Transactions of the ASME

Journal of Engineering for Gas Turbines and Power

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 109 • NUMBER 2 • APRIL 1987

TECHNICAL PAPERS

ANNOUNCEMENTS

Technical Editor
ARTHUR J. WENNERSTROM
Senior Associate Editor
G. K. SEROVY
Associate Editors
Advanced Energy Systems
S. I. FREEDMAN
Environmental Control
H. E. HESKETH
Fuels and Combustion Technologies
R. E. BARRETT
Gas Turbine
S. KUO
Internal Combustion Engine
K. J. SPRINGER
Nuclear Engineering
S. M. CHO
Power
R. W. PORTER

BOARD ON COMMUNICATIONS
Chairman and Vice-President
K. N. REID, JR.

Members-at-Large
W. BEGELL
J. T. COKONIS
M. FRANKE
W. G. GOTTENBERG
M. KUTZ
F. LANDIS
J. R. LLOYD
T. C. MIN
R. E. NICKELL
R. E. REDER
F. W. SCHMIDT

President, **N. D. FITZROY**
Executive Director,
PAUL ALLMENDINGER
Treasurer, **ROBERT A. BENNETT**

PUBLISHING STAFF
Mng. Dir., Publ., **J. J. FREY**
Dep. Mng. Dir., Pub.,
JOS. SANSONE
Managing Editor,
CORNELIA MONAHAN
Sr. Production Editor,
VALERIE WINTERS
Editorial Prod. Asst.
MARISOL ANDINO

Transactions of the ASME, Journal of Engineering for Gas Turbines and Power (ISSN 0022-0825) is published quarterly (Jan., Apr., July, Oct.) for $85 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017. Second class postage paid at New York, NY and additional mailing offices. POSTMASTER. Send address change to The 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.
PRICES: To members, $24.00, annually; to nonmembers, $85.00.
Add $6.00 for postage to countries outside the United States and Canada.
STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or ... printed in its publications (B 7.1, para. 3).
COPYRIGHT © **1987** by the American Society of Mechanical Engineers. Reprints from this publication may be made on condition that full credit be given the TRANSACTIONS OF THE ASME – JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, and the author, and date of publication be stated.
INDEXED by Engineering Information

Downloaded 03 Jun 2010 to 171.66.16.159. Redistribution subject to ASME license or copyright; see http://www.asme.org/terms/Terms_Use.cfm

M. R. Maltby
Performance Engineer,
Rolls-Royce Limited,
Leavesden, Watford,
Herts, England

Acceleration Performance of Helicopter Engines

The desirability of good acceleration behavior from helicopter gas turbine engines is discussed in the context of the increasing demand for helicopter maneuverability and safety. Features which result in improved engine handling qualities are discussed and reference is made to some analytical techniques which may be employed to assist in obtaining the best results.

Introduction

The importance of good acceleration and deceleration behavior, or handing, from aircraft gas turbine engines is well known throughout the industry and particular attention must be given to this subject at all phases of an engine's development, from initial concept through to the effect of deterioration during service.

Nowhere is the transient response of an engine brought more sharply into focus than in helicopter power plants. Large and rapid changes in power are demanded by the rotor system in many flight modes, both normal and emergency. These include rapid descent to the hover, quick stop maneuvers, "nap-of-earth" (NOE) flying and, in the case of multi-engined helicopters, recovery from a single-engine failure. In all but the last of these cases a power change demand is placed on the engine via a pilot input to the rotor system which affects its blade angle of attack and hence lift and drag. In a conventional rotor droop system a change in drag alters the rotor's rotational speed which signals the engine via the control system to change power. Poor engine response to this power demand will lead to excessive rotor speed variation which ultimately limits the helicopter's maneuverability. In the last case poor engine response will lead to excessive rotor droop, increased height loss and thus can endanger the aircraft itself.

Currently engine acceleration times are the biggest lag item in the response of the helicopter. The aim of the designer of a modern helicopter engine must therefore be to provide handling qualities which shift the limitation of helicopter maneuverability from the engine/control system combination onto the helicopter rotor system itself. This improvement must be achieved against a background of lighter, low-inertia, highly loaded main rotors (which respond in speed more quickly) and the requirements to obtain the best possible steady-state performance from the engine at the lowest weight and cost.

Increasing demands for helicopter maneuverability and safety are being placed on manufacturers by the customer and authorities alike. In order to provide the best possible solution some of the traditional measures of engine handling need to be reassessed in light of the helicopter's demands and some sophisticated analysis techniques are required.

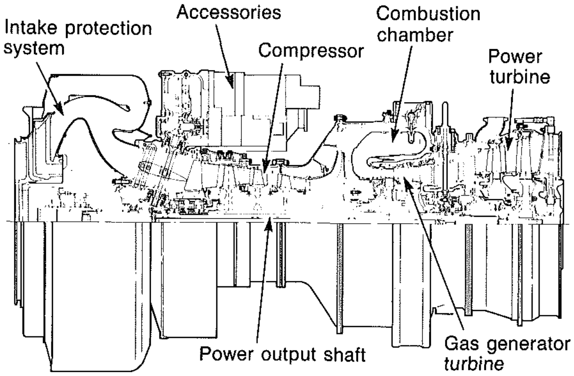


Fig. 1 Typical modern helicopter engine

The Helicopter Engine

A typical layout of a modern helicopter gas turbine is given in Fig. 1. In essence it consists of a one or two-shaft gas generator (comprising a compressor, combustion chamber, and turbine) and a free power turbine which may either provide a front drive, usually via a concentric shaft through the gas generator shaft, or a rear drive. The power turbine is controlled to a desired speed during all flight modes via the engine control system whereas the gas generator varies in speed thus altering the air mass flow through the engine and hence the power. The free power turbine configuration is chosen for helicopter engines because it isolates the gas generator speed (and hence the ability to produce power) from the rotor droop brought about by the change in power demand. Fixed shaft engines can be very dangerous in this respect: Excessive droop produces low lift for a given power and low power simultaneously, an unstable combination. The ability of the control system to maintain a constant rotor speed during load changes depends upon the rate at which the engine will respond to the load change.

What Is Good Handling?

From the point of view of the helicopter designer, ideal handling from an engine would enable the rotor to remain at a

---

Contributed by the Gas Turbine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the 31st International Gas Turbine Conference and Exhibit, Dusseldorf, Federal Republic of Germany, June 8-12, 1986. Manuscript received at ASME Headquarters January 24, 1986. Paper No. 86-GT-121.

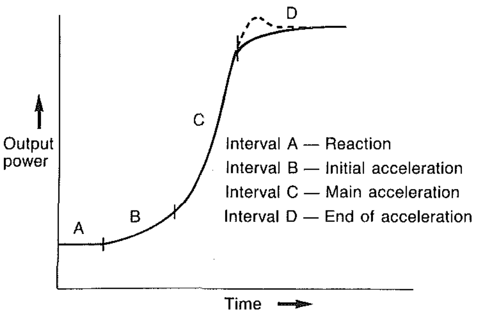


Fig. 2 Typical power response to a load demand

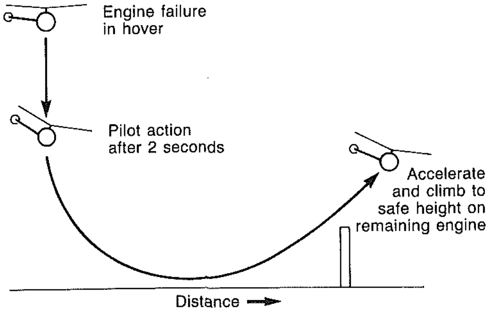


Fig. 3 Flyaway maneuver

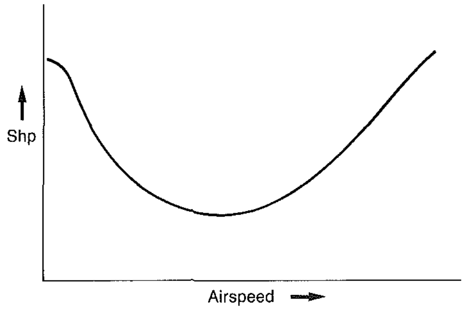


Fig. 4 Typical helicopter power requirement

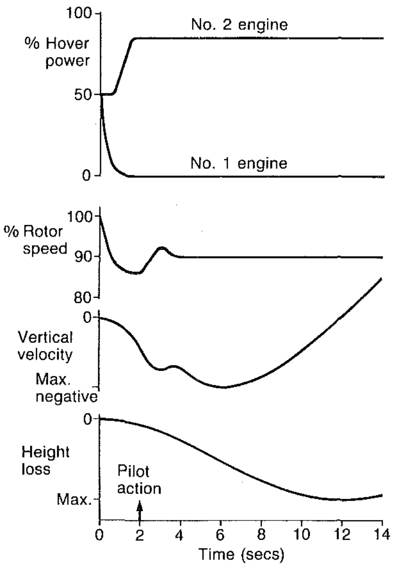


Fig. 5 Flyaway prediction: typical input and output

chosen speed irrespective of the transient loads. In other words, the power demanded by the rotor system would always be matched by the power available from the engines. It therefore follows that in practical terms the only measure of handling ability that is of use is that of rate of change of power against time. This rate of change will typically vary during an acceleration and it has been shown at Rolls-Royce that within a given overall acceleration time, the variation of rate can very significantly affect helicopter response.

A typical power versus time plot for an engine response to a load demand might be as shown in Fig. 2. The curve is characterized by a pause (interval A) during which the control system senses the load change and signals a change to the fuel flow; a relatively slow buildup of power (interval B) during which the engine moves through conditions unfavorable to rapid acceleration; a fast buildup of power (interval C) during which conditions are favorable; and arrival at the required power (interval D) which may exhibit an overshoot or an asymptotic approach as illustrated.

The effect of the magnitude of these parameters on helicopter performance during a "flyaway" maneuver has been studied at Rolls-Royce using a simple computer dynamic helicopter model. The flyaway maneuver is carried out when a multi-engined helicopter suffers a single engine failure in the hover and immediate landing is impracticable. The maneuver, illustrated in Fig. 3, takes advantage of the relatively low power required to fly at moderate forward speed in comparison with the power to hover (Fig. 4). Assuming that the power available from the remaining engine is insufficient to hover, the pilot gains forward speed by diving the helicopter, effectively trading potential energy (altitude) for kinetic energy (forward speed), while using collective pitch to reduce the rate of height loss as much as possible without reducing the rotor speed unacceptably. On reaching a forward speed at which the power available exceeds the power required the flight may be continued. The computer model simulates this maneuver and the response of the remaining engine may be varied. Typical input to and output from the model are illustrated in Fig. 5, and a diagram of the logic is shown in Fig. 6. In the example shown in Fig. 5 it was assumed that the surviving engine reacted half a second after the failure of the first engine and reached full power after a further second. The pilot was assumed to have taken action after two seconds and to have accelerated to 50 kts while endeavoring to maintain the rotor speed at 90 percent of its normal value. The effectiveness of the maneuver is determined by the overall height loss—a very significant factor if the aircraft is close to the ground or the sea.

The investigation revealed the importance of preventing rotor speed decay by applying power as quickly as possible. That is to say the time before the onset of a rapid power change illustrated in Fig. 2 as interval C must be reduced to the minimum possible. A high acceleration rate in interval C is reduced in effectiveness if it is preceded by a significant period during which the power available remains essentially static.

The effects of engine reaction time and acceleration rate for various power recoveries are illustrated in Figs. 7 and 8. As would be expected the strongest influence on the height loss is the percentage of the original hover power that the second engine is able to achieve—the higher the "contingency" power the better—but if it arrives too late the helicopter may have already hit the ground.

Factors Affecting Handling Qualities of Helicopter Engines: Definitions

Handling Qualities. The ability of an engine to change from one power condition to another rapidly and without surge.

Compressor Characteristic. A compressor characteristic relates the pressure ratio $P2/P1$, nondimensional inlet air mass flow $M\sqrt{T1}/P1$, and efficiency of a compressor for given lines of constant nondimensional speed $NG/\sqrt{T1}$ (Fig. 9), where

M is the air mass flow
$T1$ is the temperature of the air at entry to the compressor

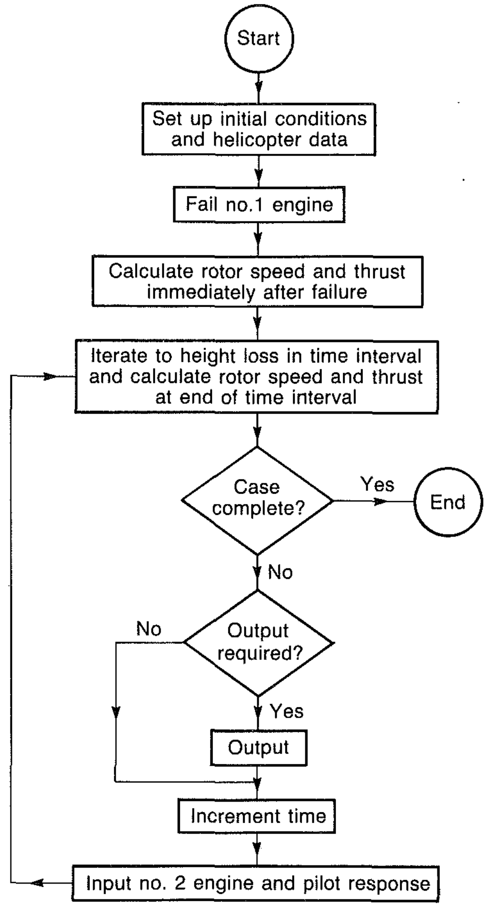


Fig. 6 Flyaway program flow diagram

*P*1 is the pressure of the air at entry to the compressor
*P*2 is the pressure of the air at exit from the compressor
NG is the compressor speed

Steady-State Running Line. The steady-state running line is the locus of points on a compressor characteristic on which the engine will operate at a steady condition (Fig. 9).

Transient Running Line. The transient running line is the path followed on the compressor characteristic during a change from one condition to another. The variation from the steady-state line depends upon the rate of that change (Fig. 9).

Surge (or Stall). Surge or stall is the breakdown of normal flow through the engine. It can lead to overtemperaturing of the turbines and structural damage because the pressure falls with reduced flow and so the flow status becomes unstable.

Surge Line. The surge line is the locus of points on the compressor characteristic at or above which the engine will surge (Fig. 9).

Surge Margin. The surge margin is a measure of the separation of the surge line from the running line. It may be defined as the ratio between the pressure ratios of the surge line and the running line of a given inlet nondimensional mass flow or speed, or a ratio between the outlet nondimensional mass flows at the surge line and running line at a given speed.

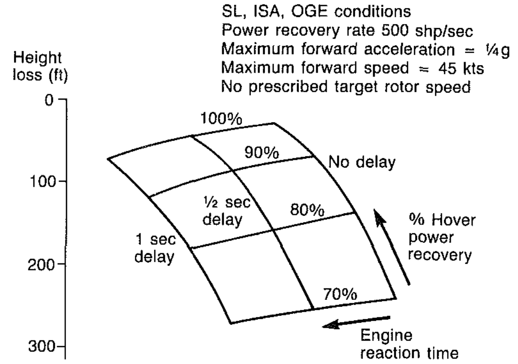


Fig. 7 Typical 12,000 lb helicopter flyaway

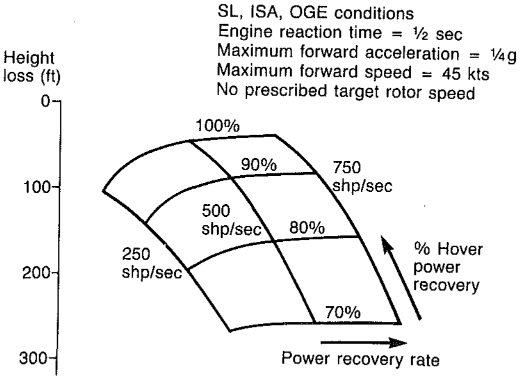


Fig. 8 Typical 12,000 lb helicopter flyaway

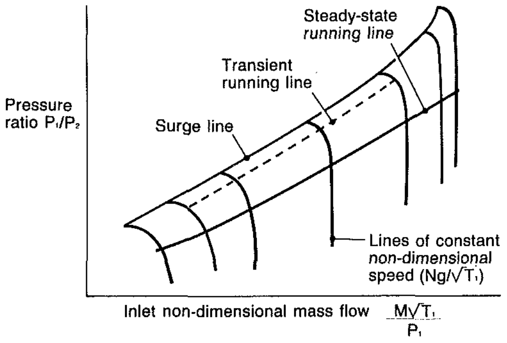


Fig. 9 Compressor characteristic

The output power from an engine depends at its most basic level on the airflow through the power turbine. The temperature and pressure of this air are obviously very important, but to obtain a significant and sustained change in power it is necessary to increase the flow through the engine and therefore through the compressor. The relationship between flow and power is direct: The temperatures and pressures remaining equal, a 1 percent change in flow will give a 1 percent change in power. Thus the most important single factor in determining the ability to change power quickly is the ability to change flow quickly. Since flow depends on engine speed the rate of change of power is primarily governed by the rate of change of flow with speed and the rate of change of speed with time. The former is governed by the flow/speed characteristic of the compressor and the latter by the spool inertia, the surge margin, and the temperature capability of the turbine. When an engine is accelerated the fuel flow and thus the turbine entry temperature is increased, thus providing an increase in turbine torque resulting in an increase in spool speed and hence air flow. However, due to the inertia of the spool, instantaneously the speed will not change. Since the turbine nozzle is choked and its capacity measured as $M\sqrt{T}/P$ is therefore fixed, the rise in turbine entry temperature (TET) must be matched by either a fall in mass flow or a rise in pressure ratio. This results in a movement on the compressor characteristic toward surge along a speed line (Fig. 10). Thus the surge margin limits the ability of the engine to accept a rapid increase in TET, assuming that the TET limit is not reached first. The inertia of the spool then determines the acceleration. Clearly then, for a given surge margin, the inertia

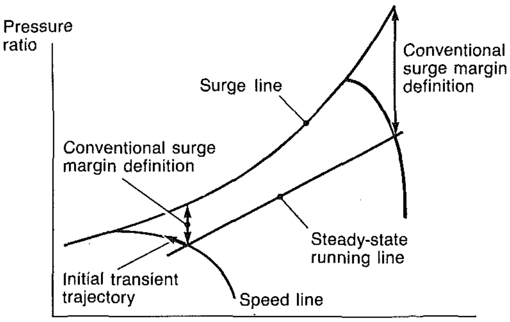


Fig. 10 Surge margin—conventional definition

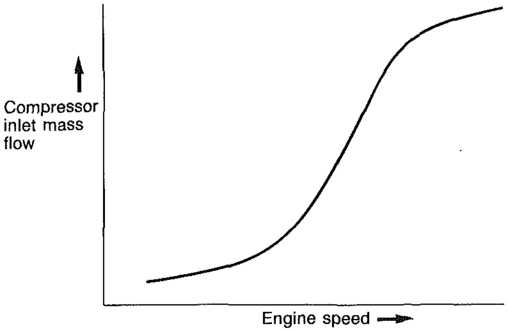


Fig. 11 Typical compressor flow/speed relationship

of the gas generator spool determines the rate of change of speed—but does this mean that low inertia should be pursued at all costs? In practice an exercise aimed specifically at reducing the shaft inertia of an engine for which overall weight will have been of prime importance might yield a reduction of, say, 5 percent. This figure would be achieved by magnifying the problems associated with stress and life leading to a reduction in acceptable speeds and temperatures or, at the very least, increasing the problems associated with developing the engine to higher powers. Against this background an acceleration which would have taken 5 s to complete would not take 4.75 s—hardly a great prize. Alternatively, should the aim be the greatest possible surge margin? The answer to this question has to be somewhat more qualified. It is true to say that an engine with a poor surge margin will not accelerate well, but the determination of what is an acceptable surge margin will depend upon other aspects of the compressor characteristic, most notably the flow/speed relationship.

Improvements to the surge margin can be made in several ways. The steady-state running line can be lowered by a suitable choice of turbine nozzle capacities but this reduces the thermodynamic efficiency of the cycle and may cause the compressor to operate at a less efficient part of the characteristic. Poor surge margin at low speed may be overcome by use of a blow-off valve which bleeds air from the compressor system to lower the working line, and the surge line position can be greatly improved by the use of casing treatments. Although both of these methods are usually effective they inevitably result in greater complexity, cost, or weight, and may serve only to provide some alleviation to a fundamentally poor position caused by a poor rate of change of flow with speed. In addition, air that is dumped through a bleed valve is not available to the power turbine to generate power and so the balance of reduced power against improved spool acceleration is a delicate one. It is by addressing the aspect of flow/speed that satisfactory handling characteristics can be assured from the outset.

It is obvious that the greater the gas generator speed range that an engine has to move through to get from one power level to another, the longer the acceleration will take. To a first order an engine accelerating from a zero power condition (no load) at 80 percent of the design speed to full power (at 100 percent) has effectively twice the inertia—in terms of power versus time—of an engine accelerating from no load at 90 percent speed. This argument may be developed further by considering the compressor flow/speed relationship—in effect the distance between the speed lines on the characteristic. Most compressors, particularly those with variable geometry, exhibit a marked change in the slope of the flow/speed line near the lower end of the operating range (see Fig. 11). Again to a first order it can be said that a doubling of the slope of the flow/speed line is equivalent for power versus time of a halving of the inertia. Thus avoiding operating in areas where the change of flow with speed is poor is of prime importance. The improvement in engine handling which results can be very great and for a desired power versus time relationship the surge margin required can actually be reduced.

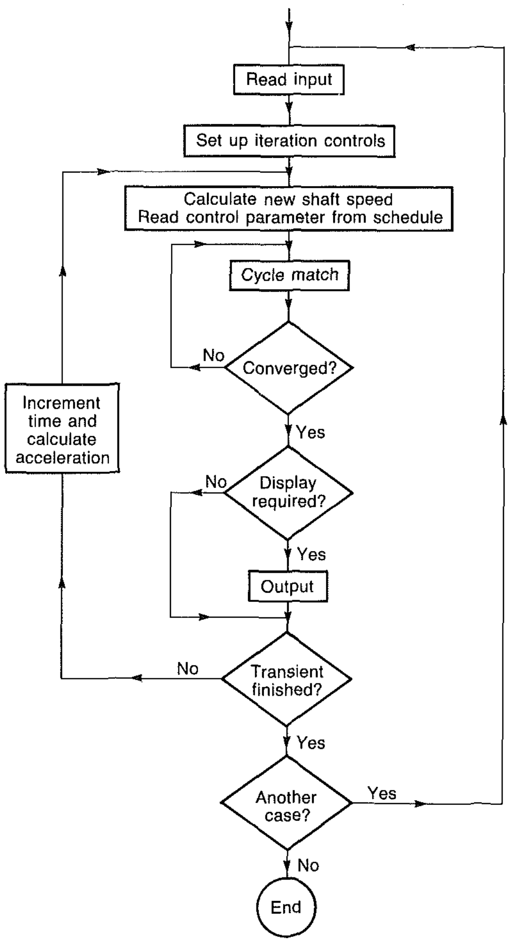


Fig. 12 Engine transient performance program—flow diagram

Transient Modeling Techniques

After the ground rules have been established, the performance engineer requires a method of ascertaining the most appropriate engine configuration from the point of view of transient operation. At Rolls-Royce non-real-time transient performance programs which are capable of simulating working line trajectories and acceleration behavior to a high degree of accuracy have been developed from the steady-state programs using a method based on very small time intervals. An outline flow diagram of the program is shown in Fig. 12.

These techniques have been applied to some engines with such success that the effect of an engine modification on transient behavior can be predicted so accurately as to make flight test virtually superfluous.

Nevertheless work is continuing so that some transient effects, at present modeled empirically, may be fully understood and scientifically applied. The work carried out to date has established the potential accuracy of the method but the engineer's task with any established engine whose transient behavior is known and has been measured is quite different from that of trying to establish desirable behavior from a new or study engine.

In the former instance the fuel schedule is known and the problem is to ensure that the computer model behaves the same way as the real engine for the same fuel flow schedule. In the latter case the actual fuel schedule is, at least at first, irrelevant. The task is to discover the potential of the turbomachinery and suggest ways in which this may be improved. When the best practicable solution has been worked out, the engineer is then in a position to provide appropriate data for the control system so that fuel scheduling laws may be established.

To this end some new techniques are being developed for new helicopter engines at RR Leavesden. The steady-state performance synthesis has been developed to a high degree of sophistication. Separate compressor characteristics for the axial and centrifugal parts of the compressor are incorporated including the effects of the variable geometry schedules in the axial compressor, thus allowing the effects of intercompressor bleed to be evaluated and enabling one to identify which part of the compressor is nearest surge at each part of the acceleration. Also included are separate characteristics of the intake protection system, installation effects, etc. The transient model has been developed from this using the small time interval technique but, in addition to several methods of scheduling fuel flow, accelerations can be run to a specified power versus time schedule or to given transient working lines on the compressor characteristics. In this latter technique an upper limit to the position of the transient working line on each characteristic is defined either as the surge line itself or, more usually, at a position a given distance away from the surge line so as to provide a safety margin (see Fig. 9). The definition of this safety margin, or transient surge margin, can present some difficulty particularly when dealing with separate axial and centrifugal characteristics on the same shaft. The most common definition of surge margin given above is that of the ratio of surge pressure ratio to working line pressure ratio at a given inlet nondimensional mass flow. From the point of view of transient behavior, this approach can present problems. First, surge occurs essentially at a constant speed rather than a constant flow and thus the important parameter is the distance between surge and the steady-state position along the speed line. Secondly, the above method takes no account of the shapes of the surge line and the speed lines (which may be vertical or near horizontal) and can thus seriously under or overestimate the margin (see Fig. 10). Consequently the definition based on a percentage of outlet flow function at a speed is used since it is closely related to overfueling margin, or the percentage increase in TET which needs to be applied at a speed to reach surge.

On this basis a transient surge margin allowance is made for inlet distortions, engine to engine variations in surge line position, deterioration, etc., and the upper limit line plotted on the characteristics and input to the program. Superimposed on these limits are maximum temperature, fuel flow, and speed limits and the program will then run to whichever is the lowest limit; providing output of power, fuel flow, limiting parameter, or indeed any other relevant parameter against time. These data can then be automatically plotted against time or each other for analysis and comparison with earlier runs.

Results from the transient testing programs of new helicopter engines have been encouraging. From the outset modifications to the engine scheduling suggested by the above techniques were incorporated and the results have been close to prediction with no-load to maximum power times of well under 2 s having been achieved without the requirement for any handling bleed. Work is now underway to provide still more sophisticated modeling of the compressor and the closest possible match between the measured behavior and the model.

In addition, it is intended that it will be integrated with a helicopter rotor model so that the full effect of rotor droop on the power turbine may be studied.

Conclusion

The increasing demand for helicopter maneuverability and safety are providing the stimulus for engine manufacturers to provide better transient response characteristics from their powerplants. In this context it is important that the correct emphasis is placed on what constitutes good handling. Analysis of the parameters involved enables the engineer to define the requirements in terms of power versus time and allows him to establish the most appropriate means of achieving them, whether it be by increasing the surge margin available or reducing the surge margin requirement by modifying the flow/speed characteristic of the compressor. A sophisticated transient modeling technique is an essential part of this process and can greatly reduce the time and effort required to obtain the best possible handling qualities from an engine.

CFD for Engine-Airframe Integration

G. C. Paynter

C. K. Forester

E. Tjonneland

Boeing Military Airplane Co.,
Seattle, WA 98124

This article provides an assessment of current CFD technology with application to propulsion integration, a definition of research and development needed to extend the technology, and a discussion of numerical error assessment and control. The CFD technology is divided into the elemental areas of the computer system, algorithms, geometry and mesh generation, turbulence modeling, and experimental validation; the current status and major issues in each of these areas are defined. Sources of numerical error are identified and some strategies for determining and controlling these are presented. CFD will have an impact on propulsion integration equivalent to that of the wind tunnel as CFD technology matures. This maturation will lead to a system which integrates the elemental areas of CFD for applications.

Introduction

The digital computer has been available to the aerospace industry for just the last two decades. Over this period, the cost of a given calculation has been reduced by a factor of ten every eight years; wind tunnel costs have continued to rise; and the available computer memory and speed have increased dramatically. In the aerospace marketplace, an important advantage is achieved through offering a superior product. In recent years, CFD has become an important tool for achieving this product superiority. These factors are responsible for the current trend toward the use of CFD to provide design information for configuration selection as illustrated in Fig. 1.

The primary role of CFD in the aircraft design process is to minimize parametric testing through the use of parametric analysis. The advantages of using CFD for design are lower design cost, time, and risk and better performance because a better understanding is obtained of the flow phenomena which have a strong influence on the performance [1, 2]. Moreover, CFD as a tool for aircraft design offers a step advance in technology similar to that offered by the wind tunnel.

Engine airframe integration can be defined as the art of combining the propulsion system components with the airframe to achieve the overall aircraft mission performance goals. This requires finding configurations with both good external aerodynamic performance and good internal performance. It is an art rather than a science because success still depends on the skill and experience of the designer rather than the systematic application of general laws or systematic reference to a data base which can be used to guide the design. There are several reasons for this. First, the scope of our interest in terms of the range of geometries and flows far outstrips our ability to acquire an adequate data base. Second, the flow phenomena associated with the installation are three dimensional and viscous. This means that the flow phenomena don't necessarily scale with the geometry and that interpretation of model scale test results is difficult. Third, these flows are difficult and expensive to predict; major reasons for this are the need for large meshes and uncertain turbulence modeling.

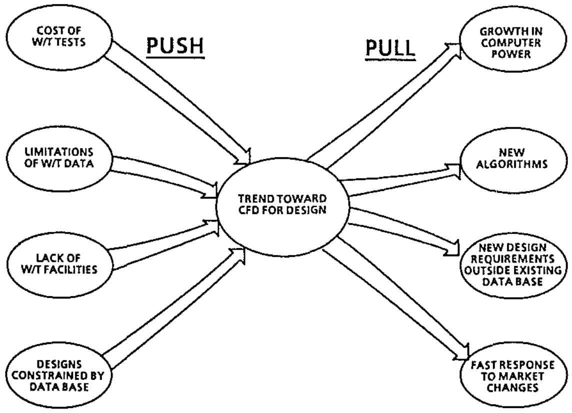


Fig. 1 CFD is an emerging technology

The typical process by which aircraft are designed is shown in Fig. 2. Aircraft design is by component buildup. This approach arose out of the way the wind tunnel is used to provide performance data for the installation components as a basis for configuration selection. Performance data are acquired for a range of configurations using either parametric analysis or testing. From this data base, configurations are selected for large or full-scale validation and optimization tests. From these results, a final design is chosen for flight test or production.

There are several problems with this process. First, interactions between components are neglected until there is evidence of a problem. For supersonic aircraft, the system must per-

---

Contributed by the Gas Turbine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the 31st International Gas Turbine Conference and Exhibit, Düsseldorf, Federal Republic of Germany, June 8–12, 1986. Manuscript received at ASME Headquarters January 24, 1986. Paper No. 86-GT-125.

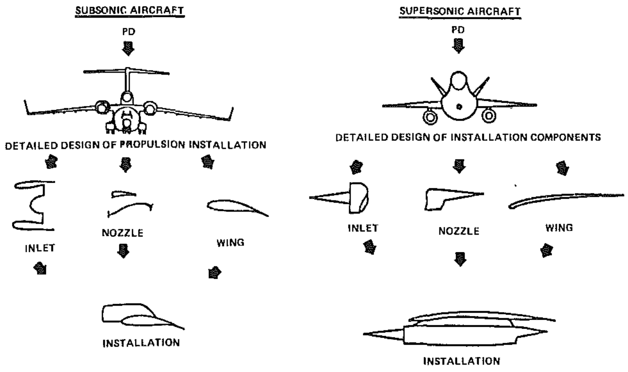


Fig. 2 Design by component buildup

form well at subsonic as well as supersonic speeds; an additional data base is therefore necessary. At supersonic speeds, adverse pressure gradients can be locally stronger and propagate farther than at subsonic speeds [3]. Neglect of the interactions between components can affect the aircraft performance in an unanticipated way from what was predicted using the performance of the isolated components. For example, for the NASA F-106 nacelle installation research program, the overall aircraft drag at $M = 0.9$ (computed from flight test) was 11 percent less than the drag predicted using the test results for the isolated nozzle boattail, as shown in Fig. 3.

If CFD is used to provide performance data for configuration selection, a number of problem elements must be addressed. These include computer hardware (storage and computational speed) and system limitations; algorithms; geometry representation; mesh generation; turbulence modeling; validation of the results; and pre- and post-processing. A user of CFD for propulsion installation design must successfully address each of these problem elements if the design process is to be successful. This is usually done within the framework of a zonal modeling strategy. The need for zonal modeling and a discussion of zonal modeling are contained in the next section of this paper.

The paper has three purposes:

1 to provide an assessment of current CFD technology for propulsion integration based on the authors' experience at Boeing,

2 to define areas where further research and development will substantially improve the current capability to use CFD for installation design,

3 to discuss numerical error assessment and control in CFD applications.

The purpose of the paper is not, however, to present the results of work completed or underway. Some examples of past or present work have been used where appropriate to illustrate a point.

The remainder of this article is a description of the zonal modeling strategy, followed by a section with an assessment of current capability and technology issues, a section on numerical error assessment and control, and concluding remarks.

Zonal Modeling

The purposes of this section are to discuss the need for a strategy which makes feasible the analysis of propulsion installation flow within the constraints of the available computer memory and speed; define the zonal modeling strategy as a way to satisfy this need; cite examples of the use of zonal modeling and provide an assessment of current capability; and make some general observations on where further research and development will improve the current capability.

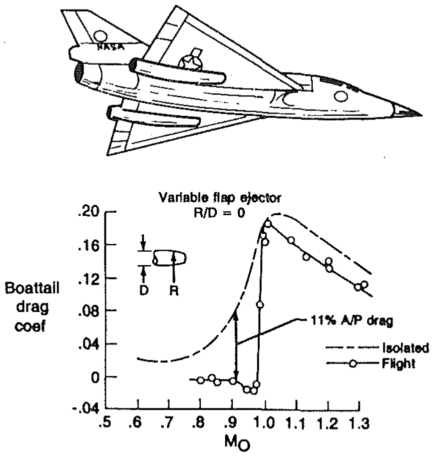


Fig. 3 Installation effect on boattail drag

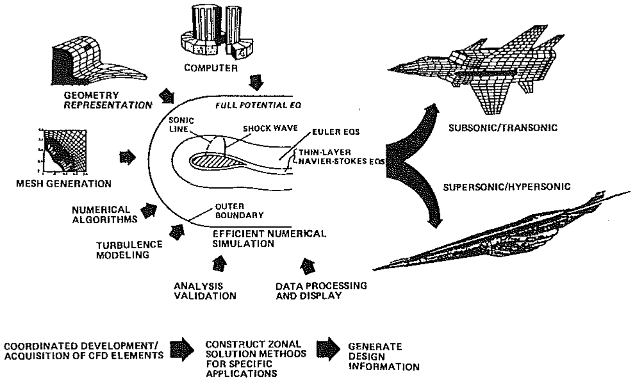


Fig. 4 Zonal modeling for design studies

The flow phenomena characteristic of a modern highly integrated propulsion installation are varied and complex. Most of these flow phenomena are three dimensional and viscous and there may be a strong interaction between the jet and inlet flows and the overall aircraft flow.

In recent years the time or mass-averaged flow three-dimensional Navier–Stokes equations have been routinely solved for specific flow phenomena; why not solve the full equations for the flows characteristic of the propulsion integration problem? It has been estimated that 10 to 100 million grid points (with up to 30 quantities stored/point) are needed for adequate resolution of a complex propulsion installation flow. A flow calculation on a grid of 10^6 points might require 10^4 iterations. A total of 10^{13} FLOP (floating point operations) would be necessary on a 10^6 grid to achieve an adequately converged solution for an algorithm which requires 10^3 FLOP/point/iteration. The current supercomputers achieve computational speeds of about 100 megaflops (10^8 FLOP/second). This suggests that 10^5 seconds (about 24 hr) on a supercomputer would be required to complete such a calculation.

A full vehicle flow simulation reported recently by Shang and Scherr [5] confirms these computer resource estimates. Shang computed the flow about the X24C re-entry vehicle at $M = 6$ with an explicit full Navier–Stokes code. This calculation was done in 20 hr on the CRAY XMP-22 on a grid of 500,000 cells. A more ambitious calculation is underway [6]; NASA-Ames and the Air Force Flight Dynamics Lab are working on a 10^6 mesh cell simulation of the F-16A aircraft. The goal is to complete the initial simulations by 1988 using the NASA Numerical Aerodynamic Simulation facility.

Computer speed and storage are still growing exponentially with time. Researchers with NASA and the Air Force have demonstrated that many complex flows of interest can be solved using three-dimensional Full Navier–Stokes (FNS)

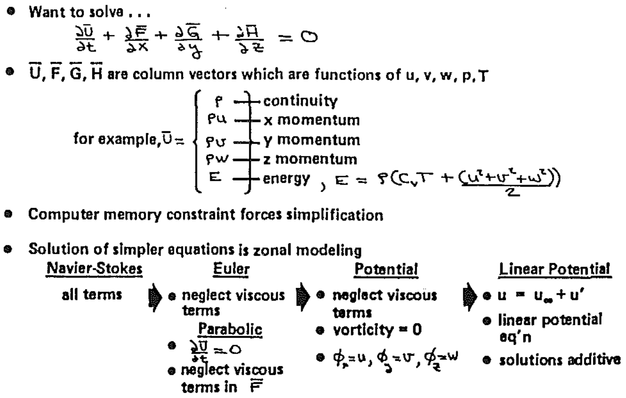


Fig. 5 Possible analyses within a zonal strategy

codes where the analysis domain is restricted in size. For the user, however, solution of the three-dimensional FNS equations over a large region is at present impractical for most design studies. Because of the rapid pace of computer technology development, computers with sufficient memory and speed for routine cost-effective solution of the full Navier–Stokes equations for large regions may become available within the next decade.

A caution to potential CFD users is necessary. It is feasible at present to do a "coarse-mesh" solution of the FNS equations for a large region. Such solutions are useful for qualitative estimates of the flow properties but, because the computed results are contaminated by an unknown truncation error which causes a diffusion effect, the quantitative accuracy is unknown. Predicted flow properties from coarse-mesh FNS calculations cannot be used to provide performance data which are reliable enough to support a design study.

Three-dimensional viscous effects which require solution of the full equations generally occur, however, in only very localized regions. About 90 or 95 percent of the flow about an aircraft is adequately described by equation sets which neglect some or all of the viscous terms. The zonal analysis strategy uses this fact to reduce the computer resources required for complex flow calculations over the large flow domains characteristic of aircraft.

Zonal analysis allows solution of complex flow problems on today's computers. Zonal analysis is the use of multiple analyses in several flow regions to take advantage of simplifications that are available through the local nature of the flow. The analyses are coupled or interacted in a way appropriate for the flow to be predicted as illustrated in Fig. 4. A reduction in the computer resources required to achieve a flow solution is traded against the increased complexity of the data management required with the zonal approach.

The set of equations which could be solved within the context of a zonal strategy ranges from three-dimensional FNS to linear potential as illustrated in Fig. 5. Use of a zonal strategy does not exclude the use of FNS as a component analysis of a zonal strategy. The zonal strategy is used to minimize the size of the solution domain to which the FNS procedure is applied and the FNS procedure is interacted in some way with less complex analysis to provide a solution for the full flow domain of interest.

Several factors are required for the successful implementation of a zonal modeling strategy. First, a hierarchy of flow codes ranging from simple to complex must exist. Second, it is important that the flow analyst understand the flows to be predicted so that a successful strategy can be developed. Finally, procedures must be developed for efficient interfacing of the codes. Desirable features of this system include use of body-fitted meshes, common input and output formats, common turbulence modeling, common code architecture, compatible numerics, and use of computer graphics for I/O. These features minimize the difficulty in coupling and interfacing the several codes used in a given analysis. These features also minimize the training, skill, and time required on the part of the user to achieve a successful result.

A number of successful applications of zonal analyses within the context of design studies can be cited. These include analysis of supersonic mixed compression inlet flows [7], subsonic propulsion installation flows [8], turbofan and turboprop installations on subsonic aircraft [9, 10], and the interaction of a canard vortex with a wing [11].

From the cited examples it should be clear that some form of zonal modeling is required today for the analysis of complex installation flows; a procedure must be constructed for each type of flow to be analyzed; and that a zonal strategy is constructed around the component flow codes available. The zonal approach is useful for dealing with the wide range of length scales characteristic of installation flows. The relative success of a procedure depends on the skill of the user in selecting an appropriate strategy and appropriate component analyses. The component analyses must accurately simulate the flow in each solution subdomain. The most successful methods represent a balance between sophistication and utility.

In the future, the rapid pace of computer hardware development should increase the sophistication of the component analyses and reduce the complexity of the zonal modeling. Increased use of computer graphics and CFD work stations should enhance the use of the zonal strategy for design. Adequate validation of zonal procedures will continue to be a problem because of the cost of acquiring high-quality experimental data suitable for analysis validation for complex installation flows.

The introduction describes the purpose of this article and this section is a discussion of zonal modeling. The remainder of the article contains sections that assess current capability and technology issues, a section on numerical error assessment and control, and a section with concluding remarks.

Assessment of Current Capability and Technology Issues

Because CFD is now an important tool for achieving superior aircraft designs, many novices are attempting to develop or acquire the technology and apply it effectively. CFD is a complex, rapidly developing technology which is easy to misuse. Such misuse results in inferior designs which are based on misleading or misunderstood information. It is also easy to use CFD inefficiently in that an inordinate amount of time and money are required to complete the work. A bewildering number of choices must be made by the user in the elemental areas which comprise the technology. A poor choice in any one of these areas or a poor overall strategy can lead to a failed or ineffective application of CFD.

As noted by Ritchie [12], the greatest invention of the Wright brothers was the process used to combine engineering, science, and testing to solve the problem of powered flight. This process which has characterized aerospace research and development (R&D) ever since is an efficient way to achieve complex and difficult objectives; it is the process which should be used to develop or apply CFD.

The R&D process is completion of the following steps:

- definition of an overall objective
- identification of key technology elements
- definition of an objective for each technology element which would lead to accomplishment of the overall objective
- preparation of a plan for accomplishing the technology element objectives

- coordination of the work toward the elemental objectives to ensure compatibility and to avoid duplication
- early identification of potential problem areas
- modification of the plan to resolve these problems within the constraints imposed.

In using this process to develop or apply CFD, the key technology elements are generally the computer hardware and system; geometry representation; mesh generation; turbulence modeling; flow codes and/or algorithms; validation; and pre- and postprocessing. In an application of CFD, definition of an appropriate zonal strategy and data handling could be added as technology elements.

For the user, the current CFD technology still has a number of problems. The codes are generally designed to operate on a single computer and are therefore not portable. Most codes are not user friendly. Both of these deficiencies reduce the usefulness of CFD to the design community. Miranda [13] has suggested that the effectiveness of CFD for design is the product of the quality of the result times acceptance by the users. Nonstandard mesh generation, geometry, algorithms, I/0, etc., and high cost reduce acceptance by the user community. These problems are the issues that need to be addressed to improve the process of using and thus the usefulness of CFD for design.

The researcher/developer works toward extending CFD for either increased accuracy or improved acceptance of the technology by users. The purposes of this section are to offer the researcher an assessment of current capability and to suggest technology issues with a high return on investment. Toward these goals, the elemental areas of CFD which are the computer system including pre- and postprocessing, algorithms, geometry and mesh generation, turbulence modeling, and experimental validation are reviewed sequentially in the order listed.

Computer System. For CFD, the primary purpose of the computer system is to process a set of instructions (code) for approximate solution of the partial differential equations which describe the fluid motion. In a field method, the flow domain of interest is divided into a finite number of cells (or points) in which the flow properties at a cell are related to the properties at neighboring cells through the algebraic equations used to approximate the partial differential equations (PDE). In the solution process, several numerical quantities must be stored in the computer memory for each cell and periodically updated through iteration of the numerical procedure. In each iteration, new values for the flow properties of a cell are computed from old values at the cell and values (old or new) from neighboring cells. The total memory required for a given calculation is thus the number of cells times the number of quantities stored per cell. The total time required for a given calculation is the number of floating point operations per cell per iteration times the number of cells times the number of iterations divided by the computer processing speed. Mesh generation, graphics, etc., if completed on the supercomputer, add to the total memory and CPU time requirements.

The parameters of the computer system which determine the maximum number of cells in a solution domain and the complexity of the PDEs solved are the size of the computer memory and the speed with which the computer can process instructions. A measure of the memory size is the number of 64-bit words which can be stored. A measure of the computing speed is the number of floating point arithmetic operations which can be completed per second (including the time required to fetch the operands from memory and return the result to memory). The speed of modern supercomputers is defined by the millions of floating point operations which can be completed per second (megaflops).

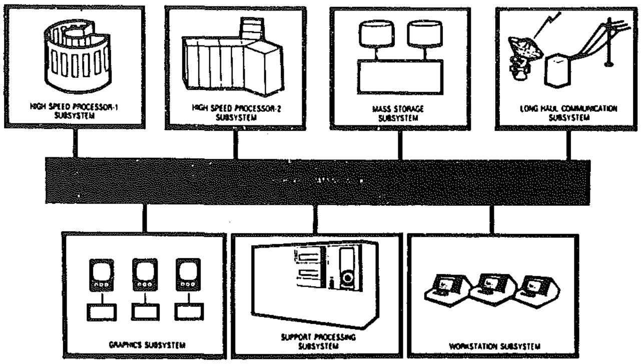


Fig. 6 Supercomputer system configuration

Computer speed and storage are still growing exponentially with time. The state of the art in supercomputer technology today is the CRAY 2 computer with 256 million words of fast memory and 250 megaflops speed. Graves [14] has suggested that computers with several billion words of memory and speeds of as much as a trillion floating point operations per second may be available in the 1990s. Cost is also an important consideration; while available speed and memory are increasing exponentially, cost is increasing linearly with time. As noted in [15], the cost to do a given calculation has decreased by a factor of ten every eight years and this trend is expected to continue.

The important issues in computer system technology arise from a desire to apply the rapidly developing technology efficiently. Efficiency is defined herein as minimizing the time and money required to apply CFD for design on a computer system as well as minimizing the cost to the user of the system.

The high volumetric through-put of data from a supercomputer requires a data management system. The system must generate, transmit, display, and manipulate the data associated with a CFD application. A modern system is illustrated in Fig. 6 from [16]. The system consists of the data communication links, terminal devices, off-line data storage, work stations, supercomputer, and the system software. If the supercomputer is to be shared by a number of users, the system must be accessed via high-speed long-haul communication links from work stations by off-site users.

Supercomputer memory and speed improvements are expected to continue. While some speed improvement is expected from faster components, new computer architectures will be more important than faster devices for achieving more speed. Two major architecture philosophies currently exist. One links a small number of very fast processors optimized with vector processing, pipelining, etc., for limited parallel operation. The second philosophy is to link a large number of low-power processors in a massively parallel design which would require multiple streams of instructions and data. For the user, the issues are the trade between cost and utility implied by the different architectures. An architecture which resulted in a very low cost but difficult to program computer could, for example, result in a higher overall cost to the user for a given application because of increased programming cost.

Artificial intelligence (AI) is expected to play an important role in improving the usefulness of CFD. CFD software is primitive and difficult to use. An expert system could automate many tasks and adaptively interact with the CFD user via a visual display [17] of menu options available on a work station terminal. AI is just beginning to be used to improve the process of applying CFD [18].

Algorithms. MacCormack [19] and Shang [20] have recently assessed the current status and trends in algorithms for solving the Navier-Stokes equations (and subsets). Current algorithms can be grouped into explicit and implicit methods which use some or all of the following steps in the solution process

- generate fluxes
- correct fluxes by a smoother or upwind limiter
- generate residuals
- generate a correction, preliminary or accepted
- generate update
- generate new cell coordinates (if multigrid and/or adaptive mesh are used)
- repeat cycle until convergence criteria are satisfied.

Desirable properties [19, 20] of an algorithm include the following:

- accuracy—second order or higher
- low number of arithmetic operations per iteration
- low number of iterations
- no dependence of the steady-state solution on the choice of time step
- fully conservative
- no added numerical dissipation for solution stability
- low numerical dissipation from truncation error
- easy implementation
- minimum storage
- easy boundary condition specification
- compatibility with available computers

Although explicit and implicit methods are both currently in use, a trend toward the use of implicit methods has evolved over the last decade in attempts to achieve algorithms with more of the above listed desirable properties.

The primary advantages of the explicit methods are ease of implementation, easy boundary condition specification, easy vectorization, and the low number of arithmetic operations per iteration. The primary disadvantage is the large number of time steps required to steady-state convergence. Wall functions to resolve turbulent boundary layers and local time stepping can be used to alleviate this problem somewhat.

The primary advantage of the implicit methods is the low iteration count to achieve a steady-state solution. Primary disadvantages include more difficult implementation, more difficult boundary condition specification, numerical dissipation due to approximate factorization and linearization error, more difficult vectorization, and a larger arithmetic operation count per iteration.

As noted by Shang [20], although the implicit procedures require more arithmetic operations per solution step than the explicit, the more favorable stability characteristics of the implicit methods allow a larger time step (toward a steady-state solution). The faster convergence rate of the implicit methods more than compensates for the overhead in the matrix inversion process to yield a substantial savings in computer resources. The relative cost of implicit and explicit solutions is dependent on both the problem solved and the accuracy required.

MacCormack [19] has recently proposed a hybrid explicit/implicit algorithm with many of the best features of each. The basic algorithm is a second-order-accurate explicit method plus options for flux splitting, local time stepping, multigrid for convergence acceleration, Newton iteration for steady-state solutions, and local implicit procedures with iterative matrix inversion. This type of algorithm is the current state of the art for solving the Navier–Stokes equations.

For the users the primary algorithm issues are cost, accuracy, and ease of implementation. In general the complexity of the algorithm increases as the order of accuracy and degree of implicitness (size of the solution step) increase. The degree of implicitness also increases the difficulty of maintaining boundary condition flexibility.

As noted by Forester [21], a trade always exists between the computer work and the engineering work required to implement and use an algorithm. With computer technology and algorithm technology both in a period of rapid development, the optimum level of algorithm sophistication is debatable. This optimum level will change with time as either new algorithms or computer hardware technology becomes available.

Geometry and Mesh Generation. Geometry programs provide an analytic representation of a geometric surface (nacelle, wing, etc.) which is a boundary of a flow region of interest. The geometry procedures which provide this analytic surface representation can be divided into two groups: analytic and numerical [2]. The analytic procedures utilize combinations of various algebraic line segments to approximate the surface. Numerical procedures use discretized data (coordinate locations of surface points), for example, digitized cross sections at a number of longitudinal stations, as input to a program which fits analytic functions (biparametric cubics, for example) to the input surface data.

The surface geometry description is used in two different ways in CFD applications and the way in which it is used depends on the flow analysis strategy. Flow analyses can be divided into two types: panel methods and field methods. In a panel method, linear superposition is used to satisfy boundary conditions imposed at the center of "panels" which are surface quadrilateral elements. The function of a geometry program for a panel method is to subdivide the surface into suitable panels and to provide the panel center locations, corner point locations, and local surface geometry properties so that the boundary conditions can be imposed. In a field method, a grid or mesh is fit about the geometry throughout the flow domain of interest and the discretized PDEs are solved for each cell or point in this domain. Field methods can be divided into two types: Cartesian and body fitted. In a Cartesian field method, the geometry description is used to provide intersections between the mesh and the bounding surfaces. In a body-fitted method, the body surface is a coordinate surface of the transformed coordinate system. The geometry description is used to help define the transformation to the body-fitted mesh system.

A number of sophisticated geometry description methods are available because of the long-term emphasis on Computer Aided Design (CAD). (If a CAD system can be used to provide the geometry description, it has the added benefit that it can generate a numerical control machine tape for machining to close tolerance the surface geometry for model tests or full-scale hardware.) A geometry description which starts from a discretized surface input is of greater utility to the CFD user because of the ability to represent very complex geometry, generate surface paneling, and extract surface "cuts." The numerical geometry description program can also be used in some cases as an algebraic mesh generation program. This is desirable because it reduces geometry-mesh interface problems and reduces by one the number of complex computer programs which must be mastered for a successful CFD application.

The primary issues in geometry description are speed, cost, and compatibility with the overall computer system. "Smart" work stations are needed with color graphics that allow shading, fast rotation, translation, and "zooming" that can quickly reveal surface flaws which would preclude a successful CFD application. The ability to quickly "merge" simple geometries (the nacelle and the wing, for example) into a more complex configuration is desirable because aircraft are designed in just this way. A display of mesh and/or flow properties on a surface or in planes intersecting a surface is important for evaluating the overall analysis result.

Mesh generation is the discretization of the flow domain into a finite number of points or cells at which the algebraic approximations to the PDEs are solved for the local flow properties. Meshes can be divided into two types: Cartesian and body fitted. With both types, coordinate line intersections form the computational mesh. With a Cartesian mesh, an arbitrary body is contained within the mesh, mesh surface intersections are irregular, and special differencing is used to compute the properties on or near the surface. In a body-fitted mesh, a coordinate line is coincident with the boundary. This allows boundary conditions to be applied without interpolation or special differencing schemes. Computations are done on a fixed square or cubical grid, depending on whether a two-dimensional or a three-dimensional transformation is used with the transformed coordinates as the independent variables. Most available programs require body-fitted meshes; the remainder of this subsection addresses the current status and issues pertaining to body-fitted mesh generation.

For the complex flow domains of aircraft and integrated propulsion installations, it appears to be impractical to discretize the necessary flow domain with a single mesh generated for the whole region. MacCormack [19] has suggested that this could be accomplished with a dozen or so subregion meshes, each fitted to an aircraft element and aligned in a natural manner with neighboring subdomain meshes. Thus a two-step process is currently used in grid generation for complex flow domains. The first step requires selecting a concept as to how the overall domain is to be divided into subdomains and how the various subdomain meshes are to interact. The second step is to select the type of grid generation scheme to be used within each subdomain. Numerous articles, for example Thompson [22], are available on the details of current mesh generation procedures.

Current mesh generation procedures for body-fitted subdomain grids are of two general types: (1) numerical solution of partial differential equations, and (2) construction by algebraic interpolation. For the first type, grid generation procedures are available based on solution of elliptic, parabolic, or hyperbolic partial differential equations. For the second type, mesh coordinates are determined by an interpolation along an algebraically determined path between two bounding surfaces.

The primary advantages of the algebraic procedures are low computational cost and explicit control of grid point distribution. A primary disadvantage is that the grids are less smooth than those generated by solution of PDEs.

The primary advantages of grids generated by solution of PDEs are smoothness and better control of grid orthogonality at boundaries. The primary disadvantages are the increased computational cost and complexity.

A large number of grid generation procedures of both types are available. For a given application, the user should select a procedure(s) which is easy to implement; provides a one-to-one mapping; is smooth; provides high grid density in regions of high gradients in flow properties; and has low "skewness." High grid skewness can result in large truncation errors if this occurs in regions where gradients in flow properties are also high.

The primary issues in grid generation are ease of implementation, compatibility with the computer system, control of "smoothness," control of "skewness," control of local orthogonality, control of surface point distributions, control of cell aspect ratios, control of local stretching, truncation error estimates, and mesh adjustment. For optimum efficiency and error control, the mesh must be controlled to obtain a high mesh density in regions of high gradients in flow properties without a priori knowledge of the solution.

The local truncation error is a function of the local grid scale, skewness, and gradients in properties. CFD solutions have one unfortunate characteristic in this regard. Solutions with large truncation errors can appear to be physically plausible. If the computed results are to be useful for design, grids must be used which result in performance quantities which are accurate enough for configuration selection. Numerical error assessment and control are specifically addressed in the next major section of this paper.

Turbulence Modeling. It is generally accepted that the unsteady Navier-Stokes, continuity, and energy equations combined with an equation of state provide an accurate description of the flows of interest (within the context of engine-airframe integration) and most of these flows are turbulent in regions which have a strong influence on the aircraft performance. Important length and time scales of the turbulent motion are so small that the grid scale and time step necessary to directly compute such flows are impractical for flows of interest. The approach used to avoid this difficulty is through solution of the Reynolds-averaged equations. The Reynolds-averaged equations are obtained by assuming that local flow properties at a given point in space and time are the sum of a mean value and a fluctuating value and then ensemble averaging or time averaging these equations.

This results in new terms which can be interpreted as stress gradients and heat fluxes associated with the turbulent motion. If these new terms along with the local mean flow properties are regarded as dependent variables, the number of unknowns outnumbers the number of equations. In order to "close" the system of equations, additional equations are necessary to relate the new turbulent stress and heat flux quantities to the mean flow properties. The turbulence model is the assumptions and approximations used to arrive at the additional equations needed for closure.

Although a number of modeling approaches are available and are being developed, for the CFD user, eddy viscosity concepts, in which the effects of the turbulence on the mean flow are represented by increased viscosity and conductivity, are the methods of choice. Within the eddy viscosity concept, a number of levels of sophistication are possible. From dimensional considerations, the eddy viscosity can be considered to be proportional to a length scale times a velocity scale. The levels of sophistication arise from the ways in which the length and velocity scales are related to the mean flow.

The two most popular eddy viscosity modeling approaches are mixing-length models and two-equation models. Mixing-length models assume the length scale is proportional to some fraction of the local shear layer thickness. The chief advantage of these models is simplicity. The primary limitations are the difficulty in switching from one type of turbulent flow to another. Adequate mixing length models do not exist for separating, reattaching, and separated flow zones. Two equation models solve transport equations (PDEs) for the local length and velocity scales. The primary advantage of two equation models is that no problem occurs in defining length and velocity scales with transitions from one type of turbulent motion to another (although the accuracy of the prediction is open to question). The primary limitations are the added computational cost and the inability of these models to predict counter gradient fluxes.

The current status of turbulence modeling is that the available modeling can predict many two-dimensional and a few three-dimensional flows of interest. Available models are based on two-dimensional data; the argument for their use for three-dimensional flow is that turbulence phenomena are three dimensional even when the mean flow is two dimensional. The assumption is that models developed for two-dimensional flows should work for at least some three-dimensional flows of interest. As noted by Hall [25], the available turbulence models work well for most flows where the turbulent structure changes slowly. Examples of these flows are boundary layers, mixing layers, wakes, or jets in mild adverse or favorable pressure gradients. The available models don't seem to work where rapid changes in turbulent structure occur. Examples of these flows are regions of separation, reattachment, shock/boundary layer, or shock/shear layer interaction.

