



Journal of Engineering for Gas Turbines and Power

Published Monthly by ASME

VOLUME 132 • NUMBER 3 • MARCH 2010

RESEARCH PAPERS

Gas Turbines: Coal, Biomass, and Alternative Fuels

- 031401 Producing Hydrogen and Power Using Chemical Looping Combustion and Water-Gas Shift
Niall R. McGlashan, Peter R. N. Childs, Andrew L. Heyes, and Andrew J. Marquis

Gas Turbines: Controls, Diagnostics, and Instrumentation

- 031601 Optimal Tuner Selection for Kalman Filter-Based Aircraft Engine Performance Estimation
Donald L. Simon and Sanjay Garg
- 031602 Application of Bayesian Forecasting to Change Detection and Prognosis of Gas Turbine Performance
Holger Lipowsky, Stephan Staudacher, Michael Bauer, and Klaus-Juergen Schmidt
- 031603 Mapping the Density Fluctuations in a Pulsed Air-Methane Flame Using Laser-Vibrometry
Fabrice Giuliani, Thomas Leitgeb, Andreas Lang, and Jakob Woisetschlager

Gas Turbines: Cycle Innovations

- 031701 A Novel Coal-Based Hydrogen Production System With Low CO₂ Emissions
Gang Xu, HongGuang Jin, YongPing Yang, Liqiang Duan, Wei Han, and Lin Gao

Gas Turbines: Manufacturing, Materials, and Metallurgy

- 032101 Advanced Braze Alloys for Fast Epitaxial High-Temperature Brazing of Single-Crystalline Nickel-Base Superalloys
Britta Laux, Sebastian Piegert, and Joachim Rösler

Gas Turbines: Microturbines and Small Turbomachinery

- 032301 Aerodynamic Design and Numerical Investigation on Overall Performance of a Microradial Turbine With Millimeter-Scale
Lei Fu, Yan Shi, Qinghua Deng, and Zhenping Feng

Gas Turbines: Oil and Gas Applications

- 032401 Influence of Blade Deterioration on Compressor and Turbine Performance
M. Morini, M. Pinelli, P. R. Spina, and M. Venturini

Gas Turbines: Structures and Dynamics

- 032501 Synchronous Response to Rotor Imbalance Using a Damped Gas Bearing
Bugra H. Ertas, Massimo Camatti, and Gabriele Mariotti
- 032502 Turbocharger Nonlinear Response With Engine-Induced Excitations: Predictions and Test Data
Luis San Andrés, Ash Maruyama, Kostandin Gjika, and Sherry Xia
- 032503 Measurement of Structural Stiffness and Damping Coefficients in a Metal Mesh Foil Bearing
Luis San Andrés, Thomas Abraham Chirathadam, and Tae-Ho Kim

(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. ©TM
♻️ 85% recycled content, including 10% post-consumer fibers.

Editor
D. R. BALLAL (2011)
Assistant to the Editor
S. D. BALLAL

Associate Editors
Gas Turbine (Review Chairs)
T. SATTELMAYER (2010)
K. THOLE (2010)

Coal, Biomass & Alternative Fuels
K. ANNAMALAI (2010)

Combustion & Fuels
N. K. RIZK (2012)
T. SATTELMAYER (2012)

Controls, Diagnostics, & Instrumentation
A. VOLPONI (2010)

Cycle Innovation
P. PILIDIS (2010)

Electric Power
P. CHIESA (2011)

Structures and Dynamics
P. S. KEOGH (2010)
J. SZWEDOWICZ (2012)

Advanced Energy Systems
J. KAPAT (2010)

Internal Combustion Engines
C. RUTLAND (2012)
J. WALLACE (2011)
M. WOOLDRIDGE (2011)

Nuclear Engineering
J. KUNZE (2011)
I. PIORO (2011)

PUBLICATIONS COMMITTEE
Chair, B. RAVANI

OFFICERS OF THE ASME
President, A. E. HOLT

Executive Director,
T. G. LOUGHLIN

Treasurer,
W. MARNER

PUBLISHING STAFF

Managing Director, Publishing
P. DI VIETRO

Manager, Journals
C. MCATEER

Production Coordinator
J. SIERANT

Transactions of the ASME, Journal of Engineering for Gas Turbines and Power (ISSN 0742-4795) is published monthly 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).

COPYRIGHT © 2010 by the American Society of Mechanical Engineers. For authorization to photocopy material for internal or personal use under circumstances not falling within the fair use provisions of the Copyright Act, contact the Copyright Clearance Center (CCC), 222 Rosewood Drive, Danvers, MA 01923. Tel: 978-750-8400, www.copyright.com. Canadian Goods & Services Tax Registration #126148048

- 032504 Computational Studies of the Unbalance Response of a Whole Aero-Engine Model With Squeeze-Film Bearings
Philip Bonello and Pham Minh Hai

Internal Combustion Engines

- 032801 A Comparison of the Different Methods of Using Jatropha Oil as Fuel in a Compression Ignition Engine
M. Senthil Kumar, A. Ramesh, and B. Nagalingam
- 032802 An Analytical Approach for the Evaluation of the Optimal Combustion Phase in Spark Ignition Engines
A. Beccari, S. Beccari, and E. Pipitone
- 032803 Simulating the Concentration Equations and the Gas-Wall Interface for One-Dimensional Based Diesel Particulate Filter Models
Christopher Depcik
- 032804 An Experimental and Numerical Investigation of Spark Ignition Engine Operation on H₂, CO, CH₄, and Their Mixtures
Hailin Li, Ghazi A. Karim, and A. Sohrabi
- 032805 Kinetic Study of the Thermo-Oxidative Degradation of Squalane (C₃₀H₆₂) Modeling the Base Oil of Engine Lubricants
Moussa Diaby, Michel Sablier, Anthony Le Negrate, and Mehdi El Fassi

Nuclear Power

- 032901 A Comparison Between Recent Advances in Cylindrical Nodal Diffusion Methods
David V. Colameco, Kostadin N. Ivanov, Rian H. Prinsloo, Djordje I. Tomasevic, and Suzanne Theron
- 032902 On the Performance of Very High Temperature Reactor Plants With Direct and Indirect Closed Brayton Cycles
Mohamed S. El-Genk and Jean-Michel Tournier
- 032903 Design Option of Heat Exchanger for the Next Generation Nuclear Plant
Chang H. Oh, Eung S. Kim, and Mike Patterson

The ASME Journal of Engineering for Gas Turbines and Power is abstracted and indexed in the following:

AESIS (Australia's Geoscience, Minerals, & Petroleum Database), Applied Science & Technology Index, Aquatic Sciences and Fisheries Abstracts, Civil Engineering Abstracts, Compendex (The electronic equivalent of Engineering Index), Computer & Information Systems Abstracts, Corrosion Abstracts, Current Contents, Engineered Materials Abstracts, Engineering Index, Enviroline (The electronic equivalent of Environment Abstracts), Environment Abstracts, Environmental Science and Pollution Management, Fluidex, INSPEC, Mechanical & Transportation Engineering Abstracts, Mechanical Engineering Abstracts, METADEX (The electronic equivalent of Metals Abstracts and Alloys Index), Pollution Abstracts, Referativnyi Zhurnal, Science Citation Index, SciSearch (The electronic equivalent of Science Citation Index), Shock and Vibration Digest

Producing Hydrogen and Power Using Chemical Looping Combustion and Water-Gas Shift

Niall R. McGlashan
e-mail: n.mcglashan@ic.ac.uk

Peter R. N. Childs

Andrew L. Heyes

Andrew J. Marquis

Department of Mechanical Engineering,
Imperial College,
South Kensington,
London SW7 2BX, UK

A cycle capable of generating both hydrogen and power with “inherent” carbon capture is proposed and evaluated. The cycle uses chemical looping combustion to perform the primary energy release from a hydrocarbon, producing an exhaust of CO. This CO is mixed with steam and converted to H₂ and CO₂ using the water-gas shift reaction (WGSR). Chemical looping uses two reactions with a recirculating oxygen carrier to oxidize hydrocarbons. The resulting oxidation and reduction stages are performed in separate reactors—the oxidizer and reducer, respectively, and this partitioning facilitates CO₂ capture. In addition, by careful selection of the oxygen carrier, the equilibrium temperature of both redox reactions can be reduced to values below the current industry standard metallurgical limit for gas turbines. This means that the irreversibility associated with the combustion process can be reduced significantly, leading to a system of enhanced overall efficiency. The choice of oxygen carrier also affects the ratio of CO versus CO₂ in the reducer’s flue gas, with some metal oxide reduction reactions generating almost pure CO. This last feature is desirable if the maximum H₂ production is to be achieved using the WGSR reaction. Process flow diagrams of one possible embodiment using a zinc based oxygen carrier are presented. To generate power, the chemical looping system is operated as part of a gas turbine cycle, combined with a bottoming steam cycle to maximize efficiency. The WGSR supplies heat to the bottoming steam cycle, and also helps to raise the steam necessary to complete the reaction. A mass and energy balance of the chemical looping system, the WGSR reactor, steam bottoming cycle, and balance of plant is presented and discussed. The results of this analysis show that the overall efficiency of the complete cycle is dependent on the operating pressure in the oxidizer, and under optimum conditions exceeds 75%. [DOI: 10.1115/1.3159371]

Keywords: chemical looping combustion, hydrogen economy, carbon capture, water-gas shift, zinc

1 Introduction

In order to implement the “hydrogen economy,” an effective and economic means of producing H₂ is required. This paper examines a means of generating H₂ while also producing power in a central station, with the added benefit of “inherent carbon capture.” The system is assumed to burn coke, which is ashless and with zero sulfur content—the fuel is taken as graphite. Overall, the system follows the reaction¹



$$\Delta G^\circ = -314.5 \text{ MJ/kmol}$$

$$\Delta H^\circ = -215.4 \text{ MJ/kmol}$$

$$\Delta H^\circ - \Delta G^\circ = T_o \Delta S^\circ = +99.1 \text{ MJ/kmol}$$

If reaction (1) is performed in a reversible, steady flow power system, with species entering or leaving the system separately,

¹For all reactions in this paper, three thermodynamic functions are given: ΔG° , ΔH° , and $T_o \Delta S^\circ$. These are indicative of, respectively, the maximum achievable work output from the reaction, the reaction’s heat release, and the required heat exchange with the environment for a reversible reaction.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 22, 2009; final manuscript received March 25, 2009; published online December 3, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Gas Turbine Technical Congress and Exposition, Orlando, FL, June 8–12, 2009.

and at T_o and P_{o_2} the maximum “overall efficiency” of the system is given by [1,2]

$$\begin{aligned} \eta_{\text{ov,rev}} &= \frac{W_{\text{rev}}}{-\Delta H^\circ} = \frac{-\Delta G^\circ}{-\Delta H^\circ} \quad (2) \\ \Rightarrow \eta_{\text{ov,rev}} &= \left(\frac{-\Delta G^\circ}{-\Delta H^\circ} \right)_{(1)} = \frac{314.5}{215.4} = 146.0\% \end{aligned}$$

Alternatively, the heat exchange with the surroundings is given by

$$Q_{\text{rej,rev}} = Q_o = \Delta H^\circ - \Delta G^\circ = T_o \Delta S^\circ \quad (\text{MJ/kmol}) \quad (3)$$

$$\Rightarrow (Q_{\text{rej,rev}})_{(1)} = +99.1 \text{ MJ/kmol}$$

Hence, a reversible machine executing reaction (1) is endothermic; i.e., an additional quantity of heat, equal to , must enter the system from the surroundings. This seemingly anomalous situation can be explained if a further power system consuming the H₂ produced in Eq. (1) is considered as part of the analysis, thus



$$\Delta G^\circ = -474.3 \text{ MJ/kmol}$$

$$\Delta H^\circ = -571.6 \text{ MJ/kmol}$$

$$\Delta H^\circ - \Delta G^\circ = T_o \Delta S^\circ = -97.4 \text{ MJ/kmol}$$

²In this work, the sink condition is assumed to be standard temperature and pressure (stp), i.e., $T_o=298.15 \text{ K}$, $P_o=1.0 \text{ bar}$.

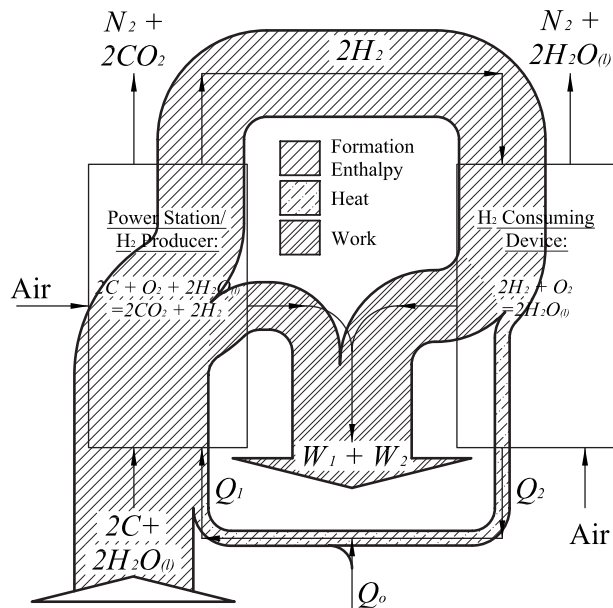


Fig. 1 Sankey diagram of a reversible, carbothermic, hydrogen economy: showing fluxes of energy flowing between coupled systems—one producing H₂ and the other consuming H₂

If reaction (4) is performed reversibly, making the same assumptions as before, a significant heat rejection is necessary. This heat rejection is almost, but not quite, equal in magnitude to the heat input required by a reversible plant executing reaction (1). Hence, added together, reactions (1) and (4) reduce to the standard carbon oxidation reaction where the theoretical maximum work output and the heating value are practically the same value, thus



$$(\Delta G^\circ)_{(5)} = (\Delta G^\circ)_{(1)} + (\Delta G^\circ)_{(4)} = -788.8 \text{ MJ/kmol}$$

$$(\Delta H^\circ)_{(5)} = (\Delta H^\circ)_{(1)} + (\Delta H^\circ)_{(4)} = -787.0 \text{ MJ/kmol}$$

$$(T_o \Delta S^\circ)_{(5)} = (T_o \Delta S^\circ)_{(1)} + (T_o \Delta S^\circ)_{(4)} = +1.8 \text{ MJ/kmol}$$

Therefore, the large $T_o \Delta S^\circ$ term of reactions (1) and (4) leads to a key thermodynamic feature of the carbothermic hydrogen economy. To illustrate this, Fig. 1 shows a Sankey diagram of two coupled reversible systems performing reactions (1) and (4), respectively. The figure shows that there are two fluxes of energy that travel between the systems: a flux of chemical energy (formation enthalpy) carried with the H₂ and a flux of heat transferred via the surroundings at T_o (the small additional heat addition, drawn from the surroundings, Q_o , is necessary to balance the $T_o \Delta S^\circ$ term of reaction (5)).

In practice, irreversibilities inherent to real systems make the prospect of an endothermic, H₂-producing power station unlikely. Considering the heat rejection (or absorption) from an irreversible system exhibiting I_{CR} lost work gives [2,3]

$$Q_{\text{rej,irrev}} = -(T_o \Delta S^\circ + I_{CR}) \quad (\text{MJ/kmol}) \quad (6)$$

where

$$I_{CR} = T_o \Delta S_{\text{gen}} \quad (\text{MJ/kmol}) \quad (7)$$

Hence, if an endothermic system is desired, it follows that

$$\Delta S_{\text{gen}} < \Delta S_{\text{reac}(1)}^\circ = +0.332 \text{ MJ/kmol} \quad (8)$$

Furthermore, if ΔS_{gen} is equal to the value specified by Eq. (8), an adiabatic system results. The work output of the power station would then equal the heating value of reaction (1), thus

$$W_{\text{irrev,adia}} = (-\Delta G^\circ)_{(1)} - I_{CR} = (-\Delta H^\circ)_{(1)} \quad (\text{MJ/kmol}) \quad (9)$$

Using Horlock's definition of the second law, or "rational" efficiency [4]:

$$\eta_{\text{rat}} = \frac{W_{\text{irrev}}}{-\Delta G^\circ} \quad (10)$$

$$\Rightarrow (\eta_{\text{rat,adia}})_{(1)} = \left(\frac{-\Delta H^\circ}{-\Delta G^\circ} \right)_{(1)} = \frac{1}{(\eta_{o,\text{rev}})_{(1)}} = 68.5\% \quad (11)$$

Therefore, an adiabatic system performing reaction (1) would not be ideal in a thermodynamic sense (although a rational efficiency of 68.5% is still higher than that achieved in any operating power plant). Instead, avoiding net heat rejection is desirable practically, as a significant proportion of the equipment found in modern power stations is engaged in rejecting heat. This includes cooling towers (or alternatively, cooling water intakes and outfalls), low-pressure turbines, steam condensers, air pumps, cooling water circulating pumps, etc. Much of this equipment is large and expensive, and in the case of cooling towers, has significant maintenance cost. In addition, in an age when power stations may have to be built on existing sites due to planning constraints, the saving of space afforded by avoiding (or at least reducing) the need for these large items is a clear advantage.

This paper explores one potential configuration of a H₂ producing power station that attempts to minimize internal irreversibilities—the goal being to approach a practical adiabatic system. The proposed configuration uses chemical looping combustion (CLC) to perform the initial oxidation of carbon to CO. The water-gas shift reaction (WGSR) is used to oxidize this CO to CO₂ and produce by-product H₂. A bottoming steam cycle is also fitted to maximize the system's overall efficiency. In Secs. 2–4, a brief review of CLC is given before the paper continues with a description of the proposed cycle, followed by a mass and energy balance.

2 Chemical Looping Combustion

Chemical looping combustion was proposed in the early 1980s as a means of augmenting power station efficiency [5]. Instead of a single combustion reaction, CLC performs hydrocarbon oxidation in two redox reactions in separate vessels. In the first reaction, a surrogate species, called the "oxygen carrier" and usually a metal, is oxidized with air at high temperature. The metal oxide produced in the oxidation is typically in a condensed state and can be separated from the residual nitrogen using physical separation methods. The oxide is then passed to a reduction vessel where the second of the redox reactions is performed—the reduction of the metal oxide back to pure metal, using a suitable hydrocarbon fuel. The regenerated metal is then separated from the product gases of the reduction process (CO_x and H₂O) and returned to the oxidizer, thereby completing the chemical loop.

CLC has two principle advantages: First, it lends itself to carbon capture as the fuel enters a reactor devoid of air, and consequently the CO_x produced in the system remains undiluted with nitrogen. This avoids the necessity for an expensive and energy intensive postcombustion scrubbing system. Hence, CLC systems are said to exhibit "inherent carbon capture" and this has resulted in increased interest in the technology in recent years [6–10]. Second, CLC enables the overall efficiency of the system to be increased for a given maximum working temperature. This is achieved by careful selection of the oxygen carrier, which enables the equilibrium temperature of the two redox reactions to be reduced. Hence, the maximum possible temperature of the working fluid can be lowered to a value below current metallurgical limits [3,5,11,12].

In a CLC system, the oxidation reaction is arranged to be exothermic, and the reduction reaction is arranged to be endothermic. This is the case for many hydrocarbon/metal combinations. Therefore, a heat engine can be straddled between the two reaction

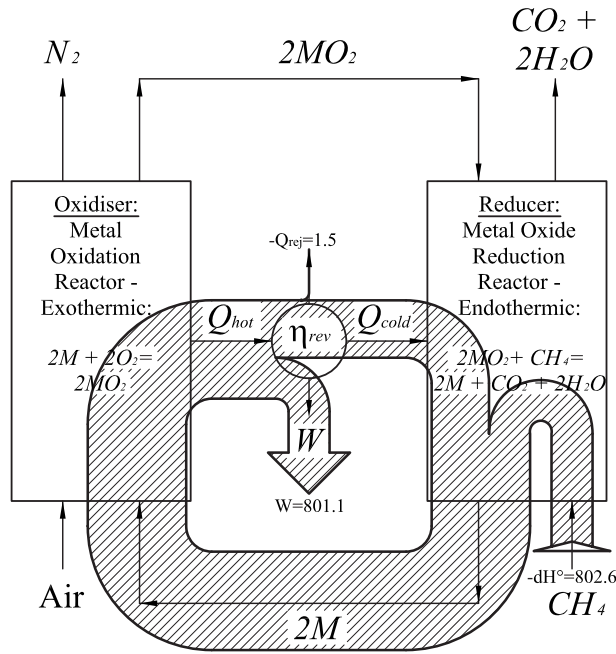
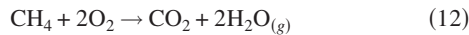


Fig. 2 Sankey diagram of energy fluxes in a reversible CLC system (units MJ/kmol)

vessels—receiving heat from the oxidizer and rejecting heat into the reducer. Figure 2 shows the CLC process diagrammatically, and includes a Sankey diagram of the energy fluxes for a reversible CLC system burning methane. The overall reaction for this system and its thermochemical values are, respectively,



$$\Delta G^o = -801.1 \text{ MJ/kmol}$$

$$\Delta H^o = -802.6 \text{ MJ/kmol}$$

$$\Delta H^o - \Delta G^o = T_o \Delta S^o = -1.5 \text{ MJ/kmol}$$

Visual examination of Fig. 2 indicates that the heat engine has a relatively low thermal efficiency. However, the work output of the system as a whole approaches the heating value of methane oxidation. This is achieved because, although the temperature of the heat fluxes exiting and entering the two reactors is low (approximately 1840 K and 1040 K in this case using a Zn based oxygen carrier leading to a relatively low reversible cycle efficiency of 43.5%), the heat flux entering the engine is much larger than the heating value of methane. Hence, McGlashan [3] showed that for hydrocarbon fuels (and any other fuel where $\Delta S^o \approx 0$ for its oxidation reaction) the overall efficiency of the system is given by

$$\eta_{\text{ov,rev,CLC}} = \Psi \eta_{\text{th,engine}} \quad (13)$$

where

$$\Psi = \frac{(\Delta H^o)_{\text{carrier}}^{\text{oxi}}}{(\Delta H^o)_{\text{fuel}}^{\text{oxi}}} \quad (14)$$

Combining Eqs. (2), (13), and (14), an expression can be found for the required thermal efficiency of the heat engine for a reversible CLC system where $\Delta S^o \approx 0$, thus³

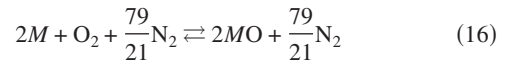
³Where $\Delta S^o \neq 0$, the Q_o term in Eq. (3) becomes significant and Eqs. (13) and (15) no longer apply. Reactions of this type generally involve a change in the number of moles of gas, and consequently pressure changes affect the equilibrium temperature. However, this effect is usually small, so the practical benefit of CLC is maintained.

$$\eta_{\text{th,engine}} = \frac{\eta_{\text{ov,rev,CLC}}}{\Psi} = \frac{(\Delta G^o)_{\text{fuel}}^{\text{oxi}}}{(\Delta H^o)_{\text{carrier}}^{\text{oxi}}} \approx \frac{(\Delta H^o)_{\text{fuel}}^{\text{oxi}}}{(\Delta H^o)_{\text{carrier}}^{\text{oxi}}} = \frac{1}{\Psi} \quad (15)$$

The value of Ψ for most combinations of oxygen carrier and hydrocarbon is greater than unity, which is why the temperature ratio over which a CLC system operates can be reduced without sacrificing efficiency.⁴ Allied to the inherent carbon capture of CLC, its potential for producing a low temperature, high efficiency system is clearly attractive.

3 Combined Chemical Looping and Water-Gas Shift System

3.1 Basic Reactions. To perform reaction (1), two subsystems can be combined: a CLC system and a WGS reactor. An additional bottoming steam cycle is specified later. Starting with the CLC system, this performs two generic reactions, using a carrier species M as follows:



Added together, and ignoring the nitrogen, these two reactions sum to the partial oxidation of carbon—the species M acting as a catalyst.



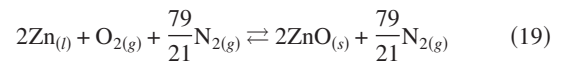
$$\Delta G^o = -274.3 \text{ MJ/kmol}$$

$$\Delta H^o = -221.1 \text{ MJ/kmol}$$

$$\Delta H^o - \Delta G^o = T_o \Delta S^o = +53.2 \text{ MJ/kmol}$$

Before proceeding, the oxygen carrier, M , must be selected. McGlashan [3] examined the limiting equilibrium temperature of the redox reactions for different oxygen carrying species and also some practical aspects of their use. It was concluded that zinc was the most suitable oxygen carrier on a combination of thermodynamic and practical grounds. An additional advantage of using Zn is that the chemical equilibrium of the carbothermic reduction in ZnO generates almost pure CO at the conditions prevailing in the reduction reactor [13,14]—see Appendix B. This enables the maximum amount of H_2 production using the WGS reactor. Conversely, if carbon is to leave the process fully oxidized to CO_2 , even if H_2 production is not desirable, the WGS (or some other use of CO) is a necessary part of any CLC system using a Zn based oxygen carrier.

Therefore, using Zn as the oxygen carrier, the CLC subsystem performs two reactions:

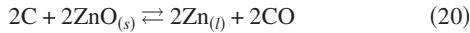


$$\Delta G^o = -649.4 \text{ MJ/kmol}$$

$$\Delta H^o = -715.6 \text{ MJ/kmol}$$

$$\Delta H^o - \Delta G^o = T_o \Delta S^o = -66.2 \text{ MJ/kmol}$$

⁴Equation (15) can also be derived from consideration of the equilibrium temperature of the two CLC redox reactions.



$$\Delta G^\circ = +375.1 \text{ MJ/kmol}$$

$$\Delta H^\circ = +494.5 \text{ MJ/kmol}$$

$$\Delta H^\circ - \Delta G^\circ = T_o \Delta S^\circ = +119.4 \text{ MJ/kmol}$$

A CLC system operating on these two reactions is described in Sec. 3.2.

3.2 General Arrangement of Proposed CLC System. To generate work from the two reactions (19) and (20), a heat engine must be installed, straddling the oxidizer and reducer. McGlashan [3] proposed a CLC system using a Rankine cycle for this task, with a working fluid of metal vapor. However, this arrangement is impractical, as the Rankine cycle must exchange heat with the two redox reactions by indirect heat transfer at a temperature above that currently achievable.

Assuming that expansion through a turbine is used to generate shaft work, what is required is a means of generating a hot pressurized working fluid in the oxidizer, without using indirect heat exchange. In addition, heat must be supplied to the reducer to drive the endothermic reduction process, again, preferably without requiring indirect heat exchange. Separation between oxygen carrier and nitrogen of any stream that ultimately ends up in the reducer is also required. This last criterion ensures that the process maintains its utility as a viable carbon capture system.

Figure 3 shows one potential configuration that meets the above criteria. The system relies on the difference in phase between the reactants and products in both oxidizer and reducer to achieve the desired physical separation of species. Both reactors are arranged so that condensed phases travel downwards with gases traveling upwards; consequently, countercurrent, direct contact heat transfer takes place, simultaneously preheating the reactants and cooling the products. Using this mode of heat transfer enables the generation of two hot pressurized working fluids in the oxidizer: one, consisting of N_2 saturated with Zn vapor, and the other a stream of liquid Zn devoid of nitrogen. Likewise, in the reducer the reactor produces a stream of liquid Zn and a stream of almost pure, hot, and pressurized CO.

3.3 The Oxidizer. Looking at the oxidizer in detail, it consists of a long, tubular pressure vessel with a reaction zone located approximately at the reactor's midriff. The expectation is that this reactor can be operated in an entrained flow mode throughout due to the high temperatures of the oxidation reaction (approximately 1840 K), and consequently rapid chemical kinetics. In operation, a stoichiometric excess of liquid Zn enters at the top of the reactor. Likewise, compressed air enters at the reactor's bottom. The stream of liquid Zn is formed into a rain of droplets that fall through the reactor, under gravity, passing a rising updraft of hot product gases coming from the reaction zone beneath. These product gases consist of nitrogen saturated with Zn vapor. Countercurrent heat transfer between the streams takes place, simultaneously preheating the liquid Zn while cooling the gas stream. Due to the fall in the temperature of the gas stream, the bulk of the Zn vapor contained by the stream, condenses on the liquid Zn droplets, refluxing Zn species back down the reactor.

Turning to the reaction zone, only a portion of the liquid Zn is converted to ZnO. This oxide agglomerates in the reaction zone, which is arranged to have a recirculating flow pattern, forming a hail of particles. Some of these ZnO particles are blown upwards, out of the reaction zone with the product gases. These particles are collected by the falling rain of liquid Zn and returned to the reaction zone. Most of the ZnO hail, however, drops out of the bottom of the reaction zone, and then falls to the bottom of the reactor through the current of compressed air rising from beneath. Once again, direct contact heat exchange takes place, simultaneously

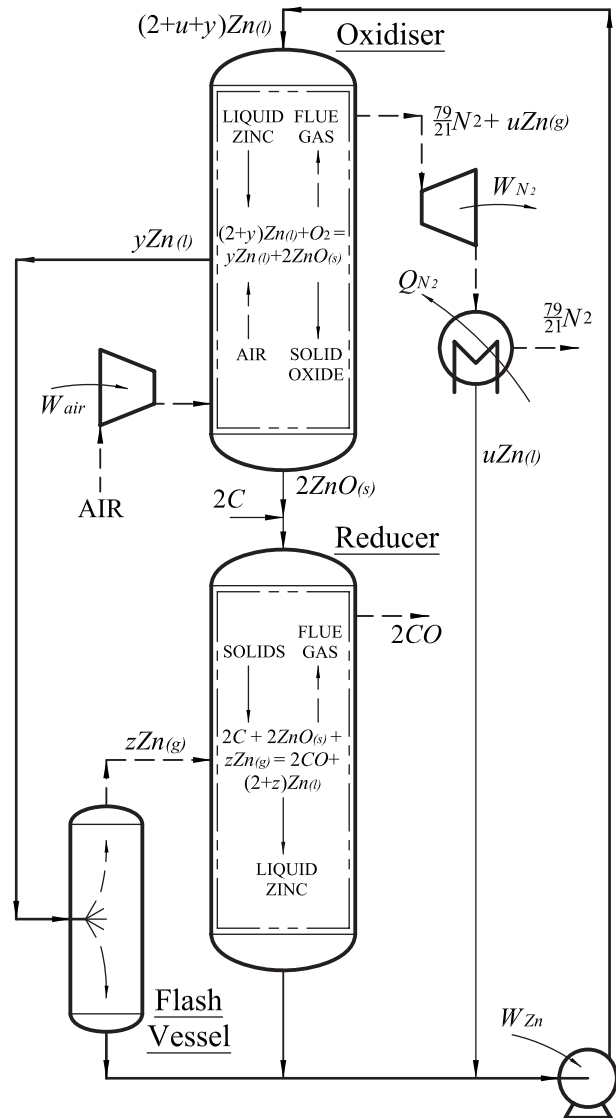


Fig. 3 Schematic diagram of CLC system avoiding indirect heat exchange and using a zinc flash vessel

cooling the ZnO hail, while preheating the air prior to its entry into the reaction zone. The resulting cooled ZnO powder passes out of the bottom of the reactor through a suitable sealing arrangement.

Considering now, the excess of liquid Zn that does not become oxidized to ZnO, after falling through the upper recuperation region of the oxidizer, the liquid is collected in an annular trough located adjacent to the reaction zone. Due to the recuperation of heat from the products in the top part of the reactor, and also heat received directly from the oxidation reaction by radiation, this Zn should be at a temperature approaching that of the oxidation reaction. The hot liquid is then passed out of the oxidizer, under pressure, and its heat content used to "drive" the endothermic reduction reaction as described later.

Returning to the top of the oxidizer, the N_2/Zn vapor stream leaving the top the reactor is at oxidizer pressure, P_{oxi} , and can be expanded through a gas turbine. Due to the fall in temperature in this process, most, but not all, of the Zn vapor in this stream will condense in the turbine forming Zn "moisture" or "wetness." This moisture is collected in a separator of some kind, placed at the exit of the turbine, and returned to the chemical loop as a stream of liquid. The exit gases from this turbine, still carrying some Zn vapor, are passed to a combined heat recovery steam generator

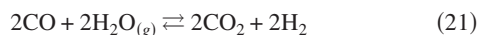
and zinc condenser (hereafter called the HRSG), where they are cooled in two stages. In the first stage, all but a vestigial amount of the Zn vapor is condensed in a cooler-condenser, which forms the superheater for the bottoming steam cycle described later. The cooling surface in this part of the HRSG is maintained at a temperature above the melting point of Zn (692 K) to avoid Zn solidifying and hence blocking the exchanger. The second part of the HRSG cools the now almost pure nitrogen leaving the first half of the exchanger. The heat released in this process performs a fraction of the boiling duties required by the steam cycle (the remainder taking place in the WGSR reactor). Using this arrangement, combined with the heat exchangers integral with the WGSR reactor, the pinch problem commonly found in bottoming steam plant of this type can be minimized [15]. Indeed, in the analysis that follows later, the pinch point was found to be at the water inlet end of the HRSG. Consequently, the nitrogen outgas from the HRSG is at a temperature approaching ambient as no feed heating is applied to the water substance entering the exchanger.

3.4 The Reducer. The reducer in this embodiment consists of a long, tubular pressure vessel, operated at a pressure markedly lower than that of the oxidizer. In operation, the vessel is fed, at its top, with the stream of solid ZnO from the oxidizer. However, before entering the reducer, the oxide is mixed with pulverized carbon as fuel. This mixture is arranged to fall through a rising updraft of product gases in similar fashion to the way liquid Zn falls through the oxidizer. As in the oxidizer, direct contact heat transfer recuperates heat from gaseous products to condensed reactants.

The reduction reaction is performed in the lower half of the reactor. Equilibrium thermodynamics pegs the temperature of this reaction to a low value determined by the operating pressure (approximately 1350 K in this case)—see Appendix B. Consequently, an entrained flow reaction zone is unlikely to affect complete conversion of ZnO to Zn due to slow kinetics at this temperature, and a fixed or fluidized bed reactor may be necessary to ensure sufficient residence time.

Regardless of how the reactor is configured, to drive the endothermic reduction reaction, the heat content of the liquid Zn exiting the oxidizer is used. The simplest arrangement is to pass this stream of hot liquid directly into the reduction zone where a release of “sensible” heat from the liquid would drive the reaction. However, this would necessitate a substantial amount of Zn flowing through the reducer. In addition, there would be a thermodynamic loss associated with the temperature drop from the liquid Zn to the reactants. Therefore, an alternative arrangement is suggested whereby Zn vapor is generated from the liquid Zn stream in an external flash vessel. This vapor can then be blown into the reducer, much as air is blown into a blast furnace, the vapor condensing within the reduction zone releasing latent heat. The liquid Zn produced inside the reducer, both by the condensation process and the reduction reaction itself, is then combined with the residual liquid exiting from the flash vessel, and recirculated back to the top of the oxidizer ready to complete another chemical loop.

3.5 The WGSR Reactor and Bottoming Cycle. The CLC system shown in Fig. 3 produces a stream of CO, which exits the reducer as a hot, pressurized gas. Therefore, a further oxidation stage converting CO to CO₂ is necessary. This can be achieved using the WGSR, thus



$$\Delta G^\circ = -57.2 \text{ MJ/kmol}$$

$$\Delta H^\circ = -82.3 \text{ MJ/kmol}$$

$$\Delta H^\circ - \Delta G^\circ = T_o \Delta S^\circ = -25.1 \text{ MJ/kmol}$$

To maximize the conversion of CO to CO₂ using the WGSR, a low reactor outlet temperature is required. Typical industrial

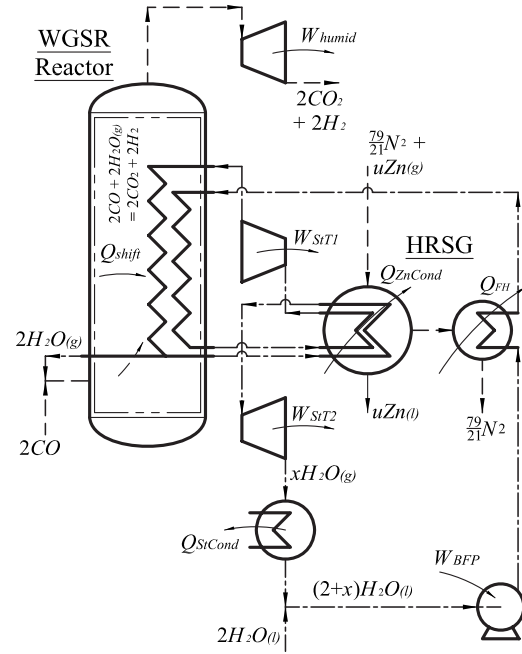


Fig. 4 Diagram of proposed bottoming steam cycle and WGSR

plants operate with an outlet temperature of approximately 460–520 K [16,17]—this compares with the temperature of the exit gas from the reducer of approximately 900 K. Added to this, the WGSR exhibits significant heat releases itself, which means that some way of cooling the reaction must be installed. This can be achieved by performing the reaction in two adiabatic stages, coupled with an interstage heat exchanger [17]; however, an alternative reactor design is suggested here which maximizes the recovery of useful energy—Fig. 4 shows the arrangement.

Looking at the WGSR reactor in detail, it consists of a tubular, catalytic reactor with two integral heat exchangers. The exchangers are arranged for countercurrent heat transfer between the catalyst bed (and hence the reacting mixture) and cooling streams consisting of either steam or water. The product gases exit the reactor and are passed directly to a dedicated exhaust gas turbine. The exit temperature from the reactor is constrained such that the exhaust temperature of this turbine is close to ambient; in the analysis described below the exhaust turbine’s exit temperature (TET) is assumed to equal T_o .

The heat rejected by the WGSR in its two heat exchangers is combined with that rejected from the HRSG cooling the N₂/Zn turbine exhaust to provide the heat addition for the bottoming steam cycle. The steam system, as shown in Fig. 4, has two turbines: a high pressure (HP) and a low-pressure (LP) cylinder. A single stage of reheat is installed between the two cylinders. The inlet pressure to the LP turbine, and hence the reheater exit pressure, is set to equal the reaction-side pressure of the WGSR reactor. This allows a fraction of the reheated steam to be extracted and mixed with the flue gas exiting the reducer, thereby providing the water substance required to complete the WGSR. The remaining steam is passed to the LP cylinder, the exhaust from which is discharged to a condenser in the conventional manner.

This completes the description of the major components of the system. Section 4 details a thermodynamic analysis of the complete cycle.

4 Mass and Energy Balance

4.1 General Approach. A mass and energy balance of the proposed system will now be described. All devices have been analyzed as a single control region and assumed to operate at

steady state. The control regions drawn within the three reactors have been selected so they are externally adiabatic—shown as double chain dotted (phantom) lines in Figs. 3 and 4. Turbomachinery, pumps, pipework, and the flash vessel are also assumed to be externally adiabatic. Heat exchangers are assumed to operate in countercurrent configuration and to exhibit zero approach temperature. However, at the hot end of the HRSG and the WGSR reactor, the steam temperature is restricted to a maximum value. This prevents the steam turbine seeing a steam temperature exceeding a maximum turbine inlet temperature (TIT)—set at 973 K in this work. Reactions are assumed to reach completion, though consideration of the chemical equilibrium of the reduction reaction is made, but only to determine the relationship between reactor pressure and temperature—see Appendix B. Pressure losses in pipework and equipment are assumed negligible.

Beginning with the “first law,” this has been applied to all control regions with the following sign convention:

$$\sum_{j=1}^{\text{no inlets}} H_j - \sum_{k=1}^{\text{no outlets}} H_k - \sum Q + \sum W = 0 \quad (\text{MJ/kmol}) \quad (22)$$

Total enthalpies, temperatures, and pressures are assumed throughout. Dynamic and hydrostatic head affects are ignored except in the zinc riser—i.e., the pipe carrying Zn exiting the zinc-circulating pump back to top of the oxidizer. However, the sensitivity of the plant’s performance to varying the height of this liquid column was examined and shown to be small, and hence an arbitrary height of 50 m was taken as an estimate of the eventual height of the plant.

4.2 Substance Thermodynamic Properties. The specific heat, Cp_i , of a given species, i , is calculated using the relationship as follows:

$$Cp_i = A_i + B_i T + C_i / T^2 \quad (\text{J/mol K}) \quad (23)$$

Values of the constants in Eq. (23) for each species have been extracted from Kubaschewski et al. [18] or from Kelley [19] and are given in Appendix A. To simplify calculations, average values of Cp_i have been taken. Hence, for species that remain gaseous throughout the process, the molar enthalpy of a species i is given by

$$h_{i,T} = \overline{Cp}_i (T - T_o) \quad (\text{J/mol}) \quad (24)$$

Equation (24) is also applied to solid ZnO and liquid Zn and H₂O. However, both Zn and H₂O change state in different parts of the process. As Zn_(g) only appears as a saturated vapor, the molar enthalpy of Zn vapor is given by the relation as follows:

$$h_{Zn,T} = \overline{Cp}_{Zn(l)} (T - T_o) + \Delta h_{Zn,T}^{\text{vap}} \quad (\text{J/mol}) \quad (25)$$

Where the latent heat of vaporization of Zn at temperature, T , is calculated using the relation of Yaws [20]:

$$\Delta h_{Zn,T}^{\text{vap}} = 137,616 \left(1 - \frac{T}{3170} \right)^{0.38} \quad (\text{J/mol}) \quad (26)$$

The enthalpy of steam is expressed in this work by the simple relationship of Moore and Siverding [21]—taking the enthalpy of liquid water at T_o as zero, it follows

$$h_{H_2O(g),T} = \Delta h_{H_2O,T_o}^{\text{vap}} + \overline{Cp}_{H_2O(g)} (T - T_o) \quad (\text{J/mol}) \quad (27)$$

where

$$\Delta h_{H_2O,T_o}^{\text{vap}} = 44,002 \quad \text{J/mol}$$

Ideal mixing is assumed for the various gas mixtures and the corresponding molar stream enthalpy and heat capacity are, respectively,

$$h_{\text{str},T} = \sum_{i=1}^s x_i h_{i,T} \quad (\text{J/mol}) \quad (28)$$

$$\overline{Cp}_{\text{str}} = \sum_{i=1}^s x_i \overline{Cp}_i \quad (\text{J/mol K}) \quad (29)$$

The saturated vapor pressure of Zn is given by the relationship of Maier [22], which (adapted to give pressure in bar) leads to

$$\ln p_{Zn,T}^{\text{sat}} = 21.0526 - \frac{15633.4}{T} - 1.051 \ln T - 2.8897 \times 10^{-4} T \quad (30)$$

Likewise, for water substance, the saturated vapor pressure (in bar) is given by adapting the relationship of Yaws [20], thus

$$\ln p_{H_2O,T}^{\text{sat}} = 62.136 - \frac{7258.2}{T} - 7.3037 \ln T + 5.5831 \times 10^{-9} T + 4.1684 \times 10^{-6} T^2 \quad (31)$$

4.3 Machinery. Starting with the two pumps in the system, liquids are assumed incompressible and, applying a hydraulic efficiency, η_{pump} , the shaftwork of the two pumps per mole of flow is given by

$$w_{\text{pump}} = \frac{\eta_{\text{pump}} \Delta P}{\rho} \quad (\text{J/mol}) \quad (32)$$

For the gas compressor, ideal gasses are assumed, and applying a polytropic efficiency, η_{comp} , the shaftwork per mole of flow is given by

$$w_{\text{comp}} = \overline{Cp}_{\text{str}} T_{\text{cold}} \left(1 - \left(\frac{P_{\text{hot}}}{P_{\text{cold}}} \right)^{\overline{Cp}_{\text{str}} \eta_{\text{comp}}} \right) \quad (\text{J/mol}) \quad (33)$$

Likewise, applying a polytropic efficiency, η_{turb} , gas turbine shaftwork per mole of flow is given by

$$w_{\text{turb}} = \overline{Cp}_{\text{str}} T_{\text{cold}} \left(\left(\frac{P_{\text{hot}}}{P_{\text{cold}}} \right)^{\overline{Cp}_{\text{str}} \eta_{\text{turb}}} - 1 \right) \quad (\text{J/mol}) \quad (34)$$

The N₂/Zn turbine expands a mixture of nitrogen saturated with Zn vapor, with Zn condensing as the expansion proceeds; consequently, the expansion is “wet” with liquid Zn droplets. Equation (32) does not apply to a wet expansion. However, to model this expansion process fully, the effect of supersaturation and droplet dynamics within the turbine would need to be considered. Constructing such a model is outside the scope of this work, so a simpler analysis is presented here.

First, it will be assumed that the Zn condensing in the turbine remains in thermal and physical equilibrium with the gas stream throughout the expansion process. Consequently, liquid Zn exits the machine as saturated liquid at the same temperature as the exhaust gas. Therefore, applying the first law to the whole turbine, the work output is given by (in J/mol)

$$w_{\text{turb}} = (n_{N_2} \overline{Cp}_{N_2} + n_{Zn}^{\text{in}} \overline{Cp}_{Zn(l)}) (T_{\text{in}} - T_{\text{out}}) + n_{Zn}^{\text{out}} \Delta h_{Zn,TET}^{\text{vap}} \quad (35)$$

From Dalton’s law, the quantity of Zn leaving the turbine as liquid is determined by the outlet saturation pressure (supersaturation is, therefore, ignored), thus

$$n_{Zn}^{\text{out}} = n_{Zn}^{\text{in}} - \frac{n_{N_2} p_{Zn}^{\text{sat}}}{P - p_{Zn}^{\text{sat}}} \quad (\text{mol/mol}) \quad (36)$$

Now, for an isentropic expansion, T_{out} can be found by equating the entropy of the gas/vapor/liquid mixture at the beginning and end of the expansion, thus

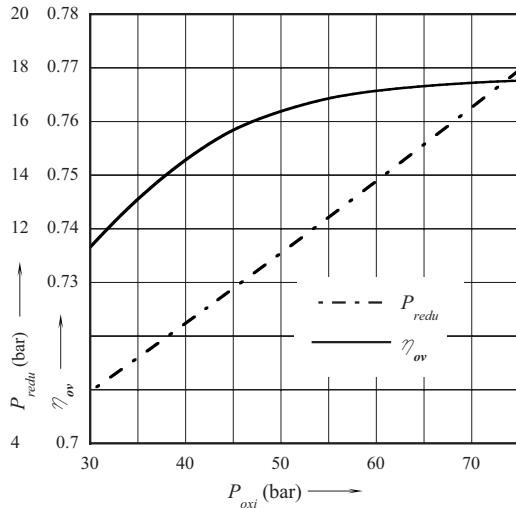


Fig. 5 Graph showing overall efficiency, η_{ov} , and reducer pressure, P_{redu} , versus P_{oxi} .

$$\begin{aligned}
 & n_{N_2} \left(\overline{Cp}_{N_2} \ln \frac{T_{in}}{T_o} - \mathfrak{R} \ln \frac{x_{N_2, in} P_{in}}{P_o} \right) + n_{Zn}^{V_{in}} \left(\overline{Cp}_{Zn(o)} \ln \frac{T_{in}}{T_o} + \frac{\Delta h_{Zn, TIT}^{vap}}{T_{in}} \right) \\
 &= n_{N_2} \left(\overline{Cp}_{N_2} \ln \frac{T_{out}}{T_o} - \mathfrak{R} \ln \frac{x_{N_2, out} P_{out}}{P_o} \right) \\
 &+ n_{Zn}^{V_{out}} \left(\overline{Cp}_{Zn(o)} \ln \frac{T_{out}}{T_o} + \frac{\Delta h_{Zn, TET}^{vap}}{T_{out}} \right) + n_{Zn}^{L_{out}} \left(\overline{Cp}_{Zn(o)} \ln \frac{T_{out}}{T_o} \right)
 \end{aligned} \quad (37)$$

Collecting terms and using an average value for the latent heat of Zn, it follows

$$\begin{aligned}
 & (n_{N_2} \overline{Cp}_{N_2} + n_{Zn}^{V_{in}} \overline{Cp}_{Zn(o)}) \ln \frac{T_{in}}{T_{out}} - n_{N_2} \mathfrak{R} \ln \frac{(P_{in} - P_{Zn}^{sat, in})}{(P_{out} - P_{Zn}^{sat, out})} \\
 &+ \Delta h_{Zn}^{vap} \left(\frac{n_{Zn}^{V_{in}}}{T_{in}} - \frac{(n_{Zn}^{V_{in}} - n_{Zn}^{L_{out}})}{T_{out}} \right) = 0
 \end{aligned} \quad (38)$$

Simultaneous solution of Eqs. (35), (36), and (38) yields the isentropic outlet temperature. The actual work output (in J/mol) and exit temperature can be found if an isentropic efficiency, η_{is} , is applied to Eq. (35), thus

$$\begin{aligned}
 w_{turb} &= \eta_{is} [(n_{N_2} \overline{Cp}_{N_2} + n_{Zn}^{V_{in}} \overline{Cp}_{Zn(o)}) (T_{in} - T_{out}^{is}) + n_{Zn}^{L_{out}} \overline{\Delta h}_{Zn}^{vap}] \\
 &= (n_{N_2} \overline{Cp}_{N_2} + n_{Zn}^{V_{in}} \overline{Cp}_{Zn(o)}) (T_{in} - T_{out}) + n_{Zn}^{L_{out}} \overline{\Delta h}_{Zn}^{vap}
 \end{aligned} \quad (39)$$

4.4 Results. Each of the components of the system has been analyzed using the assumptions and relationships given above. The gas compressor's polytropic efficiency was assumed to be 90%; all other efficiencies were taken as 85%. An exact, simultaneous, analytical solution is possible in some cases; however, the resulting equations are unwieldy and transcendental, so instead a numerical solution was obtained. To generate this solution, an equation-based process-engineering package, GPROMS, was used [23,24]. Results of this analysis are now presented; two graphs, Figs. 5 and 6, are given above, with P_{oxi} as the independent variable:

Figure 5, shows the variation in the "overall" efficiency, η_{ov} , (see Eq. (42)), and the pressure in the reducer, P_{redu} , with P_{oxi} . The pressure in the WGSR reactor and hence the steam reheater and LP cylinder inlet pressure are also equal to P_{redu} .

Figure 6 shows the variation in three temperatures with P_{oxi} , as follows.

- (i) T_{oxi} —the peak temperature occurring in the oxidizer—also

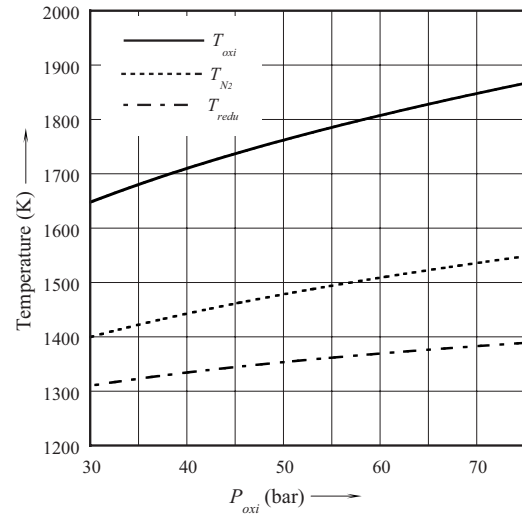


Fig. 6 Graph showing key process temperatures versus P_{oxi}

assumed to equal the temperature of liquid Zn entering the flash vessel. This is the highest temperature anywhere in the cycle.

- (ii) T_{N_2} —the TIT of the N_2/Zn turbine—assumed to equal the temperature of the Zn returned to the top of the oxidizer, T_{top} .
- (iii) T_{redu} —the peak temperature of the reducer—this affects, but does not determine T_{top} .

To illustrate the performance of the proposed arrangement, Fig. 7 shows a Sankey diagram of energy fluxes for the whole CLC system—where $P_{oxi}=35.0$ bar. To improve the readability of the figure, the enthalpy of liquid Zn was assumed to equal zero at the lowest temperature at which it occurs in the system; namely, at the exit to the HRSG. This means that only the energy added to the recirculating Zn appears on the diagram. The figure demonstrates the significant recirculation of energy within the chemical loop. It also shows the relative amount of energy converted to work compared with that leaving the system with the H_2 —equivalent to the calorific value of the H_2 . At first sight, the performance of the cycle appears to be rather poor, as the work output is not substantial by comparison with the calorific value of carbon. Indeed, using the definition of rational efficiency given earlier, for the condition shown in Fig. 7, Eq. (10) gives

$$\eta_{rat} = \frac{W_{irrev}}{(-\Delta G^o)_{(1)}} = \frac{161.1}{314.5} = 51.2\% \quad (40)$$

However, defining rational efficiency in this way punishes the work producing part of the system for the entropy generated by the system as a whole. This includes losses associated with parts of the system that exists, ostensibly, to produce H_2 only. Therefore, an alternative definition, which gives a work equivalent value to the H_2 produced in the process, can be defined as follows:

$$\eta_{H_2} = \frac{W_{irrev} + (-\Delta G^o)_{(4)}}{(-\Delta G^o)_{(5)}} = \frac{161.1 + 474.3}{788.8} = 80.6\% \quad (41)$$

Which of these two definitions of second law efficiency is used is partly a matter of whim and in this work, Haywood's first law, overall efficiency [1] has been used as the means of comparison instead—this gives:

$$\eta_{ov} = \frac{W_{irrev}}{(-\Delta H^o)_{(1)}} = \frac{161.1}{215.4} = 74.8\% \quad (42)$$

Hence, with $P_{oxi}=35.0$ bar an overall efficiency of approximately 75% has been obtained. The corresponding maximum TIT

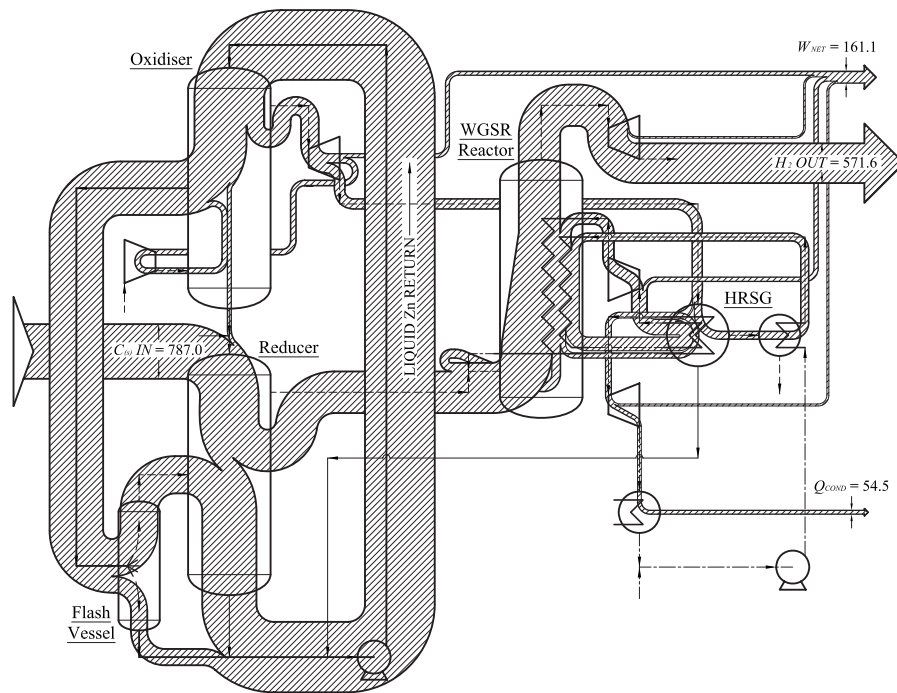


Fig. 7 Sankey diagram showing fluxes of energy in proposed system with $P_{\text{oxi}} = 35.0$ bar (W , Q , and H_{str} ; unit: MJ/kmol)

at this efficiency point is 1423 K—well below the peak temperature already achieved in gas turbine practice. Allowing for pressure losses, leakage, and finite approach temperature in heat exchangers is likely to reduce the calculated efficiency. However, further development of the system and changes in the configuration appear likely to improve the performance also. Suggested avenues for improving the cycle's performance include the following.

- (1) Optimizing the bottoming steam cycle with the addition of feedheating and perhaps using a dual pressure cycle.
- (2) Using more than one stage of flash to generate Zn vapor, perhaps with the addition of thermocompression between stages.
- (3) Avoiding, or mitigating the loss due to mixing large amounts of liquid Zn at different temperatures.

Nonetheless, despite the thermodynamic advantages of the cycle, there are a number of profound engineering difficulties that require resolution before a practical coal burning power system can be constructed. A list of the problems considered most significant is given below, in no order of importance.

- (1) Ash removal from the reducer, either molten or as “dry” powder, while avoiding fouling.
- (2) Separation of ash from liquid Zn and avoiding the loss of Zn from the cycle.
- (3) Corrosion (and erosion) caused by liquid Zn at high temperature in reactors, pipe work, the N_2/Zn vapor turbine and the HRSG.
- (4) Materials selection for high temperature reducing environments and those containing Zn vapor/liquid.
- (5) Vapor deposition of solid Zn on cool (sub 692 K) surfaces and consequent fouling of pipework, reactors and turbines.
- (6) Reducer and oxidizer reaction kinetics and hence the size, scale, and mode of operation of these reactors—entrained flow or fluidized bed?
- (7) Fuel borne sulfur removal.
- (8) Starting, stopping, and control.

5 Conclusions

A cycle capable of generating H_2 and power while capturing CO_2 has been proposed and evaluated. The cycle, based on chemical looping, uses two reactions and a recirculating oxygen carrier to oxidize hydrocarbons. Generally, in a chemical looping system the oxidation reaction is exothermic and the reduction reaction endothermic. Therefore, a heat engine can straddle the two reaction vessels, receiving heat from the oxidizer and rejecting heat to the reducer. The ratio between the standard state enthalpy of oxidation (ΔH°) of the oxygen carrier to that of the fuel, is greater than unity for most combinations of oxygen carrier and hydrocarbon. This fact enables the temperature ratio at which a chemical looping cycle operates to be reduced without sacrificing efficiency.

A system based around an oxidizing vessel, a reduction vessel, a shift reactor, a liquid metal flash vessel, and three turbines has been proposed. The system relies on the difference in phase between the reactants and products in both the oxidizer and reducer, to achieve a physical separation of species, in this case using a Zn based oxygen carrier. An overall efficiency of 75% has been obtained from the cycle analyzed without exceeding the peak temperature and pressure already achieved in current power generation practice.

Despite the thermodynamic advantages of the proposed cycle, significant challenges remain including ash removal from the reducer, carrier/ash separation, corrosion, metal vapor deposition and fouling, fuel-borne sulfur removal, and startup and shutdown processes. However, the high efficiency of the proposed cycle warrants ongoing attention to this potential technology solution for power generation with carbon capture.

Acknowledgment

The authors would like to thank the numerous colleagues who have supported and encouraged this work.

Nomenclature

Roman

- C = minimum number of species (components) from which a systems can be formed
 C_p = specific heat at constant pressure (J/mol K)
 F = degrees of freedom
 ΔG = reaction Gibbs function (MJ/kmol)
 ΔH = reaction enthalpy (MJ/kmol)
 H_i = enthalpy of stream i (MJ/kmol)
 h = specific enthalpy (J/mol)
 I = lost work
 K_p = equilibrium constant
 n_i = moles of species i , per mole of C or ZnO, respectively
 p = component pressure (bar)
 P = reactor pressure (bar)
 Q = heat flux (J/mol)
 \mathfrak{R} = universal gas constant=8.3144 (J/mol K)
 S = reaction entropy (MJ/kmol K)
 stp = standard temperature and pressure (298.15 K, 1.0 bar)
 T = thermodynamic temperature (K)
 W = shaft work (MJ/kmol)
 w = specific shaft work (J/mol)
 x = mole fraction

Greek

- Δ = change in property due to a chemical reaction—products minus reactants
 η = efficiency
 π = number of phases present at equilibrium
 ρ = molar density (mol/m³)
 δ = moles of liquid Zn present per mole of reduced ZnO
 Ψ = ratio of oxygen carrier to fuel oxidation reaction enthalpies

Subscripts/Superscripts

- adia = adiabatic
 cond = condenser
 cold = cold
 comp = compressor
 CR = control region
 eq = equilibrium condition
 (g) = gaseous state
 gen = generation
 humid = humid
 i = species index
 irrev = irreversible process
 is = isentropic
 L = liquid stream
 (l) = liquid state
 o = standard state (superscript) or sink condition (subscript)
 ov = overall
 oxi = oxidizer or oxidation
 rad = radiation
 rat = rational
 reac = reaction
 redu = reducer or reduction
 rej = rejection
 rev = reversible process
 (s) = solid state
 sat = saturation
 str = stream
 th = thermal
 turb = turbine
 V = vapor stream

- vap = vaporisation
 * = critical point

Appendix A: Values of Thermodynamic Constants

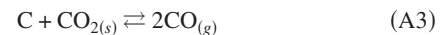
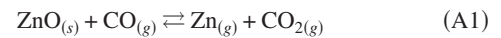
Table 1 shows the values of constants applied to Eq. (23).

Table 1 Values of constants applied to Eq. (23)

Species	A	B	C	T range (K)
C _(s)	17.166	0.00427	-879,228	298–2300
CO _(g)	28.428	0.00410	-46,055	298–2500
CO _{2(g)}	44.171	0.00904	-854,107	298–2500
H _{2(g)}	27.298	0.00327	50,242	298–3000
H _{2O_(g)}	30.019	0.01072	33,494	298–2500
H _{2O_(l)}	75.488			273–373
N _{2(g)}	27.884	0.00427		298–2500
O _{2(g)}	29.977	0.00419	-167,472	298–3000
Zn _(l)	31.401			692–1200
ZnO _(s)	49.027	0.00511	-912,722	298–1600

Appendix B: The Equilibrium of the Carbothermic Reduction in ZnO

The equilibrium of the reduction in ZnO by carbon has been investigated by a number of authors in connection with the zinc industry [13,22–26]. Conventional practice in industry is to generate Zn as vapor not liquid in coke fired furnaces. However, in the early part of the twentieth century, a number of attempts were made at producing liquid Zn by the direct reduction in zinc ores. Woolsey (as reported by Maier [22]) described one such attempt, calling it the Lungwitz process. A zinc blast furnace was used operating at a pressure of 2–3 bar as the test furnace. This and all other attempts at the direct generation of liquid Zn were fruitless. Maier [22] explained the cause of these failures using equilibrium thermodynamics and showed that to produce liquid Zn the reactor must operate above a critical pressure—notably higher than the 2–3 bar figure used by earlier workers. Hougen and Watson [13] re-examined Maier's work, they assumed the reduction to consist of three reactions:



Now, the Gibbs phase rule is given by

$$F = 2 + C - \pi \quad (\text{A4})$$

where F are the degrees of freedom, C is the minimum number of components from which the system can be formed, and π is the number of phases present at equilibrium.

The system of reactions, (A1)–(A3), can be formed from a minimum of two components, e.g., CO₂ and Zn_(g), and there are a maximum of four phases present: two solids, C and ZnO, one liquid, Zn_(l), and a vapor mixture. If all of these phases are extant simultaneously, the Gibbs phase rule gives $\rightarrow F = 2 + 2 - 4 = 0$.

This implies that there is a fixed (critical) point, invariant in composition, temperature, and pressure. To elucidate this critical point, the relevant equilibrium relations for the three reactions in terms of the reactions' equilibrium constants are as follows:

$$P = \frac{K_p^{A2}(1 + n_{\text{CO}} + 2n_{\text{Zn}})}{2n_{\text{Zn}}} \quad (\text{A5})$$

$$n_{\text{CO}} = \frac{K_p^{A2}}{K_p^{A2} + 2K_p^{A1}} \quad (\text{A6})$$

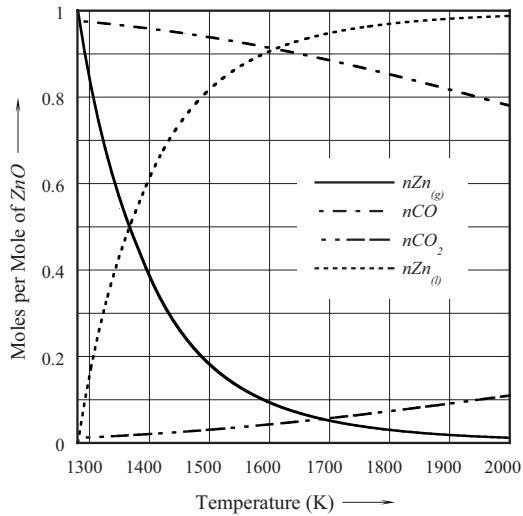


Fig. 8 Graph showing equilibrium concentrations of species in reducer versus reactor temperature

$$n_{CO_2} = \frac{K_p^{A1}}{K_p^{A2} + 2K_p^{A1}} \quad (A7)$$

$$n_{Zn} = \frac{(K_p^{A2})^3}{(K_p^{A2} + 2K_p^{A1})K_p^{A1}K_p^{A3}} \quad (A8)$$

where P is the reactor pressure, n_i is the number of gaseous moles present at equilibrium of each species i , per mole of ZnO reduced, and K_p is the equilibrium constant (superscripts correspond to equation numbers).

Now, defining δ as the moles of liquid Zn present per mole of reduced ZnO, the critical point at which liquid Zn first appears at equilibrium can be defined as follows:

$$n_{Zn} - \delta = 1 \quad (\delta \rightarrow 0) \quad (A9)$$

Designating the subscript “*” to refer to values at this critical pressure (point), Eq. (A8) can be rearranged as

$$2K_*^{A3}(K_*^{A1})^2 + K_*^{A1}K_*^{A2}K_*^{A3} - (K_*^{A2})^3 = 0 \quad (A10)$$

Solving for K_*^{A3} gives an equation relating the three equilibrium constants at the critical point, thus

$$K_*^{A3} = \frac{(K_*^{A2})^3}{2(K_*^{A1})^2 + K_*^{A1}K_*^{A2}} \quad (A11)$$

$$\Rightarrow P_* = \frac{K_*^{A2}(2K_*^{A2} + 3K_*^{A1})}{K_*^{A2} + 2K_*^{A1}} \quad (A12)$$

Using the relationships of Hougan and Watson [13] for the equilibrium constants, P_* and the corresponding temperature T_* can be found, thus

$$P_* = 5.67 \text{ bar}$$

$$T_* = 1280.8 \text{ K}$$

However, at temperatures above this point, one of the solid phases disappears as either ZnO or C is consumed completely depending on whether there is an excess of C or Zn atoms, respectively, therefore, the phase rule gives $\rightarrow F = 2 + 2 - 3 = 1$.

This leaves one degree of freedom and specifying, say, reactor pressure (as long as it is above 5.67 bar), locks the gas composition and temperature at equilibrium. It follows that the quantity of liquid Zn in the reactor has no bearing on the equilibrium concentration of any of the product gases as long as a small residue of

liquid remains. This also ensures that in principle all of the carbon entering the reactor is consumed. Hence, liquid Zn can be withdrawn from the reactor as fast as it forms.

Likewise at pressures below P_* , once again the phase rule indicates one degree of freedom, but this time there is no liquid present and both solid phases can coexist. This is not, however, a feasible operating point for the system described in this work.

In conclusion, the composition and temperature for a given reactor pressure can be found by simultaneous solution of Eqs. (A5)–(A8), using the appropriate relationships for the equilibrium constants. Figure 8 shows a plot of equilibrium product concentrations versus reactor temperature for conditions above the critical pressures, P_* .

References

- [1] Haywood, R. W., 1991, *Analysis of Engineering Cycles—Power, Refrigeration and Gas Liquefaction Plant*, 4th ed., Pergamon, Oxford, pp. 38–42.
- [2] Haywood, R. W., 1974, “A Critical Review of the Theorems of Thermodynamic Availability, With Concise Formulations—Part I: Availability,” *J. Mech. Eng. Sci.*, **16**, pp. 160–173.
- [3] McGlashan, N. R., 2008, “Chemical Looping Combustion—A Thermodynamic Study,” *Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci.*, **222**, pp. 1005–1019.
- [4] Horlock, J. H., 2002, *Combined Power Plants*, 1st ed., Krieger, Malabar, FL, pp. 35–36.
- [5] Richter, H. J., and Knoche, K. F., 1983, “Reversibility of Combustion Processes,” *Efficiency and Costing—Second Law Analysis of Processes* (ACS Symposium Series No. 235), American Chemical Society, Washington, DC, pp. 71–85.
- [6] Ishida, M., and Jin, N., 1997, “CO₂ Recovery in a Power Plant With Chemical Looping Combustion,” *Energy Convers. Manage.*, **38**, pp. S187–S192.
- [7] Brandvoll, Ø., and Bolland, O., 2004, “Inherent CO₂ Capture Using Chemical Looping Combustion in a Natural Gas Fired Cycle,” *ASME J. Eng. Gas Turbines Power*, **126**, pp. 316–321.
- [8] Naqvi, R., and Bolland, O., 2007, “Multi-Stage Chemical Looping Combustion (CLC) for Combined Cycles With CO₂ Capture,” *Int. J. Greenhouse Gas Control*, **1**, pp. 19–30.
- [9] Mattisson, T., García-Labiano, F., Kronberger, B., Lyngfelt, A., Adánez, J., and Hofbauer, H., 2007, “Chemical-Looping Combustion Using Syngas as Fuel,” *Int. J. Greenhouse Gas Control*, **1**, pp. 158–169.
- [10] Leion, H., Mattisson, T., and Lyngfelt, A., 2008, “Solid Fuels in Chemical-Looping Combustion,” *Int. J. Greenhouse Gas Control*, **2**, pp. 180–193.
- [11] Anheden, M., and Svedberg, G., 1998, “Exergy Analysis of Chemical-Looping-Combustion Systems,” *Energy Convers. Manage.*, **39**, pp. 1967–1980.
- [12] McGlashan, N. R., Heyes, A. L., and Marquis, A. J., 2007, “Carbon Capture and Reduced Irreversibility Combustion Using Chemical Looping,” *ASME Paper No. GT2007-28116*.
- [13] Hougan, O. A., and Watson, K. M., 1947, *Chemical Process Principles—Part 2 Thermodynamics*, 1st ed., Wiley, New York, pp. 737–741.
- [14] Jerndal, E., Mattisson, T., and Lyngfelt, A., 2006, “Thermal Analysis of Chemical-Looping Combustion,” *Trans. Inst. Chem. Eng., Part A*, **84**, pp. 795–806.
- [15] Winterbone, D. E., 1997, *Advanced Thermodynamics for Engineers*, 1st ed., Arnold, London, pp. 47–48.
- [16] Berkowitz, N., 1979, *An Introduction to Coal Technology*, 1st ed., Academic, New York, p. 275.
- [17] Newsome, D. S., 1980, “The Water-Gas Shift Reaction,” *Catal. Rev. - Sci. Eng.*, **21**, pp. 275–318.
- [18] Kubaschewski, O., Evans, E. LL., and Alcock, C. B., 1967, *Metallurgical Thermochemistry*, 4th ed., Pergamon, New York, pp. 303–447.
- [19] Kelley, K. K., 1949, “Critical Evaluation of High-Temperature Heat Capacities of Inorganic Compounds,” *U.S. Bureau of Mines Bulletin No. 476*.
- [20] Yaws, C. L., 1999, *Chemical Properties Handbook*, 1st ed., McGraw-Hill, New York, pp. 109–110.
- [21] Moore, J. M., and Siverding, C. H., 1976, *Two-Phase Steam Flow in Turbines and Separators*, 1st ed., Hemisphere, Washington, DC, p. 22.
- [22] Maier, C. G., 1930, “Zinc Smelting From a Chemical and Thermodynamic Viewpoint,” *U.S. Bureau of Mines Bulletin No. 324*.
- [23] Barton, P. I., and Pantelides, C. C., 1994, “Modelling of Combined Discrete/Continuous Processes,” *AIChE J.*, **40**, pp. 966–979.
- [24] Oh, M., and Pantelides, C. C., 1996, “A Modelling and Simulation Language for Combined Lumped and Distributed Parameter Systems,” *Comput. Chem. Eng.*, **20**, pp. 611–633.
- [25] Schumann, R., and Schadler, H. W., 1960, “Chemistry and Physics of Zinc Technology,” *Zinc—The Science and Technology of the Metal, Its Alloys and Compounds* (ACS Monograph No. 142), C. H. Mathewson, ed., Reinhold Publishing Corporation, New York, pp. 65–102.
- [26] Azakami, T., 1985, “Thermodynamic Studies on Reduction of Zinc Oxide,” *Zinc '85, Proceedings of the International Symposium on Extractive Metallurgy of Zinc*, Tokyo, Japan, The Mining and Metallurgical Institute of Japan, Tokyo, Japan, pp. 201–216.

Optimal Tuner Selection for Kalman Filter-Based Aircraft Engine Performance Estimation

Donald L. Simon

Sanjay Garg

NASA Glenn Research Center,
21000 Brookpark Road,
MS 77-1 Cleveland, OH 44135

A linear point design methodology for minimizing the error in on-line Kalman filter-based aircraft engine performance estimation applications is presented. This technique specifically addresses the underdetermined estimation problem, where there are more unknown parameters than available sensor measurements. A systematic approach is applied to produce a model tuning parameter vector of appropriate dimension to enable estimation by a Kalman filter, while minimizing the estimation error in the parameters of interest. Tuning parameter selection is performed using a multivariable iterative search routine that seeks to minimize the theoretical mean-squared estimation error. This paper derives theoretical Kalman filter estimation error bias and variance values at steady-state operating conditions, and presents the tuner selection routine applied to minimize these values. Results from the application of the technique to an aircraft engine simulation are presented and compared with the conventional approach of tuner selection. Experimental simulation results are found to be in agreement with theoretical predictions. The new methodology is shown to yield a significant improvement in on-line engine performance estimation accuracy. [DOI: 10.1115/1.3157096]

1 Introduction

An emerging approach in the field of aircraft engine controls and health management is the inclusion of real-time on-board models for the in-flight estimation of engine performance variations [1–3]. This technology, typically based on Kalman filter concepts, enables the estimation of unmeasured engine performance parameters, which can be directly utilized by controls, prognostics, and health management applications. A challenge that complicates this practice is the fact that an aircraft engine's performance is affected by its level of degradation, generally described in terms of unmeasurable health parameters such as efficiencies and flow capacities related to each major engine module. Through Kalman filter-based estimation techniques, the level of engine performance degradation can be estimated, given that there are at least as many sensors as parameters to be estimated. However, in an aircraft engine the number of sensors available is typically less than the number of health parameters presenting an underdetermined estimation problem. A common approach to address this shortcoming is to estimate a subset of the health parameters, referred to as model tuning parameters. While this approach enables on-line Kalman filter-based estimation, it can result in "smearing" the effects of unestimated health parameters onto those that are estimated and, in turn, introduce error in the accuracy of overall model-based performance estimation applications. Recently, Litt [4] presented an approach based on singular value decomposition (SVD) that selects a model tuning parameter vector of low-enough dimension to be estimated by a Kalman filter. The model tuning parameter vector, q , was constructed as a linear combination of all health parameters, h , given by

$$q = V^*h \quad (1)$$

where the transformation matrix, V^* , is selected, applying singular value decomposition to capture the overall effect of the larger set

of health parameters on the engine variables as closely as possible in the least-squares sense. In this paper a new linear point design technique, which applies a systematic approach to optimal tuning parameter selection will be presented. This technique, like the one presented in Ref. [4], also defines a transformation matrix, V^* , used to construct a tuning parameter vector that is a linear combination of all health parameters, and of low-enough dimension to enable Kalman filter estimation. The new approach optimally selects the transformation matrix, V^* , to minimize the theoretical steady-state estimation error in the engine performance parameters of interest. There is no known closed form solution for optimally selecting V^* to satisfy this objective. Therefore, a multivariable iterative search routine is applied to perform this function.

Sections 2–5 of this paper are organized as follows. First, the mathematical formulation of the parameter estimation problem is presented, and theoretical estimation error values are derived assuming linear steady-state operating conditions. The theoretical estimation error information is directly used by the iterative search routine applied to optimally select the Kalman filter tuning parameter vector, which is described next. Example estimation results from the application of the new methodology to an aircraft turbofan engine simulation are then presented and compared with the conventional approach of tuning parameter selection, the SVD tuner selection approach presented in Ref. [4], and the maximum a posteriori (MAP) performance estimation approach commonly applied for off-line (ground-based) aircraft engine gas path analysis applications [5,6]. After the example, there is a discussion of practical considerations for applying the method, and a discussion of future work. Finally, conclusions are presented.

2 Problem Formulation

The discrete linear time-invariant engine state-space equations about a linear design point are given as

$$\begin{aligned} x_{k+1} &= Ax_k + Bu_k + Lh_k + w_k \\ y_k &= Cx_k + Du_k + Mh_k + v_k \\ z_k &= Fx_k + Gu_k + Nh_k \end{aligned} \quad (2)$$

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 20, 2009; final manuscript received April 27, 2009; published online December 2, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Gas Turbine Technical Congress and Exposition, Orlando, FL, June 8–12, 2009.

where k is the time index, x is the vector of state variables, u is the vector of control inputs, y is the vector of measured outputs, and z is the vector of auxiliary (unmeasured) model outputs. The vector h represents the engine health parameters, which induce shifts in other variables, as the health parameters deviate from their nominal values. The vectors w and v are uncorrelated zero-mean white noise input sequences. Q will be used to denote the covariance of w , and R to denote the covariance of v . The matrices A , B , C , D , F , G , L , M , and N are of appropriate dimension. The health parameters, represented by the vector h , are unknown inputs to the system. They may be treated as a set of biases, and are thus modeled without dynamics. With this interpretation Eq. (2) can be written as

$$\begin{aligned} \begin{bmatrix} x_{k+1} \\ h_{k+1} \end{bmatrix} &= \underbrace{\begin{bmatrix} A & L \\ 0 & I \end{bmatrix}}_{A_{xh}} \underbrace{\begin{bmatrix} x_k \\ h_k \end{bmatrix}}_{x_{xh,k}} + \underbrace{\begin{bmatrix} B \\ 0 \end{bmatrix}}_{B_{xh}} u_k + \underbrace{\begin{bmatrix} w_k \\ w_{h,k} \end{bmatrix}}_{w_{xh,k}} \\ &= A_{xh} x_{xh,k} + B_{xh} u_k + w_{xh,k} \\ y_k &= \underbrace{\begin{bmatrix} C & M \end{bmatrix}}_{C_{xh}} \underbrace{\begin{bmatrix} x_k \\ h_k \end{bmatrix}}_{x_{xh,k}} + D u_k + v_k \\ &= C_{xh} x_{xh,k} + D u_k + v_k \\ z_k &= \underbrace{\begin{bmatrix} F & N \end{bmatrix}}_{F_{xh}} \underbrace{\begin{bmatrix} x_k \\ h_k \end{bmatrix}}_{x_{xh,k}} + G u_k = F_{xh} x_{xh,k} + G u_k \end{aligned} \quad (3)$$

The vector w_{xh} is zero-mean white noise associated with the augmented state vector, $[x^T \ h^T]^T$, with a covariance of Q_{xh} . w_{xh} consists of the original state process noise, w , concatenated with the process noise associated with the health parameter vector, w_h

$$w_{xh,k} = \begin{bmatrix} w_k \\ w_{h,k} \end{bmatrix} \quad (4)$$

The eigenvalues of A_{xh} consist of the original eigenvalues of A , plus an additional $\dim(h)$ eigenvalues located at 1.0 on the unit circle due to the augmentation. Thus, the new augmented system given in Eq. (3) has at least as many eigenvalues located on the unit circle as there are elements of h . Once the h vector is appended to the state vector, it may be directly estimated, provided that the realization in Eq. (3) is observable. Using this formulation, the number of health parameters that can be estimated is limited to the number of sensors, the dimension of y [7]. Since in an aircraft gas turbine engine there are usually fewer sensors than health parameters, the problem becomes one of choosing the best set of tuners for the application. This paper presents a systematic methodology for the optimal selection of a model tuning parameter vector, q , of low-enough dimension to be estimated by a Kalman filter, while minimizing the estimation error in the model variables of interest. Sections 2.1–2.4 will cover the steps in the problem setup. This includes construction of the reduced order state-space model, formulation of the Kalman filter estimator, calculation of the mean sum of squared estimation errors, and optimal selection of the transformation matrix to minimize the estimation error.

2.1 Reduced-Order State-Space Model. The model tuning parameter vector, q , is constructed as a linear combination of all health parameters, h , given by

$$q = V^* h \quad (5)$$

where $q \in \mathbb{R}^m$, $h \in \mathbb{R}^p$, $m < p$, and V^* is an $m \times p$ transformation matrix of rank m , applied to construct the tuning parameter vector.

An approximation of the health parameter vector, \hat{h} , can be obtained as

$$\hat{h} = V^{*\dagger} q \quad (6)$$

where $V^{*\dagger}$ is the pseudo-inverse of V^* . Substituting Eq. (6) into Eq. (3) yields the following reduced-order state-space equations that will be used to formulate the Kalman filter:

$$\begin{aligned} \begin{bmatrix} x_{k+1} \\ q_{k+1} \end{bmatrix} &= \underbrace{\begin{bmatrix} A & L V^{*\dagger} \\ 0 & I \end{bmatrix}}_{A_{xq}} \underbrace{\begin{bmatrix} x_k \\ q_k \end{bmatrix}}_{x_{xq,k}} + \underbrace{\begin{bmatrix} B \\ 0 \end{bmatrix}}_{B_{xq}} u_k + \underbrace{\begin{bmatrix} w_k \\ w_{q,k} \end{bmatrix}}_{w_{xq,k}} \\ &= A_{xq} x_{xq,k} + B_{xq} u_k + w_{xq,k} \\ y_k &= \underbrace{\begin{bmatrix} C & M V^{*\dagger} \end{bmatrix}}_{C_{xq}} \underbrace{\begin{bmatrix} x_k \\ q_k \end{bmatrix}}_{x_{xq,k}} + D u_k + v_k = C_{xq} x_{xq,k} + D u_k + v_k \\ z_k &= \underbrace{\begin{bmatrix} F & N V^{*\dagger} \end{bmatrix}}_{F_{xq}} \underbrace{\begin{bmatrix} x_k \\ q_k \end{bmatrix}}_{x_{xq,k}} + G u_k = F_{xq} x_{xq,k} + G u_k \end{aligned} \quad (7)$$

The state process noise, w_{xq} , and its associated covariance, Q_{xq} , for the reduced-order system are calculated as

$$\begin{aligned} w_{xq,k} &= \begin{bmatrix} I & 0 \\ 0 & V^* \end{bmatrix} w_{xh,k} = \begin{bmatrix} I & 0 \\ 0 & V^* \end{bmatrix} \begin{bmatrix} w_k \\ w_{h,k} \end{bmatrix} \\ Q_{xq} &= \begin{bmatrix} I & 0 \\ 0 & V^* \end{bmatrix} Q_{xh} \begin{bmatrix} I & 0 \\ 0 & V^* \end{bmatrix}^T \end{aligned} \quad (8)$$

2.2 Kalman Filter Formulation. In this study, steady-state Kalman filtering is applied. This means that while the Kalman filter is a dynamic system, the state estimation error covariance matrix and the Kalman gain matrix are invariant—instead of updating these matrices each time step they are held constant. Given the reduced-order linear state-space equations shown in Eq. (7), the state estimation error covariance matrix, P_∞ , is calculated by solving the following Riccati equation [8]:

$$P_\infty = A_{xq} P_\infty A_{xq}^T - A_{xq} P_\infty C_{xq}^T (C_{xq} P_\infty C_{xq}^T + R)^{-1} C_{xq} P_\infty A_{xq}^T + Q_{xq} \quad (9)$$

The steady-state Kalman filter gain, K_∞ , can then be calculated as follows [8]:

$$K_\infty = P_\infty C_{xq}^T (C_{xq} P_\infty C_{xq}^T + R)^{-1} \quad (10)$$

and assuming steady-state open-loop operation ($u=0$), the Kalman filter estimator takes the following form:

$$\hat{x}_{xq,k} = A_{xq} \hat{x}_{xq,k-1} + K_\infty (y_k - C_{xq} A_{xq} \hat{x}_{xq,k-1}) \quad (11)$$

The reduced-order state vector estimate, \hat{x}_{xq} , produced by Eq. (11) can be used to produce an estimate of the augmented state vector, and the auxiliary parameter vector as follows:

$$\begin{aligned} \hat{x}_{xh,k} &= \begin{bmatrix} I & 0 \\ 0 & V^{*\dagger} \end{bmatrix} \hat{x}_{xq,k} \\ \hat{z}_k &= [F \quad N V^{*\dagger}] \hat{x}_{xq,k} \end{aligned} \quad (12)$$

2.3 Analytical Derivation of Estimation Error. The estimation errors in $\hat{x}_{xh,k}$ and \hat{z}_k are defined as the difference between the estimated and actual values

$$\begin{aligned} \tilde{x}_{xh,k} &= \hat{x}_{xh,k} - x_{xh,k} \\ \tilde{z}_k &= \hat{z}_k - z_k \end{aligned} \quad (13)$$

Due to the underdetermined nature of the estimation problem, it will be impossible for the Kalman filter estimator to completely restore all information when transforming \hat{q} into \hat{h} . As such, the Kalman filter will be a biased estimator (i.e., the expected values of $\tilde{x}_{xh,k}$ and \tilde{z}_k will be nonzero). The estimation errors can be considered to consist of two components: an estimation error bias and an estimation variance. The estimation error bias vectors are equivalent to the mean estimation error vectors defined as

$$\begin{aligned}\bar{\tilde{x}}_{xh,k} &= E[\tilde{x}_{xh,k}] = E[\hat{x}_{xh,k} - x_{xh,k}] \\ \bar{\tilde{z}}_k &= E[\tilde{z}_k] = E[\hat{z}_k - z_k]\end{aligned}\quad (14)$$

where the operator $E[\bullet]$ represents the expected value of the argument. The variance of the estimates can be found by constructing their respective estimation covariance matrices

$$\begin{aligned}P_{\hat{x}_{xh,k}} &= E[(\hat{x}_{xh,k} - E[\hat{x}_{xh,k}])(\hat{x}_{xh,k} - E[\hat{x}_{xh,k}])^T] \\ P_{\hat{z}_k} &= E[(\hat{z}_k - E[\hat{z}_k])(\hat{z}_k - E[\hat{z}_k])^T]\end{aligned}\quad (15)$$

Diagonal elements of the covariance matrices will reflect the variance in individual parameter estimates, while off-diagonal elements reflect the covariance between parameter estimates. The overall sum of squared estimation errors (*SSEE*) can be obtained by combining the estimation error bias and estimation variance information as

$$\begin{aligned}SSEE(\hat{x}_{xh,k}) &= \bar{\tilde{x}}_{xh,k}^T \bar{\tilde{x}}_{xh,k} + \text{tr}\{P_{\hat{x}_{xh,k}}\} \\ SSEE(\hat{z}_k) &= \bar{\tilde{z}}_k^T \bar{\tilde{z}}_k + \text{tr}\{P_{\hat{z}_k}\}\end{aligned}\quad (16)$$

where $\text{tr}\{\bullet\}$ represents the trace (sum of the diagonal elements) of the matrix. In this paper, theoretical values for each error component will be derived assuming steady-state open-loop ($u=0$) operating conditions. First, the estimation error bias is derived, followed by a derivation of the estimation variance.

2.3.1 Estimation Error Bias. The estimation error biases, $\bar{\tilde{x}}_{xh,k}$ and $\bar{\tilde{z}}_k$, can be analytically derived for an arbitrary health parameter vector, h , at steady-state operating conditions. This is done taking advantage of the following expected value properties at steady-state open-loop operating conditions:

$$\begin{aligned}E[x_{k+1}] &= E[x_k] = x_{SS} \\ E[h_k] &= h \\ E[x_{xh,k}] &= x_{xh,SS} \\ E[y_k] &= y_{SS} \\ E[z_k] &= z_{SS} \\ E[u_k] &= 0 \\ E[w_k] &= 0 \\ E[v_k] &= 0 \\ E[\hat{x}_{xq,k}] &= E[\hat{x}_{xq,k-1}] = \bar{\tilde{x}}_{xq,SS} \\ E[\hat{x}_{xh,k}] &= \bar{\tilde{x}}_{xh,SS} \\ E[\hat{z}_k] &= \bar{\tilde{z}}_{SS}\end{aligned}\quad (17)$$

where the subscript SS denotes steady-state operation. By taking expected values of Eq. (2), x_{SS} , y_{SS} , and z_{SS} can be written as functions of the health parameter vector h

$$\begin{aligned}E[x_{k+1}] &= A \cdot E[x_k] + B \cdot E[u_k] + L \cdot E[h_k] + E[w_k] \\ x_{SS} &= Ax_{SS} + Lh \\ x_{SS} &= (I - A)^{-1}Lh\end{aligned}\quad (18)$$

$$\begin{aligned}E[y_k] &= C \cdot E[x_k] + D \cdot E[u_k] + M \cdot E[h_k] + E[v_k] \\ y_{SS} &= Cx_{SS} + Mh \\ y_{SS} &= (C(I - A)^{-1}L + M)h\end{aligned}\quad (19)$$

$$\begin{aligned}E[z_k] &= F \cdot E[x_k] + G \cdot E[u_k] + N \cdot E[h_k] \\ z_{SS} &= Fx_{SS} + Nh \\ z_{SS} &= (F(I - A)^{-1}L + N)h\end{aligned}\quad (20)$$

Next, by taking expected values of both sides of Eq. (11), the expected value of $\hat{x}_{xq,k}$ can be obtained as a function of y_{SS}

$$\begin{aligned}E[\hat{x}_{xq,k}] &= A_{xq} \cdot E[\hat{x}_{xq,k-1}] + K_{\infty}(E[y_k] - C_{xq}A_{xq} \cdot E[\hat{x}_{xq,k-1}]) \\ \bar{\tilde{x}}_{xq,SS} &= A_{xq}\bar{\tilde{x}}_{xq,SS} + K_{\infty}(y_{SS} - C_{xq}A_{xq}\bar{\tilde{x}}_{xq,SS}) \\ \bar{\tilde{x}}_{xq,SS} &= (I - A_{xq} + K_{\infty}C_{xq}A_{xq})^{-1}K_{\infty}y_{SS}\end{aligned}\quad (21)$$

Then, making the substitution $y_{SS} = (C(I - A)^{-1}L + M)h$ given in Eq. (19), the expected steady-state value of $\hat{x}_{xq,k}$ can be written as a function of h

$$\bar{\tilde{x}}_{xq,SS} = (I - A_{xq} + K_{\infty}C_{xq}A_{xq})^{-1}K_{\infty}(C(I - A)^{-1}L + M)h \quad (22)$$

The steady-state augmented state estimation error bias can then be found, and partitioned into error bias information for the original state vector, $\bar{\tilde{x}}_{SS}$, and the health parameter vector, $\bar{\tilde{h}}_{SS}$, by combining Eqs. (12), (14), (18), and (22) to yield

$$\begin{aligned}\bar{\tilde{x}}_{xh,SS} &= E[\hat{x}_{xh,k} - x_{xh,k}] \\ \bar{\tilde{x}}_{xh,SS} &= \bar{\tilde{x}}_{xh,SS} - x_{xh,SS} \\ \bar{\tilde{x}}_{xh,SS} &= \begin{bmatrix} \bar{\tilde{x}}_{SS} \\ \bar{\tilde{h}}_{SS} \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & V^{*\dagger} \end{bmatrix} \bar{\tilde{x}}_{xq,SS} - x_{xh,SS} \\ \bar{\tilde{x}}_{xh,SS} &= \begin{bmatrix} \bar{\tilde{x}}_{SS} \\ \bar{\tilde{h}}_{SS} \end{bmatrix} = \underbrace{\begin{bmatrix} \begin{bmatrix} I & 0 \\ 0 & V^{*\dagger} \end{bmatrix} (I - A_{xq} + K_{\infty}C_{xq}A_{xq})^{-1} \dots \\ \times K_{\infty}[C(I - A)^{-1}L + M] \dots \\ - \begin{bmatrix} (I - A)^{-1}L \\ I \end{bmatrix} \end{bmatrix}}_{G_{xh}} h \\ \bar{\tilde{x}}_{xh,SS} &= \begin{bmatrix} \bar{\tilde{x}}_{SS} \\ \bar{\tilde{h}}_{SS} \end{bmatrix} = G_{xh}h\end{aligned}\quad (23)$$

The steady-state auxiliary parameter estimation error bias can also be derived by combining Eqs. (12), (14), (20), and (22) to yield

$$\begin{aligned}\bar{\tilde{z}}_{SS} &= E[\hat{z}_k - z_k] \\ \bar{\tilde{z}}_{SS} &= \bar{\tilde{z}}_{SS} - z_{SS} \\ \bar{\tilde{z}}_{SS} &= [F \quad NV^{*\dagger}] \bar{\tilde{x}}_{xq,SS} - z_{SS}\end{aligned}$$

$$\bar{z}_{SS} = \underbrace{\begin{bmatrix} [F \quad NV^{*\dagger}](I - A_{xq} + K_{\infty} C_{xq} A_{xq})^{-1} \cdots \\ \times K_{\infty} [C[I - A]^{-1} L + M] \cdots \\ - [F(I - A)^{-1} L + N] \end{bmatrix}}_{G_z} h \quad (24)$$

$$\bar{z}_{SS} = G_z h$$

The estimation error bias equations given in Eqs. (23) and (24) are functions of an arbitrary health parameter vector, h . As such they are representative of the parameter estimation error biases in a single engine, at a given point in its lifetime of use, where its deterioration is represented by the health parameter vector h . The average sum of squared estimation error biases across a fleet of engines can be calculated as

$$\begin{aligned} \bar{x}_{xh, \text{fleet}}^2 &= E[\bar{x}_{xh, SS}^T \bar{x}_{xh, SS}] = E[\text{tr}\{\bar{x}_{xh, SS} \bar{x}_{xh, SS}^T\}] \\ &= E[\text{tr}\{G_{xh} h h^T G_{xh}^T\}] \\ &= \text{tr}\{G_{xh} \cdot \underbrace{E[hh^T]}_{P_h} G_{xh}^T\} \\ &= \text{tr}\{G_{xh} P_h G_{xh}^T\} \end{aligned} \quad (25)$$

$$\begin{aligned} \bar{z}_{\text{fleet}}^2 &= E[\bar{z}_{SS}^T \bar{z}_{SS}] = E[\text{tr}\{\bar{z}_{SS} \bar{z}_{SS}^T\}] \\ &= E[\text{tr}\{G_z h h^T G_z^T\}] \\ &= \text{tr}\{G_z \cdot \underbrace{E[hh^T]}_{P_h} G_z^T\} \\ &= \text{tr}\{G_z P_h G_z^T\} \end{aligned} \quad (26)$$

where the matrix P_h , defined as $E[hh^T]$, reflects a priori or historical knowledge of the covariance in the health parameters across all engines. If available, it can be used to predict the sum of squared estimation errors biases, as shown in Eqs. (24) and (25).

2.3.2 Estimation Variance. Next, derivations are presented for the augmented state estimate and auxiliary parameter estimate covariance matrices, $P_{\hat{x},k}$ and $P_{\hat{z},k}$, respectively. These matrices will be calculated as a function of the reduced-order state vector estimation covariance matrix, $P_{\hat{x}q,k}$, which is defined as

$$P_{\hat{x}q,k} = E \left[\underbrace{(\hat{x}_{xq,k} - E[\hat{x}_{xq,k}])}_{\varepsilon_{xq,k}} (\hat{x}_{xq,k} - E[\hat{x}_{xq,k}])^T \right] \quad (27)$$

where the vector $\varepsilon_{xq,k}$ is defined as the residual between $\hat{x}_{xq,k}$ at time k and its expected value. Since $E[\hat{x}_{xq,k}] = \bar{x}_{xq,SS}$, $\varepsilon_{xq,k}$ can be obtained by subtracting Eq. (21) from Eq. (11)

$$\begin{aligned} \varepsilon_{xq,k} &= \hat{x}_{xq,k} - E[\hat{x}_{xq,k}] = \hat{x}_{xq,k} - \bar{x}_{xq,SS} \\ &= \underbrace{A_{xq} \hat{x}_{xq,k-1} + K_{\infty} (y_k - C_{xq} A_{xq} \hat{x}_{xq,k-1})}_{\hat{x}_{xq,k}} \cdots \\ &\quad - \underbrace{(A_{xq} \bar{x}_{xq,SS} + K_{\infty} (y_{SS} - C_{xq} A_{xq} \bar{x}_{xq,SS}))}_{\bar{x}_{xq,SS}} \\ &= (A_{xq} - K_{\infty} C_{xq} A_{xq}) (\hat{x}_{xq,k-1} - \bar{x}_{xq,SS}) + K_{\infty} (y_k - y_{SS}) \end{aligned} \quad (28)$$

Making the substitutions $\varepsilon_{xq,k-1} = \hat{x}_{xq,k-1} - \bar{x}_{xq,SS}$, and $v_k = y_k - y_{SS}$ yields

$$\varepsilon_{xq,k} = (A_{xq} - K_{\infty} C_{xq} A_{xq}) \varepsilon_{xq,k-1} + K_{\infty} v_k \quad (29)$$

The estimation covariance matrix $P_{\hat{x}q,k}$ is then calculated as

$$\begin{aligned} P_{\hat{x}q,k} &= E[\varepsilon_{xq,k} \varepsilon_{xq,k}^T] \\ &= [A_{xq} - K_{\infty} C_{xq} A_{xq}] E[\varepsilon_{xq,k-1} \varepsilon_{xq,k-1}^T] \cdots \\ &\quad \times [A_{xq} - K_{\infty} C_{xq} A_{xq}]^T \cdots \\ &\quad + [A_{xq} - K_{\infty} C_{xq} A_{xq}] E[v_k v_k^T] K_{\infty}^T \cdots \\ &\quad + K_{\infty} E[v_k \varepsilon_{xq,k-1}^T] [A_{xq} - K_{\infty} C_{xq} A_{xq}]^T + K_{\infty} E[v_k v_k^T] K_{\infty}^T \end{aligned} \quad (30)$$

The substitutions $E[\varepsilon_{xq,k-1} \varepsilon_{xq,k-1}^T] = P_{\hat{x}q,k-1}$ and $E[v_k v_k^T] = R$ can be made in the above equation. Since $\varepsilon_{xq,k-1}$ and v_k are uncorrelated, the substitution $E[v_k \varepsilon_{xq,k-1}^T] = E[v_k \varepsilon_{xq,k-1}^T] = 0$ can also be made, producing

$$P_{\hat{x}q,k} = [A_{xq} - K_{\infty} C_{xq} A_{xq}] P_{\hat{x}q,k-1} [A_{xq} - K_{\infty} C_{xq} A_{xq}]^T + K_{\infty} R K_{\infty}^T \quad (31)$$

At steady-state operating conditions $P_{\hat{x}q,k-1} = P_{\hat{x}q,k}$. Making this substitution in Eq. (31) produces the following Riccati equation that can be solved for $P_{\hat{x}q,k}$:

$$P_{\hat{x}q,k} = [A_{xq} - K_{\infty} C_{xq} A_{xq}] P_{\hat{x}q,k} [A_{xq} - K_{\infty} C_{xq} A_{xq}]^T + K_{\infty} R K_{\infty}^T \quad (32)$$

It should be noted that $P_{\hat{x}q,k}$ obtained by solving Eq. (32) will be identical to P_{∞} produced via Eq. (9) if the system's actual state process noise covariance is identical to the Q_{xq} assumed in the design of the Kalman filter. However, Q is often treated as a Kalman filter design parameter to provide acceptable dynamic response. For the purpose of this derivation, we have assumed a steady-state operating condition, where the state variables and health parameters are invariant, and thus the actual system process noise is zero (i.e., $w_{xh,k} = 0$). In this case $P_{\hat{x}q,k}$ will not equal P_{∞} . Once $P_{\hat{x}q,k}$ is obtained, it can be used to calculate $P_{\hat{x}h,k}$, the covariance of $\hat{x}_{xh,k}$, which is defined as $E[(\hat{x}_{xh,k} - E[\hat{x}_{xh,k}])(\hat{x}_{xh,k} - E[\hat{x}_{xh,k}])^T]$

$$P_{\hat{x}h,k} = \begin{bmatrix} I & 0 \\ 0 & V^{*\dagger} \end{bmatrix} P_{\hat{x}q,k} \begin{bmatrix} I & 0 \\ 0 & V^{*\dagger} \end{bmatrix}^T \quad (33)$$

The augmented state vector estimation covariance given in Eq. (33) can be partitioned into covariance information for the original state vector, $P_{\hat{x},k}$ (upper left corner of the $P_{\hat{x}h,k}$ matrix), and the health parameter vector, $P_{\hat{h},k}$ (lower right corner of the $P_{\hat{x}h,k}$ matrix)

$$P_{\hat{x}h,k} = \begin{bmatrix} P_{\hat{x},k} & \cdots \\ \cdots & P_{\hat{h},k} \end{bmatrix} \quad (34)$$

The $P_{\hat{x}q,k}$ matrix from Eq. (32) can also be used to calculate $P_{\hat{z},k}$, the covariance in the estimation of z_k , which is equivalent to $E[(\hat{z}_k - E[\hat{z}_k])(\hat{z}_k - E[\hat{z}_k])^T]$

$$P_{\hat{z},k} = [F \quad NV^{*\dagger}] P_{\hat{x}q,k} [F \quad NV^{*\dagger}]^T \quad (35)$$

The variance in the estimates $\hat{x}_{xh,k}$ and \hat{z}_k can be obtained from the diagonals of the covariance matrices produced by Eqs. (33) and (35), respectively.

2.3.3 Sum of Squared Estimation Errors. Once Eqs. (25), (26), (33), and (35) are obtained, they may be used to analytically calculate the mean sum of squared estimation errors over all engines by combining the respective estimation error bias and estimation variance information, as previously shown in Eq. (16). The mean augmented state vector sum of squared estimation errors, $SSEE(\hat{x}_{xh, \text{fleet}})$, and the mean auxiliary parameter vector sum of squared estimation errors, $SSEE(\hat{z}_{\text{fleet}})$, become

$$SSEE(\hat{x}_{xh, \text{fleet}}) = \bar{x}_{xh, \text{fleet}}^2 + \text{tr}\{P_{\hat{x}h,k}\} = \text{tr}\{G_{xh} P_h G_{xh}^T + P_{\hat{x}h,k}\}$$

$$SSEE(\hat{z}_{\text{fleet}}) = \bar{z}_{\text{fleet}}^2 + \text{tr}\{P_{z,k}\} = \text{tr}\{G_z P_h G_z^T + P_{z,k}\} \quad (36)$$

If required, a weighted sum approach can be applied to normalize the contributions of individual auxiliary parameter estimation errors. This is often necessary as there may be several orders of magnitude difference between the auxiliary parameters of interest. A weighted sum approach prevents domination by individual parameters. In this study a diagonal auxiliary parameter weighting matrix, W_z , is applied based on the inverse of auxiliary parameter variance (obtained from the main diagonal of the auxiliary parameter covariance matrix, P_z)

$$P_z = [F(I-A)^{-1}L + N]P_h[F(I-A)^{-1}L + N]^T \quad (37)$$

$$W_z = \begin{bmatrix} P_{z,11} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & P_{z,ii} \end{bmatrix}^{-1}$$

W_z is then applied to calculate of a “weighted” sum of auxiliary parameter squared estimation errors given as

$$WSSEE(\hat{z}_{\text{fleet}}) = \text{tr}\{W_z[G_z P_h G_z^T + P_{z,k}]\} \quad (38)$$

From Eqs. (23), (24), (32), and (34) it can be observed that both bias and variance are affected by the selection of the transformation matrix, V^* . The sum of squared estimation error terms derived in this section give rise to an optimization problem: selecting V^* to minimize the squared estimation error in the Kalman filter produced parameter estimates. This could include health parameter estimates, auxiliary parameter estimates, or a combination of parameters. Although there is no known closed form solution for optimally selecting the V^* matrix to satisfy the objective of minimizing estimation errors, a multiparameter iterative search method has been developed to perform this task, and will be described in Sec. 2.4.

2.4 Optimal Transformation Matrix Selection. Prior to initiating the search for an optimal V^* , specific system design information must be defined or obtained. This includes:

- specifying the auxiliary parameters to be estimated
- generating system state-space equations at a fleet average (50% deteriorated) engine trim point
- defining measurement noise covariance matrix, R
- defining augmented state process noise covariance matrix, Q_{xh}
- defining fleet average health parameter covariance, P_h

Some additional clarification is provided regarding the selection of P_h and Q_{xh} as the distinction between these two covariance matrices may not be immediately obvious. P_h defines the expected health parameter covariance across all engines. It may be based on past knowledge gained from engine gas path analysis programs and/or historical studies of engine module performance deterioration. Conversely, Q_{xh} defines the expected process noise covariance in the state variables and health parameters of an individual engine, at a single discrete time step, k . The selection of Q_{xh} will directly impact the dynamic response and the variance of the estimates generated by the Kalman filter, and to a large extent Q_{xh} is treated as a design parameter.

After the necessary system information has been obtained, the search for an optimal transformation matrix to minimize the Kalman filter sum of squared estimation errors can commence. This is performed using the *lsqnonlin* function of the MATLAB® Optimization Toolbox. This function applies an iterative search to find the least-squares solution of a user-specified multivariable optimization problem. A flow chart depicting the steps in this optimal iterative search is shown in Fig. 1, and a further description of each step is given below.

1. Upon startup, an initial random guess of V^* is generated. It is

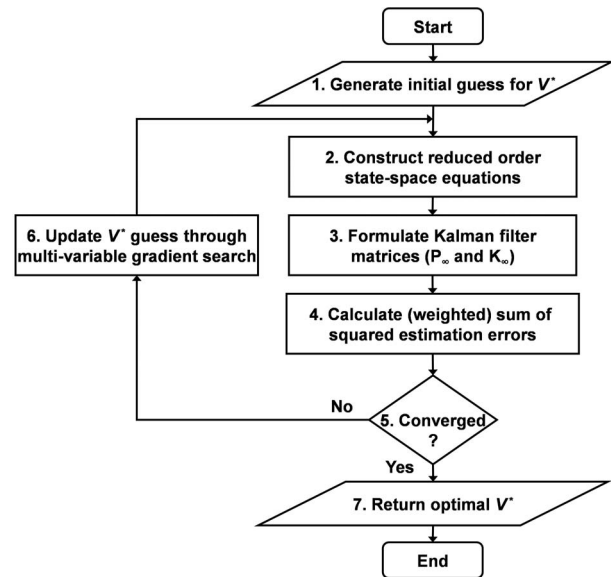


Fig. 1 Flowchart of V^* iterative optimal search

selected such that the matrix Frobenius norm $\|V^*\|_F = 1$. This requirement is applied to help prevent convergence to a poorly scaled solution.

2. Construct the reduced-order state-space model, (Eq. (7)).
3. Formulate the Kalman filter.
 - a. Calculate the estimation covariance matrix, P_∞ (Eq. (9)).
 - b. Calculate the Kalman gain matrix, K_∞ , (Eq. (10)).
4. Calculate sum of squared estimation errors (Eq. (36)), or weighted sum of squared estimation errors (Eq. (38)).
5. On each iteration the change in $SSEE$ (or $WSSEE$) relative to the previous iteration is assessed to determine if convergence within a user-specified tolerance has been achieved.
 - a. If converged, skip step 6 and proceed to step 7.
 - b. If not converged, proceed to step 6 to update V^* .
6. V^* is updated via the MATLAB® *lsqnonlin* function, again requiring that $\|V^*\|_F = 1$, and the process returns to step 2.
7. Upon convergence, the optimization routine returns the optimal value of V^* , and ends.

Experience has shown that the transformation matrix returned by the optimization routine is not unique—different matrices can be found that produce a global minimum of the objective function. Experience has also shown that the optimization routine will usually return a V^* matrix that satisfies, or nearly satisfies (i.e., within 5%), the global minimum of the objective function. However, in order to guard against potential convergence to a local minimum, it is prudent to run the optimization routine multiple times, each time starting with a different initial guess for V^* . This is only to assure the designer that the global minimum is achieved, not to produce a consistent V^* . It should be emphasized that the optimal search for V^* is only conducted off-line during the estimator design process. This calculation is not conducted as part of the on-line real-time Kalman filter implementation, and thus places no additional computational burden upon it.

3 Turbofan Engine Example

A linearized cruise operating point extracted from the NASA commercial modular aeropropulsion system simulation (C-MAPSS) high-bypass turbofan engine model [9] is used to evaluate the new systematic tuner selection methodology. The linear

Table 1 State variables, health parameters, and actuators

State variables	Health parameters	Actuators
Nf–fan speed	Fan efficiency	Wf – fuel flow
Nc–core speed	Fan flow capacity ^a	Variable stator vane (VSV)
	LPC efficiency ^a	Variable bleed valve (VBV)
	LPC flow capacity	
	High pressure compressor (HPC) efficiency ^a	
	HPC flow capacity ^a	
	High pressure turbine (HPT) efficiency ^a	
	HPT flow capacity ^a	
	LPT efficiency ^a	
	LPT flow capacity ^a	

^aHealth parameters selected as tuners in conventional estimation approach, (identical to subset of health parameters selected as tuners in Ref. [4]).

model has two state variables, ten health parameters, and three control inputs, all shown in Table 1. The model's seven sensed outputs, and corresponding sensor noise standard deviation, are shown in Table 2. The auxiliary output parameters of interest to be estimated are shown in Table 3. The linear model is used as the truth model for this application example. The model is run open-loop, so all control inputs remain at 0, i.e., they do not deviate from the trim value for the linear model and no actuator bias is present. Deviations in all ten health parameters are assumed to be uncorrelated, and randomly shifted from their trim conditions with a standard deviation of ± 0.02 ($\pm 2\%$). Since a parameter's vari-

ance is equal to its standard deviation squared, the health parameter covariance matrix, P_h , is defined as a diagonal matrix with all diagonal elements equal to 0.0004.

Next, the estimation accuracy of the systematic approach for selecting Kalman filter tuning parameters will be compared with the conventional approach of selecting a subset of health parameters to serve as tuners (the seven health parameters denoted with "a" in Table 1), and the singular value decomposition approach to tuner selection introduced by Litt in Ref. [4]. Table 4 shows a comparison of the theoretically predicted estimation errors (squared bias, variance, and total squared error) and experimentally obtained squared estimation errors for each of the three tuner selection approaches. T40 and T50 estimation errors are shown in squared degrees Rankine, and Fn and SmLPC estimation errors are shown in squared percent net thrust and squared percent stall margin, respectively. The experimental results were obtained through a Monte Carlo simulation analysis, where the health parameters varied over a random distribution in accordance with the covariance matrix, P_h . The test cases were concatenated to produce a single time history input, which was provided to the C-MAPSS linear discrete state-space model given in Eq. (2), with an update rate of 15 ms. Each individual health parameter test case lasted 30 s. At the completion of each 30 s test case, the health parameter vector input instantaneously transitioned to the next test case. A total of 375 30 s test cases were evaluated, resulting in an 11,250 s input time history. Three separate Kalman filters were implemented using the three tuner selection approaches. The experimental estimation errors were determined by calculating the mean-squared error between estimated and actual values during the last 10 s of each 30 s test case. The error calculation is based on only the last 10 s so that engine model outputs and Kalman estimator outputs have reached a quasisteady-state operating condition prior to calculating the error. This

Table 2 Sensed outputs and standard deviation as percent of operating point trim values

Sensed output	Standard deviation (%)
Nf–fan speed	0.25
Nc–core speed	0.25
P24–HPC inlet total pressure	0.50
T24–HPC inlet total temperature	0.75
Ps30–HPC exit static pressure	0.50
T30–HPC exit total temperature	0.75
T48–exhaust gas temperature	0.75

Table 3 Estimated auxiliary parameters

Auxiliary parameter
T40–combustor exit temperature
T50–LPT exit temperature
FN–net thrust
SmLPC – LPC stall margin

Table 4 Auxiliary parameter squared estimation errors

Tuners	Error	T40 (°R)	T50 (°R)	Fn (%)	SmLPC (%)
Subset of health parameters	Theor. sqr. bias	0.00	561.76	3.84	3.28
	Theor. variance	74.76	29.65	0.48	0.34
	Theor. sqr. error	74.76	591.41	4.31	3.62
	Exper. sqr. error	74.90	583.29	4.27	3.60
SVD tuner selection	Theor. sqr. bias	0.00	512.46	4.05	5.28
	Theor. variance	65.99	67.21	0.80	1.31
	Theor. sqr. error	65.99	579.67	4.86	6.59
	Exper. sqr. error	66.20	579.39	4.98	6.76
Systematic tuner selection	Theor. sqr. bias	0.00	87.81	0.66	0.95
	Theor. variance	17.49	18.55	0.13	0.35
	Theor. sqr. error	17.49	106.35	0.79	1.30
	Exper. sqr. error	17.61	106.54	0.86	1.35

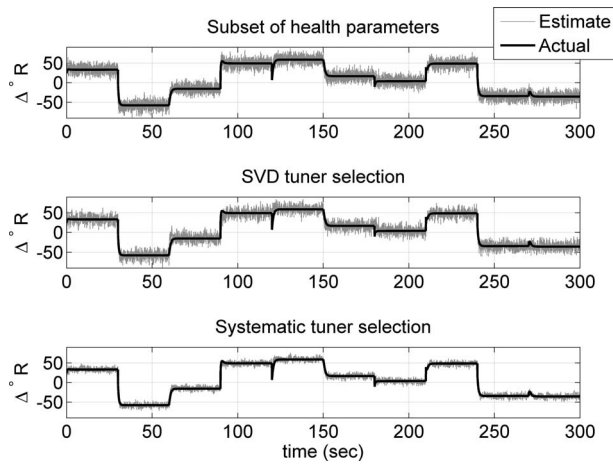


Fig. 2 T40 estimation (tuner comparison)

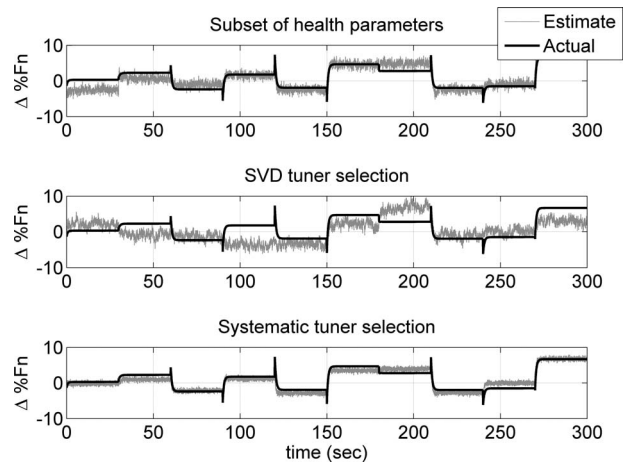


Fig. 4 Fn estimation (tuner comparison)

ensures that the experimental results are consistent with the theoretically predicted estimation errors that were derived assuming steady-state operation.

From Table 4 it can be seen that the theoretically predicted and the experimentally obtained squared estimation errors exhibit good agreement. If the number of random test cases were increased to a suitably large number, it is expected that the theoretical and experimental results would be identical. It can also be seen that all three estimators are able to produce unbiased estimates of the combustor exit temperature, T40; however, their estimates of low pressure turbine (LPT) exit temperature, T50, net thrust, Fn, and low pressure compressor (LPC) stall margin, SmLPC, are biased. The encouraging finding is that the new systematic approach to tuner selection significantly reduces the overall mean-squared estimation error compared with the other two approaches. Relative to the conventional approach of tuner selection the experimental mean-squared estimation errors in T40, T50, Fn and SmLPC are reduced to 76%, 82%, 80%, and 63%, respectively. It can also be observed that the SVD tuner selection approach, which is designed to reduce the estimation error bias, does in fact reduce the sum of squared biases relative to the subset of health parameters approach. However, the SVD approach is also found to increase the estimation variance, which contributes to its overall mean-squared estimation error.

A visual illustration of the effect that tuner selection has on Kalman filter estimation accuracy can be seen in Figs. 2–5, which show actual and estimated results for the auxiliary parameters

T40, T50, Fn, and SmLPC respectively. Each plot shows a 300 s segment of the evaluated test cases. The step changes that can be observed in each plot every 30 s correspond to a transition to a different health parameter vector. True model auxiliary parameter outputs are shown in black, and Kalman filter estimates are shown in gray. In each figure the information is arranged top to bottom according to tuner selection based upon (a) a subset of health parameters, (b) singular value decomposition, and (c) the new systematic selection strategy. The information shown in these figures corroborates the information in Table 4; namely, all three tuner selection approaches produce unbiased estimates of T40 (Fig. 2), while the systematic tuner selection strategy yields a noticeable reduction in the total squared estimation error (squared bias plus variance) of all four auxiliary parameters.

3.1 Comparison With Maximum A Posteriori Estimation.

The presented systematic tuner selection strategy minimizes the mean-squared error of the on-line estimator at steady-state operating conditions, taking advantage of prior knowledge of engine health parameter distributions. As such it is somewhat analogous to the maximum a posteriori estimation method commonly applied for ground-based aircraft gas turbine engine gas path analysis [5,6]. This leads to the question, how does the on-line Kalman filter estimation accuracy compare with MAP estimation accuracy? Prior to making this comparison the mathematical formulation of the MAP estimator is briefly introduced. Here a steady-state model of the measurement process in the following form is applied

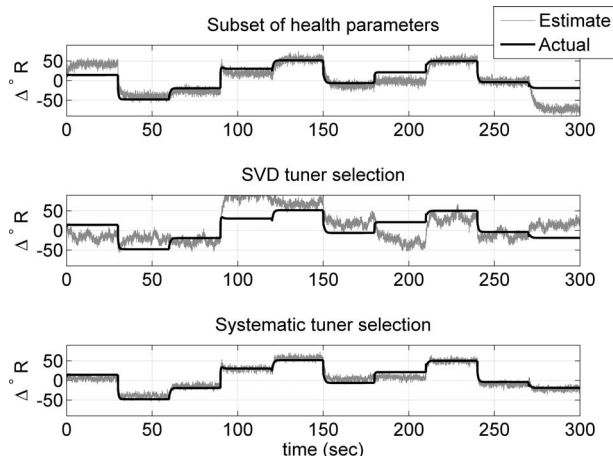


Fig. 3 T50 estimation (tuner comparison)

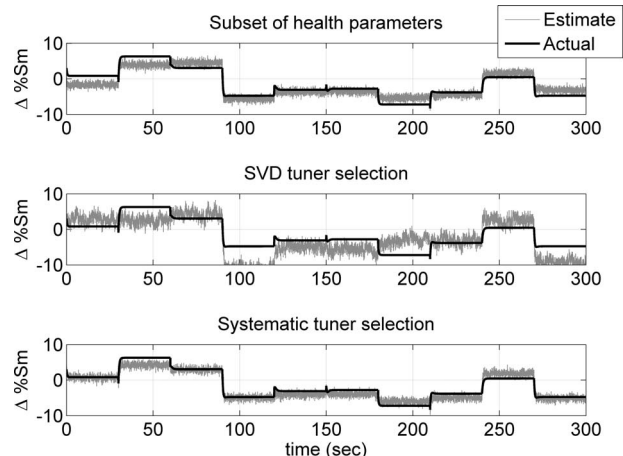


Fig. 5 SmLPC estimation (tuner comparison)

Table 5 Health parameter % squared estimation errors (nominal noise)

Estimator	Error type	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}	Sum
Kalman filter	Theor. squared bias	1.26	1.48	2.28	1.22	0.00	0.00	0.74	0.00	1.97	3.06	12.00
	Theor. variance	0.73	0.19	0.80	0.45	0.40	0.68	0.21	0.21	0.34	0.08	4.09
	Theor. squared error	2.00	1.67	3.07	1.66	0.40	0.68	0.96	0.21	2.32	3.14	16.09
	Exper. squared error	1.82	1.60	2.77	1.59	0.41	0.68	0.97	0.21	2.13	3.15	15.31
MAP estimator	Theor. squared bias	2.42	1.75	3.53	1.90	0.47	0.96	1.00	0.16	2.47	3.17	17.81
	Theor. variance	0.22	0.27	0.16	0.64	0.69	0.71	0.36	0.36	0.17	0.09	3.68
	Theor. squared error	2.63	2.02	3.69	2.54	1.16	1.67	1.36	0.52	2.64	3.26	21.48
	Exper. squared error	2.52	1.89	3.44	2.45	1.14	1.57	1.34	0.51	2.49	3.24	20.60

$$y_k = Hh_k + v_k \tag{39}$$

where the matrix H relates the effects of the health parameter vector, h , to the sensed measurements, y . From Eq. (19), it can be seen that H is equivalent to $C(I-A)^{-1}L+M$. The MAP estimator follows the closed form expression

$$\hat{h}_k = (P_h^{-1} + H^T R^{-1} H) H^T R^{-1} y_k \tag{40}$$

The MAP estimator is capable of estimating more unknowns than available measurements due to the inclusion of a priori knowledge of the estimated parameter covariance, P_h . However, the MAP estimator, unlike a Kalman filter, is not a recursive estimator and does not take advantage of past measurements to enhance its estimate at the current time step. Furthermore, the MAP estimator only considers a static relationship between system state variables and measured outputs—it does not consider system dynamics. Because of these differences, a Kalman estimator with optimally selected tuning parameters should outperform the MAP estimator. However, under steady-state conditions, with minimal sensor noise the two estimation approaches should produce similar results. To test this theory, a MAP estimator was designed, and its estimation accuracy was compared with a Kalman filter with tuning parameters optimally selected to minimize the estimation errors in the health parameter vector h . First, the two estimators were designed and evaluated using the original sensor noise levels shown in Table 2. Next, the sensor noise levels were set to 1/20th of their original levels, the estimators were redesigned, and the comparison was repeated. Monte Carlo simulation evaluations, as previously described, were applied (i.e., 375 random health parameter vectors, 30 s in duration, with estimation accuracy calculations based on the last 10 s of each 30 s test case). Theoretical and experimental estimation errors are shown in Tables 5 and 6 for the original noise and reduced noise levels, respectively. At original noise levels the Kalman estimator is able to produce smaller estimation errors. However, at the reduced noise level the two estimation approaches are found to be nearly identical. This comparison validates that the Kalman estimation approach is indeed producing a minimum mean-squared estimation error as intended, while providing the capability to support real-time on-line estimation under dynamic operating scenarios.

4 Discussion

While the systematic tuner selection approach presented here appears promising for on-line Kalman filter-based parameter estimation applications, there are several practical considerations that need to be assessed when applying such a technique. The optimization routine attempts to minimize the overall squared estimation error—both bias and variance—under steady-state operating conditions. The minimization of the estimation variance in particular can come at the expense of dynamic responsiveness of the Kalman filter. To illustrate this consider the time history plots of actual versus estimated T40 shown in Fig. 6. The top plot shows Kalman filter estimation results using a tuning parameter vector systematically selected to minimize the error in four auxiliary parameters (T40, T50, Fn, and SmLPC), as presented in Sec. 3. The bottom plot shows Kalman filter estimation results using a tuning parameter vector systematically selected to minimize the estimation error in T40 only. At time 100 s a step change in the health parameter input vector is introduced into the engine model; this allows the dynamic response of the two estimators to be compared. It can be observed that T40 estimation variance in the bottom plot is reduced, as is the mean steady-state estimation error (>300 s). This is not surprising since one would generally expect improved results when optimizing to minimize the error in a single parameter, as opposed to multiple parameters. However, the estimator shown in the bottom plot does require a significantly longer time to reach steady-state convergence. Conversely, the estimator designed to minimize the steady-state error in four auxiliary parameters (top plot) is unable to place as much emphasis on T40 estimation variance reduction, but it is able to track dynamic changes in T40 more rapidly. This example illustrates the interdependence between estimation variance and responsiveness. Therefore, it is prudent for a designer to evaluate the Kalman filter to ensure that it tracks engine dynamics acceptably. If the dynamic response is unacceptable, the optimization routine can be rerun placing more weight on estimation error bias reduction, and less weight on variance reduction.

All results presented in this paper are based on a turbofan engine linear state-space model at a single operating point. While this linear assessment enables experimental validation of theoret-

Table 6 Health parameter % squared estimation errors (reduced noise)

Estimator	Error type	h_1	h_2	h_3	h_4	h_5	h_6	h_7	h_8	h_9	h_{10}	Sum
Kalman filter	Theor. squared bias	1.26	1.48	2.28	1.22	0.00	0.00	0.74	0.00	1.97	3.06	12.00
	Theor. variance	0.02	0.00	0.03	0.01	0.01	0.02	0.00	0.00	0.01	0.00	0.10
	Theor. squared error	1.29	1.48	2.30	1.22	0.01	0.02	0.74	0.00	1.98	3.06	12.10
	Exper. squared error	1.13	1.41	1.97	1.16	0.01	0.02	0.75	0.00	1.79	3.08	11.31
MAP estimator	Theor. squared bias	1.26	1.48	2.28	1.22	0.00	0.00	0.74	0.00	1.97	3.06	12.00
	Theor. variance	0.03	0.00	0.04	0.01	0.01	0.02	0.00	0.00	0.01	0.00	0.13
	Theor. squared error	1.30	1.48	2.32	1.22	0.01	0.02	0.74	0.00	1.99	3.06	12.14
	Exper. squared error	1.14	1.41	1.98	1.17	0.01	0.02	0.75	0.00	1.79	3.08	11.34

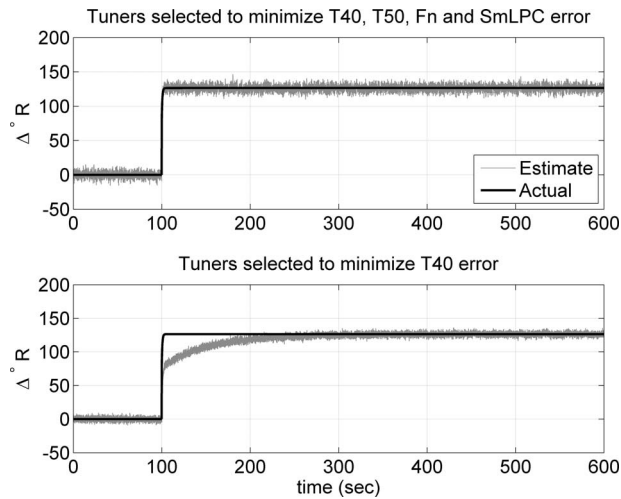


Fig. 6 Illustration of tuner impact on estimator response

ical predictions, it is not representative of actual aircraft gas turbine engine operation, which is a transient nonlinear system operating over a broad range of operating conditions. In order for the presented approach to be applicable it would need to be able to produce an optimal set of tuning parameters, not just at a single operating point, but rather a globally optimal tuning parameter vector universally applicable over the range of operating conditions that an engine is expected to experience. A potential approach to selecting a single “globally optimal” tuning parameter vector is to modify the optimization routine to minimize the combined estimation error over multiple engine operating points such as takeoff, climb, and cruise. This would be a straightforward modification to the MATLAB® optimization routine, but it would increase the computational time required to calculate the result. Since the systematic tuner selection process is only envisioned to be done once during the system design process, this will not impact the on-line execution speed of the Kalman filter. It is anticipated that the application of globally optimal tuners will result in some estimation accuracy degradation relative to tuners optimized for individual operating points, although this has not yet been verified or quantified.

5 Conclusions

A systematic approach to tuning parameter selection for on-line Kalman filter-based parameter estimation has been presented. This technique is specifically applicable for the underdetermined aircraft engine parameter estimation case where there are fewer sensor measurements than unknown health parameters which will impact engine outputs. It creates and applies a linear transformation matrix, V^* , to select a vector of tuning parameters that are a linear combination of all health parameters. The tuning parameter vector is selected to be of low-enough dimension to be estimated, while minimizing the mean-squared error of Kalman filter estimates. The multiparameter iterative search routine applied to optimally select V^* was presented. Results have shown that while the transformation matrix returned by the optimization routine is not unique (different matrices can be found that produce a global minimum of the objective function), the routine is effective in returning a transformation matrix that is optimal, or near optimal, regardless of its initial starting guess of the matrix. The efficacy of the systematic approach to tuning parameter selection was demonstrated by applying it to parameter estimation in an aircraft turbofan engine linear point model. It was found to significantly reduce mean-squared estimation errors compared with the conventional approach of selecting a subset of health parameters to serve as tuners. In some parameters the mean-squared estimation

error reduction was found to be over 80%. These estimation improvements were theoretically predicted and experimentally validated through Monte Carlo simulation studies.

The systematic approach to Kalman filter design is envisioned to be applicable for a broad range of on-board aircraft engine model-based applications that produce estimates of unmeasured parameters. This includes model-based controls, model-based diagnostics, and on-board life usage algorithms. It is also envisioned to have benefits for sensor selection during the engine design process, specifically for assessing the performance estimation accuracy benefits of different candidate sensor suites. Areas for future work include extending the technique to produce a tuning parameter vector optimal over a range of operating conditions, and evaluating the technique on a nonlinear engine model, under both steady-state and transient operating conditions.

Acknowledgment

This research was conducted under the NASA Aviation Safety Program, Integrated Vehicle Health Management Project. The authors graciously acknowledge Jonathan Litt for his support and insightful discussions that led to the development of this work.

Nomenclature

$A, A_{xh}, A_{xq}, B, B_{xh}, B_{xq}, C, C_{xh}, C_{xq}, D, F, F_{xh}, F_{xq}, G, L, M, \text{ and } N$	= system matrices
$G_{xh}, G_x, G_h, \text{ and } G_z$	= estimation bias matrices
H	= matrix that relates health parameter effects to steady-state engine outputs
I	= identity matrix
K_∞	= Kalman filter gain
$P_h \text{ and } P_z$	= health and auxiliary parameter covariance matrices
$P_{\hat{x}_{h,k}}, P_{\hat{x}_{q,k}}, \text{ and } P_{\hat{z},k}$	= covariance matrices of estimated parameters
P_∞	= Kalman filter state estimation covariance matrix
$Q, Q_{xh}, \text{ and } Q_{xq}$	= process noise covariance matrices
R	= measurement noise covariance matrix
V^*	= transformation matrix relating h_k to q_k
W_z	= auxiliary parameter weighting matrix
h_k	= health parameter vector
q_k	= Kalman filter tuning parameter vector
u_k	= actuator command vector
v_k	= measurement noise vector
$w_k, w_{h,k}, \text{ and } w_{xh,k}$	= process noise vectors
x_k	= state vector
$x_{xh,k}$	= augmented state vector (x_k and h_k)
$x_{xq,k}$	= reduced-order state vector (x_k and q_k)
y_k	= vector of measured outputs
z_k	= vector of unmeasured (auxiliary) outputs
$\epsilon_{xq,k}$	= residual vector (estimate minus its expected value)

Subscripts

k	= discrete time step index
xh	= augmented state vector (x and h)
xq	= reduced-order state vector (x and q)
SS	= steady-state value

Superscripts

\dagger	= pseudo-inverse
$\hat{}$	= estimated value
\sim	= error value
$\bar{}$	= mean value
T	= transpose

Operators

- $E[\cdot]$ = expected value of argument
 $\text{tr}\{\cdot\}$ = trace of matrix
 $SSEE(\cdot)$ = sum of squared estimation errors
 $WSSEE(\cdot)$ = weighted sum of squared estimation errors
 $\|\cdot\|_F$ = matrix Frobenius norm

References

- [1] Luppold, R. H., Roman, J. R., Gallops, G. W., and Kerr, L. J., 1989, "Estimating In-Flight Engine Performance Variations Using Kalman Filter Concepts," AIAA 25th Joint Propulsion Conference, Paper No. AIAA-89-2584.
- [2] Volponi, A., 2008, "Enhanced Self-Tuning On-Board Real-Time Model (eSTORM) for Aircraft Engine Performance Health Tracking," NASA Report No. CR-2008-215272.
- [3] Kumar, A., Viassolo, D., (2008), "Model-Based Fault Tolerant Control," NASA Report No. CR-2008-215273.
- [4] Litt, J. S., 2008, "An Optimal Orthogonal Decomposition Method for Kalman Filter-Based Turbofan Engine Thrust Estimation," ASME J. Eng. Gas Turbines Power, **130**, p. 011601.
- [5] Doel, D. L., 1994, "An Assessment of Weighted-Least-Squares-Based Gas Path Analysis," ASME J. Eng. Gas Turbines Power, **116**, pp. 336-373.
- [6] Volponi, A. J., 2003, "Foundation of Gas Path Analysis (Part I and II)," *Gas Turbine Condition Monitoring and Fault Diagnosis* (von Karman Lecture Series No. 2003-01), von Karman Institute for Fluid Dynamics.
- [7] España, M. D., 1994, "Sensor Biases Effect on the Estimation Algorithm for Performance-Seeking Controllers," J. Propul. Power, **10**, pp. 527-532.
- [8] Simon, D., 2006, *Optimal State Estimation, Kalman, H_∞, and Nonlinear Approaches*, Wiley, Hoboken, NJ.
- [9] Frederick, D. K., DeCastro, J. A., and Litt, J. S., 2007, "User's Guide for the Commercial Modular Aero-Propulsion System Simulation (C-MAPSS)," NASA Technical Memorandum No. TM-2007-215026.

Application of Bayesian Forecasting to Change Detection and Prognosis of Gas Turbine Performance

Holger Lipowsky

e-mail: lipowsky@ila.uni-stuttgart.de

Stephan Staudacher

e-mail: staudacher@ila.uni-stuttgart.de

Institute of Aircraft Propulsion Systems (ILA),
University of Stuttgart,
Pfaffenwaldring 6,
70569 Stuttgart, Germany

Michael Bauer

e-mail: michael.bauer@mtu.de

Klaus-Juergen Schmidt

e-mail: klaus-juergen.schmidt@mtu.de

Department of Performance,
TEAP,
MTU Aero Engines GmbH,
Dachauer Strasse 665,
80995 München, Germany

The performance of gas turbines degrades over time due to deterioration mechanisms and single fault events. While deterioration mechanisms occur gradually, single fault events are characterized by occurring accidentally. In the case of single events, abrupt changes in the engine parameters are expected. Identifying these changes as soon as possible is referred to as detection. State-of-the-art detection algorithms are based on expert systems, neural networks, special filters, or fuzzy logic. This paper presents a novel detection technique, which is based on Bayesian forecasting and dynamic linear models (DLMs). Bayesian forecasting enables the calculation of conditional probabilities, whereas DLMs are a mathematical tool for time series analysis. The combination of the two methods can be used to calculate probability density functions prior to the next observation, or the so called forecast distributions. The change detection is carried out by comparing the current model with an alternative model, where the mean value is shifted by a prescribed offset. If the forecast distribution of the alternative model better fits the actual observation, a potential change is detected. To determine whether the respective observation is a single outlier or the first observation of a significant change, a special logic is developed. In addition to change detection, the proposed technique has the ability to perform a prognosis of measurement values. The developed method was run through an extensive test program. Detection rates >92% have been achieved for changed heights, as small as 1.5 times the standard deviation of the observed signal (sigma). For changed heights greater than 2 sigma, the detection rates have proven to be 100%. It could also be shown that a high detection rate is gained by a high false detection rate (~2%). An optimum must be chosen between a high detection rate and a low false detection rate, by choosing an appropriate uncertainty limit for the detection. Increasing the uncertainty limit decreases both detection rate and false detection rate. In terms of prognostic abilities, the proposed technique not only estimates the point of time of a potential limit exceedance of respective parameters, but also calculates confidence bounds, as well as probability density and cumulative distribution functions for the prognosis. The conflictive requirements of a high degree of smoothing and a quick reaction to changes are fulfilled in parallel by combining two different detection conditions.

[DOI: 10.1115/1.3159367]

Keywords: engine diagnostics, failure detection, gas turbine performance

1 Introduction

The performance of gas turbines degrades over time due to deterioration mechanisms and single fault events. While deterioration mechanisms affect all components of a gas turbine, single fault events are characterized by influencing only some components and by occurring accidentally. For gas turbine monitoring, performance models are used to predict the performance of an undeteriorated engine at the desired operating point. Thus, it is possible to compare measurements with their respective predicted values. The differences between measured and predicted values (residuals) form the basis of gas turbine monitoring [1]. In case of single events, abrupt changes in the residuals are expected. Identifying these changes as soon as possible is referred to as *detection*. State-of-the-art detection algorithms are based on expert sys-

tems [2], neural networks [3], special filters, or fuzzy logic [4].

