
[15] Bradley, D., Lau, A. K. C., and Lawes, M., 1992, Philos. Trans. R. Soc. London, Ser. A, **338**, pp. 359–387.
[16] Tennekes, H., and Lumley, J. L., 1972, *A First Course in Turbulence*, MIT Press, Cambridge, MA.
[17] Kido, H., Wakuri, Y., and Nakashima, K., 1983, Proceedings of ASME-JSME Thermal Engrg. Joint Conference, Vol. 4, pp. 183–190.
[18] Chen, C., and Veshagh, A., 1991, SAE paper 910077.
[19] Boulouchos, K., Steiner, T., and Dimopoulos, P., 1994, SAE paper 940476.
[20] Chen, C., and Veshagh, A., 1992, SAE paper 920679 pp.
[21] Palm-Leis, A., and Strehlow, R., 1969, Combust. Flame, **13**, pp. 112–129.
[22] Abraham, J., Williams, F. A., and Bracco, F. V., 1985, SAE paper 850345.
[23] Trautwein, S. E., Grudno, A., and Adomeit, G., 1990, 23rd Symposium (International) on Combustion, pp. 723–728.
[24] Kwon, S., Wu, M.-S., Driscoll, J. F., and Faeth, G. M., 1992, Combust. Flame, **88**, pp. 221–238.
[25] Bielert, U., Klug, M., and Adomeit, G., 1994, SAE paper 940211.
[26] Ashurst, W. T., 1995, Combust. Sci. Technol., **104**, pp. 19–23.
[27] Groff, E. G., and Matekunas, F. A., 1980, SAE paper 800133.
[28] Rutland, C. J., and Ferziger, J. H., 1990, Combust. Sci. Technol., **73**, pp. 305–326.
[29] Bray, K. N. C., and Cant, R. S., 1991, Proc. R. Soc. London, Ser. A, **434**, pp. 217–240.
[30] Abdel-Gayed, R. G., Al-Khishali, K. J., and Bradley, D., 1984, Proc. R. Soc. London, Ser. A, **391**, pp. 393–414.
[31] Tseng, L.-K., Ismail, M. A., and Faeth, G. M., 1993, Combust. Flame, **95**, pp. 410–426.
[32] Ting, D. S-K., Checkel, M. D., Haley R., and Smy, P. R., 1994, SAE paper 940687.
[33] Irvine, T. F., and Liley, P. E., 1984, *Steam and Gas Tables with Computer Equations*, Academic, New York.
[34] Ashurst, W. T., Checkel, M. D., and Ting, D. S.-K., 1994, Combust. Sci. Technol., **99**, pp. 51–74.
[35] Checkel, M. D., 1986, ASME paper 85-WA/FE-2.
[36] Srdic, A., Fernando, H. J. S., and Montenegro, L., 1996, Exp. Fluids, **20**, pp. 395–397.
[37] Shy, S. S., I, W. K., and Lin, M. L., 2000, Exp. Therm. Fluid Sci., **20**, pp. 105–114.

A Mechanism of Combustion Instability in Lean Premixed Gas Turbine Combustors

T. Lieuwen
H. Torres
C. Johnson
B. T. Zinn

Schools of Aerospace and Mechanical Engineering,
Georgia Institute of Technology,
Atlanta, GA 30332-0150

There has been increased demand in recent years for gas turbines that operate in a lean, premixed (LP) mode of combustion in an effort to meet stringent emissions goals. Unfortunately, detrimental combustion instabilities are often excited within the combustor when it operates under lean conditions, degrading performance and reducing combustor life. To eliminate the onset of these instabilities and develop effective approaches for their control, the mechanisms responsible for their occurrence must be understood. This paper describes the results of an investigation of the mechanisms responsible for these instabilities. These studies found that combustors operating in a LP mode of combustion are highly sensitive to variations in the equivalence ratio (ϕ) of the mixture that enters the combustor. Furthermore, it was found that such ϕ variations can be induced by interactions of the pressure and flow oscillations with the reactant supply rates. The ϕ perturbations formed in the inlet duct (near the fuel injector) are convected by the mean flow to the combustor where they produce large amplitude heat release oscillations that drive combustor pressure oscillations. It is shown that the dominant characteristic time associated with this mechanism is the convective time from the point of formation of the reactive mixture at the fuel injector to the point where it is consumed at the flame. Instabilities occur when the ratio of this convective time and the period of the oscillations equals a specific constant, whose magnitude depends upon the combustor design. Significantly, these predictions are in good agreement with available experimental data, strongly suggesting that the proposed mechanism properly accounts for the essential physics of the problem. The predictions of this study also indicate, however, that simple design changes (i.e., passive control approaches) may not, in general, provide a viable means for controlling these instabilities, due to the multiple number of modes that may be excited by the combustion process. [DOI: 10.1115/1.1339002]

1. Introduction

Increasingly stringent emissions legislation has increased the demand for lean, premixed (LP) combustors (see Fig. 1) that operate at low temperatures. However, LP systems are prone to detrimental combustion instabilities in various operating ranges, including the lean conditions where they are designed to operate [1]. To develop rational approaches for preventing or controlling these instabilities, an understanding of the controlling mechanism(s) and capabilities for predicting the conditions under which they occur must be developed. Intensive experimental and theoretical work has been performed during the past few years [1–17] to understand these mechanisms and to develop effective approaches for their control. A number of experimental studies have characterized the operating conditions under which LP combustion instabilities occur [1–7]. Also, several theoretical studies have attempted to elucidate the mechanism(s) that drive the instability and to develop models that can predict their occurrence [8–13]. Yet, in spite of these efforts, there is no consensus about the mechanisms that are responsible for the onset of these instabilities.

Prior studies by the authors [10–12], Straub and Richards [3], and Peracchio and Proscia [9] suggested that heat release oscillations excited by fluctuations in the composition of the reactive mixture entering the combustion zone were the dominant mechanism responsible for the instabilities observed in these combustors. However, limited experimental evidence was available to critically assess their observations or theory. Furthermore, the theory developed in Refs. [11–12] could not account for the observed dependence of the combustor stability upon flame structure. This paper describes recent work of the authors that extends the theory to account for flame structure effects, as well as comparisons of the theory with measurements at several facilities.

Section 2 presents the mechanism of instability proposed by the authors, necessary conditions under which it will cause self-excited oscillations, and an analysis that can be used to predict the conditions under which combustors become unstable. Next, it presents results that illustrate the agreement between the predictions of the developed theory and experimental data. The paper closes with a discussion of possible approaches for passive control and of other potential instability mechanisms.

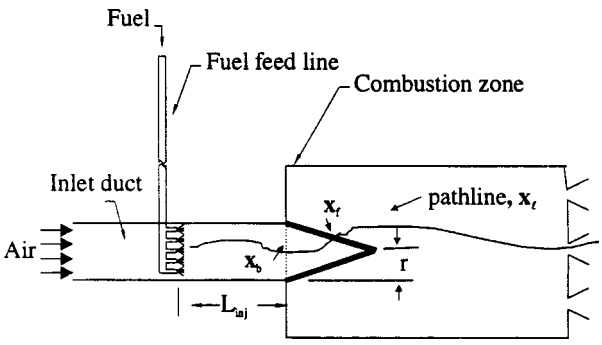


Fig. 1 Schematic of a typical LP combustor

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-3. Manuscript received by IGTI March 9, 1999; final revision received by the ASME Headquarters Apr. 2, 2000. Associate Technical Editor: J. E. Peters.

2. Mechanism of LP Combustion Instability

The primary difficulty in combustion instability studies is understanding the interactions between the unsteady heat release processes and the disturbances that drive them. This task is difficult because these heat release oscillations may be due to oscillations of the velocity, pressure, temperature, and reactants composition that are present simultaneously in combustion systems. Though heat release oscillations may be excited by a host of disturbances, there are two conditions which must be met for self-excited, combustion driven oscillations to occur. First, the unsteady heat release processes must be properly phased with the fluctuating acoustic pressure so that energy is added to the unsteady motions (i.e., Rayleigh's criterion) [10]. Second, the rate of energy addition must exceed the rate of energy dissipation.

This section describes a mechanism that appears to be responsible for LP combustion instabilities, shows what operating ranges will satisfy the conditions discussed above, and consequently, predicts the regions under which instabilities may be observed. A schematic of the feedback process upon which the proposed mechanism is based on is shown in Fig. 2. Its main elements are the generation of heat release oscillations by periodic variations in ϕ of the reactive mixture that enters the flame, and the formation of ϕ oscillations in the inlet section by velocity and pressure oscillations in the vicinity of the fuel injector. These processes are described in more detail in the following subsections.

2.1 Response of LP Systems to ϕ Oscillations. This study was partially motivated by observations that properties of premixed combustion systems, such as flame thickness, flame speed, and reaction rate become increasingly sensitive to variations in ϕ as the combustion process stoichiometry becomes leaner [10]. Moreover, systems operating near the lean limit are acutely sensitive to ϕ perturbations since they may cause periodic extinction of the combustion process. To further examine these observations, an unsteady well stirred reactor (WSR) model was developed and its response to ϕ perturbations in its inlet conditions was studied [10]. Figure 3 presents a result from this study and shows the calculated reaction rate response to ϕ perturbations as the mean equivalence ratio is decreased. This figure shows that the response of the WSR's reaction rate to perturbations in the inlet ϕ significantly increases (by 2 orders of magnitude) as the mean value of ϕ decreases. This result strongly suggests that ϕ oscillations could drive substantial heat release oscillations under lean operating conditions. It also suggests that ϕ oscillations are unlikely to drive instabilities near stoichiometric conditions because of the negligible response of the reaction rate at this operating condition. Some experimental observations that the oscillating pressure amplitude in LP combustors is relatively small under stoichiometric conditions and becomes progressively larger as ϕ is reduced [1] appear to support these assertions and the trends depicted in Fig. 3.

2.2 Formation of ϕ Oscillations. The discussion in Sec. 2.1 suggests that the sensitivity of LP systems to ϕ oscillations could be responsible for their unstable behavior. While random ϕ fluctuations undoubtedly occur in LP combustors (e.g., due to turbulent mixing), ϕ fluctuations can only play a role in an instability mechanism if they are driven by the combustion process and the resulting pressure and velocity oscillations, thus closing the feedback loop needed to maintain an instability; see Fig. 2. Consequently, it is necessary to consider how ϕ fluctuations can arise, and to elucidate the feedback mechanism between the ϕ and heat release oscillations that drive the instability. The following equation, derived from the definition of ϕ, suggests a mechanism for the formation of ϕ oscillations in the inlet section due to velocity and pressure perturbations:

$$\frac{\phi'}{\overline{\phi}} = \frac{\dfrac{m_f'}{\overline{m_f}} - \dfrac{m_o'}{\overline{m_o}}}{1 + \dfrac{m_o'}{\overline{m_o}}} \tag{1}$$

In Eq. (1), the subscripts f and o denote fuel and oxidizer (assumed to be air from this point on), respectively, and m the mass flow rate. This equation shows that ϕ oscillations can be formed by air and/or fuel flow oscillations (that are present at the fuel injector). Furthermore, for the low Mach number flows typical of these systems, this equation implies that small acoustic fluctuations can generate significant ϕ fluctuations; e.g., for $M=0.05$ and a choked fuel injector (i.e., $m_f'=0$), acoustic oscillations of 1 percent of the mean (i.e., $u'/c=0.01$) generate ϕ oscillations of 20 percent [i.e., $\phi'/\overline{\phi}=0.20$, because $\phi'/\overline{\phi}\approx(1/M)(u'/c)$]. The possibility of forming such large amplitude ϕ oscillations from modest flow perturbations suggests that even though diffusive and turbulent mixing processes will tend to homogenize the mixture as it flows from the fuel injector towards the combustor, the ϕ fluctuations may still persist at the flame.

This discussion suggests that acoustic oscillations in a combustion chamber, such as the one shown in Fig. 1, may be accompanied by oscillations in the composition of the reactive mixture in the inlet section. Furthermore, Fig. 2 shows that the presence of these ϕ oscillations in the inlet section are a critical component of this mechanism of instability. It is significant that measurements

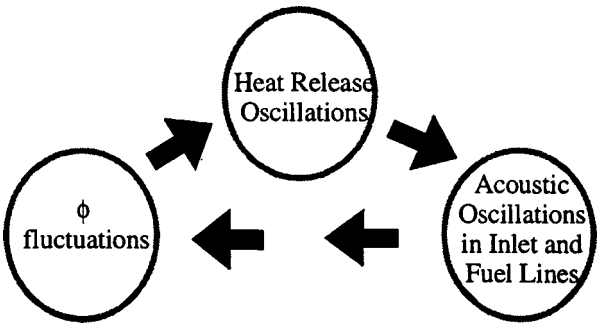


Fig. 2 Schematic of the feedback loop possibly responsible for LP combustion instabilities

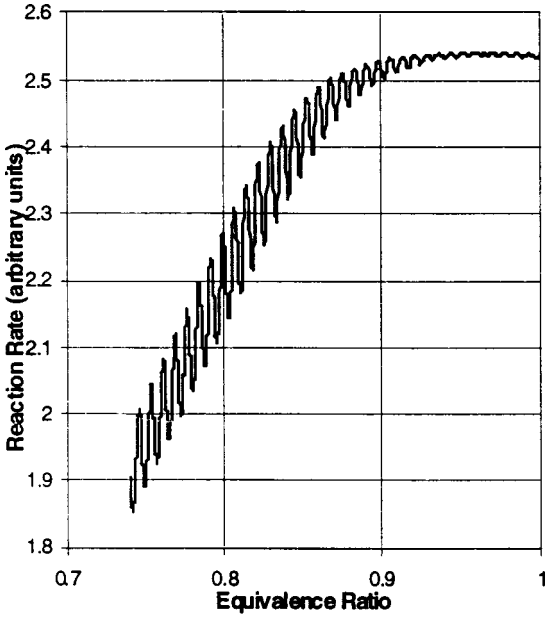


Fig. 3 Dependence of the response of the WSR reaction rate to ϕ perturbations in its inlet flow upon the mean value of ϕ

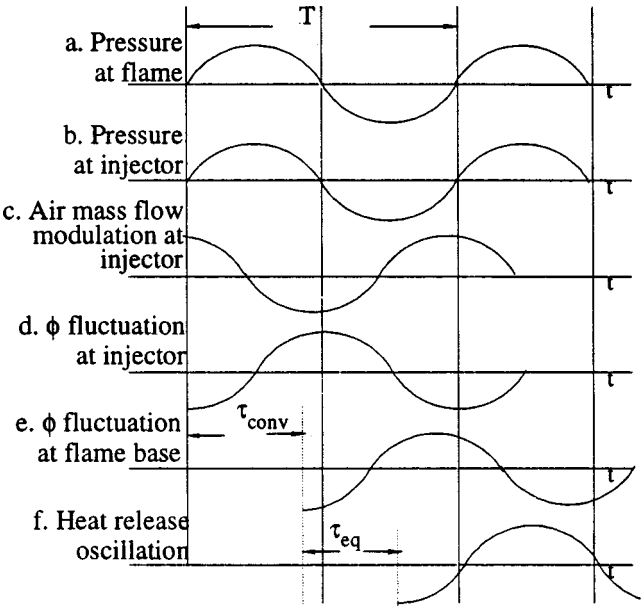


Fig. 4 Schematic showing the time evolution of disturbances responsible for a combustion instability

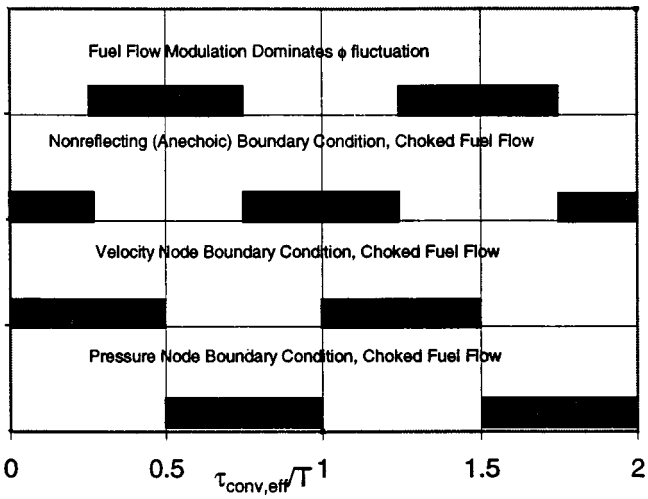


Fig. 5 Dependence of instability regions (hatched regions) upon $\tau_{conv,eff}/T$ for several different combustor configurations

of such oscillations have been recently reported [14]. Mongia et al. [14] presented measurements of fuel mass fraction oscillations in the combustor inlet section with frequencies corresponding to the instability frequency, confirming the presence of oscillations in the reactive mixture composition discussed in this section.

2.3 Conditions for Instability. The discussions in Secs. 2.1 and 2.2 have outlined the framework for a combustion instability mechanism that is potentially responsible for those observed in LP combustors. This mechanism and its controlling steps/parameters can be better understood by considering the time evolution of the various processes that sustain it, see Fig. 4 (see Ref. [18] for the application of similar methods of analysis to other combustion systems). Figure 4*a* shows the time dependence of the acoustic pressure at the flame of an unstable mode with a period T (the length of the combustion region is assumed to be very small relative to a wavelength). This disturbance propagates upstream into the inlet duct and produces pressure oscillations at the fuel injector, see Fig. 1. As long as no pressure node exists between the flame and fuel injector (implying a 180 Deg phase change across the node) and the Mach number is low, the pressure disturbances at the flame and fuel injector are nearly in phase, see Fig. 4*b*.

These pressure oscillations are accompanied by fluctuations in the velocity, and therefore the oxidizer flow rate at the fuel injector. However, the phase of these velocity oscillations relative to the pressure depends on the upstream boundary conditions (e.g., if the upstream boundary is a pressure node, the velocity leads the pressure by 90 Deg). For the sake of brevity, we carry out this analysis assuming a pressure node boundary condition and quote the results for the other cases.

For low Mach number flows, the velocity and mass flow rate are nearly in phase, see Fig. 4*c*. Assuming choked fuel flow (i.e., $m_f'=0$), Eq. (1) shows that the fluctuating ϕ of the reactive mixture formed in the vicinity of the fuel injector is out of phase with these air flow oscillations, see Fig. 4*d*. This leads to the formation of a reactive mixture with a periodically varying ϕ that is convected by the flow and reaches the base of the flame after a convective time $\tau_{convect}$, see Fig. 4*e*.

The reactive mixture is not consumed instantly when it reaches the flame base (even when any delays due to purely chemical kinetic processes are neglected) because ''different parts'' of the flow entering the combustor are consumed at different locations of the flame surface; i.e., the mixture is not completely consumed until it reaches the end of the flame region, see Fig. 1. Thus, the total (i.e., spatially integrated) heat release lags the ϕ oscillation at the base of the flame by a time, τ_{eq} (to be discussed later), see Fig. 4*f*. It follows from Rayleigh's criterion that an instability can occur if the heat release oscillations in Fig. 4*f* are in phase with the pressure oscillations at the flame in Fig. 4*a*; that is, if:

$$\textbf{Pressure node:}\quad \frac{\tau_{convect}+\tau_{eq}}{T}=C_n=n-1/4 \qquad n=1,2\ldots \tag{2}$$

Using a similar analysis, it can be shown that if the upstream inlet boundary is an acoustically rigid or nonreflecting surface, the corresponding conditions for instability become [11]:

$$\textbf{Velocity node:}\quad \frac{\tau_{convect}+\tau_{eq}}{T}=C_n=n-3/4 \tag{3}$$

$$\textbf{Nonreflecting:}\quad \frac{\tau_{convect}+\tau_{eq}}{T}=C_n=n \tag{4}$$

Finally, for situations where the fuel injector is unchoked, the dominant contribution to the ϕ fluctuation comes from fuel flow modulation (see Eq. (1)), and the fuel flow perturbation is 180 Deg out of phase with the pressure perturbation at the injection point, the instability condition is [11]:

$$\textbf{Fuel flow modulation:}\quad \frac{\tau_{convect}+\tau_{eq}}{T}=C_n=n-1/2 \tag{5}$$

Thus, instability regions are primarily a function of $(\tau_{convect}+\tau_{eq})/T$, which will be denoted as $\tau_{conv,eff}/T$. Specifically, Eqs. (2)–(5) indicate that the regions of instability approximately center about locations where $\tau_{conv,eff}/T=C_n$, where C_n is a constant that depends on the combustor configuration. In a combustion system without damping, these instability regions lie in bands where $C_n-1/4<\tau_{conv,eff}/T<C_n+1/4$ (that is, the limits occur where the phase between the acoustic pressure and heat release is 90 Deg). Figure 5 illustrates the instability regions for the combustor configurations considered above.

Note that it is possible that a region where an instability can occur will actually remain stable if the energy addition due to the unsteady heat release cannot overcome damping processes. Thus, the instability conditions of Eqs. (2)–(5) and depicted in Fig. 5 are necessary, but not sufficient conditions for an instability. For example, the negligible response of the reaction rate to ϕ fluctua-

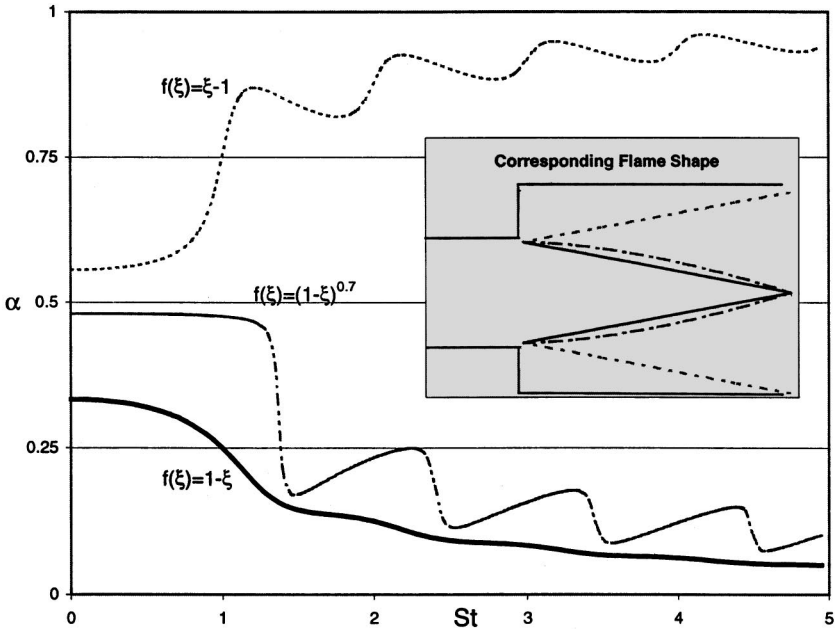


Fig. 6 Dependence of α on the flame Strouhal number, $St = fL_{flame}/u$ for three different assumed flame shapes

tions under stoichiometric conditions, see Fig. 3, suggests that even when the operating conditions are conducive to an instability, the combustor will remain stable (or else become unstable through some other mechanism). Under operating conditions where instabilities can occur, the presence of damping narrows the ''width'' of the instability regions, causing the ''edges'' of otherwise unstable regions to become stable.

2.4 Stability Calculations. In order to use these results to predict conditions for instability in real systems, the quantity $\tau_{conv,eff}/T$, and thus the convective times, $\tau_{convect}$ and τ_{eq}, must be determined. The convective time, $\tau_{convect}$, can be approximated as the distance from the fuel injector to the base of the flame, divided by the average flow velocity in the inlet, see Fig. 1:

$$\tau_{convect} = L_{inj}/\overline{u} \tag{6}$$

More sophisticated computations that better characterize the multidimensional flow dynamics in the inlet section have been reported in [9]. However, the results of this study seem to indicate that Eq. (6) is a good approximation, although additional effects, such as stratification of ϕ in the radial direction, were noted.

Determining τ_{eq} is more difficult. Since different parts of a given cross section of reactive mixture are consumed at different times and places, the effects of the structure of the flame region and the phasing of the ϕ fluctuation when it is consumed are important, see Fig. 1. Although these effects have previously been noted by the authors [11] and by Straub and Richards [3], no attempts were made to quantify them. An analysis that accounts for these effects in a slightly different context has been previously presented by Putnam [18].

A more general analysis than that given in Ref. [18] to determine this characteristic time is given in the Appendix. The analysis is general enough to account for three dimensional nonuniformities in the reactive mixture composition and a complex flowfield and flame structure. Thus, it can be used in conjunction with computational results of the mean properties of the combustor to predict its stability behavior.

A convenient relation between $\tau_{conv,eff}$ and $\tau_{convect}$ can be derived by defining a ''flame length correction coefficient,'' α (see Appendix):

$$\frac{\tau_{conv,eff}}{T} = \frac{\tau_{convect}}{T}\left(1 + \alpha \frac{L_{flame}}{L_{convect}}\right) \tag{7}$$

The coefficient α is a function of the flame Strouhal number, $St = fL_{flame}/u$, and the structure of the flame region. Physically, α represents the fractional location of a hypothetical flame sheet that consumes the entire mixture at one point downstream of the flame base. For example, a value of $\alpha = 1/3$ implies that the phase introduced between the ϕ oscillation at the base of the flame and the heat release from the distributed flame region is equivalent to that from a hypothetical flame sheet located at $L_{flame}/3$.

Figure 6 shows the results of several calculations of α, illustrating its behavior for different assumed flame shapes over a range of Strouhal numbers (α is given by Eq. (*A*16) in the Appendix). An examination of Fig. 6 reveals several noteworthy characteristics of α. First, note that at low Strouhal numbers, α is only a function of the flame shape and independent of the Strouhal number. Second, the figure illustrates the sensitivity of α to the Strouhal number when $St \sim O(1)$. For example, note that for the curved flame, a change in the value of the Strouhal number of less than 5 percent results in a change of α by 60 percent when $St = 1.4$. To obtain a feel for ranges these values may take in typical combustors, if it is assumed that $f = 200$ Hz, $L_{flame} = 2–5$ cm, $L_{convect} = 5–10$ cm, and $\overline{u} = 30–60$ m/s, then $St \sim 0.05–0.4$, implying that $\alpha \sim 0.5$ (see Fig. 6). Thus, Eq. (7) shows that $\tau_{conv,eff} \sim (1–1.5)\,\tau_{convect}$.

3. Comparisons with Experimental Results

The theory developed in Sec. 2 shows that combustion instabilities induced by ϕ oscillations will occur under conditions where the value of $\tau_{conv,eff}/T$ is within certain ranges, see Fig. 5. It follows that combustor data should correlate with this parameter when ϕ oscillations are responsible for the instabilities. However, as the analysis and results in Sec. 2 show, it may be difficult to determine $\tau_{conv,eff}/T$ because of its dependence on the length and structure of the flame region. Since the value of $\tau_{convect}$ in a given experimental configuration is much easier to determine, the theory developed in this paper suggests that unstable combustor situations occur when

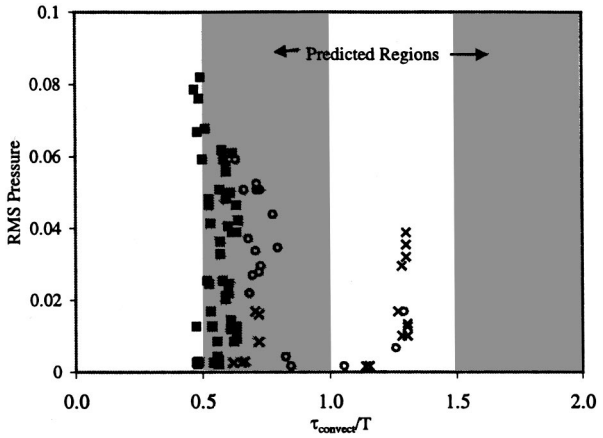


Fig. 7 Dependence of unsteady pressure amplitudes measured in a DOE combustor [3] and its predicted linear stability limits upon $\tau_{convect}/T$

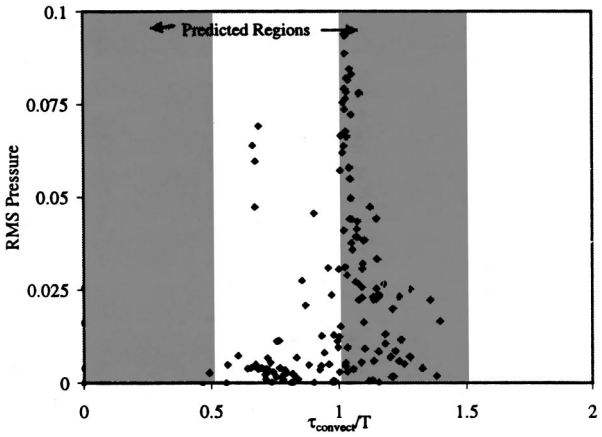


Fig. 9 Dependence of unsteady pressure amplitudes measured in a Georgia Tech combustor and its predicted linear stability limits upon $\tau_{convect}/T$

$$\frac{\tau_{\text{convect}}}{T} = \frac{C_n}{\left(1+\alpha\dfrac{L_{\text{flame}}}{L_{\text{convect}}}\right)} = \tilde{C}_n(\text{St}, f(\xi)) \tag{8}$$

The estimates of typical values in Sec. 2 suggest that instabilities will occur where $\tau_{convect}/T \sim (0.7-1)C_n = \tilde{C}_n$. Note that since $\tilde{C}_n \leqslant C_n$, measured instability regions plotted versus $\tau_{convect}/T$ will always lie to the left of those predicted in Fig. 5 or in Eqs. (2)–(5).

The rest of this section presents experimental results from several facilities showing that the regions of instability are primarily a function of $\tau_{convect}/T$, in agreement with the predictions of this paper. These results also show that the measured dependence of the instability regions on the upstream boundary condition in Fig. 5 are well described by the theory.

Figure 7 shows pressure data obtained by Straub and Richards [3] and the instability regions predicted by Eq. (2). It can be seen that their data collapses into bands when normalized by $\tau_{convect}/T$, as predicted in the preceding discussion. Furthermore, since the inlet section of the combustor was connected to a plenum (see [3]), the upstream boundary condition may be approximated as a pressure node. Then, Eq. (2) predicts that instabilities should occur in the vicinity of $\tau_{convect}/T = 0.75, 1.75, \ldots$. The figure shows that this prediction is in agreement with the measured data, particularly for the first region of instability. It can be seen that the agreement is not as good for the second band of oscillations; this may be due to distributed flame effects and is discussed further below.

Figure 8 shows unsteady pressure data obtained at Penn State [2,19]. Since the upstream boundary of the inlet section in this facility is essentially nonreflecting, Eq. (4) predicts that instabilities should occur when $\tau_{convect}/T = 0, 1, \ldots$. Figure 8 shows that the measured data agree well with this prediction.

Figure 9 shows unsteady pressure data obtained at Georgia Tech [17]. Since the upstream boundary of the inlet section is rigid, Eq. (3) predicts that instabilities should occur when $\tau_{convect}/T = 0.25, 1.25, \ldots$. Figure 9 shows that most of the large amplitude pressure oscillations occur in the predicted region. However, it should be noted that instabilities were observed in four test runs that are well outside of the predicted unstable region (i.e., $\tau_{convect}/T = 0.65$).

An examination of Figs. 7–9 reveals that some of the measured instabilities occurred in $\tau_{convect}/T$ regions that lie to the left of the predicted regions. As the discussion below Eq. (8) indicates, this consistent leftward bias is to be expected, however, because regions of instability should actually occur where $\tau_{convect}/T \sim (0.7-1)C_n = \tilde{C}_n$; i.e., between 0 and 30 percent to the left of those shown in Fig. 5. The shift of approximately 25 percent in the DOE data (see Fig. 7) and 20 percent in the Georgia Tech data (see Fig. 9) suggests that these ''discrepancies'' are simply due to correlating the measured data with $\tau_{convect}/T$ instead of $\tau_{conv,eff}/T$.

The above comparisons between the theoretical predictions and the experimental data demonstrate that LP instabilities occur at specific ranges of the parameter $\tau_{convect}/T$. Significantly, the demonstrated agreement between the theory's predictions and measurements strongly suggests that the mechanism discussed in this paper is responsible for LP instabilities.

One issue that is not clear, however, is the lack of unstable oscillations in bands that are predicted to be unstable. For example, no oscillations have been observed in the Georgia Tech facility in the $0 < \tau_{convect}/T < 0.5$ band even though the combustor has been run under operating conditions where the frequency of some of its natural modes would lie in this unstable range. Excitation of other modes that correspond to higher instability bands, such as $2 < \tau_{convect}/T < 2.5$, have been observed, however, but the dominant unstable mode still corresponds to $1 < \tau_{convect}/T < 1.5$.

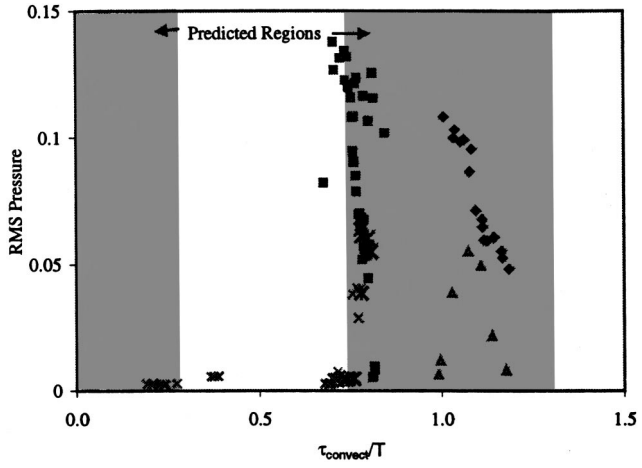


Fig. 8 Dependence of unsteady pressure amplitudes measured in a Penn State combustor [2] and its predicted linear stability limits upon $\tau_{convect}/T$

4. Passive Control Considerations

An understanding of the mechanisms responsible for LP instabilities may provide combustor designers with means for eliminat-

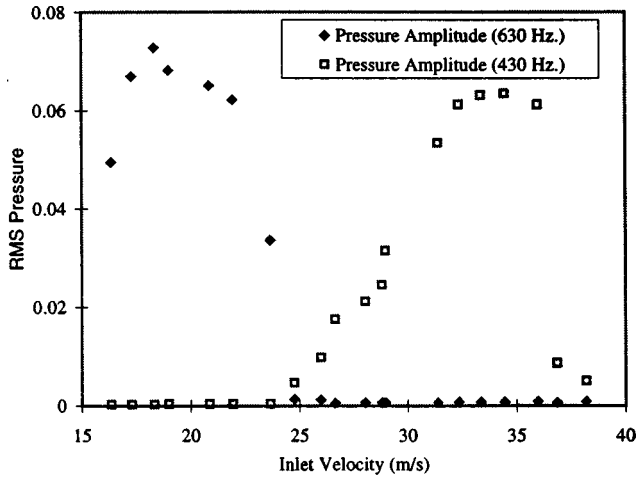


Fig. 10 Dependence of the pressure amplitudes of the 430 and 630 Hz modes, measured at Georgia Tech, upon the mean inlet velocity

ing these instabilities in the design stage. For example, one gas turbine manufacturer claims to have eliminated instabilities by designing a combustor whose operating regions lie outside of the unstable range. Specifically, they have modified τ_{convect} by changing the combustor inlet diameter and fuel injector location [20].

In spite of their success, shifting the value of $\tau_{\text{conv,eff}}/T$ outside the unstable range may be difficult to achieve in general because of design restrictions. Also, several different combustor modes may be excited, and stabilizing one of these modes may destabilize another. For example, Fig. 10 presents data obtained at the Georgia Tech facility, illustrating the dependence of the unsteady pressure amplitude upon the velocity in the inlet section. It shows that at lower velocities, the 430 Hz mode is excited, and that its amplitude decreases with increasing velocity. However, as the inlet velocity is further increased, the 630 Hz mode becomes unstable because its period T is smaller, and the parameter $\tau_{\text{conv,eff}}/T$ is within the unstable range. Figure 10 shows that the combustor is ''quiet'' only within a narrow velocity range.

Similar observations have been made by Straub and Richards [3] who moved the location of the fuel injector and found that stabilizing one mode was often accompanied by destabilization of another (they show that these problems may be dealt with by injecting the fuel at multiple locations). These examples illustrate the potential shortcomings of passive control approaches.

5. Final Remarks

This paper has shown that LP combustion instabilities are likely due to a feedback process between the heat release, acoustic pressure, and ϕ oscillations, see Fig. 2. Furthermore, it has been shown that regions of instability primarily depend upon the parameter τ_{convect}/T, implying that the frequency of oscillation, the mean velocity, and the location of the fuel injector are key parameters controlling combustor stability.

In closing, we would like to reiterate that while this paper has demonstrated that combustion instabilities in LP systems are likely due to ϕ oscillations, it is possible that other mechanisms may also be important. Thus, it should be noted that some studies have concluded that other mechanisms, such as combustion within coherent vortical structures [4,5], are responsible for instabilities in these systems. These conclusions appear to be based on observations of such structures during unstable combustion [4,5].

While it is altogether possible that other mechanisms for instability (such as combustion in vortices) are important, care must be exercised in assessing a mechanism of instability based on observations of unstable combustion at the limit cycle; that is, it is possible that some instability mechanism causes growth of the amplitude of the unsteady motions in a combustor, eventually forcing the periodic shedding of these structures. Thus the presence of these structures does not necessarily supply any information about the *origin* or *mechanism* that initiated the instability. In the same manner, the measurements of ϕ oscillations by Mongia et al. [14] during unstable combustion cannot be interpreted as an indication that ϕ oscillations are responsible for initiating the instability. While Eq. (1) shows that such ϕ oscillations can be excited by acoustic oscillations in the inlet, the instability mechanism that is responsible for exciting these oscillations could be due to a totally different mechanism (e.g., vortex shedding).

To summarize, this discussion shows that observations of combustion instabilities at the limit cycle may not provide useful information about the origin of the instability.

Nomenclature

c = speed of sound
C = stability constant, see Eqs. (2)–(5)
m = mass flow rate
M = Mach number
q = local heat release rate
Q = total rate of heat release, see Eq. (*A*1)
u = velocity
Y = mass fraction

Greek

α = flame length correction coefficient, see Eq. (7)
ϕ = equivalence ratio
τ = characteristic time
ξ = flame surface variable, see discussion above Eq. (*A*10)

Subscripts and superscripts

o = oxidizer
f = fuel
$()'$ = fluctuating quantity
$\overline{()}$ = mean quantity

Appendix

This appendix presents a derivation of an expression for the characteristic time τ_{eq} that generalizes a previous analysis of Putnam [18]. It extends the definition of the convective time to account for the fact that the combustible mixture is consumed over a finite flame region (rather than at a single point). Denoting the flame's instantaneous rate of heat release per unit area by $q'(\mathbf{x}_f,t)$, the total rate of heat release within the flame $Q'(t)$ is given by

$$Q'(t)=\oint_{\substack{\text{Flame}\\ \text{Surface}}}\oint q'(\mathbf{x}_f,t)dx_f \tag{A1}$$

where $\mathbf{x}_f$ is the flame location. Note that in writing Eq. (*A*1), it is assumed that the heat release occurs along a flame surface. However, this development can be readily generalized to a region of distributed heat release as well.

Since this study is interested in heat release fluctuations induced by mixture composition oscillations, Eq. (*A*1) is written as:

$$Q'(t)=\oint_{\substack{\text{Flame}\\ \text{Surface}}}\oint K(\phi(\mathbf{x}_f),\mathbf{x}_f,t)\phi'(\mathbf{x}_f,t)d\mathbf{x}_f \tag{A2}$$

where K is a transfer function relating the instantaneous, local fluctuating rate of heat release to the fluctuating ϕ'. If the rate of heat release is only a function of the local ϕ', then K is a constant. However, the local heat release may be a function of the heat release at other parts of the flame due to flame dynamics effects. In this case, K is a function of space, time, and the instantaneous ϕ' at other parts of the flame. These cases will be referred to as ''locally'' and ''globally'' reacting, respectively.

Neglecting molecular transport and chemical reaction, the ϕ' of the mixture entering the flame at $\mathbf{x}_f$, can be related to ϕ' at the base of the flame at $\mathbf{x}_b$ (see Fig. 1) by the following species transport equations for the fuel and oxidizer mass fractions:

$$\frac{\partial Y_F}{\partial t}+\mathbf{u}\cdot\nabla Y_F=0 \tag{A3}$$

$$\frac{\partial Y_{\mathrm{Ox}}}{\partial t}+\mathbf{u}\cdot\nabla Y_{\mathrm{Ox}}=0 \tag{A4}$$

Defining a constant k_ϕ by the relation $k_\phi Y_f/Y_{\mathrm{ox}}=\phi$, multiplying Eq. (A3) by k_ϕ/Y_{ox}, Eq. (A4) by $-k_\phi Y_f/Y_{\mathrm{Ox}}^2$, and adding the resulting equations yields the following equation for the evolution of ϕ:

$$\frac{\partial \phi}{\partial t}+\mathbf{u}\cdot\nabla\phi=0 \tag{A5}$$

Equation (A5) shows that ϕ is constant along a pathline. Consequently, ϕ at the base of the flame $\mathbf{x}_b$ can be related to ϕ at the flame surface $\mathbf{x}_f$ by the expression:

$$\phi(\mathbf{x}_f(\mathbf{u},\mathbf{x}_b,t_b),t)=\phi(\mathbf{x}_b,t_b). \tag{A6}$$

The equation of the pathline relating $\mathbf{x}_f$ at time t to $\mathbf{x}_b$ at time t_b is:

$$\frac{d\mathbf{x}_l}{dt}=\mathbf{u}(\mathbf{x}_l,t) \tag{A7}$$

Generally, Eqs. (A6)–(A7) could be used to relate ϕ_f to ϕ_b using, for example, computed or measured descriptions of the mean flow fields and mixture composition. We will now simplify these equations in order to demonstrate their application, and to gain insight into the processes that affect τ_{eq}. The following assumptions are made: (1) the flowfield is uniform, one dimensional, and axisymmetric; (2) ϕ variations are one dimensional; and (3) the flame responds only to local ϕ disturbances (i.e., it is ''locally'' reacting). It should be noted that the third assumption essentially treats the flame surface as a continuous distribution of reactors that do not interact each other.

Using assumptions (1)–(3), Eq. (A2) can be simplified to:

$$Q'(t)=\int_{\mathrm{Flame}} K\phi'(x_f,t)(2\pi y_f)dy_f \tag{A8}$$

Considering only the fluctuating component of ϕ and using assumptions (1) and (2), Eqs. (A5) and (A6) can be written as

$$\phi'(x_f,t)=\phi'\left(t-t_b-\frac{x_f-x_b}{u}\right) \tag{A9}$$

Assuming harmonic oscillations (i.e., $\phi'(x,t)=\phi'(x)e^{-i\omega t}$), defining a flame Strouhal number, $\mathrm{St}=fL_{\mathrm{flame}}/u$, and defining an equation of the flame surface, $(x_f-x_b)/L_{\mathrm{flame}}=f(y_f/r)=f(\xi)$, Eqs. (A8) and (A9) can be combined to yield the following expression:

$$Q'(t)=2\pi r^2 K\phi'(x_b,t_b)e^{-i\omega(t-t_b)}\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi e^{2\pi i\mathrm{St}f(\xi)}d\xi \tag{A10}$$

Finally, defining the quantity $Q_o(t)$ as the fluctuating heat release that would have occurred if all the reactive mixture was burned at the base of the flame, Eq. (A10) can be written as:

$$Q'(t)/Q_O(t)=2\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi e^{2\pi i\mathrm{St}f(\xi)}d\xi=\kappa e^{i\omega\tau_{\mathrm{eq}}} \tag{A11}$$

where $\kappa=|Q'(t)/Q_o(t)|$ and τ_{eq} is an ''equivalent'' time delay that accounts for the fact that different fractions of the combustible mixture that enter the combustor are burned at different instances. It follows from Eq. (11) that:

$$\omega\tau_{\mathrm{eq}}=\tan^{-1}\left[\frac{\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi\sin(2\pi\mathrm{St}f(\xi))d\xi}{\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi\cos(2\pi\mathrm{St}f(\xi))d\xi}\right] \tag{A12}$$

$$\kappa=2\left[\left[\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi\sin(2\pi\mathrm{St}f(\xi))d\xi\right]^2+\left[\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi\cos(2\pi\mathrm{St}f(\xi))d\xi\right]^2\right]^{1/2} \tag{A13}$$

Equations (A12) and (A13) can be used to determine the quantities τ_{eq} and κ for a specified flame surface $f(\xi)$. For example, if all the fuel burns in a planar flame at $x=x_b$, then the surface is described by $(x_f-x_b)/L_{\mathrm{flame}}=f(\xi)=0$, yielding:

$$Q'(t)/Q_O=2\int_0^1\xi e^{2\pi i\mathrm{St}^*0}d\xi=2\int_0^1\xi d\xi=1\Rightarrow\kappa=1,\ \ \tau_{\mathrm{eq}}=0 \tag{A14}$$

as expected.

Finally Eq. (A11) suggests that the effect of having the heat released along the flame is equivalent to having the energy released at some ''effective'' flame location. Defining an ''equivalent'' nondimensional flame location $\alpha=L_{\mathrm{eq}}/L_{\mathrm{flame}}$, then the left side of Eq. (A12) can be rewritten as follows:

$$\omega\tau_{\mathrm{eq}}=2\pi f\frac{L_{\mathrm{eq}}}{u}=2\pi\frac{fL_{\mathrm{flame}}}{u}\frac{L_{\mathrm{eq}}}{L_{\mathrm{flame}}}=2\pi\mathrm{St}\alpha \tag{A15}$$

Combining Eqs. (A12) and (A15) yields:

$$\alpha(\mathrm{St},f(\xi))=\frac{1}{2\pi\mathrm{St}}\tan^{-1}\left[\frac{\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi\sin(2\pi\mathrm{St}f(\xi))d\xi}{\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi\cos(2\pi\mathrm{St}f(\xi))d\xi}\right] \tag{A16}$$

In the limit when $\mathrm{St}\ll 1$, Eq. (A16) simplifies to:

$$\alpha(f(\xi))=\frac{\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi f(\xi)d\xi}{\int_{\substack{\mathrm{Flame}\\ \mathrm{Surface}}}\xi d\xi} \tag{A17}$$

showing that in this case α is independent of the flame Strouhal number. The nondimensional effective convective time can be then defined as:

$$\frac{\tau_{\mathrm{conv,eff}}}{T}=\frac{\tau_{\mathrm{convect}}}{T}\left(1+\alpha\frac{L_{\mathrm{flame}}}{L_{\mathrm{convect}}}\right) \tag{A18}$$

Figure 6 presents results of calculations that describe the dependence of α upon St and $f(\xi)$. These results are discussed in Sec. 2.4.

Acknowledgment

This research was supported by AGTSR under Contract No. 95-01-SR031; Dr. Daniel B. Fant, Contract Monitor.

The authors would like to acknowledge their appreciation to Dr. George Richards (U.S. Dept. of Energy) and Prof. Robert Santoro (Penn. State) for making their data available to us.

References

[1] Cohen, J., and Anderson, T., AIAA paper #96-0819 (1996).
[2] Broda, J. C., Seo, S., Santoro, R. J., Shirhattikar, G., and Yang, V., 1998, *Twenty Seventh Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA.
[3] Straub, D. L., and Richards, G. A., 1998, ASME paper #98-GT-492.

[4] Paschereit, C., Gutmark, E., and Weisenstein, W., *Twenty Seventh Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA, (1998).
[5] Venkataraman, K. K., Simons, D. W., Preston, L. H., Lee, J., and Santavicca, D. A., 1997, *Proceedings of the Eastern States Meeting of the Combustion Institute*, pp. 313–316.
[6] Anderson, T., and Morford, S., 1996, ASME paper #98-GT-568.
[7] Janus, M. C., Richards, G. A., and Yip, M. J., ASME paper #97-GT-266.
[8] Janus, M. C., and Richards, G. A., *Proc. of the 1996 AFRC International Symposium*, Baltimore, MD, 1996.
[9] Peracchio, A. A., and Proscia, W. M., ASME paper #98-GT-269.
[10] Lieuwen, T., Neumeier, Y., and Zinn, B. T., 1998, Combust. Sci. Technol., **135**, pp. 1–6.
[11] Lieuwen, T., and Zinn, B. T., 1998, AIAA paper #98-0641.
[12] Lieuwen, T., and Zinn, B. T., 1998, *Twenty Seventh Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA.
[13] Keller, J. J., 1995, AIAA J., **33**, pp. 2280–2287.
[14] Mongia, R., Dibble, R., and Lovett, J., 1998, ASME Paper #98-GT-304.
[15] Sattinger, S. S., Neumeier, Y., Nabi, A., Zinn, B. T., Amos, D. J., and Darling, D. D., 1998, ASME paper #98-GT-258.
[16] Straub, D., Richards, G., and Yip, M. J., 1998, AIAA paper #98-3909.
[17] Torres, H., Lieuwen, T., Johnson, C., Daniel, B. R., and Zinn, B. T., 1999, AIAA paper #99-0712.
[18] Putnam, A. A., 1997, *Combustion Driven Oscillations in Industry*, Elsevier, New York.
[19] Seo, S., Broda, J. C., and Santoro, R. J., 1998, personal communication.
[20] Steele, R., *1998 AGTSR Combustion Workshop*, Berkeley, CA, Mar. 25–27.

Laminar Burning Velocity of Methane–Air–Diluent Mixtures

M. Elia

M. Ulinski

M. Metghalchi

Mechanical, Industrial and Manufacturing Engineering Department, Northeastern University, Boston, MA 02115

An experimental facility for measuring burning velocity has been designed and built. It consists of a spherical constant volume vessel equipped with a dynamic pressure transducer, ionization probes, thermocouple, and data acquisition system. The constant volume combustion vessel allows for the determination of the burning velocity over a wide range of temperatures and pressures from a single run. A new model has been developed to calculate the laminar burning velocity using the pressure data of the combustion process. The model solves conservation of mass and energy equations to determine the mass fraction of the burned gas as the combustion process proceeds. This new method allows for temperature gradients in the burned gas and the effects of flame stretch on burning velocity. Exact calculations of the burned gas properties are determined by using a chemical equilibrium code with gas properties from the JANAF Tables. Numerical differentiation of the mass fraction burned determines the rate of the mass fraction burned, from which the laminar burning velocity is calculated. Using this method, the laminar burning velocities of methane–air–diluent mixtures have been measured. A correlation has been developed for the range of pressures from 0.75 to 70 atm, unburned gas temperatures from 298 to 550 K, fuel/air equivalence ratios from 0.8 to 1.2, and diluent addition from 0 to 15 percent by volume. [DOI: 10.1115/1.1339984]

Introduction

The laminar burning velocity is one of the fundamental properties of fuel/air mixtures. It is used for modeling turbulent combustion and pollutant formation, and for furthering our understanding of laminar flame theories in general. Keck [1] uses the laminar burning velocity in his model for turbulent flame propagation and pollution formation in internal combustion engines. Karpov, Lipatnikov, and Zimont [2] use the laminar burning velocity as the only input parameter in their turbulent combustion model. Recent studies into the effects of flame stretch and wrinkling on premixed turbulent combustion rely heavily on the laminar burning velocity [3,4]. These studies are aids not only to the internal combustion engine field, but also to the power generation and hazardous waste disposal industries.

Methane was chosen as the fuel is this study for three main reasons. First, it is increasingly being used as a fuel in internal combustion engines and other power plants. Second, there is a wide body of literature on methane laminar burning velocity with which to compare the results of the new model. Finally, there is a need for more laminar burning velocity data at high diluent concentrations over a wide range of pressures, temperatures, and fuel–air equivalence ratios.

Methane is the main component (85–95 percent by volume) of compressed natural gas (CNG). CNG is increasingly being used as an alternative to gasoline in internal combustion engines and industrial power plants [5]. This use requires simple and workable correlations for modeling methane combustion over a wide range of temperatures and pressures. Good experimental data is needed for use in these models.

In modeling the combustion process, burning velocity values are needed over a large range of temperatures, pressures, fuel–air equivalence ratios, and diluent additions. In internal combustion engines the unburned gas temperature can be as high as 1000 K and the pressure range from 1 to 35 atm, with variations in fuel equivalence ratios from 0.6 to 1.4. Currently, burning velocity data for methane–air mixtures in the full range of pressures and temperatures of an internal combustion engine are not fully available. Al-himary and Karim [6] have presented data in the highest pressure range, with pressures from 1 to 70 atm, temperatures from 323 to 473 K, and fuel–air equivalence ratios, ϕ, from 0.62 to 1.35. Higher temperatures have been studied by Agnew and Graiff [7] with temperatures from 328 to 890 K, but at lower pressures, 1–3 atm, and fuel–air equivalence ratios from 0.5 to 1.0. Gottegenes and Mass [8] have theoretically calculated burning velocity in the range of pressures from 1 to 40 atm, unburned gas temperatures, 290–800 K, and for lean mixtures, $\phi < 1.0$.

To reduce NO_x emissions a portion of exhaust gas is recycled into the fuel mixture. This technique, known as exhaust gas recirculation (EGR), has the effect of reducing the flame temperature, a major factor in the NO_x creation. With the addition of proper diluents the burning velocity of fuel mixtures treated with exhaust gases can be measured. (In this study a mixture of 86 percent N_2 and 14 percent CO_2 is used. This simulates the specific heat and molar mass of the exhaust gases from a typical internal combustion engine.) Exhaust concentrations can be as high as 30 percent by volume, with more typical concentrations in the range of 15 percent. Laminar burning velocity data of fuel mixtures with diluents in the range of pressures and temperatures encountered in the internal combustion engines are scarce. Clark and Stone [9] have measured the laminar burning velocity for methane–air–diluent mixtures in the range of pressures from 0.5 to 10 atm and unburned gas temperature from 298 to 450 K. Pressures of 4–18 atm and unburned gas temperatures from 470 to 600 K have been studied by Ryan and Lestz [10].

The purpose of this study was to develop a new mathematical model to calculate the laminar burning velocity and broaden the range of existing experimental data using the pressure rise from the combustion process in a constant volume vessel. The measurements have been made in the range of pressures from 0.75 to 70 atm, unburned gas temperatures from 298 to 550 K, fuel air equivalence ratio from 0.8 to 1.2, and diluent addition from 0 to 15 percent by volume.