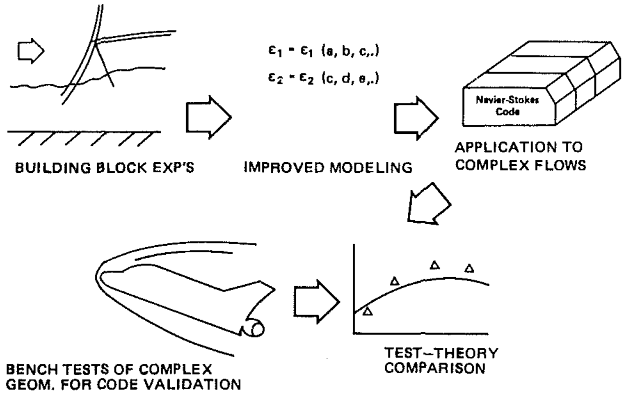


Fig. 7 Experimental validation of CFD codes

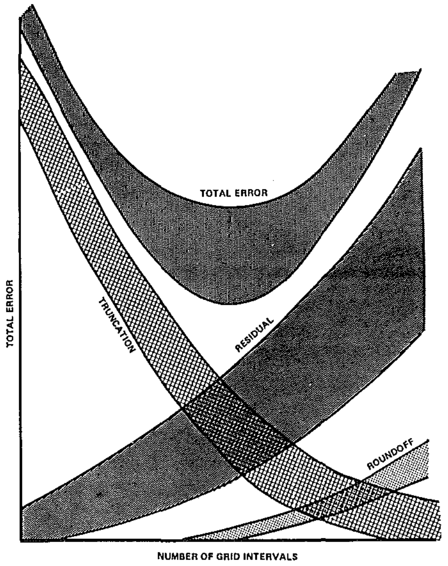


Fig. 8 Total numerical error

The most immediate issues are (1) selection of models which can be used with confidence to obtain quantitative design information, and (2) computational efficiency. A more fundamental issue is the need for turbulence models for the complex turbulent flows of aircraft. These models must be simple enough for and compatible with current numerical procedures and as general as possible. A promising approach to this does not exist at present. The problem is that an infinite number of flows must be considered on the basis of geometry. Bradshaw [26] has proposed classifying these flows on the basis of phenomena rather than geometry. The 1980–81 Stanford Meetings on Complex Turbulent Flows are believed to be an important step toward improved modeling.

Experimental Validation. Validation of a CFD analysis for design occurs at two levels as illustrated in Fig. 7. At the first level, the component analyses and modeling are shown to be capable of predicting the specific flow phenomena expected for a given application. At the second level, the selected zonal analysis strategy and the way in which it is applied are shown to be capable of providing performance data of sufficient accuracy for configuration selection.

Some detailed experiments have been completed and data are available for the first level of validation for flows of interest for aircraft design. One source of these data is the 1980–81 Stanford Meetings on Complex Turbulent Flow.

Detailed experiments are also available for the second level of experimental validation, although these are scarce. Second level experiments are done at a large scale, are very expensive, and the data from industrial experiments of this type can be unavailable for general distribution to protect proprietary interests. Data from these experiments are usually adequate for determining the overall performance of a configuration (lift, drag, thrust, etc.) and are useful for determining the accuracy to which the performance parameters can be predicted. Such data are usually inadequate for suggesting why a discrepancy exists between computed and measured quantities. It is important to understand that uncertainty in initial and boundary conditions contributes to error bands for both the computed and the predicted results. If, for a given test condition, the error bands of the test and the analysis overlap and the discrepancy between a computed and predicted property falls within this overlap, no conclusion as to the accuracy of the analysis can be drawn.

The primary issues are the availability and the quality of the data. Some data needs for the first level of validation were defined as a result of the 1980–81 Stanford Meetings on Complex Turbulent Flow. For the second level of validation, generic experiments are needed which span the range of flow and geometries of general interest for aircraft design. Almost no data of adequate quality are available for second level validation (although a number of experiments have been completed with this as the stated objective). A chronic problem with second level experiments is that these are completed without the requirements for CFD validation in mind.

The introduction describes the purpose of this article and this is followed by a section on zonal modeling. This section is an assessment of current capability and technology issues and the remaining sections are a section on numerical error assessment and control and a section with concluding remarks.

Numerical Error Assessment and Control

Types of Error Sources. The numerical methods for solving PDEs have three types of error sources: grid-related solution process (residual) errors, grid-placement related (truncation) errors, and roundoff errors. Explicit marching techniques (time or space) do not have residual errors unless there is an implicit equation imbedded in the marching scheme. Residual errors occur when implicit equations are solved by explicit marching techniques, specialized relaxation schemes, or matrix inversion processes. Truncation-error and residual-error effects produce many curious phenomena which must be understood in order to develop reliable numerical results.

Figure 8 shows a plot of the total numerical error that is associated with the numerical solution of the PDEs. Error bands for the truncation, residual, roundoff, and the total errors are used to denote the range of variation of these errors depending upon the computing environment. Note that roundoff error has been included in the chart. Roundoff error is a concern either when the significant digits are too small for the grid size or when the grid is too large for the number of significant figures carried in the computations. In the future, very large computers will permit a higher number of grid points.

In most CFD applications truncation and residual errors are more important than roundoff error. Typically the cost increases as better control of these errors is sought. Accurate results are expensive in both manpower and computer time. Computed results are difficult to interpret unless residual and truncation errors are understood in a given application. These errors are typically assessed separately. An interaction can occur between these error sources and special care is required to isolate the magnitude of each type of error. The importance of careful numerical error assessment cannot be overemphasized.

Let

$$LU = 0$$

represent the PDE system of interest.

In discretized operator notation, equation is

$$L^I U^I = T^I - R^I + (\text{Roundoff})$$

where L^I is the discretization operator, U^I is related to the discretized dependent variable vector, and T^I is the local truncation error and R^I the local residual error for each cell of the analysis domain. The grid structure index I is related to choices of the grid density distributions for each selected trial grid where in general the computational grid features coupled conformal grids with grid nesting in subregions. Ideally the grid adjustments are made in some pattern that tends toward a limiting grid configuration or "goal grid." Thus each unique grid shape is represented by an integer value of I. The goal grid is assigned the index I_g, which is known once the numerical error has been constrained to the desired bound.

Truncation (*Grid Related*) *Errors.* Truncation errors are due to the selection of the grid, the grid-related algebraic equations, the boundary and initial conditions which approximate the field equations of interest, and the smoothing operator and upwind weighting coefficients that are used.

The local truncation error is formally defined as the magnitude that the left-hand side yields for each cell when the goal grid solution is interpolated (restricted) to a trial grid difference equation minus the interpolated value of the left-hand side of the FDE from the goal grid. In practice the grid chosen for the goal grid can be any trial grid of interest. It is set at zero in conventional representations of the FDE for all values of I. Truncation error estimates are used to correct for truncation error effects. The right-hand side of equation of the FDE can also be a high-order accurate truncation error correction. Dissipative and nondissipative convective difference schemes permit expansion shocks, artificial gross separation, etc., to form under certain conditions. Artificial diffusion is added to eradicate the expansion shocks. Tuning the artificial diffusion coefficients for peak accuracy is troublesome. The goal of the tuning process is that the artificial diffusion must decay globally and locally with mesh refinement so that the accuracy of the sonic line shape and position increase with mesh refinement. This tuning process is controlled by coefficients of the upwind blending functions and the smoothing operators.

Residual Errors. The numerical modeling of the FNS equations leads to algebraic forms in which the propagation of signals is retarded as the grid density increases. This stiffness problem leads to inefficient reduction of residual errors among the simultaneous system of algebraic equations, often leading to exponential or power function decrease in convergence as the grid density increases. Three stiffness problems can appear simultaneously or separately: acoustic, convective, or diffusive stiffness; and can be aggravated by nonuniform grid. Any or all of these factors can undermine the convergence rate severely and can enlarge errors substantially. The error monitors must be designed to detect slow convergence or inefficient residual error control. For some error norms this is a severe requirement. When the local residual error is reduced to some fraction of the local truncation error, further reduction of the residual is not cost effective; determination of this fraction is a subject for further study in two-dimensional and three-dimensional FNS computations.

Error Assessment and Control

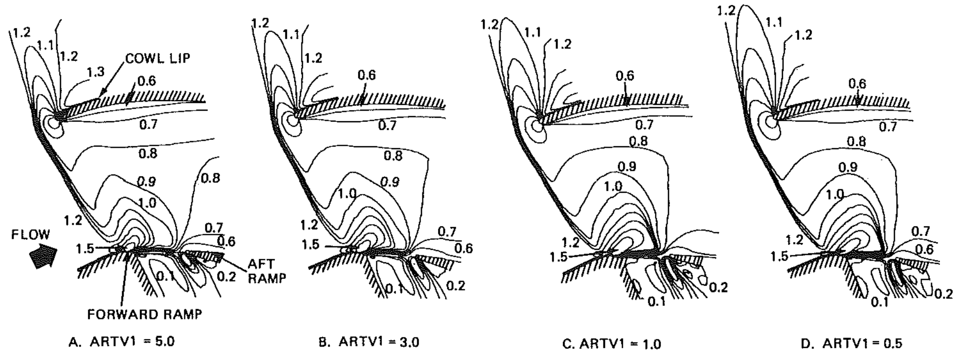
Conventional Certification Process. The process of numerical error assessment with conventional PDE modeling techniques is the following. A solution of finite difference equations (simultaneous system of algebraic equations) for a specific discretization of the analysis domain is generated for different choices of grid density and grid distribution in the analysis domain. It is common to use a sequence of grids of the same grid distribution that differ in grid count in each independent variable direction by factors of two—2, 4, 8, 16, 32, etc. The coarser grids can be generated by deleting every other point of the finer grids. The effects of the choice of grid distribution are examined by choosing sequences of grids which have different mesh distributions. The data from all of these solutions of the grid-related equations are organized by constructing an error difference table. Solution differences are posted in order of the coarse-to-fine grids for each grid sequence. The solution differences are generated by subtracting the values of adjoining pairs of grid solutions of the dependent variables at all physical locations in the analysis domain that correspond to the grid coordinates of a grid of a selected intermediate density. Interpolation is used to relate other grid solutions to these selected grid coordinates. As the grid density increases the differences should decay approximately[1] according to the formal order of accuracy for some selected mesh distribution. Iterative adjustment of the grid distribution and density is made until this type of error decay is realized. If this occurs, extrapolation may be used to solutions at infinite grid density and reliable estimates of the maximum global error may result.

The preceding process appears to work best on the modeling of parabolic and elliptic equations in smooth domains with smooth boundary conditions. For mixed elliptic/hyperbolic systems, erratic results may occur due to unresolved singularity regions and/or poor residual error control, and/or Gibbs' error effects.

A key feature of this method is that grid adjustments are made in some pattern that tends toward a limiting grid configuration. A way to think about this is to define a goal grid to which the selected grid sequences must evolve. The goal grid is a grid upon which the solution is sought to some specified er-

---

[1] Error decay according to the formal order of accuracy is expected globally but not locally in regions of singularity.



A. ARTV1 = 5.0 B. ARTV1 = 3.0 C. ARTV1 = 1.0 D. ARTV1 = 0.5

Fig. 9 Effect of fourth-order numerical smoothing term on analytical *Mach number contour results*

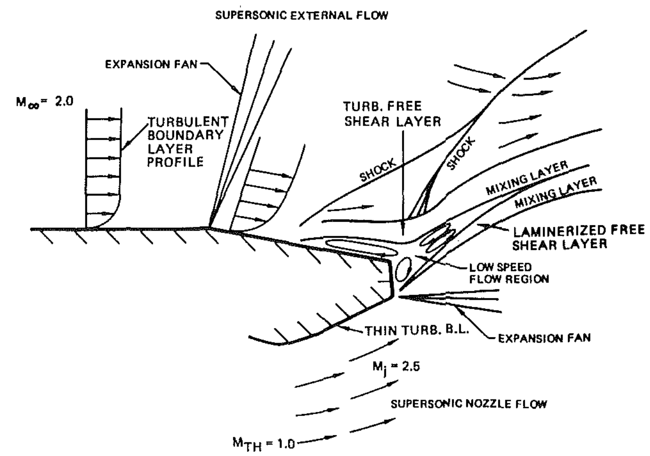
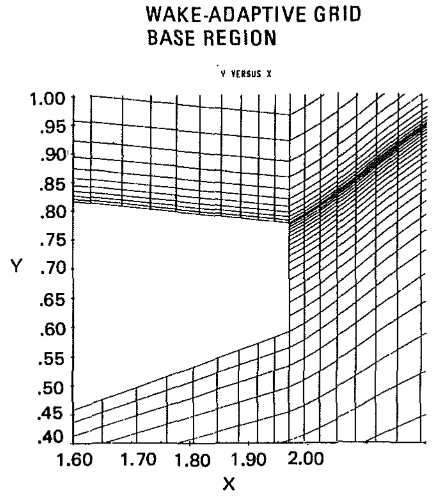
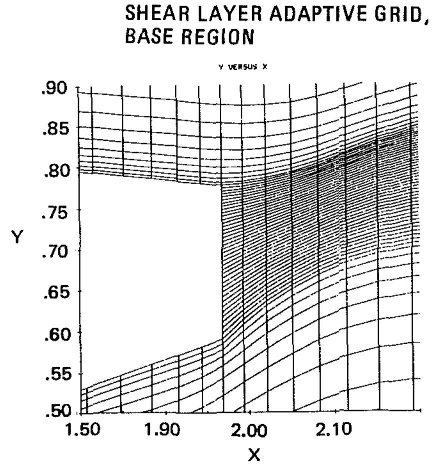


Fig. 10 Experimental validation of CFD codes


WAKE-ADAPTIVE GRID
BASE REGION

Y VERSUS X

Fig. 11(*a*)


SHEAR LAYER ADAPTIVE GRID,
BASE REGION

Y VERSUS X

Fig. 11(*b*)

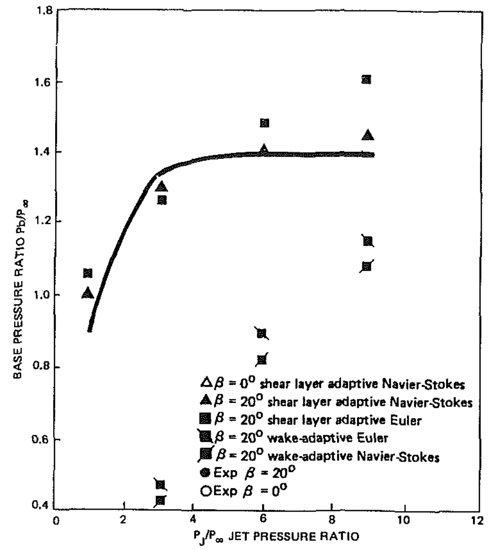


Fig. 12 Base versus jet pressure ratio

ror bound. It should be understood that the goal grid may not be unique because of grid initialization, grid generator, and grid-equation solver properties. It is assumed that adequate control of the residual error effects has been maintained in the process of assessing the truncation error effect. This is done by developing a sequence of *several solutions* on *each grid choice* with various choices of constraints on the residual tolerances that are used to terminate the computations for each solution on that grid. Because of the need to assess the contamination by residual error, the goal grid may not be the grid of greatest density but it will have the correct shape.

Example—Inlet FNS Analysis. As mentioned above the truncation error level can be influenced by the smoothing and upwind coefficient selection. Figure 9 shows an example of the variation of accuracy of a two-dimensional inlet FNS analysis when varying the smoothing coefficient. A reduction of the magnitude of the ARTV1[2] coefficient improves the accuracy of the analysis dramatically. Figures 9(*c*) and 9(*d*) agree with the shadow graph data for this inlet configuration with an upstream Mach number of 1.28. Figures 9(*a*) and 9(*b*) disagree strongly with the experimental data. These computations were performed with reasonably placed meshes of reasonable density.

Example—Nozzle FNS Analysis. Figure 10 shows a cross section of an axisymmetric nozzle and the associated flow physics of this rocket nozzle plume flow field. Several fixed grid and two adaptive grid FNS analyses of this flow field have been done by several independent investigators. The most common findings of these researchers is that the base pressure is dramatically underpredicted. The underprediction of the base pressure has been traced to the manner by which the grid is placed. Simply blanketing the regions of high gradient with grid has led to poor results. Figure 11 shows a grid which does give good results. The secret of the success of this grid is that the grid is adaptively aligned with the free-shear layers. Figure 12 shows the magnitude of the difference between this form of placing the grid over the other methods which are typified by the wake-adaptive grid. The error analysis of these FNS analyses by error norms indicates the sources of the trouble. The difficulties lay at the corners of the base of the nozzle. Large destruction of total pressure can occur at these corners. Large truncation errors due to artificial diffusion are believed to be responsible for this loss. With proper mesh alignment the truncation error can be dramatically reduced, resulting in much improved base pressure prediction.

In the preceding examples, contamination by residual error makes the solution appear as though the grid is coarser than it really is. This means that sharp detail is smeared and unfocused in the correct locations. Similar to Figs. 11(*a*) and 11(*b*) a sequence of solutions at various residual contamination levels is used to observe the influence of these errors on the solution. This work usually precedes the study of the trun-

---

[2] Artificial viscosity.

cation error effects and insures an uncluttered view of these effects.

Concluding Remarks

This article was intended to provide an assessment of current CFD technology for propulsion integration, define areas where further research and development will improve the current technology, and discuss numerical error assessment and control in the use of CFD for installation design. Aircraft design is now considered to be a mature technology (at least for conventional aircraft). Configurations are selected with a hybrid process which uses both test-derived and CFD-derived performance data. CFD is used for design when appropriate tools are available within the time, cost, compuer, etc., constraints of a project. CFD is expected to mature into a technology with an impact on aircraft design similar to that of the wind tunnel. Professor D. B. Spalding has succinctly summarized the current state of affairs: "We have learned how to solve the initial problems, but we are like the early automobile builders who got the engine to work and settled that it should have four wheels, a steering wheel and all those things. But I believe that we are now entering a new phase: Just as the motor car had to develop greatly before everybody could drive safely, economically, and reliably, so the activity of computer simulation of fluid flow and heat transfer has to be developed, streamlined, engineered, made safe, economical, and reliable, so that everybody can use it . . . it is like directing a film where many people must collaborate, where you have organizational as well as technical constraints, but here, in the end, the result must be a work of art."

References

1 Paynter, G. C., "CFD Technology for Propulsion Installation Design—Forecast for the Eighties," ASME Paper No. 82-GT-21.

2 Kutler, P., "A Perspective of Theoretical and Applied Computational Fluid Dynamics," AIAA Paper No. 83-0037, presented at the AIAA 21st *Aerospace Sciences Meeting, Reno,* NV, *Jan. 1983.*

3 Tjonneland, E., Paynter, G., and May, F., "Progress Toward Airframe/Propulsion Integration Design Analysis," Paper No. 8, AGARD Flight Mechanics Panel Symposium on Sustained Supersonic Cruise and Maneuver, Brussels, Oct. 1983.

4 Levine, R. D., "Supercomputers," *Scientific American*, Vol. 246, No. 2, Jan. 1982.

5 Shang, J. S., and Scherr, S. J., "Navier-Stokes Solution of the Flow Around a Complete Aircraft," AIAA Paper No. 85-1509, *AIAA 7th Computational Fluid Dynamics Conference Proceedings,* July 1985.

6 Peterson, V. L., and Richey, G. K., private communication, Sept. 1985.

7 Paynter, G. C., and Chen, H. C., "Progress Toward the Analysis of Supersonic Inlet Flows," AIAA Paper No. 83-1371, June 1983.

8 Kern, P. R. A., and Hopcroft, R. G., "Progress Toward the Analysis of Complex Propulsion Installation Flow Phenomena," AIAA Paper No. 83-1367, June 1983.

9 Kern, P. R. A., Paynter, G. C., Clark, D. R., and Dvorak, F. A., "A Review of the Status of Computational Fluid Dynamics (CFD) Application to the Installation and Integration of Turbofans and Turboprops in Subsonic Aircraft," AIAA Paper No. 84-1333, June 1983.

10 Paynter, G. C., Koncsek, J. L., Turczeniuk, B., and Dvorak, F. A., "Extension of CFD Technology Used to Design the JVX Inlet," AIAA Paper No. 85-1215, July 1985.

11 Paynter, G. C., Vaidyanathan, T. S., Maskew, B., and Dvorak, F. A., "Experience With Hybrid Aerodynamic Methods," AIAA Paper No. 83-1819, July 1983.

12 Ritchie, M. L., "The Research and Development Methods of Wilbur and Orville Wright," *Astronautics and Aeronautics*, July/Aug. 1978, pp. 56–67.

13 Miranda, L. R., "A Perspective of Computational Aerodynamics From the Viewpoint of Airplane Design Applications," AIAA Paper No. 82-0018, Jan. 1982.

14 Graves, R. A., "Computational Fluid Dynamics: the Coming Revolution," *Astronautics and Aeronautics*, Mar. 1982, pp. 20–62.

15 Chapman, D. R., "Computational Aerodynamics Development and Outlook," *AIAA J.*, Vol. 17, pp. 1293–1313.

16 Bailey, F. R., "NAS: Supercomputing Master Tool for Aeronautics," *Aerospace America*, Jan. 1985, pp. 118–121.

17 Denning, P. J., and Adams, G. B., "Wanted: Supercomputers for Engineering Design," *Aerospace America*, Jan. 1985, pp. 122–123.

18 Connor, R. S., Purdon, D. J., and Wamsley, F. K., "PANAIR Consultation Using Expert System Techniques," AIAA Paper No. 85-4094, presented at the AIAA 3rd Applied Aerodynamics Conference, Oct. 14–16, 1985.

19 MacCormack, R. W., "Current Status of Numerical Solutions of the Navier-Stokes Equations," AIAA Paper No. 85-0032, Jan. 1985.

20 Shang, J. S., "An Assessment of Numerical Solutions of the Compressible Navier-Stokes Equations," *AIAA J. of Aircraft*, Vol. 22, No. 5, May 1985.

21 Forester, C. K., "Private Communication on Implementing Implicit Methods," Oct. 1985.

22 Thompson, J. F., "A Survey of Grid Generation Techniques in Computational Fluid Dynamics," AIAA Paper No. 83-0447, Jan. 1983.

23 Campbell, A. F., and Forester, C. K., "Evaluation of a Method for Analyzing the Aperture Region of Two-Dimensional External Compression Inlets," AIAA Paper No. 85-3072, Oct. 1985.

24 Forester, C. K., "Navier-Stokes and Experimental Modeling of Blunt Based Rocket Nozzle Flows," presented at IUTAM meeting on Turbulent Shear Layer/Shock Wave Interactions, Sept. 1985.

25 Hall, M. G., "Computational Fluid Dynamics—A Revolutionary Force in Aerodynamics," AIAA Paper No. 81-1014.

26 Bradshaw, P., "Review—Complex Turbulent Flows," ASME *J. of Fluids Engineering*, June 1975, pp. 146–154.

Engine Control Reliability and Durability Improvement Through Accelerated Mission Environmental Testing

W. J. Davies
Pratt & Whitney,
Engineering Division,
West Palm Beach, FL

R. W. Vizzini
Naval Air Propulsion Center,
Mechanical Components and
Controls Division,
Science and Technology Group,
Trenton, NJ

The integration of aircraft control systems for future weapon systems will require the engine control system to meet the mission reliability of the flight control system. This will be accomplished through system redundancy and verified by accelerated environmental testing. Combined environment reliability testing (CERT) will assure control system structural integrity and reliability growth of engine-mounted digital electronic controls. Pratt & Whitney, under contract to the U.S. Navy, has recently completed a 10,000-hr CERT program. Dual full authority digital electronic controls (FADEC), connected by a fiber optic data link, were subjected to environmental tests simulating a composite F-14 mission profile. The FADEC units were also exposed to periodic high vibration levels which would be experienced after foreign object damage and salt spray testing to simulate aircraft carrier environment. The test results are reported herein providing insight not only into the reliability and durability of digital electronic controls but also into the equipment and procedures required for testing of future military and commercial engine control systems.

Introduction

The use of electronic controls on aircraft gas turbine engines has demonstrated substantial cost, weight, operability, and reliability benefits in both commercial and military applications. The last of these benefits, reliability of electronic engine controls, is the subject of this paper.

In 1979 Pratt & Whitney (P&W), under U.S. Navy contract, successfully engine tested a dual full authority digital electronic control system (FADEC) [1]. Subsequent to this engine testing P&W presented the Navy with a test plan to subject the two FADEC units to an extensive combined environment reliability test (CERT). This test would subject the units to an operating environment typical of actual Navy operating conditions.

CERT is a tool which has been successfully applied by P&W to establish the reliability of digital electronic engine controls on the F100-220 (F-15), JT9D-DR4 (B-767), PW2037 (B-757, C-17), and PW4000 (B-747 and A-320). This type of testing provides the ability to advance FADECs through the "infant mortality" stage where control failures might be caused by design flaws or inadequate component specification. In addition, extended testing of FADEC establishes the trends which may be used to project the reliability at maturity (i.e., after 1,000,000 engine flight hours).

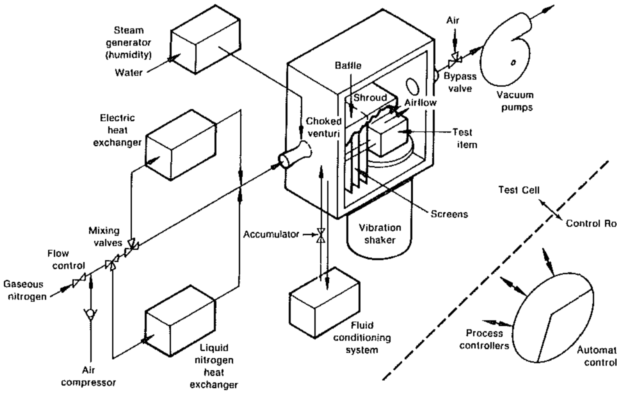


Fig. 1 CERT facility block diagram

The CERT facility, FADEC hardware, and results of 10,000 hours of testing on the Navy FADEC units are discussed below.

CERT Facility Description

The P&W CERT facility, capable of simulating a wide environment range, represents the state of the art in test capability. The facility's primary purpose is to simulate the nacelle environment of high-performance aircraft over various mission profiles to estimate the reliability of gas turbine engine control components (Fig. 1). The facility's flexible design allows a variety of components to be tested to establish their

---

Contributed by the Gas Turbine Division of The American Society of Mechanical Engineers and presented at the 31st International Gas Turbine Conference and Exhibit, Düsseldorf, Federal Republic of Germany, June 8–12, 1986. Manuscript received at ASME Headquarters January 10, 1986. Paper No. 86-GT-52.

Table 1 CERT facility variables

	Range	Rate of change
Chamber temperature	−65° to +475°F	3°F/s
Altitude	Sea level to 67,000 ft	2400 ft/s
Humidity	0 to 98 percent	100 percent/s
FADEC cooling fluid	−65° to +250°F	6°F/s
Vibration: 150 pound test item up to 50 g and 3 kHz		

reliability and performance under various environmental conditions encountered in flight.

The CERT facility features a multi-environment test chamber designed to subject engine control system components to the stress imposed on flight hardware under actual operating conditions. The facility has the capability to simultaneously vary ambient pressure, temperature, humidity, unit vibration, and cooling medium temperature in accordance with the desired duty cycle.

The environmental test chamber, housed in the CERT facility, is 3 by 3 by 4 ft. The airflow and cooling fluid conditioning, vacuum pumps, steam generator, vibration shaker, and environmental chamber are under shelter outside the control room. Capabilities of the P&W facility are listed in Table 1.

These capabilities exceed present-day military and commercial aircraft stress levels allowing accelerated testing of jet engine components and extreme testing of most other test items, if required. The CERT facility can produce all associated environment levels required by MIL-STD-810 and is compatible with MIL-STD-781.

FADEC System Description

The FADEC system used for the CERT was composed of four separate items: (1) two FADEC units, (2) FADEC breadboard, (3) verifier, and (4) optic data link. The first three of these items, including one FADEC, are shown in Fig. 2.

The engine-mounted FADEC electronic units perform all sensor signal conditioning, computation, and output signal processing functions with the capability of controlling a joint technology demonstrator engine (JTDE) engine. The control unit contains the primary system digital processor, input and output circuitry, pressure transducers, and power supply (Fig. 3). The FADEC system operates in a master/standby mode. That is, the primary FADEC normally controls the engine while the secondary monitors the system. In the event of a primary FADEC failure (see [2] for failure logic details) control transfers to the secondary FADEC.

The environment within the hardware package for the FADEC electronic units was designed to protect the digital electronic circuits while operating within the service environments projected for Navy fighter aircraft. The external configuration of the control is shown in Fig. 4. This photograph shows the pin-fin heat exchanger, fuel inlet and outlet ports, vibration isolators in a four-point mounting arrangement, the optic transmitter and receiver ports, electrical connectors, and the pressure relief valve.

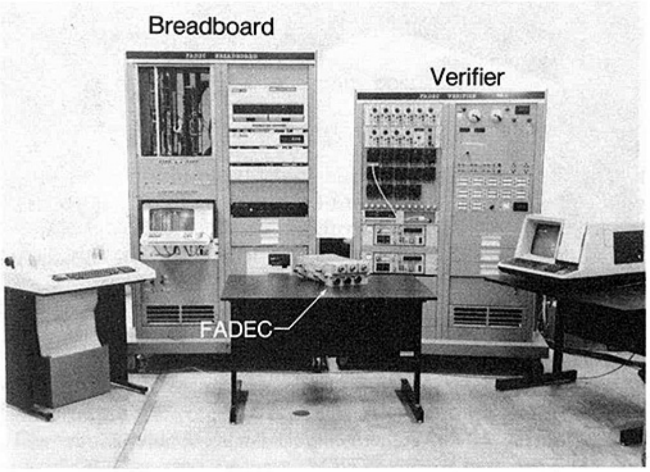


Fig. 2 FADEC control system components

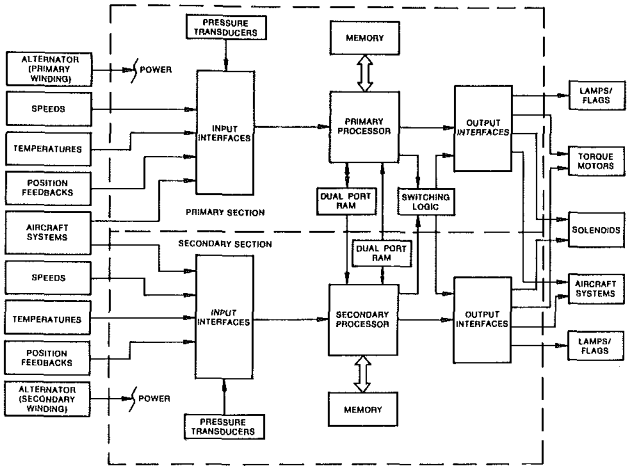


Fig. 3 FADEC block diagram

During the 1979 FADEC engine test the breadboard control acted as the secondary engine control. The breadboard processor and its sensor and output interface circuits are identical to those which are used in the engine-mounted controls. Therefore, they operate to the same accuracies and with the same dynamic response rate. For the FADEC/CERT Program

Nomenclature

AMSAA = Army Material Systems Analysis Agency
ATDE = advanced technology demonstrator engine
B-747 = Boeing 747 aircraft
B-757 = Boeing 757 aircraft
B-767 = Boeing 767 aircraft
C-17 = transport
CERT = combined environment reliability test
CPU = central processing unit
$C(t)$ = cumulative failure rate at time t
EMI = electromagnetic interference
F-14A = Navy fighter
F-15 = Air Force fighter
FADEC = full authority digital electronic control
FOD = foreign object damage
g = force of gravity
Hz = Hertz (cycles per second)
k = kilo
mil = thousandths of an inch
MIL-STD = military standard
NAPC = Naval Air Propulsion Center
NASA = National Aeronautics and Space Administration
NHPP = nonhomogeneous Poisson process
$N(t)$ = cumulative failures by time t
P&W = Pratt & Whitney
RAM = random access memory
VDC = volts direct current
$\lambda, \beta, \alpha, \gamma$ = parameters of growth model
$\rho(t)$ = intensity function

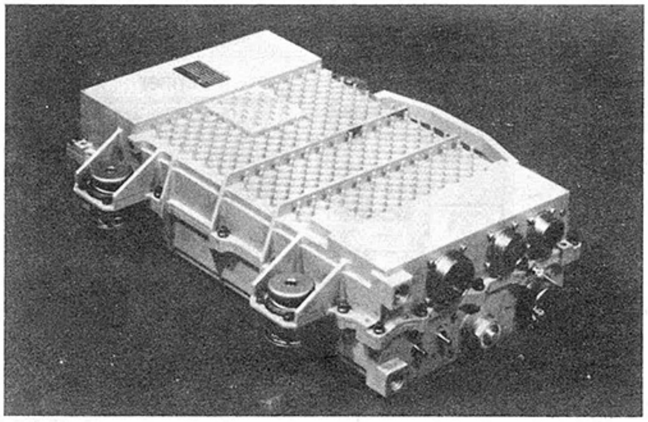


Fig. 4 Full authority digital electronic control

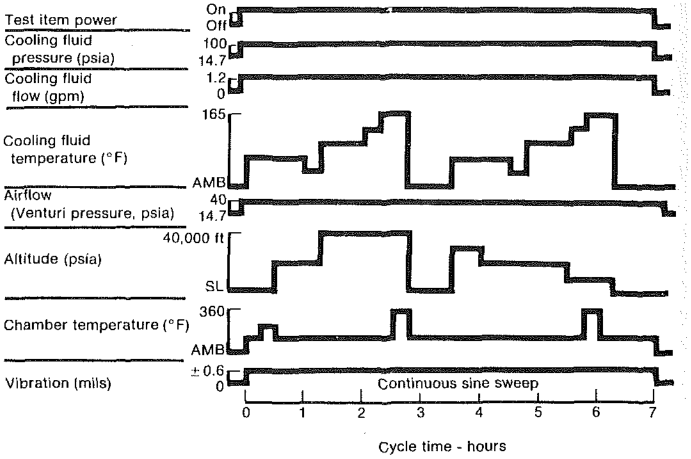


Fig. 5 FADEC test cycle

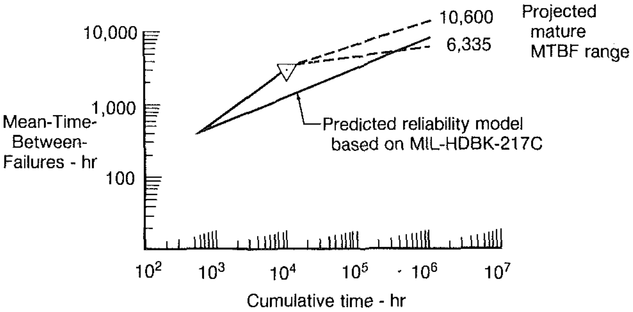


Fig. 6 FADEC reliability projections

two engine-type units were installed in the test chamber, one primary and the other a secondary control. The breadboard control was modified to act as an interface to these two operating units. In this capacity it monitored both controls for operational and fault status information.

The verifier provides simulated engine sensor and actuator signals, with manual selection, control, and display of various parameters. The rig is designed to simulate the characteristics of a particular engine through manual selection panel on the verifier. The verifier is capable of providing the following outputs:

- Electrical resolvers
- Speeds, pressures, and temperatures
- Discrete 0 or 28 vdc signal

The torque motors are simulated by resistive loads. The solenoids are simulated by discrete lights. Discrete inputs are simulated by panel-mounted switches.

Fiber optic cables were used to communicate between the FADEC primary and secondary controls through a dual-port random access memory (RAM) serial communication link. The optical fibers were in a cable constructed of a multiple glass fiber bundle in a teflon tube, wrapped with a Kevlar braid, and enclosed in a teflon casing.

Pre-FADEC/CERT Testing

In addition to the sea level and altitude F401 engine tests, conducted during the original FADEC Program, the electronic units also underwent a series of environmental tests to assess functional and environmental capability of the units. The environmental tests were conducted at test conditions similar to the conditions, procedures, and durations specified by MIL-E-5007D. These tests consisted of the following: torque motor and solenoid compatibility test, alternator simulator compatibility test, thermal mapping test, high-temperature endurance test, low-temperature endurance test, vibration test, and electromagnetic compatibility test. The results of these tests, which were not conducted in the FADEC/CERT Program, may be found in the FADEC Final Report [1].

FADEC/CERT Test Cycle

The intent of the test cycle is to subject two FADEC, engine-mounted, prototype controls to a simulated engine environment using the CERT facility. The environmental sequencing chosen is based on a composite of seven typical F-14A missions:

- Field carrier landing practice
- Familiarization
- Fighter intercept
- Instrument
- Air combat maneuver
- Conventional weapons delivery
- Post main check flight

There are five environmental variables: vibration, pressure (altitude), humidity, liquid coolant temperature, and air temperature. Figure 5 shows the FADEC test cycle. These environmental values are derived from information contained in [3-8]. The vibration cycle for the FADEC CERT was established at +/−0.6 mil from 40 to 260 Hz, sinusoidal sweep, throughout the test. This value was established through FADEC vibration measurements on the F401 engine [1].

In addition to the basic test cycle two additional tests were run: foreign object damage (FOD) vibration simulation and a salt fog test. The intent of the FOD test was to simulate the vibration which might occur in the event of FOD to the engine. Once every 1000 hr the FADEC unit vibration was increased to the level of +/−6 mils from 40 to 260 Hz for a period of 1 hr to simulate the time period required for the aircraft to return to base. One FADEC unit was run to MIL-STD-810C, method 509.1/Salt Fog. This 48-hr salt fog test was run at the completion of the 10,000 hr CERT to establish the resistance of the FADEC units to the effects of carrier operating environment. After the salt fog test the FADEC was recalibrated to verify no change in its operating condition prior to final teardown and inspection of the FADEC units.

Testing History

During the 10,000-hr period of performance for the FADEC CERT there were three accountable failures. The first of these occurred on the secondary FADEC after 758 hr of testing. The serial digital diagnostic output to the breadboard became intermittent as cooling fluid temperature was raised above 120°F. Normal operation returned as fluid temperature decreased. The central processing unit (CPU) diagnostic interface was removed and replaced and satisfactory operation restored. The second failure occurred at 760 hr of testing when a CPU chip in the primary FADEC failed. The chip was removed and replaced and normal operation restored. This second failure may be due to troubleshooting techniques employed to establish the cause of the first failure. The final failure occurred in the secondary control at 5800 hr of testing. This unit exhibited an intermittent fault in the resolver circuitry. This fault was determined to be a faulty chip in the resolver conversion circuitry.

When the primary CPU failed during the FADEC CERT, control was immediately transferred to the secondary FADEC. Although the two secondary failures did not interfere with system operation they were flagged for maintenance action through the diagnostic system. None of the failures occurred during the high-vibration FOD test periods.

The salt fog test was conducted after 10,000 hr of testing. The FADEC primary was calibrated before and after the salt fog test and exhibited no change in operation. During the final visual inspection of the FADEC no indication of salt fog intrusion into the test unit was observed. The final visual inspection of both FADEC units established that they were in excellent condition both externally and internally.

Reliability Model

Based on the results of [9], the AMSAA/Duane model was selected to analyze the FADEC/CERT results. Following is a brief description of the AMSAA/Duane model. Details and comparisons with other reliability models may be found in [9].

Model Descriptions. The Duane model is a curve of cumulative failure rate as a function of cumulative test time in the development program. $C(t) = \gamma t^{\alpha}$ where $C(t)$ is cumulative failure rate at time t; γ is the scale parameter; α is the growth parameter.

Taking the log of both sides of the equation, the model becomes linear on log-log paper with slope α.

$$\log(C(t)) = \log\gamma + \alpha\log(t)$$

The AMSAA model is a statistical model of the same form as the Duane model. It is different in that it is modeling the reliability within a test phase as a function of time accumulated in the test phase. $N(t) = \lambda t^{\beta}$; $N(t)$ is cumulative failure by time t; λ is the scale parameter; β is the growth parameter.

Model Assumptions. The Duane model assumes that the cumulative failure rate grows along the curve γt^{α} as a function of development time.

The AMSAA model assumes that the failures are from a nonhomogeneous Poisson process (NHPP) with a Weibull intensity function. The intensity function is as follows

$$\rho(t) = \lambda\beta t^{\beta-1}$$

It is further assumed that if no further effort is made to improve the reliability, the failure rate will remain constant (time between failures will follow an exponential distribution). The following assumptions are a result of the NHPP assumption:

- $N(O) = O$ (no failure can occur without some test time).
- The probability of a failure at time t is not affected by previous failures
- Simultaneous failures are not possible.

FADEC/CERT Test Results

The three failures which occurred during the CERT test of the FADEC system were used to establish a curve projecting reliability growth. These failures are those which resulted from design, quality, or procedure deficiencies. The slope of the curve showing reliability growth was determined from the AMSAA/Duane growth model [9] and is based on all events which occurred after the start of the CERT test. This slope was used based on the assumption that the definition of a failure changes the failure rate, but the slope, representing the growth rate, is affected very little by the failure identification. The results of this calculation indicate that the instantaneous mean time between failures (MTBF) ranges between 6335 and 10,600 hr (Fig. 6). This variation is due to the selection of the slope of the projection line through the last failure or through the average of the three failures. These instantaneous MTBF values compare favorably with the calculated hardware reliability from MIL-HANDBOOK-217C, which predicts a mature reliability of 8000 hr. It should be noted that the instantaneous MTBF is the reliability value projected for the FADEC units if they were to be released for production with no further environmental testing or hardware changes. Details of the reliability calculation may be found in the FADEC/CERT Final Report [10].

Summary

- CERT has been demonstrated to be an effective tool to establish the long-term durability of digital electronic engine controls, such as FADEC.

The FADEC design is projected to meet the required reliability when operating in a typical Navy operational environment.

- In the event of a component failure in a FADEC unit, the system is capable of continued operation.
- Troubleshooting procedures must be established and followed to preclude the inadvertent introduction of failures to a FADEC unit.

References

1 Lenox, T. G., "Full Authority Digital Electronic Control Phase II Final Report," Pratt & Whitney, FR-12378, Mar. 1980.
2 Lenox, T. G., "Full Authority Digital Electronic Control Phase I Final Report," Pratt & Whitney, Vols. I and II FR8652, June 1978.
3 Environmental sequence, power condition percent time F14A/TF30 DUTY CYCLE, NAVAIRSYSCOM, AIR 53602, dated Mar. 11, 1977.
4 Fuel Temperature – Model Specification N-6191, FF30-P-412A.
5 Vibration – Model Specification N-6191, TF30-P-412A.
6 Fuel Rate – FR-11910-FADEC open loop bench tests, dated Oct. 16, 1979.
7 Air Temperature – Grumman Aerospace Corporation data, memo from D. R. Smith, dated Mar. 3, 1973.
8 Relative Humidity – established by altitude, effective only at sea level conditions.
9 McGlone, M. E., "Reliability-Growth Assessment, Prediction, and Control For Electronic Engine Control," AFWAL-TR-84-2024, Apr. 1984.
10 Davies, W. J., "FADEC/CERT Draft Final Report," NAPC-PE-130C, Sept. 1985.

Development of HIDEC Adaptive Engine Control Systems

R. J. Landy

W. A. Yonke

McDonnell Aircraft Company,
McDonnell Douglas Corporation,
St. Louis, MO

J. F. Stewart

NASA—Ames/Dryden,
Edwards Air Force Base, CA

The NASA Ames/Dryden Flight Research Facility is sponsoring a flight research program designated Highly Integrated Digital Electronic Control (HIDEC), whose purpose is to develop integrated flight-propulsion control modes and evaluate their benefits in flight on NASA F-15 test aircraft. The Adaptive Engine Control System (ADECS I) is one phase of the HIDEC program. ADECS I involves uptrimming the P&W Engine Model Derivative (EMD) PW1128 engines to operate at higher engine pressure ratios (EPR) and produce more thrust. In a follow-on phase, called ADECS II, a constant thrust mode will be developed which will significantly reduce turbine operating temperatures and improve thrust specific fuel consumption. A performance seeking control mode is scheduled to be developed. This mode features an onboard model of the engine that will be updated to reflect actual engine performance, accounting for deterioration and manufacturing differences. The onboard engine model, together with inlet and nozzle models, are used to determine optimum control settings for the engine, inlet, and nozzle that will maximize thrust at power settings of intermediate and above and minimize fuel flow at cruise. The HIDEC program phases are described in this paper with particular emphasis on the ADECS I system and its expected performance benefits. The ADECS II and performance seeking control concepts and the plans for implementing these modes in a flight demonstration test aircraft are also described. The potential payoffs for these HIDEC modes as well as other integrated control modes are also discussed.

Introduction

By utilizing the high throughput and large memory of current onboard computers, along with high-speed communication multiplex buses, it will be possible to implement a wide range of integrated flight/propulsion control modes on board future aircraft. Such modes will increase total weapon system effectiveness without significant weight penalties. To support such efforts, NASA has instituted the HIDEC program for the development and flight test of technologies required to realize the full potential of integrated controls.

The HIDEC program is divided into five phases. Phase I, which was completed in 1985, involved flight testing the Digital Electronic Flight Control System (DEFCS) in the NASA F-15 airplane. The DEFCS consists of a higher order language digital flight control computer (DFCC) and two modified control augmentation system (CAS) analog computers used for interfacing with sensors and actuators. As part of Phase I, the DEFCS software has been programmed in PASCAL to provide a "flutter exciter" function that enables the pilot to select precise frequency sweep and dwell inputs to the horizontal tails. This feature allows the acquisition of data used to obtain mathematical models of flight control components and aircraft rigid and structural modes.

Figure 1 shows the schedule for Phases II, III, IV, and V of the HIDEC program. Phase II, designated the Adaptive Engine Control System I (ADECS-I) [1], consists of the design, implementation, and flight test of an integrated flight/propulsion control mode. This mode uses flight control information to determine the amount of uptrim in engine pressure ratio (EPR) necessary to improve engine performance.

Phase III will consist of the development and evaluation of trajectory guidance algorithms. These algorithms will minimize either fuel or time to fly to prescribed terminal conditions. The algorithms will be tested with and without the ADECS features.

Additional ADECS modes and enhancements will be developed and tested during Phase IV, which will be called ADECS-II.

Phase V, performance seeking control, involves the integration of the aircraft inlets with the engine and flight control systems to optimize total propulsion system performance.

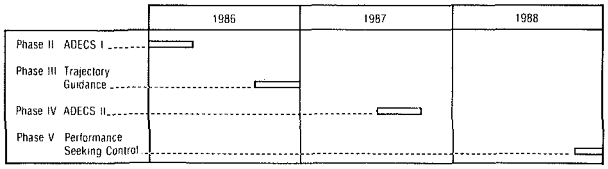


Fig. 1 HIDEC flight test schedule

Contributed by the Gas Turbine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the 31st International Gas Turbine Conference and Exhibit, Düsseldorf, Federal Republic of Germany, June 8–12, 1986. Manuscript received at ASME Headquarters March 7, 1986. Paper No. 86-GT-252.

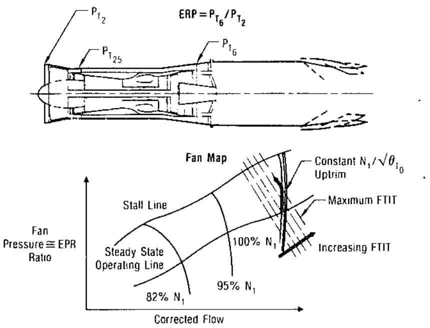


Fig. 2 Adaptive engine control system; EPR uptrim

Phase II – Adaptive Engine Control System I (ADECS-I). The ADECS-I phase will demonstrate improved engine performance based on using flight control information which had not been previously available for engine control operation. Information exchange is facilitated by the use of the airplane digital systems: DEFCS for flight control and digital electronic engine control (DEEC) for engine control. In the ADECS-I uptrim mode, additional thrust is obtained at power settings of intermediate and above by decreasing nozzle throat area to increase EPR.

ADECS-I Mode Description. A typical fan map (Fig. 2) illustrates the EPR uptrim concept. Uptrim is initiated by closing the nozzle throat area slightly, which increases fan pressure ratio (FPR = PT25/PT2) and engine pressure ratio (EPR = PT6/PT2). At the high corrected fan speeds typical of subsonic intermediate power, EPR uptrim occurs at near constant airflow.

EPR is allowed to increase until the fan turbine inlet temperature (FTIT) limit is approached. At the FTIT limit, additional thrust can still be obtained by uptrimming EPR at constant FTIT. However, once the FTIT limit is reached, airflow begins to decrease, diminishing the effect of higher pressure ratio. In most cases, after the FTIT limit is reached, only small improvements in thrust are possible. Because the normal FTIT limit is not exceeded during EPR uptrim, this mode is expected to have minimal adverse impact on engine life.

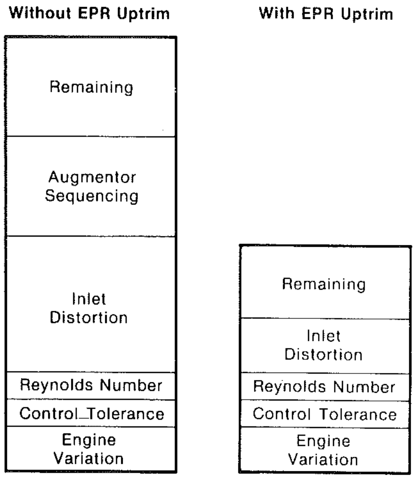


Fig. 3 Stall margin available for EPR uptrim

When EPR is uptrimmed, thrust is increased at the expense of reduced fan stall margin. Substantial stall margin is normally built into the engine control schedules to accommodate the inlet distortion produced at high angles of attack or sideslip. In the ADECS-I mode, some of the stall margin reserved for extreme inlet distortion is used to increase thrust in regions of low distortion. At flight conditions which produce high inlet distortion (i.e., maneuvering at large angle of attack or sideslip angle), the amount of EPR uptrim is reduced to restore stall margin.

A typical stability audit with and without EPR uptrim is

Nomenclature

ADC = air data computer
ADECS = adaptive engine control system
AOA = angle of attack
BIT = built-in test
BUC = hydromechanical back-up engine control
CAS = control augmentation system
CC = central computer
CP = cockpit
DEEC = digital electronic engine control
DEFCS = digital electronic flight control system
DFCC = digital flight control computer
EMD = engine model derivative
EPR = engine pressure ratio
FPR = fan pressure ratio
FTIT = fan turbine inlet temperature
HIDEC = highly integrated digital electronic control
IGV = inlet guide vanes
IFIM = *in-flight integrity management*
INS = inertial navigation set
KA2 = fan distortion
F_{lat} = lateral stick force
F_{long} = longitudinal stick force
F_{rud} = rudder pedal force
M = Mach
MUX = multiplex
NASA = National Aeronautics and Space Administration
n_y = lateral acceleration
n_z = normal acceleration
N1 = engine fan speed
N1C2 = engine fan speed, corrected to engine inlet
$N1/\sqrt{\theta_{to}}$ = engine fan speed, corrected
NCI = navigation control indicator
p = roll rate
P = pitch CAS computer
PASCOT = programmable asynchronous serial communication translator
PLA = power lever angle
PSC = performance seeking controls
PT2 = fan inlet total pressure
PT25 = fan discharge total pressure
PT6 = turbine discharge total pressure
P&W = Pratt & Whitney
q = pitch rate
$\dot{q}$ = pitch rate change
r = yaw rate
R/Y = roll/yaw CAS computer
TT2 = total temperature engine face
TF/TA = terrain following terrain avoidance
TSFC = thrust specific fuel consumption
WAC = engine airflow, corrected
ΔWAC2 = engine airflow, corrected to engine inlet
α = angle of attack
α_p = angle of attack, predicted
β = sideslip angle
β_p = sideslip angle, predicted
δ_s = stabilator deflection
δ_r = rudder deflection
θ_{to} = free-stream total temperature/518.7

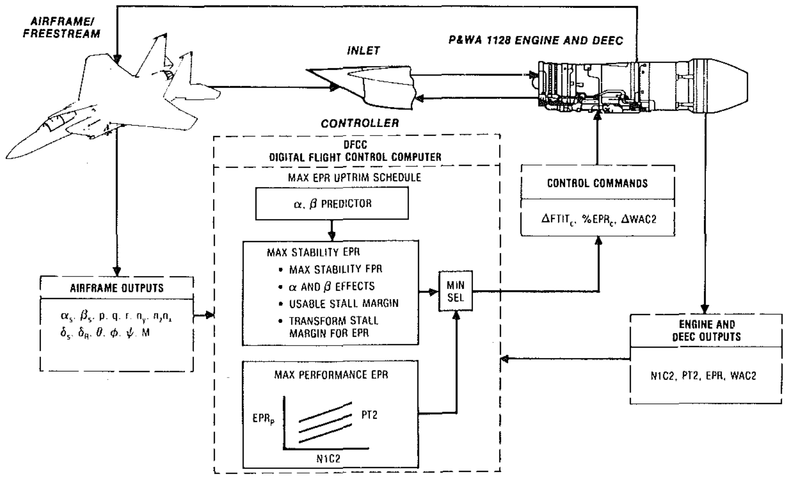


Fig. 4 Adaptive engine control system

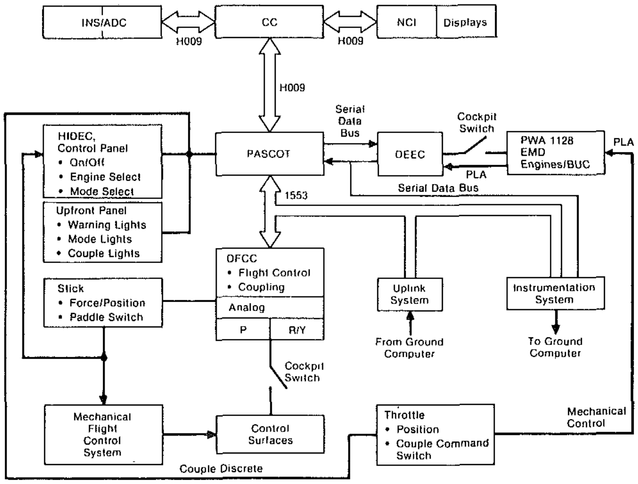


Fig. 5 HIDEC system architecture; ADECS modes

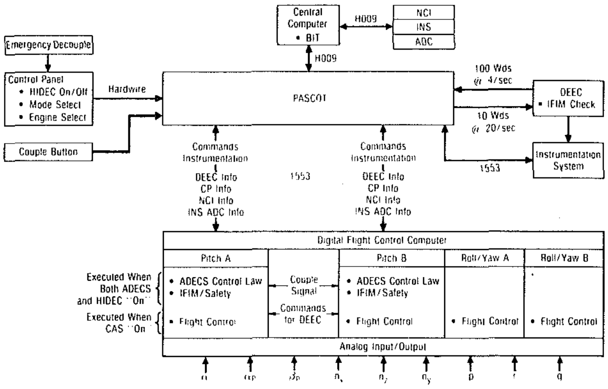


Fig. 6 ADECS architecture

shown in Fig. 3. Additional information on the stability audits and definition of the amount of stall margin available are found in [1, 2].

EPR uptrim is implemented as shown in Fig. 4. The ADECS-I control laws are incorporated into the digital flight control computer (DFCC). When predicted angle of attack (AOA) and sideslip angles are moderate, the controller issues an EPR command to the engine, causing the engine to operate closer to the stall line.

The controller will sense airframe pitch, roll and yaw rates, and normal, lateral, and axial accelerations to predict AOA and sideslip angles in the immediate future. This prediction is required in order to provide some lead to compensate for computation delays, transport lags, and response time for the nozzle actuator. As these predicted angles become large, the controller decreases the uptrim signal in a well-behaved transient, without stalling the engine.

The details of the EPR uptrim logic implemented in the DFCC are described in [2]. The amount of EPR uptrim is based on the maximum amount of EPR useful for increased performance, limited by engine stall considerations. The implementation of this control law requires information from the airframe (angle of attack, sideslip, and Mach number) and the engine (airflow, fan speed, and pressure at the engine face).

ADECS-I Mode Implementation. The HIDEC architecture for the ADECS-I mode is shown in Fig. 5. The programmable asynchronous communication translator (PASCOT) acts as a multiplex bus interface unit to transfer information among the aircraft sensors and central computer (H009 bus) to the digital flight control computer (1553 bus) and the digital electronic engine controller (serial data bus). It also acts as an interface unit for the HIDEC control and display panels in the cockpit.

Figure 6 shows the detailed architecture of the ADECS-I system, including location of the ADECS-I control laws and in-flight integrity management (IFIM) software, and gives the data rates for communication among the key ADECS system computers. The control laws will be computed in both the Pitch A and B channels within the digital flight control computer (DFCC) when the HIDEC modes are selected. The software for some IFIM tests and related safety tests for the ADECS-I mode are also programmed in the DFCC. Other IFIM tests and ADECS built-in test (BIT) are programmed in the central computer. The DEEC also contains unique IFIM tests for the engine.

The IFIM tests determine whether the propulsion and flight control hardware are operating and communicating properly. These tests consist of self-checks of the DEFCS, PASCOT, CC and DEEC and are performed continuously when the HIDEC mode is selected. Other checks verify the operation of H009 multiplex bus, 1553 multiplex bus, and serial data bus to the DEEC by monitoring the return of wraparound words sent from the digital flight control computer. If all integrity management tests are passed successfully, the EPR uptrim command is issued by the flight control computer through the PASCOT to the engine control, when so commanded by the pilot.

The crew station configuration is shown in Fig. 7. The HIDEC control panel is used to engage the HIDEC mode,

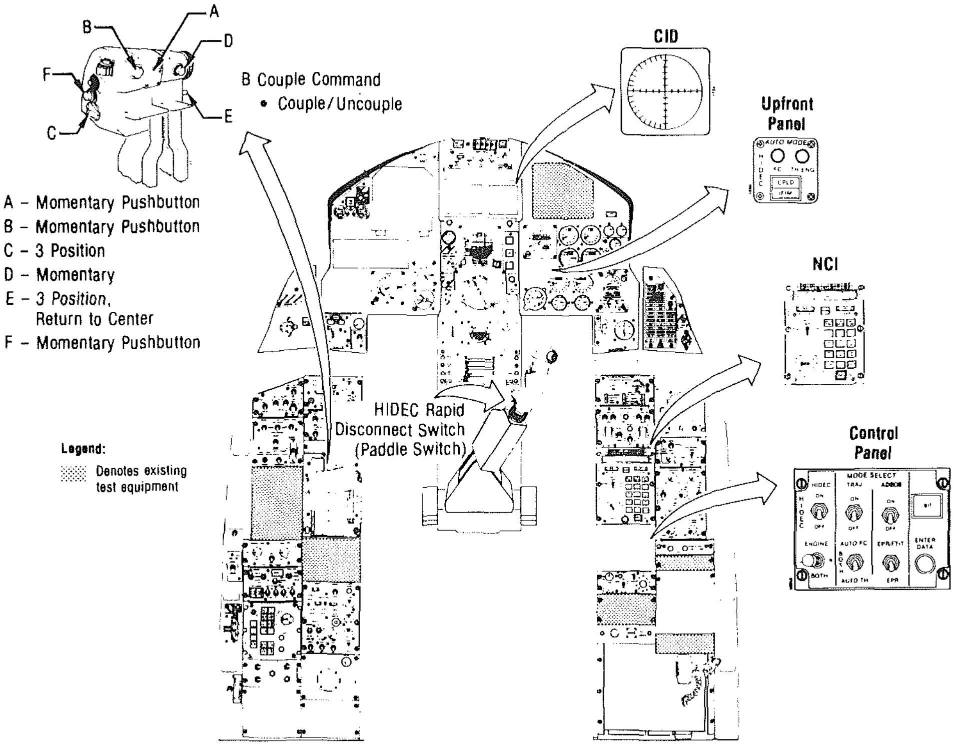


Fig. 7 *HIDEC crew station orientation*

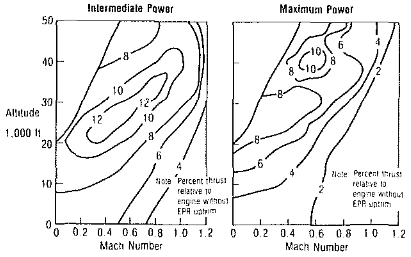


Fig. 8 ADECS mode thrust increase; PW1128 engine; Standard Day

select the desired engines to receive the commands, and select the ADECS-I mode and its submodes (EPR uptrim or a combination of EPR and FTIT uptrims). The navigation control indicator (NCI) panel is used to initiate BIT tests and select new constants for use in the DFCC logic. This latter feature increases flight test flexibility.