If a change has been detected, the gas turbine's components that are affected by the fault can be identified, and their deviation from nominal state can be assessed. These tasks are referred to as *identification* and *diagnosis*. Modern systems apply techniques like neural networks [5], fuzzy logic [6,7], nonlinear optimization [8,9], genetic algorithms [10], pattern recognition [11,12], or hybrid approaches [13–15].

If no abrupt change has been detected, the rate of gradual change can be assessed in order to estimate when predefined limits of particular parameters will be exceeded, which is referred to as *prognosis*.

The top-level flow chart of the overall monitoring process is shown in Fig. 1. The diagnosis is based on in-service measurements of the gas turbine, e.g., speeds, temperatures, and pressures. The measurements are prechecked for sensor failures by applying a filter logic; the details of which can be found in Ref. [16]. If not filtered out, the measurements are compared with predicted values (process 1). Based on this comparison, some detection logic is needed to decide whether the degradation is gradual or caused by a single fault event. Process 2 contains the diagnosis of gradual

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 22, 2009; final manuscript received March 23, 2009; published online December 3, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Gas Turbine Technical Congress and Exposition, Orlando, FL, June 8–12, 2009.

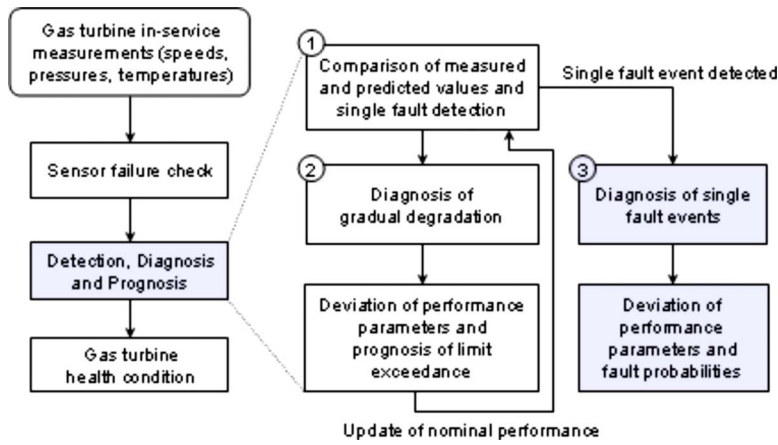


Fig. 1 Health monitoring process [15]

degradation, and yields the deviation of performance parameters as a result. Moreover, it is possible to predict limit exceedances of performance parameter deviations, which can be helpful in terms of maintenance planning. The derived performance parameters are fed back to process 1 as an update of the nominal performance for the next diagnosis. Once a single fault event has been detected, its diagnosis is performed in process 3. The details of the single fault event diagnosis were presented by Lipowsky et al. [15]. One of the key findings of their work was the importance of the correct detection, with respect to the diagnostic result. Simply spoken, a correct detection is absolutely crucial for a correct diagnosis. Hence, a lot of effort has been put into the development of a detection algorithm.

This paper discusses the development of a detection algorithm that is deemed to be new in the field of gas turbine monitoring. The method is based on the application of DLMs and Bayesian forecasting, and can be used for the offline monitoring of measurements collected periodically, e.g., once per flight. The theoretical background of the method is given in Sec. 2. The approach incorporates logics for outlier detection, change detection, and prognosis, which are described in Sec. 3. The developed method is evaluated in terms of its key abilities in a statistical manner. The results of this evaluation are shown in Sec. 4. Finally, the conclusions are drawn in Sec. 5.

2 Theoretical Background

2.1 DLMs. Dynamic linear models (DLMs) form the basis of many time series analysis methods [17]. The idea of describing a time series with a DLM is to break the observed signal down into its mathematical elements, e.g., value, gradient, curvature, etc. Figure 2 shows the modeling of a process with a second order DLM, i.e., using the elements such as *value* and *gradient*.

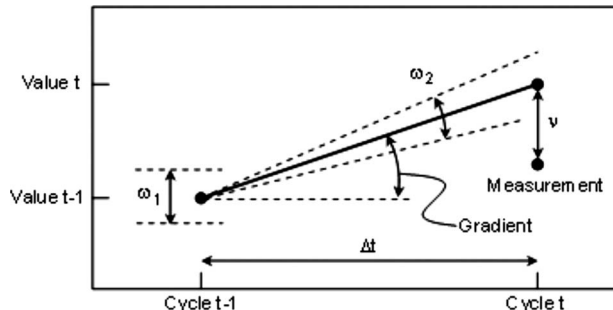


Fig. 2 Second order DLM process modeling

Based on the information available at time $t-1$, the guess values for the parameter's value and gradient, as well as the guess value for the observed measurement at time t , can be calculated by the equations

$$\text{value}_t = \text{value}_{t-1} + \Delta t \text{ grad}_{t-1} + \omega_{1,t-1} \quad (1)$$

$$\text{grad}_t = \text{grad}_{t-1} + \omega_{2,t-1} \quad (2)$$

$$\text{measurement}_t = \text{value}_t + \nu_t \quad (3)$$

where ω_1 , ω_2 , and ν are the variances in value, gradient, and measurement, respectively. For the sake of simplicity, it is assumed that observations are available in an equidistant sequence of one cycle ($\Delta t=1$). In that case, the general DLM matrix equations

$$\Theta_t = G_t \cdot \Theta_{t-1} + \omega_t \quad \text{with} \quad \omega_t \sim N(0, W_t) \quad (4)$$

$$Y_t = F_t^T \cdot \Theta_t + \nu_t \quad \text{with} \quad \nu_t \sim N(0, V_t) \quad (5)$$

apply with the following matrices [17]:

$$\Theta = \begin{pmatrix} \mu \\ \beta \end{pmatrix}, \quad F = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad G = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}, \quad W = \begin{pmatrix} \omega_1 & 0 \\ 0 & \omega_2 \end{pmatrix}, \quad V = \sigma^2 \quad (6)$$

The guess value of the parameter value is denoted by μ , and the guess value of the parameter gradient is denoted by β . The measurement variance is given by the square of its standard deviation σ .

2.2 Bayesian Forecasting. The method of Bayesian forecasting is based on the description of the observed process, by the use of probability density functions (PDFs). The following notation is used:

$$(\Theta_t | D_t) \sim N(m_t, C_t) \quad \text{with} \quad D_t = Y_1, Y_2, \dots, Y_t \quad (7)$$

Equation (7) denotes that the PDF of the parameter Θ at time t , based on the measurements D , is normally distributed with a mean of m and a variance of C . The approach of the Bayesian forecasting is similar to the Kalman filter applied to the time series [18], with respect to the fact that both methods are comprised of predictor and corrector elements. That means that the value of the next measurement is predicted and subsequently corrected when the measurement is available. However, Bayesian forecasting extends the approach by the use of PDFs. As the name implies, the method of the Bayesian forecasting is based on the application of Bayes' theorem, which enables the calculation of conditional probabilities and reads as follows:

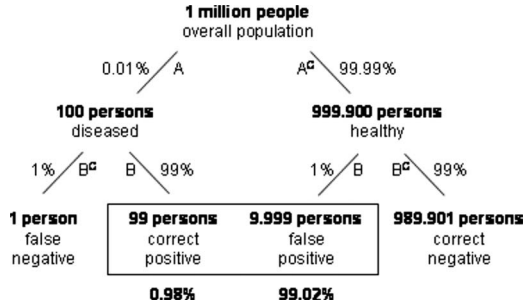


Fig. 3 Decision tree explanation of Bayes' theorem

$$p(A|B) = \frac{p(B|A) \cdot p(A)}{p(B)} \quad (8)$$

The notation $p(A|B)$ denotes the probability of event A , occurring conditional to the probability of event B having occurred. An illustrative example of Bayes' theorem states as follows: Let us assume that the probability to suffer from a certain disease be $p(A)=0.01\%$ (A denotes the event of actually suffering from the disease). Moreover, let us suppose that a clinical test exists, that is able to detect the disease with a certainty of $p(B|A)=99\%$ (B denotes the event of a positive test result). Surprisingly, the probability of actually suffering from the disease, given a positive test result, is only $p(A|B)=0.98\%$. Hence, the probability of being healthy, despite a positive test result, is $p(A^C|B)=99.02\%$ (A^C denotes the inversion of event A , i.e., being healthy). This result seems stunning, but it is due to the fact that the probability of actually suffering from the disease $p(A)=0.01\%$ is a hundred times smaller than the probability of a false test result $p(B^C|A) + p(B|A^C)=1\%$. The result can be visualized using a decision tree (see Fig. 3). To simplify matters, the calculation is based on a population of one million people. The probability of actually suffering from the disease, given a positive test result $p(A|B)=0.98\%$, represents the ratio of correct positive detections and overall positive detections.

The Bayesian forecasting equations are listed hereinafter. Their derivation is beyond the scope of this paper and can be found in Ref. [19]. Given the state at time $t-1$

$$(\Theta_{t-1}|D_{t-1}) \sim N(m_{t-1}, C_{t-1}) \quad (9)$$

the predictions for time t are given by the one-step forecast equations

$$(\Theta_t|D_{t-1}) \sim N(a_t, R_t) \quad (10)$$

$$(Y_t|D_{t-1}) \sim N(f_t, Q_t) \quad (11)$$

where means and variances can be calculated as follows:

$$a_t = G_t \cdot m_{t-1}, \quad f_t = F_t^T \cdot a_t \quad (12)$$

$$R_t = G_t C_{t-1} G_t^T + W_t, \quad Q_t = F_t^T R_t F_t + V_t \quad (13)$$

Taking into account the measurement at time t , the corrections are given by the update equations

$$(\Theta_t|D_t) \sim N(m_t, C_t) \quad (14)$$

where means and variances are given by

$$m_t = a_t + A_t e_t \quad \text{with} \quad A_t = R_t F_t / Q_t, \quad e_t = Y_t - f_t \quad (15)$$

$$C_t = R_t - A_t A_t^T Q_t \quad (16)$$

Finally, if a forecast of the next k time steps is of interest, the k -step forecast equations have to be applied

$$(\Theta_{t+k}|D_t) \sim N(a_t(k), R_t(k)) \quad (17)$$

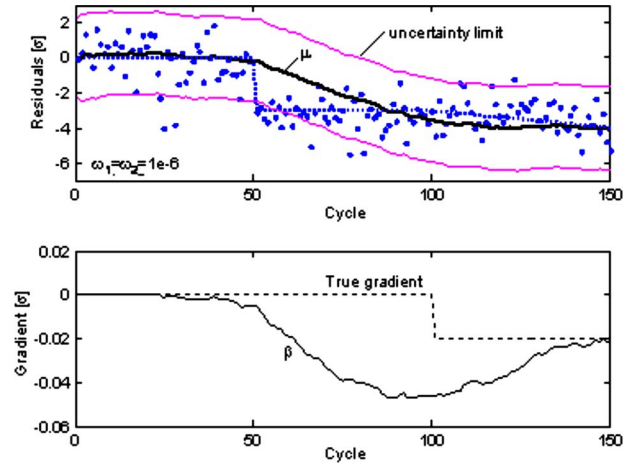


Fig. 4 Application to residuals with an implanted change height of $\Delta\mu = -3\sigma$ at cycle $t=51$

$$(Y_{t+k}|D_t) \sim N(f_t(k), Q_t(k)) \quad (18)$$

with the parameters

$$a_t(k) = G_{t+k} \cdot a_t(k-1), \quad f_t(k) = F_{t+k}^T \cdot a_t(k) \quad (19)$$

$$R_t(k) = G_{t+k} R_t(k-1) G_{t+k}^T + W_{t+k} \quad (20)$$

$$Q_t = F_{t+k}^T R_t(k) F_{t+k} + V_{t+k} \quad (21)$$

3 The Approach

3.1 Application to Residuals. Bayesian forecasting acts like a filter when applied to time series in its basic form (described in Sec. 2.2). By choosing the variances ω , the behavior of the filter can be adjusted between the two following extremes:

- (1) *High degree of smoothing.* Choosing small variances leads to a very smooth result, but also to a very low sensitivity with respect to changes.
- (2) *High degree of sensitivity.* Choosing large variances leads to a very sensitive result with a fast response to changes but also to a low rate of smoothing.

The application of Bayesian forecasting to residuals is shown in Fig. 4. The example data is given in an equidistant sequence of one measurement per cycle. For the sake of simplicity, the standard deviation is $\sigma=1$. A change of the mean value is implanted at cycle $t=51$ ($\Delta\mu = -3\sigma$), and a degradation with a constant gradient of $\beta = -0.02$ is implanted, starting at cycle $t=101$. The variances are set to $\omega_1 = \omega_2 = 1 \times 10^{-6}$. The upper diagram comprises the measurements (\bullet), the result of Bayesian forecasting (guess for mean value μ), and the uncertainty limit. How the uncertainty limit is defined and calculated will be explained in Sec. 3.2. The lower diagram shows the true and guess value (β) for the gradient. It can be seen that the guess value meets the true value at $t=150$, which is 50 cycles after its initiation.

In order to develop a method that meets the requirements of both, a high degree of smoothing and a high degree of sensitivity to changes, a special logic was developed, which will be described in the following sections in detail.

3.2 Outlier Detection. The first step in the process of developing a detection logic is the detection of potential outliers. This is carried out, based on the calculation of Bayes factors, which essentially are ratios of two different PDFs. At every time step t , the method of Bayesian forecasting yields a PDF for the next observation, which will be referred to as *model 0* in the following discussion. The basic idea is to check the current measurement

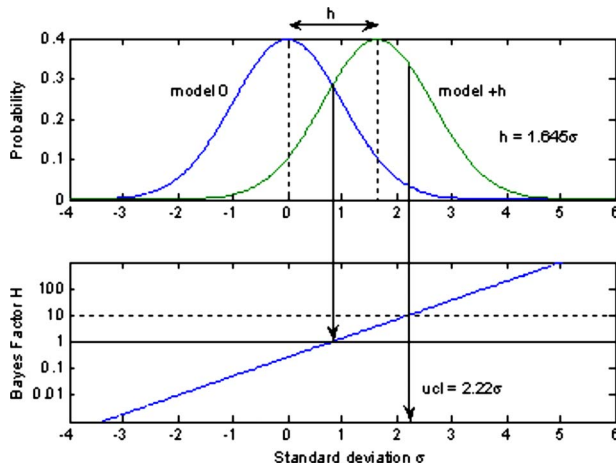


Fig. 5 Derivation of the Bayes factor and the ucl from the parameter h

against this model, and against an alternative model, where mean value is shifted by $+h$. Therefore, the alternative model is referred to as *model +h*. The Bayes factor is the ratio of the PDF values of the two models [17]:

$$H_{1,t} = \frac{\text{PDF value of model } +h}{\text{PDF value of model } 0} \quad (22)$$

In case of Gaussian distributions, the Bayes factor results in

$$H_{1,t} = \exp\left(\frac{2h \cdot (Y_t - f_t) - h^2}{2Q_t^2}\right) \quad (23)$$

The upper diagram of Fig. 5 shows the two PDFs with an offset of $h=1.645\sigma$. The lower diagram of Fig. 5 shows the Bayes factor. It can be seen that the Bayes factor is a straight line in logarithmic scaling, and therefore, monotonic, i.e., the greater the Bayes factor, the better the fit between the measurement and the alternative model. A Bayes factor of $H=1$ indicates that the probability of the measurement, deriving from the PDF of model 0, is equal to the probability of it, deriving from the PDF of model $+h$. Jeffreys [20] suggested to introduce a limit of $H_{\min}=10$ for outlier detection. In conjunction with a shift value of $h=1.645\sigma$, this yields an uncertainty limit (ucl) of 2.22σ . That means that a measurement is assessed as the outlier if its deviation from the mean value of model 0 is larger than the respective uncertainty limit.

Figure 5 implies that the uncertainty limit is a function of the shift parameter h and the limit parameter H_{\min} . The following equation applies:

$$\text{ucl} = \frac{\ln(H_{\min})}{h} + \frac{h}{2} \quad (24)$$

This relationship is shown in Fig. 6 for a given value of $H_{\min}=10$. It can be seen that a minimum exists at $h=2.15\sigma$ with a minimum uncertainty limit of $\text{ucl}=2.146\sigma$.

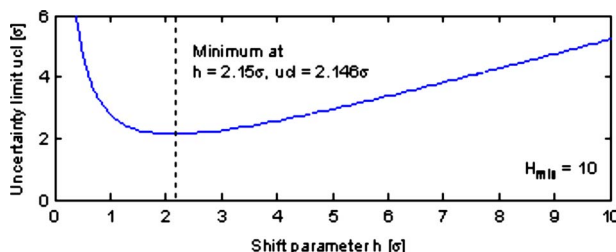


Fig. 6 Uncertainty limit as a function of h

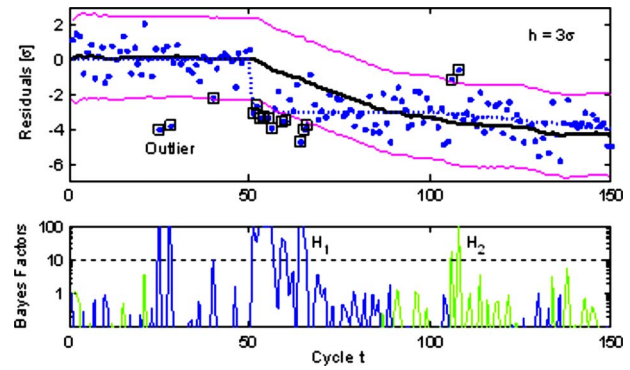


Fig. 7 Application of outlier detection to sample data (outliers are marked with a square \square)

Based on the logic explained above and the Bayes factor defined in the equation, only such outliers can be detected that exhibit a positive deviation. In order to enable the detection of outliers with a negative deviation, a second Bayes factor has to be defined as follows:

$$H_{2,t} = \exp\left(\frac{-2h \cdot (Y_t - f_t) - h^2}{2Q_t^2}\right) \quad (25)$$

Both Bayes factors H_1 and H_2 have to be checked in parallel to detect positive and negative outliers. The application of the described outlier detection to the sample data introduced in Sec. 3.1 is shown in Fig. 7. The shift parameter was set to $h=3\sigma$, resulting in an uncertainty limit of $\text{ucl}=2.27\sigma$. The lower diagram of Fig. 7 comprises the two Bayes factors H_1 (blue) and H_2 (green), plotted on top of each other. It can be seen that measurement residuals with either H_1 or $H_2 > H_{\min}$ are marked as outliers (\square).

3.3 Change Detection. The logic for outlier detection has to be extended in order to distinguish between single outliers and the beginning of a change. This is done by introducing cumulative Bayes factors, which are defined as the product of the preceding k consecutive Bayes factors

$$H_i(k) = \prod_{t=k+1}^i H_t \quad \text{with } k=1,2,\dots,l_{\max} \quad (26)$$

where l_{\max} denotes the maximum number of Bayes factors taken into account. The maximum cumulative Bayes factor is given by

$$L_t = H_i(l_t) = \max(H_i(k)) \quad \text{with } 1 \leq l_t \leq l_{\max} \quad (27)$$

where l_t is referred to as the run length. It describes the number of preceding consecutive Bayes factors leading to a maximum cumulative Bayes factor defined by Eq. (26). It shall be pointed out that the maximum cumulative Bayes factor has to be calculated based on both Bayes factors H_1 and H_2 , leading to L_1 and L_2 . Figure 8 shows the Bayes factors H , the maximum cumulative Bayes factors L , and the run length l_t for the sample data. It can be seen that the maximum cumulative Bayes factors L exceed the given threshold of $H_{\min}=10$ in case of changes only ($t=51$), and not in case of single outliers. It is also clearly visible that the run length holds the information of the time when the change took place (time of occurrence). The given run length threshold of $l_{\min}=4$ (introduced, as supposed by Pole et al. [17]) is exceeded at $t=54$ (time of notification, TNO). The time of occurrence (TOC) calculates to $t-l_t+1=51$. The difference between TOC and TNO is referred to as retrospectivity of the detection.

In contrast to the notification of outliers, the notification of changes can be triggered by two following reasons:

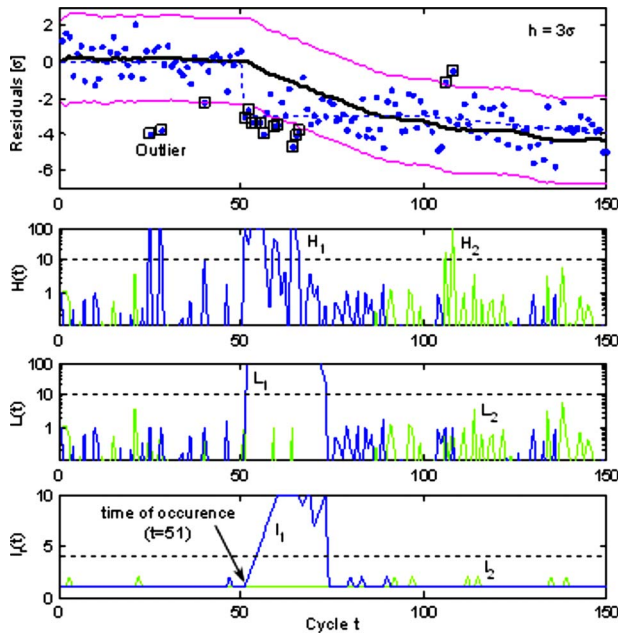


Fig. 8 Bayes factor H , maximum cumulative Bayes factor L , and run length l_r of the sample data

- (1) The occurrence of two consecutive Bayes factors of $H > H_{min}$, which is equal to the occurrence of a maximum cumulative Bayes factor of $L > H_{min}^2$.
- (2) The simultaneous occurrence of a run length of $l_r \geq l_{min}$ and a maximum cumulative Bayes factor of $L > H_{min}$.

These two conditions are the key elements of the detection logic developed. The first condition applies to changes with a height of approximately more than three times of the residuals' standard deviation, whereas the second one applies to changes with heights in the range of approximately one to three times of the residuals' standard deviation. The flow chart of the detection algorithm is shown in Fig. 9.

In case of a change in detection, the retrospective adjustment is carried out by resetting the time to the time of occurrence, and by setting the mean value of model 0 to the measurement residual at time t , i.e., $m(t) = Y(t)$. Subsequently, the analysis is continued, starting at the TOC with the adjusted model.

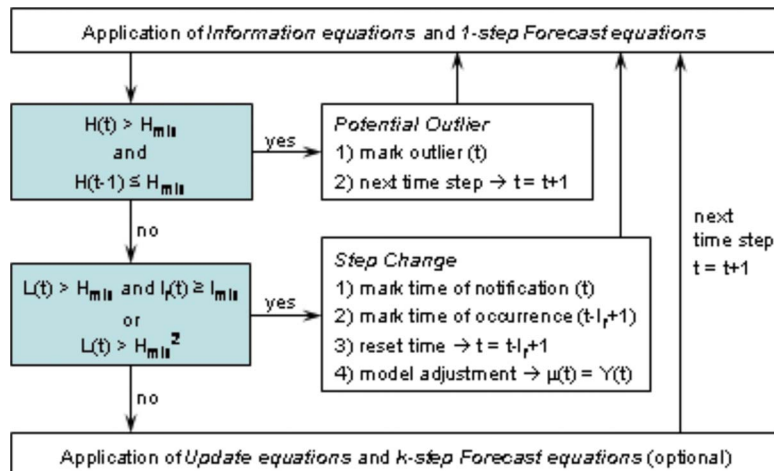


Fig. 9 Flow chart of the detection algorithm

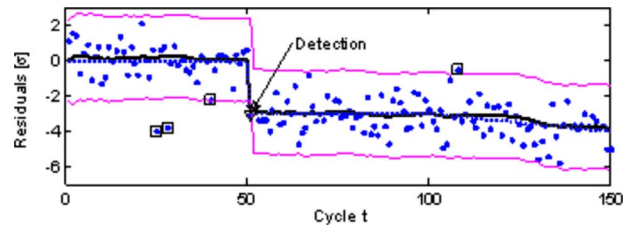


Fig. 10 Application of the method to a change with a change height of $\Delta\mu = -3\sigma$ (retrospectivity=1)

The application of the detection logic to sample data is shown in Figs. 10 and 11, with change heights of $\Delta\mu = -3\sigma$ and $\Delta\mu = -1.5\sigma$, respectively.

The example shown in Fig. 10 exhibits a retrospectivity of only one cycle, i.e., the condition $L > H_{min}^2$ that takes hold. In contrast, Fig. 11 shows an example where the conditions $l_r \geq l_{min}$ and $L > H_{min}$ are fulfilled. The retrospectivity of the detection is five cycles. It must be pointed out that the detected TOC ($t=57$) does not coincide with the real TOC ($t=51$). Hence, the actual retrospectivity of the detection is 11 cycles.

3.4 Prognosis. As the method of Bayesian forecasting provides guesses for both the value of the observed process and its dynamics (gradient, curvature, etc.), it is well suited to perform a prognosis of the measurement residuals. Therefore, the k -step forecast Eqs. (17)–(21) have to be applied. When using a second order DLM (where only value and gradient are calculated), the prognosis is essentially a continuation of the gradient, calculated at the last observed cycle. Thus, the prognosis is linear. Figure 12 shows the application of the k -step forecast equations to sample data. The upper diagram contains the sample data (\bullet), the prognosis ($-$), and its uncertainty limits ($--$). The lower diagram shows the evolution of the gradient guess value β .

It can be seen that the prognosis is carried out with a gradient guess value of $\beta = -0.0235$, which is 17.5% lower than the actual gradient value of -0.02 . Another important finding is that the uncertainty limit increases overtime, i.e., the longer the forecast period, the wider the model PDF. This can be explained by the lack of new information leading to greater uncertainty about the process developing in the respective period. The degree of the uncertainty limit enlargement also depends on the variances chosen for the process modeling.

Based on the prognostic results, two important functions can be derived, concerning the exceedance of given limits for the mea-

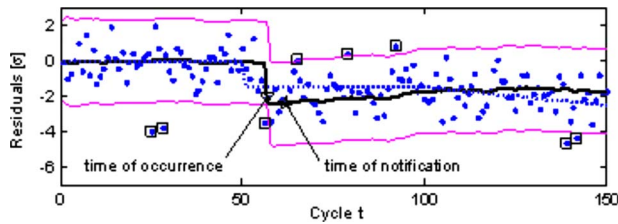


Fig. 11 Application of the method to a change with a change height of $\Delta\mu = -1.5\sigma$ (retrospectivity=5)

surement residuals, the limit PDF, and the limit cumulative distribution function (CDF). Figure 13 shows the derivation of the two functions for sample data.

The limit PDF results by evaluating the model PDF at the limit value overtime. If no events occur in the vicinity of the given limit, its PDF is of the Gaussian shape. The integration over the limit PDF yields the limit CDF, which specifies at what time and with which probability the respective limit will be exceeded. In order to obtain probability values in the interval of [0,1], the integral is divided by its respective infinite integral

$$CDF(t) = \frac{\int_0^t PDF(x)dx}{\int_0^\infty PDF(x)dx} \quad (28)$$

As it is computationally impossible to take into account an infinite number of cycles, the infinite integral is replaced by the integral

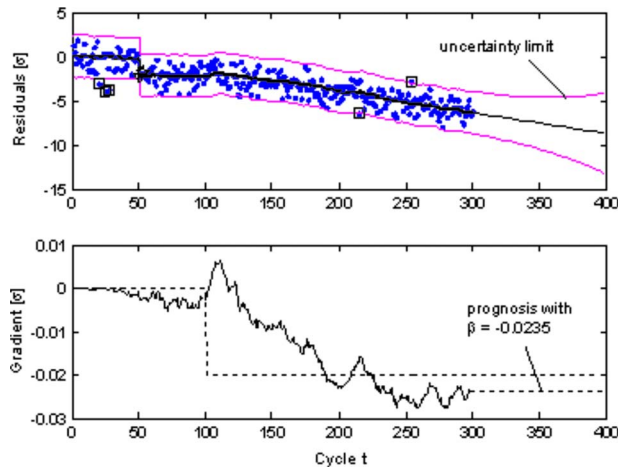


Fig. 12 Prognosis example with constant gradient

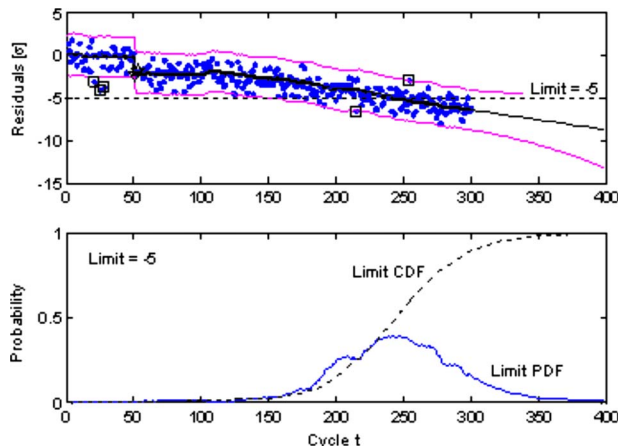


Fig. 13 Derivation of the limit PDF and the limit CDF

Table 1 Details of the evaluation test program

Parameter	Values	Number
$\Delta\mu[\sigma]$	0.7, 0.8, 0.9, 1.0, 1.1, 1.2, 1.5, 2.0, 2.5, 3.0, 4.0, 5.0, 6.0, 7.0, 8.0, 9.0	16
Runs	20 cycles with a step change at cycle 10 and the respective height $\Delta\mu[\sigma]$	99
Total runs	99 runs times 16 change heights	1,584
Total cycles	1584 runs times 20 cycles	31,680

over all the available cycles, including the forecasted cycles. This approach is only feasible if the considered period of time is long enough. As a rule of thumb the maximum of the limit PDF should be exceeded by at least three times its standard deviation.

4 Statistical Evaluation of the Method

The developed method was run through an extensive test program comprising 31,680 sample cycles.

- (1) Changed heights of 16 different magnitudes have been investigate from $\Delta\mu = 0.7\sigma$ to $\Delta\mu = 9.0\sigma$.
- (2) 99 runs have been performed per changed height, where at every run, it consisted of 20 cycles, and incorporated a change at cycle 10 with the respective changed height.

A detailed summary of the test program, especially the values chosen for changed height, can be found in Table 1.

The performance of the method was assessed using the following three detection rates:

$$\text{detection rate} = \frac{\text{number of correct detections}}{\text{number of existing changes}} \quad (29)$$

$$\text{missed detection rate} = \frac{\text{number of missed detections}}{\text{number of existing changes}} \quad (30)$$

$$\text{false detection rate} = \frac{\text{number of incorrect detections}}{\text{possible incorrect detections}} \quad (31)$$

The detection rates can be visualized in the matrix, as shown in Fig. 14.

In addition to the detection rates, a detection limit is introduced, denoting a period of time within which a detection is defined as correct. A detection is assessed as correct if the detected time of occurrence falls within the range of the actual time of occurrence \pm the detection limit (see Fig. 15 for a visualization of this definition). It has to be pointed out that using the above definition, a detection can be assessed as correct even if its detected TOC is earlier than its actual TOC. This especially applies for events with changed heights $< 3\sigma$, i.e., when measurement uncertainty can affect the detected TOC.

It shall be clearly stated that using the above definition of the detection limit, a detection is assessed as a false detection if the

		Change detected	
		Yes	No
Change existent	Yes	detection rate (correct positive)	missed detection rate (false negative)
	No	false detection rate (false positive)	n/a

Fig. 14 Definition of the different detection rates

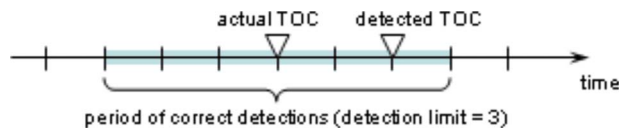


Fig. 15 Definition of a correct detection using the detection limit parameter

calculated TOC lies outside the defined period of correct detections (see Fig. 15). This definition leads to two effects, which shall be clarified at this point.

- (1) The false detection rate varies with the changed height.
- (2) The false detection rate is inflated, compared with other definitions, not taking into account this effect.

Although the used false detection rate definition differs from the ones used by other researchers in the field, it was chosen as detecting a change; as soon as possible is deemed crucial for engine diagnostic systems. A change that is detected too late is not of much value, and can even result in a hazardous engine failure. Therefore, the detection limit has been introduced to define a measure of time within which a change has to be found to be assessed as correct detection.

The data sets used for the statistical assessment of the method comprise a variation of the changed height between 0.7σ and 5σ ; a variation of the detection limit in the range of $[0,1,2,3]$ and three different values for the shift parameter h , which are shown in Table 2.

The variation in the shift parameter h leads to a variation in the ucl. Among the definition of ucl in terms of multiples of the standard deviation, it can also be interpreted as the percentage of the measurement residuals that fall within the range $[-ucl, +ucl]$. This value can be calculated by the integration of the model 0 PDF over the aforementioned range. The resulting ucl values (both the multiples of σ , as well as the percentage values) are given in Table 2.

Figures 16 and 17 show the detection rate and the false detection rate, respectively. Both rates are plotted as functions of changed height and uncertainty limit for a detection limit of three cycles.

It can be derived from Fig. 16 that a detection rate of 100% is achieved for changed heights as small as 2σ , if the uncertainty

Table 2 Variation in shift parameter h in the data sets and depending values for ucl

h (σ)	ucl (σ)	ucl (%)
2.493	2.170	97
3.225	2.326	98
4.0	2.576	99

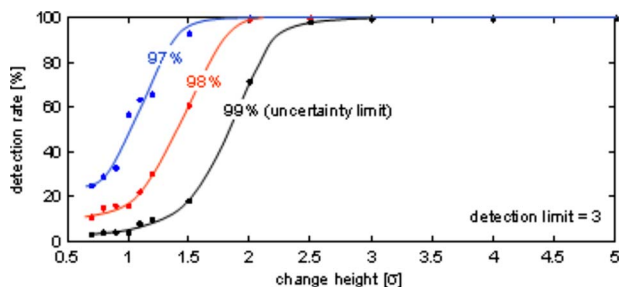


Fig. 16 Detection rate as functions of change height and uncertainty limit (detection limit=3)

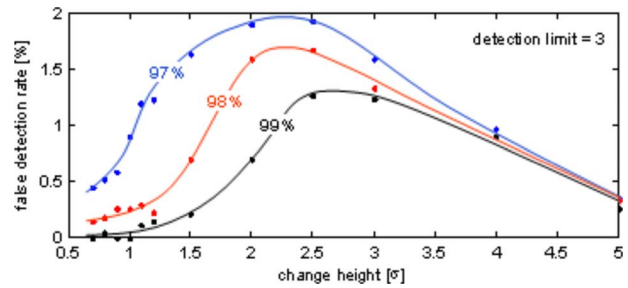


Fig. 17 False detection rate as functions of change height and uncertainty limit (detection limit=3)

limit is set to 97%. On the other hand, Fig. 17 shows that this is gained by a false detection rate as high as 2%. Another finding, based on Fig. 16 and Fig. 17, is that the method is able to reliably detect changes down to a changed height of 1.5σ to 2σ , and that the false detection rate has a maximum for changed heights between 2σ and 2.5σ . Increasing the uncertainty limit decreases both detection rate and false detection rate. An optimum must be chosen between a high detection rate and a low false detection rate. The definition of this optimum should consider a weighting of the two rates, especially the false detection rate that should be weighted with respect to the detected changed height. Furthermore, in application, the choice of the uncertainty limit will depend on the individual characteristics of the observed process.

5 Conclusions

A similar detection logic could have been developed based on other methods as well, but the method of Bayesian forecasting applied to DLMs provided several advantages over other methods of which the most important one is its statistical nature. The use of probability density functions enables the calculation of uncertainty limits, and therefore, the design of a statistics based detection logic.

5.1 Key Abilities of the Method. The developed method holds the following benefits compared with detection algorithms that are currently in use.

- (1) The conflictive requirements of a high degree of smoothing, and a quick reaction to changes are fulfilled in parallel by the special detection logic.
- (2) The dynamics of the process is modeled by DLMs, where the order can be adjusted to satisfy the prognostic needs (linear, quadratic, etc.).
- (3) Due to the statistical nature of the method, uncertainty limits can be given for the prognosis. Furthermore, limit PDFs and limit CDFs can be calculated.
- (4) By combining two different detection conditions, the method is able to reliably detect changes with a changed height of less than 3σ , but also to react to detections with larger changed heights with a delay of only 1 cycle.

5.2 Implementation of the Method. The developed method was incorporated into a gas turbine monitoring software, which was also developed by the leading author. In addition to the detection algorithm described in this paper, this software comprises diagnostic algorithms for the assessment of both single events and gradual degradation on a component level. Details of the algorithm used for single event diagnosis can be found in Ref. [15]. The diagnostic software also contains a simple observability analysis of the instrumentation, which is based on the calculation of the system's condition number [21]. The software will be used at the Institute of Aircraft Propulsion Systems for future monitoring work.

5.3 Future Work. A comparison between the developed method and other state-of-the-art detection algorithms, especially linear and nonlinear filter techniques, is the topic of the current work. First results show that the method outperforms linear filter techniques by far, but at this stage, no statement can be given with regard to nonlinear filter algorithms.

A statistical evaluation of the performance abilities of the method (similar to the statistical evaluation of the detection abilities presented in Sec. 4) would also be very beneficial. Figure 12 indicates that the method slightly overestimates the gradient. But as Fig. 12 represents only one specific example, and the estimation accuracy depends on many factors, e.g., the type of the gradient (constant or changing), the value of the gradient, the forecast horizon, and the time between the gradient has started, and the time the forecast is calculated, a statistical evaluation of many test cases should be performed, taking into account the variation in all relevant parameters.

Another very important investigation would be the comparison of the developed method against field data. Unfortunately, no such data is available at present, but the method will be applied to real-world data as soon as respective data becomes available to the authors.

In terms of improving the developed detection algorithm, a logic for the interaction of detections of multiple measurements could be implemented. At present, every measurement is run through the detection algorithm separately, and a detection flag is set if a change is detected in one of the parameters. One simple possibility of an enhanced logic would be to set a detection flag, only if a certain percentage of the observed parameters show a detection within a defined period of time.

Acknowledgment

The work of this paper has been carried out in the frame of the project KW21 (Power Plants for the 21st Century). The financial support of the MTU Aero Engines GmbH and the Ministry of Science, Research and Culture of the Federal State Baden-Württemberg is gratefully acknowledged.

Nomenclature

DLM	=	equation parameters
Θ	=	system vector comprising μ and β
Y	=	measurement residual
μ	=	process value estimation
$\Delta\mu$	=	process changed height
β	=	process gradient estimation
F, G	=	system specific matrices
V, W	=	variance matrices
ω_1, ω_2	=	process value/gradient variance
ν	=	measurement residual variance

Bayesian Equation Parameters

σ	=	standard deviation
h	=	shift parameter
ucl	=	uncertainty limit
m, a, f	=	mean values

C, R, Q	=	variance matrices
A	=	matrix for predictor correction
H	=	Bayes factor
L	=	maximum cumulative Bayes factor
l_r	=	run length
l_{\min}	=	run length threshold

References

- [1] Lipowsky, H., and Staudacher, S., 2008, "Method and Apparatus for Gas Turbine Monitoring," German Patent No. DPMA 10 2008 022 459.6.
- [2] DePold, H. R., and Gass, D. F., 1999, "The Application of Expert Systems and Neural Networks to Gas Turbine Prognostics and Diagnostics," ASME J. Eng. Gas Turbines Power, **121**, pp. 607–612.
- [3] Brotherton, T., and Johnson, T., 2001, "Anomaly Detection for Advance Military Aircraft Using Neural Networks," *Proceedings of the IEEE Aerospace Conference*.
- [4] Ganguli, R., 2002, "Data Rectification and Detection of Trend Shifts in Jet Engine Path Measurements Using Median Filters and Fuzzy Logic," ASME J. Eng. Gas Turbines Power, **124**, pp. 809–816.
- [5] Ghoreyshi, M., Pilidis, P., and Ramsden, K. W., 2005, Diagnostics of a Small Jet Engine-Neural Networks Approach," ASME Paper No. GT-2005-68511.
- [6] Eustace, R. W., 2007, "A Real-World Application of Fuzzy Logic and Influence Coefficients for Gas Turbine Performance Diagnostics," ASME Paper No. GT2007-27442.
- [7] Ganguli, R., 2002, "Fuzzy Logic Intelligent System for Gas Turbine Module and System Fault Isolation," J. Propul. Power, **18**(2), pp. 440–447.
- [8] Kamboukos, P., Mathioudakis, K., and Stamatis, A., 2003, "A Comparative Study of Optimization Methods for Jet Engine Condition Diagnosis," 16th International Symposium on Air-Breathing Engines, Cleveland, OH, Aug. 31–Sept. 5.
- [9] Mathioudakis, K., Kamboukos, P., and Stamatis, A., 2002, "Turbofan Performance Deterioration Tracking Using Non-Linear Models and Optimization Techniques," ASME Paper No. GT-2002-30026.
- [10] Gulati, A., Taylor, D., and Singh, R., 2001, "Multiple Operating Point Analysis Using Genetic Algorithm Optimisation for Gas Turbine Diagnostics," 15th International Symposium on Air-Breathing Engines, Bangalore, India, Sept. 3–7.
- [11] Aretakis, N., Mathioudakis, K., and Stamatis, A., 2004, "Identification of Sensor Faults on Turbofan Engines Using Pattern Recognition Techniques," Journal of Control Engineering Practice, **12**(7), pp. 827–836.
- [12] Loukis, E., Mathioudakis, K., and Papailiou, K., 1994, "Optimizing Automated Gas Turbine Fault Detection Using Statistical Pattern Recognition," ASME J. Eng. Gas Turbines Power, **116**(1), pp. 165–171.
- [13] Sampath, S., and Singh, R., 2006, "An Integrated Fault Diagnostics Model Using Genetic Algorithm and Neural Networks," ASME J. Eng. Gas Turbines Power, **128**, pp. 49–56.
- [14] Kobayashi, T., and Simon, D. L., 2001, "A Hybrid Neural Network-Genetic Algorithm Technique for Aircraft Engine Performance Diagnostics," J. Propul. Power, **21**(4), pp. 751–758.
- [15] Lipowsky, H., Staudacher, S., Nagy, D., and Bauer, M., 2008, "Gas Turbine Fault Diagnostics Using a Fusion of Least Squares Estimations and Fuzzy Logic Rules," ASME Paper No. GT2008-50190.
- [16] Bauer, M., and Staudacher, S., 2006, "Fully Automated Model Based Performance Analysis Procedure for Online and Offline Applications," ASME Paper No. GT2006-91050.
- [17] Pole, A., West, M., and Harrison, J., 1994, *Applied Bayesian Forecasting and Time Series Analysis*, Chapman and Hall, New York.
- [18] Provost, M. J., 2003, "Kalman Filtering Applied to Time Series Analysis," VKI Lecture Series, Gas Turbine Condition Monitoring and Fault Diagnosis, Brussels, Belgium, Jan. 13–17, Paper No. LS-2003-01.
- [19] West, M., and Harrison, J., 1997, *Bayesian Forecasting and Dynamic Models*, 2nd ed., Springer, New York.
- [20] Jeffreys, H., 1961, *Theory of Probability*, Oxford University Press, London.
- [21] Roesnick, M., 1984, "A System Theory Based Solution of the Failure Diagnosis Problem Applied to a Jet Engine," Doctoral thesis, Institute of Automation Engineering, University of the German Federal Armed Forces, Hamburg, Germany.

Mapping the Density Fluctuations in a Pulsed Air-Methane Flame Using Laser-Vibrometry

Fabrice Giuliani¹

e-mail: fabrice.giuliani@tugraz.at

Thomas Leitgeb

Andreas Lang

Jakob Woisetschläger

Department for Gas Turbine Combustion,
Institute for Thermal Turbomachinery and
Machine Dynamics,
Graz University of Technology,
Inffeldgasse 25A,
A-8010 Graz, Austria

Laser vibrometry (LV) is originally a laser-based, line-of-sight measurement technique dedicated to the analysis of surface vibrations. It was lately adapted at TU Graz for monitoring the stability of an air-methane flame (Giuliani, et al., 2006, ASME Turbo Expo, ASME Paper No. GT2006-90413). This paper reports on the mapping of density fluctuations measured with LV in a premixed air-methane flame (free jet; swirl stabilized) with a forced flow modulation (quarter-wave resonator; amplification with a siren). In order to correlate the density fluctuations with the jet aerodynamics and turbulent flame shape, stereoscopic particle image velocimetry and high-speed schlieren visualizations were used. This paper addresses issues regarding the estimate of density fluctuations, the transform from line-of-sight to local measurement with tomographic methods, and the potential of the method for detailed description of thermoacoustic couplings. One emphasized application of LV is its ability to perform precise and low-cost benchmark stability tests on a combustor during the design phase (time-resolved measurement, high frequency and phase resolution on the 5 Hz–20 kHz range with the present equipment and settings, near-constant spectral sensitivity over a large bandwidth, and no seeding required; measurement possible over the whole combustion volume).

[DOI: 10.1115/1.3159373]

1 Introduction

The well-known industrial problem that motivates this work is the limitation on performance of power and propulsion gas turbines due to the presence of combustion instabilities. These are detrimental—if not critical—for the system operability and integrity. The common research effort focuses on the development of control strategies and design guidelines toward a combustor ensuring stable combustion, with extended operation range in the lean combustion region for environmental issues. The development of ad hoc combustion stability monitoring techniques, dedicated to be implemented in an active control loop [1], is an issue of great importance.

The present work is a proof of concept of a measurement technique based on laser interferometry, designed to observe density fluctuations and able to monitor in real time the stability of continuous combustion processes. This technique, laser vibrometry (LV), represents a convenient alternative to the Rayleigh scattering method [2–4] since it offers a large spectral sensitivity that fits the frequency bandwidth where thermoacoustic couplings usually take place (e.g., 5 Hz–20 kHz for this application). This is a new application of this instrument, which was originally dedicated to surface vibration analysis [5], and that was also shown to detect coherent turbulence in a turbine flow [6,7].

While most wall-mounted sensing techniques such as fast pressure sensors placed in the diffuser or connected to the flame tube may suffer from side-effect disturbances (e.g., cooling jet noises), line-of-sight measurement techniques have the possibility to detect combustion instabilities at the very core of the flame, by aiming, for instance, through cooling holes so that sensitive parts are not directly exposed to the flame. LV offers the ability not only to detect density fluctuations over a wide spectral range but also to

measure precisely their amplitude and phase. This property is essential for combustion stability control purposes. LV allows also the establishment of detailed density fluctuation maps with the help of a reconstruction technique described in this paper.

The physical quantity of interest is the local and temporal density fluctuation ρ' of the fresh mixture, as well as of the burnt gases. ρ' is coupled to both pressure fluctuations p' and heat release fluctuations q' . As far as the Rayleigh criterion is concerned [8], a strong amplification of ρ' is a potential marker for instability detection.

In the present experiment, a premixed methane-air burner is placed at the end of a quarter-wave resonator, so that the flame resonant modes are known. A siren amplifies these modes and enhances the injector's jet dynamics, so that a vortex-driven combustion instability [9] is produced. As a result of this study, a strong correlation between the density fluctuation maps, the flame dynamics, and the jet aerodynamics is established.

The real-time analysis of ρ' is performed by a set of two laser vibrometers (see Ref. [10]). The method for the phase-averaged ρ' mapping is described. LV can perform a refined description of flame-vortex interaction in time and space, outside and inside the flame, and offers the advantage of an easy setup and simple use. In order to support these observations, high-speed schlieren visualizations and stereoscopic particle image velocimetry (PIV) measurements are presented. LV measurements are highly sensitive and nonintrusive. Embedding this sensing technique in gas turbines is a work in progress, since much research and development is required regarding the sensor's optics and drift compensation in the presence of aggressive flow conditions and vibrations. Eventually state-of-the-art LV is recommended as a burner test bench method for combustion stability analysis.

2 Setup and Operation

The atmospheric single-burner sector that was fully engineered at TU Graz (see Ref. [10] for more details) allows the analysis of a premixed air-methane flame under controlled inlet pulsed flow

¹Corresponding author.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 23, 2009; final manuscript received March 26, 2009; published online December 3, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Gas Turbine Technical Congress and Exposition, Orlando, FL, June 8–12, 2009.

Table 1 Configuration and operating conditions

Configuration	
Resonator length	3.08 m
Amplified modes	$25 + N \times 50$ Hz, $N = \{1, 2, 3, \dots\}$
Axial swirl angle	45 deg
Number of blades	3
Venturi ratio	2/5
Operating conditions	
P and T_{inlet}	Room conditions, 1 bar, 300 K
Air feed	10 g/s
Methane feed	0.47 g/s
Equivalence ratio	$\Phi = 0.8$
Axial inlet velocity	$U_{mean} = 12.6$ m/s
Injector pressure loss	$\Delta P_{inj} = 240$ Pa
Swirl number S	0.8
Siren excitation range	110–120 dB SPL

2.1 Premixed Air-Methane Burner. The injection module combines a Venturi nozzle to accelerate the flow with an axial swirler to swirl stabilize the flame, keeping the flame front away from the front plate and preventing flashback. Premixing takes place in the supply pipe (Fig. 1, a sketch of the injector is also visible in Fig. 7). Axial instead of radial swirler blades were chosen so that the injector can be considered as acoustically neutral (where pulsation damping and/or phase shift can be neglected over the frequency range of interest) at the cost of a relatively low swirl number ($S=0.8$, with S as defined by Beér and Chigier [11], where the pressure term is neglected).

2.2 Resonator and Siren. The burner is mounted at the end of an ensemble siren and quarter-wave resonator. The siren modulates the air mass flow rate (Fig. 2), and was provided to us by ONERA (Centre de Toulouse, France) [12]. The 3.08 m long pipe that follows is used as a quarter-wave resonator (closed-open resonator with a critical nozzle upstream, and an opened-end condition downstream), see Fig. 1. It enhances the fundamental mode 25 Hz

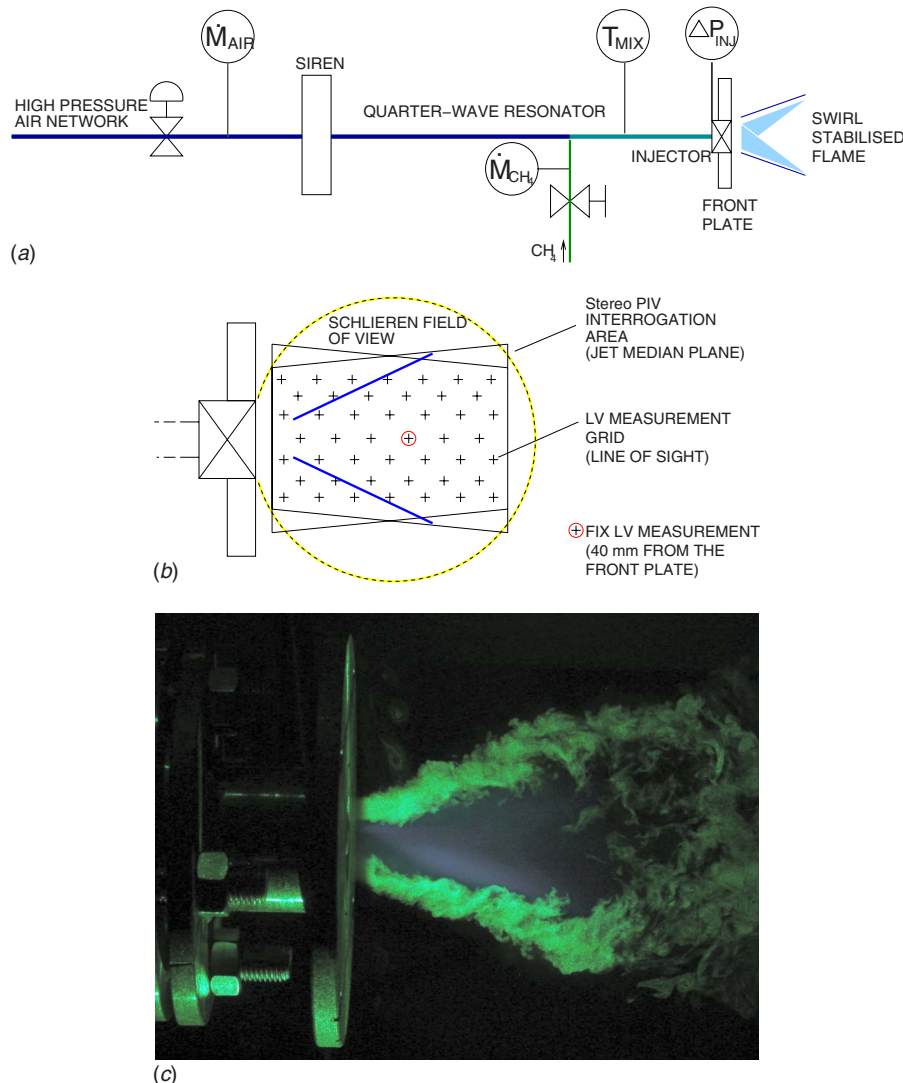


Fig. 1 Test rig. Top: ensemble siren, resonator and burner. Middle: measurements performed on the free-jet flame per specific technique. Bottom: flame aspect during the PIV measurements.

conditions. In this study, the burner power was set to 23.5 kW. All operation and configuration parameters of importance are reported in Table 1.

and all of its odd multiples (75 Hz, 125 Hz, 175 Hz, etc.). The siren amplifies these significant resonant modes (up to 120 dB sound pressure level (SPL) [13]) to achieve pressure fluctuation

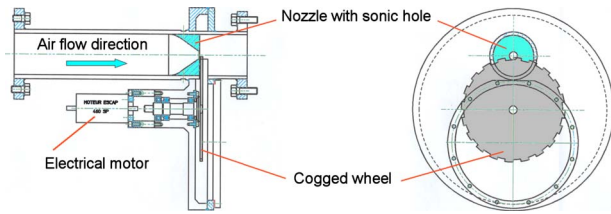


Fig. 2 ONERA siren, where a sonic jet sheared by a cogged wheel rotating at controlled velocity sets an air flow modulation at the desired frequency [12]

levels, comparable with those of a real thermoacoustic coupling. The methane nozzle (placed 0.6 m upstream the front plate) and the injector itself are assumed to be acoustically transparent.

2.3 Operating Conditions and Flame Aspect. The reactant mixture is set in the upper lean range ($\Phi=0.8$) to avoid flame blowout because of the pulsation. The fresh gases are perfectly premixed.

The flame pattern in the near injection zone is visible in Fig. 1, bottom. This picture was taken during the PIV measurements, using a seeding that burns. On the flame picture, the blue flame core is visible, surrounded by the tracer. In free-jet mode, the flame attaches to the tip of the injector, where ambient air is aspired and entrained. After that, the fresh injected mixture ignites, and the flame expands along a cone, surrounding the central recirculation zone generated by the swirler (the cone mean solid angle is 30 deg). The presence of the tracer downstream the flame marks the internal recirculation breakdown zone, where the highly turbulent flame ends.

3 Measurement Techniques

3.1 LV. The principle of LV is shown in Fig. 3, top. A laser interference pattern between a reference beam and an object beam (reflecting on the studied surface) is analyzed. If the surface moves, the path difference between the two beams changes, and

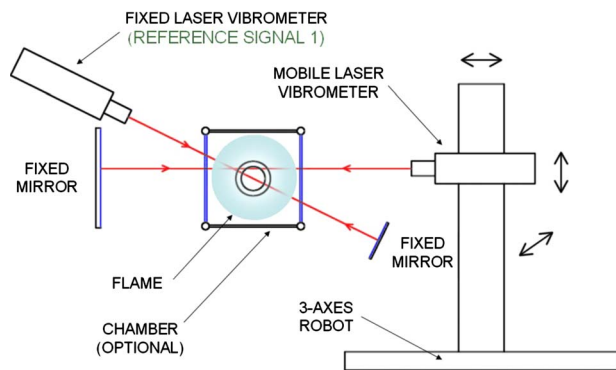
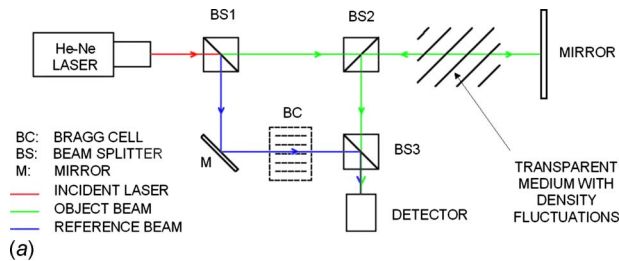


Fig. 3 Principle of laser vibrometry and dual-LV arrangement

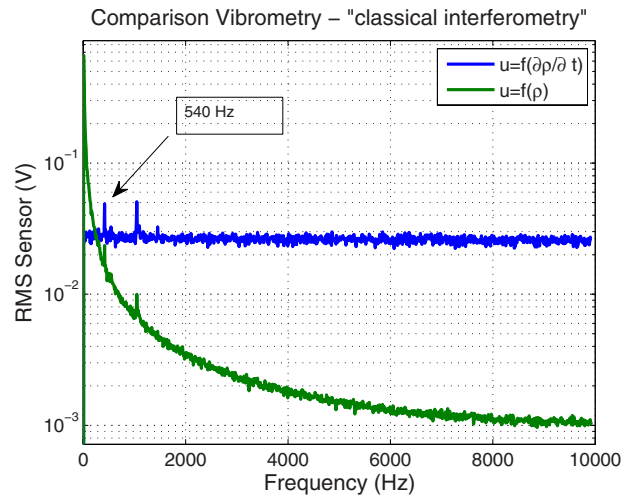


Fig. 4 Comparison between the LV output signal related to the time derivative of the density—or flat response, and the density itself—or $1/f$ response (the LV processor has an embedded integrator). Excitation at 540 Hz

so does the interference. In order to detect motion amplitudes greater than the laser wavelength, and to distinguish a forward from a rearward movement, a Bragg cell is used to slightly shift the wavelength of the reference beam. A stationary surface is then represented by a fixed frequency corresponding to the modulated interference. Any surface motion will provoke a Doppler effect on this carrier wave. A frequency demodulation allows to derive the vibration frequency and the motion amplitude of the object surface.

Mayrhofer and Woisetschläger [6] used a variant of this technique for density fluctuation analysis of ambient air, keeping the geometrical path constant (the object surface is a stationary mirror) so that mainly density fluctuations alter the optical path, advancing (negative gradient) or retreating (positive gradient) the phase front of light. As a result, LV can detect the presence of passing-by coherent structures. The relationship between LV signal voltage u' and density fluctuation ρ' is derived from Refs. [6,7] as follows:

$$u'_f = \frac{G}{k} \int_{-L/2}^{L/2} \frac{\partial \rho'_f}{\partial t} dz \quad (1)$$

where G is the Gladstone–Dale factor, related to the refraction index of the medium ($G=2.5 \cdot 10^{-4} \text{ m}^3/\text{kg} \pm 3\%$ for air-CH₄ mixture over an equivalence ratio interval varying from 0.4 to 1, as well as for the resulting burnt gases, computed after the method of Gardiner et al. [14]). k is an instrument gain factor, and L is the penetration length of the laser through the medium where the density fluctuates. Subscript f means phase averaged at a fixed frequency, so that u'_f and ρ'_f correspond to the narrowly band-passed signals—or phase averaged signals—of the time signals $u'(t)$ and $\rho'(t)$ at frequency f .

The benefit of using LV is that this instrument was designed to offer a nearly-constant spectral sensitivity over a large bandwidth [5]. While most time-resolved techniques such as Rayleigh scattering have a typical $1/f$ spectral response, LV offers a quasi-constant sensitivity since its output voltage is physically related to the time derivative of the density fluctuation, and not to the density fluctuation itself (see the two types of responses on Fig. 4). The Fourier transform of a time-derivative function is equal to the Fourier transform of the function, times the frequency: $F((d\rho'(t))/dt) = i\omega F(\rho'(t))$. The signal-to-noise ratio remains high over a large bandwidth, and the phase measurement, precise.

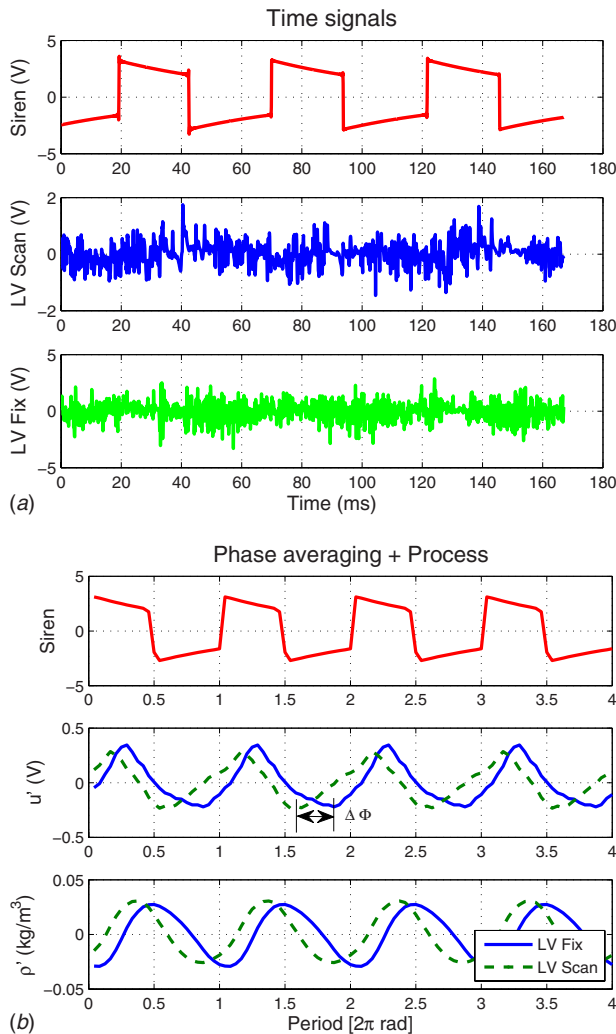


Fig. 5 Phase averaging process for LV signals, over a 25 Hz pulsation. Top: time signals. Bottom: phase-averaged signals displayed over four pulsation cycles, with the scanning LV voltage u'_i , and resulting ρ'_i computed after Eq. (2).

Therefore, LV can report qualitatively on density fluctuation using fixed instrument settings, since no extra gain compensation is required as a function of the observed frequency.

Two vibrometer models, Polytech OVD 353 with an OFV-3001 controller (Polytec, Waldbrunn, Germany), were used. One instrument provided a fixed-point reference measurement, while the second one scanned through the flame over a predefined grid (see Fig. 3, bottom, for LV arrangement and Fig. 1, bottom, for the specific measurement grid).

The determination of the phase-defined u'_i signal at frequency f based on the correlation between the scanning LV time signal and the reference LV time signal recorded simultaneously is described in Ref. [10] for a naturally resonant flame. When forced pulsation is studied, the reference signal is a transistor-transistor logic (TTL) emitted by the siren. This allows to perform phase averaging (Fig. 5). Using MATLAB user-defined postprocessing routines, the time signal is resampled over single pulsation periods, marked by an upward front from the TTL signal (Fig. 5, top). All subsamples are averaged, resulting in a single phase-defined period u'_i (Fig. 5, bottom).

Once u'_i is determined, the relative density fluctuation ρ'_i is derived from Eq. (1) as

$$\rho'_{iLS}(t+d\tau) \approx \rho'_{iLS}(t) + \frac{k}{GL} u'_i(t+d\tau) d\tau \quad (2)$$

with end condition $\bar{\rho}'_{iLS}=0$, where LS means integrated over the line-of-sight depth L . In the following, the factor term k/GL is assumed to be constant.

The mapping method consists of reorganizing all phase-defined quantities as a function of their coordinates (Fig. 6(a)), sorted with an increasing phase angle over one pulsation period. A smoothing interpolation (MATLAB v4) is used on the u'_i grid to provide a voltage map (Fig. 6(b)). The ρ'_i map computed with Eq. (2) is represented in Fig. 6(c). Finally, due to the nearly axial symmetrical geometry of the jet, an Abel transform is possible, in order to retrieve the local density fluctuation ρ' . One can derive a quantitative measurement of ρ' , provided that the integration path L was measured precisely. Although the Abel inversion generates its major uncertainty on the centerline (this underlines the lack of axial symmetry of the jet due to the injector design), it also positions the observed structures more precisely at the jet periphery than the integral measurement does (Fig. 6(d), where the Abel transform was performed on the two half planes of the jet).

In order to demonstrate the capacity of LV to perform a refined analysis of the flame dynamics, supporting techniques such as high-speed schlieren visualization and phase-averaged PIV were applied. The setup for the schlieren technique with high-speed camera is described in Ref. [10].

3.2 Stereoscopic PIV. PIV [15] measures the velocity flow field in a plane, illuminated by two consecutive laser pulses. In order to measure quantities such as the swirl number S , the measurement of the tangential velocity, or out-of-plane component is required. Stereo-PIV requires two cameras shifted from a given angle (45 deg in our case) that focus on the same measurement plane (Fig. 7), and perform simultaneously a PIV measurement. The third velocity component is derived from the analysis of the astigmatism between the two cameras. The laboratory's PIV system consists of a double-cavity Nd:YAG-laser (from New Wave Gemini, Sunnyvale, CA, 532 nm—green light, 120 mJ per pulse, pulse duration 3–5 ns, 15 Hz), two camera models: Dantec 80C60 HiSense (1280×1024 pixels, 12 bit grayscale), and a PIV processor model: Dantec FLOWMAP 1500 (Dantec Dynamics, Skovlunde, Denmark). The raw images are processed with the Dantec FLOWMANAGER PIV-software (v4.60.28). The field of view common to both cameras covers 75×60 mm² (Fig. 7). To generate vector plots out of the double images, interrogation areas of the size of 32×32 pixels, with a window overlap of 50% in both horizontal and vertical directions, were used.

The air flow is seeded with an aerosol of di-ethyl-hexyl-sebacat (DEHS) (specific particle size 0.7–1 μm) particles. This seeding was used for both isothermal jet characterization and for flame dynamics description (Fig. 1), focusing on the peripheral coherent structures and on the flame boundaries where the particles burn. The mass flow of injected DEHS is small, compared with the methane mass flow (mass ratio less than 1%), thus, its effect on combustion can be neglected.

The phase-averaging technique used for time-triggered measurements such as PIV is similar to the one described in Ref. [12]. Concerning the identification of vortices on the PIV velocity maps, the criterion called second invariant or Λ_2 from Jeong and Hussain [16] was used to sort out the multiple shear layers from the turbulent flow. To reduce the measurement uncertainty to less than 5%, a minimum of 200 phase-locked recordings with the stereo-PIV setup are required per measurement.

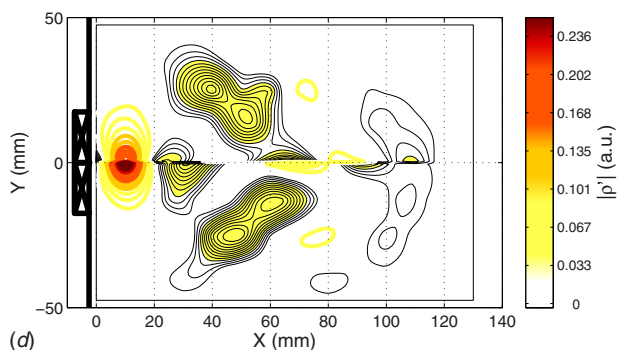
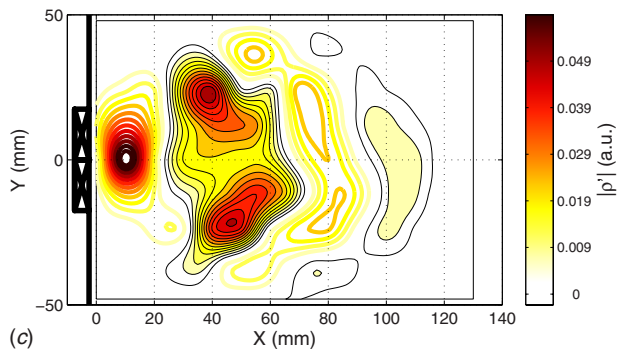
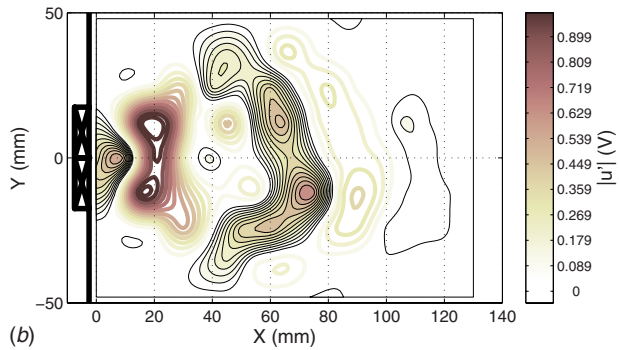
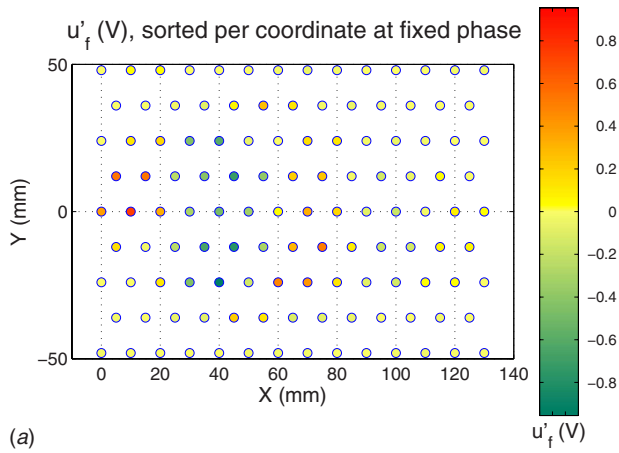


Fig. 6 Fluctuating density mapping process. The measurement grid is displayed in Fig. 1. (a) Map of the scanning LV filtered voltage $u'_f(t)$ at frequency $f=175$ Hz and (b) phase sub-period $t=d\tau$. (c) Map of the relative density fluctuation $\rho'_{f,S}(t)$. (d) Abel transform of the latter. Light contours: negative part of the fluctuation, line contours: positive part.

4 Results

4.1 Recurrent “Mushroom-Shaped” Flame. The shape modulation of a flame, governed by a vortex-driven instability, is

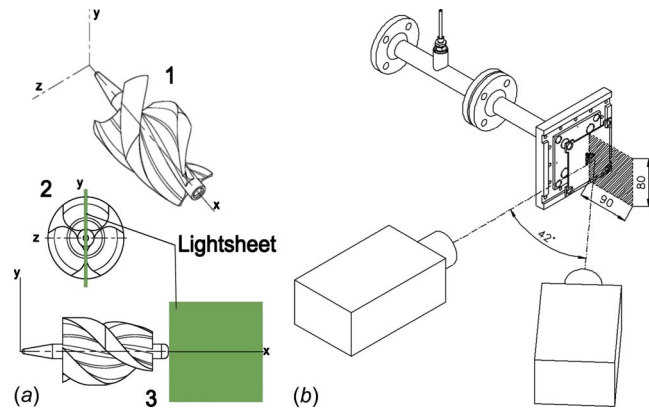


Fig. 7 Stereo-PIV arrangement and measurement plane

reproduced with the help of the ensemble siren and resonator pipe. The flame is attached forming a “tube” at the tip of the injector (Fig. 8). After that, a mushroom-shaped pattern appears around this tube, with the development of a 2D axisymmetric Kelvin–Helmholtz instability. This structure is advected by the flow, while the flame rolls up and closes on itself, and the next instability appears near the front plate. This process repeats in a recursive way. The siren excitation frequency drives the mixing layer instability.

4.2 Detailed Jet Aerodynamics. The pulsed jet dynamics (Fig. 9) are first analyzed in isothermal conditions, with the help of phase-averaged PIV measurements. On these vector plots, the superimposed contours correspond to the presence of vortical structures identified with the Λ_2 criterion. What is remarkable here is the detachment of ring vortices surrounding the jet. These are generated by the mass flow modulation at the injector outlet. They develop at the burner front plate, detach when the injection velocity U_{inj} reaches its maximum (second plot of the sequence, at $\phi=2\pi/3$), and are transported by the flow. In the meantime, the internal recirculation zone performs a back-and-forth motion. This global dynamic has been previously observed by Giuliani et al. [12] on an air-blast injector excited by the same siren. The advection velocity of these structures is usually about half of the injection velocity U_{mean} at isothermal conditions [17]. This fact was also observed during the present experiments.

The same experiment with flame is shown in Fig. 10. Due to the burning of the seeding, the valid velocity measurements take place in the flame periphery. The velocity vectors placed along the flame expansion cone indicate a redundant bottleneck shape, at the shoulder of which, a pair of clockwise (line contours on the vorticity maps, unit is 2π rad/s) and anticlockwise (filled contours) ring vortices are evolving—one situated upstream outside the jet envelope, following one situated downstream inside the jet envelope. By comparison with the schlieren visualization from Fig. 8, the flame front is sandwiched between the vortex pair and rolls up around the outer vortex. These vortices seem to act like “roller bearings” placed around the flame instability.

Applying the process of Fig. 6, one derives with LV not only the ρ' phase-defined values, but also the turbulent flame front dynamics (Fig. 11). The description of the mushroomlike roll-up motion is remarkable and observable within the depth of the flame, where both schlieren and DEHS-seeded PIV techniques are unable to show. The highest $|\rho'|$ values are achieved in the flame cone. When setting a constant flame depth $L=100$ mm, the estimated density fluctuation measurements are consistent with the ones found in the literature [2–4]. For similar turbulent flames (premixed air methane and atmospheric pressure), the greatest order of magnitude for ρ' is 0.1 kg/m³. The zone where the mushroom-shaped flame appears consists of a negative density fluctuation (line contours on the plots), due to the local sudden

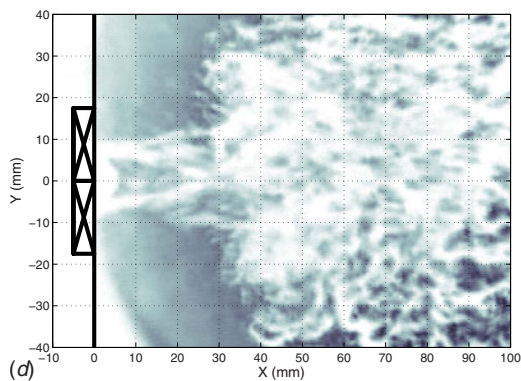
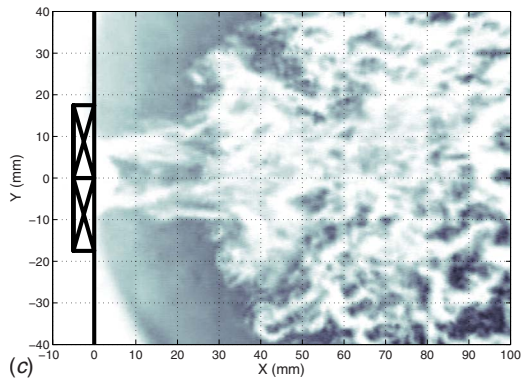
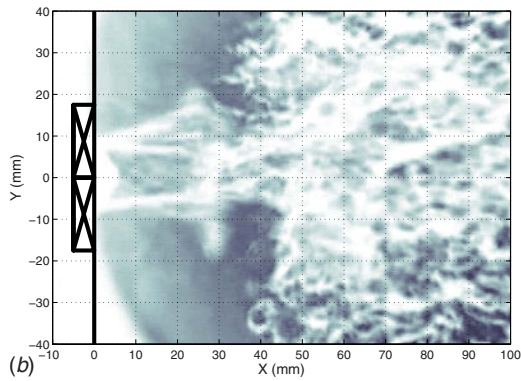
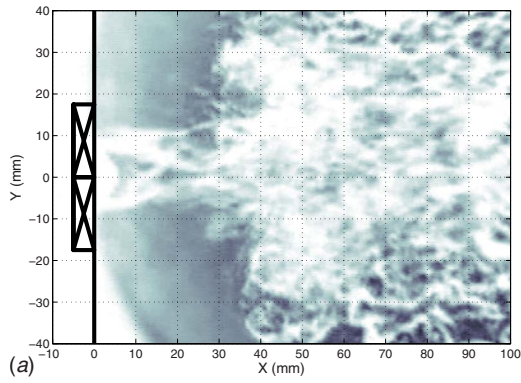


Fig. 8 Recurrent mushroom-shaped flame observed with schlieren technique. The field of view is defined in Fig. 1. Siren pulsation frequency 175 Hz. Four phase-locked subperiods are displayed with incremental time step $\tau/6=1.9$ m s from top to bottom.

heat release. This zone is pushed forward by the incoming of a positive density fluctuation zone (filled contours) that coincides in time and space with the outer vortex ring from Fig. 10. The nega-

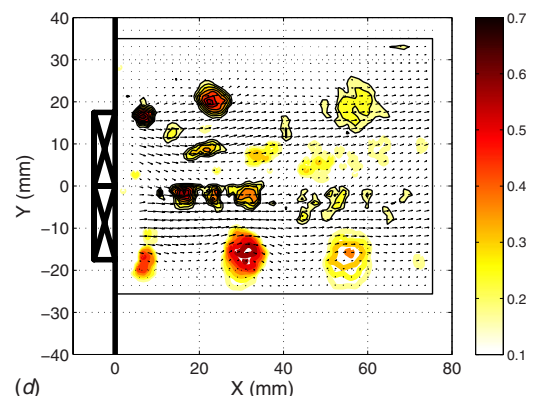
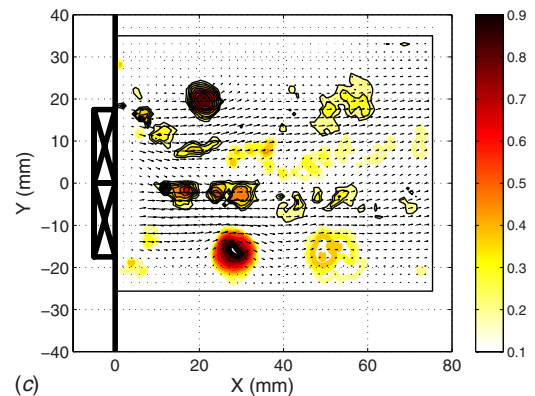
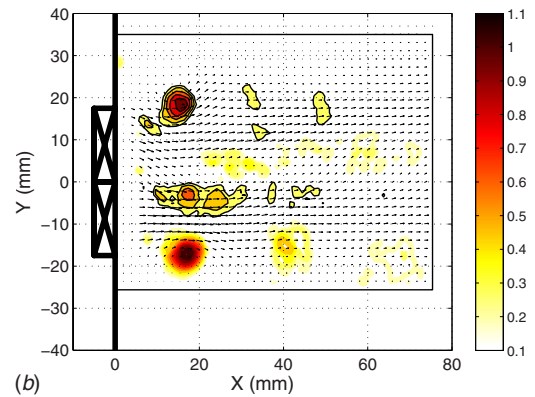
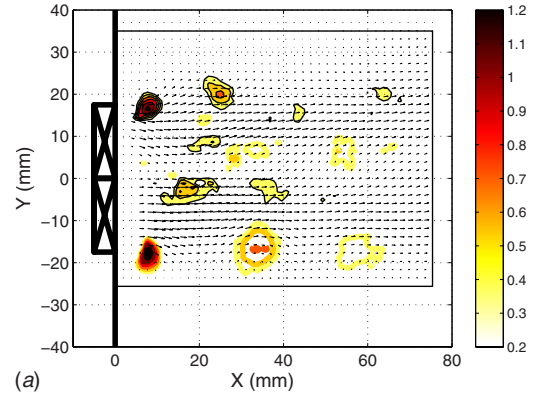


Fig. 9 Pulsed jet description at isothermal conditions. The measurement plane is defined in Figs. 1 and 7. Siren pulsation frequency 175 Hz. Vectors: PIV measurements in the jet median plane. Contour plots represent the vorticity (unit is 2π rad/s). Filled contours: anticlockwise rotation. Line contours: clockwise motion.

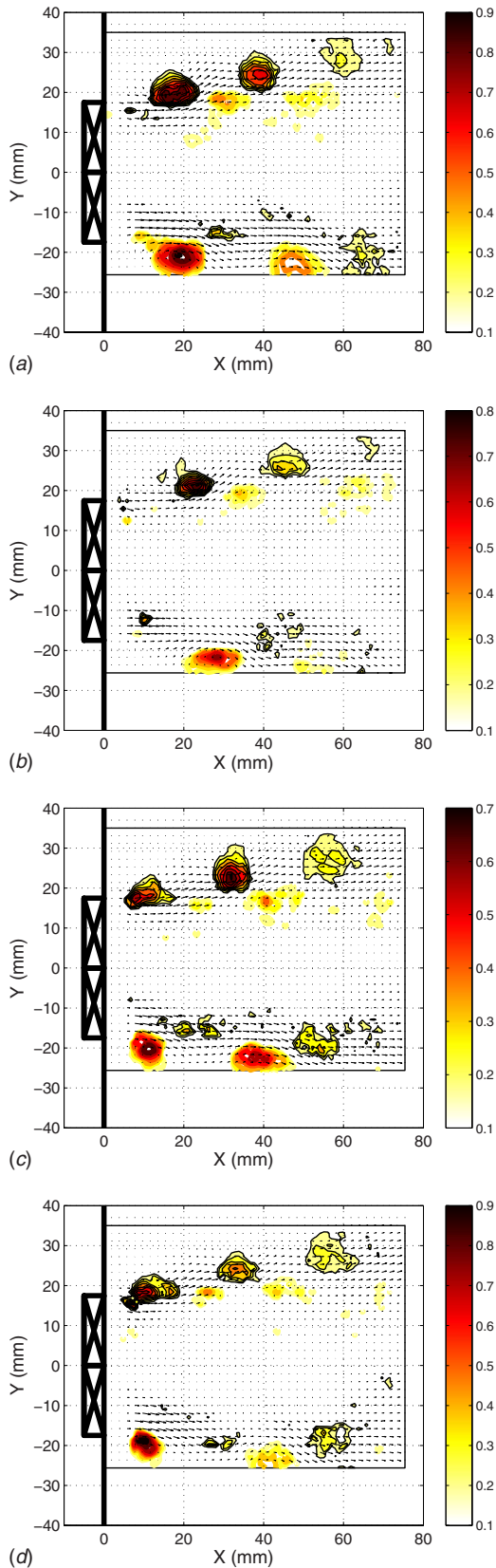


Fig. 10 Pulsed flame periphery description. The measurement plane is defined in Figs. 1 and 7. Siren pulsation frequency 175 Hz. Vectors: PIV measurements in the flame median plane. Filled contours: anticlockwise rotation. Line contours: clockwise rotation. The plots are synchronous with the ones from Fig. 8.

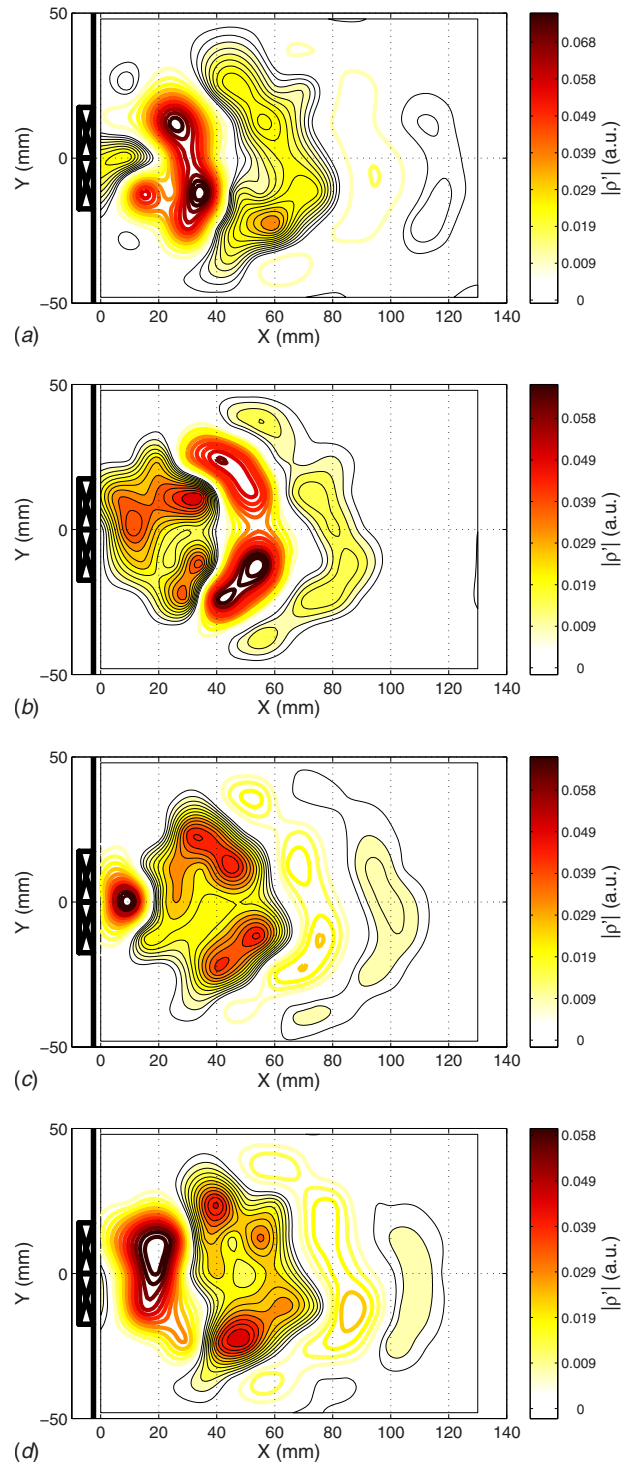


Fig. 11 Density fluctuation maps covering one pulsation cycle, based on phase-averaged LV signals, performed on a pulsed flame (free jet, 175 Hz). Relative density fluctuation $\rho'_{LS}(d\tau)$ (line-of-sight, computed with Eq. (2)). Filled contours: positive ρ' values. Line contours: negative ρ' values. Contour color: amplitude of density fluctuation $|\rho'|$. The plots are synchronous with the ones from Figs. 8 and 10.

tive ρ' zone expands around the positive ρ' zone (roll-up and closing of the flame), while the positive density zone drops in intensity. The further down the flame, the more the fluctuations are damped or dissolved. Nevertheless they remain visible up to the end of the measurement area.

5 Discussion

LV performs the fast recognition of resonant modes in a turbulent flame due to its high sensitivity, as well as it can map the density fluctuations in line-of-sight (Abel transform allowed if the jet geometry permits it). This method is lightweight in terms of instrumentation and data quantity, and also relatively low cost. It is nonintrusive, requires no seeding, and is extremely sensitive. The line-of-sight measurement provides information on the very core of the flame, and not just a side information as any surface sensor would do. Therefore, a single LV can deliver a proper reference signal for time-triggered instruments. LV provides information with a satisfactory signal-to-noise ratio from naturally resonant flames, as observed on measurements performed without activating the siren, in free-jet and enclosed jet configurations.

Current investigations at TU Graz focused on the quantification aspects, based on a precise measurement of the flame depth L , as well as a detailed analysis of the flame-vortex interaction. In the near future, an industrial natural gas burner, as well as a LPP module (premixed air-kerosene mixture), shall be tested with LV. From the fundamental research point of view, real time determination of the local Rayleigh criterion is also an important issue.

6 Conclusion

A proof of concept on the ability of LV to perform advanced turbulent combustion analysis was presented. The test case was a resonant air-methane premixed flame, where the aerodynamics and flame dynamics are known and described. The capacity of LV was established to detect the main modes of resonant flame dynamics, to determine the density fluctuations ρ' , and to plot in time and space the dynamics of the instability. The potential and relative simplicity of this technique for flame coherent turbulence analysis are emphasized. State-of-the-art LV is recommended as a test bench method at ambient conditions for stability qualification of natural gas burners.

Acknowledgment

The present work was supported by the European Commission as part of the Integrated Project "New Aeroengine Core Concepts" (NEWAC, Grant No. AIP5-CT-2006-030876), which is gratefully acknowledged.

Thanks to ONERA and PIEZOCRYST for providing us with hardware and instrumentation specific to this study.

Special thanks to Barbara Fuchs, Franz Heitmeir, Cornelia Santner, Bernhard Wagner, and Daniela Winkler.

Nomenclature

Main Scripts

f	=	frequency (Hz)
F	=	Fourier transform
G	=	Gladstone constant (m^3/kg)
k	=	LV processor gain factor (mm V/s)
L	=	penetration length of the laser through the medium where the density fluctuates (m)
\dot{M}	=	mass flow rate (kg/s)
S	=	swirl number

T = temperature (K)

t = time (s)

P = pressure (bars or Pa)

U = axial velocity component (m/s)

u' = output voltage signal of the LV (V)

X, Y, Z = Cartesian coordinates, as defined in Fig. 7 (m)

Δ = variation

Λ_2 = second invariant criterion for vortex identification

Φ = equivalence ratio

ϕ = phase angle (rad)

ρ' = density fluctuation (kg/m^3)

τ = oscillation period (s)

Subscripts

f = at frequency f

mean = averaged at the burner outlet

inj = at injector level

LS = line-of-sight

' = fluctuating (periodic) quantity

References

- [1] Dowling, A., and Morgans, A., 2005, "Feedback Control of Combustion Oscillations," *Annu. Rev. Fluid Mech.*, **37**, pp. 151–182.
- [2] Gouldin, F. C., and Dandekar, K. V., 1984, "Time-Resolved Density Measurements in Premixed Turbulent Flames," *AIAA J.*, **22**(5), pp. 655–663.
- [3] Bill, R. G. J., 1985, "Flame Structures in the Near Wake of Circular Cylinders," *AIAA J.*, **23**(12), pp. 1872–1876.
- [4] Namazian, M., Shepherd, I., and Talbot, L., 1986, "Characterisation of the Density Fluctuations in Turbulent V-Shaped Premixed Flames," *Combust. Flame*, **64**, pp. 299–308.
- [5] Lewin, A., 1999, "Compact Laser Vibrometer for Industrial and Medical Applications," *Proc. SPIE*, E. P. Tomasini, ed., **3411**, pp. 61–67.
- [6] Mayrhofer, N., and Woisetschlager, J., 2001, "Frequency Analysis of Turbulent Compressible Flows by Laser Vibrometry," *Exp. Fluids*, **31**, pp. 153–161.
- [7] Hampel, B., and Woisetschlager, J., 2006, "Frequency and Space Resolved Measurement of Local Density Fluctuations in Air by Laser Vibrometry," *Meas. Sci. Technol.*, **17**, pp. 2835–2843.
- [8] Putnam, A., and Dennis, W., 1953, "A Study of Burner Oscillations of the Organ-Pipe Type," *Trans. ASME*, **75**, pp. 15–28.
- [9] Yu, K. H., Troune, A., and Daily, J. W., 1991, "Low-Frequency Pressure Oscillations in a Model Ramjet Combustor," *J. Fluid Mech.*, **232**, pp. 47–72.
- [10] Giuliani, F., Wagner, B., Woisetschlager, J., and Heitmeir, F., 2006, "Laser Vibrometry for Real-Time Combustion Stability Diagnostic," *ASME Paper No. GT2006-90413*.
- [11] Beér, J., and Chigier, N., 1972, *Combustion Aerodynamics*, Applied Science, London, UK, pp. 100–146.
- [12] Giuliani, F., Gajan, P., Diers, O., and Ledoux, M., 2002, "Influence of Pulsed Entries on a Spray Generated by an Air-Blast Injection Device—An Experimental Analysis on Combustion Instability Processes in Aeroengines," *Proc. Combust. Inst.*, **29**, pp. 91–98.
- [13] Giuliani, F., Schrickler, A., Lang, A., Leitgeb, T., and Heitmeir, F., 2007, "High-Temperature Resistant Pressure Transducer for Monitoring of Gas Turbine Combustion Stability," *Proceedings of the 18th ISABE Conference, International Symposium on Air Breathing Engines*, Beijing, China, Paper No. ISABE-2007-1111.
- [14] Gardiner, Jr., W. C., Hidaka, Y., and Tanzawa, T., 1981, "Refractivity of Combustion Gases," *Combust. Flame*, **40**, pp. 213–219.
- [15] von Karman Institute, 1994, *Measurement Techniques in Fluid Dynamics—An Introduction* (Lecture Series of the von Karman Institute), Rhode-St-Genève, VKI LS 1994-01.
- [16] Jeong, J., and Hussain, F., 1995, "On the Identification of a Vortex," *J. Fluid Mech.*, **285**, pp. 69–94.
- [17] Panda, J., and McLaughlin, K., 1990, "Experiments on the Instabilities in a Free Swirling Jet," *Proceedings of the 28th Aerospace Sciences Meeting*, AIAA Paper No. 1990-0506.

Gang Xu

Beijing Key Laboratory of Energy Safety and
Clean Utilization,
Key Laboratory of Condition Monitoring and
Control for Power Plant Equipment of Ministry of
Education,
School of Energy and Power Engineering,
North China Electric Power University,
Beijing 102206, China

HongGuang Jin

Institute of Engineering Thermophysics,
Chinese Academy of Sciences,
Beijing 100080, China
e-mail: hgjin@mail.etp.ac.cn

YongPing Yang

Liqiang Duan

Beijing Key Laboratory of Energy Safety and
Clean Utilization,
Key Laboratory of Condition Monitoring and
Control for Power Plant Equipment of Ministry of
Education,
School of Energy and Power Engineering,
North China Electric Power University,
Beijing 102206, China

Wei Han

Lin Gao

Institute of Engineering Thermophysics,
Chinese Academy of Sciences,
Beijing 100080, China

A Novel Coal-Based Hydrogen Production System With Low CO₂ Emissions

In this paper, we have proposed a novel coal-based hydrogen production system with low CO₂ emission. In this novel system, a pressure swing adsorption H₂ production process and a CO₂ cryogenic capture process are well integrated to gain comprehensive performance. In particular, through sequential connection between the pressure swing absorption (PSA) H₂ production process and the CO₂ capture unit, the CO₂ concentration of PSA purge gas that enters the CO₂ capture unit can reach as high as 70%, which results in as much as 90% of CO₂ to be separated from mixed gas as liquid at a temperature of -55°C. This will reduce the quantity and quality of cold energy required for cryogenic separation method, and the solidification of CO₂ is avoided. The adoption of cryogenic energy to capture CO₂ enables direct production of liquid CO₂ at low pressure and thereby saves a lot of compression energy. Besides, partial recycle of the tail gas from CO₂ recovery unit to PSA inlet can help enhance the amount of hydrogen product and lower the energy consumption for H₂ production. As a result, the energy consumption for the new system's hydrogen production is only 196.8 GJ/tH₂ with 94% of CO₂ captured, which is 9.2% lower than that of the coal-based hydrogen production system with Selexol CO₂ removal process and is only 2.6% more than that of the coal-based hydrogen production system without CO₂ recovery. More so, the energy consumption of CO₂ recovery is expected to be reduced by 20–60% compared with that of traditional CO₂ separation processes. Further analysis on the novel system indicates that synergetic integration of the H₂ production process and cryogenic CO₂ recovery unit, along with the synthetic utilization of energy, plays a significant role in lowering energy penalty for CO₂ separation and liquefaction. The promising results obtained here provide a new approach for CO₂ removal with low energy penalty. [DOI: 10.1115/1.3159369]

1 Introduction

The increasing concern on global warming makes the CO₂ control technology urgent in recent years [1]. CO₂ capture and storage (CCS) is a technically feasible method of making deep reductions in carbon dioxide emissions from fossil energy utilization systems. At present, there are three basic technologies for capturing CO₂ from energy systems: postcombustion capture, oxy-fuel combustion capture, and precombustion capture [1]. For postcombustion capture, the CO₂ is separated from flue gases. Since the CO₂ in flue gases was diluted by nitrogen and other gases, its concentration is always less than 21%. However, the resulting amount of gases to be dealt with is rather large along with extremely high energy penalty for removing CO₂. Generally speaking, employing this technology, the efficiency of the power generation system with CCS will drop 9–15% points [2–5]. In oxy-fuel combustion capture, nearly pure oxygen is used for combustion instead of air. The result would show that the products of oxy-fuel combustion are mainly CO₂ and H₂O, from which CO₂ can be separated conveniently through condensation. However, in order to acquire pure

oxygen, a significant amount of energy will be consumed by the air separation unit (ASU). As a whole, the efficiency of the CO₂ abatement system using this technology will drop 8–12% points [4–7]. In precombustion capture, fossil fuels are first converted to fuel gases, which are mainly CO₂ and H₂, through coal gasification (or NG reforming) and shift reaction. After that, CO₂ is separated through absorption or membrane separation process, resulting in a hydrogen-rich fuel, which can be sent to the combustion chamber of the power generation unit. For precombustion capture technology, the energy consumption for CO₂ capture is relatively lower resulting from CO₂ separation before combustion instead of being diluted by N₂. The efficiency of the CO₂ capture system adopting this technology will drop 5–8% points [7–10]. Obviously, most energy systems have an overall thermal efficiency penalty with nearly 5–15% points for CO₂ capture. Thus, how to reduce the efficiency penalty is of importance for the future research and development of technologies for CO₂ capture, which is addressed in this paper.

Most efforts for capturing CO₂ from an energy system are focused on system integration and CO₂ separation process. For system integration, the major efforts are to integrate the CO₂ capture process with the whole energy system so that the energy used for CO₂ capture can be best supplied by suitable chemical or thermal energy within the system. Jin et al. [11] proposed the integration principles for energy systems with CO₂ control including cascade

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 22, 2009; final manuscript received April 5, 2009; published online December 2, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Gas Turbine Technical Congress and Exposition, Orlando, FL, June 8–12, 2009.

utilization of both chemical energy and thermal energy, step-by-step conversion of effective components and integration of energy utilization, and pollutant control. Chiesa et al. [9] proposed a series of hydrogen-electricity polygeneration systems with CO₂ capture and made comparative analysis of these systems. Besides, various clean fuel (such as hydrogen, methanol, and dimethyl ether (DME)) production systems with CCS were proposed in recent years, including methanol-electricity polygeneration system with CO₂ capture [12], multifunctional energy system with CO₂ removal [13], and coal-based DME production system with CO₂ capture [14]. These efforts conclude that the CO₂ capture process should be integrated with the energy conversion system, which has been proved to be effective in the study of clean fuel production systems with CO₂ capture.

As for the CO₂ separation process, chemical and physical absorption have attracted most attentions. The chemical absorption (e.g., amine and hot K₂CO₃ aqueous solution) is suitable for removing CO₂ at low concentrations, therefore, it is widely used in flue gas treatment [2–5]. For the physical absorption (e.g., Selexol solvent), since it has good absorption property at high pressure from the viewpoint of energy consumption, it is favorable for removal of CO₂ before combustion [7–10]. Recently, with the development of CCS technologies, other CO₂ separation methods, such as the cryogenic separation of CO₂, are attracting more and more attention. Several authors have conducted the studies of recovery of CO₂ by cold energy. Staicovici [15] proposed a zero CO₂ emission power plant, called “COOLENERG,” which uses advanced absorption refrigeration to capture CO₂. Deng et al. [16] proposed a novel gas turbine (GT) cycle with integration of CO₂ recovery and liquefied natural gas (LNG) cryogenic exergy utilization. Zhang and Lior [17] were devoted to the research of LNG fueled quasicombined power plant with a supercritical CO₂ Rankine-like cycle and a CO₂ Brayton cycle. Wang et al. [18] proposed a new type of integrated gasification combined cycle (IGCC) system with CO₂ recovery, which employed the cryogenic energy of the air separation unit. As a whole, recovering CO₂ by using cryogenic energy has several advantages. (1) It enables direct production of liquid CO₂ at low pressure so that the liquid CO₂ can be stored or sequestered via liquid pumping instead of compression of gaseous CO₂ to very high pressure, thereby saving on compression energy. (2) Chemical absorbent is not required so the cost on operating is low. However, challenges also exist for the cryogenic separation method, which are (a) large amount of energy consumption for refrigeration that will lead to high energy penalty of CO₂ capture and (b) CO₂ solidification under low temperature, which will bring severe operational problems such as the blockage of pipes or heaters.

The aims of this paper are as follows: (1) to propose a novel coal-based hydrogen production (CHP) system to capture CO₂ with lower energy penalty, (2) to introduce a new idea of integrating the hydrogen production process with CO₂ capture process so that the energy consumption for hydrogen product and CO₂ capture are both decreased, and (3) to seek for an approach to improve the performance of the cryogenic separation of CO₂.

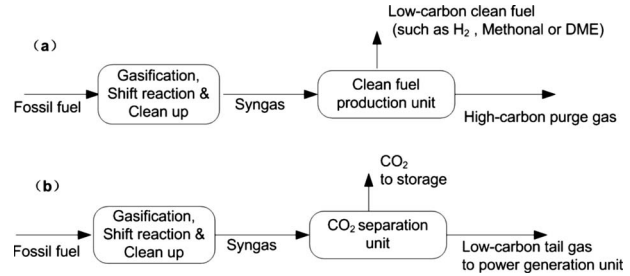


Fig. 1 Flow diagram of single production processes

2 Novel Hydrogen Production System With CO₂ Recovery

2.1 Concept of System Integration for CO₂ Recovery and Clean Energy Production.

In the clean fuel production process, syngas (mainly of H₂, CO, and CO₂) produced from fossil fuels can be converted into the low carbon clean fuels, such as hydrogen, methanol, or DME as is shown in Fig. 1(a). In the process, with more hydrogen transferred into clean fuels, the CO₂ concentration of purge gas is noticeably increased. While in the CO₂ capture process most of the CO₂ can be separated from syngas, the remaining low carbon tail gas will be rich in hydrogen, which can be seen from Fig. 1(b).

Therefore, it can be found that the two processes are strongly complementary, which are as follows: (1) In clean fuel production process, the purge gas has high CO₂ concentration and is easy to achieve CO₂ separation with low energy penalty. While in the CO₂ separation process, the tail gas has high hydrogen concentration and is suitable to produce low carbon clean fuel. Therefore, the two processes have the potential to enhance the energy utilization efficiency through integration. (2) There are similar units in the two processes, such as the syngas preparation unit and shift reaction unit, so the integration of the two processes is possible.

Based on the above analysis, we proposed a new integrated process of clean fuel production and CO₂ separation, which is conceptually represented in Fig. 2. Compared with the single process, the new integrated process applies the sequential connection between the clean fuel production unit and the CO₂ separation unit, so that when H₂ or other clean fuels are obtained in the clean fuel production unit, the CO₂ concentration of the purge gas entering separation unit will be increased effectively. As a result, the energy penalty for CO₂ separation can be decreased significantly. On the other hand, part of the low carbon tail gas discharged from the CO₂ separation unit is recycled to the entrance of clean fuel production unit, which will increase the yield of clean fuel and can also reduce the energy consumption for clean fuel production to some degree. Therefore, in the integrated process, the performance of the clean fuel production unit and the CO₂ separation unit can promote mutually and eventually lower the total energy consumption of the whole process.

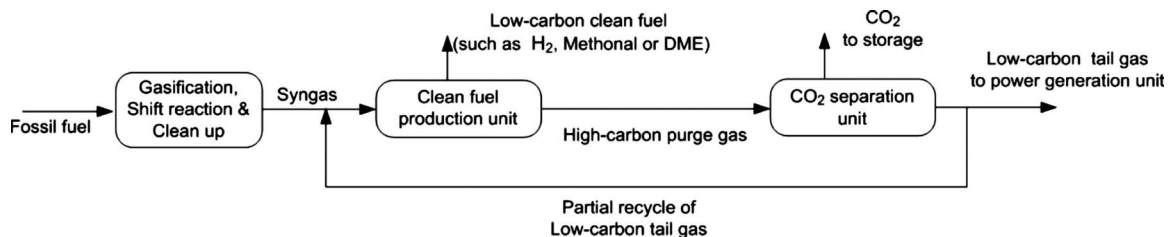


Fig. 2 Integrated process of clean fuel production and CO₂ separation

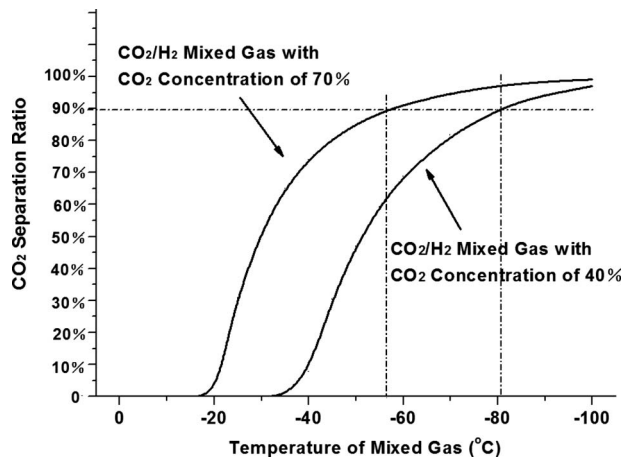


Fig. 3 Variation in CO₂ and separation ratio of CO₂/H₂ mixed gas with its temperature

2.2 Approach to Improve the Performance of CO₂ Cryogenic Separation Method. It can be seen from the temperature dropping characteristics of CO₂/H₂ mixed gas (shown in Fig. 3) that the concentration of CO₂ in the mixture is the key factor to solving the main problems (i.e., high energy consumption for refrigeration and CO₂ solidification at low temperature) encountered by cryogenic separation method. There are two CO₂ concentrations (i.e., 70% and 40%) investigated in Fig. 3, which are typical CO₂ concentrations with and without H₂ recovery prior to the CO₂ capture unit. As is shown in Fig. 3, under the pressure of 30 bars, when the CO₂ concentration of mixed gas is 40%, the temperature must be dropped to about -80°C to separate 90% of CO₂ from mixed gas. If the concentration of CO₂ in mixed gas is up to 70%, then 90% of CO₂ can be separated from mixed gas at the temperature of -55°C. Such temperature is above the triple point of CO₂, resulting in no CO₂ solidification. In addition, the cryogenic energy within the range of 0--55°C (which can be obtained by conventional single-stage or two-stage absorption refrigeration cycle with a relatively high energy efficiency) is much easier to be obtained than the cryogenic energy below -55°C (which should be prepared by more complicated refrigeration cycles with lower energy efficiency). Thus, it will be helpful to lower the energy consumption for refrigeration [19].

Therefore, the above analysis provides an important approach for improving the cryogenic separation method. That is, if the concentration of CO₂ in the mixture is increased as high as 70% or even higher before entering the separation process, the solidification of CO₂ can be avoided and the consumption of cryogenic energy, especially cryogenic energy of very low temperature can

be decreased. In a word, if the concentration of CO₂ in the mixed gas can be increased effectively, the performance of the cryogenic separation method could be improved significantly.

Obviously, in the new integrated process of clean fuel production and CO₂ separation (shown in Fig. 2), the CO₂ concentration of the purge gas from the clean fuel production unit is noticeably increased because a large amount of hydrogen have been transferred into clean fuels. If the CO₂ concentration of that purge gas can reach about 70% or even higher, the cryogenic separation method will be a suitable method to recover CO₂ from mixed gas. In this way, on the one hand, the problem of high energy consumption for refrigeration can be relieved and the emergence of solid CO₂ will be avoided. Alternatively, the advantages of the cryogenic separation method, such as producing liquid CO₂ at low pressure directly can help improve the comprehensive performance of the whole process.

2.3 Description of Reference Systems. In order to make a comparison for the performance analysis among different systems, a traditional CHP system without CO₂ capture and a CHP system with conventional Selexol CO₂ capture process are selected as the reference systems.

Figure 4 illustrates a flow sheet of the CHP system with Selexol CO₂ capture. The system adopts a Texaco O₂-blown and slurry-feed gasifier, with oxygen supplied by ASU; N₂ is directly discharged. After gasification, the crude syngas (which are mainly CO, CO₂, H₂, and H₂O) enters the shift unit in which most of the CO is converted to CO₂. Then, the clean shifted gas composed of H₂ and CO₂ is obtained after cleanup unit. In the CO₂ recovery unit, about 90% of CO₂ in shifted gas is separated and compressed to about 100 bars for storage and the remaining H₂-rich gas is sent to pressure swing absorption (PSA) unit for H₂ production. The purge gas from PSA is sent to the power generation unit as fuel.

For CHP system without CO₂ capture, there are no CO₂ separation unit and CO₂ compression unit, which are illustrated as dashed line frames in Fig. 4.

2.4 Description of the Novel CHP System. According to the above analysis, a novel CHP system with cryogenic CO₂ recovery is proposed. A conceptual configuration of the new CHP system is shown in Fig. 5. The syngas produced by a Texaco gasifier is quenched by water to lower its temperature from 1346°C to about 400°C. Then the saturated syngas is sent to a two-stage shift unit to convert 97% of CO to CO₂. The gas exit from the second stage of the shift reactor is heated to 380°C by the heat of the first stage shift reactor and then sent to an expand turbine. After that, the syngas is fed to clean up the unit to remove most of the sulfides. Leaving the cleaned up unit, the fresh gas (S4), mainly of H₂ (about 56%) and of CO₂ (about 42%), is sent to the pressure swing adsorption unit.

In the PSA unit, the high purity (>99.9 vol %) hydrogen product is extracted and the remaining purge gas in which the concen-

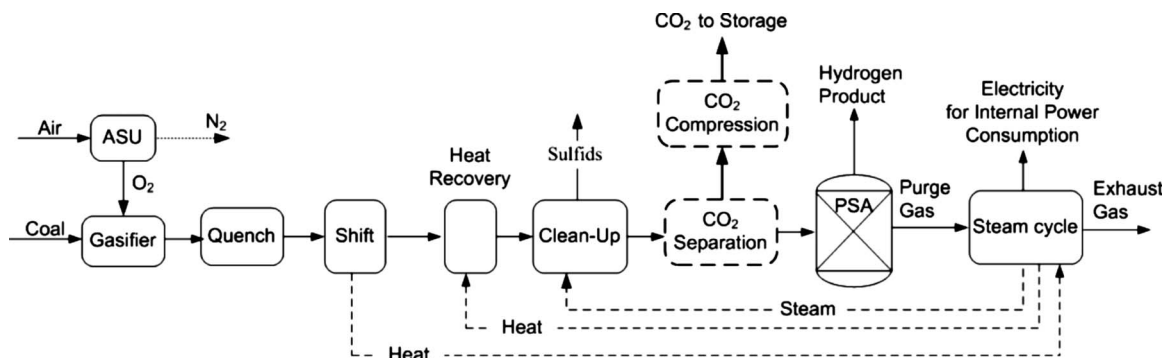


Fig. 4 CHP system with Selexol CO₂ capture process

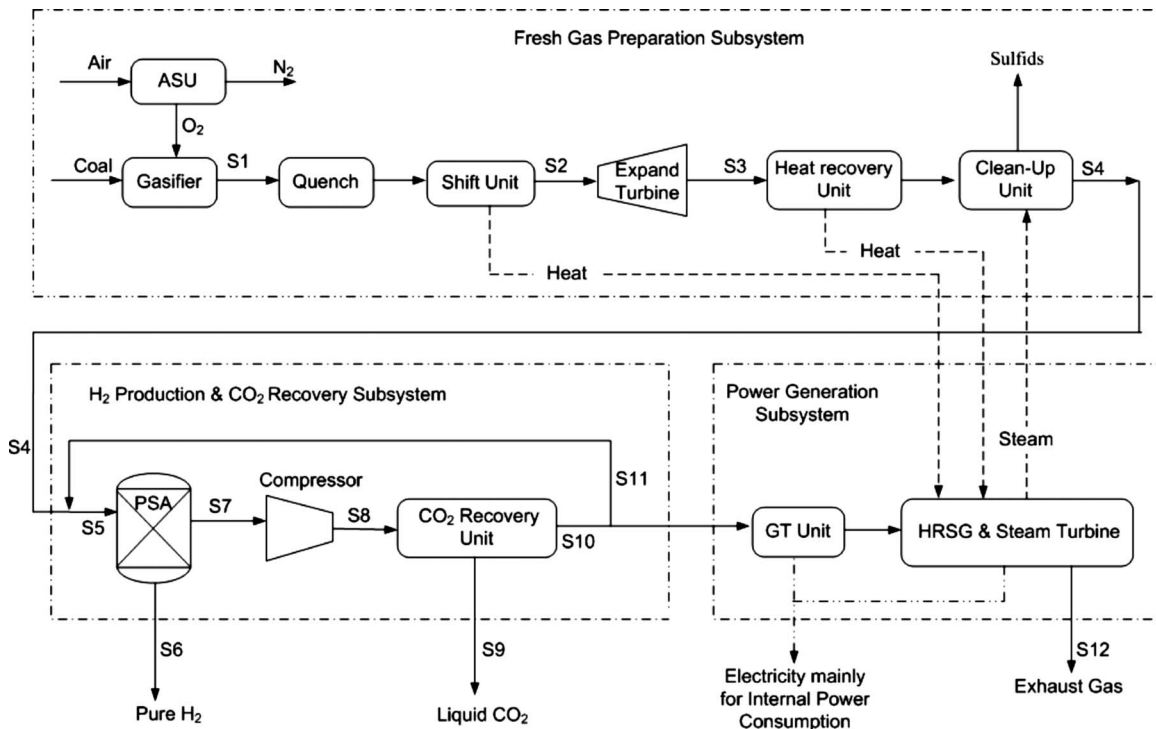


Fig. 5 New CHP system with cryogenic CO₂ recovery

tration of CO₂ is up to 70% approximately, is sent to CO₂ recovery unit after compressing. In the CO₂ recovery unit, the cryogenic energy is used to separate and liquefy CO₂ simultaneously. When entering the CO₂ recovery unit, the compressed PSA purge gas will be refrigerated by the cold energy generated by an absorption refrigeration system. Here, the absorption refrigeration cycle is adopted to take full advantage of the abundant heat within system. Because the concentration of CO₂ in the PSA purge gas is as high as 70% or even higher, once the temperature of the PSA purge gas drops to -55 °C, 90% of CO₂ will be separated as liquid from the mixed gas. The tail gas of the CO₂ recovery unit (S10) is still the mixture with main components of H₂, CO₂, and CO, with the concentration of H₂ and CO₂ at approximately 70% and 20%, respectively, part of which will be recycled to PSA to increase the quantities of hydrogen product and the captured CO₂, then the remaining tail gas of the CO₂ recovery unit will be sent to the combustion chamber of GT as fuel.

3 Evaluation of the New CHP System

In the novel CHP system, the power subsystem is based on a gas turbine of current technology with a turbine inlet temperature of 1200 °C and a pressure ratio of 15. The steam, which is saturated at a pressure of 98 bars and reheated at a pressure of 36 bars, is generated in the heat recovery steam generator (HRSG), and the temperatures of both the superheated and reheated steam are 540 °C. The total power output of the combined cycle (including gas turbine and steam turbine) is 59.3 MW. Datong coal was assumed as the basis for this study, whose lower heating value (LHV) is 26,710 kJ/kg. The coal analysis data, along with some other basic conditions, are given in Table 1.

The novel CHP system, along with two reference CHP systems, was studied by means of the commercial software ASPEN PLUS. The pressure, temperature, mole flow, and the composition of each stream corresponding to the points indicated in Fig. 3 are listed in Table 2. And the simulation results of the novel CHP system and the reference systems are presented in Table 3.

In Table 3, the overall efficiency is defined as the ratio of the total output (including net power output and hydrogen output) to

the coal input of the system. While energy consumption for hydrogen production represents the energy consumed per unit product, which can be obtained by the dividing total coal input by hydrogen output. Here, the net power output is near zero and its influence on energy consumption for hydrogen production is omitted. For all of the three systems, hydrogen is the main product and the power generated by system is primarily consumed by various equipments within the plant. For the CHP system without CO₂ capture, the internal power consumption mainly comes from ASU. For the CHP system with Selexol CO₂ capture, the internal power consumption increase greatly because of CO₂ separation and com-

Table 1 Basic conditions for simulation

Base coal analysis (wt %)	
C^f	68.54
O^f	6.85
S^f	1.08
W^f	8.84
H^f	3.97
N^f	0.74
Ash ^f	9.98
Gas turbine	
Parameter	Value
Turbine inlet temperature (°C)	1200
Pressure ratio	15
Isentropic efficiency of air compressor (%)	90
Isentropic efficiency of gas turbine (%)	87
Steam cycle	
Turbine inlet temperature (°C)	540
Turbine inlet pressure: high/middle/low pressure (bar)	98/36/6
Isentropic efficiency: high/middle/low pressure (%)	87/89/86
Pinch point in HRSG (°C)	15
Condensation pressure (bar)	0.07

Table 2 Parameters of main points of new CHP system

Streams	Temperature (°C)	Pressure (bar)	Mass flow (kg/s)	Mole frac (%)						
				N ₂	O ₂	CO	CO ₂	H ₂ O	H ₂	Others
S1	1346	35	73.9	0.9	0.0	44.0	11.7	12.6	30.5	0.3
S2	379.8	32.0	118.3	0.5	0.0	0.6	32.3	22.9	43.5	0.2
S3	275.2	11.5	118.3	0.5	0.0	0.6	32.3	22.9	43.5	0.2
S4	37.7	10.5	93.1	0.7	0.0	0.8	42.0	0.0	56.5	0.0
S5	38.0	10.0	104.6	1.5	0.0	1.8	39.2	0.0	57.6	0.0
S6	38.0	9.8	4.7	0.0	0.0	0.0	0.0	0.0	100.0	0.0
S7	38.0	1.2	99.9	2.7	0.0	3.1	68.9	0.0	25.3	0.0
S8	38.0	30.0	99.9	2.7	0.0	3.1	68.9	0.0	25.3	0.0
S9	8.8	100.0	82.0	0.2	0.0	0.3	98.9	0.0	0.6	0.0
S10	80.0	16.5	17.9	6.5	0.0	7.6	22.0	0.0	64.0	0.0
S11	80.0	16.5	11.5	6.5	0.0	7.6	22.0	0.0	64.0	0.0
S12	125.2	1.0	79.1	70.5	13.1	0.0	4.5	12.0	0.0	0.0

pression. As to the novel CHP system, the power required for syngas compressor leads to an increase in internal power consumption.

It can be seen from Table 3 that, for the traditional CHP system without CO₂ capture, the thermal efficiency is 62.5% with CO₂ emission rate of 18.05 tCO₂/tH₂. With the Selexol capture process, 89.2% of CO₂ can be captured but over 40 MW power consumed by CO₂ separation and compression, which leads to 7.1% points of efficiency penalty and 13% growth of the energy consumption for hydrogen production (from 191.8 GJ/tH₂ to 216.7 GJ/tH₂). Here, we can find that the CO₂ compression process has great influence on the overall efficiency penalty of CO₂ capture, whose energy consumption is 2.6 times of the energy consumed for CO₂ separation and accounts for over 70% of the total energy penalty of CO₂ capture.

For the new CHP system with cryogenic CO₂ recovery, when 94% of CO₂ is captured, its energy consumption for hydrogen production (196.8 GJ/tH₂) only increases by 2.6% than that of the CHP system without CO₂ recovery (191.8 GJ/tH₂), which is still 9.2% lower than that of the CHP system with Selexol CO₂ recovery (216.7 GJ/tH₂). This indicates that the performance of the new CHP system is superior to that of the CHP systems with

Selexol CO₂ recovery. Meanwhile, the aim of low energy penalty for CO₂ capture is achieved in the new CHP system.

To reveal the internal phenomena of the new CHP system, an exergy analysis is performed for both the new CHP system and the reference CHP system with Selexol CO₂ recovery. The results are listed in Table 4. The exergy analysis is also based on the assumption that the same quantity of coal was consumed in the new CHP system and the reference system.

As shown in Table 4, the exergy efficiency of the new system is 59.44%, which is 5.29% points higher than that of the reference system. Comparing the exergy distributions of the new CHP system with the reference system in Table 4, we find that the exergy of the hydrogen production has increased by 50.23 MW and the total exergy loss of the new CHP system is obviously decreased. The main reduction in exergy loss comes from the CO₂ recovery unit and power generation subsystem. And the detailed distribution of exergy loss of these two units is given in Table 4.

Compared with the reference CHP system with Selexol CO₂ recovery, the exergy loss of the CO₂ compression of the new CHP system is decreased remarkably (12.24 MW). The reason lies in that the new CHP system captures CO₂ directly as liquid by using cryogenic energy. In comparison, the exergy loss of the CO₂ sepa-

Table 3 Results for new CHP system and reference systems

	CHP system without CO ₂ capture	CHP system with Selexol CO ₂ capture	New CHP system
Coal input (MW)	927.5	927.5	927.5
Power generation (MW)			
Gas turbine	0.0	0.0	31.9
Steam turbine	48.8	80.4	27.5
Expand turbine	0.0	10.6	25.6
Subtotal	48.8	91.0	84.9
Internal power consumption (MW)			
ASU	45.7	45.7	45.7
Syngas compression	0.0	0.0	36.0
CO ₂ separation	0.0	11.1	0.0
CO ₂ compression	0.0	29.2	0.7
Other auxiliaries	2.9	4.8	2.4
Subtotal	48.7	90.9	84.8
Net power output (MW)	0.1	0.1	0.1
Hydrogen output (MW)	580.0	513.4	565.4
Hydrogen output (t/d)	417.7	369.8	407.2
CO ₂ recovery ratio (%)	0.0	89.2	94.0
CO ₂ emission rate (tCO ₂ /tH ₂)	18.05	2.21	1.11
Overall efficiency (%)	62.5	55.4	61.0
Energy consumption for hydrogen production (GJ/tH ₂)	191.8	216.7	196.8

Table 4 Exergy analysis of new CHP system and reference system

	New CHP system		CHP system with Selexol CO ₂ recovery	
	MW	%	MW	%
Exergy input of coal	947.76	100	947.76	100
Exergy output				
Hydrogen	563.30	59.43	513.07	54.14
Net electricity	0.09	0.01	0.15	0.02
Separated CO ₂	59.11	6.24	56.11	5.92
Exergy loss				
Gasifier	132.17	13.95	132.17	13.95
Air separating unit	28.04	2.96	28.04	2.96
Syngas cooling	45.40	4.79	48.26	5.09
Shift unit	15.00	1.58	18.50	1.95
Expand turbine	2.89	0.31	1.77	0.19
Heat recovery unit	12.01	1.27	8.10	0.85
Clean up unit	6.02	0.63	5.32	0.56
PSA unit	11.25	1.19	3.96	0.42
Syngas compressor	8.96	0.95	0.00	0.00
CO ₂ recovery unit:				
CO ₂ separation	22.08	2.33	16.16	1.71
CO ₂ compression	0.23	0.02	12.47	1.32
Subtotal	22.31	2.35	28.63	3.02
Power generation subsystem:				
Fuel combustion	20.24	2.14	55.83	5.89
GT compressor	1.69	0.18	0.00	0.00
GT turbine	2.53	0.27	0.00	0.00
Steam turbine	4.20	0.44	16.42	1.73
Heat exchangers	8.63	0.91	25.81	2.72
Subtotal	37.29	3.93	98.06	10.35
Exergy of exhaust streams	4.15	0.44	4.35	0.46
Exergy efficiency (%)	59.44		54.15	

ration of the new CHP system (including the exergy loss of refrigeration) is 5.92 MW higher than that of the reference system. Thus, the exergy saving on CO₂ compression is the main reason leading to the reduction in the exergy loss in CO₂ recovery unit.

Compared with the reference system, the exergy loss of the power generation subsystem of the new CHP system was decreased by 60.77 MW. The reason for such reduction may lie in two aspects. One is the significant reduction in the energy consumption for CO₂ recovery, which results in that the power generated by the combined cycle of the novel CHP system is lower than that of the reference system. The less the inner power consumption of the system is, the less the fuel required for power generation becomes, which will lead to the reduction in the exergy loss of power generation. The other is the cascade utilization of energy in power generation subsystem, such as better heat integration and using combined cycle instead of steam cycle. It is shown in Table 4 that, the exergy loss of various components of the power generation subsystem of the new CHP system, such as combustion chamber, steam turbine, and heat exchangers, are notably lower than that of the reference system, which leads to the reduction in the total exergy loss.

From the above analysis, we conclude that through synergetic integration of hydrogen production and CO₂ recovery and cascade utilization of energy, the key problem of high energy penalty for CO₂ capture may be improved in the novel CHP system and favorable thermal and environment performances can be achieved. However, some challenges still exist, such as the complexity of the system, as well as the possible high investment of the system, which will be further studied in our following work.

4 Discussion

4.1 Impact of Hydrogen Production Process on CO₂ Recovery. In PSA, when H₂ is extracted from the mixed gas, the CO₂ concentration in the mixed gas is increased significantly. As

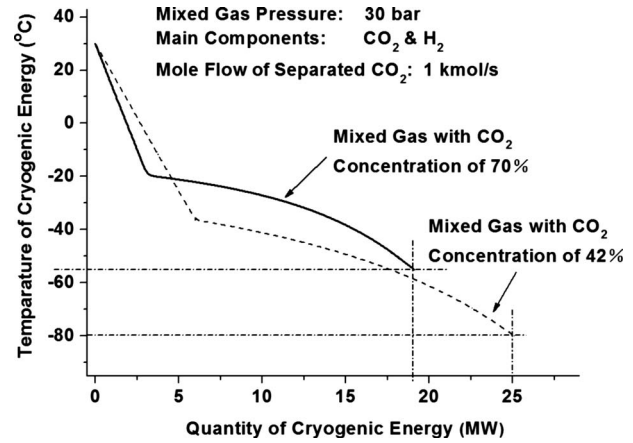


Fig. 6 Temperature and quantity of cryogenic energy required for separating 1 kmol/s CO₂ from mixed gases with different CO₂ concentrations

is shown in Fig. 5 and Table 2, the CO₂ concentration in the fresh gas from the cleaned up unit is 42%, while the concentration is up to about 70% in PSA purge gas. Figure 6 shows the cryogenic energy requirements for separating 1 kmol/s CO₂ from mixed gases with different CO₂ concentrations. In the figure, the pressure of the two kinds of mixed gases is set to be at 30 bars, the main components of which are CO₂ and H₂, and the recovery ratio of CO₂ is 90%. For the mixed gas with 70% of CO₂ concentration, the cryogenic energy required for mixed gas refrigeration is approximately 19.1 MW, which is 24% lower than that of the mixed gas with 42% CO₂ concentration. Besides, the cryogenic energy required for the former is totally above -55 °C, while nearly one-third of the cryogenic energy required for the latter is from -55 °C to -80 °C. It means that compared with the fresh gas from the cleaned up unit, to separate the same amount of CO₂ by cryogenic energy, the cold energy required for the PSA purge gas can be significantly decreased and easier to be obtained, thus leading to the energy consumption for CO₂ recovery being reduced noticeably.

The conceptual configuration of the CO₂ recovery unit is shown in Fig. 7, including four heat exchangers (H1–H4), two separators (Sep1 and Sep2), and two liquid pumps (P1 and P2). In the CO₂ recovery unit, the PSA purge gas, with temperature of 38 °C and pressure of 30 bars, is first cooled at -38 °C after passing through H1 and H2. Then some of the CO₂ is separated in Sep1 as liquid, and the rest of the mixed gas will be further cooled at -55 °C after passing through H3 and H4. In Sep2, another part of the liquid CO₂ is separated from the mixed gas to make the total CO₂ recovery ratio up to 90% approximately. The cryogenic energy possessed by the liquid CO₂ and the remaining H₂-rich tail gas is supplied for cooling the PSA purge gas through H3 and H1. And the cryogenic energy required by H2 and H4 are approximately 18.5 MW and 4.2 MW, respectively, which are provided by the refrigeration cycle. To utilize the cryogenic energy to a reasonable extent, the following measures are adopted in the CO₂ recovery process: (1) make full use of the cryogenic energy of the product streams, and (2) multistage CO₂ separation is applied to reduce the cryogenic energy requirement in the following heat transfer process.

The performance comparison of the cryogenic CO₂ recovery unit for the new CHP system with several typical CO₂ recovery processes are shown in Table 5. The CO₂ recovery process with ASU cryogenic energy is reported by Wang et al. [18], which proposes to integrate the CO₂ cryogenic recovery process with ASU and employ the cold energy of ASU to separate and liquefy CO₂ simultaneously. Mono-ethanolamine (MEA) absorption process is a widely used chemical absorption process to separate CO₂

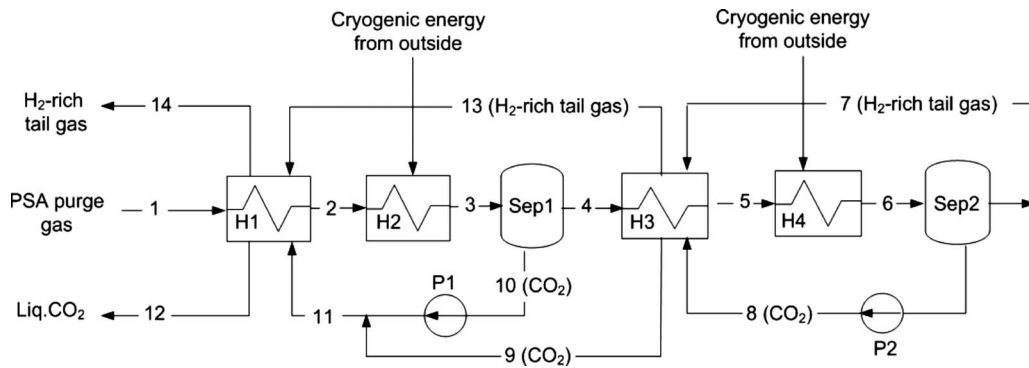


Fig. 7 Flow diagram of cryogenic CO₂ recovery process

from mixed gas [1,20]. MEA absorption process can separate CO₂ under the low CO₂ partial pressure conditions and reduce CO₂ concentration in the exit gas down to very low values. MEA solvent removes CO₂ from mixed gas by means of a chemical reaction, which can be reversed by pressure reduction and heating. The amount of heat consumed for the regeneration of MEA solvent, together with the power used for CO₂ compression, lead to a significant energy penalty. Selexol absorption process is a kind of physical absorption process, which is mostly applied to gas streams with a high CO₂ partial pressure [1,21]. Selexol absorption process has relatively low energy consumption, since the regeneration of Selexol solvent is achieved by the release of pressure without heating.

From Table 5, we can find that the unit energy consumption for CO₂ recovery in the new CHP system is only 0.395 MJ/kg CO₂, which is 37% less than that of the CO₂ recovery process with ASU cryogenic energy proposed in Ref. [18], and also 20% less than that of the conventional Selexol absorption process (0.497 MJ/kg CO₂). Therefore, PSA unit increases the concentration of the CO₂ in purge gas to make the cryogenic energy required for CO₂ recovery decrease significantly and be obtained easily, and eventually to achieve noticeable reduction in energy penalty for CO₂ recovery.

4.2 Impact of CO₂ Recovery Process on Hydrogen Production. In the new CHP system, part of the tail gas of CO₂ recovery unit is sent back to the PSA inlet (see Fig. 4). Here, we defined the recycle ratio as the mole flow rate of the recycled gas (S11 in Fig. 5) over the total tail gas of CO₂ recovery unit (S10 in Fig. 5). When the recycle ratio is positive, it means that part of the CO₂ recovery unit tail gas will enter PSA and affect the performance of the PSA unit. The larger the recycle ratio, the more CO₂ recovery unit tail gas will enter the PSA unit and the stronger the effect of CO₂ recovery unit on PSA.

Figure 8 illustrates the impact on H₂ yield and CO₂ recovery ratio imposed by varying recycle ratios. With the recycle ratio

increasing, the H₂ yield and CO₂ recovery ratio increase linearly as is shown in the curve. Here, the CO₂ recovery ratio represents the global CO₂ recovery ratio of the whole system, as is the same with the H₂ yield. When the recycle ratio increases, though the quantity of the fresh gas (S4 in Fig. 5) is constant, the quantities of the mixed gases passing through PSA and the CO₂ recovery unit (i.e., S5 and S8 in Fig. 5) will definitely increase. Since a large part of H₂ are extracted in PSA, the CO₂ concentration of the mixed gas entering the CO₂ recovery unit (S8 in Fig. 5) is kept at a high level, as is the same with the H₂ concentration of the mixed gas entering PSA. Thus the local recovery ratio of H₂ and CO₂ can also be kept at a high level. Eventually, when the tail gas is recycled, the quantities of the mixed gases passing through PSA and the CO₂ recovery unit will increase, together with the high level of the local recovery ratio of H₂ or CO₂, which results in the global CO₂ recovery ratio and the global H₂ yield significantly increased.

In the new CHP system, since most of the CO₂ in mixed gas is removed in the CO₂ recovery unit, CO₂ will not accumulate in the PSA unit. Thus, with the recycle ratio increasing, the gas flow rate of the mixed gas entering PSA increases slowly (shown in Fig. 9). At the same time, the H₂ concentration of recycled gas (S11 in Fig. 5) and the mixed gas (S5 in Fig. 5) are kept at a high level as well. As shown in Fig. 9, when the recycle ratio varies from 0 to 0.8, the mole flow rate of the mixed gas only increases 17% (from 3.6 kmol/s to 4.2 kmol/s), and the H₂ concentration of the recycled gas and the mixed gas is higher than that of the fresh gas (which is 56.5%).

Figure 10 shows the impact of the recycle ratio on equivalent energy consumption for H₂ production of the new system. With the recycle ratio increasing, the energy consumption for unit H₂ production is decreased to some extent, which means the design of feeding the CO₂ recovery unit tail gas to the PSA unit has a positive impact on reducing energy consumption for H₂ production.

Table 5 Performance comparison of typical CO₂ recovery processes

Items	CO ₂ recovery process in new CHP system	CO ₂ recovery process with ASU cryogenic energy [18]	MEA absorption process [20]	Selexol absorption process [21]
Energy penalty for CO ₂ recovery (MW)				
Refrigeration	20.59	21		
Extra compression work	12.61	47.8		
Absorbent regeneration			78.3	17.3
CO ₂ compression	0.47	0.6	35.7	67.7
Total	33.67	69.4	114	85
Quantity of CO ₂ captured (kg/s)	81.81	95.2	93.0	176.4
Energy penalty for recovering unit CO ₂ (MJ/kg CO ₂)	0.412	0.73	1.23	0.48

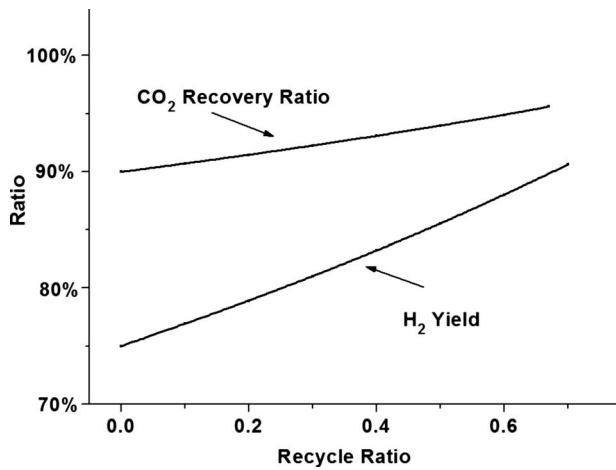


Fig. 8 Impact of recycle ratio on H₂ yield and CO₂ recovery ratio

5 Conclusion

In this paper, a novel coal-based hydrogen production system with cryogenic CO₂ capture has been proposed. Through the sequential connection between the H₂ production process and CO₂ recovery process, the utilization of cryogenic energy in CO₂ recovery and the partial recycle of the tail gas of the CO₂ recovery unit, the new system can produce the high purity hydrogen with

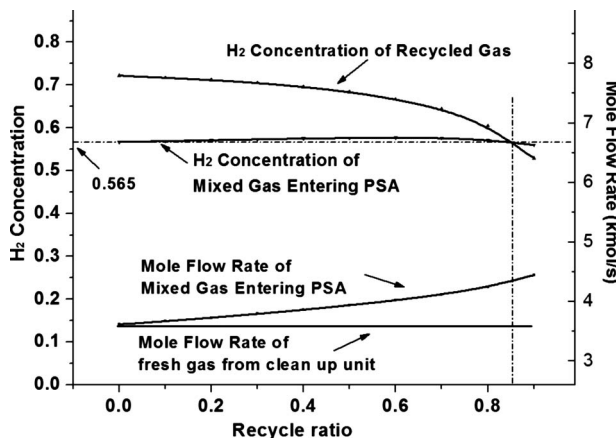


Fig. 9 Impact of recycle ratio on mole flow rate and H₂ concentration of key streams

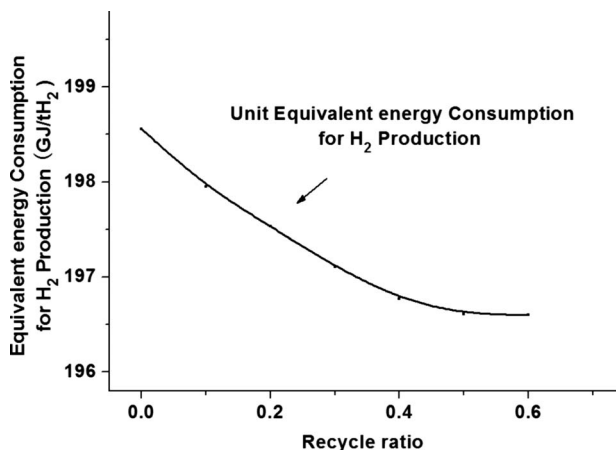


Fig. 10 Impact of recycle ratio on equivalent energy consumption for H₂ production

low energy consumption, and at the same time, concentrate the CO₂ in PSA purge gas to lower the energy penalty for CO₂ capture. The synergetic integration of clean fuel production and CO₂ capture leads to the breakthrough: With 94% of CO₂ captured, the equivalent energy consumption for hydrogen production of the new system is only 196.8 GJ/tH₂, which is only 2.6% more than that of CHP system without CO₂ recovery (191.8 GJ/tH₂). The high efficiency of the new system results from the synergetic integration of clean fuel production process and CO₂ recovery unit, and the synthetic utilization of energy.