Experimental Apparatus

The experimental facility consists of a combustion vessel, charging apparatus, and a data acquisition system as shown in Fig. 1. The combustion vessel is constructed from two hemispheres

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, June 15, 2000; final revision received by the ASME Headquarters, June 23, 2000. Associate Editor: D. Assanis.

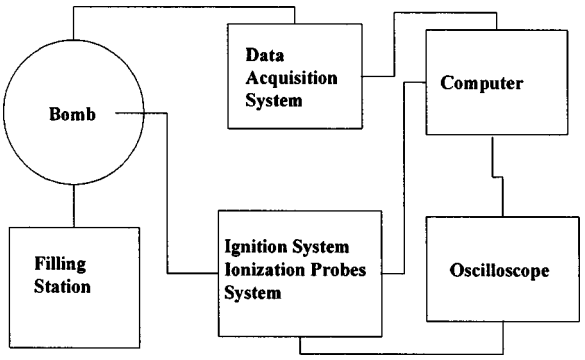


Fig. 1 Schematic of the experimental setup

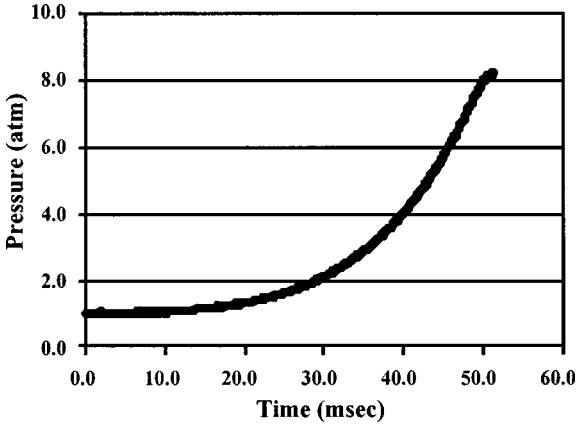


Fig. 3 Pressure as a function of time for the combustion of stoichiometric methane–air with P_i=1 atm and T_i=298 K

that are bolted together to provide a 15.24 cm (6 in) spherical vessel shown in Fig. 2. The vessel is made from 4140 alloy steel which can withstand up to 425 atm of pressure. The vessel is fitted with two center ignition electrodes, two ionization probes, a thermocouple, a pressure transducer, and a port for vacuuming and charging. Ignition is controlled by a capacitor discharge ignition system. The two ionization probes are located at the top and the bottom of the vessel and determine the arrival time of the flame front. The ionization probe's signals also provide information regarding the buoyancy effects on flame propagation. The thermocouple is used to measure the initial temperature of the mixture. A Kistler 603B1 piezoelectric quartz pressure transducer is used with a Kistler 5010B charge amplifier to measure the dynamic pressure in the vessel.

The charging apparatus consists of a fuel/air/diluent mixture and vacuum system. The fuel/air/diluent mixture is determined using the partial pressures method. Partial pressures are measured using Kulite silicon on silicon strain gauge transducers. Transducers in the ranges of 0–35, 0–175, and 0–1750 kPa (5, 25, 250 psia) were used. A vacuum pump capable of maintaining a vacuum of 100 μm in the combustion vessel is used to ensure that the partial pressures are measured correctly. The low pressure (vacuum) measurement is made using a thermocouple gauge.

The data acquisition system consists of a Data Translations 16-bit data acquisition card, which records the pressure change of the combustion event at a rate of 250 kHz. The analog-to-digital converter (AD) card receives the pressure signal from the charge amplifier and the signals from the ionization probes. All signals are recorded by a personal computer and an output data file is automatically generated. The output data file contains information about the partial pressures and initial temperature of all the gases used: fuel, air and diluents. It records the pressure-time data of the combustion process. An oscilloscope is also used to monitor the signals of the ionization probes, the pressure transducer and the ignition signal to insure that the various sensors are working and that the system has fired properly.

The experimental procedure consists of vacuuming the combustion vessel to a pressure of 100 μm. Once good vacuum has been achieved, the proper mixture of fuel/air/diluent is introduced, starting with the gas with the lowest partial pressure. After gas is charged into the vessel, the vessel is closed and the charging system is vacuumed down and the next gas is introduced into the vessel. When the vessel is filled with the proper mixture, the gas in the vessel is allowed to reach equilibrium by waiting several minutes before igniting the mixture. Upon ignition, a data file is written to the computer and the pressure signal and ionization probes signals are captured by the oscilloscope. Figure 3 shows the result of a typical experiment

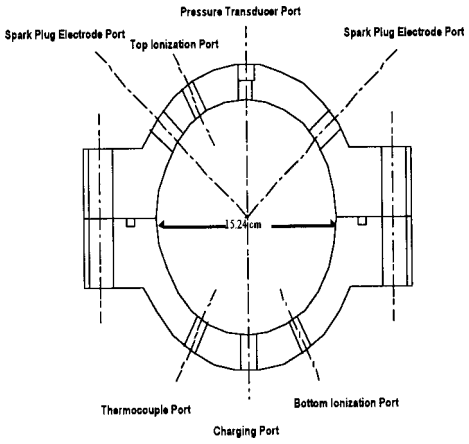


Fig. 2 Schematic of the combustion vessel

Computational Model

The new model to calculate burning velocity from the pressure measurement of the combustion process in a constant volume combustion vessel is based on the one developed by Metghalchi and Keck [11,12]. In this model, gases in the vessel are composed of two regions. One part is the burned gas in the center of the vessel. The other part is the unburned gas, which is consumed as the burned gas expands to it. The burned and unburned gases are assumed to be ideal. It is also assumed that flame thickness is negligible for a flame radius greater than 2 or 3 cm. The flame front is assumed to be smooth and spherical. The unburned gas consists of two regions: the unburned gas core, which is frozen in its chemical composition and compressed isentropically by the burned gas, and the gases within the thermal boundary layer at the vessel wall, where temperature decreases to the wall temperature. It is assumed that there is a temperature gradient within the burned gases. This gradient is caused by the compression process due to the combustion of the unburned gas, which results in a pressure increase of the gases that have previously burned. Burned gases are assumed to be in different shells, where temperature is uniform in each shell, but it varies from shell to shell. Gases in each shell are assumed to be in chemical equilibrium states. The burned gas states have been calculated using 18 species of methane combustion. Thermodynamic properties of all gases have been calculated using JANAF Tables [13]. The laminar burning velocity is determined by [14]

$$S_u = m v_u \dot{x} / A_f \tag{1}$$

where

S_u = laminar burning velocity,
m = mass of the gas mixture in the vessel,
v_u = specific volume of the unburned gas,
$\dot{x}$ = rate of mass fraction of the gas burned, and
A_f = the area of the flame $=4\pi r_f^2$, with r_f the radius of the flame front.

The mass of the gas mixture can be determined by direct calculation using the initial condition of the mixture. The specific volume of the unburned gas is determined using the pressure data from the combustion process and JANAF Tables. The properties of the burned gases are also determined using the JANAF Tables and equilibrium code. These properties are used to determine the mass fraction of gas burned, from which the rate of mass fraction burned is calculated. The mass fraction burned is calculated using the conservation of mass and energy equations. The radius of the flame front, used to determine the flame front area, is determined from the mass fraction burned calculations.

The unburned gas in the vessel is assumed to be compressed isentropically during the combustion process. Properties of the unburned gas are calculated using JANAF Tables. The temperature of the unburned gas is determined from the pressure history using variable specific heat assumptions. The specific energy and volume of the unburned gases are then calculated from the unburned gas temperature.

Internal energy and specific volume of the burned gases are calculated knowing the temperature of the gas in each shell. In this model, it is assumed that the burned gas consists of a spherical region of shells generating from the center of the vessel. At the outmost edge of this region is the shell which has just burned. This shell surrounds other shells of burned gases, which burned at earlier time steps, the innermost shell being the one which burned first. The compression of the gas from the combustion process causes the temperature of the gas in the innermost shell to be higher than the temperature of the gas in the surrounding shells. Gases in each shell are assumed to have uniform properties and to be in chemical equilibrium. By using this shell method, the burned gas temperature gradients can be accounted for, where previously a lumped method was used and the burned gas was assumed to be at an average temperature.

Burned gas properties are calculated using equilibrium code. Eighteen species have been used in the equilibrium code for methane combustion. These are H, O, OH, H_2, O_2, H_2O, HO_2, H_2O_2, CH_4, CH_3, CH_2O, CH_3OH, CO, CO_2, N_2, N, NO, and N_2O. The code, given temperature and pressure data, calculates all the thermodynamic properties of the mixture at equilibrium conditions.

The burned and unburned gas properties are used to determine the mass fraction of the burned gas. Two nonlinear equations, derived from the conservation of mass and energy equations, are solved to yield the mass fraction:

$$\frac{V}{m}+\frac{A\delta}{m}=\int_0^x v_b\,dx'+\int_x^1 v_u\,dx' \tag{2}$$

$$\frac{E}{m}-\frac{Q}{m}=\int_0^x e_b\,dx'+\int_x^1 e_u\,dx' \tag{3}$$

where

A = combustion vessel wall area,
e = specific internal energy,
E = total initial energy of gas in the vessel,
Q = total energy transfer from the boundary layer displacement thickness to the vessel wall,
x = mass fraction burned,
x' = integration variable,
v = specific volume,
V = combustion vessel volume,
δ = boundary layer displacement thickness,

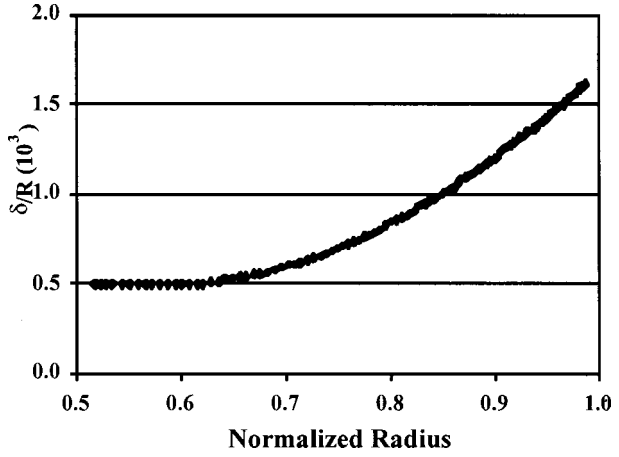


Fig. 4 Normalized displacement thickness (δ/R) as a function of normalized flame radius

and the subscripts b and u refer the to the burned and unburned gas, respectively.

The boundary layer displacement thickness, δ, is defined as

$$\delta=(1/\rho_\infty)\int_0^\infty(\rho-\rho_\infty)\,dr \tag{4}$$

where

ρ = density of unburned gas within the boundary layer displacement thickness and
ρ_∞ = density of that portion of unburned gas that is compressed isentropically.

Energy transfer at the boundary layer of the vessel wall causes the displacement thickness temperature to be lower than the rest of the unburned gas. Therefore, the displacement thickness gas density is higher. With the assumption that the displacement thickness has the same properties as the core gas, this density difference is accounted for by assuming volume of the vessel is increased by $A\delta$.

Boundary layer displacement thickness is determined from the pressure data by [12]

$$\delta(t)=\left(\frac{\mu_i}{\pi\rho_i}\right)^{1/2}\left(\frac{\rho_i}{\rho}\right)^{1/\gamma_u}\int_0^t\left[\left(\left(\frac{p'}{p_i}\right)-\left(\frac{p'}{p_i}\right)^{1/\gamma_u}\right)\times\left(\int_{t'}^t\frac{p''}{p_i}\,dt''\right)^{1/2}\right]dt' \tag{5}$$

where

ρ_i = the initial density of the gas mixture,
μ_i = the viscosity of the unburned gas at the initial temperature and pressure,
γ_u = the specific heat ratio of the unburned gas,
p_i = the initial pressure of the gas mixture,
p',p'' = instaneous pressure,
t = time, and
t',t'' = integration variables.

Figure 4 shows the variation of normalized displacement thickness as a function of normalized flame radius for a stoichiometric methane–air mixture, $P_i=1$ atm and $T_i=298$ K. The displacement thickness can be seen to be very small relative to the vessel radius.

Total energy transfer to the bomb wall is calculated as [12]

$$Q=A\int_0^\delta p\,d\,\delta' \tag{6}$$

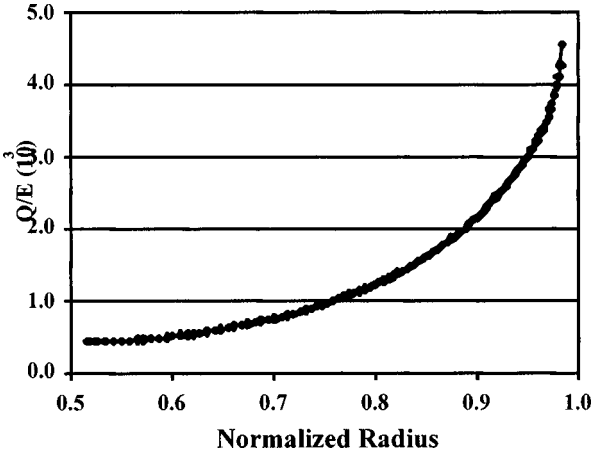


Fig. 5 Normalized heat transfer (Q/E) as a function of normalized flame radius

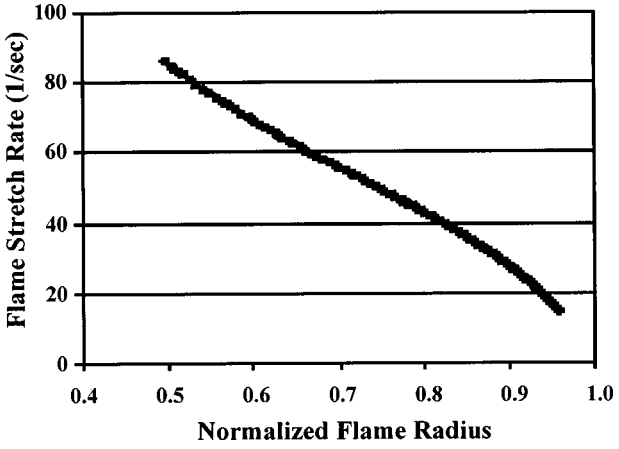


Fig. 6 Flame stretch rate as a function of normalized flame radius for the combustion of stoichiometric methane–air with $P_i = 1$ atm and $T_i = 298$ K

where δ' is the integration variable. Figure 5 shows the normalized heat transfer (Q/E) as a function of normalized radius. It also shows that the heat transfer to the vessel wall is less than 1 percent of total energy of gases in the bomb.

The conservation equations are solved for the two unknowns: the burned gas temperature of the outermost shell and the mass fraction burned at each time step. The two equations are solved by using the two-dimensional Newton–Raphson method. For this method, incorporating temperature gradient shells, the two equations are rewritten as

$$\frac{V}{m} + \frac{A\delta}{m} = \sum_{i=0}^{n-1} v_{bi}x_i + v_{bn}x_n + \left(1 - \sum_{i=0}^{n-1} x_i - x_n\right) v_u \tag{7}$$

$$\frac{E}{m} - \frac{Q}{m} = \sum_{i=0}^{n-1} e_{bi}x_i + e_{bn}x_n + \left(1 - \sum_{i=0}^{n-1} x_i - x_n\right) e_u \tag{8}$$

where x_n, v_{bn}, and e_{bn} are the mass fraction, specific volume, and specific internal energy of the outermost shell in the burned gas region, and x_i, v_{bi}, and e_{bi}, the mass fraction, specific volume, and the specific internal energy in the shells of the previously burned gas.

Equations (7) and (8) are solved using the Newton–Raphson method for the properties of the nth or outermost shell, where e_{bn} and v_{bn} are functions of the temperature of the nth shell. The nth shell is the last shell that has been burned up to the time step of consideration. In the first time step, $n=1$, the summations equal 0 and the two equations are solved for x_1, and T_{b1}. In the next time step, $n=2$, e_{b1} and v_{b1} are reevaluated to account for the compression due to the burning of the outer shell (shell 2) and Eqs. (7) and (8) are solved for x_2 and T_{b2}. This procedure is followed until the last time step, the time of maximum pressure in the bomb.

The mass fraction burned is then calculated as a function of time. The rate of the mass fraction burned, $\dot{x}$, is determined using the mass fraction calculation. It is determined by fitting a parabola to three consecutive mass fraction points and then differentiating to find the slope of the midpoint. Each point is used only once to minimize amplification of any possible errors due to the differentiation process.

The flame radius, r_f, is calculated as a function of time from the mass fraction burned calculations and the burned gas properties:

$$V_b = \sum v_{b_i} m_{b_i} = \frac{4}{3}\pi r_f^3 \tag{9}$$

An important consideration in the calculation of the laminar burning velocity is the flame stretch rate. In a constant volume combustion vessel, with a spherical outwardly propagating combustion wave, flame stretch affects the calculation of the burning velocity. The effect of stretch is highest when the flame radius is small and it diminishes as the flame radius increases. Bradley [15] and Aung and Faeth [16] have investigated the effect of the stretch on flame speed. Using Bradley's method, the burning velocity of upstretched flame is calculated using

$$S_L = S_u + C\alpha \tag{10}$$

where

S_L = the unstretched burning velocity,
S_u = the stretched burning velocity,

and C is related to the Markstein length, the values which have been taken from Bradley's work [15]. α is the flame stretch rate defined as

$$\alpha = \frac{1}{A_f}\frac{dA_f}{dt} \tag{11}$$

In a spherical combustion vessel, the stretch rate can be calculated through the determination of the flame radius:

$$\alpha = 2\frac{\dot{r}_f}{r_f} \tag{12}$$

This is shown in Fig. 6 for a stoichiometric methane–air mixture with $P_i = 1$ atm and $T_i = 298$ K. In a spherical vessel of the size used, a 10 percent increase in initial pressure increases the flame radius by about 50 percent of the total radius (flame radius of 4 cm), thus reducing the effect of stretch on much of the data collected.

The unstretched burning velocity has been determined using Eq. (10). The correction was about 1.0 cm/s when the flame radius was small and it decreased to 0.2 cm/s when the flame was close to the vessel wall.

Results and Discussion

The laminar burning velocity of methane–air–diluent mixtures in the range of pressures from 0.75 to 70 atm, unburned gas temperatures from 298 to 550 K, fuel air equivalence ratio from 0.8 to 1.2, and diluent addition from 0 to 15 percent by volume have been determined using the method described. The burning velocities have been fitted to the following relation:

Table 1 Parameters for burning velocity of methane–air–diluent mixtures, $\phi=0.8$ to 1.2, $P=0.75$ to 70 atm, $T=298$ to 550 K, and $D=0$ to 15 percent

a_0	a_1	a_2
−5.883	14.003	−7.115
b_1	b_2	b_3
4.829	−7.778	0.003
α	β	S_{uo}
1.857	−0.435	37.5

$$S_L = S_{uo}(a_o + a_1\phi + a_2\phi^2)[1-(b_1 D + b_2 D^2 + b_3 D^3)] \times \left(\frac{T_u}{T_0}\right)^{\alpha}\left(\frac{P}{P_0}\right)^{\beta} \quad (13)$$

where

S_L = the laminar burning velocity (cm/s),
S_{uo} = the laminar burning velocity at the reference point (1 atm pressure and 298 K) (cm/s),
D = the percentage (by volume) of diluent in the mixture,
T_u = the temperature of the unburned mixture (K),
T_0 = the reference temperature = 298 K,
P = the pressure of the mixture (atm), and
P_0 = the reference pressure = 1 atm.

The parameters, S_{uo}, a_0, a_1, a_2, b_1, b_2, b_3, α, and β are listed in Table 1.

The measured values of laminar burning velocities of methane–air mixtures from this study have been compared to the results of other investigators. Figure 7 shows the burning velocity of methane–air mixtures at a pressure and temperature of 1 atm and 298 K as a function of fuel–air equivalence ratio. The results are compared to those of Clark and Stone [9], Iijima and Takeno [17], Egolfopoulos and Law [18], and Yamaoka and Tsuji [19]. Clark and Stone and Iijima and Takeno have used a spherical constant volume chamber using burning velocity models similar to Metghalchi and Keck's, but without considering temperature gradients or flame stretch effects. Egolfopoulos and Law and Yamaoka and Tsuji have used the experimental results of counterflow flames. Egolfopoulos and Law extrapolate the flame speed measurement to include stretch effects. This comparison shows that, with the exception of the results measured by Yamaoka and Tsuji, all other measurements are very close to each other. The maximum laminar burning velocity of methane–air mixtures at atmospheric pressure and temperature is shown to be at a fuel air equivalence ratio of 1.05–1.10.

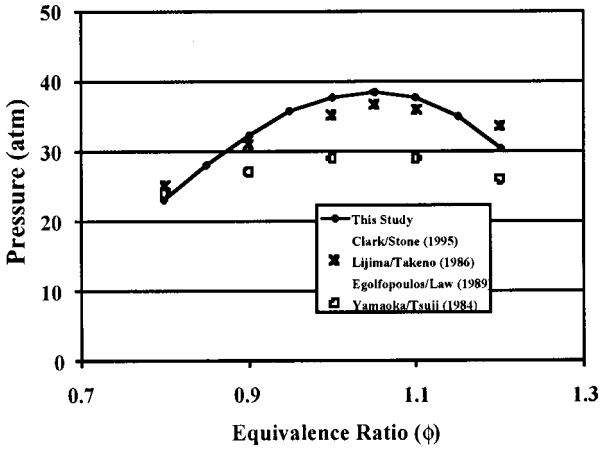


Fig. 7 Laminar burning velocity as a function of equivalence ratio for methane–air mixtures at 1 atm pressure and 298 K

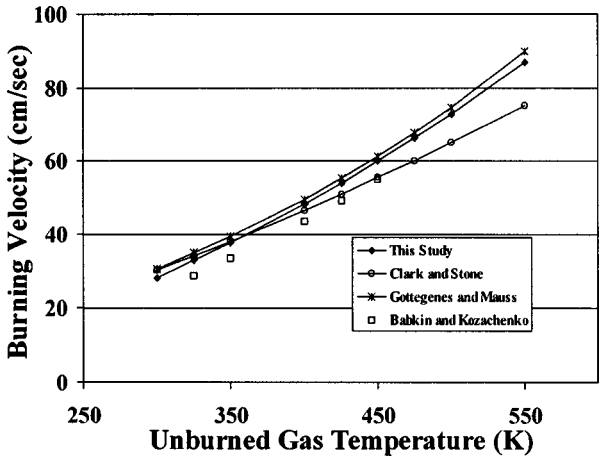


Fig. 8 Comparison of laminar burning velocity as a function of temperature for stoichiometric methane–air at 2 atm pressure

Figure 8 shows laminar burning velocity for stoichiometric methane–air mixtures at a pressure of two atmospheres as a function of temperature. The results of this study are compared with those calculated by Clark and Stone [9], Gottegenes and Mauss [8], and Babkin and Kozachenko [20]. The results show similar trends for burning velocity, with some variation in slope.

Figure 9 shows the laminar burning velocity as a function of temperature at a pressure of 15 atm. The results are compared to those of Clark and Stone [9], Gottegenes and Mauss [8], Babkin and Kozachenko [19], and Ryan and Lestz [10]. It can be seen that the results are similar, except Ryan and Lestz predict a steeper temperature dependency than the others.

The laminar burning velocity of a lean mixture ($\phi=0.83$) of methane–air determined in this study is compared to other studies, shown in Fig. 10.

Figure 11 shows the pressure dependency of laminar burning velocity of a stoichiometric methane–air mixture. The results are compared to other studies, all showing the negative dependency, with good agreement to each other.

Figures 12 and 13 show laminar burning velocity of stoichiometric methane–air–diluent mixtures for various diluent concentrations, both as a function of pressure (Fig. 12) and temperature

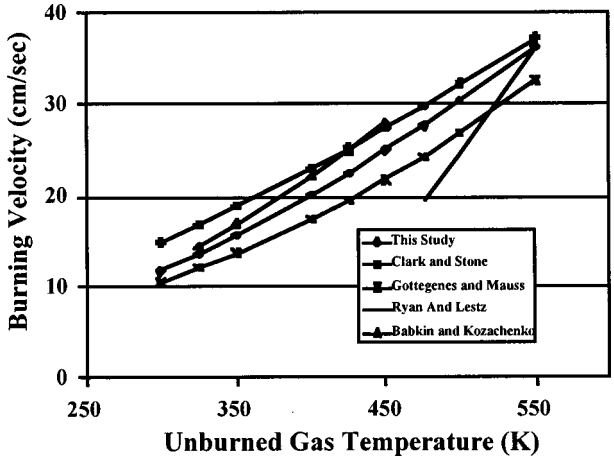


Fig. 9 Comparison of laminar burning velocity as a function of temperature for stoichiometric methane–air mixture at 15 atm pressure

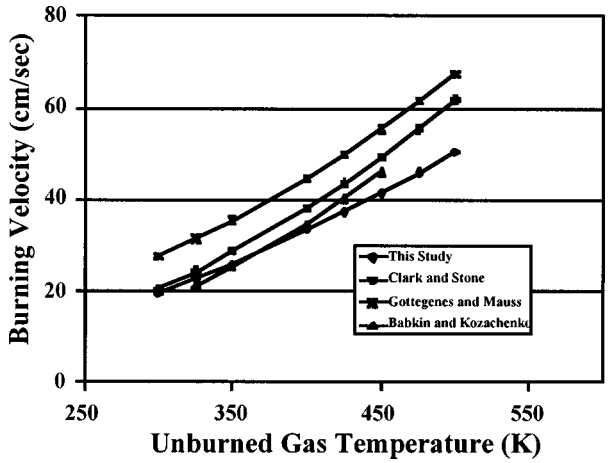


Fig. 10 Comparison of laminar burning velocity as a function of temperature for methane–air at 2 atm pressure and $\phi = 0.83$

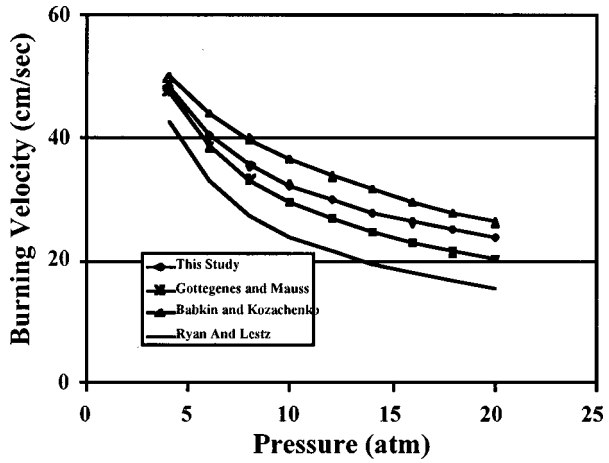


Fig. 11 Comparison of laminar burning velocity as a function of pressure for stoichiometric methane–air at 470 K

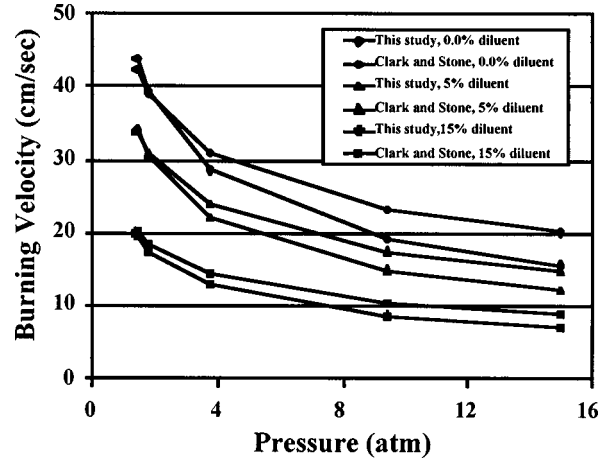


Fig. 12 Comparison of laminar burning velocity of methane–air–diluent mixtures as a function of pressure at 350 K and $\phi = 1.0$

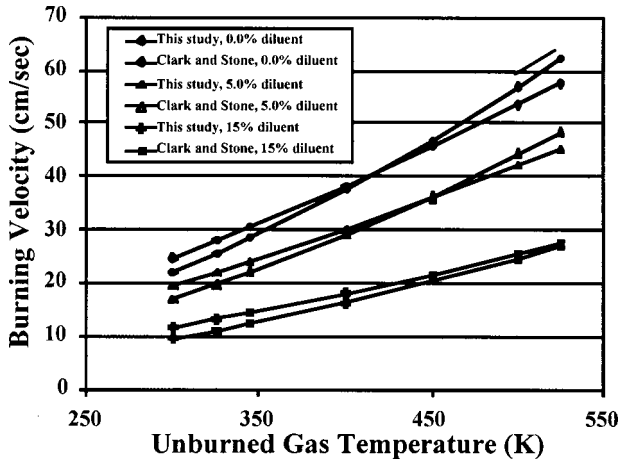


Fig. 13 Comparison of laminar burning velocity of methane–air–diluent as a function of unburned gas temperature at 3.5 atm pressure and $\phi = 1.0$

(Fig. 13). The results are compared to those reported by Clark and Stone [9]. Burning velocity is shown decreasing as diluent concentrations increase, due to the lower flame temperature. The high diluent concentrations simulate those found during idle conditions of today's internal combustion engines.

Conclusions

A new experimental facility has been built to determine the laminar burning velocity of fuels. Using a spherical constant volume combustion vessel, pressure data has been collected over a wide range of pressures, unburned gas temperatures, and mixtures. A new model has been developed. The new model incorporates temperature gradients in the burned gas, as well as the use of equilibrium calculations in the determination of the mass fraction of gas burned, from which burning velocity is determined. Flame stretch effects are also incorporated. The results of this study show good agreement with other studies in the range of equivalence ratios, 0.8 to 1.2, pressures, 1 to 70 atm, unburned gas temperatures, 298 to 550 K, and diluent concentrations up to 15 percent.

Acknowledgments

The authors would like to acknowledge Ford Motor Company for their generous financial support. Particular thanks go to Dr. Thomas Kenney of Ford Motor Company and Dr. James Keck of MIT for their encouragement and advice.

References

[1] Keck, J. C., 1982, ''Turbulent Flame Structure and Speed in Spark-Ignition Engines,'' *Nineteenth International Symposium on Combustion*, pp. 1451–1466.

[2] Karpov, V., Lipatnikov, A., and Zimont, V., 1996, ''Test of an Engineering Model of Premixed Turbulent Combustion,'' *26th International Symposium on Combustion*, pp. 249–257.

[3] Menon, S., and Kerstein, A., 1992, ''Stochastic Simulation Of The Structure And Propagation Rate Of Turbulent Premixed Flames,'' *24th International Symposium on Combustion*, pp. 443–450.

[4] Shy, S. S., Ronney, P. D., Buckley, S. G., and Yakhot, V., 1992, ''Experimental Simulation Of Premixed Turbulent Combustion Using Aqueous Autocatalytic Reactions,'' *24th International Symposium On Combustion*, pp. 543–551.

[5] Whalen, P., Kelly, K., Motta, R., and Broderick, J., 1996, ''Summary of Results From the National Renewable Energy Laboratory's Vehicle Evaluation Data Collection Efforts,'' NREL/SP-425-20821, Golden, CO, prepared by NREL for DOE.

[6] Al-Himary, T. J., and Karim, G. A., 1987, ''A Correlation for Burning Velocity of Methane-Air Mixtures at high Pressures and Temperatures,'' ASME J. Eng. Gas Turbines Power, **109**, pp. 439–442.

[7] Agnew, J. T., and Graiff, L. B., 1961, ''The Pressure Dependence of Laminar

Burning Velocity by the Spherical Bomb Method,'' Combust. Flame, **5**, pp. 209–219.
[8] Gottegenes, J., and Mass, F., 1992, ''Analytic Approximation of Burning Velocities and Flame Thickness of Lean Hydrogen, Methane, Ethylene, Ethane, Acetylene, and Propane Flames,'' *Twenty-fourth Symposium (International) on Combustion*, pp. 129–135.
[9] Clarke, A., Stone, R., and Beckwith, P., 1995, ''The Measurement of Laminar Burning Velocity of Methane/Air/Diluent Mixtures in a Constant Volume Combustion Bomb in a Micro-Gravity Environment,'' J. Inst. Energy, **68**, pp. 130–136.
[10] Ryan, T. W., and Lestz, S. S., 1980, ''The Laminar Burning Velocity of Isooctane, N-Heptane, Methanol, Methane, and Propane at Elevated Temperature and Pressures in the Presence of a Diluent,'' SAE Technical Paper Series, 800103.
[11] Metghalchi, M., and Keck, J., 1980, ''Burning Velocity Of Propane-Air Mixtures At High Temperature And Pressure,'' Combust. Flame, **38**, pp. 143–154.
[12] Metghalchi, M., and Keck, J., 1982, ''Burning Velocities of Mixtures of Air with Methanol, Isooctane, and Indolene at High Pressure and Temperature,'' Combust. Flame, **48**, pp. 191–210.
[13] *JANAF Thermochemical Tables*, Third Edition, 1986, American Chemical Institute, American Institute of Physics, National Bureau of Standards.
[14] Kuo, K., 1986, *Principles of Combustion*, John Wiley and Sons, New York.
[15] Bradley, D., Gaskel, P. H., and Gu, X. J., 1996, ''Burning Velocities, Markstein Lengths, And Flame Quenching For Spherical Methane-Air Flames: A Computational Study,'' Combust. Flame, **104**, pp. 176–198.
[16] Aung, K. T., Hassan, M. T., and Faeth, G. M., 1997, ''Flame Stretch Interactions of Laminar Premixed Hydrogen/Air Flames At Normal Temperature And Pressure,'' Combust. Flame, **109**, pp. 361–384.
[17] Iijima, T., and Takeno, T., 1986, ''Effects Of Temperature And Pressure On Burning Velocity,'' Combust. Flame, **65**, pp. 35–43.
[18] Egolfopoulos, F. N., Cho, P., and Law, C. K., 1989, ''Laminar Flame Speeds of Methane-Air Mixtures Under Reduced and Elevated Pressures,'' Combust. Flame, **76**, pp. 375–391.
[19] Yamaoka, I., and Tsuji, H., 1984, ''Determination Of Burning Velocity Using Counterflow Flames,'' *20th Symposium (International) on Combustion*, pp. 883–1892.
[20] Babkin, V., and Kozacheenko, L., 1966, ''Study of Burning Velocity in Methane-Air Mixtures At High Pressure (English Translation),'' Vzyuva, **3**, pp. 77–856.

Analysis and Modeling of the Fluid-Dynamic Effects in Branched Exhaust Junctions of ICE

F. Payri

E. Reyes

J. Galindo

CMT—Departamento de Máquinas
y Motores Térmicos,
Universidad Politécnica de Valencia,
P.B. 22012,
E-46071 Valencia, Spain
e-mail: cmt@mot.upv.es

The influence of exhaust junction geometry on flow-dynamics of exhaust gas is analyzed. The authors propose an experimental characterization method based on the measurement of the instantaneous pressure in the junction operating with engine exhaust flow and solving the problems posed for the accurate instantaneous pressure measurements under the adverse temperature condition. In this paper, the method is applied to two ''Y'' type junctions, with a reed being the unique difference between them, to determine the influence of this element on the junction behavior. The analysis of the experimental results denotes two major differences: the characteristics of the wave reflected at the junction, and the energy of the pulse transmitted to the lateral branch of the junction. The results of the analysis are introduced in the junction modeling used in a one-dimensional gas dynamic model with an important improvement in the agreement of the modeled predictions with the experimental measurements. [DOI: 10.1115/1.1339988]

Introduction

The dynamics in the exhaust system has a significant influence on several aspects of engine performance. Short circuit and back flow during the valve overlap are determined by its geometry, influencing the engine volumetric efficiency and residual mass fraction. In fact, that phenomenon is of utmost importance in the scavenging of two stroke engines. Pumping work is also greatly influenced by these dynamics and therefore by the geometry of the exhaust manifold and pipes. Additionally, in multicylinder engines the exhaust pulses of a given cylinder interfere with the gas exchange process and pumping work of the previous or next firing cylinders. There are several studies that analyze in depth all these dynamic effects [1–6].

In turbocharged engines the objectives at the design stage of the exhaust system are more complex. There is an additional goal in obtaining the maximum amount of mechanical energy with the turbine from the available exhaust gas energy. Several strategies, depending of the size of the engine, have been developed to provide not only the maximum amount of energy to the turbine, but also in the best way to optimize its efficiency. This has been studied for a long time [7–10].

The behavior of the junction also has an important influence on the dynamic effects described above. The flow in exhaust junctions is very complex and junction geometry determines the energy losses produced and how the incident exhaust pulses are transmitted to the exit duct, reflected to the exhaust port and derived to the lateral branch or branches.

There are important studies proposing mathematical models of the flow dynamics at the junctions [11–16]. But the references in the literature about the influence of the junction geometry on its dynamic behavior are limited. One traditional difficulty is to obtain accurate experimental data with actual exhaust flow conditions. The strong pressure pulses with high temperature make it extremely difficult to measure instantaneous pressure or temperature at the boundaries of the junction.

Three-dimensional (3D) modeling of junctions is still a time consuming task and the results are dependant on computing assumptions, such as the turbulence model and the size of the mesh due to the presence of eddies within the junction [17,18].

The objective of this work is to report results of flow analysis in exhaust junctions based on experimental measurements. They are obtained with a characterization method, developed by this research group, that accurately measures the instantaneous pressure in several critical sections of the junction with actual exhaust flow conditions.

This analysis identifies two important flow patterns in the junction and derives meaningful conclusions about the influence of the junction geometry on its behavior.

The method has been applied to ''Y'' type of junctions with different dimensions and configurations, although the discussion in this paper, for sake of brevity, is limited to two junctions with simple geometry.

To complete the work, conclusions are introduced in a mathematical model of the junction to consider these dynamic effects. The result is a significant improvement in the agreement between the predicted results and the experimental measurements.

The method has also been applied to 4-1 junctions. In this case, two cylinders are in the exhaust process at the same time and therefore additional aspects such as their relative position are relevant for the dynamic effects [19]. The analysis of the flow dynamics in this type of junction will be the subject of a future publication.

Experimental Process

The experimental characterization of the ''Y'' junction has been performed by connecting one of their inlet branches (active branch) to the exhaust port of an engine as described in Figs. 1 and 2.

In the real engine conditions there are many wave superpositions. The proposed configuration allows the measurement of the response of the junction to exhaust pulses avoiding interference with the pulses from other cylinders of the engine. The lateral branch has been closed since this is the normal condition in actual engine operation when the junction is connecting two contiguous cylinders. The length of this branch includes that of the engine port. The exit branch has a significant length to ensure that the reflected pulse, produced in the discharge section to a stagnation

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, Apr. 27, 2000; final revision received by the ASME Headquarters, Nov. 3, 2000. Associate Editor: D. Assanis.

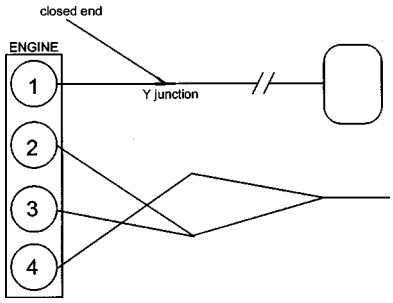


Fig. 1 Experimental setup

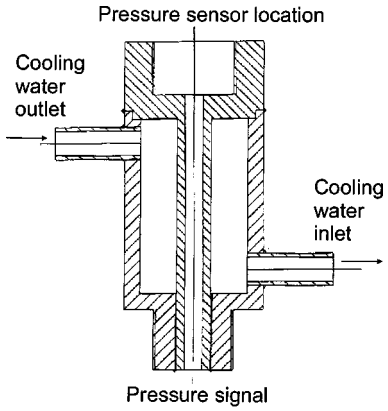


Fig. 3 Pressure transducer cooler

tank, arrives at the junction when the exhaust process in the cylinder has finished, thus avoiding undesired wave superposition.

The characterization is based on the measurement of the instantaneous pressure at least in one section of every branch of the junction. This section has to be at an adequate distance from the junction to ensure a fully developed one-dimensional flow. The instantaneous pressure has also been measured in two additional sections. One is at the port outlet to better characterize the wave reflected back to the active branch. The second is in the closed end of the lateral branch to characterize the directionality of the junction by the energy transmitted to this branch. To allow the decomposition of the incident and reflected waves, the lateral branch has been thermally isolated so that the process is close to adiabatic. The instrumentation of the junction is detailed in Fig. 2.

Measurement of instantaneous pressure with hot gas flow has been traditionally performed with piezoelectric transducers. The measured signal is then referenced with the mean pressure obtained from another pressure measurement device (liquid column or magnetic sensor). Usually this device is linked to the exhaust duct with a long pipe to avoid contact of the hot exhaust gas with the sensing surface.

The pressure losses in the junction might be smaller than the error derived from the determination of the mean pressure with that auxiliary device. In fact, it is obvious that in the case of pulsating flow the volume between the duct and the measuring device transfers the pulses to the sensing surface in such a way that the measured value is not necessarily equal to the mean pressure in the duct. The deviation function of the pipes and measurement device geometry and the flow characteristics (such as frequency and amplitude of the pressure pulses) is not easy to predict.

To avoid this inconvenience, the experimental measurements have been performed with double cooled piezoresistive transducers. The thermal shock in the sensing membrane, and its corresponding errors, have been avoided with a small cooled volume (length 37 mm and diameter 1.5 mm) introduced between the duct surface and the membrane of the sensor. This volume, strongly cooled, prevents direct contact of the membrane with the extremely hot exhaust gas and therefore eliminates this source of measurement error. Obviously this volume introduces a ripple in the measured signal, but its mean frequency is much higher than those of the exhaust pulses. Additionally, this ripple can be removed by the decomposition of the measured pressure in the sensing surface to the duct surface by the application of the transfer function of the cooler volume to the measured signal, in order to transport the signal from the cooling volume to the real measurement section. This transfer function is obtained with specific experiments and the corresponding mathematical analysis. The transducer is additionally equipped with a second cooler to protect the sensing device.

This pressure measurement process takes profit from the main advantage of the piezoresistive sensors (provides instantaneous and absolute pressure) and avoids its inconveniences (thermal sensitivity). In Fig. 3 a draft of the sensor cooler is shown.

Measurement of the temperature is not critical with this approach. Once the behavior of the junction is modeled, and the predicted pressure is in good agreement with the accurate pressure measurements, the temperature predictions of the model must also be accurate. An inaccurate temperature in the model will induce an incorrect wave propagation speed and the predicted wave dynamics will not be in phase with the experimental values.

Analysis of the Flow in a ''Y'' Junction

The experimental process has been applied to two simple ''Y'' junctions. Their geometry is shown in Fig. 4. The only difference between them is the presence or the absence of a small reed in the area connecting the two inlet branches. The reed induces a significant change in the contact surface between the lateral branches and in the way that the pressure waves discharge into the junction. This will translate into an important difference in the dynamic effects produced in the junction.

The diameter of the inlet branches of the junction is equal to the valve port to avoid sudden area changes. The area ratio between one inlet duct and the outlet duct is 1.25. Figure 5*a* represents a

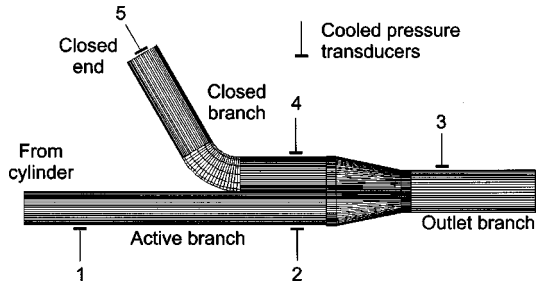


Fig. 2 Exhaust junction instrumentation

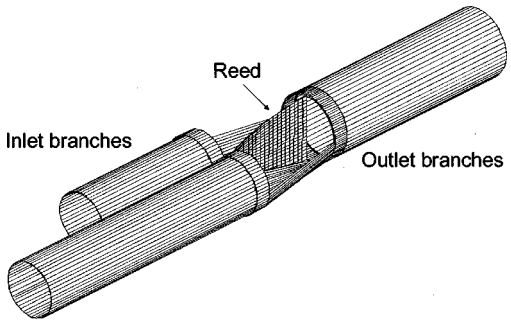


Fig. 4 Detail of the junction geometry

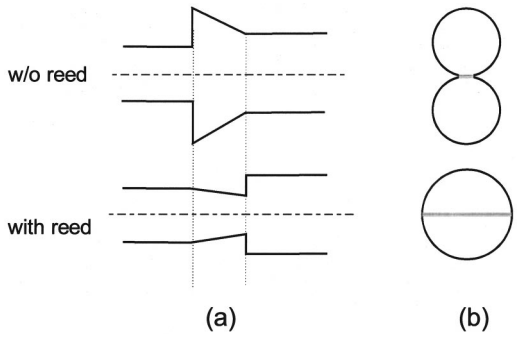


Fig. 5 Cross-section diagram of the junctions (*a*) and interface between active and lateral branch (*b*)

simple draft of the transversal areas along the junction for both junctions. Flow in the junction without the reed expands first to a section double that of the inlet duct. From this maximum the section is reduced to the exit duct area (1.25 times the entry section). In the junction with the reed the section is reduced first to the half of the exit section (0.625 times the entry area) and then expanded to a double area.

The contact surface between the lateral branches is very different in the junctions. As represented in Fig. 5*b*, it is a single point in the case of the junction without the reed and a complete diameter of the exit duct in the case with reed.

The reed has been widely used in the junctions of exhaust systems of heavy-duty turbocharged diesel engines with the aim of increasing the energy available in the exhaust gas at the turbine inlet and of reducing the interference among cylinders [7,10].

The engine used in this study was a 2 l. four in line automotive spark ignition engine. The operating condition were from 2500 up to 5500 rpm every 1000 rpm and full load. These conditions provide very strong exhaust pulses, in mass flow and pressure amplitude, ensuring the clear presence of the main dynamic effects in the junction as it will be seen below.

The experimental pressure measurements of both junctions at 4500 rpm and full load are plotted in Fig. 6.

In the junction without the reed (solid line), there is a big difference in the pressure profile between measurement locations 1 and 2. This is due to the different superposition of the strong overpressure pulse from the cylinder with a rarefaction pulse traveling upstream originated by the expansion of the first one in the junction volume. The exhaust pulse travels downstream through the exit branch (location 3) and upstream to the lateral branch (location 4) doubling the pressure amplitude in its closed end (location 5).

It can be observed that the pulse transmitted to the lateral branch is weaker than the one in the exit branch. This is due to the production of an eddy induced by the discharge flow velocity of the active branch that partially blocks the entrance of the flow into the lateral branch. Figure 7 represents the steady flow velocity field, obtained with a 3D model, in a junction with the tested geometry. The eddy in front of the lateral branch is clearly visible.

The comparison of the instantaneous pressure in location 1 (engine port exit) for the two junctions denotes a larger peak pressure for junction 2 (with reed) due to its behavior as a *partially closed end* in comparison with junction 1 (without the reed). As described earlier, this pressure is the superposition of the incident and reflected waves. Since the engine is at the same operating condition in both cases, the incident wave is similar and the high peak pressure in junction 2 is the result of the reflected overpressure wave from the junction due to the flow restriction at the junction inlet (see Fig. 5). This characteristic of the junction might have a strong influence on the pumping losses and the volumetric efficiency of the engine.

The pressure measured at location 5 for both junctions (closed

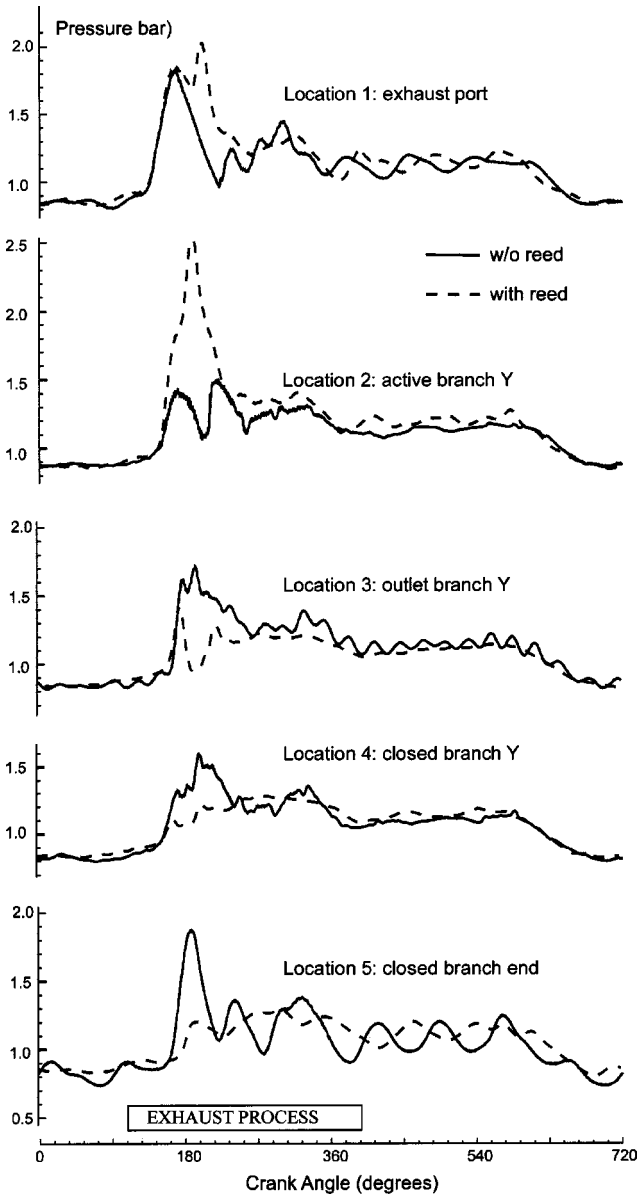


Fig. 6 Measured pressures

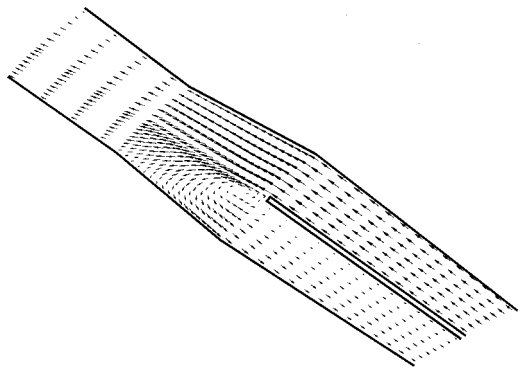


Fig. 7 Velocity vectors at junction symmetry plane calculated by 3D modeling

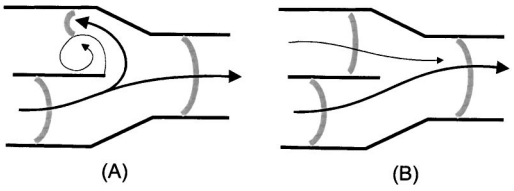


Fig. 8 Flow patterns at junction

end of the lateral branch) shows a much lower peak value for the junction with the reed that suggests the presence of a stronger eddy. This could be due to the lower discharge section (and therefore higher flow speed) and the bigger contact surface between both branches.

The comparison of the instantaneous pressure at the discharge section (3) shows a smoother and lower pressure pulse for the junction with reed due to the lamination imposed by this device.

The amount of energy transmitted upstream through the lateral branch is a key element in the interference with the exhaust process of other cylinders. It also influences the energy flow downstream to the turbine in turbocharged engines. This important flow pattern in the junction is represented in Fig. 8*A*.

In spite of this directional effect, a pressure pulse is transmitted into the lateral branch and is reflected at the closed end. The discharge of this reflected pulse into the junction occurs when the exhaust incident pulse is still blowing into the junction through the active branch. This situation determines another important flow pattern with both branches discharging to the exit branch as depicted in Fig. 8*B*. The flow of the active branch, with high flow velocity, transfers momentum to the lateral branch flow and helps its discharge into the exit branch.

In Fig. 9 the measured pressure at location 3 (outlet branch) and location 4 (closed branch Y) for the junction without reed and for several engine speeds is plotted against crank angle during the blow-down process. It is important to remark that the two measured pressures have the same trend except for the period when dynamic effects appear. When the pressure pulse is crossing the junction an important pressure drop arises due to the eddy formation. The pressure drop increases obviously with engine speed up to 0.4 bar at 5500 rpm. Later, when the transmitted pulse is reflected at the closed end there is an increase of pressure at location 4 that allows the branch emptying.

The results presented in Fig. 9 show that even if there are six possible flow patterns in a three-branch junction, only the patterns shown in Fig. 8 are relevant in an exhaust junction. Therefore the proposed characterization test can advantageously replace the tedious characterization in a test flow rig. These two main flow patterns can be quantified by computing the instantaneous mass flow in the lateral branch. Its value is obtained from the instantaneous pressure measured in the closed end of the lateral branch (location 5).

The equation of the superposition of two waves is expressed in Eq. (1), where the subindices i and r correspond to the incident and reflected waves

$$P^{(\gamma-1)/2\gamma}=P_i^{(\gamma-1)/2\gamma}+P_r^{(\gamma-1)/2\gamma}-1 \tag{1}$$

In this equation P is the pressure ratio of the absolute pressure with a reference pressure p_{ref}.