The indicator lights on the upfront panel illuminate when the HIDEC system is coupled or when an IFIM failure exists. Other lights indicate which modes have been selected, i.e., engine control, flight control, or both. The couple command button is on the throttle lever and the emergency disconnect for the system is the paddle switch located on the control stick.

ADECS-I Mode Benefits. The principal benefit of the ADECS-I mode is increased thrust, primarily in subsonic flight. Figure 8 shows the thrust improvement available at intermediate and maximum power across the flight envelope. The thrust improvement falls off at the lower altitudes and higher Mach numbers as the engine FTIT limit is reached since further EPR uptrim results in less airflow.

Current estimates of ADECS-I performance improvements are based on the actual control logic that will be programmed in the flight control computer. These data supersede the projected benefits described in [3-5].

At intermediate power the EPR uptrim is limited by the physical limits of the exhaust nozzle. At some maximum power flight conditions, additional augmentor segments can be lit because EPR uptrim increases pressure in the augmentor duct and improves augmentor efficiency.

At intermediate power, EPR uptrim increases thrust specific fuel consumption (TSFC) by 1-3 percent. This is because additional fuel is required to enable the engine to pump the same airflow against a higher back pressure. However, at maximum afterburning power, EPR uptrim reduces the TSFC, as shown in Fig. 9. The TSFC reduction is due to higher efficiency of the augmentor and fan.

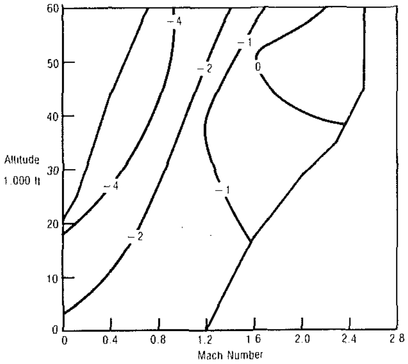


Fig. 9 ADECS mode TSFC decrease; PW1128 engine; Standard Day; maximum power

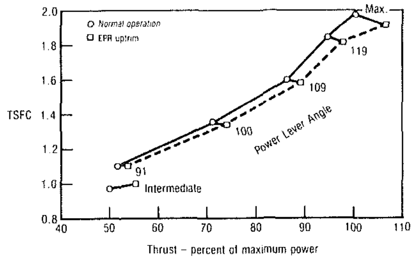


Fig. 10 Potential reduction in TSFC; Mach 0.9; 40,000 ft; Standard Day

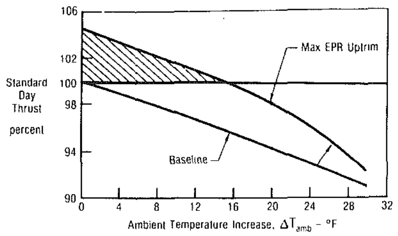


Fig. 11 Hot day performance improvement; Mach 0.6; 20,000 ft; PW1128 engine; maximum power

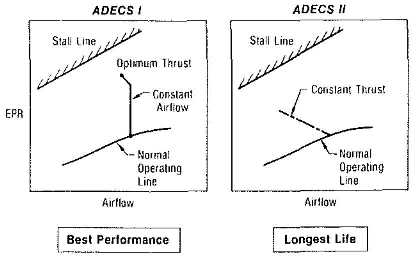


Fig. 12 EPR uptrim modes

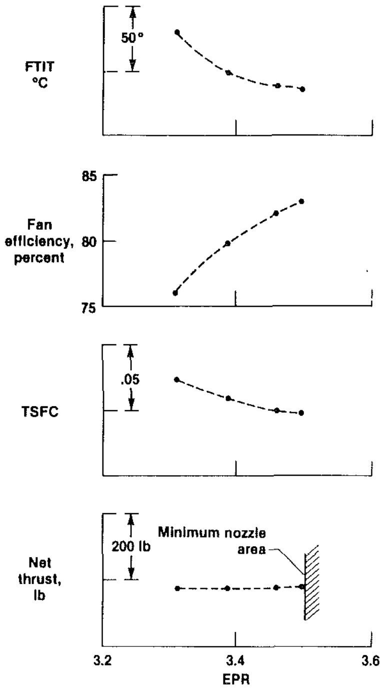


Fig. 13 ADECS constant-thrust mode uptrim; Mach 0.6; 20,000 ft; intermediate power

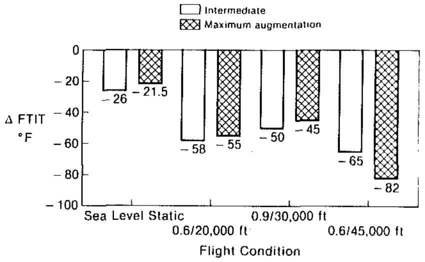


Fig. 14 ADECS-II—FTIT reduction with constant thrust mode

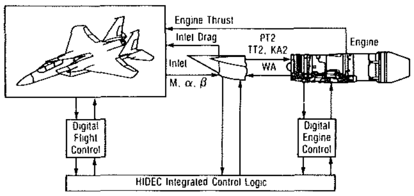


Fig. 15 HIDEC system with inlet control added

The increased thrust available from EPR uptrim can be used to reduce intercept time, increase acceleration, and increase turn capability in air-to-air combat. Figure 10 shows that EPR uptrim can also be used to produce a constant thrust at a lower specific fuel consumption during augmented operation, for a potential benefit in supersonic cruise. This is primarily due to improved fan and augmentor efficiency. For example, in the PW1128 engine the fan operating line is nominally lower than the locus of peak efficiency, due to stability considerations. Uptrim of the fan pressure ratio increases fan efficiencies, and reduces fuel consumption. This reduced fuel consumption can be used to increase combat range or to increase fuel available for combat.

One other important benefit to the ADECS mode is the capability of maintaining Standard Day thrust levels during warm day operations by offsetting the thrust lapse of the engine. Figure 11 shows that, by utilizing the EPR uptrim capability of the ADECS mode, Standard Day thrust levels can be maintained up to ambient temperatures that are 15°F warmer.

Phase IV – ADECS-II Mode. ADECS-II is a constant-thrust mode that will improve the engine TSFC and will increase engine life by reducing turbine temperature. This mode increases engine pressure ratio while reducing engine airflow to maintain constant thrust for a given power setting. This constant-thrust mode is illustrated in Fig. 12 and a typical reduction in turbine temperature and fuel consumption is shown in Fig. 13. The near 3.0 percent improvement in TSFC is significant and can be translated into increases in aircraft range. The reduction of 50° in fan turbine inlet temperature will increase the life of the engine. The turbine temperature reduction at other selected conditions is shown in Fig. 14.

Phase V – Performance Seeking Control. The integration of the propulsion and flight control system provides many new options for meeting new mission requirements. The next logical element to be integrated is the inlet control system. On the F-15, the variable geometry inlet not only determines the quantity and quality of the air delivered to the engine, but also affects the lift, drag, and stability of the airplane.

Figure 15 indicates how the inlet might be integrated by means of HIDEC logic. The inlet variables would be commanded to positions that produce the highest net propulsive force by maximizing pressure recovery and minimizing inlet drag and aircraft trim drag while maintaining inlet/engine compatibility.

The concept of a performance seeking control will be used to maximize either excess thrust or aircraft range. Figure 16 is a block diagram of the logic, using a state-variable model of the engine, together with models of the inlet and exhaust nozzle. The optimization logic will compute commands for the propulsion system variables depending on speed, acceleration, altitude, attitude, and dynamic pressure, as well as the aircraft configuration. These commands are sent to the various propulsion system elements and the resultant measured outputs are used to update the models.

This update logic makes the PSC system capable of accommodating engine degradation, non-Standard-Day atmospheric conditions, and various external stores configurations. The technology will likely be of great importance to supersonic cruise vehicles that are highly sensitive to small changes in propulsion system performance.

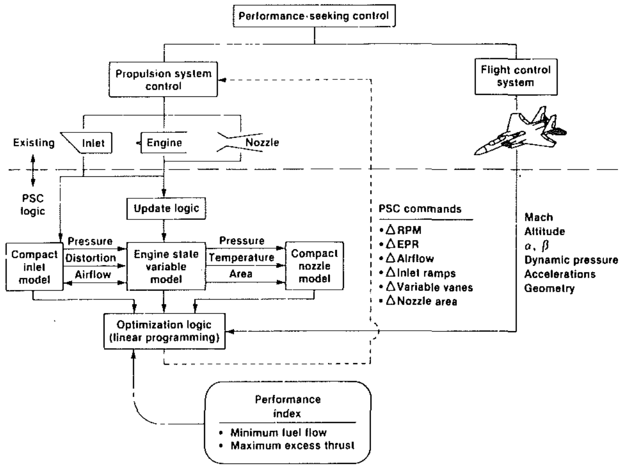


Fig. 16 Performance-seeking control

Since PSC incorporates a model of the engine that is updated to reflect actual engine operation, this same model can be used for other applications as shown in Fig. 17. It can be used to determine engine component efficiencies for trending purposes of identifying when a particular component needs replacement. The engine model can be used to monitor and detect problems in the actual engine control system. The outputs of the onboard engine model can also provide redundancy for a failed sensor. By including onboard models of the airframe with the propulsion system, the entire mission can be planned and optimized. Finally, total aircraft performance can be optimized accommodating for variations in gross weight, store configurations, ambient temperature, and engine health.

The Future of Integrated Flight/Propulsion Controls. Significant benefits for future aircraft can be gained by integrating the flight and propulsion controls. To summarize, candidate integrated control modes and their objectives are as follows:

- Stall margin control
  Uptrim for:
    – supersonic dash and for greater acceleration and turn rates
    – decreased fuel consumption and decreased turbine operating temperatures
  Downtrim for:
    – subsonic and supersonic maneuvering
    – weapon delivery
    – short field landing
- Rapid thrust modulation for:
    – improved terrain following/terrain avoidance/threat avoidance TF/TA2 (vary nozzle area and inlet guide vanes, combined potentially with vectoring nozzle)
- Combat mode for:
    – short-term combat rating featuring higher fan turbine inlet temperature (pilot selectable)

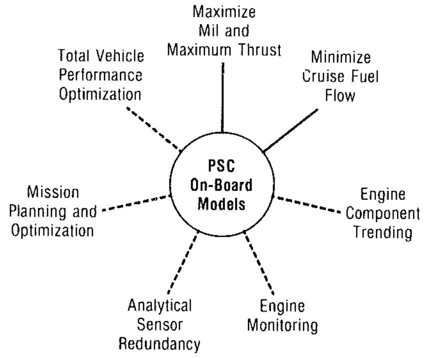


Fig. 17 PSC growth avenues

Control Mode	Performance	Supersonic Persistence	STOL	Survivability	Maneuverability	Range	Weapon Delivery
Stall Margin Control	X	X	X	X	X	X	X
Rapid Thrust Modulation	X			X	X		X
Combat	X			X	X		X
Thrust Reversing/Vectoring			X	X	X		X
Flow Modulation	X	X				X	
Auto Throttle				X		X	X
Performance Seeking	X	X		X	X	X	

Fig. 18 Integrated flight/propulsion control payoff potential

- Thrust reversing/vectoring for:
    – greater maneuverability or decreased trim drag at cruise
    – fuselage aiming for weapon delivery
    – approach mode for landing
- Flow modulation for:
    – improved inlet/engine matching through engine/nozzle variable geometry
- Autothrottle for:
    – trajectory control for energy management, TF/TA, and weapon delivery
- Performance seeking control for:
    – maximum thrust in accelerations
    – minimum cruise fuel flow

The mission areas where payoffs are expected are shown in Fig. 18. To fully realize such payoffs, integrated flight/propulsion control technology must be developed in research flight test programs such as HIDEC and addressed early in the aircraft design stages.

References

1 Putnam, T. W., Burcham, F. W., Jr., Andries, M. G., and Kelly, J. B., "Performance Improvements of a Highly Integrated Digital Electronic Control System for an F-15 Airplane," NASA TM-86748, Aug. 1985.

2 Yonke, W. A., Terrell, L. A., and Myers, L. P., "Integrated Flight/Propulsion Control: Adaptive Engine Control Mode," AIAA Paper No. 85-1425, June 1985.

3 Yonke, W. A., Landy, R. J., and Cushing, J. M., "Integrated Flight/Propulsion Control: HIDEC Modes," NAECON Paper, May 1984.

4 Burcham, F. W., Jr., Meyers, L. P., and Ray, R. J., "Predicted Performance Benefits of an Adaptive Digital Engine Control System on an F-15 Airplane," AIAA Paper No. 85-0255, Jan. 1985.

5 Baer-Riedhart, J. L., and Landy, R. J., "Highly Integrated Digital Electronic Control: Digital Flight Control, Aircraft Model Identification and Adaptive Engine Control," AIAA Paper No. 85-1877, Aug. 1985.

Maximum-Slew-Rate Control for Hydrogenerators[1]

J. Jaeger[2]
Graduate Student.

L. Wozniak, P.E.
Associate Professor.
Mem. ASME

University of Illinois,
Department of General Engineering,
Urbana, IL 61801

Traditionally, speed control of hydrogenerators has been implemented by analog proportional-plus-integral control of speed error. With digital controllers, nonlinear, adaptive control strategies may be implemented to achieve improved systems response. This paper develops a control scheme which responds to perturbations by moving wicket gates at their maximum slew rate. For a simulated 10 percent load rejection, the new controller reduced settling time by over 50 percent as compared to the equal-root-case PI controller. This was achieved while maintaining the same peak overshoot and keeping pressure within design limits. The deadbeat nature of the transient response promises decreased wear on mechanical components while substantially improving overall performance of hydrogenerators.

Introduction

The hydropower generation industry is relatively mature. The problem of feedback control to maintain output frequency has been traditionally implemented through the use of analog electromechanical devices that sense turbine speed and position the wicket gates with a signal proportional to speed error and its integral (PI control). The widespread use of PI control in hydropower, as well as other fields, has been attributed to the linear operation limitations of available hardware. Interesting control strategies that do not fall into the linear response category would be difficult, if not impossible, using this equipment.

With the introduction of the microprocessor into real time control operations many of the previous limitations have been removed, and a wide range of control possibilities can now be realistically considered. Literature is beginning to reflect the growing use of digital controllers as they demonstrate their advantages and reliability. Some present research is investigating the application of microprocessors to the control of hydrogenerators [1-3], and in one case in Scotland a digital speed controller is already in use [4].

In fixed blade angle type turbines the wicket gate is the sole control variable. The factor that limits the speed at which the gate can be opened or closed, and hence the responsiveness of the system, is the allowable pressure deviation that the conduit can withstand. This pressure limitation is achieved by setting the maximum slew rate of the gate. However, PI control generally moves the gate at a rate of speed less than this allowable limit, thus the system is responding more slowly than if the controller were always to move the gate at its maximum slew rate. An on-off, constant speed gate controller can achieve the desired result. Since the microprocessor provides the tools necessary to implement such a control, it is now appropriate to investigate this controller with the intention of its eventual utilization.

Model

This paper considers a linearized model for fixed blade turbines given by the following set of differential equations [5]

$$T_a D x = m - l \tag{1}$$

$$T_w D q = -h \tag{2}$$

$$m = \frac{\partial m}{\partial x} x + \frac{\partial m}{\partial y} y + \frac{\partial m}{\partial h} h \tag{3}$$

$$q = \frac{\partial q}{\partial x} x + \frac{\partial q}{\partial y} y + \frac{\partial q}{\partial h} h \tag{4}$$

The variables are nondimensionalized about an operating point and have the following definitions: h = per unit pressure deviation; l = per unit load demand; m = per unit turbine torque; q = per unit flow deviation; x = per unit rotor speed deviation; y = per unit gate deviation; T_a = rotor time constant; T_w = conduit water time constant; D = derivative operator; $\partial m/\partial x$, $\partial m/\partial y$, $\partial m/\partial h$, $\partial q/\partial x$, $\partial q/\partial y$, $\partial q/\partial h$ = system partial derivatives.

As a result of linearization, the system partials remain constant for a particular turbine functioning within a small range of its operating point.

In the open-loop form this model represents a linear, continuous second-order system. By adding proportional plus integral speed control the following dynamics are added

$$y = -Px - I\int x \tag{5}$$

$$Dy = -PDx - Ix \tag{6}$$

In equation (6) Dy is the rate of change of the gate position (the slew rate), while P and I are the proportional and integral gain constants.

---

[1] Funding for this effort was provided by the Program Related Engineering and Scientific Studies (PRESS) budget of the United States Department of the Interior, Bureau of Reclamation. An abbreviated version of this paper was presented at the Waterpower '85 Conference, Las Vegas, Nevada, September 25-27, 1985.

[2] Currently with Allison Gas Turbine Division of General Motors Corporation, Indianapolis, IN.

Contributed by the Power Division for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Power Division January 8, 1986.

Table 1 Standard system partials and time constants

$$\frac{\partial m}{\partial x}=-1.0 \quad \frac{\partial q}{\partial x}=0.0 \quad \frac{\partial m}{\partial y}=1.0 \quad \frac{\partial q}{\partial y}=1.0 \quad \frac{\partial m}{\partial h}=1.5 \quad \frac{\partial q}{\partial h}=0.5$$

$$T_a=9.0 \quad T_w=1.4 \quad G_s=0.1$$

This traditional control creates a linear third-order closed-loop system which is relatively well understood. The particular selection of governor gains for desired transients has lacked a unified definition of optimality. This work is based on comparison to the equal-roots case [5, 6]. The selection produces a more damped transient than the tuning suggested by Paynter [7] and Hovey [8] and is a more appropriate measure due to the deadbeat nature of the proposed control.

This paper investigates constant slew rate control that can be written as

$$Dy=G_s u \tag{7}$$

Dy, the instantaneous rate of change of the gate position, relates to G_s, the constant value of the gate speed as determined from allowable pressure deviation. u is the control variable. For this controller u can have one of three discrete values: $+1$, -1, or 0. $u = +1$ represents opening the gate at maximum rate; $u = -1$ represents closing the gate at maximum rate; $u = 0$ represents no motion of the gate.

During closed-loop control operation the actual value of the control variable will be determined by the other system variables. Like the PI controller, this closed-loop control also creates a third-order system model, but it is inherently nonlinear with the result that many of the classical analysis techniques are not applicable. Therefore, this paper makes extensive use of numerical integration as a tool for developing and analyzing the constant gate speed controller. Computations are performed using the Advanced Continuous Simulation Language (ACSL) package [9]. The values for the constants appearing in the mathematical model are required. For this purpose, the partials and time constants of an example system [5] as shown in Table 1 are used. This example system is henceforth referred to as the "standard" system. The maximum slew rate corresponding to 10 s for full gate travel is assumed. The technique used in generating the feedback control for this standard system is then generalized for application to system models with other values of partials, time constants, and gate speed.

Definition of Control Performance Goal

For constant gate speed control, the control variable u can be $+1$, -1, or 0. Under a given load disturbance it is necessary to find a control function $u(t)$ such that at any given time $u(t)$ is an element of the set $\{+1, -1, 0\}$ and the overall system response meets some performance criteria. In this work the performance goal is defined to be achieved by the control which returns the system to its operating point utilizing a "single overtravel control" and minimizing the rotor speed settling time.

The term single overtravel control is introduced to describe a control that switches gate direction no more than once in response to a step disturbance in load. For a load disturbance occurring at time t_0, the following potential control sequence represents an example of single overtravel control

$$u(t)=\begin{array}{lll} +1\ (\text{or } -1) & \text{for} & t_0 \le t < t_1 \\ -1\ (+1) & \text{for} & t_1 \le t < t_2 \\ 0 & \text{for} & t_2 \le t \end{array} \tag{8}$$

The single overtravel control is used because it can be optimized using numerical simulation for a given load level, while it also produces a response that is a significant improvement over PI control.

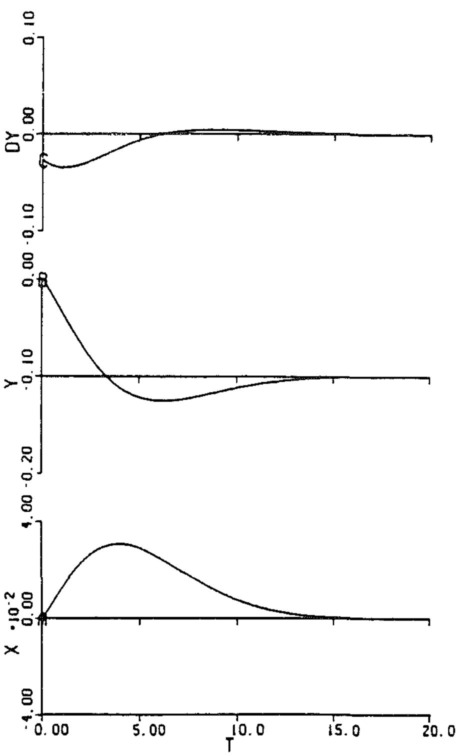


Fig. 1 Proportional plus integral control response

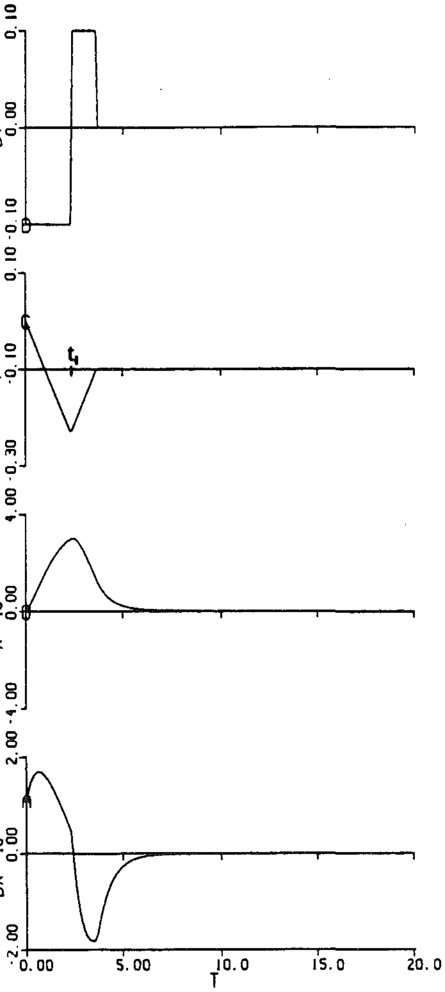


Fig. 2 Optimal single overtravel control response

The response of the PI controller with $P = 2.44$ p.u. and $I = 0.464$ p.u. tuned for the equal roots case generated by ACSL has output shown in Fig. 1. For the PI control, speed deviation returns to within 1/2 percent in 11.0 s.

For a particular step disturbance in load, there is a unique single overtravel control that will return the system to steady state in a minimum settling time. This control is called the "optimal single overtravel control" or OSOC for short. An iterative technique for determining the OSOC is presented in [10]. For the system presented earlier, the OSOC in response to a 10 percent load reduction results in $t_1 = 2.33$ and $t_2 = 3.66$ s. An ACSL program simulates the system as presented earlier and computes its response subject to the optimal control schedule. The results are shown in Fig. 2.

When compared to the PI response for the same input (Fig. 1), the single overtravel control has a shorter settling time, 4.4 s to return deviation to within 1/2 percent of its operating value as opposed to 11.0 s for the tuned PI controller. Furthermore, it is important to note that this improvement in settling time was achieved without degrading peak overshoot, which is approximately the same in both cases. The constant gate speed controller does have higher pressure deviations, but the gate speed is set to guarantee that pressure always remains within design limits. This demonstrates that constant gate speed control, utilizing a single overtravel sequence, can significantly improve system response over PI control.

Feedback Controller Development

Inputting a precalculated OSOC function to control a known disturbance on the open loop system is a preliminary step in developing a feedback controller which would be able to respond to any load disturbance on a continuous basis. A constant gate speed controller may be implemented using a microprocessor which will avail itself of the measures of the independent system states at regular sample intervals. The "best" control action is then output to the gate servomotor. An iterative technique for determining this best control at any point in the state space was devised.

Since the system with control is third order, three independent states are needed to describe its condition at any point in time. The states have been selected to be speed deviation, rate of change of speed, and gate position. Hence the state space is three dimensional, with elements $\{x, Dx, y\}$.

The location of the operating point (steady-state condition) for any given load in the state space is shown by the following equation

$$(x, Dx, y)_{ss} = (0, 0, [1/(\partial m/\partial y)]l) \tag{9}$$

The operating point moves along the gate position axis in the state space as the load level changes. In order to simplify the analysis, a new state, y_{offset}, is defined. This state represents the offset of the gate from its steady-state position

$$y_{\text{offset}} \triangleq y_{os} \triangleq y - [1/(\partial m/\partial y)]l \tag{10}$$

Now, in the new state space $\{x, Dx, y_{os}\}$, the steady-state operating point is the origin regardless of changes in load level. At any given point in the state space, it is possible to determine the complete single overtravel control that minimizes settling time by using that point as an initial condition. Then, if during real time control the system is ever at this state space point, the controller will output the initial value of the complete control sequence, namely $u(0)$. The entire optimal control for that point does not have to be used, because before it could be completely implemented the controller will have sampled the system again and measured a new state.

The technique for finding optimal $u(0)$ for a particular initial state point simulates the system response under the three controls A, B, and C of equation (11)

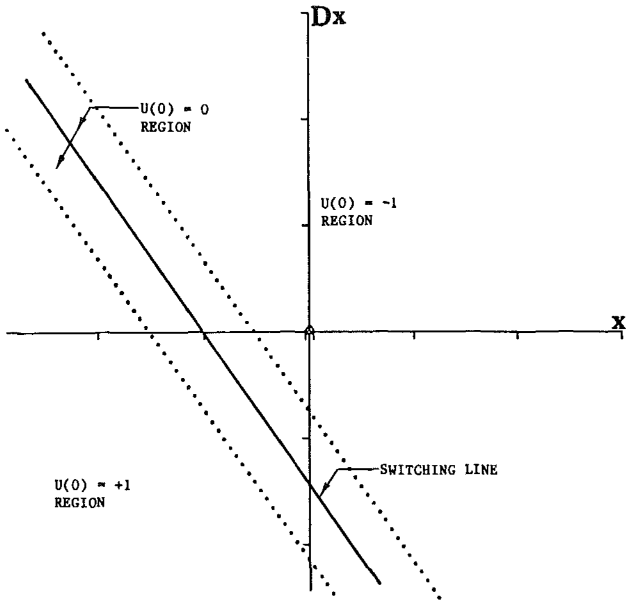


Fig. 3 State space regions where best $u(0) = +1, -1$, and 0

$$\begin{aligned} &\text{Control A:} \quad u(t) = 0 \quad \text{for } 0 \le t < \epsilon \\ &\text{Control B:} \quad u(t) = -1 \text{ for } 0 \le t < \epsilon \\ &\text{Control C:} \quad u(t) = +1 \text{ for } 0 \le t < \epsilon \end{aligned} \tag{11}$$

For $\epsilon \le t$ all controls drive gate offset directly to zero.

Determining which of the three controls yields the best result is done by measuring the speed deviation due to each control at some later observation time. Enough time must have elapsed after active control stops for the effects of that action to be demonstrated in the system states. The control sequence that results in the smallest speed deviation at this observation time is taken to represent optimal $u(0)$.

Mapping Control Onto State Space

Since there are an infinite number of points in space, actually finding and storing $u(0)$ for each one would be impractical. Fortunately bang-bang feedback controllers generally have their state space divided into regions where one or the other control is optimal. These regions are separated by a switching surface [11, 12]. The problem reduces to one of identification of the switching surface for the particular system. To gain understanding of the nature of this surface, $u(0)$ is found first for points evenly distributed on the plane of zero gate offset ($y_{os} = 0$). The number of points to be simulated is reduced by the fact that linearity requires the system to be inversely symmetric about the origin. That is, if some point (x_1, Dx_1, y_{os1}) has optimal control $u_1(0)$, the inverse point $(-x_1, -Dx_1, -y_{os1})$ has optimal $u(0) = -u_1(0)$.

For the state space of the present example system in the operating region for x, Dx, and $y_{os} = 0$, points in even increments on both the x and Dx axis were used as initial conditions and simulated under controls A, B, and C (equation (11)) using ACSL. Optimal $u(0)$ was found and plotted for a number of points on the plane of zero gate offset.

As expected, the plane $y_{os} = 0$ is divided into a region where $u(0) = -1$, and a region where $u(0) = +1$ by a curve which is the intersection of the switching surface and the plane $y_{os} = 0$, as shown in Fig. 3. Of interest here is that this curve is, or can be closely approximated as, a straight line. The result is verifiable by phase plane analysis and is treated in the appendix. The entire state space in the region where the system operates can be mapped by finding and storing equations of these switching lines for small increments in gate offset.

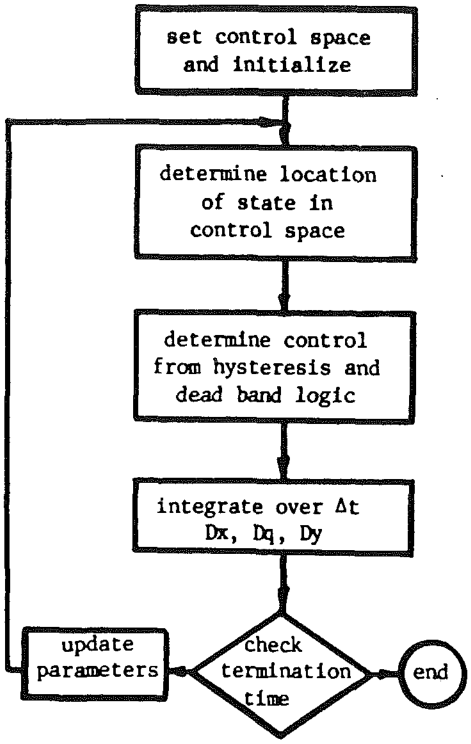


Fig. 4 ACSL program flow chart

Inspection of the control space reveals a small region around the switching surface having optimal $u(0) = 0$ (see Fig. 3). It is a result of the simulation technique. Each control A, B, and C implements $u(t) = 0, -1$, or $+1$ for a finite interval of time $0 \leq t < \epsilon$ where $\epsilon = 2T$ s has been used. Because ϵ is not infinitely small, some error will occur in identifying $u(0)$ near the switching surface, resulting in the region of $u(0) = 0$. The switching line is defined to be the center of this region. The control algorithm is implemented with a hysteresis and a dead band about the switching line.

Another useful fact is that the slope of the switching line is constant for all values of gate offset. Given this, it is necessary to determine only the value of F in the equation $Dx = Ex + F$ in order to describe the switching line for a particular gate offset level. This can be done by locating just one point on each line. To complete the control space for the standard system the switching lines are found for all gate levels from 0 to 0.5 in increments of 0.02 (one sample time times maximum gate speed). Due to the symmetry about the origin the switching lines for gate levels from 0 to -0.5 are also known. Data from the simulations are displayed in condensed form in Table 2. All intercepts are for lines of slope $= -1.43$.

The ACSL program "Feedback Control Simulation" shown in the flow chart of Fig. 4 simulates continuous feedback control of the standard system. The control space just generated is used. The feedback controller can respond to any load input, but of particular interest is its response to a 10 percent load reduction (Fig. 5) because its optimal single overtravel control has been presented earlier.

The response of the feedback controller closely matches its defined performance goal of optimal single overtravel control. For this input the optimal control returns speed deviation to within 1/2 percent of the operating value in 4.4 s, while the feedback controller does so in 4.7 s. This is slightly less than optimal, but it is still a substantial improvement over the PI controller, which required 11.0 s. Tuning of the hysteresis and dead bands may further improve the feedback controller's performance.

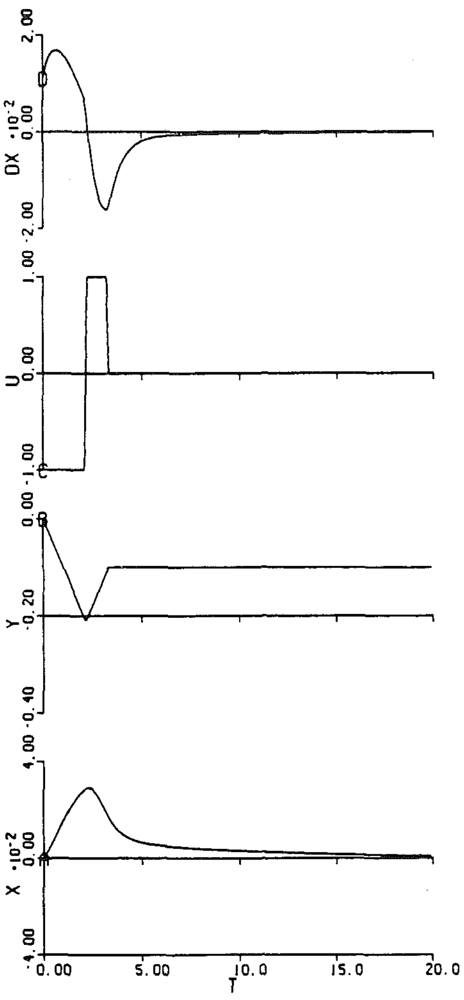


Fig. 5 Feedback controller response

Table 2 Switching line intercepts

Offsets			
gate	line	gate	line
0.02	−0.0071	0.28	−0.1682
0.04	−0.0143	0.30	−0.1878
0.06	−0.0229	0.32	−0.2075
0.08	−0.0314	0.34	−0.2289
0.10	−0.0404	0.36	−0.2511
0.12	−0.0511	0.38	−0.2743
0.14	−0.0621	0.40	−0.2993
0.16	−0.0746	0.42	−0.3250
0.18	−0.0875	0.44	−0.3521
0.20	−0.1018	0.46	−0.3807
0.22	−0.1175	0.48	−0.4100
0.24	−0.1332	0.50	−0.4411
0.26	−0.1507		

Controller Performance Simulations

The final step in the development of this controller is to apply it in a discrete manner. An ACSL program was used to simulate the controller's operation when output is updated at discrete sample times. Essentially this program simulates the feedback control strategy developed on a continuous basis with a zero order hold added. Figure 6 shows the results of an 11 percent load rejection applied to the standard system. Since a control signal must last for an entire sample time, the digital controller can not position the gate in smaller increments than gate speed times sample time. In general it will not be possible to exactly reduce gate offset to zero. With the current values of gate speed and sample time the gate can be moved in increments of 0.02, an exact multiple of -0.1, which is the

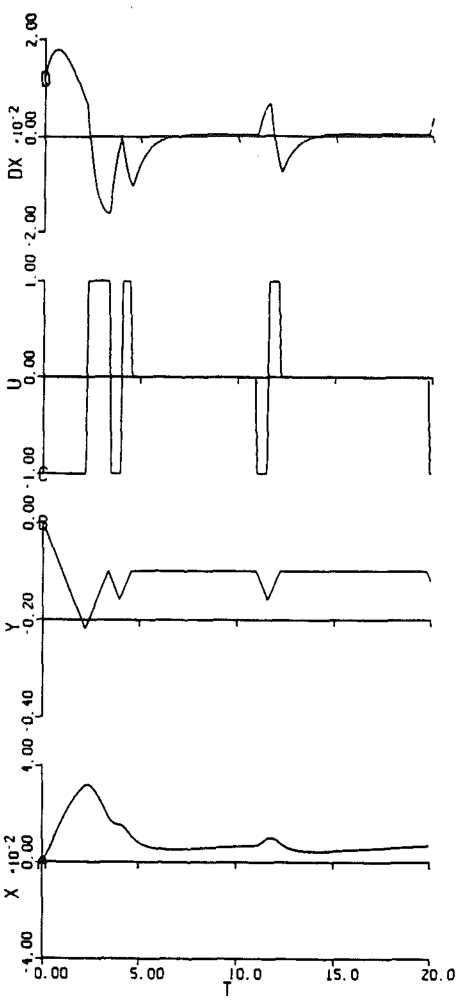


Fig. 6 System response with discrete feedback controller

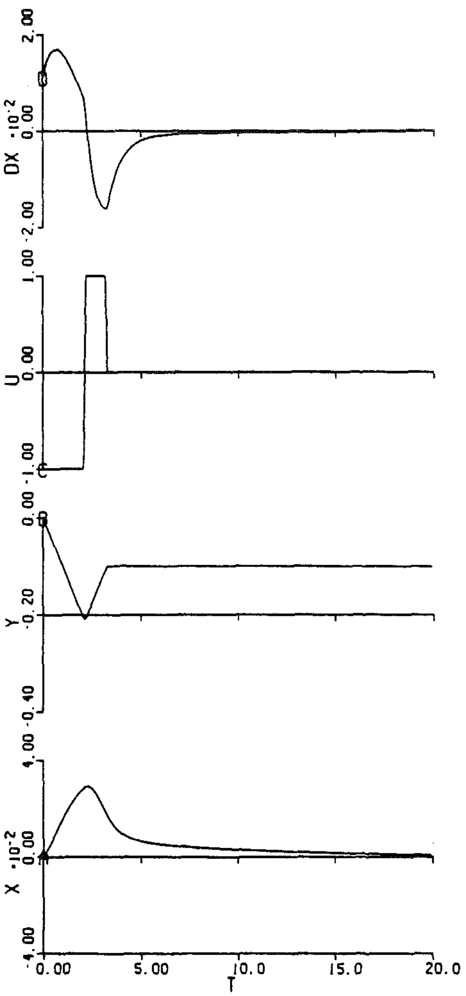


Fig. 7 System response when gate actuator dynamics are modeled

steady-state gate position for a 10 percent load reduction. Thus, to demonstrate the potential for steady-state error the load input used is 11 percent.

The initial response of the digital controller is to output a single overtravel control signal. The first drive time is equal to 2.2 s, which, although less than optimal, is within two sample times of 2.51, the optimal t_1 for an 11 percent load reduction. More serious is that when the single overtravel control would stop at 3.3 s there is a gate positioning error of 1 percent from the steady-state value. This error causes the controller to kick in further control signals in order to bring the system states to within the dead band. Since a zero gate offset can not be reached, the system again drifts out of the dead band and the controller must continue to provide gate commands to bring the system back.

In actual practice, sample times much shorter than 0.2 s are achievable, and probably would be used. This reduces the smallest possible gate position increment which will reduce steady-state error. Also, the microprocessor can switch to a small deviation control strategy in the dead band region. Such a strategy would drive the gate at speeds less than maximum, which will again reduce gate position increments, reducing steady-state error.

A digital PI control strategy would be a good alternate choice for the small deviation controller. The same microprocessor would be able to implement it with software, thus not requiring additional hardware. The PI feedback adjusts gate position based on the rotor speed error, and its integral. It can reduce steady-state error to signal quantization level.

A second point to consider is that of gate dynamics. The model used ignored the dynamic behavior of the gate actuator, and assumed the gate exactly integrated the control signal. Since these dynamics do exist, it is worthwhile to investigate their effect on the performance of the constant gate speed controller. The actual gate response to a control signal can be modeled with the second-order transfer function of equation (12) [13]. Here y represents actual gate position, and z represents the desired gate position, or the input to the hydraulic amplifier

$$y(S)=\frac{1}{(0.15S+1)^2}\,z(S) \tag{12}$$

If a maximum limit of G_s is placed on the gate speed as in equation (13), the control signal can be fed directly to the gate actuator

$$|Dy| \leq G_s \tag{13}$$

An ACSL program was written to simulate the actuator limitations using the digital feedback controller responding to a 10 percent load reduction. Figure 7 shows the standard system's response when gate dynamics are considered. As expected, the result is essentially identical to the case without dynamics, due to the relatively short time constants involved.

Conclusion

Several aspects of the constant gate speed controller still need to be addressed. This controller uses a "dead band" which is a region of the state space where no control is applied, and includes state points other than the operating point. It is generally necessary for small deviations from the operating point to be returned through control action. However, applying a maximum gate speed control for even one sampling time increment may overdrive the operating point. As a result it is desirable to switch control mode to a small deviation controller in the dead band region. It has already been mentioned that a digital PI algorithm performed by the same microprocessor is a potential candidate. Other possibilities exist. Whichever strategy is chosen, it would need to be mated with the constant gate speed controller for a smooth transition from one controller to the other to occur. The dual mode controller would use maximum gate speed control to respond to large disturbances and quickly reduce the speed deviation to within the dead band region, or alternately preset limits where the small deviation controller would act to eliminate steady-state error.

The research performed for this paper demonstrates that constant gate speed bang-bang control holds potential advantages over current control techniques. The main advantage is that speed deviation settling time can be substantially reduced without increasing its peak overshoot. In the particular example presented in this paper settling time was decreased by 60 percent. An additional benefit of reduced wear and tear may be realized due to the deadbeat nature of the system transients.

The widespread use of the three-term electric governor over the last three decades has raised much controversy about the positive and negative effects of derivative control [14]. The bang-bang strategy offers maximum control output and therefore surpasses the desired characteristics of derivative control while not being susceptible to noise. The rapidly growing use of digital controllers is demonstrating their reliability and cost as well as performance advantages over their analog counterparts. The trend in the hydroturbine industry will be to rely more and more on computer controllers, and on improved control laws. In the course of this evolution maximum gate speed control warrants further consideration.

References

1 Kopacek, P., and Zauner, E., "Governing Turbines by Microcomputers," *International Water Power and Dam Construction*, Vol. 34, 1982, pp. 26-30.

2 Ryusawa, S., Haratani, T., and Yanagisawa, T., "Microcomputer Applied Programmable Controller for Hydro Electric Power Generation," *Industrial Application of Microprocessors*, IECI Proceedings, March 20-22, 1978.

3 Sandberg, T., Karlsson, O., Forden, B., and Hartwig, S., "Computerised Governor Control for Bulb and Kaplan Turbines," *International Water Power and Dam Construction*, Vol. 31, 1979, pp. 43-45.

4 Grant, N. F., Aitken, K. H., Winning, D. J., and Davie, H., "Development of an Operational Hydroturbine Governor," *Microprocessors and Microsystems*, Vol. 4, 1980, pp. 347-351.

5 Wozniak, L., "Load Level Tuning of Governor Gains for Sub-optimal Speed Control of Hydroelectric Installations," ASME Paper No. 74-WA/FE-4, 1974.

6 Hovey, L. M., and Bateman, L. A., "Speed-Regulation Tests on a Hydro Station Supplying an Isolated Load," *Trans. IEEE* (*Power Apparatus and Systems*), Vol. 81, Pt. 3, Discussion by M. E. Leum, 1962, pp. 369-370.

7 Paynter, M., "The Analog in Governor Design, I, a Restricted Problem," Palimpsest on the Electronic Analog Art printed by G. A. Philbrick Researches, Inc., Boston, MA, 1960.

8 Hovey, L. M., "Optimum Adjustment of Governors in Hydro Generating Stations," *The Engineering Journal*, Vol. 43, No. 11, 1960.

9 *Advanced Continuous Simulation Language User Guide and Reference Manual*, Mitchell and Gauther, Inc., Concord, 1975.

10 Jaeger, J., "Maximum-Slew-Rate Nonlinear Control for Improved Transient Response in Hydrogenerators," Master's Thesis, University of Illinois, Department of General Engineering, 1985.

11 Athans, M., and Falb, P., *Optimal Control An Introduction to the Theory and Its Applications*, McGraw-Hill, New York, 1966.

12 Gibson, J. E., *Nonlinear Automatic Control*, McGraw-Hill, New York, 1963.

13 Woodward Governor Company, "Electric Governor for Hydraulic Turbines," Bulletin 07074A, Rockford, IL.

14 Ransford, G. D., "PID Regulation Revisited," *Water Power and Dam Construction*, Vol. 35, 1983, pp. 31-34.

APPENDIX

Phase Plane Analysis

A phase plane analysis has been performed to determine the trajectory $\partial Dx/\partial x$ of this system, and thus gain an insight into the nature of the switching surface. The first step is to transform the system from the equations (14) through (18) into equation (19) by elimination of all variables except speed

$$T_a Dx = m - l \tag{14}$$

$$T_w Dq = -h \tag{15}$$

$$Dy = G_s u \tag{16}$$

$$m = (\partial m/\partial x)x + (\partial m/\partial y)y + (\partial m/\partial h)h \tag{17}$$

$$q = \frac{\partial q}{\partial x}\,x + \frac{\partial q}{\partial y}\,y + \frac{\partial q}{\partial h}\,h \tag{18}$$

The symbolic equation (19) when given in terms of system parameters becomes

$$D^2x + aDx + bx + c = 0 \tag{19}$$

$$T_a T_w (\partial q/\partial h) D^2 x + \left[T_a + T_w \frac{\partial m}{\partial h} \frac{\partial q}{\partial x} - T_w (\partial q/\partial h)(\partial m/\partial x) \right] Dx - \frac{\partial m}{\partial x}\,x - \frac{\partial m}{\partial y}\,y + \left[\frac{\partial m}{\partial h} \frac{\partial q}{\partial y} - \frac{\partial q}{\partial h} \frac{\partial m}{\partial y} \right] Dy = 0 \tag{20}$$

Here the values of the x and Dx coefficients are constant, but the next two terms are dependent upon gate position and gate velocity. Restricting the analysis to the zero gate offset plane, $y = 0$. Constant y implies that $Dy = 0$, thus on the zero gate offset plane $c = 0$. To find the trajectory, the chain rule is applied to D^2x.

$$D^2x = \frac{\partial Dx}{\partial x} \frac{\partial x}{\partial t} = \frac{\partial Dx}{\partial x} Dx \tag{21}$$

The slope of the trajectory on the phase plane may be found by substituting equation (21) into (20)

$$\frac{\partial Dx}{\partial x} = \frac{-aDx - bx}{Dx} \tag{22}$$

where a and b are the coefficients of the Dx and x terms, respectively.

It is desired to find those trajectories the system will follow to the origin. Thus values of M that satisfy both equations (23) and (24) are sought

$$Dx = Mx \tag{23}$$

$$(\partial Dx/\partial x) = M \tag{24}$$

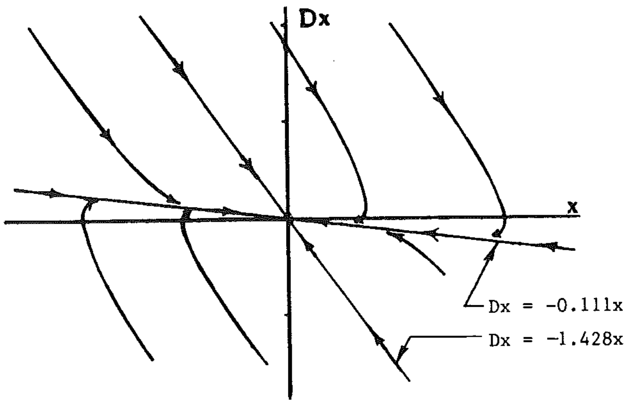


Fig. 8 Phase trajectories on zero gate offset plane

Plugging these two equations into (22) gives the following relationship.

$$M=\frac{-Mx(aMx+bx)}{Mx} \tag{25}$$

The rotor speed variable is common to all terms and may be canceled. The equation reduces to the quadratic

$$M^2+aM+b=0 \tag{26}$$

Applying the quadratic formula produces two values of the slope M

$$M=\frac{1}{2}\ (-a\pm a^2-4b) \tag{27}$$

For the standard system this reduces to the two lines of equations (28) and (29)

$$Dx=-0.111x \tag{28}$$

$$Dx=-1.428x \tag{29}$$

The feedback controller actively drives the system states to the fast trajectory and terminates gate motion, allowing the system to coast to its operating point (see Fig. 8). This analysis provides a mathematical method for calculating the slope of the switching lines. A similar analysis to find the switching line intercepts on planes of $y \neq 0$ by finding $\partial Dx/\partial x$ as performed here would not be valid. This is because when the gate offset is not zero, even though the system is on the switching surface, the controller will drive the gate to eliminate its offset. Since Dy is not zero, the system trajectory will not be planar. A three-dimensional trajectory analysis is required to provide a means of calculating switching lines on all planes of constant gate.

Periodic and Continuous Vibration Monitoring for Preventive/Predictive Maintenance of Rotating Machinery

W. C. Laws

A. Muszynska
Mem. ASME

Bently Nevada Corporation,
Minden, NV 89423-2529

The application of vibration monitoring as part of Preventive/Predictive Maintenance programs is discussed. Several alternative methods, including periodic and continuous monitoring techniques, are described. Emphasis is given to the importance of selecting vibration transducers with due regard for the specific machinery type. The equally important need to install monitoring systems which are cost effective and provide genuinely useful information for maintenance engineers and vibration analysts is also highlighted. It is argued that critical machinery should be monitored continuously, and in cases when more detailed investigation is required that high-quality Predictive Maintenance vibration analysis techniques be applied. The need is also emphasized for specialist interpretation of vibration data in order to identify specific machinery malfunctions, of which several examples are given.

Introduction

Important recent developments in instrumentation and computer-based information management systems have led to great advances in the effectiveness of Preventive/Predictive Maintenance programs. It should not be overlooked, however, that there is nothing new in the basic philosophy of Preventive Maintenance or Condition Monitoring of machinery.

For example, consider the case of a plant engineer in charge of a factory or power station or a marine engineer responsible for his engine room. Fifty or sixty years ago, Preventive Maintenance activities were carried out by the engineer through regular observation of control room gages and possibly the logging in traditional form of the operating parameters available. But more importantly, the conscientious engineer would walk about the engine room listening to the machinery, placing a finger on top of a bearing cap or at the foot of a pedestal, looking for signs of oil "squeeze" between casings and their bedplates or watching for signs of steam leaks at valve stem-seals, etc. Provided that, in the event that the engineer detected or suspected a fault was present or about to develop, he took some positive action, this traditional procedure was in every respect condition monitoring for Preventive Maintenance.

In the 1980s, more and more emphasis is being given to the use of permanently installed monitoring equipment and the remote assessment of data using computer-based systems. In many ways this is a justifiable trend – machinery is more sophisticated and expensive to repair; skilled labor is expensive, and monitoring hardware is, in real terms, very inexpensive and widely available. The cost of specialized software for Preventive/Predictive Maintenance applications is still relatively high unless it can be applied in a very standardized manner with numerous copies sold. In the next ten years, indications are that monitoring hardware generally will have even lower cost in real terms, but custom-written software programs will become relatively more expensive.

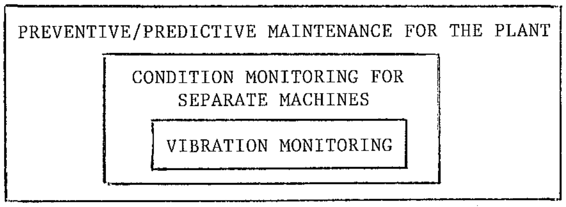


Fig. 1 Maintenance system

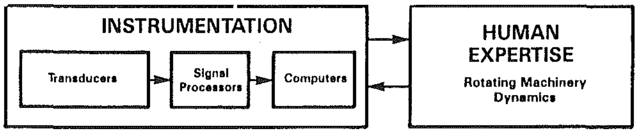


Fig. 2 Basic components of the Preventive/Predictive Maintenance systems

Predictive Maintenance should, therefore, not be thought of as something new or separate from other maintenance/cost savings schemes or individual actions; it is the systematic application of every aspect of good engineering practice (maintenance and operational), experience, good judgment, and good information management. The importance of careful selection of monitoring instrumentation and data

---

Contributed by the Nuclear Engineering Division and presented at the ASME Winter Annual Meeting, Miami Beach, Florida, November 17–21, 1985. Manuscript received at ASME Headquarters August 25, 1986.

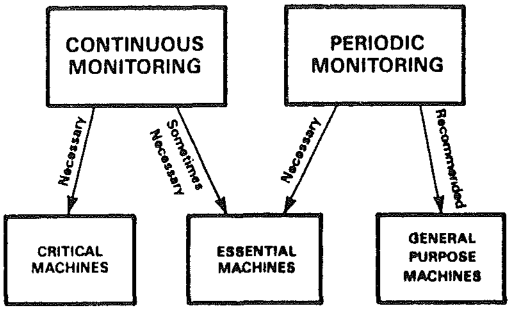


Fig. 3 Application of continuous and periodic monitoring systems

management systems is nevertheless considerable if the system is to be effective. Of course, by "effective" we mean that it saves the user money through reduced breakdowns which result in lost production or unnecessarily expensive rebuilds of machinery. Avoiding accidents (and possible injuring of personnel) is, of course, another benefit of a well-organized Preventive/Predictive Maintenance system, but even this aspect has direct financial implications.

In today's expense-conscious world, the costs of implementation of Preventive/Predictive Maintenance systems in process plants can be proved to be lower than losses caused by machine failures. Utilizing state-of-the-art techniques (both instrumentation and diagnostics), it is possible to operate machines safely from a mechanical standpoint longer than it may be economically sound from a production or energy efficiency viewpoint.

Since there exists certain confusion in terminology, Preventive/Predictive Maintenance and Condition Monitoring objectives are worthy of clarification (Figs. 1 and 2).

Preventive Maintenance activities are usually defined as "all deliberate maintenance functions intended to avoid unplanned breakdowns of machinery." For example, basic operations, such as maintaining adequate supply and quality of lubricants for bearings, mechanical adjustments, and maintenance of protective coatings on machinery casings are traditional Preventive Maintenance functions. Sometimes, regular replacement of components (typically rolling element bearings and seals) is undertaken, based on experience and "typical" operating life expectations.

Preventive Maintenance programs are based, therefore, on *rigid schedules* of activities. According to these schedules, it might happen that a machine is lubricated more often than necessary to satisfy its minimum requirements for bearing oil. Rigid preventive maintenance schedules tend, therefore, to be relatively inefficient in terms of manpower and materials.

Most operators of Preventive Maintenance schemes use some form of plan or schedule to ensure that maintenance operations are performed regularly – some are quite sophisticated programs which are optimized by continuous refinement over several years.

Electronic instrumentation for machinery monitoring added a new dimension to the Preventive Maintenance programs. From one side, based on long-term experiences of preventive activities, and from the other side on measured dynamic data of the machine performance, preventive maintenance schedules can be *optimized*. Gradually *Preventive* Maintenance programs can be developed into *Predictive* Maintenance programs.

Predictive Maintenance activities use various techniques and measurements, which are performed on a regular basis. The intention is to predict when it is necessary to take a machine out of service for adjustment and/or repair.

In some instances, Predictive Maintenance measurements (which include a wide range of parameters) can be integrated into an existing Preventive Maintenance scheme. Development of a combined Preventive/Predictive program can lead to a very effective maintenance system. This paper is concerned with the use of vibration monitoring in Predictive Maintenance operations which may or may not be part of a larger maintenance system.

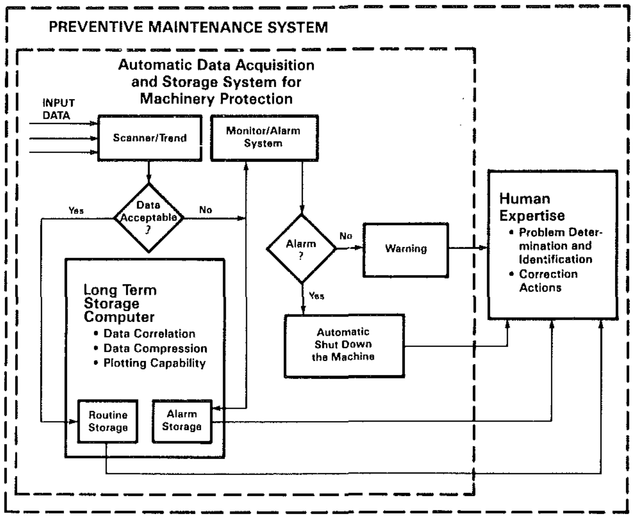


Fig. 4 On-line system for critical machines

The Predictive Maintenance objective should be zero unexpected breakdowns in the overall operation. The techniques applied enable machines to be operated to the optimum point, maximizing use of machinery and their components, permitting replacement in a predictable manner, with minimized downtime, minimized loss of production, minimized spares holding, and optimized labor scheduling.

Predictive Maintenance relates, therefore, to a group of machines or to the entire plant. Condition Monitoring, being a part of the maintenance system, relates rather to a particular machine.

Condition Monitoring aims to define the current condition of a machine and compare it with previously measured conditions. Some element of prediction is inferred by observing trends in the monitored parameters, be it efficiency, power output, delivered load, vibration, or noise level.

Condition Monitoring is, in a sense, a subset of Preventive/Predictive Maintenance. Vibration Monitoring is an important subset of both. In other words, Vibration Monitoring is a part of the database and Condition Monitoring and represents a subsystem of the overall Preventive/Predictive Maintenance, designed to save money in any way possible.

Machinery Protection is often thought to be a separate function providing a means of shutting down machinery if a serious defect results. In truth, machinery protection systems represent important parts of the Predictive Maintenance programs. Avoiding the extremely punitive results of an unpredicted, possibly catastrophic machinery failure, which may have implications for personnel safety, is a part of the equation. Catastrophic failure of large high-speed machinery is usually considered to be the most dangerous event in terms of plant personnel safety. The fact is that the risks of personal injury are much higher from the failure of much smaller auxiliary plants typified by small motor-driven pumps. High-capital-value machinery is normally monitored continuously these days; auxiliary machinery is still "ignored continuously" in many industrial environments.

Selecting Vibration Information Systems for Preventive/Predictive Maintenance (PM)

It is extremely important, when selecting systems for vibra-

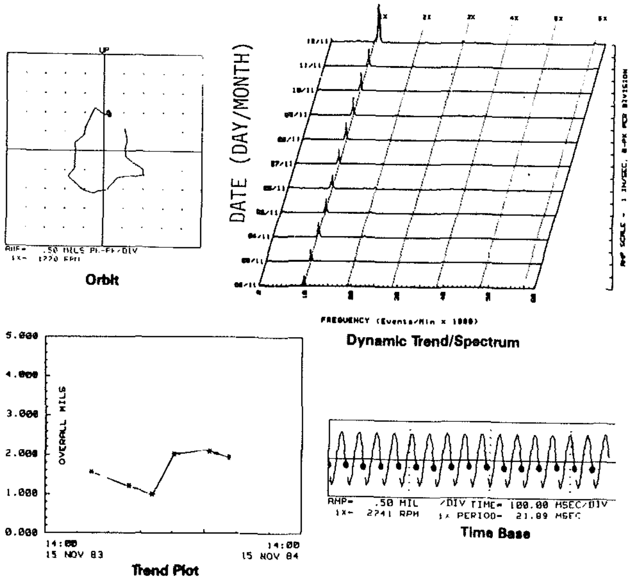


Fig. 5 Examples of displays on computer screen of a periodic vibration monitoring system

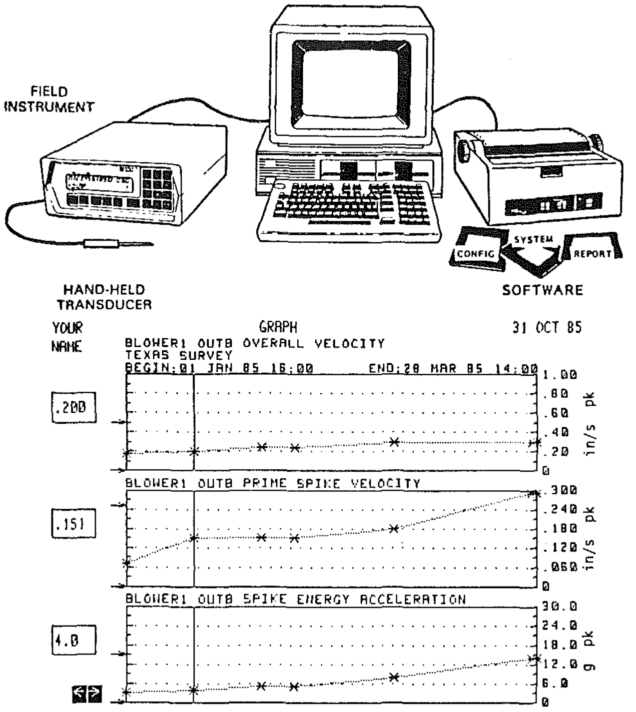


Fig. 6 Portable instrument with host computer and samples of dynamic data plots

tion data acquisition, information processing, and trend analysis/malfunction detection, to take into account present needs and possible future modification. If these decisions can be made early in a project, before rotating machinery is in operation, ensuring the effectiveness of a Preventive/Predictive Maintenance (PM) system becomes much easier. Sometimes, however, it is desirable to install transducers permanently and monitor the machine on a periodic basis, then later adopt a permanent monitoring system. It is of paramount importance to remember that no vibration monitoring/PM system can be effective unless the correct type of vibration transducer (in the correct place and orientation) is used on the specific categories of rotating machinery. Even the most sophisticated signal processing/analysis procedures (including computer-aided intelligent systems) will yield worthless information if the measured vibration signal contains no information about the condition of the machine.