Acknowledgment

This study was supported by the National Nature Science Foundation Project (Contract Nos. 50520140517, 50606010, and 50706051), the National Basic Research Program of China (Contract No. 2009CB219801), the Program for Changjiang Scholars and Innovative Research Team in University (PCSIRT0720), and the Chinese Support Project of New Century Excellent Talents in University (Contract No. NCET-05-0216).

References

- [1] Working Group III of the Intergovernmental Panel on Climate Change (IPCC), 2005, "IPCC Special Report on Carbon Dioxide Capture and Storage," Cambridge University Press, UK.
- [2] Parsons Infrastructure & Technology Group, 2002, "Updated Cost and Performance Estimates for Fossil Fuel Power Plants With CO₂ Removal," Report No. 1004483.
- [3] Rao, A. B., and Rubin, E. S., 2002, "A Technical, Economic, and Environmental Assessment of Amine Based CO₂ Capture Technology for Power Plant Greenhouse Gas Control," *Environ. Sci. Technol.*, **36**, pp. 4467–4475.
- [4] Singh, D., Croiset, E., Douglas, P. L., and Douglas, M. A., 2003, "Techno-Economic Study of CO₂ Capture From an Existing Coal-Fired Power Plant: MEA Scrubbing vs. O₂/CO₂ Recycle Combustion," *Energy Convers. Manage.*, **44**, pp. 3073–3091.
- [5] Alstom Power Inc., ABB Lummus Global Inc., Alstom Power Environmental Systems, and American Electric Power, 2001, "Engineering Feasibility and Economics of CO₂ Capture on an Existing Coal-Fired Power Plant," Report No. PPL-01-CT-09, http://www.netl.doe.gov/technologies/carbon_seq/Resources/Analysis/pubs/AlstomReport.pdf.
- [6] Dillon, D. J., Panesar, R. S., Wall, R. A., Allam, R. J., White, V., Gibbins, J., and Haines, M. R., 2005, "Oxy-Combustion Processes for CO₂ Capture From Advanced Supercritical PF and NGCC Power Plant," *Proceedings of the Seventh International Conference on Greenhouse Gas Control Technologies, Volume I: Peer Reviewed Papers and Overviews*, Oxford, UK, pp. 211–220.
- [7] Nsakala, N., Liljedahl, G., Marion, J., Bozzuto, C., Andrus, H., and Chamberland, R., 2003, "Greenhouse Gas Emissions Control by Oxygen Firing in Circulating Fluidised Bed Boilers," *Second Annual National Conference on Carbon Sequestration*, Alexandria, VA, May 5–8.
- [8] Rubin, E. S., and Rao, A. B., and Chen, C., 2005, "Comparative Assessments of Fossil Fuel Power Plants With CO₂ Capture and Storage," *Proceedings of the Seventh International Conference on Greenhouse Gas Control Technologies, Volume I: Peer Reviewed Papers and Overviews*, Oxford, UK, pp. 285–294.
- [9] Chiesa, P., Consonni, S., Kreutz, T., and Williams, R., 2005, "Co-Production of Hydrogen, Electricity, and CO₂ From Coal With Commercially Ready Technology. Part A: Performance and Emissions," *Int. J. Hydrogen Energy*, **30**(7), pp. 747–767.
- [10] Kreutz, T., Williams, R., Consonni, S., and Chiesa, P., 2005, "Co-Production of Hydrogen, Electricity, and CO₂ From Coal With Commercially Ready Technology. Part B: Economic Analysis," *Int. J. Hydrogen Energy*, **30**(7), pp. 769–784.
- [11] Jin, H., Zhang, X., and Gao, L., 2008, "Fundamental study of CO₂ Control Technologies and Policies in China," *Sci. China, Ser. E: Technol. Sci.*, **51**(5), pp. 1–14.
- [12] Jin, H., Gao, L., Han, W., and Yan, J., 2007, "A New Approach Integrating CO₂ Capture Into a Coal-Based Polygeneration System of Power and Liquid Fuel," *ASME Paper No. GT-2007-27678*.
- [13] Jin, H., Han, W., and Gao, L., 2007, "A Novel Multi-Functional Energy System (MES) for CO₂ Removal With Zero Energy Penalty," *Proceedings of ASME Turbo Expo*, Montreal, Canada, Paper No. GT-2007-27680.
- [14] Celik, F., Larson, E. D., and Williams, R. H., 2005, "Transportation Fuel From Coal With Low CO₂ Emissions," *Proceedings of the Seventh International Conference on Greenhouse Gas Control Technologies. Volume II: Papers, Posters and Panel Discussion*, Oxford, UK, pp. 1053–1058.
- [15] Staicovici, M. D., 2002, "Further Research Zero CO₂ Emission Power Production: The 'COOLENERG' Process," *Energy*, **27**(9), pp. 831–844.
- [16] Deng, S., Jin, H., Cai, R., and Lin, R., 2002, "Novel Gas Turbine Cycle With Integration of CO₂ Recovery and LNG Cryogenic Exergy Utilization," *Proceedings of ASME IMECE*, New Orleans, LA.

- [17] Zhang, N., and Lior, N., 2003, "A Novel Near-Zero CO₂ Emission Thermal Cycle With LNG Cryogenic Exergy Utilization," *Proceedings of ECOS*, Copenhagen, Denmark, pp. 1467–1478.
- [18] Wang, B., Jin, H., Han, W., and Zheng, D., 2004, "IGCC System With Integration of CO₂ Recovery and the Cryogenic Energy in Air Separation Unit," ASME Paper No. GT-2004-53723.
- [19] Wang, S. H., 2001, *Handbook Of Air Conditioning and Refrigeration*, McGraw-Hill, New York.
- [20] Parsons, E. L., Shelton, W. W., and Lyons, J. L., 2002, "Advanced Fossil Power Systems Comparison Study," report for NETL, <http://www.netl.doe.gov/publications/others/techrpts/AdvFossilPowerSysCompStudy.pdf>.
- [21] IEA GHG, 2003, "Potential for Improvements in Gasification Combined Cycle Power Generation With CO₂ Capture," Report No. PH4/19.

Advanced Braze Alloys for Fast Epitaxial High-Temperature Brazing of Single-Crystalline Nickel-Base Superalloys

Britta Laux¹

Fossil Power Generation Division,
Siemens AG,
45473 Mülheim, Germany
e-mail: b.laux@tu-bs.de

Sebastian Piegert

Siemens AG,
Energy Sector Products Gas Turbine Engineering,
Materials Applications and Joining,
45473 Mülheim, Germany
e-mail: sebastian.piegert@siemens.com

Joachim Rösler

Institut für Werkstoffe,
Technische Universität Braunschweig,
38106 Braunschweig, Germany
e-mail: j.roesler@tu-bs.de

High-temperature diffusion brazing is a very important technology for filling cracks in components from single-crystalline nickel-base superalloys as used in aircraft engines and stationary gas turbines: Alloys, which are similar to the base material, are enhanced by a fast diffusing melting-point depressant (MPD) like boron or silicon, which causes solidification by diffusing into the base material. Generally, epitaxial solidification of single-crystalline materials can be achieved by use of conventional braze alloys; however, very long hold times are necessary to provide a complete diffusion of the MPD out of the braze gap. If the temperature is lowered before diffusion is completed, brittle secondary phases precipitate, which serve as nucleation sites for stray grains and, therefore, lead to deteriorating mechanical properties. It was demonstrated in earlier works that nickel-manganese-based braze alloys are appropriate systems for the braze repair of particularly wide gaps in the range of more than 200 μm , which allow a significant shortening of the required hold times. This is caused by the complete solubility of manganese in nickel: Epitaxial solidification can be controlled by cooling in addition to diffusion. In this work, it will be shown that the nickel-manganese-based systems can be enhanced by chromium and aluminum, which is with regard to high-temperature applications, a very important aspect. Furthermore, it will be demonstrated that silicon, which could be identified as appropriate secondary MPD in recent works, can be replaced by titanium as this element has additionally a γ' stabilizing effect. Several braze alloys containing nickel, manganese, chromium, aluminum, and titanium will be presented. Previously, the influence of the above mentioned elements on the nickel-manganese-based systems will be visualized by thermodynamic simulations. Afterward, different compositions in combination with a heat treatment, which is typical for nickel-base superalloys, will be discussed: A microstructure, which is very similar to that within the base material, can be presented. [DOI: 10.1115/1.3159376]

1 Introduction

With regard to the high costs of components made from single-crystalline nickel-base superalloys, the repair of damaged parts is of outstanding economical interest. Filling cracks by high-temperature diffusion brazing is one of the most common repair technologies [1–3]. Usually, a braze system similar to the base material, which is enhanced by a fast diffusion melting-point depressant (MPD) like boron is applied to the cracks. During the brazing cycle an epitaxial solidification can be achieved by diffusing the MPD into the surrounding base material [4–11]. However, since the MPD has to be entirely diffused out of the braze gap to avoid the formation of secondary phases, very long hold times are required [12–16]. This is because the solubility of boron in nickel is very poor. In the case of insufficient hold times, the solidification terminates with an eutectic reaction, resulting in a polycrystalline microstructure with brittle phases in the center of the braze gap.

Even if small, fast diffusing elements are used, very long hold times are required. Therefore, the use of the conventional, boron-containing systems is predominantly applicable for the repair of narrow gaps up to 100 μm . Filling wider gaps with the diffusion controlled brazing technologies is economically not feasible.

To achieve a significant shortening of the required hold times, particularly for the repair of wider gaps in the range of 300 μm , new braze alloys have been developed for fast epitaxial high-temperature brazing. The approach is the use of nickel-manganese-based alloying systems, which are completely miscible. In this case, the formation of a second phase, serving as nucleation site for stray grains, is entirely avoided. The results of several brazing experiments by application of nickel-manganese based systems were demonstrated in Ref. [17]. The repair of wide gaps could be achieved within brazing times, which were up to 100 times shorter compared with conventional braze alloys. Furthermore, silicon could be identified as appropriate second MPD to reduce the amount of manganese, being necessary for an effective melting-point reduction.

Now, the nickel-manganese systems were enhanced by chromium and aluminum, since both are very important elements, acting as γ solid-solution hardeners, oxide-film formers (Al_2O_3 and Cr_2O_3), and γ' formers (Al) [18,19]. Therefore, with regard to an appropriate high-temperature stability and a sufficient oxidation and corrosion behavior, the influence of both elements has been investigated. Furthermore, the replacement of silicon by titanium was examined since the amount of γ' can be significantly increased by the addition of titanium. To identify appropriate compositions, thermodynamic simulations were conducted. Afterward, brazing experiments with four selected compositions were applied.

Several thermodynamic simulations were carried to study the influence of titanium, chromium, and aluminum on the nickel-

¹Corresponding author. Present address: Institut für Werkstoffe, Technische Universität Braunschweig, Langer Kamp 8, 38106 Braunschweig, Germany.

Manuscript received March 23, 2009; final manuscript received April 14, 2009; published online December 2, 2009. Review conducted by Dilip R. Ballal.

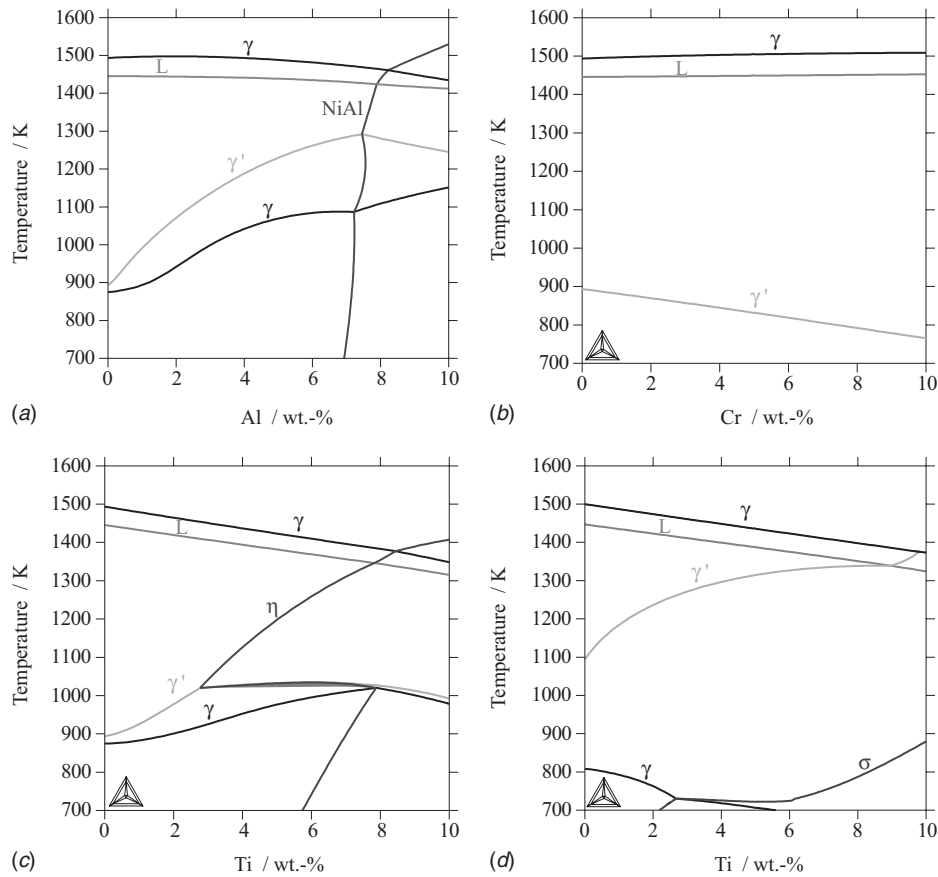


Fig. 1 Quasibinary phase diagrams with fixed fractions of manganese and varying fractions of aluminum, chromium, and titanium. In (d) the amount of manganese, chromium, and aluminum is fixed, while the amount of titanium is varied (THERMOCALC, Version TCR). (a) Ni-20Mn+Al, (b) Ni-20Mn+Cr, (c) Ni-20Mn+Ti, and (d) Ni-20Mn-5Cr-3Al+Ti.

manganese-based systems. Therefore, THERMOCALC, Version TCR, in combination with the database TTNi7, which was particularly developed for nickel-base superalloys, was used. As manganese is not an element, which is typically present in nickel-base superalloys, it is expected that the calculated phase transformation temperatures will not fit exactly with experimental data. However, in earlier works [20], the THERMOCALC simulations could be identified as a very useful tool to predict the general influence of different elements on the braze alloy nickel-manganese system, particularly, on the solidus, liquidus, and γ' solvus temperatures.

The simulation results are depicted in Fig. 1 as quasibinary phase diagrams. Differing from conventional phase diagrams, in this case, the plotted phase boundaries mark the appearance and disappearance of the phases, which form with decreasing temperatures. For example, the liquidus line is labeled by the precipitation of the γ phase (in the case of high amounts of aluminum or rather titanium, the first solid phases are NiAl and the η phase, respectively (Figs. 1(a) and 1(c)). The solidus line is marked by the disappearance of the liquid phase (L).

It can be taken from the simulation results that the amount of titanium and aluminum has to be restricted to avoid the formation of the NiAl and η phase, respectively. In both cases, the amount of γ' is reduced by the precipitation of these phases, therefore, a deteriorating high-temperature stability is expected. With regard to the melting temperatures, it is visible that chromium and aluminum have little influence on the solidus and liquidus temperatures, while, as expected, titanium significantly decreases the melting temperature. Therefore, titanium similar to silicon can be added as second MPD, while the amount of manganese can be

reduced. Since chromium decreases the γ' solvus temperature for the first enhanced alloys, the amount was restricted to 5 wt %. The influence of higher chromium fractions on the epitaxial solidification of the enhanced systems will be investigated in later experiments. The quasibinary phase diagram, which is depicted in Fig. 1(d), shows a system with 20 wt % manganese, 5 wt % chromium, 3 wt % aluminum, and varying amounts of titanium. According to conventional high-temperature braze alloys, the fraction of aluminum was restricted to 3 wt %. Based on the simulation results, four braze alloys were selected for the following experiments: one alloy with 25 wt % and three alloys with 20 wt % manganese. The fractions of chromium and aluminum were held constantly; the amount of titanium was varied between 3 wt %, 4 wt %, and 6 wt %. For the selected alloys the solidus and liquidus temperatures and the corresponding melting intervals were calculated and listed in Table 1.

Table 1 Calculated melting temperatures (in kelvins) and corresponding melting intervals for the four selected braze alloys (composition in wt %); THERMOCALC, Version TCR, database TTNi7

No.	Ni	Mn	Cr	Al	Ti	T_S	T_L	ΔT_{SL}
1	Bal.	25	5	3	3	1375	1418	45
2	Bal.	20	5	3	3	1411	1461	50
3	Bal.	20	5	3	4	1399	1448	49
4	Bal.	20	5	3	6	1375	1423	48

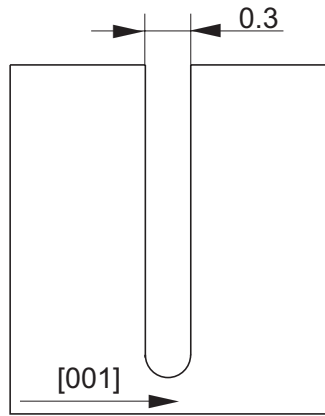


Fig. 2 Braze gap specimen (René N5, composition in wt %: Ni-7.5Co-7.0Cr-1.5Mo-5.0W-6.5Ta-6.2Al-3.0Re), gap width =0.3 mm or 300 μm , length \approx 10 mm

2 Experiments

According to earlier brazing experiments [17,20], single-crystalline samples (René N5, virgin casting) with a 300 μm wide parallel gap, which was perpendicular to the [001] direction of the single-crystal, were used (Fig. 2). The four selected alloys were produced in an arc furnace as small lenses, weighting about 10 g. Previously, the phase transformation temperatures were measured by means of differential scanning calorimetry (DSC), applying a heating/cooling rate of 10 K/min. The corresponding heating curves are depicted in Fig. 3; the measured temperatures and melting intervals are listed in Table 2.

As can be taken from Fig. 3 and Table 2, the melting behavior of alloy Nos. 1 and 4 is very similar; therefore, it is concluded that with regard to the melting temperatures, the reduction in the manganese fraction by 5 wt % can be compensated by 3 wt % of

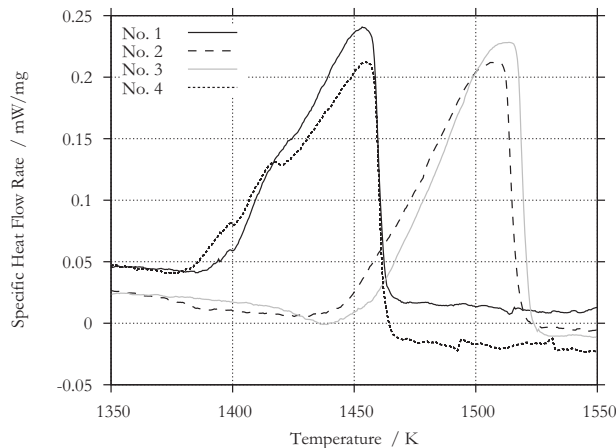


Fig. 3 Cutouts from the DSC heating curves for the chosen braze alloys

Table 2 Selected braze alloys (composition in wt %) with their corresponding solidus and liquidus temperatures as well as melting intervals (in kelvins) measured by DSC (heating curves, second run)

No.	Ni	Mn	Cr	Al	Ti	T_S	T_L	ΔT_{SL}
1	Bal.	25	5	3	3	1392	1462	70
2	Bal.	20	5	3	3	1446	1517	71
3	Bal.	20	5	3	4	1458	1521	63
4	Bal.	20	5	3	6	1379	1462	83

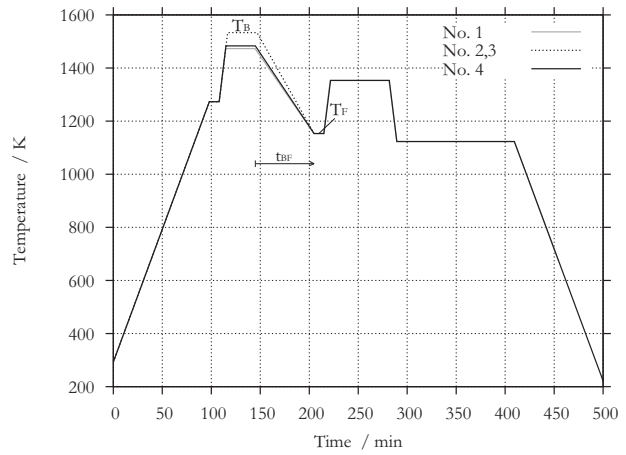


Fig. 4 Brazing cycles with following heat treatment

titanium. For braze alloy Nos. 2 and 3, relatively high melting points were measured. Surprisingly, the liquidus temperature of alloy No. 3 is slightly higher than that of alloy No. 2, although an increased amount of titanium is at hand. Furthermore, the melting interval is somewhat smaller; therefore, the solidus temperature of alloy No. 3 is also higher than in the case of alloy No. 2. Since the difference between the two alloys amounts to only 1 wt %, it is expected that due to effects of inertia, it cannot be precisely dissolved by the calorimetric measurements. However, for the selection of the brazing temperature, the exact dissolution is less important; in both cases, 1533 K were chosen for T_B .

The comparison of the simulated and measured solidus and liquidus temperatures shows that the calculated temperatures are generally too low, which is also the case for the calculated melting intervals. As expected, due to the high amounts of manganese, a systematic deviation of the simulated temperatures is at hand, whereas the mismatch of the liquidus temperatures is larger than that of the solidus temperatures. However, the brazing temperatures were chosen according to the DSC measurements; the simulation data were only used for the first alloy selection.

Brazing cycles, as depicted in Fig. 4, were conducted in a vacuum furnace with a residual pressure of less than 8×10^{-5} mbar. It could be observed in earlier works [17,20] that in the case of the manganese containing systems, even very fast cooling rates ($t_{BF}=10$ min) lead to a complete epitaxial solidification. However, it was also found that an advanced concentration homogenization can be achieved by slower cooling. Therefore, in this work 1 h was selected for the brazing time t_{BF} . The brazing temperature T_B was varied between 1473 K and 1533 K, while a constant final temperature T_F of 1153 K was chosen. A heat treatment, which consisted of a solution annealing (1 h at 1353 K) and a precipitation hardening (2 h at 1123 K), was added to all brazing cycles. Subsequently, the temperature was lowered to room temperature with a gradient of -1 K/min.

3 Results

After brazing, all samples were analyzed by means of scanning electron microscopy (SEM). To visualize the γ/γ' microstructure, molybdc-acid etchant was used for metallographic preparation. Cutouts from the center of the brazed gaps and the bonding zone are depicted in Fig. 5.

As visible in the SEM-images, a microstructure consisting of two phases is at hand. The brighter precipitations could be identified by energy dispersive X-ray spectroscopy analyses as γ' ; however, the typical cubic shape of the γ' precipitations is less pronounced. Obviously, the etching behavior of the base material and the braze gap is different. Usually, molybdc-acid etchant dissolves the γ' phase, which can also be observed in the base ma-

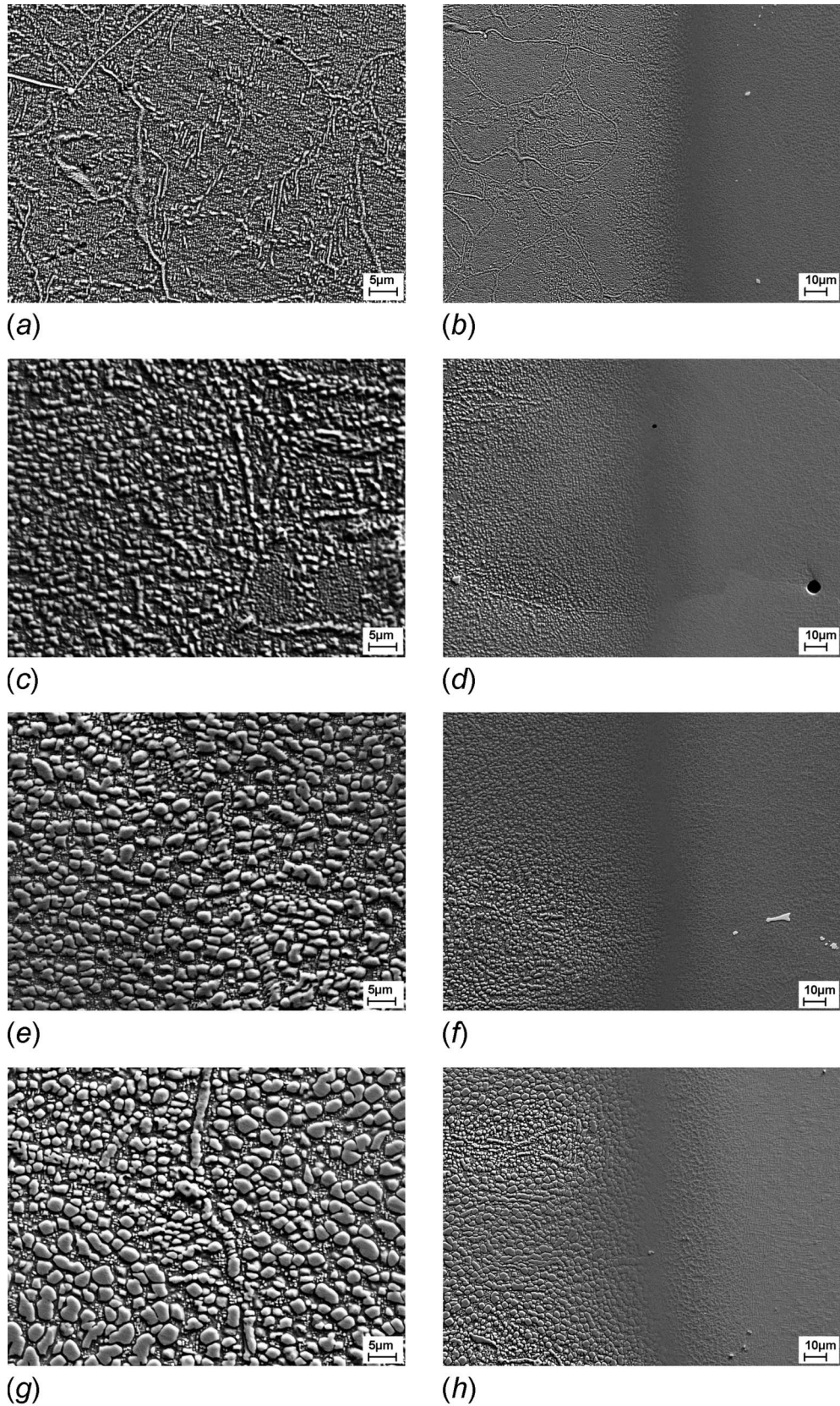


Fig. 5 SEM-images of the brazed gaps prepared with molybdic-acid etchant (in (b), (d), (f), and (h) the base material is always on the right side); (a) braze alloy No. 1: Ni-25Mn-5Cr-3Al-3Ti, gap center, $\delta = -1.131\%$; (b) braze alloy No. 1: Ni-25Mn-5Cr-3Al-3Ti, bonding zone; (c) braze alloy No. 2: Ni-20Mn-5Cr-3Al-3Ti, gap center, $\delta = -0.128\%$; (d) braze alloy No. 2: Ni-20Mn-5Cr-3Al-3Ti, bonding zone; (e) braze alloy No. 3: Ni-20Mn-5Cr-3Al-4Ti, gap center, $\delta = -0.005\%$; (f) braze alloy No. 3: Ni-20Mn-5Cr-3Al-4Ti, bonding zone; (g) braze alloy No. 4: Ni-20Mn-5Cr-3Al-6Ti, gap center, $\delta = 0\%$; and (h) braze alloy No. 4: Ni-20Mn-5Cr-3Al-6Ti, bonding zone

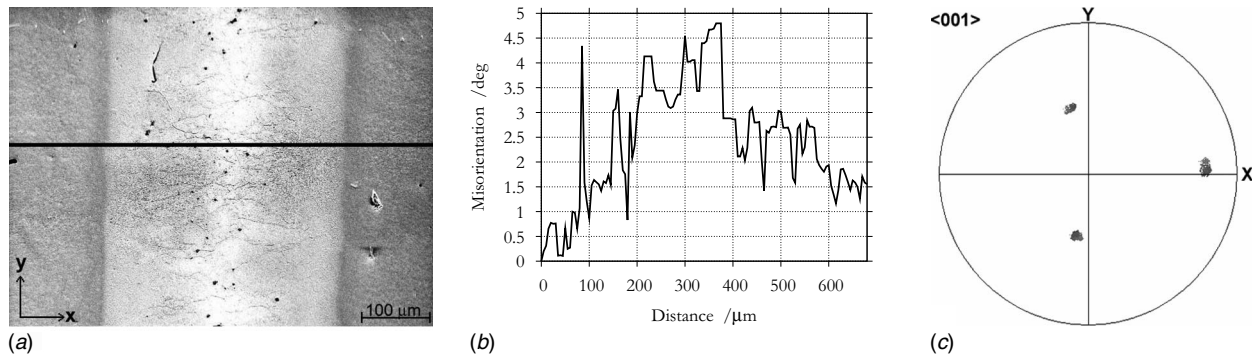


Fig. 6 Alloy No. 1: Ni-25Mn-5Cr-3Al-3Ti, $T_B=1473$ K; (a) SEM-image, (b) misorientation profile, relative to first point, and (c) $\langle 001 \rangle$ pole figure

terial. However, in the case of the braze gaps, the matrix has been dissolved. Between braze gap and base material, a transition zone exists, where no pronounced formation of γ' is visible. As predicted by THERMOCALC (Fig. 1(d)), with increasing amounts of titanium, the amount of γ' also increases; additionally, the precipitations become coarser, which can also be observed for decreasing amounts of manganese.

With regard to high-temperature applications, the misfit between γ and γ' is an important parameter to avoid γ' coarsening [18]. Therefore, the lattice parameters at room temperature were measured by means of X-ray diffraction. To eliminate the influence of microsegregations within the brazed gaps, for these measurements the braze alloy lenses were used: Previously, the lenses were heat treated according to the brazing samples (1 h solution annealing at 1353 K and 2 h precipitation hardening at 1123 K); afterward, small slices were cut for X-ray diffraction analysis. The

corresponding results are given below the SEM-images in Fig. 5. Obviously, a negative misfit is at hand, which decreases with increasing amounts of titanium and decreasing amounts of manganese.

In addition to the analyses described above, electron backscatter diffraction (EBSD) measurements were conducted to assure the epitaxial solidification and to quantify the misorientation between the braze gap and the base material. According to the evaluation of the binary and ternary nickel-manganese-based systems [17,20] misorientation profiles and pole figures were generated (Figs. 6–8). The misorientation profiles were measured perpendicular to the gap, as indicated by black lines in Figs. 6(a). The results were plotted relative to the first point, which was chosen on the left boundary within the base material.

The misorientation profiles show that the maximum misorientation, measured relative to the first point, is in the range of 6 deg

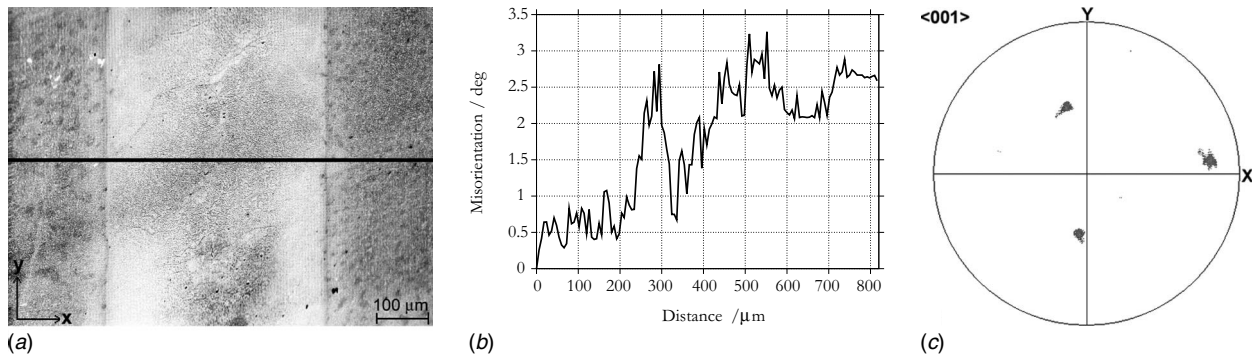


Fig. 7 Alloy No. 2: Ni-20Mn-5Cr-3Al-3Ti, $T_B=1533$ K; (a) SEM-image, (b) misorientation profile, relative to first point, and (c) $\langle 001 \rangle$ pole figure

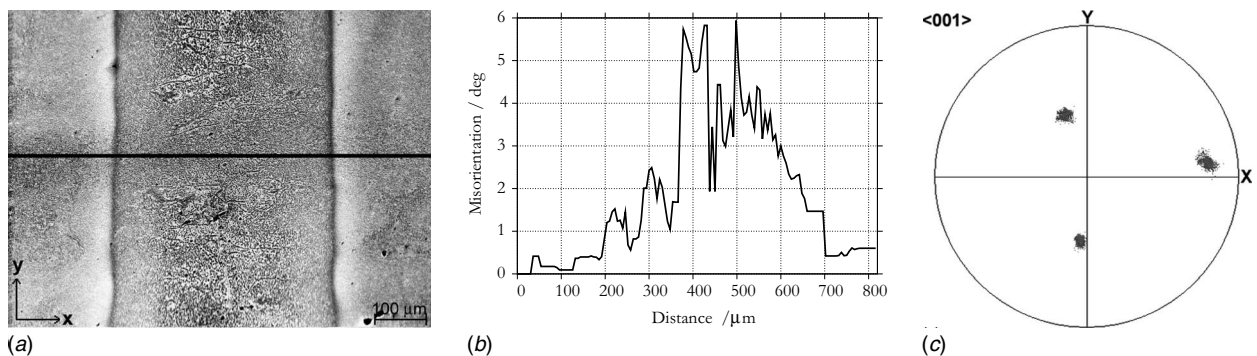


Fig. 8 Alloy No. 3: Ni-20Mn-5Cr-3Al-4Ti, $T_B=1533$ K; (a) SEM-image, (b) misorientation profile, relative to first point, and (c) $\langle 001 \rangle$ pole figure

(Fig. 8(b)). All pole figures show three discrete projection areas, which indicates that only one orientation is at hand. In each case, small rotations of the projection points can be observed: Obviously, small orientation deviations of the $\langle 001 \rangle$ from the sample x -axis direction are at hand. However, for the analysis of epitaxial solidification the exact sample orientation is not essential. Since no large angle grain boundaries with an orientation difference exceeding 10 deg [16] could be detected, it is concluded that complete epitaxial solidification in the case of all enhanced alloys occurred. The epitaxial solidification is not disturbed by the addition of chromium, aluminum, and titanium.

4 Discussion

In Refs. [17,20] nickel-manganese-systems were developed, which permit solidification times being significantly shorter than it can be achieved by conventional transient liquid phase bonding technologies. This gives the possibility of brazing even larger gaps in economically justifiable times. The main advantage is the high solubility of manganese in the solid γ phase and the low segregation tendency in the liquid phase. In contrast to the eutectic nickel-boron system, the nickel-manganese system is azeotropic. In the first case, an incomplete diffusion of boron causes a solidification, ending with the final eutectic reaction. In the second case, single phase solidification occurs even in the case of a very fast quenching, no secondary phase can precipitate. For this reason, epitaxial solidification can be accomplished within very short times without nucleation of stray grains.

However, the binary and ternary nickel-manganese-systems, developed in Refs. [17,20], contain neither any γ' formers nor elements providing a beneficial corrosion behavior. Therefore, in order to produce a γ/γ' microstructure similar to that of the base material, chromium and aluminum were added in this work. The second melting-point depressant silicon was replaced by titanium since titanium improves the wettability of the braze alloy, which is also utilized in the case of active braze alloys for producing ceramic/ceramic- or ceramic/metal-joints [21,22]. Moreover, as was visible in the simulation results, titanium increases the γ' solvus temperature.

In contrast to the nickel-manganese system, nickel-titanium and nickel-aluminum are eutectic systems. Therefore, one could expect that the epitaxial solidification is disturbed by the addition of those elements. However, compared with boron, in both cases, the solubility in nickel is much larger. In the nickel-aluminum system γ' ($\text{Ni}_3(\text{Al},\text{Ti})$) precipitates out of γ solid-solution. This is the essential effect in nickel-base superalloys, which is the reason for their excellent high-temperature properties [19]. The situation in the nickel-titanium system is similar: For higher amounts of titanium, the precipitation of the TiNi_3 phase is predicted; however, due to the high solubility of titanium, it occurs at relatively low temperatures. In contrast to boron, single phase solidification is at hand, therefore, secondary phases, which serve as nucleation sites for stray grains, cannot precipitate during solidification; the epitaxial solidification is not disturbed.

In the case of the enhanced systems, a heat treatment, which is similar to that of the superalloy CMSX-4, was added. The γ' solvus temperature of CMSX-4 is in the range of 1563–1593 K [18]. Usually the solution annealing and the precipitation hardening are conducted in several steps. Characteristic temperatures for the solution annealing are between 1473 K and 1593 K. For the precipitation hardening, temperatures between 1353 K and 1123 K are selected [18]. The γ' solvus temperatures of the braze alloys are lower. The following temperatures were calculated by THERMOCALC (see also Fig. 1(d)):

- No. 1: 1219 K
- No. 2: 1271 K
- No. 3: 1296 K
- No. 4: 1327 K

To achieve a complete dissolution of γ' , the solution annealing was conducted for 1 h at 1353 K (the precipitation hardening temperature of CMSX-4), followed by a precipitation hardening for 2 h at 1123 K [18]. Due to the lower γ' solvus temperatures, it was expected that the effect, involved by this heat treatment, is similar to that in the base material although the hold times are relatively short. The γ/γ' morphology, which can be observed within the brazed gaps, is similar to that within the base material; however, the typical cubic shape of the γ' precipitations is not at hand. In the case of 4 wt % and 6 wt % titanium, a bidisperse microstructure is visible, which is caused by the increased γ' solvus temperature. Obviously, according to the solidus and liquidus temperatures, the calculated γ' solvus temperatures were too low. γ' has not been completely dissolved: Coarse primary as well as fine secondary precipitations are at hand.

The X-ray-diffraction measurements, which were conducted at the heat treated braze alloy lenses, showed that the misfit is reduced by increasing amounts of titanium and decreasing amounts of manganese. Therefore, it is concluded that manganese is predominantly dissolved in the γ phase, whereby the γ lattice parameter is enlarged. This effect can be compensated by titanium, which is known to enlarge the lattice parameter of γ' . Due to microsegregations within the brazed gaps, the misfit will vary; however, the principal effect, which could be observed in the case of the braze alloy lenses, will be the same. In addition to the precipitation morphology discussed above, some γ' precipitations are long cord-shaped ones, which can be observed in the case of all four enhanced alloys. The formation of these precipitations still has to be examined and, where applicable, avoided by further heat treatments.

Altogether, epitaxial solidification was successful in the case of all enhanced alloys. Further heat treatments will be examined to produce cubic γ' precipitations with a mean diameter of about 0.3 μm , which is known as the ideal γ' diameter for creep resistance [18]. Moreover, high-temperature mechanical experiments such as tensile, creep rupture, and low cycle fatigue tests will be conducted to examine the high-temperature stability of the brazed gaps.

5 Conclusions

The epitaxial high-temperature brazing of 300 μm wide gaps in very short brazing times by use of enhanced nickel-manganese-based braze alloys was demonstrated. The following conclusions can be drawn.

1. A brazing process, which is not diffusion controlled, was developed. In contrast to a diffusion controlled isothermal solidification, the present brazing process is based on an epitaxial growth during cooling from brazing temperature. Therefore, compared with commercially available braze filler metals, the required solidification times could be shortened by a factor of up to 100.
2. Manganese was identified as ideal MPD because it combines a little segregation tendency with excellent solubility in the nickel-rich matrix. As a result, fast epitaxial solidification without formation of secondary phases or stray grains is possible.
3. Complete epitaxial solidification was also successful, if braze alloys, which were enhanced by chromium, aluminum, and titanium, were used: The process is robust toward changes in composition and, as demonstrated in recent works [17,20], also in temperature control.
4. Further heat treatments will be carried out to improve the microstructure within the braze gap, high-temperature stability of the joints will be assured by mechanical testing.

References

- [1] Braunny, P., Hammerschmidt, M., and Malik, M., 1985, "Repair of Air-Cooled Turbine Vanes of High-Performance Aircraft Engines—Problems and Experi-

- ence," *Mater. Sci. Technol.*, **49**(1), pp. 719–727.
- [2] Hoppe, B., 2003, *Verhalten gelöteter Nickelbasis-Superlegierungen unter thermischer Ermüdungsbelastung*, TU Braunschweig, Braunschweig, Germany.
- [3] Li, W., Jin, T., Sun, X., Guo, Y., Guan, H., and Hu, Z., 2002, "Transient Liquid Phase Bonding of Ni-Base Single Crystal Superalloy," *J. Mater. Sci. Technol.*, **18**(1), pp. 54–56.
- [4] Idowu, O. A., Richards, N. L., and Chaturvedi, M. C., 2005, "Effect of Bonding Temperature on Isothermal Solidification Rate During Transient Liquid Phase Bonding of Inconel 738LC Superalloy," *Mater. Sci. Eng., A*, **397**(1–2), pp. 98–112.
- [5] Idowu, O. A., Ojo, O. A., and Chaturvedi, M. C., 2006, "Microstructural Study of Transient Liquid Phase Bonded Cast Inconel 738LC Superalloy," *Metall. Mater. Trans. A*, **37**(9), pp. 2787–2796.
- [6] Duvall, D. S., Owczarski, W. A., and Paulonis, D. F., 1974, "TLP-Bonding: A New Method for Joining Heat Resisting Alloys," *Weld. J. (Miami, FL, U.S.)*, **53**(4), pp. 203–214.
- [7] Ellison, K. A., Lowden, P., and Liburdi, J., 1994, "Powder Metallurgical Repair of Turbine Components," *ASME J. Eng. Gas Turbines Power*, **116**, pp. 237–242.
- [8] Heikinheimo, L. S. K., Laukkanen, A., and Veivo, J., 2005, "Joint Characterization for Repair Brazing of Superalloys," *Weld. World*, **49**(5/6), pp. 5–12.
- [9] Heine, B., 1995, "Hochwärmefestigkeit von Nickelbasis-Superlegierungen für den Gasturbinenbau aus metallkundlicher Sicht," *Z. Metallkd.*, **49**(6), pp. 393–400.
- [10] Miglietti, W., and Du Toit, M., 2008, "High Strength, Ductile Braze Repairs for Stationary Gas Turbine Components—Part 1," ASME Paper No. GT2008-51133.
- [11] Nagy, D., and Huang, X., 2008, "Wide Gap Braze Repair Using Vertically Laminated Repair Scheme," ASME Paper No. GT2008-50046.
- [12] Jung, J.-P., Lee, B.-Y., and Kang, C.-S., 1993, "Liquid Phase Diffusion Bonding Using High Diffusivity- and High Melting Point-Element as Insert Material," *International Trends in Welding Science and Technology: Proceedings of an International Conference*, pp. 1101–1105.
- [13] Kim, D. U., 2001, "Isothermal Solidification and Single Crystallization Behaviors During TLP Bonding of a Ni-Base Single Crystal Superalloy," *Journal of the Korean Institute of Metals and Materials*, **39**(5), pp. 589–594.
- [14] Nishimoto, K., Saida, K., Kim, D., and Nakao, Y., 1998, "Bonding Mechanisms of Ni-Base Single Crystal Superalloy During Transient Liquid Phase Bonding," *Weld. World*, **41**(2), pp. 121–131.
- [15] Ramirez, J. E., and Liu, S., 1992, "Diffusion Brazing in the Nickel-Boron System," *Weld. J. (Miami, FL, U.S.)*, **71**(10), pp. 365–375.
- [16] Schnell, A., 2004, *A Study of the Diffusion Brazing Process Applied to the Single Crystal Superalloy CMSX-4*, EPF de Lausanne, Lausanne, Switzerland.
- [17] Laux, B., Piegert, S., and Rösler, J., 2009, "Braze Alloy Development for Fast Epitaxial High-Temperature Brazing of Single Crystalline Nickel-Based Superalloys," *Metall. Mater. Trans. A*, **40**(1), pp. 138–149.
- [18] Bürgel, R., 2001, *Handbuch Hochtemperatur-Werkstofftechnik*, Vieweg, Braunschweig, Germany.
- [19] Decker, R. F., 1969, "Strengthening Mechanisms in Nickel-Base Superalloys," *Steel Strengthening Mechanisms Symposium*.
- [20] Laux, B., Piegert, S., and Rösler, J., 2009, "Fast Epitaxial High-Temperature Brazing of Single Crystalline Nickel-Based Superalloys," *ASME J. Eng. Gas Turbines Power*, **131**(3), p. 032102.
- [21] Dupak, J., and Ustohal, V., 2001, "Soldering With Ductile Active Solders," *Mater. Manuf. Processes*, **16**(6), pp. 855–861.
- [22] Hanson, W. B., Ironside, K. I., and Fernie, J. A., 2000, "Active Metal Brazing of Zirconia," *Acta Mater.*, **48**, pp. 4673–4676.

Aerodynamic Design and Numerical Investigation on Overall Performance of a Microradial Turbine With Millimeter-Scale

Lei Fu

e-mail: leizhenlin@gmail.com

Yan Shi

Qinghua Deng

Zhenping Feng

e-mail: zpfeng@mail.xjtu.edu.cn

Institute of Turbomachinery,
School of Energy and Power Engineering,
Xi'an Jiaotong University,
Xi'an 710049, P.R.C.

For millimeter-scale microturbines, the principal challenge is to achieve a design scheme to meet the aerothermodynamics, geometry restriction, structural strength, and component functionality requirements while in consideration of the applicable materials, realizable manufacturing, and installation technology. This paper mainly presents numerical investigations on the aerothermodynamic design, geometrical design, and overall performance prediction of a millimeter-scale radial turbine with a rotor diameter of 10 mm. Four kinds of turbine rotor profiles were designed, and they were compared with one another in order to select the suitable profile for the microradial turbine. The leaving velocity loss in microgas turbines was found to be a large source of inefficiency. The approach of refining the geometric structure of rotor blades and the profile of diffuser were adopted to reduce the exit Mach number, thus improving the total-static efficiency. Different from general gas turbines, microgas turbines are operated in low Reynolds numbers (10^4 – 10^5), which has significant effect on flow separation, heat transfer, and laminar to turbulent flow transition. Based on the selected rotor profile, several microgas turbine configurations with different tip clearances of 0.1 mm, 0.2 mm, and 0.3 mm, two different isothermal wall conditions, and two laminar-turbulent transition models were investigated to understand the particular influences of low Reynolds numbers. These influences on the overall performance of the microgas turbine were analyzed in detail. The results indicate that these configurations should be included and emphasized during the design process of the millimeter-scale microradial turbines.

[DOI: 10.1115/1.3159375]

1 Introduction

Since the millimeter-scale turbomachinery for microcompact power sources and micropropulsion engines has become a new concept, e.g., Refs. [1–6], the miniaturization of gas turbines is quite meaningful. Because the proliferated portable machines, such as computers, digital assistants, cell phones, microcoolers, ultracompact unmanned aerial vehicles (UAVs), exoskeletons, and robots, require compact mobile electric equipments and energy supplies, and also because these machines demand higher power density and energy density that cannot be delivered by batteries, the microgas turbines are selected as an excellent candidate for the power source of mobile machines and portable electronics.

In recent years, research on gas turbines in millimeter-scale has been conducted in universities and research organizations all over the world. In the United States, Massachusetts Institute of Technology (MIT) aims to develop a microscale high-speed rotating machinery using the microfabricating and machining technology, which were known as microelectromechanical systems (MEMS), high aspect ratio Lithographic Galvanoformung and Abformung (LIGA) [7,8], deep-reactive ion etching (DRIE) [9,10], and wafer bonding [1]. The developed microturbine used radial turbine with a rotor diameter of 4 mm, which adopted the DRIE and wafer bonding technology [1]. Besides MIT, the microturbine developed at Stanford is a radial-axial turbine with a rotor diameter of

12 mm, and the silicon nitride rotor is produced by gel-casting technique [11].

In Asia and Europe, research teams at the University of Tokyo [12,13] and Tohoku University [14–17] have developed their microturbines. The latter used an axial-radial design with a rotor diameter of 10 mm, and the turbine wheel was machined by a five-axis numerical control (NC) milling machine. A single-stage axial microgas turbine with a rotor diameter of 10 mm made of stainless steel using die-sinking electrodischarge machining was developed at Katholieke Universiteit Leuven in Belgium [18]. The radial microgas turbine was developed with a rotor diameter of 8 mm at Office National d'Etudes et Recherches Aéropatiales (ONERA)-Applied Aerodynamic Department in France [19]. At the Singapore Institute of Manufacturing Technology, a radial microgas turbine was designed with a rotor diameter of 8.4 mm using MEMS-based micromachining technology [20]. In summary, microgas turbines have become a research hotspot with a rotor diameter of 10 mm and a power level of 50–100 W.

During the process of realizing the microgas turbine system, the bearing reliability, structural strength, and material performance are three extremely important factors. High-speed microgas bearing for microgas turbine was developed at the MIT [21,22] and Tohoku University [23]. The structural design space is defined according to the thermodynamic requirements, properties of the materials used, and manufacturing capabilities [24,25]. The selection of materials are highly dependent on their processing, structural design features, high temperature structures, and installation technology [26,27].

Manuscript received March 23, 2009; final manuscript received March 28, 2009; published online December 3, 2009. Review conducted by Dilip R. Ballal.

Table 1 Configurations and design operating conditions of the microgas turbine

	Units	Scheme 1	Scheme 2	Scheme 3	Scheme 4
TIT	K	1173	1173	1173	1173
π	-	3.0	3.0	3.0	3.0
n	rpm	800,000	930,000	930,000	930,000
G	g/s	1.5	1.4	1.6	1.6
D_1	mm	10	10	10	10
Ω	-	0.4	0.4	0.6	0.65
L	mm	1.0	1.0	1.0	1.0
P	W	253	280	349	352
eff_{it}	-	74.3%	77.5%	81.3%	85.6%
eff_{is}	-	63.6%	63.8%	68.2%	68.1%

The TurboAero team of Xi'an Jiaotong University has started the first step of realizing a 50 W microgas turbine at the smallest possible scale, with the best optional materials and existing manufacturing technologies. In this study, in consideration of the aerothermodynamic performance and geometrical restrictions, a microgas turbine has been designed and numerical simulations are performed. First, the investigation focused on the overall performance of the different rotor profiles are presented to determine the basic and suitable rotor curves for microgas turbine. Second, the several microgas turbine configurations with different tip clearances and wall temperatures are computed and analyzed. Finally, the transition from laminar to turbulent flows in the microgas turbine is simulated, and the effects of leakage flows, heat loss, and low Reynolds numbers on the overall performance are discussed in detail.

2 Aerothermodynamic Design

For millimeter-scale microgas turbines, there are many relatively high losses. In the simple cycle, the turbine losses include blade profile loss, leakage loss, bearing friction loss, exit loss, generator loss, etc. Due to the microscale, the exit losses are of the same order as that of the output power of the cycle. Compared with the conventional turbine, the micromachines are significantly different in two ways: small Reynolds number and 2D geometry limitation. The low Reynolds number (10^4 – 10^5) is a reflection of the millimeter-scale size, which places the designs in the laminar or transitional range. The lower Reynolds number will make the flow separate in the turbine passages and also imply much influence of skin friction and heat transfer than that of conventional scale gas turbines. Therefore, the extra losses of size effect should be considered through the design process of microgas turbines.

For the structure design, the stator inlet radius and exit radius are 7.0 mm and 5.6 mm, respectively. The rotor inlet radius is 5.3 mm, and the rotor outlet radius varies according to the different degrees of reaction. The particular challenge in the design of microgas turbine is the microfabrication constraints that limit the blade height to be uniform. Currently microdevices always have a constant span and 2D geometry. In this study, the blade heights are both 1.0 mm in the channel of the stator and rotor.

The microgas turbine configurations and design operating conditions are shown in Table 1. One profile of stator, four different profiles of the rotor, and the meridional view are shown in Fig. 1. The stator is designed as an aerodynamic profile of which the leading edge adapts for varying incidence angle of the incoming flow. It is worth to note that the assembly angle of the stator could be adjusted for the configurations of different rotor profiles. The rotor of scheme 1 is an aerodynamic profile, which has a good performance in axial turbine design. The rotor of scheme 2 is a lamella profile, which can be adapted to the higher rotational speed. The rotor of scheme 3, of which the camber line turns to the negative rotating direction, is able to fit with the higher rota-

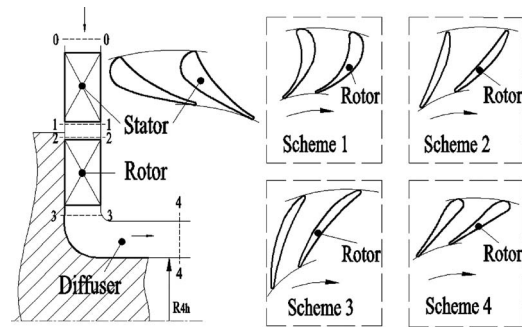


Fig. 1 Meridional view and the profiles of the turbine stator and rotor

tional speed and the angle of incoming flow better. The rotor of scheme 4 is similar to scheme 3 but the leading edge is modified to obtain higher efficiency.

The turbine inlet temperature was selected as 1600 K [1], 1545 K [19], and 1323 K [16] in past studies. In this study, however, a relatively low turbine inlet temperature is selected for several reasons given as follows.

The material restriction is the first consideration, since selecting high temperature materials for the turbine will demand high precision machining and packaging technology in millimeter-scale.

The microgas turbine needs cooling when operating under high temperature. For instance, the silicon rapidly loses strength above 900 K, so cooling is required for Si turbines. The simplest way to cool the turbine is to eliminate the shaft [1], which rejects the heat conducting to the compressor fluid. But, this method has great disadvantage of lowering the pressure ratio of the compressor, and the cycle output power decreases concomitantly.

The higher turbine inlet temperature also leads to the combustion problem in microgas turbines. The primary design requirements of combustors include large temperature rising, high efficiency, low pressure drop, structural integrity, and stable ignition. Due to the short length scale, heat transfer to the structure is large.

The last consideration of the higher temperature influence on compressors is that the temperature will make the compressor consume more power under the same expansion ratio and mass flow rate. Hence, in this study, the turbine inlet temperature is selected at relatively low level of 1173 K.

The high specific power of turbines will get the high output power of the cycle. In Refs. [1,19,16], the specific powers of turbines are 200 kJ/kg, 266 kJ/kg, and 238 kJ/kg, respectively. The current specific power of scheme 3 is 218 kJ/kg, while that will be nearly 278 kJ/kg when the turbine inlet temperature reaches to 1500 K.

3 Numerical Method

To investigate the flow characteristics of the designed scheme accurately, 3D Navier–Stokes equations are applied in the present simulation by using the commercial software NUMECA EURANUS [28]. The Spalart–Allmaras one-equation turbulence model is used to simulate the turbulent flow. This model has been used in the microgas turbine numerical simulation [17]. The simulation results were compared with the experimental data and the difference was acceptable. The four-step Runge–Kutta algorithm is adopted to ensure numerical time integration, and the implicit residual smoothing method is used to get high Courant–Friedrichs–Lewy number (CFL).

Multiblock grid is applied by the AUTOGRIDS of NUMECA with the default grid structure, and the total grid numbers are set to be 319,800 and 517,625 in the stator and the rotor, respectively. In the stator, the O-type grid number of the blade is set to be 132,225, which are composed of 25 nodes in circumferential, 41 in radial, and 129 in streamline directions. In the rotor, the same

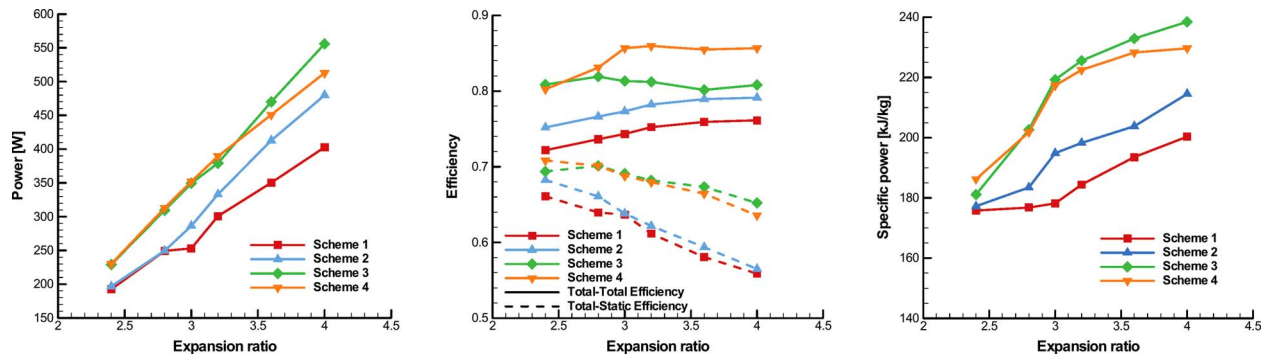


Fig. 2 Turbine power, efficiency, and specific power via stage expansion ratio for the four schemes

configuration of O-type grid is employed. The grids are refined at the endwall, near-wall, leading, and trailing edges. It is worth to note that all schemes use the same grid structure and numbers in order to compare the simulation results.

For further research, the tip clearance effect on the overall performance is taken into account in the scheme 3. The grid numbers at the rotor tip region is set to be 46,053, with grid numbers of 17 in the circumferential, 21 in radial, and 129 in streamline directions.

In order to obtain the accurate simulation and to predict the possibly existent transition, the second computational fluid dynamics (CFD) program—ANSYS CFX, Reynolds-averaged Navier-Stokes (RANS) three-dimensional viscous codes, is used. The two-equation $K-\omega$ shear stress transport (SST) turbulence model is employed. This turbulence model used in the transition of turbine has been compared with the experimental data with good prediction [29]. The grid is regenerated by the AUTOGRI5 AND CFX ICEM with the total grid numbers of 983,262 and 1,083,544, respectively. The unstructured tetrahedral mesh generated by ICEM with the total elements of 1,629,286 has also been used to predict the transition.

The near-wall Y^+ is adjusted to be 1–3 in the process of grid generating in NUMECA, while in the simulation using CFX, the near-wall Y^+ is kept between the minimal value of 0.09 and the maximal value of 2, and the average value is 1.1, which is required for the SST model in near-wall resolution.

4 Results and Discussion

4.1 Overall Performance. The variation in power, efficiency, and specific power via changes of stage expansion ratio for the four microgas turbines are shown in Fig. 2. The figures show that the power generated by the microgas turbine is proportional to the stage expansion ratio in all of four schemes. Schemes 3 and 4 have the higher power among these schemes. The slope of the curves indicates that schemes 3 and 4 gain more power with the same rise of stage expansion ratio than others. In the middle part of Fig. 2, except scheme 3, the total-total efficiency of all schemes increases proportionally to the stage expansion ratio. The total-total efficiency of scheme 3 increases in the lower stage expansion ratio range, and it decreases in the range of the higher stage expansion ratio. Scheme 4 has the highest total-total efficiency among the four schemes. The total-static efficiency is increased inversely to the stage expansion ratio for all four schemes and that of scheme 3 is highest with the stage expansion ratio above 3.0. The slope of the curves shows that the lapse rates of schemes 3 and 4 are lower than that of the other schemes, and scheme 3 has the smallest lapse rate among the four schemes. This means that it has relatively good performance in off-design conditions of microgas turbine operations. Variation in specific power presents that the four schemes have the higher specific power with the increasing stage expansion ratio, but scheme 3 has the highest specific

power and growth rate with the stage expansion ratio above 2.8.

The analyses of the overall performance described above show that the rotor profiles of scheme 3 has the most outstanding performance among the four schemes. Scheme 3 has the highest total-static efficiency of 68.2% and the smallest lapse rate of efficiency in off-design conditions. The specific power of scheme 3, which is an important parameter for the microgas turbine, is also the highest. For these reasons, scheme 3 is selected as the basic and suitable profile of the design for our millimeter-scale turbine in the subsequent research.

Even though higher specific power of the turbine presents at higher stage expansion ratio, the compressor would consume much power generated by the turbine. In order to get much output power of the cycle, the specific power is increased greatly at the cost of the slightly decreased total-static efficiency.

Along with the increase in stage expansion ratio, the change in total-total and total-static efficiency is reverse, and the difference becomes larger. That means that the leaving velocity loss should be reduced as much as possible.

4.2 Improvement of Total-Static Efficiency. It has been found that the leaving velocity loss is a large source of inefficiency in millimeter-scale turbines. This implies that the rotor exit Mach number should be reduced if possible, and the turbine would benefit from a high performance diffuser. Epstein [1] and Matsuura et al. [13] both pointed out this problem, but they did not give the modified method.

There are several ways to reduce the rotor exit Mach number. First, reducing the expansion ratio is the most direct approach, and the aerodynamic parameter is unchanged in general to keep the output power. Second, reducing the degree of reaction is another effective approach, but the change in flow speed and angle will induce variation in the flow field and reduce the total-total efficiency, correspondingly. The third method is to refine the geometric structure of rotor blades and the profile of diffuser. Although all these ways can reduce the exit Mach number, the designer has to trade-off the efficiency and the power of microgas turbine with realization of the structure. The third approach is finally used in this study.

For the geometric improvement in scheme 3, the rotor blade profile near the trailing edge is refined, which decreases the relative flow angle. The relative velocity at the rotor outlet is unchanged at a constant degree of reaction, and the circumferential speed is unchanged at the uniform rotational speed and radius. As a result, the outlet absolute velocity is decreased consequently. Furthermore, the absolute velocity of the diffuser outlet is reduced to increase the total-static efficiency. By using this modified method, the total-static efficiency has reached to 69.3%.

To design a high performance diffuser for the microgas turbine, three different geometry configurations of the diffuser at the tur-

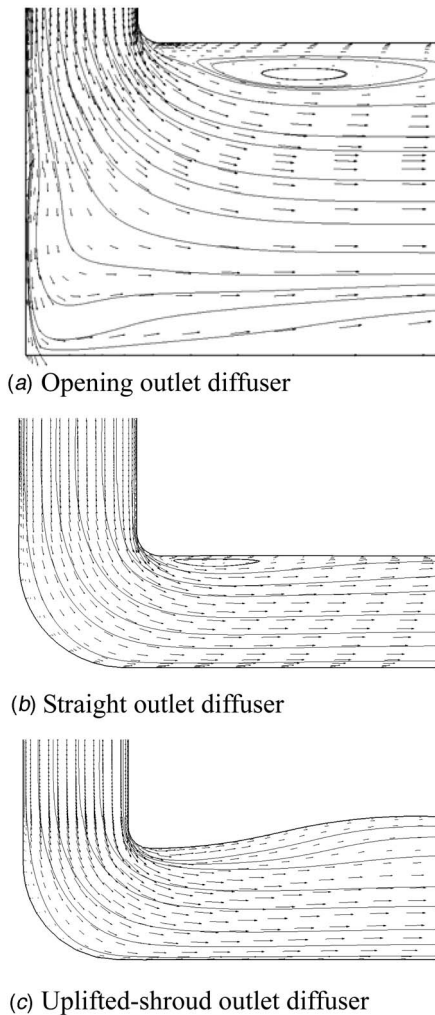


Fig. 3 Streamlines and relative velocity vectors in three kinds of diffusers

bine exit are designed and investigated. Figure 3 shows the opening outlet diffuser, straight outlet diffuser, and uplifted-shroud outlet diffuser, as well as their streamlines and relative velocity vectors.

In the opening outlet diffuser, flow separation is observed at the shroud after a 90 deg turning. Since the opening exit area is so large for the flow to fully diffuse out, there is a region of reverse flow near the exit of the shroud. Fluid is extruded and flows in a slow speed at the center line of the diffuser, which causes the total-pressure loss consequently. Without the geometrical restriction, fluid is forced to flow in axial direction by the static pressure generated by the congestion of flow in the center. The performance of this diffuser will be improved with a bulb at the center line.

In the straight outlet diffuser, flow separation is found at the same location as the opening outlet, while the range decreases. Fluid turns 90 deg uniformly and diffuses out without reverse flow. The low speed and high static pressure region exists at the hub of the 90 deg turning.

In the uplifted-shroud outlet diffuser, fluid turns 90 deg into the axial direction smoothly without reverse flow and vortex. The velocity vectors show that the velocity reduces at the exit but expands in the middle of the diffuser because of the gradual expansion of the flow channel. Accordingly, the static pressure of the rotor exit reduces when the static pressure at the diffuser exit keeps the same. As a result, the increase in the degree of reaction

Table 2 Influence of the two ratios on overall performance

	Ratio 1	Ratio 2	Ratio 3	Ratio 4	Ratio 5
	Ratio of S_{3-3}/S_{4-4}				
S_{3-3}/S_{4-4}	0.8160	0.8917	0.9775	1.1408	1.2342
eff _{ts} (%)	65.1	69.2	72.1	68.8	63.9
	Ratio of R_{4h}/R_2				
R_{4h}/R_2	0.1	0.2	0.3	0.35	0.4
P (W)	300	315	326	351	365
eff _{tt} (%)	84.2	83.7	83.9	83.1	83.2
eff _{ts} (%)	69.2	71.7	72.5	73.2	76.5

and velocity at rotor exit makes the flow expand again in the diffuser, and that expansion should be avoided in the processing of the diffuser design. In this study, the straight outlet diffuser has been used for the microgas turbine. It not only reduces the leaving velocity losses, but also decreases the exit losses, which consist of residual swirl loss and the right-angle turning loss.

In the design of diffuser, the two ratios, the area ratio of inlet area to exit area of the diffuser, and the radius ratio of outlet hub-radius to inlet radius of the turbine wheel, have significant influences on the overall performance of millimeter-scale microgas turbines. The locations of the radius used in the two ratios are shown in Fig. 1. The influence of the two ratios on the overall performance is given in Table 2.

Table 2 shows that the total-static efficiency is proportional to the area ratio from 0.8160 to 0.9775, but it varies contrarily to the area ratio from 0.9775 to 1.2342. The highest total-static efficiency appears at the area ratio around 1:1. When the $S_{3-3}/S_{4-4} < 1$, the excessively small blade height or excessively big shroud-radius of the diffuser exit will cause the total-static efficiency reduce. When the $S_{3-3}/S_{4-4} < 1$, the total-total efficiency increases and total-static efficiency decreases with the expansion of flow in the diffuser. For these reasons, the ratio of S_{3-3}/S_{4-4} should set to be around 1.0. This study takes this ratio as 0.9775, and the total-static efficiency is increased to 72.1%.

The variation in the overall performance via the ratio of R_{4h}/R_2 is shown in Table 2. The shaft power and total-static efficiency vary proportionally to the ratio of R_{4h}/R_2 . The total-total efficiency varies inversely to this ratio and the peak efficiency at the ratio of 0.1. The higher ratio of R_{4h}/R_2 cannot be selected because it will cause the shroud-radius and area of the diffuser exit higher than that of the rotor exit. The lower ratio of R_{4h}/R_2 should not be selected because of the shafting and bearing location. Then, in consideration of power, efficiency, stability of bearing and shafting, manufacturing, and installing technology, the radius ratio should be chosen as high as possible. Currently, the radius ratio is chosen as 0.35, and the total-static efficiency reaches the value of 73.2%.

As the geometry of microgas turbine is very small, the designer should consider geometry restrictions during the aerodynamic design process. That means the optimal design should be focused both on aerodynamic performance and reasonable structure of microgas turbines in such a limited space.

4.3 Effect of Tip Clearance. The tip clearance of the microgas turbine is designed as 0.1 mm, but the practical tip clearance may be more than 0.1 mm, which is determined by manufacturing and installing technology.

Hence, the tip clearances investigated in our computation have three different heights of 0.1 mm, 0.2 mm, and 0.3 mm. The tip clearances have relatively large dimensions with respect to the blade height, which are 10%, 20%, and 30%. Near the trailing edge of rotor blade with 0.1 mm tip clearance, there are scrape flow from the suction side to the pressure side and the leakage flow from the pressure side to the suction side in the upper boundary, and they form a leakage vortex near the upper pressure side.

Table 3 Effect of tip clearance on overall performance

	P (W)	Relative value (%)	eff _{is} (%)	Relative value (%)	G (kg/s)	Relative value (%)
0.0 mm	351	-	0.7322	-	1.56	-
0.1 mm	323	-7.9	0.6451	-11.9	1.64	+5.6
0.2 mm	297	-15.3	0.5903	-19.4	1.67	+8.0
0.3 mm	270	-22.9	0.5695	-22.2	1.68	+8.3

The increase in entropy in the gap is due to this friction. This means the vortex forms a high loss region in the tip clearance. In the case of the 0.2 mm and 0.3 mm tip clearances, the leakage losses are most pronounced in the gap without vortex due to the relatively large dimension.

The effect of tip clearance on the overall performance of microgas turbine is shown in Table 3. The aerodynamic performances difference between the rotor blade with and without tip clearance is evident. The tip clearance of 0.1 mm reduced the power and total-static efficiencies for 7.9% and 11.9%, respectively. These losses are too high to obtain the anticipated output power of the cycle.

Tip clearances of 0.2 mm and 0.3 mm have significant influence on power and total-static efficiency than that of 0.1 mm. With the tip clearance of 0.3 mm, the power is much lower than that of 0.2 mm but the increment in mass flow rate is almost the same. The difference in the power is derived from the decrease in relative velocity at the rotor exit with the tip clearance of 0.3 mm. From the above discussion, the large tip clearance is fatal for overall performance of the microgas turbines, and should be controlled as small as possible.

4.4 Effect of Heat Loss. In contrast with the conventional macro-turbine, the ratio of the wet area to the fluid volume, stated by the so-called cube-square law, increases with the dimension decreasing. This implies that the heat loss to the wall cannot be neglected in the micro-turbine. Therefore, heat loss is an important factor in the aerodynamic design, and its effect on overall performance should be emphasized.

Two different wall temperatures of 821 K and 703 K, which are set to be constant values of $0.7x$ and $0.6x$ turbine inlet temperature, respectively, are employed as isothermal wall boundary conditions for the microgas turbine.

The results indicate that the mass flow rates are 3.9% and 5.3% bigger than that of the adiabatic condition. The difference is attributed to the increase in the density because of the lower temperature of the fluid. Since the velocity of the fluid mainly depends on the pressure ratio, the increase in the density enhances the mass flow.

Heat loss on the turbine was calculated by the integral of heat flux on the turbine isothermal surfaces. The heat loss in the stator is more than that of the rotor because the temperature difference between the fluid and wall in the stator is larger. The heat losses are 155.8 W and 207.1 W in the stator, and 37.8 W and 79.2 W in the rotor at the wall temperature of 821 K and 703 K, respectively. Another goal in this study is to understand the effect of heat loss on the turbine efficiency. In the isothermal wall conditions, the entropy decreases so that the temperature of stage exit is lower than that of the adiabatic conditions. In order to calculate the efficiency exactly, the modified method used in Ref. [30] has been employed in this study. To recover the exit temperature to the adiabatic conditions, the fluid total temperature at the exit is modified to reflect the loss of energy through heat transfer. The change in total temperature at the exit of the stage can be computed as follows:

$$\Delta T_{\text{ex}} = \frac{\sum \int_A (\mu \partial T / \partial n|_w) dA}{\dot{m} \text{Re} \text{Pr}}$$

The integration is performed over all the heat transfer surfaces. The value of the exit total temperature is calculated as

$$T_{\text{ex}} = T_{\text{ex,iso}} + \Delta T_{\text{ex}}$$

The difference between our method and that proposed in Ref. [30] is that all the parameters computed with the units in this study. The equation of ΔT_{ex} is transformed as follows:

$$\Delta T_{\text{ex}} = \frac{\sum \int_A (\mu \lambda \partial T / \partial n|_w) dA}{GC_p \rho c l}$$

The total-total efficiencies for the microgas turbine are computed to be 77.1% and 74.2% for the wall temperatures 821 K and 703 K, respectively. The total-static efficiencies are 71.7% and 69.5% for the wall temperatures 821 K and 703 K, respectively. Failure to do so leads to an approximate error of 6–8 points in the total-total efficiency and 1–3 points in the total-static efficiency for the present computations.

4.5 Laminar-Turbulent Flow Transition. Laminar to turbulent flow transition has been recognized as one of the most difficult phenomena to simulate in fluid fields of flat plate or turbine blades. Although main stream flows are predominantly turbulent in turbomachinery, the boundary layer may be either laminar or turbulent in the millimeter-scale gas turbine as it works at low Reynolds number. The millimeter-scale microgas turbine is operated in a low Reynolds number.

Without experimental data, the current simulation results of transition should be compared between different CFD programs. As explained in the numerical method introduction, the two CFD programs evaluated in this study are commercial, RANS three-dimensional viscous codes. The first CFD program is NUMECA FINE/TURBO 7.2-1. The model used in this study is a one-equation S-A turbulence model, which uses the empirical correlations of Abu-Ghannam and Shaw. The second CFD program is ANSYS-CFX CFX11.0. The two-equation $K-\omega$ SST turbulence model [31,32] is coupled with one transport equation used for the intermittency and one for the transition onset Reynolds number criterion. The results and analyses were carried out with the inlet turbulence intensity of 1% and the varying inlet eddy viscosity ratios (turbulent dynamic viscosity to laminar viscosity) of 10, 30, 60, and 100.

Figure 4 shows the results of transition with intermittency by the FINE/TURBO of the different μ_t/μ . In different operating conditions of the microgas turbine, the Reynolds number of 7000, 5700, and 4100 are taken into account, respectively. The Reynolds number is based on the rotor exit flow conditions and the rotor chord length, and the Reynolds number is 7000 at the design point. It is worth to explain that the flows on the stator blades are laminar flows in all different conditions. The transition always takes place on the suction side of turbine blade because of the lower static pressure and the high pressure gradients, but the transition occurred in the middle range of the pressure side unexpectedly. The reason for this phenomenon may be caused by the similar pressure gradients on both sides of the blade, as well as the similar profile of the rotor in the 2D geometry.

At the Reynolds number of 7000, the laminar to turbulent flow transition occurs. In Fig. 4(a), μ_t/μ is indicated as 10, 30, 60, and 100 from the left to the right, and other pictures are placed in the same arrangement correspondingly. The visible onset of laminar to turbulent flow transition on the pressure side is at 26% chord length, and the extent of transition is approximate 2% chord length. Also, with different μ_t/μ from low to high, the transition has a small amplitude fluctuation. When μ_t/μ are 10, 30, and 60,

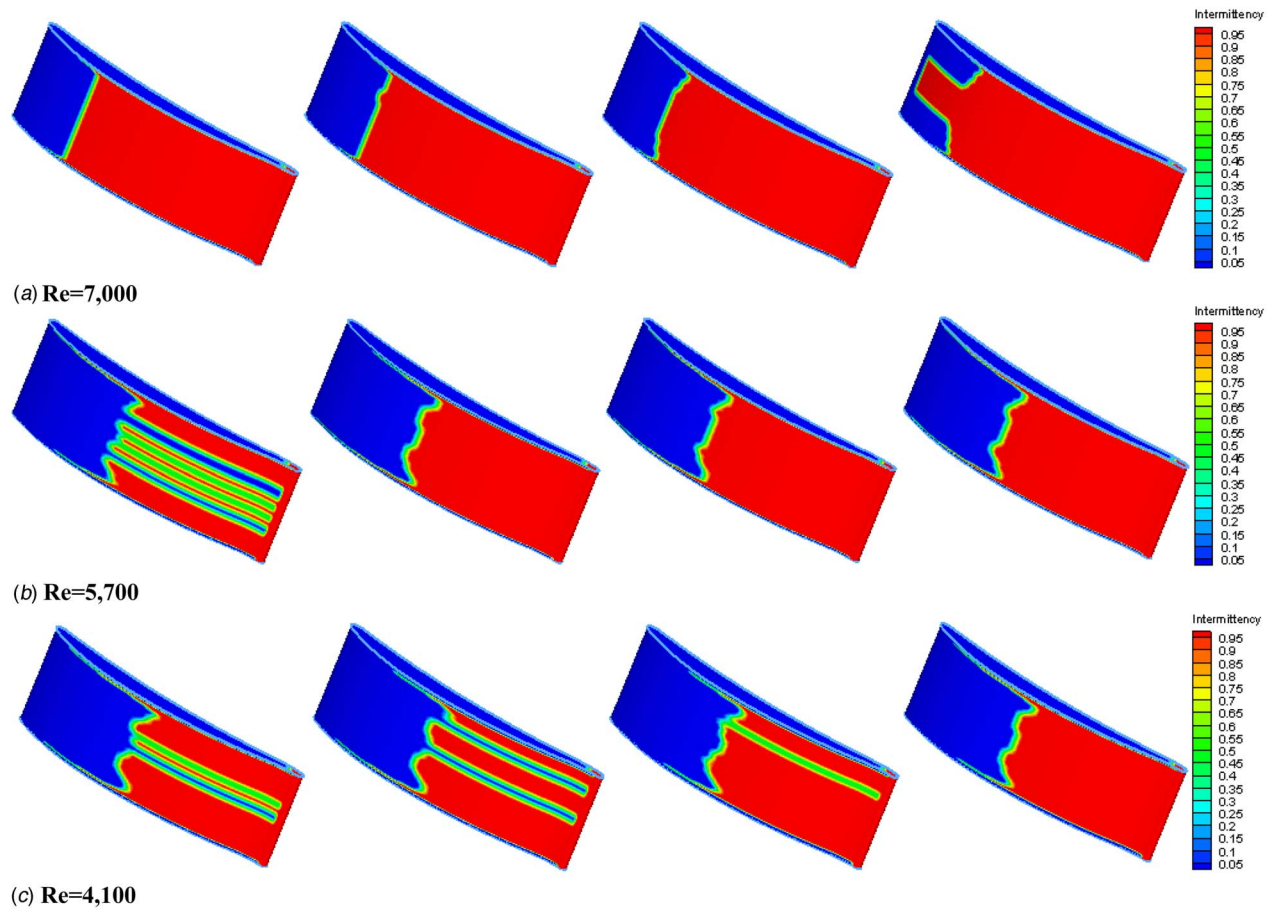


Fig. 4 Intermittency on pressure side of rotor blade by FINE/TURBO with different Reynolds numbers and various inlet eddy viscosity ratios of 10, 30, 60, and 100 (from left to right)

the transition positions are almost the same, but that moves forward in the main stream flows with the ratio of 100, and the transition onset is at 5% chord length.

At the Reynolds number of 5700, the visible onset of laminar to turbulent flow transition on the pressure side is at 34% chord length, and the extent of transition is at approximately 4% chord length. Due to the enhancement of viscous force, the transition puts off obviously. Different from the Reynolds number of 7000, the main stream is part laminar flow at μ_t/μ of 10. The transition onset moves up at approximately 2% and 4% chord length at the ratio of 60 and 100, respectively. Furthermore, near the hub and the shroud of the blade, the position of transition moves forward. This symmetrical phenomenon is attributed to the straight blade and the computation without tip clearance.

At the lowest Reynolds number of 4100, the visible onset of laminar to turbulent flow transition on the pressure side is at 38% chord length, and the extent of transition is at approximately 5% chord length. Different from Reynolds number of 7000, some legs of the main stream is still a laminar flow until the μ_t/μ of 60. When this ratio reaches to 100, the flow is fully turbulent after the transition. The symmetrical phenomenon still exists, but the transition of the flow near the bottom is moved slightly downstream. Based on the discussion above, with the Reynolds number and μ_t/μ decreasing, the transition onset is delayed and the transition extent is enhanced.

Figure 5 shows the results of transition from laminar to turbulent flows with the intermittency cloud picture solved by ANSYS CFX. At the Reynolds number of 7000, the visible onset of laminar to turbulent flow transition on the pressure side is at 62% chord length, and the extent of transition is at approximately 5% chord

length. Also, the intermittency is a zigzag fluctuation from the hub to the shroud. The transition of the flow near the endwalls is more in the front than that at the main stream. This phenomenon is the same with the results in Fig. 4, but the range is more expanded. The FINE/TURBO results show that the transition on the suction side is at 95% chord length near the trailing edge. But the CFX results show the intermittency contour is already a turbulent flow on the suction side from the 3% chord length, which is totally different from the results using FINE/TURBO in all Reynolds number conditions.

At the Reynolds number of 5700, the visible onset of laminar to turbulent flow transition on the pressure side is at 71% chord length, and the extent of transition is approximate 6% chord length. The zigzag fluctuation of the intermittency is more acute near the hub. Also the transition near the endwalls stays ahead of the blade than that of the main stream. The turbulent flow on the suction side is retarded at the 13% chord length.

At the Reynolds number of 4100, the lowest visible intermittency is over the 75% chord length on the pressure side, and the local maximum intermittency is around 85% chord length. On the suction side, the intermittency reaches to 0.6 until the 38% chord length, and this indicates that the transition is not complete yet.

In three different Reynolds number conditions, the different μ_t/μ has the same influence on the transition position compared with the results indicated in Fig. 4, but the difference is very slight. As a whole, the CFX prediction of transition position on pressure side is behind than that of the FINE/TURBO in three Reynolds number conditions. Also, the flow is already turbulent on the suction side, which is more reasonable in theory. In the lowest Reynolds number condition, the results of the simulation are very

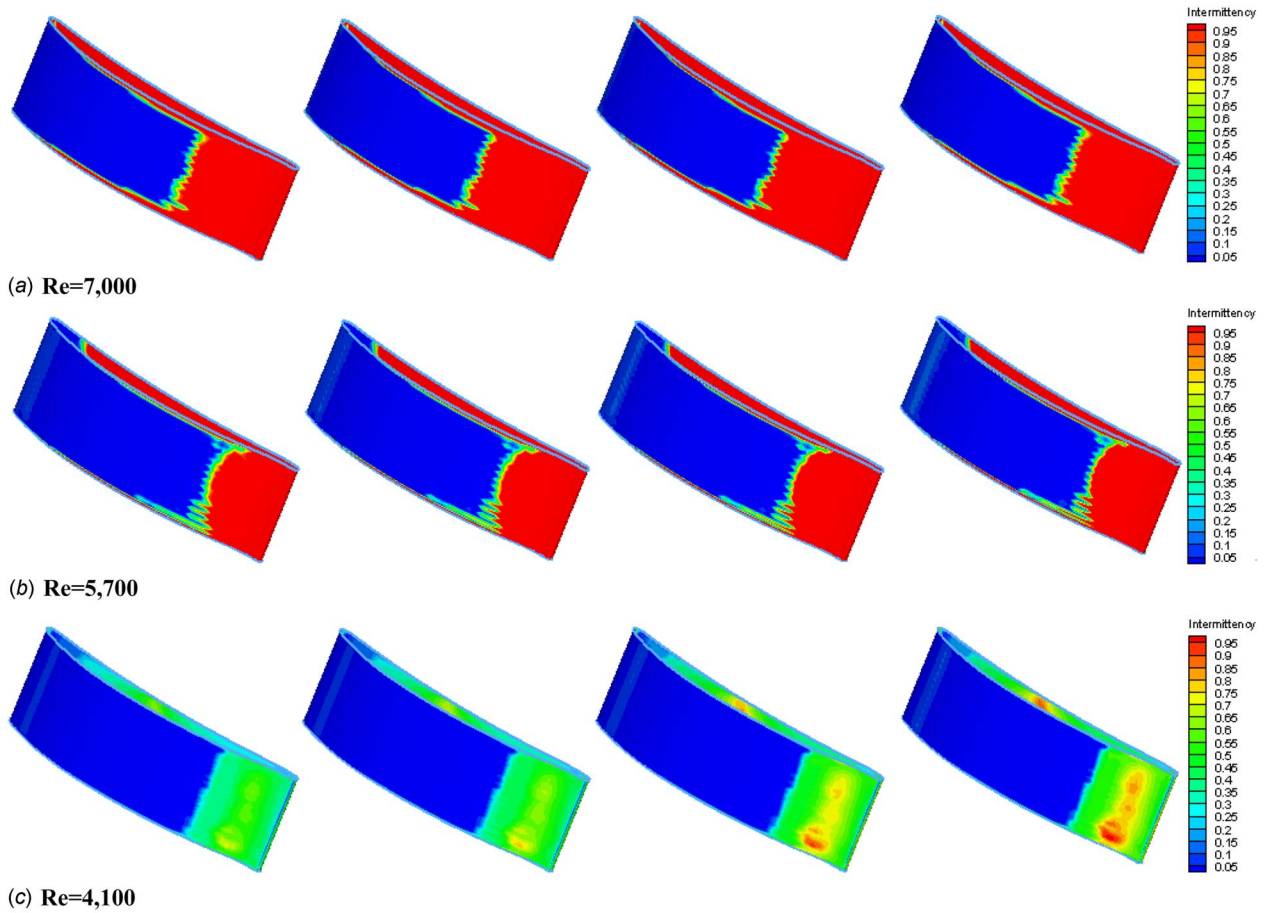


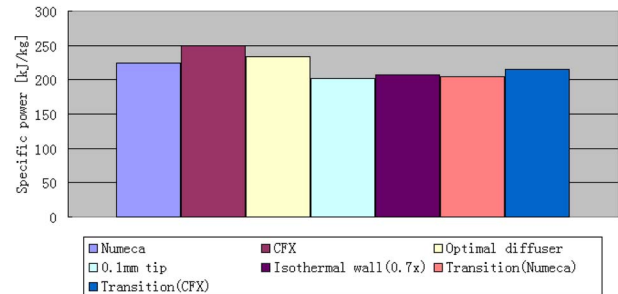
Fig. 5 Intermittency on pressure side of rotor blade by ANSYS CFX with different Reynolds numbers and various inlet eddy viscosity ratios of 10, 30, 60, and 100 (from left to right)

different. But from this study, the existence of the laminar to turbulent flow transition has been basically defined on microgas turbine blades.

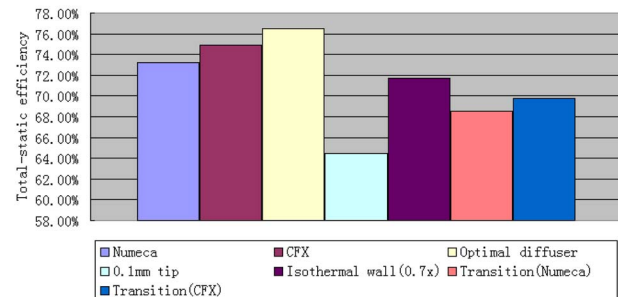
The results of the transition using the unstructured tetrahedral mesh by CFX are also studied. With the μ_t/μ of 30 and Reynolds number of 7000, the value of intermittency is in the small range of 0–0.3, which indicates that there is a laminar flow throughout the blade surface. All the simulation used the same boundary condition and Reynolds number mentioned before, but the results show that there is no laminar to turbulent flow transition. The reason for the different results obtained from using the unstructured hexahedral mesh and from using the unstructured tetrahedral mesh is not clear.

In this study, several cases of scheme 3 have been investigated and seven of them are shown in Fig. 6 in order to evaluate the influence on the overall performance directly and clearly. In these configurations, the condition of scheme 3 is only changed to measure the respective effect. These configurations include (1) the result of NUMECA in adiabatic condition, (2) the result of CFX in adiabatic condition, (3) the result of microgas turbine with optimal diffuser, (4) the result of 0.1 mm tip clearance in adiabatic condition, (5) the result of isothermal wall condition with modeling the temperature of $0.7x$ total inlet temperature, (6) the result of transition model in FINE/TURBO, and (7) the result of transition model in CFX.

In Fig. 6(a), the variety of specific power shows that the CFX results gains more value than FINE/TURBO results in adiabatic condition. With the optimal diffuser, the specific power of the microgas turbine obtains the relatively high level. Tip clearance has the most influence on the microgas turbine as that discussed in Sec.



(a) Specific power



(b) Total-static efficiency

Fig. 6 Specific power and total-static efficiency of the seven cases

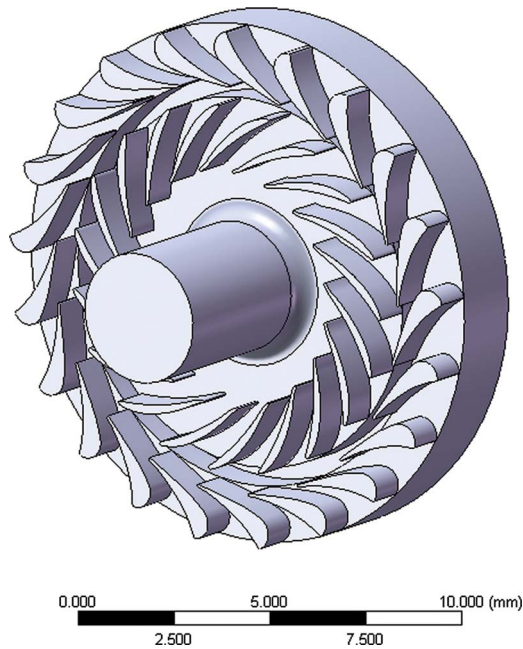


Fig. 7 Schematic diagram of final design for microgas turbines

4.3. In consideration of the transition model, the result of CFX is slightly higher than that of FINE/TURBO, the same results as the adiabatic condition.

In the Fig. 6(b), the variety of the total-static efficiency is evident. The CFX also gets the higher value than that of NUMECA, and the efficiency arrives at the highest level with the optimal diffuser. But the lowest efficiency derived from the tip clearance is 0.1 mm. The heat loss reduces the efficiency below 72%, and the transition model gets the lower value. All the configurations show that these effects should be considered carefully during the process of the microgas turbine design. Finally, the 3D design of the microgas turbine is shown in Fig. 7. The initial design of the test rig for the microgas turbine has been completed, and the aerodynamic experiments will be carried out in the near future.

5 Conclusions

The investigation on the aerothermodynamic design, geometrical design, blade profiling design, effects of the tip clearance, effects of heat loss, effects of the low Reynolds numbers, and CFD analysis for the overall performance of the microgas turbine with a rotor diameter of 10 mm are conducted and presented in this paper.

In the aerothermodynamic design, the results show that scheme 3 has the most outstanding performance among the four schemes with the power of 349 W and the eff_{ts} of 68.2%. This scheme has been taken as the basic profile of rotor for microgas turbines.

With the modified rotor profile, eff_{ts} reaches 69.3%. The straight outlet diffuser has the highest flow performance for the microgas turbine. The peak eff_{ts} appears at the area ratio of S_{3-3}/S_{4-4} around 1:1, and the radius ratio of R_{4h}/R_2 should be chosen as high as possible. After the optimization, the microgas turbine performance reaches to the power of 351 W and the eff_{ts} of 73.2%.

The tip clearance announced the fatal effect on the overall performance of microgas turbine. With the tip clearance of 0.1 mm, the power and the eff_{ts} are reduced by 7.9% and 11.9%, respectively. The heat loss causes the eff_{ts} to be reduced to 71.7% and 69.5% in the two cases. Failure to do so leads to an approximate error of 1–3 points in the eff_{ts} for the present computations. In the transition model, the existence of the laminar to turbulent flow

transition has been basically defined in microgas turbines. The decreases in power and eff_{ts} are about 8.5% and 3.9%, respectively.

Nomenclature

A	= surface area (mm ²)
c	= characteristic velocity (m/s)
C_p	= specific heat (J/kg k)
D	= diameter (mm)
eff	= isentropic efficiency
G	= mass flow rate (g/s)
L	= blade height (mm)
l	= characteristic length (mm)
N	= rotor rotational speed (rpm)
P	= power (W)
Pr	= Prandtl number
R	= radius (mm)
Re	= Reynolds number
S	= section surface area (mm ²)
TIT	= turbine inlet temperature (K)
Tu	= inlet turbulence intensity
π	= stage expansion ratio
\dot{m}	= normalized mass flow rate
λ	= thermal conductivity (W/m k)
Ω	= degree of reaction
ρ	= density (kg/m ³)
μ	= dynamic viscosity (Pa s)
μ_t/μ	= eddy viscosity ratio

Subscripts

0	= inlet of stator
1	= outlet of stator
2	= inlet of rotor
3	= outlet of rotor
4	= outlet of diffuser
ex	= exit of stage
h	= hub
iso	= isothermal condition
ts	= total-to-static
tt	= total-to-total

References

- [1] Epstein, A. H., 2004, "Millimeter-Scale, MEMS Gas Turbine Engines," *ASME J. Eng. Gas Turbines Power*, **126**(2), pp. 205–226.
- [2] Epstein, A. H., and Senturia, S. D., 1997, "Macro Power From Micro Machinery," *Science*, **276**, p. 1211.
- [3] Epstein, A. H., Senturia, S. D., Al-Midani, O., Anathasuresh, G., Ayon, A., Breuer, K., Chen, K.-S., Ehrich, F. E., Esteve, E., Frechette, L., Gauba, G., Ghodssi, R., Groshenry, C., Jacobson, S., Kerrebrock, J. L., Lang, J. H., Lin, C. C., London, A., Lopata, J., Mehra, A., Mur Miranda, J. O., Nagle, S., Orr, D. J., Piekos, E., Schmidt, M. A., Shirley, G., Spearing, S. M., Tan, C. S., Tzeng, Y. S., and Waitz, I. A., 1997, "Micro-Heat Engines, Gas Turbines, and Rocket Engines," 28th AIAA Fluid Dynamics Conference, Paper No. AIAA 97-1773.
- [4] Epstein, A. H., Senturia, S. D., Anathasuresh, G., Ayon, A., Breuer, K., Chen, K.-S., Ehrich, F. E., Gauba, G., Groshenry, C., Jacobson, S., Lang, J. H., Lin, C. C., Mehra, A., Mur Miranda, J. M., Nagle, S., Orr, D. J., Piekos, E., Schmidt, M. A., Shirley, G., Spearing, S. M., Tan, C. S., Tzeng, Y. S., and Waitz, I. A., 1997, "Power MEMS and Microengines," *Transducers '97*, The 9th International Conference on Solid-State Sensors and Actuators, Vol. 2, pp. 753–756.
- [5] Groshenry, C., 1995, "Preliminary Design Study of a Micro-Gas Turbine Engine," MS thesis, Department of Aeronautics and Astronautics, MIT, Cambridge, MA.
- [6] Jacobson, S. A., 1998, "Aerothermal Challenges in the Design of a Microfabricated Gas Turbine Engine," 29th AIAA Fluid Dynamics Conference, Paper No. AIAA 98-2545.
- [7] Wallrabe, U., Bley, P., Krevet, B., Menz, W., and Mohr, J., 1994, "Design Rules and Test of Electrostatic Micromotors Made by the LIGA Process," *J. Micromech. Microeng.*, **4**, pp. 40–45.
- [8] Guckel, H., Christenson, T. R., Skrobis, K. J., Junq, T. S., Klein, J., Hartojo, K. V., and Widjaja, I., 1993, "A First Functional Current Excited Planar Rotational Magnetic Micromotor," *Proceedings of the 1993 IEEE Micro Electro Mechanical Systems – MEMS*, pp. 7–11.

- [9] Ayón, A. A., Lin, C. C., Braff, R., Bayt, R., and Sawin, H. H., 1998, "Etching Characteristics and Profile Control in a Time Multiplexed Inductively Coupled Plasma Etcher," *Solid State Sensors and Actuator Workshop*.
- [10] Lin, C. C., 1999, "Development of a Microfabricated Turbine-Driven Air Bearing Rig," Ph.D. thesis, Department of Mechanical Engineering, MIT, Cambridge, MA.
- [11] Kang, S., Lee, S. J., and Prinz, F. B., 2001, "Size Does Matter, the Pros and Cons of Miniaturization," *ABB Rev.*, **1**(2), pp. 54–62.
- [12] Matsuo, E., Yoshiki, H., Nagashima, T., and Kato, C., 2002, "Development of Ultra Gas Turbines," *Technical Digest Power MEMS*, pp. 36–39.
- [13] Matsuura, K., Kato, C., Yoshiki, H., Matsuo, E., Ikeda, H., Nishimura, K., and Sapkota, R., 2003, "Prototyping of Small-Sized Two-Dimensional Radial Turbines," *International Gas Turbine Congress*, Paper No. IGTC-OS107, pp. 1–7.
- [14] Isomura, K., Murayama, M., Yamaguchi, H., Ijichi, N., Saji, N., Shiga, O., Tanska, S., Genda, T., Hara, M., and Esashi, M., 2002, "Component Development of Micromachined Gas Turbine Generators," *Technical Digest Power MEMS*, pp. 32–35.
- [15] Isomura, K., and Murayama, M., 2002, "Design Study of a Micromachined Gas Turbine With Three-Dimensional Impeller," *Ninth International Symposium on Transport Phenomena and Dynamics of Rotating Machinery*.
- [16] Isomura, K., Murayama, M., 2003, "Development of Micro-Turbocharger and Micro-Combustor for a Three-Dimensional Gas Turbine at Micro-Scale," *ASME Paper No. GT2003-38151*.
- [17] Isomura, K., Murayama, M., Teramoto, S., Hikichi, K., Endo, Y., Togo, S., and Tanaka, S., 2006, "Experiment Verification of the Feasibility of a 100 W Class Micro-Scale Gas Turbine at an Impeller Diameter of 10 mm," *J. Micromech. Microeng.*, **16**, pp. S254–S261.
- [18] Peirs, J., Reynaerts, D., and Verplaetsen, F., 2004, "A Microturbine for Electric Power Generation," *Sens. Actuators, A*, **113**(1), pp. 86–93.
- [19] Onishi, T., Burguburu, A., Dessorens, O., and Ribaud, Y., 2005, "Numerical Design and Study of a MEMS-Based Micro Turbine," *ASME Paper No. GT2005-68168*.
- [20] Shan, X. C., Zhang, Q. D., Sun, Y. F., and Wang, Z. F., 2006, "Design, Fabrication and Characterization of an Air-Driven Micro Turbine Device," *J. Phys.: Conf. Ser.*, **34**, pp. 316–321.
- [21] Liu, L., Teo, C. J., Miki, N., Epstein, A. H., and Spakovszky, Z. S., 2003, "Hydrostatic Gas Journal Bearings for Micro-Turbomachinery," *19th Biennial Conference on Mechanical Vibration and Noise*.
- [22] Ehrich, F. F., and Jacobson, S. A., 2003, "Development of High Speed Gas Bearings for High-Power-Density Micro-Devices," *ASME J. Eng. Gas Turbines Power*, **125**, pp. 141–148.
- [23] Isomura, K., Tanaka, S., Togo, S., and Esashi, M., 2005, "Development of High-Speed Micro-Gas Bearings for Three-Dimensional Micro-Turbo Machines," *J. Micromech. Microeng.*, **15**, pp. S222–S227.
- [24] Spearing, S. M., and Chen, K. S., 1997, "Micro-Gas Turbine Engine Materials and Structures," *21st Annual Cocoa Beach Conference and Exposition on Composite, Advanced Ceramics, Materials and Structures*.
- [25] Chen, K. S., Spearing, S. M., and Nemeth, N. N., 2001, "Structural Design of a Silicon Micro-Turbo-Generator," *AIAA J.*, **39**, pp. 720–728.
- [26] Moon, H. S., Anand, L., and Spearing, S. M., 2002, "A Constitutive Model for the Mechanical Behavior of Single Crystal Silicon at Elevated Temperature," *Mater. Res. Soc. Symp. Proc.*, **687**, pp. 279–284.
- [27] Choi, D., Shinavski, R. J., Steffier, W. S., Hoyt, S., and Spearing, S. M., 2001, "Process Development of Silicon-Silicon Carbide Hybrid Micro-Engine Structures," *Mater. Res. Soc. Symp. Proc.*, **687**, pp. 197–202.
- [28] NUMECA International, 2006, *NUMECA FINE/TURBO version 7.2-1*.
- [29] Mansour, M. L., Murthy Konan, S., Goswami, S., 2007, "Prediction of Lapse Rate in Low Pressure Turbines With and Without Modeling of the Laminar-Turbulent Transition," *ASME Paper No. GT2007-27110*.
- [30] Ameri, A. A., Steinthorsson, E., and Rigby, D. L., 1998, "Effect of Squealer Tip on Rotor Heat Transfer and Efficiency," *ASME J. Turbomach.*, **120**, pp. 753–759.
- [31] Menter F. R., Langtry, R. B., Likki, S. R., Suzen, Y. B., Huang, P. G., and Volker, S., 2004, "A Correlation Based Transition Model Using Local Variables Part I—Model Formulation," *ASME Paper No. GT2004-53452*.
- [32] Langtry, R. B., Menter, F. R., Likki, S. R., Suzen, Y. B., Huang, P. G., and Volker, S., 2004, "A Correlation Based Transition Model Using Local Variables Part II—Test Cases and Industrial Applications," *ASME Paper No. GT2004-53454*.

Influence of Blade Deterioration on Compressor and Turbine Performance

M. Morini
M. Pinelli
P. R. Spina
M. Venturini

Engineering Department in Ferrara (ENDIF),
University of Ferrara,
Via Saragat, 1,
44100 Ferrara, Italy

Gas turbine operating state determination consists of the assessment of the modification due to deterioration and fault of performance and geometric data characterizing machine components. One of the main effects of deterioration and fault is the modification of compressor and turbine performance maps. Since detailed information about actual modification of component maps is usually unavailable, many authors simulate the effects of deterioration and fault by a simple scaling of the map itself. In this paper, stage-by-stage models of the compressor and the turbine are used in order to assess the actual modification of compressor and turbine performance maps due to blade deterioration. The compressor is modeled by using generalized performance curves of each stage matched by means of a stage-stacking procedure. Each turbine stage is instead modeled as two nozzles, a fixed one (stator) and a moving one (rotor). The results obtained by simulating some of the most common causes of blade deterioration (i.e., compressor fouling, compressor mechanical damage, turbine fouling, and turbine erosion), occurring in one or more stages simultaneously, are reported in this paper. Moreover, compressor and turbine maps obtained through the stage-by-stage procedure are compared with the ones obtained by means of map scaling. The results show that the values of the scaling factors depend on the corrected rotational speed and on the load. However, since the variation in the scaling factors in the operating region close to the design corrected rotational speed is small, the use of the scaling factor as health indices can be considered acceptable for gas turbine health state determination at full load. Moreover, also the use of scaled maps in order to represent compressor and turbine behavior in deteriorated conditions close to the design corrected rotational speed can be considered acceptable.
[DOI: 10.1115/1.4000248]

1 Introduction

In gas turbines, one of the major contributions to performance decrease in both compressor and turbine (either gas generator or power turbine) is due to blade deterioration and fault. In literature, a wide number of papers deal with this subject. A recent and comprehensive review of the degradation in industrial gas turbine can be found in Ref. [1]. In particular, the main mechanisms responsible for blade deterioration and fault are (i) fouling, caused by the adherence of particles to blades, which results in an increased surface roughness and in changes to the shape of the airfoil; (ii) corrosion, which is caused by inlet air contaminants and by fuel and combustion derived contaminants, (iii) erosion, which consists of abrasive removal of material caused by particles impinging on flow surfaces; (iv) mechanical damage, caused by dramatic failure, either of the entire blade or of part of the blade due to exogenous causes (i.e., ingestion of foreign objects) or indigenous causes (i.e., part of the engine itself).

One of the main effects of blade deterioration is the modification of compressor and turbine performance maps. Since detailed information about actual modification of component maps is usually unavailable, many authors simulate the effects of deterioration and fault by scaling the map itself, i.e., by multiplying, point by point, the maps in *new and clean* condition by *scaling factors* F [2–4]. Different scaling factors can be used: compressor and turbine maps are usually scaled by multiplying efficiency and corrected mass flow rate at constant pressure ratio (or equivalent parameter, such as $\Delta h_{0c}/T$) and at constant corrected rotational

speed [2–4]. For example, the effects of a 5% reduction in the scaling factor for compressor efficiency and a 10% reduction in the scaling factor for the compressor corrected mass flow rate are shown in Fig. 1.

The modification of compressor and turbine performance maps with respect to new and clean condition due to actual deteriorations and faults can be then assessed by calculating the map scaling factors through the inverse solution of the program for gas turbine thermodynamic cycle calculation, in order to reproduce the measurements taken on the gas turbine [2–4]. If map scaling actually represents compressor and turbine map modification, then map scaling factors can be considered as gas turbine health indices, since they would be sensitive to the gas turbine health state only, while they would not be dependent on the gas turbine operating point [2–4]. Map scaling through scaling factors to match the actual gas turbine operating point implies that the *shape* of faulty curves is modified. This modification does not depend on the physics of actual deterioration, but is aimed at matching a single actual operating point, as can be seen in Fig. 1.

A different approach consists of investigating the effects of compressor and turbine stage deterioration by using stage-by-stage models [5–12].

In this paper, the effect of compressor and turbine blade deterioration is simulated through a stage-by-stage model, which uses generalized stage performance curves matched by means of a *stage-stacking* procedure [13]. In particular, compressor maps are predicted by using generalized relationships between stage efficiency, pressure coefficient, and flow coefficient (each of them normalized with respect to the respective reference value). Turbine maps are instead predicted by modeling each turbine stage by means of a series of two nozzles, a fixed one (stator) and a moving one (rotor).

The results obtained by simulating some of the most common

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received July 4, 2008; final manuscript received July 10, 2008; published online November 24, 2009. Editor: Dilip R. Ballal.

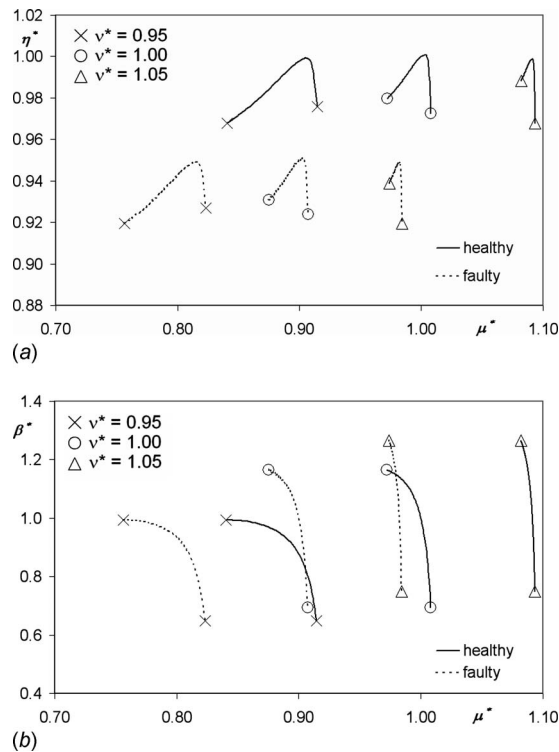


Fig. 1 Compressor efficiency and pressure ratio versus corrected mass flow: (—) new and clean condition; (---) 5% reduction in the scaling factors for efficiency and 10% for corrected mass flow

causes of blade deterioration (i.e., compressor fouling, compressor mechanical damage, turbine fouling, and turbine erosion), which occur in one or more stages simultaneously, are reported in the paper. Moreover, compressor and turbine maps obtained through the stage-by-stage procedure are compared with the ones obtained by means of map scaling.

2 Compressor and Turbine Stage-by-Stage Modeling

Thermodynamics-based models used for the simulation of gas turbine operation both in steady-state and in transient conditions require the knowledge of compressor and turbine performance maps [14–20]. These data are, however, proprietary to gas turbine manufacturers and are not usually available.

Different methods exist to predict compressor and turbine performances, such as the use of nondimensional component maps and scaling techniques [21,22] or the use of stage-by-stage compressor and turbine models, which allows the obtaining of the overall multistage compressor and turbine maps through a *stage-stacking* procedure, by using generalized stage performance curves [23–26]. This second approach, which is more general and allows a detailed stage-by-stage analysis of the flow through the compressor and the turbine, was used to set up a model for the prediction of compressor and turbine performance maps [13], and was also successfully applied for wet compression modeling [27,28]. Similar stage-by-stage compressor models were also successfully used in literature to investigate the effects of compressor stage deterioration on compressor and gas turbine performance [5–12].

2.1 Stage-by-Stage Procedure Application. Compressor and turbine modeling is carried out in this paper through a stage-by-stage procedure, which is outlined in the Appendix. This procedure was used to build compressor and turbine performance maps, in which blade deteriorations occur.

In order to build the performance maps, the knowledge of many

Table 1 Compressor parameters

Quantity	Value	Unit
No. of stages	18	
$(T_{iC})_r$	288.15	K
$(p_{iC})_r$	101.3	kPa
HR	60	%
$(M_{iC})_r$	159.0	kg/s
$(A_{iC})_r$	0.768	m ²
β_r	11	-
$(\eta_{sC})_r$	0.889	-
$\psi_{p,max}^*$	1.115	-
$\phi_{p,max}^*$	0.835	-
$(\psi_p^*/\phi^*)_{min}$	0.04	-
$\eta_{p,min}^*$	0.20	-
$(\psi_p^*/\phi^*)_{max}$	1.46	-
$\eta_{p,max}^*$	0.92	-
SF	-0.3	-

geometric parameters, which is not always available, is required. The following simplifications discussed in Ref. [13] were made in order to evaluate compressor and turbine performance maps.

- Mean diameters were supposed constant and equal to the mean diameter of the first stage both for compressor and turbine.
- For the compressor only, the geometrical similitude of the stages was assumed.

The main parameters of the compressor and the turbine are reported in Tables 1 and 2, respectively. These data refer to the TG20 FiatAvio gas turbine. The fluid at the compressor inlet was assumed humid air at ISO conditions, while the fluid at the turbine inlet was assumed to be combustion gas with natural gas as fuel.

3 Results and Discussion

3.1 Effect of Fouling and Mechanical Damage on Compressor Performance Maps. Two of the most common causes of compressor blade deterioration, i.e., fouling and mechanical damage, were simulated in one or more stages simultaneously. In case of fouling, the analyses were carried out by considering three different situations, in which the deterioration affects (i) the first stage only, (ii) all the stages gradually by imposing a deteriora-

Table 2 Turbine parameters

Quantity	Value	Unit
No. of stages	3	
$(T_{iT})_r$	1371.15	K
$(p_{iT})_r$	1080.0	kPa
$(p_{oT})_r$	104.4	kPa
$(M_{iT})_r$	162.1	kg/s
$(A_{iT})_r$	0.431	m ²
$(\eta_{sT})_r$	0.88	-
$\frac{Y_{min}^* - 1}{(i_{min}^* - 1)^2}$	0.005	-
$\frac{Y_{max}^* - 1}{(i_{max}^* - 1)^2}$	0.035	-
Ψ_1	1.297	-
R_1	0.253	-
Ψ_2	1.274	-
R_2	0.438	-
Ψ_3	1.000	-
R_3	0.500	-

Table 3 Simulated compressor deterioration

Stage No.	1	2-3	4-6	7-9	10-12	13-15	16-18
Fouling							
First stage	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	-	-	-	-	-	-
Gradual	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	$\Delta A^* = -8\%$ $\Delta \eta^* = -4\%$	$\Delta A^* = -6\%$ $\Delta \eta^* = -3\%$	$\Delta A^* = -4\%$ $\Delta \eta^* = -2\%$	$\Delta A^* = -2\%$ $\Delta \eta^* = -1\%$	-
Whole compressor	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$	$\Delta A^* = -10\%$ $\Delta \eta^* = -5\%$
Mech. damage							
First stage	$\Delta \eta^* = -5\%$	-	-	-	-	-	-

tion, which progressively decreases from the first to the last stage (the last three stages are not considered fouled), or (iii) the whole compressor. In case of mechanical damage, the analyses were carried out by considering only the situation in which the deterioration affects the first stage.

Three different corrected rotational speeds in the range 95–105% of the reference corrected rotational speed were analyzed.

According to Refs. [29–31], which summarize both operational field experiences and simulated data, fouling was modeled through a decrease in the flow passage area, coupled with a decrease in efficiency. In particular, the ratio between the variation in the flow passage area and efficiency was assumed equal to 2. Mechanical damage was instead modeled through a decrease in efficiency only.

The variation in these parameters with respect to new and clean condition values are reported in Table 3 for all the cases presented in the paper. It should be pointed out that, in case of fouling, the area variation accounts for the variation in compressor swallowing capacity, which, in turn, accounts for two contributions. The first one is a geometrical contribution (e.g., blade surface or roughness modification), while the second one accounts for fluid-dynamic blockage effects, due to the boundary layers thickening, which leads to a reduction of the actual flow passage area [32,33].

The results for the case of fouling are reported in Figs. 2–4 in terms of nondimensional pressure ratio and efficiency against nondimensional corrected mass flow rate. As can be noticed, fouling causes a shift of the pressure ratio curve toward a lower corrected mass flow rate value. The same occurs for efficiency curves. In addition, a remarkable shift of the efficiency curves toward lower efficiency values can be noticed both for the “gradual” and for the “whole compressor” deterioration (Figs. 3 and 4).

Figures 2–4 highlight that corrected mass flow rate values in choked region for the fouled compressor depends on the variation in the flow passage area and on the corrected rotational speed. A specific analysis that was conducted showed that both these dependences can be considered approximately linear. The results are reported in Fig. 5 for the case of a nondimensional corrected rotational speed equal to 1.00 only. The results for other nondimensional corrected rotational speeds are not reported, since they proved almost parallel. It can be noticed that the corrected mass flow rate in the choked region decreases linearly by decreasing the flow passage area. The rate of decrease depends on the number of stages affected by fouling: if only the first stage is fouled, the corrected mass flow rate decrease is approximately 2% for a 10% area decrease, while the decrease is about 10% for the same area decrease when fouling occurs on the whole compressor.

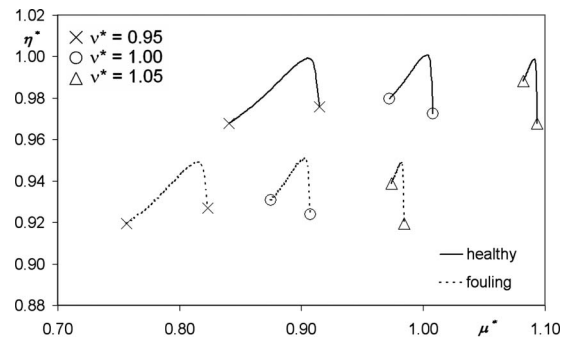
For the sake of brevity the effect of mechanical damage (simulated through a decrease in stage efficiency only) is shown in Fig. 6 in the case of damage occurring in the first stage only. It can be seen that, even if the area is not decreased, the corrected mass flow rate slightly decreases at a given corrected rotational speed due to the decrease in efficiency.

3.2 Effect of Fouling and Erosion on Turbine Performance Maps.

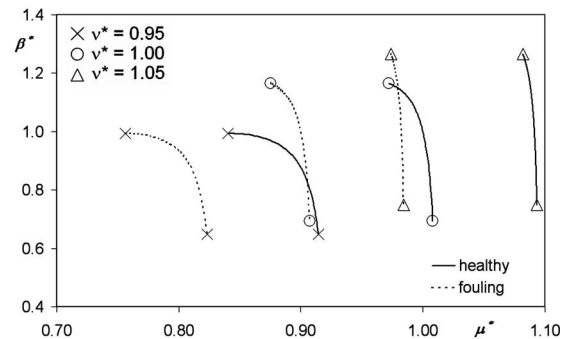
Two of the most common causes of turbine blade deterioration, i.e., fouling and erosion, were simulated by considering that deterioration affects either whole turbine or the first stator only. Three different corrected rotational speeds were analyzed from 90% to 110% of the reference corrected rotational speed.

Fouling was described by a decrease in the flow passage area and an increase in stage losses Y ($\Delta A^*/\Delta Y^* = -2$). The erosion considered in the paper only consists of the cut-back of the trailing edge by keeping the tip clearance constant, so that blade chord is reduced and the flow passage area is increased, while efficiency remains in practice unchanged [29,31]. The variation in these parameters with respect to new and clean condition values are reported in Table 4 for all the cases presented in the paper. As already underlined for compressor fouling, the area variation considered to simulate turbine fouling accounts for the variation in turbine swallowing capacity, which accounts both for a geometrical and for a fluid-dynamic flow modification.

A synthesis of the analyses performed is reported in Figs. 7–11

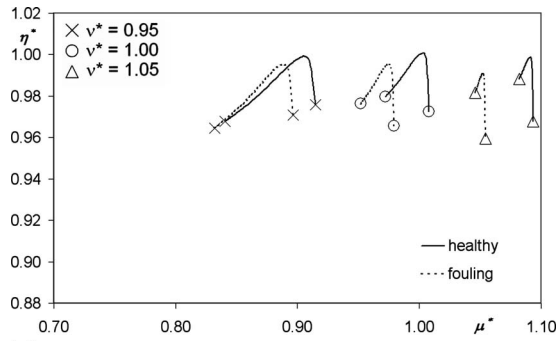


(a)

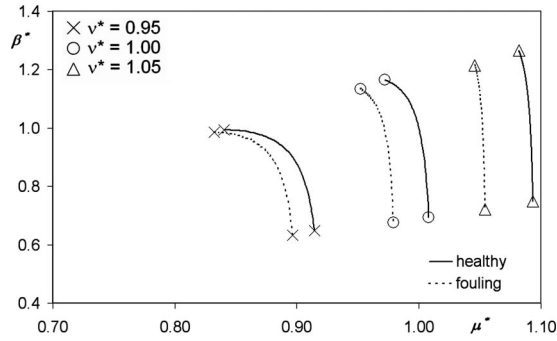


(b)

Fig. 2 Effect of fouling on the first compressor stage ($\Delta A^* = -10\%$ and $\Delta \eta^* = -5\%$)



(a)

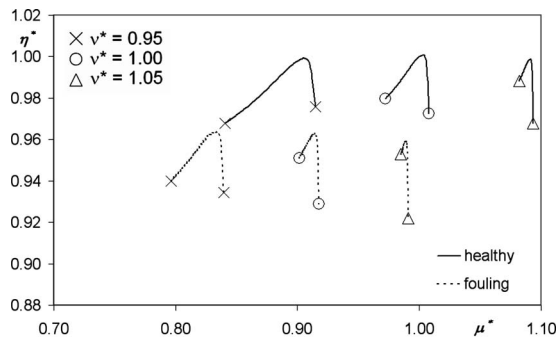


(b)

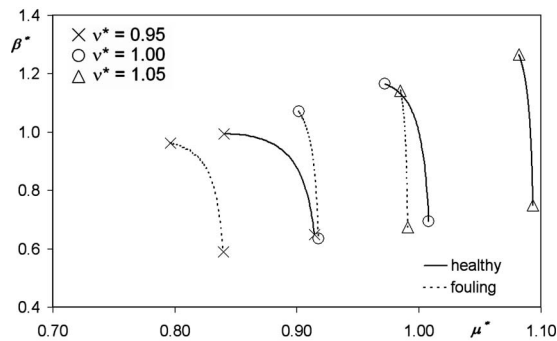
Fig. 3 Effect of gradual fouling on all the stages

in terms of nondimensional corrected mass flow rate and efficiency against the ratio between nondimensional isentropic enthalpy variation and turbine inlet temperature.

As can be noticed, fouling causes a shift of the corrected mass flow rate and efficiency curves toward lower values (Figs. 7–9).



(a)



(b)

Fig. 4 Effect of fouling on the whole compressor ($\Delta A^* = -10\%$ and $\Delta \eta^* = -5\%$)

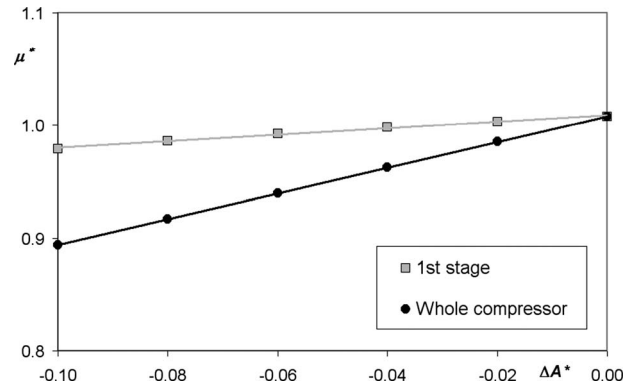
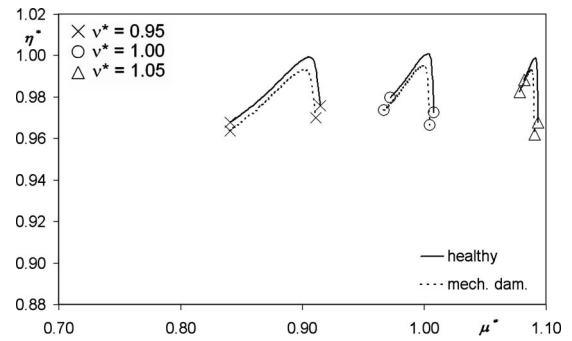


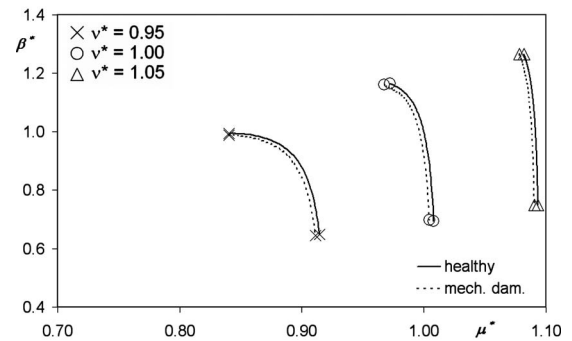
Fig. 5 Effect of fouling ($\Delta A^*/\Delta \eta^* = 2$) on the nondimensional corrected mass flow in the choked region for the reference corrected rotational speed ($v^* = 1.0$)

Otherwise, erosion causes a shift of the corrected mass flow rate curves toward higher values, while efficiency curves are in practice unaffected by this type of deterioration (Figs. 10 and 11), as reported in Refs. [29,31].

It can be also noted that, in case of fouling, the shift of the corrected mass flow rate and efficiency curves toward lower values is, in practice, independent of the number of stages affected by fouling. In fact, when the flow passage area of the first stator of the turbine reduces because of fouling, reductions in the flow passage areas of rotor, and stator cascades after the first stator negligibly affect the mass flow rate. If fouling only occurs in the first stator (Fig. 7), a saturation effect is highlighted in the choked region, where the corrected mass flow rate becomes constant and independent of the corrected rotational speed. This is due to the fact that, for the considered turbine, the flow through the first turbine nozzle is not choked in new and clean conditions. This fact is clearly highlighted by turbine performance maps in new



(a)



(b)

Fig. 6 Effect of mechanical damage on the first compressor stage ($\Delta \eta^* = -5\%$)

Table 4 Simulated turbine deterioration

Stage No.	1 stator	1 rotor	2 stator	2 rotor	3 stator	3 rotor
Fouling						
First stator	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$	-	-	-	-	-
Gradual	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$	$\Delta A^* = -4\%$ $\Delta Y^* = +2\%$	$\Delta A^* = -4\%$ $\Delta Y^* = +2\%$	$\Delta A^* = -2\%$ $\Delta Y^* = +1\%$	$\Delta A^* = -2\%$ $\Delta Y^* = +1\%$
Whole turbine	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$	$\Delta A^* = -6\%$ $\Delta Y^* = +3\%$
Erosion						
First stator	$\Delta A^* = +6\%$	-	-	-	-	-
Whole turbine	$\Delta A^* = +6\%$	$\Delta A^* = +6\%$	$\Delta A^* = +6\%$	$\Delta A^* = +6\%$	$\Delta A^* = +6\%$	$\Delta A^* = +6\%$

and clean conditions, in which corrected mass flow rate values in the choked region depend on the corrected rotational speed. In case of fouling in the first stator only, the decrease in the first nozzle flow passage area leads to choked conditions in the turbine first nozzle, so that corrected mass flow rate values become independent of the corrected rotational speed.

As regard turbine erosion (Figs. 10 and 11), the shift of the corrected mass flow rate curves slightly depends on the number of stages affected by erosion, as also happens in case of fouling (Figs. 7–9). However, differently from fouling, when erosion only occurs in the first stator (Figs. 10), the dependence of μ^* in the choked region with respect to the corrected rotational speed is more marked than in the case in which erosion occurs in the whole turbine (Figs. 11).

Finally, a specific analysis was conducted and showed that the dependence of μ^* in the choked region with respect to the variation in the flow passage area is approximately linear, as shown in Fig. 12, only for the case of a nondimensional corrected rotational speed v^* equal to 1.0 and area variation occurring only in the turbine first stator. The rate of increase in the corrected mass flow

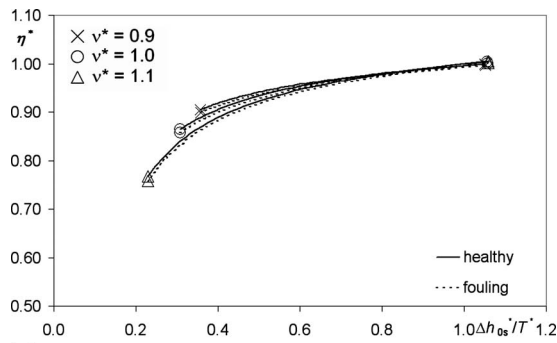
rate in the choked region due to erosion is about 0.8% for a 1.0% area increase, while fouling causes a decrease of μ^* in the choked region equal to about 1.0% for a 1.0% area decrease.

3.3 Comparison Between Stage-by-Stage Analysis and Map Scaling. Map scaling is usually adopted to evaluate the machine health state through gas path analysis (GPA) techniques [2–4,30]. As already outlined in the Introduction, scaling consists of multiplying the healthy compressor and turbine maps by a constant scaling factor F , in order to match the actual operating point. As already observed, this implies a deformation of the curve.

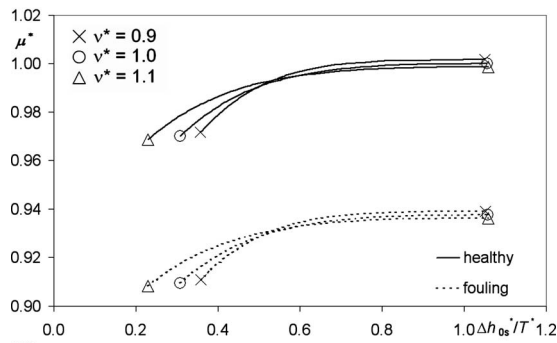
The performance maps obtained through the stage-by-stage analysis previously shown were compared with the ones obtained through map scaling. In particular, the maps built by means of stage-by-stage modeling were assumed to be the actual deteriorated curves.

Healthy curves are then scaled to match the deteriorated curves in order to:

- match the operating point on the deteriorated curve characterized by the maximum pressure ratio (compressor) or the

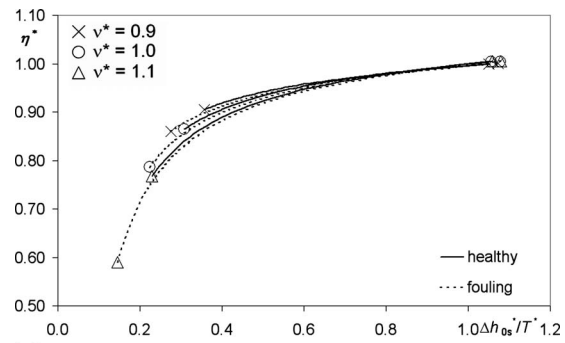


(a)

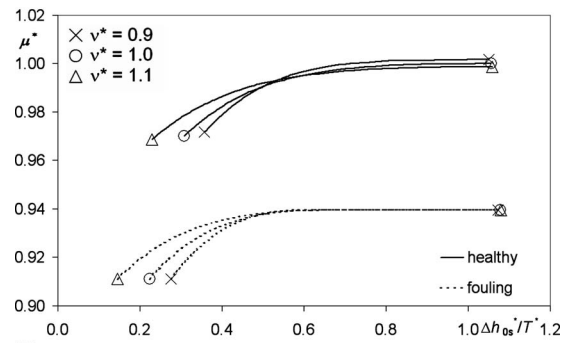


(b)

Fig. 7 Effect of fouling on turbine first stator ($\Delta A^* = -6\%$ and $\Delta Y^* = +3\%$)



(a)



(b)

Fig. 8 Effect of gradual fouling on the whole turbine

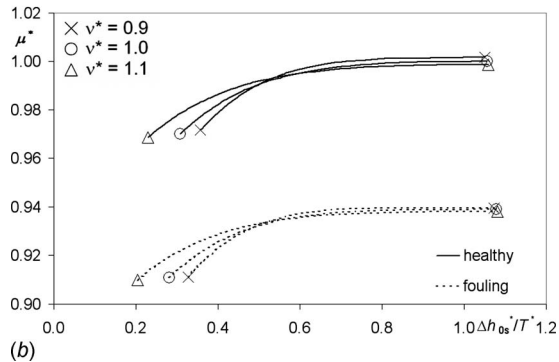
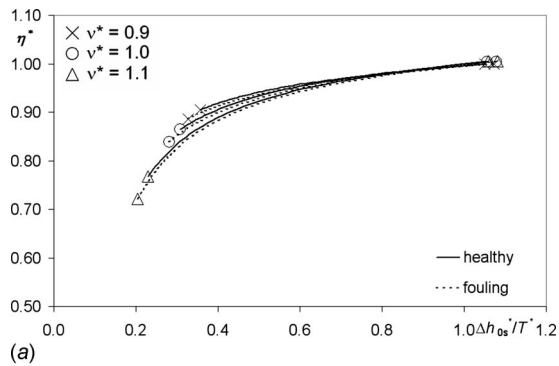


Fig. 9 Effect of fouling on the whole turbine ($\Delta A^* = -6\%$ and $\Delta Y^* = +3\%$)

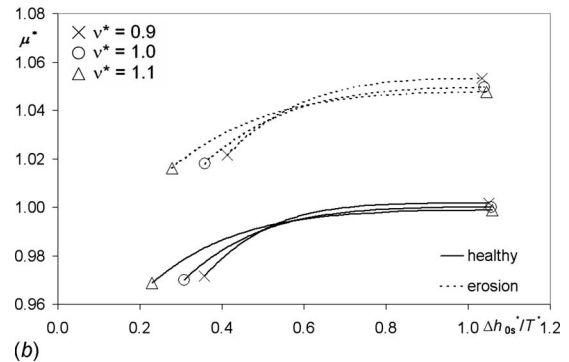
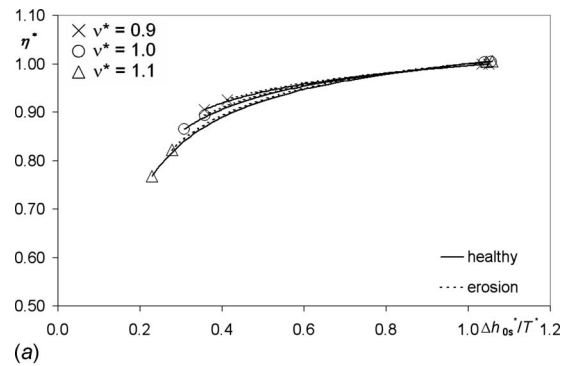


Fig. 11 Effect of erosion on the whole turbine ($\Delta A^* = +6\%$)

maximum ratio between nondimensional isentropic enthalpy variation and turbine inlet temperature (turbine) (S1)

- match the operating point on the deteriorated curve characterized by the minimum pressure ratio (compressor) or the minimum ratio between nondimensional isentropic enthalpy variation and turbine inlet temperature (turbine) (S2)

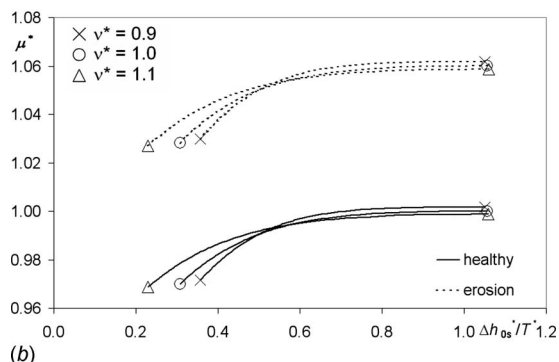
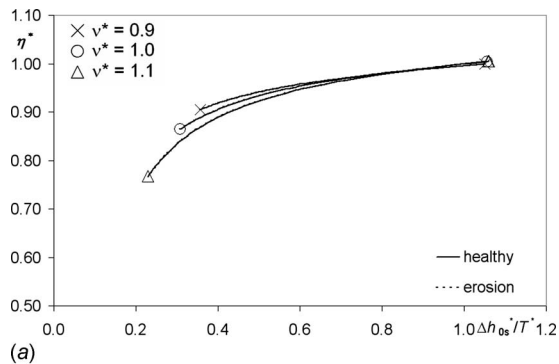


Fig. 10 Effect of erosion on turbine first stator ($\Delta A^* = +6\%$)

- minimize the root mean square error (RMSE) between the healthy and faulty map (S3)

The RMSE values for both mass flow rate and efficiency curves were calculated on 100 operating points (for each curve), with the same pressure ratio for the compressor (Figs. 13(a) and 13(b)), and the same $\Delta h_{0s}^*/T^*$ value for the turbine. Figure 13(c) shows the usual representation of the efficiency against the corrected mass flow rate, for a given value of the corrected rotational speed.

In addition to map scaling, the translation of the healthy curve was also considered, in order to match the deteriorated curve and, consequently, the curve shape was not varied. In this case, the three operations considered above (left-hand and right-hand extreme matching and the root mean square error minimization) are identified as T1, T2, and T3, respectively (not reported in Figs. 13(a)–13(c)).