In location 5 (closed end) Eq. (1) can be simplified to

$$P_{5r}=P_{5i}=\left[\frac{1}{2}\left(P_5^{(\gamma-1)/2\gamma}+1\right)\right]^{2\gamma/(\gamma-1)} \tag{2}$$

The particle velocities of the incident and reflected waves are

$$u_{5r}=-u_{5i}=\left(\frac{a_5-a_{\text{ref}}}{\gamma-1}\right) \tag{3}$$

where a_{ref} is a reference speed of sound. The local sound speed can be obtained from Eq. (4), where p_{ini} is an initial pressure and a_{ini} is obtained from an average temperature.

$$a_5=a_{\text{ini}}\left(\frac{p_5}{p_{\text{ini}}}\right)^{(\gamma-1)/2\gamma} \tag{4}$$

With the values of the incident and reflected wave velocities, the particle velocity and mass flow at the inlet of the lateral branch of the junction (section $4'$) can be computed. It is obtained by the composition (addition), at that section, of the time translated velocity components from Eq. (3), assuming isentropic flow and knowing its short length

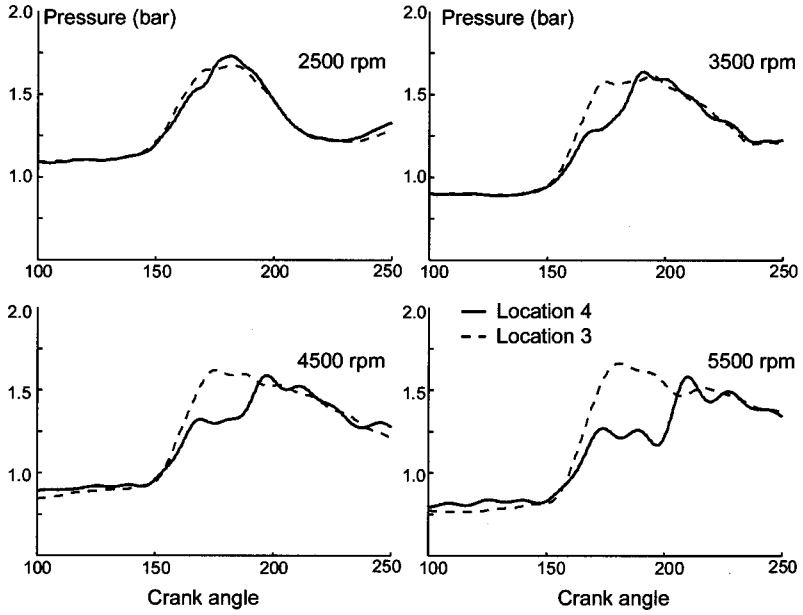


Fig. 9 Measured pressure downstream (location 3) and upstream (location 4: closed branch) the junction for several engine speeds

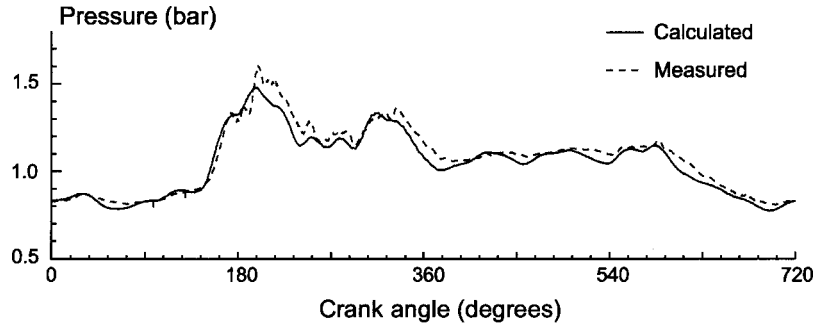


Fig. 10 Experimental and calculated pressure in location 4

$$u_{4'}(t) = u_{5i}(t-\Delta t) + u_{5r}(t+\Delta t) \quad (5)$$

The time interval is calculated from the length of the lateral duct and the mean temperature

$$\Delta t = \frac{L}{\bar{a}} = \frac{L}{\sqrt{\gamma R \bar{T}}} \quad (6)$$

The mass flow is finally computed as:

$$\dot{m}_{4'} = u_{4'}\rho_{4'}A \quad (7)$$

The determination of the density requires the temperature, which can be obtained from the modeled results. As described earlier a good pressure agreement (in phase and amplitude) guarantees the accuracy of the temperature results.

This process can be validated by a comparison of the experimental pressure measured in section 4 with the one obtained following the same method with the instantaneous pressure measured in section 5 and the corresponding time translation, assuming isentropic flow in the lateral branch

$$P_4^{(\gamma-1)/2\gamma}(t) = P_{5i}^{(\gamma-1)/2\gamma}(t-\Delta t) + P_{5r}^{(\gamma-1)/2\gamma}(t+\Delta t) - 1 \quad (8)$$

The agreement of both values, plotted in Fig. 10, confirms the validity of the process.

The instantaneous mass flow in the lateral branch obtained for the two junctions is presented in Fig. 11. It can be observed how the junction with the reed reduces the gas flow upstream of the lateral branch due to its stronger directionality in the section where the two lateral branches converge. Additionally, the discharge of the pulse, which starts at about a 200 Deg crank angle, is faster in the junction with the reed probably due to the higher momentum transfer from the active branch flow.

The actual mass flow can be also compared with the corresponding isentropic value obtained from the experimental instantaneous pressures in sections 3 and 4 with time translated to the inlet and outlet sections of the lateral branch (this assumes that the reflected pressure upstream in section 3 is negligible). These mass flows are plotted in Fig. 12 for the junction without the reed as a function of the crank angle. It can be observed how the lateral branch receives a mass flow much smaller than that obtained with the isentropic assumption during the angular period that the gas flows into the lateral branch.

This phenomenon can also be evaluated by computing, during this angular period, the discharge coefficient as the ratio between the actual and the isentropic mass flow at the inlet of the lateral branch. Its instantaneous value can be correlated with the flow velocity in the active branch.

In Fig. 13 the results at different engine speeds are compared. It can be observed that this coefficient can be characterized as a function of the flow velocity in the active branch. The trend is a reduction of the discharge coefficient with the intake flow speed. This method, proposed by Reyes et al. [20] and Galindo [21], requires an iterative process for the determination of the active branch velocity using the model proposed in the following section.

The observation of Fig. 13 denotes that there are other variables on top of the active branch velocity influencing the discharge coefficient; for instance the pressure ratio at the closed branch. The introduction of this parameter in the characterization procedure would improve the correlation but lose the benefits of the simplicity.

In conclusion, this analysis denotes that the introduction of a reed in the junction reduces the amount of energy transmitted to the lateral branch due to a higher directionality. It also increases the discharge flow of this branch due to an increased surface contact with the active branch that reflects a stronger overpressure pulse due to the flow restriction imposed by the reed. All this affects the instantaneous pressure in the exhaust port of the active cylinder and the one connected by the lateral branch of the junction and therefore the pumping work, volumetric efficiency, and residuals of the engine. Corresponding with the dynamics

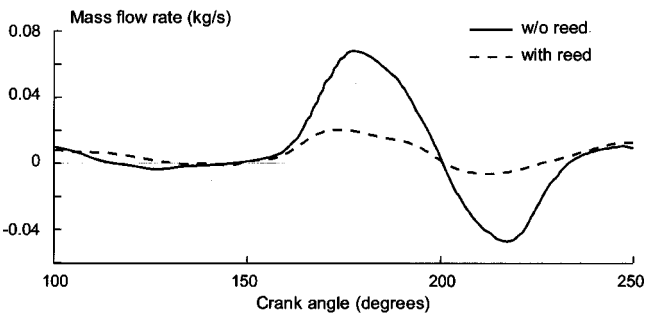


Fig. 11 Mass flow rate in the lateral branch

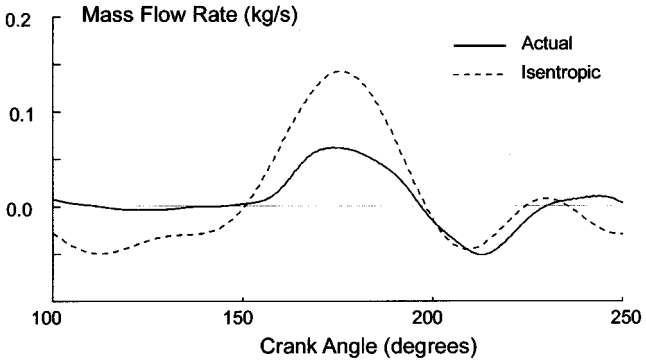


Fig. 12 Actual and isentropic mass flow rate in the lateral branch

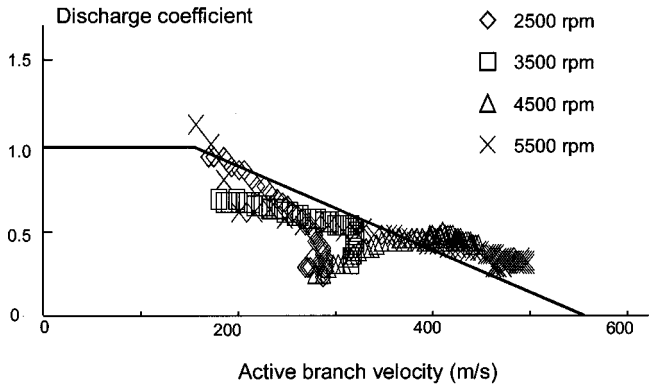


Fig. 13 Lateral branch discharge coefficient

in the lateral branches, the instantaneous gas energy in the exit duct is also affected and thus the mechanical work produced in the turbine.

An effective and efficient way to evaluate and optimize the influence of junction geometry on engine performance is an interactive process of experimental pressure measurement and modeling of the flow dynamics. The evaluation with experimental data only has the problem that the accuracy of the engine performance measurements (*bmep* or *bsfc*) is of the same order of magnitude as the junction geometry influence.

Modeling of the ''Y'' Junction

The results of the analysis can be used to improve the predictions of the junction model. The base model used in this case has been a set of ducts connected with a small volume. The volume model is the same that Benson [12] proposed for the engine cylinder because it allows any number of ducts connected to the volume with a nozzle in the discharge section. The nozzle has an effective area that takes into consideration the irreversibility of the process with a discharge coefficient.

The directional effect can be considered with a variable discharge coefficient in the lateral branch dependent on the active branch flow velocity as described in Fig. 13. This will take into consideration the partial reduction of the effective section of that branch due to the eddy produced in the discharge section of the junction. Even if the correlation of this function is limited, the model predictions using this linear function fit very well with the experimental measurements as presented below.

An important limitation of the model comes from the consideration that there is no flow speed in the junction volume. Since the pressure in that volume is computed from the balance of mass, the assumption translates in an isobaric transformation of the kinetic energy of the exhaust gas flowing into the junction in thermal energy. The assumption of total recovery of the kinetic energy through an isentropic process is not valid either. Therefore the alternative is to assume a partial recovery of that energy in the volume (which is equivalent to assume a non-negligible speed in the volume). This speed can be expressed by a fraction of the incoming flow speed:

$$K = \frac{u_{\text{volume}}}{u_{\text{inlet}}} \tag{9}$$

This coefficient is mainly related to the expansion rate from the active pipe to the junction. Therefore, the value of K is independent of the engine speed.

Figures 14 and 15 provide a comparison of the experimental measurements in sections 1 and 5 in both junctions with the results predicted by the initial model and the updated version (variable discharge coefficient in the lateral branch and flow speed in the volume). The value of K used in this case is 0.95. This value

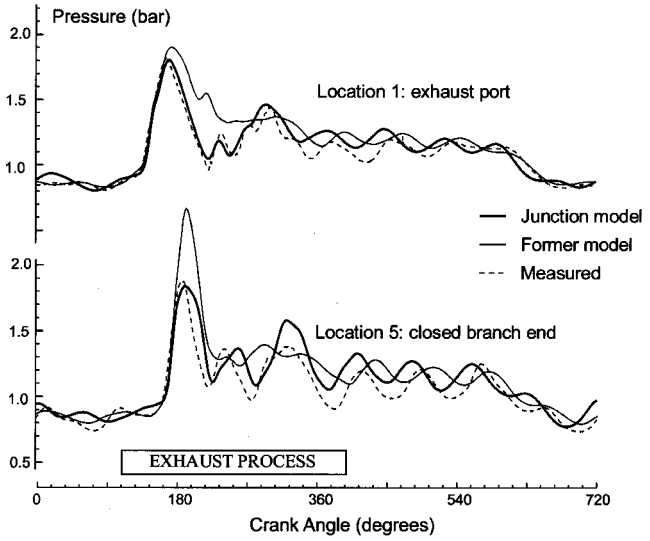


Fig. 14 Experimental and modeled pressure. Junction without reed

is defined to adjust the modeled predictions to the experimental measurements in the angular period of the crank angle where the reflected pulse in the active branch and the emptying of the lateral branch are important.

The improvement of the agreement is important. In the junction without reed, the reflected pulse is more realistic due to the consideration of the speed in the volume. This reduces the superposition pressure level close to the measured results. In the pulse transmitted to the lateral branch the variable discharge coefficient and the lower static pressure in the volume, reduce the pulse transmitted to a value close to the experimental results.

In the junction with reed, the comments are similar (Fig. 15). It is important to note how the modeled overpressure wave traveling upstream in the active branch is now very much in line with the experimental results. The improvement of modeling in these two junctions with a much different geometry shows the validity of the characterization procedure and the proposed model.

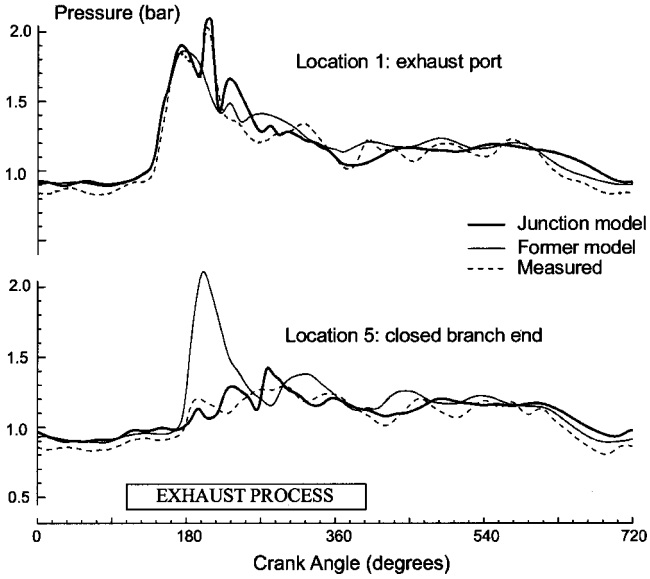


Fig. 15 Experimental and modeled pressure. Junction with reed

The implementation of that junction boundary condition in the one-dimensional engine model allows a better evaluation of the dynamics established in the engine, manifold length and junctions. Tuning conditions and effects can be predicted and analyzed.

Conclusions

- The proposed method is able to characterize the flow dynamics in an exhaust junction. The objective is the identification of the influence of the junction geometry on its dynamic behavior.
- The method has been applied to a ''Y'' junction. The influence of a reed in the junction section on the reflected wave upstream of the active branch and the energy transmitted to the lateral branch have been determined.
- The directional effect in the lateral branch can be quantified by a variable discharge coefficient, function of the velocity of the flow in the active branch.
- The results have been used to improve the junction modeling in a one-dimensional gas dynamic model. The agreement of the predicted results with the experimental measurements has largely improved.

Acknowledgment

The authors would like to express their recognition of the ''Direction des Etudes'' of RENAULT for the financial support to the broad project that has been the base of the results presented in this work.

Nomenclature

A = pipe cross-section area (m^2)
L = length (m)
M = mach number (−)
P = pressure ratio (−)
T = temperature (K)
a = speed of sound (m/s)
$\dot{m}$ = mass flow rate (kg/s)
p = static pressure (Pa)
t = time (s)
u = velocity (m/s)
γ = ratio of specific heats (−)
ρ = density (kg/m^3)

Subscripts

0 = related to stagnation conditions
r = reflected
i = incident

References

[1] Tabaczynski, R., 1982, ''Effect of the Inlet and Exhaust System Design on Engine Performance,'' SAE paper 821577.

[2] Winterbone, D. E., 1982, ''The Use of Gas Dynamics to Improve Reciprocating Engine Performance,'' Isr. J. Technol., **20**, pp. 71–80.

[3] Winterbone, D. E., 1987, ''The Application of Gas Dynamic for the Design of Engine Manifolds,'' Congress on Modelling of ICE, Valencia, Paper No. CMT8701.

[4] Blair, G. P., 1996, *Design and Simulation of Two-Stroke Engines*, Society of Automotive Engineers, Warrendale, PA.

[5] Benajes, J., Torregrosa, A. J., Reyes, M., and Brunel, J. P., 1993, ''The Effect of Exhaust Pipe Geometry on the Engine Performance,'' International Congress ''Diesel Engines,'' SIA 93081.

[6] Benajes, J., Reyes, E., Bermudez, V., and Serrano, J. R., 1998, ''Pre-Design Criteria for Exhaust Manifolds in IC Automotive Engines,'' SAE paper 980783.

[7] Janota, M. S., and Watson, N., 1973, ''Pulse Converters-A Method of Improving the Performance of the Turbocharged Diesel Engine,'' Proc. Inst. Mech. Eng., **1**, pp. 635–647.

[8] Benson, R. S., and Alexander, G. I., 1977, ''The Application of Pulse Converters to Automotive Four Stroke Cycle Engines. Part 1: Gas Dynamic Application of the Pulse Converters to Automotive Four-Stroke Cycle Engines. Part II: Optimization of Pulse-Converter-Turbo-Charger Combination,'' SAE paper 770034.

[9] Winterbone, D. E., Nichols, J. R., and Alexander, G. I., 1985, ''Efficiency of the Manifolds of Turbocharged Engines,'' Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.), **199**, pp. 137–149.

[10] Winterbone, D. E., Alexander, G. I., and Nichols, J. R., 1985, ''Developments in Methods of Considering Wave Action in Pipes Connected to I.C. Engines,'' International Simposium On Flows in Internal Combustion Engines-III, ASME Winter Annual Meeting, pp. 71–77.

[11] Benson, R. S., Woollatt, D., and Woods, W. A., 1963–64, ''Unsteady Flow in Simple Branch Systems,'' Proc. Inst. Mech. Eng., **178**, pp. 24–49.

[12] Benson, R. S., 1982, *The Thermodynamics and Gas Dynamics of Internal Combustion Engines*, Vol. 1, Clarendon Press, Oxford.

[13] Bingham, J. F., and Blair, G. P., 1985, ''An Improved Branched Pipe Model for Multicylinder Automotive Engine Calculations,'' Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.), **199**, pp. 65–77.

[14] Blair, G. P., 1994, ''Non-Isentropic Analysis of Branched Flow in Engine Ducting,'' SAE Paper 940395.

[15] Bulaty, T., and Widenhorn, M., 1993, ''Unsteady Flow Calculation of Sophisticated Exhaust Systems Using a Multibranch Junction Model,'' Trans. ASME: J. Eng. Gas Turbines Power, **115**, pp. 756–760.

[16] Endo, M., and Iwamoto, J., 1995, ''Numerical Analysis of Pulsatile Flow in Pipe System With Three-Way Junction,'' SAE paper 952068.

[17] Dimitriadis, C., Leschziner, M. A., and Winterbone, D. E., 1985, ''Computation of Three Dimensional Flow in Manifold-Type Junctions,'' International Simposium on Flows in Internal Combustion Engines III, ASME Winter Annual Meeting, pp. 57–62.

[18] Kuo, T., and Chang S., 1993, ''Three-Dimensional Steady Flow Computations in Manifold-Type Junctions and a Comparison With Experiment,'' SAE Paper 932511.

[19] Flamang, P., and Sierens, R., 1988, ''Study of the Steady-State Flow Pattern in a Multipulse Converter by LDA,'' J. Eng. Gas Turbines Power, **110**, pp. 515–522.

[20] Reyes, M., Espinoza, H., Reyes, E., and Galindo, J., 1997, ''Una Condición de Contorno para el Modelado del Efecto Eyector en Uniones de Conductors de Escape de Motores de Combustion Interna,'' Congreso Iberoamericano de Ingeniería Mecanica, La Habana, Cuba.

[21] Galindo, J., 1998, ''Design of Exhaust Manifold Junctions of Reciprocating Engines,'' (in Spanish) Servicio de publicaciones de la Universidad Politécnica de Valencia, Spain, ISBN 84-7721-614-2.

J. T. Lee

Y. Y. Kim

C. W. Lee

School of Mechanical Engineering,
Sungkyunkwan University,
Jangan-gu, Suwon 440-746, Korea

J. A. Caton

Department of Mechanical Engineering,
Texas A&M University,
College Station, TX 77840

An Investigation of a Cause of Backfire and Its Control Due to Crevice Volumes in a Hydrogen Fueled Engine

To understand the occurrence of backfire in hydrogen fueled engines using an external (inducted) fuel supply, a fundamental study was completed using a modified experimental engine. A relation was found between the crevice volume in the combustion chamber and the occurrence of backfire. The results showed that the crevice around the spark plug electrode was not a major cause of backfire, but the combustion state of the mixture in the piston top land crevice, second land, and ring groove did have a direct affect on backfire occurrence. By increasing the top land crevice volume and the amount of blow-by gas, the equivalence ratio before backfire occurred was extended. [DOI: 10.1115/1.1339985]

Introduction

Advantages of hydrogen-fueled engines include potentially low emissions and the use of nonpetroleum fuel [1]. Hydrogen fueled engines may be divided according to the fuel supply method into external (inducted) fuel mixture and direct-cylinder fuel injection [2]. The use of an external fuel mixture has the advantages of high thermal efficiency, simplicity, and avoidance of a high pressure fuel injection system [3]. So this version is relatively more feasible to put into practice than that of direct-cylinder fuel injection with its complex injection system. The use of an external fuel mixture, however, has the tendency to be susceptible to backfire under certain conditions [4]. To increase the feasibility of a hydrogen engine using external mixture and dual injection which is being investigated by the authors [5], it is necessary to control backfire occurrence.

The backfire phenomenon in engines is generally attributed to a fresh mixture in the cylinder being ignited while the intake valve is open, and before the normal combustion event. This flame then may flow ''backwards'' into the intake manifold during the intake process. The cause of the ignition is often attributed to an ignition source in the cylinder such as hot spots similar to that of the gasoline engine [6]. This possibility would be expected to occur during high temperature engine operation. But in fact, the backfire in hydrogen fueled engines frequently happens during idling or cold start conditions [7].

Since hydrogen fuel has ultralean limits of flammability and short quenching distances [8], hydrogen may be expected to burn very slowly in small and narrow volumes where other fuels such as gasoline may not be able to burn [9]. One possible cause of backfire, consequently, is that the hydrogen–air mixture could be burning slowly in the crevice volumes of an engine (such as the crevice volumes between the piston and cylinder and around the spark plug), and this slow burning could continue up to and during the intake process [10]. During the intake process, the hot burning gases could flow out of the crevice volumes and ignite the intake charge thus causing backfire. The observations of backfire during idling and cold start conditions, therefore, can be explained.

The major objective of this study is to examine this possibility. For this purpose, a systematic study using a modified, experimental hydrogen fueled engine was conducted that included artificial and actual crevice volumes. An artificial crevice was installed in the cylinder head (described below). In addition, the effect of the actual spark plug and piston–cylinder crevices in the engine on the occurrence of backfire was determined. Finally, the possibility of expansion of the equivalence ratio before backfire occurred is investigated in detail for the size of the piston top land crevice volume and by increasing the amount of blow-by gas.

Experimental Apparatus and Method

Experimental Apparatus. Figure 1 is a photograph of the experimental hydrogen fueled engine using an external fuel mixture. A single-cylinder, 433 cc, gasoline engine with a 5:1 compression ratio was modified to operate on hydrogen fuel. The hydrogen supply pipe (5 mm in diameter) was installed in the intake pipe which replaced the gasoline fuel supply system. The location for the hydrogen supply is 43 mm from the center of the intake valve stem. The cylinder head and the cylinder block were modified with a special water cooling system.

The spark ignition system was of some concern. Since hydrogen has a low flame ion density at spark ignition, residual electric energy may exist in the ignition coil [11]. In case this residual energy discharges abnormally during the intake process, this could cause backfire. To avoid this possibility, a full transistor type ignition system, with the secondary coil connected to earth, was used to eliminate any abnormal discharging.

Figure 2 shows a schematic diagram of the experimental apparatus which includes the hydrogen engine, dynamometer, hydro-

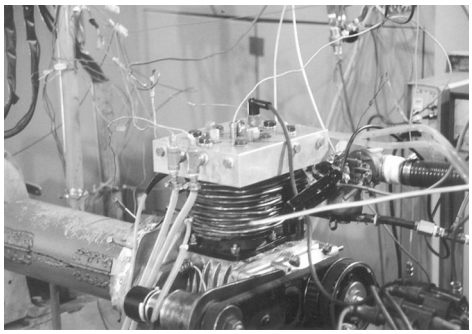


Fig. 1 Photograph of the hydrogen fueled engine

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, July 11, 2000; final revision received by the ASME Headquarters, Nov. 15, 2000. Associate Editor: D. Assanis.

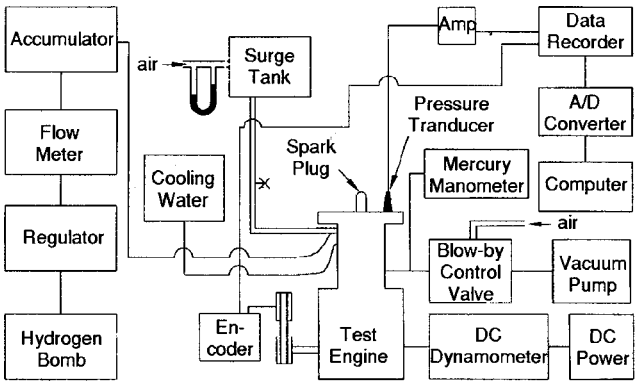


Fig. 2 Schematic diagram of the experimental apparatus

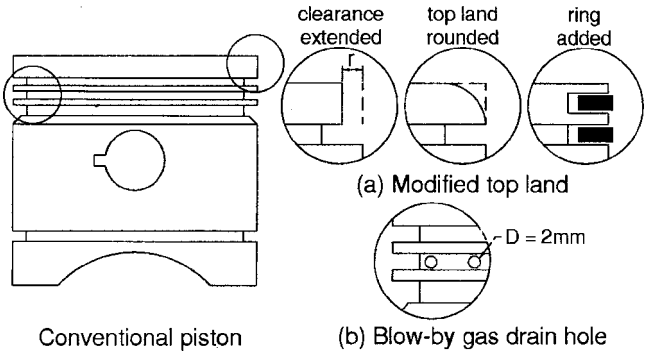


Fig. 4 Schematic of the piston geometries

gen gas supply system, intake and exhaust systems, cooling device, and data acquisition system. Hydrogen gas, obtained from a high pressure (12 MPa) tank, is decompressed to 150 kPa by high and low pressure regulators, and flows continuously into the intake pipe. The hydrogen fuel and air flow rates were measured by thermal type mass flow meters and by an orifice flow meter installed in front of surge tank. To control the mass of blow-by gas, a vacuum pump and control valve were connected with a blow-by gas outlet on the crank case.

The coolant was supplied to the cylinder head and block separately. The amount of coolant flow was controlled by a proportional control type solenoid valve installed at the coolant passage outlet. For measuring the cylinder gas pressure and the fluctuating pressure in the intake pipe, piezo type pressure transducers were inserted in the cylinder head and the intake pipe. Measured data were stored in a data recorder and calculations were completed with a personal computer.

Experimental Method. The experiment was to measure the fuel–air equivalence ratio when backfire occurs. This was done by increasing the equivalence ratio until backfire was detected by sound and by the indication from the intake system pressure. The equivalence ratio that permitted normal engine operation just before backfire occurs is called the backfire limit (BFL) equivalence ratio. The BFL equivalence ratio was determined as a function of a number of engine design variables. These variables included the volume of the artificial crevice installed in the cylinder head, the crevice of the spark plug, the crevice volume between the piston and cylinder liner in the base engine, and the amount of blow-by gas.

Figure 3 shows the location and detail of the artificial crevice installed in the cylinder head. The crevice has a cylindrical shape with a 10 mm diameter and its depth is changeable using the crevice volume control plug. The depth of the crevices studied were 0 (no crevice), 3 mm, and 6 mm. These dimensions result in crevice volumes of 0, 236, and 471 mm^3, respectively.

The actual crevices in the combustion chamber are due to the volumes around the spark plug, between the piston and cylinder liner, and due to the head gasket cutout. The spark plug was fitted to the cylinder head surface to eliminate the effect due to the spark plug threads. The effect of the spark plug crevice due to the space around the plug center electrode, however, is examined. For certain cases of a spark plug with no crevice volume, the space between the center electrode and threads is filled with a ceramic bond material.

The crevice between the piston and the cylinder liner includes the crevice of the piston land and of the ring groove. The size of top-land crevice volume is the most important since this volume is the largest among these spaces, and the mixtures in these regions are expected to flow back into the cylinder first. To study this aspect, the size of this volume was varied. Figure 4(*a*) and Table 1 show the dimension and the size of these crevice volumes for each piston. As shown, these crevice volumes are associated with the base piston, and modified pistons with distances between the piston top land and the cylinder wall of 0.9, 1.4, and 1.9 mm. A fifth configuration was a piston with an additional ring at the top land. A sixth configuration was a piston with a rounded top land. These pistons are called conventional, A, B, C, ring-added and rounded types, respectively. For these cases, the constant compression ratio was used in order to avoid the compression ratio effect on combustion.

The head gasket cut out crevice and the piston crevice are so close that they interact with each other. So for purposes of this work, the crevice due to the head gasket was eliminated by using a head gasket fitted to the cylinder head.

The amount of blow-by gas was changed by decreasing the pressure in the crank case. The pressure in the crank case is decreased from 0 mm of Hg to −500 mm of Hg by increments of 100 mm of Hg using a vacuum pump. In this portion of the study,

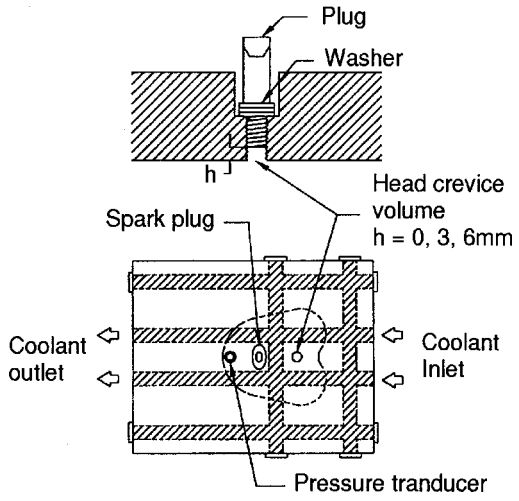


Fig. 3 The location and detail of the artificial crevice volume installed in the cylinder head

Table 1 The size of each piston

	Top land clearance (mm)	Crevice volume (mm^3)
Conventional	0.4	900.7
A piston	0.9	2015.3
B piston	1.4	3117.4
C piston	1.9	4206.9

three types of pistons were examined: the conventional piston, a piston without the second ring, and a piston with blow-by gas drain holes. To restrict the burning gas in second land and ring groove crevice volume flowing back into cylinder, the blow-by gas drain used 12 holes (2 mm diameter) located at the second ring groove as shown in Fig. 4(*b*).

The coolant temperature was varied from 90 to 40°C in this experiment. Engine speed and spark timing are fixed at 1600 rpm and minimum spark advance for best torque (MBT), respectively.

Results and Discussion

The Effect of the Artificial Crevices. Figure 5 shows the BFL equivalence ratio as a function of the size of the artificial crevice installed in the cylinder head for a coolant temperature of 50°C. As mentioned above, the BFL equivalence ratio is defined as the equivalence ratio that permitted normal engine operation just before backfire occurs. As shown in this figure, the BFL equivalence ratio decreases as the crevice volume increases. For this low coolant temperature, a hot spot is not expected to form. These results, therefore, mean that the crevice volume affects backfire occurrence. An obvious explanation is that the fresh mixture inducted during the intake process mixes with the hot gases exiting the crevice volume and this causes the backfire as was mentioned previously.

The quenching distance of the mixture in the crevice volume, which affects the occurrence of backfire, depends on the crevice wall temperature, and the crevice wall temperature is a function of the coolant temperature. Figure 6 shows the BFL equivalence ratio as a function of the coolant temperature for the three crevice volumes. For the case of no crevice volume, the BFL equivalence ratio decreases with increasing coolant temperature as is well known. For the 3 mm depth crevice, the BFL equivalence ratio decreased slightly with increasing coolant temperature. But for the 6 mm depth crevice, the BFL equivalence ratio increased with increasing coolant temperature. As the coolant temperature increases, the unburned mixture temperature increases and quenching distance decreases. These conditions affect the combustion state of the mixture in the crevice volume. For the 6 mm depth crevice, therefore, a possible reason for the slight differences of the BFL equivalence ratio between 70 and 90°C is explained by the fact that the difference of the combustion state in the crevice volume decreases gradually with the increase of coolant temperature.

Considering the above results, it is expected that the BFL equivalence ratio in cases with and without crevice volumes is similar under high coolant temperature. For the higher coolant temperatures, Fig. 6 shows that the BFL equivalence ratio increases with increasing crevice volume. The mixture leaving the crevice may be regarded as a jet flame injected from the crevice volume into the combustion chamber, like an auxiliary combustion chamber or torch combustion.

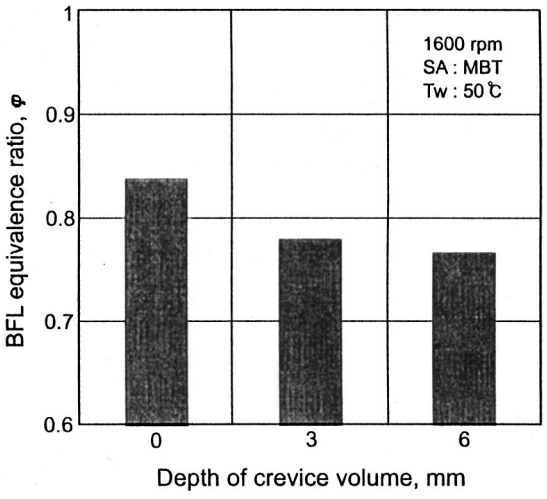


Fig. 5 The BFL equivalence ratio for each artificial crevice volume

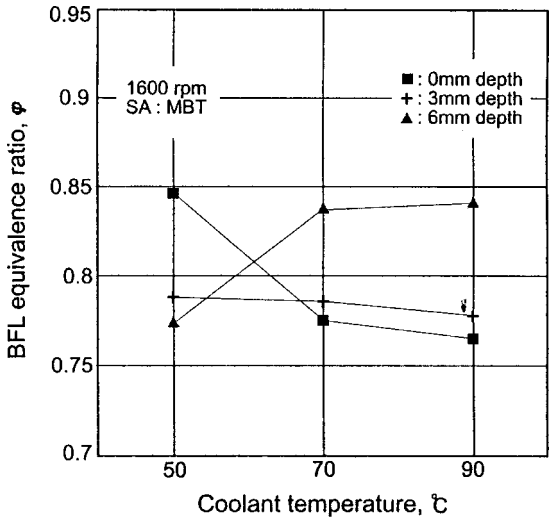


Fig. 6 The BFL equivalence ratio as a function of coolant temperature for artificial crevice size

The BFL equivalence ratio for the 6 mm depth crevice (471.2 mm^3) is roughly lower by about 8.5 percent at 50°C, and higher by about 8 percent at 90°C than those for the case with no crevice volume.

From the above results, it was found that the existence of a crevice volume and the size of that volume may affect backfire occurrence. Based on these facts, the effect of the actual crevice volumes in the engine was investigated next.

The Effect of Actual Crevices

Spark Plug Crevice. Figure 7 shows the BFL equivalence ratio for the cases with and without a crevice volume around the spark plug center electrode. In this case, the conventional type of piston was used. The figure shows that the BFL equivalence ratio is not too dependent on this clearance volume for either coolant temperature. As is well known, the spark plug is exposed to hot burned gas for a long time. This results in any unburned mixture around the spark plug center electrode burning out quickly before the intake process. This may account for the reason why the effect of the crevice around the spark plug is not too important. Since the difference of the BFL equivalence ratio is slight although the coolant temperature is as cold at 50°C, it is likely that the crevice volume around the spark plug does not contribute to backfire occurrence.

Piston Top Land Crevice. Figure 8 shows the BFL equivalence ratio as a function of the top land crevice volume for a coolant temperature of 60°C. As the crevice volume increases, the BFL equivalence ratio increases. This indicates that the crevice volume of the piston top land contributes to backfire occurrence. The heat loss from the mixture in the crevice volume to the piston and cylinder wall surfaces decreases as the distance between the piston top land and cylinder wall increases. This heat loss and the related temperature change will affect the combustion tendency of the mixture in these regions. One explanation for the above results, therefore, may be this secondary effect of a change in the ignition characteristics of the escaping mixture.

Also, Fig. 8 indicates that the increasing rate of the BFL equivalence ratio is not necessarily linear with increasing crevice

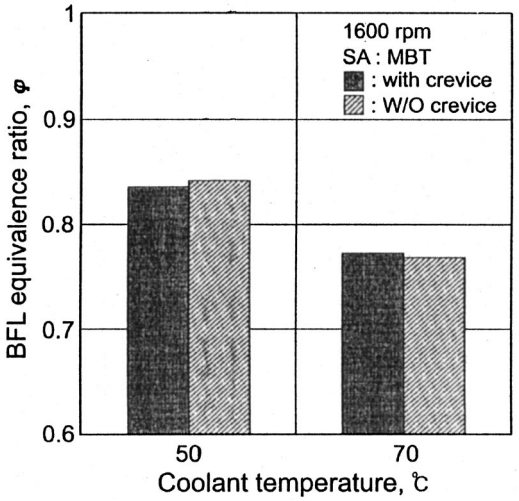


Fig. 7 The BFL equivalence ratio with and without crevice volumes around the spark plug for two coolant temperatures

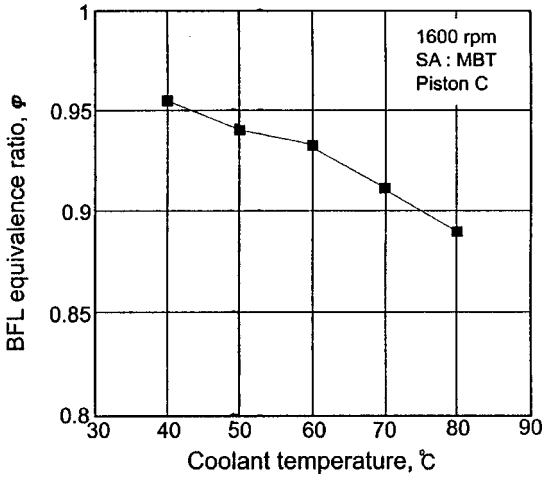


Fig. 9 The BFL equivalence ratio as a function of coolant temperature for the type C piston

volume. That is, the BFL equivalence ratio increases more for the increase to the B piston, but then increases less for the increase to the C piston. This means that increasing the crevice volume has a limitation for increasing the BFL equivalence ratio.

The top land crevice clearance of the C piston is 1.9 mm. This distance is much wider than the quenching distance of hydrogen so that the mixture in these spaces might burn rapidly. But, considering that the BFL equivalence ratio does not increase more than a certain extent, it is possible that the cause of backfire exists at another place. Regarding this, the effect of the mixture in the piston second land and ring groove was studied. Although the mass of the mixture in this small volume is low, some of this mixture will flow back into the cylinder during or up to the expansion and the exhaust processes [12]. Since the mixture in these spaces cannot burn out thoroughly in spite of the increase of the top land crevice volume, it could be a cause of backfire occurrence. If this assumption is valid, then the lower the coolant temperature is, the more the BFL equivalence ratio will increase. Temperatures in the vicinity of the second land and the ring groove space depend on that of the combustion chamber, which in turn is influenced by the coolant temperature.

To verify these tendencies, the BFL equivalence ratio was investigated as a function of coolant temperature for the C piston with the largest crevice volume. Figure 9 shows that the BFL equivalence ratio increases as the coolant temperature decreases. As mentioned above, the coolant temperature changes the conditions in the combustion chamber, and affects the quenching distance of the hydrogen. Since the clearance of the second land and the ring groove was not modified, the noted tendency of the increase of the BFL equivalence ratio with decreasing coolant temperature may be caused by changes in the quenching distances.

From the above results, it was found that the mixture in the second land and the ring groove as well as the top land affect backfire occurrence.

Expansion of BFL Equivalence Ratio

Changes to the Piston Crevice Volume. Figure 10 shows the BFL equivalence ratio as a function of the crevice volume of the piston top land for five coolant temperatures. The BFL equivalence ratio increases with the increase of the crevice volume, and it increases as the coolant temperature decreases. On average, the BFL equivalence ratio increases 5.7 percent as the size of the crevice volume increases about 4.7 times larger than the conventional condition. The BFL equivalence ratio is higher by about 1.7

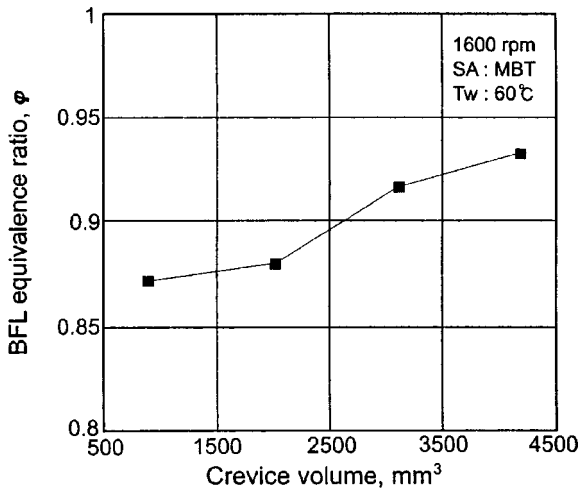


Fig. 8 The BFL equivalence ratio as a function of the crevice volume for a coolant temperature of 60°C

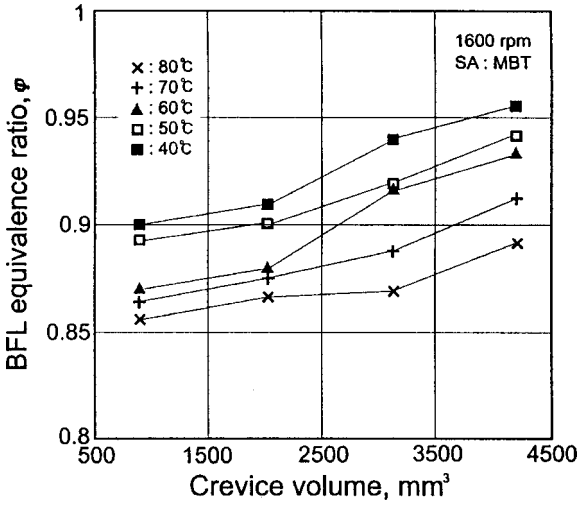


Fig. 10 The BFL equivalence ratio as a function of the piston top land crevice volume for four coolant temperatures

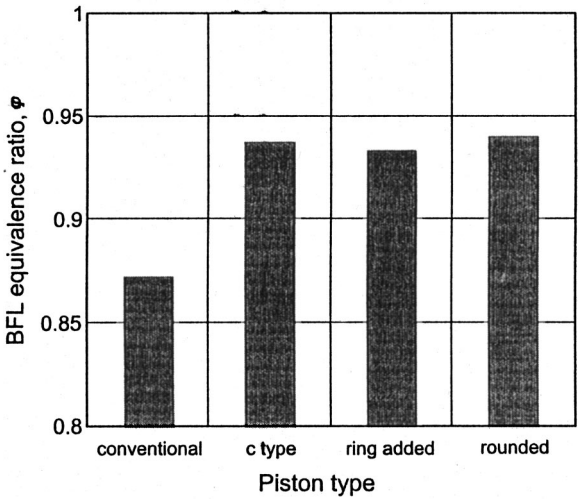


Fig. 11 A comparison of the BFL equivalence ratio for the four piston types

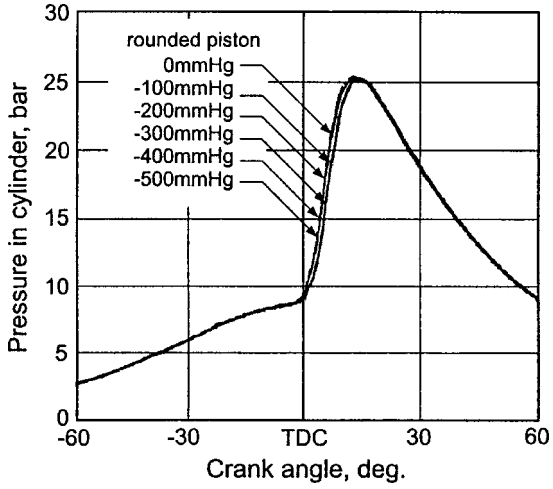


Fig. 12 The cylinder pressure as a function of crank angle for the rounded piston top for six crank case pressures

or 3.6 percent than the above case, when the coolant temperature decrease from 80 to 70°C or 60°C for each piston.

If the crevice volume of the piston top land affects the backfire occurrence, a method for controlling backfire may be to remove these volumes as much as possible. To examine this case, the BFL equivalence ratio was determined for the piston with the rounded top, and for the ring-added piston (designs which minimize the top clearance volume). For the ring-added piston, an additional ring is inserted into the upper top land to minimize the size of top-land crevice volume and to prevent possible mixture flowing into these spaces. Figure 11 shows the results that compare the BFL equivalence ratio for these two pistons with the conventional and the C pistons. The BFL equivalence ratio for the rounded, ring-added type and C piston are nearly the same and all three have BFL equivalence ratios that are remarkably higher than the BFL equivalence ratio of the conventional piston.

Since the top land crevices for the rounded and ring-added type pistons are minimized as much as possible, it is unlikely that the mixture in the top land crevice is burned slowly. The results of these three kinds of pistons indicate some possibilities to control backfire are to reconfirm that the mixture in the second land and the ring groove crevice can affect backfire occurrence.

Increase the Amount of Blow-By Gas for Preventing Back Flow. A portion of cylinder gas flows into the second land and the ring groove, and then back into the cylinder or into the crank case depending on the cylinder pressure. It is difficult for the mixture in these regions to burn out wholly. But by increasing the amount of blow-by gas, these mixtures may be prevented from flowing back into the cylinder.

The amount of blow-by gas increases due to the clearance between the piston ring and cylinder liner and due to the passage between the piston ring and the ring groove. The former has the sealing effect by the oil film. The latter has a very short opening time, since it is opened instantaneously by the inertia force of the ring at top dead center (TDC) or bottom dead center (BDC). For these reasons, it is expected that the amount of blow-by gas does not increase, even though the pressure in the crank case decreases. This is confirmed by examining the cylinder pressures for varying crank case pressure as shown in Fig. 12.

To understand the effect of increasing blow-by gas, the BFL equivalence ratio as a function of the crank case pressure is shown in Fig. 13 for the case with the rounded type piston (which minimizes the effect of the top land crevice volume). As shown, the BFL equivalence ratio is essentially constant in spite of the change of pressure in the crank case. This result confirms that the amount of blow-by gas is not increased even though the crank case pressure is decreased, as mentioned above.

To further study the blow-by gas effect, a case was completed where the second ring was removed. Figure 14 shows the BFL equivalence ratio and maximum cylinder pressure as a function of crank case pressure for this case. It follows from the slight decrease of the maximum in-cylinder pressure that the amount of blow-by gas is effected by the method of removing the second ring. The BFL equivalence ratio increases gradually and then decreases slightly with decrease of in-crank case pressure. It may be explained by the fact that back flow of hot burning gas from the second land and the ring groove is prevented by the increased flow of the blow-by gas. On the other hand, the amount of hot gas induced into these regions is also increased as the pressure in the crank case decreases. So the mixture quenched by the cooling effect of the piston and cylinder liner can be burned rather slowly. If a portion of these mixtures flows back to the cylinder, the possibility of backfire occurrence becomes higher. The tendency for decrease of the BFL equivalence ratio may be explained by these reasons.

These observations suggested another experiment. For this case, drain holes are installed at the second ring groove as illus-

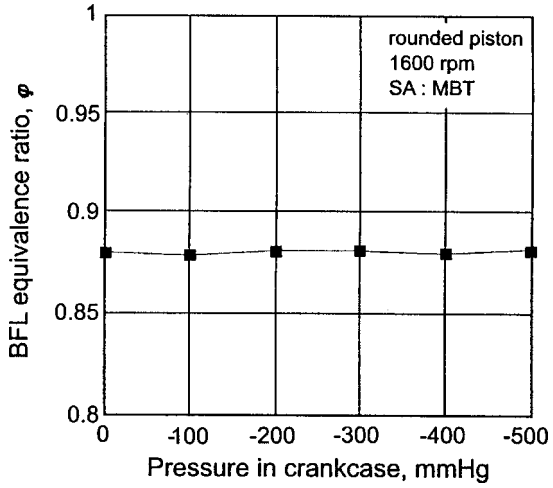


Fig. 13 The BFL equivalence ratio as a function of the crank case pressure for the piston with the rounded

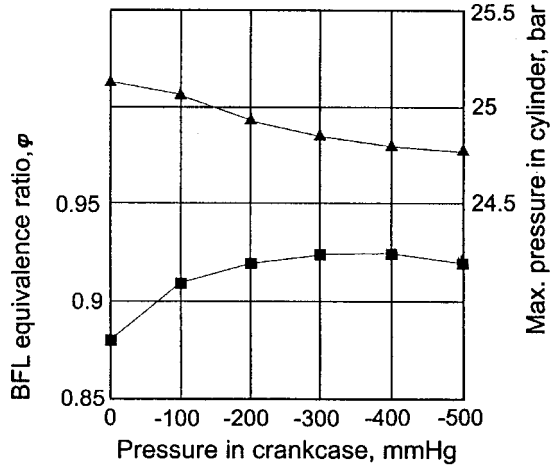


Fig. 14 The BFL equivalence ratio and the maximum cylinder pressure as a function of the crank case pressure for the case where the second ring is removed

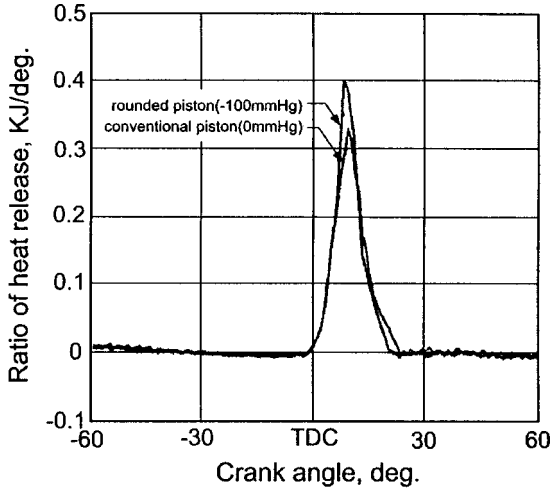


Fig. 16 The rate of heat release for the cases with the rounded piston with drain hole ($\phi_{BFL}=0.88$) and for the conventional piston ($\phi_{BFL}=0.81$)

trated in Fig. 4(*b*). Figure 15 shows the BFL equivalence ratio and maximum in-cylinder gas pressure for this case. The decrease of maximum in-cylinder gas pressure indicates an increase of blow-by gas. The BFL equivalence ratio increases as the pressure in the crank case is decreased by 200 mm Hg. The BFL equivalence ratio increases by 11 percent compared to the conventional piston under the above conditions although this is not a practical approach.

Further decreases of the crank case pressure did not increase the BFL equivalence ratio any further. It may be that the increasing method of the amount of blow-by gas with drain hole has a limit to prevent mixture in these regions from flowing back instantaneously, or there is another cause of backfire occurrence.

It is also possible that the BFL equivalence ratio increases by the decrease of combustion chamber surface temperature due to the reduction of heat release, which is caused by the increasing amount of blow-by gas. To examine these possibilities, the heat release for the rounded piston with drain hole is compared to the conventional piston. Figure 16 shows the rate of heat release as a function of crank angle for the two cases. The BFL equivalence ratio for the rounded and conventional pistons are $\phi=0.88$ and $\phi=0.81$, respectively. The heat release rate for the case with the rounded piston is larger than for the case with the conventional piston due to the increase of the supplied energy by the higher BFL equivalence ratio, even though the amount of blow-by gas is increased.

Judging from the above, it might be thought that the expansion of the BFL equivalence ratio with the increase of blow-by gas is caused mainly by the extraction of mixture existing in the piston crevice volume than the decrease of supplied energy.

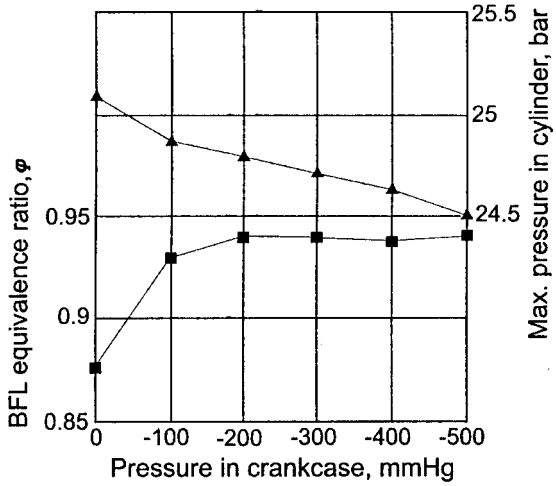


Fig. 15 The BFL equivalence ratio and maximum cylinder pressure as a function of crank case pressure for the case using the piston with drain holes for the blow-by gas

Summary and Conclusions

An investigation was conducted to determine the cause of backfire in hydrogen fueled engines which use external mixture presentation. An experimental engine was modified to isolate the importance of crevice regions on causing backfire. Crevice regions were expected to be important since hydrogen–air mixtures could burn slowly in these small regions and this burning could continue up to and during the intake process. During the intake process, the hot burning gases could flow out of the crevice volumes and ignite the intake charge thus causing backfire. The observations of backfire during idling and cold start conditions, therefore, can be explained.

The experiments conducted included the use of an artificial crevice region which was installed in the cylinder head to examine this effect. The results of this experiment demonstrated the importance of crevice volumes on causing backfire.

For the actual crevices in the engine, it was found that the crevice around the spark plug electrode did not appear to cause backfire, but the crevice of the piston top land remarkably affected the occurrence of backfire. It was also found that the crevices of the second land and the ring groove may be related to backfire occurrence.

The BFL equivalence ratio was increased about 5.7 percent compared to the conventional piston by increasing the piston top land crevice, and it was increased by about 11 percent by increasing the blow-by, which minimized the flow of hot gases from the ring pack to the cylinder, although this is not a practical approach.

In summary, this study has shown that the occurrence of backfire can be related to piston–cylinder crevices. By careful design of the crevice regions such as minimizing the piston crevice volume, it is expected that the occurrence of backfire could be controlled. Further detailed work will be needed to quantify these effects.