Selection of information management systems (including diagnostic instrumentation) can be made on a broader basis, generally by dividing rotating machinery into categories such as "critical," "essential," and "balance-of-plant." Large and expensive elements of plant machinery which cannot be spared, as well as those machines which would create major hazard or production loss if they suddenly become unoperational – potential weak points of the system – are classified as critical machines. The main factor in all PM decisions should be the vulnerability of production to failure of a given machine (Fig. 3). Experience suggests that maintenance engineers tend to overemphasize the capital cost factor when making these categorizations; the degree of criticality of certain relatively inexpensive machines within a power-generation cycle or chemical process can therefore be understated.

Vibration monitoring and information processing systems for Preventive/Predictive Maintenance can be broadly categorized into the following generic types:

1 Portable vibration meters and analyzers – stand-alone instruments not associated with an external data storage or analysis system.

2 Portable field data acquisition instruments with digital memory, associated with an information management computer system and software.

3 Permanently installed vibration monitoring systems with on-line computer-based data storage, analysis and trending capabilities; top end systems can include sophisticated data analysis/diagnostic functions and automatic shutdown devices (Fig. 4).

All three generic types of vibration data-acquisition systems follow a similar philosophy of generating information which will permit experienced engineers to detect machinery problems before mechanical failure occurs, but late enough to optimize the maintenance function. Except in cases where only a very small number of machines is included in a periodic monitoring program, a computer-based system which manages the database and plots the vibration trends is likely to be more cost effective. The real cost of manually logging and plotting data acquired with simple hand-held vibration meters is often overlooked by plant management.

Portable field data acquisition systems can be very effective for periodic monitoring of noncritical rotating machinery. It is important, though, to design the system software carefully to avoid the pitfall of generating huge quantities of paperwork unnecessarily.

Such equipment really falls into two distinct types – portable field data acquisition instruments for condition monitoring by trend analysis alone is one alternative at lowest cost. Similar systems for trend analysis and direct diagnostic functions are another type. The latter provides orbit/waveform, spectrum analysis, and spectral trending capabilities.

Figure 5 illustrates output of a trend analysis instrument system working with a low-cost computer system and easily used software. Using a specially developed hand-held accelerometer probe, vibration measurements can be taken in the field on a predefined route that is programmed into the field instrument. The computer, with appropriate data compression/correlation software algorithms, stores large trend files and outputs mixed parameter trend graphs that indicate changes in overall vibration level in two distinct frequency bands. Most rotating machine malfunctions are characterized by low frequency. The high-frequency band contains information on rolling element bearing condition. Machine parameters and process variable data can be entered into the vibration database using a keypad on the field instrument, permit-

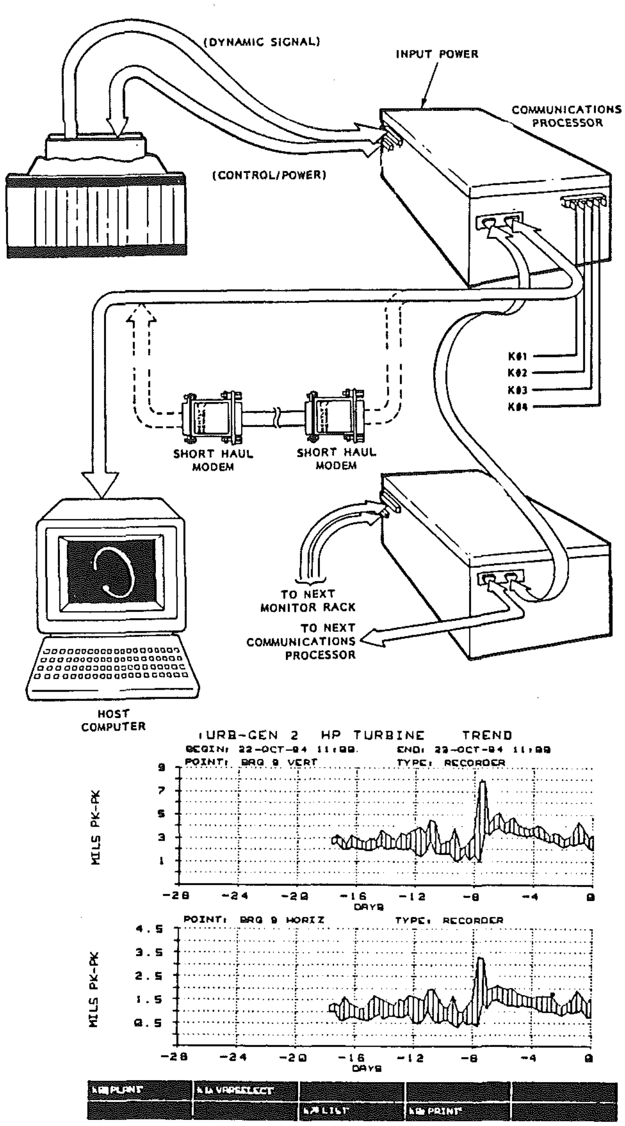


Fig. 7 Continuous on-line vibration monitoring system with example of trend display on computer screen

- Condition Monitoring
  - Process Condition Monitoring
    - At Operating Conditions
      - Efficiency
      - Flow Rate
      - Gram Mol Weight/ Viscosity/Density
      - Rotative Speed
      - Electric Current/ Voltage
      - Suction/Discharge Pressure
      - Temperature
  - Machine Condition Monitoring
    - At Start-up and Shutdown
      - Shaft Radial Vibrations
      - Shaft Torsional Vibrations
      - Shaft Radial Position
      - Thrust Position
      - Casing Vibrations
      - Torque
      - Horsepower
      - Heat Distribution
    - At Operating Conditions
      - Lube/Seal Oil Pressure, Flow and Temperature
      - Cracking
      - Noise
      - Oil Analysis (Wear Particles)
    - At Nonoperating Conditions
      - Nondestructive Testing
      - Impulse Testing
      - Perturbation Testing

Signals Qualifying for Warning/ Alarm Devices

Fig. 8 Condition monitoring parameters

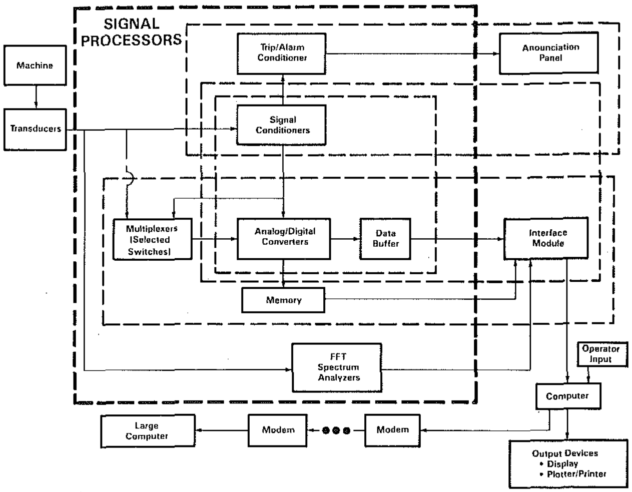


Fig. 9 Condition monitoring system

ting valuable correlations to be made between vibration trends in other variables.

Figure 6 shows a periodic monitoring and diagnostic system using a portable field data-acquisition instrument and computer-based information-processing facility. This system allows analysis to be performed on overall vibration values (displacement and velocity). However, significant vibration analysis functions are also provided, including orbit/ waveform analysis and spectrum analysis using a true Fast Fourier Transformation (FFT) method. It is also possible to measure and trend the synchronous (1 ×), as well as 2 ×, 3 ×, etc., vibration component amplitudes and phases when a Keyphasor is used. Correctly applied, this system provides a means of detecting and analyzing specific machinery problems at a very cost-effective level.

Figure 7 presents an example of continuous (on-line) vibration monitoring system with an example of the computer display. Figure 8 gives a list of condition monitoring parameters. Figure 9 illustrates (schematically) the flow of information in a modern continuous vibration monitoring sysytem with on-line computer. The long-term storage capacity of the computer has to be big enough to store large amounts of data if dynamic vibration samples are collected for spectrum analysis during machinery startup and shutdown. Systems of this type, therefore, become very expensive, especially if custom-written software is required.

Periodic monitoring systems cannot provide any continuous machine protection capability. The decision to stop the machine must be taken by the maintenance personnel on the basis of periodic measurements and their analysis. Continuous monitoring systems used for critical machines (Fig. 4) usually include troubleshooting devices which are activated when vibration amplitudes exceed the preset levels. Using recommended vibration tolerance standards as guidelines, the trip warning/alarm levels can be specified in advance only in very crude terms. It is better practice to establish normal vibration levels during early operation of the machine when it has a known good condition. Using this approach, the influence of the machine's foundation and variables, such as load, can be fully accounted for. Manufacturer's recommendations for alert/trip vibration settings are an important basis for establishing alert/trip values, but experience may lead to refinement of these settings, especially regarding trip levels. If active trips are used, special attention should be paid to the integrity of the system which initiates the trip signal. In the experience of the authors, designs favoring the total integration of the computer into the machine protection function are not sufficiently reliable using current computer technology.

Table 1 Baseline data

WHAT?	WHEN?	HOW?
Basic Geometrical/ Kinematical/Physical Parameters Analytical (Modal) Models Dynamic Parameters • Natural Frequencies • Mode Shapes • Damping • Sensitivity to Various Influences • Stability Margin • Load-to-Vibration Ratios • Vibration Limits	• Prototype Testing • Acceptance Testing • Commissioning • Normal Operation Continuous or Periodic Monitoring	• Transient Data (Start-up, Shut-down) • Synchronous Perturbation (Controlled Unbalance) • Steady-State Data (Operational Speed) • Nonsynchronous Perturbation • Impulse Testing

A better approach is to use conventional well-proven analog monitors with analog-to-digital signal conversion at the monitor rack and a communications processor (with memory) feeding information to the computer. Methods have been developed to permit inexpensive conversion of existing monitoring hardware. This would allow upgrading in the field for users who wish to install an on-line continuous vibration monitoring system for Preventive/Predictive Maintenance without necessitating complete replacement of existing monitors. It must be remembered that in some instances effective diagnosis of machinery problems, especially on complex critical machines, can require powerful dedicated analysis techniques which are difficult to achieve with on-line monitoring systems.

The use of dedicated data reduction equipment and data recorded with high-quality tape recorders continues to be an essential function. In the future, the powerful abilities of such equipment will be available on-line. At present, there is some compromise involved if these features are integrated into an on-line system.

It can often be demonstrated that even for large turbine-generator sets on conventional and nuclear power plants, the use of cost-effective on-line monitoring systems with good operator information, hard-wired primary alarms, and trend analysis, together with off-line periodic analysis using a specialized diagnostic system based on tape-recorded transient data, offers the best solution.

Machine Malfunction Diagnosis

Advances in instrumentation and computer information processing technologies offer machinery users an added dimension for their maintenance programs. As more and more information on the machine condition is available, the proper interpretation of this information is becoming increasingly important for accurately analyzing machine behavior, diagnosing malfunctions, and undertaking corrective action. Interpretation of operating data and optimization of monitoring procedures can only result from a thorough understanding of machine dynamic behavior. At this stage, the level of "know how" of the technical personnel determines success or failure of the whole Predictive Maintenance programs.

There are six basic principles creating the foundation for machine malfunction diagnosis methodology, applicable for both periodic and continuous monitoring:

1 Know the basic mechanical characteristics of the machine.

2 Know the types of malfunction mechanisms that the machine is likely to incur and their corresponding symptoms.

3 Monitor the key parameters that will indicate a change in condition of the machine.

4 Reduce diagnostic data into interpretable formats.

5 Understand historical and related events that may have caused the change in machine condition.

6 Take prompt action, based on the machine diagnosis.

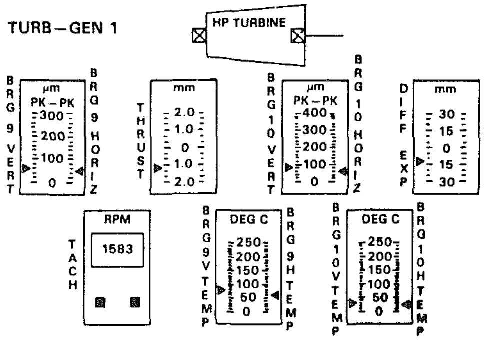


Fig. 10 Typical display of overall values of vibration, position, temperature, speed, and other parameters monitored continuously

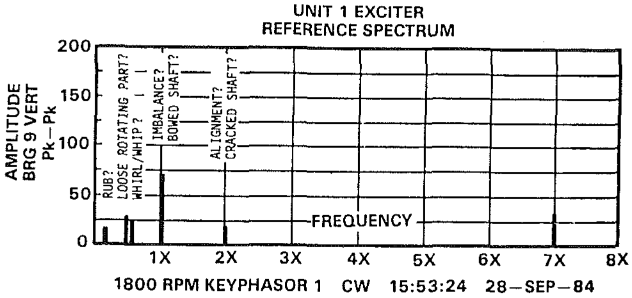


Fig. 11 Spectrum display of rotor response amplitude versus frequency of vibration, indicating existence of subsynchronous and supersynchronous components (1 × is synchronous response with frequency equal to the shaft rotative speed)

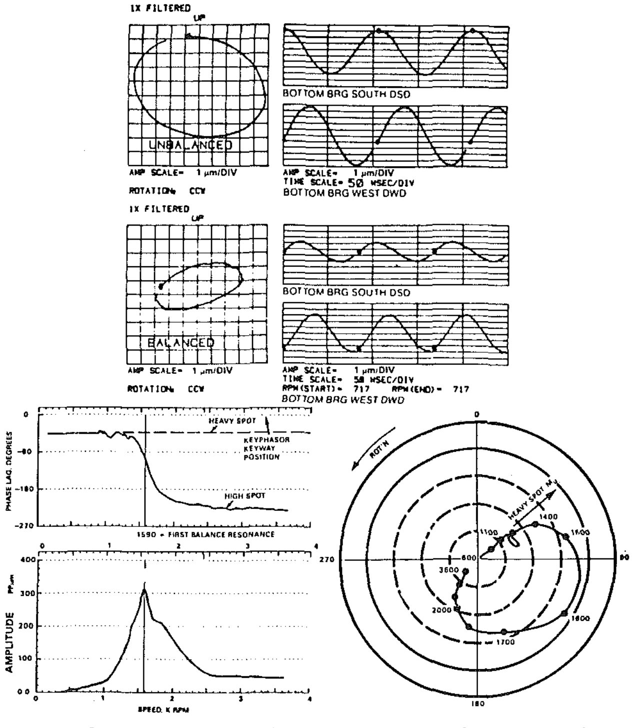


Fig. 12 Orbits, Bodé, and polar plots of phase and amplitude of vibration filtered to 1 × (synchronous) frequency—useful tools for rotor balancing

Basic Characteristics of the Machine. Any machinery analysis requires that the diagnostician first be familiar with the physical parameters of the machine, i.e., size, weight, clearances, etc. However, understanding the basic mechanical factors of the machine goes beyond knowing these physical

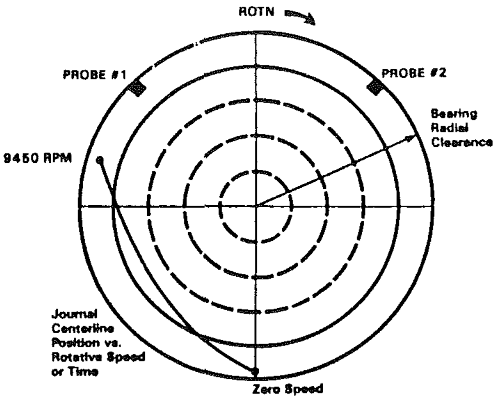


Fig. 13 Journal centerline positions versus rotative speed

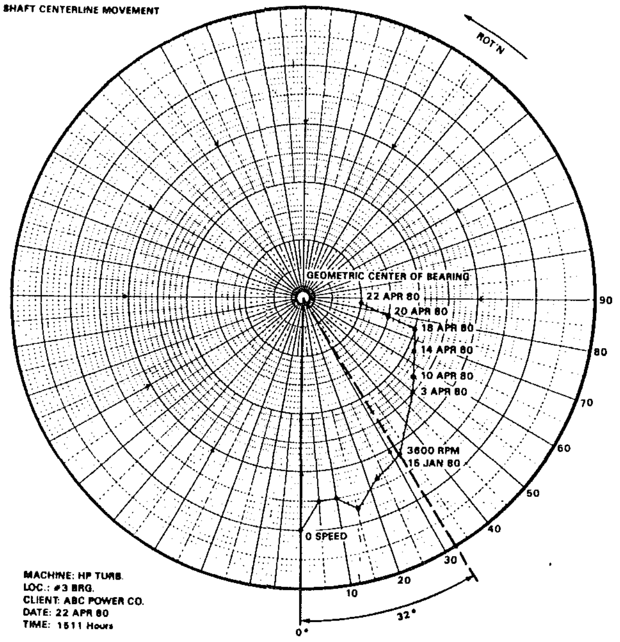


Fig. 14 Journal centerline positions versus time

characteristics. It involves understanding the machine's fundamental dynamic behavior during steady-state and transient conditions. To achieve this goal, process data and vibration data (Fig. 8) must be acquired during acceptance testing and commissioning, by using various methods, such as unbalance response testing (synchronous perturbation), nonsynchronous perturbation testing, and impulse testing, to document the rotor response during steady-state (on-line with partial and full load) and transient (startup and shutdown) conditions. This process provides baseline data (Table 1) for insight to:

1 the natural frequencies and mode shapes of the machine train;
2 the amount of damping in the system;
3 the machine's susceptibility to instability and sensitivity to various influences; and
4 the load-to-vibration relationships and vibration limits.

Analytical models of the dynamic behavior of the machine – particularly the models based on a modal approach – are of tremendous value in future diagnostic procedures. On large, critical machines, or machines that have had a history of vibration problems, it is desirable to incorporate these analytical models into the machine's database.

The analytical and testing baseline information is necessary for comparing the machine's previous and present condition when machine malfunctions occur. It also can be used after a problem is fixed to verify the validity of the correction.

Most of the baseline data and process/operational requirements as well as a comprehensive prediction profile of the machine should be provided by the machine manufacturer. Also, a monitoring system, or at least the prediction of transducer locations for diagnostic purposes, should be included as a part of the design of all rotating machines. Monitoring programs alone do not improve the reliability of machines. Design improvements based on knowledge gained from monitoring, however, can contribute positively to Preventive/Predictive Maintenance systems.

Key Monitoring Parameters. Monitoring of the machine operation and its process variables should be closely connected with the simultaneous monitoring of the machine vibrational behavior. The operation of machines is a dynamic process and the significant process functional data must include dynamic components of the quantities monitored.

Changes in process variables or changes in bearing temperature may represent the first warning signals of impending vibrational problems. Vibrational responses, which are symptoms of machine malfunctions, carry a lot of information about the character of these malfunctions. Both process variables and vibration data should be trended, ideally in cross-correlated formats.

Knowledge of the basic machine characteristics and the types of malfunctions that may occur makes it possible to determine the monitoring parameters that will provide the most insight into the condition of the rotating machine. Five vibration-related parameters are commonly measured on critical rotating machines under steady-state (on-line) and transient (startup and shutdown) conditions:

1 magnitude of dynamic parameters (Figs. 7, 10) for indicating presence and severity of a vibration problem;
2 frequencies of dynamic motion components for insight into the forcing functions, and the character – self-excited or excited – of the observed motion (Fig. 11);
3 phase angle of the rotor precessional motion relative to rotative motion. Phase angle is one of the most important and meaningful parameters for rotating machine malfunction identification (Fig. 12);
4 journal centerline position for insight into the specific location of the rotor with respect to the stationary components of the machine (Figs. 13, 14);
5 wave form and orbital path of dynamic motion (rotor precession and its direction) for insight into the nature of the malfunction (Figs. 5, 12).

Instrumentation for Obtaining Information on the Key Monitoring Vibrational Parameters. Instrumentation is presently available for measuring the key monitoring parameters. The basic component of the instrumentation is the vibration transducer, which measures the mechanical motion of the machine. A variety of transducers are available, but the purpose of any transducer is the same: to provide vibration measurements that give the most meaningful and accurate information on the dynamic condition of rotating machinery.

There are three widely used vibration transducer types: eddy current noncontacting proximity probes, velocity transducers (usually based on the electromagnetic principle), and accelerometers (most often of piezoelectric type). The ranges of application of these transducers are presented in Fig. 15. An example is used for a rotor displacement of 50 μm and machine operational speed 3000 rpm. High-frequency vibration components yield very strong acceleration signals. However, for low-frequency vibrations, the accelerometer

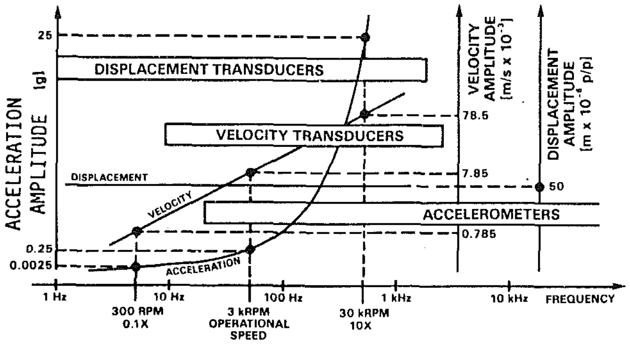


Fig. 15 Sensitivity and ranges of application of displacement, velocity, and acceleration transducers versus frequency

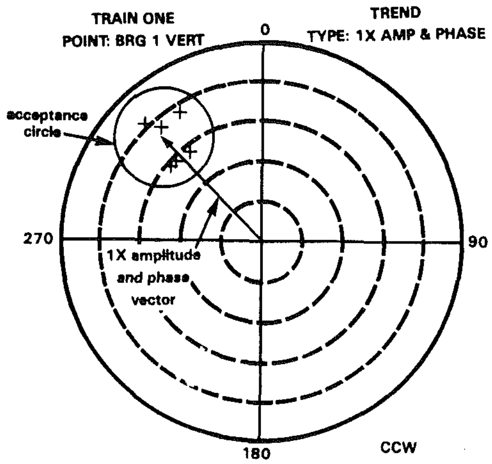


Fig. 16 Trend monitoring of maximum or operational response vector (for 1 × or 2 ×) and acceptance circle on the polar plot

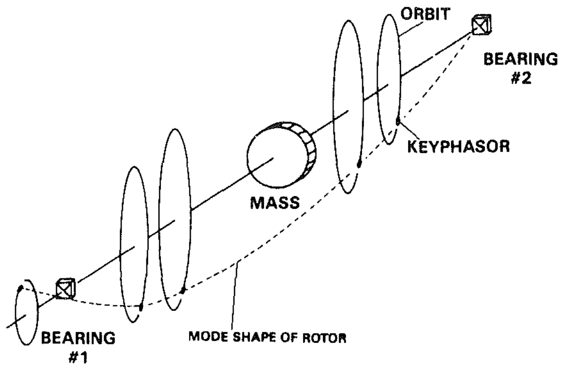


Fig. 17 Application of orbit presentation for rotor mode shape identification

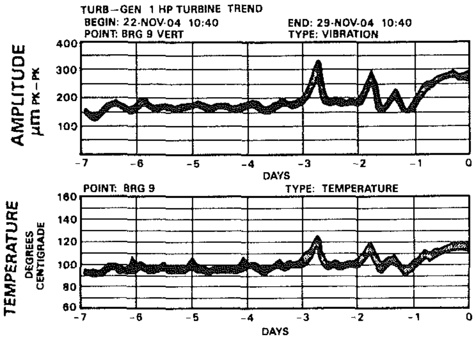


Fig. 18 Trend plots of shaft 1 × amplitude of vibration and bearing temperature *(maximum and minimum values)*

signal is too weak to be meaningful, especially when electronic integration has to be applied. Since many damaging malfunctions of rotating machines are associated with subsynchronous vibrations, which occur in low-frequency range, the application of proximity transducers is the most appropriate for machinery diagnostics and protection. Proximity transducers also provide vital information on the steady-state position and slow-roll motion of the shaft in its bearings.

The generally accepted practice is to utilize two XY proximity probes, mounted 90 deg apart, on or near each radial bearing for measuring shaft vibration and position. A proximity probe is a noncontacting device, which measures the displacement motion or position of an observed surface relative to its mounting surface. In the United States, the American Petroleum Institute (API) has adopted a standard, entitled "Noncontacting Vibration and Axial Position Monitoring System" (Standard Number 670). This standard outlines the system requirements for installing transducers in the 90 deg configuration on process compressors and their drivers. In addition to the radially oriented vibration transducers, the standard calls for two axially oriented proximity transducers to monitor and alert about machine thrust bearing problems. These transducers often are tied to automatic trip when a danger condition exists. The transducer installation practices outlined in the standard are appropriate for the monitoring and protection of turbine generators and other critical machinery as well.

One of the very important transducers called for in the standard is a phase angle reference transducer Keyphasor®. This transducer is mandatory to facilitate the machine diagnostic process. It is used to monitor rotative speed, 1 × (synchronous) response phase, as well as the phases of the fractional and/or multiple-type components of the vibrational response. The Keyphasor utilizes a proximity transducer to observe a once-per-revolution event, usually a notch (keyway) or projection, on the shaft. The transducer provides a voltage pulse, which then can be used to trigger the Z-intensity axis of an oscilloscope, a vector filter, tachometer, and other diagnostic instruments. The Keyphasor transducer ties the lateral vibration measurements to the rotational motion of the machine. Since the orbit is presented on an oscilloscope as simply a rapidly moving dot, representing the centerline motion of the shaft, the addition of a Keyphasor pulse interposed on the orbit tells the diagnostician the location of the shaft at the precise moment the keyway is under the Keyphasor transducer. This information is extremely valuable for balancing a machine and in documenting the rotor response with Bodé and polar presentations (Figs. 12, 16).

The XY proximity transducers help to determine the mode shape of the rotor. With "mode identification transducers," located at various longitudinal positions along the rotor, and a Keyphasor reference, it is possible to identify nodal point regions and rotor mode shapes for various rotative speeds. The use of mode identification proximity transducers ensures adequate machine protection by overcoming the potential danger, when only a single set of XY proximity probes is used, of locating the probes at a nodal point along the rotor.

Very flexible support structures may be encountered on large fans and some gas turbines. For such machines, both the shaft motion and the casing motion should be monitored. In these circumstances, it is possible to use an eddy current displacement probe and a seismic transducer in combination as a "Dual Probe." The output from the seismic transducer (usually a velocity transducer) is electronically integrated and vectorially summed with the eddy current transducer signal, producing a rotor absolute vibration signal. The rotor relative

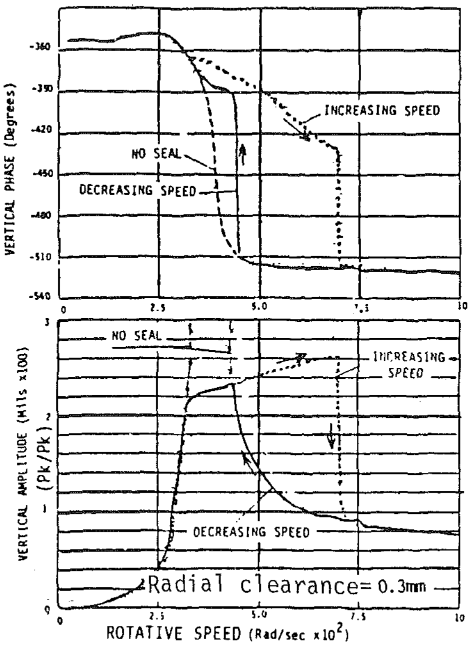


Fig. 19 Bodé plots of the synchronous (1 ×) response of a rotor rubbing against the stator during startup and shutdown [1]

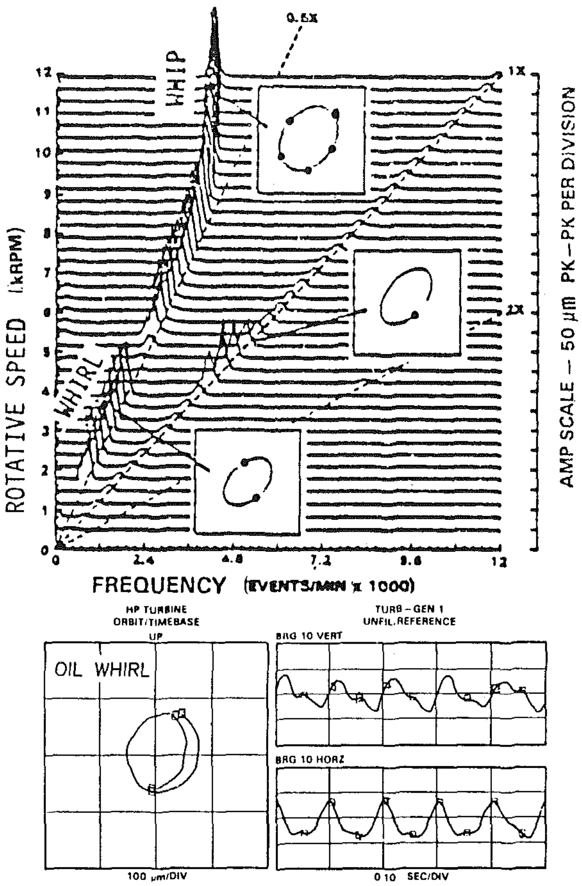


Fig. 20 Spectrum cascade, orbits, and time base presentation of rotor vibrational response, indicating synchronous (1 ×) vibrations due to unbalance, oil whirl (with frequency close to 1/2 ×), and oil whip with constant frequency close to the first balance resonance value [2]

vibration signal and the seismic casing vibration signal are also available for diagnostic purposes.

Data Reduction Into Interpretable Formats. The manner in which vibration data are processed and reduced into interpretable formats can determine the success or failure of any malfunction identification effort. Computerized diagnostic instrumentation provides reduced steady-state (on-line) and transient (startup and shutdown) vibration data in a variety of formats, including both time domain and frequency domain.

Steady-state (on-line) vibration data can be reduced in the following formats:

1 Orbits and time base plots for examining the magnitude, frequency, phase angle, and shape of the shaft precessional motion and its filtered frequency components. Proper interpretation of these plots yields insight into the nature of the vibration and makes it possible to determine whether the precession is of a forward or reverse nature. For any steady-state or transient conditions, the orbit and time base plots can be observed on the oscilloscope or obtained later during data reduction procedure (Figs. 5, 12, 20, and 22).

2 Comparative spectrum plots for tracking changes in spectral content over a period of time. New frequency components or changes in the magnitude of previous data at a particular frequency are provided by these graphs (Figs. 5 and 11).

3 Journal centerline position for observing changes in the steady-state position of the rotor in the bearings. This information can provide warning of major changes in the alignment state of the machine and gives an indication of bearing wear (Fig. 14).

4 Mode shape plots for providing estimates for the internal clearances between the rotor and stator as well as estimates of the nodal points along the rotor (Fig. 17).

5 Trend plots for analyzing changes of observed data as a function of time to observe the rates of modification in the operating parameters of the machine. Trends of vibration should be correlated with process variables to determine the effects of process variable changes on the dynamic motion of the machine (Figs. 6, 18). Trending of vibration data is also vital to detect cracked shafts.

Transient vibration data taken during startup and shutdown can be reduced in the following formats:

1 Bodé and polar plots of 1× synchronous vibrations for documenting the location of the balance resonances, mode shapes of the rotor, and the synchronous amplification factor of the machine provide a data base for solving complex multiplane field balancing problems and assist in the diagnosis of rotor-to-stator rubs and cracked shafts (Figs. 12, 16, 19). The plots present 1× data in uncompensated or compensated form (compensation eliminates slow-roll response vector). 2× polar plots are used for cracked shaft detection (Fig. 16). Polar plots of rotating machinery vibration data have a major advantage in the emphasizing of phase shifting trends. Both types of plot can be used to assess the synchronous damping of the rotor/bearing system.

2 Spectrum cascade for examining the synchronous vibration components as well as the sub- and supersynchronous components at various machine speeds. Certain types of instabilities, such as oil whirl/whip, and symptoms of other malfunctions, including cracked shafts and rubs, can be detected from this presentation (Figs. 20, 21, and 22).

3 Journal centerline position for tracking the location of the shaft centerline with respect to the geometric center of a bearing for observation of shaft alignment (Fig. 13).

Conclusions

1 Vibration monitoring for Predictive Maintenance is an important part of the overall condition monitoring process. It is vital to integrate vibration measurements with other machine parameter and process variable measurements.

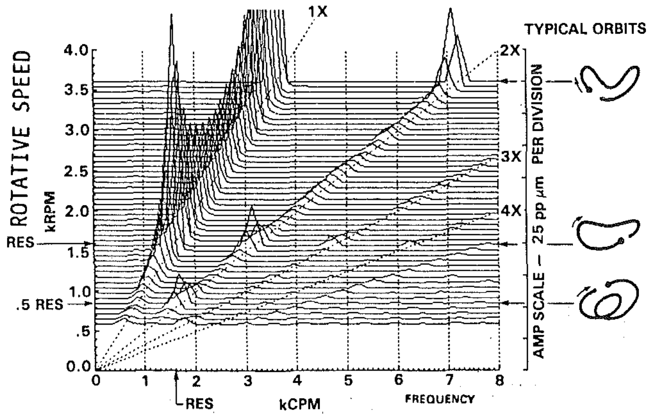


Fig. 21 Spectrum cascade and orbits of a cracked shaft [3]

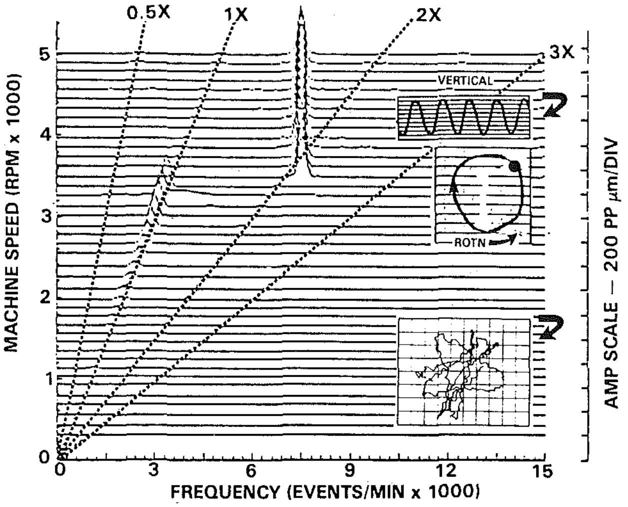


Fig. 22 Spectrum cascade and shaft orbit patterns showing multiple impacting rotor/seal rub at low speed (1 × rub) and a full annular self-excited backward rub at higher speeds when the contact with seal is maintained continuously. Frequency of full annular backward rub corresponds to the natural frequency of the rotor/bearing/seal system [1].

2 Advances in the sophistication of computer-based vibration monitoring systems in the near future will enable maintenance engineers to predict machinery problems with increased effectiveness. The role of machinery vibration specialists will continue to be important to ensure that obscure machinery problems are correctly identified.

3 When setting up a vibration monitoring program, the first step should always be to categorize various types of rotating machines, study the design features, and use this information when deciding on the monitoring philosophy.

4 Vibration monitoring transducers should be selected very carefully, paying particular attention to the expected vibration characteristics of the machine, especially the frequency components and the relative amplitudes between rotor and casing vibration.

5 The decision whether or not to monitor smaller, less critical machinery permanently should be made after considering the true economics of employing a portable intermittent monitoring scheme and the personnel required to operate it.

6 It should be recognized that continuous monitoring of machinery will always provide better protection than any type of periodic scheme; however, the adoption of computer-based information management systems and portable digital data sampling devices can significantly improve the effectiveness of the periodic monitoring approach.

7 Whether continuous or periodic monitoring schemes are adopted, using the best available hardware and software can generate detailed information at very little extra cost in the long term when compared with labor intensive manual schemes. It should always be remembered that using a simple data-collection philosophy, typified by the logging of overall vibration level on a machine, will generate data which cannot be used for diagnostic purposes other than trend analysis. Collection of vibration dynamic data typified by the examples given in this paper can enable precise identification of machinery defects to be achieved. Thus, specialized diagnostic equipment does not need to be employed except when very complex problems arise and repetition of data acquisition can be minimized.

8 In the future, vibration monitoring for Predictive Maintenance will involve the application of more computer-based systems with developments in software occurring very rapidly. Although software production will become more expensive, real hardware costs will decrease. As machinery repair costs increase and the impact of lost production through unplanned machinery outages becomes more expensive, the application of well-engineered condition monitoring schemes can be predicted to increase in importance.

9 Future computerized vibration monitoring systems used in Preventive/Predictive Maintenance programs will include more elements of diagnostics, leading eventually to fully automated machinery protection systems. It is unlikely, however, that sufficiently powerful computerized systems will be affordable for use in automated diagnostics for several years, and even then they will have limitations.

References

1 Muszynska, A., "Synchronous and Self-Excited Rotor Vibrations Caused by a Full Annular Rub," *Proceedings of the 8th Machinery Dynamics Seminar on Modal Analysis and Vibration Diagnostics in Machinery*, Halifax, Nova Scotia, Canada, Oct. 1984.

2 Muszynska, A., "Whirl and Whip – Rotor/Bearing Stability Problems," *Journal of Sound and Vibration*, Vol. 110, No. 3, 1986.

3 Bently, D. E., and Muszynska, A., "Detection of Rotor Cracks," *Proceedings of Texas A&M 15th Turbomachinery Symposium*, Corpus Christi, TX, Nov. 1986.

Aeroelastic Behavior of Low-Aspect-Ratio Metal and Composite Blades

J. F. White, III[1]
Research Assistant.

O. O. Bendiksen
Assistant Professor.

Department of Mechanical and Aerospace Engineering,
Princeton University,
Princeton, NJ 08544

The aeroelastic stability of titanium and composite blades of low aspect ratio is examined over a range of design parameters, using a Rayleigh-Ritz formulation. The blade modes include a plate-type mode to account for chordwise bending. Chordwise flexibility is found to have a significant effect on the unstalled supersonic flutter of low-aspect-ratio blades, and also on the stability of tip sections of shrouded fan blades. For blades with a thickness of less than approximately 4 percent of chord, the chordwise, second bending, and first torsion branches are all unstable at moderately high supersonic Mach numbers. For composite blades, the important structural coupling between bending and torsion cannot be modeled properly unless chordwise bending is accounted for. Typically, aft fiber sweep produces beneficial bending-torsion coupling that is stabilizing, whereas forward fiber sweep has the opposite effect. By using crossed-ply laminate configurations, critical aeroelastic modes can be stabilized.

Introduction

Aircraft gas turbine engine designers continually strive to improve the performance of new and existing engines. Advances in aircraft engine technology, especially in the area of materials technology, have provided designers with many ways of improving engine performance. Fan and compressor stages in high-performance engines can be improved by incorporating thinner blades, higher tip speeds, higher pressure ratios, and advanced materials. Such improvements make fan and compressor stages susceptible to various aeroelastic problems, especially the destructive phenomenon known as flutter. The prediction of flutter boundaries is therefore very important and has become a critical part of the design process of high-performance engines.

Researchers traditionally have modeled fan blades as one-dimensional beam structures. However, thin low-aspect-ratio blades in newer engine designs exhibit vibration characteristics that cannot be modeled accurately using beam theory. Flat plate theory has proven useful in determining the free vibration characteristics of metal and composite blades [1-4]. Other investigators have used thin shell theory for studies of blades with camber and twist [5-7]. Both methods reveal the existence of plate-type chordwise bending modes that are thought to be important in the analysis of flutter and fatigue problems in low-aspect-ratio blades.

For most turbomachinery aeroelastic studies, however, the structural modeling of the blades is kept very simple because of the complexity of the aerodynamic calculations. For example, several studies have considered only single-degree-of-freedom oscillations [8-10]. Analytical studies of blade flutter based on beam theory and uncoupled normal modes were considered in [11, 12]. Coupled mode effects were studied by Bendiksen and Friedmann [13, 14] with a simple typical section model for the blades. Their typical section model showed that coupling between bending and torsion has a pronounced effect on the flutter boundaries.

In the present paper, flat plate theory is used to study the aeroelastic stability of both metal and composite fan and compressor blades. Of particular interest is the interaction between different modes of motion and the effects of the interaction on the aeroelastic behavior of typical low-aspect-ratio blades. The results from this study illustrate the importance of including chordwise deformation (camber bending) in the structural model.

Equations of Motion

It is assumed that the blades can be modeled as fully cantilevered flat plates. Figure 1 shows a typical blade and the axis system to be used throughout the analysis. A short study of the effects of relaxed stiffness at the blade root will be presented later, in which three additional degrees of freedom are added to the the blade root. A brief description of the fully cantilevered model is presented here. For a more detailed description of the aeroelastic model including the additional degrees of freedom at the root, the reader is referred to [15].

A Rayleigh-Ritz procedure with five elastic degrees of freedom is used to formulate the equations of motion for the rotating blade system. The deflection of the flat plate is approximated by a series of assumed mode shape functions

---

[1] Presently with Boeing Vertol Company, Philadelphia, PA.

Contributed by the Gas Turbine Division of The American Society of Mechanical Engineers and presented at the 31st International Gas Turbine Conference and Exhibit, Düsseldorf, Federal Republic of Germany, June 8-12, 1986. Manuscript received at ASME Headquarters February 21, 1986. Paper No. 86-GT-243.

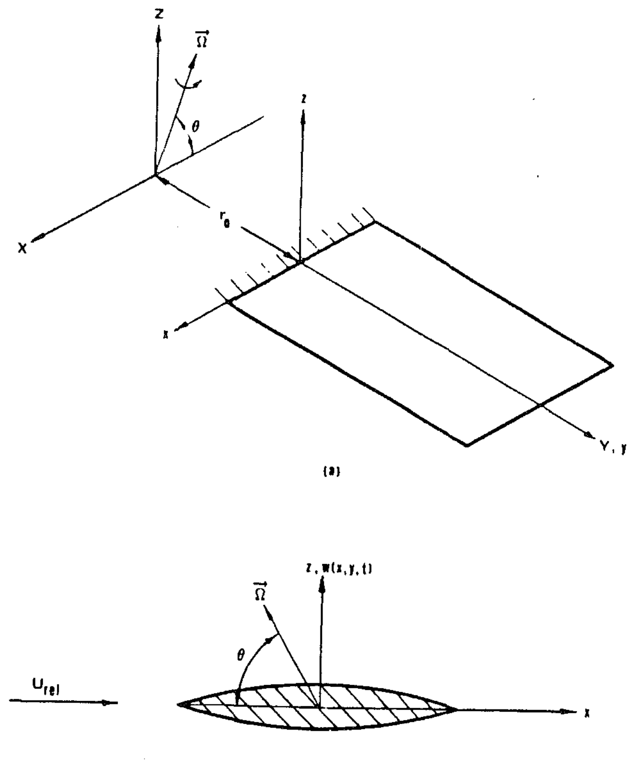


Fig. 1 (*a*) Rotating cantilevered blade system; (*b*) blade section (looking outboard) with *w*(*x*, *y*, *t*) positive up

$$w(x, y, t) = \sum_{i=1}^{5} \gamma_i(x, y) q_i(t) \tag{1}$$

The five assumed mode shapes include functions for first and second bending, first and second torsion, and a plate-type chordwise bending mode. The use of five assumed mode shapes allows the first two or three frequencies and mode shapes to be predicted accurately. This is desirable since the low-frequency bending and torsion modes are most often associated with aeroelastic problems.

The specific mode shape functions used in this analysis are shown below. Note that each mode shape function consists of explicit functions for the chordwise and spanwise distributions of the mode shape, $\gamma_i(\bar{x}, \bar{y}) = \psi_i(\bar{x})\,\phi_i(\bar{y})$.

First bending (1B):

$$\psi_1(\bar{x}) = 1$$

$$\phi_1(\bar{y}) = \cosh(\epsilon_1\bar{y}) - \cos(\epsilon_1\bar{y}) - \eta_1[\sinh(\epsilon_1\bar{y}) - \sin(\epsilon_1\bar{y})]$$

Second bending (2B):

$$\psi_2(\bar{x}) = 1$$

$$\phi_2(\bar{y}) = \cosh(\epsilon_2\bar{y}) - \cos(\epsilon_2\bar{y}) - \eta_2[\sinh(\epsilon_2\bar{y}) - \sin(\epsilon_2\bar{y})]$$

First torsion (1T):

$$\psi_3(\bar{x}) = -\bar{x}$$

$$\phi_3(\bar{y}) = \sin(g_3\bar{y}) + (g_3/f_3)[\cosh(f_3\bar{y}) - \sinh(f_3\bar{y}) - \cos(g_3\bar{y})]$$

Second torsion (2T):

$$\psi_4(\bar{x}) = -\bar{x}$$

$$\phi_4(\bar{y}) = \sin(g_4\bar{y}) + (g_4/f_4)[\cosh(f_4\bar{y}) - \sinh(f_4\bar{y}) - \cos(g_4\bar{y})] \tag{2a}$$

First chordwise (1C):

$$\psi_5(\bar{x}) = \cosh[\tfrac{1}{2}\epsilon_5(\bar{x}+1)] + \cos[\tfrac{1}{2}\epsilon_5(\bar{x}+1)] - \eta_5\{\sinh[\tfrac{1}{2}\epsilon_5(\bar{x}+1)] + \sin[\tfrac{1}{2}\epsilon_5(\bar{x}+1)]\}$$

$$\phi_5(\bar{y}) = \sin(g_5\bar{y}) + (g_5/f_5)[\cosh(f_5\bar{y}) - \sinh(f_5\bar{y}) - \cos(g_5\bar{y})]$$

$$\begin{aligned}
\epsilon_1 &= 1.875104 & \eta_1 &= 0.734096 \\
\epsilon_2 &= 4.694091 & \eta_2 &= 1.018466 \\
g_3 &= 1.902914 & f_3 &= 4.492336 \\
g_4 &= 4.936090 & f_4 &= 6.397264 \\
\epsilon_5 &= 4.730041 & \eta_5 &= 0.982502 \\
g_5 &= 1.750805 & f_5 &= 9.266354
\end{aligned} \tag{2b}$$

The first two functions are the first and second cantilevered beam vibration modes, respectively. The two torsional func-

Nomenclature

a_∞ = speed of sound, standard sea level value
A_{ij} = element of aerodynamic matrix
b, c = blade semichord and chord, respectively
D_{ij} = flexural rigidities of laminated composite plate
E = modulus of elasticity
E_x, E_y, E_s = moduli of elasticity of composite plate (longitudinal, transverse, and shear, respectively)
h = blade thickness
k = $\omega b/U$ = reduced frequency
K_{ij} = element of stiffness matrix
l = blade span
m = mass per unit length of blade
M_{ij} = element of mass matrix
M = Mach number
N_b = number of blades in rotor disk
N_x, N_y, N_{xy} = initial in-plane stress resultants
$\bar{p}$ = $\bar{p}_R \pm i\bar{\omega}$ = complex time dependence exponent (nondimensional)
$[\bar{p}_0(\bar{x}, \bar{y})]$ = complex pressure jump across the reference (zeroth) blade, positive up, nondimensionalized by $\rho_\infty U^2$
$q_i(t)$ = generalized displacements
$\bar{q}_i$ = $\bar{q}_i/b$ = nondimensional amplitude of q_i
Q_i = generalized nonconservative force
r_0 = blade root offset from axis of rotation
s = blade spacing along leading edge locus
U_∞ = air velocity
$w(x, y, t)$ = blade displacement in the z direction
$\bar{x}, \bar{y}$ = nondimensionalized blade coordinates x/b and y/l, respectively
$\gamma_i(\bar{x}, \bar{y})$ = $\psi_i(\bar{x})\phi_i(\bar{y})$ = assumed mode shape functions
θ = blade stagger angle
λ = eigenvalue
μ = $m/\pi\rho_\infty b^2$ = blade mass ratio
ν, ν_x = Poisson ratio
ρ = blade mass density
ρ_∞ = air density
ω = blade vibration frequency, rad/s
Ω = angular velocity of rotor

tions are the same mode shapes as derived by Jensen and Crawley [3]. The last function is a chordwise bending mode consisting of the free–free beam vibration mode $\psi_5(x)$, and a function $\phi_5(y)$ derived by the first author [15].

Using the deflection series given by equation (1), expressions for the kinetic and potential energies of the rotating blade system are formed. The kinetic energy of the rotating cantilevered plate, neglecting Coriolis effects, is

$$T=\frac{1}{2}\iint\limits_{\text{Area}} \rho h\{\dot{w}^2+\Omega^2[(r_0+y)^2+(w\cos\theta+x\sin\theta)^2]\}dxdy \tag{3}$$

The total potential energy of the rotating blade system is assumed to consist of the strain energy due to bending and the potential energy due to forces in the plane of the plate middle surface. Stretching of the plate middle surface is not considered. The strain energy of bending for an anisotropic plate of uniform thickness is [16]

$$\begin{aligned} V_1 = \frac{1}{2}\iint\limits_{\text{Area}} \{&D_{11}(w,_{xx})^2+2D_{12}(w,_{xx})(w,_{yy})+D_{22}(w,_{yy})^2 \\ &+4D_{13}(w,_{xx})(w,_{xy})+4D_{23}(w,_{yy})(w,_{xy}) \\ &+4D_{33}(w,_{xy})^2\}dxdy \end{aligned} \tag{4}$$

Equation (4) specifically applies to composite plates in which the fibers are oriented symmetrically about the middle surface of the plate. Equation (4) is also applicable to metal blades since isotropic materials are special (degenerate) cases of anisotropic materials.

The potential due to in-plane forces is [18]

$$\begin{aligned} V_2=\iint\limits_{\text{Area}} &[N_x(w,_x)^2+N_y(w,_y)^2 \\ &+2N_{xy}(w,_x)(w,_y)]dxdy \end{aligned} \tag{5}$$

Following the reasoning of Leissa et al. [7], the expressions for the initial in-plane stress resultants due to rotation may be approximated by

$$\begin{aligned} N_x &= \frac{1}{2}\rho h\Omega^2\sin^2\theta(b^2-x^2) \\ N_y &= \frac{1}{2}\rho h\Omega^2[(r_0+l)^2-(r_0+y)^2] \\ N_{xy} &\approx 0 \end{aligned} \tag{6}$$

The total potential energy of the system is then

$$V=V_1+V_2 \tag{7}$$

The kinetic and potential energies may be expressed in the form of double summations as

$$T=\frac{1}{2}\sum_{i=1}^{5}\sum_{j=1}^{5}M_{ij}\dot{q}_i\dot{q}_j \tag{8}$$

$$V=\frac{1}{2}\sum_{i=1}^{5}\sum_{j=1}^{5}K_{ij}q_iq_j \tag{9}$$

where M_{ij} and K_{ij} are elements of the system mass and stiffness matrices, respectively.

The equations of motion are formed using equations (8) and (9) in Lagrange's equations, resulting in the following matrix equation

$$[\mathrm{M}]\{\ddot{q}\}+[K]\{q\}=\{Q\} \tag{10}$$

The elements Q_i of the nonconservative force vector are calculated using the unsteady pressure jump distribution over the surface of the airfoil. The external virtual work is

$$\delta W_{\text{ext}}=\sum_{i=1}^{5}Q_i\delta q_i=\rho_\infty U^2\iint\limits_{\text{Area}}[\bar{p}_0(\bar{x},\bar{y})]\delta w\, dxdy \tag{11}$$

and therefore, the nonconservative forces become

$$Q_i=\rho_\infty U^2\iint\limits_{\text{Area}}[\bar{p}_0(\bar{x},\bar{y})]\gamma_i(\bar{x},\bar{y})\, dxdy \tag{12}$$

The unsteady pressure jump distribution $[\bar{p}_0(\bar{x},\bar{y})]$ is given as a complex number to account for both amplitude and phase in a convenient form.

The pressure jump is calculated for unstalled supersonic flow through a cascade of oscillating flat plate airfoils, operating with a subsonic leading edge locus. This flow regime is commonly encountered at the tip sections of large fan stages. The pressure jump distribution over the airfoil surface is obtained from Bendiksen's two-dimensional aerodynamic theory [14, 17] by using a strip theory approach. The aerodynamic theory provides pressure jump distribution for a given cascade configuration, Mach number, frequency of oscillation, interblade phase angle, and a specific mode of motion. The total pressure jump distribution can be written as a series of modal pressure jumps

$$[\bar{p}_0(\bar{x},\bar{y})]_{\text{tot}}=\sum_{j=1}^{5}[\bar{p}_0(\bar{x},\bar{y})]_jq_j \tag{13}$$

With the pressure jump in this form, the nonconservative force terms can be expressed in terms of an aerodynamic matrix

$$\{Q\}=\rho_\infty U^2[A]\{q\} \tag{14}$$

where the elements of matrix A are given by

$$A_{ij}=\iint\limits_{\text{Area}}\gamma_i(\bar{x},\bar{y})[\bar{p}_0(\bar{x},\bar{y})]_j dxdy \tag{15}$$

Substituting equation (14) into equation (10) results in the following system equations of motion

$$[\mathrm{M}]\{\ddot{q}\}+[K]\{q\}=\rho_\infty U^2[A]\{q\} \tag{16}$$

Method of Solution

Equation (16) is transformed into an aeroelastic eigenvalue problem by assuming simple harmonic motion ($q_j(t)=\tilde{q}_je^{i\omega t}$) and introducing the reduced frequency $k=\omega b/U$. The equations of motion are then nondimensionalized and rearranged into the following form

$$[\bar{K}]\{\bar{q}\}=\lambda^2([\bar{\mathrm{M}}]+[\bar{A}])\{\bar{q}\} \tag{17}$$

The eigenvalues $\lambda^2=(\omega/\omega_{\text{ref}})^2$ may be complex. The reference frequency for all calculations is chosen as the first torsion-free vibration frequency at the reference rotational speed. The integrations for all the matrix elements in equation (17) are done numerically using Gaussian integration techniques.

The assumption of harmonic motion is only valid on the flutter boundary where the eigenvalues of equation (17) are real. If the time dependence is assumed to be of the form e^{pt} (where $p=p_R \pm i\omega$) rather than simple harmonic, it can be shown that

$$\bar{p}=(p/\omega_{\text{ref}})=i\lambda=\bar{p}_R\pm i\bar{\omega} \tag{18}$$

Although the aerodynamic theory used is, strictly speaking, only valid for simple harmonic motion, $|p_R|<<1$ for turbomachinery flutter and p_R as calculated is therefore a good representation of the damping in the system. The stability of the system may be determined from the sign of the real part of the exponent p, with flutter occurring if $p_R>0$, $\omega\neq0$. If $p_R>0$ and $\omega=0$, divergence occurs.

The solution of the complex aeroelastic eigenvalue problem

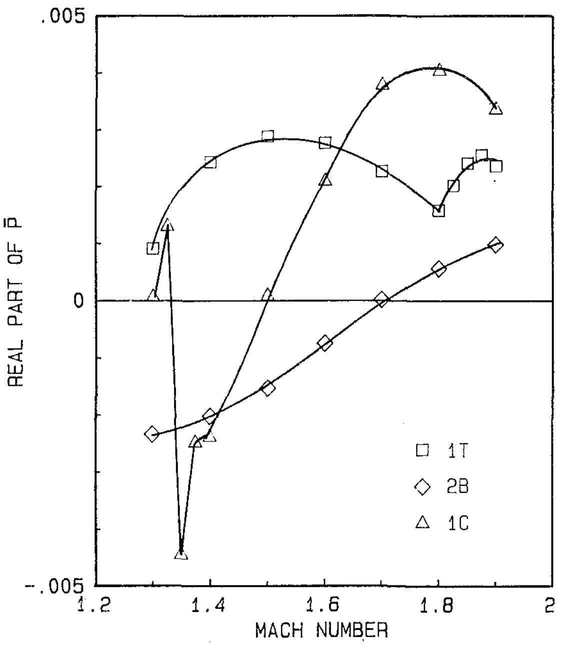


Fig. 2 Real part of exponent $\bar{p}$ versus Mach number for the first torsion, second bending, and chordwise bending branches; titanium blade, 3 percent thickness, aspect ratio = 1

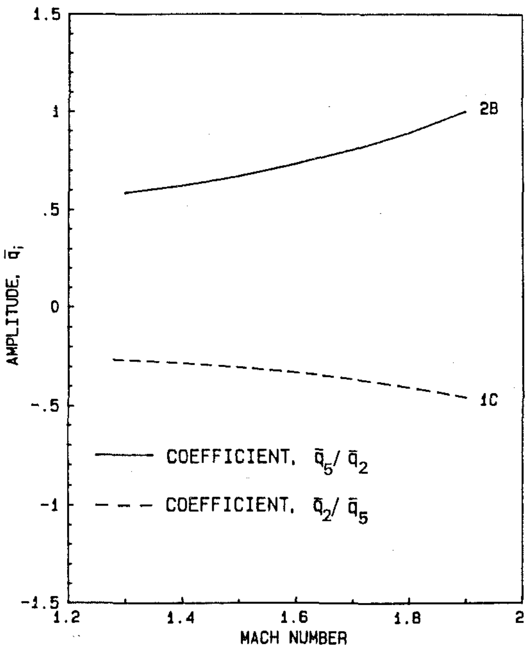


Fig. 3 Mode shape variation versus Mach number of second bending branch and first chordwise bending branch; titanium blade, 3 percent thickness, aspect ratio = 1

is determined as follows: First, the cascade characteristics and rotational speed are specified. It is assumed that the rotational speed varies linearly with the blade tip Mach number as $\Omega = \mathrm{M}_{\mathrm{tip}}\ (\Omega_{\mathrm{ref}}/\mathrm{M}_{\mathrm{ref}})$. It is also assumed that the blade spacing and Mach number do not vary along the span of the blade, and that the angle of attack is small. For low-aspect-ratio blades or blade tip sections these assumptions are believed to lead to reasonable approximations.

Next, an initial value of the reduced frequency for the mode in question is specified. The initial value of the reduced frequency is calculated using the free vibration frequency of the mode at the reference rotational speed. Thus

$$k_0 = (\omega_0 b/U) = (\omega_0 b/\mathrm{M}a_\infty) \tag{19}$$

where the standard sea level value of the speed of sound a_∞ is used in all calculations. Once the initial guess for the reduced frequency has been determined, the aerodynamic matrix is formed according to the procedure outlined previously. The eigenvalue problem of equation (17) is then solved using a standard IMSL subroutine "EIGZC" for complex eigenvalue problems.