This analysis was carried out at three different rotational speeds for fouling occurring in the compressor, both on the first stage only and on the whole compressor. The results of this analysis for

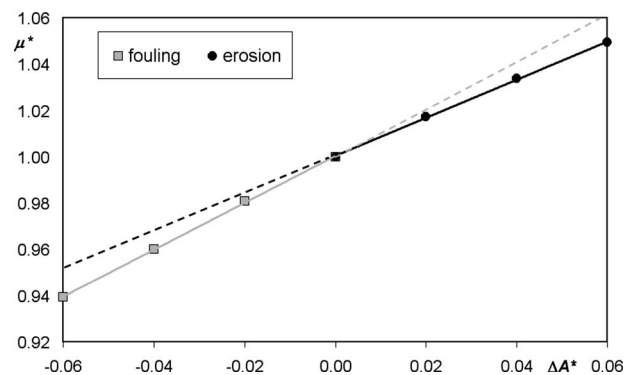
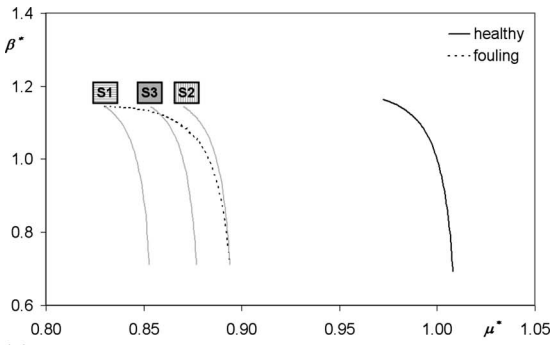
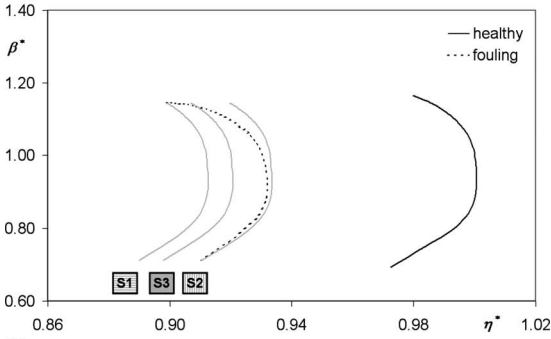


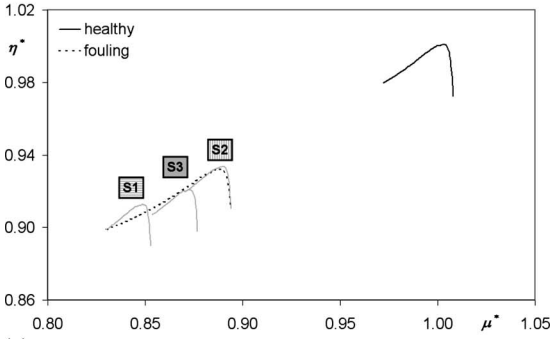
Fig. 12 Effect of turbine first stator fouling and erosion on turbine nondimensional corrected mass flow in the choked region ($v^* = 1.0$)



(a)



(b)



(c)

Fig. 13 (a) Compressor pressure ratio curve scaling, (b) compressor efficiency curve scaling ($\beta^* - \eta^*$ coordinates), and (c) compressor efficiency curve scaling ($\eta^* - \mu^*$ coordinates)

the case of fouling occurring on the first stage only are reported in Figs. 14 and 15.

Figure 14 shows the RMSE values on compressor mass flow rate: it can be noticed that the error decreases when the corrected

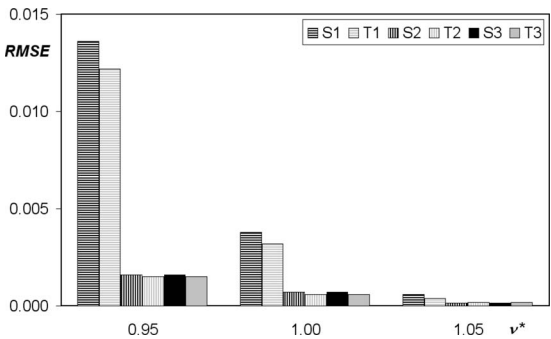


Fig. 14 Error on mass flow rate for scaling and translation for fouling ($\Delta A^* = -10\%$ and $\Delta \eta^* = -5\%$) occurring in the compressor first stage

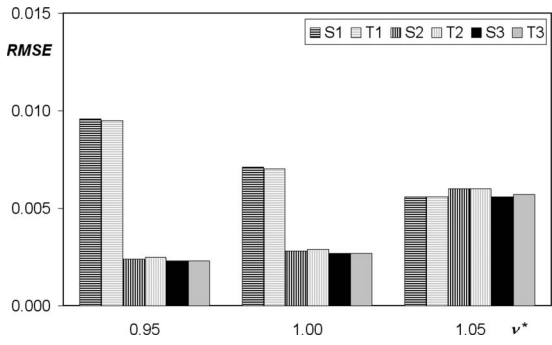


Fig. 15 Error on efficiency for scaling and translation for fouling ($\Delta A^* = -10\%$ and $\Delta \eta^* = -5\%$) occurring in the compressor first stage

rotational speed is increased. This result may be due to the fact that at high rotational speeds the curves are steeper and so they are less affected by the deformation consequent to the scaling. Moreover, translation and scaling are in practice equivalent.

In Fig. 15, the RMSE values on compressor efficiency are reported. In this case, the decrease of RMSE values by increasing the corrected rotational speed cannot be observed, since the efficiency trend is almost independent of the corrected rotational speed.

Figures 14 and 15 highlight that scaling factors depend on the corrected rotational speed and on the load (i.e., on the operating point chosen for map scaling). However, since the variation of the scaling factors F in the operating region close to the design corrected rotational speed is small, the use of the scaling factors as health indices can be considered acceptable for compressor health state determination close to the design condition. Moreover, also the use of scaled maps in order to represent compressor behavior in deteriorated conditions close to the design corrected rotational speed can be considered acceptable since, in this condition, the RMSE is lower than about 1%. Otherwise, these maps should not be used for the simulation of compressor behavior in deteriorated conditions at low corrected rotational speed, e.g., during gas turbine start-up [19], since RMSE values increase by decreasing the corrected rotational speed.

A similar analysis was carried out for the turbine. Figures 16 and 17 refer to the results that can be obtained in the case of fouling only affecting the first turbine stator. As usual, the results report the effects on both mass flow rate and efficiency.

Also in this case the scaling factors depend on the corrected rotational speed and on the load. In any case, the use of the scaling factors as health indices can be considered once again acceptable for turbine health state determination, since the variation in the scaling factors F is small. Moreover the use of scaled maps in order to represent turbine behavior in deteriorated conditions can

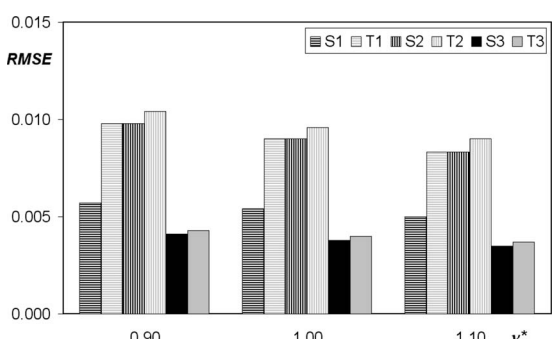


Fig. 16 Error on mass flow rate for scaling and translation for fouling ($\Delta A^* = -6\%$ and $\Delta Y^* = +3\%$) in the first turbine stator

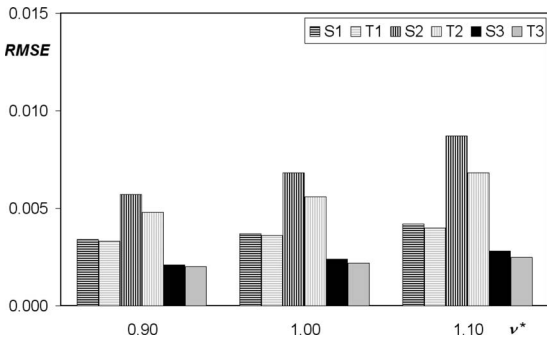


Fig. 17 Error on efficiency for scaling and translation for fouling ($\Delta A^* = -6\%$ and $\Delta v^* = +3\%$) in the first turbine stator

be considered acceptable since the *RMSE* is lower than about 1%.

Finally, it has to be underlined that the numerical results presented in this paragraph are strictly dependent on the shape of the characteristic curves. In particular, the compressor curves at constant corrected rotational speed are characterized by a quite wide range of corrected mass flow. This type of compressor maps was used in order to analyze one of the worst cases.

4 Conclusions

In this paper, stage-by-stage models of the compressor and turbine were used in order to assess the actual modification of compressor and turbine performance maps due to blade deterioration. The results obtained by simulating some of the most common causes of blade deterioration (i.e., compressor fouling, compressor mechanical damage, turbine fouling, and turbine erosion), which occur in one or more stages simultaneously, were reported in the paper.

In case of compressor fouling, the values of corrected mass flow rate in the choked region decreases linearly by decreasing the flow passage area. The rate of decrease depends on the number of stages affected by fouling: if only the first stage is fouled, the corrected mass flow rate decrease is approximately 2% for a 10% area decrease, while, when fouling occurs on whole compressor, the decrease is proportional to the flow passage area decrease.

In case of compressor mechanical damage, even if the area is not decreased, the corrected mass flow rate slightly decreases due to the decrease in efficiency.

In case of fouling and erosion affecting turbine stages, fouling causes a shift of the corrected mass flow rate and efficiency curves toward lower values, while erosion causes a shift of the corrected mass flow rate curves toward higher values with efficiency curves, which are in practice unaffected by this type of deterioration. It can be noted that, in both cases of fouling and erosion, the shift of the corrected mass flow rate and efficiency curves slightly depends on the number of stages affected by the degradation. However, when fouling or erosion only occurs in the first stator, the dependence of the corrected mass flow rate values in the choked region on the corrected rotational speed changes, decreasing in case of fouling and increasing in case of erosion. The dependence of corrected mass flow rate in the choked region from the variation in the flow passage area is approximately linear. In particular, for the case of a nondimensional corrected rotational speed equal to 1 and the area variation occurring only in the turbine first stator, the increase in the corrected mass flow rate in the choked region is about 0.8% for a 1.0% area increase due to erosion, and about 1.0% for a 1.0% area decrease due to fouling.

Finally, compressor and turbine maps obtained through the stage-by-stage procedure were compared with the ones obtained by means of map scaling. The maps built by means of the stage-by-stage compressor and turbine modeling were assumed to be the actual deteriorated curves. Healthy curves were then scaled in order to match the deteriorated curves. The results show that root

mean square errors for the cases analyzed are usually lower than about 1%. This analysis also highlights that scaling factors depend on the corrected rotational speed and on the load (i.e., on the operating point chosen for the map scaling). However, since the variation of the scaling factors in the operating region close to the design corrected rotational speed is small, the use of the scaling factors as health indices can be considered acceptable for gas turbine health state determination at full load. Moreover, also the use of scaled maps in order to represent compressor and turbine behavior in deteriorated conditions close to the design corrected rotational speed can be considered acceptable since, in this condition, the root mean square error is lower than about 1%.

It should be noted that the numerical results are specific to the engine considered, since they depend on the shape of the characteristic curves. However, since the compressor curves at constant corrected rotational speed are characterized by a quite wide range of corrected mass flow, the analyzed case represents one of the worst cases. In any case, the directional results should hold.

One of the main achievements of this paper is that stage-by-stage performance map modeling allowed the analysis of indirect and coupling effects of faults. For instance, compressor mechanical damage also leads to a decrease of the mass flow rate, though it is usually simulated by means of a variation in compressor efficiency only. On the other hand, stage-by-stage performance map modeling needs a huge amount of information, rarely available to GT users, about stage geometry and performance.

Acknowledgment

The work was carried out with the support of the MiUR (Italian Ministry of University and Research).

The authors gratefully acknowledge Professor Roberto Bettocchi for the suggestions provided during the work.

Nomenclature

A	=	area
c_p	=	specific heat at constant pressure
c_v	=	specific heat at constant volume
F	=	scaling factor
h	=	specific enthalpy
i	=	incidence angle
k	=	c_p/c_v
M	=	mass flow rate
n	=	number of stages
N	=	rotational speed
p	=	pressure
R	=	$(\Delta h)_{rot}/(\Delta h)_{stage}$ degree of reaction, gas constant
RH	=	relative humidity
RMSE	=	root mean square error
SF	=	shape factor
T	=	temperature
U	=	blade speed at the mean radius
V	=	absolute flow velocity
W	=	relative flow velocity
Y	=	total pressure loss coefficient
β	=	pressure ratio
Δ	=	variation
ϕ	=	V_a/U flow coefficient
Φ	=	$1/R \int_{T_r}^T c_p(T)(dT/T) + \Phi(T_r)$
η	=	efficiency
μ	=	$(M\sqrt{T_0})/p_0$ corrected mass flow rate
ν	=	$=N/\sqrt{T_0}$ corrected rotational speed
ψ_p	=	$\Delta h_{0s}/U^2$ pressure coefficient
Ψ	=	$\Delta h_0/U^2$ stage aerodynamic loading coefficient

Subscripts and Superscripts

- * = normalized value with respect to the reference value
- 0 = total physical state
- A = ambient, axial
- c = compressor
- l* = inlet conditions to *i*th stage, inlet section
- i*+1 = outlet conditions from *i*th stage=inlet conditions to (*i*+1)th stage
- max = maximum
- min = minimum
- O* = outlet section
- R* = reference value
- rot = rotor
- S* = isentropic
- stage = stage
- stat = stator
- t* = turbine
- th = throat

Appendix

Compressor. The overall multistage compressor performance maps (which relate overall pressure ratio β_C , efficiency η_C , corrected mass flow μ_C , and corrected rotational speed ν_C) were obtained through a stage-stacking procedure by using generalized stage performance curves (which, in turn, link together pressure coefficient ψ_p , stage efficiency η , and flow coefficient ϕ), which allow the stage-by-stage evaluation of the outlet conditions from the knowledge of those at inlet. The methodology was presented in detail in Ref. [13] to predict gas turbine performance maps starting from experimental data.

In order to model each compressor stage, generalized relationships between $\psi_p^* = \psi_p / (\psi_p)_r$, $\eta^* = \eta / (\eta)_r$, and $\phi^* = \phi / (\phi)_r$ were used. These relationships allow the complete evaluation of stage characteristics once the stage reference point $((\psi_p)_r, (\eta)_r, (\phi)_r)$ is known.

The first generalized relationship, $\psi_p^* = \psi_p^*(\phi^*, SF)$, was set up based on the work of Muir et al. [34], who obtained a single generalized curve by fitting experimental data points over a large number of compressor stages. In order to account for different stage characteristics based on different stage types (including transonic and supersonic stages), a parameter called shape factor SF was introduced to obtain a number of generalized curves covering all experimental data reported by Muir et al. [34]. In this manner, the equation of generalized curves becomes [13]

$$\psi_p^* = \psi_{p,\max}^* - \frac{(\psi_{p,\max}^* - 1) \cdot [\phi^* |_{\psi_{p,\max}^*} + SF \cdot (\phi^* |_{\psi_{p,\max}^*} - 1) - \phi^*]^2}{[\phi^* |_{\psi_{p,\max}^*} + SF \cdot (\phi^* |_{\psi_{p,\max}^*} - 1) - 1]^2} \quad (A1)$$

In Fig. 18 four curves $\psi_p^* = \psi_p^*(\phi^*, SF)$, obtained by using SF equal to -0.5 , -0.3 , 0.0 , and 1.0 , respectively, are shown, together with the experimental data points and the curve reported by Muir et al. [34]. The value of $SF = -0.3$ was adopted in this paper, since it allows a good approximation of the experimental stage data represented by a cross in Fig. 18 [35].

The second generalized relationship, $\eta^* = \eta^*(\psi_p^* / \phi^*)$, was obtained from the generalized stage efficiency curve proposed by Howell and Bonham [36]

$$\eta^* = 1 - \frac{1 - \eta^* |_{(\psi_p/\phi)_{\min}}}{\left[1 - \left(\frac{\psi_p^*}{\phi^*}\right)_{\min}\right]^{3.5}} \left(1 - \frac{\psi_p^*}{\phi^*}\right)^{3.5}, \quad \frac{\psi_p^*}{\phi^*} \in \left[\left(\frac{\psi_p^*}{\phi^*}\right)_{\min}, 1\right] \quad (A2)$$

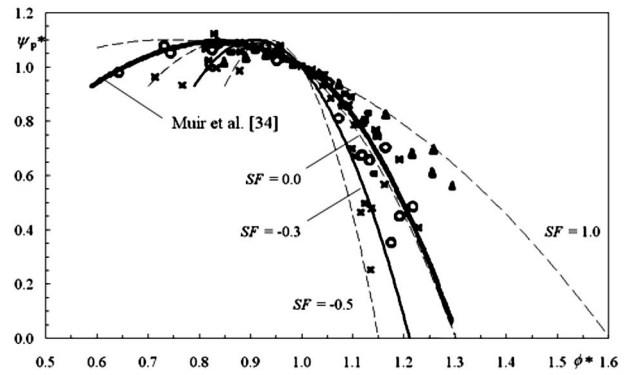


Fig. 18 Generalized stage characteristics $\psi_p^* = \psi_p^*(\phi^*, SF)$ [13] and experimental data [34]

$$\eta^* = 1 - \frac{1 - \eta^* |_{(\psi_p/\phi)_{\max}}}{\left[\left(\frac{\psi_p^*}{\phi^*}\right)_{\max} - 1\right]^2} \left(\frac{\psi_p^*}{\phi^*} - 1\right)^2, \quad \frac{\psi_p^*}{\phi^*} \in \left[1, \left(\frac{\psi_p^*}{\phi^*}\right)_{\max}\right] \quad (A3)$$

Figure 19 shows the generalized stage efficiency curve obtained by using Eqs. (A2) and (A3) compared with data obtained from the curve proposed by Howell and Bonham [36]. This single curve, together with the curves $\psi_p^* = \psi_p^*(\phi^*, SF)$, allows relationships $\eta^* = \eta^*(\phi^*, SF)$ to be established for all types of compressor stages.

The generalized stage characteristics allow the stage-by-stage evaluation of the outlet conditions starting from the inlet ones through the use of the following relationships:

$$h(T_{0(i+1)s}) = h(T_{0i}) + U_i^2 \psi_{pi} \quad (A4)$$

$$h(T_{0(i+1)}) = h(T_{0i}) + \frac{U_i^2 \psi_{pi}}{\eta_i} \quad (A5)$$

$$\frac{P_{0(i+1)}}{P_{0i}} = e^{[\Phi(T_{0(i+1)s}) - \Phi(T_{0i})]} \quad (A6)$$

$$\phi_{i+1} = \phi_i \frac{U_i}{U_{i+1}} \frac{A_i}{A_{i+1}} \frac{P_{0i}}{P_{0(i+1)}} \frac{T_{0(i+1)}}{T_{0i}} \quad (A7)$$

By “stacking” all the n stages, it is possible to evaluate

- the overall pressure ratio $\beta_C = P_{0(n+1)} / P_{01}$
- the overall total enthalpy variation $\Delta h_{0C} = h(T_{0(n+1)}) - h(T_{01})$

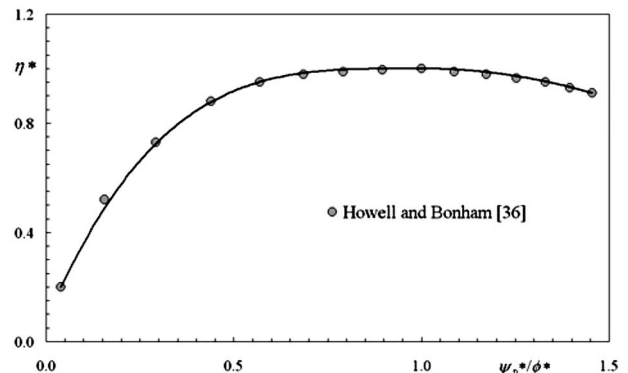


Fig. 19 Generalized stage efficiency curve

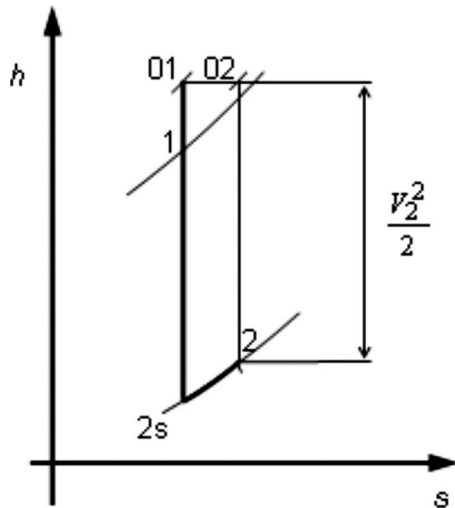


Fig. 20 Nozzle thermodynamic transformation

- the overall compressor isentropic efficiency

$$\eta_C = \frac{h(T_{0(n+1)s}) - h(T_{01})}{h(T_{0(n+1)}) - h(T_{01})}$$

where $T_{0(n+1)s}$ is determined by $\Phi(T_{0(n+1)s}) = \Phi(T_{01}) + \ln \beta_C$

Turbine. The procedure used to obtain the performance maps of a multistage turbine, which link together expansion ratio β_T (or the equivalent parameter $(\Delta h_{0s})_T / T_{0iT}$), efficiency η_T , corrected mass flow μ_T , and corrected rotational speed ν_T , is based on the following main assumptions, which were discussed in Ref. [13]:

- the splitting of the overall turbine enthalpy variation among turbine stages in design conditions based on the aerodynamic loading coefficient (Ψ_i) of each stage

$$(\Delta h_0)_i = \frac{U_i^2 \Psi_i}{\sum_{i=1}^n U_i^2 \Psi_i} \cdot (\Delta h_0)_T \quad (A8)$$

- the splitting of the stage enthalpy variation between stator and rotor in design conditions based on the degree of reaction of each stage (R_i), by assuming that $(\Delta h)_i \cong (\Delta h_0)_i$

$$(\Delta h)_{\text{rot}} = R_i \cdot (\Delta h)_i \quad (A9)$$

$$(\Delta h)_{\text{stat}} = (\Delta h)_i - (\Delta h)_{\text{rot}} \quad (A10)$$

- the calculation of the absolute and relative flow angles in design conditions as a function of ϕ , Ψ , and R , on the hypothesis that axial velocity is constant through the stage
- the mathematical approximation of turbine cascade losses as a function of incidence angle i (which, in turn, depends on flow and stagger angles) by means of the following generalized relationships:

$$Y^* = 1 + \frac{Y_{i_{\min}}^* - 1}{(i_{\min}^* - 1)^2} (i^* - 1)^2, \quad i^* \in [i_{\min}^*, 1] \quad (A11)$$

$$Y^* = 1 + \frac{Y_{i_{\max}}^* - 1}{(i_{\max}^* - 1)^2} (i^* - 1)^2, \quad i^* \in [1, i_{\max}^*] \quad (A12)$$

- the approximation of the mass flow rate characteristics of each turbine cascade by means of nonisentropic converging nozzle characteristics

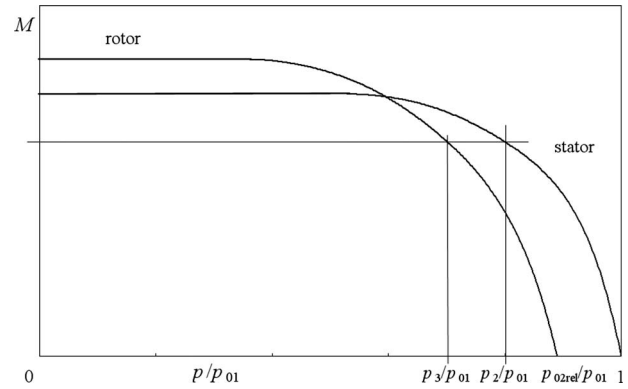


Fig. 21 Example of stator and rotor matching

$$M = \frac{p_2}{RT_2} \cdot V_2 \cdot A_{\text{th}} \quad (A13)$$

where p_2 , T_2 , and V_2 are calculated starting from the physical state 01, the pressure ratio p_{01}/p_2 and the losses Y across the cascade, by using the following relationships (see Fig. 20):

$$\frac{p_{01}}{p_2} = e^{[\Phi(T_{01}) - \Phi(T_{2s})]} \quad (A14)$$

$$Y = \frac{p_{01} - p_{02}}{p_{02} - p_2} \quad (A15)$$

$$\frac{p_{02}}{p_2} = e^{[\Phi(T_{01}) - \Phi(T_2)]} \quad (A16)$$

$$V_2 = \sqrt{2(h_{01} - h_2)}, \quad \text{where } V_2 \leq \sqrt{kRT_2} \quad (A17)$$

Choking conditions are reached when the corrected mass flow μ_1 assumes its maximum value.

It can be noted that stator and rotor cascades can be mathematically modeled in the same manner, if absolute and relative velocities are considered for stator and rotor, respectively.

The overall performance maps of the multistage turbine are obtained by matching the mass flow characteristics of each turbine cascade [13]. An example of stator and rotor matching is sketched in Fig. 21. The stator characteristic is plotted as

$$M = \frac{(A_{\text{th}})_{\text{stat}} \cdot p_{01}}{\sqrt{T_{01}}} \cdot F\left(\frac{p}{p_{01}}\right) \quad (A18)$$

For each M value, the static pressure at the stator exit (p_2) and, from the velocity triangle, the total relative physical state at the rotor inlet ($T_{02\text{rel}}, p_{02\text{rel}}$) can be evaluated. This state is used to evaluate the rotor mass flow characteristic (which, reported in $(p/p_{01}, M)$ coordinates, appears as in Fig. 21), which, in turn, allows the evaluation of the static pressure at the rotor exit p_3 .

References

- [1] Kurz, R., Brun, K., and Wollie, M., 2008, "Degradation Effects on Industrial Gas Turbines," ASME J. Eng. Gas Turbines Power, **131**(6), p. 062401.
- [2] Stamatis, A., Mathioudakis, K., and Papailiou, K. D., 1990, "Adaptive Simulation of Gas Turbine Performance," ASME J. Eng. Gas Turbines Power, **112**, pp. 168–175.
- [3] Bettocchi, R., and Spina, P. R., 1999, "Diagnosis of Gas Turbine Operating Conditions by Means of the Inverse Cycle Calculation," ASME Paper No. 99-GT-185.
- [4] Gulati, A., Zedda, M., and Singh, R., 2000, "Gas Turbine Engine and Sensor Multiple Operating Point Analysis Using Optimization Techniques," Proceedings of the 36th AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, Huntsville, AL, Jul. 16–19, Paper No. AIAA 2000-3716.
- [5] Saravanamuttoo, H. I. H., and Lakshminarasimha, A. N., 1985, "A Preliminary

- Assessment of Compressor Fouling," ASME Paper No. 85-GT-153.
- [6] Aker, G. F., and Saravanamuttoo H. I. H., 1988, "Predicting Gas Turbine Performance Degradation due to Compressor Fouling Using Computer Simulation Techniques," ASME Paper No. 88-GT-206.
- [7] Seddigh, F., and Saravanamuttoo, H. I. H., 1990, "A Proposed Method for Assessing the Susceptibility of Axial Compressors to Fouling," ASME Paper 90-GT-348.
- [8] Tabakoff, W., Lakshminarasimha, A. N., and Pasin, M., 1990, "Simulation of Compressor Performance Deterioration Due to Erosion," ASME J. Turbomach., **112**, pp. 78–83.
- [9] Massardo, A., 1991, "Simulation of Fouled Axial Multistage Compressors," IMechE Paper No. C423/048.
- [10] Cerri, G., Salvini, C., Procacci, R., and Rispoli, F., 1993, "Fouling and Air Bleed Extracted Flow Influence on Compressor Performance," ASME Paper No. 93-GT-366.
- [11] Lakshminarasimha, A. N., Boyce, M. P., and Meher-Homji, C. B., 1994, "Modeling and Analysis of Gas Turbine Performance Deterioration," ASME J. Eng. Gas Turbines Power, **116**, pp. 46–52.
- [12] Procacci, R., and Rispoli, F., 1995, "Off Design Performance Evaluation of Deteriorated Variable Geometry Axial Flow Compressors," ASME Paper No. 95-CTP-35.
- [13] Spina, P. R., 2002, "Gas Turbine Performance Prediction by Using Generalized Performance Curves Of Compressor And Turbine Stages," ASME Paper No. GT-2002-30275.
- [14] Hale, A. A., and Davis, M. W., 1992, "Dynamic Turbine Engine Compressor Code DYNTECC—Theory and Capabilities," *Proceedings of the 28th AIAA/SAE/ASME/ASEE Joint Propulsion Conference and Exhibit*, Nashville, TN, Jul. 6–8, Paper No. AIAA-92-3190.
- [15] Schobeiri, M. T., Attia, M., and Lippe, C., 1994, "GETRAN: A Generic, Modularly Structured Computer Code for Simulation of Dynamic Behavior of Aero- and Power Generation Gas Turbine Engines," ASME J. Eng. Gas Turbines Power, **116**, pp. 483–494.
- [16] Owen, A. K., Daugherty, A., Garrard, D., Reynolds, H. C., and Wright, R. D., 1999, "A Parametric Starting Study of an Axial-Centrifugal Gas Turbine Engine Using a One-Dimensional Dynamic Engine Model and Comparisons to Experimental Results: Part I—Model Development and Facility Description," ASME J. Eng. Gas Turbines Power, **121**, pp. 377–383.
- [17] Theotokatos, G., and Kyratos, N. P., 2003, "Investigation of a Large High-Speed Diesel Engine Transient Behaviour Including Compressor Surging and Emergency Shutdown," ASME J. Eng. Gas Turbines Power, **125**, pp. 580–589.
- [18] Morini, M., Pinelli, M., and Venturini, M., 2009, "Analysis of Biogas Compression System Dynamics," *Appl. Energy*, **86**(2009), pp. 2466–2475.
- [19] Morini, M., Cataldi, G., Pinelli, M., Venturini, M., 2007, "A Model for the Simulation of Large-Size Single-Shaft Gas Turbine Start-Up Based on Operating Data Fitting," ASME Paper No. GT2007-27373.
- [20] Thermoflow Inc., 2007, THERMOFLOW 17, Release 1, Sudbury, MA.
- [21] Saravanamuttoo, H. I. H., and Mac Isaac, B. D., 1983, "Thermodynamic Models for Pipeline Gas Turbine Diagnostics," ASME J. Eng. Gas Turbines Power, **105**, pp. 875–884.
- [22] Kurzke, J., and Riegler, C., 2000, "A New Map Scaling Procedure for Preliminary Conceptual Design of Gas Turbines," ASME Paper No. 2000-GT-0006.
- [23] Stone, A., 1958, "Effects of Stage Characteristics and Matching on Axial Flow Compressor Performance," *Trans. ASME*, **80**, pp. 1273–1293.
- [24] Doyle, M. D., and Dixon, S. I., 1962, "The Stacking of Compressor Stage Characteristics to Give an Overall Compressor Performance Map," *Aeronaut. Q.*, **13**(4), pp. 349–367.
- [25] Robbins, W. H., and Dugan, J. F., 1965, "Prediction of Off-Design Performance of Multi-Stage Compressors," NASA Report No. SP-36.
- [26] Howell, A. R., and Calvert, W. J., 1978, "A New Stage Stacking Technique for Axial-Flow Compressor Performance Prediction," ASME J. Eng. Power, **100**, pp. 698–703.
- [27] Bagnoli, M., Bianchi, M., Melino, F., and Spina, P. R., 2008, "Development and Validation of a Computational Code for Wet Compression Simulation of Gas Turbines," ASME J. Eng. Gas Turbines Power, **130**, p. 012004.
- [28] Bagnoli, M., Bianchi, M., Melino, F., Peretto, A., Spina, P. R., Bhargava, R., and Ingistov, S., 2008, "Application of a Computational Code to Simulate Interstage Injection Effects on GE Frame 7EA Gas Turbine," ASME J. Eng. Gas Turbines Power, **130**, p. 012001.
- [29] Zhu, P., and Saravanamuttoo, H. I. H., 1992, "Simulation of an Advanced Twin-Spool Industrial Gas Turbine," ASME J. Eng. Gas Turbine Power, **114**, pp. 180–186.
- [30] Zwebek, A., and Pilidis, P., 2001, "Degradation Effects on Combined Cycle Power Plant Performance Part 1: Gas Turbine Cycle Component Degradation Effects," ASME Paper No. 2001-GT-0388.
- [31] Pinelli, M., and Venturini, M., 2002, "Application of Methodologies to Evaluate the Health State of Gas Turbines in a Cogenerative Combined Cycle Power Plant," ASME Paper GT-2002-30248.
- [32] Suder, K. L., Chima, R. V., Strazisar, A. J., and Roberts, W. B., 1995, "The Effect of Adding Roughness and Thickness to a Transonic Axial Compressor Rotor," ASME J. Turbomach., **117**(4), pp. 491–505.
- [33] Morini, M., Pinelli, M., Spina, P. R., and Venturini, M., 2009, "CFD Simulation of Fouling on Axial Compressor Stages," ASME Paper No. GT2009-59025.
- [34] Muir, D. E., Saravanamuttoo, H. I. H., and Marshall, D. J., 1989, "Health Monitoring of Variable Geometry Gas Turbines for the Canadian Navy," ASME J. Eng. Gas Turbines Power, **111**, pp. 244–250.
- [35] Budinger, R. E., and Kaufman, H. R., 1955, "Investigation of the Performance of a Turbojet Engine with Variable-Position Compressor Inlet Guide Vanes," Report No. NACA RM E54L23a.
- [36] Howell, A. R., and Bonham, R. P., 1950, "Overall and Stage Characteristics of Axial Flow Compressors," *Proc. Inst. Mech. Eng., IMechE Conf.*, **163**, pp. 235–248.

Synchronous Response to Rotor Imbalance Using a Damped Gas Bearing

Bugra H. Ertas
Rotating Equipment Group,
Vibration and Dynamics Laboratory,
GE Global Research Center,
Niskayuna, NY 12309
e-mail: ertas@research.ge.com

Massimo Camatti
Conceptual Design,
Advanced Technology,
GE Oil and Gas,
Florence, Italy FI-50127
e-mail: massimo.camatti@ge.com

Gabriele Mariotti
Mechanical Design,
Turboexpander Division,
GE Oil and Gas,
Florence, Italy FI-50127
e-mail: gabriele.mariotti@ge.com

One type of test performed for evaluating bearings for application into turbomachinery is the synchronous bearing response to rotor imbalance. This paper presents rotordynamic tests on a rotor system using a 70 mm diameter damped gas bearing reaching ultra-high speeds of 50,000 rpm. The main objective of the study was to experimentally evaluate the ability of the damped gas bearing to withstand large rotor excursions and provide adequate damping through critical speed transitions. Two critical speeds were excited through varying amounts and configurations of rotor imbalance while measuring the synchronous rotordynamic response at two different axial locations. The results indicated a well-damped rotor system and demonstrated the ability of the gas bearing to safely withstand rotor vibration levels while subjected to severe imbalance loading. Also, a waterfall plot was used to verify ultra-high-speed stability of the rotor system throughout the speed range of the test vehicle. In addition to the experimental tests, a rotordynamic computer model was developed for the rotor-bearing system. Using the amplitude/frequency dependent stiffness and damping coefficients for the ball bearing support and the damped gas-bearing support, a pseudononlinear rotordynamic response to imbalance was performed and compared with the experiments. [DOI: 10.1115/1.3157097]

1 Introduction

Rotor-bearing systems in high-speed turbomachinery are typically required to traverse several natural frequencies before reaching nominal operating speed ranges. Safe and reliable machine operation above and through critical speeds strongly depends on the bearings that support the rotating assembly. Historically, high performance turbomachines have used oil-based bearing and damper technologies due to requirements for load capacity and damping. Aircraft gas-turbine engines rely on soft-mounted rolling element bearings using squirrel cage supports in combination with squeeze film dampers to bound rotor vibration through critical speeds and to provide useful subsynchronous damping at high operating speeds [1,2]. Turbomachinery in the oil and gas industry, such as high pressure centrifugal compressors and turboexpanders, typically incorporate oil-lubricated tilting pad journal bearings and sometimes use squeeze film bearing support dampers to provide rotordynamic reliability [3,4]. Ever increasing demands for power density and efficiency are pushing oil-requiring bearing technologies to their limits. One approach to meeting industry demands is by replacing oil-requiring bearings with oil-free bearings that use the working gas of the machine as lubrication. Gas bearings, when designed properly, can allow for very high rotor speeds while generating negligible power loss. Gases also can operate at extreme temperatures whereas oil has temperature limitations. Additionally, with the removal of the oil system, machine design is simplified reducing part count and enhancing overall system reliability. These advantages of gas lubrication can allow for nonstandard rotor architectures and operating conditions that result in better efficiency, higher power density, and lowered cost. However, when considering application into high performance turbomachinery, traditional gas-bearing systems [5] have inherent disadvantages such as low load capacity, insufficient damping, and/or compromised tolerance to misalignment and changes in

rotor geometry. To address these risks with traditional bearing designs, a compliant hybrid gas bearing using oil-free bearing support dampers was developed [6]. The intent of this bearing concept was to enable the oil-free operation of heavy loaded turbomachinery operating through a wide speed range. Therefore, the bearing design incorporates external pressurization for 0 rpm lift, compliant bearing pad surfaces for tolerance to misalignment, and dampers in the bearing support for providing enhanced damping. Initial studies [6] focused on component level testing involving the measurement of rotordynamic force coefficients, determination of load capability for different modes of operation, and verification of the design concept. The next step was to integrate the gas bearing into a rotor system as the main support bearing.

The present study evaluates a compliant hybrid gas bearing on a system level by experimentally observing the influence of rotor imbalance on vibration magnitudes, critical speed Q factors, and system stability. Experiments were performed using moderate to extreme levels of rotor imbalance. A secondary objective was to develop a rotordynamic model that predicts the dynamics of the rotor system using nonlinear bearing stiffness and damping coefficients.

2 Experimental Setup

The experiments were conducted on an existing test rig initially developed for evaluating the performance of rolling element bearings and squeeze film dampers. The test rig was modified to allow testing of a damped gas lubricated journal bearing in the aft location of the rotor. The modified test rig design and cross section for supporting an oil-free bearing is shown in Figs. 1 and 2. The test rotor was operated through a two set gear train powered by a 75 hp dc motor, where it was limited to a speed of 50,000 rpm and driven by a tight fitting spline coupling. Rotor thrust was generated using a thrust cavity, shop air was used to preload the ball bearing (350 lb axial preload), and a different air supply was used to provide external pressurization (185 psi (gauge)) to the hybrid gas bearing. Rotor vibration was measured at two different axial locations. As shown in Fig. 1, both horizontal and vertical vibration probes were installed at the forward (fwd) and aft end. A light probe at the end of the shaft was used as the key-phaser signal for

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 19, 2009; final manuscript received March 22, 2009; published online December 1, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Gas Turbine Technical Congress and Exposition, Orlando, FL, June 8–12, 2009.

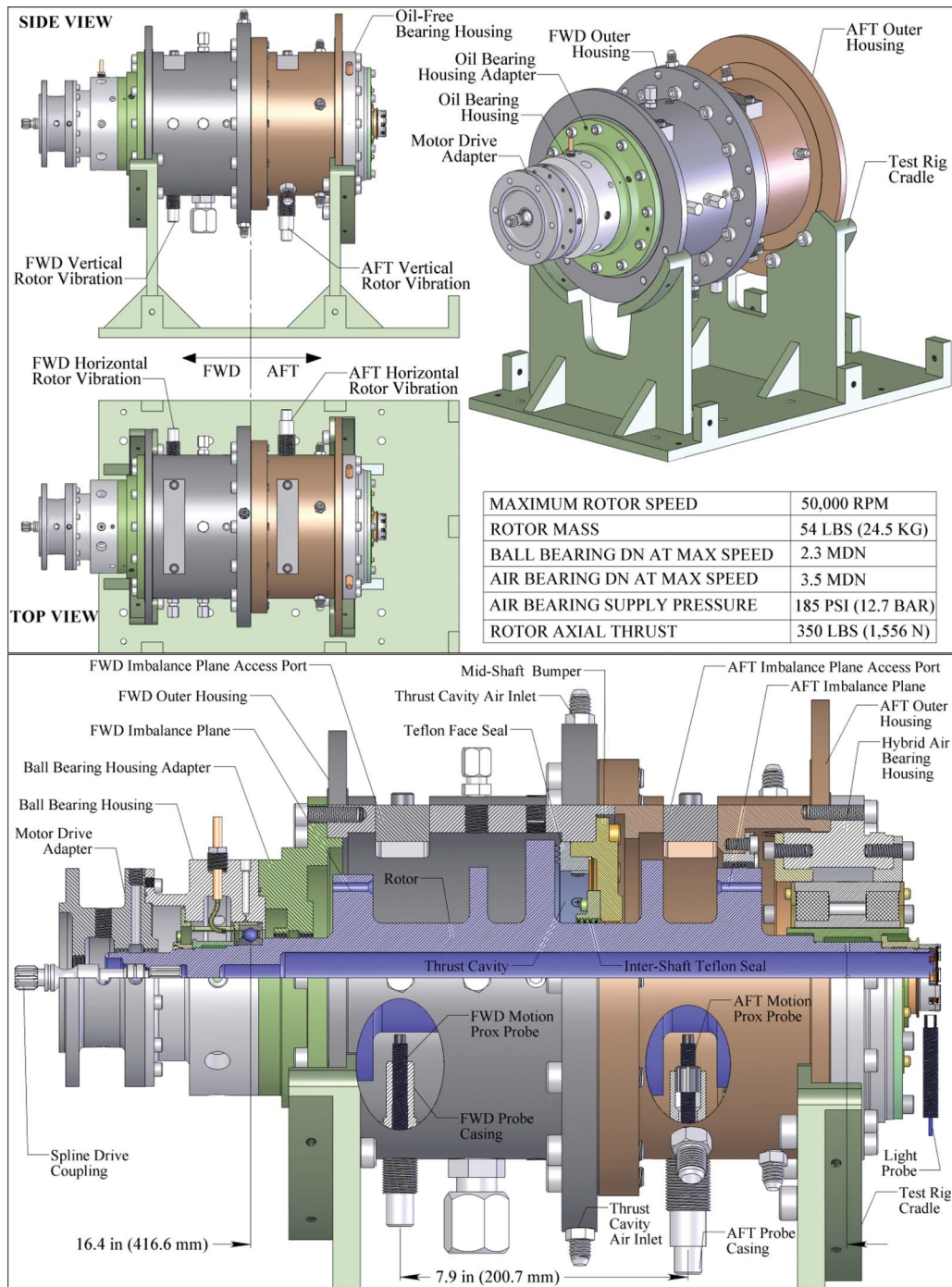


Fig. 1 Test rig layout and cross section

monitoring rotational speed. Two ports on the top of the fwd and aft outer housings were designed as access to the threaded imbalance holes on two disks for exciting the first two critical speeds of the rotor.

The forward end rotor support used an angular contact ball bearing mounted in a squirrel cage centering spring with an unsealed center-fed squeeze film damper (SFD). The squirrel cage, which is in series with the ball bearing and in parallel with the squeeze film damper, is shown in Fig. 2. The squirrel cage provides the rotor with static centering support and is an essential component for locating rotor criticals. For this bearing support to work effectively the squirrel cage radial stiffness is designed to be significantly lower than the radial stiffness of the ball bearing. By

“soft-mounting” the ball bearing in series with the SFD the equivalent damping of the bearing system is maximized, whereas bearing dynamic loads and stresses are minimized [1]. Also, soft-mounting (for straddle mounted rotors) locates the first two rotor criticals (rigid body modes) at low operating speeds, while positioning the third rotor critical (bending mode) at a high speed, potentially allowing for a wide operating speed margin without encountering rotor resonances.

The aft end of the rotor is supported by a compliant hybrid gas bearing using oil-free wire mesh dampers (WMDs) [6] in the bearing support. The general bearing design is shown in Fig. 3 and is comprised of four individual bearing pads that interface with the bearing housing through two integral “S” springs. Each pad is

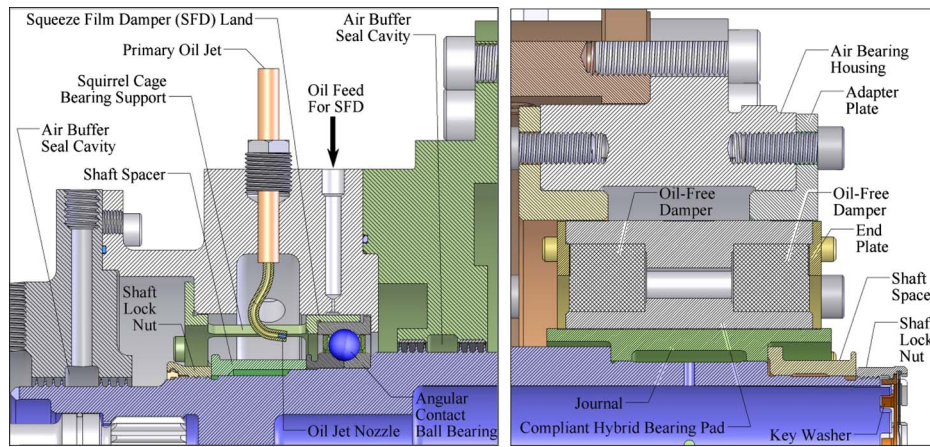


Fig. 2 Rotor support configurations

“soft-mounted” to the bearing outer rim and has the capability of being hydrostatically pressurized. The external hydrostatic pressure through the bearing pads provides 0 rpm rotor lift, high load capacity at low shaft speeds where hydrodynamic effects are weak, and a stiff gas film in comparison to the integral spring radial stiffness. The integral centering springs provide compliance to misalignment, rotor excursions, and deformation in the rotor geometry. The springs also enable tilting pad motion, which, in turn, minimizes stiffness cross-coupling and maximizes the onset speed of instability. One important function of the integral springs is to provide centering of the rotor assembly. Due to the geometric variability of the wire mesh and the tendency to creep over time, the two integral centering springs are a crucial design feature that ensures consistent and reliable centering of the rotating assembly. This helps maintain operating clearance closures, which can influence the power output and efficiency of the machine.

To introduce damping into the bearing support two integral wire mesh dampers are inserted on each axial side while being clamped in place between the end plate and the damper bridge. Wire mesh dampers have primarily been used as dampers for rolling element

bearings [7–9] but recently have been incorporated into gas lubricated bearing systems [6]. These dampers are an oil-free damping solution that can operate at extreme temperatures [10,11], lack requirements for auxiliary or ancillary devices, do not require the sealing that an SFD necessitates, and are low cost.

3 Rotor Imbalance

Imbalance in rotating systems is the most common cause of machine vibration and results from the center of mass of the rotating system being offset from the axis of rotation [12]. Selectively or collectively this uneven radial mass distribution can materialize from several sources such as manufacturing, nonhomogeneity of the material, offset of the geometric centerline with respect to the axis of rotation, rotor asymmetry, thermal distortion, stack-up errors, bent shaft, and corrosion. Several balance specifications have been developed, the most common of which are the International Standards Organization (ISO) [13] and the American Petroleum Institute (API) [12] as shown in Fig. 4. The ISO standard was developed to incorporate a wide range of rotat-

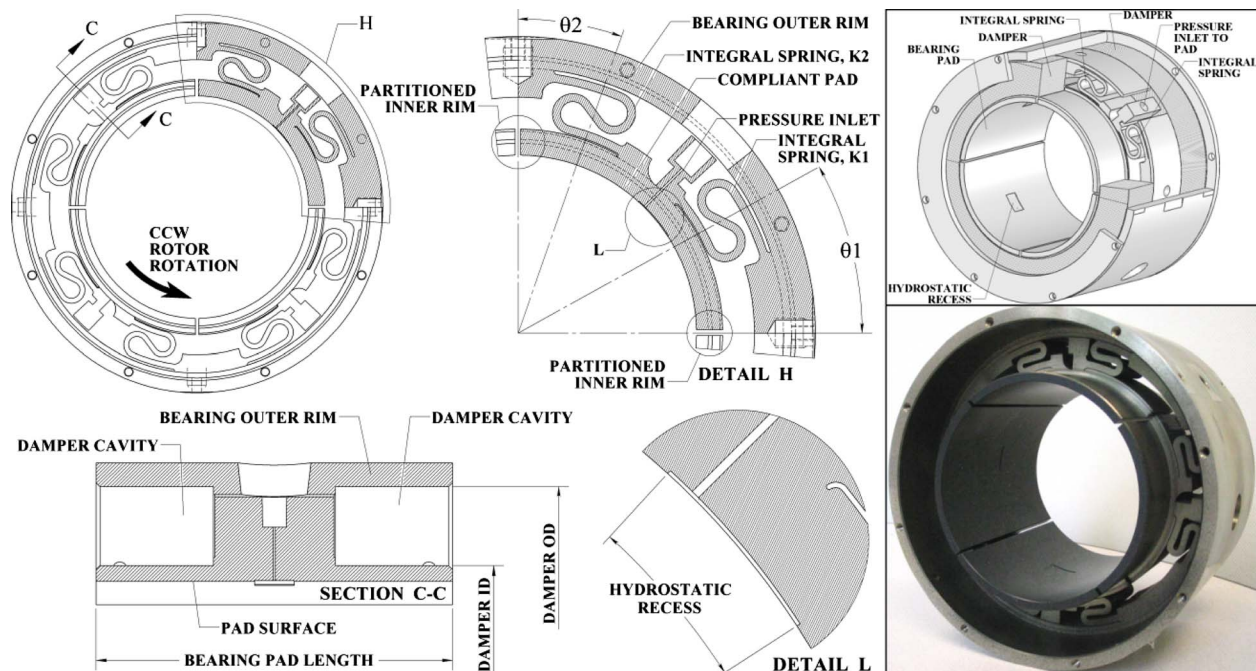


Fig. 3 Compliant hybrid gas bearing using integral wire mesh bearing support dampers [6]

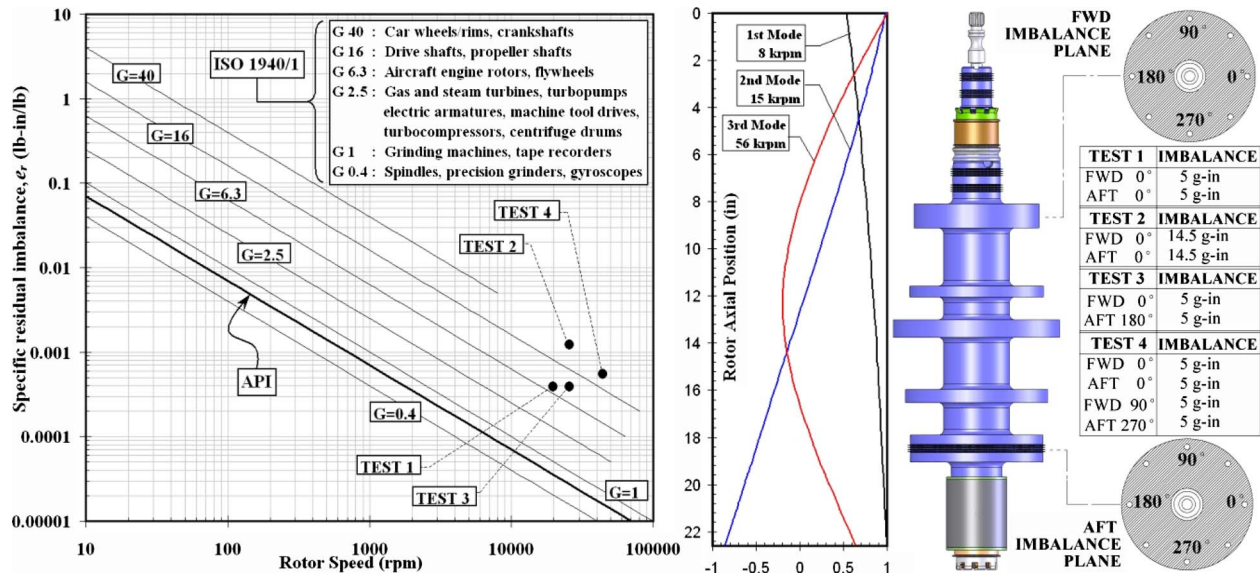


Fig. 4 Machine imbalance grades, rotor modes, and experimental test matrix

ing equipment from car wheels to gyroscopes, where turbomachinery falls between balance grades G2.5 and G6.3. The API specification on the other hand puts forth a more stringent requirement on residual imbalance in turbomachinery, which has an equivalent ISO imbalance grade of G0.7. Nevertheless, the objective of both standards is the same, which is to provide the end user with a reliable mechanical system throughout the life of the machine.

The test matrix is shown in Fig. 4 and defines four different imbalance configurations. The rotor system has two critical speeds in the operating speed range. Both modes are virtually rigid body modes (very little shaft bending) where the first critical is a bounce mode and the second is a conical mode with antinodes of vibration at the ends of the rotor. The third critical speed or bending mode was calculated to be at 56 krpm, well beyond the speed capability of the test rig. The first and second tests involved exciting the first bounce mode using two different levels of imbalance as shown in Fig. 4. By placing unbalanced weights at the same phase angle on both the aft and fwd imbalance planes the bounce mode was excited. The third test required placing weights 180 deg apart for exciting the conical mode. The final test (Test 4) used a combinational imbalance configuration so that both the first and second modes were excited. The four test points are plotted with respect to the ISO and API standards. Considering helicopter turboshaft engine rotor systems, the tests shown in Fig. 4 represent a range of typical operating imbalances over time (Test 1) to severe imbalance loading (Test 2). For example, a typical high-speed gas generator spool weighing 50 lbs is built to have a residual imbalance balance of 0.8 g in. and the low speed power turbine spool weighing 47 lbs is balanced to 0.4 g in. For oil and gas turbomachinery the API specification drives residual imbalance, however during the design phase, rotor systems are often engineered to take 4–20 times API depending on the application. Centrifugal compressors are designed using 4 times API and over time can experience 10 times API rotor imbalance limits in extreme field cases. Turboexpanders working with multiphase fluids often use imbalance magnitudes reaching 20 times API to verify designs. The tests shown here subject the gas bearing to imbalance loading up to 50 times API.

Before performing the four imbalance tests listed in Fig. 4, a baseline test was required (Fig. 5). The top two plots in Fig. 5 illustrate the horizontal and vertical synchronous vibration responses of the rotor at the fwd and aft probe locations of up to 50,000 rpm with slow roll compensation. The vibration was mea-

sured to be low due to the low residual baseline imbalance of 0.20 g in., making it difficult to discern critical speed peaks. The lower plot shows the waterfall of the aft vertical probe and shows that the main vibration is the synchronous vibration. Note that the modes at 8 krpm and 15 krpm are not subsynchronously excited, indicating an inherently stable bearing system. There is however a small $\frac{1}{4}$ tracking subharmonic frequency, which is hypothesized to come from the gearbox drive.

4 Rotordynamic Model

This section describes the rotor model, bearing support parameters, and the method used to generate the rotordynamic response. The finite element model of the rotor is shown in Fig. 6, where station 22 represents the location of the fwd probe and station 39 represents the aft probe location. The rotor model interfaces with ground through the two bearings and a spline coupling (station 3). The spline coupling was modeled with only a stiffness ($k_S = 20$ klb/in.) calculated using a separate finite element model.

The aft end of the rotor is supported at station 49 by the oil-free bearing. The fwd end of the rotor has a two level support, where the ball bearing is linked from station 13 to station 55 and the squirrel cage bearing support is linked from station 55 to ground. The aft and fwd bearing support parameters are shown in Figs. 7 and 8. Note that the equivalent fwd bearing support stiffness k_{FWD} is the sum of the squirrel cage radial stiffness k_{SC} and the squeeze film stiffness k_{SFD} (Eq. (1)). Similarly, the equivalent aft bearing support stiffness k_{AFT} is the sum of the gas-bearing integral spring stiffness values ($k_{IS} = k_1 + k_2$) and the WMD stiffness k_{WMD} (Eq. (2)). Since the squeeze film damper has a clearance, which can collapse at high vibration amplitudes, it was necessary to link the ball bearing stiffness and damping ($k_{BB} = 250$ klb/in. and $c_{BB} = 3$ lb s/in.) in series with the bearing support (squirrel cage and SFD). The imbalance loading for the forced synchronous response calculation was based on the experimental test matrix (Fig. 4) and linked to stations 20 and 42 of the rotor model (Fig. 6).

Figures 9 and 10 show the fwd and aft damper force coefficients used in the analysis. The squeeze film coefficients were generated using Eqs. (3) and (4). Equations (3) and (4) are derived from short bearing theory and are valid for short length unsealed (open ends) dampers with centered circular orbits using a half circumferential film cavitation assumption and no inertia effects [14]. Squeeze film dampers are sometimes referred to as possessing “dynamic stiffness,” meaning that “stiffness” like behavior is

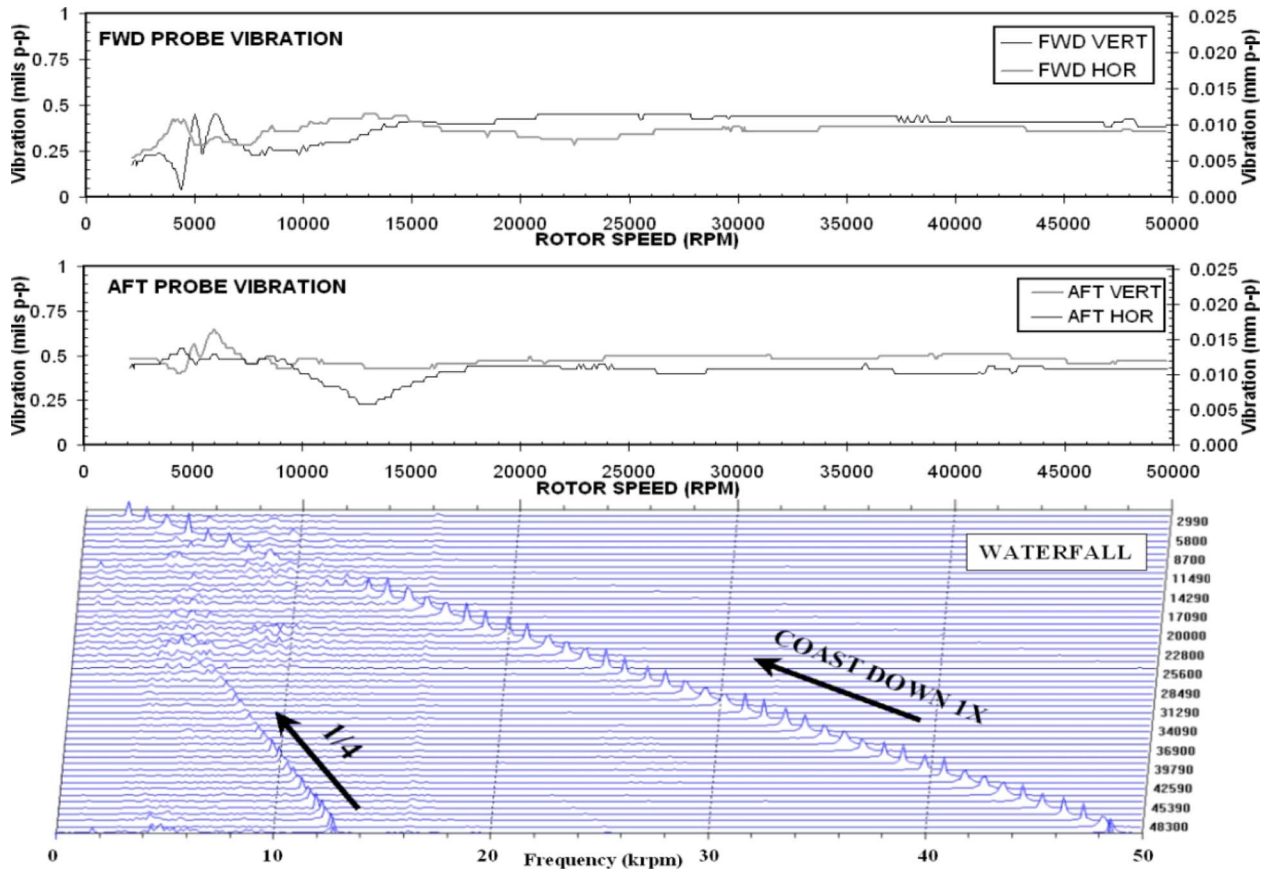


Fig. 5 Baseline test with 0.20 g in. rotor imbalance: 8.1×10^{-6} lb in./lb specific residual imbalance

generated with damper whirl velocity and therefore is frequency dependent. Dynamic stiffness k_{SFD} as shown in Eq. (3) is really cross-coupled damping C_{RT} [14] multiplied by the frequency and portrays a hardening spring behavior with increasing amplitudes

and frequencies. The squeeze film damping is defined in Eq. (4) and is frequency independent but has a nonlinear dependence in relation to whirl amplitude.

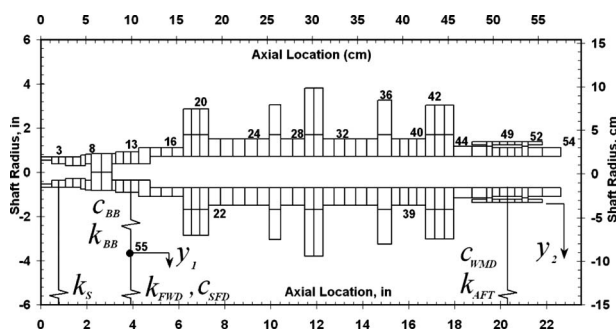


Fig. 6 Rotor model

Type	Center Feed, Unsealed
Damper radius, R	1.530 in (38.9 mm)
Effective length, L	0.500 in (12.7 mm)
Damper clearance, C	0.0032 in (0.081 mm)
Oil type	MIL PRF 23698
Oil temperature	165 F
Oil viscosity @ 165 F, μ	2.6×10^{-6} lbf-s/in ²
Squirrel cage stiffness, k_{sc}	20 klb/in (3.5 MN/m)

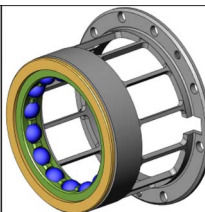


Fig. 7 SFD and squirrel cage design parameters

Material	Copper
Mesh Density	25%
Damper OD	4.75 in (121 mm)
Damper ID	3.15 in (80 mm)
Damper width, W	0.89 in (22 mm)
Qty per bearing	2
Spring stiffness, K_1	11.5 klb/in (2 MN/m)
Spring stiffness, K_2	11.5 klb/in (2 MN/m)

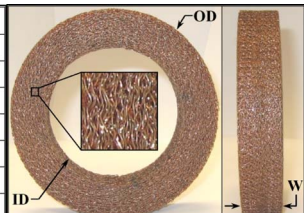


Fig. 8 WMD and integral spring design parameters

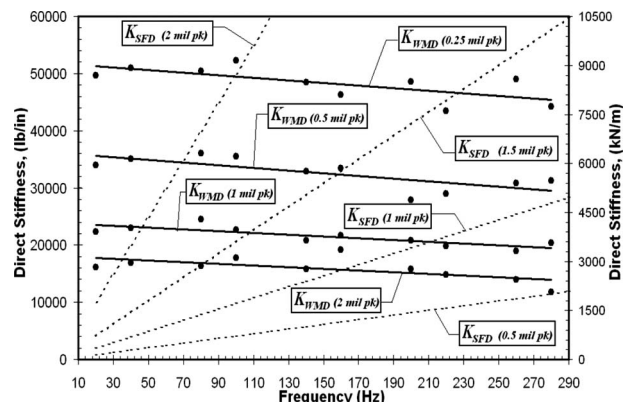


Fig. 9 SFD and WMD nonlinear stiffness coefficients

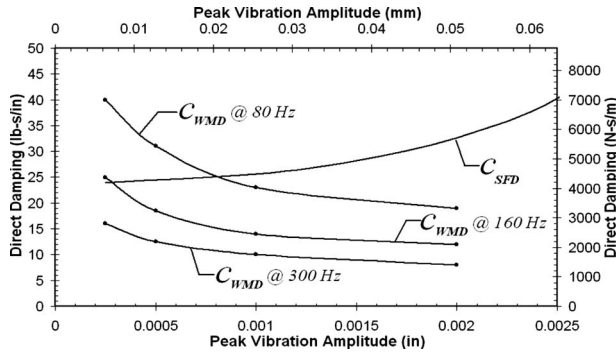


Fig. 10 SFD and WMD nonlinear damping coefficients

The WMD stiffness and damping coefficients for quantity 2 dampers were generated through bench-top testing [15] while controlling frequency and vibration amplitude. The gas film force coefficients were not taken into account, therefore the aft bearing support is modeled only with the WMD and integral springs. The integral springs are modeled to have a constant radial stiffness adding to the WMD stiffness while contributing no damping. As shown in Fig. 9 the WMD were measured to possess a softening spring characteristic with increasing vibration amplitude while showing slight dependence on excitation frequency. This is consistent with past studies using copper WMD [16,17,9]. The copper wire mesh damping (Fig. 10) shows both amplitude and frequency dependences where the damping decreases with increasing vibration amplitude and frequency. It has been shown that traditional wire mesh materials such as stainless steel, Inconel, and copper portray this type of stiffness and damping behavior [15]. However, it is important to stress that the reduction in damping with increasing amplitude is not linear. Meaning that there is a minimum damping value that is reached with increasing vibration amplitude and frequency (the damping does not become zero or negative with increasing vibration amplitude/frequency). In fact, preliminary laboratory tests have indicated a reversal in damping behavior for large amplitudes, most likely due to the appreciable change in mesh density at high vibration amplitudes. Also, note that as the damping drops with vibration so does the stiffness. Therefore, impact on the modal damping ratio, with increasing vibration amplitudes, is small.

$$k_{FWD} = k_{SC} + k_{SFD} \quad (1)$$

$$k_{AFT} = k_{IS} + k_{WMD} = k_1 + k_2 + k_{WMD} \quad (2)$$

$$k_{SFD} = C_{RT} \cdot \omega = \frac{2 \cdot \mu \cdot R \cdot L^3 \cdot \omega \cdot \varepsilon}{C^3 \cdot (1 - \varepsilon^2)^2} \quad (3)$$

$$c_{SFD} = C_{TT} = \frac{\mu \cdot R \cdot L^3 \cdot \pi}{2 \cdot C^3 \cdot (1 - \varepsilon^2)^{3/2}} \quad (4)$$

Although the rotor finite element model and synchronous forced response calculation are linear, the bearing stiffness and damping coefficients are nonlinear, as they are functions of vibration amplitude and frequency. Therefore, to capture the nonlinear behavior an iterative method was used [18] for predicting the rotordynamic response at stations 22 and 39. Figure 11 shows the algorithm flow chart used to iterate between the force response and bearing force coefficients. The analysis is run at steady state conditions and begins with seeding the iteration for each rotor speed by guessing at vibration amplitudes at stations 55 and 49. This establishes stiffness and damping coefficients for the squeeze film and wire mesh. Next, the force response is calculated and the new vibration at stations 55 and 49 are used to calculate new stiffness and damping coefficients. The single speed forced response loop is performed until the difference in vibration magni-

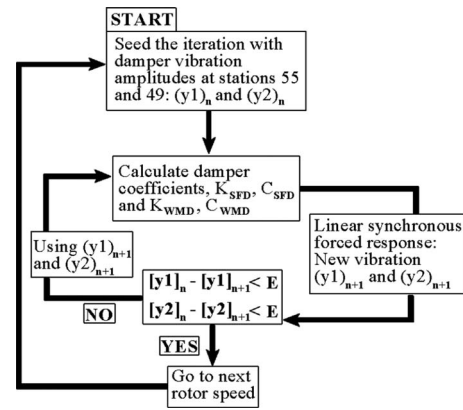


Fig. 11 Iteration scheme using nonlinear coefficients

tudes between successive iterations falls below a specified error value. Once this condition is satisfied the rotor speed is changed and the iteration between force coefficients and vibration amplitudes starts again. An important part in the solution algorithm was to account for the damper clearance by modeling the SFD and ball bearing as a bilinear spring set. Bilinear stiffness [14] can behave as a hardening spring when traversing a critical speed and materializes when the SFD land bottoms out on the support housing, which then results in a sudden change in rotor support stiffness (now controlled by the ball bearing spring rate). The other type of nonlinearity can come from the squeeze film cross-coupled damping term, which is defined in Eq. (3). The results in Sec. 5 will show that the hardening spring behavior was simulated and experimentally observed for the fwd end vibration of the rotor.

5 Experimental Results and Predictions

Each of the four imbalance test conditions described in Fig. 4 were simulated and compared with the experimental results and are shown in Fig. 12. Each plot in Fig. 12 shows experimental results that are representative of the maximum synchronous vibration component with run-out compensation measured during the coast down from the maximum test operating speeds. Also, the slow roll compensated synchronous orbit at the critical speed for each test is shown. In these orbit plots, the imbalance location is represented by a black dot. Each plot additionally has the imbalance test configuration shown for reference.

Test 1 is shown to have a first critical speed at 8000 rpm and representative Q levels (amplification factor at the critical speed) of 3.5 on the ball bearing end and 3 on the gas-bearing end. In this case the fwd SFD end response shows slightly lower vibration levels when compared with the vibration on the gas-bearing end. Predictions for this case matched well with the experimental data. For the fwd end response the simulation underpredicts the damping at the critical speed but shows to follow the same trend with rotor speed when compared with the experiments. For the aft vibration, the softening spring behavior is seen in vicinity to the critical speed.

The rotor imbalance in Test 2 was three times the imbalance used in Test 1. This test demonstrates the ability of the bearing to withstand rotor excursion roughly ten times a typical gas film thickness. This ability comes from the compliance of the bearing pads allowing the pad to conform to large rotor motion both translation and rotation. The external pressurization to the gas bearing also contributes to sustain the severe imbalance loading, especially at the low critical speed of 8 krpm where hydrodynamics are weak. Also, Q values are measured to be lower than the Q values in Test 1 that used three times less imbalance. Up to 13 krpm the agreement between experiments and simulations is generally good, as the simulations predict the location of the critical speed well and show moderate agreement on the vibration levels

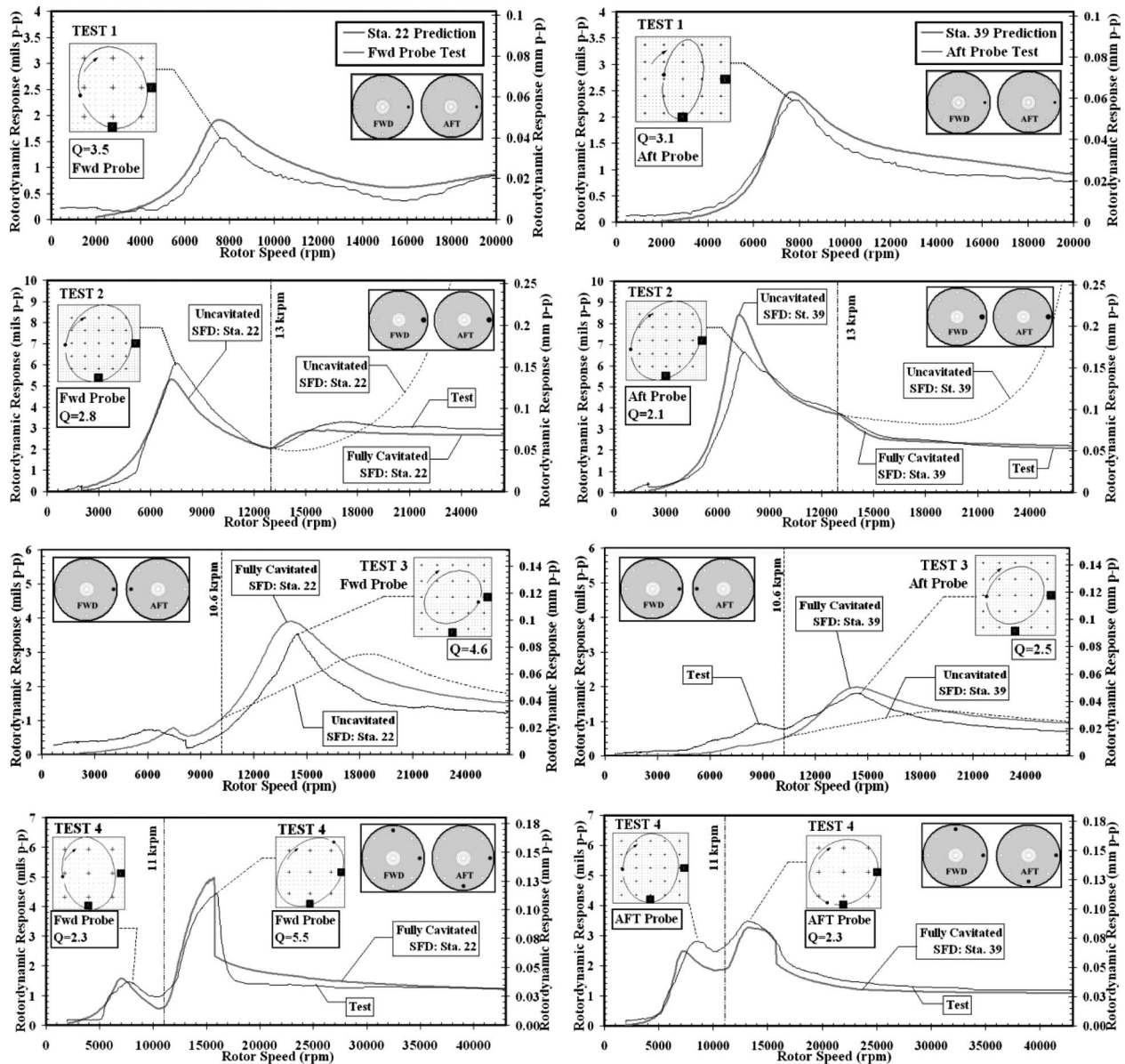


Fig. 12 Experimental results versus rotordynamic simulations

at resonance. The discrepancy on the aft vibration at the critical speed is hypothesized to be due to the change in WMD behavior for large vibrations, as testing experience is limited to lower vibration amplitude testing. Past 13 krpm there are three different curves. The dotted line represents the uncavitated or π cavitating SFD meaning that the squeeze film provides stiffness and damping as expressed in Eqs. (3) and (4). Note that this line significantly deviates from the experimental test results. It was suspected that air entrainment [19] might have occurred due to the unsealed damper configuration. Therefore, another simulation was performed with a fully cavitating SFD. To obtain good agreement with the experimental data several different SFD cavitation speeds were tried. It was found that 13 krpm resulted in the best agreement between experiments and predictions.

Test 3 excited a conical mode critical speed at 15 krpm. The orbit measurements at the critical speed show the 180 deg imbalance phase difference between the fwd and aft probes, verifying the excitation of a conical rotor mode. For this test the fwd end vibration near the SFD and ball bearing was measured to be higher in magnitude compared with the aft vibration. This is different from Test 1 to Test 2, as the fwd end vibration is lower than

the aft vibration. It was suspected that air entrainment could have compromised the damping capability of the SFD at the higher critical speed location. Test 2 rotordynamic simulations considering a fully cavitating SFD supports this theory. Therefore, the simulations performed used a fully cavitating SFD starting at 10.6 krpm and also varied the cavitation speed until good agreement was observed between tests and predictions. Even though the critical speed was most likely crossed with a nonfunctioning SFD, the Q values are acceptable, primarily due to the oil-free damping from the gas bearing.

Test 4 was configured to excite both the critical speeds at 8 krpm and 15 krpm. The first critical speed response, both on the fwd and aft ends, shows similar behavior to Test 1, except the measured rotor criticals are slightly higher in speed than Test 1. The ball bearing response (fwd probe) shows a good Q value whereas a Q value could not be calculated for the aft end due to the high damping in combination with the close vicinity to the second critical speed peak. The second critical speed for the aft end probe shows a very good Q value of 2.3 whereas the fwd probe vibration shows a less impressive Q factor of 5.5. The sec-

ond critical speed on the fwd end also shows a classic nonlinear jump at system resonance. To simulate this behavior the SFD was fully cavitated at 11 krpm in combination with modeling the damper clearance in the solution algorithm. It was specified in the iteration process that when the vibration at station 55 reaches the damper clearance the stiffness at that rotor location becomes the ball bearing stiffness resulting in a bilinear stiffness model. Using this approach, both fwd and aft probe measurements were predicted with good accuracy. It is important to note that the SFD used in these experiments is an unsealed damper. State of the art sealed SFD systems can produce higher damping and are less prone to air entrainment [20,21]. Additionally, the rudimentary SFD model used in this work can be improved by implementing existing models that have more sophisticated methods for predicting the effects of squeeze film air entrainment and cavitation [22,23].

6 Conclusions

The work presented in this paper focused on the synchronous rotordynamic response performance of a rotor system using a damped compliant hybrid gas bearing. The tests were performed to reduce risk associated with application of a compliant hybrid gas bearing into mission critical turbomachinery. It was shown through experimental observation that the damped gas bearing safely sustained significant vibration amplitudes through critical speed transitions and also generated well-damped critical speeds with low Q factors for the imbalance loading presented. Imbalance magnitudes imposed on the gas bearing far exceed API specifications by a factor of 50 and are considered to be representative of extreme and rarely seen cases for centrifugal compressors. For aircraft engines the imbalance magnitudes ranged from typical to severe cases. Based on the aft and fwd location vibrations, the damped gas bearing showed to have better performance at the higher second critical speed location compared with the ball bearing and SFD whereas the ball bearing and SFD showed better performance at the first critical speed. It was shown that the damping and stiffness of the gas bearing was primarily controlled by the bearing support parameters and can be modeled with reasonable accuracy only considering the WMD and integral springs. Another important conclusion was the verification of ultra-high-speed stability of the gas bearing while most likely operating with a fully cavitated SFD. Based on the experimental testing, the compliant hybrid gas bearing shows promise for application in several different types of machines that can benefit from oil-free operation ranging from aircraft gas-turbine engine applications to industrial size centrifugal compressors.

Acknowledgment

The authors would like to thank the General Electric Co. for sponsoring the present study and Matt Kaminske and Steve Leone from GE Aviation for their suggestions and guidance through experimental tests and rotordynamic modeling. Also, the authors would like to recognize Dr. Fouad Zeidan and Dr. Jongsoo Kim from KMC Bearings for providing the hybrid gas bearing, and Dr. Brian Murphy of Rotating Machinery Analysis for his support with the XLROTOR.

Nomenclature

C	= SFD radial clearance, mm
C_{BB}	= ball bearing damping, N s/m
C_{RT}	= SFD cross-coupled damping, N s/m
c_{SFD}	= squeeze film damping, N s/m
C_{TT}	= SFD direct damping, N s/m
c_{WMD}	= damping of two wire mesh dampers, N s/m
D	= rotor diameter, mm
E	= error criterion for a single speed forced response loop
k_{AFT}	= stiffness of aft end bearing support, N/m

k_{BB}	= ball bearing stiffness, N/m
k_{FWD}	= stiffness of fwd end bearing support, N/m
k_{IS}	= stiffness of gas-bearing integral springs, N/m
k_S	= bending stiffness of spline coupling, N/m
k_{SC}	= squirrel cage radial stiffness, N/m
k_{SFD}	= squeeze film damper dynamic stiffness, N/m
k_{WMD}	= stiffness of two wire mesh dampers, N/m
k_1	= stiffness leading edge integral spring, N/m
k_2	= stiffness of trailing edge integral spring, N/m
L	= SFD land length, mm
L_B	= gas-bearing pad axial length, mm
Q	= quality factor
R	= SFD radius, mm
y_1	= predicted peak vibration at station 55, mm
y_2	= predicted peak vibration at station 49, mm
ε	= SFD dynamic eccentricity ratio= y_1/C
θ_1	= offset angle for leading edge integral spring, deg
θ_2	= offset angle for trailing edge integral spring, deg
μ	= squeeze film viscosity, N s/m ²

Appendix

Table 1 shows the hybrid gas-bearing details.

Table 1 Hybrid gas-bearing details

Rotor diameter D	69.85 mm
Bearing length L_b	74.53 mm
Pad set-bore diameter	69.82 mm
Pad geometric preload coefficient	0.4
Supply pressure P_S	12.8 bars
Pad boundary pressure P_b	atm
Pad preload force P_{PR}	80 N
Recess depth	0.28 mm
Recess axial width	10.2 mm
Recess angular span	17 deg
Recess offset toward leading edge	60%
Capillary diameter	0.56 mm
Capillary length	14.8 mm
Bearing OD	136.6 mm
Damper OD	121 mm
Damper ID	80 mm
Damper width	22 mm
Damper mesh density	25%
Damper material	Cu
Damper wire gage diameter	0.23 mm
Integral spring stiffness K_1	2013 kN/m
Integral spring stiffness K_2	2013 kN/m
Leading edge integral spring angle θ_1	24 deg
Trailing edge integral spring angle θ_2	12.5 deg
Pad coating to 450°F	Teflon

References

- [1] Magge, N., 1975, "Philosophy, Design, and Evaluation of Soft-Mounted Engine Rotor Systems," *J. Aircr.*, **12**(4), pp. 318–324.
- [2] Vance, J., and Royal, A., 1975, "High-Speed Rotor Dynamics—An Assessment of Current Technology for Small Turbohaft Engines," *J. Aircr.*, **12**(4), pp. 295–305.
- [3] Memmott, E. A., 1990, "Tilting Pad Seal and Damper Bearing Applications to High Speed and High Density Centrifugal Compressors," *IFTOMM, Proceedings of the 3rd International Conference on Rotordynamics*, Lyon, pp. 58–590.
- [4] Zeidan, F. Y., and Paquette, D. J., 1994, "Application of High Speed and High Performance Fluid Film Bearings in Rotating Machinery," *Proceedings of the 23rd Turbomachinery Symposium*, Houston, TX, pp. 209–232.
- [5] Powell, J. W., 1970, "A Review of Progress in Gas Lubrication," *Rev. Phys. Technol.*, **1**, pp. 96–129.
- [6] Ertas, B., 2009, "Compliant Hybrid Journal Bearings Using Integral Wire Mesh Dampers," *ASME J. Eng. Gas Turbines Power*, **131**(2), p. 022503.
- [7] Zarzour, M., and Vance, J., 2000, "Experimental Evaluation of a Metal Mesh Bearing Damper," *ASME J. Eng. Gas Turbines Power*, **122**, pp. 326–329.

- [8] Al-Khateeb, E. M., and Vance, J. M., 2001, "Experimental Evaluation of a Metal Mesh Bearing Damper in Parallel With a Structural Support," ASME Paper No. 2001-GT-0247.
- [9] Ertas, B., and Luo, H., 2008, "Nonlinear Dynamic Characterization of Oil Free Wire Mesh Dampers," ASME J. Eng. Gas Turbines Power, **130**, p. 032503.
- [10] Okayasu, A., Ohta, T., Azuma, T., Fujita, T., and Aoki, H., 1990, "Vibration Problems in the LE-7 Liquid Hydrogen Turbopump," *Proceedings of the 26th AIAA/SAE/ASME/ASEE Joint Propulsion Conference*, pp. 1–5.
- [11] Ertas, B. H., Al-Khateeb, E. M., and Vance, J. M., 2003, "Rotordynamic Bearing Dampers for Cryogenic Rocket Engine Turbopumps," J. Propul. Power, **19**(4), pp. 674–682.
- [12] American Petroleum Institute, 1996, "Tutorial on the API Standard Paragraphs Covering Rotor Dynamics and Balancing: An Introduction to Lateral Critical and Train Torsional Analysis and Rotor Balancing," Standard Paragraphs API Publication 684, Feb.
- [13] International Organization for Standardization, 2003, "Mechanical Vibration—Balance Quality Requirements for Rotors in a Constant (Rigid) State—Part 1: Specification and Verification of Balance Tolerances," ISO 1940-1.
- [14] Zeidan, F. Y., San Andres, L., and Vance, J. M., 1996, "Design and Application of Squeeze Film Dampers in Rotating Machinery," *Proceedings of the 25th Turbomachinery Symposium*, Houston, TX, Sept. 17–19, pp. 169–188.
- [15] Ertas, B., Luo, H., and Hallman, D., 2009, "Dynamic Characteristics of Shape Memory Alloy Wire Mesh Dampers," *Proceedings of the 50th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference*, Palm Springs, CA, May 4–7, Paper No. AIAA-2009-2196.
- [16] Al-Khateeb, E. M., 2002, "Design, Modeling, and Experimental Investigation of Wire Mesh Vibration Dampers," Ph.D. thesis, Texas A&M University, College Station, TX.
- [17] Choudhry, V. V., and Vance, J. M., 2005, "Design Equations for Wire Mesh Bearing Dampers in Turbomachinery," ASME Paper No. GT2005-68641.
- [18] Hibner, D. H., 1975, "Dynamic Response of Viscous-Damped Multi-Shaft Jet Engines," J. Aircr., **12**(4), pp. 305–312.
- [19] Zeidan, F. Y., and Vance, J. M., 1989, "Cavitation Regimes in Squeeze Film Dampers and Their Effect on Pressure Distribution," STLE Annual Meeting, Atlanta, GA, STLE Paper No. 89-AM-4B-1.
- [20] Ide, R. D., and Zeidan, F. Y., 1995, "Integral Squeeze Film Damper Bearings," U.S. Patent No. 5,421,655.
- [21] San Andres, L., and Delgado, A., 2008, "Squeeze Film Damper With a Mechanical End Seal: Experimental Force Coefficients Derived From Circular Centered Orbits," ASME J. Eng. Gas Turbines Power, **130**, p. 042505.
- [22] Diaz, S., and San Andres, L., 2001, "A Model for Squeeze Film Dampers Operating With Air Entrainment and Validation With Experiments," ASME J. Tribol., **123**, pp. 125–133.
- [23] Diaz, S., and San Andres, L., 1999, "Reduction of Dynamic Load Capacity in a Squeeze Film Damper Operating With a Bubbly Lubricant," ASME J. Eng. Gas Turbines Power, **121**, pp. 703–709.

Luis San Andrés
Mast-Childs Professor
Fellow ASME

Ash Maruyama
Research Assistant

Department of Mechanical Engineering,
Texas A&M University,
College Station, TX 77843

Kostandin Gjika
Honeywell Turbo Technologies,
Zone Industrielle Inova 3000,
2 rue de l'Avenir,
88155 Thaon-les-Vosges Cedex, France

Sherry Xia
Rotordynamics Manager
Honeywell Turbo Technologies,
No. 430 Li Bing Road,
201203 Shanghai, China

Turbocharger Nonlinear Response With Engine-Induced Excitations: Predictions and Test Data

Turbochargers (TCs) aid to produce smaller and more fuel-efficient passenger vehicle engines with power outputs comparable to those of large displacement engines. This paper presents further progress on the nonlinear dynamic behavior modeling of rotor-radial bearing system by including engine-induced (TC casing) excitations. The application is concerned with a semifloating bearing design commonly used in high speed turbochargers. Predictions from the model are validated against test data collected in an engine-mounted TC unit operating at a top speed of 160 krpm (engine speed = 3600 rpm). The bearing model includes noncylindrical lubricant films as in a semifloating-ring bearing with an antirotation button. The nonlinear rotor transient response model presently includes input base motions for the measured TC casing accelerations for increasing engine load conditions. Engines induce TC casing accelerations rich in multiple harmonic frequencies; amplitudes being significant at two and four times the main engine speed. Fast Fourier transform frequency domain postprocessing of predicted nonlinear TC shaft motions reveals a subsynchronous whirl frequency content in good agreement with test data, in particular, for operation at the highest engine speeds. Predicted total shaft motion is also in good agreement with test data for all engine loads and over the operating TC shaft speed range. The comparisons validate the rotor-bearing model and will aid in reducing product development time and expenditures.

[DOI: 10.1115/1.3159368]

1 Introduction

Turbochargers enable smaller and more fuel-efficient passenger vehicle internal combustion (IC) engines with power outputs comparable to those of large displacement engines. In a TC, exhaust gases drive a turbine wheel, which is connected to a compressor wheel by a thin shaft, as shown in Fig. 1. The compressor forces a denser charge of air (i.e., larger air mass) into the engine cylinders, improves combustion efficiency, thus producing more power. While simple in concept, TC design and operation are complex. As a TC rotor spins, it vibrates due to several causes (imbalance, engine-induced vibrations, fluid film bearing self-excitations, aerodynamic loading, pressure fluctuations, etc.). The vibration amplitudes need to be controlled to ensure TC reliability.

The bearings support system is a key feature of a reliable turbocharger design since bearings greatly affect the TC static and dynamic forced performances. Turbochargers are usually supported by one of three bearing types: ball bearings, floating-ring bearings (FRBs), or semifloating-ring bearings (SFRBs). FRBs consist of a floating ring loosely fitted within the shaft and TC housing, thereby creating two oil films (inner and outer) in series. The ring is free to whirl while rotating at a fraction of the shaft speed due to viscous drag shear from the inner and outer films. SFRBs are similar in geometry to FRBs; however, the ring is prevented from rotation by a locking pin (or button). Thus, the outer film acts solely as a squeeze-film damper. Ball bearings are usually backed by a squeeze-film damper. In many TCs, a single

SFRB supports the rotor. The long ring incorporates both fluid film bearings (compressor and turbine) at its ends.

Rotor-ball bearing systems have a limited lifespan and yield the smallest TC subsynchronous vibrations but the largest synchronous ones [1]. FRBs and SFRBs are more economic and use the engine lubricant but need careful design to limit shaft motion amplitudes without affecting system reliability or reducing its thermal efficiency. Traditionally, TC development consists of costly gas test stand iteration. The present research focuses on the development of fast and accurate computational tools to accurately predict TC shaft motions, reducing (or eliminating) the need for repetitive test stand certification and qualification. FRBs offer decreased power losses, lower operating temperatures, and superior stability compared with conventional cylindrical journal bearings [2–6]. However, accurate prediction of FRB and TC (supported on FRBs) dynamic forced response has proven difficult since the rotor-bearing system is highly nonlinear.

Tatara [3] conducts experiments to determine the stability characteristics of a rotor supported on a number of FRBs combining various inner and outer clearances. Test results show that the rotor initially unstable operation becomes stable at high shaft speeds, lending credibility to the concept of a FRB superior stability over plain journal bearings. However, the ring speed reaches a nearly constant magnitude at shaft speeds greater than 10 krpm. This result is contrary to the familiar assertion that the ring speed increases with shaft speed. Trippett and Li [6] find the rationale for this apparently unusual behavior. The authors compared ring speed measured data to predictions derived from isothermal and thermal bearing flow analyses (lubricant viscosity and clearances varying with temperature). Test data show that the ring reaches a nearly constant speed as shaft speed increases. Isothermal model predictions, however, show a linear increase in ring speed. Only

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 21, 2009; final manuscript received March 25, 2009; published online December 1, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Gas Turbine Technical Congress and Exposition, Orlando, FL, June 8–12, 2009.

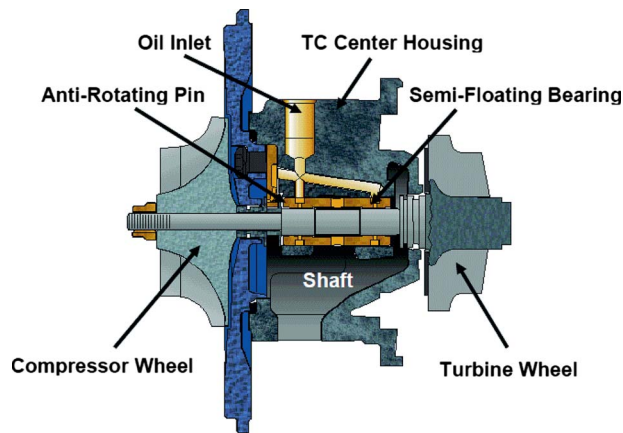


Fig. 1 Turbocharger rotor assembly with SFRB support

the thermal model results deliver good correlation with the test data. Hence, since 1984, an isothermal flow model is not suitable for accurate FRB predictive performance.

Experimental and numerical analyses of TCs supported on FRBs also show TC stability depends greatly on the FRB geometry. Tanaka and Hori [7] present a stability analysis of a short length rigid rotor supported on FRBs. Stability charts for various bearing outer (C_o) to inner clearance (C_i) ratios ($0.8 \leq C_o/C_i \leq 4.0$) and floating-ring outer (R_o) to inner (R_i) radius ratios ($1.20 \leq R_o/R_i \leq 1.44$) showed that the TC stable operating speed region increases as the clearance ratio and radii ratio increase. The stability analysis also showed that the rotor system can traverse several regions of stable and unstable behaviors, in agreement with experimental results given in Ref. [3]. Also in Ref. [7] there are experimental results for a rotor operating with FRBs supplied with oil at various feed pressures. Two subsynchronous frequencies of rotor motion were observed depending on the feed pressure. At higher supply pressures, a frequency approximately equal to half of the ring speed was observed. At lower supply pressures, a frequency of approximately half of the ring plus journal speeds appeared. In Ref. [5], Li confirms these findings theoretically and attributed the “half frequency whirl” to instabilities in the outer and inner films, respectively.

Li and Rohde [4] perform linear and nonlinear analyses of the steady state and dynamic forced performances of rotors supported on FRBs. Interestingly, the (numerical integration) nonlinear analysis shows that the rotor-FRB system reaches stable limit cycles regardless of the initial conditions used. Moreover, stable limit cycles with subsynchronous frequencies occur even outside the predicted linear stability range. The authors caution, however, that the bearing outer film to inner film clearance ratio must be carefully chosen to allow for such operation. Li [5] identifies three major types of rotor frequency response: half ring speed frequency (instability in the outer film), half ring-plus-journal speeds frequency (instability in the inner film), and synchronous frequency (response to imbalance). The ring speed increases with decreasing inner film clearance and decreases with decreasing outer film clearance. Similarly, power consumption decreases with increasing inner film clearance and is relatively unaffected by changes in the outer film clearance. Recently, Holmes et al. [8] reproduced well-known linear eigenvalue analysis of a TC supported on FRBs attempting to explain subsynchronous frequencies observed experimentally.

The increase in computational power allows contemporary numerical analyses to focus on predicting TC nonlinear motion responses. Holt et al. [9] present measurements of housing acceleration on an automotive TC supported on FRBs. Lubricant feed temperature and pressure are controlled to determine their effect on system rotordynamic response. Test data waterfalls show two

subsynchronous frequencies, which track the rotor speed, indicating a persistent condition of oil whirl below 90 krpm. Lubricant feed temperature shows minimal effect on the onset and end speeds of instabilities, whereas increasing feed pressure tends to delay the onset of instability. Holt et al. [9] also detail linear and nonlinear rotordynamic models to predict TC forced response with comparisons to the test data. The linear model uses linearized bearing force coefficients and predicts well the (first) onset speed of instability. The nonlinear model predicts limit cycle orbits with two subsynchronous frequencies of approximately 50% of ring speed and 50% of ring-plus-journal speeds. The nonlinear model also demonstrates the importance of rotor imbalance in ameliorating subsynchronous motion amplitudes at high speeds.

San Andrés and Kerth [10] detail a thermohydrodynamic flow model to predict FRB forced response. The lumped-parameter thermal energy balance model estimates the inner and outer film lubricant viscosities and changes in clearances due to thermal growth of the journal, ring, and bearing. Ring speed predictions show fair estimation of actual test data, displaying the same decreasing trend, albeit at a lower rate than test data indicates. Predictions of power losses show good correlation with test data at low speeds (<35 krpm) but slightly underpredict values at high speeds (>55 krpm, 65 krpm maximum). The novel FRB model is also fully integrated into a finite element-based rotordynamics model, which predicts well the subsynchronous whirl frequency content of the rotor response but tends to overpredict rotor motion amplitudes.

San Andrés et al. [11] continue to refine the FRB flow model and presented comparisons of predictions to test data of a TC supported on a SFRB. The bearing flow model includes hydrostatic (side) loads due to external pressurization. Predictions reveal large subsynchronous vibrations throughout the TC speed range to a maximum of 244 krpm. Subsynchronous amplitudes and frequencies correlate well with test data. Predicted total motions at the compressor end, calculated as the maximum whirl orbit at a given shaft speed, are generally overpredicted. However, predictions are well below the physical limit, which still serves to lend credibility to the analysis. Nonlinear predictions of the synchronous response to imbalance agree well with test data over the entire speed range. Discrepancies between the two are attributed to uncertainty in the imbalance distribution of the actual TC. Predicted whirl frequency ratios (WFRs) of the subsynchronous vibrations are also presented and correlate well with test data, though a third higher WFR is predicted but not evidenced by the test data. Test data (and predictions) indicate that the WFRs tend to coincide with the rotor conical mode of vibration and its harmonics.

San Andrés et al. [12,13] detail comparisons of test data to predictions based on the models in Refs. [9,10] for a FRB-supported TC. The FRB model now includes centrifugal pressure losses into the inner film arising from the lubricant circumferential velocity tendency to force the lubricant out of the inner film. After validating the rotor model against free-free mode tests, the authors present results of a linear eigenvalue analysis. Three critical speeds are predicted at 4 krpm, 20 krpm, and 67 krpm corresponding with the conical, cylindrical, and first elastic bending modes of the rotor, respectively. Nonlinear predictions show that subsynchronous amplitudes and frequencies are well predicted; however, the predictions show an absence of subsynchronous activity above 55 krpm. On the other hand, test data displays subsynchronous vibrations throughout the operating speed range. Predicted total shaft motions show excellent correlation with test data, especially at shaft speeds above 25 krpm.

For a TC supported on SFRBs, Gjika and co-workers [14,15] compare shaft test data to predictions from the model of San Andrés and Kerth [10]. Both the nonlinear analysis and test results show large subsynchronous vibrations throughout the TC operating speed range. The predicted subsynchronous frequency content also agrees well with test data. While predicted subsynchronous

amplitudes are smaller than test results indicate, both are larger than synchronous amplitudes. The tests and analysis reveal the paramount importance of oil feed pressure on the onset, persistence, and severity of subsynchronous whirl motions in the studied turbocharger.

Prior art (Refs. [9–15]) presents direct comparisons of predictions to measured shaft motions and TC casing accelerations. No other concerted effort in the open literature has documented the nonlinear rotordynamics of TCs with such detail. Predicted TC shaft motion is most often compared with test data obtained in test stands. In this setup type, the TC housing is not subjected to any external loading. However, automotive TCs are subjected to a wide range of engine vibrations that can drastically affect TC dynamic response. Sahay and LaRue [1] present a comparison of recorded total shaft motion for identical automotive TCs mounted on a test stand and on an IC engine. Engine vibrations (and pulse loading in the turbine) cause TC shaft motion amplitudes to be typically 20–30% larger than those recorded in a test stand. Kirk and co-workers [16,17] perform shaft motion measurements of an automotive FRB-supported TC mounted on a 3.9 l four-cylinder diesel engine. Waterfalls of shaft motion clearly show two subsynchronous frequencies and several lower frequency vibrations attributed to engine vibration. Interestingly, with no load placed on the engine, the TC shows, at the engine frequency, vibration amplitudes much smaller than those caused by TC subsynchronous motions. These results are equivocal. However, at full engine load, engine-induced vibrations become comparable in magnitude to TC shaft vibrations, demonstrating the dramatic effect engine motion and engine loading condition can have on a TC response. Unfortunately, no comparison of the total TC motions from each test is provided.

Other studies also show that the frequency of external vibration affects the rotor-bearing system response. Hori and Kato [18] present a numerical analysis of a Jeffcott rotor supported on plain cylindrical film journal bearings and subjected to external shocks. Eigenvalue analysis indicates that the otherwise stable rotor can become unstable at shaft speeds above twice the critical speed due to sufficiently large external shocks. In addition, unstable operation results when the external shock frequency is close to that of the natural frequency of the RBS. Lee et al. [19] also demonstrate both analytically and experimentally that vibration amplitudes were amplified when the external shock frequency was close to a system natural frequency.

This paper presents progress on the refinement of the linear and nonlinear rotordynamic models developed in Refs. [9–13] to predict the response of a TC supported on a SFRB due to specified base excitations that simulate engine-induced vibrations.

2 Turbocharger Structural Model and SFRB Thermal Model

Figure 1 displays the TC rotor assembly with the SFRB as a one-piece design integrating the compressor and turbine bearings at its ends. A button pin prevents ring rotation. Figure 2 shows the TC rotordynamic model consisting of 43 finite elements for the rotor, including the thrust collar and spacer, and 13 finite elements for the SFRB. Childs [20] presents the theoretical aspects employed for rotor lateral bending and associated finite element modeling. Validation of the rotordynamics model requires good correlation between measured and predicted rotor physical properties, as well as free-free natural frequencies and mode shapes. Figure 3 overlays on the rotor geometry the measured and predicted mode shapes corresponding to the first and second natural frequencies, respectively. The good agreement validates the rotor structural model. Reference [21] details the laboratory procedure for measurement of TC rotor free-free modes.

The TC housing incorporates a half-moon groove at the turbine bearing end and an antirotation button pin at the compressor bearing end. The groove is a manufacturing feature and its geometry is taken into account in the bearing model. The thin films of the

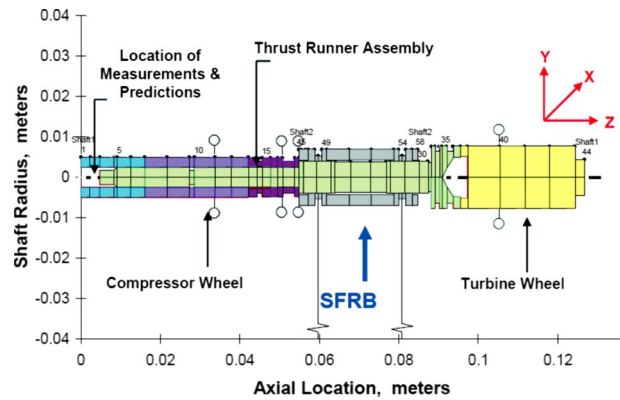
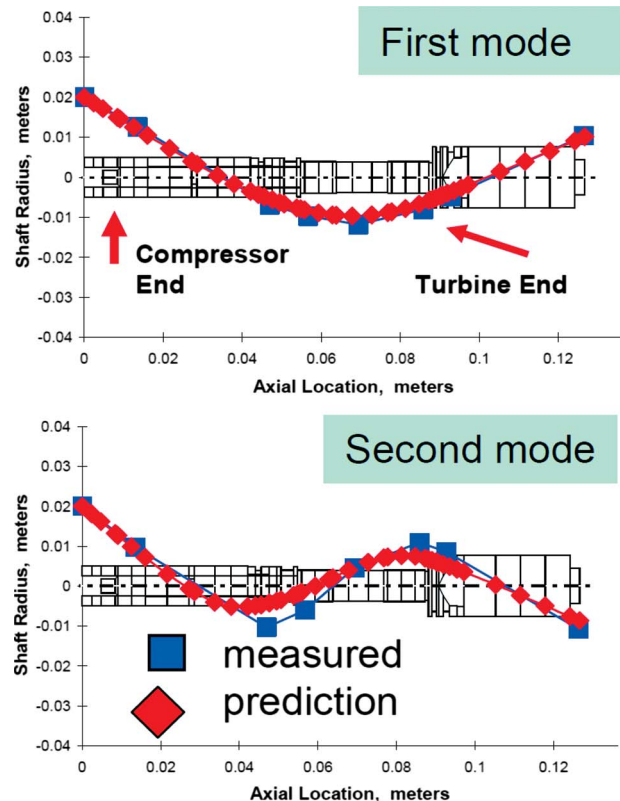


Fig. 2 Turbocharger rotor and semifloating-ring bearing structural models

SFRB are modeled using the thermal flow model detailed in Refs. [10,12]. The model calculates bearing static performance parameters, incorporating a thermohydrodynamics fluid flow model predicting bearing forced response. Along with bearing geometry, the computational analysis requires as input TC shaft speeds, oil inlet conditions, side loads, bearing material properties, and shaft, housing, and ring temperatures.

Test data for the TC was collected by the manufacturer on an IC engine test stand at three engine loads: 25%, 50%, and 100% of



	Measured [kHz]	Predicted [kHz]	% difference
First Freq.	1.799	1.823	1.3
Second Freq.	4.938	4.559	7.7

Fig. 3 Measured and predicted first and second free-free natural frequency mode shapes of TC rotor at room temperature

Table 1 TC operating conditions for 100% IC engine load condition

Engine speed (rpm)	TC speed (rpm)	Oil supply pressure (bar)	Oil supply temperature (°C)
1000	48,504	1.3	98
1250	69,366	1.7	97
1750	120,216	2.2	106
2000	134,298	2.4	109
2250	141,762	2.5	111
2500	144,036	2.6	115
2750	146,106	2.8	116
3000	148,722	2.7	120
3250	152,532	2.7	123
3600	157,674	2.7	126

full engine load. Table 1 lists the engine speed, TC shaft speed, oil supply pressure, and inlet temperature to the FRB for the 100% engine load condition. The lubricant is a typical light multigrade engine oil.

Bearing material properties are important in determining clearance changes due to material thermal growth. During testing, shaft, housing, and ring temperatures at the bearing locations were not measured. For operation with a turbine inlet temperature of 700°C, an empirically determined defect temperature ratio provides a reasonable estimation of the shaft temperature at ~205°C. TC housing temperature estimates at 150°C follow from direct measurements in a passenger vehicle [21]. Pertinent bearing parameters including effective lubricant viscosity, actual clearances, inlet film pressures, side loads, and ring speed ratios (nonzero for FRBs) are input into a linear rotordynamics model to predict the rotor-bearing system damped natural frequencies and mode shapes (eigenvalue analysis) as well as the linear response to imbalance. The Appendix presents some of the predicted parameters for the inner and outer films of the SFRB.

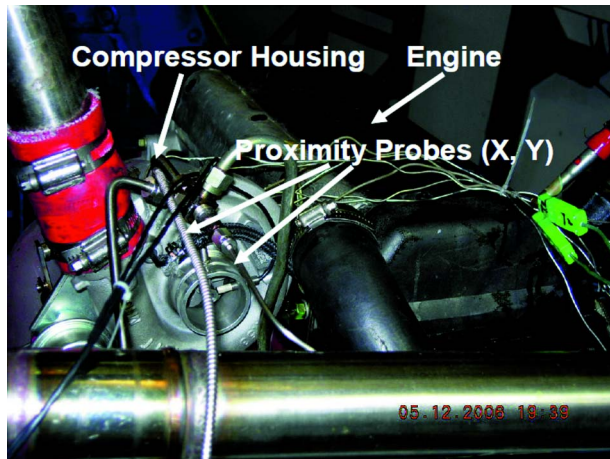


Fig. 4 Turbocharger engine test facility stand

3 TC Housing Acceleration Test Data

Measurements of TC shaft motions are routinely conducted on an engine test stand (Fig. 4). Two eddy current displacement transducers, affixed to the compressor housing and facing the TC shaft, are positioned in the horizontal and vertical planes. TC center housing and compressor housing accelerations are collected with three-axis accelerometers.

Figure 5 shows waterfalls of center housing and compressor housing accelerations along the vertical direction for 100% engine load. In the graphs, the horizontal axis represents orders (e) (or multiples) of main engine speed. Initially, the TC housing and its components were assumed to move as a rigid body due to their apparent rigid connection. However, the data acquired demonstrates that this assumption is not adequate. Typically, center housing vibration amplitudes are much smaller than compressor housing amplitudes due to the high stiffness of the center housing. The ovals in the figures denote the fundamental natural frequencies of the combined TC and manifold system.

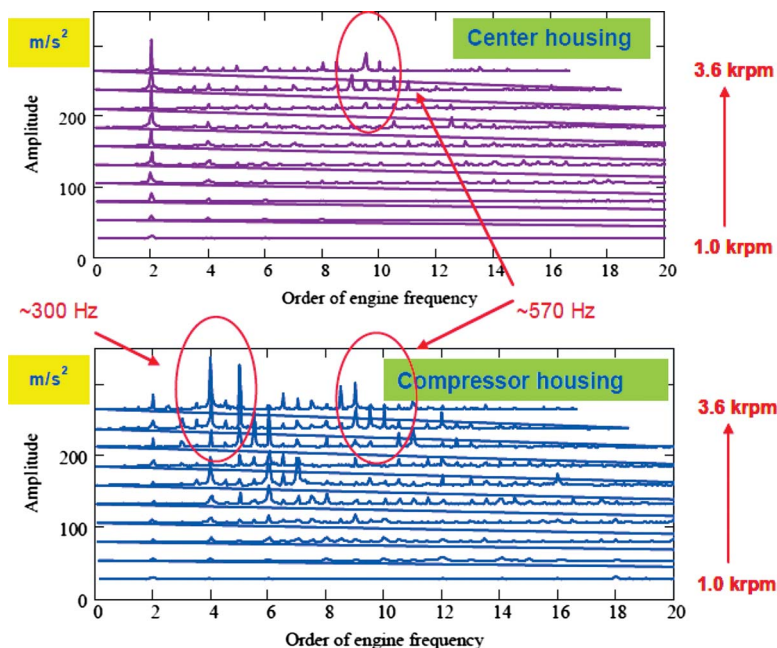


Fig. 5 Waterfalls of center housing (top) and compressor housing (bottom) acceleration test data versus orders of engine frequency (100% engine load, vertical direction)

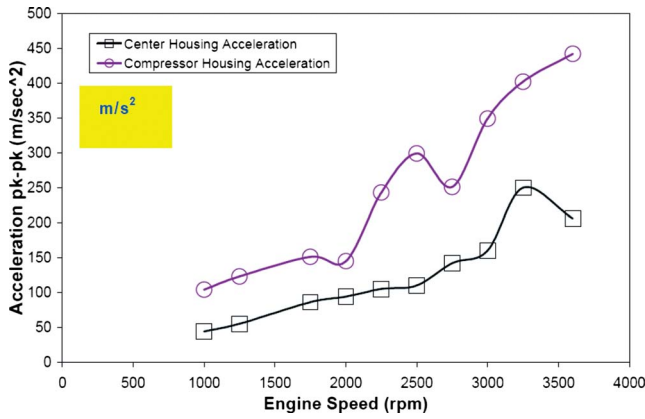


Fig. 6 Peak acceleration of TC center housing and compressor housing versus engine speed (100% load, measurements along vertical direction)

Figure 6, comparing housings' total accelerations in the vertical direction, illustrates that compressor housing acceleration is approximately twice that of the center housing. Figure 5 illustrates that $2e$, $4e$, and $6e$ order frequencies contribute significantly to the TC compressor housing motion. This is most likely due to the engine firing frequency equaling two times the engine rotational speed. Interestingly, the engine fundamental frequency does not appear in the waterfall spectra.

4 Integration of Test Data Into Rotordynamic Analysis

Integration of the test data into the rotordynamic analysis program requires accurate modeling of the housing acceleration. As simplification, the housing is assumed to move as a rigid body due to its heavy weight and thick-walled construction. Housing vibrations are transmitted to the rotor only through the bearing connections. Also, the respective bearing connections are assumed to transmit identical housing vibrations, i.e., the center housing moves only vertically or horizontally (no pitching or tilting). Figure 7 portrays graphically the assumptions above, i.e., motions of the TC rotor model base or foundation to represent engine-induced excitations.

Given the recorded TC housing acceleration as a function of time $a(t)$ with a fundamental engine frequency (ω), a Fourier coefficient decomposition of the time data yields the base displacement

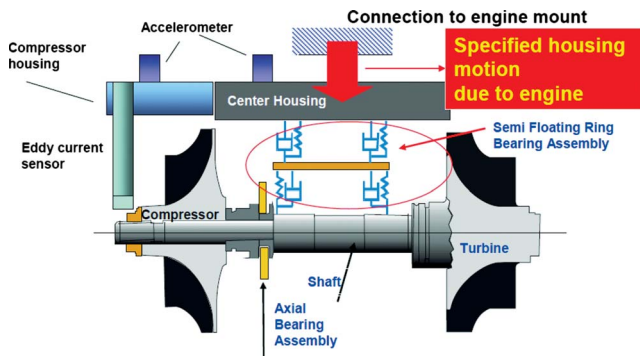


Fig. 7 Depiction of TC center housing motions for rotordynamics analysis (taken from Ref. [15])

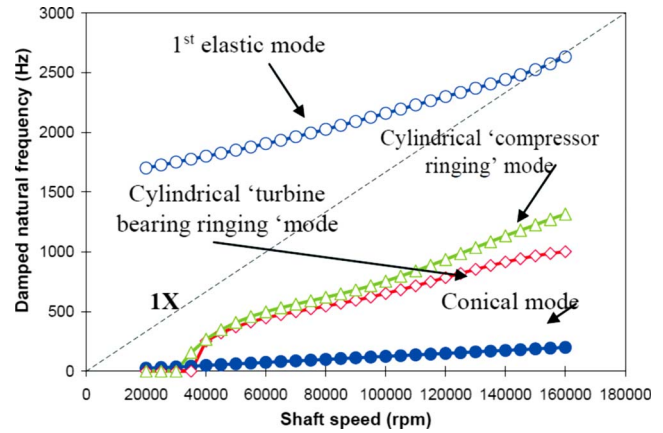


Fig. 8 TC-SFRB system damped natural frequency map for 100% engine load

$$x(t) = \sum_{n=1}^{N_F} -\frac{A_n}{(n\omega_e)^2} \cos(n\omega_e t + \phi_n) \quad (1)$$

where A_n and ϕ_n are the magnitude of acceleration and phase of the n th order of the fundamental engine frequency (maximum frequency order= N_F). Note that the shaft motion test data is actually motion relative to the compressor housing, as schematically shown in Fig. 7. Thus, shaft motion relative to the TC housing must be calculated to properly quantify the effect of housing motion on TC shaft motion and to accurately reproduce test conditions. Assuming the phase difference between the shaft and compressor housing displacements is negligible,¹ calculation of TC shaft motion relative to its housing follows as absolute shaft displacement minus housing displacement. The first ten Fourier coefficients (including half-orders) are used to model the center housing accelerations ($N_F=10$), vertical and horizontal.

5 Predictions and Comparisons to Test Data

Predictions are obtained for 25%, 50%, and 100% of full engine load [21]. The TC rotor imbalance is determined at the manufacturer facility using a proprietary method. Note that the linear and nonlinear analyses for all engine loads show either nearly identical results or similar trends, respectively. Thus, for brevity, only results for 100% engine load are discussed. Nonlinear predictions (and test data) in Figs. 10–15 shown below correspond to vertical motion (y -direction) at the compressor nose. In addition, all shaft motion amplitudes are made dimensionless relative to the static play² at the compressor nose. Nonlinear predictions presented are motion relative to the compressor housing with engine-induced excitations unless otherwise specified. Red solid diamonds represent the test data (measurement) in all figures below.

The nonlinear rotordynamic computational analysis integrates the rotor equations of motion with each bearing modeled as a nonlinear impedance with reaction forces depending on the instantaneous shaft and ring displacements and velocities, geometry, and operating conditions determined from the fluid film bearing analysis program. The parameters of interest are effective lubricant viscosities and clearances for the inner and outer films, as well as actual inlet feed pressures³ [15]. In the analysis, an unbalance

¹In the tests, the measurements of shaft displacements and TC housing acceleration were not recorded simultaneously.

²The static play refers to the maximum diametral displacement of the shaft compressor wheel nose or nut of the rotating group determined from the output of the shaft motion measurement instrumentation by manually moving the rotating group through a conical motion.

³These pressures change from their supply value measured upstream due to losses at the inlet port as well as centrifugal pressure losses due to the ring rotation and feed hole distribution and configuration.

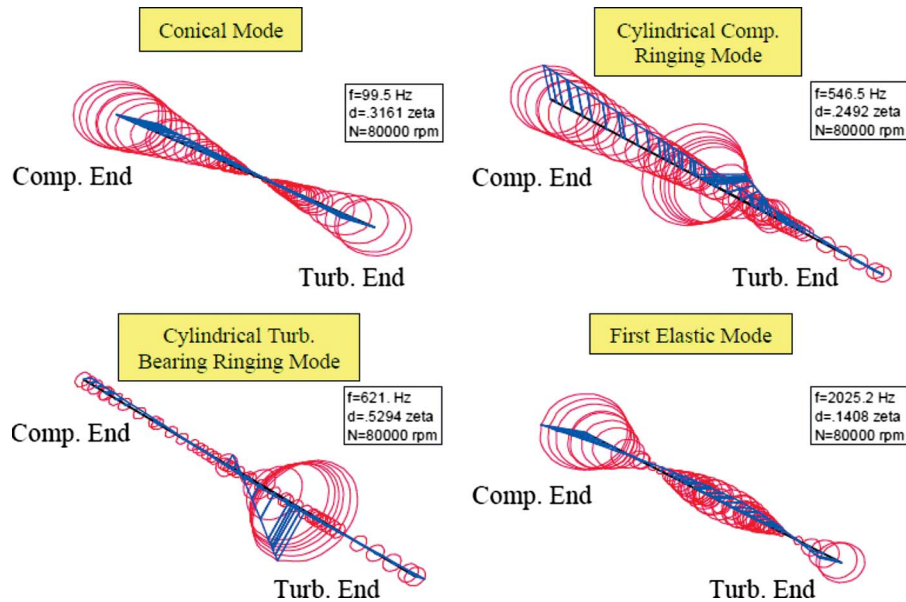


Fig. 9 Predicted rotor-SFRB natural mode shapes at 80 krpm (100% engine load)

distribution is specified, as well as the *side* loads produced by the supply pressure.

5.1 Eigenvalue Analysis. The linear eigenvalue analysis gives the system natural frequencies and corresponding rotor mode shapes. The analysis uses linearized bearing stiffness and damping force coefficients produced by the SFRB model. Figure 8 shows the predicted damped natural frequency map with four forward precessional rotor natural frequency mode shapes: a rigid conical mode, two cylindrical mode shapes, and the first elastic mode. However, only the first elastic mode shape corresponds with a rotor critical speed at ~ 150 krpm. For operation at 80 krpm, Fig. 9 shows the predicted rotor mode shapes, along with their natural frequency and damping ratio (ξ). The lowest frequency is a well-damped rigid rotor with conical mode shape ($f = 99$ Hz, $\xi = 0.32$). The next mode is a cylindrical-bending shape characterized by a large compressor wheel and a compressor bearing end motions ($f = 546$ Hz, $\xi = 0.25$). Closely following is another rigid cylindrical mode characterized by large motions at the turbine bearing end. This mode shape is well-damped ($f = 621$ Hz, $\xi = 0.53$). The highest frequency mode shape corresponds to a lightly-damped first elastic mode ($f = 2025$ Hz, $\xi = 0.14$).

5.2 Nonlinear Analysis. Linear rotordynamics analyses are ill-suited to predict the response of rotors supported on floating-ring bearing systems. Presently, the rotor and ring equations of motions with the thin film bearings acting as nonlinear force elements are integrated in time. The analysis is conducted for the engine and TC shaft speeds noted in Table 1. For numerical integration, 12,800 steps with a time step of 78.12×10^{-6} s are obtained. These parameters render FFTs with a maximum frequency span of 6.4 kHz and a frequency step of 4 Hz.

Figure 10 displays waterfalls of shaft motion (relative to compressor housing) for predictions that include or disregard engine-induced housing accelerations and test data at the compressor shaft end. Predictions with no housing accelerations show very little subsynchronous response, in direct contrast to predictions with housing accelerations and test data, which show a broad, low frequency response. Clearly, housing (i.e., engine) vibrations are the source of the low frequency response. It should be noted that test data has a frequency step size of 10 Hz, resulting in a very rough spectrum.

Accurate prediction of TC total shaft motion is of the utmost importance in industry. The total motion refers to the largest (peak to peak) amplitude of the complex orbital shape described by the turbocharger shaft at the location of measurement. Figure 11 shows very good correlation of the predicted total shaft motion to the measured shaft motions at the compressor end. The comparison serves to benchmark the accuracy of the predictive nonlinear RBS model.

Figure 12 compares predicted and measured subsynchronous amplitudes at each engine speed. Peak amplitude values are also denoted in the graph to indicate whether the synchronous or subsynchronous vibration is the dominant amplitude at that particular engine speed. At the majority of engine speeds, engine-induced TC shaft vibrations contribute significantly to the total shaft motion. Predictions agree remarkably well with test data, in particular, for engine speeds ranging from 1750 rpm to 3000 rpm.

Figure 13 depicts predicted and measured subsynchronous amplitudes in terms of orders of engine speed. Recall that only $10e$ order frequencies (including half-orders) are used to model housing vibrations and that $2e$ and $4e$ order frequencies are typically the dominant frequencies. As expected, $2e$ and $4e$ order frequencies induce marked shaft motion amplitudes, although predictions evidence vibrations up to only $7e$ order frequencies. Test data, however, only shows minimal vibration amplitudes below $6e$ order frequency.⁴ Note that predicted amplitudes at $\sim 14e$ order frequency are *not* due to engine-induced housing excitations. Rather, the amplitudes are attributed to TC shaft self-excited subsynchronous vibrations, independent of its housing motion.

Figure 14 shows comparisons of predicted and measured non-synchronous frequencies versus TC shaft speed. The two solid black lines trace the $2e$ and $4e$ engine order frequencies, i.e., two and four time multiples of the fundamental engine frequency. Recall that $2e$ and $4e$ order frequencies contribute significantly to engine motion. Predictions only evidence low frequency vibrations, albeit not predicting frequencies higher than 300 Hz. Rotor damped natural frequencies are also plotted to determine if these

⁴It is not possible to replicate the test data as housing accelerations (center housing and compressor housing) were not collected simultaneously with shaft vibration measurements. Also, the frequency step size used (10 Hz) is too large and fails to capture the shaft vibration amplitudes at frequencies coinciding with exact orders of engine speed.

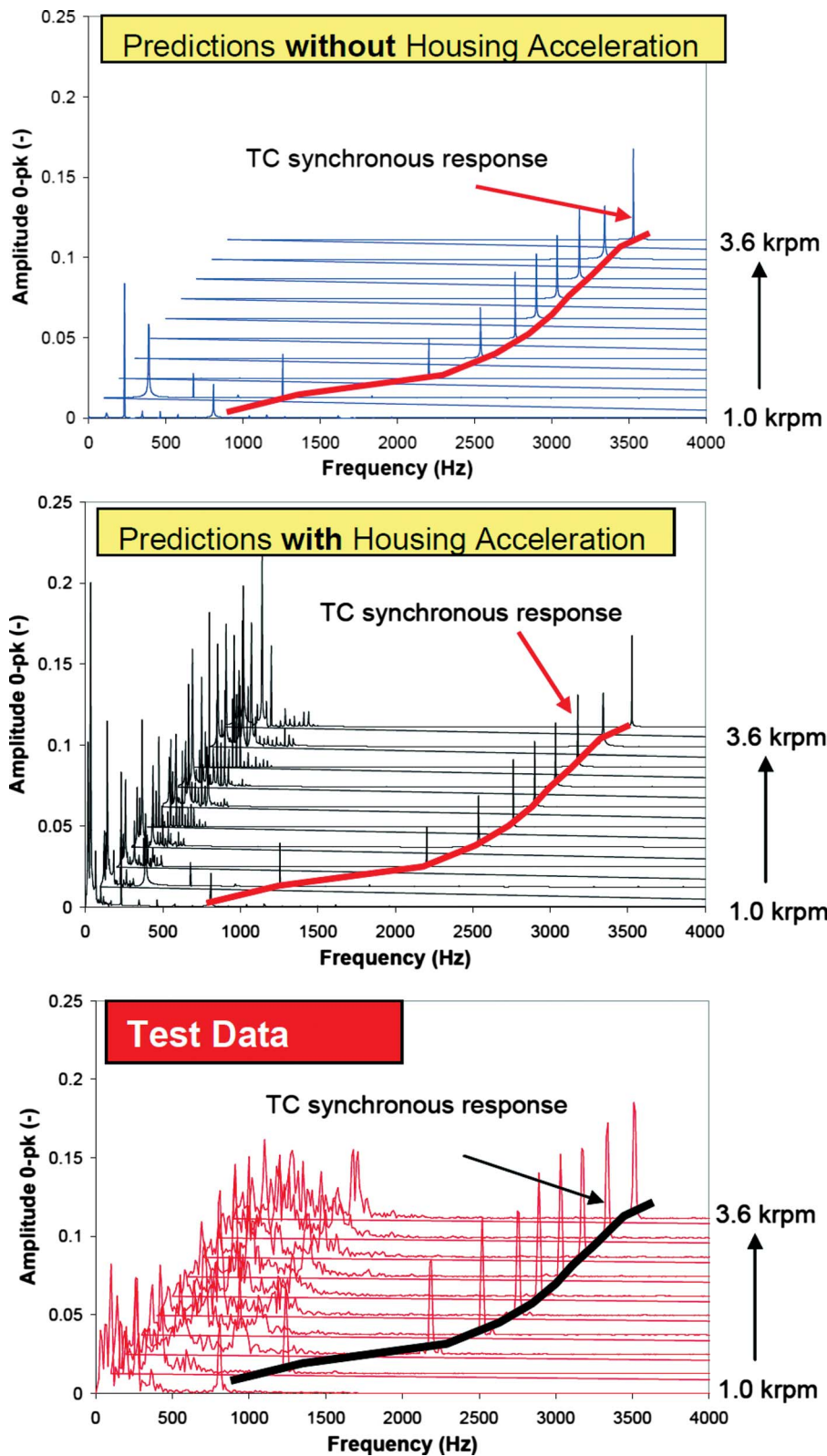


Fig. 10 Waterfalls of shaft motion (compressor nose, vertical direction). Predictions with no housing accelerations (top), predictions with housing accelerations (middle), and test data (bottom) (100% engine load)

frequencies are excited by any nonsynchronous vibrations. Although the $4e$ order frequency tracks the rotor conical mode, no nonsynchronous vibrations are evidenced since the mode is well-damped. Still, the results demonstrate that engine superharmonics

can excite rotor natural frequencies.

However, there are several groupings of nonsynchronous frequencies that track harmonics of the rotor conical mode: Group 1 at one-half rotor conical mode (~ 30 – 120 Hz), Group 2 at two

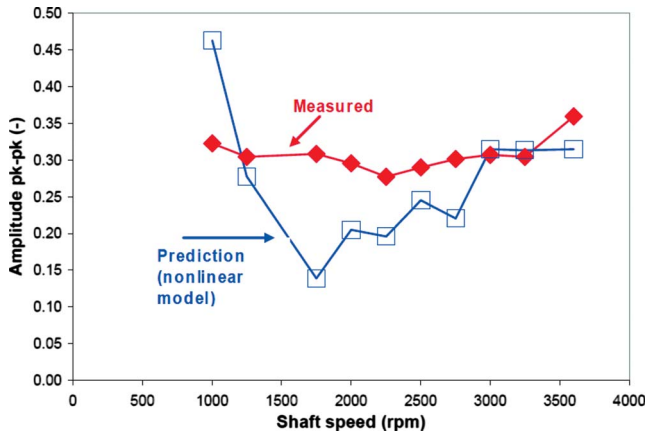


Fig. 11 Total TC shaft motion amplitude, predicted and measured, versus shaft speed (100% engine load)

times the rotor conical mode ($\sim 200\text{--}500$ Hz), and Group 3 at four times the rotor conical mode ($\sim 650\text{--}800$ Hz). Group 1 is attributed to $2e$ engine order frequency vibrations. Group 2 is actually a mix of higher engine order frequency vibrations. Group 3 is clearly a superharmonic of the second grouping.

Figure 15 also shows several nonsynchronous frequencies at 1000 rpm (engine speed). Measurements by the TC manufacturer

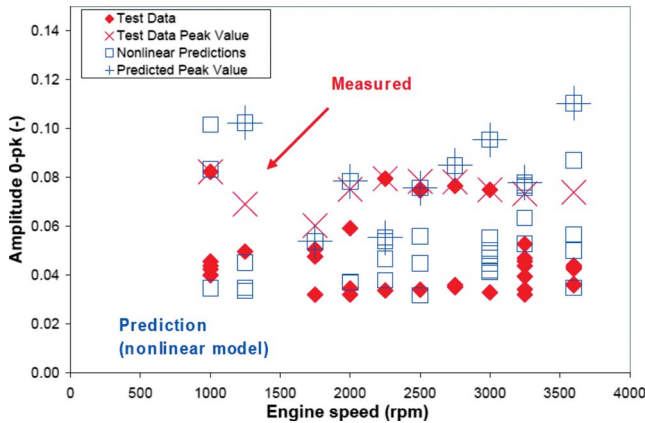


Fig. 12 Predicted and measured subsynchronous shaft motion amplitudes versus engine speed (compressor nose, vertical direction)

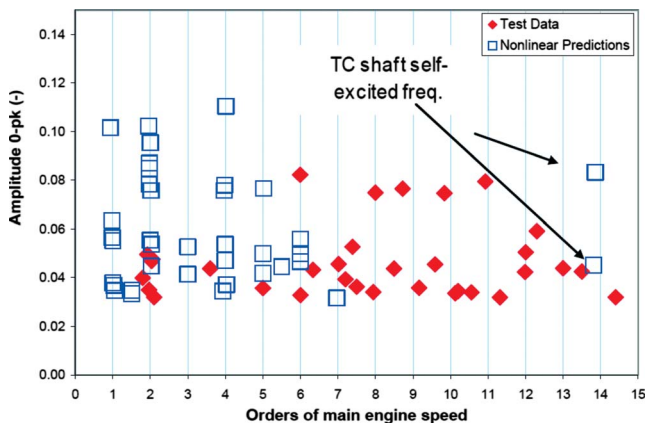


Fig. 13 Predicted and measured subsynchronous shaft motion amplitudes versus orders of engine speed (compressor nose, vertical direction)

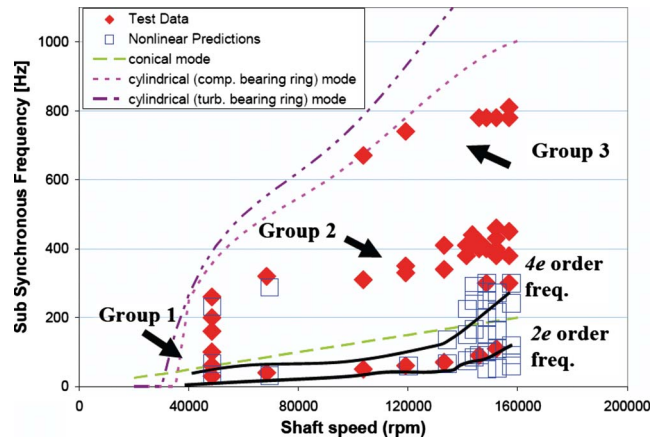


Fig. 14 Predicted and measured shaft subsynchronous whirl frequencies versus shaft speed (compressor nose, vertical direction, 100% IC engine load)

indicate an IC engine operating at a natural frequency of 18 Hz, very close to 1000 rpm (16.6 Hz). Recent finite element analysis of the combined TC and exhaust manifold conducted by the TC manufacturer show system natural frequencies at ~ 300 Hz and ~ 570 Hz. These frequencies are apparent on the spectra of accelerations recorded at the TC center housing and compressor housing (see Fig. 5). Further modal testing and mode shape measurements will continue to advance the current knowledge.

6 Conclusions

The paper reports progress on the refinement of linear and nonlinear rotordynamics models to accurately predict TC shaft nonlinear forced response with engine-induced housing excitations. The results, measurements, and predictions demonstrate that engine vibrations cause a complex low frequency TC shaft response with significant frequencies at several multiple integers of engine speed.

Analysis of measured engine-induced housing accelerations shows that the center housing and compressor housing do not vibrate as a rigid body. The compressor housing deflects nearly twice as much as the center housing. However, both housings show marked amplitude responses at $2e$ and $4e$ engine order frequencies, most likely due to the engine firing frequency. Other order frequencies, including $5e$ and $6e$, also contribute significantly to housing motion.

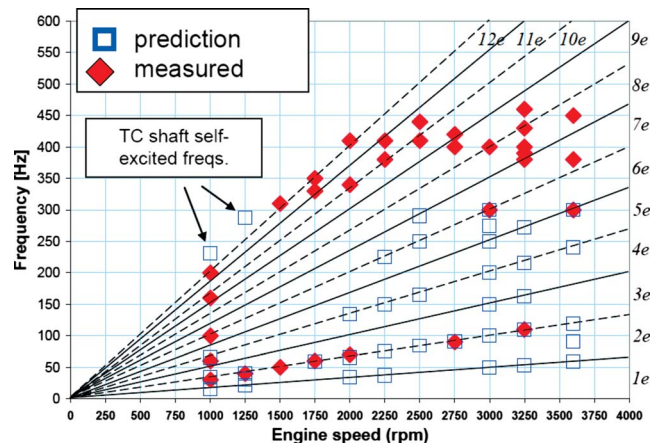


Fig. 15 Predicted and measured subsynchronous whirl frequencies versus engine speed (compressor nose, vertical direction, 100% IC engine load)

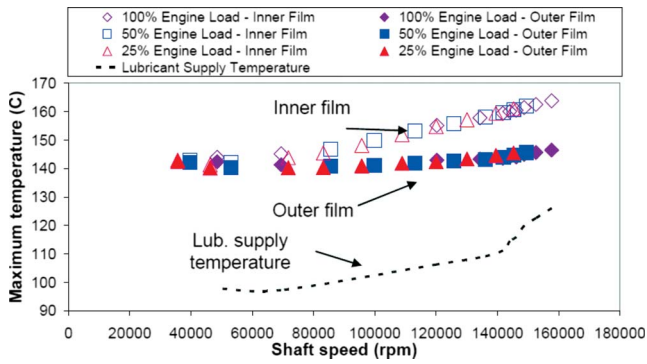


Fig. 16 Predicted exit inner and outer film temperatures at compressor bearing for 25%, 50%, and 100% engine loads

Linear eigenvalue predictions evidence four damped natural frequency TC shaft mode shapes, of which only the first elastic mode is lightly-damped (critical speed=150 krpm). The TC top speed in the tests is 160 krpm.

Nonlinear predictions of shaft motions without accounting for housing accelerations show that the TC is stable (i.e., no subsynchronous whirl) at all but the lowest operating speeds (<70 krpm). However, predictions of shaft motions including housing accelerations, as well as shaft motion test data, reveal shaft motions rich in subsynchronous activity throughout the entire speed range. The differences noted demonstrate the dramatic effect of engine-induced housing accelerations on TC shaft motions. Predicted total shaft motions agree remarkably well with test data, differing by a small amount. Also, total shaft motion amplitudes (relative to compressor housing) remain nearly constant with increasing engine speed, even though housing acceleration amplitudes increase with engine speed. This indicates that the TC shaft is isolated from housing (and engine) motion. Otherwise, engine load conditions would induce larger and more harmful motion amplitudes, even at high shaft speeds. Agreement with test data is particularly good above engine speeds greater than 3000 rpm (shaft speed ~148 krpm). Predicted TC subsynchronous amplitudes as well as peak vibration amplitudes show a fair agreement with the test data, especially between 1750–3000 rpm engine speed. However, predictions indicate TC shaft vibrations are attributed to engine order frequencies $6e$ and below, whereas measurements show marked amplitude responses at order frequencies greater than $6e$. The analysis also reveals that superharmonics of the engine speed can excite TC damped natural frequencies, although in this instance the mode shape (rotor conical mode) is well-damped. Nonetheless, the results demonstrate the importance in including housing accelerations in TC system models.

Most importantly, the comparisons of predictions to test data serve to validate this first step in incorporating engine-induced housing excitations into the rotordynamics model. The research furthers the goal to create a virtual laboratory aiding a manufacturer to reduce TC development time and costs.

Acknowledgment

The generous support and technical guidance from Honeywell Turbo Technologies as well as the permission to publish this work are gratefully acknowledged.