References

[1] Kukkonen, C. A., 1981, ''Hydrogen as an Alternative Fuel,'' Society of Automotive Engineers, Technical Paper No. 810349, pp. 127–149.

[2] Furuhama, S., 1991, ''Trend of Social Requirements and Technical Development of Hydrogen-Fueled Automobiles,'' JSME Rev., **13**, pp. 4–13.

[3] Mathur, H. B., 1985, ''Hydrogen Fueled Internal Combustion Engines,'' *Proceedings of the National Workshop on Hydrogen Energy*, New Dehli, July, pp. 159–177.

[4] Kim, J. M., Kim, Y. T., Lee, J. T., and Lee, S. Y., 1995, ''Performance Characteristics of Hydrogen Fueled Engine with the Direct Injection and Spark Ignition System,'' Society of Automotive Engineers, Technical Paper No. 952488, pp. 162–175.

[5] Kim, Y. Y., Park, J. B., and Lee, J. T., 1997, ''A Study on Development of Hydrogen Fueled Engine With High Power and High Efficiency,'' *Proceedings of 4th Japan-Korean Joint Symposium*, Yokohama, Aug. 29, pp. 62–71.

[6] Lynch, F. E., 1974, ''Backfire Control Techniques for Hydrogen Fueled Internal Combustion Engines,'' *Proceedings of Hydrogen Energy*, Part B, Miami, pp. 686–696.

[7] Furuhama, S., 1977, ''Combustion Improvement in a Hydrogen Fueled Engine,'' Int. J. Hydrogen Energy, **2**, pp. 329.

[8] Lewis, B., and von Elbe, G., 1987, *Combustion, Flames and Explosions of Gases*, Academic Press, New York.

[9] Saika, T., and Korematuse, K., 1986, ''Flame Propagation into the Ring Crevice of a Spark Ignition Engine,'' Society of Automotive Engineers, Technical Paper No. 861528, pp. 1–8.

[10] Koyanagi, K., Hiruma, M., and Furuhama, S., 1994, ''Study on Mechanism of Backfire in Hydrogen Engines,'' Society of Automotive Engineers, Technical Paper No. 942035, pp. 99–106.

[11] Kondo, T., Hiruma, M., and Furuhama, S., 1996, ''A Study on the Mechanism of Backfire in External Mixture Formation Hydrogen Engines,'' *Proceedings of WHEC*, Vol. III, Stuttgart, June, pp. 1547–1556.

[12] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York, pp. 601–608.

Experimental Study of a Hydrogen-Fueled Engine

R. Sierens
e-mail: roger.sierens@rug.ac.be

S. Verhelst

Laboratory of Transport Technology,
Ghent University,
Sint Pietersnieuwstraat 41,
B-9000 Gent, Belgium

The Laboratory of Transport Technology (Ghent University) converted a GM/Crusader V-8 engine for hydrogen use. The engine is intended for the propulsion of a midsize hydrogen city bus for public demonstration. For a complete control of the combustion process and to increase the resistance to backfire (explosion of the air–fuel mixture in the intake manifold), a sequential timed multipoint injection of hydrogen and an electronic management system is chosen. The results as a function of the engine parameters (ignition timing, injection timing and duration, injection pressure) are given. Special focus is given to topics related to the use of hydrogen as a fuel: ignition characteristics (importance of electrode distance), quality of the lubricating oil (crankcase gases with high contents of hydrogen), oxygen sensors (very lean operating conditions), and noise reduction (configuration and length of intake pipes). The advantages and disadvantages of a power regulation only by the air-to-fuel ratio (as for diesel engines) against a throttle regulation (normal gasoline or gas regulation) are examined. Finally, the goals of the development of the engine are reached: power output of 90 kW, torque of 300 Nm, extremely low emission levels, and backfire-safe operation. [DOI: 10.1115/1.1339989]

Introduction

Hydrogen is seen as one of the important energy vectors of the next century. Hydrogen as a renewable energy source provides the potential for a sustainable development particularly in the transportation sector. Hydrogen-driven vehicles reduce both local as well as global emissions.

Hydrogen-fueled engines are known for many advantages, among which is the very low concentration of pollutants in the exhaust gases compared to internal combustion engines using traditional or other alternative fuels. Further on, because of the wide flammability limits and the high flame propagation speed of hydrogen, a hydrogen-fueled engine is capable of very lean combustion (up to air-to-fuel ratios of 5).

To be able to run a hydrogen engine, the mixture formation of air and hydrogen does not need precise control [1]. Consequently, simple systems such as an external mixture system with a gas carburettor (venturi type) can be used for the fuel supply. Such a system is first implemented on the tested engine. However, a complete control of the combustion process is only possible with an injection system and an electronic control unit (electronic management system), as used for all new gasoline and diesel engines. Therefore, the carburettor is discarded to be replaced by a low-pressure gas injection system in the intake manifold, allowing multi-point sequential injection of the gaseous hydrogen fuel in each intake port just before the inlet valve.

Such an injection system, as applied to liquid fuels (gasoline, liquid LPG, . . .), has several advantages including the possibility to tune the air fuel ratio of each cylinder to a well-defined value, increased power output, and decreased cyclic variation of the combustion process in the cylinders. Timed injection also has an additional benefit for a hydrogen-fueled engine, as it imparts a better resistance to backfire. All these advantages are well known [2–5].

The disadvantage of low-pressure sequential gas injection is the low density of the gas. For smaller engines running at high speeds (traction application), the injectors have to deliver a high volume of gas in a very short time. Other problems may arise with the durability of the injectors and possible leaks.

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, April 27, 2000; final revision received by the ASME Headquarters, November 3, 2000. Associate Editor: D. Assanis.

In the period 1993–1995, different types of electromagnetic gas injectors were tested in detail [6,7]. Leakage, unequal response time (opening delay), and low durability were the main shortcomings. In the mean time, the research on gaseous injection systems (natural gas, LPG, . . .) has increased enormously by the specialized companies.

As mentioned earlier, sequential timed injection increases the resistance to backfire (explosion of the air–fuel mixture in the intake manifold). In nearly all cases, backfire-safe operation imposes a limitation of the operation region of the air–fuel mixture on the ''rich'' side, thus for high load conditions. This restriction is decreased by the use of a multi-point sequential injection system. Direct injection in the combustion chamber, cryogenic storage (LH_2 tank), and pump is even better, but not technically available for mass production [8].

Description of the Test Rig

Engine. A GM 454 spark-ignited engine (commonly known as the Chevrolet ''Big Block'') is adapted to gaseous fuels. The engine specifications are

- 8 cylinders in V
- bore: 107.95 mm
- stroke: 101.60 mm
- swept volume: 7.41 (454 in^3)
- compression ratio: 8.5:1
- engine speed: 750–4000 rpm
- ignition sequence: 18436572
- EVO 93° c.a. before BDC
- EVC 62° c.a. after TDC
- IVO 42° c.a. before TDC
- IVC 95° c.a. after BDC

The engine is connected to a water (Froude) brake.

The Fuel Supply System. As mentioned, the engine is first equipped with a gas carburettor. This gas carburettor allows testing with different fuels: pure hydrogen, natural gas, and hythane (a mixture of hydrogen and natural gas) [9]. The fuel supply system with mass flow meter, mass flow controller, and control unit allows experimenting with natural gas/hydrogen mixtures in variable proportion, regulated independently of the engine operating conditions.

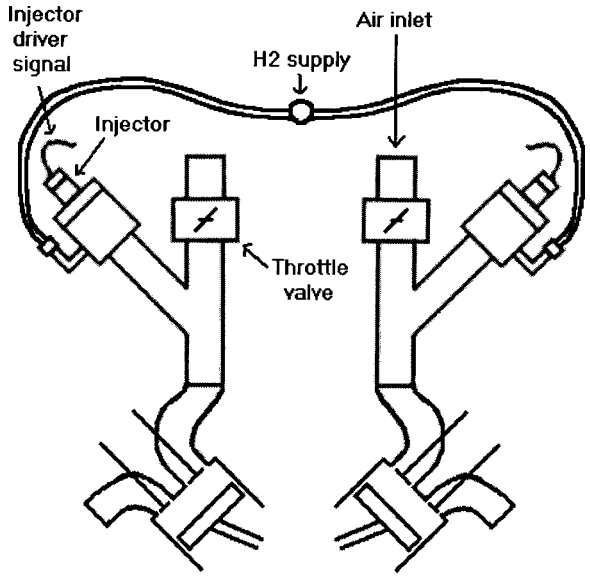


Fig. 1 Injection system setup

A multi-point sequential injection system is then implemented to take advantage of its controlling possibilities. The fuel is supplied from steel bottles with compressed hydrogen at 200 bar. After a pressure-reducing valve that expands the hydrogen to a pressure of about 3 bar, the hydrogen is admitted to a common rail system. From the common rail, eight tubes deliver the hydrogen to the eight individual injectors.

The injectors are prototypes, originally developed for use with natural gas (further information cannot be given at this time due to trade secret). In idling conditions, problems arose with deviations in injection duration between the individual injectors. This is due to the small reproducibility of the injection durations applied during idle run (of the order of 3 ms). New injectors are mounted with a shorter length of stroke to ensure good reproducibility with these short injection durations. Second, the injector needle cone angle is made more obtuse, to allow a greater fuel flow for a smaller lift of the injector needle.

Each cylinder has a short intake runner (no common intake manifold), and the injector is located at 12 cm from the cylinder head at an angle of 45°. This location and angle is studied with a CFD code to optimize the mixing of the hydrogen with air. Figure 1 gives a view of the installation of the injectors.

Apparatus. The engine is fully equipped with the usual sensors. The measurement/control signals are read and controlled by a PLC system (programmable logic controller). This system monitors engine speed, oil and coolant temperature, exhaust gas temperatures, etc., and shuts down the engine when necessary (by cutting off the hydrogen supply). All values are visible on a computer screen and can be stored in a Microsoft Excel worksheet.

The exhaust temperature and exhaust gas composition can be measured at the exhaust of each cylinder and at the end of each bank (V engine). Two oxygen sensors are installed at the common exhaust pipe of each bank, which allows an immediate reading of the air-to-fuel ratio of each bank. The oxygen sensors together with the exhaust temperatures give the possibility to check differences in mixture richness between the cylinders.

The exhaust gas components are measured with the following methods of measurement: $CO–CO_2–NO–NO_2$ (Multor 610, nondipersive infrared); O_2 (Servomex model OA 1100, paramagnetic); HC (Signal model 3000, flame ionization); and H_2 (Thermor 615, thermal conduction).

A high-pressure transducer (type AVL QC32) is located in one cylinder head (mounted flush with the combustion chamber wall of cylinder 1) giving in-cylinder pressure measurements, used for the calculation of, e.g., heat release analysis.

Experimental Program

An extensive test program is set up in different steps:

Step 1. Adaptation of the engine for hydrogen fuel with a carburetted fuel preparation system.

Step 2. For this carburetted version examination of variable compositions of hydrogen-natural gas mixtures (hythane) to obtain an increased engine efficiency and decreased emissions.

Step 3. The installation of a hydrogen timed injection system. Tests have to point out if the injection system is reliable and produces sufficient power and torque for traction applications, without backfire occurrence.

Step 4. Optimization of the intake manifold, the injection characteristics (pressure, timing), and the management system for the whole speed-load range of the engine.

The tests with the gas carburettor (venturitype mixing) are completely finished (step 1). The results on the effects of the use of hythane, step 2, were presented by Sierens and Rosseel [9]. The most important conclusions were as follows.

- For hythane with a low hydrogen content (up to 20 percent), a limited improvement in emissions can be obtained.
- Because of conflicting requirements for low hydrocarbon and low NO_x emissions, extremely low emissions are not possible without aftertreatment: to reduce hydrocarbon emission, λ must be less than 1.3, while for low NO_x emission, λ must be at least 1.5.
- For lower bmep, high efficiency can be achieved by increasing the hydrogen content and thus avoiding throttling losses. At the same time, unburned hydrocarbon emissions are minimized, while (for lean mixtures) NO_x emissions stay limited. This means that for optimal results the composition of the fuel should depend on the load.

The first results with the multi-point sequential injection system, step 3, are already given by Sierens [10]. This paper now gives part of the optimization of the engine parameters, step 4, and some problems arising from the use of hydrogen as a fuel in internal combustion engines. Further optimization is in progress.

Optimization of the Engine Parameters

One of the main problems in running a hydrogen-fueled engine is backfire. Backfire occurs when the air–fuel mixture ignites before IVC, with subsequent explosion in the intake manifold. This is caused by hot spots in the combustion chamber (residual gases, surface deposits, valves, . . .). To avoid backfire, the engine is run with a lean mixture. Several tests have shown that with an air-to-fuel ratio λ of 2, backfire-safe operation is obtained [11]. But with such lean mixtures, the power output of the engine decreases [2]. As the engine has to be built in a city bus, a power output of 90 kW and a torque of 300 Nm are the minimum conditions.

The main objective of the optimization, step 4, is thus to obtain maximum engine torque and power over the whole of the speed range (750–4000 rpm). This optimization is done with a fixed air-to-fuel ratio λ of 2. Figure 2 shows the power output (kW) and torque (Nm) for the speed range with $\lambda = 2$ and the ignition timing (IT) set to 20° c.a. BTDC. Figure 3 shows the corresponding brake mean effective pressures (bmep) and brake specific fuel consumptions (bsfc).

These are the initial and starting settings of the engine, which were the results of the third step in the experimental program. For comparison, these figures also show the results with the carburetted fuel mixing system (results of first step). The increase of the power output and torque for the injection version is mainly due to the better filling of the engine. These tests are done with wide

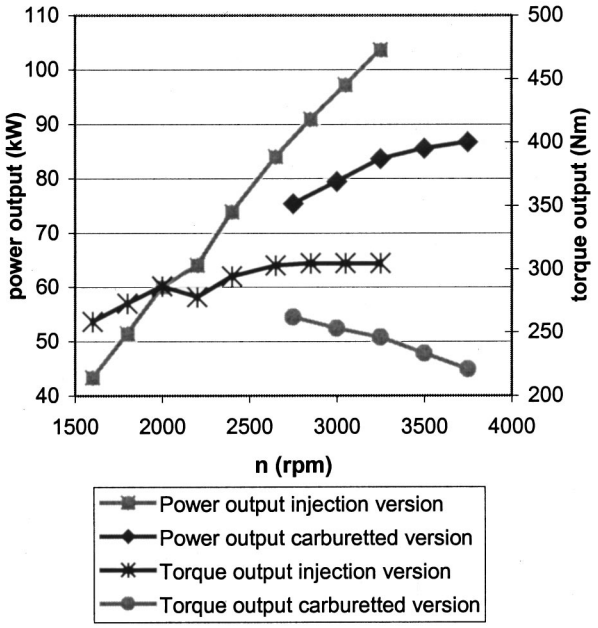


Fig. 2 Power and torque output

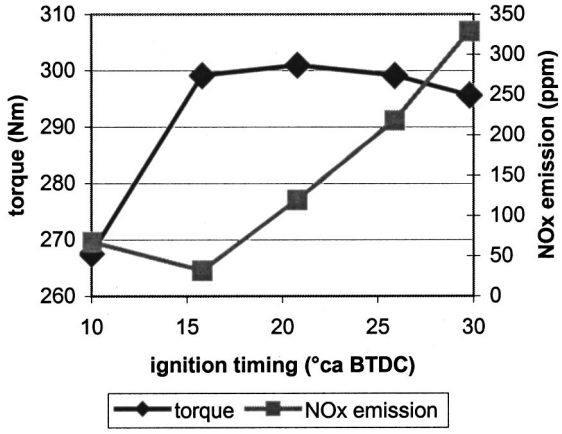


Fig. 4 Torque and NO_x emission versus IT

open throttle (WOT). For part load conditions the mixture is set leaner and leaner ($\lambda = 5$ is possible), as is done for diesel regulations, except for idling conditions.

Another possible optimization strategy is towards minimum exhaust gas pollution (second part of step 4). Although a hydrogen engine naturally is a very low emission engine, problems arose with the amount of unburned hydrogen in the exhaust gases during idle run, as will be discussed later. The possible optimization of this emission is currently being researched.

The main engine parameters suitable for optimization are the ignition timing, the injection pressure, the injection timing, and the injection duration.

When the injection pressure is raised, the power output will rise due to the higher amount of hydrogen in the engine (if injection durations are fixed). However, the possibilities of variations in injection pressure are limited according to the chosen means of storage of the hydrogen. When the hydrogen is stored in liquid form, the pressure in the cryogenic tanks is restricted. For this reason, a constant injection pressure of 3 bar was respected. In case of gaseous storage in pressurized form, it would be possible to vary the injection pressure according to the desired power output (but keeping the limitations of the air-to-fuel ratio $\lambda = 2$).

The control scheme of the motor management system is given in the Appendix. The various parts are examined in the following paragraphs.

Ignition Timing. The ignition advance is normally set to the minimum value for best torque (MBT timing). This is the compromise between a high power output (necessary due to the losses in volumetric efficiency) and a minimum ignition advance to decrease NO_x values. For the basic parameter setting ($n = 3500$ rpm, full load), as an example, the influence of the ignition timing on the torque output is given in Fig. 4.

For lean mixtures (low loads and speeds), the optimum ignition timing is early, up to 50° ca BTDC (power cycle). The engine load is the main influence. For high loads and speeds (maximum power output) the optimum ignition timing is about 20° BTDC. This is shown in Fig. 5, with the ignition timing (in ° ca BTDC) as a function of engine load (with the engine load proportional to the reading of a simulated MAP sensor, where 0 mbar represents idle conditions and 2000 mbar represents full power) and engine speed.

The efficiency of a hydrogen-fueled engine is very dependent on an optimally adjusted ignition timing as a function of the richness of the mixture (i.e., the load, as mentioned above).

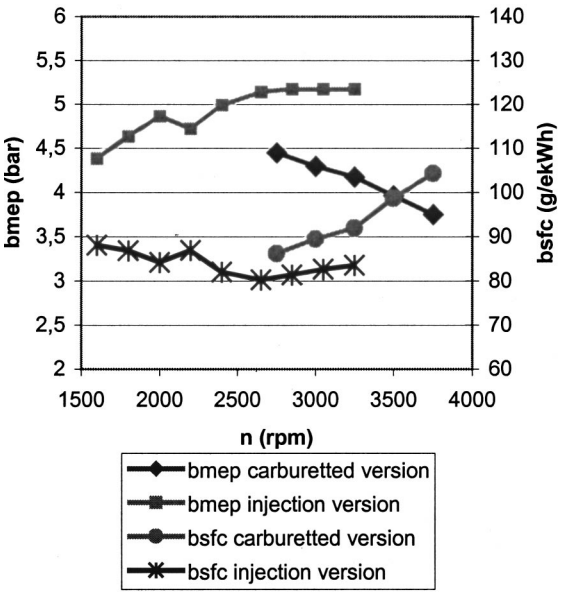


Fig. 3 Bmep and Bsfc

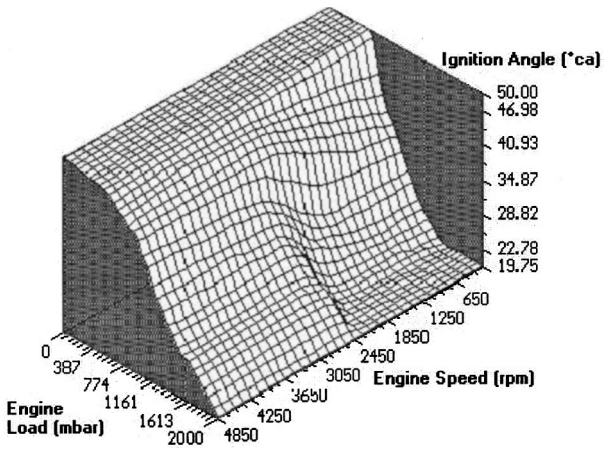


Fig. 5 Ignition map

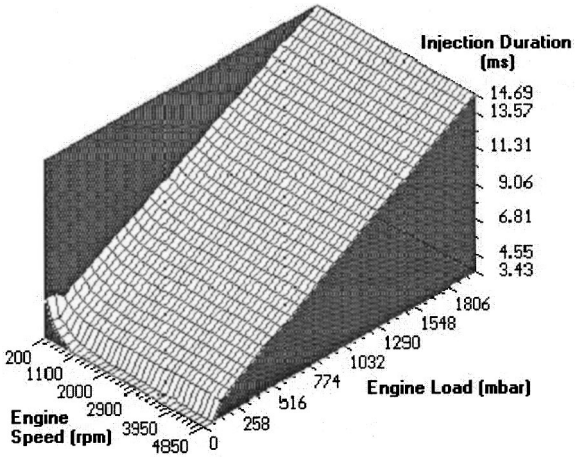


Fig. 6 Injection duration map

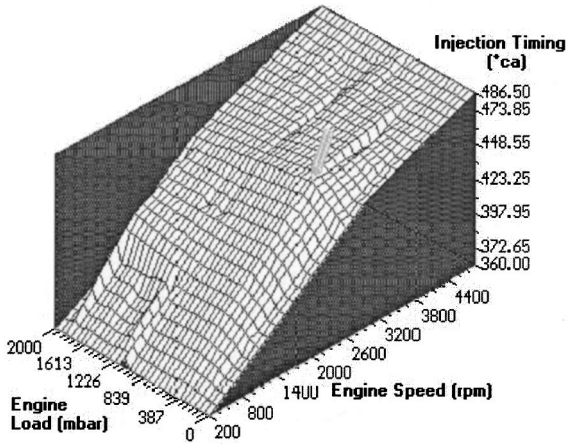


Fig. 7 Injection timing map

Figure 5 clearly shows that the influence of the load is much more important than the engine speed.

Even with this MBT timing the exhaust gases are very clean. The only noxious exhaust emission to consider for a hydrogen engine is NO_x. As an example, the influence of the ignition timing on the NO_x emissions for the conditions of Fig. 4 ($\lambda = 2$, $n = 3500$ rpm) is a minimum measured NO_x emission of 32 ppm for an ignition timing of 15.8° BTDC and a maximum of 329 ppm for the ignition timing of 29.8° BTDC. The maximum NO_x emission over the whole speed-load region was found to be about 750 ppm, occurring at a low speed, high load setting (1000 rpm, with a torque of 256 Nm). An adequate strategy for NO_x management is thus the use of MBT timing and a minimum air-to-fuel ratio of 2.

Injection Duration. The engine is operated as a diesel engine: it is a spark-ignited engine but load variations are captured through variations in the richness of the hydrogen–air mixture. As a consequence, the injection duration (in degrees crank angle) is proportional to the engine load. Thus, in idling conditions, injection durations of about 3 ms are required, corresponding to 13.5 degrees c.a. with an engine speed of 750 rpm. Under high load conditions, injection durations of up to 14 ms and more are applied, corresponding to 315 degrees c.a. with an engine speed of 3750 rpm. For comparison, the inlet valve opening time is 317 degrees c.a. A more stable idle run is reached by programming a longer injection duration when the engine speed drops below the idle speed (which allows the engine to speed up to the idle speed again) (see Fig. 6).

Injection Timing. This parameter has a great impact in the lower range of engine loads and speeds. In this region, differences in power output, by varying the injection timing, of up to 20 percent are no exception. All optimum injections start at or before TDC (gas exchange), Fig. 7, and should be advanced with speed increase. For example, during idling conditions (low speed) the injection starts at TDC and in high speed conditions the injection timing is advanced up to 105° c.a. BTDC (thus before the inlet valve opens, because of the time needed for the fuel to travel from the injector to the inlet valve, as a consequence the injection ends well before the inlet valve closes). In the higher range of engine loads and speeds, the differences in power output are still noticeable, but minimal. All injections should end before the inlet valve closes (95° c.a. after BDC).

Trims. The control system (motor management) allows corrections on the values for ignition timing and injection timing and duration as fixed in the 3D maps when the environment conditions change. Thus, changes in fuel pressure and temperature, combustion air temperature, and cooling water temperature can be automatically compensated for. The calculation of the changes in density of the hydrogen fuel as a function of the fuel's temperature and pressure is taken into account in order to apply the correct injection duration. A correction of the injection duration as a function of the combustion air intake temperature is also done.

The corrected values can differ from the programmed values by a maximum of ± 50 percent.

Other possibilities include changes in the ignition timing when the combustion air intake temperature or cooling water temperature changes. The motor management also allows the regulation of a stoichiometric mixture, but it is clear that this is not an option for a hydrogen-fueled engine.

The positions in the control scheme where the trims are applied can easily be seen in the diagram included in the Appendix.

Hydrogen-Engine-Specific Properties

Ignition Characteristics. Hydrogen under high pressure is commonly used as an insulator (e.g., in the alternator of a power plant). This results in a high ignition voltage of the hydrogen–air mixture. This is solved by choosing the spark plug gap smaller than usual in classic gasoline engines [12]. This is possible because of the smaller amount of deposits on the electrodes (only from impurities and lubricating oil). Measurements are done to define the optimal spark gap to cover the full load and speed range: testing during idle run is necessary to ensure a stable idle run, testing during full load has to be done to make sure the arc is not blown out. The tests consist of pressure measurements in cylinder nr. 1 for different spark plug gaps. Cylinder nr. 1 is the furthest cylinder of the left bank of the engine with regard to the flywheel. Thirty cycles are measured in each working point, with one sample per degree crank angle. The mean pressure curve is determined, and the mean square deviations with regard to this mean pressure curve of the measured points are calculated. The mean value of these deviations is the criterion that is used to judge the stability of the combustion: the lower this value, the more stable the combustion. The spark plug gap corresponding to the most stable combustion is considered optimal.

An optimum of 0.4 mm is found, this in comparison with the spark gap of 0.9 mm before optimization.

This previous setting of 0.9 mm is responsible for problems due to spark discharges through the air outside the cylinders. The voltage peaks on the secondary side (>40 kV) exceed the insulation possibilities of the spark plug cables, causing spark discharges between the spark plug heads and the cylinder head. These problems are completely solved with the optimized spark gap.

Lubricating Oil. During measurements of the composition of the gases in the crankcase, a very high percentage of hydrogen is noticed ($+5$ vol percent, out of range of testing equipment). The

very low density of hydrogen is responsible for this, causing high blow by volumes. The composition of the lubricating oil (semi-synthetic ''universal'' oil, viscosity class 15W50) is investigated and compared to that of the unused oil.

It appears that the properties of the oil have strongly changed with a serious decrease of the lubricating qualities. The concentration of various additives (both lubricating and wear-resisting, e.g., zincdialkyldithiophosphate) is greatly decreased, and esters appearing in the unused oil have almost completely disappeared in the used oil. These conclusions are drawn from the difference in absorption of the various elements in an infrared spectrum. This is understandable when one knows that hydrogen is used in the industry to harden oils to fats (breaking open the double C–C bonds).

The viscosity of the oil in atmospheric conditions has increased (causing more friction during starting) and decreased more quickly when the temperature rose (causing poor lubrication when the engine is at operating temperature). The kinematic viscosity at 40°C of the used oil is 141.9 mm^2/s, as compared to the value for the unused oil of 111.8 mm^2/s. At 100°C these values are respectively 14.33 mm^2/s versus 17.25 mm^2/s. The viscosity index of the used oil thus amounts to 99, substantially lower than that of the unused oil which is 163.

An x-ray fluorescence spectrometry shows no substantial engine part wear, which is normal considering the limited amount of testing time of the engine. This means that all changes of the oil characteristics are to be ascribed to the influence of the blow down gases.

Solutions to this problem are currently sought after. One possible solution is the combination of forced ventilation of the crank case, followed by an oil separator and a catalyst to convert the hydrogen to water, after which the gases can be carried off towards the atmosphere or to the intake manifold, depending on the composition of the gases after the catalyst. Copper catalysts are known to convert hydrogen to water, but research has to be done into a practical solution permitting the implementation to be built in. Another possibility is the application of special motor oils for usage in hydrogen engines. However, at the moment these oils are not available on the market (as far as the authors know).

Oxygen Sensors. Air–fuel ratios of $\lambda = 5$ and higher are no exception on this engine. However, the manufacturers of oxygen sensors consider an air-to-fuel ratio of $\lambda = 1.7$ already an extremely lean mixture. Consequently, attention must be paid to an accurate calibration of the sensors along the entire range of used richnesses. Correct calibration is necessary to ensure a correct reading of the air-to-fuel ratio, important for correct measurements as well as to be able to imply safety measures: as mentioned above, backfire-safe operation is only guaranteed if the air-to-fuel ratio is greater than or equal to 2. A lesser accuracy with lean mixtures must also be taken into account. The relation between the voltage given off by the sensor and the concentration of oxygen in the measured gases, as provided by the manufacturer, must certainly be substituted by an adjusted calibration curve (e.g., a third degree polynomial). This is because of the strong influence of a hydrogen–air mixture (as occurring with lean mixtures) on the voltage given off by the sensor.

Noise Reduction. Because of the very high noise levels of the original engine setup (up to 110 dB), tests are done with various materials and lengths of the intake pipes [13]. With metal pipes based on exhaust pipes (concentric pipes, inner pipe perforated, and damping material between the pipes) with a total intake length of 0.9 m, a noise reduction of 10 dB is reached. This means that the noise level is halved. A second benefit of this configuration is a higher torque in the working region for city-bus application (around 2000 rpm).

However, this configuration cannot be built in a bus because of the space needed and because of the rigid construction, sensitive to fatigue cracking due to engine vibrations.

Throttle Valve or Diesel Principle? The broad flammability limits of hydrogen in air (lower limit 4 percent upper limit 75 percent) allow load variations through variations in the richness of the hydrogen–air mixture, thus omitting a throttle valve. The greatest benefit is of course a better engine efficiency (no flow losses around the throttle valve).

A disadvantage is noticed when measuring the exhaust gas composition during idle run, where a high concentration of unburned hydrogen is recorded (up to 3 vol percent). The extremely lean mixtures used here are responsible for this. The high percentage of hydrogen poses a problem in connection with safety issues.

In looking for methods to solve this, throttle valves (every cylinder has its own intake) are mounted. These allow us to reduce the hydrogen percentage to about 1 vol percent. A greater reduction is probably possible with an intake manifold with one central

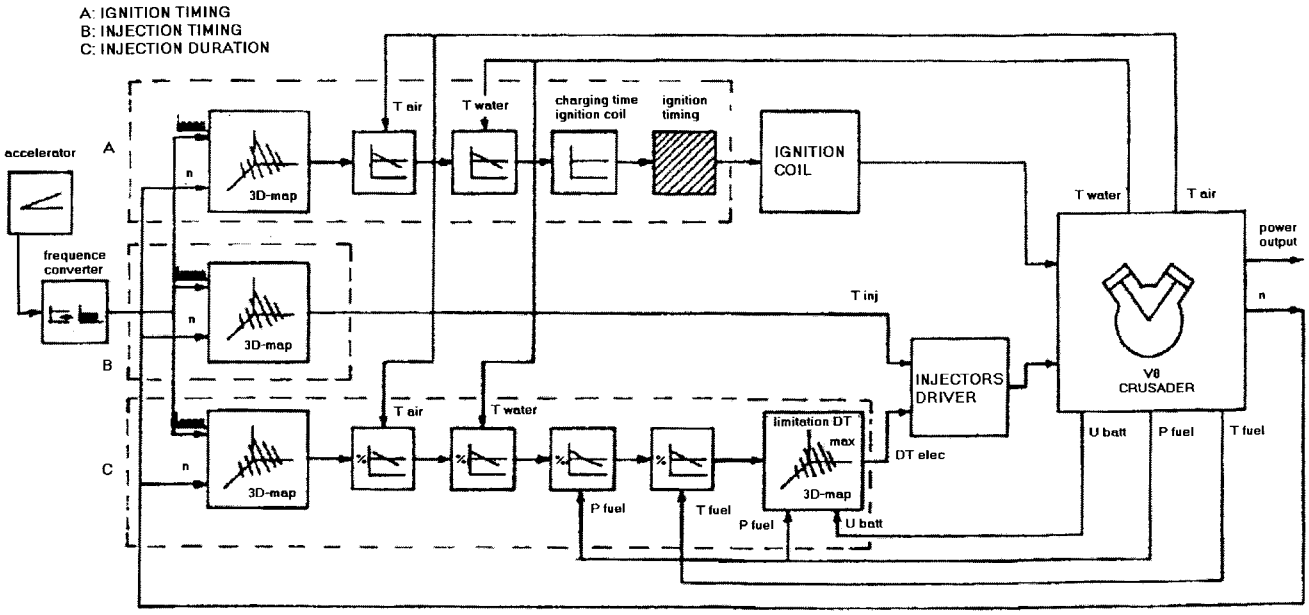


Fig. 8 Control scheme of the motor management system

throttle valve, because of the more accurate regulation possible with one big valve (mechanical problems aligning eight separate valves). Thus, a compromise could be to use a throttle valve during idle run and to use wide open throttle (WOT) in all other cases, with variation of the mixture's richness. As discussed above, the effect of the engine parameters on this emission is currently being researched.

Conclusions

- A V-8 spark-ignited engine is adapted for gaseous fuels. The first tests are done with an external mixture formation system (venturi type) for natural gas, hydrogen, and hythane.
- With a sequential timed multipoint injection of hydrogen and the corresponding electronic management system, the power output of the engine is increased without danger of backfire.
- The optimization of the engine parameters is discussed.
- Specific features of the use of hydrogen in IC engines is analyzed:
- the necessity of smaller spark plug gaps,
- the deterioration of the lubricating oil,
- additional calibration of the oxygen sensors for the extremely lean mixtures, and
- the advantage of lean mixtures to operate at low load conditions without a throttle valve, but with the disadvantage of increased hydrogen concentration in the exhaust gases at idling.

Acknowledgments

The work described in this paper is partly sponsored by the Commission of the European Union in the framework of the CRAFT action (contract BRST-CT98-5349). Other parties involved in this contract are Hydrogen Systems n.v. (BE), Vialle b.v. (NL), Trivia Technologies Int. (LU), Betronic b.v. (NL), Berkhof b.v. (NL), and CES-Continental Energy Systems (BE).

Appendix

For clarity, the control scheme of the motor management system is given here (see Fig. 8). The values for the ignition timing, the injection timing, and the injection duration are determined from three-dimensional maps, and are a function of the engine speed and the position of the accelerator. The values obtained from these 3D maps are then corrected for changes in the combustion air pressure and temperature, the fuel pressure and temperature, the cooling water temperature, and the battery voltage.

References

[1] Das, L. M., 1990, ''Fuel Induction Techniques for a Hydrogen Operated Engine,'' Int. J. Hydrogen Energy, **15**, pp. 823–842.
[2] Sorusbay, C., Veziroglu, T. N., 1988, ''Mixture Formation Techniques for Hydrogen-Fueled Internal Combustion Engines,'' *Proceedings of the 7th World Hydrogen Energy Conference*, Vol. 3, Moscow, pp. 1909–1921.
[3] Kondo, T., Hiruma, M., and Furuhama, S., 1996, ''A Study on the Mechanism of Backfire in External Mixture Formation Hydrogen Engines,'' *Proceedings of the 11th World Hydrogen Energy Conference*, Vol. 2, Stuttgart, June 23–28. pp. 1547–1556.
[4] Lee, S. T., Yi, H. S., and Kim, E. S., 1995, ''Combustion Characteristics of Intake Port Injection Type Hydrogen Fueled Engine,'' Int. J. Hydrogen Energy, **20**, pp. 317–322.
[5] Guo, L. S., Lu, H. B., and Li, J. D., 1999, ''A Hydrogen Injection System With Solenoid Valves for a Four-Cylinder Hydrogen-Fueled Engine,'' Int. J. Hydrogen Energy, **24**, pp. 377–382.
[6] Sierens, R., and Rosseel, E., 1995, ''Development of a Multi-Point Timed Injection S.I. Natural Gas Engine,'' *Natural Gas and Alternative Fuels for Engines*, Vol. 24, ASME, New York, pp. 99–104.
[7] Sierens, R., and Rosseel, E., 1995, ''Sequential Injection of Gaseous Fuels,'' 5th Int. EAEC Congress, Strasbourg, June 1995, Paper No. SIA 9506A03.
[8] Furuhama, S., 1995, ''Problems of Forecasting the Future of Advanced Engines and Engine Characteristics of the Hydrogen Injection With LH_2 Tank and Pump,'' Calvin Rice Lecture, ASME, 23 April 1995.
[9] Sierens, R., and Rosseel, E., 1998, ''Variable Composition Hydrogen/Natural Gas Mixtures for Increased Engine Efficiency and Decreased Emissions,'' ASME 1998 Spring Engine Technology Conference, Fort Lauderdale, FL, Apr. 26–29, Paper No. 98-ICE-105.
[10] Sierens, R., 1999, ''The Development of a Hydrogen Fueled V-8 Engine,'' 7th Int. EAEC Congres, Barcelona, June 1999, Paper No. STA99C435.
[11] Sierens, R., and Rosseel, E., 1998, ''Backfire Mechanism in a Carburetted Hydrogen Fueled Engine,'' *Proceedings of the 12th World Hydrogen Energy Conference*, Vol. 2, Buenos Aires, June 21–25, pp. 1537–1546.
[12] Payvey, M., 1988, ''Fuel From Water,'' Louisville, Merit Products.
[13] Van Boxlaer, L., and Poot, T., 1998, ''Optimalisatie van een inlaatsysteem voor een waterstofmotor (The optimization of the intake manifold of a hydrogen fueled engine),'' End of year thesis, Laboratory of Transporttechnology, University of Gent, Belgium.

Scroll Compressor Simulation Model

C. Schein

R. Radermacher
e-mail: rader@eng.umd.edu

Department of Mechanical Engineering,
University of Maryland,
College Park, Md 20742-3035

A computer model has been developed in Visual Basic which predicts the power, capacity, and efficiencies of a scroll compressor for given operating conditions and scroll designs. The operating conditions and data related to the compressor's geometry are inputs for the model. The model includes the effects of internal leakage and over and under compression. The compression process is broken up into a series of six degree compression steps. The work and thermodynamic state of each pocket of refrigerant are calculated as it travels from the suction to the discharge reservoir. Once the compression cycle is completed, the compressor's volumetric and isentropic efficiencies are calculated. Lastly, the work and capacity of the compressor are calculated through energy balances using the appropriate inlet and outlet refrigerant conditions. [DOI: 10.1115/1.1335483]

1 Introduction

A Visual Basic computer model was developed to simulate the performance and operation of a scroll compressor. The model is intended to calculate the following parameters:

- effects of leakage on compressor efficiencies
- primary leakage path
- effects of over and under compression on compressor efficiencies
- power
- capacity
- volumetric flow rate

The compressor model was to be used as an analytic tool to predict the performance of different positive displacement compressors. Initially, the model was developed for a scroll compressor; however, in the future it will be used to predict the performance of various types of positive displacement compressors assuming the compressor's geometry is defined. In addition to comparing the performance of different compressors, the model will eventually be used to simulate the effects of using different working fluids.

Before developing the computer model, it was necessary to understand the basic operating principles of the scroll compressor. The compressor consists of two scroll members each with a spiral shaped wrap and corresponding end plate. The two scroll members are placed at the top of the compressor housing. The top scroll member is fixed to the compressor housing while the lower member is free to orbit about the center of the fixed scroll. The moving scroll is driven by a motor and is attached to a motor shaft (see Fig. 1) [1]. As the moving scroll orbits about the fixed scroll, the wraps of the two scroll members form line points of contact. These line points of contact form crescent shaped symmetric pockets. The gas enters the first pair of symmetric pockets at the outer periphery of the scroll set and continues to travel towards the center. As the gas moves towards the center, the volume of the pockets decreases thus compressing the trapped gas. Once the gas reaches the center of the scroll set, the tip of the moving scroll begins to uncover the discharge port located in the center of the endplate of the fixed scroll member. Once the discharge port becomes uncovered, the discharge process begins (see Fig. 2) [1]. Because of the scroll's geometry, no discharge valve is needed; instead the tip of the moving scroll wrap uncovers the discharge port. It is important to note that there are always several pockets of gas being compressed at one time, and, therefore, the compression process lacks pronounced force maxima.

2 Literature Review

An extensive literature search was performed on the scroll compressor which resulted in over 65 technical papers and 334 patents. Most of the technical papers were found in the International Compressor Engineering Conference at Purdue. Most of the papers were related to compressor performance, internal leakage, multi staged designs, and scroll wrap designs. Because of the importance of internal leakage, a brief summary of the current work related to leakage prevention is presented below.

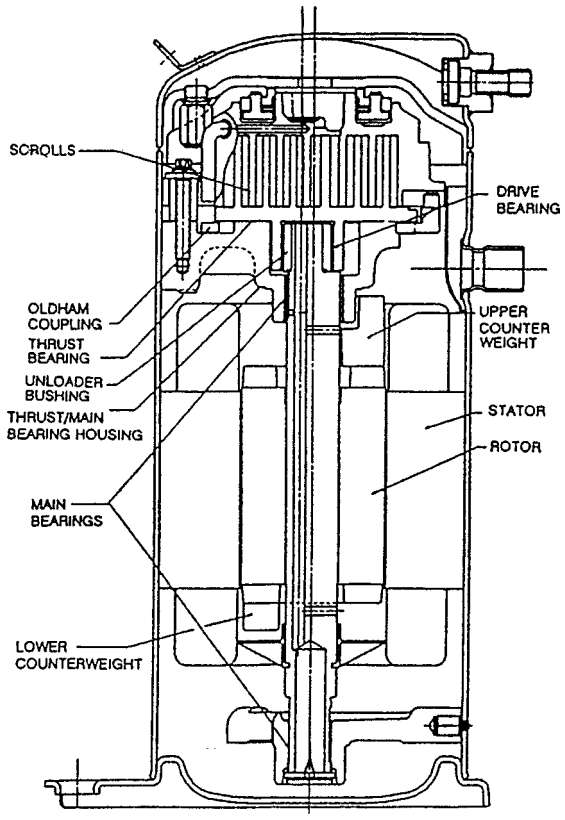


Fig. 1 Diagram of scroll compressor [1]

Contributed by the Advance Energy Systems Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received January 18, 1999; final revision received by the ASME Headquarters February 28, 2000. Technical Editor: M. J. Moran.

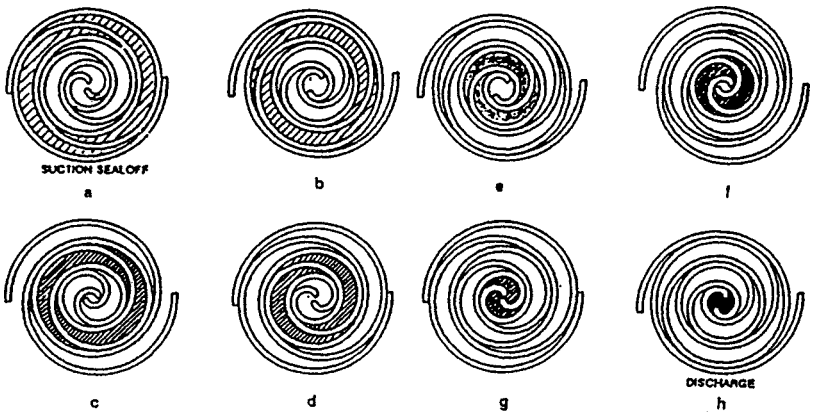


Fig. 2 Compression process [1]

The internal leakage reduces the net mass flow rate through the compressor which ultimately reduces the cooling capacity of the entire system. However, the scroll compressor has inherently less leakage than other compressors in its class. The capacity of a scroll compressor with a compression ratio of 8 is 30 percent greater than that of a reciprocating compressor [2]. One reason that the scroll compressor has a higher mass flow rate is because there is no direct re-expansion of gas from the discharge pressure to the suction pressure. In a reciprocating compressor there is a clearance volume of gas which is compressed but remains trapped inside the cylinder during the discharge process. This volume of gas then expands as the new fresh gas is entering the cylinder. This type of re-expansion does not occur in the scroll compressor. In the scroll compressor there is no direct path for the discharge gas to mix with the suction gas.

Many efforts have been made to reduce internal compressor leakage [3–5]. One method is to implement sealing elements at the locations of the leakage paths. A second approach is to use a back pressure chamber of gas which can be located above the endplate of the fixed scroll or below the endplate of the orbiting scroll [6]. This back pressure chamber would then communicate with one of the compression pockets. It is possible to calculate the force needed to ensure proper sealing through a force balance of all the gas forces [7].

Multi stage compressor designs have also been found to reduce leakage [8]. This two stage design consists of three scroll members all stacked at the top of the compressor housing. Two of the scroll members are identical to conventional scroll compressors such that each consists of an endplate and a single spiral shaped wrap. The third scroll member consists of an endplate and a spiral shaped wrap protruding through each side of the endplate. The dual sided scroll member is located in between the other two scroll members thus forming a ''sandwich'' like configuration. Gas enters into the first scroll set and then is discharged into the suction of the second scroll set. Therefore, any leaked gas which mixes with fresh suction gas occurs from the first compressor stage rather than from both compressor stages. The leakage back to the suction reservoir of the second stage mixes with gas which has been previously compressed rather than the fresh suction gas. The claim is that this two stage design reduces the leakage back to the suction reservoir which therefore produces a higher volumetric efficiency.

A final approach to reducing leakage is through oil injection [9]. Oil injected into the compression pockets helps to fill gaps where leakage would otherwise occur. However, this design also requires an oil-refrigerant separator in order to ensure that the excess oil does not enter into the system. The lower internal leakage allows for a higher mass flow rate of refrigerant and thus a larger cooling capacity of the system. Results of similar studies have shown that oil injection can increase the volumetric efficiency by 4.3 percent as compared to designs which do not use oil injection techniques [10].

3 Computer Model: Architecture

The architecture of the model is discussed below including information regarding the software, model inputs and outputs.

3.1 Software. The software used to develop the thermodynamic model was MATLAB, EXCEL/Visual Basic and REFPROP [11]. The actual computer code was written in Visual Basic, while the user input, summary, geometry, and graph sheets were displayed in EXCEL.

3.2 Model Inputs. All user inputs were found on the inputs page and are shown in Table 1. A separate geometric model was written in MATLAB and the results were imported into the geometry page of the thermodynamic model [12]. The geometry file contained information such as pocket volumes and pocket perimeters for every six degrees of rotation. The equations used to generate the curve of the scroll wrap were based on the involute of a circle. This circle is often referred to as the base circle. The scroll geometry can be understood if one were to consider an imaginary string unwrapped from the circumference of the circle. As the string is unwrapped its starting point will form the shape of a spiral. The pitch or distance between turns of the scroll is equal to the circumference of the base circle. Likewise, the entire geometry of the scroll wrap can be defined from the geometry of the base circle [12].

In order to replicate the performance of a realistic compressor, a 5-ton, R22 commercially available scroll compressor was disassembled. Internal measurements were taken including the height, thickness and pitch of the scroll members. This data was then used to generate the geometry file of the thermodynamic model.

3.3 General Description of Code. Once all the user inputs were defined, a Visual Basic code was written to determine the

Table 1

System Inputs	Compressor Inputs
Refrigerant type	Motor speed
Saturated evaporator temperature	Scroll height
Saturated condenser temperature	Scroll thickness
Evaporator Superheat	Radial clearance gap
Condenser Subcooling	Axial clearance gap

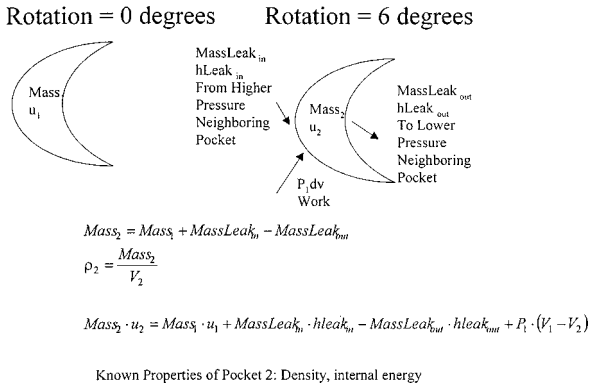


Fig. 3 Schematic of energy and mass balance on pocket

state of each pocket of vapor for every six degrees of rotation. A series of mass and energy balances were performed on each pocket of the rotation starting with the suction process and ending with the discharge process. The model follows a single pocket moving from the outer periphery of the scroll set to the most inner pocket. The model steps through the compression and discharge process for every six degrees of rotation. Although there is a time step associated with the six degree rotation, the equations are based on a steady state assumption. It is for this reason that the model is considered a quasi-steady-state compressor simulation.

For the initialization process, energy and mass balances are performed on each of the pockets starting with the suction process and ending with the discharge process. When calculating the energy balance, the change in internal energy for each pocket was simply equal to the integration of PdV. At the completion of the initialization run there existed pressure gradients between adjacent pockets. From the model initialization, the state of each pocket during the compression process was known. Because a pressure gradient and clearance existed, it was possible to calculate the leakage flow that would occur between the neighboring pockets. A mass balance was performed which accounted for the leakage into and out of a pocket as shown in Eq. (1).

$$\text{Mass}_2 = \text{Mass}_1 + \text{MassLeak}_{\text{in}} - \text{MassLeak}_{\text{out}} \tag{1}$$

Subscript 1 denotes the original mass inside the pocket. Subscript 2 denotes the mass in the pocket after the effects of the leakage are applied. The terms MassLeak refer to the net change of mass due to leaked refrigerant into and out of the pocket of interest. This calculation of MassLeak is discussed in the Section 4.2. Similarly an energy balance was performed on each pocket which accounted for the energy associated with the leakage as shown below.

$$\text{Mass}_2 \cdot u_2 = \text{Mass}_1 \cdot u_1 + \text{MassLeak}_{\text{in}} \cdot h\text{leak}_{\text{in}} - \text{MassLeak}_{\text{out}} \cdot h\text{leak}_{\text{out}} + P_1 \cdot (V_1 - V_2) \tag{2}$$

An illustration of this equation is shown in Fig. 3. One assumption made in using this equation was that the value of the pressure in the work term remains constant for every six degrees of rotation.

The fluid properties of the leakage including the enthalpy were taken from pocket from which it leaked, i.e., the higher pressurized pocket. Once the energy and mass balances were calculated the internal energy and specific volume of the pocket were the only two fluid properties known. The specific volume is determined from the Mass_2 term of Eq. (1), and the physical volume of the pocket determined from the geometry of the compressor. Using the internal energy term from Eq. 2, and the specific volume, it should have been possible to determine all of the thermodynamic properties of the gas inside the pocket. However, because of software limitations, it was not possible to use internal energy as one of the two properties needed to define the thermodynamic

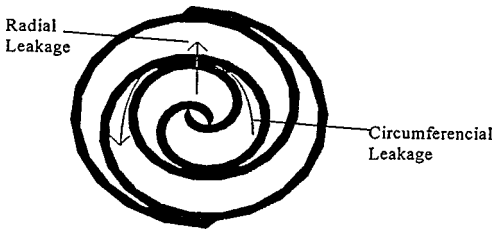


Fig. 4 Diagram of internal leakage path

state of the gas inside the pocket. Because the software could only use one of the two calculated properties a third property was needed. The enthalpy was approximated by applying a linear expansion of the enthalpy around a nearby point as shown below.

$$h = h_{\text{nearby}} + \left(\frac{dh}{dp}\right)_v \cdot (P - P_{\text{nearby}}) \tag{3}$$

Plotting enthalpy versus pressure for various specific volumes of refrigerant gas (see Fig. 5) validated the linear expansion. As shown from Fig. 5, the curves associated with each density or specific volume could be represented by a linear approximation. The error associated with the linear expansion was calculated by considering the deviation in the internal energy values. A comparison was made between the internal energy value calculated from Eq. (2) and the value of the internal energy associated with the enthalpy from the linear expansion. The error associated with this deviation of internal energy was at most .01 percent for each six degrees of compressor rotation. A second comparison was made between the difference in internal energy values calculated from Eq. (2) and that associated with Eq. (3). When considering the difference in the internal energy terms between two pockets throughout the entire compression process, the error of the linear expansion was between three to five percent.

3.4 Model Outputs. Once the model converged, the model will output the following results based on the user specified system inputs and the geometry of the compressor.

1 volumetric, isentropic, and motor efficiencies
2 compressor work
3 mass flow rate
4 compressor capacity
5 thermodynamic state of refrigerant within the compressor

The volumetric efficiency was calculated as shown in Eq. (6). The isentropic efficiency was calculated from the definition of isentropic efficiency as shown below.

$$\eta_{\text{isentropic}} = \frac{(h_{\text{outi}} - h_{\text{in}})}{(h_{\text{out}} - h_{\text{in}})} \tag{4}$$

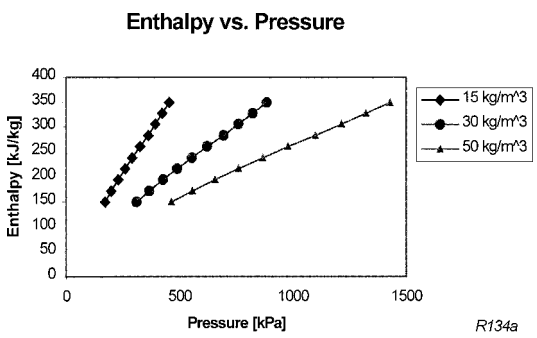


Fig. 5 Validation of linear expansion

The denominator represents the actual change in enthalpy across the compressor while the numerator was calculated from the isentropic enthalpy change across the compressor. Lastly, the motor efficiency was calculated from a motor curve supplied by the compressor manufacturer. The motor curve was valid up to 4000 RPM. For speeds greater than this limit, a 70 percent motor efficiency default value was assumed in the model.

The mass flow rate was determined from the fixed volumetric flow rate of the compressor and the volumetric efficiency. The volumetric flow rate can be calculated from the speed of the motor and the actual volume filled by the first set of symmetric pockets or the displacement volume. The final result is shown in the equation below.

$$Q = \text{MotorSpeed} \cdot \text{Displacement\_Volume} \tag{5}$$

The mass flow rate is calculated by multiplying the density of the fresh suction gas by the volumetric flow rate, Q, and the volumetric efficiency. The mass flow rate is used to calculate the cooling capacity of the compressor. The user specifies the system operating conditions on the inputs sheet, and a vapor compression cycle analysis is performed in order to determine the load on the evaporator. Similarly, the compressor power is calculated by performing an overall energy balance on the compressor.