For each eigenvalue λ^2 there are two corresponding complex conjugate roots $p_{1,2} = p_R \pm i\omega$. These roots are examined to see if the magnitude of the imaginary part matches the value of the vibrational frequency used to calculate the initial value of the reduced frequency. If the two frequencies do not match, the new calculated frequency is used to update the value of the reduced frequency. The eigenvalue problem is solved repeatedly in this manner until the frequencies converge. The process normally converges within two to three iterations.

The above process is repeated for each mode, each Mach number, each interblade phase angle, and each blade configuration of interest. The stability of each mode can be determined from the real part of the exponent p as discussed previously. For every mode of motion there exists a value of interblade phase angle that produces the least stable value of p_R. Ideally, the entire range of possible interblade phase angles should be searched to determine the least stable value, which would require N_b solutions corresponding to the interblade phase angles $\sigma = 2\pi n/N_b$; $n = 0, 1, 2, \ldots, N_b - 1$. For reasons of computational cost, an automatic search is not incorporated into the present computer program. The least stable values of interblade phase angle were determined manually using the knowledge gained from interblade phase angle studies of a baseline blade configuration (see [15]). All numerical results presented in the next section correspond to data points for least stable interblade phase angles.

Numerical Results

Titanium Blade. The material properties used for the titanium blade in this study are those of the titanium alloy Ti-6Al-4V [19]

$$E = 1.60 \times 10^4 \text{ Ksi}$$
$$\nu = 0.31$$
$$\rho = 0.160 \text{ lb/in}^3$$

The baseline blade configuration is a blade tip section of 3 percent thickness and aspect ratio of one. This configuration is believed to be representative of the class of large fans in present use. Other pertinent dimensions and reference values used are given below:

blade chord c = 0.75 ft
tip radius R_{tip} = 3.489 ft
Ω_{ref} = 400 rad/s
$\mathrm{M}_{\mathrm{ref}}$ = 1.45
N_b = 24
mass ratio μ = 138

Note that only the tip section of the blade is considered here. The tip section is treated as a fully cantilevered flat plate and the inboard blade dynamics is neglected for the time being. The boundary condition at the blade root will be relaxed later in this paper.

The cascade configuration used in computing the aerodynamic matrix is known as Verdon's Cascade A [10]. The cascade has a blade spacing $s \cos\theta/b = 0.8$ (normal to the chord), a blade stagger angle $\theta = 59.5$ deg, and a design Mach number $\mathrm{M} = 1.345$. In the present study the Mach number will be varied from 1.3 to 1.9 while keeping the cascade geometry fixed. It is assumed that the flow direction is always parallel to the blade chords.

The baseline blade configuration was found to be unstable

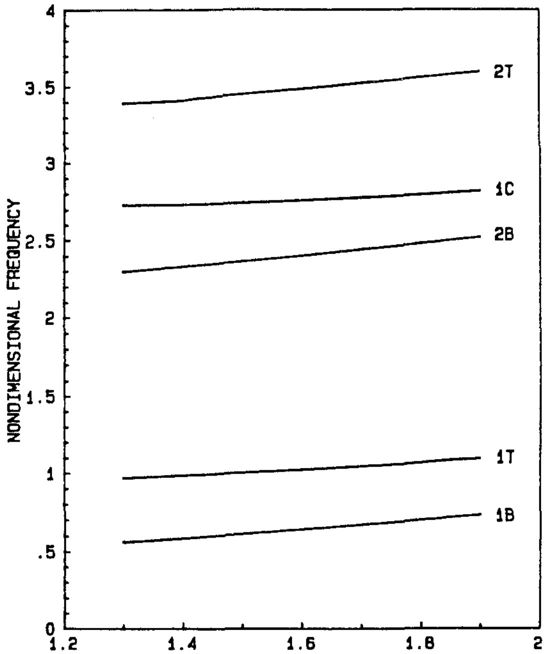


Fig. 4 Nondimensional frequency variation with Mach number of all five modes; titanium blade, 3 percent thickness, aspect ratio = 1

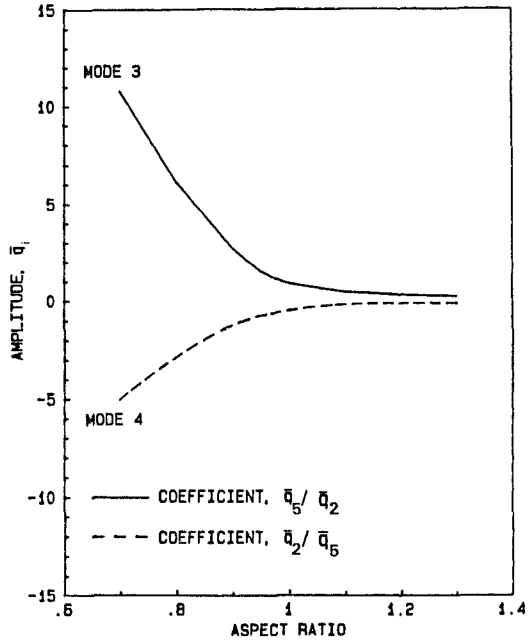


Fig. 6 Mode shape variation of modes 3 and 4 with aspect ratio (second bending and chordwise bending components); titanium blade M = 1.8, 3 percent thickness

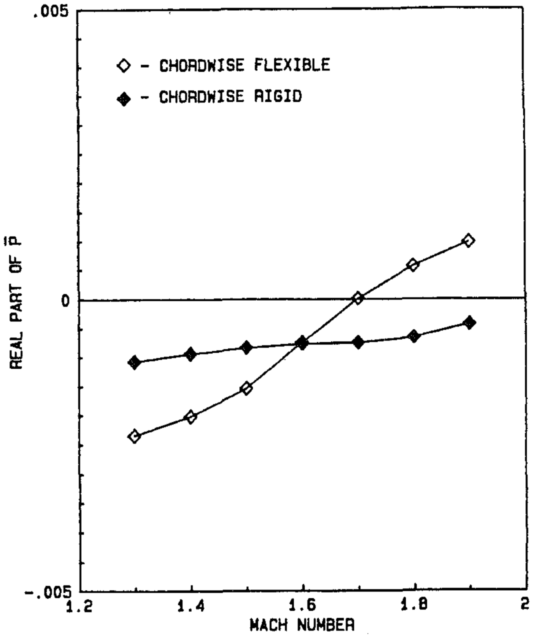


Fig. 5 Effect of chordwise flexibility on the real part of the exponent $\bar{p}$ versus Mach number for the second bending branch; titanium blade, 3 percent thickness, aspect ratio = 1

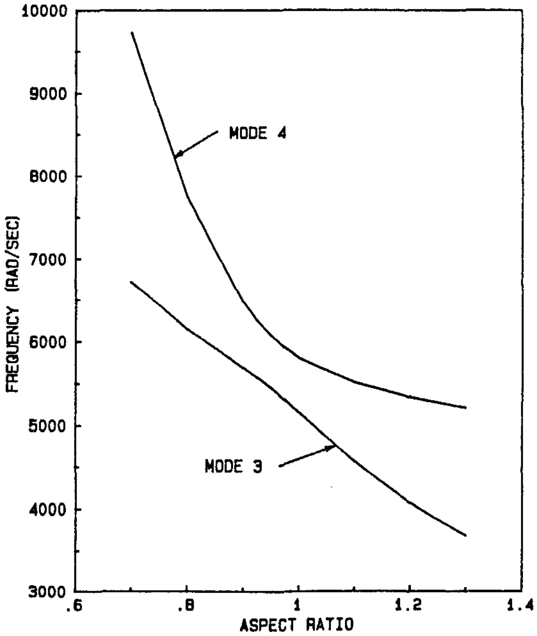


Fig. 7 Frequency variation with aspect ratio of modes 3 and 4; titanium blade, M = 1.8, 3 percent thickness

in the first torsion, second bending, and chordwise bending branches over the Mach number range. These unstable modes are shown in Fig. 2. The first torsion branch is unstable over the entire Mach number range and is essentially a pure torsion mode. It should be noted that in this study the blade is modeled as a uniform plate with the center of gravity at midchord, resulting in a relatively weak coupling between first torsion and the bending modes. The second bending and chordwise bending branches are highly coupled with each other however. The mode shape variations of these two branches are shown in Fig. 3. This coupled mode instability may appear to be similar to classical frequency coalescence flutter of isolated wing theory but this is not the case. Figure 4 shows a plot of the frequency variations with Mach number for all five branches. No signs of frequency coalesence are apparent. It is apparent that the addition of chordwise flexibility to the blade model has a destabilizing effect on the second bending branch of this low-aspect-ratio blade.

As further evidence that the second bending instability is greatly influenced by the chordwise bending participation, calculations are shown in Fig. 5 for a four-degree-of-freedom model. The reduced model does not include an assumed mode for chordwise bending, but is otherwise identical to the five-degree-of-freedom model. The chordwise rigid blade does not show an instability in the second bending branch over the Mach number range of interest.

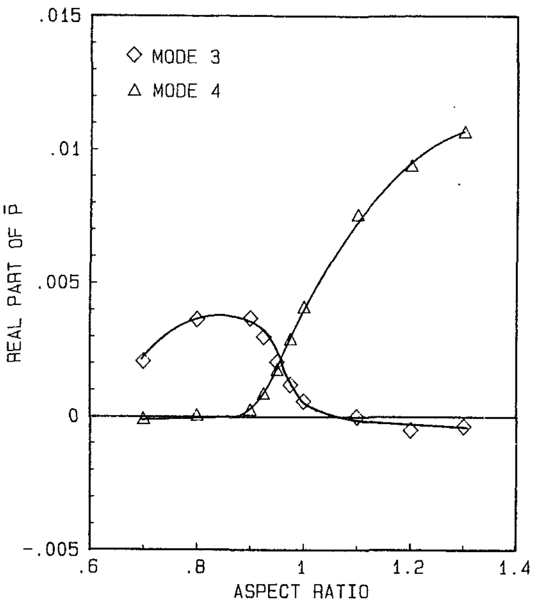


Fig. 8 Real part of the exponent $\bar{p}$ versus aspect ratio for modes 3 and 4 (coupled 2B-1C branches); titanium blade, aspect ratio = 1; M = 1.8, 3 percent thickness

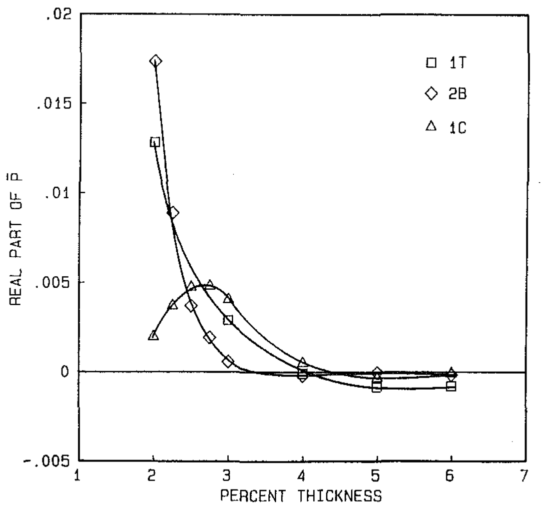


Fig. 9 Real part of the exponent $\bar{p}$ versus thickness for the first torsion, second bending, and chordwise bending branches; titanium blade, aspect ratio = 1; M = 1.8 (except for 1T, M = 1.5)

The effect of blade aspect ratio on the second bending and chordwise bending branches at high Mach number is shown in Fig. 6. The figure shows that the two branches significantly change mode shapes over a small range of aspect ratios near unity. The second bending and chordwise bending branches are obviously highly coupled in this low-aspect-ratio range. Figure 7 shows the frequency variations of these coupled modes with aspect ratio. The two frequencies approach each other but never coincide. This region of high coupling between modes of motion without a merging of frequencies is similar to a so-called "modal transfer" phenomenon reported by Weisshaar and Foist [20] in connection with their fiber sweep studies of composite wings.

Figure 8 gives a measure of the relative stability of the coupled second bending and chordwise bending branches for varying aspect ratio. Chordwise bending flexibility definitely has a destabilizing effect on the second bending branch for these low-aspect-ratio blades. The chordwise bending branch also seems to be quite unstable over a range of aspect ratio near unity.

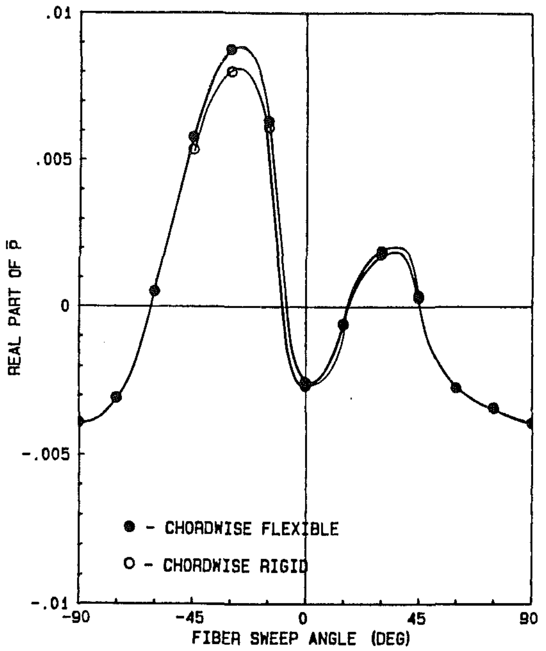


Fig. 10 Real part of the exponent $\bar{p}$ versus fiber sweep angle (positive aft) for the first bending branch; composite blade, M = 1.5, 3 percent thickness, aspect ratio = 1

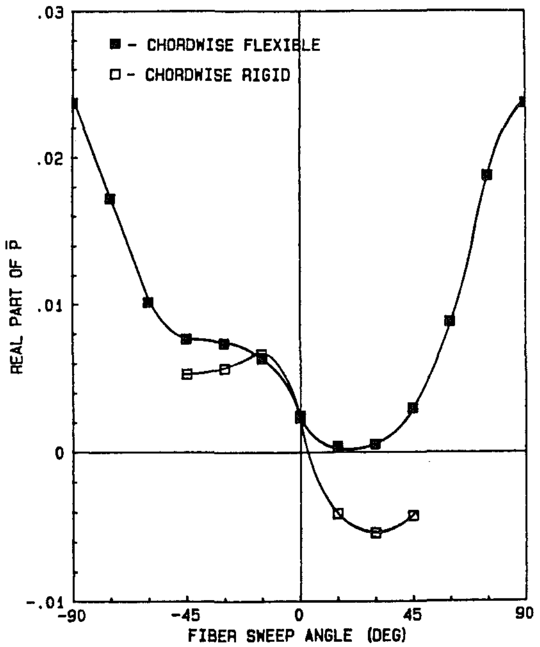


Fig. 11 Real part of the exponent $\bar{p}$ versus fiber sweep angle (positive aft) for the first torsion branch; composite blade, M = 1.5, 3 percent thickness, aspect ratio = 1

Blade thickness also affects the aeroelastic behavior of low-aspect-ratio blades, often profoundly. Figure 9 shows the relative stability of the coupled second bending and chordwise bending branches at high Mach number over a range of thickness. Both branches are unstable below 4 percent thickness. Also shown is a curve for the first torsion branch at a Mach number of 1.5. This Mach number was chosen since it corresponds to the Mach number of the maximum relative instability of the first torsion branch of the baseline case (see Fig. 2). This branch is also unstable below 4 percent thickness. It is curious that the three branches do not become increasingly stable above 4 percent thickness but remain almost neutrally

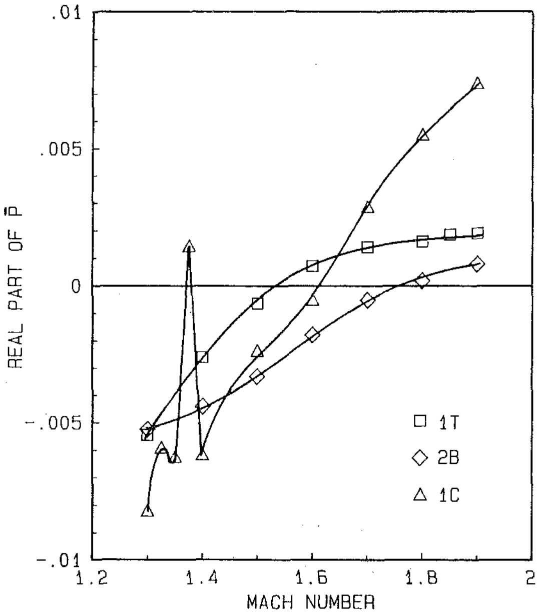


Fig. 12 Real part of the exponent $\bar{p}$ versus Mach number for the first torsion, second bending, and chordwise bending branches; composite blade, 3 percent thickness, aspect ratio = 1

stable. The reason for this behavior is unknown and is a subject of continuing study.

Laminated Composite Blade. It was decided to try to stabilize the baseline titanium configuration using the unique properties of fiber composite materials. The composite material chosen is AS/3501 graphite/epoxy and has the following properties [21]

$$E_x = 1.38\times10^{11}\ \mathrm{N/m^2}$$

$$E_y = 8.96\times10^{9}\ \mathrm{N/m^2}$$

$$\nu_x = 0.30$$

$$E_s = 7.1\times10^{9}\ \mathrm{N/m^2}$$

$$\text{specific gravity} = 1.6$$

Two different layup configurations are considered. The first configuration is one in which the plies are oriented such that the fibers are all swept in the same direction, at an angle to a spanwise reference. The second configuration is a crossed-ply configuration in which alternating plies have crossed fiber directions. For both configurations, the middle one third of the plies is always oriented such that the fibers are aligned with the spanwise direction to provide longitudinal stiffness. The outer two layers of plies are oriented in one of the two configurations.

Figures 10 and 11 show stability plots for the first bending and first torsion branches for varying fiber sweep angle. The blade geometry is the same as the baseline titanium blade. The figures show that fiber sweep produces bending-torsion coupling that can be either stabilizing or destabilizing at a moderate Mach number of 1.5. Forward sweep (negative) is generally destabilizing while moderate aft sweep (positive) can be stabilizing. Also note that chordwise flexibility has a pronounced effect on the stability of the first torsion branch. The frequency and structural coupling characteristics of thin low-aspect-ratio blades cannot be modeled correctly without including the effects of chordwise flexibility.

Numerical results obtained during this study indicate that crossed-ply configurations are somewhat more effective than swept fiber configurations in stabilizing the first bending and first torsion branches. Recall that the baseline titanium blade has an unstable first torsion branch over the entire supersonic

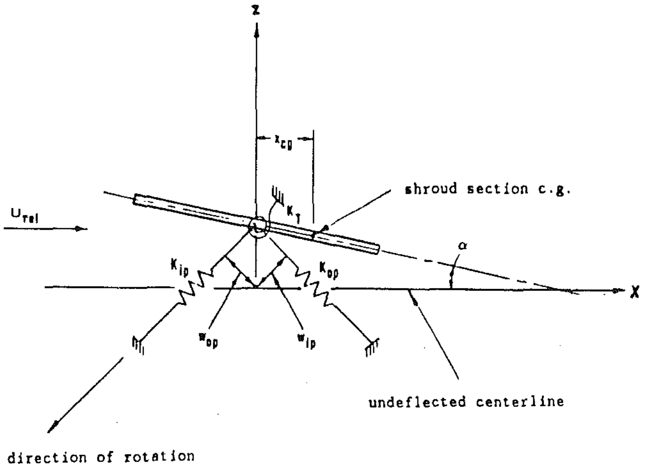


Fig. 13 Blade root section with elastic boundaries and shroud section mass

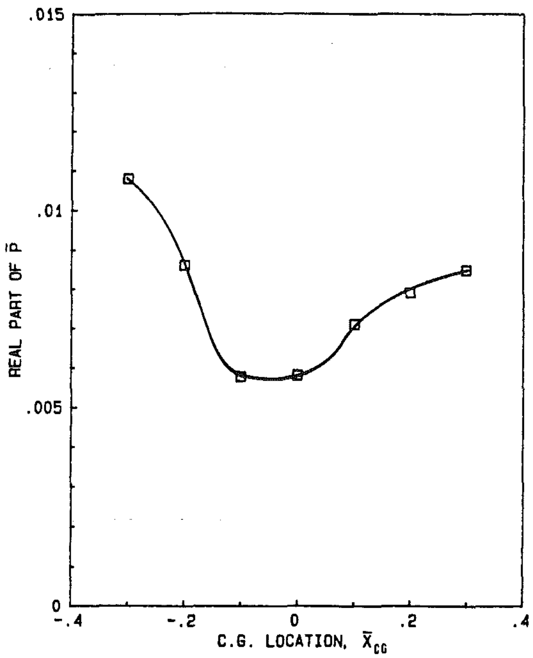


Fig. 14 Real part of the exponent $\bar{p}$ versus shroud section center of gravity location for mode 2; 8DOF model, M = 1.45, titanium blade

Mach number range (Fig. 2). Figure 12 shows that the same blade with a ±45 deg cross-ply composite section is stable up to a Mach number of about 1.55, except for a narrow instability below M = 1.4. Comparing this with Fig. 2 it is evident that the composite blade configuration is significantly more stable than its titanium counterpart. The unique characteristics of the fiber composite material can thus be taken advantage of by "tailoring" the blade to meet aeroelastic design constraints. Similar conclusions were reached in an earlier study of composite inlays in titanium blades [22].

Eight-Degree-of-Freedom Model. Figure 13 shows a blade that has three added degrees of freedom w_{ip}, w_{op}, and α at the blade root. The corresponding blade root stiffness in the plane of rotation, out of the plane of rotation, and in torsion are modeled by springs K_{ip}, K_{op}, and K_T, respectively. This model has been used to evaluate some of the effects of inboard blade motions on the outboard blade section of shrouded fans. In this case the springs model the stiffness of the shroud and inboard blade section, while the inertial properties are modeled

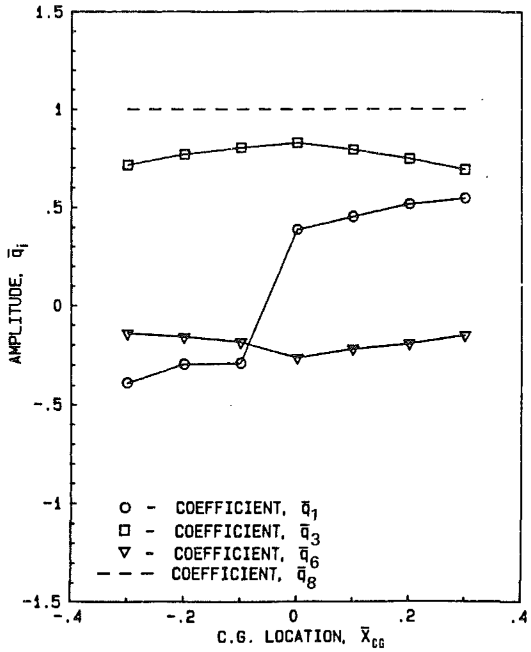


Fig. 15 Mode shape variation of mode 2 versus center of gravity location; 8DOF model, M = 1.45, titanium blade

by a lumped mass with a given center of gravity and radius of gyration about the center of gravity. Details of the necessary modifications to the equations of motion can be found in [15].

Illustrative results for a shrouded fan are shown in Figs. 14 and 15 for the baseline titanium blade tip section studied earlier. The base springs and inertial properties are adjusted to give the following blade frequency ratios for the base modes

$$\omega_{\text{in-plane}}/\omega_{\text{out-of-plane}} \simeq 4.5$$

$$\omega_{\text{out-of-plane}}/\omega_{\text{torsion}} \simeq 0.5$$

The in-board section mass ratio is assumed equal to the tip section mass ratio. Figure 14 shows the dependence of the center of gravity of the shroud and inboard blade section on the stability of the second mode. This mode is primarily first torsion, q_8 $(=\alpha)$, with a large contribution from plate bending q_1 and plate torsion q_3; see Fig. 15. There is also a significant contribution from the out-of-plane bending mode q_6. It is evident that the aeroelastic stability and mode shape are quite sensitive to the c.g. location along the blade chord.

Conclusions

The main conclusions that can be drawn from this study are:

1 Chordwise flexibility has significant effects on the aeroelastic behavior of low-aspect-ratio fan and compressor blades. A strong coupling between chordwise bending and second bending is observed, which is destabilizing on the second bending branch.

2 Fan blade aeroelastic stability is very sensitive to blade aspect ratio and thickness. For blades of aspect ratio near unity and thickness below 4 percent of chord, the first torsion, second bending, and first chordwise branches are all unstable at moderately high supersonic Mach numbers.

3 Fiber sweep in a laminated composite blade results in bending-torsion coupling that can be used as a design tool to produce stable blade configurations. Forward fiber sweep is generally destabilizing while moderate values of aft sweep can be stabilizing.

4 Crossed-ply configurations in a composite blade can be used to prevent aeroelastic instabilities without producing destabilizing bending-torsion coupling. This type of configuration is also effective in increasing the torsional and chordwise stiffness.

Acknowledgments

A part of this research was supported by NASA Lewis Research Center under Grant NAG 3-308, with Drs. G. V. Brown and R. E. Kielb as technical monitors. The authors gratefully acknowledge this support.

References

1 Barton, M. V., "Vibration of Rectangular and Skew Cantilever Plates," ASME *Journal of Applied Mechanics*, Vol. 18, 1951, pp. 129-134.

2 Dokainish, M. A., and Rawtani, S., "Vibration Analysis of Rotating Cantilever Plates," *International Journal of Numerical Methods in Engineering*, Vol. 3, 1971, pp. 233-248.

3 Jensen, D. W., and Crawley, E. F., "Frequency Determination Techniques for Cantilevered Plates With Bending-Torsion Coupling," *AIAA Journal*, Vol. 22, No. 3, 1984, pp. 415-420.

4 Hollowell, S. J., and Dugundji, J., "Aeroelastic Flutter and Divergence of Stiffness Coupled, Graphite/Epoxy, Cantilevered Plates," *AIAA Journal*, Vol. 21, No. 1, 1984, pp. 69-76.

5 Petricone, R., and Sisto, F., "Vibration Characteristics of Low Aspect Ratio Compressor Blades," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 93, No. 1, 1971, pp. 103-112.

6 Leissa, A. W., Lee, J. K., and Wang, A. J., "Vibrations of Twisted Rotating Blades," ASME *Journal of Vibration, Acoustics, Stress, and Reliability in Design*, Vol. 106, No. 2, 1984, pp. 251-257.

7 Leissa, A. W., Lee, J. K., and Wang, A. J., "Rotating Blade Vibration Analysis Using Shells," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 104, No. 2, 1982, pp. 296-302.

8 Whitehead, D. S., "Torsional Flutter of Unstalled Cascade Blades at Zero Deflection," Great Britain A.R.C. R&M 3420, 1964.

9 Snyder, L. E., and Commerford, G. L., "Supersonic Unstalled Flutter in Fan Rotors; Analytical and Experimental Results," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 96, No. 4, 1974, pp. 379-386.

10 Verdon, J. M., and McCune, J. E., "Unsteady Supersonic Cascade in Subsonic Axial Flow," *AIAA Journal*, Vol. 13, No. 2, 1975, pp. 193-201.

11 Carta, F. O., "Coupled Blade-Disk-Shroud Flutter Instabilities in Turbojet Rotors," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 89, 1967, pp. 419-426.

12 Mikolajczak, A. A., Arnoldi, R. A., Snyder, L. E., and Stargardter, H., "Advances in Fan and Compressor Blade Flutter Analysis and Predictions," *Journal of Aircraft*, Vol. 12, No. 4, 1975, pp. 325-332.

13 Bendiksen, O. O., and Friedmann, P., "Coupled Bending-Torsion Flutter in Cascades," *AIAA Journal*, Vol. 18, No. 2, 1980, pp. 194-201.

14 Bendiksen, O. O., and Friedmann, P., "Coupled Bending-Torsion Flutter in a Supersonic Cascade," *AIAA Journal*, Vol. 19, No. 6, 1981, pp. 774-781.

15 White, J. F., "Aeroelastic Behavior of Low Aspect Ratio Metal and Composite Turbomachinery Blades," Princeton University, Department of Mechanical and Aerospace Engineering, MSE Thesis, July 1985.

16 Calcote, L. R., *The Analysis of Laminated Composite Structures*, Van Nostrand Reinhold Company, New York, 1969.

17 Bendiksen, O. O., and Friedmann, P. P., "Coupled Bending-Torsion Flutter in Cascades With Applications to Fan and Compressor Blades," UCLA Report ENG-8072, Mechanics and Structures Department, Aug. 1980.

18 Timoshenko, S. P., and Woinowsky-Krieger, S., *Theory of Plates and Shells*, 2nd ed., McGraw-Hill, New York, 1959.

19 *Military Standardization Handbook—Metallic Materials and Elements for Aerospace Vehicle Structures*, MIL-HDBK-5D, Vol. 2 of 2, June 1983, pp. 5-59.

20 Weisshaar, T. A., and Foist, B. L., "Vibration Tailoring of Advanced Composite Lifting Surfaces," *AIAA Journal*, Vol. 22, No. 2, 1985, pp. 141-147.

21 Tsai, S. W., and Hahn, H. T., *Introduction to Composite Materials*, Technomic Publishing Company, Westport, CT, 1980.

22 Troha, W., and Swain, K., "Composite Inlays Increase Flutter Resistance of Turbine Engine Fan Blades," ASME Paper No. 76-GT-29.

Dr. D. A. Wilson
Tennessee Technological University

D. P. Deluca

B. A. Cowles
Pratt & Whitney Aircraft

M. A. Stucke
Air Force Wright Aeronautical Laboratories

Fatigue Crack Growth Resistance of Advanced Blade Materials

The best measure of performance in a jet engine is the thrust-to-weight ratio. Cast single-crystal superalloys provide higher temperature capability and offer opportunities for significant improvements in future jet engine performance and durability. The highly anisotropic behavior of these advanced materials, which includes modulus, strength, and ductility variation with crystallographic orientation, pose potentially significant problems for design and life prediction of advanced turbine blades. The objective of this paper is to provide a better understanding of the crack growth behavior in this advanced turbine blade material by evaluating the effect of temperature, frequency, orientation, and thickness.

Introduction and Background

Advances in turbine engine thrust, efficiency, and durability have historically been directly related to improvements in turbine airfoil materials and cooling schemes. Prior to the 1970s turbine blade and vane materials were conventionally cast, equiaxed structure materials, such as IN100 and Rene '80. A major advancement occurred in the late 1960s with the introduction of directionally solidified (DS) MAR-M200 + Hf airfoils in the F100 engine.

Directional solidification is a process which generally produces a [001] crystallographic direction, which will be illustrated later. This direction in conventional superalloys exhibits excellent strain-controlled (but relatively poor load-controlled) fatigue resistance because of its low elastic modulus (≃40 percent below that of equiaxed MAR-M200). The DS alloy also exhibits good creep and rupture strength.

The improved creep resistance, absence of transverse grain boundaries, and the low longitudinal modulus and corresponding improved thermal fatigue response of the alloy resulted in significant increases in allowable metal temperatures and blade stresses. In the mid-1970s further developmental work showed that substantial benefits could be attained over the DS materials by using the single-crystal casting approach. Objectives and guidelines for the single-crystal alloy development program then undertaken are described by Gell et al. [1]. The resulting material, Alloy 454, has a composition considerably different from and simpler than preceding cast nickel blade alloys because of the absence of grain boundary strengthening elements.

Deleting the grain boundary strengthening elements, which are also melting point depressants, increased the incipient melt temperature beyond 1288°C (2370°F) and allowed nearly complete γ' solutioning during heat treatment and a reduction in dendritic segregation. This resulted in significantly improved properties over conventionally cast and directionally solidified blade materials.

As with the DS processing, single-crystal castings usually have a [001] orientation with corresponding low elastic modulus in the longitudinal (solidification) direction. A photograph of turbine blade castings with conventionally cast, directionally solidified, and single-crystal structures is presented in Fig. 1.

The face-centered cubic cast single-crystal structure which characterizes Alloy 454 exhibits pronounced directionality in its mechanical behavior. As an example, consider the variation in elastic modulus encountered at 427°C (800°F) with a [001] crystallographic direction, 115 GPa (16.7×10^6 psi), compared to the [111] crystallographic direction elastic modulus of 274 GPa (39.7×10^6 psi).

Similar variations exist for the tensile, creep, fatigue, and fracture mechanics behavior of the alloy. Because the potential is great for improved durability and higher operating temperatures through the use of single-crystal components, a

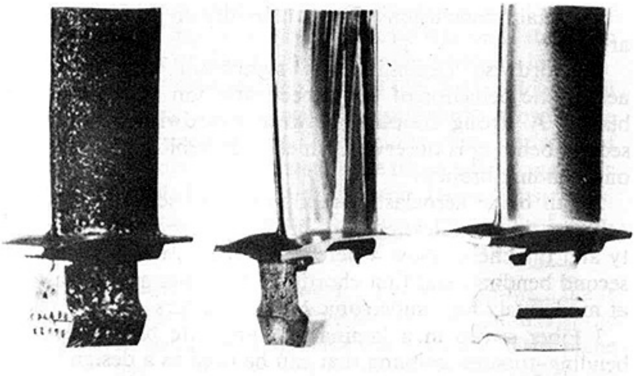

Conventional Casting
Directional Solidification
Equiaxed Grain
Columnar Grain
Single Crystal

Fig. 1 Advances in cast turbine airfoil materials

Contributed by the Gas Turbine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the 31st International Gas Turbine Conference and Exhibit, Düsseldorf, Federal Republic of Germany, June 8-12, 1986. Manuscript received at ASME Headquarters March 7, 1986. Paper No. 86-GT-253.

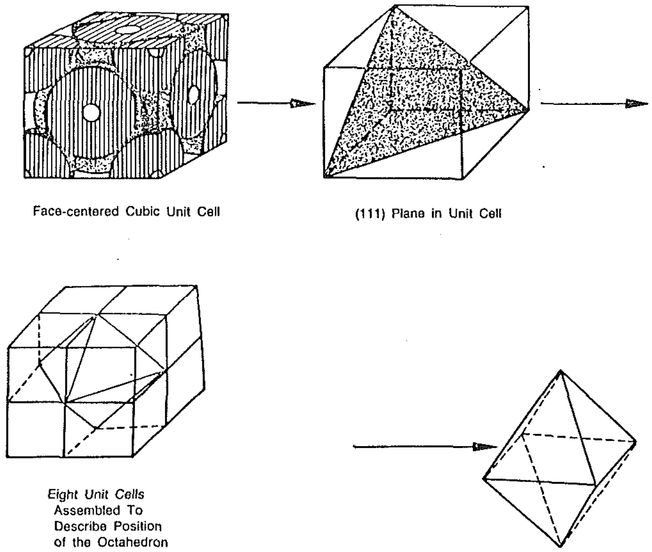


Fig. 2 Situation of the (111) octahedron in the FCC cubic structure

program sponsored by the Air Force Wright Aeronautical Laboratories (AFWAL) entitled "Fatigue and Fracture of Advanced Blade Materials" was established to consider the impact of single crystals to the already difficult problem of predicting service lives of cooled turbine airfoils.

Fatigue Crack Growth Mechanisms

The most severe cyclic stresses that turbine blades encounter are thermally induced by extreme temperature gradients and transients. These thermally induced stresses, which are sometimes combined with the mechanically induced centrifugal and gas bending loads, produce thermal mechanical fatigue (TMF) of the blade.

The mode of fatigue crack initiation and subsequent propagation in cast single-crystal alloys differs from conventional nickel superalloys, and is related to the effects of anisotropy in the directional microstructures of the face-centered cubic (FCC) cast single crystal.

In general, crack initiation and propagation for polycrystalline nickel-based superalloys at elevated temperatures can be described by three modes: Stage I, crystallographic cracking along preferred slip planes; Stage II, noncrystallographic, but generally transgranular cracking normal to the applied stress direction; and, finally, intergranular cracking. Variables which control the cracking mode for both initiation and crack growth in any equiaxed nickel-based alloy include the temperature, cyclic frequency or hold time, strain range and rate, and mean stress.

In the cast single-crystal nickel alloy, two basic types of cyclic fracture modes, crystallographic and noncrystallographic, are seen in combination in varying proportions upon temperature. In the 427°C (800°F) area small amounts of noncrystallographic propagation are seen immediately surrounding initiation sites. This propagation has the appearance of Stage II transgranular fatigue in a fine grain equiaxed alloy. This propagation changes almost immediately to a highly crystallographic mode along (111) shear planes (Fig. 2) frequently exhibiting prominent striations. This mode continues to an overstress condition. At intermediate temperatures, 760°C (1400°F), this first mode of propagation is more pronounced and may continue until tensile overstress along (111) planes, or it may change to cyclic propagation along these planes. At 982°C (1800°F) propagation is almost all noncrystallographic, similar to a Stage II condition.

In studying the alloy's cyclic behavior, tests addressing the effects of temperature and anisotropy on propagation were conducted. Crack propagation behavior in response to varying frequency, including hold times, stress ratios, and thickness effects, was also examined.

Temperature. Actual service temperatures which may result in fatigue crack initiation in turbine blades range from the temperature of high-pressure turbine cooling air on the low end – 427°C (800°F) to 540°C (1000°F) – to the maximum gas path metal temperatures on the high end – 815°C (1500°F) to more than 1093°C (2000°F).

In addition, fatigue crack initiation mechanisms of advanced nickel superalloys were expected to alter from planar slip type (Stage I) at low temperatures to a more homogeneous Stage II type deformation and, probably, a creep failure mode at the extreme upper temperatures. Test temperatures of 427°C (800°F), 760°C (1400°F), and 982°C (1800°F) span both the majority of the engine operating envelope and expected changes in fatigue behavior of the material.

Frequency. Cyclic frequency and hold time effects combine with temperature to determine the fracture mechanism which will predominate during fatigue crack growth. The baseline tests for reference data were conducted at 0.17 Hz (10 cpm) to simulate a simple low-frequency fatigue cycle, with two tests each at 427°C (800°F), 760°C (1400°F), and 982°C (1800°F). Higher frequency cyclic tests at 20 Hz were conducted at each temperature to determine the effect of a much higher frequency on crack growth rate and failure mode. Turbine blade resonant frequencies are generally somewhat higher (greater than 500 Hz); however, the test frequency was adequate for assessing frequency effects on the material behavior. Effects of crack tip oxidation on growth rate and changes in failure modes were evaluated. Since the turbine blade airfoils are exposed to elevated temperatures for extended periods, a dwell time cycle was employed, with a 120-s dwell at maximum load. Four dwell tests were conducted, two at 760°C (1400°F), and two at 982°C (1800°F). Hold time tests were not planned at 427°C (800°F) since hold time effects for the selected alloy should be considerably less at 427°C (800°F) than at the higher temperatures [2, 3].

Orientation. To maximize the effects of anisotropy, crystallographic orientations representing extreme variations in elastic modulus were chosen. These modulus extremes are also accompanied by strength and ductility variations and were expected to result in fracture mechanics behavior differences. Schematic representations of the orientations chosen are shown in Figs. 3 and 4. For baseline crack growth the [001] crystallographic direction was chosen as the loading axis. This direction is, by convention, parallel to the solidification or primary dendrite growth direction and exhibits the lowest elastic modulus and ductility, and the highest yield strength. In the case of crack growth testing, the [010] direction was used as the baseline crack path direction. This direction is perpendicular to the alloy's solidification direction and allows a crack to propagate transdendritically. It also provides the same mechanical properties as the [001] direction. Crack growth tests of this orientation are denoted [001]/[010], indicating a [001] loading direction and a [010] crack path. For high modulus isothermal crack growth testing, the [110] family of directions was chosen as the loading axis and crack path, ($[011]/[0\bar{1}1]$), representing the highest modulus in the blade's transverse plane.

Thickness. The thickness of the material influences crack growth by changing the stress state at the tip of the crack [4]. For thick specimens, a condition of plane strain exists at the crack tip, and the crack will grow normal to the applied stress. This is true in equiaxed materials, except in a region near the surface where a condition of plane stress exists and shear lips are formed by the crack propagating in shear at a 45-deg angle

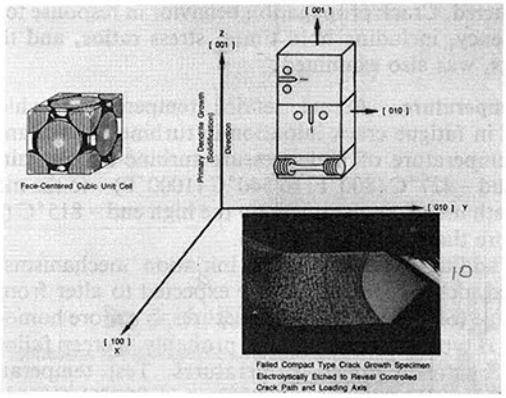


Fig. 3 Schematic of seeded [001] slabs used in isothermal crack growth and LCF testing

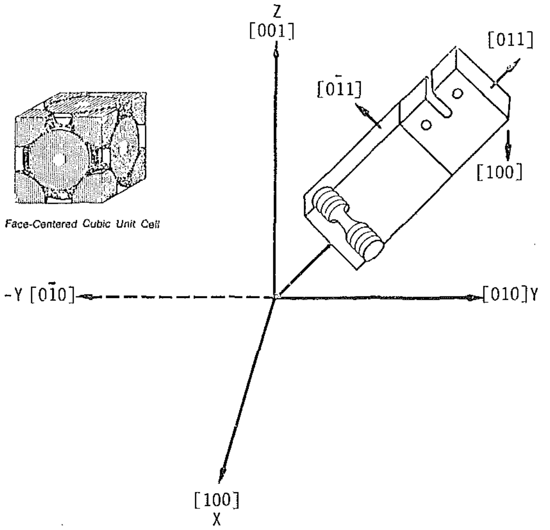


Fig. 4 Schematic of seeded [011] slabs used in isothermal crack growth

with the loading axis. In single crystals, however, the preferred slip plane (111) is the dominant shear plane.

For thin sections, as in blades, a condition of plane stress exists through the thickness encouraging the crack to grow out of plane along the plane of maximum shear mobility. This corresponds to a Mode II type of growth controlled by the shear stress. Limited testing was conducted to investigate thickness effects at 760°C (1400°F) with specimens oriented in the baseline configuration [001]/[010].

Material

Castings in the macrograin etched condition were visually inspected for subgrains or defects. Crystal orientation deviations were held at ±8 deg. Schematics showing the crystallographic orientations produced are shown in Figs. 3 and 4.

Orientations of [001] and [011] were chosen. Further versatility in orientation options was provided by controlling specimen alignment with respect to casting geometry to obtain crack paths between [010] and [110] in center notch crack growth specimens and [010] or [001] in compact tension crack growth specimens.

The test material was supplied in solution. The melt consisted of 50 percent revert and 50 percent virgin material. Chemical composition was verified via mass spectrometer analysis, and creep capability at 982°C (1800°F) for the solution was checked and found to be greater than 400 hr at 172.4 MPa (25 ksi). Microstructures were examined and optical and transmission electron microscopy was conducted on the solution material. Structures were considered typical of Alloy 454 given a step solution heat treatment. Typical cuboidal gamma prime size was found to be approximately 0.5 μ, close to the expected size for material in the solutioned condition. The castings subsequently received the coating diffusion and precipitation thermal cycles. Microstructures were again examined after application of the full heat treatment and were found to be typical for fully solutioned and aged Alloy 454.

Table 1 Basic crack growth evaluation test matrix – temperature and frequency

	Temperature (°F)			
Frequency/hold time	800	1400	1800	Total
20 Hz	1	2	1	4
0.17 Hz	2	2	2	6
120-s hold		2	2	4
			Total	14

Notes: Tests were isothermal, load controlled, fixed thickness, $R=01$. Specimen orientation was along the primary crystal direction [001] for cast single crystal.

Table 2 Supplementary crack growth test matrix – effects of orientation, thickness, and stress ratio

	Temperature (°F)			
Variable	800	1400	1800	Total
Alternate orientation [110]	2	2	2	6
Alternate thickness				
T2		1		2
T3		1		
Alternate stress ratio −1.0		2		2
			Total =	10

Notes: Tests were conducted for comparison with 0.17 Hz tests in basic crack growth test matrix.

Experimental Program

The objective of this paper is to provide a better understanding of the crack growth behavior in this advanced turbine blade material by evaluating the effects of temperature, frequency, orientation, and thickness. Therefore, the crack growth tests summarized in Tables 1 and 2 were conducted.

Test Procedures

Isothermal crack growth rate testing was conducted using ASTM E647 procedures [5]. The specimens used were the standard compact type (CT) shown in Fig. 5. They were chosen for the large amount of crack propagation data yielded. The center-cracked-tension (CCT) specimen (Fig. 6) was employed for negative stress ratio crack growth rate testing. Test specimens were precracked using procedures outlined in ASTM E647. Precracking was performed at an elevated temperature at a cyclic frequency of 20 Hz. This helped to maintain the crack propagation perpendicular to the loading direction. Cyclic tests were performed using isosceles triangular load waveforms. Specimen heating was provided by resistance, clamshell furnaces having windows which allowed observation of crack growth at the test temperature.

Using a traveling microscope, crack lengths were measured on both surfaces of the specimen at the mean test load. This procedure held the specimen rigid while increasing crack tip visibility. A high intensity light was used to provide oblique illumination to the crack and further increase crack visibility. In general, crack length measurements were taken at increments no larger than 0.25 mm (0.010 in.).

Several specimens exhibited out-of-plane growth along

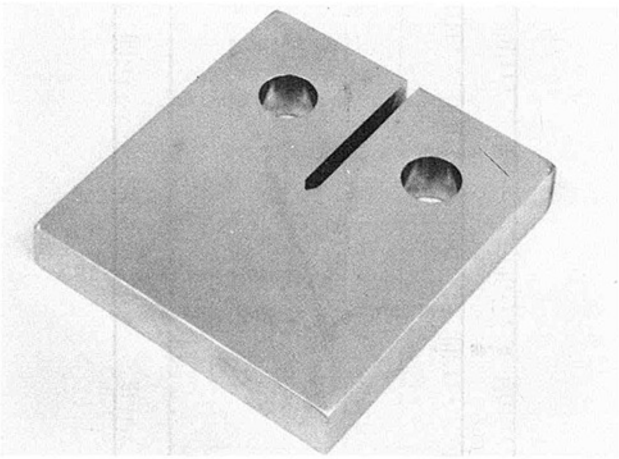
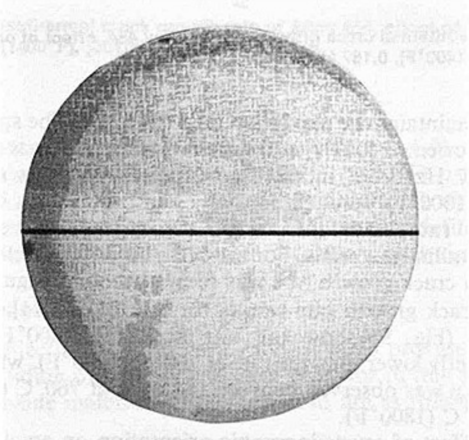
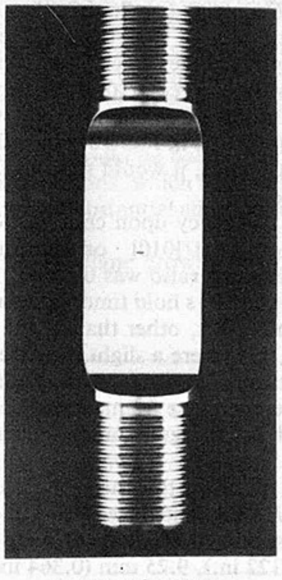


Fig. 5 Compact type (CT) specimen



A. Dendrite Orientation



B. Finished Specimen

Fig. 6 Center-cracked-tension (CCT) specimen

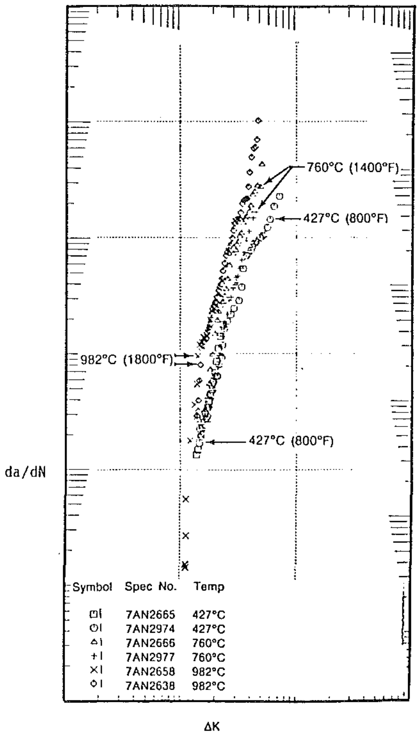


Fig. 7 Isothermal crack growth rate of PWA 1480, effect of temperature at 0.167 Hz (10 cpm), [001]/[010], $R = 0.1$

primary slip planes at an angle greater than 5 percent. The projected crack length was then used for data analysis. Although this procedure does not meet ASTM standard E647, the test specimen fracture surfaces show the crack in the bulk of the material to be growing in plane and the projected crack length accurately described this growth. This was the situation in most, if not all, of the out-of-plane specimens.

Regression of the crack length versus cycle (a, N) data to produce crack growth rate versus applied stress intensity range (da/dN, ΔK) data was accomplished with ASTM recommended data reduction techniques. In general, the seven-point incremental polynomial technique was used. When only a small number of data points were available, the direct secant method was used.

The longitudinal axis and the casting solidification direction of the specimen was [001] in all cases. Because cracking in the center notch crack growth specimen is normal to the longitudinal axis, the crack path was perpendicular to the primary dendrite growth direction. The crack path (secondary orientation of the casting) was controlled by specifying rotation of the specimen with respect to the cast bars at the time of machining.

Results

Information from this project is subject to export control laws. Consequently, absolute values of the data cannot be shown. Only relative comparisons are presented in this paper.

The convention used for describing the orientations of crack growth specimens states the loading axis crystallographic direction and the crack propagation crystallographic direction. As an example, [001]/[010] describes an [001] load axis and [010] crack growth direction. Further, turbine airfoils are solidified in the Z direction which is also the vertical axis of the Miller indices system. Consequently, [001] is considered the Z or growth direction of the castings employed in this work.

The dependence of isothermal crack growth rate on temperature was investigated over the 427°C (800°F), 760°C

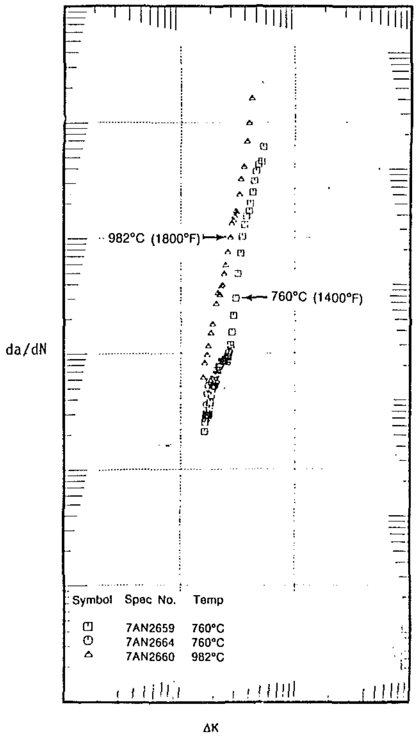


Fig. 8 Isothermal crack growth rate of Alloy 454, effect of temperature [001]/[010], 2 min dwell, $R = 0.1$

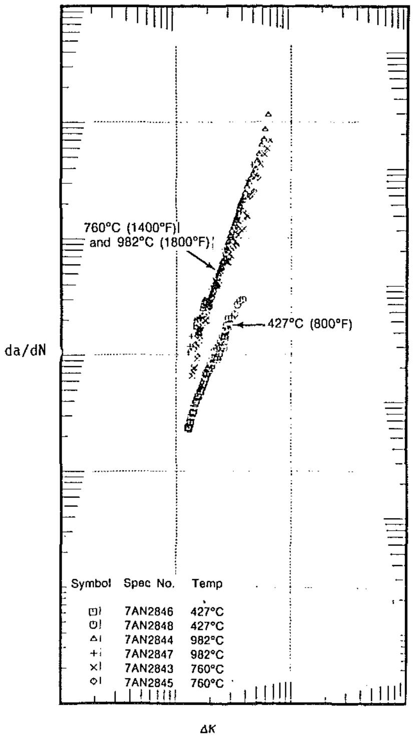


Fig. 9 Isothermal crack growth rate of Alloy 454—effect of temperature at 0.167 Hz (10 cpm), [011]/[011], $R = 0.1$

(1400°F), and 982°C (1800°F) range. The alloy responses to this increasing temperature at frequencies of 0.167 Hz (10 cpm) and 120-s peak load hold time are illustrated in Figs. 7 and 8. The hold time cycle is the same ramp time as the 0.167

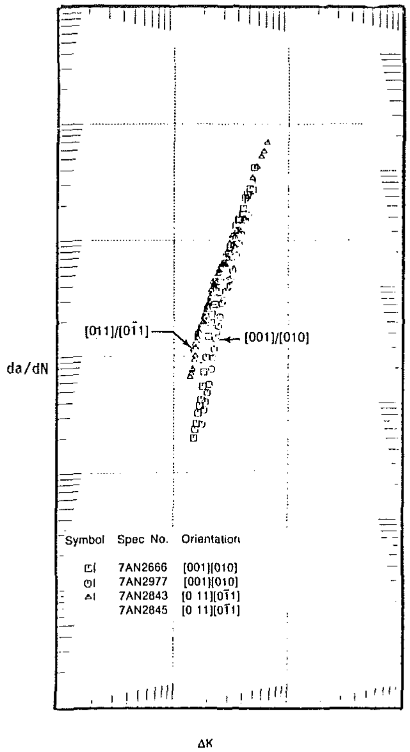


Fig. 10 Isothermal crack growth rate of Alloy 454, effect of orientation at 760°C (1400°F), 0.167 Hz (10 cpm), $R = 0.1$

Hz but maintains the maximum load for 120 s. The specimens were all oriented [001]/[010] and the stress ratio was 0.1. For the 0.167 Hz cycle, increasing temperatures by as much as 555°C (1000°F) resulted in only slight increases in crack growth rate, as compared to many conventional polycrystalline materials. For the 120-s hold time cycle, the increase in crack growth rate was somewhat more significant.

The crack growth rate results for the [011]/[01̄1] oriented material (Fig. 9) show the rate at 427°C (800°F) to be significantly lower than the rate at 760°C (1400°F), while little difference was observed between the rates at 760°C (1400°F) and 982°C (1800°F).

The effect of crystallographic orientation on crack growth rate at 0.167 Hz, $R = 0.1$ was investigated at the three temperatures 427°C (800°F), 760°C (1400°F), and 982°C (1800°F). Orientations of [001]/[010] were compared to [011]/[01̄1]. A slightly higher crack growth rate at lower stress intensity ranges was observed at all three temperatures for the [011]/[01̄1] orientation. This trend is best illustrated by the results at 760°C (1400°F) in Fig. 10. Since this trend is at the low stress intensity range, it would significantly affect the life of a turbine blade.

The effects of frequency upon crack growth rate were examined for the [001]/[010] orientation at all three temperatures. The stress ratio was 0.1 and the frequencies of 20 Hz, 0.167 Hz, and 120 s hold time were applied. The results showed little to no effect, other than at the intermediate test temperature (Fig. 11) where a slight increase in growth rate is evident with increasing frequency. This result is quite different from most high-temperature frequency effects, where low frequencies or hold times significantly increase crack growth rates [2, 3].

The effects of specimen thickness on crack growth rate at 760°C (1400°F), $R = 0.1$, and a frequency of 0.167 Hz were investigated. The specimen thicknesses were 1.88 mm (0.074 in.), 3.10 mm (0.122 in.), 9.25 mm (0.364 in.), and 11.70 mm (0.460 in.). No significant trend was observed from these results.

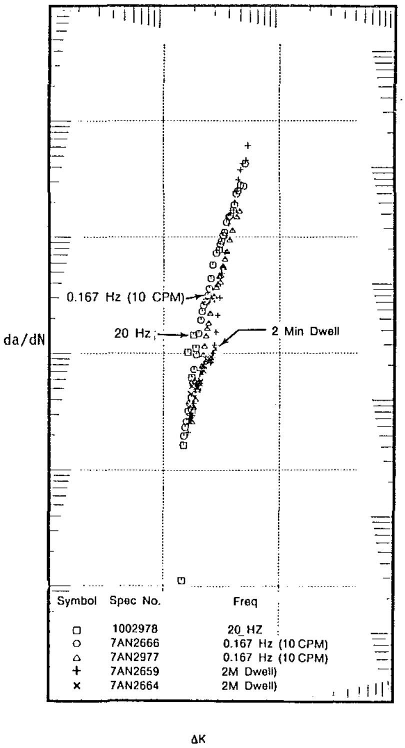


Fig. 11 Isothermal crack growth rate of Alloy 454, effect of frequency at 760°C (1400°F), [001]/[010], $R = 0.1$

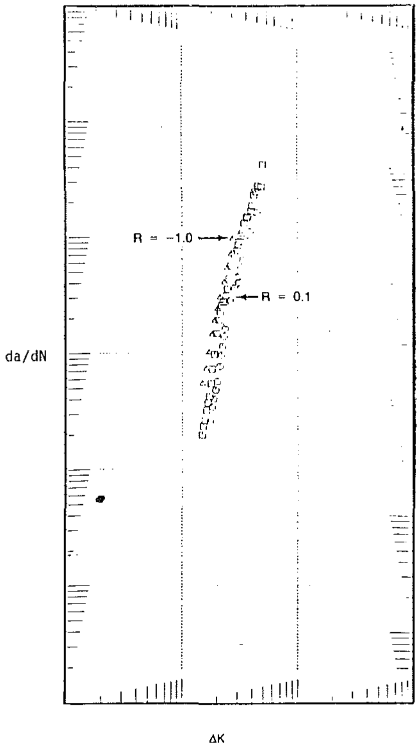


Fig. 12 Isothermal crack growth rate of Alloy 454, effect of negative stress ratio at 760°C (1400°F), [001]/[010], 0.167 Hz (10 cpm), $R = 0.1$

The effects of negative stress ratio ($R = -1.0$) on crack growth rates were evaluated at 760°C (1400°F) using CCT specimens. A comparison of these results to those obtained at $R = 0.1$ is shown in Fig. 12. The ASTM standard definition for ΔK ($\Delta K = K_{\max}$ for $R > 0$) was used for data analysis.

Conclusion

The objective of this investigation was to provide a better understanding of the crack growth behavior in this advanced turbine blade material. The conclusions are as follows:

1 The applicability of linear elastic fracture mechanics assumptions to single-crystal crack growth rate evaluation was found to be acceptable. This was demonstrated by excellent agreement between duplicate tests conducted with specimens with significant variations in dimensions, load, and crack length and was observed over a variety of test frequencies and temperatures.

2 For the future it should be noted that both test stresses and stress intensity solutions, which were used, were based on isotropic assumptions. Substantial additional work in this area is recommended.

3 Significant out-of-plane growth was observed at some conditions, but deviations from Mode I growth rates were very slight. Additional work is recommended in this area, particularly for thin-wall sections representative of turbine airfoils.

4 Single-crystal behavior in propagation was found to be predictable and repeatable, although strong anisotropic and crystallographic effects were evident.

References

1 Gell, M., Duhl, D. N., and Giamei, A. F., "The Development of Single Crystal Superalloy Turbine Blades," *Superalloys 1980, Proceedings of the Fourth International Symposium on Superalloys*, American Society for Metals, Metals Park, OH, 1980, pp. 205–214.

2 "Application of Fracture Mechanics at Elevated Temperatures," Air Force Contract No. F3361575-C-5097, AFWAL-ML, WPAFB, OH.