Appendix: Predictions of Floating-Ring Bearing Performance Parameters

The thin film flow model predicts the static and dynamic forced performances of the SFRB films at the compressor and turbine sides as a function of the operating speed, lubricant supply conditions, and thermal boundary conditions. Figure 16 depicts predicted exit film temperatures for the inner film and outer film of

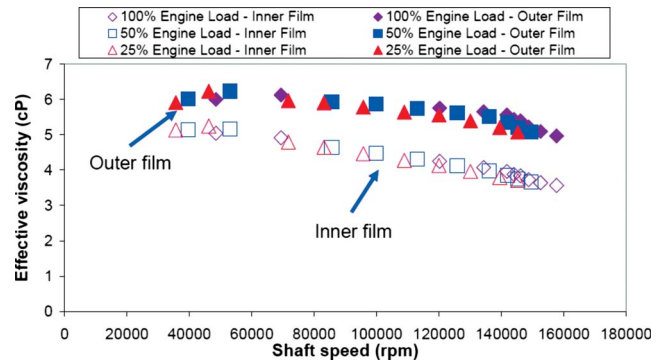


Fig. 17 Predicted effective lubricant film viscosities of the compressor bearing end for 25%, 50%, and 100% engine loads

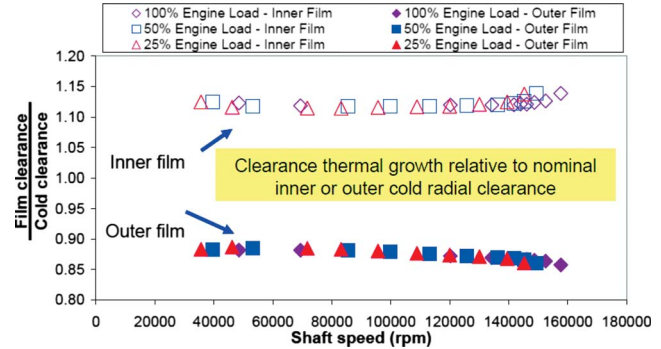


Fig. 18 Predicted inner and outer film clearances relative to nominal cold radial clearances of the compressor bearing end for 25%, 50%, and 100% engine loads

the compressor bearing end at three engine loads. The lubricant supply temperature is shown to depict film temperature increase. Turbine bearing end predictions are nearly identical to compressor bearing end predictions and not shown for brevity. Expectedly, the inner film temperature is greater than the outer film temperature since the inner film dissipates more power. Note that the temperature increase, defined as the difference between exit film temperature and supply temperature, decreases for the outer film with increasing shaft speed.

Figure 17 illustrates predicted lubricant effective viscosities (at effective film temperature) for increasing shaft speeds. The lubricant viscosity decreases with increasing shaft speed due to increased film temperatures and an increase in drag power loss (i.e., higher shear forces). Note that the inner film viscosity decreases more rapidly than the outer film viscosity.

Thermal growth of the bearing (shaft, ring, and housing) due to increasing lubricant temperatures causes the bearing clearances to change significantly, as shown in Fig. 18. The figure shows bearing clearances with respect to nominal cold temperature radial clearances. As shaft speed increases, inner film clearance increases while outer film clearance decreases, denoting an expansion of the ring. Note that a thermohydrodynamic fluid flow model is essential to reproduce actual operating conditions in a turbocharger.

References

- [1] Sahay, S. N., and LaRue, G., 1997, "Turbocharger Rotordynamic Instability and Control," NASA CP3344, Cleveland, OH, pp. 247–257.
- [2] Shaw, M., and Nussdorfer, T., 1949, "An Analysis of the Full-Floating Journal Bearing," NACA Report No. 866.
- [3] Tataru, A., 1970, "An Experimental Study of the Stabilizing Effect of Floating-Bush Journal Bearings," Bull. JSME, 13(61), pp. 858–863.
- [4] Li, C., and Rohde, S., 1981, "On the Steady State and Dynamic Performance

- Characteristics of Floating Ring Bearings.” ASME J. Lubr. Technol., **103**(3), pp. 389–397.
- [5] Li, C., 1982, “Dynamics of Rotor Bearing Systems Supported by Floating Ring Bearings,” ASME J. Lubr. Technol., **104**(4), pp. 469–477.
- [6] Trippett, R., and Li, D., 1984, “High-Speed Floating-Ring Bearing Test and Analysis,” ASLE Trans., **27**(1), pp. 73–81.
- [7] Tanaka, M., and Hori, Y., 1972, “Stability Characteristics of Floating Bush Bearings,” ASME J. Lubr. Technol., **94**(3), pp. 248–259.
- [8] Holmes, R., Brennan, M. J., and Gottrand, B., 2004, “Vibration of an Automotive Turbocharger—A Case Study,” *Proceedings 8th International Conference on Vibrations in Rotating Machinery*, Swansea, UK, Paper C623/011/2004, pp. 445–450.
- [9] Holt, C., San Andrés, L., Sahay, S., Tang, P., LaRue, G., and Gjika, K., 2005, “Test Response and Nonlinear Analysis of a Turbocharger Supported on Floating Ring Bearings,” ASME J. Vibr. Acoust., **127**, pp. 107–212.
- [10] San Andrés, L., and Kerth, J., 2004, “Thermal Effects on the Performance of Floating Ring Bearings for Turbochargers,” Proc. Inst. Mech. Eng., Part J: J. Eng. Tribol., **218**(5), pp. 437–450.
- [11] San Andrés, L., Rivadeneira, J., Chinta, M., Gjika, K., and LaRue, G., 2007, “Nonlinear Rotordynamics of Automotive Turbochargers: Predictions and Comparisons to Test Data,” ASME J. Eng. Gas Turbines Power, **129**(2), pp. 488–493.
- [12] San Andrés, L., Rivadeneira, J. C., Gjika, K., Groves, C., and LaRue, G., 2007, “Rotordynamics of Small Turbochargers Supported on Floating Ring Bearings—Highlights in Bearing Analysis and Experimental Validation,” ASME J. Tribol., **129**, pp. 391–397.
- [13] San Andrés, L., Rivadeneira, J. C., Gjika, K., Groves, C., and LaRue, G., 2007, “A Virtual Tool for Prediction of Turbocharger Nonlinear Dynamic Response: Validation Against Test Data,” ASME J. Eng. Gas Turbines Power, **129**(4), pp. 1035–1046.
- [14] Gjika, K., and Groves, C., 2006, “Nonlinear Dynamic Behavior of Rotor-Bearing Systems Involving Two Hydrodynamic Oil Films in Series: Prediction and Test Application to High-Speed Turbochargers,” *Proceedings of the 8th Biennial ASME Conference on Engineering Systems Design and Analysis*, ASME Paper No. ESDA2006-95792.
- [15] Gjika, K., Groves, C., San Andrés, L., and LaRue, G., 2007, “Nonlinear Dynamic Behavior of Turbocharger Rotor-Bearing Systems With Hydrodynamic Oil Film and Squeeze Film Damper in Series: Prediction and Experiment,” ASME Paper No. DETC2007-34136.
- [16] Kirk, R. G., Alsaeed, A. A., and Gunter, E. J., 2007, “Stability Analysis of a High-Speed Automotive Turbocharger,” STLE Tribol. Trans., **50**(3), pp. 427–434.
- [17] Kirk, R., Alsaeed, A., Liptrap, J., Lindsey, C., Sutherland, D., Dillon, B., Saunders, E., Chappell, M., Nawshin, S., Christian, E., Ellis, A., Mondschein, B., Oliver, J., and Sterling, J., 2008, “Experimental Test Results for Vibration of a High Speed Diesel Engine Turbocharger,” STLE Tribol. Trans., **51**(4), pp. 422–427.
- [18] Hori, Y., and Kato, T., 1990, “Earthquake-Induced Instability of a Rotor Supported by Oil Film Bearings,” ASME J. Vibr. Acoust., **112**(2), pp. 160–165.
- [19] Lee, A., Kim, B., and Kim, Y., 2006, “A Finite Element Transient Response Analysis Method of a Rotor-Bearing System to Base Shock Excitations Using the State-Space Newmark Scheme and Comparisons With Experiments,” J. Sound Vib., **297**(3–5), pp. 595–615.
- [20] Childs, D., 1993, *Turbomachinery Rotordynamics*, Wiley, New York, Chap. 4.
- [21] Maruyama, A., 2007, “Predictions of Automotive Turbocharger Nonlinear Dynamic Forced Response With Engine Induced Housing Excitations: Comparisons to Tests Data,” MS thesis, Texas A&M University, College Station, TX.

Measurement of Structural Stiffness and Damping Coefficients in a Metal Mesh Foil Bearing

Luis San Andrés

Mast-Childs Professor
e-mail: lsanandres@tamu.edu

**Thomas Abraham
Chirathadam**

Research Assistant

Tae-Ho Kim¹

Postdoctoral Research Associate

Department of Mechanical Engineering,
Texas A&M University,
College Station, TX 77843-3123

Engineered metal mesh foil bearings (MMFBs) are a promising low cost bearing technology for oil-free microturbomachinery. In a MMFB, a ring shaped metal mesh provides a soft elastic support to a smooth arcuate foil wrapped around a rotating shaft. This paper details the construction of a MMFB and the static and dynamic load tests conducted on the bearing for estimation of its structural stiffness and equivalent viscous damping. The 28.00 mm diameter 28.05 mm long bearing, with a metal mesh ring made of 0.3 mm copper wire and compactness of 20%, is installed on a test shaft with a slight preload. Static load versus bearing deflection measurements display a cubic nonlinearity with large hysteresis. The bearing deflection varies linearly during loading, but nonlinearly during the unloading process. An electromagnetic shaker applies on the test bearing loads of controlled amplitude over a frequency range. In the frequency domain, the ratio of applied force to bearing deflection gives the bearing mechanical impedance, whose real part and imaginary part give the structural stiffness and damping coefficients, respectively. As with prior art published in the literature, the bearing stiffness decreases significantly with the amplitude of motion and shows a gradual increasing trend with frequency. The bearing equivalent viscous damping is inversely proportional to the excitation frequency and motion amplitude. Hence, it is best to describe the mechanical energy dissipation characteristics of the MMFB with a structural loss factor (material damping). The experimental results show a loss factor as high as 0.7 though dependent on the amplitude of motion. Empirically based formulas, originally developed for metal mesh rings, predict bearing structural stiffness and damping coefficients that agree well with the experimentally estimated parameters. Note, however, that the metal mesh ring, after continuous operation and various dismantling and re-assembly processes, showed significant creep or sag that resulted in a gradual decrease in its structural force coefficients. [DOI: 10.1115/1.3159379]

1 Introduction

Oil-free microturbomachinery (<250 kW) implements gas bearings to ensure low friction and long operating life with little maintenance [1]. Due to their simple construction and low cost, metal mesh foil gas bearings (MMFBs) aspire to satisfy some of the requirements for high speed machinery, such as in passenger vehicle turbochargers. MMFBs, as shown in Fig. 1, have a top foil supported on a ring shaped metal mesh (MM) that provides structural stiffness and mechanical energy dissipation with material hysteresis [2]. As the rotor speed increases, a hydrodynamic gas film pressure builds up and separates the rotor from the top foil, thus ensuring low friction operation and adequate load capacity [3].

Since 1980s, metal mesh material have been employed as vibration isolators [4] in gas turbine engines. For example, a ring shaped metal mesh damper, installed in series with the outer race of a rolling element bearing, can provide a soft support to a jet engine turbine shaft, thus removing the additional space and com-

plexity of a squirrel cage. Exhaustive experimental work in Ref. [5] also demonstrates that ring shaped metal mesh dampers provide equivalent viscous damping as large as that of oil-lubricated squeeze film dampers. The major advantage of a metal mesh damper is its ability to withstand either high or low temperatures, as opposed to lubricants. The following summarizes literature on metal mesh vibration isolators and dampers.

In an early application, Okayasu et al. [6] used a metal mesh as a flexible bearing support in a liquid hydrogen turbopump to reduce high amplitude rotor vibrations while traversing critical speeds and to control subsynchronous instability that arise due to dry friction between rotating parts. The turbopump was designed to operate above its third critical speed of 46,139 rpm. In the absence of metal mesh dampers, high vibration levels (80–150 μm pk-pk) were measured while crossing the first and second critical speeds. Introduction of metal mesh dampers in the rotor-bearing system attenuated synchronous and subsynchronous rotor motions.

Zarour [5] introduced a stainless steel mesh damper replacing a squeeze film damper (SFD) as an effective viscous damping element. Experiments in a test rig replicating a midsize power gas turbine demonstrate that the viscous damping coefficients of the metal mesh dampers are comparable to those of similar sized SFDs. The effective damping of the metal mesh damper does not change significantly in tests covering a temperature range of 54–99°C, proving its reliability as a damping element in this temperature range. However, the stiffness decreases (softening effect) with an increase in the operating temperature. Importantly

¹Presently at Korea Institute of Science and Technology, Energy Mechanics Research Center, 39-1 Hawolgok-dong, Songbuk-gu, Seoul, Korea 136-791, Senior Research Scientist, e-mail: thk@kist.re.kr.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 23, 2009; final manuscript received March 27, 2009; published online December 3, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Gas Turbine Technical Congress and Exposition, Orlando, FL, June 8–12, 2009.

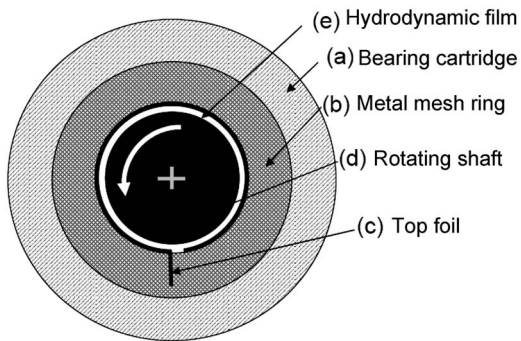


Fig. 1 Metal mesh foil bearing

enough, the amplitude of synchronous response of the rotor while crossing its critical speed does not change when the metal mesh is drenched in oil. Note that the metal mesh damper shows during static load tests a deflection proportional to the applied force. However, a nonlinear deflection follows in the unloading process.

Metal mesh dampers retain its damping when assembled in conjunction with other stiffening elements. Al-Khateeb and Vance [7] reported that a ring shaped metal mesh damper in parallel with a squirrel cage, a stiffening member, allows varying the support stiffness without any effect in its equivalent viscous damping. Ertas et al. [8] investigated the applicability of metal mesh as a bearing damper at cryogenic temperatures. The authors reported an increase in metal mesh damper stiffness with a decrease in temperature for both steel and copper meshes. The equivalent viscous damping coefficient of a steel mesh damper does not appear to change with temperature. However, the equivalent viscous damping increases significantly for a copper mesh damper with a decrease in temperature, achieving its highest value at the lowest test temperature of -190°C . The selection of the mesh material depends mainly on the damping it provides. Burshid [9] reported using a copper mesh for auxiliary sealing in a hybrid damper seal design. Tests show that a copper wire mesh exhibits more damping than a stainless steel wire mesh.

Knowledge of the physical parameters affecting the viscous damping and stiffness of a metal mesh damper is necessary for its engineered design and scalability. Al-Khateeb [2] found that dry friction and at least one more source of damping, possibly material hysteresis, influence the overall viscous damping of a metal mesh ring damper. Dynamic load tests on a metal mesh damper evidence frequency and motion amplitude dependent stiffness and equivalent viscous damping coefficients, i.e., a nonlinear forced performance. Choudhry and Vance [10] developed design equations, empirically based, predicting stiffness and viscous damping coefficients of metal mesh ring dampers. These design equations incorporate the metal mesh ring geometry and material compactness ratio, and rely on simple tests to identify the metal mesh effective Young's Modulus, at any specified excitation frequency.

Metal mesh dampers display nonlinear stiffness and viscous damping properties, which make their prediction difficult. Ertas and Luo [11] reported characterizing the nonlinear stiffness and damping of metal mesh rings by varying the vibration amplitude, excitation frequency, and static eccentricity due to an applied load. The tests show that the eccentricity has no influence on the damper force coefficients. Also, a maximum viscous damping coefficient occurs at the lower frequency ranges (0–50 Hz). The viscous damping coefficient reduces from ~ 29 kN s/m (at 20 Hz) to ~ 3.5 kN s/m (at 300 Hz). In a high speed rotordynamic test rig implementing a metal mesh ring as a damping element, the authors successfully demonstrate low rotor vibration levels while it traverses the first two critical speeds.

Recently, Ertas [12] reported a novel compliant hybrid journal bearing for enabling oil-free operation in high speed turbomachinery. The design integrates a multiple pad hybrid bearing with flex-

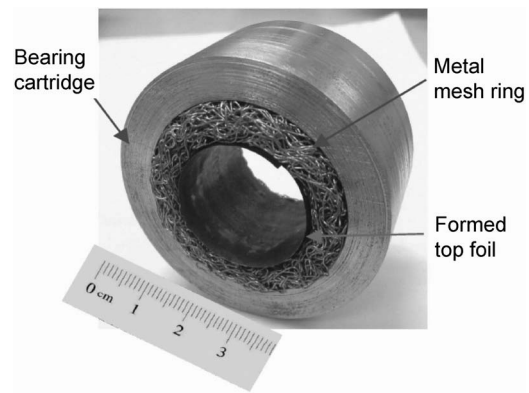


Fig. 2 Photograph of a metal mesh foil bearing

ural type supports and two metal mesh rings inserted under the pads at the bearing sides. The composite bearing maximizes load carrying capacity and effective damping without sacrificing compliance to misalignment between the rotor and bearing supports. Test data show that the viscous damping decreases with increasing excitation frequency. The bearing direct stiffness decreases as the excitation frequency increases when there is no external pressure applied into the bearing pads. However, with an applied hydrostatic pressure, the bearing direct stiffness shows an increasing trend with excitation frequency. Also, subsynchronous vibration components with a frequency equal to the system natural frequency, present in rotordynamic tests without the dampers, are absent with the metal mesh dampers installed.

Gas foil bearings (GFBs) [1], for instance, are a reliable and proven oil-free bearing technology. However, each foil bearing is custom designed for the application, leading to variability even among similar sized bearings and limited scalability [13]. Introducing an inexpensive metal mesh ring as the elastic support structure in a foil bearing provides a tunable structural stiffness plus large inherent material damping. This paper presents a metal mesh foil bearing, details its construction method, and advances experimentally identified structural stiffness and viscous damping coefficients with comparisons to predictions. This bearing type is suited to readily replace oil-lubricated floating ring bearings in automotive turbochargers, for example. MMFBs are simple to construct, inexpensive, and made of readily available materials. Hence the work is also of significance to bearing manufacturers.

2 Description of Metal Mesh Foil Bearing

Figure 2 shows a prototype of MMFB constructed in the laboratory, and Table 1 describes the main components comprising the bearing. The bearing uses commercially available ring shaped metal mesh as the elastic support under the top foil. The top foil, a smooth arcuate surface 0.127 mm in thickness, is in contact with the shaft. The metal mesh under the top foil is an elastic support that also provides mechanical energy dissipation from material structural damping and dry friction [2].

The top foil is made from a cold rolled steel strip with significant resilience, inhibiting its permanent deformation into a curved shape. Hence, the steel strip is first baked at a temperature of 927°C and then allowed to cool for ~ 10 h. Furthermore, a tempering process at $\sim 400^{\circ}\text{C}$ recovers the elasticity of the formed top foil. The process follows recommendations in Ref. [14] that detail the heat treatment processes typically employed in the manufacturing of foils in commercial bearings.

Figure 3 illustrates the installation of the top foil inside the metal mesh ring. The top foil has two identical shaped tabs at one end, which are bent and fit into the two apertures inside the ring shaped metal mesh during assembly, as shown in Figs. 3(a) and 3(b). The preformed top foil is supported on the inner surface of the metal mesh ring that is concentric with the bearing cartridge.

Table 1 Components of metal mesh foil bearing

	Description	Role
Bearing cartridge	Stainless steel cylindrical ring	Holds metal mesh ring and smooth top foil
Metal mesh ring	Ring shaped compressed woven wire mesh	Provides elastic support and energy dissipation mechanism
Smooth top foil (annealed and tempered)	Thin stainless steel sheet, curved and preformed, with one end affixed to the MM ring and the other end free	A hydrodynamic film pressure builds up within the gap between the rotating shaft and the top foil.

After affixing the top foil into the metal mesh, the assembly is inserted into the bearing cartridge to complete the construction of the MMFB. Note that a stringent control of the bearing top foil dimension will be required for reliable application of the bearing in actual machinery.

The overall dimensions of the MMFB follow those of GFBs designed for use in automotive turbochargers. Reference [15], for example, details a patented GFB supported oil-free turbocharger. Table 2 shows the dimensions and material specifications for the test MMFB.

3 Measurements of Bearing Structural Stiffness

A series of static load versus bearing deflection tests aid to estimate the static load performance of the constructed MMFB. The tests are conducted at room temperature. Figure 4 shows a schematic representation of the static load test setup with a

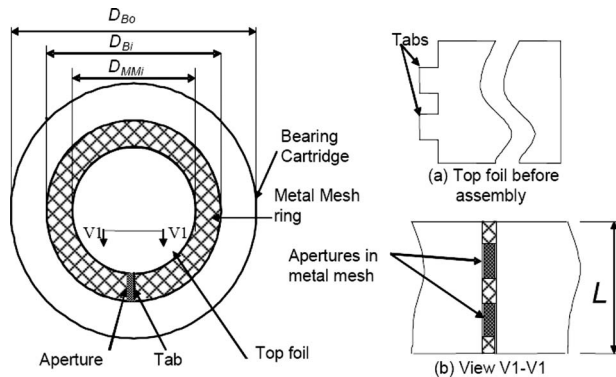


Fig. 3 Dimensions in a MMFB and views of (a) tabs at one end of top foil before assembly and (b) two apertures in metal mesh ring

Table 2 Nominal dimensions and material specifications for test MMFB

Parameter name and physical dimension	Magnitude
Bearing cartridge outer diameter, D_{BO} (mm)	58.15 ± 0.02
Bearing cartridge inner diameter, D_{BI} (mm)	42.10 ± 0.02
Bearing axial length, L (mm)	28.05 ± 0.02
Metal mesh outer diameter, D_{MMO} (mm)	42.10 ± 0.02
Metal mesh inner diameter, D_{MMI} (mm)	28.30 ± 0.02
Metal mesh mass, M_m (kg)	0.0391
Metal mesh density, ρ_{MM} (%)	20
Top foil thickness, T_{TF} (mm)	0.127
Wire diameter, D_w (mm)	0.30
Copper Young modulus, E (GPa), at 21 °C [16]	114
Copper Poisson ratio, ν [17]	0.33
Bearing mass (cartridge+mesh+foil), M (kg)	0.318

Manufacturers define the density of metal mesh as the ratio of the ring mass to its volume times the metal material density.

Uncertainty in mass measurement = ± 0.0001 kg.

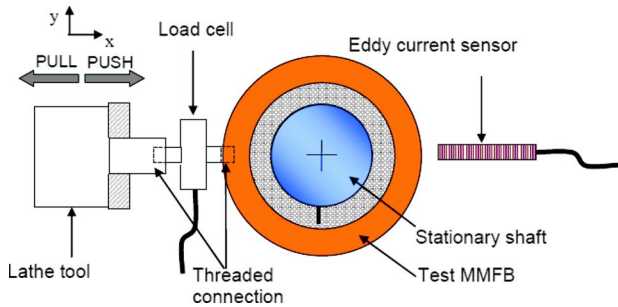


Fig. 4 Schematic view of setup for static load versus deflection tests in MMFB

Table 3 MMFB static load versus deflection tests: coefficients of polynomial curve fit

Loading direction	Load F (N) versus deflection X (mm)
Push load	$F = 35448X^3 - 2411X^2 + 891X + 48.4$
Pull load	$F = 39765X^3 + 3635X^2 + 818X - 39.1$

X range $[-0.1, 0.1]$ mm, $R^2 = 0.997$ is the correlation factor indicating the goodness of the curve fit to test data.

MMFB supported on a rigid shaft, one end affixed in a lathe chuck and the other end supported by the lathe center.² An eddy current sensor (ECS), rigidly affixed to the lathe base with a sensitivity of 8.24 V/mm, measures horizontal displacements (X) of the bearing, and a strain gauge type load cell, with a sensitivity of 0.225 V/N, records the static load. One end of the load cell is threaded directly to the bearing midspan, and the other end to a cylindrical steel rod affixed in a lathe tool holder. The forward and backward motions of the lathe tool holder transmit, through the load cell, push and pull loads to the bearing cartridge (along the x direction).

Three cycles of push and pull loads (F) are applied to the MMFB, and displacements (X) are measured. The test data, modeled as a cubic polynomial in the displacement (X), provide a relation between the static load (F) and measured displacement (X). The MMFB static stiffness coefficient ($K = \partial F / \partial X$) follows from differentiation of the polynomial found. Table 3 presents the coefficients of the cubic polynomials and the goodness of fit. Note that a cubic polynomial is the lowest order nonlinear odd function that represents a physically correct force versus the structural displacement relationship.

Figure 5 shows the applied static load (F) versus the measured MMFB displacement (X). Markers and lines represent test data

²The shaft can be characterized as rigid for the range of loads applied. The procedure calls for the operator to first apply loads on the shaft and record its deflection. Typically shaft deflections are two orders of magnitude smaller than those from the MMFB.

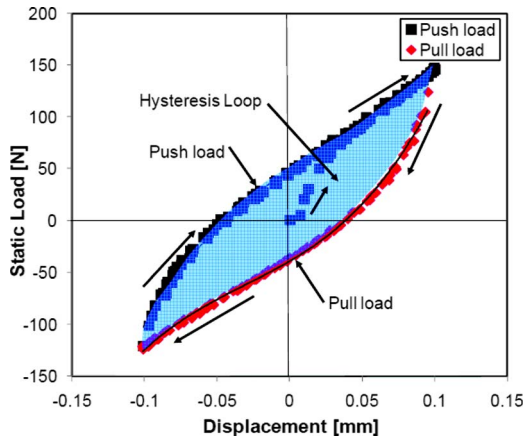


Fig. 5 Typical applied static load versus measured MMFB displacement

and their respective cubic polynomial curve fit. Note the hysteresis loop, representing mechanical energy dissipation, due to the nonlinear load versus deflection behavior of the MMFB during the push-pull loading tests. The sizeable loop evidences large mechanical energy dissipation, and thus anticipates the bearing capability to quickly dampen vibration amplitudes.

Figure 6 shows the estimated MMFB structural stiffness (K) versus bearing static deflection (X). Push and pull loads result in two different curves for the estimated stiffness coefficients of MMFB. A test starts with no load ($F=0$) and zero displacement ($X=0$) and continues in the forward direction until the maximum deflection point (A). Then, as the load reversal occurs, the stiffness instantaneously shifts to point B and continues along the backward stiffness curve till point C. Again, as the loading reverses to the forward direction, the stiffness instantly shifts to point D. The sudden changes in stiffness at the points of load reversal are thought to be related to hysteresis and dry-friction effects of the metal mesh.

4 Dynamic Load Tests on Metal Mesh Foil Bearing

Dynamic load tests aid to identify the structural stiffness and equivalent viscous damping of the MMFB for increasing bearing displacements and excitation frequencies. Figure 7 shows a schematic view of the setup for the dynamic load tests on the MMFB supported on a rigid shaft of diameter equal to 28.00 mm. A fixture, clamped to a rigid steel table, holds the shaft firmly. The fixture minimizes the motion between the shaft and the ground so that the bearing is the only component in motion during shaker

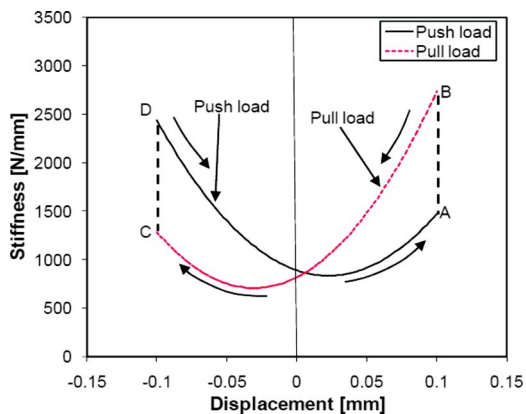


Fig. 6 Estimated structural stiffness versus MMFB displacement

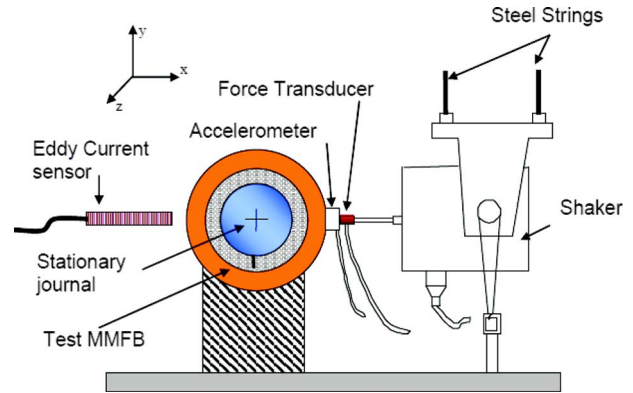


Fig. 7 Schematic view of setup for dynamic load tests with MMFB supported on rigid shaft

testing. The electromagnetic shaker, suspended with cables from the ceiling of the test facility, provides MMFB excitation forces along the horizontal direction for increasing excitation frequencies. An eddy current sensor measures the MMFB deflection. Reference [18], also identifying structural force coefficients of bump-type foil bearings, details further the test setup and procedure.

The load, displacement, and acceleration are taken as inputs to the data acquisition system. Two eddy current sensors (sensitivity 8.24 V/mm) measure the displacement of the bearing and the shaft, a piezoelectric accelerometer (sensitivity 10.73 mV/g), magnetically mounted on the bearing cartridge, measures the acceleration of the bearing cartridge, and a force transducer (sensitivity 11.2 mV/N), affixed into the bearing cartridge at its axial midspan, measures the dynamic load acting on the MMFB.

The experiments for identification of MMFB structural force coefficients involve single frequency dynamic load excitations with three prescribed MMFB motion amplitudes ($12.7 \mu\text{m}$, $25.4 \mu\text{m}$, and $38.1 \mu\text{m}$). The dynamic load is applied with frequencies ranging from 25 Hz to 400 Hz, varying in 25 Hz increments. All test conditions are repeated five times and averages of the identified parameters are presented. A LABVIEW® in-house program controls inputs to the shaker, acquires outputs from sensors, and displays real-time test results. The user inputs the desired amplitude of dynamic displacement, and a PC algorithm controls the voltage amplitude to the electromagnetic shaker to get the desired load.

Figure 8 shows a waterfall plot of the dynamic loads for excitation frequencies from 25 Hz to 400 Hz. The amplitude of dynamic load to maintain fixed bearing amplitude of motion de-

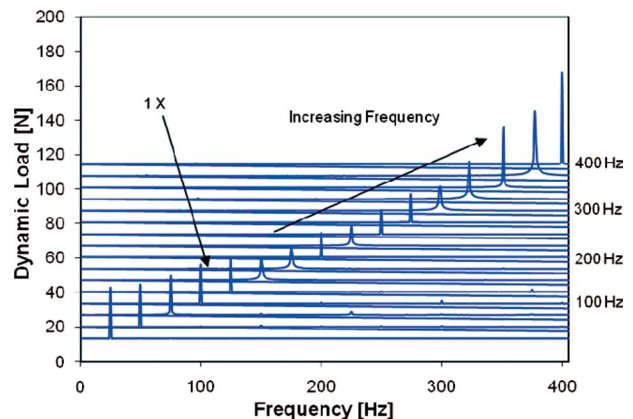


Fig. 8 Waterfall plot of dynamic load for bearing with motion amplitude of $38.1 \mu\text{m}$ and excitation frequencies from 25 Hz to 400 Hz

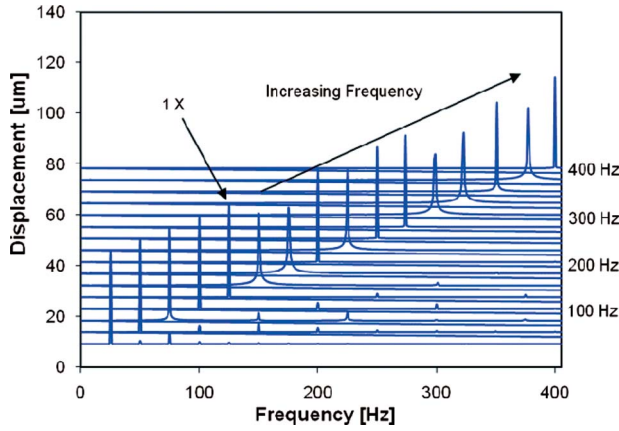


Fig. 9 Waterfall plot of bearing displacements for motion amplitude of 38.1 μm and excitation frequencies from 25 Hz to 400 Hz

increases with increasing frequency until ~ 200 Hz, and increases for higher frequencies (>300 Hz). Figure 9 shows a waterfall plot of the MMFB dynamic displacement response to the forced excitations as frequency increases from 25 Hz to 400 Hz. Data show dominant MMFB motion amplitudes (1X) synchronous to the excitation frequency. Motion amplitudes in other harmonics are negligibly small. In general, the amplitude of dynamic load decreases as the displacement amplitude decreases for all frequencies. The dynamic load is the lowest near the system natural frequency at ~ 250 Hz. For higher excitation frequencies (>300 Hz), the dynamic load increases rapidly with increasing frequency.

5 Estimation of Mechanical Parameters in Metal Mesh Foil Bearing

A single degree of freedom mechanical system with constant parameters represents the nonrotating test system [18]. The MMFB motions $X(t)$ due to an excitation force $F(t)$ is governed by

$$M\ddot{X} + C\dot{X} + KX = F(t) \quad (1)$$

where M is the bearing mass, and C and K are the bearing equivalent viscous damping and structural stiffness coefficients, respectively. For a single frequency (ω) forced excitation, $F = \bar{F}e^{i\omega t}$, the bearing response is $X = \bar{X}e^{i\omega t}$. Hence, Eq. (1) in the frequency domain becomes

$$(K - \omega^2 M)\bar{X} + i\omega C\bar{X} = \bar{F} \quad (2)$$

Ready evaluation of the stiffness (K) and viscous damping (C) coefficients follows from curve fits of the test data to the real and imaginary parts of the impedance function Z , i.e.,

$$Z = \frac{\bar{F}}{\bar{X}} \Rightarrow K - \omega^2 M = \text{Re}(Z), \quad C = \frac{\text{Im}(Z)}{\omega} \quad (3)$$

In addition, from fundamental vibrations knowledge [19], the relationship between the viscous damping coefficient C and the structural loss factor (γ) representing material damping is

$$\gamma = \frac{C\omega}{K} \quad (4)$$

Reference [10] reports empirical equations for predicting stiffness and damping coefficients, derived from structural dynamic load test data in four metal mesh dampers (MMDs) of various radial thicknesses. In the reference, dynamic load tests quantify

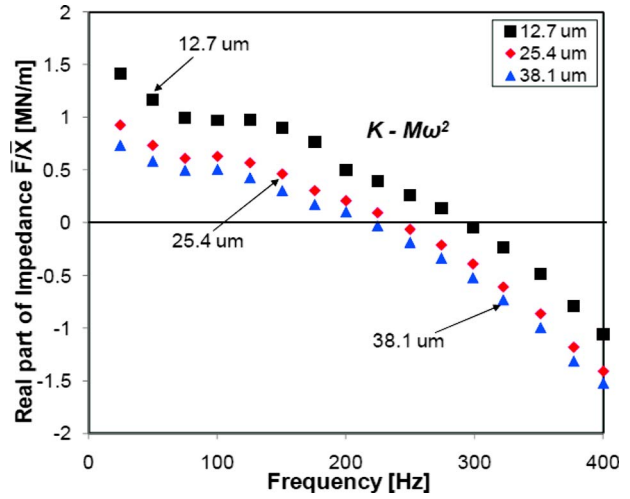


Fig. 10 Estimated MMFB dynamic stiffness ($K - \omega^2 M$) versus excitation frequency for three motion amplitudes (12.7 μm , 25.4 μm , and 38.1 μm)

the effects of axial compression, radial interference, and motion amplitude on the stiffness and damping coefficients of the tested metal mesh dampers. The tested MMDs have a wire diameter (D_w) of 0.32 mm and a density (ρ_{MM}) of 42.7%. Reference [2] provides further test data for metal mesh ring force coefficients for increasing excitation frequencies and extends the predictive equations to include the effects of frequency.

The predictive formulas for estimation of the metal mesh damper stiffness (K) and viscous damping (C) coefficients are [2,10]

$$K = E_{\text{equiv}} \cdot \frac{L}{(R_o - R_i)} \cdot \left[1 + 4 \cdot 10^{-5} \left(\frac{C_A}{L} \right)^2 \right] \cdot [1.0002\omega] \cdot \left[1 + 0.0296 \left(\frac{R_p}{R_o - R_i} \right) \right] \cdot \left[1 + 0.112 \cdot \left(\frac{A}{R_o - R_i} \right)^{-2/3} \right] \quad (5)$$

$$C\omega = H_{\text{equiv}} \frac{L}{(R_o - R_i)} \cdot \left[1 + 8.7 \cdot 10^{-5} \left(\frac{C_A}{L} \right)^2 \right] \cdot \left[\left(\frac{\omega}{\omega_n} \right)^{-4/5} \right] \cdot \left[1 + 1.8 \cdot 10^{-2} \left(\frac{R_p}{R_o - R_i} \right)^{3/2} \right] \cdot \left[1 + 0.518 \cdot \left(\frac{A}{R_o - R_i} \right)^{-2/3} \right] \quad (6)$$

where E_{equiv} and H_{equiv} are equivalent elastic moduli (direct and quadrature) determined empirically. Above L is the ring axial length with inner and outer radii (R_i and R_o), C_A is an axial compression ratio, and R_p is the radial interference. The force coefficients are functions of the motion amplitude (A) and excitation frequency (ω).

Experimentally estimated K and C at an arbitrary frequency of 100 Hz for the highest motion amplitude of 38.1 μm are input into the equations above to extract the empirical coefficients $E_{\text{equiv}} = 31$ kN/m and $H_{\text{equiv}} = 6.6$ kN/m. Note that $C_A = 0$ and $R_p = 0$ for the current test bearing. Subsequently, Eqs. (5) and (6), with the derived E_{equiv} and H_{equiv} , predict the MMFB stiffness and equivalent viscous damping coefficients for various motion amplitudes and increasing excitation frequencies.

Figure 10 shows the bearing dynamic stiffness $\text{Re}(Z) = K - \omega^2 M$ decreasing with either an increase in excitation frequency or an increase in amplitude of motion. Note that the system natural frequency decreases as the motion amplitude increases, as is typical in a nonlinear system with a softening stiffness. The

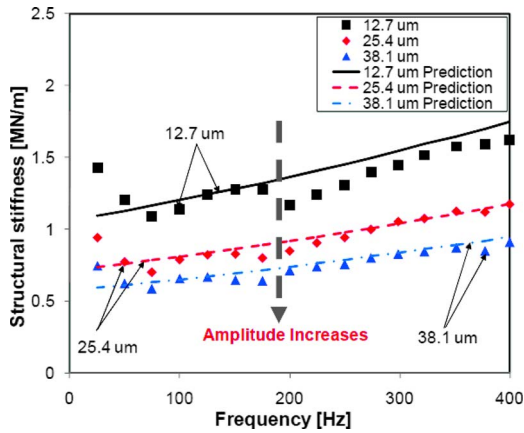


Fig. 11 Experimental and predicted MMFB structural stiffness (K) versus frequency. Results for three motion amplitudes ($12.7 \mu\text{m}$, $25.4 \mu\text{m}$, and $38.1 \mu\text{m}$).

MMFB structural stiffnesses identified are comparable with those obtained in similar sized bump-type foil bearings [18]. Figure 11 shows the test identified and predicted MMFB structural stiffness (K) versus frequency. The bearing structural stiffness decreases with increasing motion amplitudes. In the range of low to moderate frequencies (25–100 Hz), the stiffness K decreases rapidly ($\sim 30\%$). However, for larger frequencies, K increases gradually. The predictions show a good correlation with identified structural stiffness values.

Figure 12 shows the test identified and predicted MMFB equivalent viscous damping coefficient (C) decreasing rapidly with excitation frequency and with amplitude of motion, i.e., $C \approx 1/(\omega\bar{X})$. Note the logarithmic scale for the vertical axis. At the highest test frequency of 400 Hz, the damping magnitude is two orders less than that at 25 Hz. The predictions follow closely the test identified coefficients.

The test results show that the mechanical energy dissipation mechanism in a MMFB is not of viscous type. A more realistic model is material hysteresis or structural damping. Figure 13 depicts the estimated structural loss factor, $\gamma \sim C\omega/K$, versus frequency of excitation for the various load tests. The loss factor (γ) is as large as 0.70, a significant magnitude for the simple mechani-

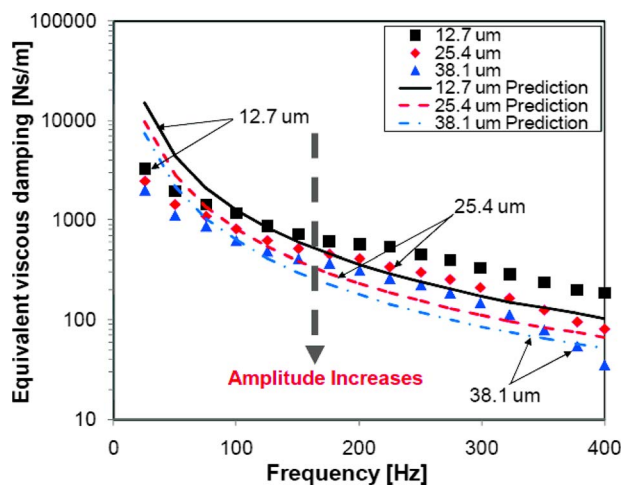


Fig. 12 Experimental and predicted MMFB viscous damping (C) versus excitation frequency for three motion amplitudes ($12.7 \mu\text{m}$, $25.4 \mu\text{m}$, and $38.1 \mu\text{m}$)

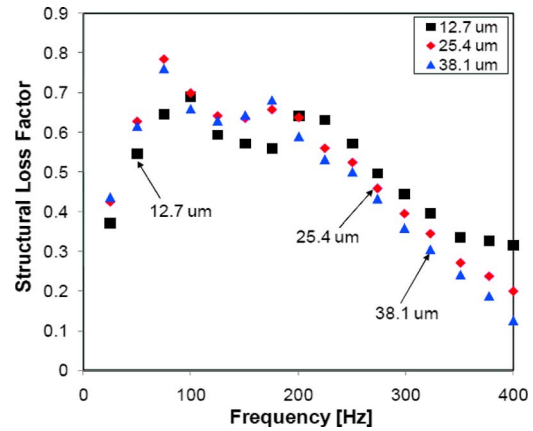


Fig. 13 MMFB structural loss factor (γ) versus excitation frequency for three motion amplitudes ($12.7 \mu\text{m}$, $25.4 \mu\text{m}$, and $38.1 \mu\text{m}$)

cal system. The loss factor, achieving its highest value near the system natural frequency, is a desirable feature to readily attenuate amplitudes of motion.

Prior MMFB experimental force coefficients in Ref. [20], obtained 3 months earlier, are larger in magnitude than the current ones. After undergoing continuous operation and multiple dismantling and reassembly processes, the stiffness and equivalent viscous damping coefficients show a degradation of $\sim 60\%$ and $\sim 50\%$, respectively. The estimated loss factor increased from 0.5 to 0.7 in average. The changes in the force coefficients are mainly due to variations in clearances resulting from creep or sag, and by multiple dismantling and reassembly processes. Al-Khateeb [2] made a similar observation, reporting a 57% and 74% reduction in stiffness and damping, respectively, over a period of 12 months. Al-Khateeb [2] recommended an axial compression and positive radial interference to avoid relaxation of the MM bearing force coefficients.

6 Conclusions

The paper details the construction of a MMFB and static and dynamic load tests for the estimation of its structural stiffness and material damping coefficients. Metal mesh, a compressed weave of metal wires, is a simple and inexpensive compliant support structure in gas foil bearings. A metal mesh foil bearing is assembled using a ring shaped metal mesh, a formed and annealed top foil, and a bearing cartridge. The MMFB, 28.00 mm in diameter and 28.05 mm in length, is made of 0.3 mm copper wire with a mesh density (compactness) of 20%.

Static load versus bearing deflection measurements show a non-linear load-deflection relationship with a sizable hysteresis loop, indicating large mechanical energy dissipation. Dynamic load tests aid to the estimation of the bearing mechanical impedance, the real and imaginary parts determining the bearing structural stiffness, and viscous damping coefficients, respectively. The bearing stiffness increases ($\sim 15\%$) with an increase in frequency from 25 Hz to 400 Hz and decreases ($\sim 50\%$) with an increase in amplitude of motion from $12.7 \mu\text{m}$ to $38.1 \mu\text{m}$. The test frequency significantly affects the MMFB equivalent viscous damping. At the lowest test frequency of 25 Hz, the equivalent viscous damping coefficient attains the highest magnitude ($\sim 3.5 \text{ kN s/m}$ and 2 kN s/m for $12.7 \mu\text{m}$ and $38.1 \mu\text{m}$ motion amplitudes, respectively). The equivalent viscous damping coefficient decays rapidly as the excitation frequency increases. For the entire range of test frequencies, the MMFB equivalent viscous damping also reduces with an increase in motion amplitude. The behavior noted is typical of mechanical systems characterized by material or structural damping, hence, the mechanical energy dissipation in a

metal mesh bearing is best described with a loss factor. From the measurements, the test identified loss factor is as high as 0.70, a large magnitude for the simple bearing configuration, showing little dependency on the excitation frequency. The loss factor is the highest around the system natural frequency of ~ 250 Hz, a desirable feature to control large vibration amplitudes at the natural frequency. Predicted MMFB structural stiffness and viscous damping coefficients, derived from empirically based equations for metal mesh dampers, show very good agreement with the experimentally derived force coefficients. However, the experiments also show a measureable reduction in MMFB stiffness and damping coefficients after multiple dismantling, reassembly processes, and continuous operation. The finding compromises the reliability of MMFBs for actual implementation in commercial machinery.

Presently, the MMFB undergoes rotordynamic tests in a high speed test rig. See Ref. [21] for details on the lift off conditions of the MMFB operation to a top speed of 40 krpm.

Acknowledgment

The support of the Turbomachinery Research Consortium (TRC) and Honeywell Turbo Technologies is gratefully acknowledged. The authors also thank Kinetic Structures Corporation for donating the metal mesh rings.

Nomenclature

A	= motion amplitude (m)
C	= equivalent viscous damping coefficient (N s/m)
C_A	= axial compression ratio
D_{BI}	= bearing cartridge inner diameter (m)
D_{BO}	= bearing cartridge outer diameter (m)
D_{MMI}	= metal mesh ring inner diameter (m)
D_{MMO}	= metal mesh ring outer diameter (m)
D_W	= metal wire diameter (m)
E	= Young's modulus (N/m ²)
E_{equiv}	= equivalent elastic modulus (N/m)
F	= force (N)
H_{equiv}	= equivalent quadrature elastic modulus (N/m)
K	= bearing structural stiffness (N/m)
L	= bearing axial length (m)
M	= bearing mass (kg)
M_m	= metal mesh ring mass (kg)
R_i	= metal mesh ring inner radius (m)
R_o	= metal mesh ring outer radius (m)
R_p	= radial interference (m)
T_{TF}	= top foil thickness (m)
X	= bearing displacement (m)
Z	= \bar{F}/\bar{X} , bearing impedance (N/m)
ρ_{MM}	= wire density = metal mesh mass/(metal mesh volume \times density of metal)

ν	= Poisson's ratio
γ	= metal mesh bearing structural loss factor
ω	= frequency of excitation (rad/s)

References

- [1] Agrawal, G., 1997, "Foil Air/Gas Bearing Technology—an Overview," ASME Paper No. 97-GT-347.
- [2] Al-Khateeb, E. M., 2002, "Design, Modeling and Experimental Investigation of Wire Mesh Vibration Dampers," Ph.D. thesis, Texas A&M University, College Station, TX.
- [3] Lee, Y. B., Kim, C. H., Jo, J. H., and Ryu, K., 2006, "Air Foil Bearing Having a Porous Foil," International Patent No. WO 2006/043736 A1.
- [4] Ao, H., Jiang, H., Wei, W., and Ulanov, A. M., 2006, "Study on the Damping Characteristics of MR Damper in Flexible Supporting of Turbo-Pump Rotor for Engine," First International Symposium Systems and Controls in Aerospace and Astronautics, Harbin, China, pp. 618–622.
- [5] Zarzour, M. J., 2000, "Experimental Evaluation of a Metal-Mesh Bearing Damper in a High Speed Test Rig," ASME J. Eng. Gas Turbines Power, **122**(2), pp. 326–329.
- [6] Okayasu, A., Ohta, T., Azuma, T., Fujita, T., and Aoki, H., 1990, "Vibration Problems in the LE-7 Liquid Hydrogen Turbopump," AIAA Paper No. AIAA-1990-2250.
- [7] Al-Khateeb, E. M., and Vance, J. M., 2001, "Experimental Evaluation of a Metal Mesh Bearing Damper in Parallel with a Structural Support," ASME Paper No. 2001-GT-0247.
- [8] Ertas, B. H., Al-Khateeb, E. M., and Vance, J. M., 2003, "Rotordynamic Bearing Dampers for Cryogenic Rocket Engine Turbopumps," J. Propul. Power, **19**(4), pp. 674–682.
- [9] Burshid, S. M., 1990, "Experimental Evaluation of Rotordynamic Coefficients for Hybrid Metal Mesh Pocket Damper Seals in Turbomachinery," MS thesis, Texas A&M University, College Station, TX.
- [10] Choudhry, V., and Vance, J. M., 2005, "Design Equations for Wire Mesh Bearing Dampers in Turbomachinery," ASME Paper No. GT 2005-68641.
- [11] Ertas, B. H., and Luo, H., 2008, "Nonlinear Dynamic Characterization of Oil-Free Mesh Dampers," ASME J. Eng. Gas Turbines Power, **130**(3), p. 032503.
- [12] Ertas, B. H., 2008, "Compliant Hybrid Journal Bearings Using Integral Wire Mesh Dampers," ASME Paper No. GT 2008-50984.
- [13] San Andrés, L., and Kim, T. H., 2008, "Forced Nonlinear Response of Gas Foil Bearing Supported Rotors," Tribol. Int., **41**, pp. 704–715.
- [14] DellaCorte, C., Radil, K. C., Bruckner, J. R., and Howard, S. A., 2007, "Design, Fabrication and Performance of Open Source Generation I and II Compliant Hydrodynamic Gas Foil Bearings," Report No. NASA/TM-2007-214691.
- [15] LaRue, G. D., Kang, S. G., and Wick, W., 2006, "Turbocharger With Hydrodynamic Foil Bearings," US Patent No. 7108488 B2.
- [16] Dodge, H. L., 1913, *The Physical Review*, AIP, New York, p. 439.
- [17] Boyd, J. E., 1917, *Strength of Materials*, McGraw-Hill, New York, p. 16.
- [18] Breedlove, A. W., 2007, "Experimental Identification of Structural Force Coefficients in a Bump-Type Foil Bearing," MS thesis, Texas A&M University, College Station, TX.
- [19] Ginsberg, J. H., 2001, *Mechanical and Structural Vibration—Theory and Application*, Wiley, New York, pp. 135–137.
- [20] San Andrés, L., Kim, T. H., Chirathadam, T. A., and Martinez, A., 2008, "Measurement of Structural Stiffness and Damping in a Metal Mesh Foil Bearing and Development of a Test Rig for Gas Foil Bearings," Texas A&M University, College Station, TX, Technical Report No. TRC-B&C-5-08.
- [21] San Andrés, L., Kim, T. H., Ryu, K., Chirathadam, T. A., Hagen, K., Martinez, A., Rice B., Niedbalski, N., Hung, W., and Johnson, M., 2009, "Gas Bearing Technology for Oil-Free Microturbomachinery—Research Experience for Undergraduate (REU) Program at Texas A&M University," ASME Paper No. GT 2009-59920.

Computational Studies of the Unbalance Response of a Whole Aero-Engine Model With Squeeze-Film Bearings

Philip Bonello

Pham Minh Hai

School of Mechanical, Aerospace and Civil Engineering,
University of Manchester,
Manchester, M60 1QD, UK

The computation of the unbalance vibration response of aero-engine assemblies fitted with nonlinear bearings requires the retention of a very large number of modes for reliable results. This renders most previously proposed nonlinear solvers unsuitable for this application. This paper presents three methods for the efficient solution of the problem. The first method is the recently developed impulsive receptance method (IRM). The second method is a reformulation of the Newmark-beta method. In addition to these two time-domain methods, a whole-engine receptance harmonic balance method (RHBM) is introduced that allows, for the first time, the frequency domain calculation of the periodic vibration response of a real engine. All three methods use modal data calculated from a one-off analysis of the linear part of the engine at zero speed. Simulations on a realistically-sized representative twin-spool engine model with squeeze-film damper bearings provide evidence that the popular Newmark-beta method can be unreliable for large-order nonlinear systems. The excellent correlation between the IRM and RHBM results demonstrates the efficacy of these two complementary tools in the computational analysis of realistic whole-engine models. [DOI: 10.1115/1.3159381]

1 Introduction

Aero-engine assemblies typically have at least two nested rotors mounted within a flexible casing via squeeze-film damper (SFD) bearings. In leading European designs most of the SFDs are unsupported, as can be seen in Fig. 1, where a parallel retainer spring (“squirrel cage”) is only used with one SFD at the end of each rotor for axial location. Such a deployment of SFDs is highly cost effective. However, it requires careful unbalance response calculations that take account of the SFDs’ nonlinearity to ascertain smooth running particularly at the unsupported SFDs [1].

Such calculations can be done in either the time or frequency domain. The former approach uses a numerical technique to march the equations forward in time past the transient stage to predict the steady-state response. Depending on the governing dynamics, this may not necessarily be periodic. The frequency domain approach, based on harmonic balance or trigonometric collocation, is much faster but is restricted to a pre-assumed periodic steady-state response. An efficient computational facility takes advantage of the relative merits of both approaches through an integrated strategy that makes effective use of both [2].

Most proposed computational techniques, although ostensibly generic, have almost invariably been illustrated on simple rotor-bearing systems, e.g., Refs. [2–5]. In fact, they are not tractable to real engine structures due to their complexity and large number of physical degrees of freedom. Moreover, even a transformation of the problem to modal space would require the retention of a large number of modes [1]. This is illustrated in Fig. 2, which shows a typical frequency response function (displacement per unit force or “receptance”) of the linear part of the engine in Fig. 1 at zero rotational speed, computed from a finite element (FE) model supplied by a leading engine manufacturer. “Linear part” means the structure that remains in Fig. 1 when the SFDs are replaced by gaps.

Figure 2 clearly shows a high modal density at relatively low

frequencies. One notes that the sudden reduction in modal density beyond 500 Hz is merely an artifact of the degree-of-freedom reduction technique used by the FE modelers.

The standard time-domain solution approach is to integrate the modal equations of motion with the external forces and nonlinear motion-dependent forces on the right hand side [2,5]. Conventional implicit integration schemes (e.g., Refs. [4–6]) transform these equations into an equal number of algebraic equations which then have to be solved iteratively at each time step to obtain the current modal state variables. Hence, for a whole-engine model, the time-marching process slows down to impractical levels due to the large number of modes (and hence, equations). This problem has been recently resolved by the authors through the novel implicit IRM [1]. The IRM equations relate the relative displacements and velocities at the SFDs with the motion-dependent forces and other excitations acting on the linear part. Hence, the IRM’s computational efficiency is largely immune to the number of modes since the number of equations to be solved at each time step is dependent only on the number of nonlinear elements. This method was validated in Ref. [1] against a standard generic form implicit (SGFI) solver for a realistic engine problem and found to be about 40 times faster. The term “SGFI” means an implicit integration routine that requires expression of the equations of motion in standard first order (state-space) form, or second order form, prior to integration, e.g., the modified Rosenbrock algorithm used by MATLAB solver ode23s© [6]. The Newmark-beta method (NBM) [7] is an implicit alternative to SGFI solvers. In its conventional formulation the NBM suffers from the same inefficiency problem when subjected to a large-order system. However, it will be shown in this paper that the NBM can be reformulated to avoid this problem. Nonetheless, it will be shown that the NBM can be unreliable for a real engine-type problem.

The RHBM was developed in Ref. [2] for the frequency domain analysis of a rotating system comprising a single shaft with negligible gyroscopic effect running on flexibly-housed SFDs. This approach is ideal for a large-order system since it simplifies the problem by using the receptances of the linear part to extract equations governing the harmonics of the response at the SFDs

Manuscript received March 23, 2009; final manuscript received March 27, 2009; published online December 3, 2009. Review conducted by Dilip R. Ballal.

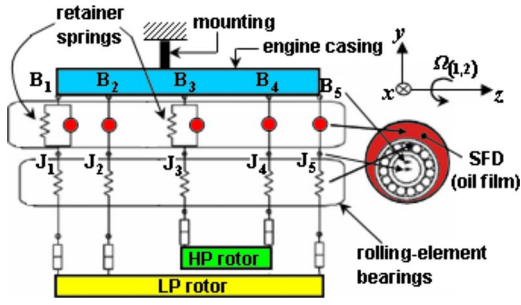


Fig. 1 Schematic of a typical twin-spool engine [1]

only. The present paper resolves outstanding issues with the RHBM, which are essential for its application to whole-engine models. Foremost among these are (a) the consideration of statically indeterminate problems in the “zeroth” harmonic of the vibration [2] and (b) gyroscopic effects. Receptances are computed from a modal series summation [7]. Hence, gyroscopy is an issue since, for RHBM to be feasible, the receptances should pertain to the linear part under nonrotational conditions, making them independent of the rotors’ rotational speeds, thus allowing a one-off eigenproblem. Hence, a means has to be found of including gyroscopic effects into the nonlinear problem without adding to the number of unknowns.

Section 2 of this paper presents the theory behind the three methods presented (IRM, fast Newmark-beta method (FNBM), and RHBM). In Sec. 3 these methods are tested on a realistically-sized representative twin-spool aero-engine model.

2 Computational Analysis

2.1 System Description. The engine is assumed to have J rotors each with speed $\Omega_{(j)}$ and unbalance excitation distributed in vector (column matrix) $\mathbf{u}_{(j)}(t)$, $j=1, \dots, J$ as follows:

$$\mathbf{u}_{(j)}(t) = \mathbf{u}_{(j)\cos} \cos \Omega_{(j)}t + \mathbf{u}_{(j)\sin} \sin \Omega_{(j)}t \quad (1)$$

If the unbalance of the engine is confined to only one rotor, then it is referred to as “single frequency unbalance” (SFU). Otherwise it is multifrequency unbalance (MFU).

The governing equation of motion is

$$\ddot{\mathbf{q}} + \Lambda \mathbf{q} = \mathbf{H}_f^T \mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t) + \mathbf{H}_g^T \mathbf{g}(\dot{\boldsymbol{\theta}}) \quad (2)$$

In Eq. (2), \mathbf{q} is the $R \times 1$ vector of modal coordinates and Λ is a diagonal matrix of the squares of the natural frequencies $\omega_1, \dots, \omega_R$. The modal parameters in all the theory pertain to the linear part of the assembly under nonrotating conditions. The damping in the linear part of an engine is commonly regarded as negligible [2]. The linear part is acted on by the force and moment vectors \mathbf{f} and \mathbf{g} , respectively. \mathbf{H}_f is the matrix whose columns are the mass-normalized eigenvectors $\boldsymbol{\psi}_f^{(r)}$ ($r=1, \dots, R$) taken at the degrees-of-freedom corresponding to the directions and locations of the elements of \mathbf{f} . Similarly for \mathbf{H}_g .

\mathbf{f} comprises the unbalance forces, the static loading distribution \mathbf{w} on the rotors, and the SFD forces $\boldsymbol{\rho}$ and $-\boldsymbol{\rho}$ as follows:

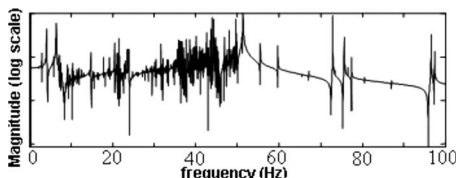


Fig. 2 Point receptance at B_1 in the y direction [1]

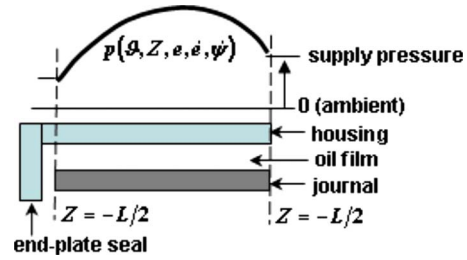


Fig. 3 Axial cross-sectional view of SFD

$$\mathbf{f}(\mathbf{x}, \dot{\mathbf{x}}, t) = [\boldsymbol{\rho}(\mathbf{v}, \dot{\mathbf{v}})^T - \boldsymbol{\rho}(\mathbf{v}, \dot{\mathbf{v}})^T \mathbf{u}_{(1)}^T(t) \cdots \mathbf{u}_{(j)}^T(t) \cdots \mathbf{u}_{(N)}^T(t) \mathbf{w}^T]^T \quad (3)$$

$\boldsymbol{\rho}(\mathbf{v}, \dot{\mathbf{v}}) = [\boldsymbol{\rho}_{(1)}^T \cdots \boldsymbol{\rho}_{(i)}^T \cdots \boldsymbol{\rho}_{(N)}^T]^T$ where $\boldsymbol{\rho}_{(i)}$ contains the Cartesian (x, y) forces on the journal J_i of SFD no. i , $i=1, \dots, N$. The vector $-\boldsymbol{\rho}(\mathbf{v}, \dot{\mathbf{v}})$ comprises the SFD forces on the bearing housings B_i of the SFDs (Fig. 1), i.e., SFD fluid inertia is neglected. The vector \mathbf{v} contains the instantaneous Cartesian (x, y) displacements of all the journal centers relative to the instantaneous positions of the bearing housing centers:

$$\mathbf{v} = \mathbf{v}_s + \mathbf{x} \quad (4)$$

where \mathbf{x} is the component of \mathbf{v} induced by \mathbf{f} and \mathbf{v}_s is a vector defining the static offsets of the SFD journals relative to their housings, in the x, y directions, under no rotor loading. The transformation from \mathbf{q} to \mathbf{x} is given by

$$\mathbf{x} = \mathbf{H}_x \mathbf{q} \quad (5)$$

\mathbf{H}_x being the matrix whose columns $\boldsymbol{\psi}_x^{(r)}$ define the values of \mathbf{x} in the respective mass-normalized modes. For known \mathbf{v} and $\dot{\mathbf{v}}$, the forces within $\boldsymbol{\rho}$ are calculated by numerical integration of the instantaneous pressure distribution across each oil film. As in Refs. [1,8], the pressure distribution p is assumed to be

$$p(\vartheta, Z, e, \dot{e}, \dot{\psi}) = p_{\text{short}}(\vartheta, Z, e, \dot{e}, \dot{\psi}) + \lambda p_{\text{long}}(\vartheta, e, \dot{e}, \dot{\psi})(0.5 - Z/L) \quad (6)$$

p_{short} and p_{long} are the short and long bearing approximations, respectively [8], ϑ and Z are the spatial coordinates local to the oil film, and e and ψ are the polar coordinates of the journal center relative to the instantaneous position of the bearing housing center. λ is an “end-leakage factor” that quantifies the degree of sealing at the outlet of the SFD, which could be fitted with an end-plate (as in Fig. 3) or other form of sealing. Prior to integration, the pressure distribution p in Eq. (6) is truncated below a minimum level (-101.325 kPa in this work) to account for cavitation [1,2].

The gyroscopic effect on the rotors is assumed to be concentrated at a number of locations. \mathbf{g} is the vector containing the gyroscopic moments on all the rotors. Hence,

$$\mathbf{g}(\dot{\boldsymbol{\theta}}) = \mathbf{P} \dot{\boldsymbol{\theta}} \quad (7)$$

where \mathbf{P} is a diagonal matrix of polar angular momenta [1] and $\boldsymbol{\theta}$ is the vector of instantaneous flexural rotations at the gyroscopic locations (where $\dot{\boldsymbol{\theta}} = \mathbf{H}_\theta \dot{\mathbf{q}}$, similar to Eq. (5)).

2.2 Time Domain Solution. The integrators described in this section calculate \mathbf{q}_k and $\dot{\mathbf{q}}_k$ for known \mathbf{q}_{k-1} and $\dot{\mathbf{q}}_{k-1}$, where $\mathbf{q}_k = \mathbf{q}(t_k)$, $\dot{\mathbf{q}}_k = \dot{\mathbf{q}}(t_k)$, and $t_k = t_{k-1} + h$.

2.2.1 IRM. In this outline of the IRM, the gyroscopic effect is excluded for ease of presentation. The addition of the gyroscopic effect to the following analysis is detailed in Ref. [1].

Using the Duhamel integral method [7] over the interval $[t_{k-1}, t_k]$,

$$\mathbf{q}_k = \dot{\mathbf{q}}_{k-1} \cdot \sin(\omega h) \cdot \omega + \mathbf{q}_{k-1} \cdot \cos(\omega h) + \int_0^h \{\mathbf{H}_f^T \mathbf{f}(\tau)\} \cdot \sin(\omega h - \omega \tau) \cdot \omega d\tau \quad (8a)$$

$$\dot{\mathbf{q}}_k = \dot{\mathbf{q}}_{k-1} \cdot \cos(\omega h) - \mathbf{q}_{k-1} \cdot \omega \cdot \sin(\omega h) \int_0^h \{\mathbf{H}_f^T \mathbf{f}(\tau)\} \cdot \cos(\omega h - \omega \tau) d\tau \quad (8b)$$

For the purpose of evaluating the above integrals, $\mathbf{f}(\tau)$ is approximated over the interval $0 \leq \tau \leq h$ as

$$\mathbf{f}(\tau) = \mathbf{f}_{k-1} + (\mathbf{f}_k - \mathbf{f}_{k-1})\tau/h \quad (9)$$

where $\mathbf{f}_k = \mathbf{f}(\mathbf{x}_k, \dot{\mathbf{x}}_k, t_k)$ and $\mathbf{x}_k = \mathbf{x}(t_k)$, $\dot{\mathbf{x}}_k = \dot{\mathbf{x}}(t_k)$. Hence, after substituting Eq. (9) into Eqs. (8a) and (8b) and evaluating the integrals and simplifying, one obtains

$$\mathbf{q}_k = \hat{\mathbf{q}}_k + (\mathbf{H}_f^T \mathbf{f}_k - \mathbf{H}_f^T \mathbf{f}_{k-1}) \cdot \left\{ \omega - \frac{1}{h} \sin(\omega h) \right\} \cdot \omega^3 \quad (10a)$$

$$\dot{\mathbf{q}}_k = \hat{\dot{\mathbf{q}}}_k + (\mathbf{H}_f^T \mathbf{f}_k - \mathbf{H}_f^T \mathbf{f}_{k-1}) \cdot \{\mathbf{n} - \cos(\omega h)\} \cdot \omega^2 \quad (10b)$$

where

$$\hat{\mathbf{q}}_k = \dot{\mathbf{q}}_{k-1} \cdot \sin(\omega h) \cdot \omega + \mathbf{q}_{k-1} \cdot \cos(\omega h) + \mathbf{H}_f^T \mathbf{f}_{k-1} \cdot \{\mathbf{n} - \cos(\omega h)\} \cdot \omega^2 \quad (11a)$$

$$\hat{\dot{\mathbf{q}}}_k = \dot{\mathbf{q}}_{k-1} \cdot \cos(\omega h) - \mathbf{q}_{k-1} \cdot \omega \cdot \sin(\omega h) + \mathbf{H}_f^T \mathbf{f}_{k-1} \cdot \sin(\omega h) \cdot \omega \quad (11b)$$

In Eqs. (10b) and (11a), \mathbf{n} is a vector of ones ("1s") of the appropriate size. Multiplying both sides of Eqs. (10a) and (10b) by \mathbf{H}_x , noting Eq. (5), and simplifying

$$\mathbf{x}_k = \hat{\mathbf{x}}_k + \mathbf{R}_{xf} \{\mathbf{f}_k(\mathbf{x}_k, \dot{\mathbf{x}}_k, t_k) - \mathbf{f}_{k-1}\} \quad (12a)$$

$$\dot{\mathbf{x}}_k = \hat{\dot{\mathbf{x}}}_k + \mathbf{S}_{xf} \{\mathbf{f}_k(\mathbf{x}_k, \dot{\mathbf{x}}_k, t_k) - \mathbf{f}_{k-1}\} \quad (12b)$$

where

$$\hat{\mathbf{x}}_k = \mathbf{H}_x \hat{\mathbf{q}}_k \quad (13a)$$

$$\hat{\dot{\mathbf{x}}}_k = \mathbf{H}_x \hat{\dot{\mathbf{q}}}_k \quad (13b)$$

and

$$\mathbf{R}_{xf}(h) = \sum_{r=1}^R \left\{ \boldsymbol{\psi}_x^{(r)} \boldsymbol{\psi}_f^{(r)T} \left(\omega_r - \frac{1}{h} \sin \omega_r h \right) / \omega_r^3 \right\} \quad (14a)$$

$$\mathbf{S}_{xf}(h) = \sum_{r=1}^R \left\{ \boldsymbol{\psi}_x^{(r)} \boldsymbol{\psi}_f^{(r)T} (1 - \cos \omega_r h) / (h \omega_r^2) \right\} \quad (14b)$$

Equations (12a) and (12b) are a set of nonlinear algebraic equations with \mathbf{x}_k and $\dot{\mathbf{x}}_k$ as unknowns. Hence, one can use an iterative method (like the Newton–Raphson method [9] used here) to solve for \mathbf{x}_k and $\dot{\mathbf{x}}_k$. The number of equations to be solved at each time-step is only four times the number of SFDs. Once \mathbf{x}_k and $\dot{\mathbf{x}}_k$ are found, \mathbf{f}_k is determined, hence from Eqs. (10a) and (10b) one can determine \mathbf{q}_k and $\dot{\mathbf{q}}_k$, and progress the solution.

2.2.2 FNBm. The gyroscopic effect is again excluded merely for ease of presentation. Applying the Newmark-beta method to the modal coordinate space, it is assumed that [7]

$$\mathbf{q}_k = \mathbf{q}_{k-1} + h \dot{\mathbf{q}}_{k-1} + h^2 \{(0.5 - \beta) \ddot{\mathbf{q}}_{k-1} + \beta \ddot{\mathbf{q}}_k\} \quad (15a)$$

$$\dot{\mathbf{q}}_k = \dot{\mathbf{q}}_{k-1} + h \{(1 - \gamma) \ddot{\mathbf{q}}_{k-1} + \gamma \ddot{\mathbf{q}}_k\} \quad (15b)$$

where β and γ are constants that are arbitrarily chosen with a view to accuracy and stability of the solution. For $\beta=0.25$ and $\gamma=0.50$, the method becomes the constant average acceleration method [7]. Rearranging Eq. (15a),

$$\ddot{\mathbf{q}}_k = h^{-2} \beta^{-1} (\mathbf{q}_k - \mathbf{q}_{k-1}) - (h\beta)^{-1} \dot{\mathbf{q}}_{k-1} - (0.5\beta^{-1} - 1) \ddot{\mathbf{q}}_{k-1} \quad (16)$$

Substituting Eq. (16) into Eq. (15b) gives

$$\dot{\mathbf{q}}_k = \gamma(h\beta)^{-1} (\mathbf{q}_k - \mathbf{q}_{k-1}) + (1 - \gamma\beta^{-1}) \dot{\mathbf{q}}_{k-1} + h(1 - 0.5\gamma\beta^{-1}) \ddot{\mathbf{q}}_{k-1} \quad (17)$$

Substituting Eq. (16) into Eq. (2) and solving for \mathbf{q}_k gives

$$\mathbf{q}_k = \hat{\mathbf{q}}_k + \mathbf{D} \mathbf{H}_f^T \{\mathbf{f}_k(\mathbf{x}_k, \dot{\mathbf{x}}_k, t) - \mathbf{f}_{k-1}\} \quad (18)$$

where

$$\mathbf{D}(h) = \{h^{-2} \beta^{-1} \mathbf{I} + \boldsymbol{\Lambda}\}^{-1} \quad (19)$$

$$\hat{\mathbf{q}}_k = \mathbf{D} \{h^{-2} \beta^{-1} \mathbf{q}_{k-1} + (h\beta)^{-1} \dot{\mathbf{q}}_{k-1} + (0.5\beta^{-1} - 1) \ddot{\mathbf{q}}_{k-1} + \mathbf{H}_f^T \mathbf{f}_{k-1}\} \quad (20)$$

Substituting for \mathbf{q}_k in Eq. (18) into Eq. (17) gives

$$\dot{\mathbf{q}}_k = \hat{\dot{\mathbf{q}}}_k + \gamma(h\beta)^{-1} \mathbf{D} \mathbf{H}_f^T \{\mathbf{f}_k(\mathbf{x}_k, \dot{\mathbf{x}}_k, t) - \mathbf{f}_{k-1}\} \quad (21)$$

where

$$\hat{\dot{\mathbf{q}}}_k = \gamma(h\beta)^{-1} (\hat{\mathbf{q}}_k - \mathbf{q}_{k-1}) + (1 - \gamma\beta^{-1}) \dot{\mathbf{q}}_{k-1} + h(1 - 0.5\gamma\beta^{-1}) \ddot{\mathbf{q}}_{k-1} \quad (22)$$

It is noted that Eqs. (15a), (15b), and (16)–(22) are basically the same as those in Ref. [7] except that they have been adapted to the modal coordinate space.

In Eqs. (18) and (21), it is noted that $\mathbf{x}_k = \mathbf{H}_x \mathbf{q}_k$ and $\dot{\mathbf{x}}_k = \mathbf{H}_x \dot{\mathbf{q}}_k$. Hence, Eqs. (18) and (21) comprise a set of $2R$ nonlinear algebraic equations in an equal number of unknowns contained in state variables \mathbf{q}_k and $\dot{\mathbf{q}}_k$. In the conventional implementation of the Newmark-beta method to nonlinear systems (e.g., Ref. [5]) it is state-variable equations such as these that are solved at each time-step, making the process unwieldy for a very large number of modes. This problem is overcome in this paper through the following slight modification to the formulation. Multiplying both sides of Eqs. (18) and (21) by \mathbf{H}_x yields

$$\mathbf{x}_k = \hat{\mathbf{x}}_k + \mathbf{R}'_{xf} \{\mathbf{f}_k(\mathbf{x}_k, \dot{\mathbf{x}}_k, t) - \mathbf{f}_{k-1}\} \quad (23a)$$

$$\dot{\mathbf{x}}_k = \hat{\dot{\mathbf{x}}}_k + \mathbf{S}'_{xf} \{\mathbf{f}_k(\mathbf{x}_k, \dot{\mathbf{x}}_k, t) - \mathbf{f}_{k-1}\} \quad (23b)$$

where

$$\hat{\mathbf{x}}_k = \mathbf{H}_x \hat{\mathbf{q}}_k \quad (24a)$$

$$\hat{\dot{\mathbf{x}}}_k = \mathbf{H}_x \hat{\dot{\mathbf{q}}}_k \quad (24b)$$

$$\mathbf{R}'_{xf}(h) = \mathbf{H}_x \mathbf{D} \mathbf{H}_f^T \quad (25a)$$

$$\mathbf{S}'_{xf}(h) = \gamma(h\beta)^{-1} \mathbf{R}'_{xf}(h) \quad (25b)$$

The analogy with the IRM equations (12a) and (12b) is evident. Equations (23a) and (23b) comprise only $4N$ equations in an equal number of unknowns in \mathbf{x}_k and $\dot{\mathbf{x}}_k$. It is these equations that are solved at each time step. The definitive value of \mathbf{f}_k is then determined, enabling the updating of \mathbf{q}_k and $\dot{\mathbf{q}}_k$ through Eqs. (18) and (21).

The gyroscopic effect can be added to the above FNBm analysis in a similar manner to the IRM [1]. It is noted that all simulations presented in Sec. 3 include the gyroscopic effect.

2.3 Frequency Domain Solution: RHBm. In the RHBm, the vibration is assumed to be periodic with a fundamental frequency of ω , where

$$\varpi = \Omega_{(j_{\text{ref}})}/Q \quad (26)$$

Q is a positive integer and $\Omega_{(j_{\text{ref}})}$ is termed the “reference unbalanced shaft speed.” In the case of SFU, $\Omega_{(j_{\text{ref}})}$ is the speed of the (only) unbalanced shaft and Q is commonly taken as 1. In the case of MFU, $\Omega_{(j_{\text{ref}})}$ is the speed of the slowest unbalanced shaft and the value of Q will be chosen according to the ratio of the speeds of the unbalanced shafts. Both the system response and the motion-dependent forces can be represented by a Fourier series, e.g.,

$$\mathbf{v} = \bar{\mathbf{v}} + \sum_{k=1}^K \{ \mathbf{v}_{\text{cos}}^{(k)} \cos k\varpi t + \mathbf{v}_{\text{sin}}^{(k)} \sin k\varpi t \} \quad (27)$$

For each nonzero harmonic of \mathbf{v} one can write the following force-response relationship:

$$\begin{aligned} \begin{bmatrix} \mathbf{v}_{\text{cos}}^{(k)} \\ \mathbf{v}_{\text{sin}}^{(k)} \end{bmatrix} &= \begin{bmatrix} \mathbf{C}_{\text{xx}}(k\varpi) & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\text{xx}}(k\varpi) \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho}_{\text{cos}}^{(k)} \\ \boldsymbol{\rho}_{\text{sin}}^{(k)} \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{\text{xg}}(k\varpi) & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\text{xg}}(k\varpi) \end{bmatrix} \\ &\times \begin{bmatrix} \mathbf{g}_{\text{cos}}^{(k)} \\ \mathbf{g}_{\text{sin}}^{(k)} \end{bmatrix} + \sum_{j=1}^J \delta_{k\varpi, \Omega_{(j)}} \begin{bmatrix} \mathbf{C}_{\text{xu}_{(j)}}(k\varpi) & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\text{xu}_{(j)}}(k\varpi) \end{bmatrix} \\ &\times \begin{bmatrix} \mathbf{u}_{(j)\text{cos}} \\ \mathbf{u}_{(j)\text{sin}} \end{bmatrix}, \quad k = 1, \dots, K \end{aligned} \quad (28)$$

where the matrices \mathbf{C}_{xx} , etc., are receptance (or “compliance”) matrices relating the relative x and y displacements at the SFD terminals to the various excitations. From modal theory [7],

$$\mathbf{C}_{\text{xx}}(\omega) = \sum_{r=1}^R \frac{\boldsymbol{\psi}_x^{(r)} \boldsymbol{\psi}_x^{(r)T}}{\omega_r^2 - \omega^2} \quad (29a)$$

$$\mathbf{C}_{\text{xg}}(\omega) = \sum_{r=1}^R \frac{\boldsymbol{\psi}_x^{(r)} \boldsymbol{\psi}_g^{(r)T}}{\omega_r^2 - \omega^2} \quad (29b)$$

$$\mathbf{C}_{\text{xu}_{(j)}}(\omega) = \sum_{r=1}^R \frac{\boldsymbol{\psi}_x^{(r)} \boldsymbol{\psi}_{\mathbf{u}_{(j)}}^{(r)T}}{\omega_r^2 - \omega^2} \quad (29c)$$

The gyroscopic terms $\mathbf{g}_{\text{cos}}^{(k)}$ and $\mathbf{g}_{\text{sin}}^{(k)}$ can be eliminated from the right hand side of Eq. (28) by considering the force-displacement relationship for each harmonic $k\varpi$ of the flexural angular displacement $\boldsymbol{\theta}$ at the gyroscopic locations,

$$\begin{aligned} \begin{bmatrix} \boldsymbol{\theta}_{\text{cos}}^{(k)} \\ \boldsymbol{\theta}_{\text{sin}}^{(k)} \end{bmatrix} &= \begin{bmatrix} \mathbf{C}_{\boldsymbol{\theta}\mathbf{x}}(k\varpi) & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\boldsymbol{\theta}\mathbf{x}}(k\varpi) \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho}_{\text{cos}}^{(k)} \\ \boldsymbol{\rho}_{\text{sin}}^{(k)} \end{bmatrix} + \begin{bmatrix} \mathbf{C}_{\boldsymbol{\theta}\mathbf{g}}(k\varpi) & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\boldsymbol{\theta}\mathbf{g}}(k\varpi) \end{bmatrix} \\ &\times \begin{bmatrix} \mathbf{g}_{\text{cos}}^{(k)} \\ \mathbf{g}_{\text{sin}}^{(k)} \end{bmatrix} + \sum_{j=1}^J \delta_{k\varpi, \Omega_{(j)}} \begin{bmatrix} \mathbf{C}_{\boldsymbol{\theta}\mathbf{u}_{(j)}}(k\varpi) & \mathbf{0} \\ \mathbf{0} & \mathbf{C}_{\boldsymbol{\theta}\mathbf{u}_{(j)}}(k\varpi) \end{bmatrix} \\ &\times \begin{bmatrix} \mathbf{u}_{(j)\text{cos}} \\ \mathbf{u}_{(j)\text{sin}} \end{bmatrix}, \quad k = 1, \dots, K \end{aligned} \quad (30)$$

where the matrices $\mathbf{C}_{\boldsymbol{\theta}\mathbf{x}}$, etc., are calculated in a similar manner to the matrices in Eq. (29). From Eq. (7),

$$\mathbf{g}_{\text{cos}}^{(k)} = k\varpi \mathbf{P} \boldsymbol{\theta}_{\text{sin}}^{(k)} \quad (31a)$$

$$\mathbf{g}_{\text{sin}}^{(k)} = -k\varpi \mathbf{P} \boldsymbol{\theta}_{\text{cos}}^{(k)} \quad (31b)$$

Hence, by substituting for $\boldsymbol{\theta}_{\text{cos}}^{(k)}$ and $\boldsymbol{\theta}_{\text{sin}}^{(k)}$ in Eq. (31) into Eq. (30), one can solve the resulting equation for $\mathbf{g}_{\text{cos}}^{(k)}$ and $\mathbf{g}_{\text{sin}}^{(k)}$ in terms of $\boldsymbol{\rho}_{\text{cos}}^{(k)}$ and $\boldsymbol{\rho}_{\text{sin}}^{(k)}$ and $\mathbf{u}_{(j)\text{cos}}$ and $\mathbf{u}_{(j)\text{sin}}$. Hence, substituting the resulting expressions for $\mathbf{g}_{\text{cos}}^{(k)}$ and $\mathbf{g}_{\text{sin}}^{(k)}$ into Eq. (28), one obtains equations of the form

$$\begin{aligned} \begin{bmatrix} \mathbf{v}_{\text{cos}}^{(k)} \\ \mathbf{v}_{\text{sin}}^{(k)} \end{bmatrix} &= \begin{bmatrix} \mathbf{F}(k\varpi) & \mathbf{G}(k\varpi) \\ -\mathbf{G}(k\varpi) & \mathbf{F}(k\varpi) \end{bmatrix} \begin{bmatrix} \boldsymbol{\rho}_{\text{cos}}^{(k)} \\ \boldsymbol{\rho}_{\text{sin}}^{(k)} \end{bmatrix} + \sum_{j=1}^J \delta_{k\varpi, \Omega_{(j)}} \\ &\times \begin{bmatrix} \mathbf{T}_{(j)}(k\varpi) & \mathbf{Q}_{(j)}(k\varpi) \\ -\mathbf{Q}_{(j)}(k\varpi) & \mathbf{T}_{(j)}(k\varpi) \end{bmatrix} \begin{bmatrix} \mathbf{u}_{(j)\text{cos}} \\ \mathbf{u}_{(j)\text{sin}} \end{bmatrix}, \quad k = 1, \dots, K \end{aligned} \quad (32)$$

where the matrices \mathbf{F} , \mathbf{G} , $\mathbf{T}_{(j)}$, and $\mathbf{Q}_{(j)}$ are obtained through the above-described manipulation. Equation (32) constitutes the *dynamic* block of the RHBM equations.

The existence of the mean or zero-frequency harmonic in the response (i.e., $\bar{\mathbf{v}}$ in Eq. (27)) necessitates an additional block of equations. The determination of this block requires a different analysis to that used for the dynamic block due to the presence of rigid body modes in the linear part of the system, which introduce eigenfrequencies $\omega_r=0$, resulting in indeterminate receptances at zero-frequency, i.e., $\mathbf{C}_{\text{xx}}(0) \rightarrow \pm \infty$ in Eq. (29a). In the case of the twin-spool engine of Fig. 1, there are four rigid body modes—two for each rotor, defining rigid body pivoting motion of a rotor about its only sprung journal (J_1 for low-pressure (LP) rotor and J_3 for high-pressure (HP) rotor) in the planes xz and yz , respectively. The mean components $\bar{\boldsymbol{\rho}}$ of the SFD forces (which are a consequence of the vibration) are in equilibrium with the static loading distribution \mathbf{w} . However, in the case of the LP rotor it is clear that the equilibrium problem is statically indeterminate. This problem can be overcome as follows. Let the system have P rigid body modes, and let the mean component of the modal coordinate vector be partitioned thus:

$$\bar{\mathbf{q}} = [\bar{\mathbf{q}}^T \quad \bar{\tilde{\mathbf{q}}}^T]^T \quad (33)$$

where $\bar{\mathbf{q}}^T$ and $\bar{\tilde{\mathbf{q}}}^T$, respectively, contain the rigid and flexible mean modal coordinates. Combining Eqs. (4) and (5), one can write, for the zeroth harmonic

$$\bar{\mathbf{v}} = \mathbf{v}_s + \check{\mathbf{H}}_x \bar{\mathbf{q}} + \tilde{\mathbf{H}}_x \bar{\tilde{\mathbf{q}}} \quad (34)$$

where the modal matrix \mathbf{H}_x has been partitioned into rigid and flexible modal matrices $\check{\mathbf{H}}_x$ and $\tilde{\mathbf{H}}_x$. Equation (34) can be expressed as follows:

$$\bar{\mathbf{v}} = \mathbf{v}_s + \check{\mathbf{H}}_x \bar{\mathbf{q}} + \tilde{\mathbf{C}}_{\text{xx}}(0) \bar{\boldsymbol{\rho}} + \tilde{\mathbf{C}}_{\text{xw}}(0) \mathbf{w} \quad (35a)$$

where the matrices $\tilde{\mathbf{C}}_{\text{xx}}(0)$ and $\tilde{\mathbf{C}}_{\text{xw}}(0)$ are computed as for Eq. (29), but with the omission of the first P rigid body modal terms in their series expressions.

In Eqs. (32) and (35a), $\bar{\boldsymbol{\rho}}$, $\boldsymbol{\rho}_{\text{cos}}^{(k)}$, and $\boldsymbol{\rho}_{\text{sin}}^{(k)}$ are each functions of $\bar{\mathbf{v}}$, $\mathbf{v}_{\text{cos}}^{(k)}$, and $\mathbf{v}_{\text{sin}}^{(k)}$. Hence, in Eqs. (32) and (35a), there are P more unknowns than there are scalar equations due to the unknowns in $\bar{\mathbf{q}}$. The required additional P scalar equations can be obtained by considering the “pseudostatic” equilibrium of the rotors under $\bar{\boldsymbol{\rho}}$ and \mathbf{w} . This principle can be applied in a systematic fashion as follows:

$$\mathbf{0} = \mathbf{A}_{\text{x}_{\text{sel}x}}(0) \bar{\boldsymbol{\rho}} + \mathbf{A}_{\text{x}_{\text{sel}w}}(0) \mathbf{w} \quad (35b)$$

where the acceleration matrices $\mathbf{A}_{\text{x}_{\text{sel}x}}(\omega) = -\omega^2 \mathbf{C}_{\text{x}_{\text{sel}x}}(\omega)$ and $\mathbf{A}_{\text{x}_{\text{sel}w}}(\omega) = -\omega^2 \mathbf{C}_{\text{x}_{\text{sel}w}}(\omega)$, respectively, relate the forces in $\bar{\boldsymbol{\rho}}$ and \mathbf{w} to the relative acceleration response at *selected* SFDs. The journal of a selected SFD should not coincide with any node of the corresponding rotor’s rigid body modes. The method of Eq. (35b) yields a maximum of P independent scalar equations. Since each selected SFD produces two equations (one for each plane xz , yz), the number of selected SFDs is then $P/2$.

Equations (32), (35a), and (35b) constitute the full set of RHBM equations totaling $2N(2K+1)+P$ scalar equations in an equal number of unknowns contained in the vector

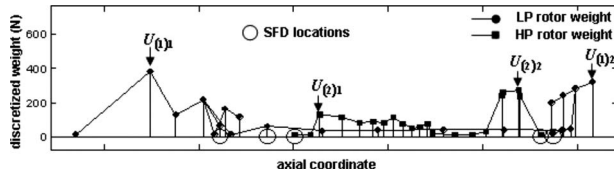


Fig. 4 Rotor weight distribution, SFD bearing locations, and unbalance locations for twin-spool engine

$$\mathbf{z} = [\bar{\mathbf{q}}^T \quad \bar{\mathbf{v}}^T \quad \mathbf{v}_{\cos}^{(1)T} \quad \mathbf{v}_{\sin}^{(1)T} \quad \dots \quad \mathbf{v}_{\cos}^{(K)T} \quad \mathbf{v}_{\sin}^{(K)T}]^T \quad (36)$$

For given $\Omega_{(j_{ref})}$ and Q (Eq. (26)), the RHBM can be solved using an iterative scheme like Newton–Raphson [9]. For any iterate of \mathbf{z} , the time history of \mathbf{v} is determined (Eq. (27)). This allows the calculation of the terms $\bar{\rho}$, $\rho_{\cos}^{(k)}$, and $\rho_{\sin}^{(k)}$ through a Fourier analysis of the time history of $\rho(\mathbf{v}, \dot{\mathbf{v}})$ evaluated at a suitable number of points over one period using the physical model of the SFD. A predictor-corrector continuation scheme is used to advance the solution process over a range of speeds $\Omega_{(j_{ref})}$ for given Q , tracing out a “speed-response” curve [2]. The initial approximation (“predictor”) to the Newton–Raphson solver at an attempted solution point on the speed-response curve is estimated from the previous solution points. The initial approximation to the *first* solution point (only) is provided by the Fourier coefficients of a time-domain solution.

Once the response at the SFDs is determined, the response at any arbitrary degree-of-freedom can be readily calculated (without further solution) by applying the appropriate receptance relationships [2].

3 Simulations on a Whole-Engine Model

The above-developed methods were applied to a representative twin-spool aero-engine having the schematic layout in Fig. 1, using a realistically-sized whole-engine FE model provided by a leading engine manufacturer [10].

Prior to the nonlinear computation, a one-off eigenvalue analysis was performed on the linear part of the whole-engine model at zero rotational speed. A typical receptance function computed from the results of this eigenproblem (as in Eq. (29)) is illustrated in Fig. 2. All 934 modes over the range 0–1 kHz were included in the subsequent nonlinear analysis due to high shaft speeds and harmonics in the response.

The three methods were applied for two cases A and B, respectively, involving SFU and MFU. For each rotor the unbalance was concentrated at two locations. In case A the unbalance was restricted to the LP rotor only. Figure 4 indicates the positions of these locations, as well as the locations of the SFDs and the distribution of the weights of the two rotors.

The gyroscopic effect was discretized at 7 points on the LP rotor and 12 points on the HP rotor. The SFDs considered for this study were single-land and end-fed with oil of viscosity of 0.0049 N s m⁻² at a pressure of 3 bars (gauge). The bearing diameters and radial clearances were typically 200 mm, 0.1 mm respectively and the land lengths ranged from 16 mm to 34 mm. The end-leakage factor $\lambda=0.03$, as typically used in industry. The bearing housings were assumed to be perfectly aligned with each other prior to rotor assembly i.e., $\mathbf{v}_s=0$ in Eq. (4).

For each cases A and B, the nonlinear analysis was performed for a fixed speed ratio $\Omega_{(2)}/\Omega_{(1)}=1.2$. All simulations were performed in MATLAB on a standard desktop PC with Intel® Pentium® D CPU 3 GHz processor.

3.1 Case A: SFU. In this case, $U_{(1)1}$ and $U_{(1)2}=6.3$ kg mm in phase and $U_{(2)1}$ and $U_{(2)2}=0$. Figures 5(a)–5(d) compare the results obtained by the three methods for the steady-state SFD orbits for LP and HP rotor speeds of 10,000 rpm and 12,000 rpm, respectively. The time-domain results required up to 4000 LP shaft

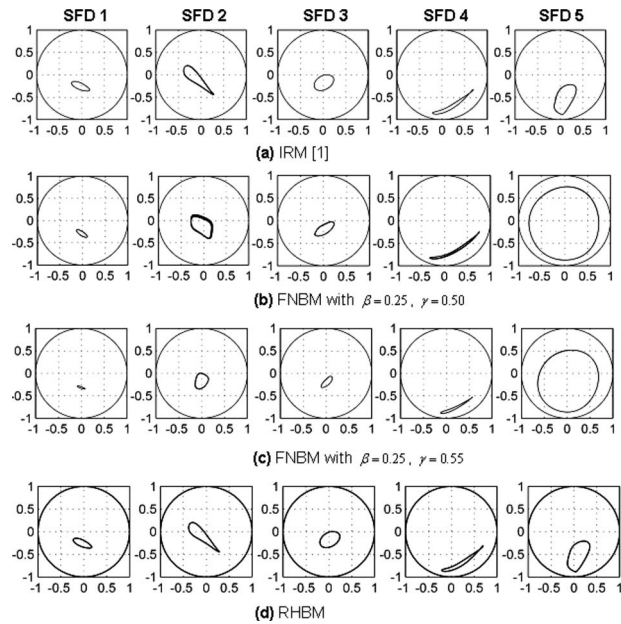


Fig. 5 Computed periodic steady-state orbits of bearing journals relative to their housings for SFU (orbits normalized with respect to the respective radial clearances). (a) IRM [1]; (b) FNBM with $\beta=0.25$, $\gamma=0.50$; (c) FNBM with $\beta=0.25$, $\gamma=0.55$; and (d) RHBM.

revolutions to reach steady-state from default initial conditions (corresponding to zero relative displacements and velocities at each SFD). Figures 5(a) and 5(b) show that there was significant disagreement between the IRM and FNBM orbits, particularly at the unsupported SFD No. 5. The FNBM result was also found to be sensitive to slight variations in β and γ , without improving the correlation with the IRM (Fig. 5(c)). It is noted that, for both IRM and FNBM, a variable step size h was used to efficiently maintain the accuracy within a preset tolerance, and that h was capped at a maximum value of 1/250 of the period of rotation of the unbalanced shaft. The Fourier coefficients of the FNBM result were used to provide the initial approximation for the RHBM iterative process, which subsequently converged to a solution (Fig. 5(d)) that was in virtual perfect agreement with the IRM. It is noted that, for RHBM, eight harmonics of the fundamental frequency $\Omega_{(1)}$ were used, giving $5 \times 2(2 \times 8 + 1) + 4 = 174$ unknowns.

The RHBM convergence to the IRM result proved conclusively that the FNBM results were erroneous. Similarly erroneous FNBM results were observed at other speeds. The use of the FNBM was therefore discontinued for the remaining simulations presented here.

Having generated this first RHBM solution point, a speed-response curve was constructed using a rudimentary continuation technique with $\Omega_{(1)}$ as the control parameter. The results are shown in Fig. 6 for SFD Nos. 4 and 5. The steady-state IRM results at discrete speeds are overlaid on the same axes and show excellent correlation with the RHBM.

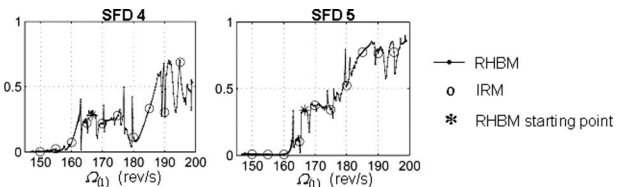


Fig. 6 SFU speed-response curves of y relative displacements at SFDs 4 and 5 (vertical axes give half peak-peak amplitude normalized by radial clearance)

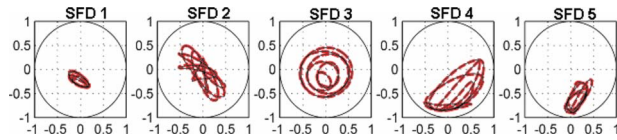


Fig. 7 Computed periodic steady-state orbits of bearing journals relative to their housings for MFU (RHBM (—), IRM (---)); orbits normalized with respect to the respective radial clearances)

3.2 Case B: MFU. In this case, $U_{(1)1}$ and $U_{(1)2}=6.3$ kg mm in phase and $U_{(2)1}$ and $U_{(2)2}=5$ kg mm in phase. The angle between $U_{(1)1}$ and $U_{(2)1}$ at $t=0$ was taken to be zero. Figure 7 shows the steady-state SFD orbits computed by IRM and RHBM for LP and HP rotor speeds of 10,000 rpm and 12,000 rpm, respectively. The IRM result was achieved after a total of 4000 LP shaft revolutions from default initial conditions. The initial approximation for the RHBM solution was provided by the Fourier coefficients of the transient time-domain solution over only the first 100 LP shaft revolutions. Since $\Omega_{(2)}/\Omega_{(1)}=1.2=6/5$, the fundamental frequency of the RHBM was taken as $\Omega_{(1)}/5$ (i.e., $Q=5$). 33 harmonics of this fundamental were used, giving a total of 674 unknowns. The excellent agreement between IRM and RHBM was maintained over a range of speeds, as indicated by the speed-response curves in Fig. 8.

3.3 Discussion. Prior to analyzing the whole-engine model, the three methods were tested on the simple system shown in Fig. 9, which was a simplification of the real twin-spool engine depicted in Fig. 1. A total of ten modes was used to analyze this simple system, and excellent agreement was generally achieved between the IRM and FNBM. The FNBM was also observed to be robust with respect to slight changes in the values of β and γ about their typical values of 0.25 and 0.50, respectively.

The discrepancy between IRM and FNBM in the subsequent real engine analysis leads to the conclusion that the underlying assumptions of the Newmark-beta method, expressed by Eqs. (15a) and (15b), may become problematic in the presence of a large number of modes and several nonlinear elements. The relationships in Eqs. (15a) and (15b) effectively prelock the state-variables into a relationship, and this restriction is carried through the computation. In fact, comparing the FNBM equations (23a), (23b), (24a), (24b), (25a), and (25b) with the IRM equations

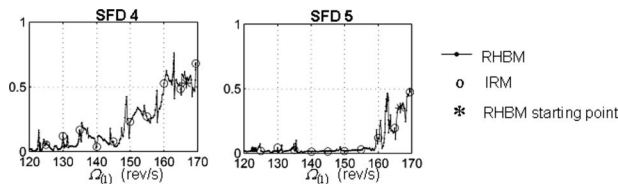


Fig. 8 MFU speed-response curves of y relative displacements at SFDs 4 and 5 (vertical axes give half peak-peak amplitude normalized by radial clearance)

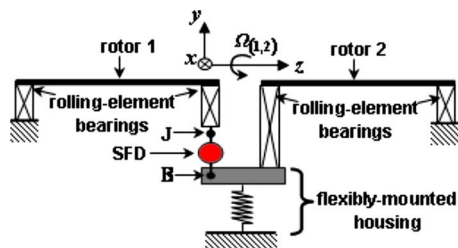


Fig. 9 Schematic of the simple twin-shaft system

(12a), (12b), (13a), (13b), (14a), and (14b), it is evident that the matrices \mathbf{R}'_{xf} and \mathbf{S}'_{xf} differ by just a scalar factor, in contrast to \mathbf{R}_{xf} and \mathbf{S}_{xf} .

Returning to the engine predictions (Figs. 5 and 7), it is noted that the offsets at the spring supported SFD orbits (Nos. 1 and 3) were due to the static loading. The prediction of the response at the remaining (unsupported) SFDs (particularly Nos. 4 and 5) posed the major computational challenge for the RHBM iterative process, particularly in the presence of a large number of harmonics. The typical solution times per point of the RHBM speed-response curves (Figs. 6 and 8) were 5 s for SFU and 50 s for MFU. In general, for MFU, if $\Omega_{(2)}/\Omega_{(1)}=H/Q$, where H and Q are integers with no common factor and $H \geq Q$, then the RHBM computation time increases with Q since a greater number of harmonics K has to be included. At certain, very narrow, regions of the speed range the RHBM computation time increased considerably from the values quoted since the rudimentary continuation process used here struggled to provide an adequate initial approximation to the iteration. However, even in the absence of a continuation scheme, the RHBM is highly useful at rapidly completing unfinished time-domain solutions through to periodic steady-state. Further computational details and results for the whole-engine RHBM will be available in a separate paper [11].

Finally, it is noted that the observed agreement between IRM and RHBM is a consequence of the existence of stable periodic vibration for the operating regime considered. As observed in Ref. [1], reducing the degree of sealing to a very low level ($\lambda=0.010$), all other parameters kept the same, resulted in a steady-state quasiperiodic response which could only be computed by a time-domain technique.

4 Conclusions

This paper presented three methods for the efficient computation of the unbalance response of a whole aero-engine model with nonlinear bearings. An overview of the recently developed rapid impulsive receptance method was given. An alternative solver was proposed through the reformulation of the Newmark-beta method for rapid implementation to large-order systems. In addition to these two time-domain methods, a whole-engine receptance harmonic balance method has been introduced, which has allowed, for the first time, the frequency domain analysis of a real whole-engine model. Simulations on a whole-engine model with squeeze-film damper bearings provided evidence that the popular Newmark-beta method can be unreliable for large-order nonlinear systems. The excellent correlation between the IRM and RHBM results demonstrated the efficacy of these two complementary tools in the computational analysis of realistic whole-engine models.

Acknowledgment

The authors acknowledge the support of the Engineering and Physical Sciences Research Council (EPSRC) of UK and the support of Rolls-Royce plc.

Nomenclature

- $()^T$ = matrix/vector transpose
- $\dot{}$ = derivative
- $()$ = $d()/dt$
- $(\tilde{})$ = component of $()$ pertaining to the flexible body modes
- $(\check{})$ = component of $()$ pertaining to the rigid body modes
- $\mathbf{a} = [a_1 \cdots a_n]^T$
- $\mathbf{b} = [b_1 \cdots b_n]^T$
- $\mathbf{a} * \mathbf{b} = [a_1 b_1 \cdots a_n b_n]^T$
- $\mathbf{a} ./ \mathbf{b} = [a_1 / b_1 \cdots a_n / b_n]^T$

$\mathbf{a}^m = [a_1^m \cdots a_n^m]^T$
 $\sin(\mathbf{a}) = [\sin a_1 \cdots \sin a_n]^T$
 $\cos(\mathbf{a}) = [\cos a_1 \cdots \cos a_n]^T$
 $\overline{(\)}, (\)_{\cos}^{(k)}, (\)_{\sin}^{(k)} =$ Fourier coefficients of vector $(\)$
 $\mathbf{A}_{\text{sef}}(\omega) =$ accelerance matrices, Eq. (35b)
 $\mathbf{C}_{\text{xx}}(\omega) =$ receptance matrices, Eq. (29)
 $\mathbf{D} =$ diagonal matrix defined in Eq. (19)
 $\mathbf{f} =$ vector of squeeze-film forces, unbalance forces, static loading, Eq. (3)
 $\mathbf{f}_k = \mathbf{f}(\mathbf{x}_k, \dot{\mathbf{x}}_k, t_k)$
 $\mathbf{F}(k\varpi) =$ matrix used in Eq. (32)
 $\mathbf{g} =$ vector of gyroscopic moments (Eq. (7))
 $\mathbf{G}(k\varpi) =$ matrix used in Eq. (32)
 $h =$ time-step size
 $\mathbf{H}_f, \mathbf{H}_x, \mathbf{H}_\theta,$ and $\mathbf{H}_g =$ modal matrices with columns $\psi_f^{(r)}, \psi_x^{(r)}$, etc.
 $i =$ counter for nonlinear bearings
 $\mathbf{I} =$ identity matrix
 $j =$ counter for rotors
 $J =$ total number of rotors
 $k =$ time-step counter (Secs. 2.2.1 and 2.2.2)
 $=$ harmonic counter (RHBM, Sec 2.3)
 $K =$ order of highest harmonic in RHBM
 $L =$ SFD land length
 $\mathbf{n} =$ vector of ones
 $N =$ total number of nonlinear bearings (SFDs)
 $p(\vartheta, Z, e, \dot{e}, \dot{\psi}) =$ SFD pressure distribution, Eq. (6)
 $P =$ total number of rigid body modes
 $\mathbf{P} =$ diagonal matrix of polar angular momenta [1]
 $\mathbf{q} =$ vector of modal coordinates (Eq. (2))
 $\mathbf{q}_k = \mathbf{q}(t_k)$
 $\dot{\mathbf{q}}_k = \dot{\mathbf{q}}_k(t_k)$
 $\hat{\mathbf{q}}_k$ and $\hat{\dot{\mathbf{q}}}_k =$ defined in Eqs. (11a) and (11b)
 $\hat{\mathbf{q}}_k$ and $\hat{\dot{\mathbf{q}}}_k =$ defined in Eqs. (20) and (22)
 $Q =$ positive integer, Eq. (26)
 $\mathbf{Q}_{(j)}(k\varpi) =$ matrix used in Eq. (32)
 $r =$ counter for modes
 $R =$ total number of modes considered
 \mathbf{R}_{xf} and $\mathbf{S}_{\text{xf}} =$ matrices defined by Eqs. (14a) and (14b)
 \mathbf{R}'_{xf} and $\mathbf{S}'_{\text{xf}} =$ matrices defined by Eqs. (25a) and (25b)
 $t =$ time
 $t_k =$ discrete time, $=t_{k-1}+h$
 $\mathbf{T}_{(j)}(k\varpi) =$ matrix used in Eq. (32)
 $\mathbf{u}_{(j)} =$ unbalance excitation on rotor no. j
 $U_{(j)1}$ and $U_{(j)2} =$ unbalance mass-radius products at locations 1 and 2 of rotor no. j (Fig. 4)
 $\mathbf{v} =$ vector of instantaneous x and y displacements of the SFD journal centers relative to their bearing housing centers
 $\mathbf{v}_s =$ vector of static x and y offsets of the SFD journals relative to their housing

$\mathbf{w} =$ vector of static loading on all rotors
 $\mathbf{x} = \mathbf{v} - \mathbf{v}_s$
 $\mathbf{x}_k = \mathbf{x}(t_k)$
 $\dot{\mathbf{x}}_k = \dot{\mathbf{x}}(t_k)$
 $\hat{\mathbf{x}}_k$ and $\hat{\dot{\mathbf{x}}}_k =$ defined by Eqs. (13a) and (13b)
 $\hat{\mathbf{x}}_k$ and $\hat{\dot{\mathbf{x}}}_k =$ defined by Eqs. (24a) and (24b)
 $\mathbf{z} =$ vector of RHBM unknowns (Eq. (36))
 $x, y,$ and $z =$ Cartesian frame, Fig. 1
 $\Lambda =$ diagonal matrix of $\omega_1^2, \dots, \omega_R^2$
 β and $\gamma =$ parameters of Newmark-beta method
 $\lambda =$ end-leakage factor of SFD (Eq. (6))
 $\psi_x^{(r)} =$ value of \mathbf{x} in r th mass-normalized mode
 $\psi_f^{(r)}$ and $\psi_g^{(r)} =$ mass-normalized eigenvectors at degrees of freedom corresponding to directions and locations of elements in \mathbf{f} and \mathbf{g}
 $\theta =$ vector of flexural rotations (Eq. (7))
 $\omega_r =$ circular frequency of r th undamped mode
 $\boldsymbol{\omega} = [\omega_1 \cdots \omega_r \cdots \omega_R]^T$ (Sec. 2.2.1)
 $\omega =$ generic circular frequency (Sec. 2.3)
 $\varpi =$ fundamental circular frequency of RHBM
 $\delta_{k\varpi, \Omega_{(j)}} = 1(k\varpi = \Omega_{(j)})$ or $0(k\varpi \neq \Omega_{(j)})$
 $\Omega_{(j)} =$ rotational speed of rotor no. j (rad/s)
 $\Omega_{(j_{\text{ref}})} =$ speed of reference shaft (Eq. (26))
 $\tau =$ local time over interval $[t_{k-1}, t_k]$
 $\boldsymbol{\rho} =$ vector of x and y SFD forces on all journals
 $\boldsymbol{\rho}_{(i)} =$ vector of x and y SFD forces on journal of SFD no. i

References

- [1] Hai, P. M., and Bonello, P., 2008, "An Impulsive Receptance Technique for the Time Domain Computation of the Vibration of a Whole Aero-Engine Model With Nonlinear Bearings," *J. Sound Vib.*, **318**(3), pp. 592–605.
- [2] Bonello, P., Brennan, M. J., and Holmes, R., 2002, "Non-Linear Modelling of Rotor Dynamic Systems With Squeeze Film Dampers—An Efficient Integrated Approach," *J. Sound Vib.*, **249**(4), pp. 743–773.
- [3] Lim, C. L., Kik, D. F., and Levesley, M. C., 2004, "Analysis of a Flexible Rotor System Including Non-Linear Squeeze Film Damping Using a Finite Element Method," *IMEchE Conference Transactions*, 8th International Conference on Vibrations in Rotating Machinery, Professional Engineering, Bury St. Edmunds, UK, pp. 131–140.
- [4] Chu, F., and Holmes, R., 1998, "Efficient Computation on the Nonlinear Responses of a Rotating Assembly Incorporating the Squeeze-Film Damper," *Comput. Methods Appl. Mech. Eng.*, **164**, pp. 363–373.
- [5] Armentrout, R. W., and Gunter, E. J., 1999, "Transient Modal Analysis of Nonlinear Rotor-Bearing Systems," *Proc. SPIE*, **3727**(1), pp. 290–296.
- [6] Shampine, L. F., and Reichelt, M. W., 1997, "The Matlab ODE Suite," *SIAM J. Sci. Comput. (USA)*, **18**(1), pp. 1–22.
- [7] Petyt, M., 1998, *Introduction to Finite Element Vibration Analysis*, Cambridge University Press, Cambridge, UK, pp. 392 and 430–432.
- [8] Dede, M. M., Dogan, M., and Holmes, R., 1985, "The Damping Capacity of a Sealed Squeeze Film Bearing," *ASME J. Tribol.*, **107**, pp. 411–418.
- [9] Dahlquist, G., 1974, *Numerical Methods*, Prentice-Hall, Englewood Cliffs, NJ.
- [10] Rolls-Royce plc, 2007, Data for EPSRC Project No. EP/D054575/1, confidential communication, Derby, UK.
- [11] Bonello, P., and Hai, P. M., 2009, "A Receptance Harmonic Balance Technique for the Computation of the Vibration of a Whole Aero-Engine Model With Nonlinear Bearings," *J. Sound Vib.*, **324**, pp. 221–242.

A Comparison of the Different Methods of Using Jatropha Oil as Fuel in a Compression Ignition Engine

M. Senthil Kumar¹

Département Systèmes Energétiques et
Environnement,
Ecole des Mines de Nantes,
4 rue Alfred Kastler,
BP 20722,
44307 Nantes, Cedex 03, France
e-mail: mskit@m@yahoo.com

A. Ramesh

B. Nagalingam

Department of Mechanical Engineering,
Indian Institute of Technology Madras,
Chennai 600 036, India

Different methods to improve the performance of a jatropha oil based compression ignition engine were tried and compared. A single cylinder water-cooled, direct injection diesel engine was used. Base data were generated with diesel and neat jatropha oil. Subsequently, jatropha oil was converted into its methyl ester by transesterification. Jatropha oil was also blended with methanol and orange oil in different proportions and tested. Further, the engine was modified to work in the dual fuel mode with methanol, orange oil, and hydrogen being used as the inducted fuels and the jatropha oil being used as the pilot fuel. Finally, experiments were conducted using additives containing oxygen, like dimethyl carbonate and diethyl ether. Neat jatropha oil resulted in slightly reduced thermal efficiency and higher emissions. Brake thermal efficiency was 27.3% with neat jatropha oil and 30.3% with diesel. Performance and emissions were considerably improved with the methyl ester of jatropha oil. Dual fuel operation with methanol, orange oil, and hydrogen induction and jatropha oil injection also showed higher brake thermal efficiency. Smoke was significantly reduced from 4.4 BSU with neat jatropha oil to 2.6 BSU with methanol induction. Methanol and orange oil induction reduced the NO emission and increased HC and CO emissions. With hydrogen induction, hydrocarbon and carbon monoxide emissions were significantly reduced. The heat release curve showed higher premixed rate of combustion with all the inducted fuels mainly at high power outputs. Addition of oxygenates like diethyl ether and dimethyl carbonate in different proportions to jatropha oil also improved the performance of the engine. It is concluded that dual fuel operation with jatropha oil as the main injected fuel and methanol, orange oil, and hydrogen as inducted fuels can be a good method to use jatropha oil efficiently in an engine that normally operates at high power outputs. Methyl ester of jatropha oil can lead to good performance at part loads with acceptable levels of performance at high loads also. Orange oil and methanol can be also blended with jatropha oil to improve viscosity of jatropha oil. These produce acceptable levels of performance at all outputs. Blending small quantity of diethyl ether and dimethyl carbonate with jatropha oil will enhance the performance. Diethyl ether seems to be the better of the two.

[DOI: 10.1115/1.3155400]

Keywords: diesel engine, jatropha oil, biodiesel, dual fuel engine, emission control

1 Introduction

In the context of fast depletion of petroleum reserves associated with rising vehicle population and stringent emission standards, the development of alternative energy sources has become important. Vegetable oils hold good promise as alternative fuel for diesel engines. They are biodegradable and renewable fuels. They have a reasonably high cetane number. The flash point of vegetable oils is high and hence it is safe to use them. Vegetable oils typically have large molecules with carbon, hydrogen, and oxygen being present [1]. They have a higher molecular mass and viscosity. Contrary to fossil fuels, vegetable oils are free from sulfur and heavy metals [2].

A number of vegetable oils like rapeseed oil, neem oil, palm oil, karanja oil, coconut oil, cottonseed oil, jatropha oil, etc., were tested to evaluate their performance in diesel engines [3]. Among these, jatropha oil was found as the most suitable for diesel.

As a compression ignition engine fuel, jatropha oil has a high cetane number, which is very close to diesel [4]. The flash point of jatropha oil is high 240°C as compared with 75°C for diesel. Due to its high flash point, jatropha oil has certain advantages like greater safety during storage, handling, and transport. However, this may create problems during starting. The viscosity of jatropha oil is less as compared with other vegetable oils but is higher than diesel. Some of the important properties of different fuels are compared in Table 1.

The main problem associated with the use of vegetable oils is their high viscosity and poor volatility. Different methods have been tried to use vegetable oils efficiently. Some of them are as follows:

1. transesterification with alcohols
2. blending with diesel/alcohol
3. dual fueling with gaseous and liquid fuels, and
4. use of additives, etc.

Transesterification is the process of converting the triglycerides in vegetable oils to their monoesters by reacting them with alcohols in the presence of a catalyst. Transesterification of vegetable oil results in a significant reduction in viscosity, thereby enhanc-

¹Corresponding author.

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received June 4, 2008; final manuscript received May 7, 2009; published online November 24, 2009. Review conducted by Margaret Wooldridge.

Table 1 Properties of fuels [6,13,14,16–23]

Properties	Diesel	JO	MEJO	H ₂	METH	Org. oil	DEE	DMC
Formula	C ₁₃ H ₂₄	C ₂₀ H ₂₂ O ₃	N/A	H ₂	CH ₃ OH	C ₁₀ H ₁₆	C ₂ H ₃ O C ₂ H ₃	C ₃ H ₆ O ₃
Density (kg/m ³)	840	918.6	880	0.089	790	820	713	1070
Calorific value (kJ/kg)	42,490	39,774	38,450	119,810	19,674	42,000	33,897	15,780
Viscosity (cS)	4.59	49.93	5.65	N/A	0.59	0.95	0.23	0.625
Cetane number	45–55	40–45	50	N/A	3–5	10	>125	40
Auto ign. temp. (°C)	280	340	300	585	478	N/A	160	220
Vapor pr. at 20°C (mm Hg)	10	N/A	N/A	N/A	97	4	440	53
Oxygen (% by wt)	0	15.4	11	0	50	0	21	53.3
Stoch. A/F ratio (kg/kg)	14.6	10.8	N/A	34.4	6.45:1	14.21	11.2	4.65
Flash point (°C)	75	240	170	N/A	12.2	46.1	19	17

JO, jatropa oil; MEJO, methyl ester of jatropa oil; METH, methanol; H₂, hydrogen; Org. oil, orange oil; DEE, diethyl ether; DMC, dimethyl carbonate; and N/A, not available data.

ing the physical properties. The cetane number is also improved [5]. It was reported that the methyl ester of vegetable oils offers low smoke levels and high thermal efficiencies than neat vegetable oils [6]. Methyl ester of vegetable oil also leads to improved heat release rates [7]. Power output was found to be superior to pure vegetable oils. However, the transesterification process requires approximately 3 h for making the ester and a further 12 h for separation. It also results in byproducts such as fatty acids and glycerol [8]. These cannot be used as fuels in engines though they have other uses.

Vegetable oils offer the advantage of freely mixing with alcohols, and these blends can be used in the existing diesel engines without modifications. This is a simple process. Blending of vegetable oils with methanol results in significant improvement in physical properties [9]. Viscosity and density are considerably reduced. Volatility is also improved. Vegetable oils in varying proportions in the fuel blend were tried by a number of investigators. Results obtained from the experiments on a diesel engine using a blend of vegetable oil and alcohol showed improved brake thermal efficiency and reduced exhaust smoke emissions than neat vegetable oils. However, the maximum quantity of alcohol that can be blended is limited by the presence of water in alcohol. High quantities lead to separation.

The dual fueling is a well-established technique to use different types of fuels in diesel engines. A conventional diesel engine can be easily modified to operate in this mode. This engine can accept a wide range of liquid and gaseous fuels. The dual fuel engine can result in good thermal efficiency and extremely low smoke emissions particularly at high power outputs. In a dual fuel engine, a volatile liquid or gaseous fuel with a high octane number is inducted along with air through the intake manifold. The resulting homogeneous mixture is compressed to a temperature below its self-ignition point. A pilot fuel (with a high cetane number) is injected through the standard injection system. This self-ignites and initiates the combustion in the primary fuel air mixture [10]. Alcohols were widely used as inducted fuels in the dual fuel engine. However, dual fuel operation with alcohol induction results in higher hydrocarbon and carbon monoxide emissions [11]. Use of hydrogen due to its high flame velocity can enhance the combustion rate of vegetable oils and reduce emissions. Dual fuel operation normally poses problems of low efficiency when the concentration of the inducted fuel is low. This is because the lean mixture of the inducted fuel with air does not burn well [12]. Use of fuels with wide flammability limits and high flame velocity can overcome such problems in the dual fuel engine. Since vegetable oils produce high smoke emissions, dual fuel approach can be a viable option for improving their performance. Vegetable oils can be used as the pilot fuel and alcohol can be the inducted fuel. The high flame velocity of alcohols and hydrogen can improve the overall combustion process.

Adding some oxygenated fuels to diesel to reduce emissions is attractive. It was found that the addition of dimethyl carbonate

(DMC), which is an oxygenate, to diesel resulted in a shorter diffusion combustion phase, improved thermal efficiency, and lower carbon monoxide, hydrocarbon, and smoke emissions [13]. Diethyl ether (DEE) is a renewable oxygenate, which has a very high cetane number and is also volatile [14]. Hence, it can be used as an additive to reduce the ignition delay and improve mixture formation with vegetable oils.

In this work different methods to improve the performance of a jatropa oil fueled diesel engine were studied experimentally. The optimum results obtained with different methods have been compared. The comparison has been made at peak power output. The following were tried:

- neat jatropa oil and methyl ester of jatropa oil
- dual fuel operation with jatropa oil as the injected fuel and methanol, orange oil, and hydrogen as the inducted fuels
- blends of jatropa oil with methanol and orange oil, and use of oxygenates (diethyl ether and dimethyl carbonate) as additives to jatropa oil

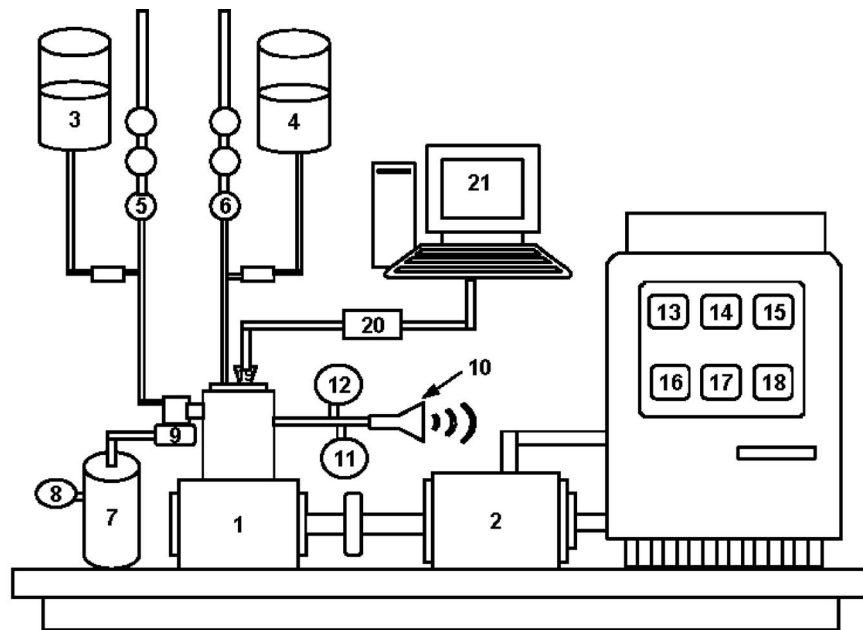
Where secondary fuels are used, the comparison is made at the optimal energy share. Comparison is made at optimum blend/oxygenate fraction when blends/oxygenates are used.

2 Experimental Setup and Experiments

A single cylinder four-stroke water-cooled diesel engine developing 3.7 kW at 1500 rpm was used for the research work. Engine details are given in Table 2. The schematic diagram of the experimental set up is shown in Fig. 1. An electrical dynamometer was used for loading the engine. A carburetor varied the amount of fuel in the dual fuel mode. A turbine type meter connected to a large tank was attached to the engine to make air flow measurements. The fuel flow rate was measured on a volumetric basis using a burette and a stopwatch. Chromel alumel thermocouples in conjunction with a digital temperature indicator were used for measuring the exhaust gas temperature. A high-speed digital data acquisition system in conjunction with a piezoelectric transducer was used for the measurement of cylinder pressure history. An

Table 2 Engine details

Name of the engine	KIRLOSCAR AV1.
General details	Four stroke, CI, water-cooled, single cylinder
Bore and stroke	80 × 120 mm ²
Compression ratio	15:1
Rated output	3.68 kW at 1500 rpm
Fuel injector opening pressure	170 bars
Injection timing	27 deg before TDC (diesel), 29 deg before TDC (jatropa oil)



- | | | |
|-----------------------|---------------------------------------|-------------------------------------|
| 1.Engine | 10. Silencer | 17. Lube oil Temperature Indicator |
| 2.Dynamometer | 11. Smoke Pump | 18. Water flow rate Indicator |
| 3.Alcohol Tank | 12. HC/CO Analyzer | 19. Pressure Sensor |
| 4.Diesel Tank | 13. Stop Watch | 20. Charge Amplifier |
| 5.Burette (Alcohol) | 14.RPM Indicator | 21. Digital Data Acquisition System |
| 6.Burette (Diesel) | 15. Exhaust gas Temperature Indicator | |
| 7. Air Tank | 16. Coolant Temperature Indicator | |
| 8. Air flow meter | | |
| 9. Carburetor | | |

Fig. 1 Experimental setup

infrared exhaust analyzer was used for the measurement of HC/CO in the exhaust. For measuring NO_x , a Rosemount Analytical Model 951 A chemiluminescent NO/NO_x analyzer was utilized. Smoke levels were obtained using a Bosch smoke meter.

Experiments were initially carried out on the engine with diesel and jatropha oil for providing baseline data. Jatropha oil was converted into its methyl ester by the transesterification process. Performance, fuel economy, combustion, and emission characteristics were obtained. Subsequently jatropha oil was blended with diesel, alcohol, and orange oil in different proportions and tested. In the next phase, the engine was modified to work in the dual fuel mode with methanol, orange oil, and hydrogen being used as the inducted fuels while jatropha oil and diesel were used as pilot fuels in turn. In the last phase of the work, experiments were conducted by mixing oxygenates like dimethyl carbonate and diethyl ether to jatropha oil. In all cases comparison was made at peak power output. Where secondary fuels are used, the comparison was made at the optimal energy share.

3 Results and Discussion

3.1 Performance Parameters. Neat jatropha oil as indicated in Fig. 2 results in reduced brake thermal efficiency due to poor mixture formation. However, there is an improvement in brake thermal efficiency with all methods as compared with neat jatropha oil operation. Highest brake thermal efficiency is obtained in the dual fuel operation as compared with other methods. Brake thermal efficiencies as indicated in Table 3 are 30.7% with methanol, 29.4% with orange oil, and 29.3% with hydrogen in the dual fuel mode. The best energy shares are 46% with methanol, 31% with orange oil, and 18% with hydrogen. The increase in brake

thermal efficiency in dual fuel operation is due to high combustion rate with all inducted fuels. As the inducted charge gets ignited the amount of heat release occurring by flame propagation is increased and that leads to better thermal efficiency [12,15]. Methanol induction results in the highest brake thermal efficiency. It can be noted that with methanol induction the brake thermal efficiency is even better than neat diesel operation at the optimum energy share. Methyl ester of jatropha oil also shows a higher thermal efficiency as compared with neat jatropha oil and blends with methanol and orange oil. The methyl ester of jatropha oil has lower viscosity, which results in better atomization of the fuel as compared with neat jatropha oil [16,17]. Methanol blend shows higher thermal efficiency (28.5%) than the orange oil blend (28.3%) due to superior physical characteristics of methanol than orange oil such as low viscosity (0.6 cp) and high vapor pressure (126 mm Hg at 30°C). DEE shows higher thermal efficiency than the DMC. Since the volatility of DEE is higher than DMC, the mixture formation is better and this leads to higher brake thermal efficiency.

The exhaust gas temperature shown in Fig. 3 is very high with neat jatropha oil due to slow combustion as a result of poor volatility and high viscosity [18]. With all methods there is a reduction in exhaust gas temperature. Dual fuel operation with methanol induction shows the highest reduction in exhaust gas temperature (i.e., 404°C at optimum energy share of 46%) whereas it is 428°C with neat jatropha oil. This reduction in exhaust gas temperature is due to a reduction in charge temperature as a result of vaporization of methanol. Similar trend is seen in the case of orange oil also. With hydrogen induction exhaust gas temperature is higher than neat jatropha oil due to rapid combustion that leads

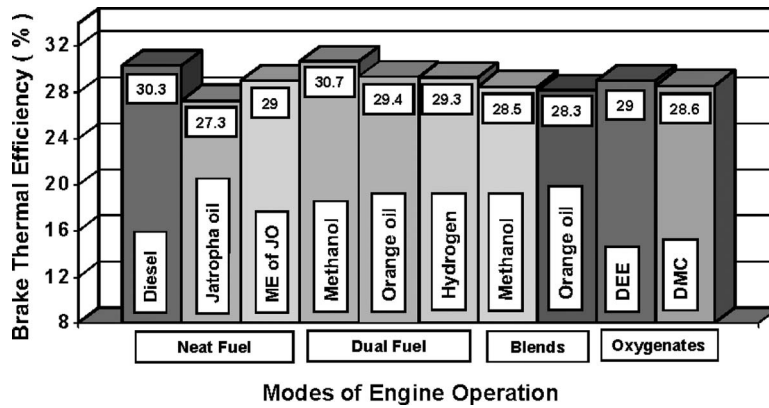


Fig. 2 Variation in brake thermal efficiency with different modes of engine operation

to high charge temperature. Blends of methanol and orange with jatropha oil also reduce the exhaust gas temperature due to fast combustion rate of methanol and orange oil in the blend. In addition to that, the high latent heat of vaporization of methanol reduces the charge temperature. Due to the fast diffusion combustion rate, exhaust temperature is reduced with both oxygenates.

3.2 Emission Parameters. The smoke level is indicated in Fig. 4. Neat jatropha oil due to poor atomization as a result of heavier molecular structure increases smoke emission. However, smoke reduces with all methods. Greatest reduction in smoke emission is seen in the dual fuel operation than the other methods. The homogeneous combustion of methanol and orange oil air mixture in the case of the dual fuel mode lowers the smoke level.

Since hydrogen has no hydrocarbon smoke emission is reduced significantly with the hydrogen dual fuel mode. It is noted from Table 3 that the smoke levels in dual fuel operation with jatropha oil are much lower than neat diesel (3.8 BSU). The lowest smoke level (2.6 BSU) is obtained with methanol jatropha oil dual fuel operation. Blends of methanol and orange oil also show lower smoke levels than neat jatropha oil. Methanol blend shows more reduction than orange oil due to increased premixed combustion rate as a result of increased ignition delay. DEE and DMC show a greater reduction than the other blends and esters due to their high oxygen content. In general, dual fuel operation is much better than the others as regards smoke reduction.

The variation in hydrocarbon and carbon monoxide emissions

Table 3 Results of different methods

Methods	Neat fuels			Dual fuel operation			Blends		Oxygenates	
	Diesel	JO	MEJO	DF1	DF2	DF3	B1	B2	DEE	DMC
	Brake thermal efficiency (%)									
	30.3	27.3	29	30.7	29.4	29.3	28.5	28.3	29	28.6
	Exhaust gas temperature (°C)									
	402	428	415	404	416	430	410	417	406	408
	Smoke number (Bosch smoke units)									
	3.8	4.4	4	2.6	3.2	3.3	4.1	4.2	3.9	4.1
	Hydrocarbon (ppm)									
	100	130	110	190	160	100	110	120	110	110
	Carbon monoxide (%)									
	0.20	0.26	0.22	0.32	0.30	0.17	0.24	0.24	0.22	0.22
	Nitric oxide (ppm)									
	795	740	765	702	730	870	697	765	713	765
	Ignition delay (deg CA)									
	9	11	10	16	14	12	12	11	9	11
	Peak pressure (bar)									
	63.5	61.9	62.6	64.3	65.5	64.3	62.7	62.3	60.9	62.4
	Maximum rate of pressure rise (MRPR) (bar)									
	4.1	3.8	4	4.8	4.9	4.8	3.9	3.9	3.7	4
	Combustion duration (deg CA)									
	42	46	44	57	46	44	45	45	42	44

JO, jatropha oil; MEJO, methyl ester of jatropha oil; DF1, methanol; DF2, orange oil; DF3, hydrogen; B1, methanol; B2, orange oil; DEE, diethyl ether; and DMC, dimethyl carbonate.

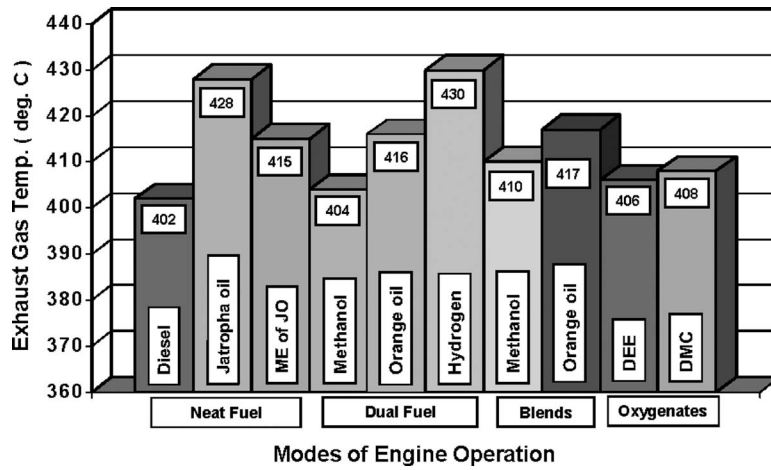


Fig. 3 Variation in exhaust gas temperature with different modes of engine operation

at peak power output is shown in Figs. 5 and 6. It is seen that in single fuel operation with methyl ester of jatropha oil, blends with methanol and orange oil and oxygenates, lower hydrocarbon and carbon monoxide levels are observed than neat jatropha oil. However, hydrocarbon and carbon monoxide emissions increase in

dual fuel operation with methanol and orange oil as inducted fuels. In the case of dual fuel operation with methanol and orange oil, the quenching of the flame as it propagates through the homogeneous inducted fuel air mixture leads to relatively high hydrocarbon levels [19]. The inducted fuel also lowers the net air sup-

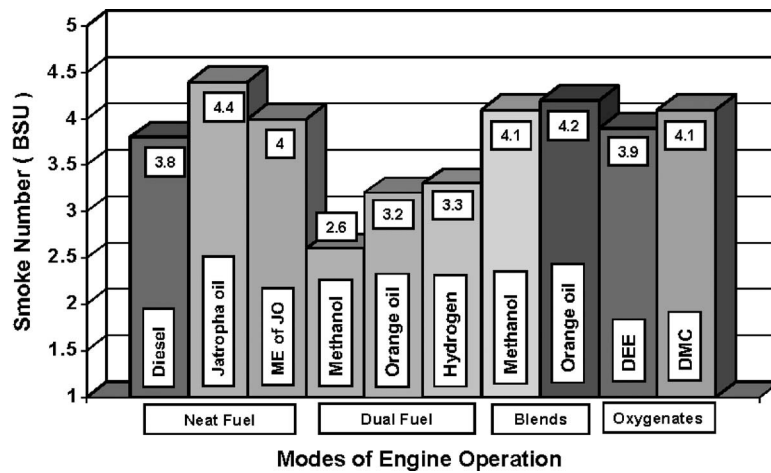


Fig. 4 Variation in smoke number with different modes of engine operation

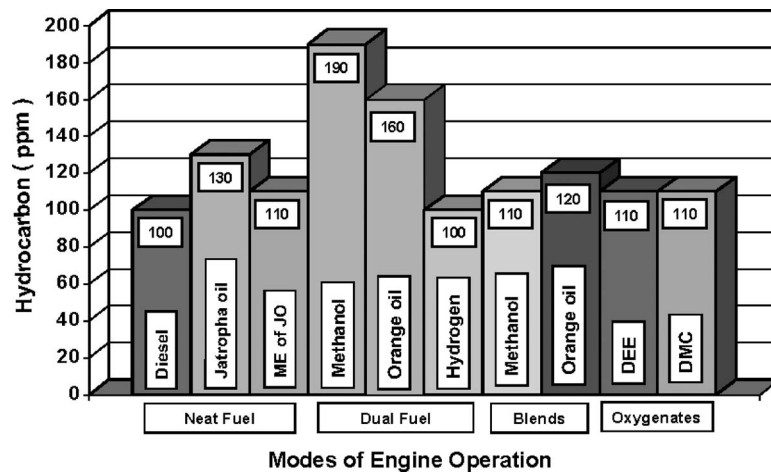


Fig. 5 Variation in hydrocarbon level with different modes of engine operation

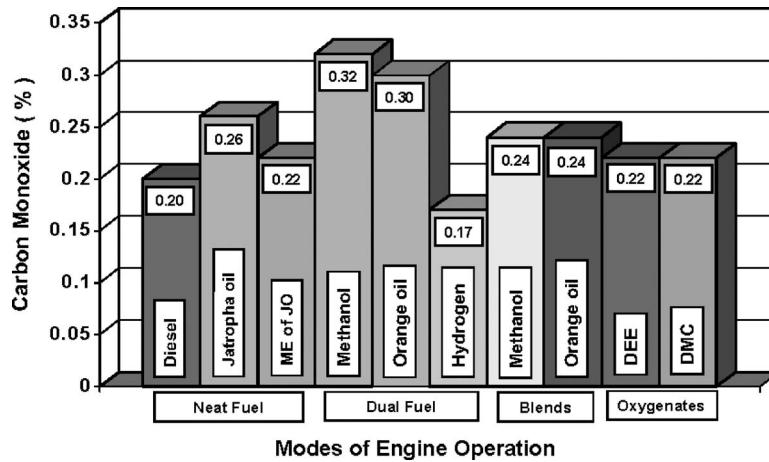


Fig. 6 Variation in carbon monoxide with different modes of engine operation

plied. This coupled with partial combustion at fuel lean pockets may be the reason for the high carbon monoxide levels. With hydrogen induction, the hydrocarbon and carbon monoxide are low. This is due to the reduction in the overall amount of carbon admitted into the engine [15].