4 Details of Model

The details of the model are addressed in the following sections, which include internal leakage, over and under compression, and the model convergence criterion.

4.1 Internal Leakage. One of the major performance losses associated with a positive displacement compressor is internal leakage. Internal leakage occurs whenever a gap exists between regions of different pressures. There are several reasons that internal leakage degrades the performance of the compressor. One reason is the displacement of fresh gas into the compressor's suction reservoir as a result of the space or volume taken up by the leaked gas. In addition to reducing the quantity of fresh refrigerant, the leakage process also increases the temperature of the gas prior to the compression process. This temperature increase is a result of the mixing of the fresh suction gas with leaked gas. The density and mass flow rate of the gas are reduced as a result of the temperature increase. The reduced mass flow rate will ultimately decrease the compressor's cooling capacity. The performance impacts associated with internal leakage are measured by the volumetric efficiency. Volumetric efficiency was defined in the model as shown in the equation below.

$$\eta_{\text{volumetric}} = \frac{\text{Mass}_{\text{real}}}{\text{Mass}_{\text{ideal}}} \tag{6}$$

$\text{Mass}_{\text{ideal}}$ refers to the mass of refrigerant which would enter the first compression pocket from the suction reservoir if there were no internal leakage. $\text{Mass}_{\text{real}}$ is defined as the actual mass of refrigerant inside the first pocket minus the mass of the leakage back to the suction reservoir. In addition to degrading the volumetric efficiency, the re-expansion and re-compression associated with the leakage process degrades the compressor's isentropic efficiency.

A mixing process must occur between the fresh incoming gas and the gas that has leaked in from the higher pressurized pockets. The control volume used for this energy analysis was taken as the inside of the suction reservoir. Two assumptions were made in this mixing process analysis. The first was that the mixing process occurs at constant pressure namely the suction pressure. A second assumption was needed to determine the mass of the fresh suction gas entering into the compressor. This quantity of gas was defined as shown in the equation shown below.

$$\text{Mass}_{\text{FreshGas}} = \rho_{\text{FreshGas}} \cdot V_{\text{pocket1}} \tag{7}$$

Using this assumption the mass inside the control volume was calculated as shown below.

$$\text{Mass}_{\text{mix}} = \text{Mass}_{\text{FreshGas}} + \sum \text{Mass}_{\text{Leak}} \tag{8}$$

An energy balance was performed on the control volume to determine the mixed enthalpy of the gas mixture that enters the first compression pocket. The equation for the energy balance is shown below.

$$h_{\text{mix}} = \frac{H_{\text{FreshGas}} + \sum H_{\text{Leak}}}{\text{Mass}_{\text{mix}}} \tag{9}$$

In the scroll compressor there are two primary leakage paths namely, radial and circumferential leakage. Radial leakage occurs at the gap found in between the top of the scroll wrap and the end plate of the second scroll member (see Fig. 4). The second leakage path, circumferential leakage, occurs as a result of the clearance found between the two mating scroll walls. This type of clearance is found at the beginning and end of each pocket (see Fig. 4).

4.2 Leakage Model Assumptions. Various fluid flow models have been used to simulate internal leakage including nozzle flow, compressible flow, CFD, and incompressible pipe flow [13,14]. The flow model selected for this model was one-dimensional pipe flow with the governing equation listed below.

$$\Delta P = \frac{f \rho L V^2}{2 D_h} \tag{10}$$

In order to use the equation above, it was necessary to define the flow regime. In the model, the flow was defined as turbulent, incompressible, and fully developed.

As mentioned above, the model was based on a turbulent assumption in determining the friction factor used in Eq. (10). To validate the turbulent assumption, the velocity of the leakage was calculated using both a turbulent and a laminar correlation for the friction factor. The velocity was then used to calculate the Reynolds number. For some cases applying the laminar correlation for the friction factor yielded a turbulent Reynolds number. Likewise, applying the turbulent correlation for the friction factor yielded a laminar Reynolds number. Because of this discrepancy, it was suggested that the leakage was in a transition regime of flow [15,16]. The two Reynolds numbers from the turbulent and laminar friction factor correlation were plotted versus the compressor's angle of rotation. The intersection point of the two Reynolds number curves was taken as the cutoff point between laminar and turbulent flow. The intersection point occurs within less than the first quarter of rotation of the compressor and therefore it was assumed that the majority of the leakage flow path was turbulent. The turbulent flow friction factor correlation was taken from a technical paper [13,17] related to leakage in scroll compressors. The correlation is shown below.

$$f = 0.35 \cdot \text{Re}^{-0.25} \tag{11}$$

This correlation was validated by comparing it to the correlation derived from an established two-layer model for turbulent flow [18]. Because of the similarity between the two correlations, it was decided to use Ishii's correlation because it was applied to test data taken from internal leakage.

For simplicity, the leakage flow was treated as incompressible flow. In general when the Mach number of the leakage flow is greater than 0.3, the effects of compressibility should not be neglected. It was evident from the model results that for some cases the Mach number was greater than 0.3. These findings suggest that more work is needed in this area.

Finally, it was assumed that the flow was fully developed. A flow can be defined as fully developed once it has passed its

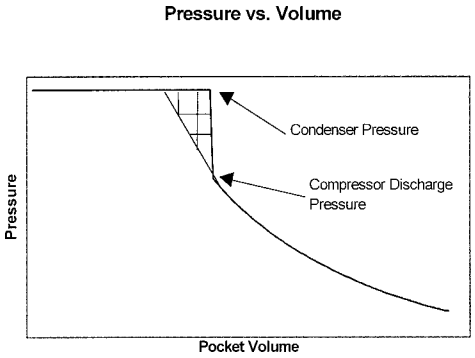


Fig. 6 Diagram of under compression

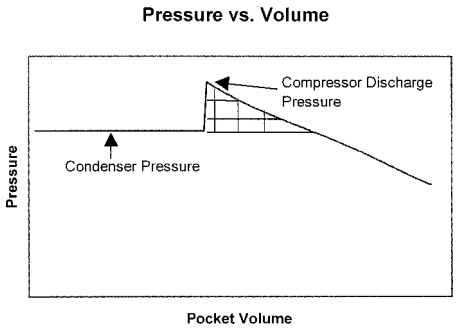


Fig. 7 Diagram of over compression

entrance length and the velocity profile no longer changes. The entrance length was calculated for turbulent flow using the equation below [19].

$$\mathrm{Le} = 4.4 \cdot \mathrm{Re}^{1/6} \tag{12}$$

The entrance length was calculated for each increment of the compressor and compared to the length of the leakage path. It was found that the length of the leakage path was 3.5 times greater than the calculated entrance length, and, therefore, the leakage flow was treated as fully developed.

Based on the assumption of fully developed, turbulent, incompressible flow the velocity of the leakage flow was calculated using Eq. (10). The velocity was multiplied by the density and cross sectional area of the leakage path to calculate the mass flow rate of the leakage.

4.3 Over and Under Compression. Based on the geometry of the scroll compressor, there exists a fixed volume ratio which is defined in the equation below.

$$\mathrm{Vol}_{\mathrm{ratio}} = \frac{V_1}{V_2} \tag{13}$$

V_1 and V_2 correspond to the volume of the pockets at the beginning and end of the compression process. The pressure ratio can be derived from the volume ratio if the compression process is assumed to be polytropic. This relationship is shown below.

$$\frac{P_2}{P_1} = \left(\frac{V_1}{V_2}\right)^{\gamma} \tag{14}$$

From Eq. (14), it is evident that the pressure ratio is a function of the fixed volume ratio of the compressor. Therefore, when the system pressure ratio, which is defined as the ratio of the saturated condenser pressure to the saturated evaporator pressure, does not match that of the compressor there is a loss in performance. This mismatch in pressure ratio leads to two possible cases, over compression and under compression (see Figs. 6 and 7). It is important to re-emphasize the fact that there are no valves which open at the end of the compression process. Instead the discharge process is initiated by the uncovering of the discharge port.

Under compression occurs when the pressure ratio of the compressor is less than the system pressure ratio. For the case of under compression the pressure of the final compression pocket is less than that of the discharge reservoir. In order for the discharge reservoir and the final compression pocket to be in equilibrium, the pressure inside the final pocket must increase. This pressure increase is achieved at the instant the discharge port becomes uncovered and there is a sudden back flow of gas from the discharge reservoir into the final compression pocket. The back flow of gas continues until the pressure inside the final compression pocket is equal to that of the discharge reservoir. An energy balance is performed to determine the thermodynamic state of the gas which actually exits the compressor.

$$\mathrm{Mass}_1 \cdot u_1 + \mathrm{Mass}_2 \cdot h_2 = (\mathrm{Mass}_1 + \mathrm{Mass}_2) \cdot u_2 \tag{15}$$

Equation (15) can be expressed in words as the internal energy of the final compression pocket plus the enthalpy of the gas which flows into the pocket from the discharge reservoir is equal to the new adjusted internal energy of the final compression pocket. Through simple manipulation, Eq. (15) can be re-written to directly solve for the enthalpy of the gas inside the discharge reservoir.

$$h_2 = h + (P_2 - P_1) \cdot \frac{V_1}{\mathrm{Mass}_1} \tag{16}$$

One of the penalties associated with under compression is the extra work used to compress the gas. As shown in Fig. 6, the compressor design discharge pressure is lower than that of the gas inside the condenser. As a result of this pressure ratio mismatch refrigerant gas flows back into the final compression pocket (backflow). In the model the backflow phenomenon occurs within one six degree-rotation step of the compressor. It should be noted that the actual duration of time for this pressure equalization process is unknown. As shown in Fig. 6, the hatched area under the curve illustrates the extra work associated with under compression.

Analogous to the case of under compression is the case of over compression. Over compression occurs when the pressure of the final compression pocket is greater than that of the condenser. In order for the gas inside the discharge reservoir (at condenser pressure) to be in equilibrium with the gas in the final compression pocket there must be a drop in pressure inside the final compression pocket. This pressure drop is achieved by gas flowing from the final compression pocket into the discharge reservoir. The gas continues to flow until the gas inside the final compression pocket is equal to that of the discharge reservoir. The discharge process can not begin until there is equalization between the gas inside the discharge reservoir and the gas inside the final compression pocket. An energy balance similar to Eq. (15) is performed in order to determine the state of the gas exiting the compressor.

Similar to the case of under compression, there is a work penalty associated with over compression. For the case of over compression, the pressure of the gas inside the condenser is lower than the gas inside the final compression pocket. However because the scroll compressor is a fixed volume ratio machine, it will compress the gas to its design point regardless of the high end pressure of the system. As a result, extra work is used to compress the gas than would be needed if there were a match between the system and the compressor's compression ratio. The extra work is illustrated on a pressure volume diagram in Fig. 7. The triangular shaped region adds to the area under the pressure volume curve which is used to calculate compressor work.

4.4 Model Convergence. Once the mass of the leakage was calculated, the mass and energy balances were re-calculated for each pocket until the model converged for all pockets. The program iterates until the change in pressure for each pocket for the current and previous runs is below a set convergence criterion. A

comparative analysis was performed on the change of pocket pressure with the number of iterations. From this analysis it was decided to use a convergence criteria of 100 Pa such that the model would iterate until the pressure difference between the same two pockets of consecutive runs was less than 100 Pa.

4.5 Model Summary. The sequence and coupling of equations used in the model is described below. The steps below summarize the sequence of equations used in the model starting with the initialization of the model and ending with the final convergence of the model.

4.5.1 Model Initialization

1 The inputs related to the compressor's geometry and the state of the gas entering into the compressor are read into the model from the spreadsheet.

2 The pocket volumes and leakage areas for every 6 degrees of rotation are dimensioned into an array.

3 Setting the cross sectional areas of the leakage paths equal to zero (no leakage) initializes the model.

4 The mass and enthalpy of the first pocket of compression, V_1, are calculated using Eqs. (7), (8), and (9) (the leakage terms drop out as a result of previous step).

5 The next value from the pocket volume array is retrieved, V_2.

6 Using the volume and thermodynamic properties from V_1, an energy and mass balance are performed on V_2 using Eqs. 1 and 2.

7 The enthalpy of pocket V_2 is approximated using Eq. (3).

8 The pressure of pocket V_2 was calculated from the approximated enthalpy, internal energy, and specific volume.

9 Steps 4–6 are repeated for every six degrees of rotation until the thermodynamic state of the final compression process has been calculated.

4.5.2 Model Simulation

10 After the initialization process is complete, the thermodynamic state of each pocket is known.

11 The cross sectional areas of the radial and circumferential leakage paths are no longer set to zero, but instead are set to the values supplied on the geometry page of the spreadsheet.

12 Because a pressure gradient and gap exists between neighboring pockets, it is possible for radial and circumferential leakage to occur.

13 The leakage mass flow rate is calculated for each of the neighboring pockets using Eq. (8) and the continuity equation.

14 The mass and enthalpy of the first pocket of compression, V_1, are calculated using Eqs. (7), (8), and (9).

15 The next value from the pocket volume array is retrieved, V_2.

16 Using the volume and thermodynamic properties from the previous pocket, an energy and mass balance are performed on pocket V_2 using Eqs. (1) and (2).

17 The enthalpy for pocket V_2 is approximated using Eq. (3).

18 The pressure of pocket V_2 is calculated from the approximated enthalpy, internal energy, and specific volume.

19 Steps 16–18 are repeated for every six degrees of rotation until the thermodynamic state of the final compression process has been calculated.

20 The pressures calculated for each pocket are compared to the corresponding pressures from the previous run.

21 If the difference in pressures between corresponding pockets is greater than 100 Pa, the simulation process is repeated starting with step 10.

22 If the difference in pressure between corresponding pockets is less than 100 Pa, the model has converged.

4.5.3 Model Simulation: Discharge Process

23 A logic statement is used to determine if the discharge pressure is above (over compression) or below (under compression) the condensing pressure.

24 Gas flows from the discharge reservoir into or out of the compressor depending on whether the compressor has over or under compressed the gas. Equation (16) is used to determine the enthalpy of the gas exiting the compressor after the necessary mixing process has occurred.

4.5.4 Model Simulation: Final Calculations

25 The isentropic efficiency, the volumetric efficiency, are calculated using Eqs. (4) and (6), respectively.

26 The volumetric flow rate is calculated using Eq. (5). The mass flow rate is also calculated using the density of the suction gas, and the volumetric efficiency and flow rate.

27 The motor efficiency is calculated from a curve fit supplied by the vendor. The motor efficiency is a function of the motor speed.

28 The work is calculated using two different approaches to ensure that the model results are valid. The work is calculated by an energy balance and by calculating the area under the pressure volume curve.

29 The capacity of the compressor is calculated using user-defined inputs and the calculated mass flow rate.

5 Model Results: Parametric Studies

After completing a detailed literature search, the computer model was developed and tested. A series of parametric studies were performed related to clearance gap sizes, operating conditions, and motor speed. A base case was designed for comparison, which is shown below.

- saturation temperature of the evaporator: −1.11°C (30°F)
- saturation temperature of the condenser: 24.4°C (76°F)
- superheat: 5.6°C (10°F) and subcooling: 4.3°C (7.7°F)
- the motor speed: 3500 RPM (100 percent efficiency)
- the leakage clearance: 0 (no leaks)
- refrigerant: R134a

The saturation temperatures of the evaporator and condenser were selected so that the pressure ratio across the system would match that of the compressor.

The first set of studies was related to the effects of over and under compression. All the parameters were held fixed in the base case except the condenser saturation temperature. The saturated condenser temperature was varied from 10°C (50°F) to 37.8°C (100°F). As shown in Fig. 8, the isentropic efficiency increases as the pressure ratio approaches that of the compressor design. Similarly, the isentropic efficiency decreases as the system pressure ratio increases beyond that of the compressor design. From Fig. 8, it is clear that the slopes of the curve before and after the cross

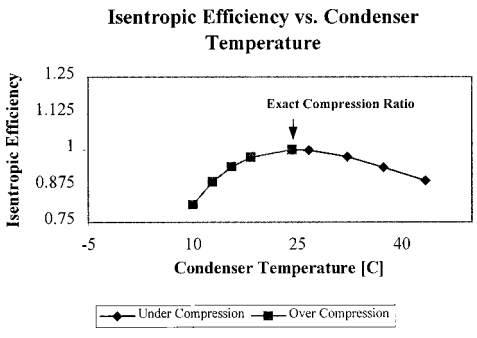


Fig. 8 Effects of over and under compression on performance

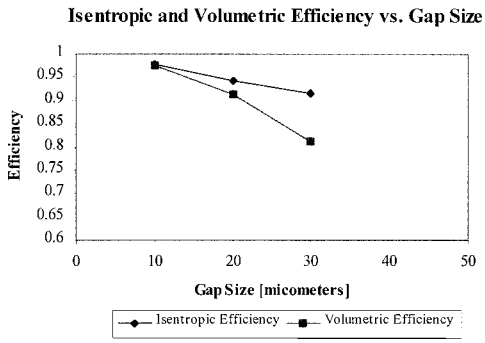


Fig. 9 Effects of gap size on compressor efficiency

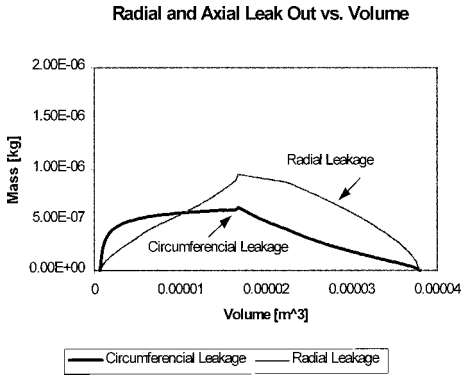


Fig. 10 Comparison of leakage out of a pocket

over point are not the same. Hence, the decrease in isentropic efficiency occurs more rapidly for over compression than for the under compression cases. As a result of the lower isentropic efficiency, the compressor work increases at a faster rate for over compression than for under compression. Based on these model results, it can be concluded that the work penalty for over compression is greater than for under compression.

The second set of studies related to the effects of the leakage clearance size. The parameters were set to the base case values except for the leakage clearance. The gap size was varied from 10 to 30 micrometers which were typical values found in the literature. As the gap size increased, the volumetric and isentropic efficiencies decreased. The decrease in isentropic efficiency results in an increase in compressor work. The compressor work increases as a result of the higher discharge pressures associated with the larger gap sizes. Intuitively it makes sense that as the clearance size increases more gas can leak from the higher pressure pockets and mix with the lower pressure pocket. While the leakage of gas to the lower pressure pocket is small due to small pressure difference, the pressure inside the pocket after the mixing process would be greater than before the mixing process. Therefore, as the mixing process continues for each pocket along the compression process, the final pressure at the end of the compression process is greater with increased leakage. Likewise the decrease in volumetric efficiency results from the fact that more gas will leak back into the suction reservoir as the gap size is increased. The leaked gas mixes with the fresh gas and thus decreases its density. In addition to decreasing the density of the suction gas, the leaked gas also displaces volume to be occupied by the fresh suction gas.

Although both efficiencies decreases with increasing gap sized it is interesting to compare the rates at which the efficiencies decreased. The volumetric efficiency decreased at a more rapid rate than the isentropic efficiency with the increase in the leakage clearance (see Fig. 9). This result implies that for the set operating conditions, the effects of displaced suction volume by the leaked gas dominates over the effects associated with the expansion and re-compression of the leaked gas.

In addition to performing sensitivity studies on the size of the leakage clearance, it was also desired to know which leakage path dominates for a fixed clearance size. The gap size was set to 20 micrometers and the mass of the radial and circumferential leaks were plotted vs. pocket volume. As can be seen from Fig. 10, the radial leakage is greater than the circumferential leakage over approximately 70 percent of the compression process. The radial leakage is greater initially because the cross sectional area for this type of leakage is greater than for the circumferential leakage. Following the pocket towards the center of the scroll set, the perimeters decrease and circumferential leakage dominates.

Upon completion of the single parameter variations, the next step was to look at the combined effects of under compression and leakage. For this study, the saturated condenser temperature was set to 37.4°C (100°F) and the gap size was varied from 10–30 micrometers. As compared to the single parametric study of under compression (no leakage), the volumetric efficiency dropped from 100 percent to 77.8 percent. Similarly, the isentropic efficiency decreased from 91.8 percent to 83 percent. From this finding it was concluded that the leakage process effects the volumetric efficiency more than the isentropic efficiency even when the two events (under compression and leakage) are combined. The results were also compared to the single parametric study of the size of the leakage clearances. When these two cases were compared, the volumetric efficiency dropped from 81.1 percent (exact compression ratio) to 77.8 percent for the combined case. Similarly, the isentropic efficiency decreased from 91.7 percent for the exact compression case to 83 percent for the combined case. This result implies that the effects of under compression are more sensitive to the isentropic efficiency than the volumetric efficiency.

The combination of leakage and under compression also impacts the compressor work. The compressor work increases with increase in gap size for the under compressed case. One way to visualize the effects of under compression and leakage on compressor work is to plot a pressure volume curve for the compression process. As shown in Fig. 11 the effects of under compression are evident by the vertical line at the completion of the compression process. Also from Fig. 11, it is evident that the work or area under the curve increases as the gap size increases. The pressure and mass inside each pocket increased including the pocket of the final compression pocket. This pressure increase resulted in an increase in work. More gas must leak into the pockets than leaks out because there is an increase in the mass inside each pocket with the larger gap sizes.

In addition to analyzing the case of under compression with internal leakage, the case of over compression with internal leakage was also studied. For this parametric study, the condenser temperature was set so that the system pressure ratio was less than that of the compressor design. The leakage gap was varied from 10–30 micrometers. As was expected, as the gap size increased

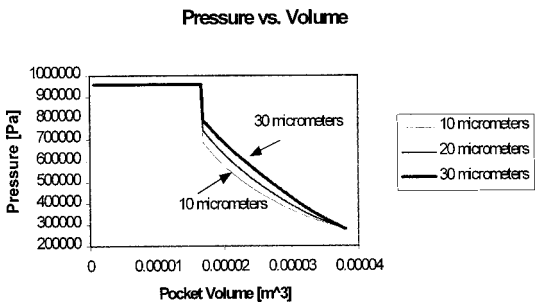


Fig. 11 Pressure volume curve for case of under compression and leakage

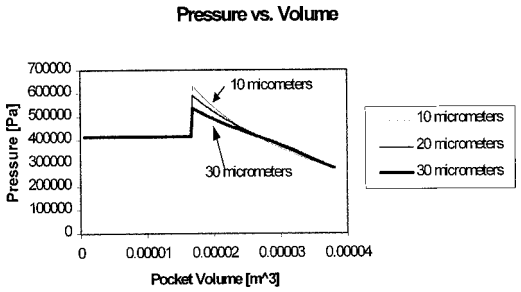


Fig. 12 Pressure volume curve for case of leakage and over compression

the volumetric efficiency decreased. Unlike the previous case, the isentropic efficiency increased with the larger gap sizes. This unexpected result occurs because the change in enthalpy across the compressor decreases with the increase in the gap size. The decrease in enthalpy change is due to the fact that the discharge pressure decreases as the gap size increases. As the discharge pressure decreases, the pressure ratio of the compressor approaches that of the system pressure ratio. The pressure volume curve associated with this study was plotted and is shown in Fig. 12.

The question then arises why does the discharge pressure decrease with increasing gap size. From the single parametric study of gap size variation, it was concluded that as the leakage increases the pocket pressure increases as a result of the mixing process with the leaked gas. One reason the pressure decreases is because with larger gap sizes there is less mass inside the final compression pockets which implies that more must leak out than leaks into these pockets. A ''reverse leakage'' phenomenon that occurs in the model explains this unpredicted result. At the beginning of the compression process the cases with larger gap sizes have higher corresponding pocket pressures. At the middle to end of the compression process, the pressure inside the pockets begins to decrease for the larger gap sizes. For the case of over compression, the gas pressure inside the final compression pockets can be greater than the gas which is about to be expelled from the inner portion of the scroll set. Therefore, the gas in these pockets can leak into the outer periphery pockets AND into the inner portion of the scroll set. Because of the two leakage paths, the mass of gas inside the pockets at the end of the compression process decreases with larger gap sizes. This phenomenon of reverse leakage has not yet been measured or verified so more testing is needed in this area.

6 Conclusion

In conclusion a detailed computer model was developed to predict the performance of a scroll compressor including its efficiencies, power, capacity, and volumetric flow rate. Some of the major findings obtained from the results of the computer model and the literature review are listed below.

- The scroll compressor consists of two spiral shaped scroll wraps which when mated together form successive crescent shaped pockets. Through the orbiting motion of the moving scroll, the volumes of vapor filled pockets are continuously reduced and the vapor becomes compressed.
- The two primary leakage paths inside the scroll compressor are radial and circumferential leakage.
- Radial leakage dominates for most of the compression process while circumferential leakage dominates at the end of the process.
- For the case of under compression more gas leaks into a pocket than leaks out.
- For the case of over compression gas leaks from one pocket into the outer periphery pockets and into the inner pockets of the scroll set (reverse leakage).
- The size of the clearances associated with internal leakage influence the volumetric efficiency more than the isentropic efficiency.
- The effects of over and under compression influence the isentropic efficiency more than the volumetric efficiency.
- Isentropic efficiency decreases at a more rapid rate for over compression than for the case of under compression.
- Over and under compression increase the work of the compressor as compared to a compressor with a pressure ratio equal to that of the system.

Nomenclature

ρ = density
f = friction factor
η_{motor} = motor efficiency
γ = polytropic exponent
ΔP = pressure difference
$\dot{m}$ = refrigerant mass flow rate
$\eta_{\text{isentropic}}$ = isentropic efficiency
$\eta_{\text{volumetric}}$ = volumetric efficiency
D_h = hydraulic diameter
h_1, h_2 = specific enthalpy before and after an event
$h_{\text{in}}, h_{\text{out}}$ = specific enthalpy of vapor entering and exiting compressor
H_{leak} = enthalpy of leakage back to suction chamber
$h\text{Leak}_{\text{in}}$ = specific enthalpy of refrigerant leaking into a pocket
$h\text{Leak}_{\text{out}}$ = specific enthalpy of refrigerant leaking out of a pocket
h_{mix} = specific enthalpy of gas mixture in suction chamber
h_{nearby} = specific enthalpy of nearby point for linear expansion
Le = entrance length
$\text{Mass}_1, \text{Mass}_2$ = mass of refrigerant inside pocket
$\text{Mass}_{\text{FreshGas}}$ = mass of fresh gas entering into the compressor
$\text{Mass}_{\text{ideal}}$ = ideal mass of refrigerant in the first compression pocket
$\text{Mass}_{\text{leak}}$ = mass of refrigerant leakage back to suction reservoir
$\text{MassLeak}_{\text{in}}$ = mass of refrigerant leaking into a pocket
$\text{MassLeak}_{\text{out}}$ = mass of refrigerant leaking out of a pocket
Mass_{mix} = assumed mass of mixture
$\text{Mass}_{\text{real}}$ = actual mass of refrigerant in the first compression pocket
P_1, P_2 = pressure before and after an event
P_{nearby} = pressure of nearby point for linear expansion
Power = compressor power
Q = volumetric flow rate
Re = Reynolds number
RPM = revolutions per minute
u_1, u_2 = specific internal energy of pocket before and after an event
V = velocity of leakage
V_1, V_2 = volume of pocket before and after an event
$\text{Vol}_{\text{ratio}}$ = fixed volume ratio of compressor
V_{pocket1} = volume of first compression pocket

References

[1] American Society of Heating, Refrigeration, and Air Conditioning (ASHRAE), 1988, *Equipment Handbook*, Atlanta, GA.
[2] Inaba, T., Sugihara, M., Nakamura, T., Kimura, T., and Morishita, E., 1986, ''A Scroll Compressor With Sealing Means and Low Pressure Side Shell,'' *Proceedings of the 1986 International Compressor Engineering Conference at Purdue*, Vol. 3, pp. 887–900.
[3] Chen, Z., Qiao, Z., and Xiam J., 1994, ''The Influence of Leakage on the Performance of Scroll Compressor With Self Adjusting Back Pressure,'' *Pro-*

ceedings of the 1994 International Compressor Engineering Conference at Purdue, Vol. 1, July 19–22, pp. 211–216.
[4] Lee, J., and Kim, S., 1996, ''Investigation of Axial Compliance Mechanism in Scroll Compressor,'' *Proceedings of the 1996 International Compressor Engineering Conference at Purdue*, Vol. 2, July 23–26, pp. 459–464.
[5] Qian, Z., and Xian J., 1992, ''Back Pressure Mechanism of Scroll Compressor,'' *Proceedings of the 1992 International Compressor Engineering Conference at Purdue*, Vol. 5, July 14–17, pp. 1149–1156.
[6] Tojo, K., Ikegawa, M., Maeda, N., Machida, S., Shiibayashi, M., and Uchikawa, N., 1986, ''Computer Modeling of Scroll Compressor With Self Adjusting Back Pressure Mechanism,'' *Proceedings of the 1986 International Compressor Engineering Conference at Purdue*, Vol. 3, Aug., pp. 872–886.
[7] Ikegawa, M., Sato, E., Tojo, K., and Arai, A., 1984, ''Scroll Compressor with Self Adjusting Back Pressure Mechanism,'' ASHRAE Trans., No. 2846.
[8] Shibamoto, Y., 1994, ''*Scroll Compressor of Two Stage Compression Type Having an Improved Volumetric Efficiency*,'' U.S. Patent Number: 5,304,047, Date of Patent: April 19, 1994.
[9] Drost, R., and Debois, R., 1996, ''Scroll Compressor Performance With Oil Injection/Separation,'' *Proceedings of the 1996 International Compressor Engineering Conference at Purdue*, Vol. 1, July 23–26, pp. 329–334.
[10] Huiqing, L., Dishceng, W., Huanan, W., and Penaggo, C., 1992, ''Research of Oil-Injected Scroll Compressor Working Process,'' *Proceedings of the 1992 International Compressor Engineering Conference at Purdue*, Vol. 1, July 14–17, pp. 118b1–118b14.
[11] REFPROP/Version 5.0 National Institute of Standards and Technology (NIST), U.S. Department of Commerce, Gaithersburg, MD 20899.
[12] Lindsay, D., 1997, private contractor.
[13] Ishii, N., Bird, K., Sanno, K., Oono, M., Sanno, K., and Iwamura, S., 1996, ''Refrigerant Leakage Flow Evaluation for Scroll Compressors,'' *Proceedings of the 1996 International Compressor Engineering Conference at Purdue*, Vol. 2, July 23–26, pp. 633–638.
[14] Huang, Y., 1992, ''Leakage Calculation Through Clearances,'' *Proceedings of the 1992 International Compressor Engineering Conference at Purdue*, Vol. 1, July 19–22, pp. 35–40.
[15] Dimas, A., 1997, Sept. 17, University of Maryland College Park, College Park, MD.
[16] Kiger, K., 1997, Sept. 17, University of Maryland College Park, College Park, MD.
[17] Ishii, N., Sakai, M., Sano, K., Yamamoto, S., and Otokura, T., 1996, ''A Fundamental Optimum Design for High Mechanical and Volumetric Efficiency of Compact Scroll Compressors,'' *Proceedings of the 1996 International Compressor Engineering Conference at Purdue*, Vol. 2, July 23–26, pp. 639–644.
[18] Kays, W. M., Crawford, M. E., 1993, *Convective Heat and Mass Transfer*, McGraw-Hill, New York, pp. 75–85 and pp. 193–252.
[19] White, F., 1986, *Fluid Mechanics*, 3rd ed., McGraw-Hill, New York.

Determination of Propane and Air Maximum Mixing Times

D. Brasoveanu

A. K. Gupta
e-mail: akgupta@eng.umd.eud

The Combustion Laboratory,
Department of Mechanical Engineering,
University of Maryland,
College Park, MD 20742

The model of gaseous fuel and air mixing, developed by the authors, is applied here to calculate maximum mixing times of propane and air. The degree of mixing is determined using the mass fraction of fuel while the rate of mixing is determined from the rate of this mass fraction. The values of both these parameters are local, i.e., measured within an infinitesimal element of fluid. A Eulerian representation is used. The model is based on the assumption that both fuel and air behave as a single chemical species. It is further assumed that pressure is low and only fuel and air are present within the fluid element. Under nonreacting conditions, the model is valid anywhere in the combustor. Under reacting conditions, the model is valid within those combustor regions where the fuel–air mixture is not flammable. The results of this analysis show that mixing times of propane and air are most reduced by high gradients of temperature and velocity, as long as these gradients provide in phase contribution. To a lesser degree, high gradients of pressure also help reduce mixing times. High initial pressure and temperature increase mixing time. Mixing with air penetration into the fuel flow is slower than with propane dispersion into the surrounding air. In general, the exact mixing time has to be determined numerically. Nevertheless, the analytical solutions included here provide maximum mixing times of propane and air under most conditions. These results provide important guidelines for the development of high intensity, high efficiency, and low emission combustors.
[DOI: 10.1115/1.1338946]

Introduction

A model that unifies the treatment of mechanical and diffusive mixing of fuel and air is needed in order to solve the practical problems of combustor design [1]. The authors have presented such a mixing model in their previous publication [2–7]. The model is enhanced and applied here to examine the mixing of propane and air.

The model quantifies the degree and rate of mixing within an infinitesimal element of fluid using the local mass fraction of fuel and the rate of this fraction, i.e., the total derivative of fuel mass fraction with respect to time, respectively. The fluid element is analyzed in an Eulerian frame of reference. The model uses the ideal gas law and is therefore limited to combustors operating at low pressures. It is further assumed that fuel and air behave like a single chemical species and other chemical species are present only in trace amounts within the infinitesimal fluid element. Chemical reactions do not occur within the mixture element being analyzed but may occur elsewhere in the combustor and subsequently change the local temperature.

The Model

Assume a fluid element contains only fuel and air. The local mass fraction of fuel within the element, based on the ideal gas law [2,6], is given by:

$$y_f = \frac{\rho_f}{\rho} = \frac{M_f}{M_f - M_a}\left[1 - M_a \frac{P}{RT\rho}\right] \tag{1}$$

The total derivative of Eq. (1) with respect to time provides the rate of mixing, dy_f/dt, (and therefore defines mixing in a straightforward manner [8]), i.e.,

$$\frac{dy_f}{dt} = \frac{M_f M_a}{M_a - M_f}\frac{1}{M}\left(\frac{1}{P}\frac{dP}{dt} - \frac{1}{T}\frac{dT}{dt} + \operatorname{div}\vec{U}\right) \tag{2}$$

Because Eq. (2) combines both mechanical and diffusive mixing in a single step [6,7] dy_f/dt is the overall rate of mixing as sought by Guyon et al. [1]. The total derivative of pressure with respect to time dP/dt will be called rate of pressure. The total derivative of temperature with respect to time dT/dt will be called rate of temperature. The total derivative of density with respect to time $d\rho/dt$ will be called rate of density. The rate of pressure can be written as:

$$\frac{dP}{dt} = \frac{\partial P}{\partial x}u + \frac{\partial P}{\partial y}v + \frac{\partial P}{\partial z}w + \frac{\partial P}{\partial t} \tag{3}$$

The rate of temperature can be determined in a similar manner. The equation of continuity [8] show that the total derivative of fluid density with respect to time $d\rho/dt$ is given by

$$\frac{d\rho}{dt} = -\rho\vec{\nabla}\cdot(\vec{U}) \tag{4}$$

Equation (2) also shows that mixing is due to rates of pressure (i.e., total derivative of pressure with respect to time), rates of temperature, or velocity divergence, whenever the model assumptions are applicable.

The rates of pressure and temperature are correlated with gradients of pressure and temperature, respectively. The overall velocity divergence is correlated with velocity gradients. The overall velocity includes both flow and diffusion velocity. The molecular weight of fuel, initial pressure, temperature, and molecular weight of the mixture also affects the rate of mixing. The molecular weight of mixture M can be determined from:

$$M = x_f M_f + (1 - x_f)M_a \tag{5}$$

where the mole fraction of fuel x_f is given by:

$$x_f = \frac{n_f}{n} = \frac{m_f}{M_f}\frac{M}{m} = \frac{y_f}{M_f}M$$

Therefore, Eq. (5) can be written as:

$$M = M_a/[1 + y_f(M_a/M_f - 1)] \tag{6}$$

Contributed by the Fuels & Combustion Technology Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the FACT Division, Jan. 31, 2000; final revision received by the ASME Headquarters, Feb. 15, 2000. Associate Editor: S. R. Gollhalli.

Based on Eqs. (2) and (4), the rate of fuel mass fraction can be written as:

$$\frac{dy_f}{dt}=\frac{M_f}{M_a-M_f}\left[1+y_f\left(\frac{M_a}{M_f}-1\right)\right]\left(\frac{1}{P}\frac{dP}{dt}-\frac{1}{T}\frac{dT}{dt}+\text{div}\,\overline{U}\right) \tag{7}$$

and therefore,

$$\frac{dy_f/dt}{1+y_f\left(\frac{M_a}{M_f}-1\right)}=\frac{M_f}{M_a-M_f}\left(\frac{1}{P}\frac{dP}{dt}-\frac{1}{T}\frac{dT}{dt}+\text{div}\,\overline{U}\right) \tag{8}$$

By definition, the average value of a parameter $\kappa(t)$ over a time duration t is given by:

$$\langle\kappa(t)\rangle\equiv\frac{1}{t}\int_0^t\kappa(\iota)d\iota \tag{9}$$

From Eqs. (8) and (9):

$$\int_0^t\left[1+y_f(\iota)\left(\frac{M_a}{M_f}-1\right)\right]^{-1}\frac{dy_f}{d\iota}=at \tag{10}$$

where

$$a=\frac{M_f}{M_a-M_f}\left(\left\langle\frac{1}{P}\frac{dP}{dt}\right\rangle-\left\langle\frac{1}{T}\frac{dT}{dt}\right\rangle+\langle\text{div}\,\vec{U}\rangle\right)$$

For any gaseous fuel, the fuel-lean mixing time t_1 can be calculated using Eq. (10), i.e.,

$$t_l=\frac{\ln\left[\frac{|y_f(t_l)+M_f(M_a-M_f)^{-1}|}{|M_f(M_a-M_f)^{-1}|}\right]}{\langle P^{-1}dP/dt\rangle-\langle T^{-1}dT/dt\rangle+\langle\text{div}\,\vec{U}\rangle} \tag{11}$$

Equation (10) also provides the fuel-rich mixing t_r:

$$t_r=\frac{\ln\left[\frac{|y_f(t_r)+M_f(M_a-M_f)^{-1}|}{|1+M_f(M_a-M_f)^{-1}|}\right]}{\langle P^{-1}dP/dt\rangle-\langle T^{-1}dT/dt\rangle+\langle\text{div}\,\vec{U}\rangle} \tag{12}$$

For propane and air, Eq. (11) becomes:

$$t_l=\frac{-0.011}{\langle P^{-1}dP/dt\rangle-\langle T^{-1}dT/dt\rangle+\langle\text{div}\,\vec{U}\rangle} \tag{13}$$

and Eq. (12) becomes:

$$t_r=\frac{0.3749}{\langle P^{-1}dP/dt\rangle-\langle T^{-1}dT/dt\rangle+\langle\text{div}\,\vec{U}\rangle} \tag{14}$$

Two mixing mechanisms that can be distinguished are the penetration of air into the fuel flow and the dispersion of fuel into the surrounding air [2–6]. The times required to obtain a flammable mixture within fuel and within the surrounding air are, called the fuel-rich mixing time t_r and fuel-lean mixing time t_l, respectively [5]. Propane–air mixtures reach the upper and lower flammability limit when the mass fraction of fuel is 0.1382 and 0.0317, respectively [8]. Therefore, t_r is the time required to reduce the fuel mass fraction from 1 to 0.1382 within the fuel flow, while t_l is the time required to raise the fuel mass fraction from 0 to 0.0319 within the surrounding air.

In general, the average logarithmic derivatives of pressure, $\langle P^{-1}dP/dt\rangle$, and temperature, $\langle T^{-1}dT/dt\rangle$, and the average velocity divergence, $\langle\text{div}\,\vec{U}\rangle$, depend on mixing times. Therefore, the calculation of exact fuel-rich and fuel-lean mixing times involves numerical iteration. Nevertheless, if the range of velocity divergence and logarithmic rate of pressure and temperature are known, Eqs. (13) and (14) provide analytical solutions for maximum mixing times. Assume two variables, w and z, so that they obey the following constraints: $w>0$, $z>0$, and $\delta z\geqslant 0$. Then

$$\frac{w}{z}\geqslant\frac{w}{z+\delta z}$$

Therefore, whenever the distributions of pressure, temperature, and velocity within the fuel flow obey the following restrictions:

$$d(\ln P)/dt\geqslant[d(\ln P)/dt]_{\min}>0$$

$$dT/dt\leqslant 0$$

$$\text{div}\,\vec{U}\geqslant 0$$

the fuel-rich mixing time t_r is smaller than

$$t_r\leqslant\frac{0.3749}{[d(\ln P)/dt]_{\min}} \tag{15}$$

For distributions of pressure, temperature, and velocity that satisfy the conditions:

$$dP/dt\geqslant 0$$

$$d(\ln T)/dt\leqslant[d(\ln T)/dt]_{\max}<0$$

$$\text{div}\,\vec{U}\geqslant 0$$

the fuel-rich mixing time t_r is smaller than:

$$t_r\leqslant\frac{0.3749}{-[d(\ln T)/dt]_{\max}} \tag{16}$$

Distributions of pressure, temperature, and velocity, as given below,

$$dP/dt\geqslant 0$$

$$dT/dt\leqslant 0$$

$$\text{div}\,\vec{U}\geqslant(\text{div}\,\vec{U})_{\min}>0$$

provide the fuel-rich mixing time limited by:

$$t_r\leqslant\frac{0.3749}{(\text{div}\,\vec{U})_{\min}} \tag{17}$$

Within air, distributions of pressure, temperature, and velocity that obey the following conditions:

$$d(\ln P)/dt\leqslant[d(\ln P)/dt]_{\max}<0$$

$$dT/dt\geqslant 0$$

$$\text{div}\,\vec{U}\leqslant 0$$

provide a fuel-lean mixing time t_l that is smaller than

$$t_l\leqslant\frac{-0.011}{[d(\ln P)/dt]_{\max}} \tag{18}$$

For distributions of pressure, temperature, and velocity that satisfy the conditions:

$$dP/dt\leqslant 0$$

$$d(\ln T)/dt\geqslant[d(\ln T)/dt]_{\min}>0$$

$$\text{div}\,\vec{U}\leqslant 0$$

the fuel-lean mixing time t_l is less than

$$t_l\leqslant\frac{-0.011}{[d(\ln T)/dt]_{\min}} \tag{19}$$

Finally, if

$$dP/dt\leqslant 0$$

Table 1 The range of pressure, temperature, and velocity distributions within the mixing region

Minimum value of logarithmic rate of pressure within fuel, $[d(\ln P)/dt]_{min}$ (s^{-1})	0 to 10
Maximum value of logarithmic rate of pressure within air, $[d(\ln P)/dt]_{max}$ (s^{-1})	-10 to 0
Maximum value of logarithmic rate of temp. within fuel, $[d(\ln T)/dt]_{max}$ (s^{-1})	-100 to 0
Minimum value of logarithmic rate of temp. within air, $[d(\ln T)/dt]_{min}$ (s^{-1})	0 to 100
Minimum velocity divergence in fuel, $(\text{div } \vec{U})_{min}$ (s^{-1})	0 to 100
Maximum velocity divergence in air, $(\text{div } \vec{U})_{max}$ (s^{-1})	-100 to 0

$$dT/dt \geq 0$$

$$\text{div } \vec{U} \leq (\text{div } \vec{U})_{max} < 0$$

then

$$t_l \leq \frac{-0.011}{[\text{div } \vec{U}]_{max}} \tag{20}$$

The mixing model presented here is applicable to any combustor geometry and operational conditions, as long as pressure does not exceed 10 atm.

Diffusion Modeling

Previous mixing studies [8] isolate convective and diffusion transport for each species. For example, the ordinary diffusion law is written as [9]:

$$\frac{d(n_i/n)}{dt} = D\Delta\left(\frac{n_i}{n}\right) \tag{21}$$

where, D (s^{-1}) is the diffusion coefficient, n_i is the number of moles of specie i, n is the total number of moles, and t is time (s).

The diffusion law is considered in conjunction with the conservation law of mass, species, momentum, and rates of energy and chemical reactions in a complex system using partial differential equations. Usually, the system cannot be solved without some approximations [8] which restrict the applicability of the resulting model. As an example, a typical approximation is:

$$\frac{d(\rho_i/\rho)}{dt} \approx \frac{1}{\rho}\frac{d\rho_i}{dt}$$

where ρ is the mixture density and ρ_i is the density of specie i.

The inclusion of this approximation produces a diffusion model that is adequate only for trace species. In addition, a complex model is more difficult to use. To simplify the mixing model, diffusion should be included in another manner. The total number of moles present in a gaseous mixture of fuel and air is:

$$n = n_f + n_a \tag{22}$$

where n_f is the number of moles of fuel and n_a is the number of moles of air.

Using the ideal gas law:

$$\frac{P}{\rho} = \frac{RT}{M} \tag{23}$$

The molecular weight of mixture (kg/mole) M is:

$$M = \frac{n_f}{n}M_f + \frac{n_a}{n}M_a \tag{24}$$

If pressure, temperature, and total density are constant, the molecular weight of mixture is constant, see Eq. (23). Equation (22) and (24) can be written as:

$$\frac{n_f}{n} + \frac{n_a}{n} = 1$$

$$\frac{n_f}{n}M_f + \frac{n_a}{n}M_a = M$$

Therefore,

$$\frac{d(n_f/n)}{dt} + \frac{d(n_a/n)}{dt} = 0 \tag{25}$$

$$M_f\frac{d(n_f/n)}{dt} + M_a\frac{d(n_a/n)}{dt} = \frac{dM}{dt} = 0 \tag{26}$$

From Eqs. (25) and (26)

$$M_f\frac{d(n_f/n)}{dt} + M_f\frac{d(n_a/n)}{dt} = 0 \tag{27}$$

$$M_f\frac{d(n_f/n)}{dt} + M_a\frac{d(n_a/n)}{dt} = 0 \tag{28}$$

Subtracting Eq. (27) from Eq. (28):

$$(M_a - M_f)\frac{d(n_a/n)}{dt} = 0 \tag{29}$$

Assuming that fuel and air molecular weights are different Eq. (29) shows that:

$$\frac{d(n_a/n)}{dt} = 0$$

Substituting this result in Eq. (25) yields

$$\frac{d(n_f/n)}{dt} = \frac{d(n_a/n)}{dt} = 0$$

i.e., any diffusion process is linked to a rate of pressure, temperature, or overall velocity divergence. These rates are in turn correlated with pressure and temperature gradients, see Eq. (3). In other words, the unified model implicitly includes all types of diffusion.

Numerical Approach

Maximum values of fuel-rich mixing time have been calculated using Eqs. (15)–(17). Maximum values of fuel-lean mixing time were calculated using Eqs. (18)–(20). The range of pressure, temperature, and velocity distributions within the mixing region that were considered here is given in Table 1. Such distributions are characteristic of low pressure combustors. Calculations were performed using MATHCAD 6.0.

Results and Discussion

Results are provided here to show the effect of distributions of pressure, temperature, and velocity on propane and air mixing times. Figure 1 shows the effect of pressure distribution on maximum values of both fuel-rich and fuel-lean mixing times. The left curve shows the fuel-lean mixing time while the right curve shows the fuel-rich mixing time. Both of these mixing times are inversely proportional to the logarithmic rate of pressure. The pressure distribution facilitates mixing with fuel dispersion into the surrounding air and air penetration into the fuel flow

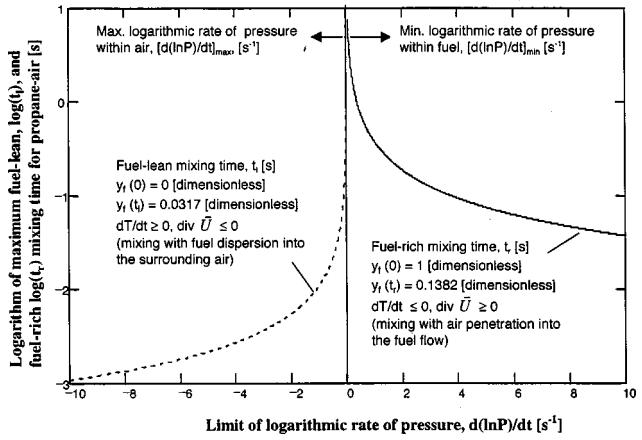


Fig. 1 The effect of pressure distribution on both fuel-lean and fuel-rich mixing times

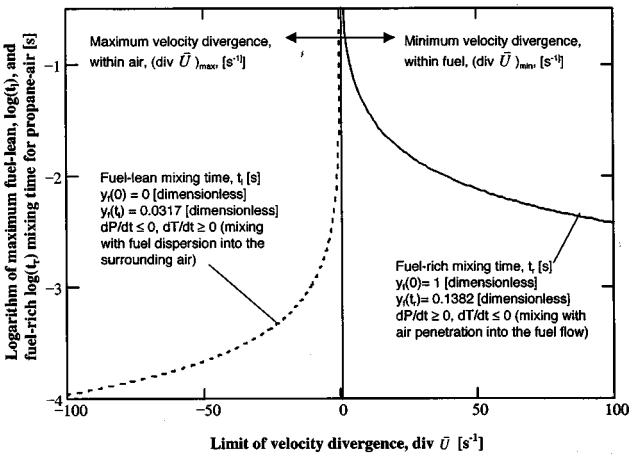


Fig. 3 The effect of velocity distribution on fuel-lean and fuel-rich mixing times

whenever $d(\ln P)/dt<0$ within air and $d(\ln P)/dt>0$ within fuel. For negligible temperature and velocity gradients, both mixing times are inversely proportional to the rate of pressure dP/dt and increase with high initial pressure. Therefore, both mixing times tend to infinity when the rate of pressure tends to zero. For similar rates of pressure, fuel-lean mixing times are about 20 times smaller than fuel-rich mixing times. In most combustors, pressure is quite constant. Only high amplitude and high intensity sound waves are expected to cause logarithmic rates of pressure that exceed 1 s^{-1} within fresh mixture. Normal rates of pressure do not reach 100,000 Pa/s and would provide fuel-lean and fuel-rich mixing times that exceed 0.01 and 0.2 s, respectively. Typical gas residence times range from about a millisecond in a gas turbine combustor to few seconds in furnaces. For enhanced mixing, residence time should exceed mixing time by at least an order of magnitude, i.e., mixing times should be less than 0.002 in gas turbine combustors and 0.2 s in furnaces. As shown in Fig. 1, usual rates of pressure may enhance fuel dispersion in furnaces, but not in gas turbine combustors. Mixing with air penetration into the fuel that is linked to pressure gradients is too slow for both gas turbine and furnace applications. Similar results have been noted for methane–air mixtures [3,5].

Figure 2 shows the effect of temperature distribution on both fuel-lean and fuel-rich mixing times. The left curve shows the fuel-rich mixing time, while the right curve shows the fuel-lean mixing time. For negligible rate of pressure and velocity divergence, both these mixing times are inversely proportional to the rate of temperature. Both mixing times tend to infinity when the rate of temperature tends to zero and increase at high initial temperature. Fuel cooling promotes air penetration into the fuel flow. Air preheating promotes fuel dispersion into the surrounding air. Therefore, propane preheating enhances both mixing mechanisms. For fuels that are lighter than air, air preheating is recommended [5]. Analogous to pressure gradients, temperature gradients provide fuel-lean mixing times that are 20 times smaller than the fuel-rich mixing times. In furnaces, logarithmic rates of temperature having a magnitude of 10 or more can enhance both fuel dispersion into the surrounding air and air penetration into the fuel. For turbine gas combustors, logarithmic rates of temperature that exceed 80 provide fuel-lean mixing times of less than 0.0002 s. Comparable rates of temperature provide fuel-rich mixing time of more than 0.001 s.

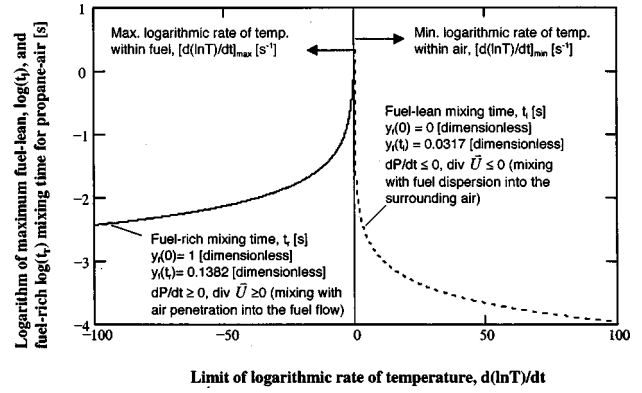


Fig. 2 The effect temperature distribution on both fuel-lean and fuel-rich mixing times

The effect of velocity divergence on both mixing times is shown in Fig. 3. The left curve shows the fuel-lean mixing time while the right curve shows the fuel-rich mixing time. High velocity divergence reduces both mixing times. Air pressure may be decreasing (i.e., $dP/dt\leqslant 0$) and fuel pressure may be increasing (i.e., $dP/dt\geqslant 0$). Air and fuel may also be subject to heating and cooling, respectively. As indicated before, the fuel-rich mixing times are shown to be about 20 longer than the fuel-lean mixing times for similar distributions of velocity. As an example, a velocity divergence having a magnitude of 50 s^{-1} provides fuel-rich and fuel-lean mixing time of about 0.01 and 0.0002 s, respectively, thereby enhancing mixing with fuel dispersing into the surrounding air, but not with air penetrating into the fuel. To achieve comparable fuel-rich and fuel-lean mixing times, velocity gradients should be about 20 greater within fuel than air. Therefore, mixing with fuel dispersion into the surrounding air is the primary mixing mechanism. Operating conditions should be designed to enhance fuel dispersion. For example air, and not propane should be subject to negative velocity divergence and heating. To enhance the secondary mixing mechanism (i.e., air penetration into the fuel) propane should be subjected to positive velocity divergence and cooling. Other combinations of velocity divergence and rates of temperature reduce mixing and should be avoided.