3 "Cumulative Damage Fracture Mechanics Under Engine Spectra," Air Force Contract No. F33615-77-C-5153, AFWAL-ML, WPAFB, OH.

4 Harris, J. A., and White, R., "Engine Component Retirement for Cause," *Interim Technical Report, FR-16511 Contract No. F33615-80-C-5160*, Aug. 1982.

5 American Society for Testing and Materials, Section 3, Volume 03.01, E647.

6 Wilson, D. A., and Warren, J. R., "Thermal Mechanical Crack Growth Rate of a High Strength Nickel Base Alloy," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 108, 1986, pp. 396–402.

Quality Assurance in Manufacturing High-Performance Ceramics

M. Srinivasan
Niagara Falls R & D Center.

M. Earl
Structural Ceramics Division.

Standard Oil Engineered
Materials Company,
Niagara Falls, NY 14302

Quality assurance of ceramic components during various processing stages is an important ingredient in assuring the reliability in engineering applications. The current advances which are made in advanced nondestructive evaluation of finished components have been very useful. However, the importance of in-process quality control in closed-loop iterative process and product development has to be understood in terms of materials science aspects. This paper addresses some of the tools in nondestructive evaluation which are available in ceramic manufacture and gives illustrations of successful implementation of those techniques in two categories of ceramic applications, current and advanced. Further research and development needs in the NDE are also briefly addressed.

Introduction

Reliability and reproducibility are the key to the successful application of sintered alpha silicon carbide as engineered material. The reliability of the finished product is a strong function of the quality of the starting powders and additives as well as the control used in processing parameters throughout the manufacturing operation. An understanding of the chemistry of the raw materials, various ceramic forming steps, the sintering phenomenon, product strength and the nature of its variability, and the ubiquitous nonhomogeneity is very important in instituting appropriate quality checks at various stages.

Generalization with respect to the degree of control required is not possible because it is very much dependent on the product and its use. For example, quality requirements for wear components are very different from those for components slated for engine applications such as turbocharger or gas turbine rotors. For several heat exchanger applications the requirements may be "essentially round and essentially straight" tubes. Therefore, inspection procedures can be expected to be different for different components.

At present satisfactory accept/reject criteria do not exist for high-performance ceramics. These are established through mutual negotiations between the user and the supplier dependent upon cost and *expected* performance. In fact, in a majority of circumstances there is insufficient product performance history related to product quality to justify rigorous specification writing. This applies to both wear components and engine components.

On the basis that assurance of raw material quality and process control within specified limits is expected to yield good-quality product, it is wiser to institute adequate in-process quality control in addition to final inspection of the finished component. This practice is being followed at the Structural Ceramics Division. The in-process quality checks during all manufacturing steps are very critical in ensuring the quality of the final product as well as in reducing the cost by eliminating rejects at the earliest manufacturing step.

In this paper, we will address these concerns in the commercial manufacture of ceramic components for current and advanced applications. We will consider steps involved in the in-process quality checks of sintered alpha silicon carbide (Hexoloy SA-80) components. The importance and the requirements of the use of quality control in closed-loop iterative product manufacture with high quality, low cost as the objective will be emphasized. The synergism between the use of both commercially available nondestructive evaluation techniques and emerging techniques in manufacture will be addressed. The limitations in some of the current practices which need further research and development will be pointed out for future work.

Type of Inspection

The inspection process can be broadly divided into two segments, as shown in Fig. 1. The in-process inspection includes both raw materials and processing parameters in all aspects of the processing steps. Postfabrication inspection is really quality assurance of the final product.

Further classification of the inspection of fabricated ceramic articles generally falls into two categories: destructive and nondestructive. Some of the ways of inspecting silicon carbide ceramics are indicated in Fig 2.

It should be noted that not all of the inspection techniques are applicable to all product lines. Depending upon the requirements of product performance and history, a few of the above methods are selected and agreed upon mutually by the supplier and the user.

Contributed by the Gas Turbine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS and presented at the 31st International Gas Turbine Conference and Exhibit, Düsseldorf, Federal Republic of Germany, June 8–12, 1986. Manuscript received at ASME Headquarters February 21, 1986. Paper No. 86-GT-248.

INSPECTION

IN-PROCESS POST-FABRICATION

Fig. 1 Division of inspection process

INSPECTION

DESTRUCTIVE	NONDESTRUCTIVE
Microstructure	Density
Hardness	Ultrasonic velocity
Fracture Toughness	Young's modulus
Strength	Dimensional Examination
Flaw Fractography	Contour analysis (Warpage)
XRD Phase Analysis	Penetrant Inspection
Elemental Analysis (via mass spectroscopy)	X-ray radiography
	Surface wave ultrasonics

Fig. 2 Some inspection techniques to establish reliability of silicon carbide ceramic

Rigorous use of statistical quality control is lacking at present for high-performance ceramics (except ceramic chips) because of inadequate product volume and field experience. However, this is changing as the structural ceramics business matures, especially in wear-resistant applications.

In-Process Quality Control

The in-process quality control (Q.C.) begins with the characterization of the incoming powder. The properties which are examined are (Appendix I):

- particle size
- particle size distribution
- surface area
- bulk composition
- surface chemistry
- crystalline form
- degree of agglomeration
- tap density

In production, all of these are *not* usually examined. However, a sintering qualification run is made to certify all incoming batches of powders. Here density is used as a measure of sinterability. Samples of the powder are mixed with suitable binders and pressed into disks 1.25 in. diameter × 0.25 in. thick. After firing, the density and open porosity are measured by using ASTM Procedure C-20. Selected pieces are also polished and etched for microscopic determination of grain size, grain shape, and porosity.

Other in-process quality checks are usually associated with processing parameters. In the case of injection molding, the shear properties of the compound are tested to insure that these are within specific range. Likewise in slip casting, viscosity, solids content, and pH are measured for the slurry to insure that the values are within specified range. *In all processing, visual inspection of the finished component plays a predominant role.* For example, in slip casting, the surface finish of the drain cast ceramic is visually inspected for quality by experienced operator. Visual inspection for cracks and blemishes is performed after casting, baking, and sintering. In injection molding, until the machine parameters are optimized, the as-molded components are inspected for cracks, blemishes, blisters, and flow lines. The components are also dissected in different planes to examine for voids (unfilled areas), delaminations, and plastic-rich regions. Occasionally x-radiography is also performed on components such as rotors in the as-molded condition (Fig. 3). Such examinations have been very useful in adjusting processing and machine parameters to produce components of acceptable quality and increase the yield.

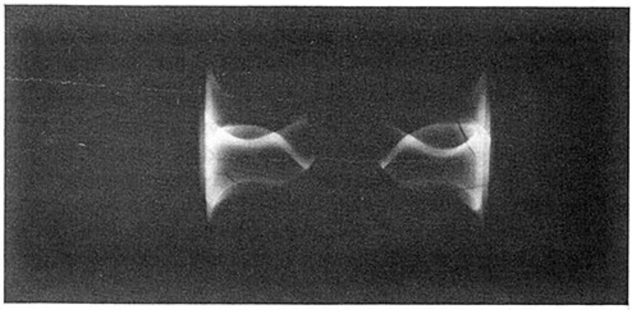


Fig. 3 Cracks in the hub and blade of rotors after baking in the injection-molded silicon carbide rotors

Quality Assurance of Finished Components

Quality assurance requires the capability to detect flaws in fabricated components which arise from process deviation. Standard NDE methods have been optimized for testing Standard Oil's ceramic materials.

- *Fluorescent Penetrant Inspection.* A Zyglo ZA-28-W semiproduction inspection unit has been used in our present facility. Samples are cleaned using a Branson B-92 ultrasonic cleaner. The Zyglo line uses ZL-22C penetrants, ZE-4A emulsifier, and ZP-4B developer to produce maximum sensitivity for small surface defects.
- *X-Radiography.* A Magnaflux Model MXK-100M Microfocus x-ray source, exposure room, and film development dark room have been established for high-resolution radiographs. Other available x-ray facilities include G.E. KXD 325 II and Baltospot 11/5 units. In 1986, we will also be installing, within Standard Oil, a real-time microfocus x-ray system.
- *Ultrasonic Testing.* A state-of-the-art high-frequency ultrasonic system consisting of a broad band 50 MHz transducer, microprocessor-controlled scanning stage, A-scan and C-scan output, HP 1740A recorder is available. In addition, a Panametrics 100 ultrasonic analyzer with a variety of ultrasonic probes is available.
- *Visual Inspection.* Examination of components for surface defects is performed using the low-power magnification of a Wild binocular microscope with photographic capability.
- *Surface Finish.* Determined using a Talysurf 10.
- *Dimensional Measurements.* An array of micrometers, calipers, gages, guillontine gages, an optical comparator, and TalyRand 110 are used for dimensional inspection.

These facilities provide a comprehensive NDE capability to be used in quality assurance. The QA engineer uses these tools and works with customer NDE/QA personnel to establish and ensure conformance to specifications. As new developments, improved equipment or techniques become available, they are added to this facility.

Present NDE Capabilities

The present NDE capabilities are different for research and development and manufacturing groups. *For simple shapes* such as test bars, the following capabilities exist.

(*a*) *Development:*

  (*i*) Can detect surface defects of sizes greater than or equal to 50 μm (0.002 in.).

  (*ii*) Can detect bulk defects of sizes greater than or equal to 150 μm (0.006 in.) depending upon part thickness.

(*b*) *Manufacturing:*

  (*i*) Can detect surface defects of sizes greater than or equal to 150 μm (0.006 in.).

(*ii*) Can detect bulk defects of sizes greater than or equal to 250 μm (0.010 in.) depending upon part thickness.

The above capabilities are valid assuming that an experienced engineer examines components via ultrasonics in the development group, and an experienced inspector examines components via visual, penetrant, and x-ray examination in manufacturing.

The Challenges in NDE and Research Needs

Although considerable progress has been made in NDE development over the last 5-6 years for silicon carbide ceramics [1-4], it is only the beginning, and much more careful and well-coordinated research efforts are needed to bring NDE efforts to a meaningful production environment. Some of the important tasks needing further development are:

1. Unambiguous characterization of detected defects such as type (chemistry), size, shape, and orientation
2. Complex shape examination in detail
3. Performance prediction via NDE signals
4. Development of accept/reject criteria

The NDE research must not only pursue activities with respect to defect detection but must explore classifying potential harmful defects. Considerable research activity and field experience are needed in order to establish reliable and acceptable accept/reject criteria for a given application [5]. At present, we do not know unequivocally whether rejected components (based on present NDE information) would in fact have performed satisfactorily in the field. We do know, however, that some accepted components (based on present NDE methods) do in fact perform less than satisfactorily in some applications. Note that both these situations are not acceptable and add to cost.

Current In-Process NDE Practice

In-process quality control usually is custom tailored to individual product lines and expected performance. As examples, we will elucidate in this paper two product categories: one for current applications and the other for advanced applications such as gas turbine or turbocharger rotors.

Current Applications. In the majority of instances, spray-dried premix is used in the manufacture of various components. A premix is defined as a mixture of submicron alpha silicon carbide powder, sintering additives, and some carbon source. Initially the premix is characterized for its flow density, volatiles (percent), and vibrated density. These characteristics have been found, by experience, to be important in forming operations. For isopressed components, manufacturing isopressing trials are performed. Two 14-inch tubes are isopressed and green machined into four tubes approximately 6 in. long. Inspection is done at this stage and o.d. and i.d. (top and bottom), length, and green weight are all recorded.

After sintering, the qualifier tubes are checked for density (3.10 g/cc minimum), dimensions—o.d and i.d. (top and bottom), length and extent of a warpage. Each tube is then sliced into five to six ring sections and each sliced surface is inspected with a low-power microscope (at 70× magnification). The maximum flaw size per surface and number of flaws greater than 0.005 in. but less than 0.010 in. within one circumferential inch are recorded.

Flaw type characteristics are also noted using a letter code. Certain types of flaws are generated by different processing variables. For example, long narrow circumferential flaws stem from the isopressing operation, while circular or half-moon flaws of random orientation are often related to spherical agglomerates produced in material spray drying. Documentation of flaw type can help identify problems and/or potential process improvements. The measured density, dimensions, and flaw size data are recorded into computer. A section of a ring is polished and an optical micrograph at 200× is obtained. The section is then etched and another optical micrograph at 200× is recorded for crystal size and shape information.

The isopress qualification report data are then examined and the premix is released for manufacture only if the qualification data comply with product requirements. For most applications, the density should be at least 3.10 g/cc and the shrinkage should be appropriate for tooling. Flaw size statistics are monitored in correspondence with customer requirements based on a minimum allowable quality level. This quality level is determined by the quantity of samples free of rejectable flaws (per customer specification) over the total samples examined. An example of statistical data which is used in production qualification including flaw characterization is shown in Appendix II.

Advanced Applications. A summary of current practice in testing injection molded material is given below: After mixing the ceramic powder with plastics and compounding, relative viscosity is measured by using Brabender Plastic-Corder. Sample pucks are made by compression molding in order to check for metal contamination and high-density inclusions. After molding, a few rotors are sliced in half while still warm to check for obvious voids and flaws. All rotors are inspected visually for knit lines primarily using low-power magnification. The weight of the part is also determined. After binder removal, the part is visually examined, sometimes also via x-radiography and weighed. The weight loss is used to determine the extent and success of binder removal. If the molding parameters and binder removal conditions are well optimized then the x-ray inspection is unnecessary since the parts are usually x-ray defect free.

Components such as rotors, after binder removal, are very fragile and the blades break easily, so any examination sequence at this state should consider this aspect. The defect-free components are then sintered. They are post-inspected by using visual, low-power microscopy, x-radiography (sometimes real time), and fluorescent penetrant inspection and ultrasonic NDE as well as density determinations.

Destructive evaluation is also performed, on a selected basis, which includes data acquisition including microstructure, ultrasonic velocity; and on bars cut from components, flexural strength, Weibull modules, and fracture toughness are all obtained. Some rotors are spin tested, as well.

Summary

In this paper we have discussed the currently practiced quality assurance for sintered alpha silicon carbide components. Adequate practices exist although quality control criteria are not formally established in many applications. It is shown that practices differ primarily because of application and expected performance differences. Finally, the progress being made in advanced NDE methods for examination and characterization of complex components can be expected to have significant impact in manufacturing application only when these methods are simple, cost effective, and identify specific accept/reject quality characteristics.

Acknowledgments

Assistance by Mr. Les Carlysle, Mr. Roger Ohnsorg, Mr. Charles Lake, Dr. William Binnie, and Ms. Diana Rebon in providing data for this manuscript is gratefully acknowledged.

References

1 Srinivasan, M., Lawler, D., Inglehart, L. J., Thomas, R. L., and Yuhas, D., "Comparison of NDE Techniques for Sintered SiC Components," *Ceramics Engineering and Science Proceedings*, Vol. 3, Nos. 9-10, 1982, pp. 654-679.

2 Rogers, W. P., and Srinivasan, M., "Ultrasonic Detection of Surface Flaws in Sintered Alpha Silicon Carbide," *Ceramic Engineering and Science Proceedings*, Vol. 5, Nos. 7–8, 1984, pp. 603–613.

3 Friedman, W. D., Bhagat, A. R., Srinivasan, M., and Wilson, J., "An Assessment of Various Methods Which Detect Critical Surface Flaws in Sintered SiC," *5th Review of Progress in Quantitative Non-destructive Evaluation*, Williamsburg, VA, June 1985.

4 Khandelwal, P. K., Kinnick, R. R., and Heitman, P. W., "Photoacoustic Microscopy of Ceramic Turbine Blades," *Am. Ceramic Bulletin*, Vol. 64, No. 8, 1985, pp. 112–115.

5 Evans, A. G., "Non-destructive Failure Prediction in Ceramics," Tech. Rept. No. SC5064, Rockwell Science Center, 1978.

APPENDIX A

Powder Characterization Methods

1 **Pressureless Sinterability.** Density is used as a measure of the sinterability. Samples of the powder are mixed with suitable binders and pressed into discs 1.25 in. in diameter × 0.25 in. thick. After firing, the density and open porosity are measured using ASTM Procedure C-20. Selected pieces are also polished and etched for microscopic determination of grain size, grain shape, and porosity.

2 **Bulk Composition.** Previously established instrumental and chemical methods are employed to determine the general chemistry of the powder product. Free carbon and free silicon, total carbon, nitrogen and oxygen, and silica are measured to a fraction of a percent by the following methods.

2a Free Carbon. The free carbon analysis is done by heating a weighed sample in an oxygen stream at 800°C for a specified time. The CO_2 generated by oxidation of the free carbon is absorbed in a glass bulb and weighed. In the case of submicron powders, oxidation of part of the SiC must be considered. A correction for this is made from the sample weight before and after the test. The analysis is based on the premise that clean SiC is used. Appreciable amounts of oxidizable metals or silicon would alter the results.

2b Total Carbon. Total carbon analysis is done using a Leco WR-12 or CS-244. The sample is combusted in an oxygen atmosphere in a small induction furnace, with accelerators added to ensure complete destruction of the sample. The carbon content is read directly from the instrument.

2c Free Silicon. Free silicon is determined by the silver fluoride method. The sample is first treated with HCl to remove iron and is then washed until it is free of chlorides. Silver fluoride is added to the sample in a plastic funnel and allowed to react for a short period of time. The free silicon reacts with the silver fluoride to form SiF_4, which volatilizes, and metallic silver. The excess silver salts are washed from the filter paper with water and discarded. The metallic silver is collected using hot nitric acid and titrated with $0.1N$ potassium thiocyanate. The corresponding amount of free silicon is then calculated.

2d Free Silica. A weighed sample is fired at 800°C to constant weight. The residue is then carefully treated with HF and H_2SO_4, which volatilize the free SiO_2. The crucible is again ignited at 800°C, and the loss in weight caused by the acid treatment is taken as the free silica content.

2e Oxygen. Oxygen is measured using a Leco RO-116 or TC-136. The oxygen in the sample is converted to CO, which is analyzed by an infrared detector. The oxygen content is read directly from the instrument. Up to 6 percent O_2 can be measured, with a detection limit of 0.1 ppm, for a 1 gram sample.

2f Nitrogen. Nitrogen analysis is done with a Leco TC-136, using standard procedures for that instrument. The nitrogen content is measured with a thermal conductivity cell, and is read directly from the TC-136.

3 **Particle Size/Size Distribution.** The particle size distribution is measured using one of several Leed & Northrup Microtrac units. The small Particle Analyzer (SPA) covers the 0.1–20 μm range using scattered laser light. With proper dispersion, the SPA can be used to determine the size distribution of individual powder particles. Samples may be introduced in a wet or dry state. Data collected in 14 size channels are summed to give a cumulative distribution curve. However, if the sample contains particles which are 0.1 μm in diameter, these would not be detected by the system.

The standard Microtrac covering the 2–176 μm range or one with an extended range of 3–300 μm might be useful in gaining information about agglomeration. Here efforts to disperse the discrete particles would be minimized. Thirteen size channels are available on this instrument.

Both of these instruments also give other statistical information.

4 **Particle Shape.** The shape and morphology of powder particles are studied by scanning electron microscopes and, where appropriate, the information will be supplemented by that obtained from the microprobe mode of such instruments. Microscope available is the AMR Model 1400. The resolution and depth of field offered by such instruments make them invaluable for direct viewing of submicron particles and of small agglomerates by using dispersed and nondispersed materials. Sufficient materials are examined in order to eliminate sampling problems, and the number of samples viewed is determined by the particle size distribution.

5 **Phase Composition.** The polytypes of SiC are identified and levels determined by employing general automated x-ray diffractometers and, if required, by powder Gandolfi and single-crystal film cameras. X-ray diffraction procedures offer sensitivities in the range of a few percent to several percent, depending on which polytypes are present. Precision and accuracy are also in the few percent level. In addition, limited information regarding minimum crystallite size and average shape is also available from diffraction patterns. It is not anticipated that low-trace levels of crystalline phases other than SiC will be readily identified without the application of selected particle diffraction.

6 **Tap Density.** Tap density is a measure of the packing characteristics of the powder. Specifications for this test are found in ASTM-B 527, a test designed for refractory metals and compounds. A weighed sample is tapped mechanically for specified number of taps. The volume occupied by the powder is read from the cylinder, and weight divided by this volume is the tap density. Tap density is calculated in g/cm^3 to the nearest hundredth. The average of three tests is reported. This capability exists within Standard Oil Co., but it has not been applied to submicron powders. The fineness of the powder could be a problem.

7 **Surface Area.** Surface area is determined by standard gas adsorption techniques, using BET method, with an accuracy of ±10 percent. The in-house instrument is a Micromeritics 2100 D. Multipoint nitrogen adsorption at 77 K is the usual technique. Successive amounts of gas are adsorbed by the sample and the changes in pressure in the system are used to calculate the amount of gas adsorbed. Reproducibility is ± 0.2 m^2/g for the surface area range of interest, or approximately ±2 percent of the value. Also available is a Quantachrome unit that uses a single point method to determine the BET surface area. Reproducibility is 3–5 percent of the value measured.

APPENDIX B

Statistical Analysis Flaw Size

VARIABLE	N	MEAN	STANDARD DEVIATION (INCH)	MINIMUM VALUE (INCH)	MAXIMUM VALUE (INCH)	STD ERROR OF MEAN (INCH)
LENGTH	40	0.0021	0.001	0.001	0.006	0.001
WIDTH	40	0.0016	0.0006	0.001	0.003	0.000

FLAW SIZE MEASUREMENTS

TUBE ID	LENGTH (INCH)	WIDTH (INCH)	TYPE	WITHIN	TUBE ID	LENGTH (INCH)	WIDTH (INCH)	TYPE	WITHIN
13	0.001	0.001	C	0	33	0.002	0.002	C	0
15	0.001	0.001	C	0	33	0.002	0.001	A	0
24	0.001	0.001	C	0	35	0.002	0.002	C	0
25	0.001	0.001	C	0	36	0.002	0.001	A	0
31	0.001	0.001	C	0	42	0.002	0.001	A	0
32	0.001	0.001	C	0	43	0.002	0.001	A	0
34	0.001	0.001	C	0	43	0.002	0.002	C	0
34	0.001	0.001	C	0	44	0.002	0.001	A	0
45	0.001	0.001	C	0	44	0.002	0.001	C	0
11	0.002	0.002	C	0	45	0.002	0.002	C	0
13	0.002	0.002	C	0	12	0.003	0.002	A	0
14	0.002	0.002	C	0	14	0.003	0.003	C	0
15	0.002	0.002	C	0	22	0.003	0.002	A	0
16	0.002	0.002	C	0	35	0.003	0.003	C	0
22	0.002	0.001	A	0	41	0.003	0.002	A	0
23	0.002	0.002	C	0	42	0.003	0.002	A	0
24	0.002	0.002	C	0	46	0.003	0.002	A	0
25	0.002	0.001	A	0	12	0.004	0.002	A	0
26	0.002	0.002	C	0	21	0.004	0.002	A	0
32	0.002	0.002	C	0	23	0.006	0.003	A	0

Flaw Type Code: A = Long, narrow flaw concentric to tube diameter.
C = Round flaw with no specific orientation.

On the Premixed Combustion in a Direct-Injection Diesel Engine

A. C. Alkidas
General Motors Research Laboratories,
Engine Research Department,
Warren, MI 48090-9055

The factors influencing premixed burning and the importance of premixed burning on the exhaust emissions from a small high-speed direct-injection diesel engine were investigated. The characteristics of premixed and diffusion burning were examined using a single-zone heat-release analysis. The mass of fuel burned in premixed combustion was found to be linearly related to the product of engine speed and ignition-delay time and to be essentially independent of the total amount of fuel injected. Accordingly, the premixed-burned fraction increased with increasing engine speed, with decreasing fuel-air ratio and with retarding injection timing. The hydrocarbon emissions did not correlate well with the premixed-burned fraction. In contrast, the oxides of nitrogen emissions were found to increase with decreasing premixed-burned fraction, indicating that diffusion burning, and not premixed burning, is the primary source of oxides of nitrogen emissions.

Introduction

Combustion in diesel engines principally consists of an initial premixed-combustion stage, followed by the main diffusion-combustion stage [1-4]. As shown in Fig. 1, the premixed-combustion stage commences at the end of the ignition delay period and is associated with a short-duration high rate of heat release. The heat-release history during the premixed-combustion stage is strongly dependent on the amount of fuel that mixes with air to form a combustible mixture during the ignition-delay period.

In the direct-injection diesel the diffusion-combustion stage is generally associated with lower heat-release rates than those of the premixed-combustion stage. The heat-release history during this stage is controlled by the mixing rate since during this period the characteristic time of combustion is much shorter than the characteristic time of fuel–air mixing [3, 4].

The premixed-combustion stage is of great importance because of its high rate of heat release, which results in a rapid rate of pressure rise. This in turn affects the noise level of the engine. In addition, recent studies [5, 6] suggested that the premixed-combustion stage strongly influenced the exhaust emissions from diesel engines.

Wade and Hunter [5] showed, from a relatively limited amount of data, that both the hydrocarbon (HC) and oxides of nitrogen (NO_x) emissions increased with increasing premixed-burned fraction (fraction of fuel burned in premixed combustion). Furthermore, they stated that the reason that direct-injection diesel engines produce higher HC and NO_x emissions than divided-chamber diesel engines is that the former engines have larger premixed-burned fractions.

The magnitude of the premixed-burned fraction was also used by Wade and co-workers [6] to explain the trends of HC and NO_x emissions observed from an uncooled direct-injection diesel engine. In contrast to several other studies (e.g., [7-9]), Wade and co-workers [6] found that their uncooled diesel engine produced lower NO_x and HC emissions than the corresponding water-cooled engine. They argued that the lower NO_x and HC emissions were caused by a reduction in the amount of premixed combustion that resulted from the shortened ignition-delay period of the uncooled engine.

Several investigators [3, 10-12] have dealt with the premixed-combustion stage of diesel engines. Watson et al. [3] and Miyamoto and co-workers [10, 11] developed empirical relationships between the premixed-burned fraction and parameters such as ignition-delay period and equivalence ratio. The premixed-burned fraction was found to increase with increased engine speed, decreased fuel–air ratio, and lengthened ignition-delay period. Gerpen and co-workers [12] found that, in general, the premixed-burned fraction increased with increasing air swirl and increasing injection pressure.

The aims of the present study were to investigate the factors which influence the premixed-combustion characteristics of a

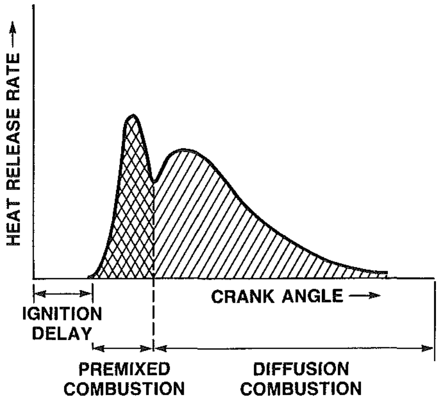


Fig. 1 Typical heat-release history of a direct-injection engine

---

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources and Technology Conference and Exhibition, New Orleans, Louisiana, February 23-27, 1986. Manuscript received at ASME Headquarters December 23, 1985. Paper No. 86-ICE-4.

Table 1 Engine description

Bore, mm	85
Stroke, mm	94
Connecting rod length, mm	160
Displacement, L	0.533
Compression ratio	21.7:1

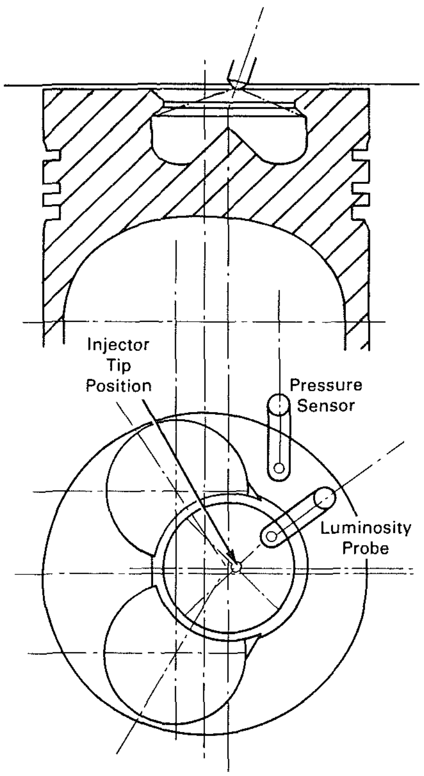


Fig. 2 Combustion-chamber geometry

small high-speed direct-injection engine and to evaluate the role of premixed combustion as a primary source of exhaust emissions.

The combustion characteristics of the engine at various engine conditions were examined using a heat-release analysis. This analysis was a single-zone thermodynamic model which utilized measured cylinder-pressure histories. The engine operational parameters examined were fuel–air ratio, engine speed, and injection timing.

Apparatus and Procedure

The measurements were made in an experimental single-cylinder 0.53-L DI diesel engine. The principal dimensions of the engine are listed in Table 1. A cross-sectional view of the piston geometry of the engine is shown in Fig. 2. The piston bowl was a reentrant type with its axis offset 6.2 mm from the cylinder axis. The fuel injector was a mechanical unit injector. It had a VSO (Valved-Sealed Orifice) tip with four 0.20-mm-dia holes. The opening pressure of the injector was nominally 18 MPa.

The engine tests were performed for a range of fuel–air ratios (0.018 to 0.040), engine speeds (800 to 3300 rpm), and injection timings (9 to 0° BTDC). The corresponding baseline values of these parameters were: fuel–air ratio = 0.022, engine speed = 1500 rpm, injection timing = 6° BTDC. In all tests, the intake air was maintained at ambient pressure and at 310 K temperature.

Heat-Release Analysis

The heat-release model used is a single-zone model similar to that developed by Krieger and Borman [13] that utilizes the measured pressure history of the gases in the combustion chamber to calculate the heat-release history of the fuel. The model is based on the first law of thermodynamics and considers the cylinder contents to be a homogeneous mixture of ideal gases (air and products of combustion) which are always in thermodynamic equilibrium. The products of combustion are assumed to be in chemical equilibrium, and their thermodynamic properties are computed based on the temperature, pressure, and the overall equivalence ratio of the mixture.

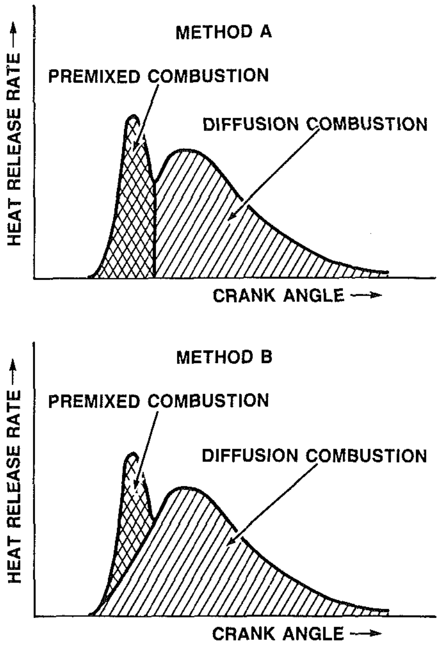


Fig. 3 Definitions for the premixed-combustion and diffusion-combustion stages

The instantaneous surface heat transfer, $Q_w(\theta)$, is described by an equation of the form:

$$Q_w = h_g(\theta) \sum A_i(\theta)[T_g(\theta) - T_{w,i}] \qquad (1)$$

where $T_g(\theta)$ and $h_g(\theta)$ are the gas temperature and convective heat transfer coefficient, respectively, at crank angle θ; A_i represents the surface areas of the cylinder head, piston, and liner; and $T_{w,i}$ represents their corresponding time-averaged temperatures. At each crank angle the convective heat transfer coefficient is computed from an empirical correlation developed by Woschni [14]. This correlation is scaled (constant multiplier) such that the computed final fuel-burned fraction is equal to a specified value which is estimated from exhaust-emissions measurements.

The fraction of fuel burned in premixed combustion, and consequently the fraction of fuel burned in diffusion combustion, was calculated from the computed heat-release history of the fuel. Because the definition of premixed burning from a heat-release history is somewhat arbitrary, in the present study premixed-fuel-burned fractions were computed using two definitions which are demonstrated in the schematics of Fig. 3. The first definition (see also Wade and Hunter [5]) implies that diffusion burning starts at the end of premixed burning, whereas the second definition implies that premixed and diffusion burning occur concurrently. It is obvious that the amount of fuel burned in the premixed mode computed by method A would always be larger than that computed by method B. However, the question to be answered is whether the results of the two methods show similar trends when the influences of operational parameters of the engine on premixed combustion are examined.

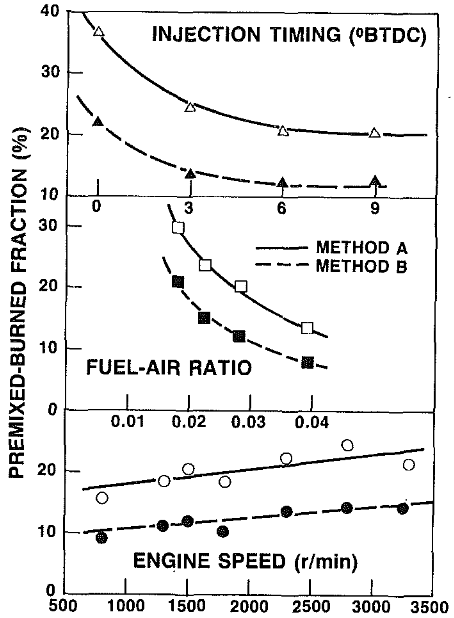


Fig. 4 Premixed-burned fractions as functions of engine speed, fuel-air ratio, and injection timing

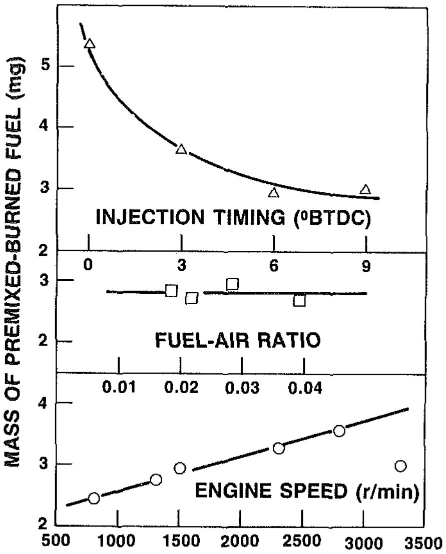


Fig. 5 Variations of the mass of premixed-burned fuel with engine speed, fuel-air ratio, and injection timing

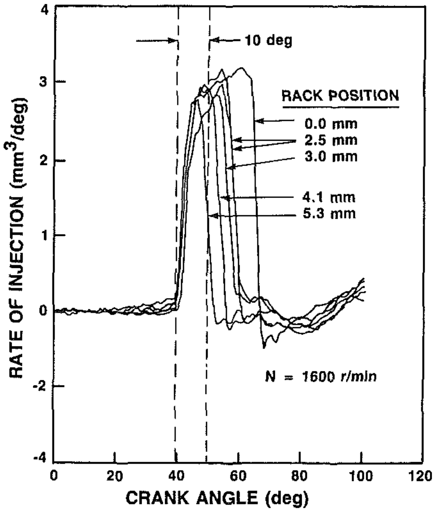


Fig. 6 Rate-of-injection histories for different rack positions [15]

Premixed-Combustion Characteristics

With the other operating parameters held at their baseline values, the variations of the premixed fractions computed by methods A and B with engine speed, fuel-air ratio, and injection timing are shown in Fig. 4. The premixed fractions computed by the two methods showed similar behaviors with changes in engine speed, fuel-air ratio, and injection timing. Because of this similarity and the somewhat arbitrary definition of the premixed mode of combustion, in the subsequent sections the premixed combustion was computed by method A only. This is also the method used by other investigators [5, 10].

Figure 4 shows that the premixed-burned fraction increases with increasing engine speed, with decreasing fuel-air ratio, and with retarding injection timing. The premixed-burned fraction appears to be very sensitive to variations in fuel-air ratio and injection timing. This sensitivity increases with decreasing fuel-air ratio and retarding injection timing. In contrast, engine speed appears to have a moderate effect on the premixed-burned fraction. The premixed-burned fraction varies linearly with engine speed.

To understand the above trends it may be instructive to examine the influence of the operating parameters of the engine on the mass of fuel that burns in the premixed mode of combustion. Figure 5 shows the variations of the mass of premixed-burned fuel with engine speed, fuel-air ratio, and injection timing. For a constant fuel-air ratio and injection timing, the mass of premixed-burned fuel increased linearly with engine speed. For a constant engine speed and injection timing, the mass of premixed-burned fuel remained relatively unchanged with fuel-air ratio. Lastly, for a constant engine speed and fuel-air ratio, the mass of premixed-burned fuel increased dramatically with retarded injection timing.

The fact that the amount of fuel that burns in premixed combustion is not significantly affected by load is not entirely unexpected since the quantity of fuel injected during the ignition-delay period does not change significantly with load [15]. This is demonstrated in Fig. 6, which shows rate-of-injection measurements for different rack positions. For a period of 10 crank-angle degrees during the initial period of injection, which corresponds to about the longest ignition-delay period encountered in these tests, the rate-of-injection history appears to be essentially independent of rack position.

The above trends are also depicted for a speed of 1500 rpm in Fig. 7, which is a plot of the mass of premixed-burned fuel versus the mass of fuel injected. From these results one may make the following deductions on the factors that influence premixed burning in direct-injection diesel engines. The mass of fuel burned in the premixed mode of combustion (*a*) is essentially independent of the mass of fuel injected,[1] (*b*) increases with increasing mixing rate, which is believed to be the primary cause of the observed increase in the mass of premixed-burned fuel with increasing engine speed, and (*c*) increases with increasing ignition-delay period, which is believed to be the primary cause of the observed increase in the mass of premixed-burned fuel with retarding injection timing.

Figure 8 shows the mass of fuel burned in the premixed mode of combustion m_{PR} as a function of the ignition-delay period θ_{ID}. The ignition delay period is expressed in crank-angle degrees, which is directly proportional to the product of engine speed N and ignition-delay time τ_{ID}.

In general, the mass of premixed-burned fuel increases with increasing ignition-delay period. With the exception of the

---

[1] This observation may be applicable only to engines with injection-system characteristics similar to those of the engine used in this study.

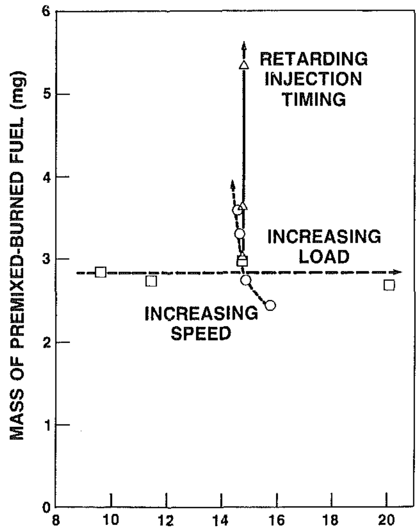


Fig. 7 Mass of premixed-burned fuel versus mass of fuel injected

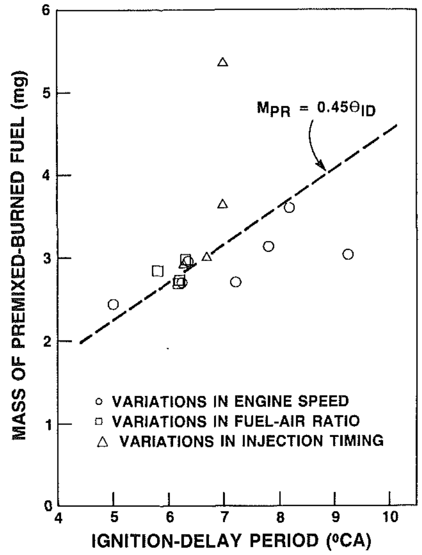


Fig. 8 Mass of premixed-burned fuel versus ignition-delay period

most retarded injection-timing condition (θ_{inj} = 0° BTDC) and of the highest engine-speed condition (N = 3300 rpm), the mass of premixed-burned fuel appears to vary linearly with the ignition-delay period. The straight line shown in Fig. 8 results from a linear regression of the data with an added condition that when the ignition-delay period is zero, the mass of premixed-burned fuel is zero.

From the linear relationship between m_{PR} and θ_{ID} and from the experimental result that m_{PR} is independent of the total mass of fuel injected m_F, one may derive the following empirical relationship for the fraction of premixed-burned fuel, β:

$$\beta = 0.45\ \theta_{ID}/m_F \tag{2}$$

where m_F, the total mass of fuel injected, is expressed in units of mg, and θ_{ID}, the ignition-delay period, is in crank-angle deg.

Empirical equations for the fraction of premixed-burned fuel also were obtained by Watson and co-workers [3] and Miyamoto and co-workers [10].

Miyamoto and co-workers [10] developed heat-release correlations (Wiebe functions) for the premixed and diffusion combustion using direct-injection and divided-chamber single-cylinder four-stroke naturally aspirated diesel engines. They found that the mass of premixed-burned fuel varied linearly with the ignition-delay time for constant engine speed and that the mass of premixed-burned fuel increased with increasing engine speed for a fixed ignition-delay time. These results are generally in agreement with the results of the present investigation. Miyamoto and co-workers tested several fuel blends (diesel, rapeseed oil, palm oil, ethanol, and methanol) and found that the amount of premixed combustion correlated best with the amount of fuel injected during the ignition-delay period. For the DI engine their correlation may be reduced to the form:

$$\beta = 0.50\ Y_{ID} \tag{3}$$

Equation (3) states that 50 percent of the fuel injected during the ignition-delay period in a direct-injection engine burns in the premixed mode. The corresponding number for a divided-chamber engine was 88 percent. The data for the direct-injection engine were not well correlated by equation (3). In fact, a close examination of their data (see Fig. 10 of [10]) shows that for diesel fuel only, equation (3) is a poor cor-

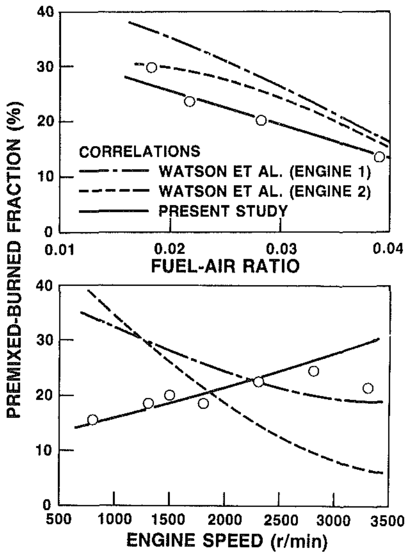


Fig. 9 Comparisons of Watson's correlations with the correlation and data obtained in the present study

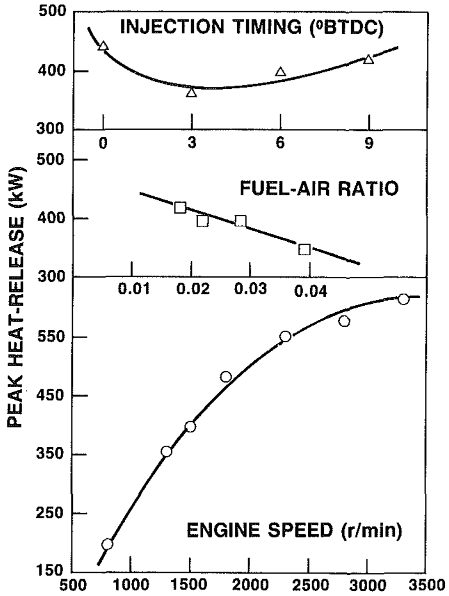


Fig. 10 Variations of the peak rate of heat release with engine speed, fuel-air ratio, and injection timing

Table 2 Empirical constants of equation (2)

Engine*	a	n	m
1	0.926	0.37	0.26
2	0.950	0.41	0.28
3	0.810	0.28	0.51

*Watson et al. [3] gave little description of their three four-stroke direct-injection diesel engines. Engine 1 was a six-cylinder in-line diesel with inlet-port swirl generation. Engine 2 was a V-8 turbocharged and intercooled diesel. Engine 3 was a turbocharged and intercooled diesel but much more highly rated than engines 1 and 2.

relation of the data. In addition, Miyamoto and co-workers [10] expressed the amounts of premixed combustion in units of MJ/kmole, which when converted to appropriate mass units result in unrealistically large quantities of premixed-burned fuel.

Watsons's correlation is of the form

$$\beta = 1.0 - a\,F^{n}\,(\tau_{ID})^{-m} \qquad (4)$$

where F is the equivalence ratio. The values of the empirical constants a, n, and m for the three turbocharged four-stroke open-chamber diesel engines tested are shown in Table 2.

Figure 9 shows a comparison of the correlations obtained by Watson et al. [3] for engines 1 and 3 with the data and correlation (equation (2)) obtained in the present study. The two Watson correlations[2] show large disagreements at high speeds and light loads. However, most distressing is the fact that both correlations predict a decrease in the premixed-burned fraction with increasing engine speed, which contradicts the findings of this study and of Miyamoto and co-workers [10].

The influence of fuel-air ratio, engine speed, and injection timing on the peak rate of heat release due to premixed burning is shown in Fig. 10. The peak rate of heat release increases with increasing engine speed and decreasing fuel-air ratio; the relationship between this rate and injection timing exhibits a minimum for an injection timing of 3° BTDC.

Influence of Premixed Burning on Emissions

As stated in the Introduction, Wade and co-workers [5, 6] have suggested that premixed burning in diesel engines strongly influences the hydrocarbon (HC) and oxides of nitrogen (NO_x) emissions. Both HC and NO_x emissions were found to increase with increasing premixed-burned fraction.

Considering first the hydrocarbon emissions, Fig. 11 shows the emission index of hydrocarbons, EIHC, versus the premixed-burned fraction β for the various operational conditions examined. Increasing the load (while keeping injection timing and engine speed constant) decreased both EIHC and β. Increasing the engine speed increased β and produced a minimum EIHC in the midspeed range. Retarding the injection timing had little effect on EIHC but increased the premixed-burned fraction dramatically.

The above results suggest that, in contrast to the findings of Wade and Hunter [5], hydrocarbon emissions do not correlate well with the premixed-burned fraction. This is not too surprising, however, because a number of investigators [16-19] showed that there are several sources of hydrocarbon emissions from diesel engines. These sources are: (*a*) air-fuel mixing beyond the limit of combustion during the ignition-delay period, (*b*) fuel emptying from the volume of the injector sac, (*c*) bulk quenching of both flame and oxidizing nonflame reactions, (*d*) wall quenching of the flame reactions, and (*e*) lubricating oil sources. For example, Alkidas [19] found that bulk quenching and lubricating oil were the primary sources of HC emissions from a divided-chamber diesel engine. Mixing beyond the lean limit of combustion during the ignition-delay period and wall quenching were secondary sources of HC emissions.

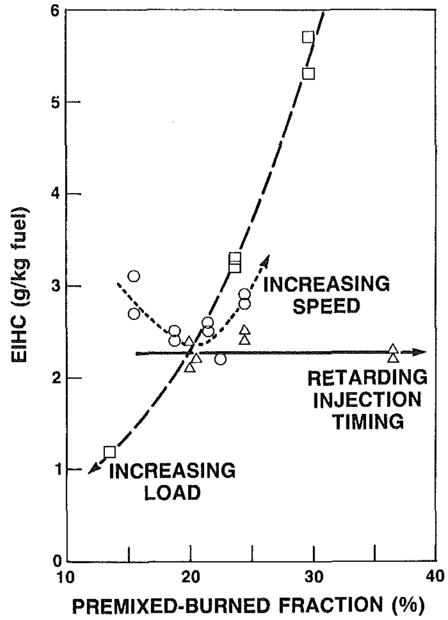


Fig. 11 Emission index of HC as a function of the premixed-burned fraction

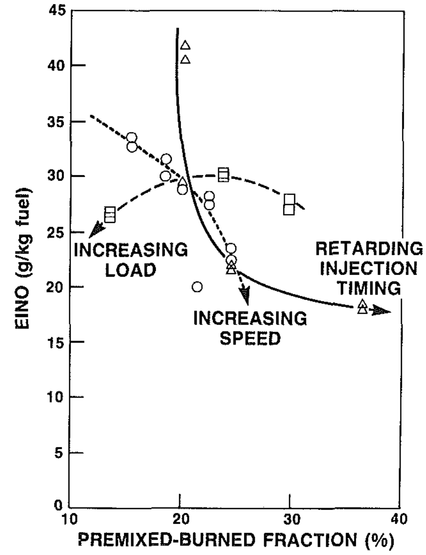


Fig. 12 Emission index of NO as a function of the premixed-burned fraction

Considering next the nitric oxide emissions, Fig. 12 shows the emission index of nitric oxide, EINO, versus the premixed-burned fraction, β, for the various operational conditions examined. Increasing the load (while keeping other parameters constant) decreased β, but EINO typically (see, e.g. [20]) increased at light loads, reaching a maximum, and then decreased. In comparison to β, however, the variation of EINO with load was not large (less than about 10 percent). Increasing the engine speed increased β and produced a peak in EINO in the mid-speed range. Retarding the injection timing decreased EINO while β was increased.

---

[2] Watson's correlation for engine 2 is not presented because it was found to be in close agreement with the correlation for engine 1 for the present test conditions.

The results of Fig. 12, which show that the nitric oxide emissions generally decreased with increasing premixed-burned fraction, suggest that, in contrast to the contentions of Wade and co-workers [5, 6], diffusion burning rather than premixed burning is the principal source of nitric oxide emissions in diesel engines.

Closure

The fraction of fuel burned in premixed combustion in the small high-speed direct-injection diesel engine used in this investigation may be computed with the aid of the established empirical equation (2). In general, the premixed-burned fraction may be expected to be linearly related to the product of engine speed and ignition-delay time and to be inversely proportional to the amount of fuel injected in the cylinder. However, equation (2) may not be directly applicable to other direct-injection engines because the amount of fuel burned in premixed combustion may be affected by several parameters which are characteristic of a particular engine-combustion system. In addition to the combustion-chamber-geometry parameters, other important parameters are air swirl, rate of fuel injection, and number of injector holes [4]. The amount of fuel burned in premixed combustion is expected to increase with increasing air swirl, increasing rate of fuel injection and increasing number of injector holes.

Conclusions

The following conclusions are based on results of this investigation, which was performed on a small single-cylinder high-speed direct-injection diesel engine:

1 The mass of fuel burned in the premixed mode of combustion was found to be essentially independent of the mass of fuel injected and, in general, to increase linearly with the product of engine speed and ignition-delay time.

2 In agreement with the above observations, the fraction of fuel burned in premixed combustion (premixed-burned fraction) increased with increasing engine speed, with decreasing fuel–air ratio, and with retarding injection timing.

3 The hydrocarbon emissions did not correlate well with the premixed-burned fraction. This indicates that air–fuel mixing beyond the lean limit of combustion during the ignition-delay period may not be a major source of hydrocarbon emissions in this engine.

4 The nitric oxide emissions were found to increase with decreasing premixed-burned fraction. This indicates that diffusion burning, rather than premixed burning, may be the primary source of oxides of nitrogen emissions.

Acknowledgments

The experimental program was initiated and completed by J. P. Myers, presently of the General Motors CPC Group.

References

1 Lyn, W. T., "Study of Burning Rate and Nature of Combustion in Diesel Engines," *Ninth Symposium (International) on Combustion*, Academic Press, 1963, pp. 1069–1082.

2 Meguerdichian, M., and Watson, N., "Prediction of Mixture Formation and Heat Release in Diesel Engines," SAE Paper No. 780225, 1978.

3 Watson, N., Pilley, A. D., and Marzouk, M., "A Combustion Correlation for Diesel Engine Simulation," SAE Paper No. 800029, 1980.

4 Plee, S. L., and Ahmad, T., "Relative Roles of Premixed and Diffusion Burning in Diesel Combustion," *SAE Transactions*, Vol. 92, 1983, pp. 4.892–4.909.

5 Wade, W. R., and Hunter, C. E., "Analysis of Combustion Performance of Diesel Fuel," CRC Workshop on Diesel Fuel Combustion Performance, Atlanta, GA, Sept. 14, 1983.

6 Wade, W. R., Havstad, P. H., Ounsted, E. J., Trinker, F. H., and Garwin, I. J., "Fuel Economy Opportunities With an Uncooled DI Diesel Engine," *Fuel Efficient Power Trains and Vehicles*, Inst. Mech. Engrs., Paper No. C432/84, 1984, pp. 59–72.

7 Toyama, K., Yoshimitsu, T., Nishiyama, T., Shimauchi, T., and Nakagaki, T., "Heat Insulated Turbocompound Engine," *SAE Transactions*, Vol. 92, 1983, pp. 3.1086–3.1097.

8 Belardini, P., Bertoli, C., Corcione, F. E., and Police, G., "Thermal Barriers Adoption in D. I. Diesel Engines: Effect of Smoke and Gaseous Emissions," SAE Paper No. 840995, 1984.

9 Cole, R. M., and Alkidas, A. C., "Evaluation of an Air-Gap-Insulated Piston in a Divided-Chamber Diesel Engine," *Advances in Adiabatic Engines*, SAE SP-610, Paper No. 850359, 1985, pp. 37–48.

10 Miyamoto, N., Chikahisa, T., Murayama, T., and Sawyer, R., "Description and Analysis of Diesel Engine Rate of Combustion and Performance Using Wiebe's Functions," SAE Paper No. 850107, 1985.

11 Murayama, T., Miyamoto, N., Yamada, T., Kawashima, J., and Itow, K., "A Method to Improve the Solubility and Combustion Characteristics of Alcohol-Diesel Fuel Blends," *SAE Transactions*, Vol. 91, 1982, pp. 3485–3494.

12 Van Gerpen, J. H., Huang, C., and Borman, G. L., "The Effects of Swirl and Injection Parameters on Diesel Combustion and Heat Transfer," SAE Paper No. 850265, 1985.

13 Krieger, R. B., and Borman, G. L., "The Computation of Apparent Heat Release for Internal Combustion Engines," ASME Paper No. 66-WA/DGP-4, 1966.

14 Woschni, G., "A Universally Applicable Equation for the Instantaneous Heat Transfer Coefficient in the Internal Combustion Engine," *SAE Transactions*, Vol. 76, 1967, pp. 3065–3083.

15 Henningsen, S., General Motors Research Laboratories, unpublished results, 1981.

16 Greeves, G., Khan, I. M., Wang, C. H. T., and Fenne, I., "Origins of Hydrocarbon Emissions From Diesel Engines," *SAE Transactions*, Vol. 86, 1977, pp. 1235–1251.

17 Yu, R. C., Wong, V. W., and Shahed, S. M., "Sources of Hydrocarbon Emissions From Direct-Injection Diesel Engines," *SAE Transactions*, Vol. 89, 1980, pp. 324–335.

18 Henningsen, S., "Hydrocarbon Emissions From the Ignition-Delay Period in a Direct-Injection Diesel Engine," SAE Paper No. 841381, 1984.

19 Alkidas, A. C., "Hydrocarbon Emissions From a Single-Cylinder, Divided-Chamber Diesel Engine," SAE Paper No. 841380, 1984.

20 Alkidas, A. C., and Cole, R. M., "Gaseous and Particulate Emissions From a Single-Cylinder, Divided-Chamber Diesel Engine," *SAE Transactions*, Vol. 92, 1983, pp. 3.802–3.814.

Heat Release Studies in a Divided-Chamber Diesel Engine

A. C. Alkidas
Engine Research Department,
General Motors Research Laboratories,
Warren, MI 48090-9057

The influences of operational parameters on the heat release and heat transfer characteristics of a divided-chamber diesel engine were examined. Increasing the fuel-air ratio increased the heat release rate, as expected, and increased the duration of combustion. Near the beginning and end of combustion the mass-burned rate was found to increase in direct proportion to an increase of engine speed. In contrast, in the central part of the combustion duration, the mass-burned rate was found to increase at a higher rate than engine speed. For motored conditions, the computed area-averaged heat-flux histories were found to be in reasonable agreement with the corresponding measured local heat-flux histories.

Introduction

Diesel engine combustion is an exceedingly complex phenomenon which involves the burning of a high-pressure transient liquid-fuel spray in an unsteady turbulent flow which is confined by the walls of a combustion chamber whose geometry is generally complex and varies with time. The combustion process is characterized by the nonhomogeneous fuel distribution in the combustion chamber and by the strong influence of air-fuel mixing [1, 2].

In view of the above, it is not surprising that detailed modeling of diesel engine combustion is in its infant stage, although research activity in this area has substantially increased [3]. Presently most modeling studies of diesel combustion use engine heat release and simulation analyses, which in general require either pressure-time data or an empirical burning-rate history. The simulation analyses, despite their limitations, are well suited for parametric studies [2].

The present investigation utilized a heat release analysis in conjunction with heat-flux measurements to analyze details of the combustion processes and to examine the influence of operational parameters (engine speed, fuel-air ratio, and combustion timing) on the combustion processes in a divided-chamber diesel engine. This type of diesel engine provided the added complication that combustion occurs in two chambers, the antechamber and the main chamber, which are separated by a short, narrow connecting passage.

Heat release studies in divided-chamber diesel engines have been performed by several investigators [4-13]. Of these investigators, Heywood and co-workers [12] used a stochastic mixing model, and Leone et al. [10] used a simple air-entrainment model to simulate combustion in the two chambers of the engines. The accuracy of these simulation analyses is limited by the accuracy with which all the physical and chemical processes are modeled. In contrast, both Anisits [6] and Watson and Kamel [7, 8] utilized measured pressure histories in the main chamber and antechamber to compute the heat release rates in the individual chambers. However, because of the small difference in pressures in the two chambers, these analyses require tedious matching of the two pressure transducers, and often the required accuracy of the measurement is beyond the capability of the available pressure transducers. To circumvent this problem Terada [9, 11] used measured main-chamber pressures and ionization measurements in the connecting passage to compute the heat release rates in the two chambers.

In this study the heat release analysis used was a single-zone single-chamber thermodynamic analysis based on main-chamber pressure-time data. The computed heat release rate is an overall rate in which no distinction is made between the two separate combustion chambers. Single-chamber heat release analyses of divided-chamber diesel engines were also performed by Bowden et al. [4] and Kort et al. [13]. The heat release analysis of Kort and co-workers [13] utilized main-chamber and antechamber pressure-time data as input. However, the calculated rate of heat release included the effect of surface heat transfer.

Apparatus and Diagnostics

Engine. The engine used in this study was an experimental single-cylinder divided-chamber diesel engine whose principal specifications are summarized in Table 1. The cylinder head of the engine was instrumented with three surface heat-flux probes and a pressure transducer. Two of the surface heat-flux probes, designated TM1 and TM2, were located in the main chamber. The other probe, designated TS1, was located in the antechamber. The locations of these probes relative to the valves, cylinder, and antechamber are shown in Fig. 1.

The fuel injector was instrumented to monitor the poppet-lift profile. The glow plug was replaced with a luminosity sensor, of the same external geometry, to detect the start of combustion. The sensing element of the luminosity sensor was a photodiode.

Pressure Measurements. The main-chamber pressure history was measured using a water-cooled piezoelectric pressure transducer. At each test condition, the pressure was

---

Contributed by the Internal Combustion Engine Division for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received at ASME Headquarters February 20, 1986. Paper No. 85-DGP-22.

Table 1 **Engine specifications**

Antechamber: double-throat swirl chamber	
Bore, mm	103.0
Stroke, mm	85.6
Connecting rod length, mm	223.5
Displacement, L	0.718
Antechamber volume, L	0.0147
Compression ratio	19.2:1
Intake valve opens, CA	6 BTC
Intake valve closes, CA	38 ABC
Exhaust valve opens, CA	64 BBC
Exhaust valve closes, CA	17 ATC
Injection system:	
Rotary pump	
Poppet injector (opening pressure = 7.9 MPa)	

recorded at each crank angle (CA) for 96 consecutive cycles. From these data the average pressures were calculated, together with the corresponding standard deviations at each CA.