NO emission shown in Fig. 7 is low with neat jatropha oil as compared with neat diesel due to reduced premixed burning rate as a result of slow combustion. It further reduces in the dual fuel operation with methanol and orange oil as inducted fuels. It is seen from Table 3 that NO is reduced from 740 ppm with neat jatropha oil to 702 ppm with methanol and 730 ppm with orange oil at optimum energy shares. This reduction in NO emission is due to the drop in charge temperature as a result of vaporization of methanol and orange oil. However, NO increases with methyl ester of jatropha oil due to better combustion. Methanol blend also shows lower NO emission than neat jatropha oil. The presence of methanol in the blend lowers the temperature after injection due to its higher latent heat of vaporization and this reduces NO levels. Hydrogen induction shows the highest NO level as compared with all other methods due to rapid combustion. This is the main disadvantage of using hydrogen as the inducted fuel. Hence, measures have to be taken to control NO emission with hydrogen [15].

3.3 Combustion Parameters. Ignition delay shown in Fig. 8 is high with neat jatropha oil as compared with neat diesel due to the low cetane number. It increases further in the dual fuel mode with all inducted fuels. Methanol induction shows higher ignition

delay than orange oil and hydrogen. This is due to the highest latent heat of vaporization of methanol. However, ignition delay is reduced with methyl ester of jatropha oil. With DEE addition it is reduced further and is almost equal to diesel (i.e., 9 deg) due to the increase in cetane number.

Peak pressure and rate of pressure rise are higher in dual fuel operation as compared with neat jatropha oil, blends, and oxygenates (Figs. 9 and 10). The increase in peak pressure and rate of pressure rise is due to enhanced combustion rate on account of flame propagation through the inducted charge. Methyl ester of jatropha oil and the blends of jatropha oil with methanol and orange oil also show increase in peak pressure and rate of pressure rise. However, DEE addition to jatropha oil results in reduced peak pressure and rate of pressure rise than others due to the increase in cetane number. An increase in the cetane number reduces the amount of fuel participating in the uncontrolled combustion phase.

Jatropha oil results in higher combustion duration than neat diesel as seen in Fig. 11 due to increase in the quantity of fuel injected. With methyl ester of jatropha oil, blends of methanol and orange oil and oxygenates like DEE and DMC with jatropha oil, the combustion duration is reduced. However, dual fuel operation with methanol induction results in high combustion duration due to reduction in charge temperature. This indicates late burning of the fuel. Dual fuel operation with hydrogen and orange oil showed low combustion duration due to their high combustion rate.

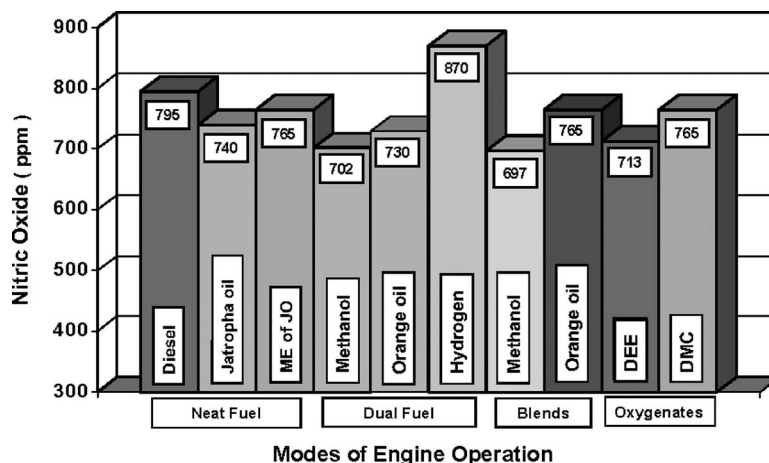


Fig. 7 Variation in nitric oxide with different modes of engine operation

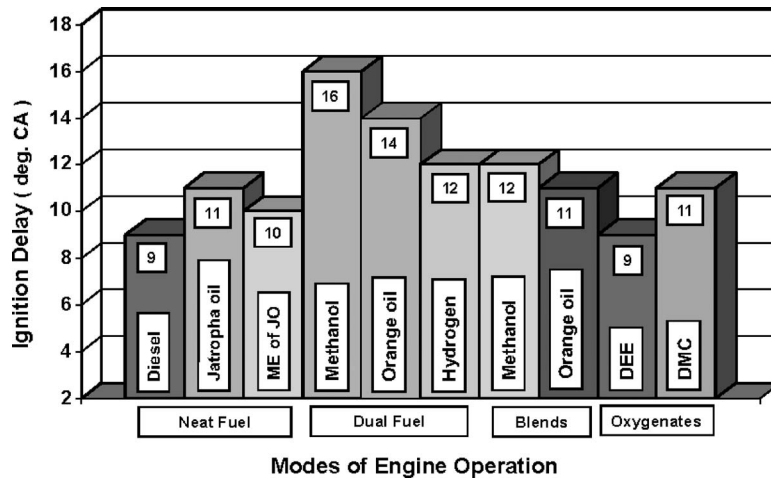


Fig. 8 Variation in ignition delay with different modes of engine operation

A comparison of heat release rate with jatropha oil, methyl ester of jatropha oil, and diesel is shown in Fig. 12. It is seen that the premixed burning is most significant with diesel. The diffusion-burning phase indicated under the second peak is greater for jat-

ropha oil. This is due to vegetable oil viscosity on the fuel spray and reduction in air enrichment and fuel air mixing rates. This leads to less fuel being prepared for rapid combustion with jatropha oil after the ignition delay. Therefore more burning occurs in

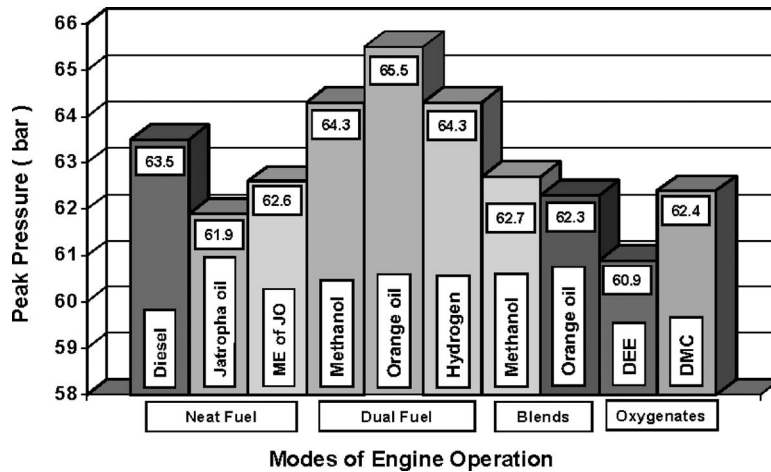


Fig. 9 Variation in cylinder peak pressure with different modes of engine operation

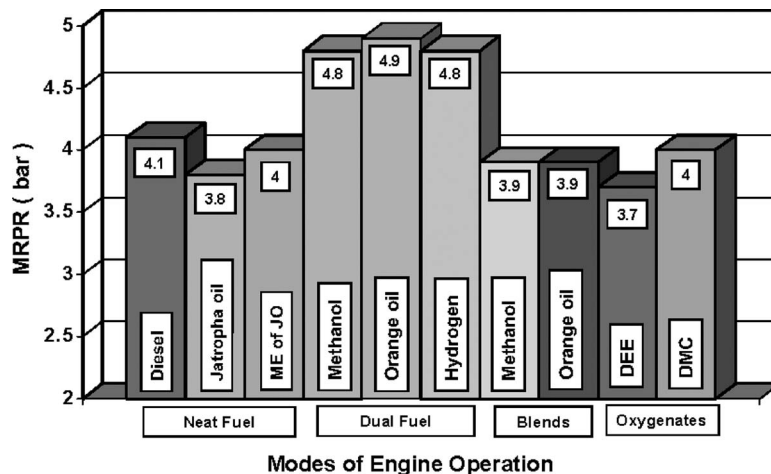


Fig. 10 Variation in maximum rate of pressure rise with different modes of engine operation

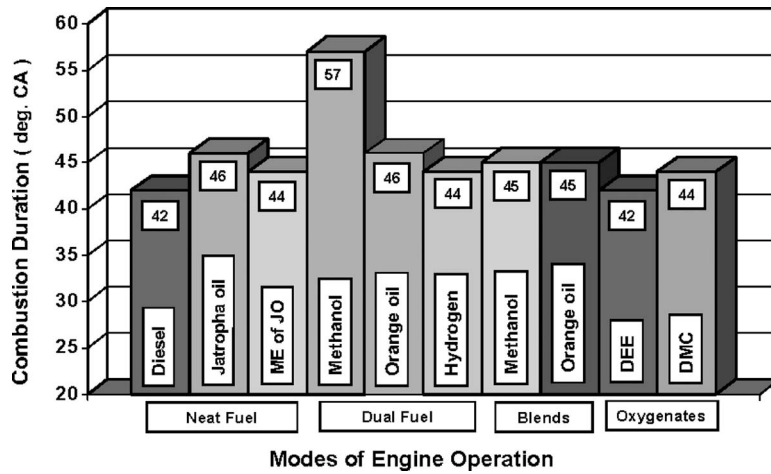


Fig. 11 Variation in combustion duration with different modes of engine operation

the diffusion phase rather than in the premixed phase with jatropha oil. The significantly higher combustion rates during the later stages with jatropha oil lead to high exhaust temperatures and lower thermal efficiency. However, methyl ester of jatropha oil shows improvement in heat release rate compared with neat jatropha oil.

In dual fuel operation, sharp increase in the heat release rate in the premixed combustion phase is observed with all inducted fuels at peak power (Fig. 13). This is due to rapid combustion as a result of high flame speed of all the inducted fuels. This is the reason for highest brake thermal efficiency in dual fuel operation than other methods.

With methanol and orange oil blends also there is a rise in premixed phase of the heat release rate than neat jatropha oil (Fig. 14). Methanol blend shows higher premixed combustion rate than orange oil blend due to higher ignition delay and large mass of fuel being accumulated during this period.

In the case of oxygenates, the premixed combustion phase with DMC is higher and more pronounced (Fig. 15). However, the premixed phase is lower with DEE than the neat oil due to re-

duced ignition delay. The end point of combustion also arrives earlier than that of neat jatropha oil. The good atomization and vaporization of DEE promote rapid mixing with the surrounding air, and oxygen available in the fuel improves the rate of diffusion combustion. With DMC blend the premixed phase is higher due to rapid combustion as a result of low cetane number.

Table 4 indicates the summary of results of all the methods attempted. It is seen that the use neat jatropha oil results in inferior performance and increased emissions indicated as "×" (except low NO emission), which are not good results to use neat jatropha oil in diesel engines. Methyl ester of jatropha oil resulted in considerable improvement in brake thermal efficiency and reduced smoke, hydrocarbon, and carbon monoxide emissions, which are the good results (indicated as "✓") to use methyl ester of jatropha oil in diesel engines. Dual fuel mode of operation showed higher brake thermal efficiency and lower smoke emissions indicated as ✓, which are the favorable results for diesel engines. However, increased hydrocarbon and carbon monoxide emissions with methanol and orange oil induction represent an unfavorable result of dual fuel operation indicated as ×. Blends of

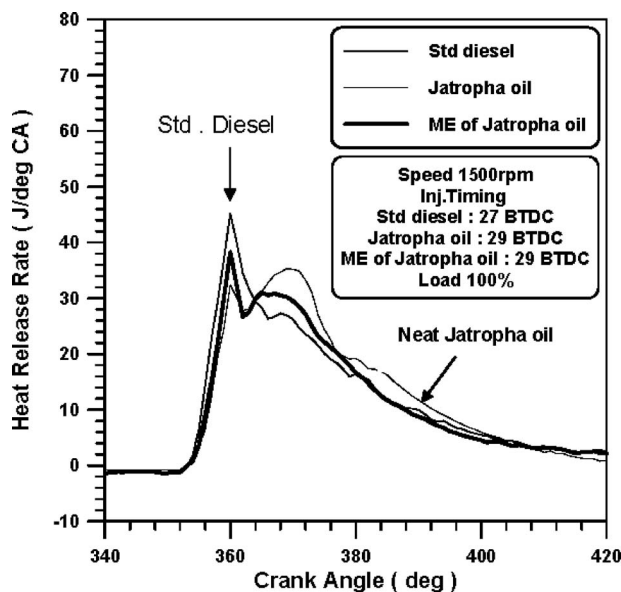


Fig. 12 Variation in heat release rate with jatropha oil and ME of jatropha oil at maximum power

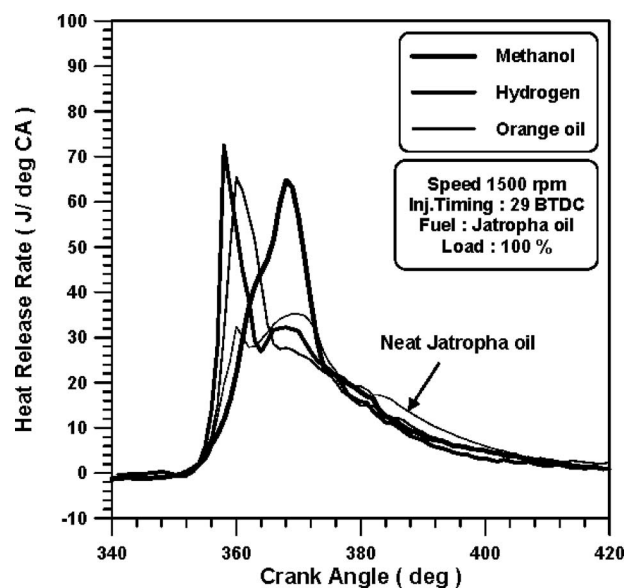


Fig. 13 Variation in heat release rate at maximum efficiency points

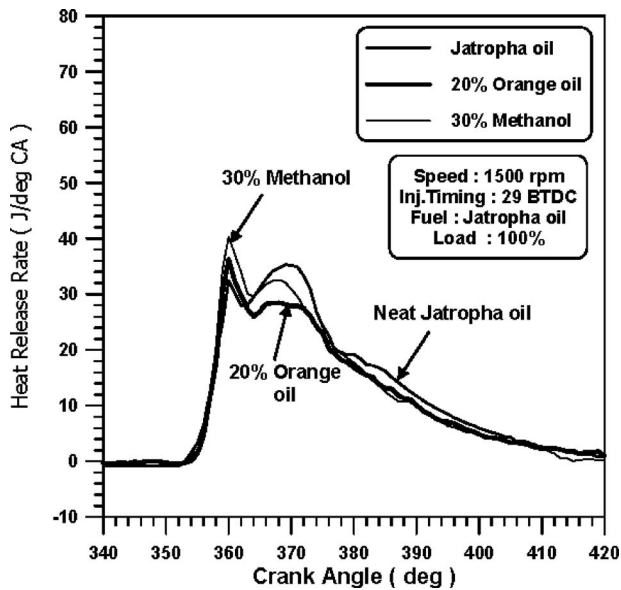


Fig. 14 Variation in heat release rate with different blends at optimum blend ratio

methanol and orange oil and oxygenates also improved the efficiency of the engine, as indicated in Table 4. However they increased the NO emission.

4 Conclusion

Methyl ester of jatropha oil, dual fuel operation, etc.:

- Increases the brake thermal efficiency as compared with neat jatropha oil at all outputs. However, the performance is still inferior to diesel. At full load (3.7 kW), the brake thermal efficiency is 27.3% with jatropha oil, 29% with methyl ester of jatropha oil, and 30.3% with diesel.
- Reduces the hydrocarbon, carbon monoxide, and smoke emissions as compared with neat jatropha oil. The smoke level for jatropha oil is 4.4 BSU and 4 BSU for its ester at full load. In the case of diesel it is 3.8 BSU.
- Increases NO emission but is lower than diesel.

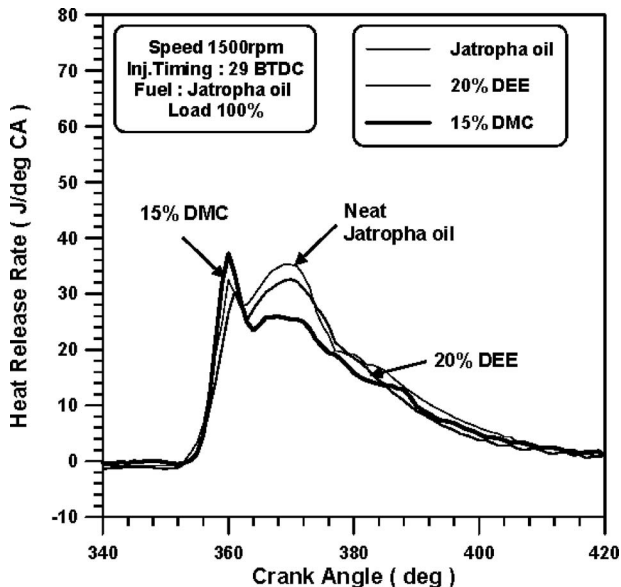


Fig. 15 Variation in heat release rate with different oxygenates

Table 4 Comparison of different methods

Methods	Neat fuels		Dual fuel operation			Blends		Oxygenates	
Fuels	JO	MEJO	DF1	DF2	DF3	B1	B2	DEE	DMC
Brake thermal efficiency									
High		✓				✓	✓	✓	✓
Higher			✓	✓	✓				
Low	×								
Lower									
Smoke number									
High	×								
Higher						✓	✓		
Low		✓	✓	✓	✓			✓	✓
Lower									
Nitric oxide									
High		×					×		×
Higher					×				
Low	✓					✓		✓	
Lower			✓	✓					✓
Hydrocarbon									
High	×								
Higher			×	×					
Low		✓				✓	✓	✓	✓
Lower					✓				
Carbon monoxide									
High	×								
Higher			×	×					
Low	✓					✓	✓	✓	✓
Lower					✓				
Ignition delay									
High	×					×	×		×
Higher			×	×	×				
Low		✓							
Lower								✓	
Peak pressure									
High		✓				✓	✓		✓
Higher			✓	✓	✓				
Low	×								
Lower								×	

JO, jatropha oil; MEJO, methyl ester of jatropha oil; DF1, methanol; DF2, orange oil; DF3, hydrogen; B1, methanol; B2, orange oil; DEE, diethyl ether; DMC, dimethyl carbonate; ×, not a good result; and ✓, good result.

- Reduces ignition delay and combustion duration as compared with neat jatropha oil but values are higher than diesel.
- Increases the heat release rate as compared with neat jatropha oil resulting in better brake thermal efficiency.

In dual fuel operation, the following were achieved.

- Increases the brake thermal efficiency considerably with all inducted fuels. At peak power output the maximum brake thermal efficiencies are 30.7%, 29.4%, and 29.3% at energy shares of 46%, 31%, and 18%, respectively, when methanol, orange oil, and hydrogen are inducted.
- Reduces the smoke level significantly from 4.4 BSU with neat jatropha oil to 2.6 BSU with methanol, 3.2 BSU with orange oil, and 3.3 BSU with hydrogen at corresponding maximum efficiency points. Hydrogen induction results in the lowest smoke emission for a given energy share.
- Increases HC and CO levels with methanol and orange oil as inducted fuels. However, hydrogen admission reduces these emissions.
- Increases NO emission level with hydrogen from 740 ppm

to 870 ppm at the maximum efficiency point. Methanol and orange oil induction reduces NO emission to 702 ppm and 734 ppm, respectively.

- Increases peak pressure, rate of pressure rise, and heat release rate due to rapid combustion with all the inducted fuels.

In blending of jatropha oil with orange oil and methanol, the following were achieved.

- Increases brake thermal efficiency. The maximum brake thermal efficiency is 28.5% with methanol and 28.3% with orange oil at optimum blends at peak load.
- Reduces smoke, hydrocarbon, and carbon monoxide emissions. Smoke is reduced from 4.4 BSU with neat jatropha oil to 4 BSU with methanol and 4.2 BSU with orange oil.
- Increases peak pressure, rate of pressure rise, and premixed combustion rate.
- Methanol blend reduces NO level and increases the ignition delay.

In oxygenates (diethyl ether and dimethyl carbonate) to jatropha oil, the following were achieved.

- Increase brake thermal efficiency. The maximum brake thermal efficiency is 29% with 20% diethyl ether and 28.6% with 15% DMC by volume.
- Reduce smoke, hydrocarbon, and carbon monoxide emissions
- Reduce NO levels with DEE but increase with DMC
- Reduce ignition delay with DEE
- Reduce in combustion duration
- Reduce peak pressure, rate of pressure, and heat release rate

The above results are made based on the tests on a single cylinder, water-cooled, constant speed diesel engine running at a rated power output of 3.68 kW. Hence these results are more applicable for the diesel engines of same configurations. It can be concluded from the above results that dual fuel operation with methanol, orange oil, and hydrogen can be a good method to use jatropha oil efficiently in an engine that normally operates at high power outputs. Induction of hydrogen will be very effective if measures are taken to maximize its amount without knock. Methyl ester of jatropha oil can lead to good performance at part loads with acceptable levels of performance at high loads also. However, preparation of the methyl ester is a cumbersome process and also leads to certain byproducts that cannot be used in the engine. Orange oil, methanol, diethyl ether, and dimethyl carbonate can also be blended with jatropha oil. These produce acceptable levels of performance at all outputs.

References

- [1] Barsic, N. J., and Humke, A. L., 1981, "Performance and Emission Character-

istics of a Naturally Aspirated Diesel Engine With Vegetable Oils," SAE Paper No. 810262.

- [2] Bari, S., Lim, T. H., and Yu, C. W., 2002, "Effect of Preheating of Crude Palm Oil (CPO) on Injection System, Performance and Emission of a Diesel Engine," *Int. J. Renewable Energy Technol.*, **27**, pp. 339–351.
- [3] Gopalakrishnan, K. V., and Rao, P. S., 1996, "Use of Non-Edible Vegetable Oils as Alternate Fuels in Diesel Engines," DNES Project Report, I.C.E.lab, Indian Institute of Technology Madras-36.
- [4] Vellguth, G., 1983, "Performance of Vegetable Oils and Their Monoesters as Fuel for Diesel Engines," SAE Paper No. 831358.
- [5] Senthil Kumar, M., Ramesh, A., and Nagalingam, B., 2001, "Complete Vegetable Oil Fuelled Compression Ignition Engine," SAE Paper No. 2001-28-0067.
- [6] Basker, T., 1993, "Experimental Investigation on the Use of Vegetable Oil and Vegetable Oil Ester in a Low Heat Rejection Engine," MS thesis, Indian Institute of Technology Madras, Chennai, India.
- [7] Masjuki, H., Abdulmuin, M. Z., and Sii, H. S., 1996, "Indirect Injection Diesel Engine Operation on Palm Oil Methyl Esters and Its Emulsions," *Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.)*, **211**, pp. 291–299.
- [8] Senthil Kumar, M., Ramesh, A., and Nagalingam, B., 2000, "Performance Studies on a CI Engine Using Methyl Ester of Jatropha Oil as Fuel," 16th National Conference on I.C. Engines and Combustion, Calcutta, India, pp. 89–94.
- [9] Takagi, N., and Itow, K., 1984, "Low Carbon Flower Buildup Low Smoke and Efficient Diesel Operation With Vegetable Oils by Conversion to Monoesters and Blending With Diesel Oil or Alcohols," SAE Paper No. 841161.
- [10] Karim, G. A., 1987, "The Dual Fuel Engine," *Automotive Engine Alternatives*, R. L. Evans, ed., Plenum, New York, pp. 83–104.
- [11] Naeser, D., and Bennet, K. F., 1980, "The Operation of Dual Fuel Compression Ignition Engines Utilizing Diesel and Methanol," *Proceedings of the Fourth International Symposium on Alcohol Fuels Technology*, Brazil.
- [12] Karim, G. A., and Amoozegar, N., 1983, "Determination of the Performance of a Dual Fuel Diesel Engine With Addition of Various Liquid Fuels to the Intake Charge," SAE Paper No. 830265.
- [13] Murayama, T., Zheng, M., and Chikahisa, T., 1995, "Simultaneous Reduction of Smoke and NO_x From a Diesel Engine With EGR and Dimethyl Carbonate," SAE Paper No. 952518.
- [14] Bailey, B., Eberhardt, J., and Goguen, S., 1997, "Diethyl Ether as a Renewable Diesel Fuel," SAE Paper No. 972978.
- [15] Prabhu Kumar, G. P., 1985, "Theoretical and Experimental Investigation on the Performance of Hydrogen Fueled Internal Combustion Engine," Ph.D. thesis, Indian Institute of Technology Madras, Chennai, India.
- [16] Patil, V., and Singh, K., 1991, "Oil Gloom to Oil Boom (Jatropha Curcas)," *Agro-Forestry Federation Maharashtra*.
- [17] Foidl, N., Foidl, G., Sanchez, M., Mittelbach, M., and Hackel, S., 1996, "Jatropha Curcas L. as a Source for the Production of Biofuel in Nicaragua," *Bioresour. Technol.*, **58**, pp. 77–82.
- [18] Pramanik, K., 2003, "Properties and Use of Jatropha Curcas Oil and Diesel Fuel Blends in a Compression Ignition Engine," *Int. J. Renewable Energy Technol.*, **28**, pp. 239–248.
- [19] Ramesh Babu, P., 1993, "Experimental Investigations on a Two Stroke SI Engine to Achieve Improved Performance," Ph.D. thesis, Indian Institute of Technology Madras, Chennai, India.
- [20] Prabhukumar, G. P., Swaminathan, S., Nagalingam, B., and Gopalakrishnan, K. V., 1987, "Water Induction Studies in a Hydrogen-Diesel Dual-Fuel Engine," *Int. J. Hydrogen Energy*, **12**(3), pp. 177–186.
- [21] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill Ltd.
- [22] Srinivasaraghavan, S., 1992, "Performance Studies on a Small Output Two Stroke S.I. Engine With Orange Oil and Eucalyptus Oil Blends," B.Tech. Project Report, I.C. Engines Laboratory, Indian Institute of Technology Madras, Chennai-36, India.
- [23] Thomas, C., 1999, "Experimental Investigation on the Performance, Fuel Economy, Combustion and Emission Characteristics of a Diesel Hydrogen Dual Fuel Engine," M.Tech. Project Report, I.C. Engines Laboratory, Indian Institute of Technology Madras, India.

An Analytical Approach for the Evaluation of the Optimal Combustion Phase in Spark Ignition Engines

A. Beccari

S. Beccari

E. Pipitone¹

e-mail: pipitone@dima.unipa.it

Department of Mechanics,
University of Palermo,
Palermo 90128, Italy

It is well known that the spark advance is one of the most important parameters influencing the efficiency of a spark ignition engine. A change in this parameter causes a shift in the combustion phase, whose optimal position, with respect to the piston motion, implies the maximum brake mean effective pressure for given operative conditions. The best spark timing is usually estimated by means of experimental trials on the engine test bed or by means of thermodynamic simulations of the engine cycle. In this work, instead, the authors developed, under some simplifying hypothesis, an original theoretical formulation for the estimation of the optimal combustion phase. The most significant parameters involved with the combustion phase are taken into consideration; in particular, the influence of the combustion duration, of the heat release law, of the heat transfer to the combustion chamber walls, and of the mechanical friction losses is evaluated. The theoretical conclusion, experimentally proven by many authors, is that the central point of the combustion phase (known as the location of the 50% of mass fraction burnt, here called MFB50) must be delayed with respect to the top dead center as a consequence of both heat exchange between gas and chamber walls and friction losses.

[DOI: 10.1115/1.3155395]

1 Introduction

The phasing of the combustion process with respect to the piston motion is one of the most important parameters influencing the torque provided by a spark ignition (SI) engine. Since this combustion phase depends on the spark ignition timing, located at the so called ignition angle, the control upon this angle is very important to obtain the best performance in every operative condition.

Two different approaches are commonly followed for the determination of the maximum brake torque (MBT) spark timing: The first is based on experimental trials at the engine test bed, which can be performed by either maximizing the engine torque (see, e.g., Refs. [1–3]) or setting a “combustion phase indicator” (which is a parameter derived from in-cylinder pressure analysis and assumes reference fixed values when the combustion timing is optimal [4]) to its best value [5–7]. The second approach, instead, proceeds by means of simulations based on a thermodynamic model of the engine [8–11], which, endowed of appropriate sub-models for combustion [12], heat transfer [13], and friction loss [8,14,15] modeling, allow the brake mean effective pressure (bmep) estimation once the spark advance is fixed. For each engine operative condition, the best combustion phase can be found by means of successive trials. In this paper, instead, the authors propose an original theoretical approach to the problem of the determination of the optimal combustion phase in spark ignition engines.

With reference to the working cycle of the engine (compression and expansion strokes during a crankshaft rotation from -180 deg to $+180$ deg with respect to the top dead center (TDC)), the angular phase of the heat introduction Q_1 that grants the maximum torque first depends on the way the combustion takes place (i.e.,

on the heat release law) and then on the effects of the heat Q_1 upon the various thermodynamic variables involved.

The theoretical approach proposed can be carried out at different approximation levels with respect to the following hypothesis:

- unsteady, ideal, zero-dimensional evolution of a perfect gas in the combustion chamber
- adiabatic engine (i.e., no heat exchange between the gas and the chamber walls)
- constant specific heat capacity (c_v) of the gas
- instantaneous combustion
- combustion length $\vartheta_c = \vartheta_b - \vartheta_a$, being ϑ_a and ϑ_b the crank angular positions at the starting and ending of the combustion process
- presence of heat transfer between the gas and the chamber walls
- presence of friction losses

2 Gas Temperature Trend During Combustion

Considering that previous hypotheses (a)–(c) and (e) are true, it is quite simple to evaluate the gas temperature trend versus the crank angular position ϑ . Ignoring the gas speed and the subsequent viscous friction losses, it is possible to write the first and the second laws of thermodynamics during the infinitesimal combustion time dt , corresponding to the crank rotation angle $d\vartheta$, in which the gas receives the combustion specific heat dQ_{in} :

$$dQ_{in} = c_v dT + p dv = c_v dT + RT \frac{dv}{v} = T dS \quad (1)$$

where T , v , p , S , c_v , and R are the gas temperature, specific volume, pressure, specific entropy, constant volume specific heat, and gas constant, respectively.

Assuming that

$$\varphi(\vartheta) = \frac{R}{c_v} \frac{1}{v} \frac{dv}{d\vartheta} = \frac{k-1}{v} \frac{dv}{d\vartheta}$$

¹Corresponding author.

Contributed by the Internal Combustion Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received September 11, 2008; final manuscript received May 21, 2009; published online November 24, 2009. Review conducted by Christopher J. Rutland.

$$\psi(\vartheta) = -\frac{1}{c_v} \frac{dQ_{in}}{d\vartheta}$$

where k is the isentropic coefficient equal to the ratio between the constant pressure (c_p) and constant volume (c_v) specific heat, expression (1) assumes the form of a linear differential equation

$$\frac{dT}{d\vartheta} + T \cdot \varphi(\vartheta) + \psi(\vartheta) = 0 \quad (2)$$

whose integral is

$$T = e^{-\int \varphi(\vartheta) d\vartheta} \left[\text{const} - \int \psi(\vartheta) e^{\int \varphi(\vartheta) d\vartheta} d\vartheta \right]$$

The solution is then

$$T v^{k-1} = \text{const} - \int \psi(\vartheta) \cdot v^{k-1} d\vartheta$$

which applied from the start ϑ_a to the end ϑ_b of the combustion process gives

$$T_b v_b^{k-1} = T_a v_a^{k-1} + \frac{1}{c_v} \int_{\vartheta_a}^{\vartheta_b} v^{k-1} \frac{dQ_{in}}{d\vartheta} d\vartheta = \zeta_b \quad (3)$$

From Eq. (1), the whole specific entropy increase ΔS_{ab} , corresponding to heat introduction $Q_1 = \int_{\vartheta_a}^{\vartheta_b} dQ_{in}$, is

$$\begin{aligned} \Delta S_{ab} &= \int_a^b \left[c_v \frac{dT}{T} + R \frac{dv}{v} \right] = c_v \left[\ln \frac{T_b}{T_a} + \frac{R}{c_v} \ln \frac{v_b}{v_a} \right] \\ &= c_v \ln \frac{T_b v_b^{k-1}}{T_a v_a^{k-1}} = c_v \ln \frac{\zeta_b}{\zeta_a} \\ &= c_v \ln \left[1 + \frac{1}{c_v T_a v_a^{k-1}} \int_{\vartheta_a}^{\vartheta_b} v^{k-1} \frac{dQ_{in}}{d\vartheta} d\vartheta \right] \\ &= c_v \ln \left[1 + \frac{1}{c_v T_M v_M^{k-1}} \int_{\vartheta_a}^{\vartheta_b} v^{k-1} \frac{dQ_{in}}{d\vartheta} d\vartheta \right] \end{aligned} \quad (4)$$

where T_M and v_M are the gas temperature and specific volume when the piston is at the bottom dead center (BDC) related to T_a and v_a by the isentropic law; these variables are independent of the combustion phase length and position, i.e., from ϑ_a and ϑ_b .

The minimum entropy increase ΔS_{ab} corresponds to the maximum efficiency of the thermodynamic cycle if compression and expansion phases are adiabatic; in fact, in this case, $-\Delta S_{ab}$ is also the entropy decrease corresponding to the heat subtraction Q_2 , at constant volume v_M , needed to close the thermodynamic cycle of the gas.

With reference to the thermodynamic cycle (A_0EFGHH_0) of Fig. 1, assuming that $x(\vartheta) = Q_{in}(\vartheta)/Q_1$ is the heat fraction released during the crank rotation from ϑ_a to ϑ (being $\vartheta_a < \vartheta < \vartheta_b$) with respect to the total $Q_1 = Q_{in}(\vartheta_b)$,

$$Q_2 = c_v T_{A_0} (e^{\Delta S_{ab}/c_v} - 1) \quad (5)$$

$$\Rightarrow \frac{\Delta S_{ab}}{c_v} = \ln \left(1 + \frac{Q_2}{c_v T_{A_0}} \right) \quad (6)$$

Then, from relation (4), it follows that

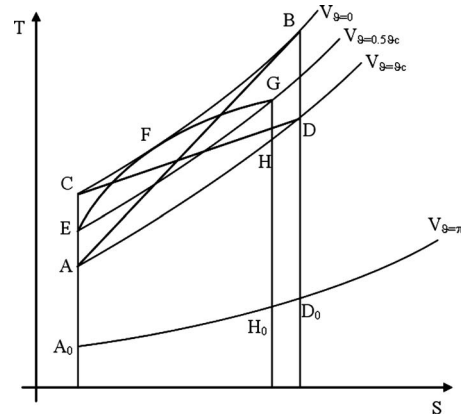


Fig. 1 Thermodynamic cycle with noninstantaneous combustion (AB, CD, and EFG have the same duration)

$$\begin{aligned} \frac{\Delta S_{ab}}{c_v} &= \ln \left[1 + \frac{1}{c_v T_C} \int_{\vartheta_a}^{\vartheta_b} \left(\frac{v}{v_C} \right)^{k-1} \frac{dQ}{d\vartheta} d\vartheta \right] \\ &= \ln \left[1 + \frac{Q_1}{c_v T_C} \int_{\vartheta_a}^{\vartheta_b} \left(\frac{v}{v_C} \right)^{k-1} \frac{dx}{d\vartheta} d\vartheta \right] \\ &\Rightarrow \frac{Q_2}{c_v T_{A_0}} = \frac{Q_1}{c_v T_C} \int_{\vartheta_a}^{\vartheta_b} \left(\frac{v}{v_m} \right)^{k-1} \frac{dx}{d\vartheta} d\vartheta \end{aligned} \quad (7)$$

$$1 - \eta = \frac{Q_2}{Q_1} = \frac{T_{A_0}}{T_C} \int_{\vartheta_a}^{\vartheta_b} \left(\frac{v}{v_m} \right)^{k-1} \frac{dx}{d\vartheta} d\vartheta = \frac{1}{\rho_0^{k-1}} \int_{\vartheta_a}^{\vartheta_b} \left(\frac{v}{v_m} \right)^{k-1} \frac{dx}{d\vartheta} d\vartheta$$

where $v_m = v_C$ is the gas specific volume at TDC, v_M is the gas specific volume at BDC, $\rho_0 = V_M/V_m$ is the volumetric compression ratio, and η is the thermodynamic cycle efficiency.

The above relations confirm that the combustion with a finite time ($\vartheta_c = \vartheta_b - \vartheta_a$), involving gas specific volumes higher than the TDC one v_m , leads to a lower efficiency than the instantaneous combustion, in which $v = v_m$, hence

$$1 - \eta = \frac{1}{\rho_0^{k-1}}$$

being, in this case,

$$\int_{\vartheta_a}^{\vartheta_b} \left(\frac{v}{v_m} \right)^{k-1} \frac{dx}{d\vartheta} d\vartheta = \int_{\vartheta_a}^{\vartheta_b} \frac{dx}{d\vartheta} d\vartheta = 1$$

3 Optimal Combustion Phase in the Adiabatic Engine

In the hypothesis d with instantaneous combustion,

$$\vartheta_a = \vartheta_b \Rightarrow \vartheta_c = 0$$

$$v = v_a = v_b = \text{const}$$

relation (6) yields

$$\Delta S_{ab} = c_v \ln \left(1 + \frac{Q_1}{c_v T_a} \right) \quad (8)$$

The entropy increase ΔS_{ab} reaches the minimum when T_a reaches the maximum, which happens at TDC ($T_a = T_C$ in Fig. 1).

When the combustion takes place during a finite phase $\vartheta_c = \vartheta_b - \vartheta_a$, the evaluation of ΔS_{ab} by means of relation (4) requires

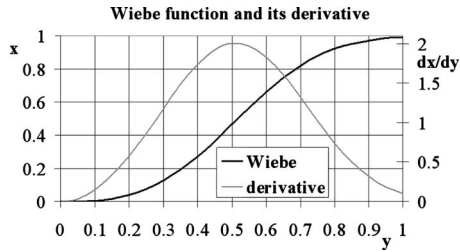


Fig. 2 The Wiebe function and its derivative, which is not a symmetric curve

the knowledge of the two functions $V=V(\vartheta)$ and $dQ_{in}(\vartheta)/d\vartheta = Q_1(dx/d\vartheta)$.

Assuming that $V_0=V_M-V_m$ is the engine displacement, $\mu = r/c$ the ratio between the connecting rod length (r) and the crank radius (c), and ρ the compression ratio, the first function can be evaluated as follows:

$$V(\vartheta) = V_0 \left[\frac{1}{\rho-1} + \frac{\mu}{2} \left(1 + \frac{1 - \cos \vartheta}{\mu} - \sqrt{1 - \frac{\sin^2 \vartheta}{\mu^2}} \right) \right] \quad (9)$$

which is a symmetrical function with respect to the TDC position $\vartheta=0$.

Concerning the "heat release" function $Q_{in}(\vartheta)$ and its derivative $dQ_{in}(\vartheta)/d\vartheta$, it is opportune to evaluate some specific quantities such as $x=Q_{in}(\vartheta)/Q_1$ (fraction of heat released) and $y=(\vartheta - \vartheta_a)/\vartheta_c$ (fraction of combustion arc).

The function $x=x(y)$ has two constrained points, the origin O (0; 0) and the final point P (1; 1), because at the beginning of combustion ($\vartheta=\vartheta_a; y=0$), there is no heat released ($x=0$), while at the end ($\vartheta=\vartheta_b; y=1$), all the heat Q_1 has been released ($x=1$).

The Wiebe function (see Fig. 2), commonly used to model the heat released by the combustion in a SI engine, is

$$x = 1 - e^{-5y^3} \quad (10)$$

$$\frac{dx}{dy} = 15y^2(1-x) \quad (11)$$

The starting point of this function is ($y=0; x=0$) and for $y=1$, x is very close to unit (0.99) so it almost respects the two conditions mentioned above, while derivative (11) is not symmetric with respect to abscissa $y=0.5$, i.e., to the middle of the combustion arc.

A more simple function is

$$x = y \quad (12)$$

$$\frac{dx}{dy} = 1 \quad (13)$$

which have the extreme points O ($x=0; y=0$) and P ($x=1; y=1$), and a symmetric derivative.

The constant derivative (which means constant combustion velocity) is not plausible in a spark ignition engine; hence a better approximation is obtained by the following polynomial function:

$$x = 3y^2 \left(1 - \frac{2}{3}y \right) \quad (14)$$

$$\frac{dx}{dy} = 6y(1-y) \quad (15)$$

which, as shown in Fig. 3, passes through the extreme points O ($x=0; y=0$) and P ($x=1; y=1$) and has a symmetric derivative (combustion speed with parabolic trend).

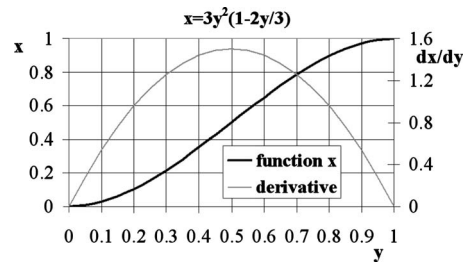


Fig. 3 The function (14) and its derivative

With respect to function (13), function (15) (as well as the classical Wiebe function (11)) has a more reasonable characteristic: The combustion rate is zero at the beginning ($y=0$) and at the end of the combustion process ($y=1$).

The function V^{k-1} is symmetric with respect to the TDC ($\vartheta=0$); if the function $dQ_{in}/d\vartheta$ (that is, dx/dy) is also symmetric with respect to the middle of the combustion angle ($y=0.5$), then, in Fig. 1, the two combustions AB, entirely located before the TDC ($\vartheta_a=-\vartheta_c; \vartheta_b=0$), and CD, entirely located after the TDC ($\vartheta_a=0; \vartheta_b=\vartheta_c$), produce the same entropy increase ΔS_{ab} (see Eqs. (3) and (4) and Fig. 4).

The combustion EFG (as shown in Figs. 1 and 4), which is symmetrically located with respect to the TDC ($\vartheta_a=-\vartheta_b=-0.5\vartheta_c$), represents instead the optimum because it produces the minimum entropy increase: The function $dQ_{in}/d\vartheta$, in fact, is the same as before (combustions AB and CD) but the function V^{k-1} has a smaller average value (the piston position, on the average, is closer to the TDC than before). It is important to notice that, as shown in Eqs. (4) and (7), the optimum condition does not depend on the whole heat released Q_1 but rather on the heat release law as a function of the time or crank angle (CA) rotation. An analytical approach to the determination of the optimal combustion phase for the adiabatic and frictionless engine is shown in the Appendix, also for the case of a not symmetrical heat release rate function.

4 Optimal Combustion Phase Taking Into Account Heat Exchanges and Friction Losses

Taking into account the heat exchanges between in-cylinder gas at temperature T and combustion chamber walls at temperature T_w , according to the hypothesis f in Sec. 1, Eq. (1) becomes

$$dQ_{in} - dQ_w = c_v dT + RT \frac{dv}{v} \quad (16)$$

where dQ_w is the specific heat subtracted, in the time interval dt , from the gas by the chamber walls, whose mean temperature T_w can be assumed to be constant:

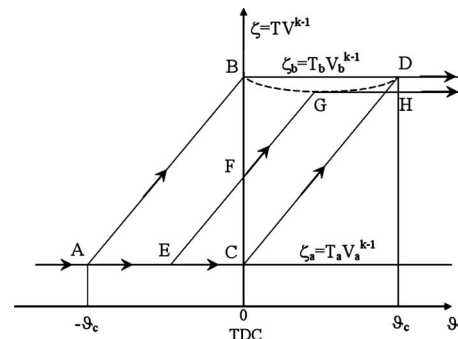


Fig. 4 Progress of $\zeta = TV^{k-1}$ as a function of the crank position ϑ for three combustions with different phase and same heat release law (see Eq. (3))

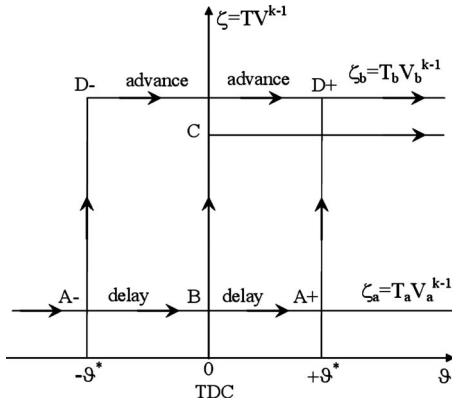


Fig. 5 Progress of $\zeta = TV^{k-1}$ as a function of the crank position ϑ for three instantaneous combustions with different phase

$$m \cdot \dot{Q}_w = h \cdot S' \cdot \Delta T = h(\vartheta) \cdot S'(\vartheta) \cdot (T(\vartheta) - T_w) \quad (17)$$

where m is the gas mass, h is the convective heat exchange coefficient, S' is the combustion chamber surface, and $\dot{Q}_w = dQ_w/dt$. Both Q_{in} and Q_w are functions of the crank position ϑ , but \dot{Q}_w is certainly not symmetric with respect to the TDC ($\vartheta=0$).

In this case the differential equation cannot be solved analytically and only qualitative considerations can be made.

Taking into account, for a first approximation, the temperature progress related to the adiabatic engine (see Figs. 1 and 4), it is possible to state that the combustion CD, delayed with respect to the symmetrical one EFG, causes, for each piston position, lower gas temperatures than the advanced one AB, producing thus lower heat exchanges (Q_w) and probably reducing the whole heat subtracted from the gas during its thermodynamic cycle $Q_{2total} = Q_2 + Q_w$.

Obviously a reduction in Q_{2total} , at constant introduced heat Q_1 , means an improvement of the cycle efficiency.

As a matter of fact, a combustion delayed by the angle ϑ with respect to the one symmetrically placed around the TDC produces an increase in ΔS_{ab} (see Eq. (4)) that means an increase in Q_2 according to Eq. (5), while, on the other hand, it produces a decrease in Q_w so when the two effects (the first tends to worsen the engine efficiency, while the second tends to improve it) compensate each other ($\Delta Q_2 = \Delta Q_w$), the optimal combustion phase is reached:

$$\Delta Q_{2total} = \Delta Q_2 + \Delta Q_w = 0$$

In the case of the heat release function (11), asymmetric with respect to the middle of the combustion arc, the above conclusions remain valid because heat transfer causes anyway the optimum condition to be forward shifted with respect to the adiabatic case.

Some quantitative conclusions can be drawn taking again into consideration the temperature progress of the adiabatic engine in the limit case of an instantaneous combustion (i.e., with $\vartheta_c=0$).

To this purpose, Fig. 6 shows the combustion at the TDC (line BC), the advanced combustion (line A-D-), and the delayed combustion (line A+D+). The last two combustions are symmetrically placed around the TDC. In Fig. 5, the same combustion lines are represented in a plane with coordinates $[\vartheta, \zeta = TV^{k-1}]$.

First, it is clear that an advanced or delayed instantaneous combustion brings a smaller volumetric compression ratio with respect to ρ_0 (TDC instantaneous combustion).

Second, from the analysis of Fig. 5, traced on the basis of Eq. (3), it is clear that an advanced combustion (line A-D-) produces higher mean temperatures than the delayed combustion (line A+D+), and so in the case of a delayed instantaneous combustion, the heat exchanges (Q_w) are smaller than in the case of an ad-

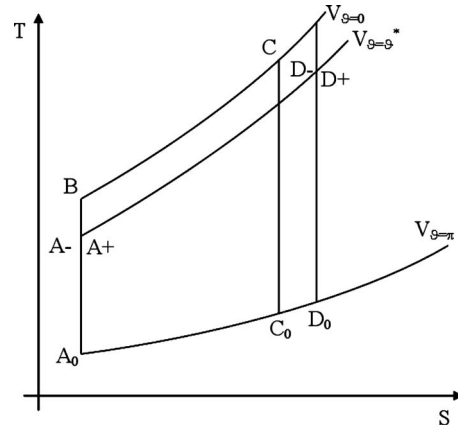


Fig. 6 Effect of instantaneous combustion phase changes on the Otto cycle

vanced one: This leads to the conclusion that the optimum phase of the combustion will be delayed with respect to the TDC.

When also the friction losses are taken into account, the bmep has to be considered as the engine output performance, rather than the indicated mean effective pressure (imep), being

$$\text{bmep} = \text{imep} - \text{fmep} \quad (18)$$

where fmep is the friction mean effective pressure, which accounts for the energy drained for different reasons: the mechanical friction between piston rings and cylinder liner, the mechanical friction in the loaded bearings (crankshaft and connecting rod), and the energy required to run the auxiliary components. The maximum engine efficiency then corresponds, for a constant introduced heat Q_1 , to the maximum bmep.

In Sec. 5, the optimal instantaneous combustion phase will be searched for now: For a given increment $d\vartheta^*$ of the instantaneous combustion delay (with respect to the TDC) ϑ^* , the variation $d(\text{bmep})$ will be evaluated, so that the condition $d(\text{bmep})=0$ will indicate the optimum delay value ϑ^*_{opt} of the instantaneous combustion.

In an ideal engine (adiabatic and without friction losses), an instantaneous combustion delayed or advanced with respect to TDC has the same effect of a volumetric compression ratio decrease on the thermodynamic cycle and its efficiency, as can be seen in Fig. 6; hence the variation in the heat subtracted from the expansion end (dQ_2), due to a given instantaneous combustion delay increment $d\vartheta^*$, can be evaluated as follows:

$$Q_1 = Q_2 \rho^{k-1}$$

For the constant Q_1 condition it follows that

$$\frac{dQ_1}{Q_1} = 0 = \frac{dQ_2}{Q_2} + (k-1) \frac{d\rho}{\rho}$$

$$\Rightarrow dQ_2 = - (k-1) \frac{d\rho}{\rho} Q_2 = - \frac{(k-1) d\rho}{\rho^{k-1}} \frac{Q_1}{\rho} \quad (19)$$

By definition, the imep is

$$\frac{\text{imep}}{m} = \frac{Q_1 - Q_2}{V_0} = \frac{Q_1(1 - Q_2)}{V_0} = \frac{Q_1}{V_0} \left(1 - \frac{1}{\rho^{k-1}} \right) \quad (20)$$

where m is the in-cylinder gas mass. Concerning the real engine (i.e., in the presence of heat exchanges with the chamber walls Q_w and friction losses), the variation in imep and bmep caused by a given instantaneous combustion delay increment, at constant Q_1 , can be written as

$$\frac{d(\text{imep})}{m} = \frac{d(Q_1 - Q_{2total})}{V_0} = - \frac{(dQ_2 + dQ_w)}{V_0} \quad (21)$$

$$\frac{d(\text{bmep})}{m} = \frac{d\left(Q_1 - Q_{2\text{total}} - \frac{V_0}{m} \cdot \text{fmep}\right)}{V_0} = -\frac{\left(dQ_2 + dQ_w + \frac{V_0}{m} d(\text{fmep})\right)}{V_0} \quad (22)$$

and the condition of maximum efficiency, $d(\text{bmep})=0$, gives

$$dQ_2 = -dQ_w - \frac{V_0}{m} d(\text{fmep})$$

In the above relation, only the components of fmep that are actually affected by the combustion position delay must be taken into account, namely, the friction at the piston-cylinder interface and in the loaded bearings (crankshaft and connecting rod); neither the energy required by the engine accessories nor the friction caused by the inertia forces of the moving masses must be considered because these are not affected by combustion phase variations.

The two considered friction components are then fmep_1 (related to the piston-cylinder friction) and fmep_2 (related to the loaded bearings), the condition of maximum efficiency thus becomes

$$dQ_2 = -dQ_w - \frac{V_0}{m} d(\text{fmep}_1) - \frac{V_0}{m} d(\text{fmep}_2) \quad (23)$$

The heat transferred from the gas to the chamber walls Q_w is assumed to be proportional to the exposition time of the gas to the cylinder walls and to the gas temperature (as a first approximation, the increase in heat exchange surface S' during the gas expansion is assumed to be compensated by the decrease in heat exchange coefficient h); then

$$Q_w = \text{const} \cdot \int_{\text{EC}}^{\text{BDC}} T d\vartheta$$

where T_{EC} is the gas temperature at the end of the instantaneous combustion (which remains almost unchanged for small combustion retard ϑ^*).

The increase in the heat exchanged with the chamber walls due to a delay increment $d\vartheta^*$ (with respect to the TDC) of the instantaneous combustion is then

$$dQ_w \approx -\text{const} \cdot T_{\text{EC}} \cdot d\vartheta^*$$

The last relation is quite acceptable considering that a combustion delay increment $d\vartheta^*$ around the TDC (where the in-cylinder volume is minimum) does not have a considerable effect on the temperature at the end of combustion (EC) so that $T_{\text{EC}} \approx \text{const}$ (see Figs. 5 and 6).

Since, as a first approximation, once the combustion has been completed, it can be considered as

$$TV^{k-1} = \text{const}$$

then

$$Q_w = \text{const} \cdot \int_{\text{EC}}^{\text{BDC}} T d\vartheta = \text{const} \cdot T_{\text{EC}} \int_{\text{EC}}^{\text{BDC}} \left(\frac{V^*}{V}\right)^{k-1} d\vartheta$$

where V is the instantaneous chamber volume (and V^* is the value it assumes at the end of combustion when $\vartheta = \vartheta^*$), and k is the gas isentropic coefficient. The mean crank rotation $d\vartheta$ during the expansion from the EC to the BDC can be considered, as a first approximation, proportional to the volume variation dV .

Hence

$$d\vartheta \approx \frac{\pi}{V_0} dV \Rightarrow Q_w = \text{const} \cdot T_{\text{EC}} \frac{\pi}{V_0} \int_{\text{EC}}^{\text{BDC}} \left(\frac{V^*}{V}\right)^{k-1} dV$$

being

$$\int_{\text{EC}}^{\text{BDC}} \left(\frac{V^*}{V}\right)^{k-1} dV \equiv \int_{\text{TDC}}^{\text{BDC}} \left(\frac{V_{\text{TDC}}}{V}\right)^{k-1} dV \quad (24)$$

it follows that

$$Q_w \equiv \text{const} \cdot T_{\text{EC}} \frac{\pi}{V_0} \int_{\text{TDC}}^{\text{BDC}} \left(\frac{V_{\text{TDC}}}{V}\right)^{k-1} dV = \text{const} \cdot T_{\text{EC}} \cdot \pi \frac{\rho^{2-k} - 1}{(2-k)(\rho-1)} \quad (25)$$

and then

$$\frac{dQ_w}{Q_w} = -\frac{d\vartheta^* (2-k)(\rho-1)}{\pi \rho^{2-k} - 1} \quad (26)$$

As can be observed in Eq. (26), a delay increment $d\vartheta^*$ of the instantaneous combustion produces a decrease in Q_w (hence, from Eq. (21), an imep increase) and an increase in Q_2 (see Fig. 6 and Eq. (19) where $d\vartheta^* > 0$ involves $d\rho < 0$), which instead causes an imep decrease in Eq. (21); the maximum of imep then implies a null variation $d(\text{imep})$ for a given increment $d\vartheta^*$, i.e., from Eq. (21), $-dQ_w = dQ_2$.

The losses due to the piston-cylinder friction during the gas expansion fmep_1 are here supposed, as a first approximation, to be proportional to the in-cylinder pressure and to the piston stroke, and hence to the chamber volume variation:

$$\text{fmep}_1 = \text{const} \cdot \int_{\text{EC}}^{\text{BDC}} p \tan(\varphi) \cdot dV$$

where φ is the angle between the connecting rod and the cylinder axis, which is a function of the crank angle ϑ and assumes the value φ^* when the combustion delay angle is ϑ^* . If p_{EC} represents the gas pressure at the end of combustion, the increase in friction losses due to the delay increment $d\vartheta^*$ (corresponding to a volume variation dV^*) of the instantaneous combustion is

$$d(\text{fmep}_1) \approx -\text{const} \cdot p_{\text{EC}} \tan(\varphi^*) \cdot dV^*$$

It is also

$$\tan(\varphi^*) \approx \sin(\varphi^*) = \frac{1}{\mu} \cdot \sin(\vartheta^*) \approx \frac{\vartheta^*}{\mu} \Rightarrow$$

$$d(\text{fmep}_1) \approx -\text{const} \cdot p_{\text{EC}} \cdot \frac{\vartheta^*}{\mu} \cdot dV^*$$

Hence the mean value of the variable $\tan(\varphi)$ during the crank rotation from the EC to the BDC can be found as follows:

$$\overline{\tan(\varphi)} \approx \overline{\sin(\varphi)} = \frac{1}{\mu} \overline{\sin(\vartheta)} = \frac{1}{\mu} \cdot \frac{1}{\pi} \cdot \int_0^\pi \sin(\vartheta) d\vartheta = \frac{2}{\mu \cdot \pi}$$

Assuming, as a first approximation, an isentropic expansion from the EC to the BDC, then

$$\begin{aligned} \text{fmep}_1 &= \text{const} \cdot \int_{\text{EC}}^{\text{BDC}} p \tan(\varphi) \cdot dV \\ &\approx \text{const} \cdot p_{\text{EC}} \overline{\tan(\varphi)} \cdot \int_{\text{EC}}^{\text{BDC}} \left(\frac{V^*}{V}\right)^k dV \\ &\approx \text{const} \cdot p_{\text{EC}} \cdot \frac{2}{\mu \cdot \pi} \cdot \int_{\text{EC}}^{\text{BDC}} \left(\frac{V^*}{V}\right)^k dV \end{aligned}$$

where the gas pressure p_{EC} has been assumed to remain almost constant for small combustion delay ϑ^* around the TDC; under the same approximation of Eq. (24), it follows that

$$\begin{aligned} \text{fmep}_1 &\approx \text{const} \cdot p_{\text{EC}} \cdot \frac{2}{\mu \cdot \pi} \cdot \int_{\text{TDC}}^{\text{BDC}} \left(\frac{V_{\text{TDC}}}{V} \right)^k dV \\ &= \text{const} \cdot \frac{2}{\mu \cdot \pi} \cdot p_{\text{EC}} V_0 \frac{\rho^{1-k} - 1}{(1-k)(\rho-1)} \end{aligned} \quad (27)$$

$$\frac{d(\text{fmep}_1)}{\text{fmep}_1} = - \frac{\vartheta^* \cdot \pi}{2} \cdot \frac{dV^*}{V_0} \cdot \frac{(1-k)(\rho-1)}{\rho^{1-k} - 1} \quad (28)$$

As can be observed in Eq. (28), a delay increment of the instantaneous combustion $d\vartheta^*$ (to which corresponds a volume change dV^*) produces a decrease in fmep_1 : This should imply a bmep increase (from Eq. (22)); the delay increment $d\vartheta^*$ however causes also a decrease in imep (21), for the shorter working stroke of the gas after the combustion (see Eq. (19)), and consequently a decrease in bmep in Eq. (22). In conclusion the maximum bmep condition can be pursued by a null variation $d(\text{bmep})$ for a given increment $d\vartheta^*$.

In a similar way to the losses fmep_1 , the friction losses fmep_2 (caused in the loaded bearings) are supposed to be proportional to the in-cylinder pressure and to the crank rotation angle during the gas expansion; hence

$$\text{fmep}_2 = \text{const} \cdot \int_{\text{EC}}^{\text{BDC}} p d\vartheta$$

$$d(\text{fmep}_2) \approx - \text{const} \cdot p_{\text{EC}} d\vartheta^*$$

where $d(\text{fmep}_2)$ is the increase in fmep_2 due to the delay increment $d\vartheta^*$ of the instantaneous combustion. Under the same hypothesis made for the evaluation of $d(\text{fmep}_1)$, it follows that

$$\text{fmep}_2 \cong \text{const} \cdot p_{\text{EC}} \int_{\text{TDC}}^{\text{BDC}} \left(\frac{V_{\text{TDC}}}{V} \right)^k d\vartheta$$

$$\begin{aligned} d\vartheta &\approx \frac{\pi}{V_0} dV \Rightarrow \text{fmep}_2 = \text{const} \cdot p_{\text{EC}} \frac{\pi}{V_0} \int_{\text{TDC}}^{\text{BDC}} \left(\frac{V_{\text{TDC}}}{V} \right)^k dV \\ \Rightarrow \text{fmep}_2 &= \text{const} \cdot p_{\text{EC}} \cdot \pi \frac{\rho^{1-k} - 1}{(1-k)(\rho-1)} \end{aligned} \quad (29)$$

$$\Rightarrow \frac{d(\text{fmep}_2)}{\text{fmep}_2} = - \frac{d\vartheta^* (1-k)(\rho-1)}{\pi \rho^{1-k} - 1} \quad (30)$$

As can be observed in Eq. (30), a delay increment $d\vartheta^*$ of the instantaneous combustion produces a decrease in fmep_2 , likewise for fmep_1 , with similar consequences.

The maximum engine efficiency condition (23)

$$dQ_2 = -dQ_w - \frac{V_0}{m} d(\text{fmep}_1) - \frac{V_0}{m} d(\text{fmep}_2)$$

together with Eqs. (19), (26), (28), and (30), leads to the following equation:

$$\begin{aligned} - \frac{(k-1) d\rho}{\rho^{k-1}} Q_1 &= w Q_1 \frac{d\vartheta^* (2-k)(\rho-1)}{\pi \rho^{2-k} - 1} \\ &+ \frac{\pi \cdot \vartheta^*}{2} v_1 \cdot (\text{imep}) \frac{dV^* (1-k)(\rho-1)}{V_0 \rho^{1-k} - 1} \frac{V_0}{m} \\ &+ v_2 \cdot (\text{imep}) \frac{d\vartheta^* (1-k)(\rho-1)}{\pi \rho^{1-k} - 1} \frac{V_0}{m} \end{aligned} \quad (31)$$

where $w = Q_w / Q_1$, $v_1 = \text{fmep}_1 / \text{imep}$, and $v_2 = \text{fmep}_2 / \text{imep}$.

Being

$$\begin{aligned} V_0 &= V_{\text{BDC}} - V_{\text{TDC}} = V_{\text{TDC}}(\rho_0 - 1) \approx V^*(\rho - 1) \Rightarrow V^* \approx \frac{V_0}{\rho - 1}, \quad \rho \\ &= \frac{V_{\text{BDC}}}{V^*} \end{aligned}$$

the equivalent compression ratio variation due to the instantaneous combustion phase shift is then

$$\frac{d\rho}{\rho} = - \frac{dV^*}{V^*} = - \frac{dV^* V_0}{V_0 V^*} = - \frac{dV^*}{V_0} (\rho - 1) \Rightarrow \frac{dV^*}{V_0} = - \frac{1}{\rho - 1} \frac{d\rho}{\rho} \quad (32)$$

and remembering Eq. (9),

$$\begin{aligned} \frac{d\rho}{\rho} &= - \frac{\rho - 1}{2} \sin \vartheta^* \left(1 + \frac{1}{\mu} \cos \vartheta^* \right) \cdot d\vartheta^* \Rightarrow \left(\frac{d\rho}{\rho} \right)_{\text{near TDC}} \\ &= - \frac{1 + \mu}{2\mu} (\rho - 1) \cdot \vartheta^* \cdot d\vartheta^* \end{aligned} \quad (33)$$

Hence, from Eqs. (31), (20), (32), and (33), the instantaneous combustion delay, which maximizes the engine efficiency, can be evaluated by the following relation:

$$\vartheta^*_{\text{opt}} = \frac{2\mu}{\pi(1+\mu)} \frac{w \frac{2-k}{(k-1)(\rho^{2-k} - 1)} + v_2}{\rho^{1-k} - v_1 \cdot \frac{\pi \cdot \vartheta^*}{2}} \quad (34)$$

In a SI engine, approximately one-third of the initial energy of the fuel becomes available as mechanical energy and the rest is lost, one-half in the form of heat through the cylinder surfaces (cooling losses), and one-half as waste heat through the exhaust pipe. In the analysis that leads to Eq. (34) only the expansion phase is concerned; hence only about two-thirds of the total cooling losses (which take place during compression, expansion, and exhaust phases) is considered in the evaluation of coefficient w ; this means that $w = 1/3 \times 2/3 = 2/9 \approx 0.22$ as a value of first approximation.

A great portion of the initial fuel energy fraction that becomes mechanical work (approximately one-third) is available as an output at the crankshaft and the rest is dissipated by friction or used by accessories. A reasonable value of mechanical efficiency for a SI engine, at wide open throttle (WOT), is around 80%, which means that $\text{fmep}_{\text{total}} = 0.2(\text{imep})$, where $\text{fmep}_{\text{total}}$ (the total friction mean effective pressure) accounts for the friction between piston and cylinder, the friction in the loaded bearings, and the energy needed to move the engine accessories [8].

In the present analysis, only the friction components influenced by a combustion phase shift are taken into account, i.e., the piston-cylinder and the bearing friction due to the gas pressure, while the friction due to the inertia forces of the moving masses can be neglected, so the sum of fmep_1 and fmep_2 is supposed to be equal to one-half of $\text{fmep}_{\text{total}}$ [8].

In a spark ignition engine the friction at the piston-cylinder interface is approximately four to five times the friction in the loaded bearings (crankshaft and connecting rod bearings) [8]; consequently it can be stated that

$$\text{fmep}_1 = 4 \cdot \text{fmep}_2 \Rightarrow v_1 = 4v_2 \quad (35)$$

and

$$v_1 + v_2 = 0.1$$

hence $v_1 = 0.02$ and $v_2 = 0.08$.

Assuming that $k=1.3$, $\rho=10$, $\mu=3.18$, and $w=0.22$, Eq. (34) yields the results shown in Table 1. As is shown, a 22% of heat exchange with chamber walls (with respect to the introduced heat Q_1) is responsible for an instantaneous combustion optimal delay of about 7 crank angle degrees (CAD). Moreover the delay angle ϑ^*_{opt} exhibits a linear dependence on the heat ratio $w (= Q_w / Q_1)$.

Table 1 Optimal instantaneous combustion delay ϑ^*_{opt} (from Eq. (34)) for different combinations of friction and heat transfer coefficient

v_1	v_2	w	ϑ^*_{opt} CAD ATDC
0.08	0.02	0.22	8.6
0.00	0.00	0.22	7.1
0.08	0.02	0.00	1.1
0.00	0.02	0.22	8.2
0.08	0.00	0.22	7.4
0.08	0.02	0.30	11.3
0.00	0.00	0.30	9.6
0.16	0.04	0.22	10.3
0.16	0.04	0.00	2.3

As regards friction effect, a 10% loss (with respect to imep) produces a combustion optimal delay of about 1 CAD, and once more the delay angle ϑ^*_{opt} is linearly related to the sum (v_1+v_2). The two phenomena (heat transfer and friction losses) slightly influence each other (as in a real engine), even if their effects do not add. Moreover the friction losses at the piston-cylinder interface (coefficient v_1), although greater than the losses in the loaded bearings (coefficient v_2), have a minor influence on the optimal delay ϑ^*_{opt} (as can be deduced, for example, by rows 4 and 5 in Table 1). This can be explained considering that the friction forces between piston and cylinder approach zero near TDC, so the f_{mep_1} is nearly unaffected by an instantaneous combustion phase shift in the close proximity of TDC.

5 Evaluation of the Optimal Combustion Phase by Means of Thermodynamic Simulations

In the previous paragraph, the authors showed that if the heat release function has a symmetrical derivative the optimal combustion phase in an adiabatic engine without friction losses is symmetrical with respect to the TDC (i.e., the middle of the combustion angle must be located at the TDC). The authors also arrived to a simple expression, Eq. (34), for the evaluation of the optimal instantaneous combustion phase in a real engine, that is, in the presence of heat exchanges with chamber walls and friction losses; the equation found demonstrated a good agreement with the well known experimental evidence according to which the maximum bmep is reached when the 50% of heat released by combustion is located approximately 8 CAD ATDC, a part from engine speed, load, and mixture strength [4].

If the combustion is not considered instantaneous and its heat release rate is an asymmetrical function with respect to the middle of the combustion arc, then the optimal combustion phase can be evaluated by means of thermodynamic simulations of the compression-expansion process, using heat transfer and friction loss models (e.g., as shown in Refs. [8,9]). In this kind of simulations the start and the duration of the combustion can be set as input variables and both imep and bmep can be evaluated as output quantities.

A zero-dimensional thermodynamic model based on the first law of thermodynamics (Eqs. (16) and (17)) has been realized to predict the in-cylinder pressure and temperature during the compression and expansion phases of the engine (i.e., the constant mass phases).

The first law of thermodynamics, applied to the constant mass of gas in the cylinder, yields

$$dQ_{in} - dQ_w - pdv = du = c_v \cdot dT \quad (36)$$

where c_v is the gas specific heat at constant volume.

The gas here considered (a stoichiometric mixture of C_8H_{16} and air) is supposed to respect the perfect gas law:

$$pv = RT$$

Table 2 Engine dimensions used in the simulation

Compression ratio, ρ	10
Rod to crank ratio, μ	3.18
Cylinder bore, B (mm)	79.5
Piston stroke, s (mm)	80.5

$$k = \frac{c_p}{c_v} \Rightarrow R = c_v(k - 1)$$

where $R=274.7$ J/kg K is the gas constant and k is a function of temperature obtained by Ceviz and Kaymaz [10]:

$$k = 1.338 - 6 \times 10^{-5}T_{(K)} + 10^{-8}T_{(K)}^2$$

Wall heat transfers have been taken into account by means of the Woschni model, which allows the evaluation of the instantaneous convection heat exchange coefficient h [8,11,13] that figures in Eq. (17):

$$h = F_w \cdot B^{-0.2} p^{0.8} \left[C_1 v_p + C_2 V_0 \frac{T_r}{P_r V_r} (p - p_m) \right]^{0.8} T^{-0.55} \quad (37)$$

where h is the heat exchange coefficient ($W/m^2 K$), F_w is the gain factor (allows to fit experimental data), B is the cylinder bore (m), T is the gas temperature (K), p is the gas pressure (kPa), v_p is the mean piston speed (m/s), V_0 is the cylinder displacement, T_r , p_r , and V_r are the gas temperature, pressure, and volume at the inlet valve closure, p_m is the motored gas pressure (evaluated by means of a polytropic law), C_1 is the first model constant ($=2.28$), and C_2 is the second model constant ($=3.24 \times 10^{-3}$).

The ratio $Q_w/Q_1=w$ can be tuned by means of the gain factor F_w , and the heat released by the combustion has been assumed to follow the Wiebe function [11]

$$dQ_{in}(\vartheta) = dx(\vartheta)Q_1 \quad \text{and} \quad x(\vartheta) = 1 - e^{-ay(\vartheta)^{m+1}} \quad (38)$$

where $x(\vartheta)$ is the fraction of heat released ($\vartheta_a < \vartheta < \vartheta_b$), and $y(\vartheta) = (\vartheta - \vartheta_a) / (\vartheta_b - \vartheta_a)$ is the combustion angle fraction. $m=2$ and $a=5.33$ are the values chosen for the two constants: The latter in order to obtain an almost symmetric derivative function dx/dy with its maximum in the middle of the combustion arc; it results in $x(\vartheta_a)=0$ and $x(\vartheta_b)=0.995$.

Equations (36), (17), and (38), together with the perfect gas law, lead to the differential equation

$$\frac{dp}{d\vartheta} = -\frac{k}{V} \frac{dV}{d\vartheta} \cdot p + \frac{m}{V}(k-1) \left(\frac{dQ_{in}}{d\vartheta} - \frac{dQ_w}{d\vartheta} \right) \quad (39)$$

which, numerically solved, provide the gas pressure as a function of crank position. The engine data used in the simulation are resumed in Table 2.

The friction losses at the piston-cylinder interface (f_{mep_1}) and in the loaded bearings (f_{mep_2}) have been evaluated by means of a simplified model based on Rezek and Henein's work [15] as follows:

$$f_{mep_1}(\vartheta^*) = c_1 \cdot \int_{-180}^{+180} (1 - |\sin(\vartheta)|) \cdot p(\vartheta) \cdot |\tan(\varphi)| \cdot ds \quad (40)$$

$$f_{mep_2}(\vartheta^*) = c_2 \cdot \int_{-180}^{+180} |\cos(\vartheta)| \cdot \frac{2 \cdot p(\vartheta)}{1 + \cos(\varphi)} \cdot d\vartheta \quad (41)$$

where c_1 and c_2 are proportionality constants depending on variables not affected by the combustion phase (geometrical data and engine speed), p is the relative gas pressure ($p = |p_{gas} - 1|(\text{bar})$), ds is the piston movement related to the crank rotation $d\vartheta$, and φ is

Table 3 Optimal center combustion delay evaluated by means of the thermodynamic simulation

Combustion angle CAD	$\vartheta_{opt,1}^*$ (max imep) CAD	$\vartheta_{opt,2}^*$ (max bmep) CAD
60	6.7	7.5
70	7.1	7.9
80	7.5	8.5
Mean values	7.1	8.0

the angle between connecting rod and cylinder axis, function of the crank angle ϑ .

As regards the f_{mep_1} in Eq. (40), the product $p(\vartheta) \cdot |\tan(\varphi)|$ is proportional to the normal trust between the piston and the lateral surface of the cylinder; the lubricating condition between the two surfaces can be either mixed and boundary or hydrodynamic: When the piston is in the neighborhood of the firing TDC, lubrication is mixed or boundary (maximum friction coefficient) because the relative speed between the surfaces is small. Near ± 90 deg CA ATDC instead, the lubricating conditions become hydrodynamic; hence the friction coefficient drops down significantly [15] due to the relative speed increase that means much lower friction forces. Rezek and Henein's [15] model accounts for the two lubrication modes by means of two different correlations, while in the simplified model proposed here, both the lubrication modes are represented by means of the same term $(1 - |\sin(\vartheta)|)$, which is maximum near the TDC (only mixed and boundary lubrication, and maximum friction coefficient) and minimum at ± 90 deg CA ATDC where the friction forces are supposed to be zero.

In the same way, in Eq. (41) the term $2 \cdot p(\vartheta) / (1 + \cos(\varphi))$ is proportional to the mean friction force inside the loaded bearings, while the term $|\cos(\vartheta)|$ is proportional to the friction coefficient [15]. Now the trust discriminates between the two lubricating conditions because the relative speed between journals and bearings is almost constant.

The operative conditions chosen to perform a first comparison between the results obtained by the use of Eq. (34) and those obtained by simulations are manifold absolute pressure (MAP) = 1 bar, engine speed = 2000 rpm, and chamber wall temperature = 200 °C. In these conditions w has been set to 0.22 (by means of the gain factor F_w), while v_1 and v_2 , by means of the two constants c_1 and c_2 , have been set in order to respect the conditions of Eq. (35):

$$f_{mep_1} = 4 \cdot f_{mep_2} \quad \text{and} \quad f_{mep_1} + f_{mep_2} = 0.1 \cdot \text{imep}$$

Three different values of combustion duration ($\vartheta_b - \vartheta_a$) have been considered in the simulations, namely, 60, 70, and 80 CAD. For each combustion duration, the start of combustion (i.e., the angle ϑ_a) has been phased in order to find both the maximum imep (which takes into account only heat exchange effect) and the maximum bmep (which instead is influenced also by friction

losses). The related optimal delays of the center of the combustion arc with respect to the TDC, $\vartheta_{opt,1}^*$ and $\vartheta_{opt,2}^*$, are reported in Table 3. As can be seen, the results (mean values) are in a fine agreement with those obtained by the use of the approximated formula (34) listed in Table 1; namely, the delay due only to heat exchanges is nearly the same (i.e., 7.1 CAD, second row in Table 1), and the delay obtained taking into account also friction losses corresponds (in Table 3 $\vartheta_{opt,2}^* = 8.0$ CAD \approx 8.6 CAD reported in the first row of Table 1).

Some other simulations have been carried out, considering a fixed combustion duration of 70 CAD for different engine speeds at WOT, and the results are reported in Table 4. The values of v_1 and v_2 shown here are taken from data available in literature [8] and were used to fix the values of the two constants c_1 and c_2 in Eqs. (40) and (41), while the values of the heat ratio w descend from the simulation since the gain factor F_w has been fixed to obtain $w = 0.22$ at 2000 rpm and MAP = 1 bar.

As is reasonable, increasing engine speed causes an increase in the sum ($v_1 + v_2$) and a decrease in the heat ratio w .

In particular, the sum ($v_1 + v_2$) undergoes an increase of about 70%, and this leads to an increase in ($\vartheta_{opt,2}^* - \vartheta_{opt,1}^*$) of about 1.3 CAD according to simulations and of about 1.6 CAD using the formula; however, the delay $\vartheta_{opt,2}^*$ is almost unaffected by engine speed variations since the rising friction losses are counterbalanced by the decreasing heat transfer. This behavior is confirmed by experimental results, which show that the MBT spark timing is obtained when the 50% of MFB is located approximately 8 CAD ATDC regardless of the engine and of its operative conditions [4]. Table 4 also shows a perfect agreement between the simulation results and formula results as regards the maximum imep combustion center delay $\vartheta_{opt,1}^*$: This means that the formula in Eq. (34) adequately takes into account heat transfer effects. Concerning the bmep combustion center delay $\vartheta_{opt,2}^*$, there is a small difference between the two values. Equation (34) hence slightly overestimates the effects of friction losses.

Another series of simulations has been performed for different engine loads at constant speed (3000 rpm) and combustion duration (70 CAD), whose results are exposed in Table 5. Here the values of the friction loss constants c_1 and c_2 have been chosen so as to maintain coefficients v_1 and v_2 almost unchanged since it is assumed that with decreasing manifold pressure, both f_{mep_1} and f_{mep_2} reduce in the same percentage of imep so that the ratios v_1 and v_2 are poorly influenced. On the other hand the coefficient w (calibrated by means of the gain factor F_w at 2000 rpm at WOT) increases about 16%, thus causing an angle $\vartheta_{opt,1}^*$ increase of about 1 CAD, confirmed by the use of the formula. Again, as regards the optimal combustion phase $\vartheta_{opt,1}^*$, a fine agreement was found between the simulations outputs and formula results, whereas persists the small difference of about 1 CAD in the evaluation of the combustion phase for the maximum bmep $\vartheta_{opt,2}^*$.

6 Conclusions

This paper deals with the analytical determination of the best combustion phase in a spark ignition engine. The authors first

Table 4 Optimal combustion delay angles obtained at WOT for different engine speeds (combustion arc = 70 deg CA)

Engine speed (rpm)	MAP (bar)	$w (Q_w/Q_1)$	$v_1 + v_2$	v_1/v_2	$\vartheta_{opt,1}^*$ (CAD) (simulation)	$\vartheta_{opt,2}^*$ (CAD) (simulation)	$\vartheta_{opt,1}^*$ (CAD) (formula)	$\vartheta_{opt,2}^*$ (CAD) (formula)
2000	1	0.220	0.100	4.00	7.1	7.9	7.1	8.6
3000	1	0.200	0.116	3.14	6.3	7.5	6.4	8.3
4000	1	0.187	0.132	3.12	6.0	7.3	6.0	8.1
5000	1	0.177	0.152	2.80	5.8	7.3	5.7	8.3
6000	1	0.170	0.172	2.58	5.5	7.6	5.5	8.6

Table 5 Optimal combustion delay angles obtained for different loads at 3000 rpm (combustion duration=70 deg CA)

Engine speed (rpm)	MAP (bar)	$w (Q_w/Q_1)$	v_1+v_2	v_1/v_2	$\vartheta_{opt,1}^*$ (CAD) (simulation)	$\vartheta_{opt,2}^*$ (CAD) (simulation)	$\vartheta_{opt,1}^*$ (CAD) (formula)	$\vartheta_{opt,2}^*$ (CAD) (formula)
3000	1.0	0.200	0.116	3.14	6.3	7.5	6.4	8.3
3000	0.8	0.208	0.115	3.10	6.6	7.8	6.7	8.6
3000	0.6	0.218	0.113	3.07	7.1	8.2	7.0	8.9
3000	0.4	0.232	0.110	3.01	7.2	8.2	7.4	9.3

analyzed the optimal phasing of the heat release in an ideal adiabatic engine, showing that the best phase of the combustion depends on the “shape” of the heat release rate law: If this is symmetrical with respect to its maximum, then the best spark timing is obtained when the middle point of the combustion arc is located at the TDC. The case of a not symmetrical heat release rate function is also considered and solved in the Appendix. The authors then took into consideration both the effects of friction losses and heat transfer to the wall on the optimal phasing of the instantaneous combustion, thus arriving to a formula for the calculus of the best combustion phase. The results obtained by the use of this formula under different operative conditions were then compared with the results obtained by means of zero-dimensional thermodynamic simulations performed using the Wiebe function to represent the heat release law (three different combustion durations were considered). A good agreement was found in any of the case considered.

Heat exchanges with the chamber walls, which are extremely asymmetrical and occur mainly during the expansion stroke because of the higher gas temperatures, were observed to cause a mean delay of the optimal combustion phase of the order of 6 CAD with respect to the adiabatic engine.

Friction losses due to the gas pressure are also asymmetrical and occur mainly during the expansion stroke because of the higher gas pressure. They produce a smaller delay, about 1 CAD, of the optimal combustion phase position with respect to the frictionless engine.

Nomenclature

Q = whole heat received by the gas
 L = mechanical work made on the gas
 U = internal energy of the gas
 Q_1 = whole combustion specific released by combustion
 c_p = constant pressure specific heat of the gas
 c_v = constant volume specific heat of the gas
 ϑ = generic crank angle
 ϑ_a = crank angle corresponding to combustion start
 ϑ_b = crank angle corresponding to combustion end
 ϑ_c = combustion arc
 ϑ^* = instantaneous combustion delay with respect to the TDC
 ϑ_{opt}^* = optimal instantaneous combustion delay with respect to the TDC
 Q_{in} or $Q_{in}(\vartheta)$ = specific heat released by combustion at the crank position ϑ
 T = gas temperature
 T_w = chamber wall temperature
 $V(\vartheta)$ = combustion chamber (function of the crank position ϑ)
 m = constant mass of the gas in the cylinder
 v = gas specific volume
 p = gas pressure
 S = gas specific entropy
 R = gas constant

k = isentropic coefficient
 c_p = constant pressure specific heat of the gas
 Q_2 = heat subtracted from the gas, at constant volume, as needed to close the thermodynamic cycle
 $x(\vartheta)$ = heat fraction released, during the crank rotation from ϑ_a to ϑ , with respect to the total $Q_1=Q_{in}(\vartheta_b)$
 y = $(\vartheta - \vartheta_a)/\vartheta_c$ fraction of crank rotation angle
 ρ or ρ_0 = volumetric compression ratio of the engine
 η = thermodynamic cycle efficiency
 V_0 = engine displacement
 μ = rod to crank ratio
 Q_w = heat subtracted to the gas by the chamber walls
 $S'(\vartheta)$ = surface of the combustion chamber
 h = convection heat exchange coefficient
 Q_{2total} = the whole heat subtracted to the gas, sum of the Q_2 and the Q_w $\zeta = Tv^{k-1}$
 f_{mep} = friction mean effective pressure
 f_{mep1} = friction mean effective pressure due to the friction between piston and cylinder
 f_{mep2} = friction mean effective pressure due to the friction in the loaded bearings
 $f_{mep_{total}}$ = $f_{mep1} + f_{mep2}$
 i_{mep} = indicated mean effective pressure
 T_{TDC} = gas temperature at TDC
 T_{EC} = gas temperature at the end of combustion
 V_{TDC} = cylinder volume at TDC
 p_{TDC} = gas pressure at TDC
 φ = angle between connecting rod and cylinder axis
 φ^* = the angle φ when $\vartheta = \vartheta^*$
 w = Q_w/Q_1
 v_1 = f_{mep1}/i_{mep}
 v_2 = f_{mep2}/i_{mep}
 B = cylinder bore
 u_m = mean gas velocity
 s = piston stroke
 n = engine speed
 ϑ_1^* = optimum instantaneous combustion delay considering heat exchanges
 ϑ_2^* = optimum instantaneous combustion delay considering both heat exchanges and friction losses

Appendix

The determination of the optimal combustion phase for the adiabatic and frictionless engine consists in the minimization of the entropy variation ΔS_{ab} of Eq. (4):

$$\begin{aligned}\Delta S_{ab} &= \int_a^b \left[c_v \frac{dT}{T} + R \frac{dv}{v} \right] = c_v \left[\ln \frac{T_b}{T_a} + \frac{R}{c_v} \ln \frac{v_b}{v_a} \right] = c_v \ln \frac{T_b v_b^{k-1}}{T_a v_a^{k-1}} \\ &= c_v \ln \frac{\zeta_b}{\zeta_a} = c_v \ln \left[1 + \frac{1}{c_v T_a v_a^{k-1}} \int_{\vartheta_a}^{\vartheta_b} v^{k-1} \frac{dQ_{in}}{d\vartheta} d\vartheta \right] \\ &= c_v \ln \left[1 + \frac{1}{c_v T_M v_M^{k-1}} \int_{\vartheta_a}^{\vartheta_b} v^{k-1} \frac{dQ_{in}}{d\vartheta} d\vartheta \right] \quad (A1)\end{aligned}$$

where T_M and v_M represent the gas temperature and specific volume at the BDC; hence Eq. (A1) becomes

$$\Delta S_{ab} = c_v \ln \left[1 + \frac{1}{c_v T_{BDC} \rho_0^{k-1}} \int_{\vartheta_a}^{\vartheta_b} \left(\frac{v}{v_{TDC}} \right)^{k-1} \frac{dQ_{in}}{d\vartheta} d\vartheta \right] \quad (A2)$$

Now, if the combustion is phased so as to put its middle point at the TDC (i.e., $\vartheta_a = -0.5\vartheta_c$ and $\vartheta_b = +0.5\vartheta_c$) then the determination of the optimal combustion phase with respect to piston motion (hence to in-cylinder volume) merely consists in the minimization of the following integral:

$$\int_{\vartheta_a}^{\vartheta_b} \left(\frac{V(\vartheta')}{V_{TDC}} \right)^{k-1} \frac{dQ_{in}}{d\vartheta} d\vartheta \quad (A3)$$

where $\vartheta' = \vartheta - \lambda$, with λ the phase difference between the combustion middle point ($\vartheta=0$) and the TDC position ($\vartheta'=0$). If the following functions f , q , and F are taken into consideration,

$$\begin{aligned}f &= \left(\frac{V(\vartheta')}{V_{TDC}} \right)^{k-1} = \left(\frac{V(\vartheta - \lambda)}{V_{TDC}} \right)^{k-1} = f(\vartheta; \lambda) \\ q &= \frac{dQ_{in}}{d\vartheta} = \text{const} \cdot \frac{dx(\vartheta)}{d\vartheta} = q(\vartheta) \quad (A4)\end{aligned}$$

$$F = f \cdot q = F(\vartheta; \lambda)$$

the integral to be minimized in Eq. (A3) becomes

$$\int_{\vartheta_a}^{\vartheta_b} f q d\vartheta = \int_{\vartheta_a}^{\vartheta_b} F d\vartheta = \int_{\vartheta_a}^{\vartheta_b} F(\vartheta; \lambda) d\vartheta \quad (A5)$$

According to the Leibniz rule [16],

$$\frac{\delta}{\delta \lambda} \left(\int_{\vartheta_a}^{\vartheta_b} F(\vartheta; \lambda) d\vartheta \right) = \int_{\vartheta_a}^{\vartheta_b} \frac{\delta F}{\delta \lambda} d\vartheta \quad (A6)$$

which, being the function q independent of the variable λ , becomes

$$\int_{\vartheta_a}^{\vartheta_b} \frac{\delta F}{\delta \lambda} d\vartheta = \int_{\vartheta_a}^{\vartheta_b} \frac{\delta(f q)}{\delta \lambda} d\vartheta = \int_{\vartheta_a}^{\vartheta_b} q \frac{\delta f}{\delta \lambda} d\vartheta = 0 \quad (A7)$$

From Eq. (A7) it follows the important conclusion that if the function q (i.e., the heat release rate) is symmetrical with respect to the middle of the combustion arc, the optimal combustion phase is obtained when $\lambda=0$, i.e., when the combustion is centered around the TDC, because the derivative $\delta f / \delta \lambda$ is antisymmetric with respect to the TDC.

In the more general case of function q not symmetrical with respect to its middle point, the integral in Eq. (A7) can be solved analytically or numerically. A fast solution can be, however, obtained by means of some approximations; taking into consideration the following series expansions $\cos \vartheta \approx 1 - 0.5\vartheta^2$ and $(1 \pm \varepsilon)^a \approx 1 \pm a\varepsilon$, the in-cylinder volume can be expressed as $V(\vartheta')/V_{TDC} \approx 1 - C \cdot \vartheta'^2 = 1 + \varepsilon$ and the function f becomes

$$\begin{aligned}f &= \left(\frac{V}{V_{TDC}} \right)^{k-1} \approx 1 + (k-1) \cdot \left(\frac{V}{V_{TDC}} - 1 \right) \approx 1 + C' \cdot \vartheta'^2 \\ &= 1 + C'(\vartheta - \lambda)^2 \Rightarrow \frac{\delta f}{\delta \lambda} \approx -2 \cdot C' \cdot \vartheta' = -2 \cdot C'(\vartheta - \lambda)\end{aligned} \quad (A8)$$

where C and C' are constants.

Then the integral in Eq. (A7) becomes

$$\begin{aligned}\int_{\vartheta_a}^{\vartheta_b} q \frac{\delta f}{\delta \lambda} d\vartheta &\approx \int_{-0.5\vartheta_c}^{0.5\vartheta_c} q C' \cdot (\vartheta - \lambda) d\vartheta = 0 \\ &\Rightarrow \int_{-0.5\vartheta_c}^{0.5\vartheta_c} (\vartheta - \lambda) \frac{dx}{d\vartheta} d\vartheta = 0 \quad (A9)\end{aligned}$$

which gives the optimal phase lag λ :

$$\lambda = \frac{\int_{-0.5\vartheta_c}^{0.5\vartheta_c} \vartheta \frac{dx}{d\vartheta} d\vartheta}{\int_{-0.5\vartheta_c}^{0.5\vartheta_c} \frac{dx}{d\vartheta} d\vartheta} = \vartheta_g \quad (A10)$$

Equation (A10) shows that the phase lag λ coincides with the centroid ϑ_g of the area subtended by the function $dx/d\vartheta$. As example, if the Wiebe function (11) represents the heat release rate, the optimum combustion phase will then be obtained with a slight advance with respect to the symmetric position.

Up to now, the entropy variation during the combustion phase has been evaluated supposing the gas specific heat c_v to remain constant; it is instead a function of the gas temperature and composition and may change appreciably, thus introducing a further asymmetrical element in the evaluation of the minimum ΔS_{ab} combustion phase.

Considering the following temperature dependence $c_v = B + A \cdot T$ (where A and $B = c_{v0} - T_0 A$ are two constants) and $k_0 = (c_{v0} + R')/c_{v0}$, Eq. (1) becomes

$$dQ_{in} = T dS = c_v dT + p dv = B dT + T \left(R \frac{dv}{v} + AdT \right) \quad (A11)$$

Assuming that $\varphi(\vartheta) = (R/B)(1/v)(dv/d\vartheta) + (A/B)(dT/d\vartheta)$ and $\psi(\vartheta) = -(1/B)(dQ_{in}/d\vartheta)$, Eq. (2) and its general solution remain valid, Eq. (3) hence becomes

$$T v^{k_0-1} e^{AT/B} = T_a v_a^{k_0-1} e^{AT_a/B} + \frac{1}{B} \int_{\vartheta_a}^{\vartheta} v^{k_0-1} e^{AT/B} \frac{dQ_{in}}{d\vartheta} d\vartheta \quad (A12)$$

and the entropy variation across the combustion phase is now

$$\begin{aligned}\Delta S_{ab} &= B \ln \left[1 + \frac{1}{B T_a v_a^{k_0-1} e^{AT_a/B}} \int_{\vartheta_a}^{\vartheta_b} v^{k_0-1} e^{AT/B} \frac{dQ_{in}}{d\vartheta} d\vartheta \right] \\ &= B \ln \left[1 + \frac{1}{B T_M v_M^{k_0-1}} \int_{\vartheta_a}^{\vartheta_b} v^{k_0-1} e^{A(T-Ta)/B} \frac{dQ_{in}}{d\vartheta} d\vartheta \right] \quad (A13)\end{aligned}$$

As shown, it mainly differs from Eq. (4) for the function $e^{A(T-Ta)/B}$ inside the integral, which is a rising function of the temperature difference $(T - T_a)$, which, in turn, increases when the combustion is advanced (see, for example, Fig. 1, the advanced combustion AB compared with the retarded one CD): That is the reason why, when the c_v variation with temperature is considered, the optimal combustion phase moves in retard of about 1 CAD, thus reducing the optimal spark advance.

References

- [1] Pestana, G. W., 1989, "Engine Control Methods Using Combustion Pressure Feedback," SAE Paper No. 890758.
- [2] Plint, M., and Martyr, A., 1999, *Engine Testing—Theory and Practice*, 2nd ed., Butterworth & Heinemann, Oxford.
- [3] Eriksson, L., 1999, "Spark Advance Modeling and Control," Ph.D. thesis, Linköping University, Linköping, Sweden, www.vehicular.isy.liu.se.
- [4] Pipitone, E., 2008, "A Comparison Between Combustion Phase Indicators for Optimal Spark Timing," *ASME J. Eng. Gas Turbines Power*, **130**(5), p. 052808.
- [5] Powell, J. D., 1993, "Engine Control Using Cylinder Pressure: Past, Present and Future," *ASME J. Dyn. Syst., Meas., Control*, **115**, pp. 343–350.
- [6] Bargende, M., 1995, "Most Optimal Location of 50% Mass Fraction Burned and Automatic Knock Detection Components for Automatic Optimization of SI-Engine Calibrations," *Motortech. Z.*, **56**, pp. 632–638.
- [7] Matekunas, F. A., Battiston, P. A., Chang, C. F., Sellnau, M. C., and Lancaster, D. R., 2000, "Cylinder-Pressure-Based Engine Control Using Pressure-Ratio Management and Low-Cost Non-Intrusive Cylinder Pressure Sensor," SAE Paper No. 2000-01-0932.
- [8] Ferguson, C. R., and Kirkpatrick, A. T., 2001, *Internal Combustion Engines Applied Thermosciences*, Wiley, New York.
- [9] Ramos, J. I., 1989, *Internal Combustion Engine Modelling*, Hemisphere, New York.
- [10] Ceviz, M. A., and Kaymaz, I., 2005, "Temperature and Air–Fuel Ratio Dependent Specific Heat Ratio Functions for Lean Burned and Unburned Mixture," *Energy Convers. Manage.*, **46**, pp. 2387–2404.
- [11] Heywood, J. B., 1988, *Internal Combustion Engines Fundamentals*, McGraw-Hill Automotive Technology Series, McGraw-Hill, New York.
- [12] Heywood, J. B., 1994, "Combustion and Its Modeling in Spark-Ignition Engines," International Symposium COMODIA 94, Yokohama, Japan, Jul.
- [13] Finol, C. A., and Robinson, K., 2006, "Thermal Modelling of Modern Engines: A Review of Empirical Correlations to Estimate the In-Cylinder Heat Transfer Coefficient," *Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.)*, **220**, pp. 1765–1781.
- [14] Horlock, J. H., and Winterbone, D. E., 1986, *The Thermodynamics and Gas Dynamics of Internal Combustion Engines*, Vol. II, Clarendon, Oxford.
- [15] Rezeka, S. F., and Henein, N. A., 1984, "A New Approach to Evaluate Instantaneous Friction and Its Components in Internal Combustion Engines," International Congress and Exposition, Detroit, MI, Feb. 27–Mar. 2, SAE Paper No. 840179.
- [16] Korn, G. A., and Korn, T. M., 1968, *Mathematical Handbook for Scientists and Engineers*, McGraw-Hill, New York.

Simulating the Concentration Equations and the Gas-Wall Interface for One-Dimensional Based Diesel Particulate Filter Models

Christopher Depcik

Department of Mechanical Engineering,
The University of Kansas,
3120 Learned Hall, 1530 West 15th Street,
Lawrence, KS 66045-7609
e-mail: depcik@ku.edu

This paper enhances an earlier publication by including the concentration equations of motion into the area-conserved one-dimensional based diesel particulate filter model. A brief historical review of the species equations is accomplished to describe this model and the pertinent physics involved. In the species equations through the wall and soot layers, the diffusion constants are modified to account for the close proximity of the porous walls and the particulate matter to the gas flowing through the accompanying layers. In addition, a review of potential options involving the diffusion velocity is accomplished to determine the effect of pressure gradients on this phenomenon. In the previous paper, the model formulation illustrated that a common assumption to make for an enthalpy difference is the use of constant pressure specific heat times a temperature difference. Because of the different heats of formation and sensible enthalpies associated with the chemical species, this assumption reviewed is found to have a related error. Finally, because each channel is treated as an open system, making the common assumption of dilute mixture simplification is reviewed and found to have an associated error. [DOI: 10.1115/1.3155792]

1 Introduction

Diesel particulate filters (DPFs) have been researched for close to three decades based on their ability to capture soot or particulate matter (PM) with near 100% efficiency [1–4]. Because of the Environmental Protection Agency (EPA) and the European Union (EU) emission standards, DPFs have become commonplace equipment on all diesel vehicles sold for on-road applications [5,6]. In addition, future off road Tier 4 emission regulations target significant reductions in problematic species. This indicates the eventual implementation of DPFs on generators and other industrial use engines [7]. Since reductions in nitrogen oxides (NO_x), carbon monoxide (CO), and hydrocarbons (HC) among others are also being targeted, a system arrangement of multiple aftertreatment devices in the exhaust will be required. Key to the implementation of a DPF in this type of system architecture is the use of fast and accurate models to help augment and guide experimental efforts. With respect to DPF devices, one-dimensional based modeling has become the industry standard for this type of multiple apparatus modeling [8–15].

Most of the one-dimensional DPF modeling efforts stem from the papers by Bissett and co-workers [8,16,17]. In a previous paper [18], a brief historical review of this modeling activity from the original works of Bissett until now was documented, and a model was presented that formulated the governing equations in the area-conserved format often used for converging-diverging nozzles [19,20] and the intake and exhaust of internal combustion engines [21,22]. This quasi-one-dimensional flow accounted for the effects of changing gaseous area in the inlet channel and varying cross-sectional areas of flow through the PM layer. One large simplification in the previous paper was the use of constant mass

fractions of the different species through the device. With the advent of catalyzed DPFs and the importance of NO_2 reduction in PM in the cake and wall layers, there is an additional requirement of 1D based DPF models to include the flow of chemical species through these layers [11–13,23–25].

As a result, this paper solves for the concentration equation of motion through the channels, PM, and wall layer in area-conserved format. A brief historical review of the species equations in these layers is first accomplished to describe the history of this model and the pertinent physics. While area conservation is not required as per the conclusions of the previous paper, it is inherently more accurate and can be utilized in all situations of PM loading. In this paper, the diffusion constants are modified to account for the close proximity of the porous walls and the PM to the gas flowing through the accompanying layers. In addition, a review of potential options involving the diffusion velocity is accomplished to determine the effect of pressure gradients on this phenomenon. Finally, in the previous paper, the model formulation illustrated that a common assumption to make for an enthalpy difference is the use of constant pressure specific heat times a temperature difference. Because of the different heats of formation and sensible enthalpies associated with the chemical species, this assumption is reviewed and a conclusion as to the model formulation is presented.

2 Governing Equations

In 1984, Bissett [8] created a model for the transport of nearly all thermodynamic variables through a wall-flow DPF while trying to avoid the complexities of multidimensional compressible fluid flow. He accomplished this by representing the flow at each axial position in one dimension and incorporating the proper physics for the flow to and from the wall system. In his original paper, Bissett assumed that the mass fractions through the inlet channel were constant and unaffected by the changing control volume due to PM loading. In Secs. 2.1–2.6, a review of the

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received November 3, 2008; final manuscript received April 26, 2009; published online November 30, 2009. Review conducted by Christopher J. Rutland.

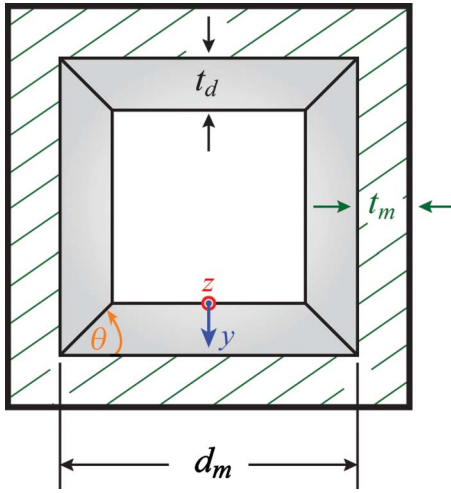


Fig. 1 Square channel schematic illustrating the important geometric and soot parameters

author's previous paper formulation is presented along with a brief historical summary, and the addition of chemical species is included in the area-conserved model.

2.1 Mass Equations. The thermodynamic properties through the PM and wall layers are allowed to change and will be calculated according to the Navier–Stokes equations of motion. As a result, the equations of mass for the channels are written using local density and velocity terms on the right-hand side of the equations as follows:

$$\frac{\partial(\rho_I u_I A_I)}{\partial z} = -\rho_s u_s S_I \quad (1)$$

$$\frac{\partial(\rho_{II} u_{II} A_{II})}{\partial z} = \rho_w u_w S_{II} \quad (2)$$

where the cross-sectional area and surface area on a per length basis are represented as

$$A_I = (d_m - 2t_d)^2, \quad S_I = 4(d_m - 2t_d), \quad A_{II} = d_m^2 \quad \text{and} \quad S_{II} = 4d_m \quad (3)$$

In this case, for the inlet channel equations the density and velocity of the gas in the PM layer used in the flux term are evaluated at $y=0$ as illustrated in Fig. 1. For the outlet channel equations, the density and velocity of the gas in the wall layer used in the flux term are evaluated at $y=t_m+t_d$.

In the PM layer, the flow is no longer one-dimensional due to the varying cross-sectional areas through the layer. Since area conservation is a methodology for incorporating multidimensional effects without having to include the other dimensions in the solution algorithm, the PM and wall layer equations are written as follows:

$$\frac{\partial(\rho_s u_s A_s)}{\partial y} = A_s \dot{S}_{\text{mass}} \quad (4)$$

$$\frac{d(\rho_w u_w A_w)}{dy} = 0 \quad (5)$$

where the right-hand side of the mass equation in the PM layer equates to a gain in mass due to PM oxidation reactions; that is, as the solid PM burns it converts into partial and complete products of combustion in gaseous form. This was previously found to be important as a significant deviation in mass flow rate can be encountered without this term. As with the previous paper, deep bed filtration is not included in this model, which would result in a

similar term in the wall equation, and is left to future efforts.

For square channels, the change in cross-sectional area per unit length through the PM layer and wall layers is calculated as

$$A_s = 4 \left[d_m - 2 \left(t_d - \frac{y}{\tan \theta} \right) \right] \quad (6)$$

where $\tan \theta = 1$.

$$A_w = 4d_m \quad (7)$$

and the instantaneous change the cross-sectional area can then be determined by taking the derivative of this value with respect to the distance through the layer, that is, $dA_s/dy=8$ and $dA_w/dy=0$.

2.2 Momentum Equations. For the momentum equations in the channels, the model for the bulk gas is the same as those used in traditional area-conserved one-dimensional flow [21,22,26].

$$\frac{\partial(\rho_I u_I^2 A_I)}{\partial z} + \frac{\partial(p_I A_I)}{\partial z} = -\frac{F\mu|u_I|}{d_m - 2t_d} S_I + p_I \frac{dA_I}{dz} \quad (8)$$

$$\frac{\partial(\rho_{II} u_{II}^2 A_{II})}{\partial z} + \frac{\partial(p_{II} A_{II})}{\partial z} = -\frac{F\mu|u_{II}|}{d_m} S_{II} + p_{II} \frac{dA_{II}}{dz} \quad (9)$$

As mentioned in the previous paper, keeping the area terms in the derivatives and solving the governing equations as indicated is more accurate with respect to mass conservation [20–22]. It is important to note that the area change term on the right-hand side is required in order to balance the pressure forces when writing in area-conserved format.

The formulation of the momentum equations through the PM and wall layers includes Darcy's law and should not include the convective term as previously discussed.

$$\frac{d(p_s A_s)}{dy} = p_s \frac{dA_s}{dy} - \frac{\mu u_s A_s}{K_s} - \beta_s \rho_s A_s u_s^2 \quad (10)$$

$$\frac{d(p_w A_w)}{dy} = p_w \frac{dA_w}{dy} - \frac{\mu u_w A_w}{K_w} - \beta_w \rho_w A_w u_w^2 \quad (11)$$

By expanding the derivative of the pressure force, the traditional equation for pressure through each layer written by Bissett will be recovered.

2.3 Energy Equations. The gas flowing through the channels transfers heat to the wall as a function of temperature difference with the filter wall and from a flow of energy to the wall. As discussed in the historical review in the previous paper, the energy equations in the channel are equal to

$$\frac{\partial(\rho_I u_I h_I A_I)}{\partial z} = -S_I [h_g(T_I - T_s) + \rho_s u_s h_I] \quad (12)$$

$$\frac{\partial(\rho_{II} u_{II} h_{II} A_{II})}{\partial z} = -S_{II} [h_g(T_{II} - T_w) - \rho_w u_w h_w] \quad (13)$$

It was found that the gas temperature in the wall quickly becomes the filter temperature. Convection within the porous membrane is at least on the order of 10^4 times larger than advection and conduction validating previous model assumptions that the gas temperature in the PM and wall layers equates rapidly to the filter temperature: $T_s = T_w = T_f$.

2.4 Species Equations. In order to model the local oxygen level at the wall along the inlet channel, Bissett assumed that oxygen transport in the porous wall is dominated by convection; hence, the depletion of oxygen in the reactive deposit layer does not influence the oxygen concentration at the upstream channel and wall interface [8]. As a result, he assumed that there is not a gradient of mass fraction in the inlet channel.

$$\frac{dY_{1,O_2}}{dz} = 0 \quad (14)$$

In Bissett's original paper, the reaction rate for the PM layer is determined from a model of the mass fraction of oxygen in the wall. This model is fundamentally expressed the following year in a paper co-authored by Bissett and Shadman [17]. To take into account the variation in oxygen across the PM and wall layers, the balance between the convective transport in the y -direction and the reaction kinetics in the particulate layer governs an O_2 species equation as follows:

$$\frac{\partial(\rho_s u_s Y_{s,O_2})}{\partial y} = -S_p k \rho_s Y_{s,O_2} \quad (15)$$

$$\frac{\partial(\rho_w u_w Y_{w,O_2})}{\partial y} = 0 \quad (16)$$

with the reaction rate expressed in Arrhenius format, $k = AT_f \exp[-E_a/(R_u T_f)]$. At this point in the literature, one-dimensional based DPF modeling consists of two synergistic models: one for the channels and one for the wall. For a while, the chemical species are modeled in depth within the wall while the constant mass fraction assumption is used for the channels.

A few years later, Garner and Dent [27] developed a model independent of Bissett applicable to both fibrous and porous wall DPFs. This model incorporates medium porosity and, more importantly, an equation for the oxygen transport to the particulate surface and its subsequent consumption.

$$\varepsilon_p \rho_1 \left(\frac{\partial Y_{1,O_2}}{\partial t} + u_1 \frac{\partial Y_{1,O_2}}{\partial z} \right) = -k \rho_w Y_{w,O_2} \quad (17)$$

This equation is reminiscent of the species equation from the Euler equations of motion that can replace Bissett's constant O_2 mole fraction assumption. However, it was never utilized along with the classical model formulation.

In 1996, Koltsakis and Stamatelos [28] revisited the Bissett and Shadman model and modified the kinetics model to account for incomplete soot oxidation as follows:

$$\frac{\partial(\rho_s u_s Y_{s,O_2})}{\partial y} = -S_p k \rho_s Y_{s,O_2} \alpha_{O_2} \quad (18)$$

where α_{O_2} is an index of completeness, which was estimated through experimental validation to be 0.8 (note that in Ref. [18] $k_{O_2} = k \alpha_{O_2}$). In this year, they utilized the same formulation for the catalytically induced regeneration of the PM by metal oxides [29].

In the following year, Jørgensen and Sorenson [30] developed a combustion model through the PM layer that includes the effects of temperature, concentrations, velocity, and diffusion. This two-film model calculates the species and temperature profiles as a function of the flame sheet where PM oxidation occurs. Species equations are written for O_2 , CO, and CO_2 ; however, the model formulation and complexity does not lend itself to merging with the classical Bissett model. The important facet to take away is the incorporation of Fick's law into the modeling of the governing species equations, but this is done as a function of a flame sheet interpretation.

In 1999, Konstandopoulos and Kostoglou [31] presented the first paper where the channel equations are shown at the same time as the species equations through the PM and wall layers (see also Ref. [32]). The merging of the models has yet to occur as the channel species equations are not present; however, further expansion of the wall species equations occurs when catalytic coating oxidation of soot is added to the model in multiple layers. The PM layer is divided into two sections with one layer (Layer I) reacting catalytically and noncatalytically through a partial fraction and the other portion (Layer II) sitting on top of the catalytic layer following the traditional thermal oxidation pathway. Catalytic addi-

tive assisted oxidation of the PM with compounds either added in the fuel or sprayed in the exhaust is also described using a factor, which can be modified as a function of catalytic additive doses. This paper is followed next year by a further inclusion of PM oxidation via NO_2 in these layers. The governing equation for NO_2 in Layer II is similar to the O_2 layer with notation revised here for consistency [33].

$$\frac{\partial(\rho_s u_s Y_{s,NO_2})}{\partial y} = -S_p k \rho_s Y_{s,NO_2} \alpha_{NO_2} \quad (19)$$

Layer I now includes the additional production of NO_2 from NO and O_2 because of the catalytic washcoat material as follows:

$$\frac{\partial(\rho_s u_s \mathbf{Y}_s)}{\partial y} = \dot{\mathbf{S}}_s + \mathbf{R}\mathbf{W} \quad (20)$$

where the bold nomenclature indicates an array of species (O_2 , NO_2), $\dot{\mathbf{S}}_s$ is an array of PM conversion in this layer (O_2 —catalytic and thermal oxidation similar to previous paper; NO_2 —thermal oxidation), and \mathbf{R} is an array that accounts for gaseous catalytic reactions via the NO oxidation reaction (O_2 —loss; NO_2 —gain). The local NO mole fraction is then obtained from nitrogen atom conservation.

It is not until the papers of Peters and co-workers [24,25] and Haralampous and co-workers [11–13,23,34] that a merging of two models is accomplished. This is done by including the wall species equations into the overall formulation instead of solving separately. Their reason for solving the wall species equations along with the channel equations is to provide the proper boundary conditions needed for the channel species equations. In the papers of Peters and co-workers [24,25], these channel equations are expressed as

$$\frac{\partial(\rho_I u_I d_I^2 \mathbf{Y}_I)}{\partial z} = -4d_I \rho_I u_s \mathbf{Y}_I \quad (21)$$

$$\frac{\partial(\rho_{II} u_{II} \mathbf{Y}_{II})}{\partial z} = \frac{4}{d_m} \rho_{II} u_w \mathbf{Y}_3 \quad (22)$$

Whereas, for the papers of Haralampous and co-workers [11–13,23,34] they are represented as

$$\frac{\partial(u_I \mathbf{Y}_I)}{\partial z} = -\frac{\bar{u}_w \mathbf{Y}_I}{d_m f^2} + \frac{k_m (\mathbf{Y}_s - \mathbf{Y}_I)}{d_m f} \quad (23)$$

$$\frac{\partial(u_{II} \mathbf{Y}_{II})}{\partial z} = \frac{\bar{u}_w \mathbf{Y}_w}{d_m f^2} + \frac{k_m (\mathbf{Y}_w - \mathbf{Y}_{II})}{d_m f} \quad (24)$$

Note that the papers of Peters and co-workers [24,25] present a variable \mathbf{Y}_3 , which is not described in their paper, and for the papers of Haralampous and co-workers [11–13,23,34], \mathbf{Y}_s are the species evaluated at $y=0$ and \mathbf{Y}_w are the species evaluated at $y = t_m + t_d$.

In the papers of Haralampous and co-workers [11–13,23,34], they included the effects of diffusing species to and from the channel to the wall similar to catalyst modeling [35]. This is because the mass transfer coefficients can be on the same order of magnitude as the wall-flow velocity. In these equations, d_I represents the hydraulic diameter of the inlet channel ($d_m - 2t_d$) and f is a geometrical parameter that accounts for the varying areas in the particulate layer. The reader is referred to the previous paper [18] for a complete description of the wall velocity terms utilized in each formulation and how the varying areas in the PM layer is simulated within this paper.

This historical summary is utilized to include all pertinent channel phenomena for the governing equations of species in area-conserved format.

$$\frac{\partial(\rho_I u_I \mathbf{Y}_I A_I)}{\partial z} = -S_{II}[k_m(\rho_I \mathbf{Y}_I - \rho_s \mathbf{Y}_s) + \rho_s u_s \mathbf{Y}_I] \quad (25)$$

$$\frac{\partial(\rho_{II} u_{II} \mathbf{Y}_{II} A_{II})}{\partial z} = -S_{II}[k_m(\rho_{II} \mathbf{Y}_{II} - \rho_w \mathbf{Y}_w) - \rho_w u_w \mathbf{Y}_w] \quad (26)$$

This use of concentration equations ensures that the equations properly conform to the mass equations presented earlier. This can be seen by summing both sides of the equations, taking into account that the sum of mass fractions equals one, and utilizing the dilute mixture simplification that is common in catalyst modeling [35]. In specific, summation of the diffusing gas species concentration difference terms is too low to affect the overall mass balance of the carrier gas: $\sum k_m(\rho_I \mathbf{Y}_I - \rho_s \mathbf{Y}_s) \approx 0$ and $\sum k_m(\rho_{II} \mathbf{Y}_{II} - \rho_w \mathbf{Y}_w) \approx 0$. This assumption will be checked later in this paper.

For the species equations through the wall, a transport-reactive coupling is utilized by both Peters and co-workers [24,25] and Haralampous and co-workers [11–13,23,34]. For simplicity, only the formulation of Haralampous and co-workers [11–13,23,34] for the soot layer is described here as follows:

$$\bar{u}_w \frac{\partial \mathbf{Y}_s}{\partial y} - \mathbf{D} \frac{\partial}{\partial y} \left(f \frac{\partial \mathbf{Y}_s}{\partial y} \right) = \dot{S}_{s,H04} \quad (27)$$

In the above equation, the first term on the left-hand side accounts for the convection of the species through the layer similar to all wall models starting with Bissett and Shadman [17]. The second term on the left-hand side accounts for diffusion through the layer. It was found that since the flow velocity is very low through the wall, diffusion of the species can be on the same order of magnitude. In fact, “back-diffusion” of NO₂ from the wall into the soot layer is prevalent and must be accounted for, especially in passive regenerated DPFs. The right-hand side of the equation accounts for PM oxidation and the effects of catalyzed walls or catalytic additive assisted oxidation (see Refs. [11–13,23,34] for its functional form).

Taking these effects into account, the area-conserved formulation of the species equation through the PM layer equals,

$$\frac{\partial(\rho_s u_s \mathbf{Y}_s A_s)}{\partial y} + \frac{\partial(\rho_s A_s \mathbf{Y}_s \mathbf{V}_s)}{\partial y} = A_s \dot{S}_s \quad (28)$$

where the source term on the right-hand side accounts for the addition or subtraction of gaseous species (such as O₂, CO, NO, etc.) from PM combustion. For example, later in this paper the loss in oxygen from the species equation incorporated here is equal to $\dot{S}_{O_2} = -S_p \rho_s Y_{s,O_2} k_{O_2}$. The second term on the left-hand side incorporates a diffusion velocity of which the calculation will be discussed in more depth below.

For the wall layer, the species equation will equal,

$$\frac{\partial(\rho_w u_w \mathbf{Y}_w A_w)}{\partial y} + \frac{\partial(\rho_w A_w \mathbf{Y}_w \mathbf{V}_w)}{\partial y} = 0 \quad (29)$$

where future work will utilize the right-hand side to account for deep bed PM oxidation and catalytic reactions as illustrated by the documented history.

These equations are subject to the following constraint:

$$\sum_{j=1}^{NM} Y_j = 1 \quad (30)$$

with the net species diffusion flux equal to zero,

$$\sum_{j=1}^{NM} V_j Y_j = 0 \quad (31)$$

The full formulation of the diffusion velocity utilizing mixture-averaged diffusion coefficients is described as [36,37]

$$\mathbf{V} = -\frac{\hat{\mathbf{D}}}{\mathbf{X}} \left[\frac{\partial \mathbf{X}}{\partial y} + \frac{(\mathbf{X} - \mathbf{Y})}{p} \frac{\partial p}{\partial y} \right] - \frac{\mathbf{D}^T}{\rho \mathbf{Y} T} \frac{\partial T}{\partial y} \quad (32)$$

Quite often the pressure derivative and temperature derivative terms are neglected. While this is true for the temperature derivative based on the conclusions from the previous paper, the pressure difference might be important. This is because the pressure change from the inlet channel to the outlet channel can be significant. Changing this equation from mole fraction to mass fraction derivatives and assuming that the molecular weight change in the y-direction is negligible (checked later) the diffusion velocity equals,

$$\mathbf{V} = -\frac{\hat{\mathbf{D}}}{\mathbf{Y}} \left[\frac{\partial \mathbf{Y}}{\partial y} + \frac{\mathbf{W} \mathbf{Y}}{p \hat{\mathbf{W}}} \left(\frac{\hat{\mathbf{W}}}{\mathbf{W}} - 1 \right) \frac{\partial p}{\partial y} \right] \quad (33)$$

Note that if the pressure derivative term was omitted, incorporation into the PM and wall species equations would result in the same formulation as that of Haralampous and co-workers [11–13,23,34] but on a concentration instead of mass fraction basis.

The diffusion velocity utilized here must take into account the porosity of the wall because of its close proximity. As a result, the effective diffusivity in the PM and wall can be calculated based on the values for the Knudsen and bulk diffusivities along with the tortuosity of the flow [38–42].

$$\frac{1}{\hat{\mathbf{D}}} = \frac{1}{\varepsilon_p} \left(\frac{\tau_{Kn}}{\mathbf{D}_{Kn}} + \frac{\tau_b}{\mathbf{D}_b} \right) \quad (34)$$

The binary diffusion constants are computed as

$$\mathbf{D}_b = \frac{1 - Y_j}{\sum_{s \neq j}^{NM} X_j / D_{sj}} \quad (35)$$

where the Knudsen and bulk tortuosities can be assumed to depend only on the porosity,

$$\tau_{Kn} = \frac{9}{8} - \frac{1}{2} \ln \varepsilon_p + \left(\frac{13}{9} - \frac{9}{8} \right) \varepsilon_p^{2/5} \quad \text{and} \quad \tau_b = 1 - \frac{1}{2} \ln \varepsilon_p \quad (36)$$

and the Knudsen diffusivity for each species is calculated as

$$\mathbf{D}_{Kn} = \frac{d_p}{3} \sqrt{\frac{8 R_u T_f}{\pi \mathbf{W}}} \quad (37)$$

2.5 Surface Particulate Mass Equation. When particulates flow into a DPF, they first load within the wall layer where the pores are small enough to capture the PM by interception via a deep bed filtration. After the wall has filled, the particulate then loads upon the surface in a growing cake layer. Previous efforts in DPF modeling have described these two phases of storage [14,33,43–45], however for simplicity, only the cake layer is modeled here with the deep bed filtration inclusion noted as future work.

In the previous paper, an assumption of a constant mass fraction of PM within the inlet channel was utilized as follows:

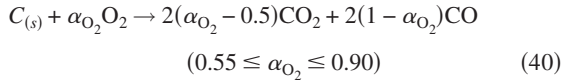
$$\frac{dY_d}{dz} = 0 \quad (38)$$

With the added effects of mass transfer within the inlet channel (Eq. (25)), this assumption may no longer be valid. However, since this mass transfer is written as a function of a concentration difference, a proper adaptation of this equation involves solving for the flow of PM within the different layers similar to the species equations. This would then allow for the calculation of the PM concentration at the interface of the inlet channel and wall or cake layer. For the purposes of this paper, this effect will be neglected as the gaseous species equations are the current focus.

This will be revisited in the future work involving deep bed filtration. As a result, a PM mass flow rate per mesh interval to the surface of inlet channel can be determined as a function of the changing surface area.

$$s = \rho_s u_s S_f Y_d \quad (39)$$

The molar combustion reaction is often written including partial oxidation factors [46–48].



indicating incomplete combustion with carbon monoxide produced during the oxidation process. PM oxidation via O_2 now utilizes an average mass fraction of oxygen within the layer, whereas before the mass fraction of oxygen was constant throughout the entire DPF.

$$\chi = - \frac{S_p \hat{\rho}_s \hat{Y}_{s,O_2} k_{O_2} W_{C_{(s)}}}{\alpha_{O_2} W_{O_2} \rho_d} \quad (41)$$

In this equation, the Arrhenius reaction rate term is equal to $k_{O_2} = A_{O_2} T_f \exp[-E_{O_2}/(R_u T_f)]$.

While the local oxygen level does vary through the PM layer, since the equation for PM mass per unit length on the surface is written over the entire layer,

$$\frac{dm_d}{dt} = s - \chi m_d \quad (42)$$

using the average oxygen concentration in the layer is the correct approach. If the PM was discretized in the y -direction, then the local oxygen concentration should be utilized. These average values are calculated by summing over the layer and then dividing by the number of discretizations. The thickness of PM on the surface is calculated from geometric principles from this mass [33].

$$t_d = \frac{1}{2} \left(d_m - \sqrt{d_m^2 - \frac{m_d}{\rho_d}} \right) \quad (43)$$

Note that the mass on the surface may vary in each numerical cell, but using the above equation normalized per length ensures that the thickness remains consistent.

For the mass equation in the PM layer (Eq. (4)), the source term stemming from PM oxidation has to be determined. Since the mass added to the gas equals the mass of PM that has been combusted, over the entire layer this source term equals,

$$\dot{S}_{C_{(s)}} = - \frac{S_p \hat{\rho}_s \hat{Y}_{s,O_2} k_{O_2} W_{C_{(s)}}}{\alpha_{O_2} W_{O_2}} \quad (44)$$

In order to use it within the mass equation, it is split by the number of discretizations in the soot layer.

$$\dot{S}_{mass} = \frac{\dot{S}_{C_{(s)}}}{N_s - 1} \quad (45)$$

While not explicitly derived, the inclusion of soot oxidation via the NO_2 pathway follows similar to the O_2 derivation presented.

For incorporation in the PM layer species equation (Eq. (28)), the O_2 , CO , and CO_2 reaction rate terms are equal to

$$\dot{S}_{O_2} = - S_p \hat{\rho}_s \hat{Y}_{s,O_2} k_{O_2} \quad (46)$$

$$\dot{S}_{CO} = - \frac{2(1 - \alpha_{O_2}) W_{CO}}{\alpha_{O_2} W_{O_2}} \dot{S}_{O_2} \quad (47)$$

$$\dot{S}_{CO_2} = - \frac{2(\alpha_{O_2} - 0.5) W_{CO_2}}{\alpha_{O_2} W_{O_2}} \dot{S}_{O_2} \quad (48)$$

The use of molar ratios via combustion reaction (Eq. (41)) and molecular weights allows the proper conversion into grams of species per unit volume and time.

2.6 Filter Energy Equation. The current version of the filter energy equation can now be seen as variants of the following full description [13,25,44,49]:

$$(\rho_f c_f V_f + \rho_d c_d V_d) \frac{\partial T_f}{\partial t} = \dot{Q}_{cond} + \dot{Q}_{conv} + \dot{Q}_{wall} + \dot{Q}_{react} + \dot{Q}_{supply} + \dot{Q}_{multi} + \dot{Q}_{rad} \quad (49)$$

where the effects of conduction (\dot{Q}_{cond}), convection (\dot{Q}_{conv}), energy flow (\dot{Q}_{wall}), PM oxidation (\dot{Q}_{react}), active heating (\dot{Q}_{supply}), multidimensional effects (\dot{Q}_{multi}), and radiation (\dot{Q}_{rad}) can be important. In this paper, the DPF modeled does not incorporate any active heating and is simulated in one dimension ignoring the effects of radiation. It is important to note that this equation sets the time-step of the simulation when running transient tests because of the incorporated time-derivative.

Since the soot loading does not vary dramatically from node-to-node in the axial direction, the heat transfer due to conduction often removes this axial dependence.

$$\dot{Q}_{cond} = (\lambda_f V_f + \lambda_d V_d) \frac{\partial^2 T_f}{\partial z^2} \quad (50)$$

Because the gas in the inlet channel sees a different surface area than the outlet channel, the model must account for these varying areas in the heat transfer due to heat transfer and energy transfer as follows:

$$\dot{Q}_{conv} = \left(\frac{N_c}{2} \right) [h_g S_I (T_I - T_s) + h_g S_{II} (T_{II} - T_w)] \quad (51)$$

$$\dot{Q}_{wall} = \left(\frac{N_c}{2} \right) (\rho_s u_s h_f S_I - \rho_w u_w h_w S_{II}) \quad (52)$$

Note that the heat transfer terms are now a function of the local gas values at the interface of the soot and inlet channels and the wall and outlet channels, respectively. In the previous papers referenced, the energy flow through the wall is often written as a function of constant pressure specific heat times a temperature difference, i.e., $h_I - h_w = c_p (T_I - T_w)$. Proper inclusion requires the use of enthalpies because chemical species will have different heats of formation and sensible components. The effect of this difference will be shown later in this paper.

Finally, the contributions due to combustion are a function of the soot oxidation as described in Sec. 2.5:

$$\dot{Q}_{react} = - \left(\frac{N_c}{2} \right) \left(\frac{m_d \Delta H_{react}}{\rho_d} \right) \dot{S}_{C_{(s)}} \quad (53)$$

where oxidation only occurs on half the number of total channels and includes the heat of reaction, which is a function of the combustion reaction modeled (Eq. (41)).

3 Boundary Conditions

At the interface between the inlet channel and PM layer, the flow of species into the wall must be balanced by the flow of the species leaving the inlet channel. In other words, the mass of each species needs to be conserved.

$$\dot{C}_j|_{in} = \dot{C}_j|_{out} \quad (54)$$

Following the methodology of Haralampous and co-workers [11–13,23,34], this balance of mass transfer, flow, and diffusion

can be written utilizing the fact that the area is the same between the inlet channel gas and the PM or porous wall.

$$k_m(\rho_I \mathbf{Y}_I - \rho_s \mathbf{Y}_s) + \rho_s u_s \mathbf{Y}_I = \rho_s u_s \mathbf{Y}_s + \rho_s \mathbf{V}_s \mathbf{Y}_s \quad (55)$$

Unlike the papers of Haralampous and co-workers [11–13,23,34], it is written on a concentration basis taking into account local densities. This ensures that the mass of each species is conserved instead of mass fraction; for example, $\rho_I Y_{1,O_2}$ equals the mass of oxygen in the inlet channel. Because of the potential for a different densities at the interface ($y=0$), conserving mass fractions would not necessarily conserve the mass of each species. Incorporating the diffusion velocity of Eq. (33), the differential of mass fraction of the species can then be written at the following interface:

$$\left. \frac{\partial \mathbf{Y}_s}{\partial y} \right|_{y=0} = - \frac{k_m(\rho_I \mathbf{Y}_I - \rho_s \mathbf{Y}_s) + \rho_s u_s (\mathbf{Y}_I - \mathbf{Y}_s)}{\rho_s \hat{\mathbf{D}}} - \frac{\mathbf{W}_s \mathbf{Y}_s}{\rho_s \hat{W}_s} \left(\frac{\hat{W}}{\mathbf{W}_s} - 1 \right) \frac{\partial p_s}{\partial y} \quad (56)$$

This can be solved implicitly for the mass fractions at the interface once Darcy's law is used to replace the pressure derivative in the last term on the right.

For the interface between the outlet channel and the wall, the balance of mass transfer, flow, and diffusion equals,

$$\rho_w u_w \mathbf{Y}_w + \rho_w \mathbf{V}_w \mathbf{Y}_w = -k_m(\rho_{II} \mathbf{Y}_{II} - \rho_w \mathbf{Y}_w) + \rho_w u_w \mathbf{Y}_w \quad (57)$$

Again by incorporating the diffusion velocity, the local derivative of mass fraction at the interface equals

$$\left. \frac{\partial \mathbf{Y}_w}{\partial y} \right|_{y=m+t_d} = \frac{k_m}{\rho_w \hat{\mathbf{D}}} (\rho_{II} \mathbf{Y}_{II} - \rho_w \mathbf{Y}_w) - \frac{\mathbf{W}_w \mathbf{Y}_w}{\rho_w \hat{W}_w} \left(\frac{\hat{W}}{\mathbf{W}_w} - 1 \right) \frac{\partial p_w}{\partial y} \quad (58)$$

To calculate the mass fraction of species at the left wall of the outlet channel, a method of characteristics (MOCs) analysis from fluid dynamics needs to be completed. Since the velocity at the wall is equal to zero, the mass fractions become a function of the flow from the wall and the left running wave ($u-a$) from the neighboring cell. From previous efforts [35,50], the following governing equation at the wall can be written:

$$\left. \frac{\partial \mathbf{Y}_{II}}{\partial t} \right|_{z=0} = \Delta_4 \quad (59)$$

To calculate the steady-state solution, the left-hand side of the equation is set equal to zero with the following components:

$$\Delta_4 = \frac{Z_j - \mathbf{Y}_{II} Z_{\text{mass}}}{\rho_{II}} = 0 \quad (60)$$

Note that all channel equations do not include time-derivatives because these terms are relatively small in comparison to the convective terms, often found in catalyst modeling [35,51].

From the DPF governing equations, the terms in the above equation are the following:

$$Z_{\text{mass}} = \frac{\rho_w u_w S_{II}}{A_{II}} \quad (61)$$

$$Z_j = \frac{S_{II}}{A_{II}} [k_m(\rho_w \mathbf{Y}_w - \rho_{II} \mathbf{Y}_{II}) + \rho_w u_w \mathbf{Y}_w] \quad (62)$$

Hence, the mass fractions of chemical species at the wall equal,

$$\mathbf{Y}_{II} = \rho_w \mathbf{Y}_w \frac{k_m + u_w}{\rho_{II} k_m + \rho_w u_w} \quad (63)$$

In the case where the dilute mixture simplification is not included, an additional term needs to be included, which equates to the sum of the mass transfer term across all species.

4 Numerical Methods

For the inlet and outlet channels, there exists a differential-algebraic system of equations. Similar to Bissett's work, the numerical code LSODI [52] can be used to solve for the resulting thermodynamic properties. Since the species equations were not included in that paper, they are solved here by expanding out the derivative in Eq. (25):

$$\frac{\partial(\rho_I u_I \mathbf{Y}_I A_I)}{\partial z} = \mathbf{Y}_I \frac{\partial(\rho_I u_I A_I)}{\partial z} + \rho_I u_I A_I \frac{\partial \mathbf{Y}_I}{\partial z} \quad (64)$$

Since the first term on the right-hand side is just the governing equation of mass, this can be substituted and the only derivative that is left is for the mass fraction in the channel,

$$\frac{d \mathbf{Y}_I}{dz} = - \frac{S_I k_m (\rho_I \mathbf{Y}_I - \rho_s \mathbf{Y}_s)}{\rho_I u_I A_I} \quad (65)$$

This, and similarly the outlet channel mass fractions, can then be included in the LSODI solver along with the other derivatives. It is interesting to note that if mass transfer to the surface is not important, the original mass fraction equation of Bissett (Eq. (14)) is recovered.

In the PM and wall layers, the need for computational speed required a numerical approximation. This is because solving the compressible version of the governing equations through these layers proved too tedious a numerical task. Using a linear density approximation in the last paper significantly decreased the computational time. In addition, because the temperature is nearly uniform in the PM and wall layers, this also simulated compressibility since the density change varied directly with the pressure drop. To solve for the species equations in the PM and wall layers, a variable differentiation formulation is utilized similar to the previous efforts for the energy equation.

Taking the PM species equation (Eq. (28)) and expanding the derivative terms on the left-hand side while incorporating the mass equation through the layer, the governing equation becomes,

$$\rho_s A_s \frac{\partial(\mathbf{Y}_s \mathbf{V}_s)}{\partial y} + \mathbf{Y}_s \mathbf{V}_s \left[A_s \frac{\partial \rho_s}{\partial y} + \rho_s \frac{\partial A_s}{\partial y} \right] + \rho_s u_s A_s \frac{\partial \mathbf{Y}_s}{\partial y} = A_s (\dot{S}_s - \mathbf{Y}_s \dot{S}_{\text{mass}}) \quad (66)$$

In this equation, the diffusion velocity in Eq. (33) can be utilized

Table 1 Parameters used for diffusion velocity comparison of Fig. 2

Variable	Value
EX-80 permeability	$2 \times 10^{-13} \text{ m}^2$
Channel diameter	2.11 mm
Wall thickness	0.432 mm
Cell density	100 cells in. ⁻²
Pore diameter	12 μm
Substrate specific heat	600 J kg ⁻¹ K ⁻¹
Inlet pressure	1.0119 bars
Inlet mole fraction N ₂	0.82
Initial loading	4 g l ⁻¹
EX-80 Forscheimer	$5 \times 10^8 \text{ m}^{-1}$
Channel length	12 in.
DPF diameter	5.66 in.
Substrate density [53]	1300 kg m ⁻³
Substrate thermal conductivity	0.5 W m ⁻¹ K ⁻¹
Porosity	50%
Inlet temperature	1000 K
Inlet mole fraction O ₂	0.18
Inlet mass flow rate	0.03 kg s ⁻¹

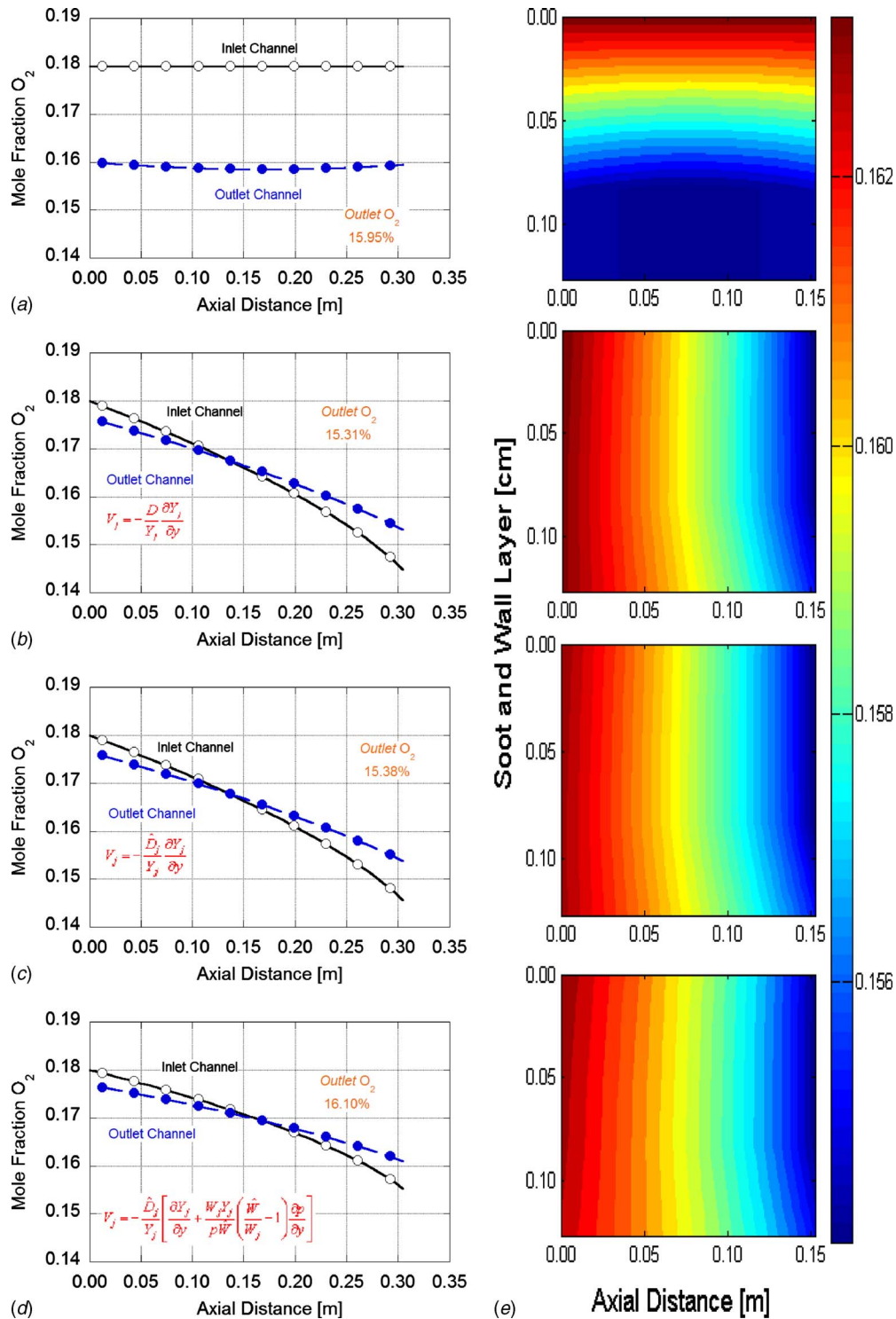


Fig. 2 Comparison of the different diffusion options

as follows:

$$\mathbf{Y}_s \mathbf{V}_s = -\hat{\mathbf{D}} \left[\frac{\partial \mathbf{Y}_s}{\partial y} + \frac{\mathbf{W}_s \mathbf{Y}_s}{p_s \hat{W}} \left(\frac{\hat{W}}{\mathbf{W}_s} - 1 \right) \frac{\partial p_s}{\partial y} \right] \quad (67)$$

along with its derivative,

$$\frac{\partial(\mathbf{Y}_s \mathbf{V}_s)}{\partial y} = -\hat{\mathbf{D}} \left\{ \frac{\partial^2 \mathbf{Y}_s}{\partial y^2} - \frac{\mu W_s}{\hat{W} K_d p_s^2} \left(\frac{\hat{W}}{W_s} - 1 \right) \left[p_s \left(\mathbf{Y}_s \frac{\partial u_s}{\partial y} + u_s \frac{\partial \mathbf{Y}_s}{\partial y} \right) - u_s \mathbf{Y}_s \frac{\partial p_s}{\partial y} \right] \right\} \quad (68)$$

Note that Forscheimer's effect has not been included in the de-

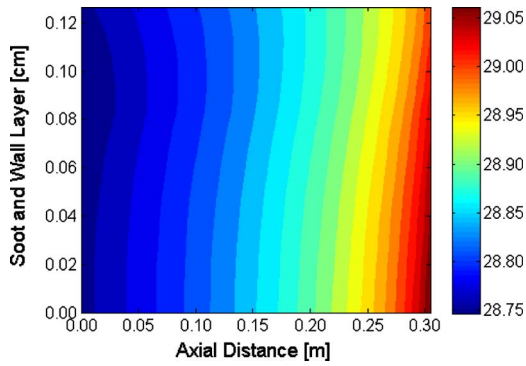


Fig. 3 Contour plot of molecular weight through PM and wall layer (dp/dy diffusion option modeled).

ivative; however, it is often neglected in DPF modeling. Variable differentiation, where adjustments are made to the finite difference expansions, was then utilized for all of the derivatives in this equation to account for the fact that the wall thickness is going to be different than the PM thickness when the same number of computed divisions for each layer are utilized ($\Delta y_w \neq \Delta y_s$). The use of variable differentiation was presented in the previous paper along with the solution of all other differential equations given in this paper.