The limit of both mixing times is inversely proportional to velocity divergence. Mixing times tend to infinity when velocity divergence tends to zero and vice versa. As shown in Figs. 1–3 mixing times tend to infinity whenever the velocity divergence and both rates of pressure and temperature are negligible. These results are valid for all gaseous fuels having a molecular weight different than air.

The effect of initial pressure and temperature on flammability limits is small [8]. Higher initial temperature increases the upper flammability limit and reduces the lower limit somewhat. As a consequence actual fuel-lean and fuel-rich mixing times are re

duced. Nevertheless, maximum mixing times do not change. A higher initial pressure does not change the fuel-lean mixing time. It reduces the fuel-rich mixing time slightly. Again, both maximum mixing times remain constant.

Figures 1–3 provide a priori information on the effect of distributions of pressure, temperature, and velocity whether the mixture would result as ''flammable mixture'' within a given time. Advance information on these distributions can reduce the detailed Computational Fluid Dynamics (CFD) simulations and/or laborious experimental studies for designing high intensity, high efficiency and low emission combustors.

Conclusions

The results obtained from the model show that maximum mixing times are inversely proportional to velocity divergence, logarithmic rate of pressure, or logarithmic rate of temperature, whenever mixing is primarily due to gradients of velocity, pressure, or temperature, respectively. Fuel dispersion is the primary mixing mechanism and therefore, should be enhanced whenever possible. Rates of pressure may provide enhanced mixing in furnaces. For gas turbine combustors, mixing times of less than 0.0002 s are desired. Only air heating at a logarithmic rate of 50 s^{-1} and velocity divergence of less than -50 s^{-1} enhance mixing in gas turbine combustors. High initial temperature will result in reduced mixing. Therefore, the higher the initial temperature, the more important velocity divergence is for mixing enhancement. Air heating should be combined with negative velocity divergence in order to enhance mixing. Under reacting conditions, mixing takes place in the vicinity of flame, and both fuel and air are subject to heating. Propane preheating is recommended in order to enhance both mixing mechanisms.

Acknowledgments

The support of this research by NASA Glenn Research Center, DoE via the South Carolina Energy Research & Development Center, and the University of Maryland is gratefully acknowledged.

Nomenclature

D = diffusion coefficient, s^{-1}
$\mathrm{div}\,\vec{U}$ = velocity divergence, s^{-1}
$\mathrm{div}\,\vec{U}_{\max}$ = maximum velocity divergence within air, s^{-1}
$\mathrm{div}\,\vec{U}_{\min}$ = minimum velocity divergence within fuel, s^{-1}
$d(\ln P)/dt$ = logarithmic rate of pressure, s^{-1}
$[d(\ln P)/dt]_{\max}$ = maximum logarithmic rate of pressure within air, s^{-1}
$[d(\ln P)/dt]_{\min}$ = minimum logarithmic rate of pressure within fuel, s^{-1}
$d(\ln T)/dt$ = logarithmic rate of temperature, s^{-1}
$[d(\ln T)/dt]_{\max}$ = maximum logarithmic rate of temperature within fuel, s^{-1}
$[d(\ln T)/dt]_{\min}$ = minimum logarithmic rate of temperature within air, s^{-1}
M = molecular weight of mixture, kg/mole
M_a = molecular weight of air, kg/mole
M_f = molecular weight of fuel, kg/mole
m_a = mass of air, kg
m_f = mass of fuel, kg
n_a = number of moles of air, dimensionless
n_f = number of moles of fuel, dimensionless
P = pressure, N/m^2
R = universal gas constant, J/mole-K
T = temperature, K
t = time, s
t_l = fuel-lean mixing time, s
t_r = fuel-rich mixing time, s
$\vec{U}$ = velocity, m/s
x_f = mole fraction of fuel, dimensionless
y_f = mass fraction of fuel, dimensionless

References

[1] Guyon, E., Nadel, J. P., and Pomeaux, Y., eds., 1987 *Disorder and Mixing*, Klumer, Dordrecht, pp. 339–341.
[2] Brasoveanu, D., and Gupta, A. K., 1995, Analysis of Gaseous Fuel and Air Mixing, Summer Fluids Engineering Meeting, Hilton Head, SC, August 13–18.
[3] Brasoveanu, D., and Gupta, A. K., 1998, 33rd Intersociety Energy Conversion Engineering Conference, Colorado Springs, August 2–6.
[4] Brasoveanu, D., and Gupta, A. K., 1998, ''Maximum Mixing Times of Methane and Air Under Non-Reacting and Reacting Conditions,'' Proc. International Joint Power Generation Conference (IJPGC), Baltimore, MD, August 24–26.
[5] Brasoveanu, D., and Gupta, A. K., 2000, ''Maximum Mixing Times of Methane and Air,'' AIAA J. Propulsion and Power, **16**, No. 6, Nov.–Dec., pp. 956–963.
[6] Brasoveanu, D., and Gupta, A. K., 2000, ''Analysis of Gaseous Fuel and Air Mixing in Flames and Flame Quenching,'' AIAA J. Propulsion and Power, **16**, No. 5, Sept.–Oct., pp. 829–836.
[7] Brasoveanu, D., and Gupta, A. K., 1999, ''Analysis of Gaseous Fuel and Air Mixing,'' Combust. Sci. Technol., **141**, pp. 111–121.
[8] Strehlow, R. A., 1989, *Combustion Fundamentals*, McGraw-Hill, New York, pp. 114–116 and 370–373.
[9] Loeb, L. B., 1961, *The Kinetic Theory of Gases*, 3rd ed. Dover, New York, pp. 252–256.

Influence of Imperfections in Working Media on Diesel Engine Indicator Process

S. N. Danov

A. K. Gupta[1]
e-mail: akgupta@eng.umd.edu

The Combustion Laboratory,
Department of Mechanical Engineering,
University of Maryland,
College Park, MD 20742

Several improvements to the mathematical model for the indicator process in a diesel engine cylinder are proposed. The thermodynamic behavior of working media is described by the equation of state valid for real gases. Analytical mathematical dependencies between thermal parameters (pressure, temperature, volume) and caloric parameters (internal energy, enthalpy, specific heat capacities) have been obtained. These equations have been applied to the various products encountered during the burning of fuel and the gas mixture as a whole in the engine cylinder under conditions of high pressures. An improved mathematical model, based on the first law of thermodynamics, has been developed by taking into account imperfections in the working media that appear under high pressures. The numerical solution of the simultaneous differential equations is obtained by Runge–Kutta-type method. The mathematical model is then used to solve the desired practical problems in two different two-stroke turbo-charged engines: 8DKRN 74/160 and Sulzer-RLB66. Significant differences between the values calculated using ideal gas behavior and the real gas at high-pressure conditions have been found. The numerical experiments show that if the pressure is above 8 to 9 MPa, the imperfections in working medium must be taken into consideration. The results obtained from the mathematical dependencies of the caloric parameters can also be used to model energy conversion and combustion processes in other thermal machines such as advanced gas turbine engines with high-pressure ratios. [DOI: 10.1115/1.1339986]

Introduction

In recent years the peak values of pressure in some turbocharged diesel engines have reached 12–15 MPa. This rise of pressure causes some difficulties in modeling the energy conversion processes in the engine cylinders. This is because in the existing mathematical models of indicator process, the thermodynamic behavior of working medium is described by equations for ideal gases, despite their modeling efforts for almost half a century. Glagolev [1] provided a differential energy equation to describe the indicator process of a diesel engine by considering the working medium as an ideal gas with constant (average) values for internal energy u, enthalpy h, and specific heat capacities c_p and c_v during the process. Woschni [2] provided some improvements by considering internal energy u and enthalpy h as a function of the temperature, i.e., $du=c_v dT$ and $dh=c_p dT$, where T is temperature. McAuley et al., [3] took into account the influence of temperature on specific heat capacities c_p and c_v, and hence on internal energy u and enthalpy h, i.e., $du=d(c_v T)=c_v dT+Tdc_v$ and $dh=d(c_p T)=c_p dT+Tc_p$. Samsonov [4] and Klaus [5] took into account the influence of both temperature and gas mixture composition as $du/dt=(\partial u/\partial T)dT/dt+(\partial u/d\Phi)d\Phi/dt$, where t is time and Φ is fuel/air equivalence ratio. In the existing mathematical models, the internal energy u, enthalpy h, and specific heat capacities c_v and c_p of the working medium in the engine cylinder are defined only as a function of temperature and the gas mixture composition. No mathematical model is available that accounts for thermodynamic properties of the working medium with respect to pressure (density or specific volume), i.e., these models do not view working medium as a real gas. This ideal behavior assumption leads to considerable errors in modeling the indicator process for the modern forced diesel engines operating at very high peak pressures. The gas parameters, therefore, must be defined by an equation of state valid for real gases at high pressures, and thermodynamic properties of the air–gas mixture defined by considering the imperfections in the working medium.

Objectives

The overall objective of this study is to develop an improved mathematical model for the indicator process of diesel engine by taking into account imperfections in the working media. The specific objectives are to

- define mathematical descriptions of the caloric parameters (u,h,c_v,c_p) as a function of the thermal parameters (temperature T, pressure P), i.e.,
  - — enthalpy: $h=h(P,T)$,
  - — internal energy: $u=u(P,T)$,
  - — specific heat capacities: $c_v=c_v(P,T)$, $c_p=c_p(P,T)$;
- create a differential equation for the energy conversion processes in the engine cylinder valid for real gasses at high pressures and temperatures;
- provide numerical algorithms and model validation of the developed differential equations;
- investigate the effects of imperfections in the working media on combustion-related parameters in a diesel engine cylinder, such as
  - — cylinder pressure,
  - — temperature in the cylinder,
  - — heat transfer coefficient,
  - — specific heat capacities c_v and c_p,
  - — gas mixture composition (mass fractions),
  - — instantaneous values for air/fuel (A/F) ratio,
  - — fuel vaporization rate,
  - — fuel combustion rate.

The approach used for achieving the above objectives involved parallel efforts between theoretical and experimental studies. The theoretical study is used to develop an improved mathematical

[1]Corresponding author.

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, October 20, 2000; final revision received by the ASME Headquarters, November 15, 2000. Associate Editor: S. Gollahalli.

model based on the realistic thermodynamic properties in the engine cylinder. The experimental study is used to provide the model validation and model calibration.

Theory

Background. For a pressure of up to 30 Mpa, the Redlich–Kwong equation of state is recommended for describing the thermodynamic behavior of gases [6]:

$$\left(P+\frac{a}{v(v+b)\sqrt{T}}\right)(v-b)=RT \tag{1}$$

where P is pressure, T is temperature; v is specific volume, and R is gas constant. Coefficients a and b can be defined, considering that all partial derivatives are equal to zero at the critical point. For instance,

$$\left(\frac{\partial P}{\partial v}\right)_{T_c,v_c}=0 \quad \left(\frac{\partial^2 P}{\partial v^2}\right)_{T_c,v_c}=0 \tag{2}$$

Besides, at the critical point the Planck–Gibbs equation is valid:

$$\left(\frac{dP}{dT}\right)_{T_c}=\left(\frac{\partial P}{\partial T}\right)_{v_c} \tag{3}$$

The values of the critical parameters P_c, T_c, and v_c for products of combustion (e.g., N_2, O_2, CO_2, CO, H_2, H_2O) are available, e.g., [7,8].

Thermodynamic Properties of Pure Real Gases. If $P=P_c$, $T=T_c$, and $v=v_c$, the solution of Eqs. (2) and (3) allows one to obtain the coefficients a and b. The values obtained for coefficients a and b for products of combustion are shown below. These coefficients are valid for the Redlich–Kwong equation of state.

Gas	a	b
Nitrogen, N_2	2042.0	0.0009969
Oxygen, O_2	1613.0	0.0006395
Carbon dioxide, CO_2	2906.0	0.0005555
Carbon oxide, CO	2018.0	0.0008635
Hydrogen, H_2	33,942.9	0.0083852
Water (steam), H_2O	33,700.5	0.0008180

The application of equation of state to real gases leads to some complications. They follow from the necessity to lay down mathematical dependencies between thermal parameters (P, T, v) and caloric parameters (internal energy u, enthalpy h, specific heat capacities c_v and c_p). These problems can be solved by using the following partial differential equations for the thermodynamic properties [6]:

$$\left(\frac{\partial u}{\partial v}\right)=T\left(\frac{\partial P}{\partial T}\right)_v-P \tag{4a}$$

$$\left(\frac{\partial h}{\partial v}\right)_T=T\left(\frac{\partial P}{\partial T}\right)_v+v\left(\frac{\partial P}{\partial v}\right)_T \tag{4b}$$

$$\left(\frac{\partial c_v}{dv}\right)_T=T\left(\frac{\partial^2 P}{\partial T^2}\right)_v \tag{4c}$$

$$c_p=c_v-T\left(\frac{\partial P}{\partial T}\right)_v^2\left(\frac{\partial P}{\partial v}\right)_T^{-1} \tag{4d}$$

The partial derivatives on the right side of Eq. (4) can be obtained by differentiating the Redlich–Kwong equation (1). After substituting partial derivatives in (4) and integrating, the following mathematical dependency on the thermophysical properties of working media have been obtained:

$$u(v,T)=u_o(T)+\frac{3}{2}\frac{a}{b}\frac{1}{\sqrt{T}}\ln\left(\frac{v}{v+b}\right) \tag{5}$$

$$h(v,T)=h_o(T)+\frac{3}{2}\frac{a}{b}\frac{1}{\sqrt{T}}\ln\left(\frac{v}{v+b}\right)+\frac{RTv}{v-b}-\frac{a}{(v+b)\sqrt{T}} \tag{6}$$

$$c_v(v,T)=c_{v_o}(T)-\frac{3}{4}\frac{a}{b}T^{-3/2}\ln\left(\frac{v}{v+b}\right) \tag{7}$$

$$c_p(v,T)=c_v(v,T)+T\left(\frac{R}{v-b}+\frac{a}{2v(v+b)T^{3/2}}\right)^2 \times\left(\frac{RT}{(v-b)^2}-\frac{a(2v+b)}{v^2(v+b)^2\sqrt{T}}\right)^{-1} \tag{8}$$

where, $u_o(T)$, $h_o(T)$, and $c_{vo}(T)$ are values for ideal gas, and they depend on the temperature only. Equations (5)–(8) enable one to calculate the values of caloric parameters (u,h,c_v,c_p) for single (i.e., not a mixture) real gases as affected by temperature and specific volume (i.e., density), which means both temperature and pressure. The first terms on the right-hand side in these equations represent the temperature-dependent value of the corresponding parameter, i.e., a value for ideal gas. These terms are defined as

$$u_0(T)=a_o+a_1T+a_2T^2, \quad h_o(T)=b_o+b_1T+b_2T^2,$$
$$c_{vo}(T)=c_o+c_1T+c_2T^2, \quad c_{po}(T)=d_o+d_1T+d_2T^2$$

Coefficients a_i, b_i, c_i, and d_i for all products of combustion ($N_2, O_2, CO_2, CO, H_2, H_2O$) are defined by means of processing the tables of experimental data [9] for u, h, c_v, and c_p using the least squares theory [10]. The second term (and next, if present) on the right side of Eqs. (5)–(8) provide the corrections caused by the imperfections of the gasses. The equations (5)–(8) are valid for single real gases only (not for a gas mixture). The real working media in the engine cylinder, however, is a mixture of real gases. This necessitates the determination of mathematical dependencies for defining the thermophysical properties for a mixture of real gases (see below).

Verification. The validity of Eqs. (5)–(8) can be accomplished from the known thermodynamic dependencies between enthalpy, internal energy, and specific heat capacity. If Eqs. (5)–(8) are applied to ideal gases, then coefficients a and b are equal to zero. Besides, by definition, a gas is considered ideal if the specific volume is very high ($v\to\infty$), or density is very low ($\rho\to 0$). Under these conditions the forces of repulsion and attraction between the molecules are negligible, and there are no additional degrees of freedom in the molecules. This means that the equations (5)–(8) should transform to equations for ideal gas. Indeed:

If $a=0$ and $b=0$, then

$$u(v,T)=u_o(T), \quad h(v,T)=h_o(T),$$
$$c_v(v,T)=c_v(T), \quad c_p-c_v=R$$

Also, if $\rho\to 0$ (i.e., $\nu\to\infty$):

$$u(v,T)=u_o(T), \quad h(v,T)=h_o(T),$$
$$c_v(v,T)=c_v(T), \quad c_p-c_v=R.$$

Besides, the derivative of the right-hand-side of Eq. (5) with respect to the temperature is equal to the right-hand side of Eq. (7), and the derivative of the right-hand-side of Eq. (6) is equal to the right-hand side of Eq. (8), i.e., $\partial u/\partial T=c_v$, $\partial h/\partial T=c_p$, which is well known, and hence illustrates the validity of Eqs. (5)–(8).

Thermodynamic Properties of a Real Gas Mixtures—a General Case. A mixture of ideal gas is an idealized picture. If $P\to 0$ (i.e., $v\to\infty$ or $\rho\to 0$), the behavior of a gas mixture is

described by equations valid for an ideal gas mixture. However, if the pressure gets higher, the equations for ideal gases are not correct. For instance, the Dalton laws ($p=\Sigma p_i$ and $p_i=pr_i$) for a mixture of real gases are not valid for high density (pressure). Mixing of real gases with equal pressure P and temperature T in a common volume V leads to some changes of pressure and temperature of the mixture in comparison with the initial values. If the temperature and pressure of the mixture are restored to the previous value of P and T, then the specific volume v will change as

$$v(P,T)=\sum\ [r_i v_i(P,T)]+\Delta v(P,T,r_1,r_2,r_3,\ldots) \quad (9)$$

where r_i $(i=1,2,\ldots)$ are the mass fractions of components in a gas mixture. This equation is valid for a mixture of real gases. Indices i assign to the ith component of the gas mixture. The caloric parameters (u, h, c_v and c_p) will also change as

$$u(P,T)=\sum\ [r_i u_i(P,T)]+\Delta u(P,T,r_1,r_2,r_3,\ldots) \quad (10a)$$

$$h(P,T)=\sum\ [r_i h_i(P,T)]+\Delta h(P,T,r_1,r_2,r_3,\ldots) \quad (10b)$$

$$c_v(P,T)=\sum\ [r_i c_{v_i}(P,T)]+\Delta c_v(P,T,r_1,r_2,r_3,\ldots) \quad (10c)$$

$$c_p(P,T)=\sum\ [r_i c_{p_i}(P,T)]+\Delta c_p(P,T,r_1,r_2,r_3,\ldots) \quad (10d)$$

The nonadditive correction for Δu, Δh, Δc_v, and Δc_p are functions of the parameters for state and mass fractions of the components. The correction Δv, Δh is interconnected. If one of them is known, the others can be determined. The differential equation of Shpillrain and Kessellman [7] can be used for this purpose:

$$\left(\frac{\partial h}{\partial P}\right)_{T,r}=v-T\left(\frac{\partial v}{\partial T}\right)_{P,r} \quad (11)$$

This equation can be used to obtain the gas mixture enthalpy, if one takes into consideration Eq. (9):

$$\left(\frac{\partial h}{\partial P}\right)_{T,r}=\left[v_1-T\left(\frac{\partial v_1}{\partial T}\right)_P\right]r_1+\left[v_2-T\left(\frac{\partial v_2}{\partial T}\right)_P\right]r_2+\ \ldots\ +\Delta v - T\left(\frac{\partial \Delta v}{\partial T}\right)_{P,r} \quad (12)$$

or

$$\left(\frac{\partial h}{\partial P}\right)_{T,r}=\left(\frac{\partial h_1}{\partial P}\right)_{T,r}r_1+\left(\frac{\partial h_2}{\partial P}\right)_{T,r}r_2+\ \ldots\ +\Delta v-T\left(\frac{\partial \Delta T}{\partial T}\right)_{P,r} \quad (13)$$

Both sides of Eq. (11) may be integrated formally for $T=\text{const}$ from $P=0$ to the current pressure P to obtain the following expression for the enthalpy:

$$h-h_0=(h_1-h_{1o})r_1+(h_2-h_{2o})r_2+\ \ldots\ +\int_0^p\left[\Delta v-T\left(\frac{\partial \Delta v}{\partial T}\right)_{P,r}\right]dP \quad (14)$$

where h_{io} and h_o are the respective enthalpies of the relative component and mixture for ideal gases. Considering that

$$h_o=h_{o_1}r_1+h_{o_2}r_2+\cdots \quad (15)$$

from (14) and (15), it follows that

$$h=h_1r_1+h_2r_2+\ \ldots\ +\int_0^p\left[\Delta v-T\left(\frac{\partial \Delta v}{\partial T}\right)_{P,r}\right]dP\cdots \quad (16)$$

Equation (16) is a concrete expression for (10b). It confirms that the correction Δh is connected with the corrections Δv:

$$\Delta h=\int_0^P\left[\Delta v-T\left(\frac{\partial \Delta v}{\partial T}\right)_{P,r}\right]dP \quad (17)$$

Using the same approach the corrections for the other caloric parameters can be obtained by means of Δv:

$$\Delta u=\Delta h-P\Delta v \quad (18)$$

$$\Delta c_p=-T\int_0^P\left(\frac{\partial^2 \Delta v}{\partial T^2}\right)_P dP \quad (19)$$

$$\Delta c_v=\Delta c_p+T\left(\frac{\partial \Delta v}{\partial T}\right)_P^2\left(\frac{\partial \Delta v}{\partial P}\right)_T^{-1} \quad (20)$$

Equations (17)–(20) show that for Δh, Δu, Δc_p, and Δc_v are valid from the same differential equations which are valid for the common properties. For instance, if $\Delta v(P,T,r_1,r_2,\ldots)$ is known, it is easy to obtain the correction for the internal energy Δu, enthalpy Δh, and specific heat capacity Δc_p and Δc_v.

Thermodynamic Properties of Real Gas Mixture in the Engine Cylinder. The numerical experiments showed that for gas mixtures in the engine cylinder at pressures up to 20 MPa and temperatures up to 3000 °C, the thermodynamic, behavior of the mixture, as a whole, subordinates to the Amago law, i.e.,

$$\Delta v(P,T,r_1,r_2,\ldots)=0 \quad (21)$$

For this reason the mathematical model of the engine indicator process can be simplified. Especially, the equations for describing the caloric parameters of real gas mixture can be presented by Dalton's laws for gas mixture, although the components are real gases. The working medium could be viewed as a real gas mixture, containing three components only: ''pure'' air, ''pure'' gases (which are products from the combustion of fuel for a stoichiometric air/fuel ratio), and fuel vapors. However, the ''pure'' air and ''pure'' gases themselves are viewed as real gas mixtures, subordinating to the Amago law. It should be noted that coefficients a and b do not depend on the correction for specific volume, Δv:

$$a=\Sigma r_i a_i, \quad b=\Sigma r_i a_i \quad (22)$$

The Differential Equation. Equations (5)–(8) and (17)–(20) presented above allow one to develop an improved mathematical model for the energy conversion process in a diesel engine cylinder. The engine cylinder is a thermodynamic system. The first law of thermodynamics can describe the various ongoing processes in the cylinder:

$$dQ=dU+P.dV \quad (23)$$

This gives the dependency between heat dQ, internal energy dU, and mechanical work $P.dV$. It is convenient to present Eq. (23) in a differential form with respect to the angle of the crankshaft rotation φ:

$$\frac{dU}{d\varphi}=\frac{dQ}{d\varphi}-P\frac{dV}{d\varphi} \quad (24)$$

The rate of changing of absolute internal energy with respect to the angle of the crankshaft rotation can be presented as

$$\frac{dU}{d\varphi}=\frac{d(Mu)}{d\varphi}=M\frac{du}{d\varphi}+u\frac{dM}{d\varphi} \quad (25)$$

where M is the current mass of the working medium at the cylinder, viewed as a mixture from pure air M_1, pure gasses M_2, and

fuel vapors M_3 remaining (i.e., not yet burned). As a result, the current value of the specific internal energy u is a function of the mass fraction (percentages) of the components r_1, r_2, r_3, temperature T, and specific volume v:

$$u = u(r_1, r_2, r_3, T, v) \tag{26}$$

The respective mass fractions are functions of the integral laws of fuel combustion $f_b = f_b(\varphi)$ and fuel vaporization $f_v = f_v(\varphi)$:

$$r_1 = r_1(f_b), \quad r_2 = r_2(f_b), \quad r_3 = r_3(f_b) \tag{27}$$

Taking into account the circumstances (26) and (27), Eq. (25) can be presented as

$$\frac{dU}{d\varphi} = \left[\frac{du}{dT}\frac{dT}{d\varphi} + \frac{\partial u}{\partial r_1}\frac{dr_1}{df_b}\frac{df_b}{d\varphi} + \frac{\partial u}{\partial r_2}\frac{dr_2}{df_b}\frac{df_b}{d\varphi}\right]M + \left[\frac{\partial u}{dr_3}\left(\frac{\partial r_3}{\partial f_v}\frac{df_v}{d\varphi} - \frac{\partial r_3}{\partial f_b}\frac{df_b}{d\varphi}\right) + \frac{\partial u}{\partial v}\frac{dv}{d\varphi}\right]M + u\frac{dM}{d\varphi} \tag{28}$$

The specific heat capacity is a function of temperature. Thus

$$\frac{du}{dT} = \frac{\partial u}{\partial T} + \frac{\partial u}{\partial c_{v_1}}\frac{dc_{v_1}}{dT} + \frac{\partial u}{\partial c_{v_2}}\frac{dc_{v_2}}{dT} + \frac{\partial u}{\partial c_{v_3}}\frac{dc_{v_3}}{dT} = c_v(v,T) + T\sum_{i=1}^{3}\frac{dc_{v_i}}{dT} \tag{29}$$

The following dependency was obtained for the current mass fraction of the gas mixture components:

$$r_1 = m_c L(\lambda - f_b)/M \tag{30a}$$

$$r_2 = m_c[\gamma\lambda L + f_b(L+1)]/M \tag{30b}$$

$$r_3 = m_c(f_v - f_b)/M \tag{30c}$$

where L is the stoichiometric mass of air for burning 1 kg of fuel; λ is the air/fuel equivalence ratio; m_c is the cycle amount of fuel injected per cycle per cylinder (kg/cyc/cyl); γ is the coefficient of residual gases; and $f_v = f_v(\varphi)$ and $f_b = f_b(\varphi)$ are current nondimensional mass fractions of fuel vaporized and burned (so-called integral laws of fuel vaporization and fuel combustion). The dynamics of change of specific volume can be presented by the rates of change of the working medium mass M and current volume V, because $v = V/M$:

$$\frac{dv}{d\varphi} = \frac{\partial v}{\partial V}\frac{dV}{d\varphi} + \frac{\partial v}{\partial M}\frac{dM}{d\varphi} = \frac{1}{M}\frac{dV}{d\varphi} - \frac{V}{M^2}\frac{dM}{d\varphi} \tag{31}$$

The heat-exchange rates can be presented as follows:

$$\frac{dQ}{d\varphi} = \frac{dQ_c}{d\varphi} + h\frac{dM}{d\varphi} - \frac{dQ_w}{d\varphi} - \frac{dQ_{ev}}{d\varphi} - \frac{dQ_d}{d\varphi} \tag{32}$$

where $dQ/d\varphi$ is the heat release rate due to the fuel combustion; $dQ_w/d\varphi$ is the heat transfer rate between gases and cylinder walls; $dQ_{ev}/d\varphi$ is the heat-exchange rate due to the evaporating of fuel; $dQ_d/d\varphi$ is the heat-exchange rate due to the dissociation of products of burning; and h is the enthalpy of the working medium viewed as a real gas mixture. The heat release rate can be given as

$$\frac{dQ_c}{d\varphi} = m_c Q_f \frac{df_b}{d\varphi} \tag{33}$$

In this equation, Q_f is the fuel calorific value and $df_b/d\varphi$ is the differential nondimensional law of fuel combustion, which corresponds to the overall combustion rate, df_b/dt.

After substituting (29) into (28), and Eq. (32) into (24), one can obtain the final differential equation for the first law of thermodynamics, which is valid for the cylinder of a diesel engine:

$$M\left[c_v(v,T) + T\sum_{i=1}^{3}\frac{dc_{v_i}}{dT}\right]\frac{dT}{d\varphi} = m_c Q_f\frac{df_b}{d\varphi} - M\left(\frac{du}{dr_1}\frac{dr_1}{df_b} + \frac{du}{dr_2}\frac{dr_2}{df_b} + \frac{du}{dr_3}\frac{dr_3}{df_b}\right)\frac{df_b}{d\varphi} - M\left[\frac{\partial u}{\partial r_3}\frac{\partial r_3}{\partial f_b}\frac{df_V}{d\varphi} - \frac{\partial u}{\partial v}\left(\frac{1}{M}\frac{dV}{d\varphi} - \frac{V}{M^2}\frac{dM}{d\varphi}\right)\right] + (h-u)\frac{dM}{d\varphi} - \frac{dQ_W}{d\varphi} - \frac{dQ_{ev}}{d\varphi} - \frac{dQ_d}{d\varphi} - P\frac{dV}{d\varphi} \tag{34}$$

Differential equation (34) describes the energy conversion process in a diesel engine cylinder from the beginning of the real compression process to opening of the exhaust valve/port. During the indicator process (compression, combustion, expansion) the gas-exchange valves/ports are closed ($dM_{in}/d\varphi = 0$; $dM_{ex}/d\varphi = 0$), i.e., there is no mass exchange and energy exchange through the inlet and exhaust valves/ports. For this case the mass balance in the engine cylinder can be presented as

$$\frac{dM}{d\varphi} = m_c\frac{df_i}{d\varphi} - \frac{dM_L}{d\varphi} \tag{35}$$

i.e., the change of mass at the cylinder is caused by the fuel injection ($df_i/d\varphi$ is the nondimensional fuel injection rate) and gas leakage through the piston rings, $dM_L/d\varphi$. It is obvious that before the beginning of fuel injection $df_i/d\varphi = 0$, $df_v/d\varphi = 0$, and $df_b/d\varphi = 0$, i.e., no fuel injection, vaporization and combustion.

In order to solve the differential equation (34) one needs

- mathematical descriptions for the caloric parameters (u, h, c_v, and c_p), i.e., the above-obtained Eqs. (5)–(8)
- mathematical model for the fuel vaporization [11],
- mathematical model for the fuel combustion rate [11],
- mathematical model for the heat transfer rate [12],
- mathematical model for the leakage of gases through the piston rings [13],
- Mathematical model for loses due to dissociation [14],
- geometrical characteristics and kinematic dependencies of crank mechanism for defining $dV/d\varphi$ via engine speed n, rpm.

Numerical Solution

An algorithm has been developed to numerically integrate the above system of nonlinear first-order differential equations. The input data includes

- geometric parameters of engine:
  - — cylinder bore
  - — stroke
  - — geometry of combustion chamber
  - — piston rings geometry
  - — characteristics of crank shaft
  - — geometric compression ratio
- operating conditions
  - — fuel amount per cycle per cylinder
  - — fuel temperature at injector outlet
  - — engine speed
  - — ambient temperature and pressure
  - — cooling fluid temperature
  - — start of fuel injection
- thermo-physical constants
  - — fuel calorific value
  - — emissivity of cylinder walls
  - — fuel latent heat (heat of vaporization)
  - — fuel activation energy
  - — Semenov's constants

The differential equations are nonlinear, with variable coefficients. The standard fourth-order Runge–Kutta process [15] has been applied to integrate these simultaneous first-order ordinary differential equations. The step $\Delta\varphi$ of integration is variable. It depends on the current truncation error of this Runge–Kutta process [10]. The current truncation error is estimated by the rule of Collatz [16]. An algorithm and FORTRAN program have been developed for use on both a PC and a UNIX workstation. The computing procedure is iterative. The solution allows one to get curves for $T=T(\varphi)$, $P=P(\varphi)$, $r_i=r_i(\varphi)$, $f_v=f_v(\varphi)$, and $f_b=f_b(\varphi)$ (see Figs. 7–16). In parallel with integration of the differential equations, a numerical process has also been provided for defining some of the integral parameters. As an example, the mean indicated pressure is defined as

$$P_i=\frac{1}{V_b-V_a}\int_{V_a}^{V_b}PdV$$

$$=0.5\Delta\varphi V_c(\varepsilon-1)\sum_{k=1}^{N}\left[P_k(\sin\varphi+0.5\lambda_c\sin 2\varphi)\right] \quad (36)$$

where λ_c is characteristic of the crank mechanism; k is the number of the current computing interval; ε is the engine compression ratio; $\Delta\varphi$ is the step of the independent variable; and V_c is the volume of the combustion chamber.

Model Validation

Figure 1 shows a comparison between the simulation, via both the ideal and real gas models, and the experimental results on pressure history in the cylinder for a two-stroke slow-speed marine turbo-charged diesel engine Sulzer, type 6RLB66. The model used the same operating conditions as that in the experimental case, i.e., engine speed $n=140$ rpm, fuel calorific value $=42.39$ MJ/kg (diesel fuel), start of fuel injection$=-11°$BTDC, fuel amount$=0.03785$ kg/cycle/cylinder, ambient pressure $=740$ mm Hg and ambient temperature$=27°$C. The ''noise'' in the measured experimental curve is filtered by means of a windowing method for filtration [17]. A comparison between the integral indicator parameters for the same engine and operating conditions are shown in Table 1. In the case of the ideal gas model, the influence of pressure (and therefore density) on the caloric parameters (u, h, c_v, and c_p) has been ignored. It can be seen that if the pressure in the engine cylinder is over 8–9 MPa, considerable differences exist between experimental and theoretical results for the ideal gas model case. These differences exceed the errors of the experiment (see Fig. 1 and Table 1). The results clearly show that the ideal gas model is not adequate to represent the real indicator process at high pressures in the cylinder.

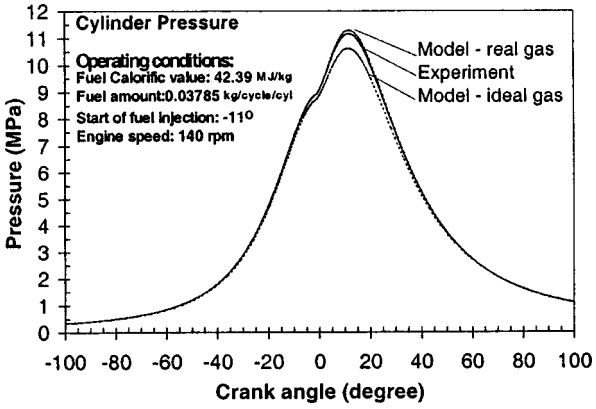


Fig. 1 A comparison between the experimental data calculated (using both real and ideal gas models) results on indicated pressure for engine 6RLB66

Table 1 A comparison between the predicted and experimental data for the integral parameters of marine diesel engine ''Sulzer-6RLB66''

Parameters	Modeling: Ideal gas model	Modeling: Real gas model	Experimental data
Mean indicated pressure (MPa)	1.451	1.572	1.554
Compression pressure (MPa)	8.752	8.970	8.953
Peak pressure of cycle (MPa)	10.609	11.284	11.150
Specific indicated fuel consumption (kg/kW·h)	0.196	0.181	0.183
Heat losses in the cooling fluid (percent)	13.88	14.97	14.50

Table 2 shows a comparison between the predicted results obtained from the real gas model and experimental data for marine diesel engine 8DKRN 74/160-3 on the indicator parameters. The differences between corresponding parameters are commensurable with the experimental error. In Table 2, P_i is the mean indicated pressure; P_c is the compression pressure; P_z is the peak pressure of the cycle; and T_g is the temperature of gases behind the exhaust valve. Both the stochastic nature of the cycle variability and the imperfections of diagnostic facilities cause the errors

Table 2 A comparison between the model results and experimental data for engine 8 DKRN 74/160-3, for $q=0.02654$ kg/cyc and $n=100$ rpm

Parameter	Model	Experiment: Mean	Experiment: Probability	Relative error of experiment (percent)
P_i (MPa)	0.70	0.70 ± 0.021	$P(0.68<P_i<0.72)=0.95$ $P(0.016<\sigma_{pi}<0.051)=0.95$	3.03
P_c (MPa)	3.62	3.58 ± 0.055	$P(3.52<P_c<3.63)=0.98$ $P(0.032<\sigma_{Pc}<0.124)=0.98$	1.55
P_z (MPa)	5.42	5.14 ± 0.175	$P(4.97<P_z<5.32)=0.99$ $P(0.083<\sigma_{Pc}<0.378)=0.99$	3.42
T_g (K)	631	623 ± 1.82	$P(612<T_g<634)=0.98$ $P(6.58<\sigma_{Tg}<25.4)=0.98$	1.82

The experiments cited in Tables 1 and 2 and Fig. 1 have been performed by one of the authors (S.D.) during his affiliation in Department of ''Thermotechnics'' at the Maritime Navigation Bulgar Co.

and uncertainties. The relative errors of experiments for the parameters listed in Tables 1 and 2 are as follows: $\delta P_i = 3.03$ percent, $\delta P_c = 1.55$ percent, $\delta P_z = 3.42$ percent, $\delta sfc = 3.1$ percent, and $\delta P = 3.05$ percent.

The comparison between the experimental data and numerical results from the real gas model shows a very good agreement on both the current value of the parameters (Fig. 1) and integral parameters (Tables 1 and 2). An examination of the results show that the relative differences between the experimental data and results from the real gas model for the integral indicator parameters are commensurable with the experimental error. The real gas model developed here, therefore, does a good job to predict the measured parameters, at least for the chosen conditions in the engine. The general applicability of the model to other engines, operating under high-pressure and high-temperature conditions, require the availability of comprehensive experimental data so that the general validity of the model can be confirmed. In the absence of other experimental data one can conclude that the model provides an opportunity of using the real gas model to solve applied engineering problems.

Results and Discussions

Figure 2 shows the computed isotherms of carbon monoxide viewed as a real gas (via the iterative solution of Eq. 1). The results show that the model is capable of providing the thermal parameters (pressure, volume, and temperature) both for a single gas and a gas mixture, viewed as real gas. Figures 3–6 show the differences between the values for real gas and ideal gas for specific heat capacities Δc_v and Δc_p, internal energy Δu, and enthalpy Δh. These examples are for carbon dioxide, but the model can predict the thermodynamic behavior of any gas under high-pressure and high-temperature conditions [Eqs. (5)–(8)].

The mathematical model provides an opportunity to investigate the effect of imperfections in the working media on various parameters in the cylinder during the indicator process. The results obtained for the cylinder pressure, average temperature in the cylinder, overall heat transfer coefficient, specific heat capacities c_v and c_p, gas mixture composition (mass fractions), instantaneous values for A/F ratio, fuel vaporization rate, and fuel combustion rate are shown in Figs. 7–16. The results show a comparison on the value of the respective parameters, obtained using the real gas and ideal gas models.

It should be noted that the calculated results for pressure and temperature, via the new mathematical model, depend not only on the accuracy of equation of state (Ridlich–Kwong equation) but also on the equations for caloric parameters of the gas mixture (u, h, c_v, and c_p) in the engine cylinder. Furthermore, in this case, the current (instantaneous) value of temperature T and pressure P influence the process of fuel vaporization and combustion and vice versa. The mathematical model for the fuel combustion rate [11] requires the instantaneous value of temperature and pressure, but during the same time the fuel combustion rate strongly influences the temperature rate (and hence pressure) via the heat re-

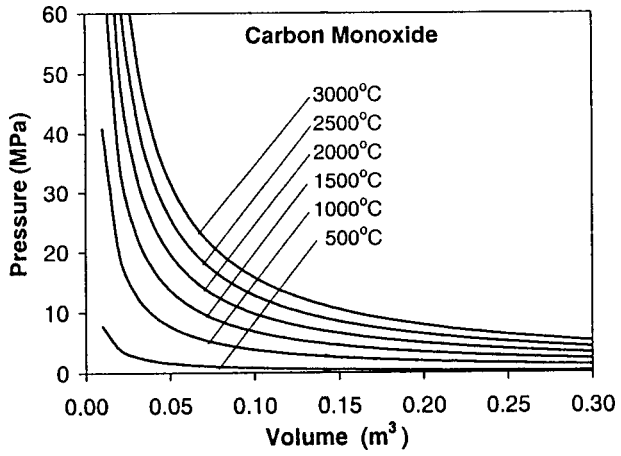


Fig. 2 Isotherms for carbon monoxide viewed as a real gas

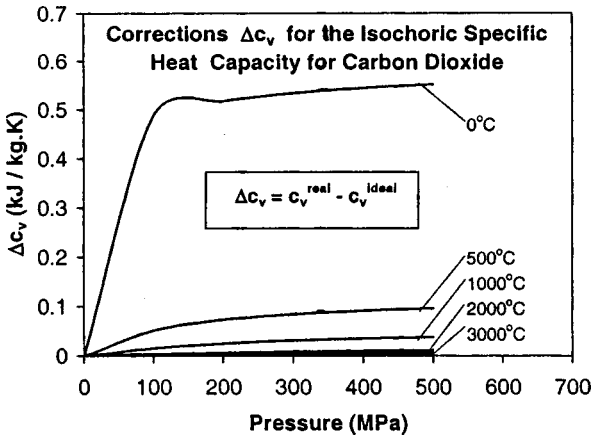


Fig. 3 Differences between the values of specific heat at constant volume Δc_v for real gas and ideal gas for carbon dioxide at various pressures and temperatures

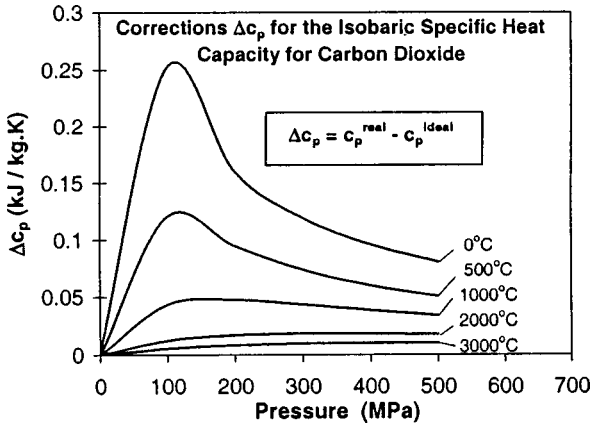


Fig. 4 Differences between the specific heat values at constant pressure Δc_p for real gas and ideal gas for carbon dioxide at various pressures and temperatures

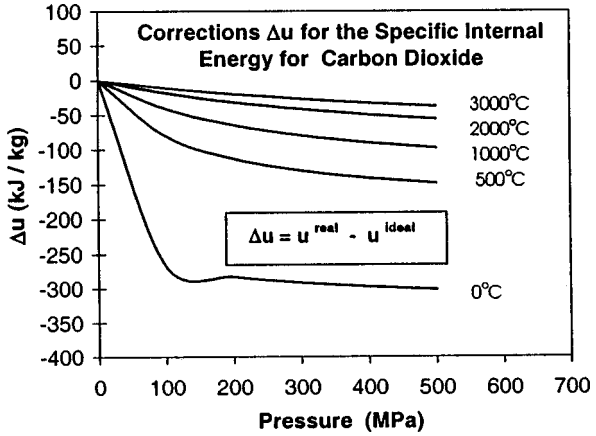


Fig. 5 Differences between the internal energy value Δu for real gas and ideal gas for carbon dioxide at various pressures and temperatures

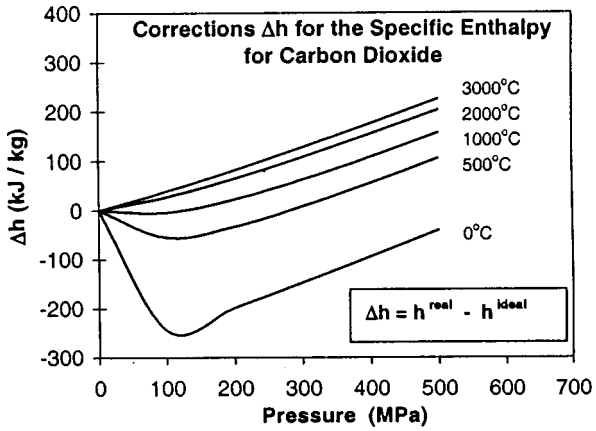


Fig. 6 Differences between the enthalpy values Δh for real gas and ideal gas for carbon dioxide at various pressures and temperatures

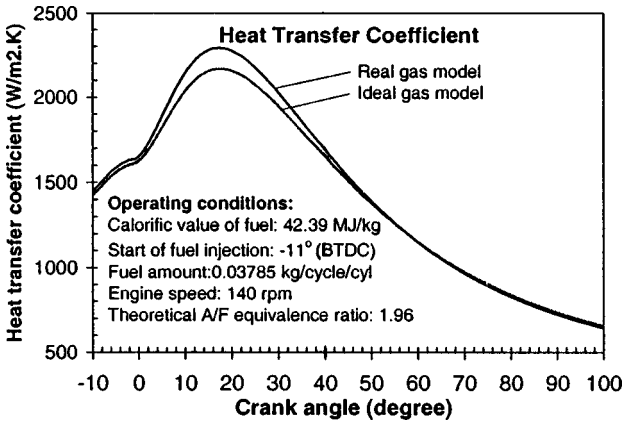


Fig. 9 Comparison between the results from the real gas model and the ideal gas model for the heat transfer coefficient during the combustion process in the cylinder for engine 6RLB66

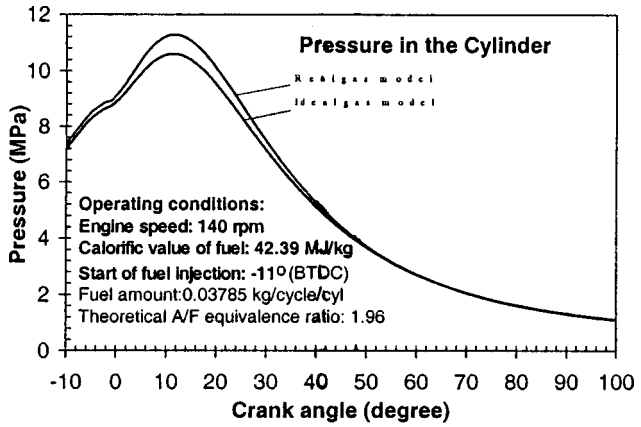


Fig. 7 Comparison between the results from the real gas model and the ideal gas model on the cylinder pressure during the combustion process for diesel engine 6RLB66

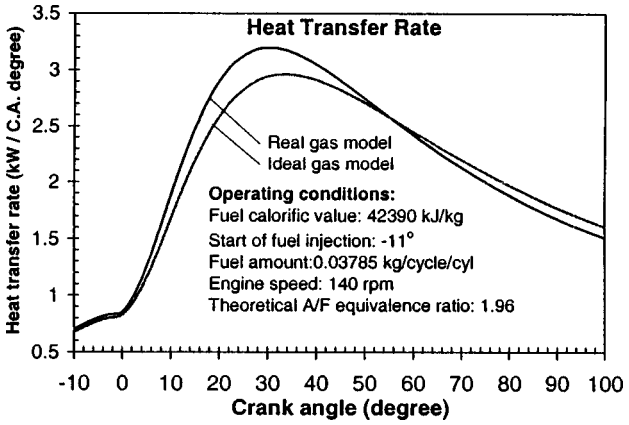


Fig. 10 Comparison between the results from the real gas model and the ideal gas model for the heat transfer rate during the combustion process in the cylinder for diesel engine 6RLB66

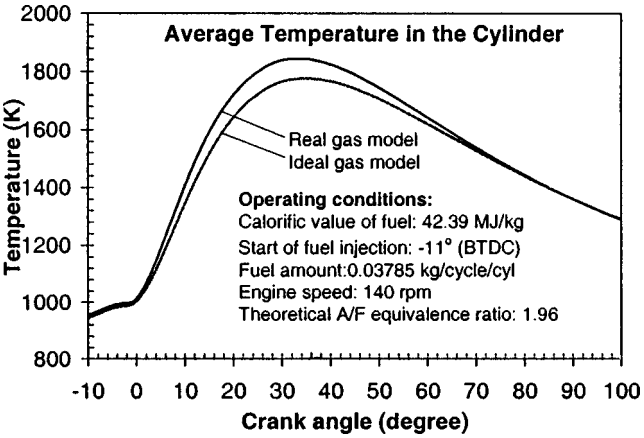


Fig. 8 Comparison between the results from the real gas model and ideal gas model for the cylinder temperature during the combustion process for diesel engine 6RLB66

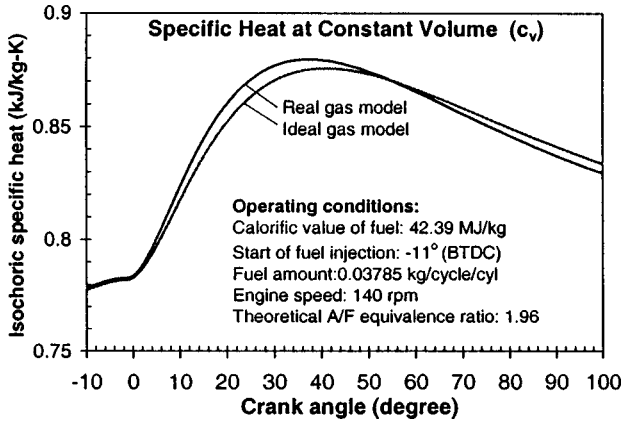


Fig. 11 Comparison between the results from the real gas model and the ideal gas model for the specific heat at constant volume c_v during the combustion process for engine 6RLB66

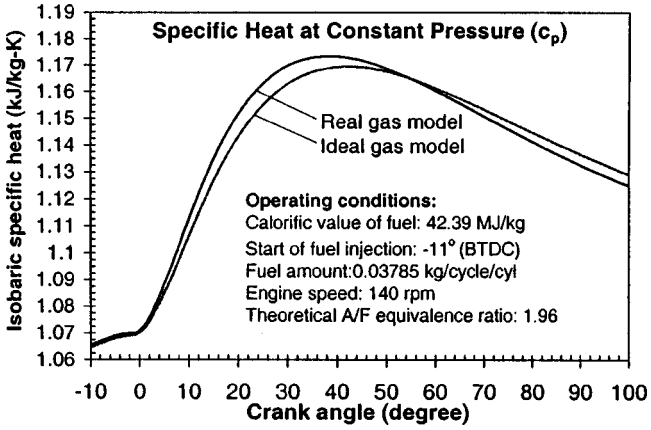


Fig. 12 Comparison between the results from the real gas model and the ideal gas model for the specific heat at constant pressure c_p during the combustion process for engine 6RLB66

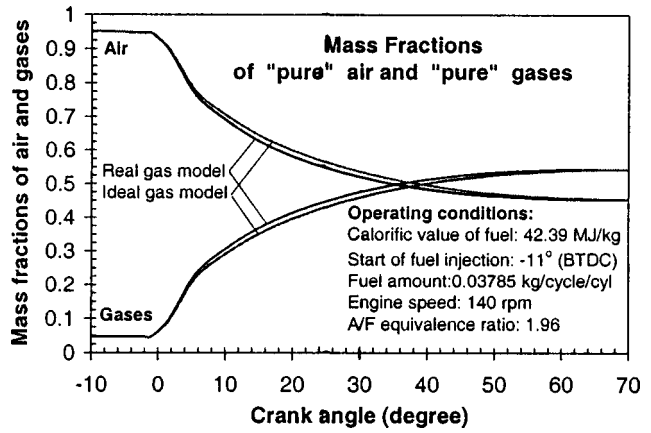


Fig. 13 Current mass fractions of pure air and pure gas for the real gas and ideal gas cases during the combustion process for diesel engine 6RLB66

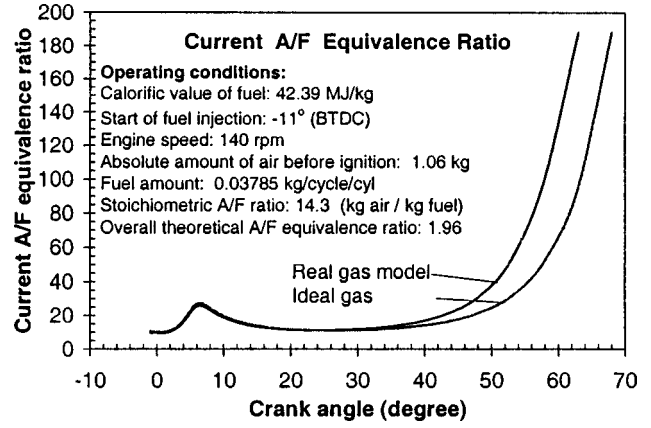


Fig. 14 Comparison of results between the real gas model and the ideal gas model on the current A/F equivalence ratio during the combustion process in the cylinder for engine 6RLB66

lease rate. Therefore, the mathematical model for fuel evaporation and combustion plays a multiplying (accelerating) effect on the temperature rate dT/dt (T is temperature, t is time). From this point of view the comparison between Fig. 1 and Tables 1 and 2 confirms that the proposed ''real gas'' model is more precise. These multiplying effects, in fact, enhance the advantages of us-

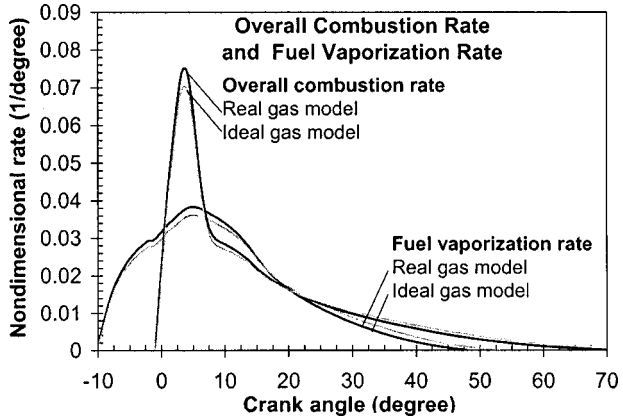


Fig. 15 Comparison of results between the real gas model and the ideal gas model on the fuel vaporization rate and overall combustion rate in the cylinder of diesel engine 6RLB66

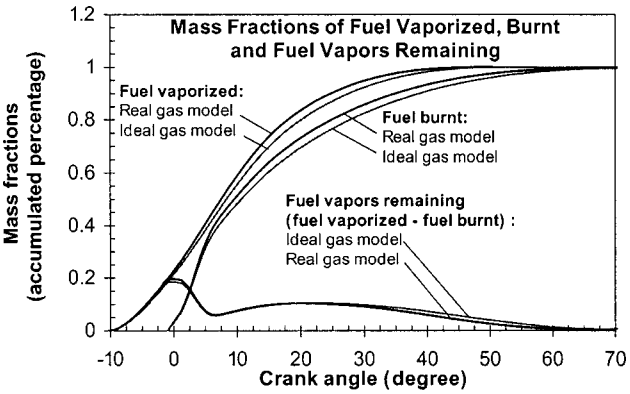


Fig. 16 Comparison of results on the fuel mass vaporized, fuel mass burnt, and fuel vapors remaining in the cylinder for the related real and ideal gas models for diesel engine 6RLB66.

ing the proposed ''real gas'' model and disadvantages of the ''ideal gas'' model, in particular when the pressure in the engine cylinder is high, in excess of 8 MPa.