Surface Heat-Flux Measurements. The heat-flux probes were modified fast-response surface thermocouples of the Bendersky type having a second thermocouple placed at a known distance from the tip of the probe [14]. The surface temperatures at locations TM1, TM2, and TS1 were recorded, simultaneously with main-chamber pressure, at each crank angle for 96 consecutive cycles. From these data the local average surface-temperature histories were computed and used to calculate the corresponding local heat-flux histories.

The surface heat flux may be separated into two components: the steady-state and the unsteady components. The steady-state component was evaluated from the time-averaged temperature difference between the surface thermocouple and the second, in-depth thermocouple. The unsteady component was evaluated from the measured cyclic surface-temperature variation. This method is described in several studies [15, 16]. Briefly, the measured cyclic surface-temperature variation is represented by a Fourier series of sine and cosine terms. The surface heat flux is derived from the solution of the one-dimensional unsteady heat-conduction problem with a surface-temperature boundary condition described by the above harmonic analysis of the measured surface-temperature history.

Heat-Release Analysis. The analytical model is a single-chamber single-zone model, similar to that developed by Krieger and Borman [17], that utilizes the main-chamber pressure history and an estimation of the trapped mass in the combustion chamber to calculate the heat release history of the fuel. Thermodynamic equilibrium and ideal gas behavior are assumed at all times. At any instant, the control volume contains a homogeneous mixture of air and combustion products. Thus the fuel is assumed to be introduced into the chamber at the same rate as it is burned, and the combustion process is uniformly distributed. The property calculations are based on the temperature, pressure, and equivalence ratio. For negligible dissociation of the products the rate of heat release Q_{HR} may be approximated by

$$Q_{HR} = \frac{dm_b}{dt}(\Delta H_0) \tag{1}$$

where dm_b/dt is the fuel mass-burned rate and ΔH_0 is the lower heating value of the fuel. A quantity used in this study is the specific mass-burned rate of fuel per unit crank-angle degree θ. This is defined as

$$\frac{1}{m_f}\frac{dm_b}{d\theta} = \frac{1}{m_f}\left(\frac{dm_b}{dt}\right)\left(\frac{60}{N}\right) \tag{2}$$

where m_f is the total mass of fuel injected per cycle and N is the engine speed in rpm.

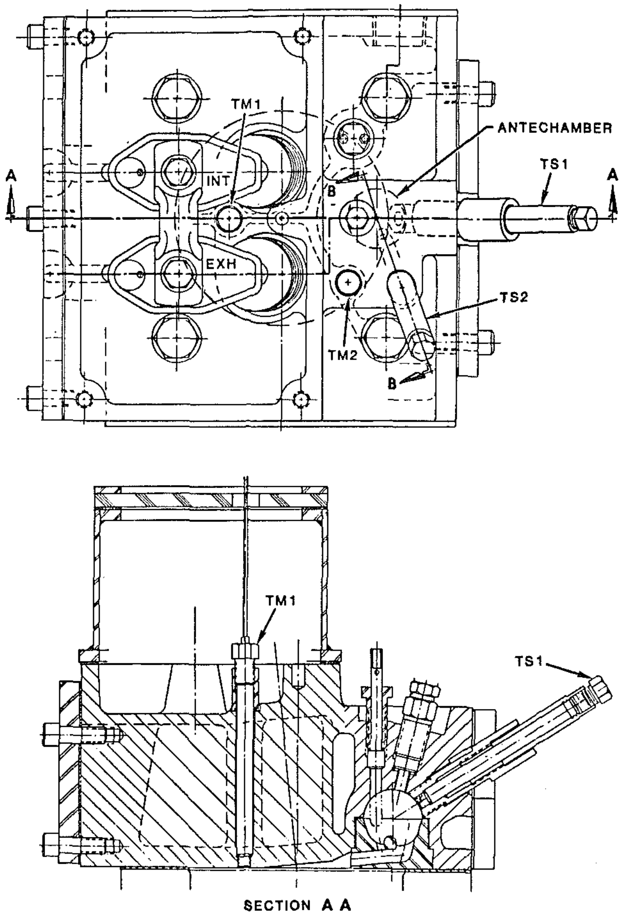


Fig. 1 Locations of heat-flux probes on the cylinder head

The instantaneous surface heat transfer $Q_w(t)$ is described by an equation of the form

$$Q_w(t) = \sum h_g(t) A_i(t) [T_g(t) - T_{w,i}] \tag{3}$$

where $T_g(t)$ and $h_g(t)$ are the gas temperature and convective heat transfer coefficients, respectively, at time t; A_i represents the surface areas of the various regions of the combustion chamber; and $T_{w,i}$ represents the corresponding time-averaged temperatures of these regions.

The heat transfer coefficient is given by

$$h_g = c(h_g)_w \tag{4}$$

where $(h_g)_w$ is the Woschni convective heat transfer coefficient [18] and c is a constant multiplier whose value is adjusted at each engine condition such that the computed final fuel-burned fraction is equal to a specified value.

Test Conditions. The engine tests were performed with diesel fuel No. 2 for a range of engine speed (1000–2500 rpm), fuel-air ratio (0.000 to 0.040), and start of combustion (0 to 10 deg ATDC). Unless otherwise specified, the start of combustion was fixed at TDC. The intake air was maintained at ambient pressure and at a temperature of 35°C. In addition, for all tests the coolant temperature (82°C), oil temperature (82°C), and fuel temperature at the pump entry (35°C) were maintained at fixed values.

Results and Discussion

Heat Release Characteristics. Figure 2 shows measured main-chamber pressure histories and corresponding calculated

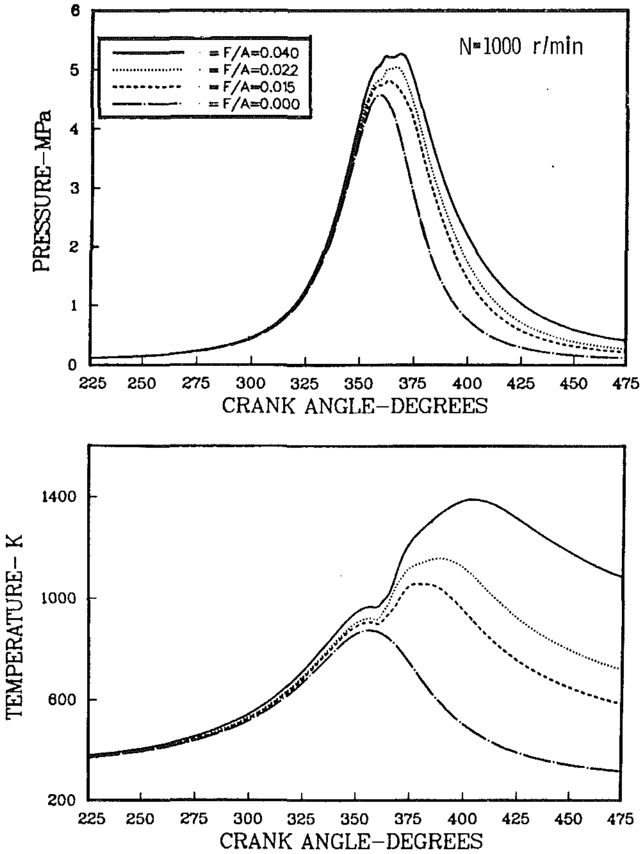


Fig. 2 Variations of pressure and temperature histories with fuel-air ratio

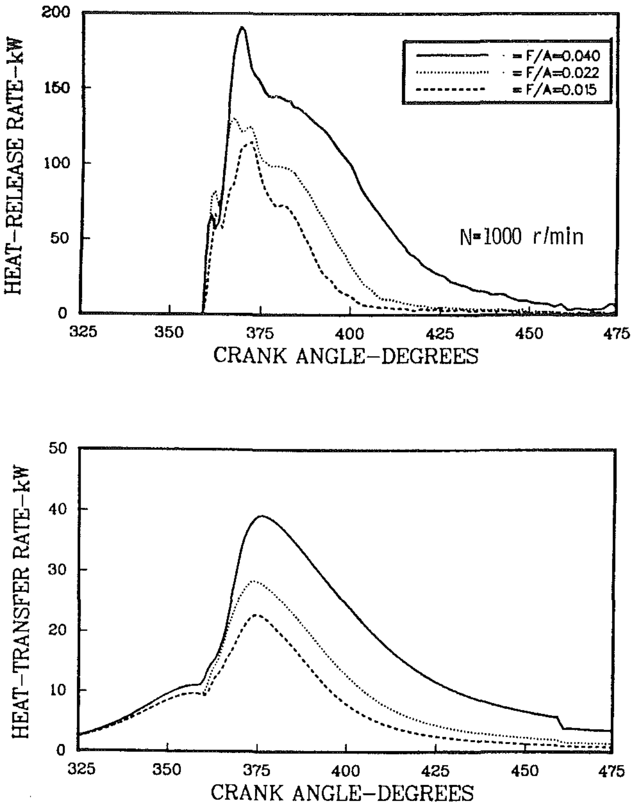


Fig. 3 The influence of fuel-air ratio on the rates of heat release and heat transfer

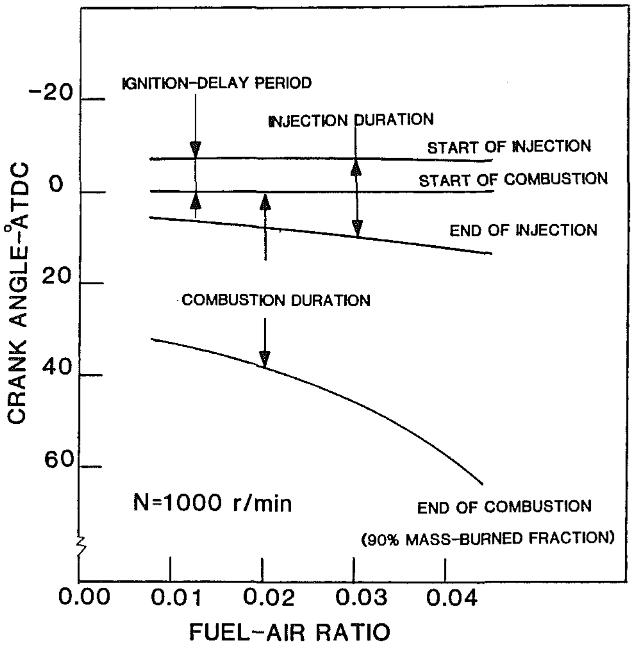


Fig. 4 The variations of injection duration, ignition delay, and combustion duration with fuel-air ratio

average gas-temperature histories for various fuel-air ratios at 1000 rpm and TDC combustion timing. Increasing the fuel-air ratio resulted not only in an increase in peak pressure, but also a broadening of the pressure history. The computed average gas temperatures increased significantly with increasing fuel-air ratio during the period of combustion because of the resultant increase in the amount of fuel burned per cycle. It must be noted that the computed gas temperatures are comparatively low because they represented spatially averaged quantities due to the assumption of spatially uniform combustion.

The influence of fuel-air ratio on the heat release history of the engine is demonstrated in Fig. 3. The upper graph shows heat release-rate histories for various fuel-air ratios, and the lower graph shows the corresponding computed combustion-chamber heat transfer-rate histories. In general, increasing the fuel-air ratio (by increasing the amount of fuel injected per cycle) increased the instantaneous heat release rate, as expected, because of the greater amount of fuel that burns during each time increment. The exception was at the beginning of heat release, where the initial rate of heat release appears to be essentially independent of fuel-air ratio. This has also been observed by Kort et al. [13]. At the bottom of Fig. 3 the observed increase in the calculated heat transfer rate with increasing fuel-air ratio is caused primarily by the higher gas temperatures. Although radiation was not accounted for in the present analysis, an increase in combustion-chamber heat transfer rate with load is also caused by the augmented radiation heat transfer due to the resulting higher concentration of soot in the combustion chamber [19, 20].

Figure 4 shows the variations of injection duration and combustion duration with fuel-air ratio. The injection duration was obtained from poppet-lift profiles of the instrumented injector. The start of combustion was detected by a luminosity sensor. The end of combustion was defined as the crank angle corresponding to a mass-burned fraction of 0.90.[1] Increasing the fuel-air ratio, or more precisely the fuel per cycle, resulted in an approximately linear increase of the period of injection, whereas it caused a much more drastic increase in the duration of combustion.

---

[1] The beginning and end of combustion, and consequently the duration of combustion, are parameters that require arbitrary definitions.

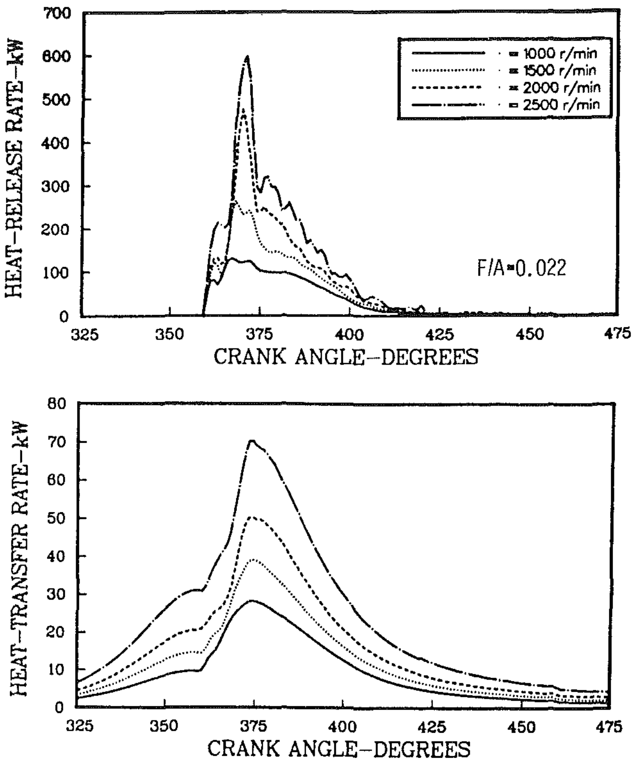


Fig. 5 The influence of engine speed on the rates of heat release and heat transfer

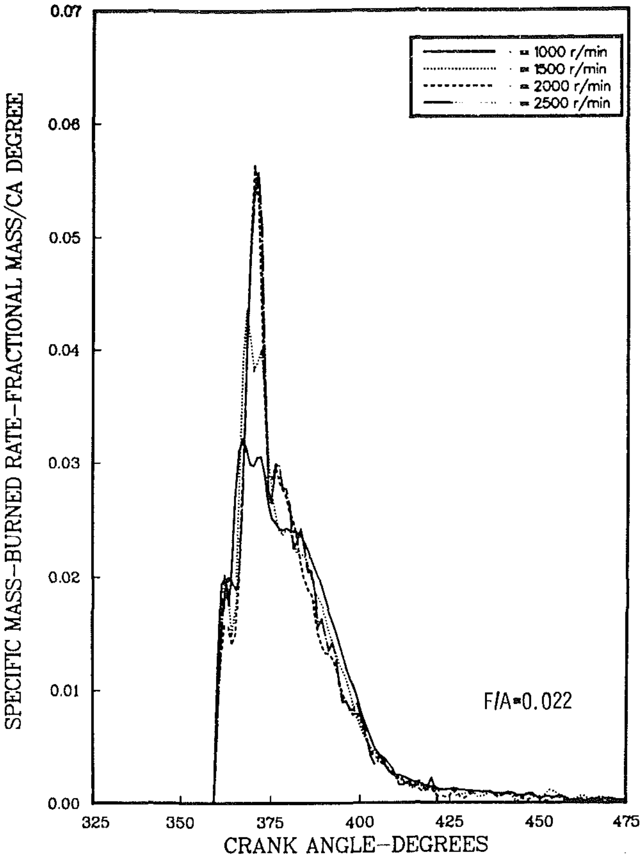


Fig. 6 The influence of engine speed on the specific mass-burned-rate histories of the fuel

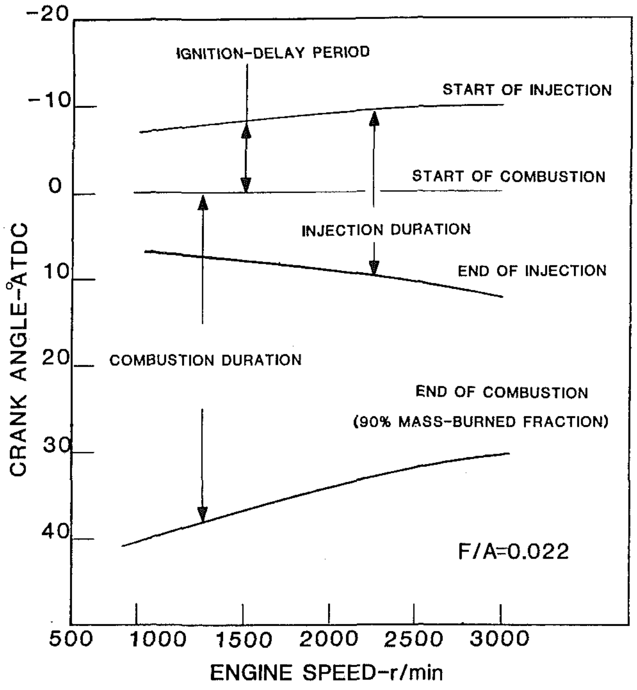


Fig. 7 The variations of injection duration, ignition delay, and combustion duration with engine speed

The influences of engine speed on the heat release-rate histories and calculated heat transfer-rate histories of the engine for a fuel-air ratio of 0.022 at various engine speeds and combustion timing at TDC are shown in Fig. 5. Increasing the engine speed generally increased the heat release rate because of the higher mixing rate achieved and increased the heat transfer rate primarily because of the resulting higher flow velocities, which augmented the convective heat transfer coefficient. The heat transfer rates may also be enhanced by the higher gas-temperature levels achieved at higher engine speeds.

Perhaps a more interesting presentation of the influence of engine speed on the heat release characteristics of the engine is provided by Fig. 6, which is a plot of the specific mass-burned-rate histories for various engine speeds. From Fig. 6, the initial stages and final stages of the specific mass-burned-rate history appear to be relatively insensitive to engine speed, whereas in the central stage, the peak values of the specific mass-burned rate increase with increasing engine speed. This suggests that during the initial and final stages of combustion the burning rate, and consequently the mixing rate, increased in proportion to engine speed. In the central region, however, the increase in magnitude of the specific mass-burned rate with engine speed suggests that higher speeds cause proportionally better dispersion of the fuel-rich regions through better mixing.

It appears in Fig. 6, as it did in Fig. 5, that the duration of combustion expressed in crank-angle degrees was not significantly influenced by engine speed. However, if one arbitrarily defines the end of combustion as the crank angle at which 90 percent of the fuel is burned (as in Fig. 4), then, as shown in Fig. 7, the duration of combustion in crank-angle degrees appears to decrease somewhat with increasing engine speed. This seeming anomaly is, in fact, consistent with Fig. 6, close inspection of which reveals that the high peak values at high speeds are balanced by subsequent lower values of specific mass-burned rate.

The influence of combustion timing on the heat release-rate history and calculated heat transfer-rate history of the engine is shown in Fig. 8 for an engine speed of 1000 rpm and a fuel-air ratio of 0.022. Retarding the combustion timing delayed the combustion schedule of the fuel, as expected. Retarding the combustion timing from 0 to 5 deg ATDC

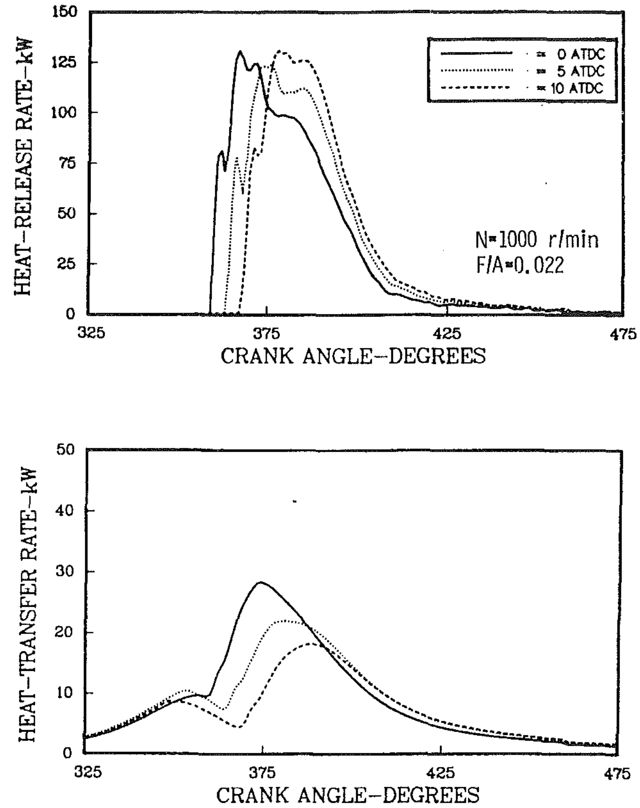


Fig. 8 The influence of combustion timing on the rates of heat release and heat transfer

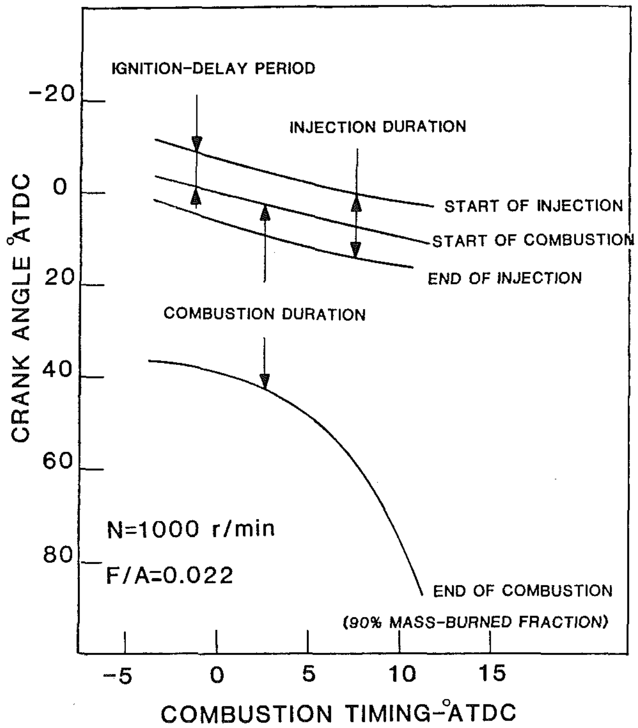


Fig. 9 The variations of injection duration, ignition delay, and combustion duration with combustion timing

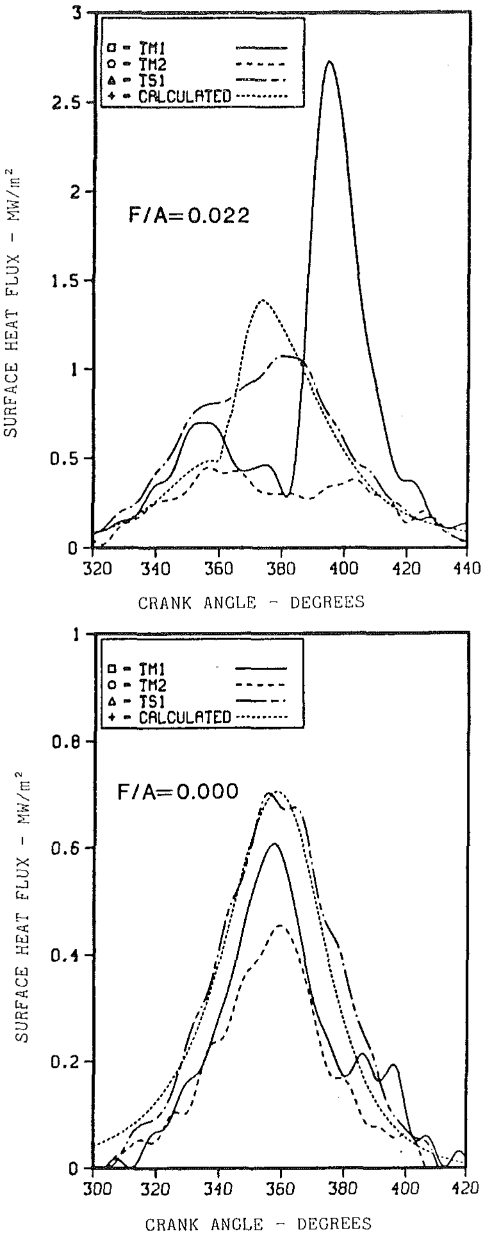


Fig. 10 Measured local surface heat-flux histories and corresponding computed area-averaged heat-flux histories for two fuel-air ratios

slightly decreased the peak value of mass-burned rate, whereas a further retarding of the combustion timing to 10 deg ATDC restored the original peak value of the mass-burned rate. In contrast, the peak heat transfer rate decreased with combustion-timing retard because of the lower pressure and temperature levels during the combustion period.

The variations of the measured injection duration and calculated combustion duration (to 90 percent mass burned) with changes in combustion timing are shown in Fig. 9. The combustion duration appears to become progressively longer when retarding combustion timing beyond TDC. This is attributed to lower pressure and temperature levels, as well as to the reduction in mixing rate because of lower flow velocities later after TDC. On the other hand, both the injection duration and ignition-delay period appear not to be influenced significantly by combustion timing within the range of combustion timings examined. This is probably because, with the ignition delay period so close to TDC at all timings investigated, cylinder-gas temperature was insensitive to injection timing.

Measured Heat Transfer Characteristics. Figure 10 shows the measured local heat-flux histories at locations TM1 and TM2 in the main chamber and at location TS1 in the antechamber for fuel-air ratios of 0.022 and of 0.000 (motored condition). These data were obtained for an engine speed of 1000 rpm, and for a combustion timing of TDC in the fired case. Also shown in Fig. 10 are the corresponding area-averaged heat-flux histories obtained from the heat release analysis.

For the motored condition (lower plot), the measured heat-flux histories peaked near TDC (360 deg). The heat fluxes are highest at location TS1, because of the high swirl flows present in the antechamber, followed by the heat fluxes at TM1 and TM2. The computed area-averaged heat-flux history appears to be in reasonably good agreement (overall shape and magnitude) with the local heat-flux measurements.

For the fired condition (upper plot), the heat-flux histories appear to be double peaked. The first peak, which in general is lower in magnitude, occurs near TDC and is due to compression. In agreement with the motored data, the peak heat flux due to compression was highest at TS1, followed by TM1 and TM2. The second peak, which occurs later in the cycle and is strongly dependent on location, is caused by combustion. The peak heat flux due to combustion was highest at TM1, followed by TS1 and TM2. The location of TM1 is between the valves and in line with the throat of the antechamber; therefore it experiences the convective action of the hot combustion gases exiting the antechamber during the early stages of the expansion stroke.

The comparison of the local heat-flux measurements with the corresponding calculated area-averaged values clearly demonstrates that for fired conditions, because the local heat-flux measurements are strongly influenced by combustion which is spatially nonuniform, a meaningful comparison of local heat-flux measurements with computed area-averaged heat fluxes is not possible.

Some insight concerning the relative magnitudes of the heat release in the two chambers may be gained by considering Fig. 11. This shows the variations of the measured local heat-flux histories at locations TS1 and TM2 with changes in the combustion timing for an engine speed of 1000 rpm and a fuel-air ratio of 0.022. At the TS1 location in the antechamber, retarding the combustion timing ATDC significantly decreased the peak heat flux due to combustion. In contrast, at the TM2 location in the main chamber, the peak heat flux due to combustion increased with retarded combustion timing. These observations indicate that as the combustion timing is retarded ATDC, a proportionally larger amount of unburned fuel enters the main chamber, where it mixes and burns.

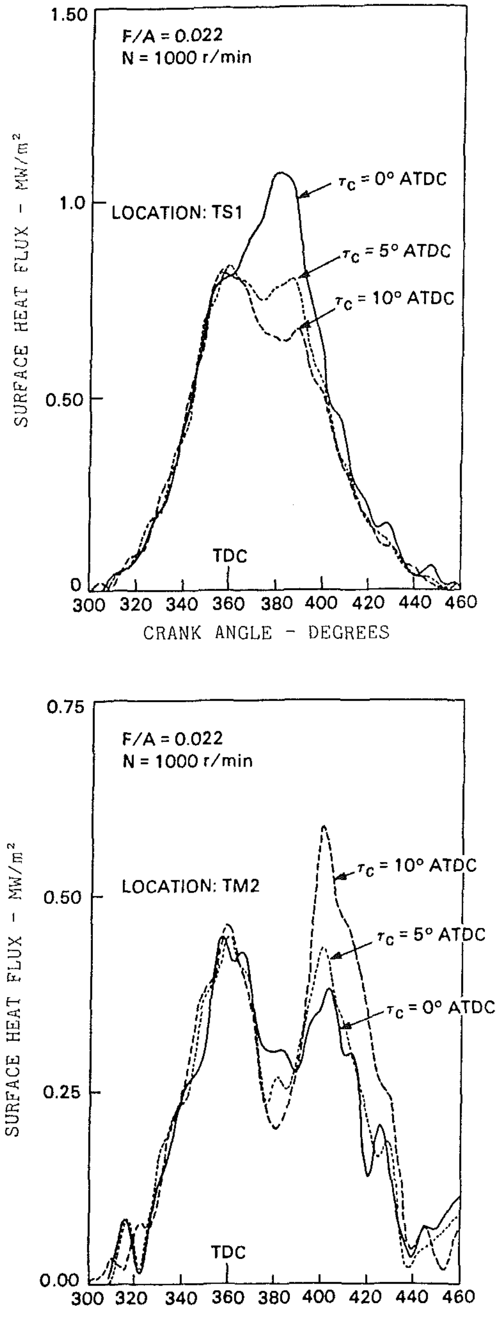


Fig. 11 Variations of the heat-flux histories at locations TS1 and TM1 with combustion timing

Closure

In the present study, a heat release analysis and local surface heat-flux measurements were employed to examine in some detail the combustion process in a divided-chamber diesel engine. The heat release model was a single-zone single-chamber thermodynamic model that utilized measured main-chamber pressure histories and the amount of trapped mass in the chamber as inputs. Although the model is incapable of distinguishing combustion in the individual chambers of the divided-chamber diesel engine, it provided useful quantitative information with regard to the overall characteristics of the combustion process of the engine. Utilizing this analytical model, the influences of engine speed, fuel-air ratio, and combustion timing on ignition delay, combustion duration, and the fuel mass-burning rate were examined.

The heat-flux measurements in the main chamber and antechamber of the engine provided additional insight on the fluid motion and combustion in the two chambers. For example, the relative magnitude of the heat flux in the antechamber during motored operation of the engine indicated the strong air swirl generated in the antechamber of the engine. Additionally, the observed variations of the heat-flux histories in the two chambers with changes in the combustion timing provided information on how the fuel-mass-burned distribution in the two chambers varied with combustion timing.

In closing one should observe that more detailed information may be obtained by employing one of the recently developed two-chamber heat-release models [6, 7]. Unfortunately, however, these models appear to be very sensitive to the accuracy of the measured pressure difference between the two chambers, since this difference controls the mass flow between the two chambers. Unlike the engines used to validate these models, the engine used in this study has a relatively large-area antechamber throat, and consequently the pressure differences between the two chambers are relatively small. This makes any heat release histories obtained by the two-chamber model suspect.

Conclusions

This study was conducted on an experimental single-cylinder divided-chamber diesel engine. The engine parameters examined were engine speed (1000 to 2400 rpm), fuel–air ratio (0.000 to 0.040), and start of combustion (0 to 10 deg ATDC). Based on the results of this investigation, the following conclusions were reached:

1 Increasing the fuel–air ratio increased the duration of combustion (0 to 90 percent mass-burned fraction) and, in general, the instantaneous fuel-mass-burned rate (which is linearly related to heat release rate). The exception was the initial mass-burned rate, which was found to be essentially independent of fuel–air ratio.

2 Increasing the engine speed increased the fuel-mass-burned rate as expected. Near the beginning and end of combustion the mass-burned rate increased in direct proportion to engine speed. In the central part of the combustion duration, the mass-burned rate was found to increase at a higher rate than engine speed. As a consequence of this, the combustion duration to 90 percent completion, expressed in crank-angle degrees, decreased somewhat with increasing engine speed.

3 Retarding the combustion timing (start of combustion) after TDC significantly increased the combustion duration to 90 percent mass burned.

4 For motored conditions, comparison of computed area-averaged heat fluxes and measured local heat fluxes showed good agreement. For fired conditions, because local heat fluxes are strongly influenced by combustion, which is spatially nonuniform, a meaningful comparison of local heat-flux measurements with area-averaged heat fluxes computed from a heat release model was not possible.

5 Heat-flux measurements in the main chamber and antechamber of the engine suggested that as the combustion timing was retarded ATDC, progressively more fuel burned in the main chamber.

Acknowledgments

The author gratefully acknowledges the valuable contribution of R. M. Cole in the acquisition of the pressure and surface-temperature measurements.

References

1 Patterson, D. J., and Henein, N. A., *Emissions From Combustion Engines and Their Control*, Ann Arbor Science Publishers, Inc., Ann Arbor, MI., 1972, pp. 231–304.

2 Heywood, J. B., "Engine Combustion Modeling – an Overview," in: *Combustion Modeling in Reciprocating Engines*, J. N. Mattavi and C. A. Amann, eds., Plenum Press, New York, 1980, pp. 1–35.

3 Mattavi, J. N., and Amann, C. A., eds., *Combustion Modeling in Reciprocating Engines*, Plenum Press, New York, 1980.

4 Bowden, C. M., Samaga, B. S., and Lyn, W. T., "Rate of Heat Release in High-Speed Indirect Injection Diesel Engines," *Proc. Instn. Mech. Engrs*, Vol. 184, 1969/70, pp. 122–129.

5 Nakajima, K., Kajiya, S., and Nagao, F., "An Experimental Investigation of the Air Swirl Motion and Combustion in the Swirl Chamber of Diesel Engines," *XII FISITA Congress*, Paper No. 1-02, 1968.

6 Anisits, F., and Zapf, H., "Auswertverfahren der Druckverlaufe und Elektronische Berechnung des Verbrennungs-Verlaufs in Dieselmotoren mit Unterteilten Brennraumen," *Motortechnische Zeitschrift*, Vol. 32, 1971, pp. 447–451.

7 Kamel, M., "Thermodynamic Analysis of Indirect Injection Diesel Engine Operation," Ph.D. Thesis, Imperial College of Science and Technology, University of London, 1977.

8 Watson, N., and Kamel, M., "Thermodynamic Efficiency Evaluation of an Indirect Injection Diesel," *SAE Transactions*, Vol. 88, 1979, pp. 158–171.

9 Terada, K., "Zur Ermittlung der Heizverlaufe in einem Nebenkammer-Dieselmotor," *Motortechnische Zeitschrift*, Vol. 40, 1979, pp. 237–243.

10 Leone, G., Danieli, G. A., and Canonaco, M., "Modelling the Performance of a Prechamber Diesel Engine," *ISATA 79 Proceedings*, Vol. 1, 1979, pp. 439–454.

11 Terada, K., "Ermittlung der Gastemperaturen in beiden Brennraumen eines Wirbelkammer-Dieselmotors," *Motortechnische Zeitschrift*, Vol. 42, 1981, pp. 19–21.

12 Mansouri, S. H., Heywood, J. B., and Radhakrishnan, K., "Divided-Chamber Diesel Engine, Part 1: A Cycle-Simulation Which Predicts Performance and Emissions," *SAE Transactions*, Vol. 91, 1982, pp. 1101–1132.

13 Kort, R. T., Mansouri, S. H., Heywood, J. B., and Ekchian, A., "Divided-Chamber Diesel Engine, Part II: Experimental Validation of a Predictive Cycle-Simulation and Heat Release Analysis," *SAE Transactions*, Vol. 91, 1982, pp. 1133–1147.

14 Alkidas, A. C., and Cole, R. M., "Transient Heat-Flux Measurements in a Divided-Chamber Diesel Engine," ASME Paper No. 84-HT-44.

15 Overbye, V. D., Bennethum, J. E., Uyehara, O. A., and Myers, P. S., "Unsteady Heat Transfer in Engines," *SAE Transactions*, Vol. 69, 1961, pp. 461–494.

16 Alkidas, A. C., "Heat Transfer Characteristics of a Spark-Ignition Engine," ASME *Journal of Heat Transfer*, Vol. 102, 1980, pp. 189–193.

17 Krieger, R. B., and Borman, G. L., "The Computation of Apparent Heat Release for Internal Combustion Engines," ASME Paper No. 66-WA/DGP-4.

18 Woschni, G., "A Universally Applicable Equation for the Instantaneous Heat Transfer Coefficient in the Internal Combustion Engine," *SAE Transactions*, Vol. 76, 1967, pp. 3065–3083.

19 Matsui, Y., Kamimoto, T., and Matsuoka, S., "A Study on the Time and Space Resolved Measurement of Flame Temperature and Soot Concentration in a D.I. Diesel Engine by the Two-Color Method," *SAE Transactions*, Vol. 88, 1979, pp. 1808–1822.

20 Fujiwara, Y., Fukazawa, S., Tosaka, S., and Murayama, T., "Formation of Soot Particulates in the Combustion Chamber of a Precombustion Chamber Type Diesel Engine," SAE Paper No. 840417, 1984.

Adapting the Heat and Mass Transfer Analogy to Model Performance of Automotive Catalytic Converters

J. R. Mondt
Engine Research Department,
General Motors Research Laboratories,
Warren, MI 48090-9055

Using theory for mass transfer of a diffusing specie with very low concentration, rate equations and a mass-balance equation can be combined to derive a differential equation for mass transfer in an automotive catalytic converter. A closed-form solution to this equation shows conversion efficiency to be a function of the dimensionless size of the converter, or the number of transfer units, Ntu_m. This mass-transfer-limited analysis does not include catalyst kinetics; hence it is limited to fully warm, fresh catalyst performance. However, a technique is developed to model lightoff of a catalytic converter by combining convective heat transfer when warming up with mass-transfer-limited conversion when fully warm. Realistic assessment of the merit of a catalytic converter must also include the influence of size and shape on flow pressure drop. Accordingly the size and shape of square-cell monoliths and packed-sphere bead beds are correlated with both conversion performance and pressure drop. Applications of these correlations are shown to compare realistically the size versus performance characteristic of monoliths with that for bead beds, in spite of drastically different flow patterns for the two converters. Model predictions are confirmed by engine-dynamometer and vehicle test results.

Introduction

Since the mid-1960s, spark-ignition engines and emission-control systems have changed markedly to meet specific emissions requirements set forth by Federal regulations, as shown in Fig. 1. Stringent emission levels legislated for 1975 were met by installing catalytic converters on all GM passenger cars [1].

Although catalyst technology was well developed for the chemical and petrochemical industries, adapting this technology to automobiles posed several new challenges. Size, shape, and mass limitations for a practical vehicle installation posed new challenges for the designed catalytic converters. Rapidly varying inlet conditions to the converter and the requirement that a catalyst must be heated before it is effective influenced the design of automotive catalytic converters.

Modeling the performance of an automotive catalytic converter [2] is further complicated by the many alternative geometries [3, 4] that could be used as substrates to support catalyst metals. These substrates can be broadly categorized as either monoliths or bead beds.

"Monolith" identifies a substrate which contains an array of parallel continuous flow channels of square, triangular, or hexagonal shape. Regardless of flow-channel shape, the small passages of a typical automotive monolithic converter cause the flow to be laminar, so the velocity profiles are established by viscous shear forces. As shown in Fig. 2, these velocity profiles are developed in the entrance region of a monolith, where thin boundary layers can significantly augment heat and mass transfer. Augmented heat or mass transfer is accompanied by augmented pressure drop required to develop the velocity profiles.

A "bead bed" identifies an aggregation of small pellets, either spheres or cylindrical extrudates, retained by a perforated grid at both the inlet and outlet faces. Flow through a bead bed follows a tortuous path, shown in Fig. 3, where eddy

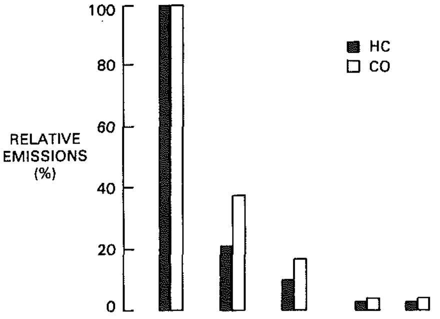
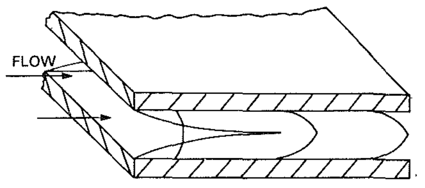
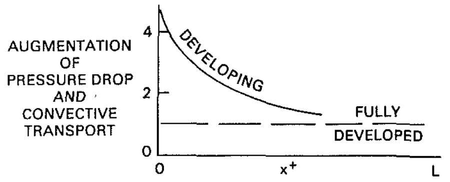


Fig. 1 **Legislated tailpipe emission levels for unburned hydrocarbons (HC) and carbon monoxide (CO) in the USA**

Contributed by the Heat Transfer Division and the Internal Combustion Engines Division and presented at the ASME Winter Annual Meeting, Miami Beach, Florida, November 17–21, 1985. Manuscript received at ASME Headquarters May 19, 1986. Paper No. 85-WA/HT-33.



a. Developing Boundary Layer



b. Augmented Pressure Drop and Convective Transport

Fig. 2 Developing flow in a monolith

and momentum forces make boundary layers very thin. These thin boundary layers result in high heat transfer and mass transfer but with very high flow pressure drop.

To compare performance of monoliths with that for bead beds, mass-transfer-limited models were developed to calculate tradeoffs of conversion efficiency, pressure drop, and overall size and shape. Use of mass-transfer-limited models simplifies modeling by excluding any limitation to conversion performance because of kinetics for the chemical reactions [2–5]. By excluding kinetics, these models calculate an upper limit for attainable conversion efficiency, or performance of a fresh catalyst. Other models have been reported to address aging and poisoning of catalysts [6].

For modeling catalyst activity on a vehicle, the temperature level establishes three regimes:

1 Warmup to initiate catalyst lightoff
2 Continued heating until all the catalyst is above the lightoff temperature
3 Fully warm operation

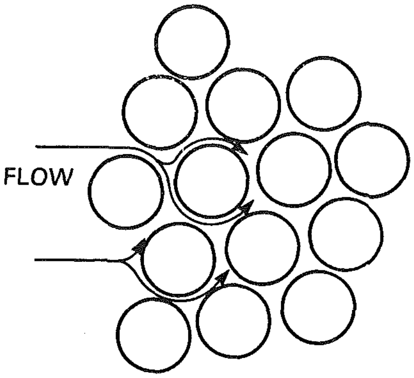


Fig. 3 Bead-bed flow schematic

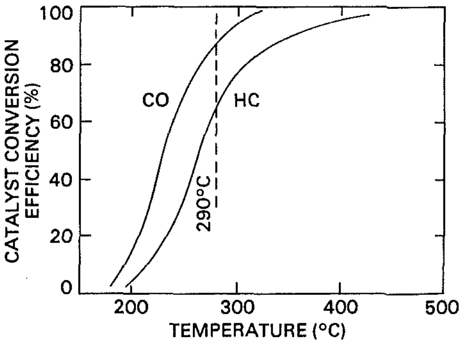


Fig. 4 Lightoff for a typical catalyst

These three temperature regimes are experienced by a catalytic converter for every cold start of an automobile. Immediately following startup of an engine, exhaust gases passing through the converter begin heating the substrate by forced convection. Some time elapses after the engine is

Nomenclature

A = surface area exposed to gas flow
A_{fr} = frontal area
A_g = augmentation factor for developing boundary layer
c = specific heat of solid
c_p = specific heat of fluid
C = constant
D = Fick's diffusion coefficient
f = friction factor
g = variable mass-transfer coefficient
$\bar{g}$ = mean mass-transfer coefficient over a prescribed flow length
G = flow mass velocity
g_c = gravitational constant
h = heat transfer coefficient for convection
j = convection factor for a bead bed
k_{ex} = thermal conductivity of exhaust gas
K_c = contraction coefficient
K_e = expansion coefficient
L = flow length
m = mass concentration of specie in a gas stream
$\dot{m}$ = mass-transfer flux by convection
$\dot{M}$ = mass flow rate of specie in a gas stream
n = exponent constant
Ntu_m = number of mass-transfer units
Nu = Nusselt number
P = pressure
Re = Reynolds number
Sc = Schmidt number
Sh = Sherwood number
St = Stanton number
$\dot{q}$ = heat transfer flux by convection
r_h = hydraulic radius of substrate matrix
T = temperature
V = exhaust gas velocity
Vol = volume of substrate matrix
$\dot{w}$ = exhaust-gas mass flow rate
X = flow direction
x^+ = dimensionless developing-flow parameter
y = direction normal to the wall
α = porosity, void fraction of substrate matrix
β = area density, surface area per volume of substrate matrix
η = conversion efficiency
μ = dynamic viscosity
ξ = dummy variable
ρ = fluid density
τ_{50} = time delay until conversion efficiency = 50 percent

Subscripts and Symbols

a = ambient
Δ = difference
j = specie
m = mass transfer
h = heat transfer
M = monolith
B = bead bed
ex = exhaust
diff = diffusion
1, 2, 3 = different constants
s = solid
f = flow friction
g = exhaust gas
0 = wall
$\overset{\Delta}{=}$ = defining
∞ = free stream

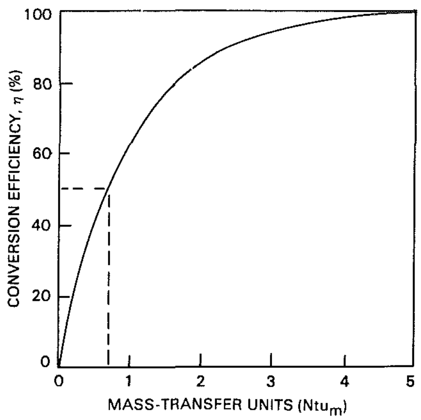


Fig. 5 **Mass-transfer-limited conversion efficiency**

started until the substrate and catalyst are hot enough to promote chemical reactions. As shown in Fig. 4, significant catalytic activity can be expected for a typical, fresh oxidizing catalyst only after the temperature reaches approximately 290°C. This threshold temperature for a converter to become active is termed "lightoff" temperature.

As soon as the leading edge of the substrate attains a temperature above lightoff, the second temperature regime begins and catalyst activity starts. Exothermic reactions that are promoted by any fraction of a catalytic converter that is above lightoff temperature can augment substrate heating, although the contribution to substrate heating from chemical reactions is small for normal engine operation. The second temperature regime terminates when the lightoff temperature wave propagates to the exit face of the substrate.

The final temperature regime occurs most of the operating time, when all of the catalyst has been heated to a temperature higher than lightoff. Then for fresh catalysts, convective mass transport limits conversion efficiency. For mass-transfer-limited conversion, adequate quantities of chemical species must be available to complete the reactions, such as adequate oxygen in the exhaust stream to oxidize unburned hydrocarbons and carbon monoxide. Thus, mass-transfer-limited theory has been successfully applied only to oxidizing converters.

Heat and Mass-Transfer Analogy

Exhaust gas flowing into an automotive catalytic converter normally contains reactants at very dilute concentrations of less than 0.1 percent by volume. These low concentrations permit simplified solutions for diffusion mass transfer based on Fick's Law [7], which is analogous to Fourier's Law for conduction heat transfer [8].

Heat Transfer:

$$\dot q'' = -k\left(\frac{dT}{dy}\right)_{y=0} \tag{1}$$

Mass Transfer:

$$\dot m_j'' = -\rho D_j\left(\frac{dm_j}{dy}\right)_{y=0} \tag{2}$$

For either convective transfer process, rate equations define the heat transfer coefficient h and mass-transfer coefficient g as follows:

Heat Transfer:

$$\dot q'' = h(T_g - T_0) \tag{3}$$

Mass Transfer:

$$\dot m_j'' = g_j(m_{j,\infty} - m_{j,0}) \tag{4}$$

The convective mass-transfer rate equation (4) can be combined with a constituent balance

$$\dot M_{j,1} - \dot M_{j,2} = dm_j \dot w \tag{5}$$

where

$$\dot M_{j,1} = m_{j,1}\dot w \tag{6}$$

$$\dot M_{j,2} = \dot w\left(m_{j,1} + \frac{\partial m_j}{\partial x}\,dx\right) \tag{7}$$

To derive an equation for mass transfer from the exhaust gas to the surface of a substrate in a catalytic converter

$$dm_j = -\frac{\bar g}{\dot w}(m_{j,\infty} - m_{j,0})dA \tag{8}$$

Rearranging equation (8) and integrating:

$$\int_{m_{j,i}}^{m_{j,2}} \frac{dm_j}{(m_j - m_{j,0})} = -\int_0^A \frac{\bar g}{\dot w}\,dA \tag{9}$$

$$\frac{m_{j,1} - m_{j,2}}{m_{j,1} - m_{j,0}} = e^{-\bar g A/\dot w} \tag{10}$$

Defining the number of mass-transfer units by

$$\mathrm{Ntu}_m \triangleq \frac{\bar g A}{\dot w} \tag{11}$$

and the conversion efficiency by

$$\eta \triangleq \frac{m_{j,1} - m_{j,2}}{m_{j,1} - m_{j,0}} \tag{12}$$

it follows that

$$\eta = 1 - e^{-\mathrm{Ntu}_m} \tag{13}$$

In equation (13) the parameter Ntu_m indicates the dimensionless size of a catalytic converter. This parameter has been used to size catalytic converters for chemical plants [7] and compact heat exchangers for many heat transfer processes [9].

Conversion efficiency for a warmed-up converter that is mass-transfer limited, from equation (13), is graphed versus Ntu_m in Fig. 5, showing efficiency increasing at a decreasing rate as Ntu_m increases. Regardless of substrate type or geometry – monolith, bead bed, packed screen, or porous mesh – a specific Ntu_m defines a specific conversion efficiency.

Monolith Catalyst Supports

Modeling of convective transport in monoliths is logically subdivided according to whether the flow is fully developed laminar flow or developing flow.

Fully Developed Laminar Flow. For fully developed laminar flow, velocity, temperature, and reactant profiles remain invariant with flow length. Closed-form solutions describe either the heat transfer or mass-transfer convective processes.

Heat Transfer [*10*]:

$$\mathrm{Nu} = C_1 = \frac{h(4r_h)}{k_{ex}} \tag{14}$$

Mass Transfer [*7*]:

$$\mathrm{Sh} = C_1 = \frac{\bar g(4r_h)}{\rho D} \tag{15}$$

In equation (14), Nu is the Nusselt number, which contains the heat transfer coefficient h. In equation (15), Sh is the Sherwood number, which contains the mass-transfer coefficient $\bar g$. In both equation (14) and equation (15), the constant C has the same value, depending only on the shape of the flow passage.

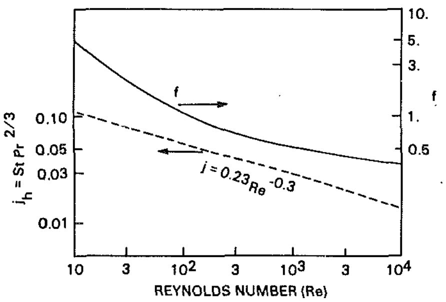


Fig. 6 Correlation of experimental friction factor *f* and Colburn factor *j* for an infinite randomly stacked sphere bed [9] (porosity varies from 0.37 to 0.39)

By combining equation (11) with equation (15), the Ntu_m for a monolith can be derived using $A = \beta \bullet \mathrm{Vol}$

$$(\mathrm{Ntu}_m)_M = \left(\frac{\mathrm{Vol}}{\dot{w}}\right)\beta\,\frac{[\rho D \mathrm{Sh}]}{4r_h} \tag{16}$$

Equation (16) shows that conversion efficiency is directly proportional to the volume of the monolith but independent of the arrangement of that volume into frontal area and flow length. This allows the designer flexibility to package monolith volume into a variety of shapes with the same fully warm conversion performance, as long as flow is distributed uniformly across the frontal area of the monolith.

Developing Laminar Flow. A thinner boundary layer in the entrance region of monoliths, shown schematically in Fig. 2, augments both convective heat transfer and mass transfer. Closed-form solutions have been derived to correlate heat transfer and mass-transfer augmentation in the entrance region of flow passages of a variety of cross sections using x_h^+ and x_m^+:

Heat Transfer [8]:

$$x_h^+ = \frac{L/4r_h}{\mathrm{RePr}} \tag{17}$$

Mass Transfer [7]:

$$x_m^+ = \frac{L/4r_h}{\mathrm{ReSc}} \tag{18}$$

In equations (17) and (18), Prandtl and Schmidt numbers are explicit in the derivation; thus there are no limitations for using these correlations for either heat transfer or mass transfer. Similarly in the entrance region, augmented pressure drop has been correlated using the parameter x_f^+ [10]:

$$x_f^+ = \frac{L/4r_h}{\mathrm{Re}} \tag{19}$$

All monoliths have developing flow in an entrance region. The influence of developing-flow augmentation on conversion efficiency depends on the relative amount of flow length in the developing region. Wendland [11] reported theory and experimental tests for monoliths segmented to exploit augmented performance by using several entrance regions in one monolith converter.

Wendland based conversion efficiency on a variable mass-transfer coefficient characterized by a variable Sherwood number $\mathrm{Sh}(x^+)$. If a mean Sherwood is defined as

$$\overline{\mathrm{Sh}} \triangleq \frac{1}{x_m^+}\int_0^{x_m^+} \mathrm{Sh}(\xi)\,d\xi \tag{20}$$

the conversion efficiency defined by Wendland simplifies to the following

$$\eta = 1 - e^{-\int_0^{x_m^+} \mathrm{Sh}(\xi)\,d\xi} = 1 - e^{-\mathrm{Ntu}_m} \tag{21}$$

Note that equation (21) is identical to equation (13).

Bead-Bed Catalyst Supports

A bed of packed pellets, or a bead bed, provides a tortuous flow path which does not allow fully developed velocity, temperature, or species profiles to be established. Boundary-layer development starts from the stagnation region on an individual pellet, but adjacent pellets interrupt development of any boundary layer. Since theoretical solutions have not been obtained for convection processes in a bead bed, heat transfer correlations have been established from experimental measurements. Convective heat transfer can be correlated by the Colburn factor (j_h), and convective mass transfer can be correlated by j_m:

Heat Transfer:

$$j_h = (\mathrm{St})(\mathrm{Pr})^{2/3} = C_2\mathrm{Re}^n \tag{22}$$

Mass Transfer:

$$j_m = (\bar{g}/G)(\mathrm{Sc})^{2/3} = C_2\mathrm{Re}^n \tag{23}$$

For convective heat transfer, the Prandtl number in the Colburn factor j_h accounts for the ratio of momentum diffusivity to thermal diffusivity. Similarly, the Schmidt number in the mass-transfer factor j_m accounts for the ratio of momentum diffusivity to mass diffusivity. Thus the use of experimental heat transfer correlations for calculating mass-transfer coefficients requires the following two substitutions and one limiting condition [12]:

1 substitution of $\bar{g}$, mass transfer, for h/c_p, heat transfer, in the Stanton number
2 substitution of Schmidt number for Prandtl number in the Colburn factor
3 a limitation that the heat transfer correlations are based on a Prandtl number value that is numerically equal to the Schmidt number to be used for the mass-transfer calculation

If this limiting condition is not met, the analogy may be applied but with some reservations about the accuracy of calculated results.

A formulation for the Ntu_m of a bead bed can be derived from equations (11) and (23)

$$(\mathrm{Ntu}_m)_B = \left[\frac{\mathrm{Vol}}{A_{fr}}\right]\left[\frac{1}{r_h}\right]\left[\frac{j_m}{\mathrm{Sc}^{0.67}}\right] \tag{24}$$

The Colburn factor for heat transfer j_h for a sphere bed varies with Reynolds number as shown in Fig. 6. Thus, assuming condition No. 3 above is satisfied,

$$(\mathrm{Ntu}_m)_B = \left[\frac{\mathrm{Vol}}{A_{fr}^{0.70}}\right]\left[\frac{1}{r_h}\right]\left[\frac{\beta\mu}{4\dot{w}}\right]^{0.30}\left[\frac{0.23}{\mathrm{Sc}^{0.67}}\right] \tag{25}$$

This formulation shows the $(\mathrm{Ntu}_m)_B$ to be a function of both frontal area and volume. Hence in contrast to the monolith, bead-bed performance is dependent on the arrangement of volume into frontal area and flow length.

Performance Modeling

Performance modeling must include flow pressure drop in tradeoffs of performance versus size for catalyst supports. High conversion efficiency lowers the amount of pollutants exhausted from a vehicle. High pressure drop, however, increases engine backpressure, which lowers both engine power and efficiency.

Monolith Pressure Drop. For a monolith substrate, pressure

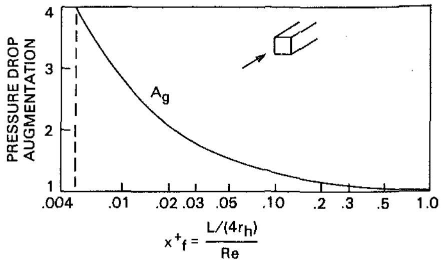


Fig. 7 Pressure drop augmentation for a developing boundary layer

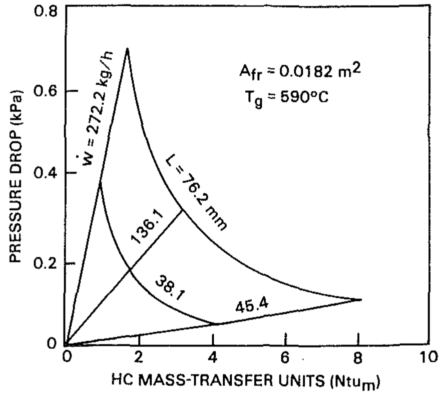


Fig. 8 Monolith performance map

Table 1 Monolith and bead-bed geometry factors

	Monolith	Bead bed
Porosity	0.620	0.380
Area density, m^2/m^3	1654	1171
Hydraulic radius, mm	0.3749	0.3245

Table 2 Typical thermodynamic properties and transport constants

Exhaust mass flow rate	136 kg/h
Exhaust temperature	590°C
Exhaust viscosity	3.84×10^{-5}Pa s
Monolith Sherwood number	2.976
Schmidt number, HC	1.77
Schmidt number, CO	0.69
Diffusion coefficient, HC	50 mm^2/s
Diffusion coefficient, CO	130 mm^2/s
Substrate specific heat	963. J/kg K

drop is required to overcome momentum losses from three sources:

1 flow contraction at the entrance minus expansion at the exit
2 fluid shear in the boundary layer
3 development of the parabolic velocity profile

A correlation for the overall pressure drop which accounts for these momentum losses, derived by Kays and London [9], is

$$\Delta P_M = \frac{\rho V^2}{2g_c}\left(K_c + K_e + A_g f \frac{L}{r_h}\right) \tag{26}$$

Entrance contraction coefficients K_c, and exit expansion coefficients K_e, are plotted in Kays and London [9]. The pressure drop augmentation A_g required to develop the flow is graphed in Fig. 7 for a monolith flow passage with square cross section. For fully developed laminar flow in a square passage the friction factor f correlates with Reynolds number [10]

$$f\mathrm{Re} = 14.23 \tag{27}$$

Bead-Bed Pressure Drop. For flow through a packed bed, all contraction, expansion, and augmentation effects are included in the experimental friction factor. A correlation [9] for this friction factor is also graphed in Fig. 6. The following equation has been derived for bead-bed pressure drop using this experimental friction factor [9]

$$\Delta P_B = \frac{\rho V^2}{2g_c} \frac{fL}{r_h} \tag{28}$$

Conversion Efficiency. Equation (13) can be used to calculate conversion efficiency for either a monolith or bead-bed geometry. Equal fully warm conversion for either a monolith or bead bed can be assured by equating their respective Ntu_m.