5 Results

As illustrated in Sec. 4, the inclusion of the pressure derivative in the diffusion velocity creates a higher level of numerical complexity. Since the flow velocity in the wall is based on the pressure difference between the inlet and outlet channels, it seems relevant that this extra term is required to more accurately model the diffusion of the chemical species. However, in order to determine whether or not this intricacy is justified, two simulation tests were accomplished.

The first test examined the effect of the oxygen level leaving the DPF through the myriad of different diffusion options. A moderately loaded DPF was utilized along with a reasonable flow rate through the DPF as documented in Table 1. Figure 2 shows the effects of this numerical test when running (a) without diffusion, (b) using one average binary diffusion constant as a numerical

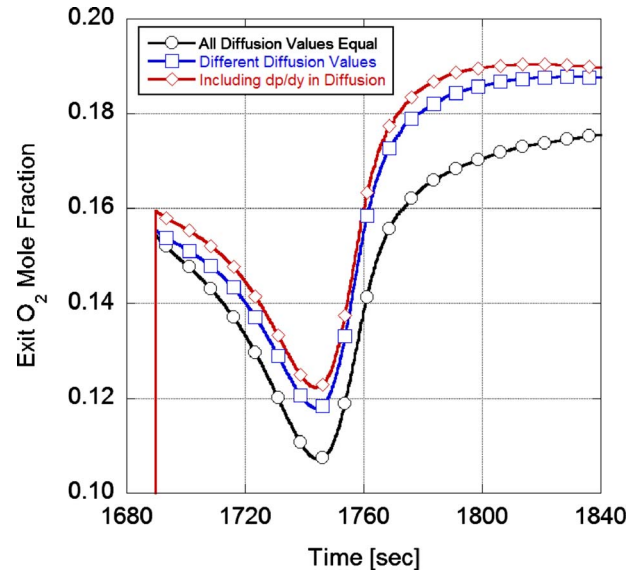
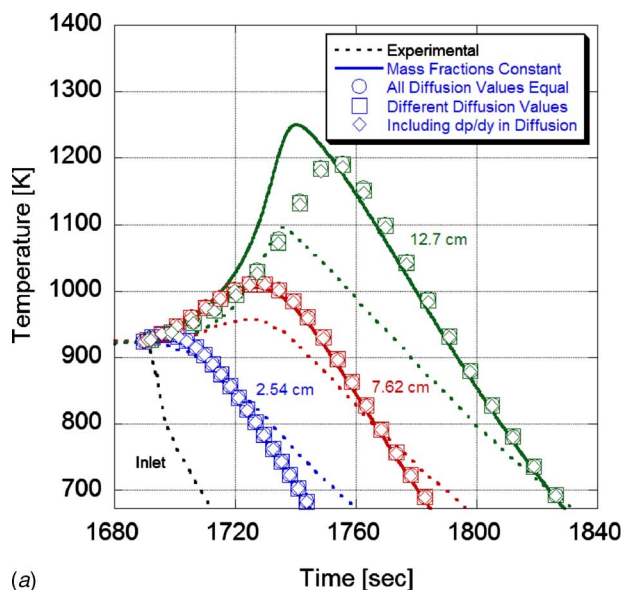


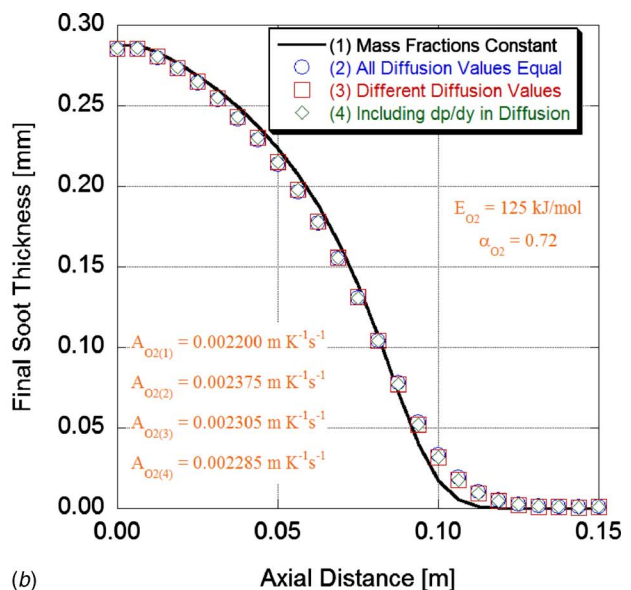
Fig. 5 Exit oxygen mole fraction comparing three different diffusion velocity options

simplification, (c) using individual diffusion constants, and (d) using individual diffusion constants along with the dp/dy effect. It was found that the influence of diffusion in the layer had a large impact as previously documented in the literature. When the pressure derivative was included, the outlet oxygen percentage changed notably. Looking at the absolute numbers for the diffusion velocity demonstrated that the pressure derivative term had around 5% effect on this value.

Regarding the pressure difference effect, earlier in the paper it was assumed that the molecular weight change in the y -direction is negligible. This was accomplished to reduce some complexity in the calculation of the diffusion velocity when this effect is included. To validate this assumption, the molecular weight was plotted in Fig. 3 for both the PM and wall layers. It was found that the maximum change in molecular weight from entrance to exit is only on the order of 0.3%. This stems from the fact that diffusion uniformly disperses the species in the y -direction.



(a)



(b)

Fig. 4 Oxidation and cool down test simulated using the different diffusion velocity options

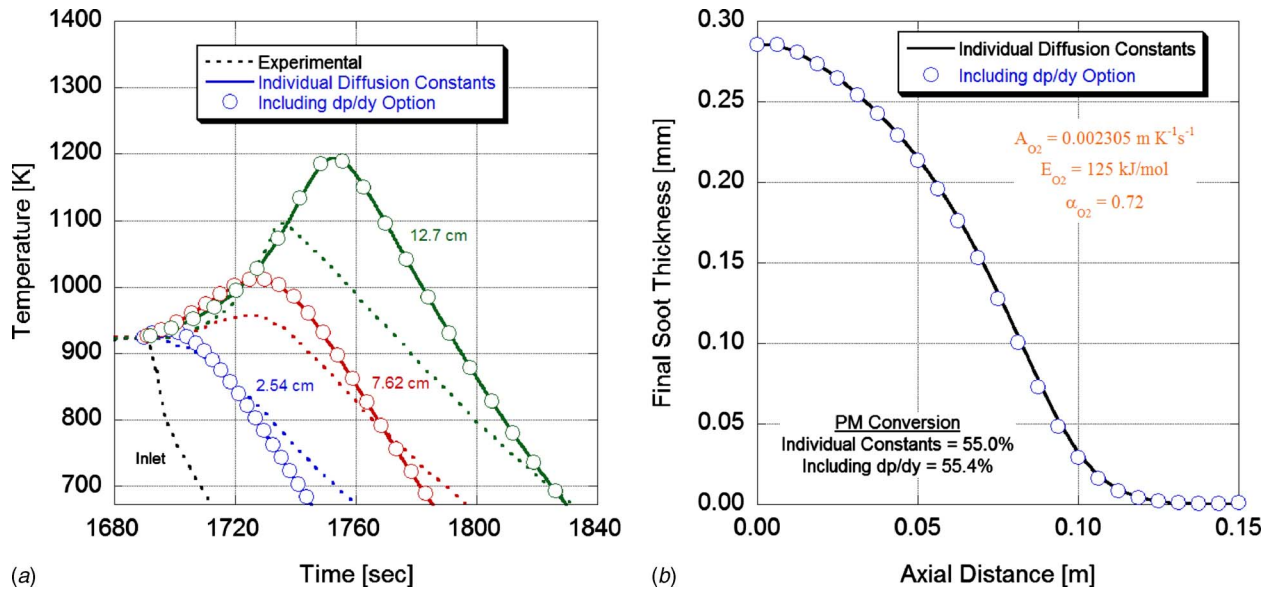


Fig. 6 Utilizing individual diffusion constant calibrated pre-exponential with dp/dy version

The second test utilized the same oxidation and cool down tests presented in the last paper to give a more realistic impact of this term. All four options were again utilized and the pre-exponential in Eq. (42) calibrated to give the experimentally found 55% conversion of PM on the surface for each option. In Fig. 4, the results show that including diffusion has a significant impact on the temperature profile, again validating what is already known in the literature. When each diffusion version was calibrated, little difference can be seen between all of the options in the resultant temperature and PM profiles. However, a noticeable difference is observed in the oxygen mole fractions exiting the DPF as shown in Fig. 5. While the impact via O_2 oxidation does not appear to be that large for regeneration, in the case of a passively-regenerated DPF using NO_2 or a catalyzed DPF, the differences in species concentrations could have a significant influence in the modeling results leading to altered regeneration strategies. A final simulation was accomplished to see what error would result when using

the pre-exponential for the individual diffusion constants with the version including the pressure difference effect. In Fig. 6, only a small difference in the final PM level on the surface is found again indicating that the impact on O_2 oxidation may be negligible given the additional complexity required.

Previous DPF modeling efforts often neglect the heats of formation and sensible enthalpy components of the chemical species within the energy equations by modeling an enthalpy difference as the constant pressure specific heat times a temperature difference for Eq. (53). Figure 7 demonstrates the effect of this assumption on the same cool down and oxidation tests. Both the temperature and PM profiles show a disparity with the final conversion of PM under predicted by 3% versus the enthalpy difference case.

In order to balance the concentration equations with the governing equation of mass, an additional term is required in the mass equations accounting for the summation of all mass transfer terms in Eqs. (25) and (26). This term is often assumed to be negligible

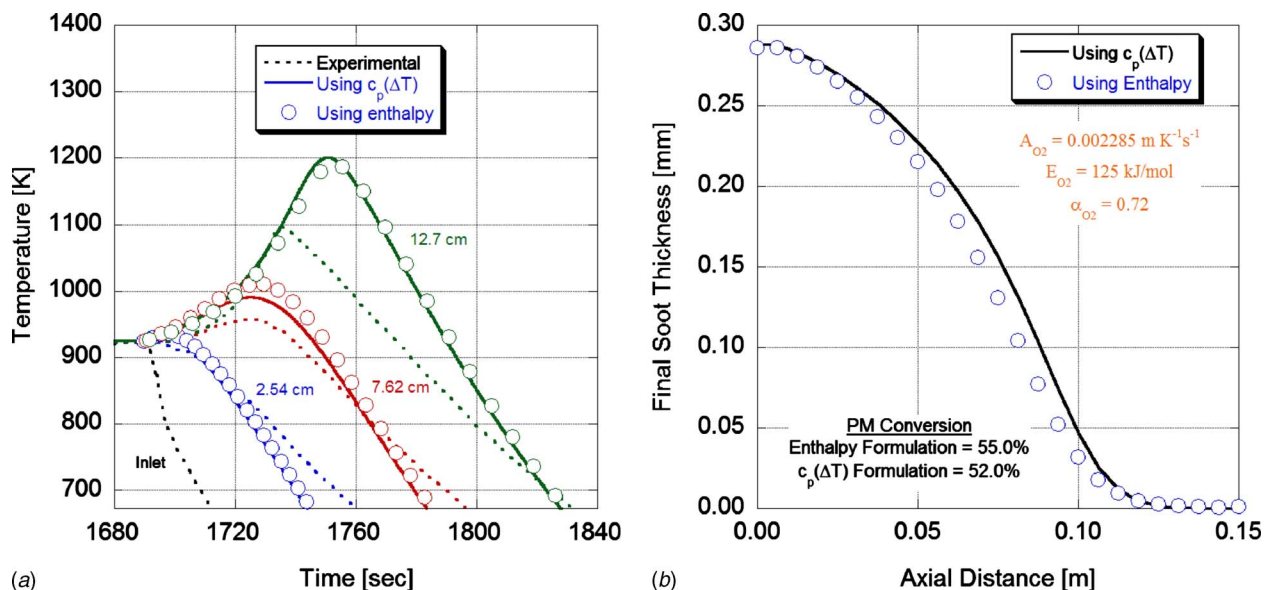


Fig. 7 Comparing the use of $c_p(\Delta T)$ for enthalpy with actual enthalpy in energy equations (dp/dy diffusion option modeled)

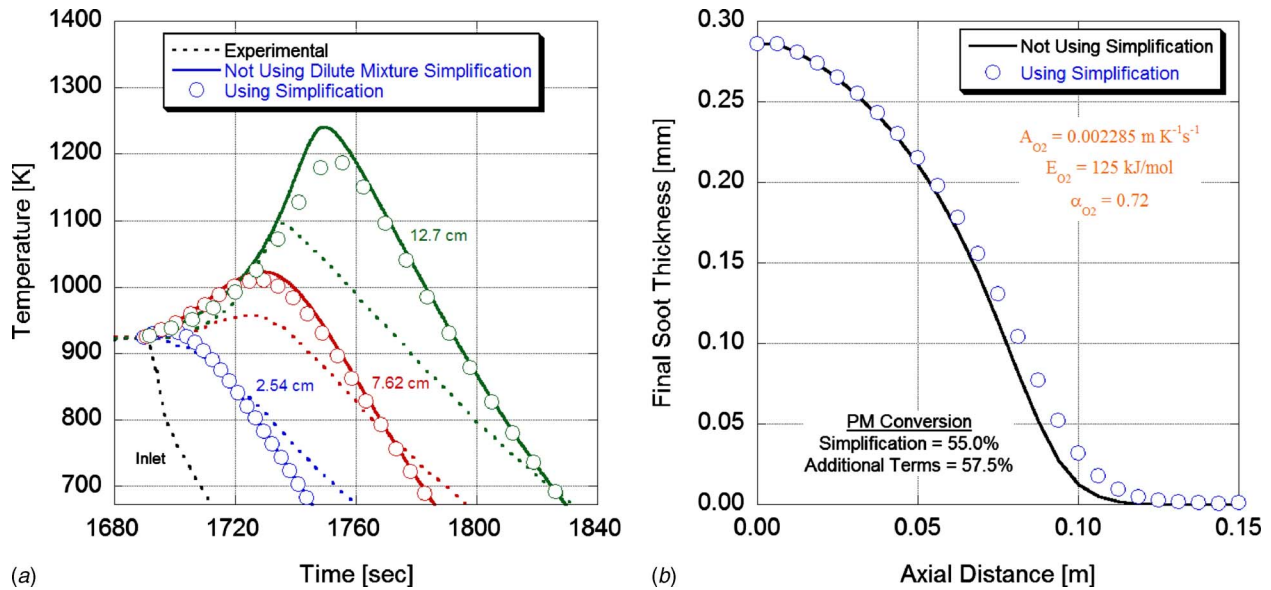


Fig. 8 Comparing the use of dilute mixture simplification (dp/dy diffusion option modeled)

in catalyst modeling via a dilute mixture simplification; however, since each channel is an open system, this assumption may be erroneous. In Fig. 8, this is indeed found to be the case when the dilute mixture simplification is not utilized. The maximum temperature differs by approximately 50 K and the final soot thickness differs in a non-negligible amount. This disparity relates to the omitted change in density leading to the overall mass flow rate leaving the DPF differing by up to 6.5%, as shown in Fig. 9. Since the mass transfer term is on the same order as the convective velocity to the surface, even a small difference between the mass fractions is enough to impact the oxidation event.

6 Conclusions

This paper enhances an earlier paper by including the concentration equations of motion into the area-conserved one-dimensional based diesel particulate filter model. A brief historical review of the species equations was presented to illustrate how this model differs with respect to the governing equations. In ad-

dition, a succinct recap of the channel and wall model species equations was presented to illustrate how the models were merged via this particular equation.

The diffusion constants were modified to include the close proximity of the wall and it was found that expanding the diffusion velocity to take into account the pressure difference through the layers changed the oxygen concentration leaving the DPF. However, this did not have a large effect on an actual oxygen regeneration event with only a small difference in final soot loading indicated. This component may need to be considered for NO_2 passively-regenerated events where small changes in NO_2 could have a relatively important impact on the amount of soot oxidized and the regeneration strategy employed.

The model formulation illustrated that a common assumption to make for an enthalpy difference is the use of constant pressure specific heat times a temperature difference. Because of the different heats of formation and sensible enthalpies associated with the chemical species, this assumption was reviewed and indicated to have some error. Finally, the commonly accepted dilute mixture simplification involving the mass transfer coefficients in the concentration equations may need to be included because of the error involved with the open system architecture of the model.

In the papers presented by the author, a number of assumptions with respect to the original model formulation by Bissett and current DPF modeling papers have been checked. While analyzing each assumption has led to small differences in results, the accumulation of ignoring all these effects together may lead to a more significant error in the model predictions. In specific, not including area-conservation, ignoring the pressure effect on diffusion, using constant pressure times a temperature difference and assuming a dilute mixture may lead to an associated error in the model. However, as the results in the paper show, none of these changes led to an order of magnitude difference in temperature that is required to match the experimental data. Hence, further expansion of the model to two and three dimensions as others have accomplished is most likely required to achieve these gains. Through the previous paper and this document, this expansion can now be accomplished understanding the impact of some widely held assumptions prevalent throughout the history of DPF modeling.

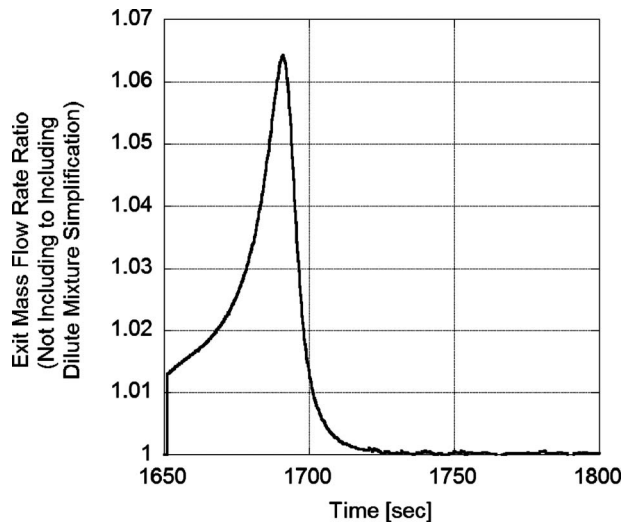


Fig. 9 Difference in the exit mass flow rates between different dilute mixture simplification options

Nomenclature

- A = cross-sectional area (m^2)
- A = pre-exponential ($\text{m K}^{-1} \text{s}^{-1}$)

A_{O_2} = pre-exponential in O_2 Arrhenius rate law ($m K^{-1} s^{-1}$)	W = molecular weight ($g mol^{-1}$)
A_s = cross-sectional area in soot layer per unit length (m)	\hat{W} = mixture molecular weight ($g mol^{-1}$)
A_w = cross-sectional area in wall layer per unit length (m)	X = mole fractions
c = specific heat ($J kg^{-1} K^{-1}$)	y = soot and wall (normal) distance (m)
\dot{C} = concentration flow rate ($kg m^{-2} s^{-1}$)	Y = mass fractions
c_p = constant pressure specific heat of gas ($J kg^{-1} K^{-1}$)	Y_3 = mass fraction array [24,25]
d = effective diameter (m)	z = channel (axial) distance (m)
d_m = channel diameter (m)	Z = right-hand side components in MOC format
d_p = pore diameter (m)	
\mathbf{D} = diffusion vector ($m^2 s^{-1}$)	
$\hat{\mathbf{D}}$ = mixture-averaged diffusion vector ($m^2 s^{-1}$)	
\mathbf{D}^T = thermal diffusion vector ($m^2 s^{-1}$)	
E_a = activation energy ($kJ mol^{-1}$)	
E_{O_2} = activation energy in O_2 Arrhenius rate law ($kJ mol^{-1}$)	
f = geometry factor [11–13,23,34]	
F = friction factor of Bissett (28.454)	
h = specific enthalpy ($J kg^{-1}$)	
h_g = convective heat transfer coefficient in the channels ($W m^{-2} K^{-1}$)	
k = rate constant ($m s^{-1}$)	
k_m = mass transfer coefficient ($m s^{-1}$)	
m = mass (kg)	
N_c = total number of channels	
N_s = number of discretizations in the soot layer	
p = pressure (Pa)	
\dot{Q}_{cond} = energy conduction within the wall ($W m^{-1}$)	
\dot{Q}_{conv} = energy convection due to gas in channels ($W m^{-1}$)	
\dot{Q}_{multi} = energy conduction due to multidimensional effects ($W m^{-1}$)	
\dot{Q}_{rad} = radiation energy within filter channels ($W m^{-1}$)	
\dot{Q}_{reac} = energy exotherm/endotherm due to soot/wall reactions ($W m^{-1}$)	
\dot{Q}_{supply} = active energy supply to wall ($W m^{-1}$)	
\dot{Q}_{wall} = energy flow through the wall ($W m^{-1}$)	
\mathbf{R} = catalytic reaction vector ($mol m^{-3} s^{-1}$)	
R_u = universal gas constant (8.3145) ($J mol^{-1} K^{-1}$)	
S = surface area per unit length (m)	
$\dot{S}_{C(s)}$ = overall soot combustion rate ($kg C_{(s)} m^{-3} s^{-1}$)	
\dot{S}_{mass} = discretized soot combustion rate ($kg C_{(s)} m^{-3} s^{-1}$)	
\dot{S}_{CO} = carbon monoxide generation rate ($kg CO m^{-3} s^{-1}$)	
\dot{S}_{CO_2} = carbon dioxide generation rate ($kg CO_2 m^{-3} s^{-1}$)	
\dot{S}_{O_2} = oxygen combustion rate ($kg O_2 m^{-3} s^{-1}$)	
S_p = specific surface area of soot (5.5×10^7) (m^{-1})	
t = time (s)	
t_d = soot thickness (m)	
t_m = filter wall thickness (m)	
T = temperature (K)	
u = velocity ($m s^{-1}$)	
\bar{u} = average velocity in soot and wall layer ($m s^{-1}$)	
V = volume per axial length (m^2)	
\mathbf{V} = diffusion velocity vector ($m s^{-1}$)	
	Greek Variables
	α_{NO_2} = NO_2 partial oxidation factor
	α_{O_2} = O_2 partial oxidation factor
	β = Forscheimer inertial coefficient (m^{-1})
	χ = inverse of soot combustion time scale (s^{-1})
	Δ_4 = source term in the method of characteristics
	ΔH_{reac} = heat of reaction ($kJ mol^{-1}$)
	ϵ_p = porosity (%)
	\mathbf{K} = Darcy's law permeability (m^2)
	λ = thermal conductivity ($W m^{-1} K^{-1}$)
	μ = dynamic viscosity ($kg m^{-1} s^{-1}$)
	θ = soot accumulation angle (deg)
	ρ = density ($kg m^{-3}$)
	s = mass flow rate of soot to the surface per mesh interval ($kg C_{(s)} m^{-1} s^{-1}$)
	τ = tortuosity
	Subscript and Sub-Subscript
	b = bulk
	$C_{(s)}$ = solid carbon, particulate, or soot
	d = soot layer (solid)
	f = filter
	i = axial index number
	I = inlet channel
	II = outlet channel
	j = species and normal index numbers
	Kn = Knudsen
	NM = total number of species
	N_2 = nitrogen
	NO_2 = nitrogen dioxide
	O_2 = oxygen
	s = soot layer (gas)
	w = wall
	References
	[1] Howitt, J. S., and Montierth, M. R., 1981, "Cellular Ceramic Diesel Particulate Filter," SAE Paper No. 810114.
	[2] General Motors, 1979, "General Motors Response to EPA Notice of Proposed Rulemaking on Particulate Regulation for Light-Duty Diesel Vehicles Submitted to Environmental Protection Agency."
	[3] Springer, K. J., and Stahman, R. C., 1977, "Removal of Exhaust Particulate From a Mercedes 300d Diesel Car," SAE Paper No. 770716.
	[4] Tessier, L. P., Sullivan, H. F., Bragg, G. M., and Hermance, C. E., 1980, "The Development of a High Efficiency Diesel Exhaust Particulate Filter," SAE Paper No. 800338.
	[5] Chatterjee, S., Walker, A. P., and Blakeman, P. G., 2008, "Emission Control Options to Achieve Euro IV and Euro V on Heavy-Duty Diesel Engines," SAE Paper No. 2008-28-0021.
	[6] Johnson, T. V., 2008, "Diesel Emission Control Technology in Review," SAE Paper No. 2008-01-0069.
	[7] Dieselnets, 2008, "Emission Standards: USA: Nonroad Diesel Engines," http://www.dieselnets.com/standards/us/offroad.html .
	[8] Bissett, E. J., 1984, "Mathematical Model of the Thermal Regeneration of a Wall-Flow Monolith Diesel Particulate Filter," Chem. Eng. Sci., 39 , pp. 1233–1244.
	[9] Koltsakis, G. C., and Stamatelos, A. M., 1997, "Modes of Catalytic Regeneration in Diesel Particulate Filters," Ind. Eng. Chem. Res., 36 , pp. 4155–4165.
	[10] Konstandopoulos, A. G., and Johnson, J. H., 1989, "Wall-Flow Diesel Particulate Filters—Their Pressure Drop and Collection Efficiency," SAE Paper No. 890405.
	[11] Haralampous, O. A., Dardiotis, C. K., Koltsakis, G. C., and Samaras, Z. C., 2004, "Study of Catalytic Regeneration Mechanisms in Diesel Particulate Filters Using Coupled Reaction-Diffusion Modeling," SAE Paper No. 2004-01-1941.

- [12] Haralampous, O. A., and Koltsakis, G. C., 2004, "Oxygen Diffusion Modeling in Diesel Particulate Filter Regeneration," *AIChE J.*, **50**(9), pp. 2008–2019.
- [13] Haralampous, O. A., Koltsakis, G. C., Samaras, Z. C., Vogt, C.-D., Ohara, E., Watanabe, Y., and Mizutani, T., 2004, "Reaction and Diffusion Phenomena in Catalyzed Diesel Particulate Filters," SAE Paper No. 2004-01-0696.
- [14] Konstandopoulos, A. G., Kostoglou, M., Vlachos, N., and Kladopoulou, E., 2005, "Progress in Diesel Particulate Filter Simulation," SAE Paper No. 2005-01-0946.
- [15] Konstandopoulos, A. G., Vlachos, N., Housiada, P., and Kostoglou, M., 2003, "Simulation of Triangular-Cell-Shaped, Fibrous Wall-Flow Filters," SAE Paper No. 2003-01-0844.
- [16] Bissett, E. J., 1985, "Thermal Regeneration of Particle Filters With Large Conduction," *Math. Modell.*, **6**, pp. 1–18.
- [17] Bissett, E. J., and Shadman, F., 1985, "Thermal Regeneration of Diesel-Particulate Monolithic Filters," *AIChE J.*, **31**, pp. 753–758.
- [18] Depcik, C., and Assanis, D., 2008, "Simulating Area Conservation and the Gas-Wall Interface for One-Dimensional Based Diesel Particulate Filter Models," *ASME J. Eng. Gas Turbines Power*, **130**(6), pp. 062807.
- [19] Lee, W.-T., 1992, "Local Preconditioning of the Euler Equations," Ph.D. thesis, University of Michigan, Ann Arbor, MI.
- [20] Roe, P. L., 1986, "Characteristic-Based Schemes for the Euler Equations," *Annu. Rev. Fluid Mech.*, **18**, pp. 337–365.
- [21] Liu, J., Schorn, N., Schernus, C., and Peng, L., 1996, "Comparison Studies on the Method of Characteristics and Finite Difference Methods for One-Dimensional Gas Flow Through IC Engine Manifold," SAE Paper No. 960078.
- [22] Onorati, A., Ferrari, G., and D'errico, G., 1999, "Fluid Dynamic Modeling of the Gas Flow With Chemical Specie Transport Through the Exhaust Manifold of a Four Cylinder SI Engine," SAE Paper No. 1999-01-0557.
- [23] Haralampous, O. A., Koltaskis, G. C., Samaras, Z. C., Vogt, C.-D., Ohara, E., Watanabe, Y., and Mizutani, T., 2004, "Modeling and Experimental Study of Uncontrolled Regenerations in Sic Filters With Fuel Borne Catalyst," SAE Paper No. 2004-01-0697.
- [24] Peters, B. J., 2003, "Numerical Simulation of a Diesel Particulate Filter During Loading and Regeneration," 2003 Spring Technical Conference of the ASME Internal Combustion Engine Division, Salzburg, Austria.
- [25] Peters, B. J., Wanker, R. J., Münzer, A., and Wurzenberger, J. C., 2004, "Integrated 1D to 3D Simulation Workflow of Exhaust Aftertreatment Devices," SAE Paper No. 2004-01-1132.
- [26] Depcik, C., Van Leer, B., and Assanis, D., 2004, "The Numerical Simulation of Variable-Property Reacting-Gas Dynamics: New Insights and Validation," *Numer. Heat Transfer, Part A*, **47**, pp. 27–56.
- [27] Garner, C. P., and Dent, J. C., 1988, "A Thermal Regeneration Model for Monolithic and Fibrous Diesel Particulate Traps," SAE Paper No. 880007.
- [28] Koltsakis, G. C., and Stamatelos, A. M., 1996, "Modeling Thermal Regeneration of Wall-Flow Diesel Particulate Traps," *AIChE J.*, **42**(6), pp. 1662–1672.
- [29] Koltsakis, G. C., and Stamatelos, A. M., 1996, "Modeling Catalytic Regeneration of Wall-Flow Particulate Filters," *Ind. Eng. Chem. Res.*, **35**, pp. 2–13.
- [30] Jørgensen, M. W., and Sorenson, S. C., 1997, "A 2-Dimensional Simulation Model for a Diesel Particulate Filter," SAE Paper No. 970471.
- [31] Konstandopoulos, A. G., and Kostoglou, M., 1999, "Periodically Reversed Flow Regeneration of Diesel Particulate Traps," SAE Paper No. 1999-01-0469.
- [32] Konstandopoulos, A. G., and Kostoglou, M., 2000, "Reciprocating Flow Regeneration of Soot Filters," *Combust. Flame*, **121**, pp. 488–500.
- [33] Konstandopoulos, A. G., Kostoglou, M., Skaperdas, E., Papaioannou, E., Zarvalis, D., and Kladopoulou, E., 2000, "Fundamental Studies of Diesel Particulate Filters: Transient Loading, Regeneration and Aging," SAE Paper No. 2000-01-1016.
- [34] Haralampous, O. A., Koltsakis, G. C., and Samaras, Z. C., 2003, "Partial Regenerations in Diesel Particulate Filters," SAE Paper No. 2003-01-1881.
- [35] Depcik, C., and Assanis, D., 2005, "One-Dimensional Automotive Catalyst Modeling," *Prog. Energy Combust. Sci.*, **31**(4), pp. 308–369.
- [36] Coltrin, M. E., Kee, R. J., and Rupley, F. M., 1990, "Surface Chemkin: A Fortran Package for Analyzing Heterogeneous Chemical Kinetics at a Solid-Surface-Gas-Phase Interface," Sandia National Laboratories, Technical Report No. SAND90-8003B.
- [37] Reaction Design, 2003, *Transport Core Utility Manual*.
- [38] Cussler, E. L., 1995, *Diffusion: Mass Transfer in Fluid Systems*, Cambridge University Press, Cambridge, UK.
- [39] Ho, F. G., and Strieder, W., 1980, "Numerical Evaluation of the Porous Medium Effective Diffusivity Between the Knudsen and Continuum Limits," *J. Chem. Phys.*, **73**(12), pp. 6296–6300.
- [40] Mason, E. A., and Malinaukus, A. P., 1983, *Gas Transport in a Porous Media: The Dusty Gas Model*, Elsevier, Amsterdam.
- [41] Reif, F. P., 1975, *Statistical and Thermal Physics*, McGraw-Hill, New York.
- [42] Smith, J. M., 1970, *Chemical Engineering Kinetics*, McGraw-Hill, New York.
- [43] Masoudi, M., Konstandopoulos, A. G., Nikitidis, M. S., Skaperdas, E., Zarvalis, D., Kladopoulou, E., and Altiparmakis, S., 2001, "Validation of a Model and Development of a Simulator for Predicting the Pressure Drop of Diesel Particulate Filters," SAE Paper No. 2001-01-0911.
- [44] Mohammed, H., Triana, A. P., Yang, S.-L., and Johnson, J. H., 2006, "An Advanced 1D 2-Layer Catalyzed Diesel Particulate Filter Model to Simulate: Filtration by the Wall and Particulate Cake, Oxidation in the Wall and Particulate Cake by NO₂ and O₂ and Regeneration by Heat Addition," SAE Paper No. 2006-01-0467.
- [45] Zhang, Z., Yang, S. L., and Johnson, J. H., 2002, "Modeling and Numerical Simulation of Diesel Particulate Trap Performance During Loading and Regeneration," SAE Paper No. 2002-01-1019.
- [46] Aoki, H., Asano, A., Kurazono, K., Kobashi, K., and Sami, H., 1993, "Numerical Simulation Model for the Regeneration Process of a Wall-Flow Monolith Diesel Particulate Filter," SAE Paper No. 930364.
- [47] Jacquot, F., Logie, V., Brilhac, J. F., and Gilot, P., 2002, "Kinetics of the Oxidation of Carbon Black by NO₂: Influence in the Presence of Water and Oxygen," *Carbon*, **40**, pp. 335–343.
- [48] Neeft, J. P. A., Nijhuis, T. X., Smakman, E., Makkee, M., and Moulijn, J. A., 1997, "Kinetics of the Oxidation of Diesel Soot," *Fuel*, **76**(12), pp. 1129–1136.
- [49] Konstandopoulos, A. G., Kostoglou, M., Housiada, P., Vlachos, N., and Zarvalis, D., 2003, "Multichannel Simulation of Soot Oxidation in Diesel Particulate Filters," SAE Paper No. 2003-01-0839.
- [50] Depcik, C., 2003, "Modeling Reacting Gases and Aftertreatment Devices for Internal Combustion Engines," Ph.D. thesis, University of Michigan, Ann Arbor, MI.
- [51] Kuo, J. C., Morgan, C. R., and Lassen, H. G., 1971, "Mathematical Modeling of CO and HX Catalytic Converter Systems," SAE Paper No. 710289.
- [52] Hindmarsh, A. C., 2002, "Serial Fortran Solvers for ODE Initial Value Problems," <http://www.llnl.gov/CASC/odepack/>.
- [53] Kandylas, I. P., and Koltsakis, G. C., 2002, "Simulation of Continuously Regenerating Diesel Particulate Filters in Transient Driving Cycles," *Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.)*, **216**, pp. 591–606.

An Experimental and Numerical Investigation of Spark Ignition Engine Operation on H₂, CO, CH₄, and Their Mixtures

Hailin Li¹

Department of Mechanical and Aerospace
Engineering,
West Virginia University,
Morgantown, WV 26506
e-mail: hailin.li@mail.wvu.edu

Ghazi A. Karim

A. Sohrabi

Department of Mechanical and Manufacturing
Engineering,
University of Calgary,
Calgary, Alberta, T2N 1N4, Canada

The knock and combustion characteristics of CO, H₂, CH₄, and their mixtures were determined experimentally in a variable compression ratio spark ignition (SI) cooperative fuel research (CFR) engine. The significant effects of gaseous fuel mixtures containing H₂ in enhancing the combustion and oxidation process of CH₄ were examined. The unique combustion characteristics of CO in dry air and its distinct performance in mixtures with H-containing fuels were investigated. The addition of a simulated synthesis gas (2H₂ + CO) to CH₄ was found to enhance the combustion process of the resulting mixture and lowers its knock resistance. The effectiveness of such an addition is slightly weaker than that of a comparable H₂ addition but much stronger than that with CO addition only. A predictive model with detailed kinetic chemistry was used successfully to simulate SI engine operation fuelled with CH₄, H₂, CO, and their mixtures. The predicted engine performance and knock limits of CH₄, H₂, CO, and their mixtures agree well with experimental data with the exception around pure CO operation in dry air with the presence of small amounts of CH₄ or H₂. A remedial approach to improve the prediction of the knock limits of fuel mixtures containing mainly CO with a small amount of H-containing fuels such as H₂ and CH₄ was proposed and discussed.

[DOI: 10.1115/1.3155795]

1 Introduction

Hydrogen has long been recognized as an energy carrier having some unique and highly desirable combustion properties for applications in spark ignition (SI) engines, such as a wide flammable mixture range, low ignition energy, and very fast flame propagation rates [1]. These attractive features make it possible to operate conventional SI engines on very lean mixtures so as to achieve both high power production efficiency and low oxide of nitrogen (NO_x) emissions [1,2]. Negligible amounts of carbon monoxide (CO), carbon dioxide (CO₂), and unburned hydrocarbon (UHC) are produced from the oxidation of the lubricating oil [3,4]. When used in SI engines for power production, H₂ can be burned either as a sole fuel or supplement to relatively slow burning fuels, noted as H₂-enrichment. The latter approach has been demonstrated to be very effective in accelerating the flame propagation rates, extending the lean operational limits and enhancing the combustion stability in the presence of heavy exhaust gas recirculation (EGR) or for extremely lean mixture operation [5–7].

The application of H₂ in the transportation sector, although it has been claimed and demonstrated to be attractive, remains limited by some lingering obstacles. These include its mass production economically, the lack of infrastructure for its transportation and distribution, concerns for safety issues, its tendency to encounter knocking operation in SI engines at high loads, the occurrence of backfire, and the inconvenience of its storage on-board [1,2,8–11]. As a compromise, there is an increasing interest in recent years to produce on-board gaseous fuel mixtures containing H₂ in the presence of a variety of concentrations of CO, CO₂,

H₂O, and N₂ such as those produced through the partial oxidation or steam reforming of hydrocarbon fuels notably natural gas, gasoline, and diesel fuels [12–14]. These H₂-rich gaseous mixtures have features reflecting the desirable combustion properties of H₂ in SI engine applications such as fast flame propagation rates combined with wide lean operational mixture region [12–17]. This would make it possible to operate SI engines with traditional fuels under either ultra lean or with a high level of EGR to enhance the combustion process and reduce exhaust emissions of NO_x by allowing extremely lean operation. For example, Andreatta and Dibble [16] and Bromberg et al. [17] examined the performance of a SI engine operated on a mixture of natural gas and its reforming products prepared by plasmacreation with the aim of improving engine performance. Allenby et al. [18] examined the potential of maintaining combustion stability in a natural gas fuelled engine with EGR through adding reformed gases. Quader et al. [19] examined the effects of reformed gases supplement on SI gasoline engine performance when operated at around the lean operational limits. These researches demonstrated the desirable features of reformed gases in enhancing flame propagation rates, improving power production efficiency, reducing UHC emissions under lean mixture operation and extending the lean operational limit. Recently, the production of H₂-rich gases has been proposed through exhaust gas reforming and partial oxidation technology aiming to improve either performance of internal combustion (IC) engines, the operation of after-treatment devices or the regeneration of diesel particulate matter (PM) filter [20–23]. The application of reformed EGR was shown to improve the performance of lean mixture SI engine operation [24]. Plus, gas reforming technology has also been used to improve the cold start performance of SI engines. For example, Isherwood et al. [25] investigated the cold start performance of a SI engine fueled with partially reformed gases. Improved cold start capability was reported while the time averaged UHC and CO emissions were significantly reduced. More recently, gas reforming technologies have been also

¹Corresponding author.

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received January 16, 2009; final manuscript received April 6, 2009; published online November 30, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Internal Combustion Engine Division 2004 Fall Technical Conference (ICEF2004), Long Beach, CA, October 24–27, 2004.

used to modify the combustion phasing of homogeneous charge compression ignition (HCCI) engine and extend its operational region [26].

Although the minimum energy needed to ignite H₂-air mixtures is the lowest among fuels suitable for SI engine application, the auto-ignition temperature of H₂ is higher than most of the fuels examined, which would imply a high resistance to knock. Some researchers even suggested that the addition of H₂ to gasoline will enhance the knock-resisting property of the resulting mixture [27]. However, others indicated that H₂ has a higher tendency to knock than most SI engine fuels [28,29] and has been used as a primary low knock-resisting reference fuel in the empirical "Methane Number" [30,31]. This is due mainly to the fast flame propagation rate of H₂-air mixtures, which makes it easier for the unburned mixtures to reach higher temperatures beyond those needed for auto-ignition [32]. The application of H₂ as SI engine fuel also has been highlighted for its wide flammable mixture range, which makes it possible to run SI engine at extremely lean mixtures to achieve high power production efficiency and negligible NO_x emissions. This raises also the issue that the onset of knock may still be encountered at very lean mixtures, which is usually not a concern with other fuels such as gasoline or natural gas [28,29,32]. Accordingly, a fuel mixture containing a large amount of H₂ produced on-board can support a fast-propagating flame but also raises concern about the onset of knock that accompanies H₂ applications in SI engines. Relatively limited research was conducted to address this issue of H₂-rich gaseous fuels. For example, Anthonissen and Wallace [33] examined the knock-free operation region of a simulated H₂-rich synthesis gases (2H₂+CO) for different equivalence and compression ratios (CRs). Experimental data demonstrated that a compression ratio of up to 10:1 can be used without encountering knock at equivalence ratios (ERs) over 0.7 while employing significantly retarded spark timing (e.g., 10° CA after top dead center (ATDC)), well beyond that required for maximum brake torque. It is expected that the compression ratio for knock-free operation will be limited to a much lower value if operated with spark timing around that optimized to obtain maximum brake torque. However, the detailed knock characteristics with changes in fuel composition and engine operating conditions were not reported.

With the increasing interest to produce H₂-rich gaseous fuels on-board to be mixed with traditional fuels, much of the research work about supplementing SI engines with H₂-rich gases focused mainly on the examination of engine performance and exhaust emissions under relatively limited operating conditions [16–19,34]. In comparison, the detailed examination of the corresponding combustion and knock characteristics when mixed with traditional fuels appears not to have been reported. It should be noted that the knock characteristics of fuel mixtures involving freshly produced H₂-rich gases tend to be more complex than other commercial fuels stored on-board. This is due to their specific physical and chemical features of the gaseous fuel mixtures produced on-board. For example, their temperature and composition can vary widely with the extent of partial oxidation or the reforming processes, which depends to a large extent on the operating conditions such as exhaust temperature, O₂ concentration, and the availability of H₂O for water-shift reactions aiming to produce more H₂. Also, the presence of N₂, CO₂, or H₂O in the H₂-rich gases will act as selective EGR and lower the tendency of the resulting fuel mixtures to knock [35,36]. These make it both unwieldy and impractical to examine the knock characteristics of these fuel mixtures entirely through experimentation. It is needed to develop a suitable numerical model capable of predicting the combustion and knock characteristics of fuel mixtures containing H₂-rich gases. The validated model can be used then to explore the combustion and knock characteristics under a wider range of operating conditions and fuel compositions, and targeting those that have not been measured previously. There are at present sufficiently reliable approaches for modeling the combustion of com-

mon gaseous fuels such as CH₄, H₂, C₂H₆, C₃H₈, and their mixtures in SI engine applications [32, 37, and 38]. The extension of the previous modeling approach to a broader area of application to include H₂-rich fuel mixtures requires experimental data such as those of the combustion duration, heat release process, and knock limits so as to provide an extended basis for the development of the model and its validation [39].

This paper presents the results of the experimental and analytical examination of the combustion and knock characteristics of H₂, CO, and CH₄ and their mixtures in SI engines. Extensive experiments were conducted to examine the effect of changes in fuel composition and operating conditions on SI engine operation. The operating parameters examined included engine performance, combustion process development and stability, and knock limits and exhaust emissions. The data obtained should contribute toward improving the knowledge of the behavior and suitability of these increasingly important fuels in SI engine applications and providing also an extended basis to support model development and its validation.

A two-zone quasidimensional model developed earlier [37,38] for CH₄ and H₂ operations was extended to consider also the operation on H₂, CO, and their mixtures with CH₄. The modification to the model included the incorporation of experimentally derived correlations for the combustion duration values with changes in fuel composition and engine operating conditions. The oxidation reactions of these fuel mixtures in air were simulated by employing a suitably detailed chemical kinetic scheme. It is to be shown that the model can predict successfully the performance, combustion, and the onset of knock in SI engines fueled with mixtures containing H₂-rich gaseous fuels produced on-board. Predicted results agreed well with experimental data obtained.

2 Experimental Setup and Operational Procedure

A single cylinder four stroke spark ignition CFR engine of 82.6 mm bore and 114 mm stroke was used while operating unthrottled under local atmospheric pressure (about 88 kPa). The engine is of variable compression ratio (from 4:1 to 16:1) and spark timing (up to 40° CA before and after top dead center). The engine was maintained at a constant speed of 900 rpm. The induction system was modified to permit the simultaneous admission of multifuels supplied from a set of high-pressure gaseous fuel tanks and metered individually using a series of precalibrated choked nozzles. The simulated gaseous fuel mixture, which was mixed on-site by precisely controlling the flow rate of each individual fuel component, was introduced into the intake manifold to mix with air just before entry into the cylinder. The flow rate of the air was determined using a laminar flow air meter of low pressure drop. The intake air whenever required could be dried with silica gel to remove moisture and heated electrically to the required temperature. The in-cylinder pressure was measured using a high frequency-response water-cooled piezoelectric type transducer that was mounted flush with the pan-shaped cylinder head. The pressure-time variation was recorded with a data acquisition system and stored in a computer for subsequent analysis.

A number of approaches have been proposed and demonstrated to be effective in detecting the onset of knock. In the present work, it was determined by a combination of the appearance of the characteristic oscillations on the cylinder pressure trace together with the detection of the specific level of the occasional audible noise related to knocking. The "knock limit" in this contribution is defined in terms of the equivalence ratio when borderline knock was observed at the specified operating condition. During the knock rating tests carried out, the engine was operated gradually from lean toward stoichiometric mixture until the onset of knock was observed while other parameters were kept constant. In this research, the engine was considered as operating regularly if the onset of knock did not develop for at least a period of 5 min.

To determine the knock characteristics, the equivalence ratio was varied gradually from lean toward stoichiometric until the

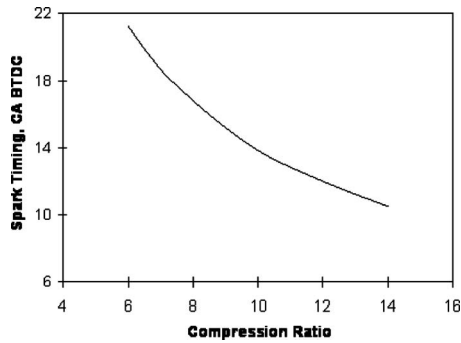


Fig. 1 Variation of spark timing employed with changes in compression ratio

onset of knock was first encountered, while all other operating parameters were kept constant. Throughout, the spark timing was varied with compression ratio according to the relationship shown in Fig. 1, which was employed previously for the knock rating of gaseous fuels and follows a similar ASTM approach [28]. The variations of the knock limited equivalence ratio with changes in fuel composition were determined at a compression ratio of 12:1 and intake temperature of 38°C for fixed spark timing of 12°CA before top dead center (BTDC). This operating condition was used for the comparison of the different fuel. The observed knock limited equivalence ratio remained relatively far from the ignition-failure limits, permitting a relatively wide knock-free operational region. This allowed stable operation and helped to evaluate the knock limits for a wide range of fuel mixtures. The employment of too high compression ratios may result in a very narrow operational range where the engine will either misfire or encounter serious knocking with minor fluctuation in engine operating conditions. The knock limits obtained then will be close to those of the misfire limits, which usually show much weaker variation compared with that of the knock limits [28,40]. In comparison, operation under too low compression ratios may make it impossible to obtain SI knock for some fuel mixtures having a relatively high knock-resisting property.

The combustion duration were determined through analysis of the instantaneous temporal variation of the apparent polytropic index of the pressure-volume record during compression, combustion, and expansion processes, as described by Bade Shrestha and Karim [39]. For any operating conditions, the start of combustion is associated with an apparent detectable increase in the effective polytropic index calculated from the variation of in-cylinder pressure with the change in volume. In comparison, the end of the combustion process is associated with the first apparent occasion when the effective polytropic index approaches a relatively constant value associated with that of the expansion. Hence, the combustion duration is defined as the period in crank angles beginning with the start of combustion to that of the end of combustion.

The commercial CH₄ employed was a pipeline natural gas that had the following average volumetric composition: 93.07%CH₄, 5.74%C₂H₆, 0.16%CO₂, and 1.03%N₂. The purity of iso-octane, CO, and H₂ used were 99.9%, 99.9%, and 99.5%, by volume, respectively.

3 Numerical Simulation Model Development

A two-zone quasidimensional model was employed to predict the performance of SI gas engines including the onset of knock and estimating its intensity whenever it occurs [32,36–38]. The homogeneous charge of the cylinder is assumed at any instant during turbulent flame propagation combustion to be divided into two zones of burned products and unburned reactants that include the end gas region. The intake stroke is simulated while accounting for the presence of residual gases and the charging efficiency. The simulation for the compression stroke is based on the state of

the mixture at the end of the intake stroke. Changes in the unburned reactants are evaluated up to the time of spark ignition while accounting for heat transfer between the bulk fuel-air mixture and surrounding walls [41]. An appropriate combustion energy release pattern developed from experimental evaluation of the fuels combustion behavior over the entire combustion period is employed to simulate the combustion and the corresponding energy release processes after ignition. The composition of the burned products following flame propagation is calculated while accounting for thermodynamic dissociation effects. Changes in the unburned reactants before being consumed by the propagating flame are evaluated using detailed chemical kinetics at the transient temperature and pressure while considering the compression of the piston and propagating flame front with the assumption of uniform pressure through the burned and unburned zones.

Combustion duration of fuel mixtures. The value of the combustion duration is a measure of the inverse of the flame propagation rate. It was shown by Bade Shrestha and Karim [39] that the values of combustion duration of a fuel mixture in the absence of direct experimental data could be estimated adequately by assuming that the individual components contribute to the total combustion duration according to their molar fraction and corresponding combustion duration as fuels on their own when observed under the same operating conditions. The assumption can be expressed as

$$\frac{1}{\Delta\theta_{c,m}} = \frac{y_1}{\Delta\theta_{c,1}} + \frac{y_2}{\Delta\theta_{c,2}} + \dots + \frac{y_i}{\Delta\theta_{c,i}} + \dots \quad (1)$$

where $\Delta\theta_{c,m}$ is the combustion duration for the fuel mixture, y_i is the molar fraction for the i th component in the fuel mixture, and $\Delta\theta_{c,i}$ is the corresponding combustion duration when operating with the i th fuel component on its own under the same operating conditions.

This approach has been successfully applied to predict the performance of SI engine operated on fuel mixtures of CH₄ and H₂ [37,38].

Knock model. Knock in SI engines is due to the auto-ignition of the end gases before being consumed fully by the propagating turbulent flame originated from the spark plug region. In order to monitor the likelihood of the onset of knock, the reactivity of the unburned end gas region of the charge must be evaluated throughout the compression and combustion processes while employing sufficiently detailed description of the reaction kinetics of the charge made up of the fuel, intake air, residual gas, and the exhaust gas recirculated if applicable. The kinetic scheme to be employed must describe the pre-ignition and oxidation reactions of common gaseous fuels such as CH₄, H₂, CO, and their mixtures at conditions relevant to those normally encountered in SI engines that involve lean to stoichiometric mixtures and relatively high temperature and pressures. The presence of other fuels such as C₃H₈ and C₂H₆ can also be accounted for. The scheme employed in this contribution consisted of 155 elementary reaction steps and involved 38 chemical species [37,38].

The net rates of formation and consumption of each species of the reactive end gas charge are calculated at every instant of time to examine the energy released by these oxidation reactions within the end gas region ahead of the propagating flame. The prediction of the onset of knock is based on establishing whether auto-ignition is to take place or not within the end gas region during the flame propagation period [37]. A knock parameter K_n is defined as the calculated temporally-varying pre-ignition energy release within the diminishing end gas mass resulting from preflame-front oxidation reactions per unit of the instantaneous cylinder volume. This energy release is normalized relative to the corresponding total amount of energy to be released through normal flame propagation per unit of cylinder swept volume. This dimensionless function represents also the total increase in-cylinder pressure due to the contribution of the preflame oxidation reaction activity of the instantaneous mass of the end gas relative to the mean effective

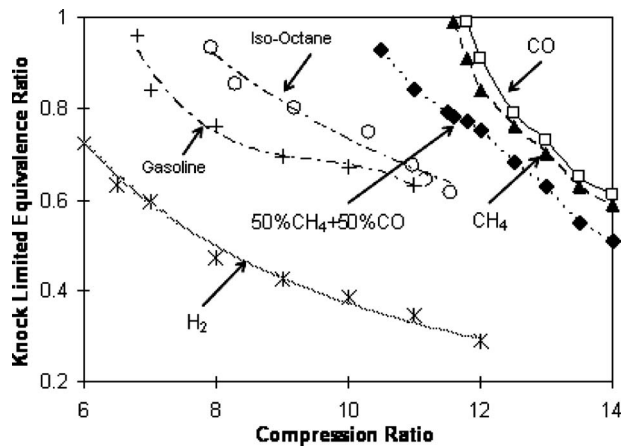


Fig. 2 Variation of the knock limited equivalence ratios with changes in compression ratio while operating on gasoline (Octane No. 92), iso-octane, CO, CH₄, and their mixtures (50%CH₄+50%CO), $T_{in}=38^{\circ}\text{C}$, spark timing as in Fig. 1

tive combustion pressure. Accordingly, the value of K_n is an integrated function of the normalized energetic consequences of any preflame-front oxidation energy release of the diminishing end gas. The knock criterion at any crank angle during flame propagation can be shown to approximate to the following equation:

$$K_n = \frac{h_{sp} - h_t}{\Delta h_o} \times \frac{m_{u,t}}{m_o} \times (CR - 1) \quad (2)$$

where h_{sp} and h_t are the specific enthalpy of the unburned end gas at spark passage and instant time t , respectively, Δh_o is the effective heating value of the fresh charge, $m_{u,t}$ is the corresponding remaining mass of the end gas at the instant time t , and m_o is the mass of the initial fresh charge. This modeling approach, which requires detailed chemical kinetics calculations to evaluate h_t , was shown to be capable of predicting the onset of knock in SI engines operated on CH₄-H₂ mixture [37,38]. It was further shown that for improved prediction with H₂, the energy release may be normalized relative to that of the stoichiometric mixture while the combustion duration normalized relative to that commonly observed for common fuel gases such as CH₄ [32,42]. This is necessary since knock with hydrogen can occur over a very wide range of equivalence ratios well beyond those of the common SI engine fuels that is usually encountered in the neighborhood of the stoichiometric value.

In an engine cycle simulation, changes with time of the values of key variables such as mixture composition, pressure, and temperature for the two zones of the cylinder charge are calculated to yield the instantaneous change in the value of the corresponding K_n of the unburned zone. Should this sequence of calculations find that up to the end of the combustion period the value of K_n never exceeded a critical value of around 1.0, then the cycle is considered to be knock-free. On the other hand, if this is reached before the entire mixture is burned by the propagating flame, then knock is considered to take place shortly after this instant. A more detailed description of the basis for this approach can be found elsewhere [37-39,42]

4 Results and Discussion

The knock limits of pure dry CO, H₂, CH₄, iso-octane, and premium gasoline ((MON+RON)/2=92) were determined experimentally over a wide range of compression ratios. As shown in Fig. 2, both CH₄ and pure CO in dry air have superior knock-resisting qualities compared with gasoline and iso-octane, the common liquid fuel that is considered to have excellent knock-resisting qualities. In comparison, H₂ has a lower resistance to

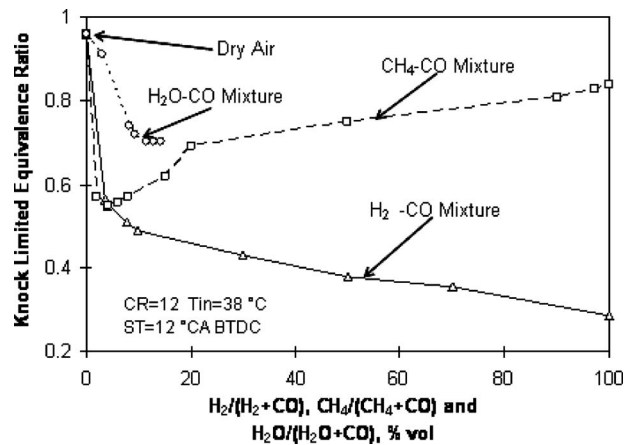


Fig. 3 Variation of the knock limited equivalence ratio with changes in composition of the corresponding binary fuel mixtures of CO with CH₄, H₂, or H₂O

knock compared with the other fuels examined with knocking observed at the compression ratio of 6:1. It is also found that pure CO in dry air has superior knock-resisting qualities to CH₄, which is normally considered to have an excellent resistance to knock among the hydrocarbon fuels usually employed in SI engines. This is mainly due to the exceedingly slow oxidation reactions of pure CO in dry air [43,44]. However, it can be seen that a mixture of CO and CH₄ in the volumetric ratio of 1:1 has significantly lower knock-resisting properties than either of the two component fuels on their own, under the same operating conditions. This is inconsistent with the general knock limit variation trend of the common gaseous fuel mixtures where the knock limits of a binary fuel mixture are normally somewhere between those of its individual components [28,29].

The variation of the experimentally derived knock limits with composition of the binary fuel mixtures of CO and CH₄ at the specified set of operating conditions are shown in Fig. 3. It can be seen that the addition of a very small amount of CH₄ (<5%) to CO burned in dry air can dramatically reduce the knock limited equivalence ratios of the resulting mixtures. The superior knock resistance of CO on its own in dry air becomes much lowered in the presence of CH₄. This is due to the H-containing species enhancing the oxidation reactions of CO with O₂ by providing H-bearing radicals such as HO₂, H, and especially OH during their own oxidation reactions with oxygen [43-45], which accelerate the oxidation reaction rates of CO in dry air and correspondingly lowering its resistance to knock. The superior knock resistance of CO on its own in dry air becomes much lowered in the presence of a relatively small amount of CH₄, a fuel having slightly weaker knock-resisting properties compared with CO as shown in Fig. 2. With the further increase in the proportion of CH₄ beyond that with the lowest knock limit, the knock limited equivalence ratio increases and gradually approach the knock limit of CH₄ when only CH₄ is supplied. The knock-resisting property of CO and CH₄ mixtures is lower than those of both fuel components on their own as shown in Fig. 3. This is not consistent to well accepted knock resistance properties of fuel mixtures [29].

Carbon monoxide is commonly present with H₂ in synthesis gases. The dramatic lowering of the knock limits of CO through the presence of some hydrogen is also shown in Fig. 3. A small amount (<5%) of H₂ addition to CO reduces greatly the knock-resisting properties of the resulting fuel mixture. Further increases in the concentration of H₂ in the binary fuel mixtures will continue to increase gradually the tendency to knock. Despite the significant differences in their corresponding flame propagation rates and knock-resisting properties, the addition of a small amount of H₂ and CH₄ shows similar effects in reducing the

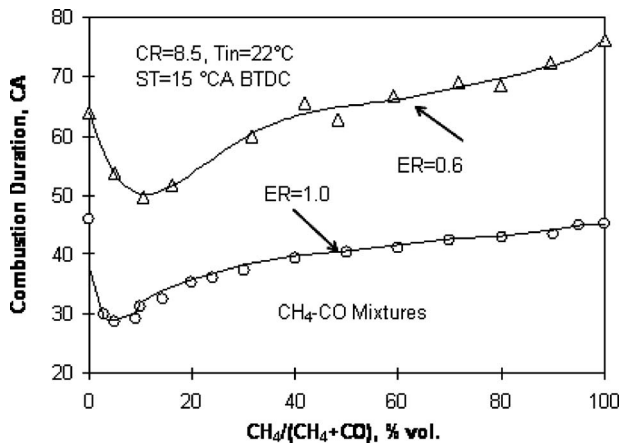


Fig. 4 Variation of the combustion duration with changes in composition of the binary fuel mixture of CO with CH₄ for stoichiometric and lean operations

knock limits of the resulting mixture with CO. This indicates that the effect of the presence of H-containing species in a small percentage wherever they come from plays a more important role than that as a supplementary fuel only such as the cases of adding H₂ to CH₄. It is also noted that a continuing increase in the presence of CH₄ or H₂ in their binary mixtures with CO will increase or continue to decrease gradually the tendency to knock, respectively, reflecting the quite different knock-resisting qualities of H₂ and CH₄ as SI engine fuels. Compared with the addition of CH₄ or H₂, the effect of H₂O on the knock-resisting properties of CO are found to be relatively weak. The addition of H₂O to CO decreases the knock limits of the resulting mixtures but at a lesser rate compared with that of CH₄ and H₂. This is mainly due to the fact that H₂ and CH₄ are fuels, which release energy during oxidation while the endothermic dissociation of H₂O requires relatively large amounts of energy.

The onset of knock in SI engines is also affected by the flame propagation rate, which is cumulatively represented by the inverse of the combustion duration and affects the temperature of the unburned mixture. As shown in Fig. 4, the addition of a small amount of CH₄ to CO reduces dramatically the combustion duration of the resulting mixture for both stoichiometric and lean operations. This would indicate that such an addition enhances the overall flame propagation rates, which increase the temperature of the unburned fuel-air mixture. The minimum combustion duration

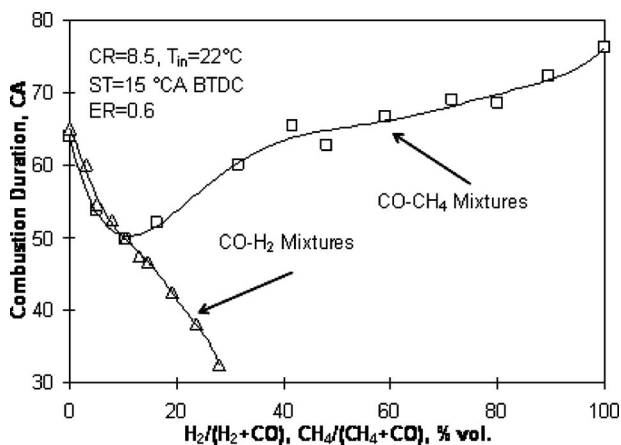


Fig. 5 Variation of the combustion duration with changes in composition of the binary fuel mixture of CO with CH₄ and CO with H₂ at lean mixture operations

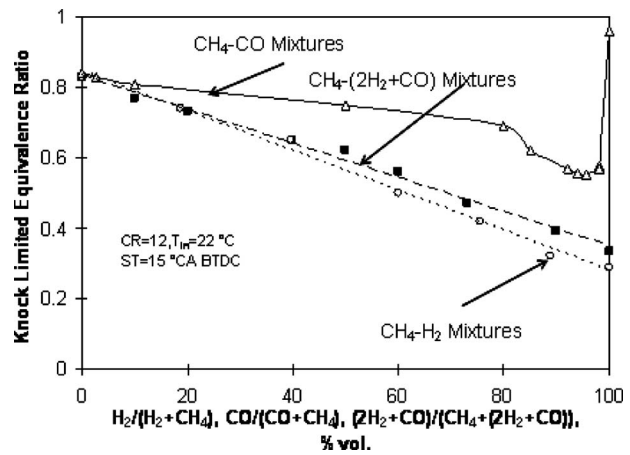


Fig. 6 Variation of the knock limited equivalence ratios with changes in composition of the fuel mixtures of CH₄ with H₂, CH₄ with CO, and CH₄ with H₂-rich gas (2H₂+CO)

was observed when around 5% CH₄ was added to CO in the binary fuel mixtures for stoichiometric operation. However, the minimum combustion duration of the lean mixture operation of an equivalence ratio of 0.6 was observed when around 10% CH₄ was added. A further increase in the relative presence of CH₄ beyond those of the minimum combustion duration increases gradually the combustion duration reflecting the increasing influence of the relatively slow burning properties of CH₄. As observed with CH₄ addition, the addition of a small amount of H₂ produces a comparable effect in speeding up the flame propagation rates, as shown in Fig. 5, for a lean operation case. Further increases in the presence of H₂ continue to enhance the flame propagation rates reflecting the fast flame propagation rates associated with H₂ over both CH₄ and CO. It is obvious that the resulting faster flames appear incapable of stopping the deterioration in the resistance of the resulting fuel mixtures to knock.

A simulated synthesis gas mixture with a volumetric composition of (2H₂+CO), which could have been produced through the partial oxidation or steam reforming of CH₄ or other hydrocarbon fuels, when added to CH₄ reduced the knock limit and combustion duration values as shown in Figs. 6 and 7, respectively. These values are also compared with those obtained when adding H₂ and CO individually to CH₄. As shown in Fig. 6, the knock limited equivalence ratio decreases gradually with the increasing presence of the synthesis gas reflecting its weak knock-resisting properties.

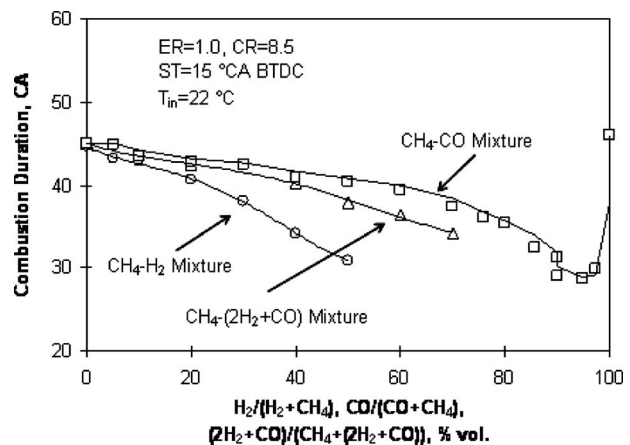


Fig. 7 Variation of the combustion durations with changes in composition of the fuel mixtures when adding H₂, CO, and H₂-rich gas (2H₂+CO) to CH₄

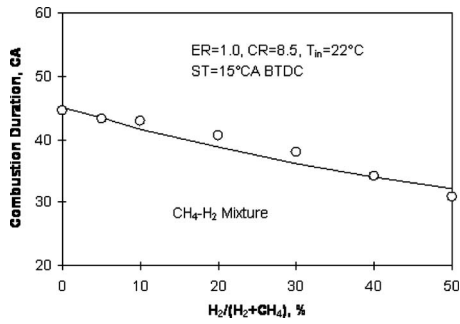


Fig. 8 Comparison of the predicted values of combustion duration with those derived experimentally for binary fuel mixture of H_2 and CH_4

Similarly, the addition of such mixtures to the CH_4 reduced the combustion durations to produce increasingly faster flame propagation, as shown in Fig. 7. The effect of the H_2 -rich synthesis gases addition tends to be relatively weak in comparison to that of pure H_2 addition but greater than that of adding pure CO to CH_4 . The addition of CO to CH_4 enhances the flame propagation rates and increases the tendency of the resulting fuel mixtures to knock, although pure CO in dry air shows a slower propagating flame and superior knock-resisting properties in comparison to those of CH_4 . Due to the likely presence of a large amount of diluents such as N_2 , CO_2 , and H_2O in the H_2 -rich gases produced on board, their addition to traditional fuels is expected to have a much weaker effect than this simulated synthesis gas in affecting the combustion process and the knock-resisting properties.

The value of the combustion duration is an important parameter in the modeling approach adopted. As indicated earlier, values of the combustion duration of gaseous fuel mixtures could be estimated following the examination of the combustion durations of the components of the fuel mixtures on their own under the same operating conditions in accordance with Eq. (1) [37,39]. As an example, Fig. 8 shows that the predicted combustion duration agrees well with the corresponding experimental values obtained for fuel mixtures of CH_4 and H_2 . As shown in Fig. 9, the knock limits of this fuel mixture can also be well predicted.

However, it was found that the behavior of CH_4 and CO cannot be considered as that of a simple case of fuel mixing. This is because of the combustion behavior of CO in binary mixture with H-bearing species is quite different from that of pure CO in dry air. When a "nondry" CO , given the symbol $CO_{n,d}$ is considered, then the apparent value of combustion duration of $CO_{n,d}$ can be derived from experimental observations using the equation developed by Bade Shrestha and Karim [39]. The averaged value of the estimated combustion durations of this nondry CO was found to be approximately 70% of that of the pure CO in dry air [46]. By

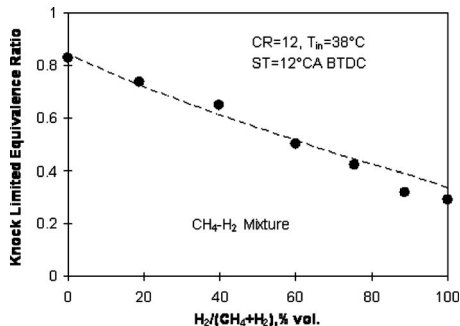


Fig. 9 Comparison of the predicted knock limited equivalence ratios with those derived experimentally for binary mixtures of H_2 and CH_4

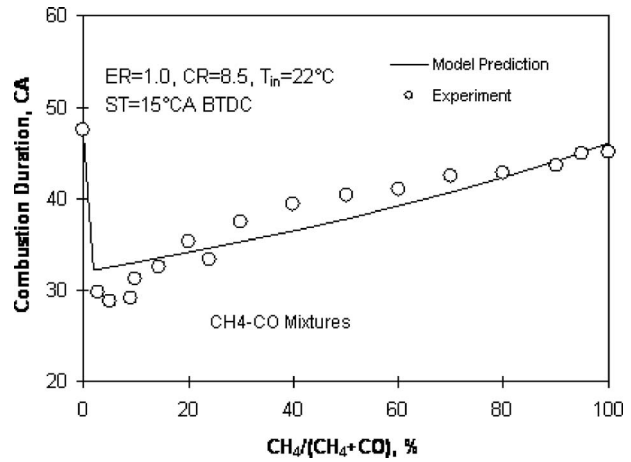


Fig. 10 Comparison of the predicted combustion duration with those determined experimentally for binary fuel mixtures of CO and CH_4 at stoichiometric operation

considering the nondry CO , the behavior of fuel mixtures involving CO and CH_4 , an H-containing fuel component, could be successfully predicted as shown in Figs. 10 and 11 for combustion duration and indicated power, respectively. As shown in Figs. 12 and 13, the predicted values of indicated work and knock limited equivalence ratios agree well with the corresponding experimental data obtained in this investigation for fuel mixtures of CO with simulated synthesis gases ($2H_2+CO$). However, there is no satisfactory agreement of the predicted knock limit values for fuel mixtures of CO with small amounts of H_2 or CH_4 as shown in Fig. 14. It is suggested that the very small amount of lubricating oil that is inevitably consumed through volatilization, cracking, and oxidation during the thermally and chemically intense combustion process under the conditions approaching the onset of knock may account significantly for this disagreement [4]. Due to the very small amounts commonly involved, the thermal consequences of lubricating oil oxidation in affecting the performance of SI engines are commonly ignored. However, these minute amounts of lubricating oil consumed may be sufficient to contribute the small amounts of H-bearing species that enhance sufficiently the oxidation of CO to produce the observed increase in the tendency to knock.

Of course, lubricating oils consist of mainly heavy hydrocarbon components of complex composition and chemical structures.

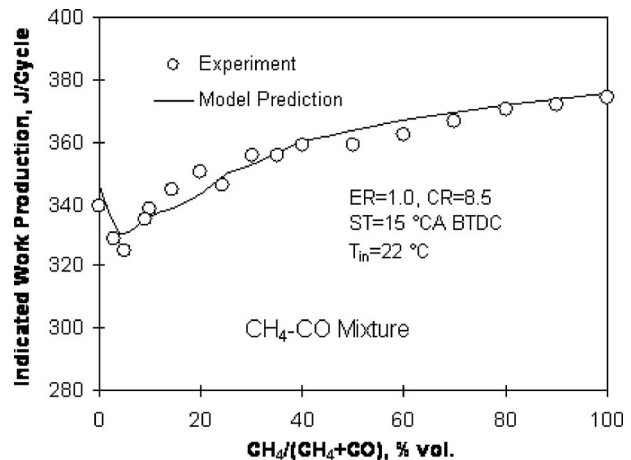


Fig. 11 Comparison of the predicted indicated work production with those determined experimentally for binary mixtures of CO and CH_4 for stoichiometric operation

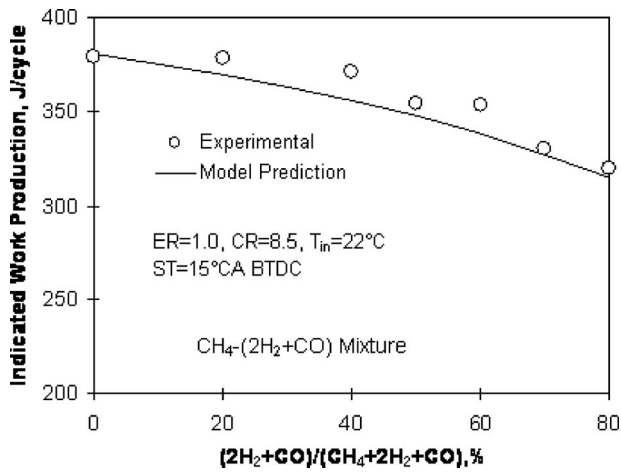


Fig. 12 Comparison of the predicted indicated power output with those derived experimentally for fuel mixtures of CH_4 with simulated synthesis gases ($2\text{H}_2+\text{CO}$) under stoichiometric operation

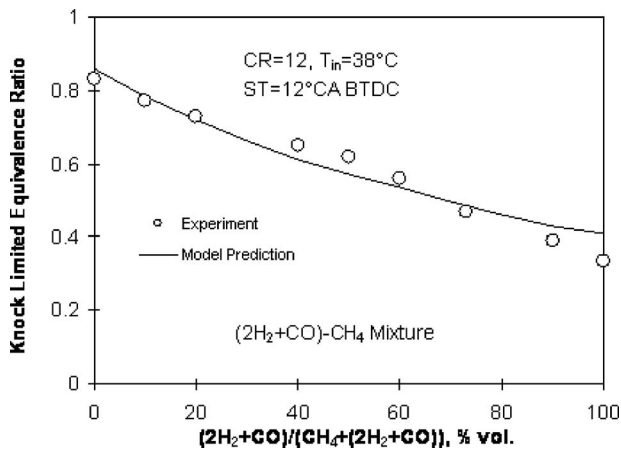


Fig. 13 Comparison of the predicted knock limited equivalence ratio with those determined experimentally for fuel mixtures of CH_4 and simulated synthesis gas ($2\text{H}_2+\text{CO}$)

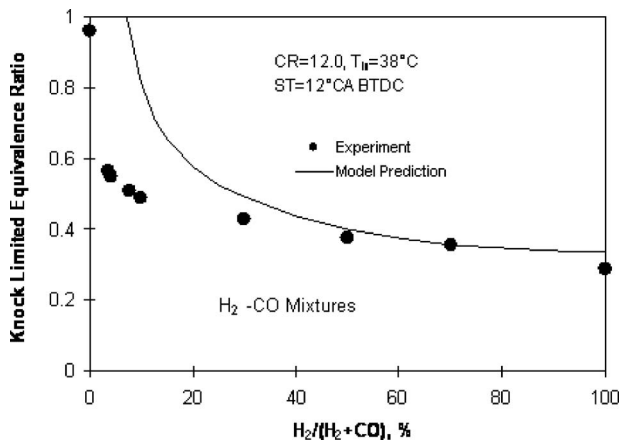


Fig. 14 Comparison of the predicted knock limited equivalence ratio with those determined experimentally for binary fuel mixtures of H_2 and CO

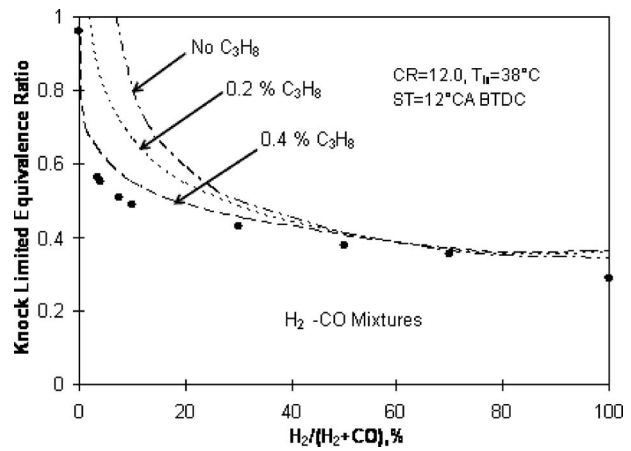


Fig. 15 Effect of adding a small amount of C_3H_8 to the intake fuels during modeling in improving the agreement of the predicted knock limited equivalence ratios with those determined experimentally for binary fuel mixtures of H_2 and CO

Their corresponding reactions with O_2 and other species are largely unknown. In order to illustrate the plausible kinetic role of lubricating oil and its dissociation and oxidation products in affecting the oxidation of dry CO , a relatively very small amount of a higher hydrocarbon that may be considered comparable to the amount of lubricating oil consumed was artificially assumed to have been also involved during the oxidation of intake fuels. Figure 15 shows the results of an example when small amounts of C_3H_8 were assumed to be present to contribute to the end gas reactions that lead to the onset of knock. Such a remedial approach could produce a better agreement between predicted knock limits and corresponding experimental values of the fuel mixtures containing mainly CO in the presence of a relatively small amount of H_2 . In comparison, their effects are quite small for fuel mixtures containing relatively large amounts of H_2 . As shown in Fig. 15, the assumed presence of merely 0.4% C_3H_8 within the fuel produced sufficiently good agreement with those derived experimentally. Similarly, the application of such an approach to fuel mixtures of CO with small amounts of CH_4 also produced improved agreement of predicted knock limits with those observed experimentally, as shown in Fig. 16.

5 Conclusions

The addition of small amounts of H-containing compounds such as H_2 , CH_4 , or H_2O to CO accelerates the overall flame propagation rates and lowers the knock-resisting properties of CO . The resulting faster flames appear incapable of stopping the deterioration in the resistance to knock of the resulting fuel mixtures.

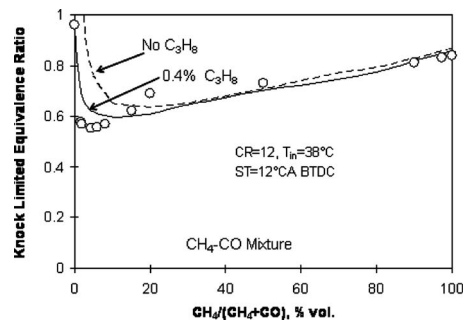


Fig. 16 Effect of adding a small amount of C_3H_8 to the intake fuels during modeling in improving the agreements of the predicted knock limited equivalence ratios with those determined experimentally for binary fuel mixtures of CH_4 and CO

The addition of a simulated synthesis gas ($2\text{H}_2+\text{CO}$) to CH_4 enhances the combustion process of the resulting mixture and lowers its knock resistance. The effectiveness of such an addition is slightly weaker than that of a comparable H_2 addition but much stronger than that with CO addition only.

Through accounting for a nondry CO in the presence of H-bearing fuels, the engine performance and knock limits could be predicted well over a wide range of composition of fuel mixtures of H_2 , CO , and CH_4 .

The prediction of the knock limits of fuel mixtures containing mainly CO with very small amounts of H-containing species could provide a better agreement with corresponding experimental values when a very small concentration of a higher hydrocarbon fuel such as propane was considered present in the end gas region. Such an approach could be considered to account for the possible contribution of the very small amounts of lubricating oil that normally get consumed at near the knock operational limits.

Acknowledgment

The financial support of the National Science and Engineering Research Council (NSERC) of Canada is gratefully acknowledged.

Nomenclature

- K_n = knock criterion
 h_{sp} = enthalpy of the mixture at spark timing, kJ/kg
 h_{st} = enthalpy of the mixture at any instant "t," kJ/kg
 m_0 = mass of the fresh charge, kg
 m_u = mass of the end gas at any instant "t," kg

References

- [1] Karim, G. A., 2003, "Hydrogen as a Spark Ignition Engine Fuel," *Int. J. Hydrogen Energy*, **28**, pp. 569–577.
- [2] White, C. M., Steeper, R. R., and Lutz, A. E., 2006, "The Hydrogen-Fueled Internal Combustion Engine: A Technical Review," *Int. J. Hydrogen Energy*, **31**, pp. 1292–1305.
- [3] Furuhashi, S., and Hiruma, M., 1977, "Some Characteristics of Oil Consumption Measured by Hydrogen Fueled Engine," *Journal of American Society of Lubrication Engineers*, **34**(12), pp. 665–675.
- [4] Li, H. L., and Karim, G. A., 2004, "Examination of the Oil Consumption in a S.I. Hydrogen Engine," SAE Paper No. 2004-01-2916.
- [5] Kido, H., Huang, S., Tanoue, K., and Nitta, T., 1994, "Improving the Combustion Performance of Lean Hydrocarbon Mixture by Hydrogen Addition," *JSAE Rev.*, **15**(2), pp. 165–170.
- [6] Bauer, C. G., and Forest, T. W., 2001, "Effect of Hydrogen Addition on the Performance of Methane-Fueled Vehicles. Part I: Effect on S.I. Engine Performance," *Int. J. Hydrogen Energy*, **26**, pp. 55–70.
- [7] Herdin, G., 2007, "Hydrogen and Hydrogen Mixtures as Fuel in Stationary Gas Engines," SAE Paper No. 2007-01-0012.
- [8] Lattin, W. C., and Utgikar, V. P., 2007, "Transition to Hydrogen Economy in the United States: A 2006 Status Report," *Int. J. Hydrogen Energy*, **32**, pp. 3230–3237.
- [9] Marban, G., and Valdes-Solis, T., 2007, "Towards the Hydrogen Economy?," *Int. J. Hydrogen Energy*, **32**, pp. 1625–1637.
- [10] Holt, D. J., 2003, "Hydrogen and Its Future as a Transportation Fuel," SAE Paper No. PT-95.
- [11] Liu, X. H., Liu, F. S., Zhou, L., Sun, B. G., and Schock, H., 2008, "Backfire Prediction in a Manifold Injection Hydrogen Internal Combustion Engine," *Int. J. Hydrogen Energy*, **33**, pp. 3847–3855.
- [12] Jamal, Y., and Wyszynski, M. L., 1994, "On-Board Generation of Hydrogen-Rich Gaseous Fuels—A Review," *Int. J. Hydrogen Energy*, **19**, pp. 557–572.
- [13] Rabinovich, A., Bromberg, L., Cohn, D. R., Surma, J., and Virden, J. W., 1998, "Onboard Plasmatron Reforming of Bifuels, Gasoline and Diesel Fuel," SAE Paper No. 981920.
- [14] Karim, G. A., and Wierzbka, I., 2008, "The Production of Hydrogen Through the Uncatalyzed Partial Oxidation of Methane in an Internal Combustion Engine," *Int. J. Hydrogen Energy*, **33**, pp. 2105–2110.
- [15] Shudo, T., Nakajima, Y., and Tsuga, K., 2001, "Combustion Characteristics of H_2 - CO - CO_2 Mixture in an IC Engine," SAE Paper No. 2001-01-0252.
- [16] Andreatta, D., and Dibble, R. W., 1996, "An Experimental Study of Air-Reformed Natural Gas in S.I. Engines," SAE Paper No. 960852.
- [17] Bromberg, L., Cohn, D. R., Rabinovich, A., and Heywood, J. B., 2001, "Emissions Reductions Using Hydrogen From Plasmatron Fuel Converters," *Int. J. Hydrogen Energy*, **26**, pp. 1115–1121.

- [18] Allenby, S., Chang, W.-C., Megaritis, A., and Wyszynski, M. L., 2001, "Hydrogen Enrichment: A Way to Maintain Combustion Stability in a Natural Gas Fueled Engine With Exhaust Gas Recirculation, the Potential of Fuel Reforming," *Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.)*, **215**, pp. 405–418.
- [19] Quader, A. A., Kirwan, J. E., and Grieve, M. J., 2003, "Engine Performance and Emissions Near the Dilute Limit With Hydrogen Enrichment Using an On-Board Reforming Strategy," SAE Paper No. 2003-01-1356.
- [20] Tsolakis, A., Megaritis, A., and Wyszynski, M. L., 2003, "Application of Exhaust Gas Reforming in Compression Ignition Engines Fueled by Diesel and Biodiesel Fuel Mixtures," *Energy Fuels*, **17**, pp. 1464–1473.
- [21] Tsolakis, A., and Megaritis, A., 2005, "Combustion Characteristics and Exhaust Gas Emissions of a Diesel Engine Supplied With Reformed EGR," SAE Paper No. 2005-01-2087.
- [22] Abu-Jrai, A., Tsolakis, A., and Megaritis, A., 2007, "The Influence of H_2 and CO on Diesel Engine Combustion Characteristics, Exhaust Emissions, and After Treatment Selective Catalytic NO_x Reduction," *Int. J. Hydrogen Energy*, **32**, pp. 3565–3571.
- [23] Zheng, M., Asad, U., Reader, G., Han, X., Pournazeri, M., Ting, T., and Wang, M., 2008, "Diesel EGR Fuel Reformer Improvement with Flow Reversal and Central Fueling," SAE Paper No. 2008-01-1607.
- [24] Ashur, M., Misztal, J., Wyszynski, M. L., Tsolakis, A., Xu, H. M., Qiao, J., and Golunski, S., 2007, "Onboard Exhaust Gas Reforming of Gasoline Using Integrated Reformer and TWC," Eighth International Conference on Engines for Automobiles, Capri (Naples), Italy, Sept., SAE Paper No. 2007-24-0078.
- [25] Isherwood, K. D., Linna, J. R., and Loftus, P. J., 1998, "Using On-Board Fuel Reforming by Partial Oxidation to Improve SI Engine Cold-Start Performance and Emissions," SAE Paper No. 980939.
- [26] Hosseini, V., and Checkel, M. D., 2008, "Reformer Gas Composition Effect on HCCI Combustion of *n*-Heptane, Iso-Octane, and Natural Gas," SAE Paper No. 2008-01-0049.
- [27] Topinka, J. A., Gerty, M. D., and Heywood, J. B., 2004, "Knock Behavior of a Lean-Burn, H_2 and CO Enhanced, SI Gasoline Engine Concept," SAE Paper No. 2004-01-0975.
- [28] Karim, G. A., and Klat, S. R., 1966, "The Measurement of the Mass Flow Rate of Different Gases Using a Choked Nozzle," *Lab. Pract.*, **15**, pp. 184–186.
- [29] Attar, A. A., and Karim, G. A., 2003, "Knock Rating of Gaseous Fuels," *ASME J. Eng. Gas Turbines Power*, **125**, pp. 41–47.
- [30] Owen, K., and Coley, T., 1995, *Automotive Fuels Reference Book*, 2nd ed., Society of Automotive Engineers, Warrendale, PA.
- [31] Ryan, T. W., Callahan, T. J., and King, S. R., 1993, "Engine Knock Rating of Natural Gases-Methane Number," *ASME J. Eng. Gas Turbines Power*, **115**, pp. 769–776.
- [32] Li, H. L., and Karim, G. A., 2006, "Hydrogen Fueled Spark-Ignition Engines Predictive and Experimental Performance," *ASME J. Eng. Gas Turbines Power*, **128**(1), pp. 230–236.
- [33] Anthonissen, E., and Wallace, J. S., 1983, "Dissociated Methanol Engine Testing Results Using H_2 - CO Mixtures," *Proceedings of the 18th Intersociety Energy Conversion Engineering Conference*, Aug. 21–26, pp. 549–556.
- [34] Kirwan, J. E., Quader, A. A., and Grieve, M. J., 1999, "Advanced Engine Management Using On-Board Gasoline Partial Oxidation Reforming for Meeting Super-ULEV (SULEV) Emissions Standards," SAE Paper No. 1999-01-2927.
- [35] Rodrigues, R., and Bade Shrestha, S. O., 2006 "Knock Rating of Gaseous Fuels in the Presence of Diluents," SAE Paper No. 2006-01-3429.
- [36] Li, H. L., and Karim, G. A., 2008, "Modeling the Performance of a Turbo-Charged S.I. Natural Gas Engine With Cooled EGR," *ASME J. Eng. Gas Turbines Power*, **130**, p. 032804.
- [37] Bade Shrestha, S. O., and Karim, G. A., 1999, "A Predictive Model for Gas Fueled Spark Ignition Engine Applications," SAE Paper No. 1999-01-3482.
- [38] Bade Shrestha, S. O., and Karim, G. A., 1999, "Hydrogen as an Additive to Methane for Spark Ignition Engine Applications," *Int. J. Hydrogen Energy*, **24**, pp. 577–586.
- [39] Bade Shrestha, S. O., and Karim, G. A., 2001, "An Experimental and Analytical Examination of the Combustion Period for Gas-Fueled Spark Ignition Engine Applications," *Proc. Inst. Mech. Eng., Part A*, **215**, pp. 1–12.
- [40] Bade Shrestha, S. O., and Karim, G. A., 2006, "The Operational Mixture Limits in Engines Fueled by Alternative Gaseous Fuels," *ASME J. Energy Resour. Technol.*, **128**(3), pp. 223–228.
- [41] Woschni, G., 1967, "A Universal Applicable Equation for the Instantaneous Heat Transfer Coefficient in the Internal Combustion Engines," *SAE Trans.*, **76**, pp. 3065–3083.
- [42] Karim, G. A., 2004, "A Dimensionless Criterion for Predicting the Onset of Knock in Spark Ignition Engine," SAE Paper No. 2004-01-1992.
- [43] Glassman, I., 1987, *Combustion*, Academic, New York.
- [44] Turns, S. R., 1996, *An Introduction to Combustion: Concepts and Applications*, McGraw-Hill, New York.
- [45] Mclean, I. C., Smith, D. B., and Taylor, S. C., 1994, "The Use of Carbon Monoxide/Hydrogen Burning Velocities to Examine the Rate of the $\text{CO}+\text{OH}$ Reaction," 25th Symposium (International) on Combustion, pp. 749–757.
- [46] Li, H. L., 2004, "An Experimental and Analytical Examination of Gas Fueled Spark Ignition Engine—Performance and Combustion," Ph.D. thesis, Department of Mechanical and Manufacturing Engineering, University of Calgary, Alberta, Canada.

Kinetic Study of the Thermo-Oxidative Degradation of Squalane (C₃₀H₆₂) Modeling the Base Oil of Engine Lubricants

Moussa Diaby

e-mail: moussa@dcmr.polytechnique.fr

Michel Sablier

e-mail: michel.sablier@dcmr.polytechnique.fr

Ecole Polytechnique,
Laboratoire des Mécanismes Réactionnels,
CNRS 91128 Palaiseau Cedex, France

Anthony Le Negrate

e-mail: anthony.lenegrate@mpsa.com

Mehdi El Fassi

e-mail: mehdi.elfassi@mpsa.com

PSA Peugeot Citroën,
Centre Technique de Vélizy,
78943 Vélizy-Villacoublay Cedex, France

On the basis of ongoing research conducted on the clarification of processes responsible for lubricant degradation in the environment of piston grooves in exhaust gas recirculation (EGR) diesel engines, an experimental investigation was aimed to develop a kinetic model, which can be used for the prediction of lubricant oxidative degradation correlated with endurance test conducted on engines. Knowing that base oils are a complex blend of paraffins and naphthenes with a wide range of sizes and structures, their chemistry analysis during the oxidation process can be highly convoluted. In the present work, investigations were carried out with the squalane (C₃₀H₆₂) chosen for its physical and chemical similarities with the lubricant base oils used during the investigations. Thermo-oxidative degradation of this hydrocarbon was conducted at atmospheric pressure in a tubular furnace, while varying temperature and duration of the tests in order to establish an oxidation reaction rate law. The same experimental procedures were applied to squalane doped with two different phenolic antioxidants usually present in engine oil composition: 2,6-di-tert-butyl-4-methylphenol and octadecyl-3-(3,5-di-tert-butyl-4-hydroxyphenyl)propionate. Thus, the effect of both antioxidants on the oxidation rate law was investigated. Data analysis of the oxidized samples (Fourier transform infrared spectroscopy and gas chromatography/mass spectrometry) allowed rationalization of the thermo-oxidative degradation of squalane. The resulting kinetic modeling provides a practical analytical tool to follow the thermal degradation processes, which can be used for prediction of base oil hydrocarbon aging. If experiments confirmed the role of phenolic additives as an effective agent to lower oxidation rates, the main results lie in the observation of a threshold temperature where a reversed activity of these additives was observed. [DOI: 10.1115/1.3155797]

1 Introduction

Lubricants are essential to the good operation of the automotive engines [1,2]. In fact, engine lubricants are complex mixtures involving base oils, which should have their characteristics as close as possible to those of the referred lubricant, and many different additives (detergent, dispersant, viscosity improvers, antioxidants, etc.) [3–5]. In the operating conditions of engines, lubricants are susceptible to oxidation, which leads to the formation of low-molecular weight materials such as aldehydes, ketones, acids, and alcohols (Fig. 1) [6–9].

These oxidation products not only affect the properties of lubricants, but also undergo polymerization reactions that produce high-molecular-weight materials resulting in the formation of sludge and varnishes responsible for the failure of lubrication by increasing the lubricant viscosity and promoting the engine wear [10–16].

Then antioxidant additives are added in the lubricant to ensure a high oil oxidative stability. Several classes of antioxidants are used during the formulation [17–21]. Two main classes inhibit the initiation of thermal oxidation: (1) peroxide decomposers and (2) metal deactivators [19]. The most important classes of antioxidants that inhibit the propagation step are hindered phenols [9,22], secondary alkyl arylamines, and diarylamines [9]. The mechanism

currently admitted [9] for inhibition consists in capturing the free radicals generated during the oxidation reaction (Fig. 2).

Previous studies have shown that the oxidative degradation was directly correlated with the consumption of phenolic antioxidant additives [23]. This fact was more noticeable at high temperature. In addition, the down-sizing and the extreme conditions of the engine operation (e.g., operational temperature and pressure of the combustion chamber, the oil sump, or the cylinder liner temperatures) could speed up the process of lubricating oil degradation and then the increase in deposit formation in the engine parts [23]. This raises the question of drain interval that should always be either identical or longer for the convenience of the customer.

The current work aimed to develop a kinetic model, which can be used for prediction of lubricant oxidative degradation. Knowing that base oils are a complex blend of linear saturated hydrocarbons (paraffins) and cyclic saturated hydrocarbons (naphthenes) with a wide range of sizes and structures, their chemistry analysis during the oxidation process can be highly convoluted. To simplify the analysis, common strategies consist in using single compounds, which mimic the lubricant. For example, Blaine and Savage [24–26] performed a detailed investigation of hexadecane oxidation under conditions similar to those under which a mineral oil or natural lubricants undergo degradation. Pfaendtner and Broadbelt [27] developed a mechanistic model of lubricant degradation by studying the auto-oxidation of decane and octane. Experimentally, condensed-phase auto-oxidation of alkanes has been studied extensively, and there are numerous studies reported in literature. Garcia-Ochoa et al. [28] studied the oxidation of *n*-octane over 135–145°C.

Similarly to this approach, our experiments were conducted on

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 9, 2009; final manuscript received April 7, 2009; published online November 30, 2009. Review conducted by Dilip R. Ballal. Paper presented at the ASME Internal Combustion Engine Division 2009, Milwaukee, WI, May 3–6, 2009.

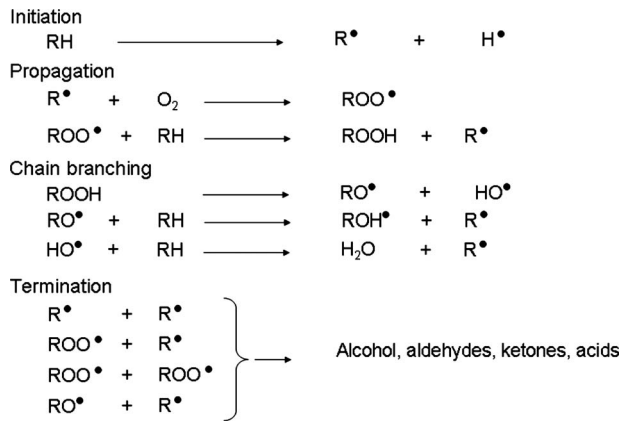


Fig. 1 Mechanism of low-molecular weight material formation [8]

a model compound: squalane ($C_{30}H_{62}$) chosen for its similarities in terms of physical and chemical similarities with the group III base oils of lubricant (grade SAE 5W-30) used during previous investigations on the degradation of engine oil [23]. Accelerated oxidation tests were carried out for squalane, without additives, at different temperatures and durations in order to establish an oxidation reaction rate law.

The same experimental procedure was applied to squalane doped with two different phenolic antioxidants (Fig. 3): (a) 2,6-di-*tert*-butyl-4-methylphenol (BHT) and (b) octadecyl-3-(3,5-di-*tert*-butyl-4-hydroxyphenyl)propionate (OBHP). The effect of both antioxidants was investigated.

2 Experiment

2.1 Model Compounds: Squalane. A model compound consisting of a single chemical has been used to simulate base oil. The conditions for the choice of this compound were its simplicity, its chemical structure preferably symmetric, its cost, and its probability to behave similarly to the base oil (group III) of the lubricant (grade SAE 5W-30) currently used during our investigations on engine oil degradation processes. Squalane, $C_{30}H_{62}$ (Fig. 4), appeared as a good model, with 30 carbons, and a highly symmetric structure. The physical characteristics of squalane are reported in Table 1.

2.2 Thermal Oxidation Tests. Thermo-oxidative degradation tests were carried out in a tubular furnace. 300 μ l of squalane was placed in a preweighed vessel. The vessel was subsequently weighed in order to measure the content weight of squalane. Then the sample was directly introduced in the furnace already heated and kept at the desired constant temperature. The arrival of oxygen in the furnace was ensured by the ambient air flow due to



Fig. 2 Antioxidant mechanism for radical scavengers

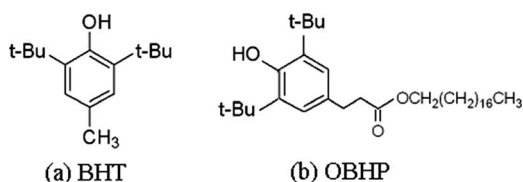


Fig. 3 Chemical structures of (a) BHT and (b) OBHP

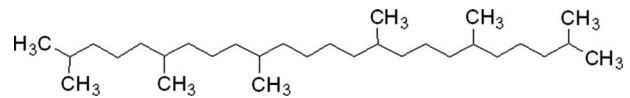


Fig. 4 Chemical structure of squalane ($C_{30}H_{62}$)

natural convection (air draft). The samples were oxidized at 150°C, 180°C, and 205°C, respectively. The temperature of 150°C was chosen to simulate the extreme conditions of operation in the engine oil sump, and the temperatures of 180°C and 205°C are the average temperature of the cylinder wall and the first ring groove of pistons. Experiments were conducted at different times from 6 h to 72 h. After each test, the vessel was removed from the furnace and weighed again. The mass loss found represented the vaporized fraction.

For the tests with antioxidants, experiments were conducted with each antioxidant separately (BHT or OBHP). About 0.5% of antioxidant was added in squalane. Then the same experimental procedure as previously described for the thermo-oxidative degradation test of squalane was applied in the presence of antioxidant.

2.3 Chemical Analyses. Fourier transform infrared (FTIR) spectroscopy and gas chromatography/mass spectrometry (GC/MS) were used to characterize the condensed-phase samples obtained during the thermo-oxidative degradation tests.

2.3.1 FTIR Spectroscopy. The presence of oxidation products was monitored by infrared spectroscopy. Oxidized samples were investigated with a Perkin Elmer Paragon 1000 FTIR spectrometer (Perkin Elmer, Courtaboeuf, France). Wavelengths were scanned in the range 600–4000 cm^{-1} . Each FTIR spectrum was the result of four co-added scans at a nominal resolution of 4.0 cm^{-1} . NaCl pellets were used for the recording of data. The oxidation of squalane was quantified by calculating the area of the spectral band around 1720 cm^{-1} , measured from 1650 cm^{-1} to 1850 cm^{-1} , and the spectral band around 3450 cm^{-1} , measured from 3200 cm^{-1} to 3600 cm^{-1} . These bands are characteristics of carbonyl bond (C=O) and hydroxyl bond (OH) vibrations present in most of the oxidation products [29–32]. This method is used and acknowledged as a satisfactory measure of the hydrocarbon degradation level, since the carbonyl and hydroxyl bands are not found in the spectrum of the fresh nonoxidized squalane used as a reference for FTIR spectra interpretation.

2.3.2 GC/MS Analysis Conditions. Characterization of chemicals and identification of oxidation products were carried out on a Varian 3800 gas chromatograph and coupled to a Varian Saturn 2000 ion trap mass spectrometer (Varian, Les Ulis, France). Chromatographic separation was performed on a Chrompack DB-5 capillary column (30 m length, 0.25 μ m film thickness, 0.25 mm ID, and 5% phenyl–95% methyl polysiloxane). The gas chromatograph was operated at a constant helium flow of 1.0 $ml\ min^{-1}$ and a 1:40 split ratio. The injector was set at 300°C. The oven programmed temperature was set at 40°C for 2 min, and ramped at a rate of 6°C min^{-1} up to 300°C where it was held for 20 min. The total duration of GC analysis was ~65 min. The manifold, ion trap source, and transfer line temperatures were set at 120°C, 220°C, and 300°C, respectively. The SATURN software was used under the following conditions. Mass spectra were scanned on the m/z 30–550 range with a rate of 1 scan s^{-1} . Two ionization modes were used: electronic ionization (EI) and chemical ioniza-

Table 1 Physical characteristic of squalane

Formula	Mol. weight (g mol^{-1})	Boiling point at 1 Torr (°C)	Density (at 20°C)	Viscosity (mPa s)
$C_{30}H_{62}$	422.83	210–215	0.809	28.254

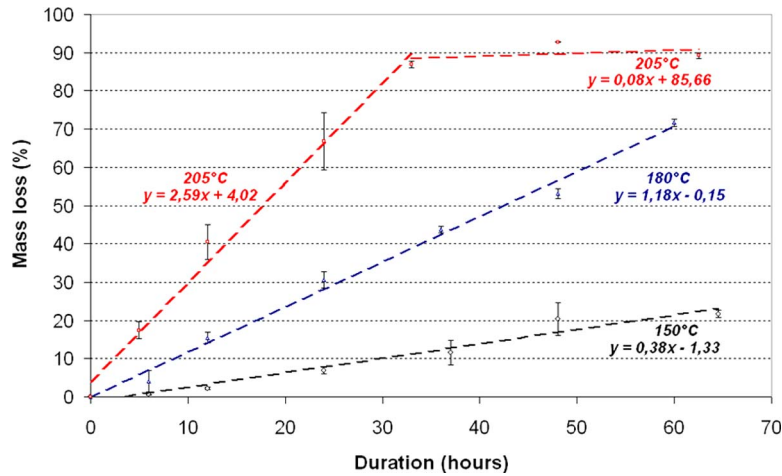


Fig. 5 Mass loss of squalane versus the test duration without antioxidant additives

tion (CI). In the first case, the electron ionization energy was 70 eV, the emission current was 10 μA , and the multiplier voltage was 2200 V; spectra were recorded using the automatic gain control (AGC) function with a target of 20,000. Chemical ionization was performed using methanol as the reactant. In this case, the emission current was set to 20 μA , the multiplier voltage to 2200 V, and the AGC target to 5000.

3 Results and Discussions

3.1 Characterization of Oxidized Samples. As previously notified, samples were weighed before and after the oxidation tests in order to determine by difference the percentage of the evaporated fraction. Figure 5 shows the evolution of mass versus the test duration for experiments carried out without antioxidant additives. One can notice a net increase in the evaporation rate with the temperature. But experiment carried out at 205°C presents two different slopes attributed to two different evaporation rates. To rationalize the lower evaporation rate at 205°C and for the longer experiment durations (above 30 h), one can assume that the formation of heavy products at this high temperature value and after a long time of experiments may result in a noticeable decrease in the evaporation rate correlated with the high-molecular weight product contents resulting from extended polymerization processes.

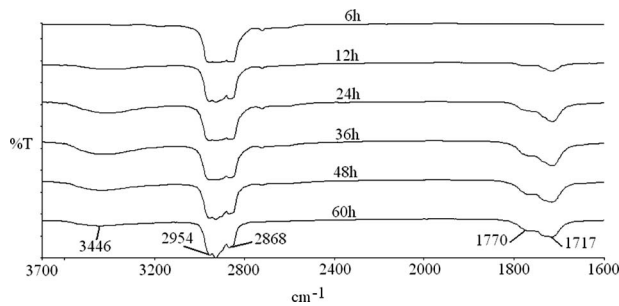


Fig. 6 FTIR spectra of oxidized sample at 180°C and different durations of the tests

The infrared spectra of the oxidized samples, as exemplified in Fig. 6 for experiments at 180°C without antioxidant additives, confirm the presence of carbonyl bands around 1720 cm^{-1} and hydroxyl bands around 3450 cm^{-1} . These characteristic bands of oxidation increase with the test duration.

Based on the measurement of IR band areas, a simple kinetic model presented by Crisostomo et al. [33] allowed the estimation

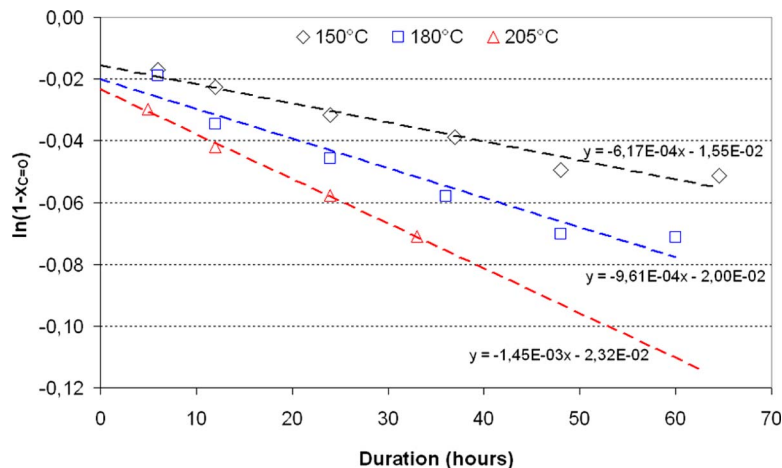


Fig. 7 Plot of $\ln(1 - X_{\text{C=O}})$ against time of oxidation tests without antioxidant additives

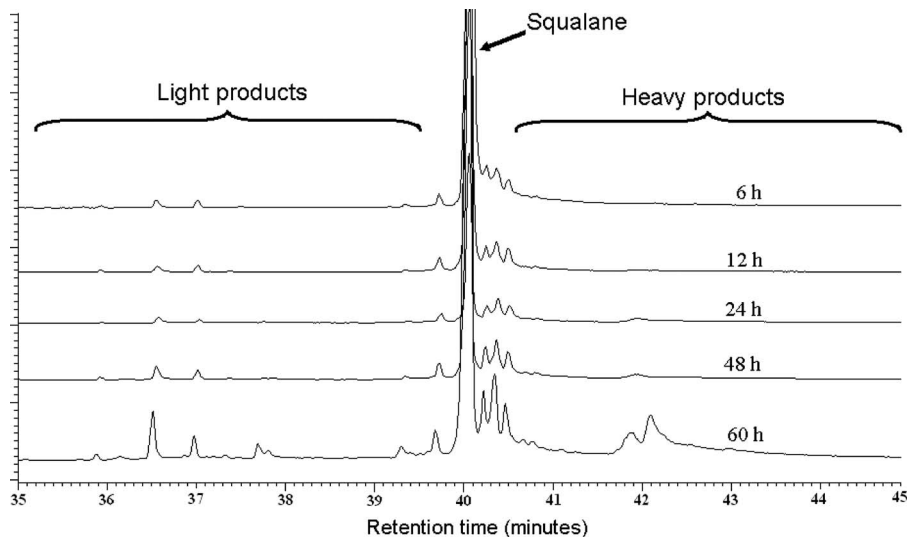


Fig. 8 Chromatogram of an oxidized sample at 180°C and different durations of the tests without antioxidant additives

of the global activation energy of the oxidation reaction using the experimentally measured areas of the carbonyl bands detected along the oxidation tests. According to this model, the plot of $\ln(1 - X_{C=O})$ versus t (where $X_{C=O}$ represents the carbonyl fraction and t the test duration) as presented in Fig. 7 for oxidation reaction carried out without antioxidants gives lines with a slope equal to the rate constant k . So, applying the Arrhenius equation, $\ln(k) = \ln(k_0) - Ea/RT$, to the rate constant values obtained for the three different temperatures, the activation energy was found to be 6 kcal mol⁻¹. This value is close to the activation energy, between 9 kcal mol⁻¹ and 15 kcal mol⁻¹, reported in literature for oxidation reaction of several lubricating oils [33–37].

In order to better characterize the oxidative degradation products, oxidized samples were analyzed by gas chromatography coupled with mass spectrometry. The analyses showed clearly several classes of products, such as alcohols, aldehydes, ketones, acids, and light hydrocarbons; as well as heavier products that appeared at an elution time longer than those of the squalane model molecule, reflecting the occurrence of polymerization processes during the thermal oxidative degradation (Fig. 8).

Product identification was achieved on the basis of their mass spectra using the NIST MS library, by using chemical ionization mass spectra for molecular weight determination, and finally by comparing their elution time with those of commercial standard if uncertainty remains in their attribution. Table 2 shows the retention time (R.T.), the corresponding molecular weight (M.W.), and the most likely structure of the main products detected.

3.2 Effect of Antioxidant Additives. As stated previously, thermal degradation experiments conducted with squalane were extended to squalane doped with phenolic antioxidants.

3.2.1 Effect of Antioxidants. The plots of squalane mass loss and the percentages of oxidation products detected versus the test duration for experiments carried out at 180°C and several contents of antioxidant OBHP are reported in Figs. 9(a) and 9(b), respectively.

One can notice a decrease in the evaporation rate when the content of antioxidant increased (Fig. 9(a)) in comparison to the previous data obtained during the thermodegradation of squalane in the absence of antioxidants (Sec. 3.1 above, Fig. 5). In the same way, experiments in the presence of antioxidants confirm the existence of an induction time during which oxidation is hardly observed (Fig. 9(b)). Furthermore, the presence of antioxidants results in a noticeable decrease in the amount of oxidation products

at the longer duration where the decrease in oxidation product ratios was lowered of ~50% (Fig. 9(b)). From these results, a correlation between evaporation process and oxidation process was tentatively assigned to the fact that oxidation of hydrocarbons produces light molecular weight products whose evaporation temperatures are lower than that of squalane. Consequently, an increase in the oxidation rate, which is the case in the absence of OBHP, favors the formation of light products more easily evaporated: The net result of this leads to an associated increase in the evaporation rate. On the contrary, the presence of antioxidants inhibits oxidation reactions, during the induction time, and limits these reactions at the higher test duration times. The resulting formation of light products is reduced, and the associated evaporation rates are decreased.

3.2.2 Effect of Temperature. The action of both antioxidants, BHT and OBHP, has been studied varying the temperature. For illustration, Fig. 10 below show the corresponding mass loss and the percentage of oxidation product formation versus the test duration for experiments carried out at 205°C with squalane samples doped with 0.5% of OBHP.

In both cases, i.e., in the presence of BHT (not shown here) and OBHP (Fig. 10), it has been noticed that at 205°C the evaporation rates are higher even in the presence of antioxidants. Moreover, above a certain duration test estimated around 20 h, the percentage of oxidation products becomes higher in the presence of antioxidants compared with experiment without an antioxidant (Fig. 10(b)). This phenomenon was also observed at 150°C, but only with squalane samples doped with BHT; noticeably this tendency was increased at high temperature for these samples. This observation is in agreement with the work of Zabarnick and co-worker [38,39] who observed, for jet fuel oxidation at 140°C in the presence of BHT, a slow initial oxidation followed by a sudden increase in the oxidation rate due to the consumption of BHT below some critical concentration.

The phenomenon could be also attributed to a possible reversed activity of antioxidants above a threshold temperature (related to antioxidant structure as will be seen below), which causes an acceleration of the oxidation process. In addition, the existence of this phenomenon at 150°C only observed in the case of BHT doped squalane samples can be explained by the fact that BHT is less stable than OBHP at the highest temperatures investigated, due to the presence of a long-chain octadecyl propionate in para-position for OBHP leading to a lower volatility compared with

Table 2 Identification of the main products by GC/MS analysis

Retention time (min)	M.W. (g mol ⁻¹)	Formula	Attribution
Light hydrocarbons			
5.58	114	C ₈ H ₁₈	Heptane, 2-methyl
5.39	100	C ₇ H ₁₆	Hexane, 2-methyl
10.65	142	C ₁₀ H ₂₂	Octane, 2,6-dimethyl
11.52	128	C ₉ H ₂₀	Heptane, 2,6-dimethyl
16.88	156	C ₁₁ H ₂₄	Nonane, 2,6-dimethyl
22.90	192	C ₁₅ H ₃₂	Dodecane, 2,6,10-trimethyl
26.50	226	C ₁₆ H ₃₄	Tridecane, 2, 6, 10-triméthyl
Alcohol			
6.09	116	C ₇ H ₁₆ O	1-hexanol, 5-methyl
12.49	144	C ₉ H ₂₀ O	3-heptanol, 2,6-dimethyl
14.79	144	C ₉ H ₂₀ O	2-heptanol, 2,6-dimethyl
15.50	158	C ₁₀ H ₂₂ O	1-octanol, 3,7-dimethyl
17.01	186	C ₁₂ H ₂₆ O	1-decanol, 5,9-dimethyl
19.35	200	C ₁₃ H ₂₈ O	2-undecanol, 6,10-dimethyl
22.81	228	C ₁₅ H ₃₂ O	1-dodecanol, 3,7,11-trimethyl
27.56	298	C ₂₀ H ₄₂ O	1-hexadecanol, 2,7,11,15-tetramethyl
Aldehydes/ketones			
8.90	128	C ₈ H ₁₆ O	2-heptanone, 6-methyl
10.79	142	C ₉ H ₁₈ O	4-heptanone, 2,6-dimethyl
12.05	142	C ₉ H ₁₈ O	3-heptanone, 2,6-dimethyl
19.90	198	C ₁₃ H ₂₆ O	2-undecanone, 6,10-dimethyl
21.17	210	C ₁₄ H ₂₆ O	Undecanal, 2,6,10-trimethyl
29.32	268	C ₁₈ H ₃₆ O	2-pentadecanone, 6,10,14-trimethyl
Acids/esters			
5.88	116	C ₆ H ₁₂ O ₂	Pentanoic acid, 2-methyl
9.76	114	C ₆ H ₁₀ O ₂	3-pentenoic acid, 4-methyl
14.88	172	C ₁₀ H ₂₀ O ₂	Octanoic acid, 4-methyl-, methyl ester
27.80	270	C ₁₇ H ₃₄ O ₂	Tridecanoic acid, 4,8,12-trimethyl-, methyl ester

BHT where the methyl group in paraposition cannot afford such property [9]. As a consequence, the adverse effect of antioxidant volatility is likely increased at the higher temperatures of the thermal degradation processes.

3.3 Formulation of the Kinetic Model Without Antioxidant Actions. The reaction mechanisms for the oxidation of paraffin hydrocarbons involve numerous steps of free radical chain reactions that result in primary oxidation products followed by further reaction leading to heavy products (polymers), and varnish deposit formation. Numerous studies present a computational approach for modeling each step of these reactions [40,41]. However, as the experimental identification of intermediate species (radicals and peroxides) and their quantification are difficult, a

simple reaction model is required to depict as far as possible the global kinetic degradation of paraffin hydrocarbons, admitting that the details of involved reactions are gathered under the form of general transformation: oxidation, polymerization, or evaporation, respectively. Naidu et al. [42] previously proposed such a simple chemical reaction for the oxidation of hydrocarbons as shown in Fig. 11. In this model, the hydrocarbon precursor evaporates with the rate constant k_4 and reacts with oxygen to form the primary oxidation products (Ox) with the rate constant k_1 . These oxidized products are also submitted to an evaporation process with the rate constant k_5 and undergo further condensation and polymerization reactions to generate heavy oxidation products (Pol) with the rate constant k_2 . These heavy oxidation products can also evaporate

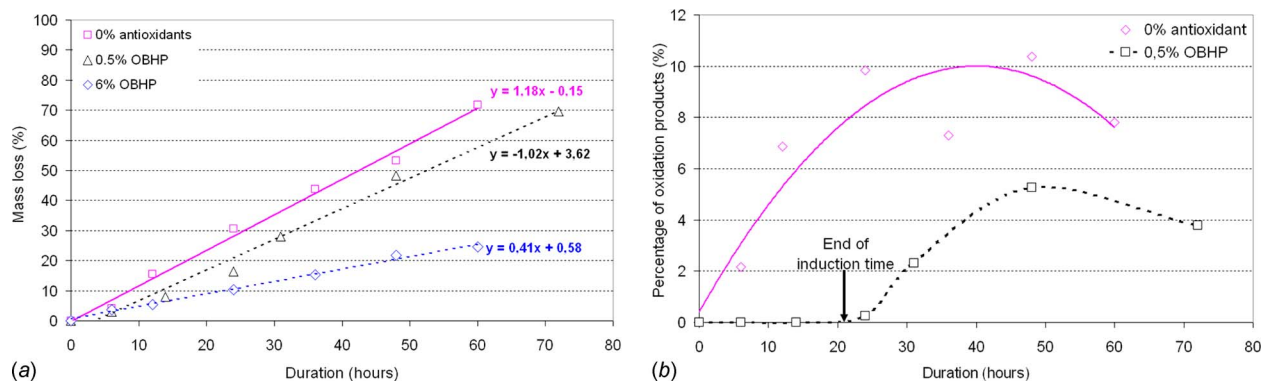


Fig. 9 (a) Mass loss and (b) percentage of oxidation products versus time for increased test duration time at 180°C with squalane doped with OBHP

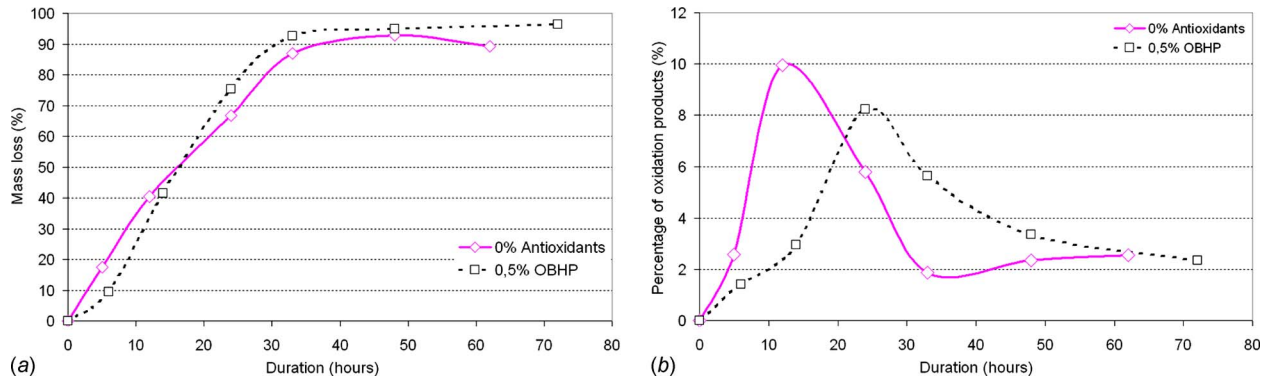


Fig. 10 (a) Mass loss and (b) percentage of oxidation products versus time for increased test temperature at 205°C with squalane doped with OBHP

(rate constant k_6) and carbonized into varnish deposits (Dep) with the rate constant k_3 . Kinetic equations [8] that describe the reaction system are shown in Fig. 11.

Data analysis of the present experiments with oxidized samples of squalane was employed to check the possibility to rationalize its thermo-oxidative degradation by using a kinetic model derived from the model of Naidu et al. [42]. Moreover, the resulting kinetic modeling might provide a practical analytical tool to follow the thermal degradation processes, and predict base oil hydrocarbon aging under engine test duration experiments.

It was experimentally observed that at the highest temperature used (205°C), no any varnish deposits were observed, whatever the test duration time. Only the presence of insoluble compounds was noticed in the oxidized samples. Consequently, it has been considered that the kinetic model describing the thermal oxidative degradation of squalane under our experimental conditions could be simplified for the formation of heavy products. The model presented here proposes to gather the reactions leading to heavy oxidation products (Pol) and varnish deposits (Dep) in a one step reaction (rate constant k_2), considering a global rate constant k_{6-7} to quantify their evaporation process (Fig. 12).

The solution of these equations provides the expression of each species at a given time. But all kinetic constants must be determined to complete these solutions.

$$[C_{30}] = [C_{30}]_0 \times e^{-(k_1+k_4) \times t}$$

$$[Ox] = \frac{k_1 \times [C_{30}]_0}{(k_2 + k_5) - (k_1 + k_4)} \times \{ e^{-(k_1+k_4) \times t} - e^{-(k_2+k_5) \times t} \}$$

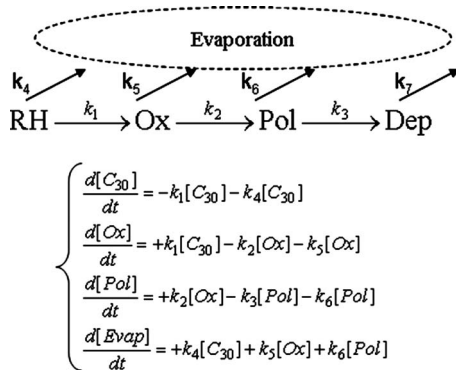


Fig. 11 The model reaction and the kinetic rate equations proposed by Naidu et al. [42] for the modeling of paraffin hydrocarbon oxidative thermal degradation

$$[Pol + Dep] = \frac{k_1 \times k_2 \times [C_{30}]_0}{(k_2 + k_5) - (k_1 + k_4)} \times \left\{ \frac{e^{-(k_1+k_4) \times t}}{k_{6-7} - (k_1 + k_4)} - \frac{e^{-(k_2+k_5) \times t}}{k_{6-7} - (k_2 + k_5)} \right\}$$

$$+ \frac{k_1 \times k_2 \times [C_{30}]_0 \times e^{-k_{6-7} \times t}}{(k_{6-7} - (k_2 + k_5)) \times (k_{6-7} - ((k_1 + k_4)))}$$

$$[E] = 100 - ([C_{30}] + [Ox] + [Pol + Dep])$$

3.4 Determination of Kinetic Rate Constants

3.4.1 Determination of k_1 and k_4 . The first step of the simplified model of oxidation reaction can be decomposed into two fictitious stages, as shown in Fig. 13.

If one considers that the initial amount of squalane (C_{30}) is evaporated in a first stage, the nonevaporated fraction (reported here as C_{30} reactant) will be oxidized in a second stage. So the remaining sample in the vessel consists of (i) residual squalane and (ii) oxidation products.

Then, the quantification of the residual squalane (C_{30}) is amenable to GC/MS analysis that provides the rate constant k_1 using Eq. (1) where α and β are the reaction orders for squalane and oxygen, respectively. Considering that the oxygen ratio is constant and in excess compared with squalane, the rate constant k_1 is expressed as follows:

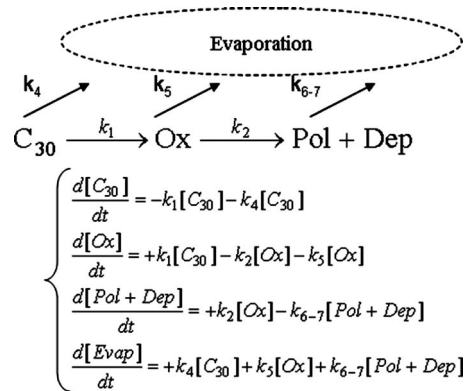


Fig. 12 The simplified model and the kinetic rate equations proposed for the modeling of squalane thermal oxidative degradation

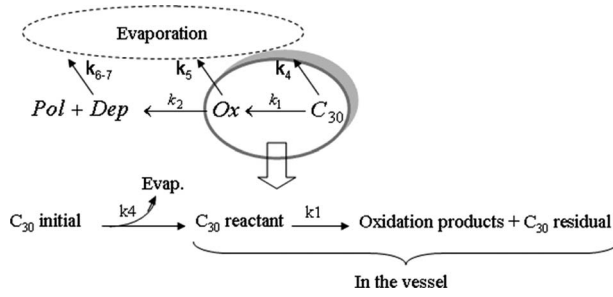


Fig. 13 Fictitious steps of oxidation products' formation for squalane oxidative thermal degradation

$$-\frac{d[C_{30}]}{dt} = k_1 \times [C_{30}]^\alpha \times [O_2]^\beta \cong k_1 \times [C_{30}]$$

$$\int_0^{[C_{30}]} \frac{d[C_{30}]}{[C_{30}]} = -k_1 \times t \Leftrightarrow \ln\left(\frac{[C_{30}]}{[C_{30}]_{\text{reac}}}\right) = -k_1 \times t \quad (1)$$

However, the residual squalane (C_{30}) measured by GC/MS from the sample present in the vessel did not include the fraction directly evaporated as squalane or light oxidation products, associated with process controlled by the rate constants k_4 and k_5 , respectively. To take into account such evaporated fractions, the percentage of each species measured by GC/MS was normalized to the initial amount of squalane. So Eq. (2) below, which expresses the normalized concentration of squalane, provides the constant k_1+k_4 . Thus, the rate constant k_4 can be easily deduced.

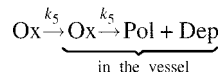
$$[C_{30}] = [C_{30}]_0 \times e^{-(k_1+k_4) \times t} \quad (2)$$

3.4.2 Determination of k_2 and k_5 . Equation (3) below giving the fraction of oxidation products (Ox) versus time (t) has a maximum, which is reached at a time t_M . The theoretical expression of this time "Eq. (4)" is achieved through the derivative of Eq. (3). The time t_M is experimentally determined. Then, as the rate constant k_1+k_4 has been already determined, one can easily calculate the constant k_2+k_5 through Eq. (4) as follows:

$$[Ox] = \frac{k_1 \times [C_{30}]_0}{(k_2+k_5) - (k_1+k_4)} \times \{e^{-(k_1+k_4) \times t} - e^{-(k_2+k_5) \times t}\} \quad (3)$$

$$t_M = \frac{1}{(k_1+k_4) - (k_2+k_5)} \times \ln\left(\frac{k_1+k_4}{k_2+k_5}\right) \quad (4)$$

By applying the same approach used for determining the rate constant k_1 , one can consider that the oxidized products are first evaporated (k_5), then the nonevaporated fraction undergo polymerization and condensation reactions that produce the high-molecular weight products (Pol) and the insoluble compound observed (Dep) as follows:



Quantification by GC/MS of the oxidized products (Ox) without including the evaporated fraction provides the constant k_2 using Eq. (5), which has a maximum reached at the time t_m "Eq. (6)." The experimental estimation of the time t_m allows the calculation of the constant k_2 since the value of the constant k_1 has been previously determined.

$$[Ox] = \frac{k_1 \times [C_{30}]_0}{k_2 - k_1} \times \{e^{-k_1 \times t} - e^{-k_2 \times t}\} \quad (5)$$

$$t_m = \frac{1}{k_1 - k_2} \times \ln\left(\frac{k_1}{k_2}\right) \quad (6)$$

Table 3 Experimental rate constant values

k_i ($\times 10^{-3} \text{ h}^{-1}$)	150°C	180°C	205°C
k_1	4	16	20
k_1+k_4	8	25	41
k_2	37	45	75
k_2+k_5	42	64	149
k_{6-7}	5	2	1

3.4.3 Determination of k_{6-7} . In the equation giving the fraction (Pol+Dep) as a function of time (t), one can suppose that the rate constant k_{6-7} for the evaporation of the fraction (Pol+Dep) is negligible compared with that of squalane (k_4) and oxidation products (k_5). As a consequence, the previous equation expressing the concentration of [Pol+Dep] can be rewritten as follows:

$$[Pol + Dep] = \frac{k_1 \times k_2 \times [C_{30}]_0}{(k_2+k_5) - (k_1+k_4)} \times \left\{ \frac{e^{-(k_1+k_4) \times t}}{-(k_1+k_4)} - \frac{e^{-(k_2+k_5) \times t}}{-(k_2+k_5)} \right\} + \frac{k_1 \times k_2 \times [C_{30}]_0 \times e^{-k_{6-7} \times t}}{(-(k_2+k_5)) \times (-(k_1+k_4))} \quad (7)$$

In addition, the curve associated with this equation has a maximum reached at the time ζ . The value of ζ can be experimentally determined through the evolution of the high weight products (polymer) fraction quantified by GC/MS. Then the following equation allows the calculation of the constant k_{6-7} since all the other data were previously determined:

$$k_{6-7} = \frac{(k_2+k_5) \times (k_1+k_4)}{(k_2+k_5) - (k_1+k_4)} \times \{e^{-(k_1+k_4) \times \zeta} - e^{-(k_2+k_5) \times \zeta}\} \quad (8)$$

3.5 Results and Generalization of the Model. By applying the model previously described, all the rate constants have been determined. Table 3 reports rate constants for experiments carried out without an antioxidant. According to these results, one can notice that the rate constant k_2+k_5 of the reaction "Ox \rightarrow Pol + Dep" is higher than the rate constant k_1+k_4 of the reaction "C₃₀ \rightarrow Ox." Consequently, after the formation of primary oxidation products, the degradation process becomes much faster. Then the use of antioxidant additives should lead to a decrease in the overall rate of degradation process.

As all the rate constants are known, one can easily plot the curves giving the evolution of the different fractions versus time. To illustrate it, Fig. 14 shows the evolution of the evaporated

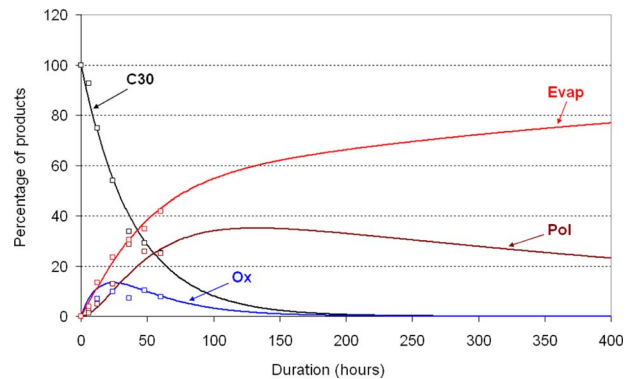


Fig. 14 Variations of squalane and degradation products' experimental percentages versus test duration time at a temperature of 180°C (dots) and their theoretical variations (thick lines) applying the simplified model presented here for the oxidative thermodegradation of squalane

fraction (Evap), the reactant squalane (C₃₀), the oxidation products (Ox), and the heavy products (Pol) as a function of the duration for oxidative test conducted at 180°C without an antioxidant. As can be seen, these graphs show a good agreement between theoretical and experimental curves.

The reactions occurring during the oxidative degradation of the squalane take place with the rate constants k_1 , k_2 , k_4 , k_5 , and k_{6-7} , set previously. Each of these rate constants could be expressed under the form of a mathematical function depending on the temperature in accordance with an Arrhenius treatment for the chemical processes involved. The general equation for each rate constant can be expressed as follows:

$$\ln k_i = \ln k_{i,0} - \frac{E_a}{RT} \quad (9)$$

In fact, the activation energy obtained for the reaction “C₃₀ → Ox” including the evaporation rate was 12 kcal mol⁻¹, when that for the reaction “Ox → Pol+Dep” was 5 kcal mol⁻¹ or 9 kcal mol⁻¹ including the evaporation phenomenon. These values are much closed to the activation energy between 9 kcal mol⁻¹ and 15 kcal mol⁻¹ reported in literature for oxidation reaction of several lubricating oils [33–37]. Consequently, the initial assumption relative to the possibility to separate the evaporation from the oxidation itself for an easier calculation of the rate constant can be validated.

4 Conclusion

The oxidative degradation of squalane (C₃₀H₆₂) has been investigated with and without phenolic antioxidants. The rate constants obtained for tests without antioxidants enable the determination of activation energy in accordance with values reported in literature. If experiments confirmed the effectiveness of antioxidants in reducing the oxidation rate, the existence of a threshold temperature, from which a reversed activity of antioxidant appeared, has been highlighted. This threshold temperature strongly depends on the stability of the antioxidants.

Data analysis allowed formulating a kinetic model, which can be used as a practical analytical tool to follow the effectiveness of thermal oxidative degradation of squalane without antioxidants. This kinetic model can be easily generalized at different temperatures. Nevertheless, in the presence of antioxidants, application of the kinetic model requires first the determination of the observed threshold temperature for the antioxidants investigated.