Modeling studies have also been carried out with all coefficients a and b ($i=1,2,3$) in Eqs. (5)–(8) for thermophysical properties and the differential equation (34) of the first law of thermodynamics set equal to zero. In this case the working medium has been considered as an ideal gas, i.e., pressure (density) and temperature influences on parameters u, h, c_v, and c_p have been ignored. The numerical experiments show that if the pressure at the engine cylinder is over 8–9 MPa, considerable differences between experimental and theoretical data appear, and these differences can exceed experimental errors. The results show that to model the working process of a forced diesel engine, at high maximum pressures (above 8–9 MPa), the imperfections in the working media must be taken into consideration.

Conclusions

• Mathematical dependencies between the caloric parameters and thermal parameters have been provided to describe the thermodynamic behavior of real gases.

• An improved mathematical model has been developed for calculating the indicator process of a diesel engine.

• The obtained mathematical dependencies for the thermophysical properties have a universal character. They can be used for both real and ideal gases (single gas or a gas mixture) without any adjustment.

• It has been demonstrated that to model the working process of diesel engines, operating at pressures above 8–9 MPa, the imperfections in working media must be taken into consideration.

• The proposed model provides an opportunity to investigate the effects of imperfections in the working media on the combustion related parameters, such as pressure, temperature, thermodynamic properties of air/gas mixture, heat transfer, fuel vaporization and combustion rate, and current A/F ratio.

• The results obtained from the mathematical model on the caloric parameters can be used to address combustion and energy conversion processes in other thermal engines, such as advanced gas turbine engines operating at high pressures.

Acknowledgments

Partial support of this research by NASA Glenn Research Center, ONR and the University of Maryland is gratefully acknowledged.

Nomenclature

Symbols

a = coefficient in the Redlich–Kwong equation of state, $N \cdot m^4 \cdot kg^2$
b = coefficient in the Redlich–Kwong equation of state, m^3/kg
c_v = specific isochoric heat capacity, J/kg.K
c_p = specific isobaric heat capacity, J/kg.K
f_b = overall nondimensional fuel mass fraction burned up to the current
f_v = nondimensional fuel mass fraction vaporized up to the current moment of time
h = specific enthalpy of air/gas mixture, J/kg
M = current mass of air–gas mixture in the cylinder, kg
m_b = mass of fuel burned up to the current moment of time, kg
m_c = total fuel mass injected per cycle per cylinder, kg
m_d = fuel mass burned up to the current moment under the conditions of diffusion flame, kg
m_p = fuel mass burned up to the current moment under the conditions of premixed flame, kg
m_v = fuel mass vaporized up to the current moment of time, kg
n = engine rotating speed, rpm
P = pressure, N/m^2
Q_d = heat loss due to the dissociation
Q_v = heat loss due to fuel vaporization
Q_w = heat loss due to the heat transfer
Q_f = fuel calorific value, J/kg
r_i = ($i=1,2,3$) mass fractions of ''pure air,'' ''pure gases,'' and fuel vapors
R = gas constant, J/kg-K
t = time, sec
T = temperature (current) in the engine cylinder, K
u = specific internal energy, J/kg
v = specific volume of air/gas mixture, m^3/kg
V = volume (current) of the engine cylinder, m^3

Greek Symbols

ρ = density of the air–gas mixture in the cylinder, kg/m^3
ρ_a = density of air, kg/m^3
λ = air/fuel equivalence ratio $\lambda = (A/F)_{actual}/(A/F)_{stoichiometric}$
λ_c = characteristic of the crank mechanism
Φ = fuel/air equivalence ratio $\Phi = (F/A)_{actual}/(F/A)_{stoichiometric}$
φ = crank angle

Subscripts

b = fuel burned
c = cycle
c = cylinder
d = diffusion
d = dissociation
ev = evaporation
f = fuel
g = gasses
I = fuel injected
o = initial, originating value
p = premixed flame
Σ = overall, sum
v = fuel vaporized
w = cylinder wall

Definitions, Acronyms, Abbreviations

(A/F) = air/fuel ratio
air/fuel equivalence ratio = $(A/F)_{actual}/(A/F)_{stoichiometric}$
fuel/air equivalence ratio = $(F/A)_{actual}/(F/A)_{stoichiometric}$
CA = crank angle

Non-dimensional fuel mass fractions (accumulated percentages):

$f_i = m_i/m_c$ = overall nondimensional fuel mass fraction injected up to the current moment of time
$f_b = m_b/m_c$ = overall nondimensional fuel mass fraction burned up to the current moment of time
$f_v = m_v/m_c$ = nondimensional fuel mass fraction vaporized up to the current moment of time

Nondimensional rates:

FIR = fuel injection rate, 1/sec $FIR = df_i/dt = d(m_i/m_c)/dt$
FVR = fuel vaporization rate, 1/sec $FVR = df_v/dt = d(m_v/m_c)/dt$
OCR = overall combustion rate, 1/sec $OCR = df_b/dt = d(m_b/m_c)/dt$

References

[1] Glagolev, N. M., 1950, *Working Processes in Internal Combustion Engines* (Mashgiz, Moscow) (in Russian).
[2] Woschni, G., 1965, ''Electronische Berechnung von Verbrenung Motor-krei-Prozessen,'' MTZ, **26**, No. 11, pp. 439–446.
[3] McAulay, et al., 1965, ''Development Evaluation of the Simulation of the Compression Ignition Engine,'' SAE paper 650452, p. 30.
[4] Samsonov, L. A., 1980, ''Modeling the Indicator Process of Marine Diesel Engines by using the Monte-Carlo Method,'' Dvigatelestroenie, **1980**, No. (4), pp. 23–25 (in Russian).
[5] Klaus, H., 1984, ''Bettbag zur Berechnung des Betriebsverhaltens Gleichstromgespulter zweitakt-Schiffs-dieselmotoren,'' MTZ, **45**, No. 9, pp. 345–352.
[6] Wark, K., 1995, *Advanced Thermodynamics for Engineers*, McGraw-Hill, New York.
[7] Shpillrain, E., and Kessellman, P., 1977, ''Bases of the Theory of Thermo-Physical Properties of Materials,'' Energy (Moscow), **1977**, p. 248 (in Russian).
[8] Bird, G. A., 1994 *Molecular Gas Dynamics and the Direct Simulation of Gas Flows*, Oxford University Press, Oxford, p. 480.
[9] Turns, S., 1996, *An Introduction to Combustion Concepts and Applications*, McGraw-Hill, New York.
[10] Dorn, W. S., and McCracken, D. D., 1972, *Numerical Methods With ''FORTRAN 4'' Case Studies*, John Wiley and Sons, Inc., New York.
[11] Danov, S., and Gupta, K., 2000, ''Effect of Sauter Mean Diameter on the Combustion Related Parameters in Heavy-Duty Diesel Engines,'' AIAA J. Propulsion and Power, **16**, No. 6, Nov.–Dec., pp. 980–987.
[12] Woschni, G., and Fleger, J., 1979, ''Auswertung gemessener Temperatur felder zur Bestimmang ortlicher warmeulergangs koeffizienten am kollben eingines schnellaufenden Dieselmotors,'' MTZ, **1979**, pp. 153–158.
[13] Danov, S., 1988, ''Mathematical Modeling of Energy Conversion Characteristics of Ship Diesel Power Plants,'' Ph.D. thesis, Varna Technological University.
[14] Danov, S., 1989, ''Modeling the Combustion Process of a Diesel Engine for Partial Operating Conditions,'' Fourth International Symposium ''PRADS'89,'' Varna, Vol. 3, pp. 156.1–156.5.
[15] Fox, L., 1962, *Numerical Solution of Ordinary and Partial Differential Equations*, Pergamon Press, Oxford.
[16] Ralston, A., 1962, ''Runge-Kutta Methods With Minimum Error Bounds,'' Math. Comput., **16**, pp. 431–437.
[17] Rabiner, L., and Gold, B., 1985, *Theory and Application of Digital Signal Processing*, Prentice-Hall, Englewood Cliffs, NJ.

K. Cen
Department of Energy Engineering,
Zhejiang University,
Hangzhou, 310027, China

P. Basu
Department of Mechanical Engineering,
DalTech Dalhousie University,
Canada B3J 2X4

L. Cheng[1]

M.-X. Fang

Z.-Y. Luo

Department of Energy Engineering,
Zhejiang University,
Hangzhou, 310027, China

Design and Performance Verification of a 62-MWt CFB Boiler

The present paper discusses the goals and methods of design of a circulating fluidized bed (CFB) boiler: A 62-MWt (75-T/h) CFB boiler was designed manually using experience and available published data. The design was regenerated by an Expert System for evaluation and confirmation of the manual design. Design choices and resulting surface areas from these two approaches were compared and validated. Final confirmation of the design can be obtained only from operation of the boiler. The boiler was built and commissioned in the Jiangshu Province, China, firing the design coal. A comprehensive test program was undertaken to monitor the performance of the boiler. Predicted performance from both the manual design and the Expert System is compared with those measured in the boiler. A reasonable agreement between the predicted performance and measured values was obtained, which confirmed both experience-based design decision and that from the Expert System. [DOI: 10.1115/1.1340621]

Keywords: Circulating Fluidized Bed Boiler, Design, Expert System

Introduction

Trigeneration is a clean coal technology in which coal is used to simultaneously generate electricity, process steam, and fuel gas. This system partially gasifies the coal in an atmospheric pressure reactor. The carbonaceous solid residue is burnt in a circulating fluidized bed (CFB) boiler to produce steam for heat and electricity generation. Thus, the boiler is an important component of a trigeneration system. A trigeneration system was designed for the Jiangshu Province of China [1]. This system required a 62-MWt boiler that had to be designed to burn Pingdingshan coal.

Conventional procedure-based design methods cannot handle all design issues involved in an emerging technology such as CFBs. A great deal of manual intervention is required to address the design issues. An Expert System [2,3], based on both rules and procedures, can perform this task efficiently. So the 62-MWt CFB boiler for the trigeneration system was designed manually first. Then an Expert System was used to evaluate the design and compare its performance with those measured after commissioning of the boiler.

A general arrangement of the boiler is shown in Fig. 1. The main features of the boiler included high recycle ratio, high-temperature cyclones, external heat exchanger, and in-duct start-up burner. Coal and limestone are fed into the furnace using two scraper conveyors. Unburned char and hot solids leave the furnace at the top. It is separated in two hot cyclones. Collected solids are returned to the furnace through a bubbling fluidized bed external heat exchanger. Ash is drained just above the air distributor at the rear wall of the boiler. Additional details are given elsewhere by Fang et al. [1]. The present paper concentrates on the design method of the furnace.

Table 1 lists the main design parameters. The coal used in the design was Chinese Pingdingshan bituminous (Table 2). Average size of the coal was 3 mm (0–8 mm). Its density was 2200 kg/m^3. Limestone was used to capture sulfur during the operation. Its size was between 0 and 2 mm and the average value was 0.25 mm. The density of the limestone was 2700 kg/m^3. The ratio of calcium to sulfur was selected as 2 in the design for a sulfur capture efficiency of 85 percent.

Expert System

An Expert System (CFBCAD) was developed by Greenfield Research Inc. (GRI) in collaboration with the DalTech Dalhousie University and Electricite de France with the support of fluidized bed operators and boiler designers from Europe, North America, and Asia. It is a commercially available Expert System for evaluation of the design of circulating fluidized bed boilers.

For given input specification steam parameters and environmental stipulations, and a given feed stock, CFBCAD© [3] optimizes the size and configuration of a CFB boiler following a default or designer-prescribed choice and methodology. Although the Expert System carries out the entire design, the designer can control and step into it while the design is in progress. The main architecture of the program is given in Fig. 2. The main elements of the design method include (1) input parameters; (2) designer choice; (3) stoichiometric calculations; (4) steam enthalpy calculations; (5) efficiency, flow rate of coal, limestone, and flue gas; (6) design constraints; (7) heat duty distribution; (8) design of cyclones; (9) design of the furnace; (10) design of the backpass; and (11) matching of the different components. More details were reported by LaFanechere et al. [4].

Design Considerations

Three major aspects, i.e., furnace cross section, furnace height, and furnace openings, are included in the design of the furnace [5]. The furnace design is based on optimizing the combustion efficiency of the fuel and the emission characteristics. The furnace cross section is chosen primarily from combustion considerations. The gas velocity and excess air level were predetermined as 6 m/s and 20 percent, respectively. The combustor temperature was set as 900°C in consideration of high combustion efficiency and low SO_2 and NO_x emission. The furnace cross section is determined from the air flow, the gas velocity, and the fuel type. It is 2.45 m×5.45 m for the present 62-MWt boiler. Furnace height is based on the need to provide adequate residence time for combustion of the fuel. It is also defined by the operating bed temperature and the requirement of heat-transfer surfaces in the combustor. These considerations led to the choice of 20 m in the design.

[1]Presently at Dalhousie University.

Contributed by the Power Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Power Division, June 18, 2000; final revision received by the ASME Headquarters, June 27, 2000. Editor: H. D. Nelson.

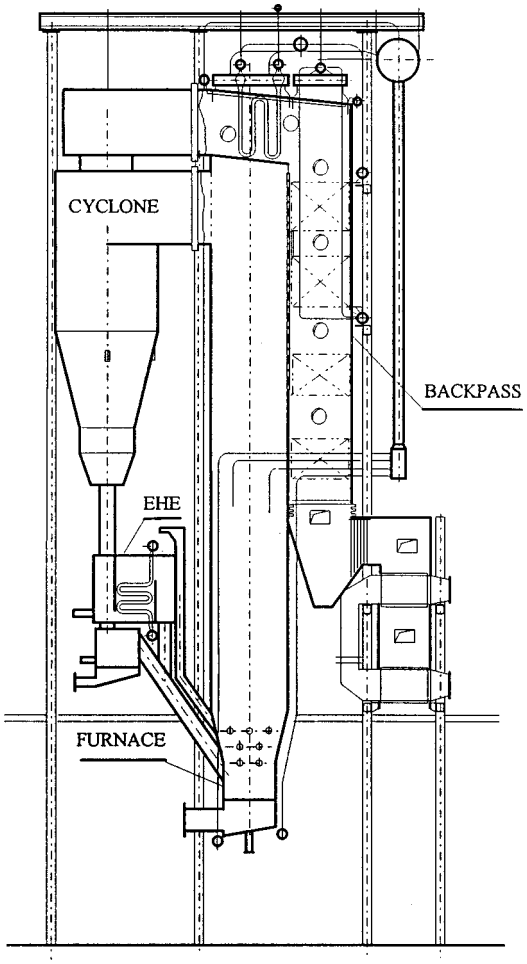


Fig. 1 62-MWt circulating fluidized bed boiler side elevation

(1) Design of the Lower Furnace Zone. The bed material in the lower furnace zone is fluidized by primary air. A narrow cross section of 1.5 m×5.45 m is used in order to keep the bed moderately fluidized, even at a low load, which helps to maintain a similar superficial gas velocity above and below the secondary air level at all operating conditions. Thus it minimizes the risk of agglomeration at low load conditions. Superficial gas velocity was preset as 6 m/s at full load condition. The other important consideration is the minimum pressure drop needed for stable fluidization, especially at part-load conditions. No exposed heat-transfer surface was included within this zone to avoid serious erosion problems. It was lined with SiC plastic castable refractory and SiC prefabricated components. The height of the lower furnace zone is 1.2 m. The heat release rate in the lower furnace was chosen to be 7.5 MW/m^2 on the basis of analyses of similar boilers.

Table 1 Main design parameters for the 62-MWt CFB boiler

No	Design Parameters	Unit	Value
1	Steam flow rate	kg/s	20.83
2	Feed water temperature	°C	150
3	Final superheater outlet temperature	°C	450
4	Final superheater steam pressure	MPa	3.82
5	Bed temperature	°C	900
6	Solids recycle ratio[@]	-	30
7	Stack gas temperature	°C	150

[@] Ratio of recycling solids mass flow to coal feed rate.

Table 2 Ultimate analysis of Pingdingshan coal (as received basis)

Content	C	H	O	N	S	A	M	V	LHV
Unit	%	%	%	%	%	%	%	%	MJ/kg
Value	63.39	3.88	6.51	1.13	0.78	21.62	2.7	25.13	24.54

Note: C, H, O, N, S, A, M, V are carbon, hydrogen, oxygen, nitrogen, sulfur, ash, moisture and volatile matter respectively.

(2) Ratio of Primary to Secondary Air. In a CFB boiler, a significant part of the combustion air is injected into the furnace as secondary air to reduce the NO_x emission. The ratio of primary to secondary was selected as 60/40 in the design. The secondary was injected into the furnace at different heights, i.e., 1.5 m, 2 m, and 2.5 m above the air distributor. Injection velocity was 20 m/s at full load condition.

(3) Design of the Upper Furnace Zone. A membrane wall was chosen as enclosure as well as heat-transfer surface. The designed surface area was 310 m^2. Superficial gas velocity was set as 6.7 m/s in this zone. Finned plates are welded onto the water tubes of the membrane wall to protect from erosion. Also a nickel-based alloy was spray-coated onto the leading row of the high-temperature superheater for erosion protection. The heat release rate chosen for the upper furnace is 5 MW/m^2. This choice was again based on experience with similar boilers.

(4) Furnace Openings. There are several openings in a CFB furnace, depending to some extent on the manufacturing preference. These are used for coal feed, sorbent feed, recycle solids feed (and/or exit to external heat exchanger), primary air inlet, secondary air inlets, start-up burner inlet, and bed drain openings. Their size, location, and number are important design parameters.

Coal Feed Openings. A reduced number of coal feed points is a major feature of a CFB boiler, which is a direct result of the CFB's superior level of mixing. Poor fuel distribution leads to a reduction in combustion efficiency through increased hydrocarbon and CO emissions, and increased calcination heat losses. Nonuniform fuel distribution may further lead to oxygen-deficient reducing zones causing bed agglomeration, slagging problems, and local hot spots within the combustor. Nonuniform fuel distribution also results in increased consumption of sorbent to achieve the same SO_2 emission level and may increase the NO_x generation rate. Therefore, too few feed points may be counterproductive.

To decide the number of coal feed points in a CFB boiler, the following factors should be considered: Fuel properties, including volatile content, reactivity, feed size distribution, and sulfur content and its forms (organic, pyritic, sulfatic). The firing rate per fuel feeder will determine the local concentration of fuel at the feeder outlet; solid mixing in the bed; fuel distribution; cyclone grade efficiency; recycle solids feed points; and cost.

Two parameters may influence the number of feed points in the CFB boilers. They are firing rate q_1 and mixing property q_2. They can be calculated as the following equations:

$$q_1 = BQ/nA_f \tag{1}$$

$$q_2 = A_f/n \tag{2}$$

where B is the coal firing rate, Q is the heating value of coal, n is the number of coal feeders, and A_f is the cross-section area of the

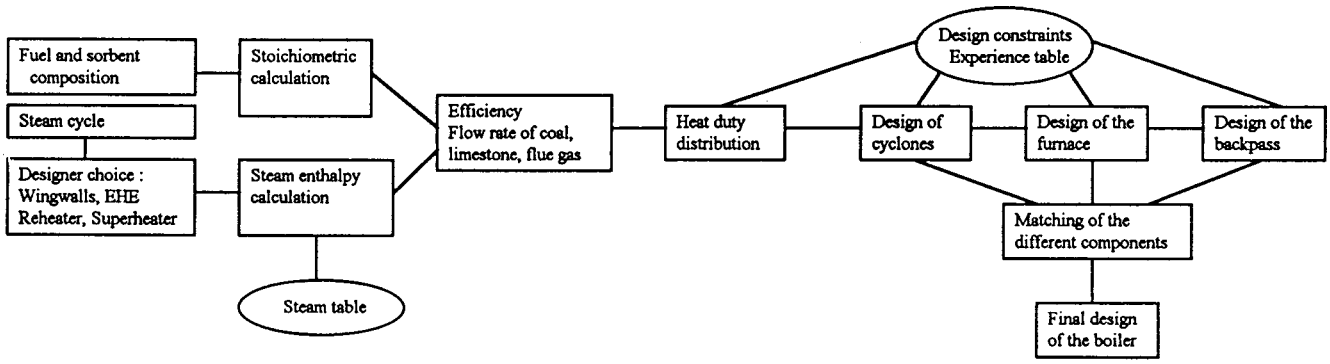


Fig. 2 Design procedure used by CFBCAD©

furnace. The feeder size is easy to determine after the amount of coal fired and the number of coal feed points is known. Based on experience, the values chosen for q_1 and q_2 were 3.75 MW/m^2 and 4.1 m^2, respectively. Two coal feed ports were chosen, positioned on the front wall of the boiler. This gave an area served and heat input per feed point as 4.1 m^2 and 30.7 MW, respectively.

The coal feed ports are located in the lower furnace zone, as far below the secondary air inlet points as possible. Also, they were distributed around the furnace for good particles' mixing and longer residence time in the lower furnace zone before the coal particles are entrained in the high-velocity gas region. The coal feed ports are located at height 1.3 m above the air distributor and 1.5 m below the first row of secondary air ports.

Sorbent Feed Openings. The sorbent feed point is less critical than the coal feed points because of the much slower reaction rate of limestone compared to that of coal combustion. However, sorbent distribution is important to ensure a uniform concentration of unreacted CaO in the CFB at the location where SO_2 is released. There are four methods of feeding sorbent into the furnace: Feed it into the coal feeders; feed it into the recycle system; feed it into the furnace directly; a combination of the above methods. The main consideration is to achieve good mixing with the coal. In the present design, the limestone is premixed with the coal before it is fed into the furnace. The mixture is then fed into the furnace in the coal feeders.

Recycle Solid Feed Openings. Two solid recycle ports underneath the cyclone were arranged on the front wall, 0.8 m above the air distributor but below the first row of secondary inlet ports. Such an arrangement allows the solids to mix well in the furnace before they are entrained by the high-velocity gas. Thus, their residence time in the lower furnace zone is increased. The solid recycle system works on the principle of a pressure balance between the solid return leg and the furnace pressure above the solid recycle point. These criteria were checked when determining the height of the solid recycle point.

The size of the solid recycle port is decided by the external solid circulating rate. Luo et al. (1989) recommended that the solid velocity, V_s, in the standpipe should be less than 0.5 m/s, where $V_s = 1.6 d_s^{0.5}$ m/s and d_s is the diameter of the standpipe.

Primary Air Inlet Port. The primary air enters the base of the bed through the air distributor. The velocity through a directional air nozzle on the distributor was designed as 40 m/s. The pressure drop of each nozzle is 1500 Pa at full load condition.

Secondary Air Inlet Port. Staged combustion is an important feature of CFB boilers. Therefore, combustion air enters the furnace as both primary and secondary supplies. It also has the advantage of reducing the fan power requirement and helps to create a thermal storage at the bottom of the furnace. Low velocity results in a dense bed. So a large fraction of the bed solids remains in this insulated lower section of the furnace. Thus, when there is load demand, an increased air velocity can send greater amounts of solids to the upper heat-absorbing section of the furnace. The best location for the secondary air inlet ports is based upon a combination of these aspects. Some designers prefer to inject the secondary air in several heights above the distributor. This may increase the oxygen concentration in stages for better coal burning and maintain a low oxygen concentration in the furnace for minimization of NO_x formation. Therefore, the secondary air is fed separately into the furnace at elevations 1.5 m, 2 m, and 2.5 m above the distributor. A total of 14 ports are used, with 7 ports on each side wall. A lower ratio of secondary to primary air would reduce the penetration of the secondary air into the furnace. The velocity of secondary air should be increased if a deeper penetration of secondary air into the furnace is desired. In the present design, the penetration depth of the secondary air was observed to be about 1 m at 20 m/s injection velocity in the tests.

Bed Drain. The bed drain extracts bed solids from the lowest section of the bed. It serves two purposes: (1) To maintain the required level of solid mass in the bed, (2) to maintain the size distribution of solids. The size of a bed drain can be determined based on the bottom ash flow produced in a CFB boiler. Each bed drain point may serve 4–9 m^2 of bed area. In the present design, two bed drains were located very close to the air distributor in the rear water wall of the furnace. Directional air nozzles oriented towards the bed drain were used for fluidizing the bed material as well as for pushing the bed ash to the drains. With this design, a bottom ash cooler was arranged at the side of the boiler instead of at the bottom. It saved the height between the bottom of the boiler and the ground.

(5) Air Nozzle. Directional air nozzles were selected in the design. Unidirectional air jets directed the ash towards the two drains at the rear of the furnace.

(6) Start-Up. Oil or natural gas can be used to start up a CFB boiler. There are two methods to start up a CFB boiler.

(a) Under Bed Start-Up. The primary air is heated by an in-duct burner. The high-temperature gas passes through the distributor and then through the solids in the bed. Since the distributor is subjected to a high-temperature stress, it must be cooled.

(b) Above Bed Start-Up. Here an oil or natural-gas flame hits the bed from the top. The hot gas does not contact the distributor. So there is no problem associated with the overheating of the distributor, but only a fraction of the burner heat reaches the solids. In the present design, an in-duct oil burner is located in the primary air duct at the entrance to the combustion chamber. It preheats the incoming air to the windbox. The fuel is fed into the furnace only when the bed temperature is above the fuel ignition temperature.

(7) Load Control and Recycle System. The solids elutriated from the furnace are separated in two streams. Each passes through a 3 m diameter (ID) high-temperature cyclone. An external heat exchanger (EHE) is used to combine the duty of heat transfer from the recycle solids and a solid recycle device. It is designed for easy control of the heat load during the boiler operation. Heat absorption by the heat exchanger tubes can be easily controlled by bypassing the solids through the EHE. Two loop-seal devices that recycle solids back to the furnace are located down stream of the EHE.

Comparison of Design Between Manual and Expert System

The Expert System, CFBCAD©2.1 [3], was used to regenerate and evaluate the manual design. Required input parameters are given in Tables 1 and 2. The Expert System developed design results for the 62-MWt CFB boiler. Design results from the Expert System are discussed below.

(1) Heat Duty Distribution. Feed water enters the economizer at 150°C. Steam is raised in an evaporator in the furnace. It is then superheated to 450°C at 3.82 MPa in a superheater. The combustion air is preheated in an air heater using the sensible heat of the flue gas leaving the economizer. Heat absorption in the furnace and EHE maintains the design furnace temperature. There are several considerations used to arrive at this design decision [4]. The most important of these is to assign the entire evaporative heat load to the furnace.

Figure 3 compares the heat duty distributions of water wall, superheater, economizer, EHE, and air heater computed on the basis of manual design and Expert System design. The Expert System made greater use of the furnace water wall and less on the superheater, economizer, and EHE. The difference is within few percentage points, which is a result of the small difference in design decisions.

(2) Furnace Cross Section and Height. The Expert System gave a cross-section area 2.54×5.5 m^2, which was 4.6 percent larger than that computed from the manual design, for a fixed furnace height at 20 m. The Expert System used a heat release rate of bed area as the guiding factor while the manual design used fluidization velocity as the guide. Both approaches are valid.

(3) Heat-Transfer Surface. Heat-transfer surface areas computed manually and by CFBCAD©2.1 [3] are compared in Fig. 4. Results for water wall and EHE area are similar. The small difference is a reflection of the difference in heat duty distribution. However, differences in heating surface areas for the superheater, economizer, and air heater are significant. These components are

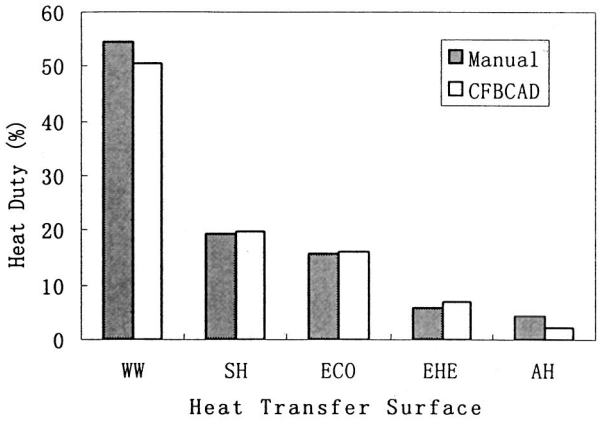


Fig. 3 Heat duty distribution by manual design and the Expert System

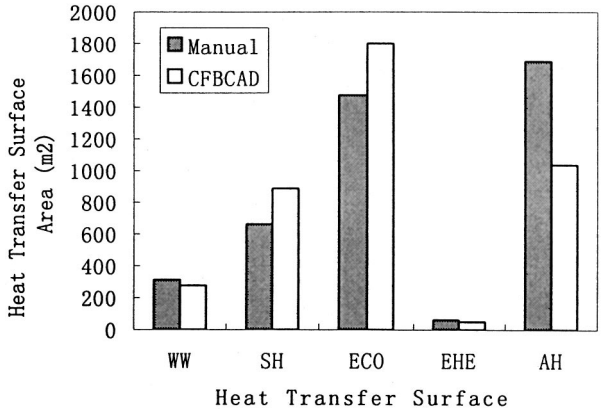


Fig. 4 Heat-transfer surface area by manual design and the Expert System

located in the backpass. Considering the potential erosion for a tube, a relatively low gas velocity (7.5 m/s) was selected in the backpass in the manual design. This gave a lower heat-transfer coefficient compared to that obtainable from 12 m/s gas velocity chosen by the Expert System. Thus a larger heat-transfer surface area is required in the manual design. But this trend is not noticed for the economizer and the superheater. The design method of a CFB furnace is less mature than that for the backpass. Some of the difference shown in Fig. 4 may come from different designer's preferences.

Operation Experience

Both manual and Expert System designs are theoretical exercises. Even a perfect agreement in the designs and performance predictions has little value if the boiler built to these designs fails to attain the specified performance. While the foregoing discussions verified details of the design, the real test of their validity lies in the actual performance of the boiler when built. The 62-MWt CFB boiler, designed above, was constructed and commissioned in 1995. It is operated in the Jiangshu Province of China on the design Pingdingshan coal. Operating results have shown the boiler to have wide fuel flexibility, easy load control, and high combustion efficiency [1]. The following section discusses the final test of the design methods. After commissioning, a comprehensive test program was undertaken in April 1996. After one year of operation, all teething problems were solved and the boiler was stable and operating continuously. A variation of different operating parameters was attempted. These included variation in coal, heat load, bed temperature, and superficial gas velocity. Five different coals were tested, with LHV ranging from 11.3 MJ/kg to 26 MJ/kg and the volatile content varying in the range of 8.7 percent to 33.95 percent. The boiler operated under different heat loads varying from 115 percent to 36 percent. The boiler efficiency was about 88 percent and the combustion efficiency reached 97 percent at full load in the tests [1].

Figure 5 shows heat duty distributions for different heat-transfer surfaces at different heat loads. The designed heat absorption in the water wall was 55 percent (Fig. 3). The measured heat absorption in the water wall at full load (100 percent) is also close to 55 percent (Fig. 5). Experiments show that heat transfer to the water wall increases, while that to the other heat-transfer surfaces (apart from the air heater) decreases with decreasing load. As the load decreased, the air flow also decreased, along with a reduction in the coal feed rate. Secondary air is the main adjusting parameter, as primary air has to be kept constant to maintain the bed at minimum fluidization. The suspension density profile along the bed height in the furnace changes as the primary to secondary air ratio changes. More solids drop out of the upper furnace zone to fall in the lower furnace zone as the gas velocity decreases. This

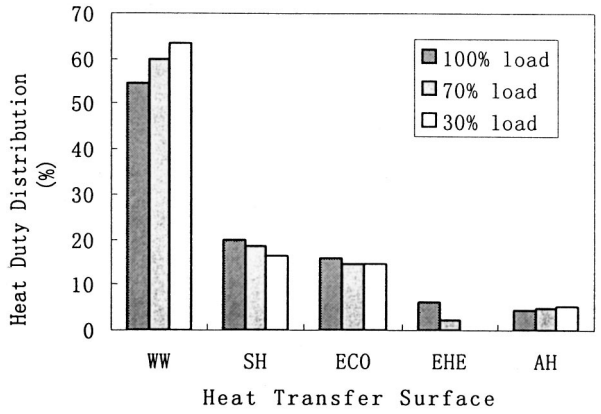


Fig. 5 Heat duty distribution to different heat-transfer surfaces at different boiler loads

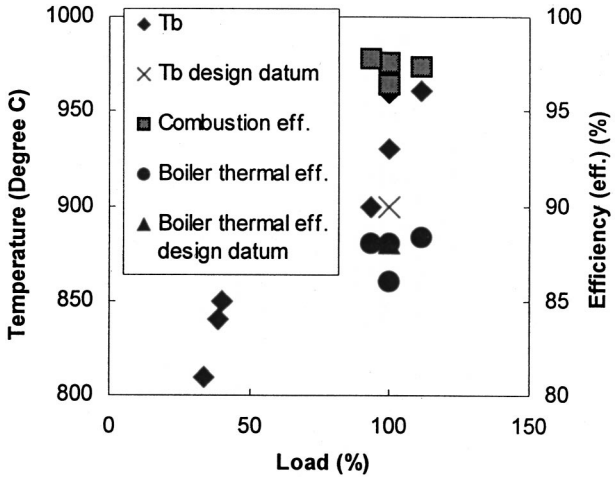


Fig. 6 Combustion efficiency, boiler efficiency, and bed temperature variations with boiler load

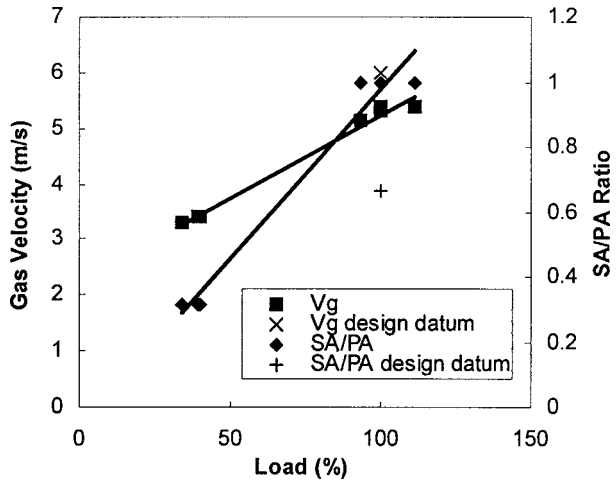


Fig. 7 Superficial gas velocity and SA/PA ratio variations with boiler load

reduces the voidage in the lower furnace and the carryover rate to the upper furnace. Since the heat-transfer coefficient depends on solid concentration, the heat-transfer coefficients in the upper furnace zone decrease when the secondary air decreases [6]. Thus, the heat absorbed in the lower furnace increases. A comparison of heat duty distribution at different loads is shown in Fig. 5.

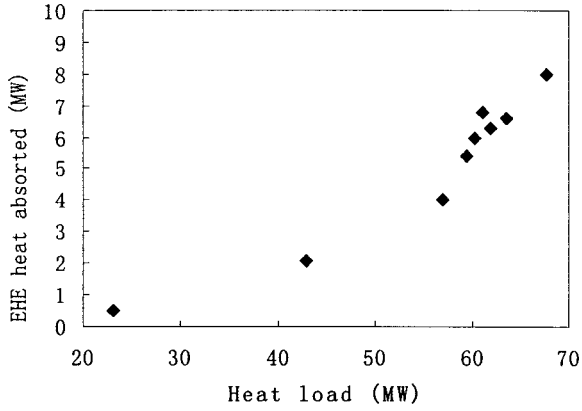


Fig. 8 Heat absorbed by the EHE during boiler load variations

The bed temperature increased with load as more coal was fed into the furnace. However, the combustion efficiency and boiler thermal efficiency did not change much in the tests. It is shown in Fig. 6.

As mentioned above, during load reduction the air flow rate decreased along with the decreasing feed rate of coal. Figure 7 shows how superficial gas velocity in the furnace decreases with reducing load. It is also apparent that the ratio of secondary to primary air decreases with reducing load as the secondary air is the main adjusting parameter in air flow control. The EHE showed its good load-adjusting capability during the comprehensive test run of the boiler. It played an important role in the boiler load variations from 115 percent to 36 percent. Figure 8 gives the heat absorbed by the EHE at different boiler load [7].

Conclusion

Design considerations and method are discussed for a 62-MWt CFB boiler based on experience and available published data. Predicted performance from a manual design procedure and from an Expert System compares favorably with those measured from the boiler in operation. This validates the usefulness of the Expert System for design of a CFB boiler. Comprehensive test results of the boiler show that the design was successful in meeting its performance objectives. Reasonable agreement between the predicted performance and measured values was obtained.

Acknowledgments

The project was supported in part by a collaborative project between DalTech Dalhousie University (formerly the Technical University of Nova Scotia) and Zhejiang University funded by the Canadian International Development Agency. The authors also acknowledge people involved in the comprehensive test program of the boiler.

Nomenclature

A_f = cross section area of furnace, m^2
B = coal firing rate, kg/s
d_s = standpipe diameter, m
n = number of coal feeders
Q = heating value of coal, kJ/kg
q_1 = firing rate, kJ/m^2
q_2 = mixing property, m^2
T_g = bed temperature, °C
V_g = superficial gas velocity, m/s
V_s = solid velocity in standpipe, m/s

References

[1] Fang, M. X., Luo, Z. Y., Li, X. T., Wang, Q. H., Shi, Z. L., Ni, M. J., and Cen, K. F., 1997, ''Development of the First Demonstration CFB Boiler for Gas and Steam Cogeneration,'' *Proceedings of the 14th International Conference on Fluidized Bed Combustion*, F. D. S. Preto, ed., Vancouver, BC, Canada, May 11–14, ASME, New York, NY, pp. 663–669.

[2] Basu, P., and Mitra, S., 1994, ''Application of an Expert System to the Design of Furnace of a Circulating Fluidized Bed Boiler,'' ASME J. Eng. Gas Turbines Power, **116**, pp. 462–467.

[3] CFBCAD©, an Expert System, 1996, Greenfield Research Inc., P.O. Box 25018, Halifax, Canada B3M 4H4.

[4] LaFanechere, L., Basu, P., and Jestin, L., 1998, ''Use of an Expert System to Study the Effect of Steam Parameters on the Size and Configuration of Circulating Fluidized Bed Boilers,'' ASME J. Eng. Gas Turbines Power, **120**, pp. 861–866.

[5] Basu, P., and Fraser, S. A., 1991, ''Design Considerations,'' *Circulating Fluidized Bed Boiler—Design and Operations*, Butterworth-Heinemann, Boston, pp. 174–184.

[6] Tang, J. T., and Engstrom, F., 1987, ''Technical Assessment on the Ahlstrom Pyroflow Circulating and Conventional Bubbling Fluidized Bed Combustion Systems,'' *Proceedings of the 9th International Conference on Fluidized Bed Combustion*, J. P. Mustonen, ed., Boston, May 3–7, ASME, New York, pp. 38–54.

[7] Luo, Z. Y., Ni, M. J., Zhou, J. H., Cheng, L. M., Chang, Z. J., and Cen, K. F., 1989, ''Solids Recycle System for Circulating Fluidized Beds,'' *Proceedings of the 10th International Conference on Fluidized Bed Combustion*, A. M. Manaker, ed., San Francisco, April 30–May 3, ASME, New York, pp. 557–562.

Dynamic Optimization of Startup and Load-Increasing Processes in Power Plants—Part I: Method

J. Bausa[1]
e-mail: bausa@olymp.fb10.tu-berlin.de

G. Tsatsaronis
e-mail: tsa@olymp.fb10.tu-berlin.de

Institute for Energy Engineering,
Technical University of Berlin,
Marchstrasse 18,
10587 Berlin, Germany

The fast startup and load-increasing process of power plants is a complex task involving several restrictions that have to be fulfilled simultaneously. An important restriction is the maximum allowed thermal stress of the steam generator pipes and the steam turbines caused by temperature gradients. In this paper the startup process is treated as a dynamic optimization problem. Any appropriate objective function can be used in this optimization problem. Examples include the minimization of fuel consumption or the minimization of the time required to reach the desired load. The maximum allowable temperature and pressure gradients in major plant components appear as additional constraints. In this paper a general method for solving these problems is presented: The dynamic process model, consisting of first-order ordinary differential equations (ODEs) and algebraic equations, is discretized over the time horizon using well established methods for the solution of ODEs. Thus, the continuous dynamic optimization problem is transformed into a large-scale non-linear parameter optimization problem with up to 20,000 optimization parameters and constraints. Such parameter optimization problems can be solved with appropriate sequential quadratic programming (SQP) methods that have become available lately. An application of this method is presented in the second part of this study by optimizing the process of rapid load increase in a single-pressure combined-cycle power plant on the basis of a simplified model. [DOI: 10.1115/1.1286728]

1 Introduction

The design of power plants focuses on steady-state operation and high efficiency at full load. However, the startup and shutdown procedures in a plant are events that occur quite frequently. Even large coal fired power plants that have been designed originally for base load operation are reported to go through about 100 startups per year [1]. The control of these startup procedures is very important because the life span of power plants is severely affected by thermal stresses caused by temperature gradients in the thick parts of the steam generator piping and in the steam turbines. Considering the plant safety as the first priority can result in a relatively slow startup procedure that takes a long time to reach the desired load. On the other hand, the startup procedure should, for economic reasons, be accomplished in the shortest time with a minimum fuel consumption. The same applies to load-increasing procedures, particularly when a significant load increase is required, as for example, in the morning after the load of a power plant has been reduced to low load over night.

This discussion shows that the startup and load-increasing processes of power plants can be treated as a dynamic optimization problem: find a control profile for each main control variable (e.g., fuel flow rate, steam turbine inlet valve position) that minimizes an appropriate objective function (e.g., time and/or fuel consumption required to reach the desired load) and satisfies the constraints on the temperature gradients.

The conventional method for generating admissible control profiles makes use of a detailed computer simulation of the power plant performance. The input-output behavior of the plant is studied for different control profiles. In an iterative manner the control profiles are modified to satisfy the constraints and to meet the objectives in a better way. Usually, the control profiles are defined over finite elements of the time horizon as ramps with different slopes [2]. To apply some numerical optimization methods, the slopes of each ramp can be treated as parameters of an optimization problem [3]. The main disadvantage of this method is that only control profiles consisting of different ramps are considered, although actually the structure of the control profiles should be left totally free. If the optimal control profile, for example, is a polynomial of second order or has a discontinuous step, it will never be found but only approximated with ramps causing the startup or load-increasing process to take longer time and to consume more fuel.

In this paper the dynamic processes of startup and load increasing is treated in the most general way as a dynamic optimization problem. In the last decade, methods have been developed that are applicable to complex nonlinear dynamic optimization problems. These methods have been developed for optimal control problems in chemical engineering like batch-reactor and batch-distillation control [4–7] but can be applied to any dynamic system described by a differential algebraic equation system (DAE) that consists of first-order ordinary differential equations (ODEs) and algebraic equations.

There are two classes of methods: the parameterization of the control variables on the one hand and the parameterization of both the control *and* the state variables on the other hand. In this work the second method, which is also known as the collocation method, is applied. After the discretization, a large scale nonlinear parameter optimization problem has to be solved. Only few computer codes are known to be able to solve problems with up to 20,000 variables and constraints efficiently (see, e.g., Gill et al. [8]; Betts and Frank [9]). This task was done with the SQP-code SNOPT developed by Gill et al. [8].

In the next section a brief description of the numerical methods is given and some new findings about appropriate collocation methods are presented. In the third section the software implementation is described. An application of this software is presented in the second part of this study [10] that discusses the optimization of the process of rapid load increase in a single-pressure combined-cycle power plant.

[1]Author to whom all correspondence should be addressed. Present address: BASF AG, DWF/AM, L440, 67056 Ludwigshafen, Germany.

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division May 26, 1999; final revision received by the ASME Headquarters December 29, 1999. Technical Editor: D. Assanis.

2 Mathematical Methods

2.1 Conditions of Optimality for a Continuous Optimization Problem. In our work we consider fully implicit DAE systems that are described by ordinary differential equations and algebraic equations. To just demonstrate some basic properties of dynamic optimization problems, we will first discuss a problem with an explicit ordinary differential equation system (see Bryson and Ho [11] for details):

$$\dot{x} = f(x,u,t), \tag{1}$$

where x is the n_x-component state vector, u is the n_u-component control vector and t is the time.

The dynamic optimization problem is now given by the minimization of the functional J

$$\min_{u(t)} J = \hat{P}(x(t_f),u(t_f),t_f) + \int_{t_0}^{t_f} L(x,u,t)\,dt \tag{2}$$

that consists of the function $\hat{P}$ at the end of the time interval being considered (t_f) and the integral of L over the time horizon.

After introducing the Hamiltonian

$$H = L + \lambda^T(t) f, \tag{3}$$

where L is expanded by the product of the system functions f and the time-dependent Lagrangian multiplier λ, we can write the Euler-Lagrange equations as follows:

$$\dot{\lambda}^T = -\frac{\partial H}{\partial x} = -\frac{\partial L}{\partial x} - \lambda^T \frac{\partial f}{\partial x} \tag{4}$$

$$\frac{\partial H}{\partial u} = \frac{\partial L}{\partial u} + \lambda^T \frac{\partial f}{\partial u} = 0. \tag{5}$$

Equation (5) gives the optimality condition for the optimization problem. Unfortunately, we only know the initial values of the state vector x and the values of the co-state vector λ at the final time t_f:

$$x(t_0) = x_0 \tag{6}$$

$$\lambda(t_f) = \left.\frac{\partial \hat{P}}{\partial x}\right|_{t_f}. \tag{7}$$

Equations (1) through (7) describe a two-point boundary-value problem that is very difficult to solve numerically because guess values must be initially assigned to the co-state vector $\lambda(t_0)$ which is an abstract variable without physical meaning. With multiple-shooting relatively small problems can be solved very precisely (see, e.g., Bulirsch [12]; Deuflhard [13]; Deuflhard and Bader [14]). However, this method does not seem to be appropriate for large nonlinear systems describing power plants.

2.2 Discretization Methods. The discretization methods can be divided into two classes: In the first class only the control variables are parameterized, whereas in the second class the state and control variables are fully discretized using collocation techniques. Both discretization methods can be applied to the dynamic optimization of fully implicit DAE systems with inequality constraints and constraints at the final time point:

$$0 = f(\dot{x},x,u,t) \tag{8}$$

$$0 \leqslant h(\dot{x},x,u,t) \tag{9}$$

$$0 = \psi(\dot{x}(t_f),x(t_f),u(t_f),t_f) \tag{10}$$

$$x(t_0) = x_0 \tag{11}$$

$$\min_{u(t)} J = P(x(t_f),u(t_f),t_f). \tag{12}$$

In this notation, f includes all ordinary differential equations and all algebraic equations of the DAE system. The state vector x consists of all variables appearing in the differential and the algebraic equation, i.e., in Eq. (8) not all components of vector x appear as derivatives. Note, that the objective function P depends only on the values of the variables at the final time point. Without any loss of generality, this is essentially the same as the functional given in Eq. (2) if we add one more state to the state vector x and the function $\dot{x} = L(x,u,t)$ to f.

Control Parameterization Versus Collocation. Control parameterization methods discretize the control variables using a finite set of parameters, for example, using low order polynomials for each finite element of the time horizon. The system of continuous equations is still integrated with an appropriate step-size controlled integration method; this is clearly an advantage of this method compared to the full discretization by collocation where the control of the integration error is difficult.

Through the parameterization of the control variables, the continuous optimization problem becomes a finite optimization problem with relatively few optimization parameters over the full discretization by collocation. The only problem is to calculate the gradients of the objective function with respect to the parameters of the control variables. This can be done by deriving the sensitivity equation system that has to be integrated simultaneously to the DAE system (see Vassiliadis et al. [6]; Barton et al. [7] for details). The sensitivity equation system is an ordinary differential equation system of high order and has to be solved with high accuracy to provide accurate gradient informations for the optimization with an SQP-solver. This can be viewed as a disadvantage because it significantly increases the calculation time. Furthermore, path constraints on the state variables, like the maximum value of a temperature gradient, are not easy to implement. For these reasons the control parameterization has not been considered, but collocation methods have been applied.

Collocation methods discretize both the state variables and the control variables over the time horizon. The main idea of collocation is that the state variables are substituted by polynomial expressions that have to satisfy the differential equations at the discrete time-steps, the so-called collocation points. This means that every continuous equation given in the system of Eqs. (8)–(9) has to be applied at each discrete time-step with a polynomial approximation of the time derivatives $\dot{x}$ in f, h and ψ. Every control and state variable at every time-step is treated as an optimization variable. Additional constraints like the maximum value of a temperature gradient can easily be added to the set of equations. In contrast to the control parameterization, where the simulation steps and the optimization steps follow each other in a sequential manner, the collocation method solves the simulation problem and the optimization problem simultaneously.

Discretization Methods for ODEs. The type of polynomials and the distribution of the discrete time-steps is strongly related to the methods for solving ODEs. For this reason we want to discuss some classes of methods that have been applied so far and to very briefly present some results. Detailed derivations will be in a subsequent paper.

Figure 1 shows the basic concept of the collocation at Gaussian points: the differential equation is collocated at some interior points and the last one is extrapolated. For the solution of ordinary differential equations each finite element is handled separately. For optimal control problems all finite elements are handled simultaneously. It can be shown that the collocation at Gaussian points is basically the same as an implicit Runge-Kutta method and gives the best accuracy for a fixed number of discrete time-steps. This collocation scheme has been mainly used for the discretization of optimal control problems. However, our findings show clearly that the extrapolation of the last point in each finite time element can cause oscillations of the control and state variables and that other discretization schemes are much more suitable for many systems.

In contrast to the collocation at Gaussian points the Radau

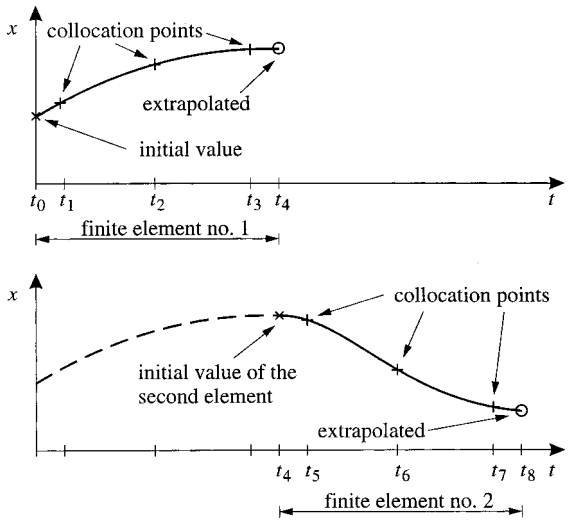


Fig. 1 Fourth order collocation at Gaussian points over two finite elements. The last point of each finite time element is extrapolated.

Runge-Kutta scheme applies the differential equation at every discrete time-step, Fig. 2. It has been found that this discretization method behaves much better than a collocation at Gaussian points and has less problems associated with oscillations. Compared to Gaussian points, the Radau method remains stable even for large step sizes and stiff differential equations [15].

Another discretization method that has been considered uses the backward differentiation formulas (BDF). This method is well known for the efficient solution of DAE-systems (Brenan et al., [16]). Unlike the single-step Runge-Kutta methods, the multi-step methods consider some values in the past, as shown in Fig. 3: the fourth-order BDF-method makes use of three points in the past. To start this method the implicit Euler method is applied in the first time-step, followed by a second and third order BDF, respectively. In the literature the BDF-method has not been applied to optimal control problems so far. Nevertheless, our studies show that BDF-methods have some advantages compared to the Runge-Kutta methods, especially if the control profile does not have too many discontinuities. Furthermore, optimization problems that

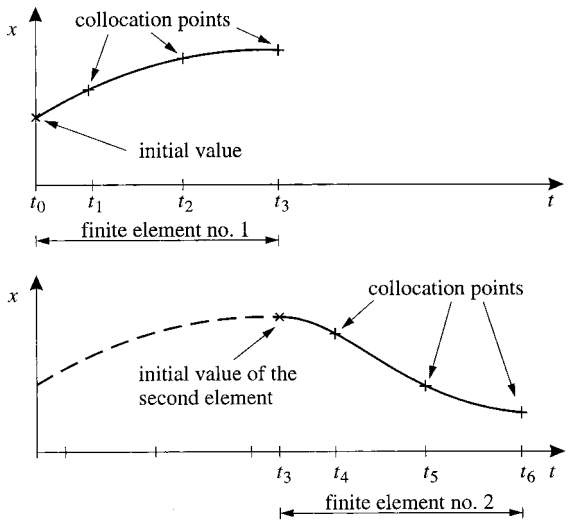


Fig. 2 Fourth order collocation at Radau points over two finite elements. All points are collocated and satisfy the differential equation.

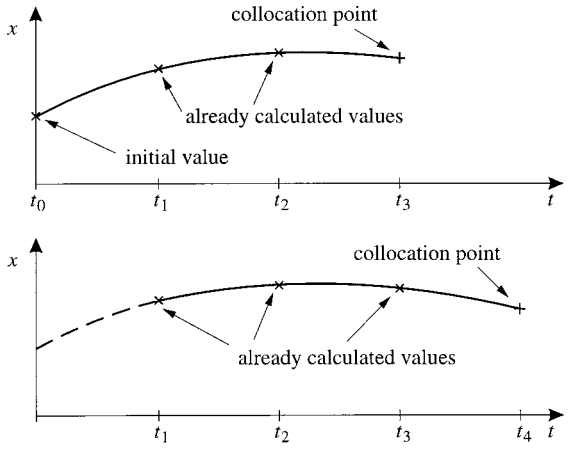


Fig. 3 Multistep BDF-method. Each new point is calculated on the base of some already calculated points.

have been discretized using the BDF-method tend to converge faster than problems discretized with the Radau-Runge-Kutta approach. The BDF-method is accurate enough, robust, and appropriate for solving the problem that is presented in the second part of this study [10]. Therefore, all optimizations in the second part are performed with second order BDF-discretization.