Further experiments are currently conducting to validate the present kinetic model in the case of more complex mixtures for the prediction of lubricating oil aging.

References

- [1] Stachowiak, G. W., and Batchelor, A. W., 2006, “Physical Properties of Lubricants,” *Engineering Tribology*, 3rd ed., Elsevier-Butterworth-Heinemann, Amsterdam, pp. 11–50.
- [2] Taylor, C. M., 1998, “Automobile Engine Tribology—Design Considerations for Efficiency and Durability,” *Wear*, **221**(1), pp. 1–8.
- [3] Gamlin, C. D., Dutta, N. K., Choudhury, N. R., Kehoe, D., and Matison, J., 2002, “Evaluation of Kinetic Parameters of Thermal and Oxidative Decomposition of Base Oils by Conventional, Isothermal and Modulated TGA, and Pressure DSC,” *Thermochim. Acta*, **392–393**, pp. 357–369.
- [4] Mortier, R. M., and Orszulik, S. T., 1994, *Chemistry and Technology of Lubricants*, VCH, New York.
- [5] Liston, T. V., 1992, “Engine Lubricant Additives What They Are and How They Function,” *Lubr. Eng.*, **48**(5), pp. 389–397.
- [6] Sasaki, A., Tobisu, T., Uchiyama, S., and Kawasaki, M., 1990, “GPC Analysis of Oil Insoluble Oxidation Products of Mineral Oil,” *Lubr. Eng.*, **47**(7), pp. 525–527.
- [7] Ayel, J., 2001, “Lubrifiants: Additifs à action chimique,” Ed. Techniques de l’Ingénieur, BM5343.
- [8] Chen, C. I., and Hsu, S. M., 2003, “A Chemical Kinetics Model to Predict Lubricant Performance in a Diesel Engine. Part I: Simulation Methodology,” *Tribol. Lett.*, **14**(2), pp. 83–90.
- [9] Mousavi, P., Wang, D., Grant, C. S., Oxenham, W., and Hauser, P. J., 2006, “Effects of Antioxidants on the Thermal Degradation of a Polyol Ester Lubricant Using GPC,” *Ind. Eng. Chem. Res.*, **45**, pp. 15–22.
- [10] Singh, S. K., Agarwal, A. K., Srivastava, D. K., and Sharma, M., 2006, “Experimental Investigation of the Effect of EGR on Lubricating Oil Degradation and Wear of a Compression Ignition Engine,” *ASME J. Eng. Gas Turbines Power*, **128**, pp. 921–927.
- [11] Singh, S. K., Agarwal, A. K., and Sharma, M., 2006, “Experimental Investigations of Heavy Metal Addition in Lubricating Oil and Soot Deposition in an EGR Operated Engine,” *Appl. Therm. Eng.*, **26**, pp. 259–266.
- [12] Wilkinson, J. J., Priest, M., Stark, M. S., Lee, P. M., and Taylor, R. I., 2004, *Understanding Piston Assembly Lubricant Degradation in Gasoline Engine*, University of York, London.
- [13] Lakatos, K. L., Jones, R. N., and Roby, S. H., 1992, *Modeling ASTM Sequence III-E Piston Ring Land Deposit Formation*, The Lubrizol Corporation.
- [14] Owrang, F., Mattson, H., Olsson, J., and Padersen, J., 2004, “Investigation of Oxidation of a Mineral and a Synthetic Engine Oil,” *Thermochim. Acta*, **413**, pp. 241–248.
- [15] Adhvaryu, A., Erhan, S. Y., Sahoo, S. K., and Singh, I. D., 2002, “Thermo-Oxidative Stability Studies on Some New Generation API Group II and Group III Base Oils,” *Fuel*, **81**, pp. 785–791.
- [16] Balster, L. M., Zabarnick, S., Striebich, R. C., Shafer, L. M., and West, Z. J., 2006, “Analysis of Polar Species in Jet Fuel and Determination of Their Role in Autoxidative Deposit Formation,” *Energy Fuels*, **20**, pp. 2564–2571.
- [17] Slade, P. E., 1998, *Handbook of Fiber Finish Technology*, Dekker, New York.
- [18] Dexter, M., 1985, *Antioxidants: In Encyclopedia of Polymer Science and Engineering*, Vol. 2, 2nd ed., J. I. Kroschwitz, ed., Wiley, New York, pp. 73–91.
- [19] Dexter, M., 1992, *Antioxidants: In Kirk-Othmer Encyclopedia of Chemical Technology*, Vol. 3, 4th ed., J. I. Kroschwitz, ed., Wiley, New York, pp. 424–447.
- [20] Zeman, A., Von Roenne, V., and Trebert, Y., 1987, “Fate of Amine Antioxidants During Thermal Oxidative Aging of Neopentylpoly Ester Oils,” *Synth. Lubr.*, **4**, pp. 179–201.
- [21] Ervin, J. S., and Zabarnick, S., 1998, “Computational Fluid Dynamics Simulations of Jet Fuel Oxidation Incorporating Pseudo-Detailed Chemical Kinetics,” *Energy Fuels*, **12**(2), pp. 344–352.
- [22] Pfaendtner, J., and Broadbelt, L. J., 2007, “Elucidation of Structure-Reactivity Relationships in Hindered Phenols Via Quantum Chemistry and Transition State Theory,” *Chem. Eng. Sci.*, **62**, pp. 5232–5239.
- [23] Diaby, M., Sablier, M., Le Negrate, A., El Fassi, M., and Bocquet, J., 2009, “Understanding Carbonaceous Deposit Formation Resulting From Engine Oil Degradation,” *Carbon*, **47**(2), pp. 355–366.
- [24] Blaine, S., and Savage, P. E., 1991, “Reaction Pathways in Lubricant Degradation. 1. Analytical Characterization of n-Hexadecane Autoxidation Products,” *Ind. Eng. Chem. Res.*, **30**, pp. 792–798.
- [25] Blaine, S., and Savage, P. E., 1991, “Reaction Pathways in Lubricant Degradation. 2. n-Hexadecane Autoxidation,” *Ind. Eng. Chem. Res.*, **30**, pp. 2185–2191.
- [26] Blaine, S., and Savage, P. E., 1992, “Reaction Pathways in Lubricant Degradation. 3. Reaction Model for n-Hexadecane Autoxidation,” *Ind. Eng. Chem. Res.*, **31**, pp. 69–75.
- [27] Pfaendtner, J., and Broadbelt, L. J., 2008, “Mechanistic Modeling of Lubricant Degradation. 2. The Autoxidation of Decane and Octane,” *Ind. Eng. Chem. Res.*, **47**, pp. 2897–2904.
- [28] Garcia-Ochoa, F., Romero, A., and Querol, J., 1989, “Modeling of the Thermal Normal-Octane Oxidation in the Liquid Phase,” *Ind. Eng. Chem. Res.*, **28**, pp. 43–48.
- [29] Lin-Vien, D., Colthup, N. B., Fateley, W. G., and Grasselli, J. G., 1991, *The Handbook of Infrared and Raman Characteristic Frequencies of Organic Molecules*, Academic, New York.
- [30] Barman, B. N., 2002, “Behavioral Differences Between Group I and Group II Base Oils During Thermo-Oxidative Degradation,” *Tribol. Int.*, **35**, pp. 15–26.
- [31] Bowman, W. F., and Stachowiak, G. W., 1996, “New Criteria to Assess the Remaining Useful Life of Industrial Turbine Oils,” *Lubr. Eng.*, **52**(10), pp. 745–750.
- [32] Coates, J. P., and Setti, L. C., 1986, “Infrared Spectroscopic Methods for the Study of Lubricant Oxidation Products,” *ASLE Trans.*, **29**(3), pp. 394–401.
- [33] Crisóstomo Rabelo Neto, R., Lima, D. O., Pinheiro, T. D. S., Almeida, R. F., and Azevedo, D. C. S., 2004, “Thermo-Oxidative Stability of Mineral Naphthenic Insulating Oils: Combined Effect of Antioxidants and Metal Passivator,” *Ind. Eng. Chem. Res.*, **43**, pp. 7428–7434.
- [34] Adhvaryu, A., Perez, J. M., Singh, I. D., Tyagi, O. S., and Sharman, Y. K., 1998, “Spectroscopic Studies of Oxidative Degradation of Base Oils,” *Energy Fuels*, **12**, pp. 1369–1374.
- [35] Koh, C. S., and Butt, J. B., 1995, “Experimental and Modeling Study of Kinetics and Selectivity in the Oxidation of a Poly(α-Olefin),” *Lubr. Ind. Eng. Chem. Res.*, **34**, pp. 524–535.
- [36] Krishnamoorthy, P. R., Vijayakumari, S., Krishnaswamy, K. R., and Thomas, P., 1991, “Effect of Benzotriazole and 2,6-Di-tert-butyl Paracresol on the Accelerated of New Reclaimed Transformer Oils a Comparative Study,” *Proceedings of the Third International Conference on Properties and Applications of Dielectric Materials*, p. 732.
- [37] Natarajan, S., Olson, W. W., and Abraham, M. A., 2000, “Reaction Pathways and Kinetics in the Degradation of Forging Lubricants,” *Ind. Eng. Chem. Res.*, **39**, pp. 2837–2842.

- [38] Zabarnick, S., and Mick, M. S., 1999, "Inhibition of Jet Fuel Oxidation by Addition of Hydroperoxide-Decomposing Spacing," *Ind. Eng. Chem. Res.*, **38**, pp. 3557–3563.
- [39] Zabarnick, S., 1998, "Pseudo-Detailed Chemical Kinetic Modeling of Antioxidant Chemistry for Jet Fuel Applications," *Energy Fuels*, **12**, pp. 547–553.
- [40] Zabarnick, S., 1993, "Chemical Kinetic Modeling of Jet Fuel Autoxidation and Antioxidant Chemistry," *Ind. Eng. Chem. Res.*, **32**(6), pp. 1012–1017.
- [41] Kuprowicz, N. J., Ervin, J. S., and Zabarnick, S., 2004, "Modeling the Liquid-Phase Oxidation of Hydrocarbons Over a Range of Temperatures and Dissolved Oxygen Concentrations With Pseudo-Detailed Chemical Kinetics," *Fuel*, **83**, pp. 1795–1801.
- [42] Naidu, S. K., Klaus, E., and Duda, J. L., 1984, "Evaluation of Liquid Phase Oxidation Products of Ester and Mineral Oil Lubricants," *Ind. Eng. Chem. Prod. Res. Dev.*, **23**(4), pp. 613–619.

A Comparison Between Recent Advances in Cylindrical Nodal Diffusion Methods

David V. Colameco

Kostadin N. Ivanov

The Pennsylvania State University,
University Park, PA 16803

Rian H. Prinsloo

Djordje I. Tomasevic

NECSA,
Pretoria 0001, South Africa

Suzanne Theron

University of Pretoria,
Pretoria 0002, South Africa

The resurgence of high temperature reactor (HTR) technology has prompted the development and application of modern calculation methodologies, many of which are already utilized in the existing power reactor industry, to HTR designs. To this end, the use of nodal diffusion methods for full core neutronic analysis is once again considered for both their performance and accuracy advantages. Recently a number of different approaches to two-dimensional and 3D multigroup cylindrical nodal diffusion methods were proposed by various institutions for use in HTR and, specifically, pebble-bed modular reactor (PBMR) calculations. In this regard, we may mention the NEM code from the Pennsylvania State University based on the nodal expansion method and the OSCAR-4 code from NECSA, utilizing a conformal mapping approach to the analytic nodal method. In this work we will compare these two approaches in terms of accuracy and performance. Representative problems, selected to test the methods thoroughly, were devised and based on both a modified version of the PBMR 400 MW benchmark problem and a “cylindrized” version of the IAEA two-group problem. The comparative results between OSCAR-4 and NEM are given, focusing on global reactivity estimation, as well as power and flux errors as compared with reference finite-difference solutions. These results indicate that both OSCAR-4 and NEM recover the global reference solution for the IAEA problem and show power errors, which are generally acceptable for nodal methods. For the PBMR problem the accuracy is similar, but some convergence difficulties are experienced at the outer boundaries of the system due to the very large dimensions of the reflector (when compared with typical water-moderated reactors). For both codes a significant performance increase was found, as compared with finite-difference calculations, which is the method currently employed by the PBMR (Pty) Ltd. In conclusion it seems that nodal methods have potential for use in the HTR analysis and, specifically, the PBMR calculational arena, although cylindrical geometry based nodal methods will have to develop toward maturity before becoming the industry standard. [DOI: 10.1115/1.3095806]

1 Introduction

The resurgence of high temperature reactor (HTR) technology has prompted the development and application of modern calculation methodologies, many of which are already utilized in the existing power reactor industry, to HTR designs. To this end, the use of nodal diffusion methods for full core neutronic analysis is once again considered for both their performance and accuracy advantages.

Recently a number of different approaches to two-dimensional (2D) and 3D multigroup cylindrical nodal diffusion methods were proposed by various institutions for use in HTR and, specifically, pebble-bed modular reactor (PBMR) calculations. In this regard, we may mention the NEM code from the Pennsylvania State University (PSU) based on the nodal expansion method (NEM) [1] and the OSCAR-4 code from NECSA, utilizing a conformal mapping approach to the analytic nodal method (ANM) [2].

This paper concisely describes the NEM and ANM nodal diffusion approaches and subsequently compares both of these against diverse numerical diffusion problems—first against a typical light water reactor (LWR) cylindrical problem and second against a cylindrical HTR problem, the published Organization for Economic Co-Operation and Development/Nuclear Energy Agency (OECD/NEA) PBMR 400 mega watt (MW) benchmark.

2 Codes and Methods

In this paper we will compare two approaches in terms of accuracy and performance. Two numerical problems, selected to test the methods thoroughly, include the PBMR 400 MW benchmark problem [4] and a “cylindrized” version of the IAEA two-group LWR problem [3]. Comparative results between OSCAR-4 and NEM will be given, focusing on global reactivity estimation, as well as power and flux errors as compared with reference finite-difference solutions. Prior to investigating the comparative results, short descriptions of the two methods are given in Secs. 2.1 and 2.2.

2.1 Cylindrical Nodal Expansion Method. The NEM code is a 3D multigroup nodal diffusion code developed at PSU for modeling both steady-state and transient core conditions based on the nodal expansion method. It utilizes a transverse integration procedure and is based on the partial current formulation of the nodal balance equations. The leakage term in the one-dimensional transverse-integrated equations is approximated using a standard parabolic expansion using the transverse leakages in three neighboring nodes. The code has options for the modeling of 3D Cartesian, cylindrical, and hexagonal geometries. The cylindrical option utilizes fourth-order polynomial expansions of the 1D transverse-integrated flux distribution in the R -, Z -, and θ -directions. The cylindrical option has been verified using the simple analytical benchmark problems. This code has also been tested for applicability in pebble-bed reactors through a number of steady-state core physics benchmarks and was shown to be suitable for the analysis of this reactor type.

Manuscript received October 21, 2008; final manuscript received October 27, 2008; published online November 24, 2009. Review conducted by Dilip R. Ballal. Paper presented at the Fourth International Topical Meeting on High Temperature Reactor Technology (HTR2008), Washington, DC, September 28–October 1, 2008.

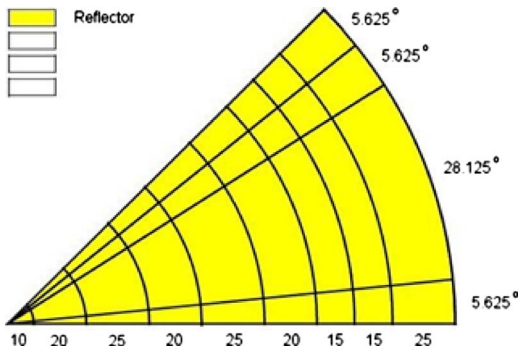


Fig. 1 Bottom reflector region, 20 cm axially

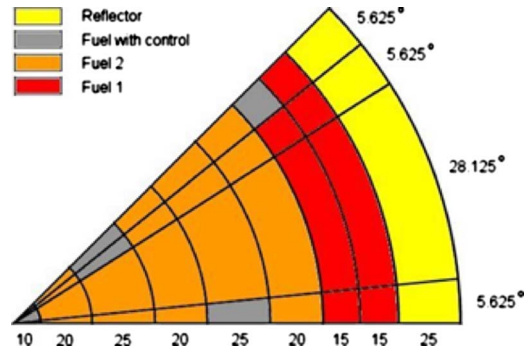


Fig. 3 Top core region, 80 cm axially

2.2 Cylindrical Analytic Nodal Method. The cylindrical test module of the OSCAR-4 code system uses an ANM to solve the 3D multigroup diffusion equation. This method employs the transverse integration procedure to produce three coupled one-dimensional equations. Unfortunately, in cylindrical geometry, the application of the transverse integration technique to the ANM cannot produce a one-dimensional equation in the θ -direction with uniquely defined transverse-integrated flux. To circumvent this, a conformal mapping technique is applied. This technique maps the cylindrical nodes into rectangular ones, after which transverse integration can be completed in the mapped Cartesian coordinate system. Leakage terms in the one-dimensional Z-equation are approximated via the standard quadratic leakage approximation. In the radial plane the leakage term is approximated within the mapped coordinate system. For full details on this approach see Ref. [2].

3 Description of Numerical Benchmark Problems

3.1 Cylindrized IAEA LWR Problem. The first chosen numerical problem addressed in this work is a cylindrized version of the 3D IAEA Power Water Reactor (PWR) benchmark problem [3]. The original benchmark has been adapted to cylindrical coordinates by closely matching the volumes of fuel, control, and reflector material and hence by preserving the nature of flux gradients in the problem. The cross sections used are identical to those described in the original benchmark specifications. The adjusted problem is graphically represented in Figs. 1–4 below. The segments in the figures represent a 1/8th cut of the core with reflective boundary conditions applied azimuthally and vacuum boundary conditions on outer radial and top and bottom axial surfaces.

Figures 1–4 illustrate four radial cuts of the IAEA 3D problem, describing the material layout in each axial segment from bottom to top. Radial and azimuthal sizes are indicated in the figures.

These segments stack onto each other, starting with the segment in Fig. 1 at the bottom and in Fig. 4 at the top, to complete one full symmetry segment of the problem.

3.2 PBMR 400 MW Benchmark Problem. The main objective of the PBMR 400 MW OECD benchmark test cases is to develop coupled kinetics-core thermal-hydraulics test problems that include both fast (reactivity insertions) and slow (thermal heat-up due to decay heat) transients. The reference model for this benchmark problem is derived from the PBMR 400 MW design. It is an annular core with an outer diameter of 3.7 m and a solid central reflector made of graphite with an outer diameter of 2 m, an effective cylindrical core height of 11 m, and a graphite side reflector of 0.9 m thickness. The design was simplified to yield a two-dimensional core design (R - Z), assuming equal speed parallel flow of pebbles within the core, a flattened pebble-bed's upper surface, and the removal of the bottom cone and de-fuel channel, resulting in a flat bottom reflector.

The simplifications were made in the benchmark specification in order to minimize the need for any further approximations to be made by participants. In the figure in the Appendix, the 2D symmetry layout of the problem is shown as defined for Exercise 1 (neutronics only), with the annular core region (materials 1–110), the two void areas (materials 111 and 189), the barrel (materials 190), the control rod gray curtain representation (materials 164–168), and the graphite reflector regions (fixed central column in the center, top, and bottom reflectors and side reflector to the right of the pebble bed).

The OECD benchmark includes steady-state and transient cases. It makes use of a common set of cross sections. The focus of the benchmark is on the modeling of the transient behavior of the PBMR core, but it was necessary to define some steady-state cases to ensure consistency between the different approaches before results of transient cases are compared. In this paper, only Exercise 1 from this benchmark is utilized. Exercise 1 is defined as a neutronics solution with fixed cross sections for an equilib-

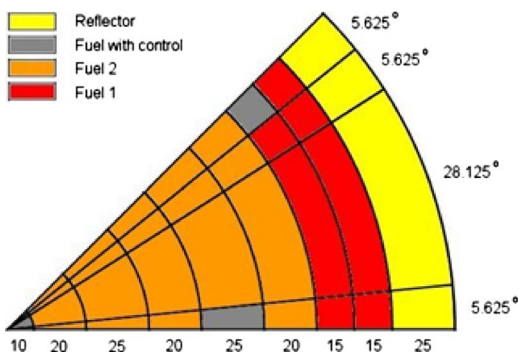


Fig. 2 Bottom core region, 260 cm axially

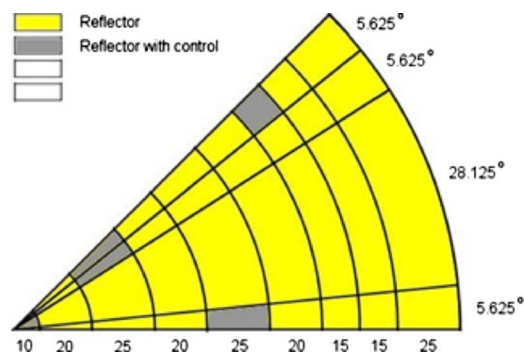


Fig. 4 Top reflector region, 20 cm axially

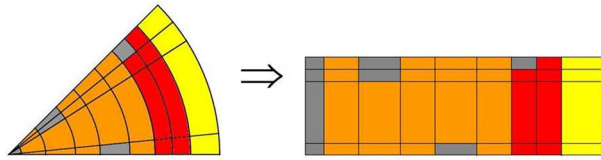


Fig. 5 Rectangular representation of the segment

rium core. Participants have to make use of the simplified set of macroscopic cross sections with no state-parameter dependence. This approach facilitates better and well-defined comparisons but also allows broader participation in the benchmark. The calculations are performed in two energy groups with a thermal cutoff of 3.059 eV. In total 190 data sets are included in the library and are assigned to positions shown in the figure in the Appendix. The five flow channels, each with 22 axial regions, make up 110 materials with the balance mainly made up by graphite at different temperatures and with different fine group spectra and therefore different few-group cross sections. The different material numbers correspond with the so-called spectrum regions and can therefore also be used as a first estimate of the mesh for determining state parameters, which can then be used to reconstruct cross sections.

The participants are requested to report k -eff, 2D R - Z power profile, two-group flux profiles, and core leakage. The calculation mesh is determined by each participant, which implies that each participant has to obtain a spatially converged solution.

4 Numerical Results and Comparisons

4.1 Cylindrized IAEA LWR Problem. This section compares the accuracy and the performance of the ANM and the NEM nodal approaches versus a finite-difference reference solution of the cylindrized IAEA problem. The reference solution is obtained via linear extrapolation of two sets of finite-difference calculations, with the following number of meshes (radial \times radial \times radial): (a) $(70 \times 32 \times 76)$ and (b) $(105 \times 48 \times 114)$. This approach is similar to that used in the original Cartesian benchmark problem. Flux results are given for a set of 39 “assemblies,” laid out as indicated in Fig. 6. Figure 5 illustrates how the zones in Fig. 6 should be interpreted since they are illustrated on a rectangular grid.

The ANM requires a small number of additional calculational meshes near the center of the system as shown in Table 1. This requirement is due to the conformal mapping approach, which

1	26	27	34	35	36	37	38	39
			28	29	30	31	32	33
	18	19	20	21	22	23	24	25
	10	11	12	13	14	15	16	17
	2	3	4	5	6	7	8	9

Fig. 6 Resulting regions or assemblies for the cylindrized LWR problem. Results are reported on 39 different edit regions, on an axial discretization of 20 cm.

Table 2 3D eigenvalue (k -eff) and timing results for the LWR problem

Calculation	k -eff (error)	Time
Extrapolated FD reference	1.03353	
Refined FD ($105 \times 48 \times 114$)	1.03348 (4 pcm)	9300 min
NEM (coarse mesh)	1.03359(6 pcm)	10 s
NEM (refined mesh)	1.03349 (4 pcm)	14 s
ANM	1.03350 (3 pcm)	57 s

places a limit on the ratio of adjacent radii (should not exceed approximately 2.5). A NEM calculation is also included on the same mesh. Using the above discretizations, a comparison of reactivity predictions is given in Table 2.

Results in Table 2 indicate that all nodal approaches produce an accurate k -eff. Relative times of the nodal approaches indicate that the NEM outperforms the ANM by approximately a factor of 4. Disregarding the fact that the codes might be differently accelerated, some performance lag is expected in the case of the ANM since the conformal mapping approach introduces an extra inhomogeneous source in the one-dimensional equations, which has to be resolved via an extra iteration level.

Since the finite-difference calculations are performed without appropriate inner iteration acceleration, the observed performance advantage factor from the nodal methods is in the order of 10,000. Appropriate acceleration would reduce the finite-difference calculation time by a factor of approximately 10, and hence we would obtain typical nodal performance advantage factors (order of 1000).

In order to investigate the accuracy of the methods further, Figs. 7 and 8 represent relative assembly averaged fluxes and relative assembly average flux errors for the various nodal approaches and meshes, respectively. Fluxes in Fig. 7 are normalized to an average core power of 1.

Results in Fig. 8 confirm that both the ANM and NEM models capture assembly averaged flux errors within the typical margins for nodal diffusion codes. In all cases, slightly larger errors are visible in the inner control rod and outer reflector zones. An interesting observation relates to the fact that the NEM code, which does not have meshing restrictions near the center as in the case of the ANM conformal mapping approach, also requires central mesh refined in order to decrease the flux errors in these regions. This observation potentially points to the fact that cylindrical geometry severely challenges the quadratic leakage approximation, especially in the radial extremities where the nodal aspect ratios become most skewed. Care should thus be taken in choosing a meshing scheme in these regions, irrespective of the nodal method under consideration.

Since Fig. 8 showed assembly averaged flux errors, maximum nodal power errors are presented in Table 3. The nodal power errors in Table 3 are larger than what would typically be expected from nodal codes. In both cases (ANM and NEM), the maximum power error occurs in relatively isolated locations. In the case of the ANM approach, the maximum error of 3.79% occurs at the axial rod tip of the central control assembly (Assembly 1), and in the case of NEM, the maximum error occurs radially adjacent to the rod tip of the partially inserted control assembly (Assembly 4).

Table 1 Radial plane mesh discretization

Code (direction)	Calculational mesh
ANM method (radial) (cm)	(1.0, 2.0, 3.0, 4.0), (8, 12), 25, 20, 25, 20, 15, 15, 25
NEM (radial), coarse (cm)	10.0, 20.0, 25, 20, 25, 20, 15, 15, 25
NEM (radial), refined (cm)	(1.0, 2.0, 3.0, 4.0), (8, 12), 25, 20, 25, 20, 15, 15, 25
ANM (azimuthal) (deg)	5.625, 14.0625, 14.0625, 5.625, 5.625
NEM (azimuthal) (deg)	5.625, 14.0625, 14.0625, 5.625, 5.625

Group 1								
28.9	36.8	39.7	39.3	34.8	25.1	14.4	9.5	1.4
			39.1	34.9	27.1	19.3	11.9	1.7
	36.8	39.9	38.2	34.2	29.6	24.6	15.4	2.1
	36.9	39.6	35.3	29.2	28.5	26.0	16.9	2.4
	36.9	39.4	32.9	22.1	25.7	25.4	16.8	2.4
Group 2								
5.2	8.9	9.4	9.6	8.5	6.1	2.6	2.8	3.8
			9.5	8.5	6.6	4.8	3.6	4.4
	8.9	9.7	9.3	8.4	7.3	6.3	4.6	5.6
	8.9	9.7	8.6	7.1	7.0	6.7	5.1	6.2
	8.9	9.6	7.9	3.9	6.2	6.5	5.1	6.2

Fig. 7 Extrapolated finite-difference reference solution

In both these cases, the origin of the problem most probably lies in the application of the quadratic transverse leakage approximation in cylindrical geometry. The verification of this statement and possible remedies in this regard are left for future work.

4.2 PBMR 400 MW Benchmark Problem. Exercise 1 of the PBMR 400 MW benchmark problem is a much larger problem compared with the IAEA problem to be analyzed. There are four reflector regions (in radial plane—central column and the outer radial reflector and two axial reflectors) with very large dimensions (when compared with typical water-moderated reactors). For the two void regions, directional dependent diffusion coefficients

NEM refined - group 1								
1.48	0.73	0.72	0.96	0.97	1.14	0.38	0.97	1.45
			1.01	0.87	0.55	0.32	0.95	1.59
	0.69	0.77	1.10	0.65	0.34	0.45	0.69	1.50
	0.68	0.89	1.34	0.69	1.03	0.20	1.12	1.56
	0.70	0.98	2.07	0.75	1.31	1.33	1.41	1.78
NEM refined - group 2								
1.45	0.74	0.72	0.94	0.98	1.12	0.41	1.53	0.34
			0.99	0.88	0.50	0.34	1.52	0.43
	0.64	0.74	1.11	0.67	0.37	0.67	1.24	0.33
	0.61	0.88	1.34	0.67	1.14	1.40	1.65	0.36
	0.62	0.99	2.16	0.57	1.38	1.49	1.94	0.56
NEM coarse - group 1								
5.60	2.05	1.18	1.28	1.04	1.07	0.46	1.14	1.59
			1.32	0.97	0.55	0.40	1.13	1.73
	2.27	1.29	1.38	0.78	0.45	0.62	0.89	1.66
	2.49	1.48	1.65	0.81	1.16	1.41	1.35	1.73
	2.58	1.58	2.42	0.88	1.45	1.54	1.65	1.97
NEM coarse - group 2								
6.32	2.01	1.18	1.26	1.08	1.04	0.54	1.71	0.48
			1.30	0.99	0.52	0.44	1.71	0.58
	2.37	1.28	1.38	0.81	0.48	0.84	1.44	0.46
	2.61	1.50	1.65	0.81	1.27	1.60	1.87	0.52
	2.70	1.63	2.51	0.68	1.52	1.70	2.17	0.77
ANM - group 1								
1.45	0.84	0.83	0.93	0.67	0.99	0.74	0.44	2.61
			0.89	0.67	0.62	0.79	0.50	2.64
	1.06	0.87	0.78	0.77	0.61	0.48	0.34	1.56
	1.15	0.81	0.71	0.71	0.63	0.48	0.35	1.60
	1.16	0.85	1.62	0.87	0.95	0.45	0.36	2.80
ANM - group 2								
1.34	0.72	0.63	0.65	0.46	0.72	0.79	0.88	0.96
			0.62	0.47	0.42	0.41	0.86	0.97
	0.94	0.60	0.58	0.55	0.40	0.30	0.74	0.37
	0.97	0.62	0.52	0.49	0.43	0.34	0.65	0.39
	1.04	0.64	1.36	0.73	0.71	0.28	0.83	1.14

Fig. 8 Relative assembly average flux errors (in %)

Table 3 Assembly-averaged power errors and nodal power errors based on the discretization in Figs. 1–4. The maximum and average thermal flux errors are edited from the full calculation, including reflector regions.

Calculation	Assembly averaged power error (ave., max.) (%)	Maximum nodal power error (%)
NEM (refined)	(0.92, 2.24)	3.68
ANM	(0.57, 1.27)	3.79

Table 4 Eigenvalue (k -eff) and timing results for the PBMR problem

Calculation	k -eff (error)	Time
Benchmark based FD reference	1.00463	~600 min
NEM (twice refined)	1.00461 (2 pcm)	361 s
ANM (coarse mesh)	1.00462 (1 pcm)	25 s

as specified in Ref. [4] are applied in all calculations. The reference solution was obtained by spatial mesh refinement studies until a spatially converged solution was obtained using the finite-difference method in BOLD VENTURE [5]. The reference ($R \times Z$) fine mesh was obtained to be (585×1450) . The ANM method uses the coarse mesh associated with the material mesh, as shown in the figure in the Appendix. The NEM method uses twice refined mesh as compared with the ANM mesh.

The use of a common cross section library without any state-parameter dependencies has resulted in a small spread in k -eff, as can be seen in Table 4. Power and flux profiles (averaged radially and axially), as predicted by ANM and NEM, compare very well with the reference solution, as can be seen in Figs. 9–12. Larger differences are observed when 2D two-group flux maps are com-

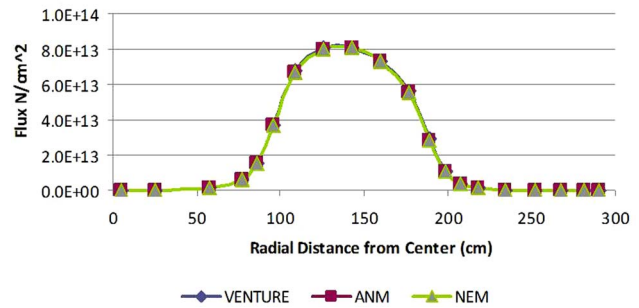


Fig. 9 Exercise 1 results for radial fast flux profiles

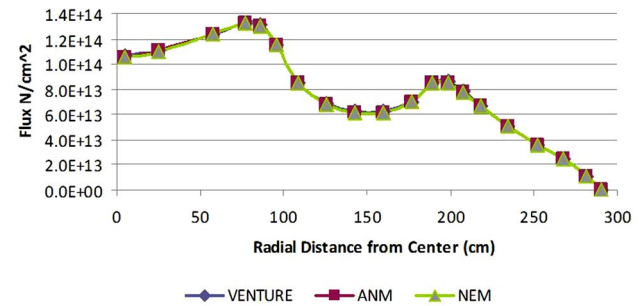


Fig. 10 Exercise 1 results for radial thermal flux profiles

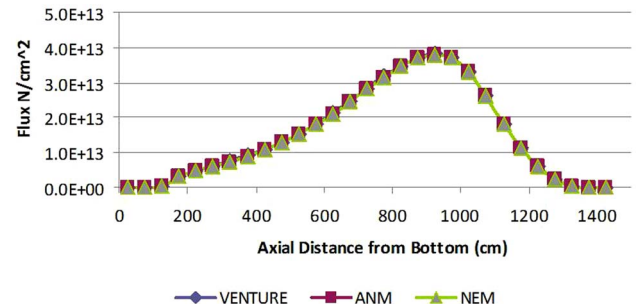


Fig. 11 Exercise 1 results for axial fast flux profiles

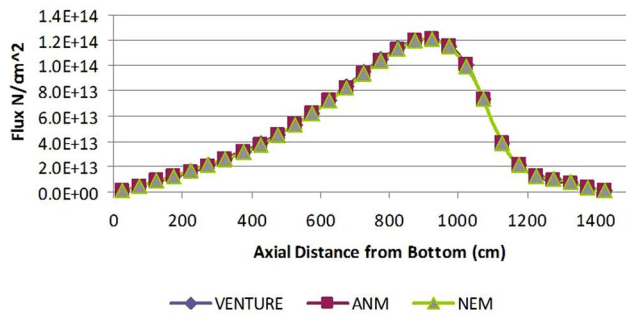


Fig. 12 Exercise 1 results for axial thermal flux profiles

pared. These comparisons are performed on the material mesh as shown in the figure in the Appendix. The summary of the comparative analysis only for the core region is shown in Table 5. The differences, which are much larger in the reflector regions up to 8%, may be caused by differences in solution methods (some convergence difficulties are experienced with NEM at the outer boundaries of the system due to the very large dimensions of the reflector—when compared with typical water-moderated reactors), boundary conditions, or remaining model approximations. One of these is the treatment used to represent the streaming ef-

Table 5 Material mesh-averaged flux errors only in the core region

Calculation	Fast flux error (ave., max.) (%)	Thermal flux error (ave., max.) (%)
NEM (twice refined)	(0.51, 0.71)	(0.51, 1.05)
ANM (coarse mesh)	(0.80, 2.52)	(0.82, 2.89)

fects in the void regions defined to be treated by directional diffusion coefficients. This is confirmed to some extent by the detailed flux results, where the largest deviations are seen close to the void areas.

The timing comparisons shown in Table 4 demonstrate that for both nodal methods a significant performance increase was found as compared with spatially converged finite-difference calculations with BOLD VENTURE (which uses the same CITATION based finite-difference formulation as the method currently employed by the PBMR (Pty) Ltd. for design analysis).

5 Conclusions

In this work we compare two nodal approaches—NEM and ANM (implemented in OSCAR)—in terms of accuracy and performance. Representative problems, selected to test the methods thoroughly, were devised and based on both a modified version of the PBMR 400 MW benchmark problem and a cylindrical version of the IAEA two-group problem. The comparative results between OSCAR and NEM are given, focusing on global reactivity estimation, as well as power and flux errors, as compared with reference finite-difference solutions.

These results indicate that both OSCAR and NEM recover the global reference solution for the IAEA problem and show power errors that are generally acceptable for nodal methods. For the PBMR problem, the accuracy is similar, but some convergence difficulties are experienced at the outer boundaries of the system due to the very large dimensions of the reflector (when compared with typical water-moderated reactors).

For both codes, a significant performance increase was found as compared with finite-difference calculations, which is the method currently employed by the PBMR (Pty) Ltd. In conclusion it seems that nodal methods have potential for use in the HTR analysis and, specifically, PBMR calculational arena, although cylindrical geometry based nodal methods will have to develop toward maturity before becoming the industry standard.

Appendix

Exercise 1: PBMR core neutronics model cylindrical R-Z view.

	0	10	41	73.6	80.55	92.05	100	117	134	151	168	185	192.95	204.45	211.4	225	243.6	260.6	275	287.5	292.5
-200	10	31	32.6	6.95	11.5	7.95	17	17	17	17	17	17	7.95	11.5	6.95	13.6	18.6	17	14.4	12.5	5
-150	50	133	133	133	133	155	116	113	113	113	113	113	135	164	144	144	152	152	152	189	190
-100	50	133	133	133	133	155	116	113	113	113	113	113	135	164	144	144	152	152	152	189	190
-50	50	133	133	133	133	155	116	112	112	112	112	112	135	164	144	144	152	152	152	189	190
0	50	133	133	133	133	155	116	111	111	111	111	111	135	165	144	144	152	152	152	189	190
50	50	134	134	134	125	156	117	1	23	45	67	89	136	166	145	145	153	153	153	189	190
100	50	134	134	134	125	156	117	2	24	46	68	90	136	167	145	145	153	153	153	189	190
150	50	134	134	134	126	157	118	3	25	47	69	91	137	168	146	146	153	153	153	189	190
200	50	134	134	134	126	157	118	4	26	48	70	92	137	169	146	146	153	153	153	189	190
250	50	134	134	134	126	157	118	5	27	49	71	93	137	170	146	146	153	153	153	189	190
300	50	134	134	134	127	158	119	6	28	50	72	94	138	171	147	147	153	153	153	189	190
350	50	134	134	134	127	158	119	7	29	51	73	95	138	172	147	147	153	153	153	189	190
400	50	134	134	134	127	158	119	8	30	52	74	96	138	173	147	147	153	153	153	189	190
450	50	134	134	134	127	158	119	9	31	53	75	97	138	174	147	147	153	153	153	189	190
500	50	134	134	134	128	159	120	10	32	54	76	98	139	175	148	148	153	153	153	189	190
550	50	134	134	134	128	159	120	11	33	55	77	99	139	176	148	148	153	153	153	189	190
600	50	134	134	134	128	159	120	12	34	56	78	100	139	177	148	148	153	153	153	189	190
650	50	134	134	134	128	159	120	13	35	57	79	101	139	178	148	148	153	153	153	189	190
700	50	134	134	134	129	160	121	14	36	58	80	102	140	179	149	149	153	153	153	189	190
750	50	134	134	134	129	160	121	15	37	59	81	103	140	180	149	149	153	153	153	189	190
800	50	134	134	134	129	160	121	16	38	60	82	104	140	181	149	149	153	153	153	189	190
850	50	134	134	134	129	160	121	17	39	61	83	105	140	182	149	149	153	153	153	189	190
900	50	134	134	134	130	161	122	18	40	62	84	106	141	183	150	150	153	153	153	189	190
950	50	134	134	134	130	161	122	19	41	63	85	107	141	184	150	150	153	153	153	189	190
1000	50	134	134	134	130	161	122	20	42	64	86	108	141	185	150	150	153	153	153	189	190
1050	50	134	134	134	131	162	123	21	43	65	87	109	142	186	151	151	153	153	153	189	190
1100	50	134	134	134	131	162	123	22	44	66	88	110	142	187	151	151	153	153	153	189	190
1150	50	132	132	132	132	163	124	114	114	114	114	114	143	188	151	151	154	154	154	189	190
1200	50	132	132	132	132	163	124	115	115	115	115	115	143	188	151	151	154	154	154	189	190
1250	50	132	132	132	132	163	124	115	115	115	115	115	143	188	151	151	154	154	154	189	190

References

- [1] 2005, NEM Theory and Input Manual, Revision 6, Pennsylvania State University.
- [2] Prinsloo, R. H., and Tomašević, D. I., 2006, "Development of the Analytic Nodal Method in Cylindrical Geometry," HTR 2006 Conference, Johannesburg, South Africa.
- [3] Lee, R., 1977, "Benchmark Problem Book, Suppl. 2," Argonne Code Center Report No. ANL-7416.
- [4] Reitsma, F., 2007, "PBMR Coupled Neutronics/Thermal Hydraulics Transient Benchmark—The PBMR-400 Core Design Benchmark Definition," Nuclear Energy Agency Draft V07.
- [5] Kamerow, S., and Gehin, J., 2007, "Steady-State Analysis of the OECD PBMR-400 Benchmark With Bold Venture," 2007 ANS Winter Meeting, Washington, DC.

On the Performance of Very High Temperature Reactor Plants With Direct and Indirect Closed Brayton Cycles

Mohamed S. El-Genk

Jean-Michel Tournier

Institute for Space and Nuclear Power Studies
and Department of Chemical and Nuclear
Engineering,
University of New Mexico,
Albuquerque, NM 87131

The performance of very high temperature reactor plants with direct and indirect closed Brayton cycles (CBCs) is compared and the effects of the molecular weight of the CBC working fluid on the number of stages and sizes of the axial flow, single shaft compressor and turbine are investigated. The working fluids considered are helium (4 g/mole), He–Xe, and He–N₂ binary mixtures (15 g/mole). Also investigated are the effects of using low and high pressure compressors with intercooling, instead of a single compressor, and changing the reactor exit temperature from 700°C to 950°C on the plant thermal efficiency, the CBC pressure ratio, and the number of stages in and size of the turbomachines. For plants with direct CBCs, the effect of cooling the reactor pressure vessel with He bled off at the exit of the compressor is also investigated. The present analyses are performed for a reactor thermal power of 600 MW, shaft rotation speed of 3000 rpm, and intermediate heat exchanger temperature pinch of 50°C.

[DOI: 10.1115/1.3126264]

1 Introduction

Helium-cooled very high temperature reactor (VHTR) concepts with direct and indirect closed Brayton cycles (CBCs) for energy conversion are being developed in the United States, Europe, Russia, Japan, and South Africa [1–4]. With reactor exit temperatures of 850–950°C the plant has a thermal efficiency >40% and can provide process heat for the cogeneration of hydrogen using thermochemical processes [1–4] and for other industrial uses. Such high temperatures, however, pose challenges to the selection of materials for the reactor pressure vessel (RPV) and the heat exchangers and to the fuel performance, to mention a few.

In order to use steel alloys, bleed cooling of RPV would be necessary to keep the vessel wall temperature below 371°C (644 K); in this paper this option is used only with helium direct CBC (Fig. 1) [5,6]. Without bleed cooling of the RPV, the returning helium gas from the recuperator cools the RPV wall only to or below 490°C (763 K) before entering the reactor core. This relatively high vessel wall temperature favors using Mo–Cr alloys, requiring further development and testing and could delay deployment date.

While helium has the best transport and thermal properties of all noble gases, its low molecular weight (4 g/mole) significantly increases the size and number of stages of the CBC turbomachines [7,8]. Mixing helium with xenon to a molecular weight of 15 g/mole increases the forced convection heat transfer coefficient by ~7% and significantly reduces the aerodynamic loading of the compressor and turbine blades [7]. Similar results have been reported for He–N₂ binary mixture with the same molecular weight [8].

Nitrogen is relatively inexpensive and more abundant than xenon, however, potential nitriding and embrittlement of the structure is a concern with high temperature operation for extended periods. The Xe and N₂ gases should be excluded as reactor coolants because they decrease the beginning of life excess reactivity

due to neutron capture and contaminate CBC loop components with irradiation products. Examples are ¹²⁵Xe ($T_{1/2}=16.9$ h) and ¹³⁴Cs ($T_{1/2}=2.06$ y), which emit high energy gammas, and ¹⁴C ($T_{1/2}=5715$ y) and ³H ($T_{1/2}=12.32$ y), which emit beta particles [9]. Therefore, the use of binary mixtures of He–Xe and He–N₂ should be limited to indirect CBCs (Fig. 1(b)).

VHTR plants with both direct and indirect CBCs for energy conversion (Figs. 1(a) and 1(b)) are being considered. With direct CBC, the reactor coolant is also the CBC working fluid, simplifying the plant layout. Also, for the same turbine inlet temperature, the reactor exit temperature is lower than in the plants with indirect CBCs. Plants with indirect CBCs could use helium for cooling the reactor and higher-molecular binary mixture CBC working fluids for reducing the size and number of stages of the turbomachines. In addition to minimizing the impact of changing the electrical load demand on the reactor operation, indirect CBC provides flexibility in maintenance and operation and for supplying process heat to a hydrogen generation facility, but decreases the plant thermal efficiency.

Using low pressure compressor (LPC) and high pressure compressor (HPC) with intercooling in direct and indirect CBCs decreases the compression work and the bleed fraction for cooling the electrical generator, turbine disks, and shaft bearings (Figs. 1(a) and 1(b)). Bleeding off helium at the exit of the compressor to cool the RPV is used in the Japanese GTHT300 plant with direct helium CBC [5,6] (Fig. 1(a)). In this plant, bleeding off only 0.2% of the helium flow was enough [5] to keep the RPV wall temperature at or below 371°C (644 K). Since RPV bleed cooling is not practically possible for the plants with indirect CBCs, the temperature of the primary loop helium entering the reactor is kept constant at or below 490°C (763 K) [10], as in the GT-MHR [4] and the NGNP [1,3] plants (Fig. 1(b)).

To determine the impact of using direct or indirect CBC with different working fluids on the plant thermal efficiency and the size and number of stages of the CBC turbomachines, detailed models of the compressor and the turbine are developed [8]. These models, only briefly summarized in this paper, have been validated for the GTHT300 helium compressor and turbine [11,12].

Manuscript received October 27, 2008; final manuscript received October 27, 2008; published online November 24, 2009. Review conducted by Dilip R. Ballal. Paper presented at the 4th International Topical Meeting on High Temperature Reactor Technology (HTR2008), Washington, DC, September 28–October 1, 2008.

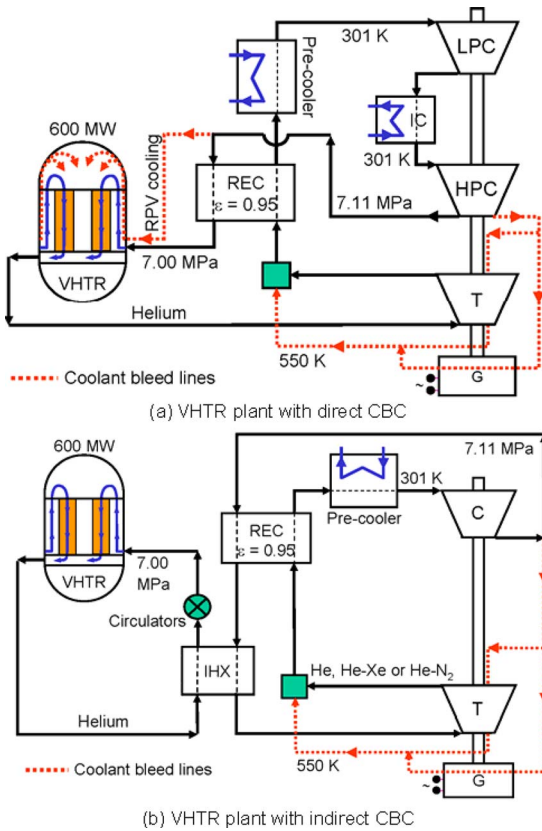


Fig. 1 Direct and indirect closed Brayton cycles

The determined number of stages, polytropic efficiencies, pressure ratios, and the hardware details were in excellent agreement with the reported values [8].

To investigate the effect of the molecular weight of the CBC working fluid on the performance of the VHTR plant with indirect CBC, the present analyses considered helium in direct CBC plants and helium and binary mixtures of He–Xe and He–N₂ (15 g/mole) in indirect CBC plants. Also investigated are the effects of using LPC and HPC with intercooling instead of a single compressor, using bleed cooling of the RPV in the plants with helium direct CBC (Fig. 1(a)), and changing the reactor exit temperature from 700°C to 950°C on the plant thermal efficiency, cycle pressure ratio, polytropic efficiency, and the number of stages in and size of the turbomachines. The present analyses are performed for a shaft rotation speed of 3000 rpm, reactor thermal power of 600 MW, and a 50°C temperature pinch in the intermediate heat exchanger (IHx) for the plants with indirect CBCs (Fig. 1(b)).

2 Model Description

The present compressor and turbine design and performance models are based on a mean-line analysis of vortex free flow along the blades [13–17]. The compressor and turbine are designed for constant axial flow velocity V_x in the stages and constant mean-line blade radius r_m . In all stages, the mean flow coefficient φ_m is kept constant, assuming uniform loading coefficient λ_m and reaction \mathfrak{R}_m . The gas flow enters the first stage axially, $V_{in} = V_x$, and exits the last stage also axially, $V_{ex} = V_x$.

Figures 2(a) and 2(b) present typical velocity triangles at the leading and trailing edges of the turbine and compressor rotor blade cascades. The x -coordinate coincides with the axis of the turbomachine, and θ is the angular coordinate. The subscripts 1 and 2 refer to the cascade inlet and exit stations. The absolute velocity in the stationary coordinate system V becomes W in the

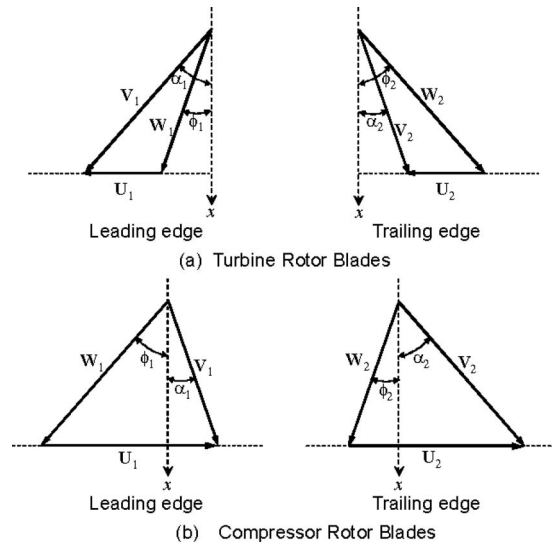


Fig. 2 Velocity triangles for turbine and compressor

rotating coordinate system. The tangential velocity of the rotating blades $U = U_\theta$ and α and ϕ are the absolute and relative velocity angles. The gas turning angle of the compressor blades is much smaller than the turbine, limiting the boundary layer separation due to a positive pressure gradient. In addition to the velocity triangles, other input to the turbomachine models includes \dot{N} , P_{in} , T_{in} , the shaft mechanical work \dot{W} , the number of rotor stages, and the shaft rotational speed ω . The blades aspect ratio for the compressor stators $H/C = 1.7$ and 1.4 for all other blades. The maximum thickness ratio for the turbine blades $t_{max}/C = 0.2$ and 0.1 for the compressor blades (Fig. 3). The trailing edge thickness ratio for all turbine cascades, except the exit guide vanes (EGVs) $t_{TE}/S = 0.02$ [16,17]; while for all compressor cascades, $t_{TE}/C = 0.00046$ (same as for the symmetric NACA 65A010 compressor airfoils [14]). When the blades are shrouded, the blade tip clearance $\tau = 1$ mm and the number of tip seals are specified in the input. For the compressor blades and the turbine EGVs, the relative position of the maximum camber is $Z/C = 0.5$. The thermodynamic and transport properties of the working fluids are calculated using reported semi-empirical relationships developed and validated successfully against a large experimental database for temperature from 300–1400 K and pressure of up to 20 MPa [7,18,19].

2.1 Compressor Model. The compressor model accounts for the profile, secondary, end wall, and tip clearance pressure losses. The profile loss coefficient is calculated using the approach by Koch and Smith [20] and the secondary flow loss coefficient is calculated using the correlation proposed by Howell [21]. The end wall losses are calculated using the approach of Aungier [14] and the tip clearance losses are calculated using the model of Yaras and Sjolander [22] for the turbine blades, after substituting $\cos(\phi_2)$ for $\cos(\phi_1)$. This is because the compressor blade losses are calculated based on the inlet flow conditions.

The values of (Z/C) , the maximum camber (b/C) , and the blade angles measured from the chord line (χ_1 and χ_2) uniquely define the parabolic-arc camberline profile [14] of the compressor blades (Figs. 3(b) and 4). The blades camber angle $\theta = \chi_1 + \chi_2 = \beta_1 - \beta_2$. The leading edge (LE) and trailing edge (TE) angles of the blades β_1 and β_2 are calculated from the determined optimum incidence and deviation angles. The compressor model uses a solidity $\sigma = 1.10$ for the inlet guide vanes (IGVs) and computes the optimum solidity for the other cascades using a developed correlation based on the optimum values of σ for minimizing the sum of the profile and the secondary flow loss coefficients.

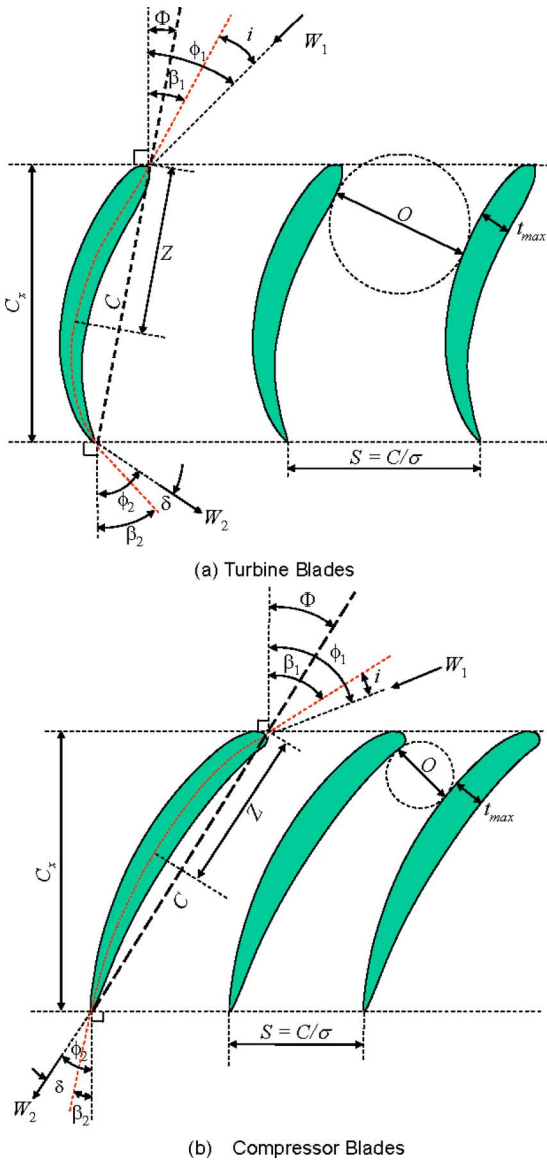


Fig. 3 Turbine and compressor blade cascades

2.2 Turbine Model. The parabolic-arc camberline profile (Figs. 3(a) and 4) is also used in the turbine model, but with caution, to predict the blade cascade geometry. The turning angle and the maximum thickness of the turbine blades are larger than those of the compressor blades, and $\beta_1 < \beta_2$. The turbine blades camber angle is $\theta = \chi_1 + \chi_2 = \beta_1 + \beta_2$. The optimum solidity for the

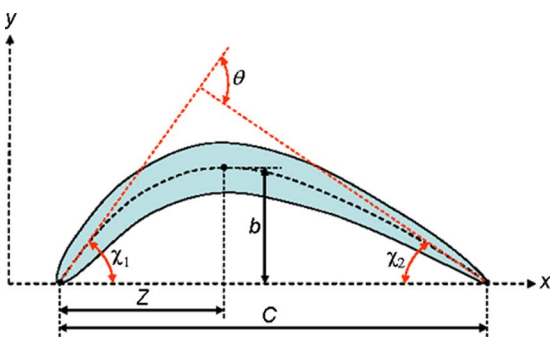


Fig. 4 Blade profile with parabolic-arc camberline

turbine blades cascade is calculated using the correlation by Aungier [15], based on matching the Ainley–Mathieson [17] minimum profile loss coefficient. The turbine blades are designed for a zero incidence angle and the deviation angle at the trailing edge is calculated using a correlation by Zhu and Sjolander [23].

Because the turning angle of turbine blades is larger than in the compressor, they are very sensitive to the value of the maximum camber position (Z/C), which may not be known accurately. The turbine model uses the values recommended by Kacker and Okapu [16] for predicting the staggered angle of typical turbine blade cascades Φ as a function of β_1 and β_2 . This staggered angle is used to determine the blade angles with respect to the chord line (Fig. 4). The location of the maximum camber [Z/C , b/C] is obtained from simultaneously solving equations describing the blade shape. The pressure losses in the turbine are calculated using recommended correlations from various sources [24–26].

3 Plant Analyses

The plant analyses assume recuperator effectiveness of 0.95, compressor inlet temperature of 301 K, and reactor inlet pressure of 7.0 MPa (Figs. 1(a) and 1(b)). The pressure losses in the cold and hot legs of the recuperator and in the pre- and intercoolers are assumed to be 1.5% and 1.9%, and 1.5% of inlet pressure (Figs. 1(a) and 1(b)). The assumed shaft mechanical efficiency is 99% and the generator efficiency is 98.7%. The analyses also assume fixed flow channels dimensions in the VHTR core.

For the plants with direct and indirect He-CBCs (Figs. 1(a) and 1(b)), the bled off flow of the working fluid at the exit of the compressor cools the turbine casing and disks, shaft bearings, and the electrical generator. For the plants with He-direct CBC, additional flow is bled off at the exit of the compressor to cool and keep the temperature of the RPV wall $\leq 371^\circ\text{C}$ (644 K) (Fig. 1(a)). With such RPV bleed cooling, the inlet temperature to the reactor core increases freely as the reactor exit temperature increases.

The bled off flow for cooling the generator and turbomachines is calculated assuming a temperature of 550 K at the inlet of the mixing chamber (Figs. 1(a) and 1(b)); it varies from 1% to 13%. This temperature is dictated by the electrical insulation of the generator. In the plants with indirect CBCs, analyses assume a 50°C temperature pinch in the IHX (Fig. 1(b)). The losses in the IHX primary leg are taken to be equal to 1.7% of the inlet pressure. The electrical power to circulate the primary loop helium (Fig. 1(b)), calculated assuming a circulator isentropic efficiency of 87%, varies from 10 MW to 35 MW, depending on the operating conditions. As a result, when heat losses are neglected the thermal power transferred to the CBC loop through the IHX varies between 610 MW and 635 MW.

4 Results and Discussion

The present analyses of the VHTR plants with direct and indirect CBCs, with and without intercooling, are performed for reactor exit temperatures of 973–1223 K (Figs. 1(a) and 1(b)). For the plant with direct He-CBC, the analyses are performed with and without RPV bleed cooling. For the plants with indirect He, He–Xe and He– N_2 CBCs, the performed analyses are for constant reactor inlet temperature of 490°C (763 K).

4.1 Direct He-CBC Plants With RPV Cooling. Without intercooling, there is only one compressor, while with intercooling, the LPC, and HPC operate at the same inlet temperature of 301 K and same compression ratio (Figs. 1(a) and 1(b)). When cooling the RPV with helium bled off at the exit of the HPC, the plant operates near its peak thermal efficiency. The low cycle pressure ratio (2.25–2.3) increases very little as the reactor exit temperature increases (Figs. 5(a) and 5(b)). The nominal operation points of the VHTR plants with intercooling are indicated by the blue circular solid symbols in Fig. 5(b). The corresponding plant thermal efficiency increases steadily with increasing the reactor exit tem-

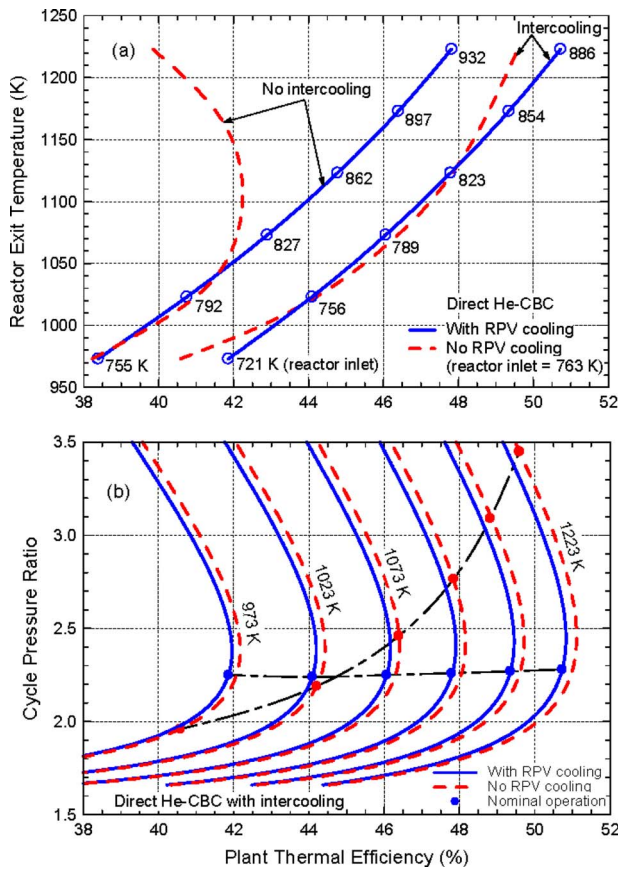


Fig. 5 Effect of RPV cooling on performance of direct He-CBC VHTR plants

perature, but remains very close to its peak value (Fig. 5(b)). It increases steadily from 41.8% to 51.5% as the reactor exit temperature increases from 973 K to 1223 K. The reactor inlet temperature also increases from 721 K to 886 K as its exit temperature increases from 973 K to 1223 K. Without intercooling, the reactor inlet temperature is slightly higher than with intercooling, increasing from 755 K to 932 K as the reactor exit temperature increases from 973 K to 1223 K (Fig. 5(a)). The plant with non-intercooled direct He-CBC has lower thermal efficiency (Fig. 5(b)), which increases steadily from 38.4% to 47.6% as the reactor exit temperature increases from 973 K to 1123 K (Fig. 5(a)).

By contrast, without bleed cooling of the RPV the inlet temperature to the reactor core is constant (490°C), and the cycle pressure ratio increases, but the plant thermal efficiency decreases as the reactor exit temperature increases (Figs. 5(a) and 5(b)). When the reactor exit temperature is >1023 K, the thermal efficiency of the plant without intercooling decreases steadily and the cycle pressure ratio increases with increasing the reactor exit temperature. The plant thermal efficiency is much lower and the cycle pressure ratio is much higher than in the plant with LPC, HPC, and intercooling (Figs. 5(a) and 5(b)). For reactor exit temperatures <1023 K, both the thermal efficiency and the cycle pressure ratio of the plant with direct He-CBC and without intercooling are lower than those with intercooling.

The differences in the thermal efficiency and the cycle pressure ratio of the plants with He-direct CBC without RPV bleed cooling and intercooling are significantly lower and higher than those of the plants with RPV bleed cooling, but without intercooling. The results in Figs. 5(a) and 5(b) clearly indicate that for the best performance of the VHTR plants with direct He-CBCs, using RPV cooling, LPC, and HPC with intercooling is recommended. With intercooling, however, the combined number of stages is

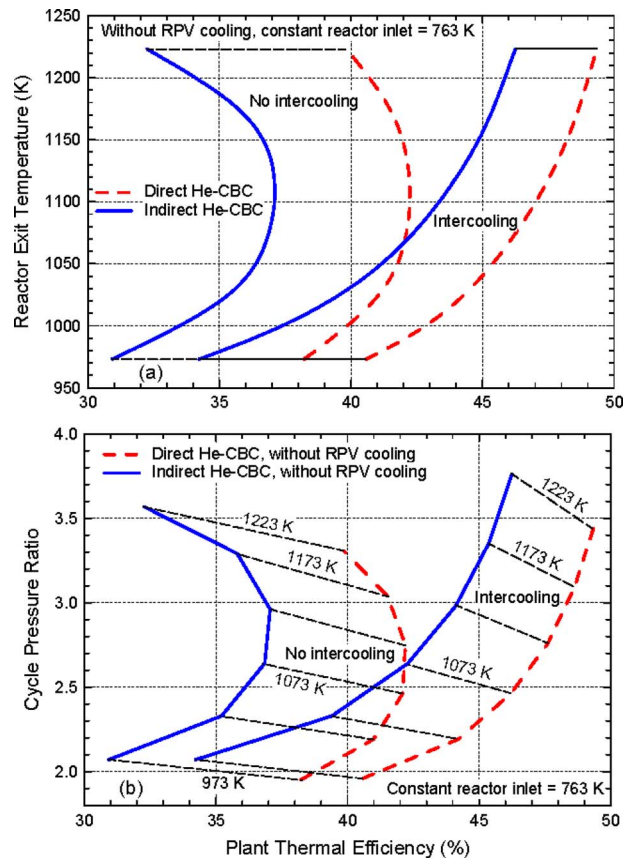


Fig. 6 Direct and indirect He-CBC VHTR plants without RPV cooling

higher and the size of the compressors is larger, as will be shown later in Figs. 8 and 9. The pressure ratio and the exit temperature of the LPC and HPC are much lower than for a single compressor, resulting in lower helium bleed fraction for cooling the generator and the shaft bearings (Fig. 1(a)).

4.2 Direct and Indirect He-CBCs Without RPV Cooling.

Figures 6(a) and 6(b) compare the performance of the VHTR plants with direct and indirect He-CBCs, with and without intercooling, but without bleed cooling of the RPV. For these plants the reactor inlet temperature is constant at 490°C (763 K). For the same reactor exit temperature, the plants with intercooling operate at slightly higher cycle pressure ratios but significantly higher thermal efficiencies. Using intercooling steadily increases the plant thermal efficiency as the reactor exit temperature increases. Without intercooling, the plant efficiency increases as the reactor exit temperature increase up to ~1123 K, then decreases with a further increase in the reactor exit temperature (Figs. 6(a) and 6(b)). The differences in the performance of the plants with indirect and direct He-CBCs are primarily due to the 50 K pinch in the IHX (Fig. 1(b)). Thus, for the same reactor exit temperature, the turbine of the indirect He-CBC operates cooler than in the direct cycle, resulting in 5.5–7.5% point decrease in the thermal efficiency of the plant without intercooling, and 3–6% point drop in the efficiency of the plant with intercooling.

4.3 Plants Performance With Indirect CBCs. This section compares the performance of VHTR plants with indirect CBCs, with and without intercooling, and without RPV bleed cooling (Figs. 7(a) and 7(b)). The results for He–Xe and He–N₂ working fluids (15 g/mole) are compared with those of helium, at the same reactor exit temperature. With He–N₂ and He–Xe binary mixture working fluids, the pressure losses in the CBC loop increase by a

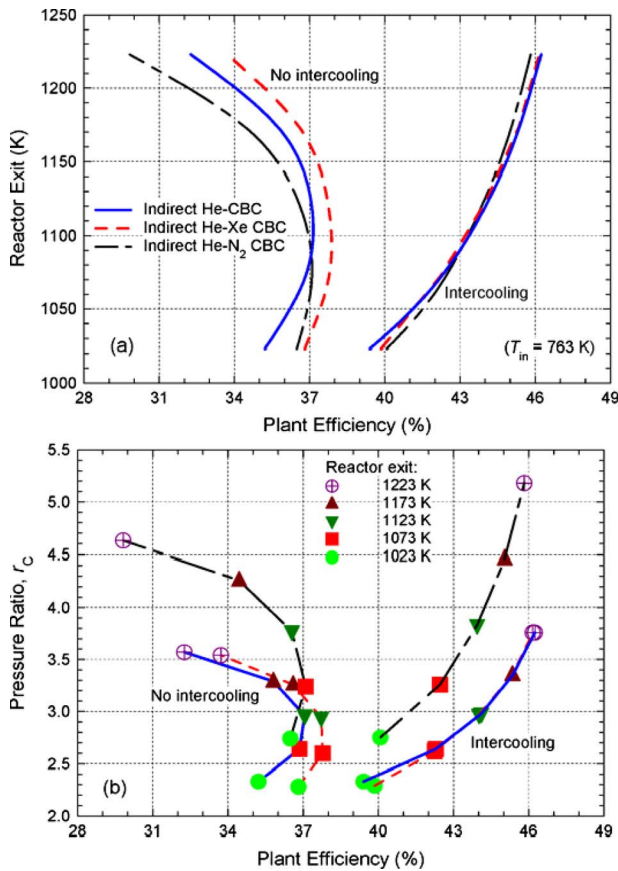


Fig. 7 Indirect CBC VHTR plants without RPV cooling

factor 2.7, compared with those in the He-CBC [7], when operating at the same temperature, pressure, and molar flow rate. Thus, with He-N₂ or He-Xe (15 g/mole), the analyses assume the diameters of the flow channels of the indirect CBC loop increase by 23% in order to maintain the same relative pressure losses as with helium [7]. The turbines for the higher-molecular weight working fluids (He-N₂ and He-Xe (15 g/mole)) are designed for a higher mean-radius flow coefficient of 0.60 (compared with 0.433 for He), which corresponds to near optimum polytropic efficiency [13,15].

Figures 7(a) and 7(b) show that for the same reactor exit temperature, the plants with indirect He and He-Xe CBCs operate at the same cycle pressure ratio (Fig. 7(b)). This is because He-Xe behaves essentially as a pure noble gas with a specific heat ratio of ~1.67. By contrast, the plant with indirect He-N₂ CBC operates at a higher cycle pressure ratio (Fig. 7(a)), because He-N₂ of the same molecular weight of 15 g/mole has a lower specific heat ratio (~1.45) [7,8,19]. The plants with intercooling operate at slightly higher pressure ratios than the plants without intercooling because of higher pressure losses in the former (Fig. 7(b)).

At a reactor exit temperatures <1090 K, the plants with He-N₂ indirect CBCs without intercooling deliver higher performance than those with He-CBCs without intercooling (Fig. 7(a)). This is because the He-N₂ turbine and compressor have higher polytropic efficiencies (93.2% and 90.9%) than their He counterparts. At higher reactor exit temperatures, however, the plant with the nonintercooled indirect He-N₂ CBC has lower thermal efficiency than that with indirect He-CBC. This is because the former operates at a higher cycle pressure ratio, which translates into higher compressor exit temperature and higher bleed flow for cooling the shaft bearings and electrical generator (Fig. 1(b)).

The thermal efficiency of the VHTR plant with indirect He-Xe CBC, but without intercooling is 1–1.6% points higher than the

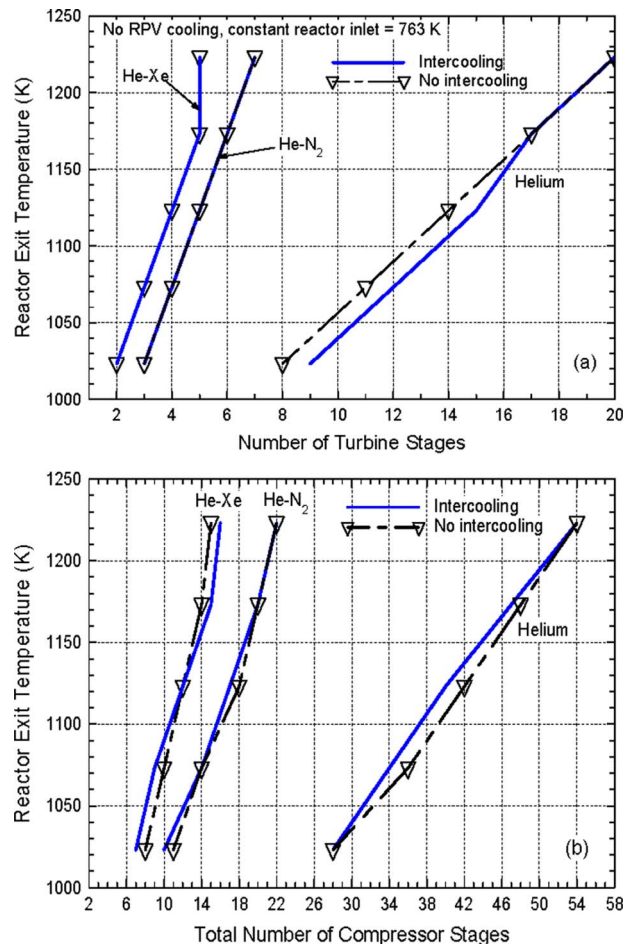


Fig. 8 Number of stages in turbomachine with He, He-Xe, and He-N₂ indirect-CBCs

plant with indirect He-CBC (Fig. 7(a)). The two CBCs operate at the same pressure ratio, but the higher-molecular weight turbine and compressor are more efficient, having polytropic efficiencies of 93.3% and 90.8%. By contrast, the helium turbine and compressor have polytropic efficiencies of 92.3% and 90.5%, which are nearly independent of the VHTR exit temperature, using the design rules of this work.

The VHTR plants with intercooled indirect CBCs operate at higher thermal efficiencies than the plants without intercooling (Fig. 7(a)). The CBC working fluid has very little effect on the thermal efficiency of the plants employing indirect CBCs with intercooling. The turbine has the same polytropic efficiency as that in the nonintercooled CBC, but the LPC and HPC are slightly less efficient than the single compressor in the CBC without intercooling. The polytropic efficiencies of the He, He-Xe, and He-N₂ LPCs are 90.3%, 90.3%, and 90.4%, and those of the HPCs are 90.1%, 90.1%, and 90.4%, respectively.

4.4 Effect of Working Fluid on Turbomachine Size. For the plants employing direct He-CBCs with RPV bleed cooling, increasing the reactor exit temperature increases both its inlet temperature and the number of stages of the turbine and compressor. In the plants with intercooling, the combined number of stages of the helium LPC and HPC are higher than of the single compressor in the CBC of the plants without intercooling. Also, the number of stages in the helium turbine for the CBC with intercooling is higher than the helium turbine in the direct-CBC without intercooling.

For the plants with direct He-CBC without intercooling, when operating at a reactor exit temperature of 1173 K the turbine has 8

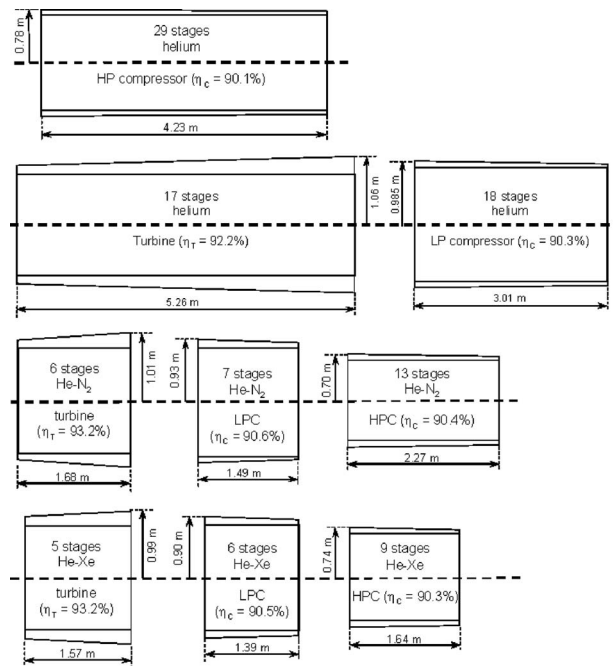


Fig. 9 Comparison of the size and dimensions of single-shaft turbomachine with intercooling

stages and the compressor has 25 stages. By contrast, for plants with the direct He-CBC with intercooling, the turbine has 11 stages, the LPC has 13 stages, and the HPC has 18 stages, for a total of 31 stages.

As shown in Figs. 8(a) and 8(b), the number of stages of the He-Xe and He-N₂ compressors and turbines for plants with the indirect CBCs is considerably smaller than in the He-CBC plants, when operating at the same reactor exit temperature. The size and the number of stages in the turbine and the compressor, however, increase as the reactor exit temperature increases (Figs. 8(a) and 8(b)). At a reactor exit temperature of 900°C (1173 K), the He-Xe turbine has only 5 stages and the He-N₂ turbine has 6 stages, compared with the 17 stages in the helium turbine (Figs. 8(a) and 9).

Similarly, the plants with He-Xe and He-N₂ indirect CBCs require 15 and 20 compression stages, compared with 47 stages for the He-CBC plants (Figs. 8(b) and 9). As a result, the rotor shafts of the He-N₂ and He-Xe turbomachines are a factor 2 and 3 shorter than that of the helium turbomachines (Fig. 9). In addition to the reduced capital cost of the turbomachines; shorter shafts are dynamically more stable and would have longer operation life. Also, for these shafts the lighter rotor's bearings are easier to design.

5 Summary and Conclusions

Design models of multistage axial flow turbomachines for CBCs are developed and the effect of the molecular weight of the CBC working fluid on the VHTR plant performance and the number of stages in and the size of the compressor and turbine are investigated. The working fluids considered are helium in the VHTR plants with direct and indirect CBCs, He-Xe, and He-N₂ with a molecular weight of 15 g/mole for the plants with indirect CBCs. Investigated are the effects on the plant thermal efficiency, cycle pressure ratio, and number of stages in and the size of the turbomachines of the following: (a) using LPC and HPC with intercooling in direct and indirect CBCs, instead of a single compressor, (b) cooling the VHTR pressure vessel with He bled off at the exit of the compressor versus keeping the helium inlet temperature to the reactor constant at 490°C in the plants with direct

He-CBCs, and (c) changing the reactor exit temperature from 700°C to 950°C. The plant analyses are performed for shaft rotation speed for the turbomachines of 3000 rpm, reactor thermal power of 600 MW, and temperature pinch of 50°C in the IHX for the VHTR plants with indirect CBCs (Fig. 1(b)).

Based on the obtained results, the following conclusions are reached.

- In VHTR plants with direct He-CBCs, when the reactor exit temperature >1023 K–1073 K (750–800°C), using bleed cooling of RPV and LPC and HPC with intercooling is highly recommended. The plant thermal efficiency is highest, and the cycle compression ratio is relatively low and increases very little as the reactor exit temperature increases.
- At reactor exit temperatures <1073 K, maintaining the reactor inlet temperature constant at 490°C, and using LPC and HPC with intercooling lowers the cycle pressure ratio and slightly lowers the plant thermal efficiency below that in (a) above. Thus, this plant with direct He-CBC and without RPV cooling may be used when the reactor exit temperature is 1023–1073 K.
- Using LPC and HPC with intercooling in VHTR plants with both direct and indirect CBCs steadily increases the cycle pressure ratio, the plant thermal efficiency, and the reactor inlet temperature as the reactor exit temperature increases. However, for the same cycle pressure ratio, intercooling significantly increases the thermal efficiency of the plant.
- The thermal efficiencies of the plants with indirect CBCs and LPC and HPC with intercooling are higher than those of the plants without intercooling, and almost independent of the molecular weight of the working fluid, but increase steadily with increased reactor exit temperature.
- Intercooling in the VHTR plants with direct He-CBCs and RPV bleed cooling decreases the reactor inlet temperature but increases the number of the compressor and turbine stages. The number of these stages also increases as the reactor exit temperature increases.
- In the helium plants employing CBCs without intercooling, the stages of the turbomachines are fewer than for the CBCs with intercooling, but the reactor inlet temperature is higher.
- In the plants with indirect CBCs and working fluids of He-Xe and He-N₂ (15 g/mole), the number of stages of the compressor and the turbine are significantly fewer and the length of the rotating shaft is much shorter than for the helium turbomachines. Such reductions in the size and mass of the turbomachines for the He-Xe and He-N₂ working fluids translate into considerable cost saving, higher operation reliability, and lower maintenance.

Acknowledgment

This research is sponsored by the University of New Mexico's Institute of Space and Nuclear Power Studies.

Nomenclature

- b = maximum camber of blade (m)
 C = actual chord length of blade (m); compressor
 C_x = axial chord length of blade (m), $C_x = C \times \cos \Phi$
 H = height of blades (m)
 M = molecular weight (kg/mole)
 \dot{N} = molar flow rate (mole/s)
 P = pressure (Pa)
 r = radius (m)
 \mathfrak{R} = stage reaction

S = pitch or distance between blades in cascade (m)
 t_{\max} = maximum blade thickness (m)
 t_{TE} = thickness of blades trailing edge (m)
 T = temperature (K); turbine
 U = rotor tangential velocity (m/s), $U=r\omega$
 V_x = gas meridional velocity component (m/s)
 V_θ = gas tangential velocity component (m/s)
 W = gas relative velocity vector (m/s), $W=V-U$
 Z = position of maximum camber (m)

Greek

α = angle between gas absolute velocity vector and meridional plane (deg)
 β = blade angle relative to meridional plane (deg)
 ΔP = total pressure loss (Pa)
 ε = recuperator effectiveness
 θ = blade camber (or turning) angle (deg)
 λ = stage loading (work) coefficient
 μ = coolant dynamic viscosity (kg/m s)
 ρ = density (kg/m³)
 σ = blade cascade solidity, $\sigma=C/S$
 τ = blades clearance gap (m)
 φ = stage flow coefficient, $\varphi=V_x/U$
 ϕ = angle between gas relative velocity vector and meridional plane (deg)
 Φ = blades staggered angle (deg)
 χ = blade angle measured from the chord line (deg)
 ω = Shaft angular speed (radians/s)

Subscripts/Superscripts

in = inlet flow
 m = median location of annular flow passage
 x, z = Axial component
 1 = inlet or leading edge of rotor blades cascade
 2 = exit or trailing edge of rotor blades cascade
 θ = tangential or "whirl" component

References

- [1] Idaho National Engineering and Environmental Laboratory, 2007, "Next Generation Nuclear Plant—Pre-Conceptual Design Report," Report No. INEEL/EXT-07-12967.
- [2] Oh, C. H., Barner, R., Davis, C., and Sherman, S., 2006, "Evaluation of Working Fluids in an Indirect Combined Cycle in a Very High Temperature Gas-Cooled Reactor," *Nucl. Technol.*, **156**, pp. 1–10.
- [3] Johnson, G. A., 2008, "Power Conversion System Evaluation for the Next Generation Nuclear Plant (NGNP)," *Proceedings of the International Congress on Advances in Nuclear Power Plants (ICAPP 08)*, American Nuclear Society, Anaheim, CA, Jun. 8–12, Paper No. 8253.
- [4] No, H. C., Kim, J. H., and Kim, H. M., 2007, "A Review of Helium Gas Turbine Technology for High-Temperature Gas-Cooled Reactors," *J. Nuclear Engineering and Technology*, **39**(1), pp. 21–30.
- [5] Yan, X., Kunitomi, K., Nakata, T., and Shiozawa, S., 2002, "Design and Development of GTHTR300," *Proceedings of the First International Topical Meeting on HTR Technology (HTR2002)*, Petten, Netherlands, Apr. 22–24, International Atomic Energy Agency, Vienna, Austria.
- [6] Yan, X., Takizuka, T., Takada, S., Kunitomi, K., Minatsuki, I., and Mizokami, Y., 2003, "Cost and Performance Design Approach for GTHTR300 Power Conversion system," *Nucl. Eng. Des.*, **226**, pp. 351–373.
- [7] El-Genk, M. S., and Tournier, J.-M., 2008, "Noble Gas Binary Mixtures for Gas-Cooled Reactor Power Plants," *Nucl. Eng. Des.*, **238**, pp. 1353–1372.
- [8] Tournier, J.-M., and El-Genk, M. S., 2008, "Alternative Working Fluids to Reduce Size of Turbomachinery for VHTR Plants," *Proceedings of the International Congress on Advances in Nuclear Power Plants (ICAPP 08)*, American Nuclear Society, Anaheim, CA, Jun. 8–12, Paper No. 8072.
- [9] Schriener, T. M., and El-Genk, M. S., 2008, "Neutronic Performance of High Molecular Weight Coolants for a Prismatic VHTR," *Proceedings of the International Congress on Advances in Nuclear Power Plants (ICAPP 08)*, American Nuclear Society, Anaheim, CA, Jun. 8–12, Paper No. 8104.
- [10] Vasyaev, A. V., Golovko, V. F., Dimitrieva, I. V., Kodochigov, N. G., Kuzavkov, N. G., and Rulev, V. M., 2005, "Substantiation of the Parameters and Layout Solutions for an Energy Conversion Unit With a Gas-Turbine Cycle in a Nuclear Power Plant With HTGR," *At. Energy*, **98**(1), pp. 21–31.
- [11] Kunitomi, K., Katanishi, S., Takada, S., Yan, X., and Tsuji, N., 2004, "Reactor Core Design of Gas Turbine High Temperature Reactor 300," *Nucl. Eng. Des.*, **230**, pp. 349–366.
- [12] Takizuka, T., Takada, S., Yan, X., Kosugiyama, S., Katanishi, S., and Kunitomi, K., 2004, "R&D on the Power Conversion System for Gas Turbine High Temperature Reactors," *Nucl. Eng. Des.*, **233**, pp. 329–346.
- [13] Fielding, L., 2000, *Turbine Design—The Effect on Axial Flow Turbine Performance of Parameter Variation*, ASME Press, New York, pp. 130–132.
- [14] Aungier, R. H., 2003, *Axial-Flow Compressor—A Strategy for Aerodynamic Design and Analysis*, ASME Press, New York, Chap. 6, pp. 118–152.
- [15] R. H. Aungier, 2006, *Turbine Aerodynamics—Axial-Flow and Radial-Inflow Turbine Design and Analysis*, ASME Press, New York, pp. 69–79.
- [16] Kacker, S. C., and Okapuu, U., 1982, "A Mean Line Prediction Method for Axial Flow Turbine Efficiency," *ASME J. Eng. Power*, **104**, pp. 111–119.
- [17] Ainley, D. G., and Mathieson, G. C. R., 1951, "A Method of Performance Estimation for Axial-Flow Turbines," *British Aeronautical Research Council, Reports and Memoranda No. 2974*.
- [18] Tournier, J.-M., and El-Genk, M. S., 2008, "Properties of Noble Gases and Binary Mixtures for Closed Brayton Cycle Applications," *Energy Convers. Manage.*, **49**(3), pp. 469–492.
- [19] Tournier, J.-M., and El-Genk, M. S., 2008, "Properties of Helium, Nitrogen, and He–N₂ Binary Gas Mixtures," *J. Thermophys. Heat Transfer*, **22**(3), pp. 442–446.
- [20] Koch, C. C., and Smith, L. H., Jr., 1976, "Loss Sources and Magnitudes in Axial-Flow Compressors," *ASME J. Eng. Power*, **A98**, pp. 411–424.
- [21] Howell, A. R., 1947, "Fluid Dynamics of Axial Compressors," *Lectures on the Development of the British Gas Turbine Jet Unit, War Emergency Proceeding No. 12*, Institution of Mechanical Engineers, London, pp. 441–452.
- [22] Yaras, M. I., and Sjolander, S. A., 1992, "Prediction of Tip-Leakage Losses in Axial Turbines," *ASME J. Turbomach.*, **114**, pp. 204–210.
- [23] Zhu, J., and Sjolander, S. A., 2005, "Improved Profile Loss and Deviation Correlations for Axial-Turbine Blade Rows," *Proceedings of the GT2005 ASME Turbo Expo 2005: Power for Land, Sea and Air*, Reno-Tahoe, NV, American Society of Mechanical Engineers, New York, pp. 783–792.
- [24] Benner, M. W., Sjolander, S. A., and Moustapha, S. H., 2006, "An Empirical Prediction Method for Secondary Losses in Turbines—Part I: A New Loss Breakdown Scheme and Penetration Depth Correlation," *ASME J. Turbomach.*, **128**, pp. 273–280.
- [25] Benner, M. W., Sjolander, S. A., and Moustapha, S. H., 2006, "An Empirical Prediction Method for Secondary Losses in Turbines—Part II: A New Secondary Loss Correlation," *ASME J. Turbomach.*, **128**, pp. 281–291.
- [26] Matsunuma, T., 2006, "Effects of Reynolds Number and Freestream Turbulence on Turbine Tip Clearance Flow," *ASME J. Turbomach.*, **128**, pp. 166–177.

Design Option of Heat Exchanger for the Next Generation Nuclear Plant

Chang H. Oh

Eung S. Kim

Mike Patterson

Idaho National Laboratory,
P.O. Box 1625,
Idaho Falls, ID 83415

The next generation nuclear plant (NGNP), a very high temperature gas-cooled reactor (VHTR) concept, will provide the first demonstration of a closed-loop Brayton cycle at a commercial scale, producing a few hundred megawatts of power in the form of electricity and hydrogen. The power conversion unit for the NGNP will take advantage of the significantly higher reactor outlet temperatures of the VHTRs to provide higher efficiencies than can be achieved with the current generation of light water reactors. Besides demonstrating a system design that can be used directly for subsequent commercial deployment, the NGNP will demonstrate key technology elements that can be used in subsequent advanced power conversion systems for other Generation IV reactors. In anticipation of the design, development, and procurement of an advanced power conversion system for the NGNP, the system integration of the NGNP and hydrogen plant was initiated to identify the important design and technology options that must be considered in evaluating the performance of the proposed NGNP. As part of the system integration of the VHTRs and the hydrogen production plant, the intermediate heat exchanger is used to transfer the process heat from VHTRs to the hydrogen plant. Therefore, the design and configuration of the intermediate heat exchanger are very important. This paper describes analyses of one stage versus two-stage heat exchanger design configurations and simple stress analyses of a printed circuit heat exchanger (PCHE), helical-coil heat exchanger, and shell-and-tube heat exchanger. [DOI: 10.1115/1.3126780]

1 Introduction

An intermediate heat exchanger (IHX) is one of the essential components in the very high temperature gas-cooled reactor (VHTR) systems, since it transfers reactor core heat to the other systems for the application of electricity generation or hydrogen production. Therefore, its effectiveness is directly related to the overall system efficiency and economics. Generally, the VHTR systems deal with a high heat flux that requires a very large heat transfer area. For this reason, a compact heat exchanger (CHE) with large surface area density is recently being considered as a potential candidate for the IHX, replacing the classical shell-and-tube design. This type of the compact heat exchanger is widely used in industry, especially for gas-to-gas or gas-to-liquid heat exchange.

The compact heat exchanger is arbitrarily referred to be a heat exchanger (HX) with a surface area density greater than $700 \text{ m}^2/\text{m}^3$ [1]. The compactness is usually achieved by fins and microchannels, and leads to the enormous heat transfer enhancement and size reduction. The surface area density is the total heat transfer area divided by the volume of the heat exchanger. In the case of printed circuit heat exchangers (PCHEs) in the category of the CHEs, the heat transfer surface area density can be as high as $2500 \text{ m}^2/\text{m}^3$ [1,2]. This high compactness implies an appreciable reduction in material cost. A major advantage of this heat exchanger is its ability to operate at high temperatures and under high pressure, while standard shell-and-tube heat exchangers would require very careful material selection and still possibly have difficulties operating at these high pressures and temperatures. The PCHE units manufactured by Heatric™ located in UK allow operation at temperatures and pressures up to 900°C and 50 MPa, respectively. Figure 1 depicts a typical core of the PCHEs. These heat exchangers are constructed from flat alloy plates with

fluid flow passages photochemically etched into them. The plates are then staked and diffusion-bonded together to form strong compact cores.

In this paper, two options of IHX arrangements were investigated: (1) single-stage and (2) two-stage [3]. The single-stage is the most common type of IHX arrangement. In this concept, one big IHX unit is connected to the reactor primary side and the secondary side to exchange heat between them. Sometimes, the size of the IHX is almost the same scale as the reactor vessel or power conversion units (PCUs). In the two-stage concept, we split one large IHX into two smaller parts: low temperature (LT) and high temperature (HT) units. The low temperature unit is designed for the full lifetime, while the high temperature unit is designed to be replaced two to three times within the plant lifetime of 30–60 years.

The rationale for this split option is that it is driven by the cost benefit to replace one small IHX unit instead of one large IHX. The only material approved by the ASME Section III is Alloy 800H, which is only good for the maximum temperature of 760°C . In the high temperature unit that is operated at 900°C , we propose to use Alloy 617, Alloy 230, or ceramics as potential candidates.

In this paper, various options of the IHXs were taken into consideration in terms of arrangement, materials, and types. Finally, the options were compared together in view of thermal design and mechanical stress, and some suitable options for the VHTR applications were recommended.

2 Thermal Design Methods

In this section, thermal designs have been performed for three types of the IHXs: shell-and-tube, PCHE, and helical-coil. The shell-and-tube heat exchanger is the most common heat exchanger type. This heat exchanger is built of round tubes mounted in a cylindrical shell with the tubes parallel to the shell. For design of the shell-and-tube IHX, we followed the general design guidelines provided by the heat transfer and heat exchanger handbooks

Manuscript received December 3, 2008; final manuscript received December 5, 2008; published online December 1, 2009. Review conducted by Dilip R. Ballal.

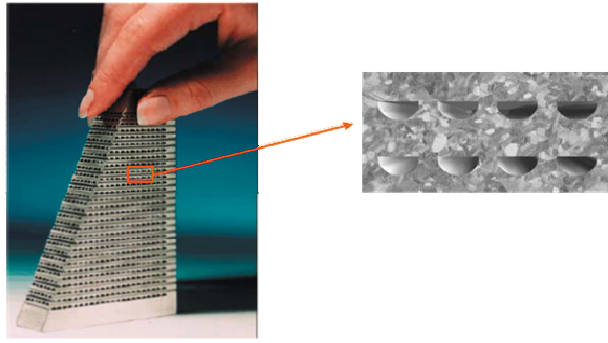


Fig. 1 Cross section through a typical PCHE core (Courtesy of Heatric Ltd.)

[1,4,5]; Table 1 summarizes the major parameters and guidelines for thermal design of the shell-and-tube heat exchangers used in this paper.

The thermal design of the IHX for the PCHE type was followed by the general compact heat exchanger design procedure, equations, and parameters summarized by Hesselgreaves [2]. The printed circuit heat exchangers are generally subjected to very few constraints compared with the shell-and-tube heat exchanger in the thermal design. Fluids may be liquid, gas, or two-phase. Multistream and multipass configurations can be assembled and flow arrangements can be truly countercurrent, cocurrent, cross-flow, or a combination of these, at any required pressure drop.

The following is the summary of the basic PCHE design constraints and guidelines used in this paper. It is based on the information provided by Gezelius [6] and Heatric [7] suggested for PCHE application to the VHTRs:

- (1) semicircular cross section
- (2) width: 0.1–0.2 (2.0 mm shows maximum thermal performance and economic efficiency but for nuclear application, 1.2 mm is suggested)
- (3) depth: 0.5–1.0 mm
- (4) carbon steel is typically not used because of the small channel diameter vulnerable to corrosion and unsuitability for diffusion bonding
- (5) average mass-to-duty ratio: 0.2 tons/MW (13.5 tons/MW in shell-and-tube design)
- (6) no constraint to the pressure drop
- (7) plate thickness: $0.8D$ (D : channel diameter)

- (8) channel pitch: $1.22D$ (D : channel diameter)
- (9) multiported PCHE module size; width: 0.5 m (1.5 m is max), height: 0.6 m, and depth: 0.4–0.6 m
- (10) fatigue can be caused by thermal transient
- (11) only pressure drop restricts the velocity
- (12) minimum life is 20 years

For thermal design of the helical-coil type IHX, we followed the methods similar to those of the shell-and-tube heat exchanger design since there is no well-published guideline for helical-coil heat exchangers. However, the heat transfer and friction loss correlations were corrected to the adequate ones to represent helical-coil flow configurations. Those correlations were obtained from the general heat transfer reference books [1,8].

Alloy 617 was selected to be the reference material based on the material selection studies for the VHTR IHXs performed by Dewson and Li [9]. They selected eight candidate materials based on ASME VIII (Boiler and Pressure Vessel Code) and compared them together. The materials include Alloy 617, Alloy 556, Alloy 800H, Alloy 880 HT, Alloy 330, Alloy 230, Alloy HX, and 253 MA. They extensively compared the mechanical properties, physical properties, and corrosion resistance for the candidate materials, and finally concluded that Alloy 617 and 230 are the most suitable materials for the IHX. The sensitivity of the material types on the heat exchanger thermal design was not considered in this paper, since the types of materials showed negligible effect on the heat exchanger thermal design according to the study of Natesan et al. [10] study. Even for the ceramic materials, only about 13% size reduction was reported [10] in their work.

3 Heat Exchanger Thermal Design

In order to design the IHX, we first need to determine design requirements and conditions. For the VHTR applications, the following are the general requirements for the IHX:

- maximum operating temperature: $\sim 1000^\circ\text{C}$
- pressure: 50–100 bars
- duty: 600 MWt
- effectiveness: 95%
- helium environment
- low pressure drop
- 30–60 year design life

In order to determine the design conditions, we referred to the reference VHTR configuration (integrated with high temperature

Table 1 Guideline for shell-and-tube type heat exchanger design parameter change [1,4]

Adjusted parameters	High pressure drop shell side	High pressure drop tube side	Low coefficient shell side	Low coefficient tube side	Temperature cross	Vibration indication
Baffle type	Double/triple segmental	-	Single segmental	-	-	Double/triple segmental
Shell type	J or X type shell	-	E or F type shell	-	E, F, or G type shell	J or X type shell
Tube pattern	Rotated square or square	-	Triangular	-	-	Rotated square
Tube diameter	Increase to 0.025 m or 0.032 m	Increase to 0.025 m or 0.032 m	Decrease to 0.016 m or 0.013 m	Decrease to 0.016 m or 0.013 m	-	Increase to 0.025 m or 0.032 m
Baffle cut	Use 30–40%	-	Use 15–20%	-	-	-
Tube pitch	Increase to 1.4 or $1.5\times$ tube outside diameter	-	Limit to TEMA std. spacing	-	-	Increase to 1.4 or $1.5\times$ tube outside diameter
Fluid allocation	Switch sides	Switch sides	Switch sides	Switch sides	-	Switch sides
Arrangement	Increase in number of exch. in parallel	Increase in number of exch. in parallel	Increase in number of exch. in series	-	Increase in number of exch. in series	Increase in number of exch. in parallel
No. of tube passes	-	Plain	-	Increase in number of tube passes	Limit to one tube pass	-
Tube type	Plain	-	Ext. enhanced	Internally enhanced	-	-

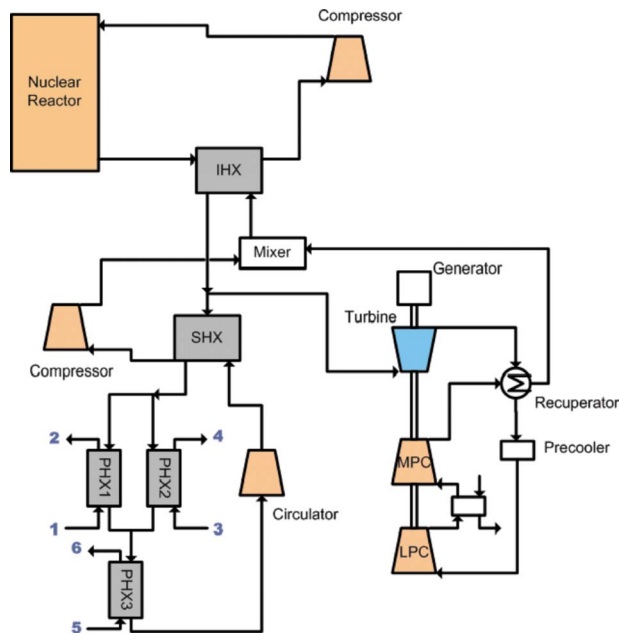


Fig. 2 Reference VHTR system (indirect parallel configuration [11])

Table 2 Design conditions

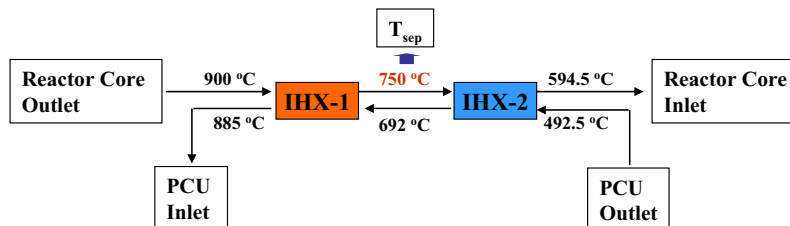
	IHX
Duty (MWt)	612
LMTD (°C) ^a	45.37
Tube side coolant	He
Shell side coolant	He
Tube inlet temperature (°C)	900
Tube outlet temperature (°C)	594.5
Tube side pressure (MPa)	7
Mass flow in tube side (kg/s)	385.3
Shell inlet temperature (°C)	492.5
Shell outlet temperature (°C)	884.8
Shell side pressure (MPa)	7.58
Mass flow in shell side (kg/s)	300

^aLog mean temperature difference.

steam electrolysis) proposed by Oh et al. [11]. Figure 2 shows the schematics of the system configuration. In this system, there are three coolant loops: (1) primary, (2) secondary, and (3) intermediate. The primary loop contains the nuclear reactor, the hot side of the IHX, and a compressor. The secondary loop consists of the cold side of the IHX, the hot side of the secondary heat exchanger (SHX), the PCU, and the connecting piping, which is assumed to be short. The IHX connects the primary and secondary loops. The intermediate heat transport loop connects the secondary coolant system and the high temperature steam electrolysis (HTSE) plant through several process heat exchangers (PHXs). The flow in the secondary coolant system is divided, with most of the flow going toward the PCU and the remainder going through the SHX that directs heat toward the HTSE plant. The flow through the hot side of the SHX is then mixed with the flow from the PCU to feed the cold side of the intermediate heat exchanger (IHX). However, some of the flow is diverted away from the PCU, which acts to decrease the efficiency of the cycle.

Oh et al. [11] estimated the operating conditions and system thermal performance of this system by using HYSYS process analysis code [12]. According to their estimation, the operating conditions of the IHX are as listed in Table 2. Therefore, all the thermal design data provided in this paper are based on the numbers in Table 2. This paper considered only IHX thermal design. The designs of the SHXs and PHXs were not taken into consideration.

In this paper, two different heat exchanger arrangement options were considered: single-stage and two-stage. The single-stage IHX is the most general type of the arrangement options. In this layout, the IHX consists of one heat exchanger unit or separate modules aligned in parallel. Therefore, there is no serial combination of the IHXs in this arrangement. On the other hand, the two-stage IHX has two heat exchangers (i.e., low temperature and high temperature units) arranged and connected in serial. This arrangement was recommended for reducing associated risk and cost of the IHX [3]. As described earlier, the low temperature unit is designed to last the full plant lifetime, while the high temperature unit is designed for replacement within the lifetime. The maximum temperature of the LT unit is proposed to be around 760 °C, allowed in ASME Section III, Subsection NH materials. The design conditions of each unit have been estimated by HYSYS code based on the overall design conditions in Table 2. Figure 3



	IHX-1	IHX-2
Duty	3.001e5 kW	3.111e5 kW
Primary Inlet Temperature	900.0 °C	750.0 °C
Primary Outlet Temperature	750.0 °C	594.5 °C
Secondary Inlet Temperature	692.3 °C	492.5 °C
Secondary Outlet Temperature	885.0 °C	692.3 °C
Primary Inlet Pressure	7.000 MPa	6.975 MPa
Secondary Inlet Pressure	7.560 MPa	7.585 MPa
Flow Rate (Primary)	385.3 kg/s	385.3 kg/s
Flow Rate (Secondary)	300.0 kg/s	300.0 kg/s
LMTD	31.69 °C	77.74 °C

Fig. 3 Design conditions for two-stage IHX

Table 3 Estimated IHX thermal design specifications for PCHE

	Two stage		
	Single stage	HT unit	LT unit
Duty (MWt)	612	300	312
LMTD (°C)	45.57	31.74	77.82
U (W/m ² K)	2313	2410	2735
A (m ²)	5805	3928	1464
Channel diameter (m)	1.20×10^{-03}	1.20×10^{-03}	1.20×10^{-03}
Channel pitch (m)	1.46×10^{-03}	1.46×10^{-03}	1.46×10^{-03}
Plate thickness (m)	9.60×10^{-04}	9.60×10^{-04}	9.60×10^{-04}
Surface area density (m ² /m ³)	2195	2195	2195
Ratio of free flow and frontal area	0.2148	0.2148	0.2148
Stack width (m)	0.6	0.6	0.16
Stack length (m)	0.43	0.29	0.14
Stack height (m)	0.6	0.6	0.6
No. of stacks	34	34	100
Total core volume (m ³)	5.29	3.569	1.334
Pressure drop (primary) (kPa)	66.83	49.2	30.45
Pressure drop (secondary) (kPa)	31.76	26.89	21.74

shows the flow schematics and the calculated design conditions of the two-stage IHX. It is notable that splitting the HXs highly affects log mean temperature differences (LMTD) required for both sides. For example, in the high temperature unit, the LMTD decreases to 31.69°C, while in the low temperature unit, it increases to 77.74°C compared with overall LMTD in the single-stage IHX, 45.37°C.

As mentioned in Sec. 2, three different HX types were considered for thermal design; PCHE, shell-and-tube, and helical-coil. Tables 3–5 summarize the design specifications estimated for each IHX type. Design conditions are based on Table 2 for the single-stage design and Fig. 3 for the two-stage design. For design and rating of the heat exchangers, HYSYS code, and Excel spreadsheet were used. PCHE heat exchanger design has been performed based on the method proposed by Hesselgreaves [2], and shell-and-tube heat exchanger by Kern [5]. Detailed design procedure and equations are not described in this paper.

Table 3 summarizes the estimated IHX design specifications for the PCHE option. The total core volumes were estimated to range from 4.78 m³ to 5.29 m³ for the given overall pressure drop requirement ($dP < 70$ kPa) for both single-stage and two-stage design. Total heat transfer surface areas are about 5800 m². In this

Table 4 Estimated IHX thermal design specifications for shell-and-tube

	Two stage		
	Single stage	HT unit	LT unit
Duty (MWt)	612	300	312
LMTD (°C)	45.54	31.74	77.83
U (W/m ² K)	474	$4.90 \times 10^{+02}$	$4.61 \times 10^{+02}$
A (m ²)	$2.83 \times 10^{+04}$	$1.93 \times 10^{+04}$	$8.69 \times 10^{+03}$
Tube inner diameter (mm)	$1.80 \times 10^{+01}$	$1.80 \times 10^{+01}$	$1.80 \times 10^{+01}$
Tube outer diameter (mm)	$2.00 \times 10^{+01}$	$2.00 \times 10^{+01}$	$2.00 \times 10^{+01}$
Tube pitch (mm)	$3.00 \times 10^{+01}$	$3.00 \times 10^{+01}$	$3.00 \times 10^{+01}$
Pitch-to-diameter ratio	1.5	1.5	1.5
Tube length (m)	15.42	10.16	4.564
Tube thickness (mm)	$1.00 \times 10^{+00}$	$1.00 \times 10^{+00}$	$1.00 \times 10^{+00}$
No. of tubes	2.924	$3.03 \times 10^{+04}$	$3.03 \times 10^{+04}$
Shell diameter (m)	5600	5700	5700
Aspect ratio	0.3631	0.56	1.249
Total core volume (m ³)	379.9	259	116.5
Pressure drop (tube side) (kPa)	73.11	53.88	26.95
Pressure drop (shell side) (kPa)	57.15	41.07	24.84

Table 5 Estimated IHX thermal design specifications for helical-coil

	Two stage		
	Single stage	HT unit	LT unit
Duty (MWt)	612	300	312
LMTD (°C)	45.58	31.74	77.83
U (W/m ² K)	1189	1249	1176
A (m ²)	$1.354 \times 10^{+04}$	7580	3405
Tube inner diameter (mm)	18	18	18
Tube outer diameter (mm)	2	2	2
Tube pitch (mm)	30	30	30
Pitch-to-diameter ratio	1.5	1.5	1.5
No. of tubes	5025	5025	5025
No. of coiled columns	3.33	3.007	1.35
Shell inner diameter (m)	0.49	0.4883	0.49
Shell outer diameter (m)	4.6	4.6	4.6
Shell length (m)	9.86	6.614	2.97
Volume (m ³)	163.8	109.9	49.38
Aspect ratio (D_s/L)	0.47	0.7	1.5
dP (tube) (kPa)	66.57	44.67	17.51
dP (shell) (kPa)	62.51	52.87	18.84

thermal design, the mechanical design factors are not considered yet. Therefore, the numbers specified here could have some discrepancies from those of the final design specifications afterward.

Table 4 summarizes the estimated design specifications of the shell-and-tube heat exchangers. The tube configuration was assumed as a U-tube shape. According to the estimation, the total volumes for this heat exchanger type range between 379 m³ and 483 m³. It is much larger than the volume for the PCHEs in Table 3. There are two reasons for this. First, the surface area densities of the PCHEs are much smaller than the shell-and-tube heat exchangers because of the small channel diameters. Second, the heat transfer coefficients are much larger in the PCHEs. The larger heat transfer coefficient is also contributed from the small channel diameters significantly reducing thermal boundary layers thickness.

Table 5 summarizes the design specifications of the helical-coil heat exchangers. The diameters, thickness, and pitches of the tubes were assumed to be the same as those of the normal shell-and-tube type shown in Table 4. However, in the practical applications, the helical-coil heat exchanger usually requires larger tube size than the normal shell-and-tube heat exchangers because of the manufacturing difficulties. Therefore, the estimated design parameters in Table 5 would be a little bit changed in the real component designs. However, the numbers summarized in Table 5 still show the reasonable initial design approximation. In our estimation, the volumes and heat transfer areas of the helical-coil heat exchanger are ranged between those of the PCHEs and the shell-and-tube heat exchangers. The reason why it shows smaller size and surface area than the normal shell-and-tube heat exchangers is because of the effect of the heat transfer enhancement by the helical-coil.

3.1 Comparisons of Heat Exchanger Types. Based on the design specifications summarized in Tables 3–5, three types of heat exchangers were compared for application of the VHTR IHXs as follows.

- The size of the HX can be represented by its total volume. In the high temperature applications, the smaller volume of the HXs is generally preferred since (1) the high alloy metals or ceramics are very costly, and (2) the smaller HXs can reduce the total system size. The PCHEs provide much smaller sizes than other types.
- In the VHTRs, the smaller heat transfer area is preferred since the large heat transfer areas can permeate tritium into the hydrogen plant, which is a very serious problem in the

Table 6 Heat exchanger combinations in the two-stage IHX

Case	HT unit	LT unit
C1	PCHE	PCHE
C2	PCHE	Shell-and-tube
C3	PCHE	Helical-coil
C4	Shell-and-tube	PCHE
C5	Shell-and-tube	Shell-and-tube
C6	Shell-and-tube	Helical-coil
C7	Helical-coil	PCHE
C8	Helical-coil	Shell-and-tube
C9	Helical-coil	Helical-coil

hydrogen production system integrated with VHTRs. PCHEs provide the lowest surface area.

- Fewer modules are preferred for design simplicity. The helical-coil type heat exchangers look to provide the simplest system configuration.
- In the high temperature applications, thermal stress is an important issue. The high compactness of the PCHE offers great heat performance, but has some potential problems on the thermal stress in the axial direction, where the sharp temperature variations occur.
- The shell-and-tube type HXs are well proven technology in the industry. Especially, the helical-coil HXs have a great operating history and their use in the High Temperature Test Reactor (HTTR) of Japan has shown a long lifetime. Technically, the PCHE types show good potential for high temperature applications, but have not been used in industry yet.
- The shell-and-tube type HXs are a proven technology. Therefore, the problems and limitations of these HXs are very well identified.
- Generally, the shell-and-tube type HXs allow in-service inspections and have well-established maintenance methods. However, regarding the PCHEs, it is difficult to find defects and failures in operation, and it is not easy to maintain compared with the shell-and-tube HXs.

3.2 Heat Exchanger Combinations and Separation Temperature for Two-Stage IHX. In order to investigate the two-stage IHX option, we came up with nine different combinations of three HX types in the HT unit and three HX types in the LT unit in a serial configuration, as shown in Table 6.

Table 7 summarizes the heat exchanger core volumes and heat transfer areas calculated from the total heat duty, the mass flow rate, and the separation temperature between the HT unit and the LT unit. The volumes specified in this table were obtained from Tables 3–5. As shown in Table 7, it is obvious that C1 (PCHE-PCHE) has the smallest total volume and heat transfer area with the highest compactness (surface area density=2195 m²/m³). This high compactness is contributed from the small channel

diameters of the PCHEs by increasing the surface area density and the heat transfer coefficient.

In the two-stage concept, the smaller size of the high temperature unit is highly recommended for cost saving and safety enhancement. Because of the material problems from severe high temperature conditions, careful considerations are necessary for the selection and design of the high temperature related components. If the high temperature unit becomes smaller, less effort and cost will be required for manufacture and maintenance. In addition, the size reduction will also enhance the component safety by reducing the probability of incidents.

In the two-stage design summarized in Table 7, the total duty was split almost in half (300MWt/312MWt). However, the high temperature unit requires more than twice of the volume and the surface area compared with the lower temperature unit. This is because of the reduced LMTD in the high temperature unit (see Fig. 3). The overall LMTD is 45–50°C, but the LMTD is decreased to about 30°C in the high temperature unit, while it is increased to 75–80°C in the low temperature unit. The decrease in the LMTD requires a larger heat transfer surface area for the same duty. It means that the high temperature unit will have a much larger size than the low temperature unit assuming they have the same duties.

The large size of the high temperature unit significantly decreases the benefits of the two-stage concept because it costs more. Therefore, to reduce the size of the high temperature units, the separation temperature between the HT unit and the LT unit has been adjusted to 800°C, which reduces the duty on the HT unit compared with the separation temperature of 750°C. Again, the increase in the separation temperature decreases the LMTD in the high temperature unit, which increases the size of the HX. However, it leads to significant reduction in the duty on the high temperature units, eventually reducing the overall size. The increased separation temperature makes the low temperature unit more vulnerable to the stress.

Table 8 summarizes the volumes and the heat transfer surface areas for the two-stage IHXs at the increased separation temperature of 800°C. The volumes and the surface areas were estimated by the same method as used for the design of the original separation temperature (750°C). When we increase the separation temperature up to 800°C, the volume of the high temperature units becomes about a half of the total volume for the combinations of the same type heat exchangers. Since the high temperature unit is replaceable during the plant lifetime, the PCHEs, which have a smaller size, will be more beneficial, leading to lower cost and easier maintenance compared with the case, $T_{sep}=750^{\circ}\text{C}$.

3.3 Comparisons of IHX Arrangement Options. In this section, the single-stage IHX and the two-stage IHX were compared based on the results summarized in Tables 3–8.

Table 7 Volume (m³) of HXs for two-stage IHX

Case	Volume (m ³)			Heat transfer area (m ²)		
	HT unit	LT unit	Total volume	HT unit	LT unit	Total area
C1	3.5	1.334	4.834	3877	1464	5341
C2	3.5	116.5	120	3877	8690	12,567
C3	3.5	49.38	52.88	3877	3405	7282
C4	259	1.334	260.3	19,300	1464	20,764
C5	259	116.5	375.5	19,300	8690	27,990
C6	259	49.38	308.38	19,300	3405	22,705
C7	109.9	1.334	111.23	7580	1464	9044
C8	109.9	116.5	225	7580	8690	16,270
C9	109.9	49.38	159.28	7580	3405	10,985

Table 8 Volume (m³) of HXs for two-stage IHX ($T_{\text{separation}}=800^{\circ}\text{C}$)

Case	Volume (m ³)			Heat transfer area (m ²)		
	HT unit	LT unit	Total volume	HT unit	LT unit	Total area
C1	2.549	2.019	4.568	2792	2216	5008
C2	2.549	173.6	176.15	2792	13,000	15,792
C3	2.549	73.63	76.179	2792	5077	7869
C4	201.9	2.019	203.92	19,300	2216	17,216
C5	201.9	173.6	375.5	19,300	13,000	28,000
C6	201.9	73.63	275.53	19,300	5077	24,377
C7	85.58	2.019	87.599	5901	2216	8117
C8	85.58	173.6	259.18	5901	13,000	18,901
C9	85.58	73.63	159.21	5901	5077	10,978

3.3.1 *Total HX Core Volume/Heat Transfer Area.* Theoretically the core volume and heat transfer area by the splitting option should be the same size compared with the single IHX. However, practically, the two-stage IHX will require more spaces for manifold of flow distribution, which may not be a major cost penalty.

3.3.2 *Size of High Temperature Section.* The smaller size of the high temperature unit is preferred because it can reduce the maintenance cost and enhance component safety. The two-stage IHX with an 800°C separation temperature has the lowest high temperature section size among the designs performed in this paper. The single-stage IHX provided the largest high temperature section.

3.3.3 *System Complexity.* The number of modules is related to the system complexity. Therefore, fewer modules are preferred for simplicity. The single-stage IHX has the simplest design because of less number of modules.

According to the above comparisons, conceptually, the two-stage PCHEs with a separation temperature of 800°C are expected to offer more advantages in system cost, safety, and maintenance compared with the single-stage IHXs. For example, material cost can be saved in the two-stage IHXs since some cheap commercial materials, such as stainless steel or Alloy 800H, can replace high alloy materials in the low temperature sections. In addition, it provides easier management and controllability of risks by concentrating safety issues on the high temperature sections. The thermal stress can also be relieved in the two-stage IHXs by reducing temperature differences in the two components. On the other hands, the single-stage PCHE provides quite small size and simplicity.

In the thermal design, the two-stage heat exchangers using the PCHEs with the separation temperature of 800°C are recommended for the IHX. Using the standard shell-and-tube heat exchanger is not recommended for the VHTR applications here. However, to avoid some risks about the lack of experience using the PCHEs as the IHXs, the single-stage helical-coil heat exchangers can be another option, since it has been proven in the operation of the HTTR.

4 Heat Exchanger Simple Stress Analysis

A simplified stress analysis was performed in order to estimate the mechanical performance of the IHX options. We estimated the required thickness in which the maximum stress was less than or equal to an assumed allowable value. The use of consistent stresses allowed identification of limiting components and a fair comparison between different configurations. In this paper, besides the IHX, the SHX was also taken into consideration for stress analysis because the SHX suffers from high pressure differences, as well as high temperature.

Since the IHX and the SHX are operated in the high temperature environment, creep deformation is very important, and Subsection NH of ASME Code, Section III is applicable. However,

there are several problems with directly applying ASME Section III code rules at this time. The use of the primary candidate, structural Alloy 617, is currently not approved in Subsection NH. So, we used the draft code case for the designs using Alloy 617 to obtain the allowable stress for our stress analyses. The primary stress limits of the Alloy 617 draft code case were well summarized by Natesan et al. [10].

Figure 4 shows allowable stress (S_{mt}) versus lifetime for Alloy 617. These data were obtained from Alloy 617 draft code case [10]. The plot shows the lifetime data as a function of allowable stress for three different temperatures: 750°C, 800°C, and 900°C. This graph can be interpreted to be the maximum operating lifetime at the given stress load. As the stress increases, the lifetime decreases exponentially. Since the available lifetime data was limited to 1×10^5 h (about 10 years), the values beyond this lifetime were predicted by a simple linear extrapolation method. For more accurate analyses, available experimental data or proven extrapolation methods are recommended to be measured or developed afterward.

Table 9 summarizes the assumed allowable stresses for Alloy 617 predicted by interpolation and extrapolation of Fig. 4. The stress analyses of this paper have been carried out based on the allowable stresses assumed in this table.

The required thickness for the IHX channel walls were estimated based on the simple stress and failure theory. For thick-walled cylinders, the tangential stress, σ , is calculated as in Ref. [13]

$$\sigma = \frac{P_i[(r_o/r_i)^2 + 1] - P_o[(r_o/r_i)^2 + (r_o/r)^2]}{(r_o/r_i)^2 - 1} \quad (1)$$

where r is the radius, P is the pressure, and the subscripts i and o refer to the inner and outer surface, respectively. The stress is

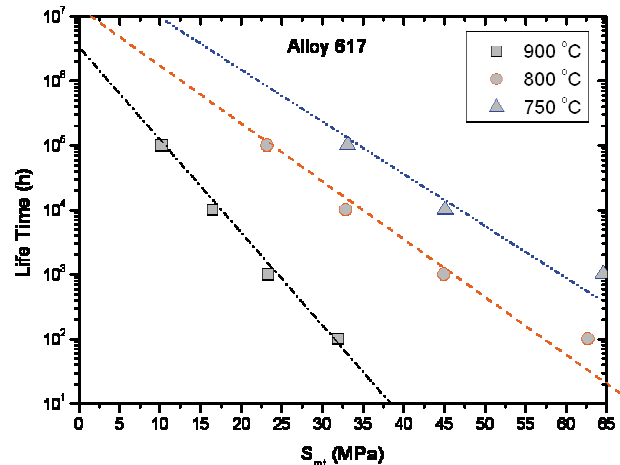


Fig. 4 S_{mt} (MPa) versus lifetime (h) for Alloy 617

Table 9 Assumed allowable stresses for Alloy 617

Temperature (°C)	Allowable stress (MPa)		
	10 ⁵ h	5 × 10 ⁵ h	10 ⁶ h
750	33.1	25	22.5
800	23.2	15	12.3
900	10.2	5	3.5

negative if the external pressure exceeds the internal pressure, but the maximum magnitude always occurs at the inner surface. The radius ratio that causes the maximum stress to be less than or equal to the allowable stress, σ_D , can be calculated from Eq. (1). For cases where the internal pressure exceeds the external pressure, the limiting ratio is

$$\frac{r_o}{r_i} \geq \sqrt{\frac{\sigma_D + P_i}{\sigma_D + 2P_0 - P_i}} \quad (2)$$

For cases where the external pressure exceeds the internal pressure, the maximum absolute value of the stress will be less than or equal to the allowable stress when the radius ratio is

$$\frac{r_o}{r_i} \geq \sqrt{\frac{\sigma_D - P_i}{\sigma_D - 2P_0 - P_i}} \quad (3)$$

A stress analysis was also performed for the IHX assuming that it is a compact heat exchanger of the type designed by Heatic [7]. The design of the heat exchanger channels is defined by the channel diameter, d , pitch, p , and plate thickness, t_p . Following the method used by Dostal et al. [14], the minimum wall thickness between channels, t_f , can be approximated as

$$t_f \geq \frac{P}{\frac{\sigma_D}{\Delta P} + 1} \quad (4)$$

where σ_D is the allowable stress and ΔP is the differential pressure between the hot and cold streams. Expressing Eq. (4) in terms of pitch-to-diameter ratio yields

$$\frac{p}{d} \geq 1 + \frac{\Delta P}{\sigma_D} \quad (5)$$

The required plate thickness can also be calculated based on the method of Dostal et al. [14]. The plate is assumed to be a thick-walled cylinder, with an inner radius of $d/2$ and an outer radius of t_p .

Engineering problems concerned with the design and development of structural or machine parts generally involve biaxial (occasionally triaxial) stresses. However, available strength data usually pertain to uniaxial stress, and often only to uniaxial tension. To resolve this problem, a failure theory is used in the engineering practice. The failure theories are generally based on the assumption that tensile yielding occurs as a result of exceeding the capacity of the materials in one or more respects, such as the following:

- (a) capacity to withstand normal stress (maximum normal stress theory)
- (b) capacity to withstand shear stress (maximum shear stress theory)
- (c) capacity to withstand normal strain (maximum normal strain theory)
- (d) capacity to withstand shear strain (maximum shear strain theory)
- (e) capacity to absorb strain energy (total strain energy theory)
- (f) capacity to absorb distortion energy (maximum distortion energy theory)

Hence, in the simple classical theories of failure, it is assumed that the same amount of whatever caused the selected tensile specimen to fail will also cause any part made of the materials to fail regardless of the state of stress involved. The model details are well described in Ref. [15]. Figure 5 shows the reported biaxial strength data for various ductile and brittle materials provided by Collins [15]. This figure shows that the maximum normal stress theory is appropriate for brittle behavior; however, the distortion energy or maximum shear stress theory is appropriate for ductile failure such as metal. In this paper, for conservative failure estimation, the maximum shear stress theory was used.

The lifetimes of the reference HXs designed in the previous section were estimated by the failure theory, and the results were summarized in Table 10. The lifetime was estimated by the maximum shear stress theory, which is generally the most conservative. The maximum pressure differences were set to be 1.0 MPa between the hot and cold channels. The maximum shear stress theory was applied to obtain the most conservative estimation of the lifetime. Theoretically, the PCHE has larger t/r_i values than the tubular type heat exchangers because it shows the longer lifetime. However, as the thickness-to-inner radius ratio in the tubular type heat exchanger is increased, the lifetime of the tubular heat exchanger gets closer to the PCHEs. However, the common t/r_i

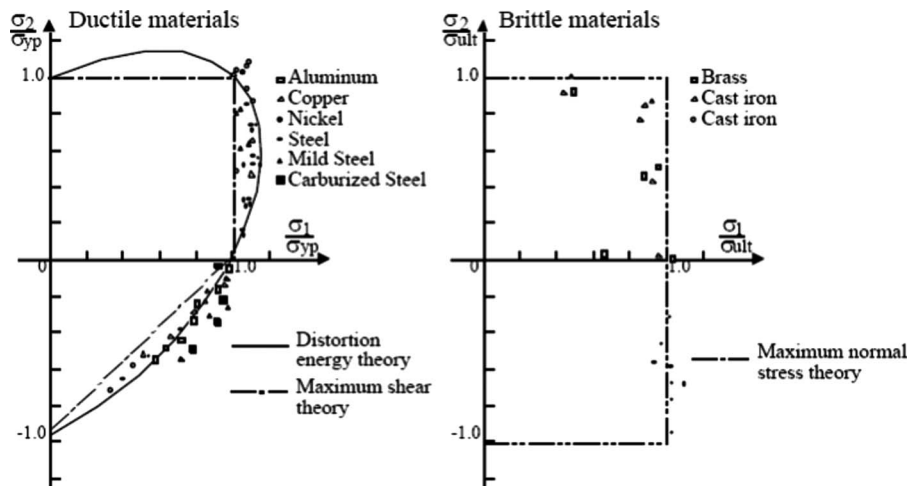


Fig. 5 Comparisons of biaxial strength data with failure theories [15]

Table 10 Estimated lifetime of reference IHXs

Temperature (°C)	Lifetime (h)	
	IHX ($\Delta P_{max}=1$ MPa)	
	PCHE	Tubular Type
750	1.34×10^7	7.42×10^6
800	2.62×10^6	1.32×10^6
900	3.12×10^5	1.16×10^5

values in commercial tubing range from 0.1 to 0.5, and most of them are within 0.3; the designed tube thickness should not be larger than this.

According to Table 10, the PCHE can be operated as an IHX in 900°C for about 35 years with the reference design. However, the tubular heat exchangers (shell-and-tube or helical-coil) can be operated only for about 13 years (1.16×10^5 h). However, if the required pressure drop between hot and cold channels is increased to 5 MPa (this is usually required for SHX application.), our estimation shows that the lifetime of the heat exchangers are significantly dropped to 1.44×10^4 h and 149 h for the PCHE and tubular types, respectively (although the detail results were not shown in Table 10). However, for 800°C, both PCHE and tubular type heat exchangers can be operated for the whole plant lifetime. It shows that the separation temperature, 800°C, is acceptable for the low temperature units in the whole plant lifetime.

Figure 6 shows the maximum effective stress of the PCHEs as a function of plate thickness-to-diameter ratio (t_p/d) for three different applications assuming the operating temperature is 900°C: (1) IHX ($\Delta P_{max}=1$ MPa), (2) SHX ($\Delta P_{max}=2$ MPa), and (3) SHX ($\Delta P_{max}=5$ MPa). The maximum effective stress was calculated by the maximum shear stress theory. The dashed lines in this graph show the rupture stress for the given lifetime (see Fig. 4). Therefore, for reliable operation, the effective maximum stress should always be designed to be lower than the rupture stress. According to these graphs, as the plate thickness-to-diameter ratio increases, the effective stress is exponentially decreased. In addition, lower pressure differences between the hot and the cold channels led to much less effective stress, meaning that the thicker the plate, the stronger the mechanical integrity gets.

Figure 7 shows the maximum effective stress for the tubular heat exchangers as a function of tube thickness-to-inner radius ratio, which is basically the same trend as the PCHEs. The required thicknesses for the IHXs and the SHXs were estimated based on these two figures (Figs. 6 and 7).

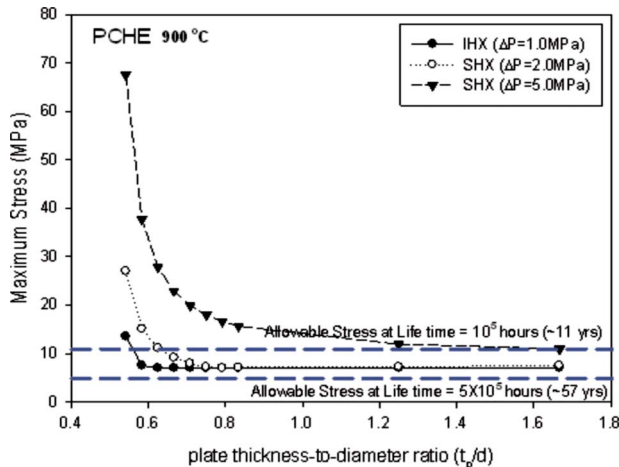


Fig. 6 Maximum stress as a function of plate thickness-to-diameter ratio (PCHE, 900°C)

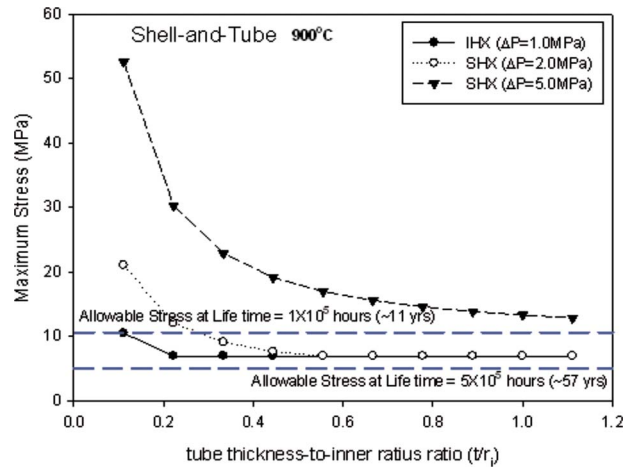


Fig. 7 Maximum stress as a function of tube thickness-to-radius ratio (shell-and-tube, 900°C)

Table 11 summarizes the required plate or tube thickness normalized by the channel size. This table shows that permanent operation (>50 years) is possible for less than 800°C for both PCHEs and tubular type under the appropriate plate and tube thickness. However, for 900°C, the full time operation seems not to be possible without replacement. For the reference thermal design, the PCHE requires to be replaced twice, and the tubular type requires to be replaced for about five times for the 900°C operation.

5 Summary and Conclusions

As part of the system integration of the VHTR and the hydrogen production plant, design and configuration of the intermediate heat exchanger have been investigated. This paper included analyses of single-stage versus two-stage heat exchanger design configurations and simple stress analyses of the printed circuit heat exchangers, the helical-coil heat exchangers, and the shell-and-tube heat exchangers.

Comparisons of the heat exchanger types showed that the PCHE type offers obvious benefits for its thermal performance, size, and compactness. However, it was estimated to have some difficulties regarding inspections and maintenances. Lack of experience for the long lifetime is another problem of the PCHEs. On the other hand, the tubular type HX is ready-to-use technology. Helical-coil HXs showed successful operating records in HTTR application. However they require too large size for high efficient operations and show poor thermal performance.

According to the overall comparisons, the two-stage IHX looks better than the single-stage IHX in cost, safety, and reliability. Conceptually, the two-stage heat exchangers using the PCHEs (for both the HT and LT units) with the separation temperature of 800°C are recommended to be the best option for the IHXs. How-

Table 11 Required normalized plate or tube thickness for IHX

Temperature (°C)	PCHE (t_p/d)		Shell-and-tube (t/r_i)	
	10^5 h	5×10^5 h	10^5 h	5×10^5 h
750	0.51	0.51	0.006	0.002
800	0.51	0.53	0.02	0.025
900	0.56	∞	0.1	∞

t_p is the plate thickness of the PCHE.

d is the channel diameter of the PCHE.

t is the tube thickness of the tubular heat exchanger.

r_i is the tube inner radius of the tubular heat exchanger.

ever, to avoid some risks from the lacks of experience in PCHE as an IHX, the helical-coil single-stage helical-coil heat exchanger can be another option.

By simple stress analyses, the lifetime and t/r_i of the reference HXs has been estimated for the reference thermal designs. As a result, full time operation (>50 years) is possible for higher than 800°C for both PCHEs and tubular type HXs with proper selection of the plate or tube thicknesses. However, for 900°C operation, both PCHEs and tubular type were estimated to require a number of replacements during the lifetime. For the reference thermal design, the PCHE requires replacement for about two times, and the tubular type requires replacement for about five times. However, the replacement time can be changed and optimized by different selection of channel size and wall thickness. We are planning to investigate further using other types of heat exchanger and different heat exchanger materials, including ceramics, in the near future.

Acknowledgment

This work was supported through the U.S. Department of Energy's NGNP-Engineering Program under DOE Idaho Operations Office Contract No. DE-AC07-99ID13727.

References

- [1] Kakac, S., and Liu, H., 2002, *Heat Exchangers: Selection, Rating, and Thermal Design*, CRC, Boca Raton, FL.
- [2] Hesselgreaves, J. E., 2001, *Compact Heat Exchangers*, Pergamon, New York.
- [3] Patterson, M., 2008, personal communication, NGNP Internal Summary, Conceptual Design Studies for NGNP, Project Status Meeting, Feb. 7.
- [4] Hewitt, G. F., 2002, *Heat Exchanger Design Handbook*, Begell House, New York.
- [5] Kern, D. Q., 1991, *Process Heat Transfer*, McGraw-Hill, New York, p. 143.
- [6] Gezelius, K., 2004, "Design of Compact Intermediate Heat Exchangers for Gas Cooled Fast Reactors," MS thesis, MIT, Cambridge, MA.
- [7] Dewson, S. J., and Grady, C., 2003, Heatric™ Workshop at MIT, Cambridge, MA, Oct. 2.
- [8] Bejan, A., and Kraus, D., 2003, *Heat Transfer Hand Book*, Wiley, New York.
- [9] Dewson, S., and Li, X., 2005, "Selection Criteria for the High Temperature Reactor Intermediate Heat Exchanger," *Proceedings of ICAPP' 05*, Seoul, South Korea, May, pp. 15–19.
- [10] Natesan, K., Moisseytsev, A., Majumdar, S., and Shankar, P., 2006, "Preliminary Issues Associated With the Next Generation Nuclear Plant Intermediate Heat Exchanger Design," Report No. ANL/EXT-06/46.
- [11] Oh, C. H., Kim, E. S., Sherman, S. R., Vilim, R., Lee, Y. J. and Lee, W. J., 2007, "HyPEP FY-07 Annual Report: A Hydrogen Production Plant Efficiency Calculation Program," Report No. INL/EXT-07-13078.
- [12] Aspen Technology, 2006, HYSYS Process Version 3.2.
- [13] Crandal, S. H., Dahl, N. C., and Lardner, T. J., 1972, *An Introduction to the Mechanics of Solids*, 2nd ed., McGraw-Hill, New York.
- [14] Dostal, V., Driscoll, M. J., and Hejzlar, P., 2004, "A Supercritical Carbon Dioxide Cycle for Next Generation Nuclear Reactors," Report No. MIT-ANP-TR-100.
- [15] Collins, J. A., 1981, *Failure of Materials in Mechanical Design—Analysis, Prediction, Prevention*, Wiley, New York.