Lagrangian Polynomials. As mentioned before, the state variables must be expressed by polynomials on finite elements. For the discretization of optimal control problems it is common to use Lagrangian polynomials. The formulas for the first finite time element are

$$\phi_j(t)=\prod_{\substack{k=0\\k\neq j}}^{K}\frac{t-t_k}{t_j-t_k},\quad j=0,1,\ldots,K \tag{13}$$

$$x(t)=\sum_{k=0}^{K}\phi_k(t)x_k, \tag{14}$$

where $K+1$ is the order of the polynomial. The main advantage of this expression compared to the conventional expression $x(t)=a_0+a_1t+a_2t^2+\ldots+a_{K-1}t^{K-1}$ is that the parameters of the polynomials are the values of x at the discrete times t_k. Unlike the conventional parameters a_i, these values have physical meaning and the upper and lower bounds for each variable can be given easily. To have proper bounds for each optimization parameter is important for the convergence of the optimization problem.

If we apply Lagrangian polynomials, the derivatives of the state variables with respect to t can be calculated easily:

$$\dot{x}(t)=\sum_{k=0}^{K}\dot{\phi}_k(t)x_k. \tag{15}$$

At each discrete time-step these derivatives must be inserted into Eqs. (8) and (9), respectively. This results in the fully discretized optimization problem:

$$0=f_D(\hat{x},\hat{u},t_f) \tag{16}$$

$$0\leqslant h_D(\hat{x},\hat{u},t_f) \tag{17}$$

$$0=\psi_D(x_{n_k},u_{n_k},t_f) \tag{18}$$

$$x(t_0)=x_0 \tag{19}$$

$$\min_{\hat{u},\hat{x},t_f} J=P(x_{n_k},u_{n_k},t_f). \tag{20}$$

In this notation the vectors $\hat{x}$ and $\hat{u}$ consist of all vectors x and u at the n_k discrete time-steps:

$$\hat{x}=\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n_k} \end{pmatrix}, \quad \hat{u}=\begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_{n_k} \end{pmatrix}.$$

The discretized equations are of high dimensionality: if we have 10 ordinary differential equations and 10 algebraic equations, i.e., $n_f=20$, with $n_u=2$ control variables that we want to discretize with $n_k=100$ time-steps, we obtain $(n_f+n_g)\times n_k=2000$ algebraic constraints for the optimization problem. In this case we have $n_x\times n_k=2000$ variables in the vector $\hat{x}$ plus the $n_u\times n_k=200$ independent control variables in the vector $\hat{u}$.

Before the nonlinear optimization problem (Eqs. (16)–(20)) can be solved with a large-scale SQP method, the Jacobian matrix with all gradients of the functions f_D, g_D, h_D, ψ_D, and P with respect to $\hat{x}$ and $\hat{u}$ has to be calculated. For the solution of large-scale systems it is crucial to provide exact gradient information. Nevertheless, there are implementations of these methods, where the gradients are only approximated by finite differences.

In the example given above the Jacobian has $(n_f+n_g)\times n_k=2000$ rows and $(n_x+n_u)\times n_k=2200$ columns, i.e., it has 4.4×10^6 elements with a total of more than 34 MByte of memory demand with double precision if it is stored in a dense manner. Fortunately, only few elements in each row are non-zero and sparse matrix techniques can be applied.

3 Software Implementation

It has been shown that the continuous optimization problem (Eqs. (8)–(12)) can be transformed into a parameter optimization problem of high dimension (Eqs. (16)–(20)). Conducting this transformation and calculating all gradients by hand would be a tedious task with many sources of errors. To avoid these problems the software OCOMA (Optimal COntrol Code generator for MAple) has been developed. In OCOMA the discretization methods and the calculation of the exact gradients are implemented with the help of the language for symbolic mathematical calculation MAPLE [17].

During the development of the software two main objectives have been considered: (a) to achieve a maximum ease of use, and (b) to make the optimization process as efficient and robust as possible by exploiting the mathematical structure of the problem.

The basic concept of OCOMA is shown in Fig. 4. The user only has to define the continuous dynamic optimization problem. This is done easily by coding the DAE system, the inequality constraints, the constraints at the final time point and the objective function in a simple syntax. The syntax to be used for this purpose, is illustrated in Table 1 that refers to a part of the optimization problem that will be discussed in the following section: This simple continuous optimization problem describes the steam drum with an initial pressure of 50 bar. The objective is to increase the pressure to 100 bar in minimum time without violating the maximum allowed drum temperature gradient. The control variable of this optimization problem is the heat flow rate in the evaporator $\dot{Q}_{\mathrm{vap}}$. To code the continuous optimization problem, some easy to use functions have been developed to add differential equations (`add_deq`), algebraic equations (`add_aeq`), inequalities (`add_ieq`), constraints at final time (`add_tfeq`) and the objective function (`objective`). In addition to the equations shown in Table 1, bounds for every variable, the number of time-steps, the discretization method (Runge-Kutta, BDF) and the order of the discretization method have to be selected.

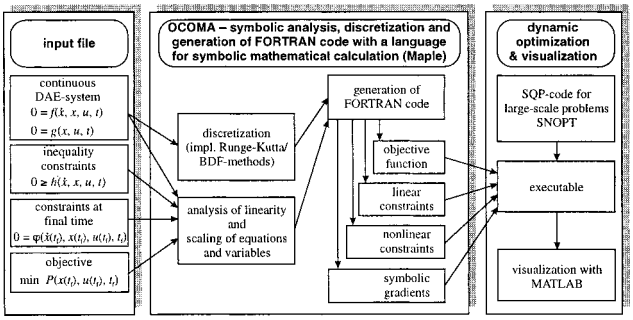


Fig. 4 Basic concept and structure of the OCOMA software developed for dynamic optimization

The main tasks in the OCOMA software are conducted by the MAPLE program (Fig. 4): in the first step all equations and variables are scaled and analyzed for linearity. It is useful to distinguish between linear and nonlinear variables and equations because they are treated differently by the optimization code SNOPT. For example, the Jacobian elements of linear variables are constant and do not need to be updated during the optimization process. Furthermore, SNOPT satisfies the linear constraints in the outer iteration before solving the nonlinear quadratic programming (QP) subproblem in the inner iteration, see Gill et al. [8]. If there are many linear equations, the size of the QP subproblem decreases significantly by distinguishing between linear and nonlinear variables and equations.

In the second step, the continuous system is discretized using the methods that have been discussed in Section 2.2. In the third step, a FORTRAN code consisting of the objective function, the constraints and the gradients is generated automatically. The re-

Table 1 Example for the syntax of the MAPLE input file

Continuous optimization problem	Syntax of the MAPLE input file
$\frac{dp_{\mathrm{drum}}}{dt} = \frac{1}{k_{\mathrm{drum}}}(\dot{m}_{\mathrm{vap}}-\dot{m}_{\mathrm{steam}})$	`add_deq( k_drum*DOT(p_drum) =`
$\dot{m}_{\mathrm{steam}} = k\sqrt{p_{\mathrm{drum}}-0.06}$	`m_vap - m_steam , p_drum );`
$T_{\mathrm{drum}} = -1.18287\times10^{-7}p_{\mathrm{drum}}^4+6.84645\times10^{-5}p_{\mathrm{drum}}^3$	`add_aeq( m_steam = k*sqrt(p_drum - 0.06) );`
$-1.63485\times10^{-2}p_{\mathrm{drum}}^2+2.41939p_{\mathrm{drum}}$	`add_aeq(T_drum =`
$+175.904$	`-1.18287e-7*p_drum^4 + 6.84645e-5*p_drum^3`
$\Delta h_{\mathrm{vap}} = -1.15929\times10^{-7}p_{\mathrm{drum}}^3+3.57007\times10^{-5}p_{\mathrm{drum}}^2$	`-1.63485e-2*p_drum^2 + 2.41939 *p_drum`
$-9.73352\times10^{-3}p_{\mathrm{drum}}+2.05198$	`+ 175.904);`
$\dot{m}_{\mathrm{vap}} = \dot{Q}_{\mathrm{vap}}/\Delta h_{\mathrm{vap}}$	`add_aeq(dh_vap =`
$p_{\mathrm{drum}}(t=0) = 50\,\mathrm{bar}$	`-1.15929e-7*p_drum^3 + 3.57007e-5*p_drum^2`
$\left\lvert\frac{dT_{\mathrm{drum}}}{dt}\right\rvert \leq 2^{\circ}\mathrm{C/min}$	`-9.73352e-3*p_drum + 2.05198);`
$\min_{\dot{Q}_{\mathrm{vap}}} J = \int_{t=0}^{t_f=1000s}(p_{\mathrm{drum}}-100)^2dt$	`add_aeq(dh_vap*m_vap = Q_vap);`
	`add_initial( p_drum , 50 );`
	`add_ieq(-2.0/60, DOT(T_drum), 2.0/60 , T_drum);`
	`add_deq( DOT(J) = (p_drum -100)^2 , J );`
	`add_initial( J , 0 );`
	`T_FINAL := 1000;`
	`objective(J);`

sulting FORTRAN files are quite large: For a problem with about 10,000 constraints and variables the FORTRAN code consists of about 200,000 lines of code.

The resulting files are compiled and linked with the optimization code SNOPT to an executable file. After the optimization, the results are printed out in a format that can be read and visualized with the numerical software package MATLAB.

4 Conclusions

In this paper some improvements to a general method for the solution of dynamic optimization problems have been presented. An efficient, robust and easy to use computer-based procedure for solving various dynamic optimization problems has been developed. The software allows to solve dynamic optimization problems consisting of a DAE-system with up to approximately 200 equations. Because of the general mathematical formulation of the problem, the software is not restricted to energy engineering, but can be applied to solve dynamic optimization problems in different engineering disciplines such as chemical and aerospace engineering. The application of the software to the process of rapid increase of power in a single-pressure combined-cycle is presented in the second part of this study.

Nomenclature

a_i = polynomial coefficient
c_p = heat capacity [MJ/(kg K)]
f = continuous function
H = Hamiltonian
h = inequality constraints
h = mass specific enthalpy [MJ/kg]
J = objective function
L = performance function
$\dot{m}$ = mass flow rate [kg/s]
n_f = number of differential equations
n_g = number of algebraic equations
n_k = number of time steps
n_u = number of control variables
n_x = number of state variables
P = performance function at t_f
p = pressure [bar]
$\dot{Q}$ = heat flow rate [MW]
T = temperature [K]
t = time [s]
u = control vector
x = state vector
λ = co-state vector
ϕ = Lagrangian polynomial
ψ = constraints at the final time point

Subscripts

D = discretized system
f = final
vap = evaporation

References

[1] Lehmann, L., Klein, M., and Wittchow, E., 1996, ''Der Block 4 im Kraftwerk Heyden-Referenzanlage für moderne 1000 MW-Steinkohle-Kraftwerke,'' VGB Kraft., **76**, No. 2, pp. 85–94.

[2] Beato, A., Bolis, V., Colombo, P., De Marco, A., Quatela, S., Zizzo, A., and Marziali, F., 1992, ''Study of the Start-Up Control of a Drum Boiler by Means of a Mathematical Model,'' *IFAC Control of Power Plants and Power Systems*, Munich, pp. 115–121.

[3] Akiyama, T., Matsumoto, H., and Asakura, K., 1996. ''Dynamic Simulation and its Application to Optimum Operation Support for Advanced Combined Cycle Plants,'' *ECOS'96: Efficiency, Costs, Optimization, Simulation and Environmental Aspects of Energy Systems*, P. Alvfors, L. Eidensten, G. Svedberg, and J. Yan, eds. Stockholm, pp. 337–344.

[4] Biegler, L. T., 1984, ''Solution of Dynamic Optimization Problems by Successive Quadratic Programming and Orthogonal Collocation,'' Comput. Chem. Eng., **8**, No. 3, pp. 243–248.

[5] Cervantes, A. M., and Biegler, L. T., 1998, ''Large-Scale DAE Optimization Using a Simultaneous NLP Formulation,'' AIChE J., **44**, No. 5, pp. 1038–1050.

[6] Vassiliadis, V. S., Sargent, R. W. H., and Pantelides, C. C., 1994, ''Solution of a Class of Multistage Dynamic Optimization Problems: Part II—Problems With Path Constraints,'' Ind. Eng. Chem. Res., **33**, pp. 2123–2133.

[7] Barton, P. I., Allgor, R. J., Feehery, W. F., and Galán, S., 1998, ''Dynamic Optimization in a Discontinuous World,'' Ind. Eng. Chem. Res., **37**, pp. 966–981.

[8] Gill, P. E., Murray, W., and Saunders, M. A., 1997, ''SNOPT: An SQP Algorithm for Large-Scale constrained Optimization,'' Tech. Rep. NA 97-2, Department of Mathematics, University of California, San Diego.

[9] Betts, J. T., and Frank, P. D., 1991, ''A Sparse Nonlinear Optimization Algorithm,'' Tech. Rep. AMS-TR-173, Boeing Computer Services.

[10] Bausa, J., and Tsatsaronis, G., 2001, ''Dynamic Optimization of Start-Up and Load-Increasing Processes in Power Plants: Part II—Application,'' ASME J. Eng. Gas Turbines Power, **123**, pp. 251–254.

[11] Bryson, A. E., and Ho, Y.-C., 1988, *Applied Optimal Control*, Hemisphere, Washington, DC.

[12] Bulirsch, R., 1971, ''Die Mehrzielmethode zur numerischen Lösung von nichtlinearen Randwert-problemen und Aufgaben der optimalen Steuerung,'' Tech. rep., Carl-Cranz-Gesellschaft, Oberpfaffenhofen, Germany.

[13] Deuflhard, P., 1974, ''A Modified Newton Method for the Solution of Ill-Conditioned Systems of Nonlinear Equations with Application to Multiple Shooting,'' Numer. Math., **22**, pp. 289–315.

[14] Deuflhard, P., and Bader, G., 1982, ''Multiple Shooting Techniques Revisited,'' Tech. Rep., Univ. Heidelberg, SFB 123.

[15] Butcher, J. C., 1986, *The Numerical Analysis of Ordinary Differential Equations: Runge-Kutta and General Linear Methods*, John Wiley and Sons, New York.

[16] Brenan, K. E., Campbell, S. L., and Petzold, L. R., 1989, *Numerical Solution of Initial-Value Problems in Differential-Algebraic Equations*, Elsevier, New York.

[17] Waterloo Maple, Inc., 1996, *Maple V Programming Guide*, 2nd ed., Springer, Berlin.

Dynamic Optimization of Startup and Load-Increasing Processes in Power Plants—Part II: Application

J. Bausa[1]
e-mail: bausa@olymp.fb10.tu-berlin.de.

G. Tsatsaronis
e-mail: tsa@olymp.fb10.tu-berlin.de

Institute for Energy Engineering,
Technical University of Berlin,
Marchstrasse 18,
10587 Berlin, Germany

In the first part of this study, a general method for solving dynamic optimization problems has been presented: the dynamic process model, consisting of first-order ordinary differential equations (ODEs) and algebraic equations, is discretized over the time horizon using well established methods for the solution of ODEs. The discretized system is then treated as large-scale non-linear parameter optimization problem. This transformation is implemented in a user-friendly software package. An application of this software is demonstrated in the present paper by optimizing the process of rapid load-increase in a single-pressure combined-cycle power plant. The power plant is described with a simplified model that consists of 18 first order ordinary differential equations and 67 algebraic equations. For this model a time-optimal operation associated with a load increase from 50 percent to 75 percent of base load is calculated by considering given restrictions on some temperature gradients. [DOI: 10.1115/1.1286729]

1 Introduction

In this paper the dynamic process of startup and load-increasing is treated as a dynamic optimization problem. In the first part of this study [1], the methods to solve such optimization problems have been discussed and the software implementation has been presented. The present part deals with the application of these methods to the process of rapid increase of load in a single-pressure combined-cycle power plant.

The aim of this study is to demonstrate an interesting application of dynamic optimization methods to mathematical models of power plants. To the best of the authors' knowledge such studies have not been presented in literature so far. Because of the complexity of a dynamic optimization, it is not possible to consider models as detailed as in a steady-state or dynamic simulation application. For this reason several details have been neglected (e.g., pressure drops and variable heat transfer coefficients) that only influence the quantitative results of the model. In this first study the focus was on developing a relatively simple model with a dynamic behavior qualitatively comparable to a real plant.

A brief description of the mathematical model of the power plant is presented next before the optimization problem is defined and the results are discussed.

2 Model

The method is applied to the rapid increase of load of a simplified single-pressure combined-cycle power plant, the dynamic behavior of which is described by a simplified mathematical model. Sketches of the power plant and the components of the mathematical model are shown in Fig. 1. The plant consists of six components: gas turbine, steam turbine, drum, economizer, evaporator and superheater. In the following, the most important equations are briefly described.

For the *gas turbine*, a Siemens V84.3 was chosen. The model consists of polynomial expressions for the exhaust gas temperature ($T_{gt,\mathrm{out}}$) and the exhaust gas mass flow rate ($\dot{m}_{\mathrm{gas}}$) as a function of the gas turbine power (P_{gt}). These polynomials have been fitted to the data presented by Maghon et al. [2]. The V84.3 has adjustable guide vanes that allow about 60 percent to 100 percent of power to be generated with an almost constant exhaust gas temperature. The mass flow rate and exhaust gas temperature characteristics have to be described each with *one* equation for the entire gas turbine power range of 68 MW to 171 MW (41 percent to 100 percent). This is necessary because at the present state of the optimization software it is not possible to implement conditional equations over more than one section. Therefore, the polynomials for the low power interval ($68\,\mathrm{MW} \leqslant P_{gt} \leqslant 102.6\,\mathrm{MW}$) and for the high power interval ($102.6\,\mathrm{MW} < P_{gt} \leqslant 171\,\mathrm{MW}$) are multiplied with the damping functions $\mu_{\mathrm{low}} = 0.5 - 0.5 \times \tanh(P_{gt} - 102.6)$ and $\mu_{\mathrm{high}} = 0.5 + 0.5 \times \tanh(P_{gt} - 102.6)$, respectively, and added together. The resulting graphical representations of $\dot{m}_{\mathrm{gas}}$ and $T_{gt,\mathrm{out}}$ as a function of P_{gt} are shown in Fig. 2. Compared with the other components of the combined cycle plant, the dynamic response of the gas turbine is much faster. Therefore, the delay between the control signal and the increase or decrease of power output has been neglected.

The *steam turbine* is modeled using the polytropic expansion of an ideal gas with a constant heat capacity and constant isentropic and polytropic exponent κ and polytropic efficiency η_{pol}:

$$\frac{T_{st,\mathrm{out}}}{T_{sh,\mathrm{steam},1}} = \left(\frac{p_{\mathrm{cond}}}{p_{\mathrm{drum}}}\right)^{(\kappa-1/\kappa)\eta_{\mathrm{pol}}}. \tag{1}$$

The power of the steam turbine is calculated by assuming a first-order delay:

$$k_{st,1}\frac{dP_{st}}{dt} = k_{st,2}c_{p,\mathrm{steam}}\dot{m}_{\mathrm{steam}}(T_{sh,\mathrm{steam},1} - T_{st,\mathrm{out}}) - P_{st}. \tag{2}$$

A typical value of 10 s was assigned to the time constant $k_{st,1}$ [3]. The maximum power of the steam turbine was set to 80 MW by adjusting the constant $k_{st,2}$. The total power generated by the plant is the sum of gas turbine and steam turbine power:

$$P_{\mathrm{tot}} = P_{gt} + P_{st}. \tag{3}$$

At full load the steam turbine power is about 32 percent of the total power.

[1]Author to whom all correspondence should be addressed. Present address: BASF AG, DWF/AM, L440, 67056 Ludwigshafen, Germany.

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division May 26, 1999; final revision received by the ASME Headquarters December 29, 1999. Technical Editor: D. Assanis.

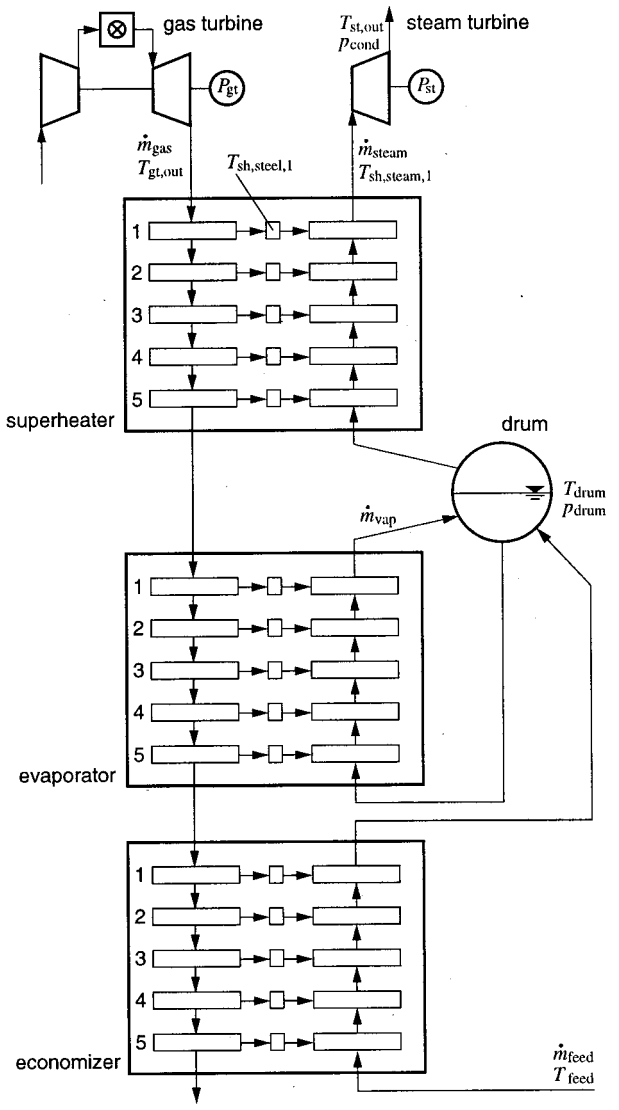


Fig. 1 Sketch of the single-pressure combined-cycle power plant

The mass flow rate of steam is calculated as a function of the drum pressure and the constant pressure in the condenser ($p_{cond}=0.06$ bar):

$$\dot{m}_{steam}=k\sqrt{p_{drum}-p_{cond}}. \tag{4}$$

The model of the drum considers the storage of steam. With the assumption of a constant density we can formulate the mass balance as a differential equation for the drum pressure:

$$\frac{dp_{drum}}{dt}=\frac{1}{k_{drum}}(\dot{m}_{vap}-\dot{m}_{steam}-\dot{m}_{approach}). \tag{5}$$

A constant value of 275 kg/bar was assigned for k_{drum} in Eq. 5 [3]. The saturation temperature T_{drum} and the enthalpy increase during vaporization Δh_{vap} are calculated as functions of the drum pressure with fourth and third order polynomials that have been developed using data from the steam tables [4]:

$$T_{drum}=-1.18287\times10^{-7}p_{drum}^4+6.84645\times10^{-5}p_{drum}^3-1.63485\times10^{-2}p_{drum}^2+2.41939p_{drum}+175.904 \tag{6}$$

$$\Delta h_{vap}=-1.15929\times10^{-7}p_{drum}^3+3.57007\times10^{-5}p_{drum}^2-9.73352\times10^{-3}p_{drum}+2.05198, \tag{7}$$

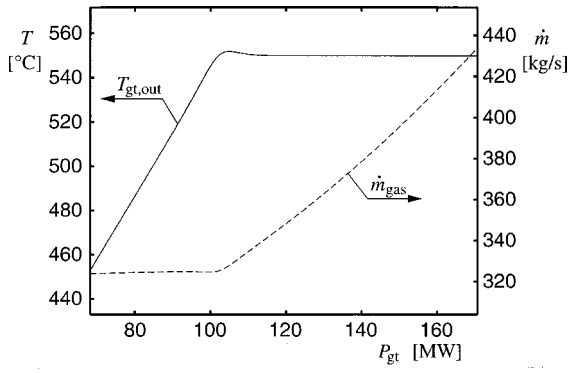


Fig. 2 Temperature and mass flow rate of the exhaust gas as a function of gas turbine power for the V84.3 gas turbine [2]

where T_{drum} and Δh_{vap} are calculated in °C and MJ/kg, respectively.

The mass flow rate of steam exiting the evaporator is calculated as a function of the heat transfer in the evaporator:

$$\dot{m}_{vap}=\dot{Q}_{vap}/\Delta h_{vap}. \tag{8}$$

The mass flow rate $\dot{m}_{approach}$ represents the steam rate that has to be condensed to heat the water exiting the economizer up to the saturation temperature:

$$\dot{Q}_{approach}=c_{p,water}\dot{m}_{steam}(T_{drum}-T_{ec,water,1}) \tag{9}$$

$$\dot{m}_{approach}=\dot{Q}_{approach}/\Delta h_{vap} \tag{10}$$

The models for the *economizer*, *evaporator*, and *superheater* are formulated similarly: all three components are discretized spatially in five sections. For each section (index *i*) a dynamic energy balance for the pipes is formulated:

$$\frac{1}{5}m_{steel}c_{steel}\frac{dT_{steel,i}}{dt}=\dot{Q}_{gas-steel,i}-\dot{Q}_{steel-steam,i}. \tag{11}$$

The heat flow rates are calculated with the assumption of uniform temperatures and heat transfer coefficients:

$$\dot{Q}_{gas-steel,i}=\frac{1}{5}U_{gas}A(T_{gas,i}-T_{steel,i}) \tag{12}$$

$$\dot{Q}_{steel-steam,i}=\frac{1}{5}U_{steam}A(T_{steel,i}-T_{steam,i}). \tag{13}$$

For simplicity reasons the heat transfer coefficients have been assumed to be constant. The temperatures of gas and steam for each section of the counterflow heat exchangers are calculated implicitly with the assumption of constant specific heats:

$$\dot{Q}_{gas-steel,i}=c_{p,gas}\dot{m}_{gas}(T_{gas,i-1}-T_{gas,i}) \tag{14}$$

$$\dot{Q}_{steel-steam,i}=c_{p,steam}\dot{m}_{steam}(T_{steam,i}-T_{steam,i+1}). \tag{15}$$

Equation (15) is only applied to the economizer and superheater. The steam temperature in the evaporator is equal to the drum temperature.

For the drum we assume a perfect control of the liquid level, i.e., the mass flow rate of the feedwater is equal to the mass flow rate of steam entering the steam turbine:

$$\dot{m}_{steam}=\dot{m}_{feed}. \tag{16}$$

The temperature of the feedwater was set to 50 °C. Pressure drops have been neglected in all heat exchangers.

To summarize, the model of the combined-cycle consists of 18 first order ordinary differential equations and 67 algebraic equa-

tions. With the a total of 68 algebraic variables there is one degree of freedom. This degree of freedom is represented by the control variable P_{gt} that has to be optimized over time.

The model developed in this section includes several simplifications and assumptions that can be classified into two groups. The first group includes simplifications such as neglected pressure drops, neglected heat loss, and constant heat transfer coefficients. These details can be easily implemented in a more accurate model. However, the maximum dimension of the DAE-system of approximately 200 [1] has to be considered.

The second group of simplifications cannot be removed at the present state of the optimization software. At present all model equations have to be implemented in a symbolic manner. Thus, it is not possible to calculate the properties of water and steam in an external software package as it is done in steady-state or dynamic simulation packages. This important feature will be implemented in the next version of the OCOMA package. The next version will make use of external routines that calculate thermodynamic properties and their exact partial derivatives.

3 Optimization Problem

The objective is to find a profile of the control variable P_{gt} that increases the total power of the plant from steady-state conditions at 50 percent (126.32 MW) of design capacity to a load of 75 percent (189.5 MW) as fast as possible. This objective is formulated with the following objective function:

$$\min_{P_{gt}(t)} J = \int_{t=0}^{t=1500s} (P_{\text{tot}} - 0.75 P_{\text{tot,max}})^2 dt. \tag{17}$$

To demonstrate the effect of constraints on the temperature gradients, maximum values for the gradients of the drum temperature and the temperature of each element of the heat exchangers (evaporator and superheater) are enforced:

$$\left| \frac{dT_{\text{drum}}}{dt} \right| \leqslant 2\,^{\circ}\text{C/min} \tag{18}$$

$$\left| \frac{dT_{\text{steel},i}}{dt} \right| \leqslant 5\,^{\circ}\text{C/min}. \tag{19}$$

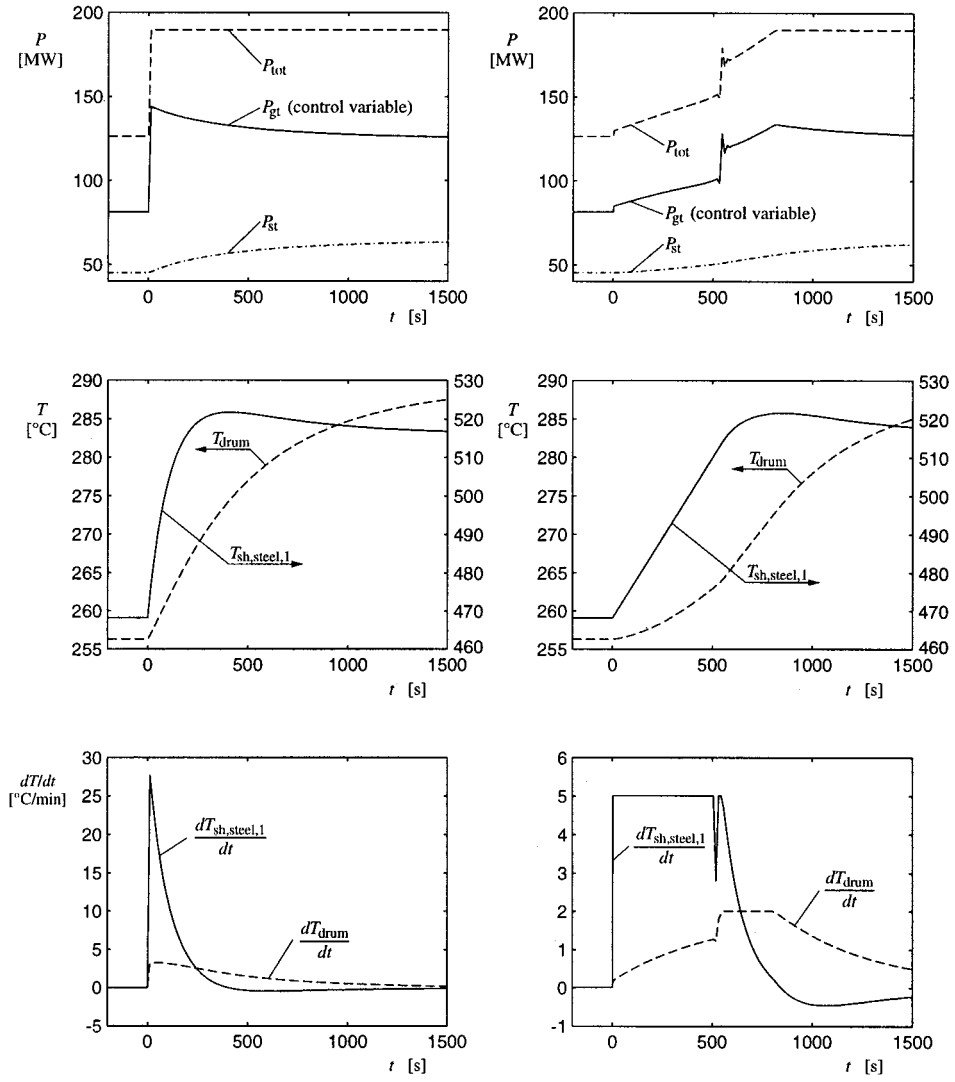


Fig. 3 Optimal profiles for the increase of power from steady-state conditions at 50 percent load to 75 percent of full load. The left side presents the results for the unconstrained case (without Eqs. (18) and (19)) whereas the right side shows the profiles for the constrained case.

The values of 2 °C/min and 5 °C/min were chosen arbitrarily. More realistic values for the maximum temperature gradients can be calculated using the method of Trautmann [5] that allows to calculate the maximum allowed temperature gradients with respect to time for a given maximum value of stress for typical plant components. As a result of this approach, the choice of maximum values for the temperature gradients is almost equivalent to restricting the maximum stress of the components.

The time interval is discretized with 120 time-steps. Thus, we get an optimization problem with 7080 optimization parameters, 6960 equality constraints and 1320 inequality constraints. The FORTRAN program that is generated consists of more than 200,000 lines of code. The generation of the FORTRAN code takes approximately 14 min on a Sun UltraSparc Workstation (300 MHz, 256 MByte Ram).

To derive appropriate guess values for the optimization variables, the time horizon has to be discretized on a coarse grid with only few time-steps. For small optimization problems with up to 500 variables the SQP software converges from almost any starting point. Guessed values for the large-scale problem are then calculated by interpolating between the time-steps of the coarse grid. With these guessed values the solution of the large-scale problem is solved in 3.4 min.

Figure 3 shows the results of the optimization. The diagrams on the left side show the optimal profile of the control variable P_{gt} for the case without the constraints given by Eqs. (18) and (19): the nominal value of 189.5 MW is reached immediately by increasing the gas turbine power from 81.3 MW to 144.0 MW in one step. This causes a rapid increase of the drum temperature and the temperature of the heat exchangers. The gradient of the drum temperature reaches a maximum value of about 4 °C/min and the gradient of the temperature of the first element of the superheater exceeds 20 °C/min.

The diagrams on the right side of Fig. 3 show the results for the constrained case: in the first time period ($0 < t \leqslant 535$ s) the increase of gas turbine load is restricted by the temperature gradient in the superheater. In this period the gas turbine exhaust gas temperature increases with power, Fig. 2.

In the second time period ($535 < t \leqslant 790$ s) the increase of gas turbine load is restricted by the maximum temperature gradient in the drum until the nominal power of 189.5 MW is reached after 810 s.

Figure 3 shows that it is possible to optimize the control profile of the gas turbine power in a way that the temperature gradients never exceed their given maximum values.

4 Conclusions

In this paper the method and software presented by Bausa and Tsatsaronis [1] have been applied to the process of rapid increase of power in a single-pressure combined-cycle. It has been demonstrated that optimal control profiles for the power of the gas turbine can easily be calculated. These control profiles are optimal in the sense that total power is increased as fast as possible without violating the constraints imposed by the maximum temperature gradients.

The method is not restricted to simplified models such as the one presented in this paper. Models with approximately 200 equations have already been solved. Future work will focus on more complex models and on the possibility to link external software packages for more accurately calculating the required thermodynamic properties (e.g., steam data).

Acknowledgment

The authors wish to thank Prof. P. E. Gill, University of California, San Diego, for providing the excellent SQP-software SNOPT.

Nomenclature

c_p = heat capacity [MJ/(kg K)]
h = mass specific enthalpy [MJ/kg]
J = objective function
k = constant
$\dot{m}$ = mass flow rate [kg/s]
P = power [MW]
p = pressure [bar]
$\dot{Q}$ = heat flow rate [MW]
T = temperature [K]
t = time [s]
U = heat transfer coefficient [MW/(m^2 K)]
η_{pol} = polytropic efficiency
κ = isentropic exponent
μ = damping function

Subscripts

cond = condensation
ec = economizer
f = final
gt = gas turbine
sh = superheater
st = steam turbine
tot = total
vap = evaporation

References

[1] Bausa, J., and Tsatsaronis, G., 2001, ''Dynamic Optimization of Start-Up and Load-Increasing Processes in Power Plants: Part I—Method,'' ASME J. Eng. Gas Turbines Power, **123**, pp. 246–250.

[2] Maghon, H., Becker, B., Schulenberg, T., Termuehlen, H., and Kraemer, H., 1993, ''The Advanced V84.3 Gas Turbine,'' in *Proceedings of the American Power Conference*, pp. 1360–1380.

[3] Klefenz, G., 1991, *Die Regelung von Dampfkraftwerken*, BI-Wiss.-Verl., Mannheim.

[4] International Formulation Commitee of the 6th International Conference on the Properties of Steam, 1967, *The 1967 IFC Formulation for Industrial Use*, Verein Deutscher Ingenieure, Düsseldorf, Germany.

[5] Trautmann, G., 1988, ''Optimierung des Anfahrvorganges in konventionellen Kraftwerken bezüglich der Wechselbeanspruchung durch Temperatur und Druck,'' Ph.D. thesis, Technische Universität Braunschweig.

Journal of Engineering for Gas Turbines and Power

Technical Brief

Analytical Prediction Techniques for Axisymmetric Flow in Gas Labyrinth Seals

U. Yucel[1] and J. Y. Kazakia
Lehigh University, Department of Mechanical Engineering and Mechanics, 19 Memorial Drive West, Bethlehem, PA 18015-3085

Labyrinth seals are commonly found in turbines and compressors. Their objective is to control gas leakage from high pressure regions to low pressure regions. The correct prediction and control of this leakage is crucial for the efficient and economic operation of turbomachinery. In this paper we present approaches for obtaining the above prediction in a simple analytical and explicit method. Both constant and pressure dependent flow coefficients are incorporated in the present study which extends to the higher inlet/outlet pressure differences. The results obtained with our methods compare favorably with the ones obtained by both numerical and experimental techniques. In many cases there is hardly a distinction between our results and the numerical prediction. [DOI: 10.1115/1.1340630]

1 Introduction

Labyrinth seals are commonly found in turbines and compressors. Their objective is to control gas leakage from high pressure regions to low pressure regions. This leakage, which depends on a great variety of parameters such as geometry of the teeth, number of cavities, pressure differences, temperature, type of gas, etc., is inevitably present even in the case of abradable seals. The correct prediction and control of this leakage is crucial for the efficient and economic operation of turbomachinery.

The air flow through the seals and its swirling motion in the cavities formed between the seal teeth may create net pressure and shear forces acting on the rotor. These forces may cause the destabilization of marginally stable rotors rotating at very high speeds.

The prediction of the leakage mass flow rate, the associated pressure distribution in the seal cavities, and the associated swirl velocities have been previously studied numerically by Childs [1], Childs and Scharrer [2], Eser and Kazakia [3], Yucel [4], and others. In addition Benvenuti et al. [5] and Brownell et al. [6] have studied these problems experimentally as well as others.

In this paper we present approaches for obtaining the above prediction in a simple analytical and explicit method. Both constant and pressure dependent flow coefficients are incorporated in the present study which extends to the higher inlet/outlet pressure differences.

The results obtained with our methods compare favorably with the ones obtained by both numerical and experimental techniques. In many cases there is hardly a distinction between our results and the numerical prediction.

2 Leakage Flow Rate Model

Consider a straight-through gas labyrinth seal with teeth being either on the rotor or on the stator. When the concentric rotor rotates undisturbed with constant speed, the gas flow in the labyrinth seal is steady-state and the continuity equation implies that the leakage mass flow rate $\dot{m}_i$ through each of the teeth must be constant as i varies from 1 to N_T (the number of teeth in a particular labyrinth seal). The mass flow rate $\dot{m}$ depends on the geometry of the seals, the inlet temperature T and the outlet to inlet pressure ratio $r = P_{\text{out}}/P_{\text{in}}$.

We assume that the gas flow can be reasonably modeled using isothermal flow with constant values for the pressure and density at each cavity of the seal. We also assume the validity of the gas law.

In many studies [1–4], a semiempirical leakage equation is used to describe the leakage flow rate. This equation is given as

$$\dot{m} = C_i \mu_0 A \left(\frac{P_{i-1}^2 - P_i^2}{RT} \right)^{1/2}, \tag{1}$$

where A represents the annular area of the seal, P_i is the pressure at cavity i, and R is the gas constant. The kinetic energy carryover coefficient μ_0 is equal to 1 for the first tooth but it is calculated from

$$\mu_0 = 1 + 0.0791(N_T - 1), \tag{2}$$

for all other teeth in the seal. Childs and Scharrer [2] used a clearance dependent coefficient μ_0, but we have found Eq. (2) to adequately describe the effect of increased flow rate at a given tooth due to the formation of a jet in the earlier tooth.

The flow coefficient C_i may depend on the pressures in the two cavities adjacent to a particular tooth i. Several investigators have used the Chaplygin formula (Gurevich [7])

$$C_i = \frac{\pi}{\pi + 2 - 5S_i + 2S_i^2}, \quad S_i = \left(\frac{P_{i-1}}{P_i} \right)^{(\gamma-1)/\gamma} - 1, \tag{3}$$

to describe this dependence. Here γ is the ratio of specific heat ($=1.4$ for air).

When Eq. (3) is used in conjunction with Eq. (1), a numerical iteration procedure must be used in order to predict the leakage flow and the associated pressure distribution. In the present paper we show that the above analysis is greatly simplified when one considers either a constant flow coefficient or one which depends

[1]Electronic mail: ugy2@lehigh.edu; jyk0@lehigh.edu

Contributed by the Structures and Dynamics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the S&D Division, April 18, 2000; final revision received by the ASME Headquarters May 12, 2000. Editor: H. D. Nelson.

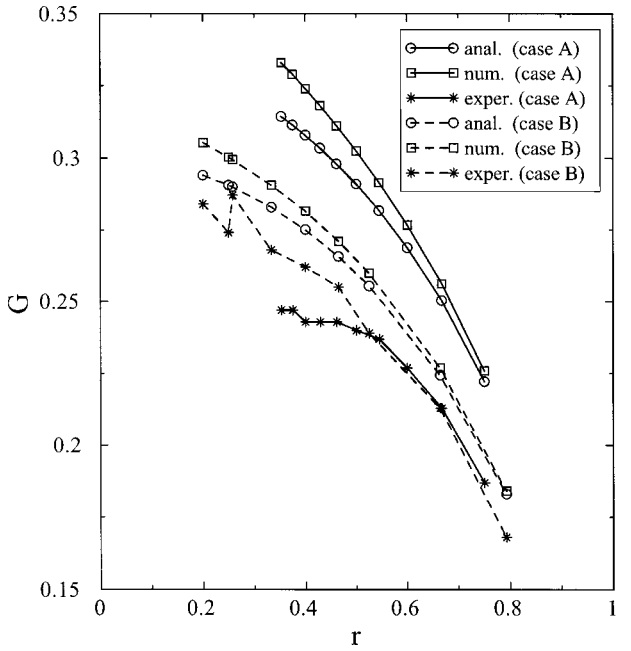


Fig. 1 The dependence of the mass flow number *G* on the pressure ratio $r=P_{out}/P_{in}$. Analytical (present), numerical and experimental results are compared for two different seals. Case A refers to a five-tooth seal used by Brownell et al. [6], case B refers to a twelve-tooth seal used by Benvenuti et al. [5].

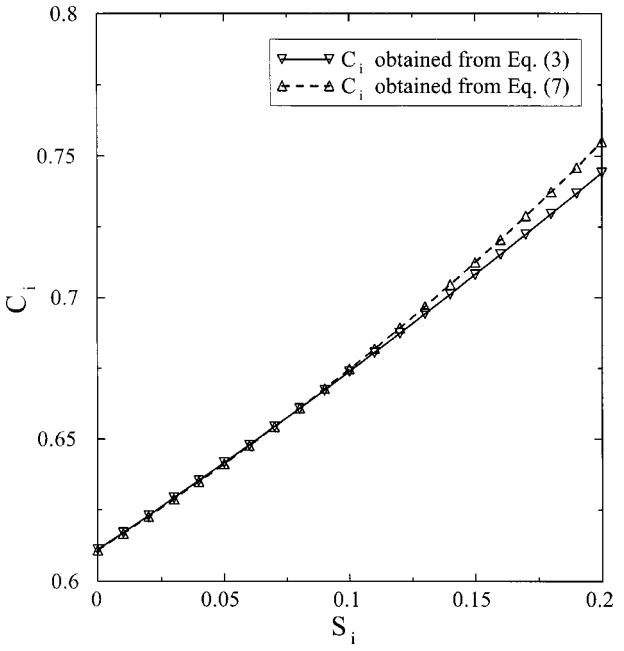


Fig. 2 The flow coefficient variation with pressure. Formula used in numerical results is compared with formula used for semi-analytical approach.

on the pressure ratio in a special way described below. This simplification occurs without any significant reduction in the accuracy of the prediction.

2.1 Constant Flow Coefficient. Using the constant value $C_0=0.611$ in Eq. (1) and introducing k by

$$k=\frac{\dot{m}(RT)^{1/2}}{AC_0}, \tag{4}$$

it can be shown that

$$k=\left(\frac{P_{in}^2-P_{out}^2}{1+(N_T-1)/\mu_0^2}\right)^{1/2}, \tag{5}$$

and

$$G=\frac{\dot{m}(RT)^{1/2}}{AP_{in}}=\frac{C_0k}{P_{in}}=C_0\left(\frac{1-r^2}{1+(N_T-1)/\mu_0^2}\right)^{1/2}. \tag{6}$$

In Fig. 1 this analytical prediction is compared to the numerical iterative solution and to experimental results for two different seals. The first seal is a five-tooth (on stator) straight-through labyrinth seal used in the experiments of Brownell et al. [6] and the second seal is a twelve-tooth (on stator) straight-through labyrinth seal used in the experiments by Benvenuti et al. [5]. As it can be seen the analytical approach used in this paper agrees very well with the more general numerical iteration solution, especially when the pressure ratio is between 0.5 and 1. Even though the analytic results overestimate the measured mass flow, they do so in a way no worse then the numerical iteration technique.

Similar comparison is made for fourteen and eighteen-tooth straight-through labyrinth seals used in the experiments of Benvenuti et al. [5] with a good agreement between the analytical approach and numerical iteration solution as well as the measurements. In addition the mass flow rate predicted by our analytic approach agrees well with the results of Kwanka and Nagel [8] and Rosen [9].

Having obtained the leakage flow rate we can now use Eq. (1) to obtain analytically the associated pressure distribution P_i. This analytical prediction compares well with the experimental results obtained by Benvenuti et al. [5] for a twelve-tooth straight-through labyrinth seal with pressure ratios $r=0.205$ and $r=0.787$, and the iterative numerical results.

2.2 Pressure Dependent Flow Coefficient. The results obtained in the earlier section are based on a constant flow coefficient. Figure 2 depicts the variation of the flow coefficient with the quantity S_i defined by Eq. (3). The range of S_i in Fig. 2 is rather wide, however, it would be desirable to use an expression for the flow coefficient which could represent this type of variation. The broken line of Fig. 2 is obtained using $\alpha=2.68$ and $\beta=-1.45$ in the following expression for the flow coefficient

$$C_i=\frac{1}{\{\alpha+\beta\gamma/(\gamma-1)\ln(1+S_i)\}^{1/2}}. \tag{7}$$

For this special dependence of C_i, one can solve for the mass flow number G and the associated pressure distribution using a semi-analytical technique. It can be shown that

$$G=\left\{\frac{1-(P_1/P_{in})^2}{\alpha-\beta\ln(P_1/P_{in})}\right\}^{1/2}. \tag{8}$$

The pressure ratio $\hat{P}=P_1/P_{in}$, needed in Eq. (8), is to be obtained by solving the equation

$$\mu_0^2\frac{\hat{P}^2-(P_{out}/P_{in})^2}{1-\hat{P}^2}\{\alpha-\beta\ln(\hat{P})\} -\left\{\alpha(N_T-1)+\beta\ln\left(\frac{\hat{P}}{P_{out}/P_{in}}\right)\right\}=0. \tag{9}$$

After solving Eq. (9) using a numerical method such as the Secant method of root finding, we can obtain G from Eq. (8) and the remaining pressures from

$$\mu_0^2(P_i/P_{in})^2+G^2\{\alpha-\beta\ln(P_i)\} =\mu_0^2(P_{i-1}/P_{in})^2+G^2\beta\ln(P_{i-1}). \tag{10}$$

The results have been compared with results obtained using either a constant flow coefficient or one of the form described in Eq. (3). The comparison is very satisfactory.

3 The Associated Swirl Velocities

In addition to the leakage flow rate and to the associated pressure distribution it is important to predict the circumferential gas velocities. These exist due to the inlet swirl of the gas entering the seal and also due to the high speed rotation of the shaft. These velocities are needed in the prediction of the pressure fluctuations which may generate sufficiently high forces on the rotor causing its destabilization.

As was the case during our discussion of the pressure distribution, we make here as well the same assumption that the circumferential velocities are different from cavity to cavity, but sufficiently similar in a single cavity to permit a bulk representation of them with constant V_i where i denotes a particular cavity. Of course the effects of the seal walls are incorporated by the introduction of boundary layers.

The circumferential momentum equation gives

$$\dot{m}(V_i - V_{i-1}) = 2\pi R_s L(\tau_{r_i} a_r - \tau_{s_i} a_s), \tag{11}$$

where a_r and a_s are the rotor and stator shear area numbers and τ_{r_i} and τ_{s_i} are the corresponding shear stresses.

Due to the complex geometry of seal walls the hydraulic diameter approximation is used in modeling the cavity flow by the Blasius correlation for turbulent flow in smooth pipes. Hence we write

$$\tau_i = -0.03955\rho_i U_i^2 \,\mathrm{Re}_i^{-0.25}\,\mathrm{sgn}(U_i), \tag{12}$$

where U_i is the gas velocity relative to the stator (i.e., V_i) when we calculate τ_{s_i} and relative to the rotor (i.e., $V_i - R_s w$) when we calculate τ_{r_i}. We also note that the sign in front of the right hand side of Eq. (12) must be changed to plus for the calculation of τ_{s_i}. In Eq. (12) the Reynolds numbers Re_i are calculated using

$$\mathrm{Re}_i = \frac{|U_i| D_h \rho_i}{\mu}, \tag{13}$$

where D_h is the hydraulic diameter of the cavity, ρ_i is the density of the gas at cavity i, and μ is the absolute viscosity.

For a given inlet swirl velocity V_0 we can use Eqs. (11)–(13) to predict the circumferential velocities starting with V_1 and progressing towards the last cavity. At each step we must solve an implicit nonlinear equation for V_i due to the character of Eq. (12). This analysis can be simplified by assuming that during the calculation of the shear stress at cavity i we can use the Reynolds numbers calculated in the previous cavity. We then obtain a quadratic equation at each cavity which can be solved analytically thus providing an explicit method for the prediction of the circumferential velocities.

A comparison of the explicit and implicit solutions for the velocity distribution across the labyrinth teeth for an eighteen-tooth seal and a pressure ratio of $r = 0.32$ is given by Yucel [4]. The biggest deviation between the implicit and explicit results occurs in the fifth cavity of the labyrinth seal and does not exceed 0.5 percent.

4 Conclusions

We conclude that for special forms of the discharge flow coefficient the leakage mass flow rate and the associated pressure distribution can be determined by analytic formulas. We also introduce a simplification which makes the explicit calculation of the circumferential velocities in the cavities of labyrinth seal possible.

Nomenclature

A = annular flow area (m^2)
C = flow coefficient
D_h = hydraulic diameter of cavity (m)
G = mass flow number
L = axial length of seal cavity (m)
N_T = number of teeth
P = pressure (N/m^2)
$\hat{P}$ = P_1 / P_{in}
R = gas constant [J/(kg K)]
Re = Reynolds number
R_s = shaft radius (m)
T = inlet temperature (K)
V = circumferential velocity (m/s)
a = shear area number
$\dot{m}$ = leakage mass flow rate (kg/s)
r = outlet to inlet pressure ratio
sgn(U) = sign of U
γ = ratio of specific heat
μ = absolute viscosity [kg/(m s)]
μ_0 = kinetic energy carryover coefficient
ρ = density of fluid (kg/m^3)
τ = shearing stress (N/m^2)
ω = angular velocity of shaft (1/s)

Subscripts

i = ith cavity value
r = rotor
s = stator
in = inlet
out = outlet

References

[1] Childs, D., 1993, *Turbomachinery Rotordynamics Phenomena, Modeling, and Analysis*, John Wiley and Sons, New York, pp. 290–354.
[2] Childs, D., and Scharrer, J. K., 1986, ''An Iwatsubo-Based Solution for Labyrinth Seals: Comparison to Experimental Results,'' J. Eng. Gas Turbines Power, **108**, pp. 325–331.
[3] Eser, D., and Kazakia, J. Y., 1995, ''Air Flow in Cavities of Labyrinth Seals,'' Int. J. Eng. Sci., **33**, No. 15, pp. 2309–2326.
[4] Yucel, U., 1996, ''Leakage and Swirl Velocities in Labyrinth Seals,'' Masters thesis, Lehigh University, Bethlehem, PA.
[5] Benvenuti, E., Ruggeri, G., and Tomasini, E. P., 1980, ''Analytical and Experimental Development of Labyrinth Seals for Process Centrifugal Compressors,'' *Performance Prediction of Centrifugal Pumps and Compressors*, ASME, New York, pp. 21–34.
[6] Brownell, J. B., Millward, J. A., and Parker, R. J., 1989, ''Nonintrusive Investigations Into Life-Size Labyrinth Seal Flow Fields,'' J. Eng. Gas Turbines Power, **111**, pp. 335–342.
[7] Gurevich, M. I., 1966, *The Theory of Jets in an Ideal Fluid*, Pergamon, New York, pp. 319–323.
[8] Kwanka, K., and Nagel, M., (1996), ''Experimental Rotordynamic Coefficients of Short Labyrinth Gas Seals,'' *NASA Conference Publication 3344: Rotordynamic Instability Problems in High-Performance Turbomachinery*, Texas A&M University, College Station, TX, pp. 135–144.
[9] Rosen, M. C., 1986, ''Calculation of Leakage and Stiffness/Damping Coefficients for Compressible Flow in Straight-Through and Interlocking Labyrinth Seals,'' Report No. UVA/643092/MAE86/346, University of Virginia, Charlottesville, VA, p. 16.