Warmup. Since a catalyst does not promote chemical reactions until heated to lightoff temperature, quick warmup is essential to control vehicle emissions to very low levels. Initial warmup of a substrate by convective heating was modeled using a finite-difference solution reported previously [13]. This model calculates the temperatures in both the solid substrate and the flowing fluid as functions of both space and time. These temperatures can be calculated if time-varying temperature and flow rates are known for the exhaust gas entering a substrate. If kinetics are assumed to be instantaneous and adequate quantities of chemical species are available to complete the chemical reactions, transient conversion efficiencies during warmup can be calculated by assuming that the instant any portion of substrate is heated above a lightoff temperature of 290°C, conversion efficiency is limited only by convective mass transfer.

The time required to warm up a converter until conversion efficiency equals 50 percent has been arbitrarily selected to compare alternative converter designs [14]. From Fig. 5, 50 percent conversion efficiency for a warmed-up converter corresponds to an Ntu_m of 0.70. Thus when a quantity of substrate corresponding to Ntu_m = 0.70 is heated above lightoff temperature, conversion efficiency should equal 50 percent.

Monolith Versus Bead-Bed Tradeoffs. Pressure drops and fully warm Ntu_m were calculated to compare monolith performance with that for a bead bed. Pertinent geometry factors for a square-cell ceramic monolith and a bed of 3.2-mm spherical pellets are listed in Table 1.

Performance was calculated for both a monolith and a bead bed for a frontal area of 0.0182 m^2. Pressure drop for the bead bed did not include a contribution from bead retainers. Contributions to pressure drop from housings were not included for either the monolith or the bead bed. Pertinent thermodynamic properties and mass-transfer constants are listed in Table 2.

Using pressure drop for the substrate only and constants in Table 2 for mass transport, both pressure drop and Ntu_m are linear functions of length for either a monolith or bead bed. Thus if both frontal area and flow rate are fixed, pressure drop and flow length are both linear functions of Ntu_m. By selecting different flow rates, a sizing map can be constructed. Figure 8 is a sizing map for a square-cell monolith. As shown in Fig. 8 at an idle flow rate of 45.4 kg/h, potential conversion efficiency or Ntu_m is high and pressure drop is low. At a relatively high flow rate of 272 kg/h, conversion efficiency decreases and pressure drop increases.

A similar sizing map for a bead bed is shown in Fig. 9. At low flow rates the bead-bed potential conversion efficiency or Ntu_m is slightly higher than that for a monolith. As flow rate is increased, bead-bed potential conversion efficiency does not decrease as rapidly as does monolith potential conversion efficiency. However pressure drop for the bead bed increases

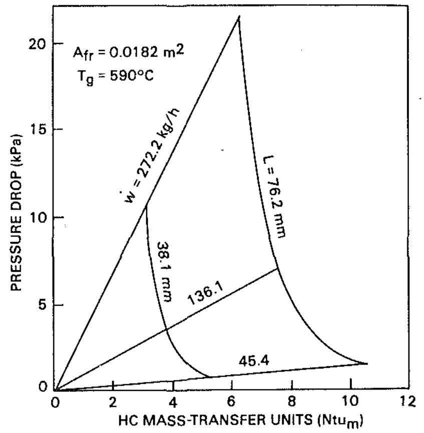
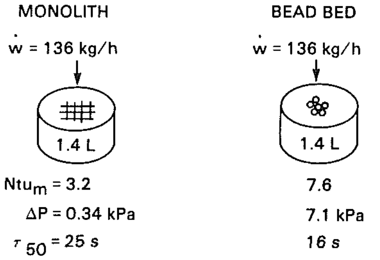
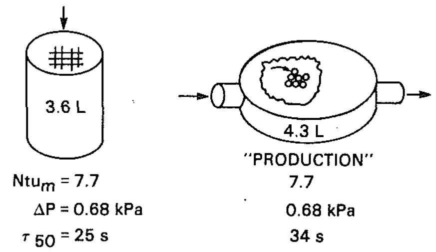


Fig. 9 Bead-bed performance map



a. Same Size



b. Same Conversion and Pressure Drop

Fig. 10 Two sizing comparisons of monoliths versus bead beds

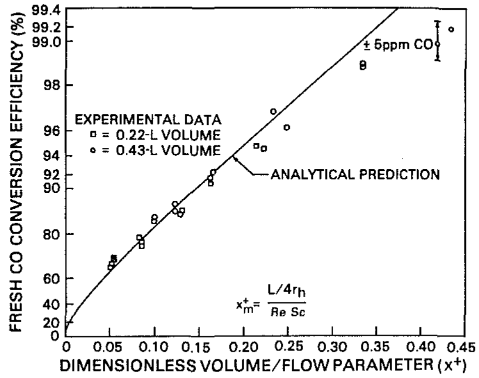


Fig. 11 CO conversion efficiency for two small monoliths

rapidly as flow rate increases, and it is 31 times that for the monolith at a flow rate of 272 kg/h.

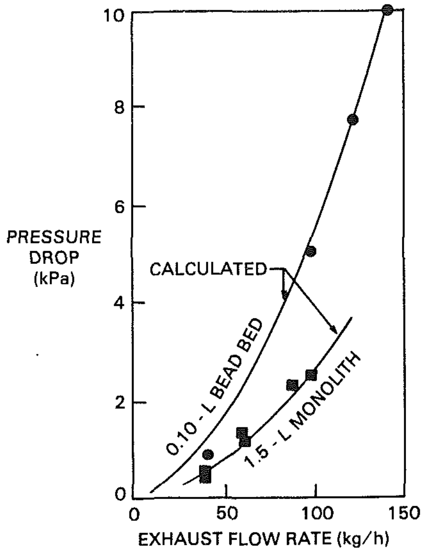


Fig. 12 Pressure drops for a monolith and a bead bed tested on a vehicle

Two examples of monolith versus bead-bed sizing tradeoffs are presented in Fig. 10, for the geometries listed in Table 1 and the operating environment represented by the constants in Table 2. In Fig. 10(*a*), a bead bed and monolith are compared with the same size and shape. The bead bed offers higher potential conversion efficiency with a calculated Ntu_m of 7.6, compared with potential conversion efficiency of a monolith with a calculated Ntu_m of only 3.2. However, the pressure drop for this small bead bed was calculated to be 21 times that for the monolith.

For warmup comparisons, time delays were calculated using the procedure discussed previously in the section on warmup. Time delays are those required to heat a volume of substrate corresponding to $Ntu_m = 0.7$ to a lightoff temperature of 290°C. For these calculations, experimental temperature data were used for the exhaust gas flow. These temperature data were measured for a standard converter warmup test, which is described in more detail in the next section of this report. Calculated time delays correspond to the time to attain 50 percent conversion efficiency, calculated as 16 s for the bead bed and 25 s for the monolith.

As shown in Fig. 10(*b*), a monolith can be sized to match conversion efficiency of a production bead bed by increasing the flow length of the monolith to match the bead bed Ntu_m of 7.7. Pressure drop calculated for this larger 3.6-*L* monolith matches the pressure drop of a 4.3-*L* bead bed. Comparing lightoffs calculated using the dynamometer warmup-test temperatures, the monolith delay time until attaining 50 percent conversion was calculated as 25 s, compared to a calculated delay time of 34 s for the bead bed.

Experimental Verification

Many experimental catalytic converters have been built and tested to confirm model predictions. Conversion efficiencies for small monoliths were measured for fully warm operation using an engine-dynamometer test. Pressure drops were measured for small monoliths and a small bead bed tested on vehicles.

Fully Warm Conversion Efficiencies for a Monolith. Two small monolith converters, 0.22-*L* and 0.43-*L*, were tested for conversion performance on an engine dynamometer. Small converters were used so differences in conversion performance could be accurately detected. Efficiencies for conversion of carbon monoxide were measured for these small converters [11], and the test results were correlated using the entrance-region parameter x_m^+. As shown in Fig. 11, measured conversion efficiencies confirm within a few percent values for conversion efficiencies calculated using mass-transfer-limited models for x_m^+ varying from 0 to 0.35.

Pressure Drop Measurements. Pressure drop data are graphed in Fig. 12 for a 1.5-*L* ceramic monolith tested on a vehicle operated at several steady-speed conditions on a

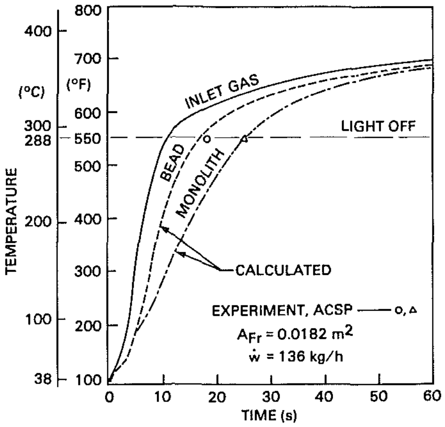


Fig. 13 Converter thermal warmup for $Ntu_m = 0.70$ on the ACSP test

chassis dynamometer. To measure pressure drop for the substrate only, probes were mounted in the gas stream immediately upstream of, and immediately downstream from, the monolith. As shown in Fig. 12, measured pressure drops confirm within a few percent pressure drops calculated using equation (26). Pressure drops were also measured for a very small 0.10-L bed of 3.2-mm spherical pellets. Pellets were retained in a housing by open-mesh screens to minimize the contribution to pressure drop by the retainers. This bead bed was tested on a vehicle operated at several controlled speeds on a chassis dynamometer. Measured data for pressure drop for this 92.0-mm diameter by 19.1-mm flow-length bead bed are also graphed on Fig. 12. Measured data confirm within a few percent pressure drop values calculated using equation (28).

Warmup. Warmup models were confirmed by comparing calculated warmup with warmup measured using an engine dynamometer test developed at AC Spark Plug Division, ACSP [14]. Exhaust from a test-stand engine can be valved either to bypass around or to flow through a test converter. By running the engine with exhaust gas bypassing the converter, very stable exhaust temperature, flow rate, and constituent concentrations can be established. Then by closing the bypass valve, the test converter is exposed to hot exhaust. Immediately after closing the bypass valve, the flow rate and temperature of the gas entering the converter increase rapidly. On this test, converter thermal response is rated by the time delay after valve closure for conversion efficiency to reach 50 percent.

To validate warmup models, a small monolith and a small bead bed were built with the dimensions shown in Fig. 10(*a*). The ACSP transient dynamometer test was used to measure warmup delay until conversion efficiency attained 50 percent for these two substrates. This time delay, or lightoff time, was compared to the lightoff time calculated using the thermal-response model discussed previously.

Lightoff time for the monolith measured 25 s, compared to a calculated lightoff time of 24 s to heat a volume of substrate corresponding to an $Ntu_m = 0.70$ to a temperature of 290°C, as shown in Fig. 13. Similarly for the small bead bed, measured lightoff of 17 s compares to 16 s calculated using the model.

Summary

Mass-transfer-limited models based on the heat and mass-transfer analogy can be used to calculate fully warm conversion performance for either monolith or bead-bed converters with fresh catalyst. These models, combined with models for convective heat transfer and flow pressure drop, are useful to expedite development of alternative catalytic converters by simplifying calculations of tradeoffs of size and shape versus conversion efficiency and pressure drop.

1 For bead beds, a prerequisite for strict application of the analogy is that the numerical value of the Schmidt number (mass transfer) must fall within the range of the Prandtl number (heat transfer) used to correlate the experimental heat transfer data.

2 For fully developed laminar flow, Nusselt number (heat transfer) equates to Sherwood number (mass transfer) independent of the ratio of Schmidt number to Prandtl number.

3 The heat and mass-transfer analogy also is applicable to developing flow in the entrance region of a monolith by substituting Schmidt number for Prandtl number in the x^+ parameter.

4 For a fixed volume, both conversion efficiency and warmup time are functions of frontal area for a bead bed, but not functions of frontal area for monoliths.

5 For an equal volume of monolith or bead bed, a bead bed can be packaged with a small frontal area to provide higher conversion efficiency and quicker warmup than for a monolith; however, bead-bed pressure drop is then unacceptably high.

6 For equal fully warm conversion and acceptable pressure drop, a monolith warms up faster than a bead bed, even if bed retainers are neglected.

Acknowledgments

The author gratefully acknowledges the contribution to this paper by D. W. Wendland of test results for CO Conversion efficiency of small monoliths. Also the Product Engineering and Test Engineering Groups at A. C. Spark Plug Division are acknowledged for providing measured data from the engine-dynamometer transient test.

References

1 Mondt, J. R., "An Historical Overview of Emission-Control Techniques for Spark-Ignition Engines," GM Research Publication GMR-4228, Dec. 23, 1982.

2 Kuo, J. C. W., Morgan, C. R., and Lassen, H. G., "Mathematical Modeling of CO and HC Catalytic Converter Systems," *SAE Transactions*, Vol. 80, Paper No. 710289, 1971.

3 Hegedus, L. L., "Effects of Channel Geometry on the Performance of Catalytic Monoliths," ACS Symposium on Catalytic Approaches to Environmental Control, Chicago, Aug. 1973.

4 Johnson, W. C., and Chang, J. C., "Analytical Investigation of the Performance of Catalytic Monoliths of Varying Channel Geometries Based on Mass Transfer Controlling Conditions," SAE Preprint No. 740196, Feb. 1974.

5 Mondt, J. R., "NO_x Catalyst Performance Comparison of 304 Stainless Steel, Inconel, and Monel in a 10°-Herringbone Foil Configuration," *AIChE Symposium Series No. 165*, Vol. 73, 1977, pp. 169–177.

6 Oh, S. H., and Cavendish, J. C., "Design Aspects of Poison-Resistant Automobile Monolithic Catalysts," *Ind. Eng. Chem. Prod. Res. Dev.*, Vol. 22, 1983, pp. 509–518.

7 Skelland, A. H. P., *Diffusional Mass Transfer*, Wiley, New York, 1974.

8 Kays, W. M., *Convective Heat and Mass Transfer*, McGraw-Hill, New York, 1966.

9 Kays, W. M., and London, A. L., *Compact Heat Exchangers*, 2nd ed., McGraw-Hill, New York, 1964.

10 Shah, R. K., and London, A. L., *Laminar Flow Forced Convection in Ducts*, Academic Press, New York, 1978.

11 Wendland, D. W., "The Segmented Oxidizing Monolith Catalytic Converter," ASME *Journal of Heat Transfer*, Vol. 102, No. 2, May 1980, pp. 194–198.

12 Moffat, R. J., "An Introduction to the Calculation of Mass Transfer Rates Using Heat Transfer Data," unpublished notes, 1973.

13 Mondt, J. R., "Vehicular Gas Turbine Periodic-Flow Heat Exchanger Solid and Fluid Temperature Distributions," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 86, No. 2, Apr. 1964, pp. 121–126.

14 Herod, D. M., Nelson, M. V., and Wang, W. M., "An Engine Dynamometer System for the Measurement of Converter Performance," SAE Preprint No. 730557, May 1973.

Fuel Oil Reburning Application for NO_x Control to Firetube Package Boilers

J. A. Mulholland

R. E. Hall

U. S. Environmental Protection Agency,
Air and Energy Engineering
Research Laboratory,
Research Triangle Park, NC 27711

Two pilot-scale (0.73 MW or 2.5 × 10^6 Btu/hr) firetube package boilers were retrofitted for fuel oil reburning application for NO_x emission control. When firing distillate fuel oil (0.01 percent nitrogen content), an overall NO_x reduction of 46 percent from an uncontrolled emission of 125 ppm (dry, at zero percent O_2) was realized by diverting 20 percent of the total boiler load to a second stage burner; a 51 percent NO_x reduction from 265 ppm was achieved in a distillate/residual fuel oil mixture (0.14 percent nitrogen content) reburning application. Nitrogen-free fuel oil reburning was found to be slightly more effective at reducing NO_x than was natural gas reburning, although longer fuel-rich zone residence times were required to allow for evaporation and mixing of the fuel oil droplets. Key parameters investigated which impact the reburning process were: primary flame NO_x, reburn zone stoichiometry, and reburn zone residence time. Reburning applied to firetube package boilers requires minimal facility modification. Reburning can be coupled with other NO_x control techniques (e.g., distributed air low NO_x burners) to achieve NO_x emissions of less than 100 ppm. However, for very low primary flame NO_x conditions (i.e., less than 200 ppm), reburning fuel nitrogen content is a limiting factor, and reburning with a low-nitrogen-content fuel, such as natural gas or nitrogen-free distillate oil, may be necessary to achieve 50 percent NO_x reduction.

Introduction

Oxides of nitrogen (NO_x) are formed during combustion by the reduction and oxidation of molecular nitrogen and nitrogen contained in the fuel. Because NO_x emitted from stationary combustion sources contributes to the degradation of air quality as well as to acid deposition and forest damage, various control technologies have been developed to reduce emissions. Many of the NO_x control technologies currently in use involve modification of the combustion process to reduce peak flame temperatures and create fuel-rich conditions by altering fuel and oxidizer mixing rates. Thus, NO_x formation, a kinetically controlled process, is minimized. However, NO_x control can also be achieved by removing NO_x from the combustion products after they are formed. Reburning, or fuel staging, is an in-furnace NO_x control strategy that accomplishes this.

As many as 30 years ago [1, 2], experimental results were available indicating that NO_x could be destroyed by reaction with hydrocarbon radicals. Commercial utilization of this reaction mechanism occurred in the late 1960s by the John Zink Corporation, which patented an NO_x control process for nitric acid plants [3] based on staged injection of natural gas. Early fundamental studies of NO_x destruction by injection of secondary fuel into the flame zone were performed by Wendt et al. [4], who coined the name "reburning" to describe the process. Concerted effort to develop reburning for application to boilers occurred in Japan during the late 1970s and early 1980s. The first reporting of these efforts was by Takahashi et al. [5], documenting extensive laboratory-, pilot-, and full-scale evaluation of a fuel staging process they refer to as Mitsubishi Advanced Control Technology, or MACT. MACT results indicate that NO_x emissions can be reduced by at least 50 percent independent of initial NO_x level and fuel type. These promising results have renewed interest in reburning, and further development is being sponsored by the U. S. Environmental Protection Agency (EPA), the Electric Power Research Institute, the Gas Research Institute, and the U. S. Department of Energy.

Application of reburning involves diverting portions of the fuel and combustion air streams from the main burner(s) for injection into the postflame gases. Thus, reburning establishes a three-stage combustion process, consisting of a fuel-lean primary zone, a fuel-rich reburning zone, and a fuel-lean burnout zone. NO_x reduction via reburning occurs in two ways: NO_x formation is inhibited due to reduced mixing intensities in the derated primary burner zone and fuel-rich conditions in the secondary combustion zone, and NO_x destruction is promoted due to secondary flame radical attack on nitric oxide (NO) produced in the primary zone to form molecular nitrogen (N_2). Maximizing the destruction component of NO_x reduction is a key to optimizing the reburn technology. Bench-scale tests [6, 7] have been used to verify the hypothesized NO_x destruction mechanism, represented in a simple form as follows:

$$\mathrm{NO} \xrightarrow{\mathrm{CH}} \mathrm{HCN} \xrightarrow{\mathrm{OH,\,H}} \mathrm{NH} \begin{array}{l} \xrightarrow{\mathrm{OH,\,H}} \mathrm{NO} \\ \xrightarrow[\mathrm{NO}]{} \mathrm{N_2} \end{array} \qquad (1)$$

Contributed by the Air Pollution Control Division and presented at the Joint Power Generation Conference, Milwaukee, Wisconsin, October 1985. Manuscript received by the Air Pollution Control Division June 18, 1985.

Full-scale application of reburning has been achieved in Japan, where NO_x emission regulations are stringent, by Hitachi Zosen [8] and Ishikawajima-Harima Heavy Industries [9], as well as Mitsubishi Heavy Industries [10]. In the U. S., bench- and pilot-scale evaluation of gas, oil, and coal reburning processes continues [11-13]. Application of reburning to a coal-fired industrial boiler, sponsored by the U. S. Environmental Protection Agency (EPA) and the Gas Research Institute (GRI), is scheduled to begin in 1986. While relatively inexpensive distributed air low-NO_x burner designs have been developed to meet current federal and state NO_x emission standards (in the 300 to 500 ppm range), reburning offers a means to reduce NO_x emissions by another 50 percent without impacting furnace combustion efficiency.

The current study utilizes two pilot-scale firetube package boiler test facilities located at EPA's Environmental Research Center in Research Triangle Park, NC. Tests were performed to evaluate fuel oil reburning application at low primary combustion NO_x levels (50 to 250 ppm[1]), which simulates a combined NO_x control strategy to achieve extremely low NO_x levels (less than 125 ppm), and low reburn zone temperatures (1300 to 1600 K), which simulates the thermal environment of a small commercial firetube package boiler. Previously documented tests on these facilities have shown that natural gas reburning effectiveness in NO_x control decreases as primary NO_x levels decrease [14, 15]. That is, NO_x reduction via natural gas reburning was shown to have an order of 1.5 with respect to primary NO_x. It has also been shown that, at low initial NO_x levels, volatile nitrogen in the reburning fuel limits reburning effectiveness [16]. This paper describes tests to firetube package boilers that evaluate these results for fuel oil reburning application.

Experiment

Apparatus. Two pilot-scale firetube package boiler test facilities, each rated at 0.73 MW (2.5×10^6 Btu/hr) heat input and capable of firing gaseous and liquid fuels, were modified to accommodate fuel oil reburning for this series of experiments. A package boiler simulator (PBS) fitted with a low NO_x precombustion chamber burner, shown in Fig. 1(*a*), was utilized to perform parametric input/output and detailed tests to evaluate fuel oil reburning effectiveness. A North American Scotch-type package boiler, shown in Fig. 1(*b*), was used to verify PBS test results and demonstrate optimized fuel oil reburning application.

The radiant section of the PBS is cooled with Dowtherm G®, with absorbed heat dumped to the atmosphere via a cooling tower. The horizontal radiant section is 0.6 m (2 ft) in diameter and 3.0 m (10 ft) in length. Combustion gas exits the boiler through a vertical stack. A commercial burner originally installed on the research boiler was replaced with a low NO_x precombustion chamber burner patterned after the unit developed by England et al. [17]. This major facility modification was made to achieve control of the NO_x concentration at the plane of secondary fuel injection, independent of boiler fluid mechanics. The prechamber itself is a refractory tunnel designed for near adiabatic conditions and contains a fuel-rich flame. At the interface between the prechamber and the radiant section, the flow path is constricted through a 25 cm (10 in.) diameter throat. Eight ports, located in this throat region, allow staged combustion air to be injected inward. This air mixes rapidly with the fuel-rich exhaust from the prechamber and produces a fuel-lean, essentially premixed flame in the radiant section. Multiple ports allow for radial and axial insertion of various sample probes and staged fuel and air injectors.

Reburning fuel oil was injected through a heat-traced line contained in a water-cooled, stainless steel jacket inserted through a side port in the PBS. A 45-deg hollow-cone fuel spray was achieved with a pressure-atomizing nozzle directed downstream into the primary combustion exhaust along the boiler centerline. Final burnout air was injected downstream of the staged fuel injection location through eight radial ports in a water-cooled boom inserted along the boiler centerline through a rear flange.

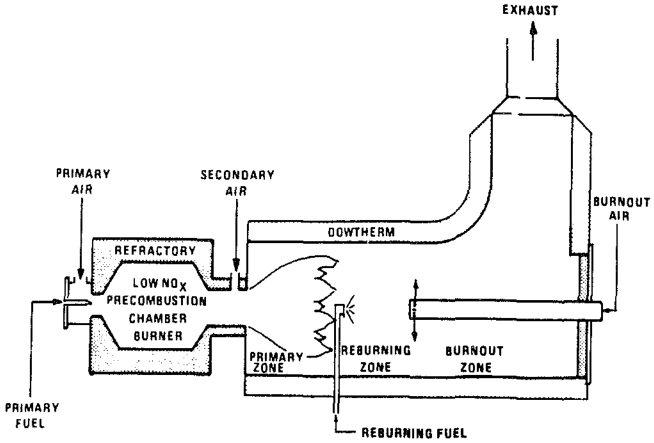


Fig. 1(*a*) Pilot-scale package boiler simulator test facility showing fuel oil reburning application

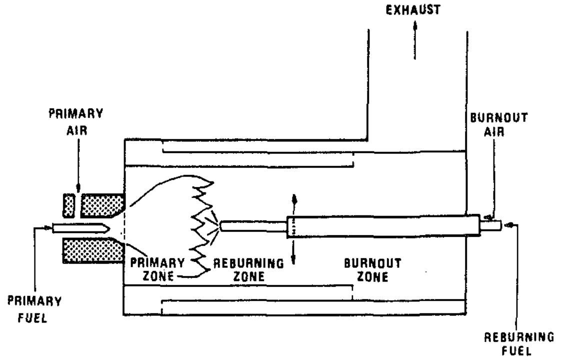


Fig. 1(*b*) Pilot-scale North American firetube package boiler showing fuel oil reburning application

The North American boiler is a three-pass unit, 0.6 m (2 ft) in internal diameter and 3.0 m (10 ft) in length, with a continuous service rating of 302 kg of steam per second (2400 lb/hr). The mechanical linkage for automatic control of air/fuel ratio with load variation was disconnected from the burner and replaced with a manual adjustment that permits independent control of primary flame excess air at any desired load.

Reburning fuel and final burnout air were injected at locations along the centerline of the North American boiler through a water-cooled boom assembly inserted through a modified maintenance access plate at the rear of the boiler. The assembly consists of two concentric telescoping injectors, with the inner injector supplying reburning fuel through a heat-traced line and the outer injector providing burnout air in eight radial jets. As in the PBS, a 45-deg hollow-cone, pressure-atomizing fuel spray nozzle was used, but here it was directed upstream into the primary combustion gas.

The experimental facilities were designed with independent control of each fuel, fuel dopant, and air stream. Natural gas, distillate fuel oil, and a distillate/residual fuel oil mixture were fired, with ammonia and pyridine dopants being used to simulate fuel-bound nitrogen. Ammonia was injected into the natural gas delivery system sufficiently upstream to ensure thorough mixing before entry to the combustion system. Distillate fuel oil, containing 0.01 percent (by weight)

---

[1] All NO_x concentrations are reported on a dry basis, and corrected to zero percent O_2, in this paper.

Table 1 Fuel analysis comparison

	Natural Gas	Distillate Oil	Distillate/Residual Oil Mixture
Gross Heating Value:			
Btu/lb$_f$ (kJ/kg$_f$)	23,563 (54,810)	19,382 (45,083)	18,734 (43,578)
Theoretical Air:			
lb$_a$/10,000 Btu (kg$_a$/kJ)	7.04 (3.03)	7.48 (3.22)	7.58 (3.26)
Encal Chromatograph Analysis, mol. %:		Ultimate Analysis, wt %:	
Nitrogen	0.35563	Carbon: 6.44	87.12
Carbon dioxide	0.91649	Hydrogen: 13.21	11.97
Methane	96.11870	Nitrogen: 0.01	0.14
Ethane	1.85666	Sulfur: 0.22	0.84
Propane +	0.75252	Ash: <0.01	0.02
		Viscosity, cs (ft^2/sec):	
		@ 38°C (100°F) 2.58 (0.0000278)	26.76 (0.0002880)
		@ 99°C (210°F) 1.08 (0.0000116)	4.68 (0.0000504)

Table 2 Gaseous emission monitor specifications

Gas	Sample Type	Instrument Type	Manufacturer/Model	Range
O_2	Continuous	Paramagnetic	Beckman/F3	0 to 5 10, 25%
CO_2	Continuous	Infrared	Beckman/864	0 to 5 10, 20%
CO	Continuous	Infrared	Beckman/865	0 to 500, 1000, 2000 ppm
NO_x	Continuous	Chemiluminescence	Teco/10A	0 to 2.5, 10, 25, 100, 250, 1000, 2500, 10000 ppm
THC	Continuous	Flame Ionization	Beckman/400	0 to 10, 100, 1000, 10000 ppm
NH_3	Batch-wet	Specific Ion Electrode	Orion/95-10	0.1 to 100 ppm
HCN	Batch-wet	Specific Ion Electrode	Orion/94-06	0.1 to 100 ppm
N_2O	Batch	Gas Chromatograph	Varian/1400	1 to 1000 ppm

nitrogen, was doped with pyridine to achieve a 0.5 percent fuel nitrogen batch mixture. The distillate/residual fuel oil mixture contained 0.14 percent nitrogen. This heavier fuel oil, with a net viscosity of a light No. 5 fuel oil, was preheated to 370 K (200°F) prior to injection. A comparison of various properties of these fuels is presented in Table 1.

Methods. All gas flows were measured independently using various types of flow meters. The fuel oil volumetric flow rate was measured with a calibrated metering pump.

Stack and in-flame combustion gas speciation, temperature, and velocity were measured. A sampling panel board provided for continuous measurement of NO_x, CO, CO_2, THC, and O_2 concentrations, as defined in Table 2. In-flame species concentrations were determined by extracting samples through a water-cooled stainless steel probe and ducting the sample gases to the continuous emission monitors. The sample probe is equipped with a water quench spray which was activated during batch sampling to collect HCN and NH_3 in solution from the reacting gas. HCN and NH_3 were measured using specific ion electrode techniques. The collection efficiency for HCN was experimentally determined to be 70 percent. Ammonia collection efficiency was found to be nearly 100 percent. Nitrous oxide (N_2O) was measured with a gas chromatograph equipped with an electron capture detector and a 3.66 m (12 ft) long stainless steel column packed with Porapak Q®.

Flame temperature measurements utilize suction pyrometers equipped with R-type (platinum-platinum/13 percent rhodium) thermocouples. A short pyrometer was inserted through sidewall ports in the furnace to measure radial temperature profiles. A long pyrometer was used to measure axial temperature profiles through the entire length of the radiant section at three different radii.

Table 3 Reburning process parameters

Primary Zone	Reburning Zone	Burnout Zone
° Initial NO_x Concentration	° Stoichiometry	° Stoichiometry
° Stoichiometry	° Residence Time	° Temperature
° Primary Fuel Properties	° Temperature	° Burnout Air Mixing Rate
	° Reburning Fuel Mixing Rate	
	° Reburning Fuel Properties (e.g., nitrogen content, volatility, C/H ratio)	

Velocity profiles at the secondary fuel injection plane are derived from measurements made with a Prandtl probe inserted through a port in the boiler sidewall. Measurement of the dynamic pressure differential in the radiant section of the firetube package boiler was particularly difficult due to the low mean velocity (less than 5.0 m/s) and the high apparent turbulent fluctuations. Pressures from 0.1 to 100 Pa (0.00071 to 0.71 in. of H_2O) were measured using a differential pressure transducer in tandem with an electronic manometer (Barocell model 1173). The electrical output signal for the dynamic pressure differential measurement was sampled by a microcomputer at an approximate rate of 15 Hz, with the square roots of approximately 1250 individual readings averaged at each point. With this technique it is possible to obtain accurate measurements (with mean and standard deviation repeatable to at least two significant digits) for boiler radiant section velocities greater than 3.0 m/s (10 ft/s). It is significant to note that the ratio of the standard deviation to the mean was often in excess of 50 percent, increasing with decreasing velocity.

Approach. This paper, describing a series of pilot-scale experiments designed to investigate fuel oil reburning application for in-furnace NO_x control, provides information bridging the gap between fundamental studies of NO/hydrocarbon reaction mechanisms and full-scale field applications of reburning. The firetube package boilers used in this study have single centerline burners and relatively simple furnace geometries (right circular cylinders), allowing for complete spatial and temporal characterization of experimental conditions while providing thermal and fluid mechanical environments typical of small commercial systems.

Three objectives were defined at the outset of the experimental program. First, parametric tests were performed in the PBS to characterize the fuel oil reburning process at a practical level. Key process parameters to be varied independently were identified, as listed in Table 3. In this paper, light and heavy fuel oil reburning test results are discussed which describe the effects of three of these parameters: primary NO_x concentration, reburn zone stoichiometry, and reburn zone residence time. Separate papers have been published, or are in preparation or review, on natural gas reburning application [14], reburning fuel mixing rate effects (in review), reburn zone thermal environment effects (in preparation), and reburning fuel nitrogen content effects [16]. A second objective of this study was to evaluate the validity of proposed reburning reaction mechanisms by making detailed measurements of intermediate nitrogenous species in the reburn zone of the PBS. Third, fuel oil reburning application was to be optimized in the North American boiler.

Results

Parametric tests, in-flame measurement tests, and optimization tests are discussed. N_2O measurements in both the stack and in-flame gases were found never to exceed 3 ppm. Therefore, these measurements are not reported here.

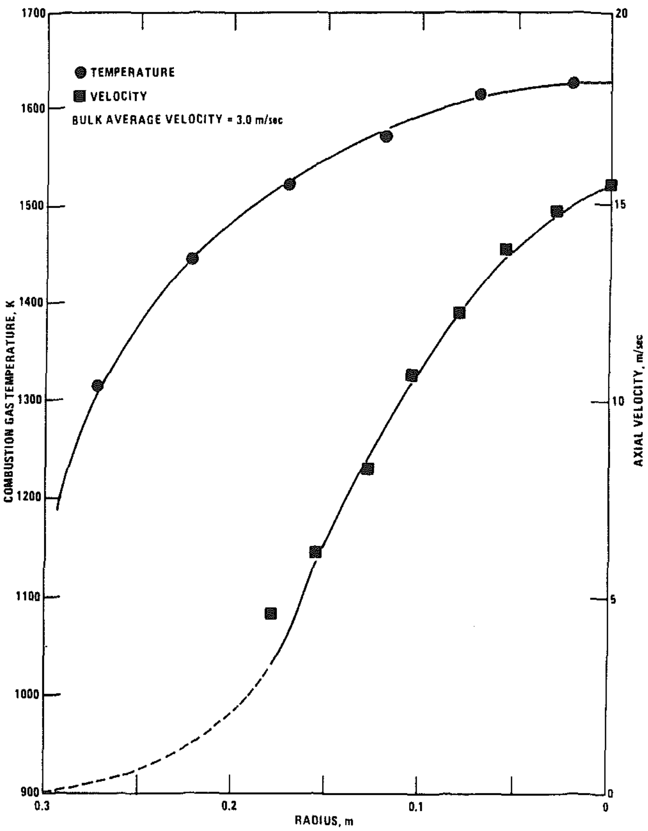


Fig. 2 Baseline radial temperature and velocity profiles measured at the secondary fuel addition location

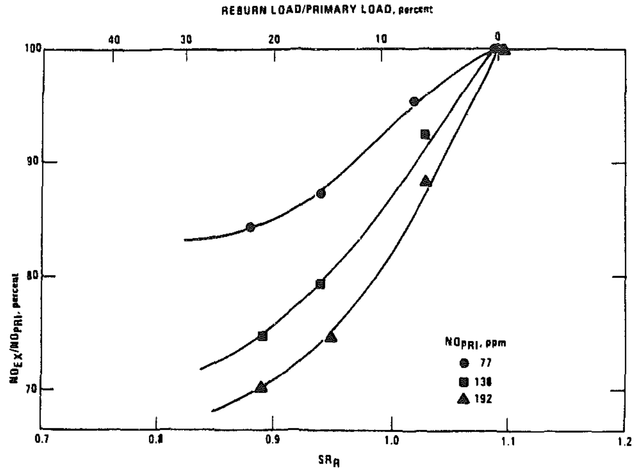


Fig. 3 Exhaust NO as a function of reburn load for distillate fuel oil reburning

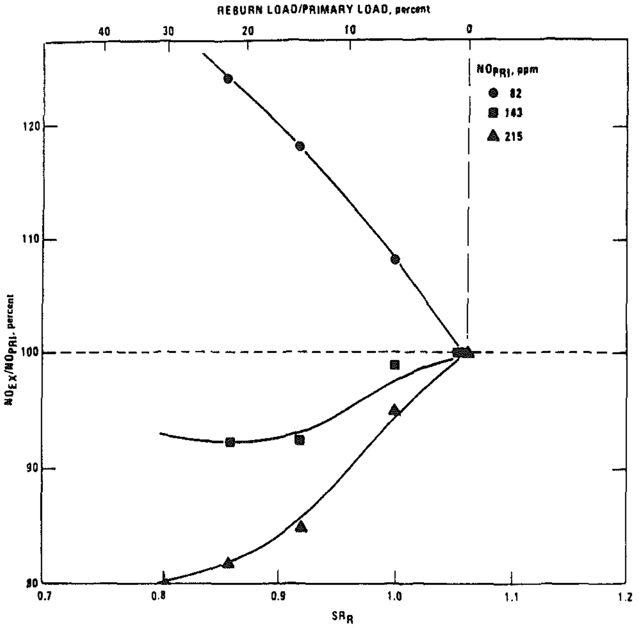


Fig. 4 Exhaust NO as a function of reburn load for distillate/residual fuel oil mixture reburning

Parametric Tests. In the PBS tests, primary burner load was held constant at 0.59 MW (2.0×10^6 Btu/hr). A precombustion chamber stoichiometry of 0.61 was maintained to allow for primary NO_x variation from 50 to 250 ppm, adjusted by varying the amount of ammonia injected into the primary natural gas stream. A secondary combustion air addition rate was established such that a primary zone excess air level of 5 to 10 percent was achieved. Under these conditions, the primary flame extended approximately 0.9 m (35 in.) into the boiler. At the 1.02 m (40 in.) axial location, the radial profile of O_2 concentration was essentially flat. Also, this location was just downstream of the point at which flow from the precombustion chamber constriction reattached to the boiler wall. Thus, a nominal location for secondary fuel injection of 1.02 m (40 in.) was chosen. Final burnout air was staged at 1.63 m (64 in.), providing a 200 ms reburn zone bulk residence time. For natural gas reburning, this reburn zone length was found to be the minimum necessary to achieve maximum NO_x reduction.

The temperature and velocity profiles measured at the location of reburn fuel injection are shown in Fig. 2. Temperatures ranged from 1200 K (1700°F) near the wall to 1600 K (2420°F) at the boiler centerline. These low radiant zone temperatures are characteristic of small package boilers. It was found that this was the low temperature limit at which the fuel oil could be injected without heavy sooting in the reducing zone. The combustion gas flow had a peak velocity of over 15 m/s (50 ft/s) and a mean velocity of 3 m/s (10 ft/s).

Input/output fuel oil reburning trials were performed at three levels of primary zone NO (NO_{PRI}) and three reburn zone stoichiometries (SR_R). Results from these tests using distillate oil and distillate/residual oil mixture as reburning fuels are shown in Figs. 3 and 4, respectively. In these figures the abscissa is SR_R, varied by adjusting the flow rate of reburning fuel, and the ordinate is the percent ratio of exhaust NO (NO_{EX}) to NO_{PRI}. Thus, the ordinate is equivalent to 100 percent minus the percent NO reduction via reburning. Each curve represents a line of constant NO_{PRI}. These results show clearly that NO_x control by reburning is a function of primary zone NO_x, reburn zone stoichiometry, and reburning fuel properties.[2]

To show the dependence of NO_x reduction on NO_{PRI}, the light and heavy fuel oil reburning data are replotted in Fig. 5 for a reburn stoichiometry of 0.90. The overall rate of change of NO in the reburning process can be expressed as

$$\frac{d[\mathrm{NO}]}{dt} = -k_1[\mathrm{NO}]^a + k_2 \qquad (2)$$

---

[2] The formation of NO in the reburning zone, via either the Fenimore "prompt" NO mechanism or fuel-bound nitrogen oxidation (reburn zone temperatures are too low to result in NO formation by the Zeldovich "thermal" NO mechanism), is not independent of NO_{PRI}. It has been shown that, for NO_{PRI} less than 250 ppm, SR_R between 0.75 and 1.0, and volatile fuel nitrogen content of up to 1 percent, reburning fuel nitrogen to exhaust NO net conversion is 30 to 40 percent [16]. Reburning prompt NO formation was found to range from 20 to 30 ppm. Therefore, over the range of NO_{PRI} considered in this paper, the assumption that reburning NO formation is independent of NO_{PRI} is approximately true. It should also be noted that, at high NO_{PRI} conditions, fuel nitrogen may result in a net increase in NO reduction due to amines formed from the fuel nitrogen reacting with NO to form N_2. It has been shown by as yet unpublished data that, for high NO_{PRI} conditions (greater than 500 ppm), reburn zone ammonia injection increases reburning NO_x reductions.

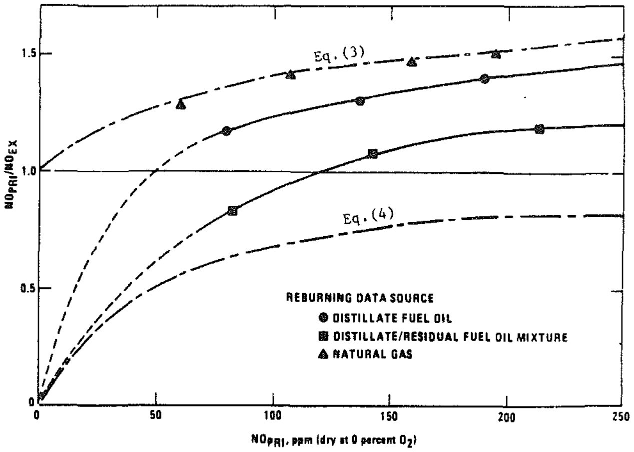


Fig. 5 Exhaust NO as a function of primary NO for a reburn load of 20 percent of primary laod

where k_1 is an overall destruction rate term, a is the overall order of destruction with respect to NO, and k_2 is a formation term which is assumed in this expression to be independent of NO concentration. Equation (2) is evaluated for two limiting cases. First, if k_2 is small relative to the k_1 term, as would be the case for high initial NO_x, the equation can be integrated to yield

$$\frac{NO_{PRI}}{NO_{EX}} = [1 + (a-1)k_1 t\, NO_{PRI}^{(a-1)}]^{1/(a-1)}, \qquad a > 1 \tag{3}$$

Using this expression, a curve fit to the natural gas reburning data is shown in Fig. 5. The natural gas data are consistent with an overall order of NO_x destruction of 1.5 with respect to NO_{PRI}, for $k_1 t = 0.032$ ppm$^{-1/2}$. The formation term is negligible for NO_{PRI} greater than 50 ppm. This result contradicts the first-order reduction ($a = 1$) claimed by the MACT researchers [5]. In a later report on MACT [10], it is suggested that the MACT process includes variable application of a selective catalytic reduction technology as well as reburning.

At low NO_{PRI} conditions, where the NO formation term dominates the NO destruction term of equation (2) (i.e., $k_2 \gg k_1$ $[NO]^a$), integration yields

$$\frac{NO_{PRI}}{NO_{EX}} = [1 + (k_2 t/NO_{PRI})]^{-1} \tag{4}$$

A curve representing equation (4) is also shown (Fig. 5), with $k_2 t = 51$ ppm.

The fuel oil reburning data plotted in Fig. 5 lie between the two limiting curves. For low NO_{PRI} conditions, the NO formation term is significant. At high NO_{PRI} conditions, the data appear to be approaching the natural gas results described by the destruction term curve with an order of 1.5 with respect to NO_{PRI}. Thus, it is shown that, as initial NO_x is decreased, it becomes more difficult to reduce NO_x via reburning, and fuel nitrogen can limit reburning effectiveness.

There exists an optimum reburn zone stoichiometry for maximum NO_x reduction by reburning. This optimum is a function of NO_{PRI} and reburning fuel nitrogen content. Figures 3 and 4 show that, for high levels of NO_{PRI} (greater than 200 ppm), maximum NO_x reductions are achieved for reburn zone stoichiometries less than 0.85 (reburn load/primary load greater than 25 percent). As the NO_{PRI} level decreases, the optimum SR_R increases. For the distillate/residual fuel oil mixture containing 0.14 percent nitrogen, at an initial NO level of 82 ppm, the minimum exhaust NO was achieved with no reburning.

Data from Figs. 3 and 4 are replotted in Figs. 6(*a*) and 6(*b*), respectively, to show the effect of SR_R on the various components of NO_x reduction by reburning. Overall NO_x reduction via reburning occurs by both the destruction of moles of NO produced in the primary zone and the "dilution" of the high NO primary zone gases by the staged fuel and air. The dilution portion of the overall NO reduction is calculated from the percent reburn load (assuming zero NO contribution by secondary combustion) and plotted as dashed lines on Figs. 6(*a*) and 6(*b*). In the case of distillate fuel oil, which contains virtually no fuel-bound nitrogen, the difference between the overall NO reduction curve and the dilution line represents the net destruction. In the case of the distillate/residual fuel oil mixture, the conversion of fuel nitrogen to N_2 is included in the net destruction component of overall NO reduction.

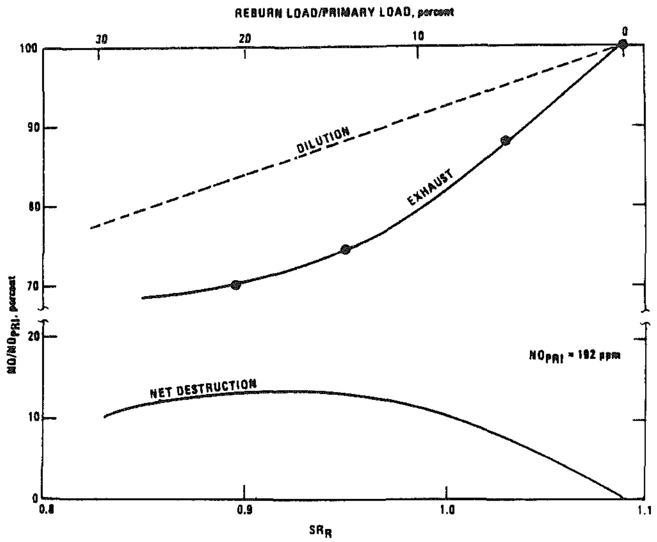


Fig. 6(*a*) Optimum stoichiometry for distillate fuel oil reburning

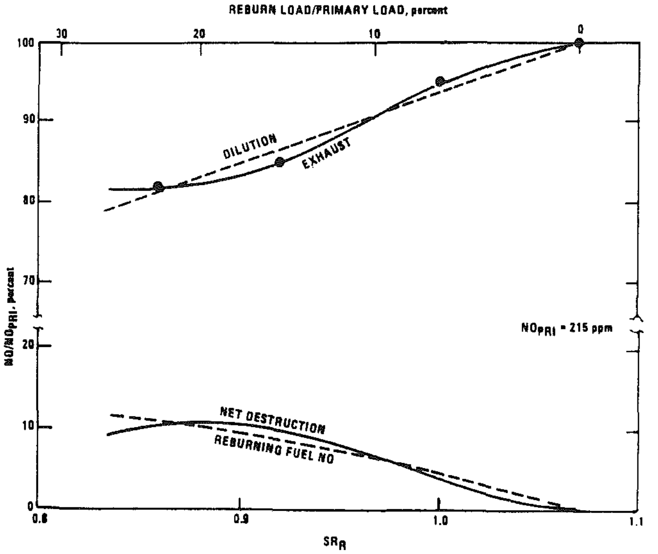


Fig. 6(*b*) Optimum stoichiometry for distillate/residual fuel oil mixture reburning

It has been shown in bench-scale tests [6], and verified in natural gas reburning tests in the PBS [15], that NO destruction is maximized for SR_R of about 0.9. This was found to be true of fuel oil reburning as well. This optimum is consistent with the minimum in the total fixed nitrogen (TFN) species equilibrium curve at the reburn zone temperature, although the experimental NO levels greatly exceed the equilibrium values under fuel-rich conditions.

Tests were conducted to compare fuel oil reburning effectiveness with natural gas reburning effectiveness. In these tests reburn zone bulk residence time was varied from 100 to 400 ms by varying the location of final air addition; the results are

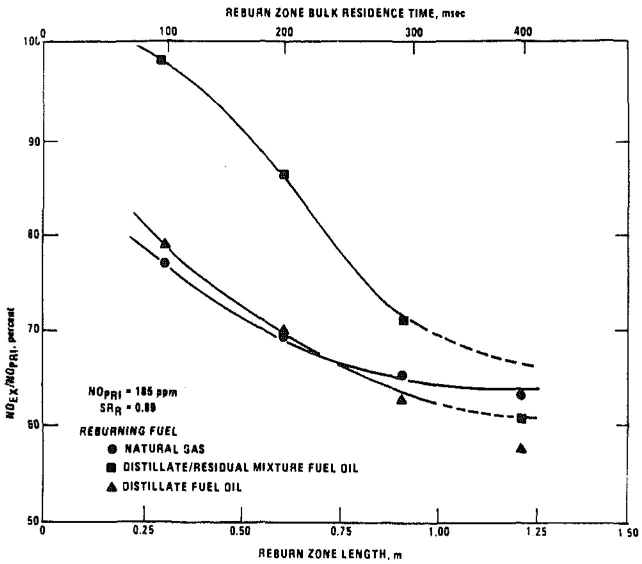


Fig. 7 Exhaust NO as a function of reburn zone residence time

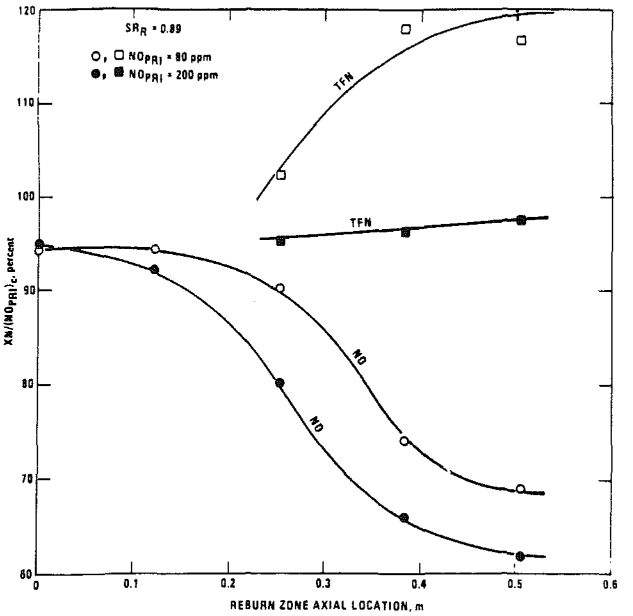


Fig. 8 Reburn zone nitrogenous species profiles for distillate fuel oil reburning at NO_{PRI} of 200 and 80 ppm

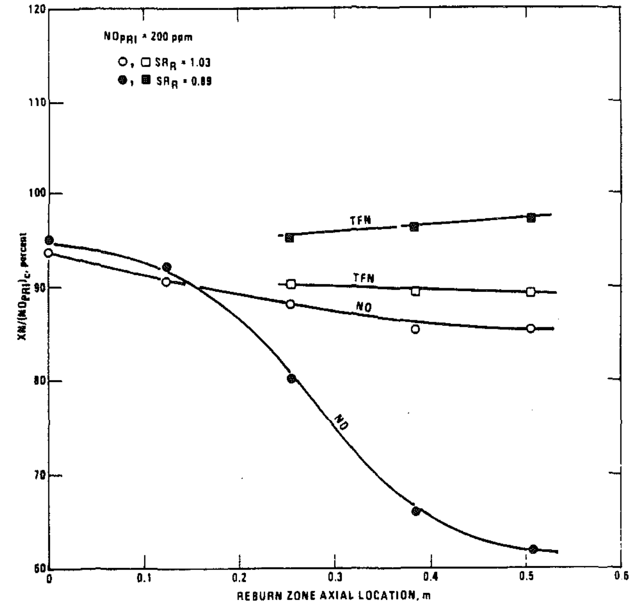


Fig. 9 Reburn zone nitrogenous species profiles for distillate fuel oil reburning at SR_R of 0.89 and 1.03

plotted in Fig. 7. At the nominal reburn zone length of 0.61 m (24 in.), natural gas reburning was slightly more effective than distillate fuel oil reburning and much more effective than distillate/residual fuel oil reburning. It was originally hypothesized that low nitrogen fuel oil reburning would be more effective than natural gas reburning due to the higher carbon-to-hydrogen ratio in fuel oils that would produce more methyldine radicals than would be produced in methane combustion. Figure 7 shows that, for high reburn zone residence times, distillate fuel oil reburning is slightly more effective than natural gas reburning at reducing NO. Oil reburning requires longer reburn zone lengths to allow for complete vaporization and combustion of the fuel droplets prior to ignition. For natural gas reburning, with the staged fuel injected circumferentially into the primary zone effluent to produce a uniform reburn zone stoichiometry, a fuel-rich zone length of 0.76 m (30 in.) was required to achieve maximum NO_x reduction. With a peak gas velocity in the reburn zone of 15 m/s (50 ft/s), as shown in Fig. 2, this suggests a characteristic time of 50 ms for gas phase reburning reactions to occur. For the distillate fuel oil, approximately 1.0 m (40 in.) reburn zone length was required to maximize NO_x reduction. For the heavy fuel oil mixture, about 1.2 m (48 in.) reburn zone length was required. Photographs of the reburning fuel oil spray confirm that the fuel oil droplets persist several inches into the reburning zone.

Before discussing in-flame measurements of nitrogenous species, a parameter that was investigated and found to have little impact on the reburning NO_x reductions is mentioned. The fuel oil nozzle spray angle was varied from 30 to 90 deg. In addition, various size nozzles were tested under identical conditions to investigate atomization effects. All of these nozzle variation tests yielded negligible effect on reburning effectiveness. Detailed measurements revealed that large-scale turbulent structures existed in the reburning zone. Apparently, the combustion gas turbulence intensity dominated reburn zone mixing rates, resulting in NO_x emissions independent of oil injection methods. This result, which simplifies application, may be specific to the firetube package boiler class.

In-Flame Measurement Tests. Detailed in-flame sampling was performed for a variety of fuel oil reburning test conditions (Figs. 8 and 9). NO, HCN, and NH_3 profiles in the reburn zone were measured, and total fixed nitrogen (TFN) was determined by summing the three measured nitrogen species. Time averaged axial combustion gas profiles were measured at a 13.3 cm (5.25 in.) radius. The species concentrations are ratioed to the primary NO corrected to reburn zone molar concentrations. These profiles were not time resolved as large radial gradients in stoichiometry were observed, as well as large-scale turbulent eddy structures. Thus, these trends are not representative of the entire boiler volume, but are useful in comparing different reburning conditions.

In Fig. 8 reburn zone profiles for distillate fuel oil reburning at a 0.89 stoichiometry were measured for 200 and 80 ppm NO_{PRI}. The curves show that reburn zone reactions are delayed until about the 0.2 m (8 in.) axial location. Natural gas reburning, on the other hand, resulted in immediate nitrogen species reactions. The curves also indicate that there is no reduction in TFN in the reburn zone for either test condition. In fact, for the 80 ppm NO_{PRI} condition, a significant net TFN formation occurred. It is hypothesized that this occurs via Fenimore's prompt NO mechanism (N_2 reaction with methyldine to form HCN), since the temperatures in the

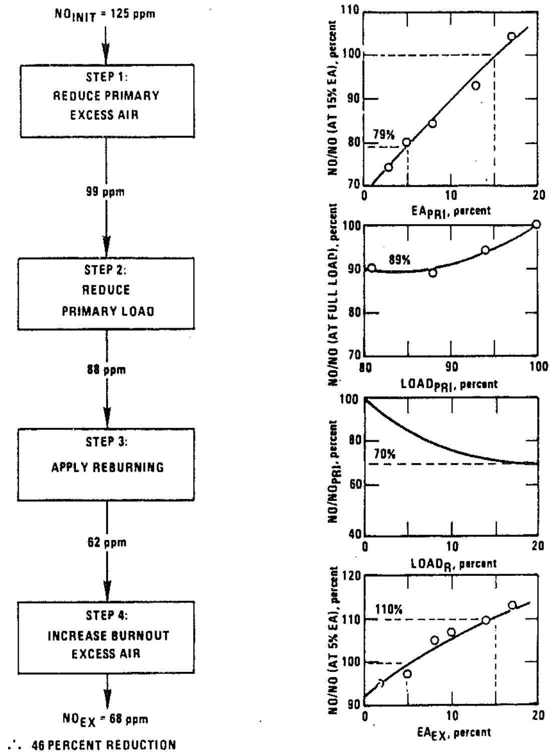


Fig. 10 Distillate fuel oil reburning applications

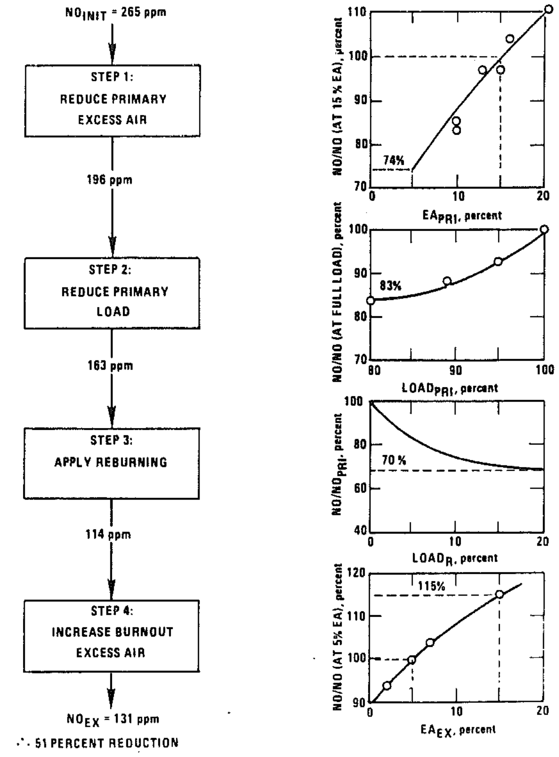


Fig. 11 Distillate/residual fuel oil mixture reburning application

reburn zone are well below the temperature necessary for Zeldovich reactions.

The difference between the TFN and NO curves represents the sum of the measured HCN and NH_3 concentrations. In all cases, NH_3 concentrations were less than 20 ppm. The low temperatures resulted in slow kinetics and no N_2 formation in the reburn zone. N_2 formation occurred in the burnout zone when the final air addition completed burnout of the hydrogen and carbon. In the reburn zone, the composition of the TFN was changed by the reaction

$$\mathrm{NO}+\mathrm{CH}\longrightarrow\mathrm{HCN}+\mathrm{O} \qquad (5)$$

The test conditions shown in Fig. 8 and other test conditions indicate that this reaction is limited by the presence of CH radicals and not NO_{PRI}. That is, the reburn zone NO curve is a function of SR_R but not of NO_{PRI}. However, the TFN curve does depend on NO_{PRI}. The net moles of N_2 destruction (by the Fenimore mechanism and reburning fuel nitrogen) in the reburn zone appears to be a relatively constant 20 ppm. Thus, at lower NO_{PRI} this becomes increasingly significant, and overall NO reduction is reduced.

In Fig. 9 reburn zone nitrogen species profiles are compared to two distillate fuel oil reburn loads: SR_R's of 0.89 and 1.03. Here it is shown that the reaction expressed by equation (5) is dependent on SR_R, with more NO being converted to HCN for more fuel-rich stoichiometries. Although the TFN is actually a greater percentage of the NO_{PRI} in the 0.89 SR_R condition, the composition of TFN is such that more N_2 formation occurs in the burnout zone, probably by the following reaction sequence

$$\mathrm{HCN}\xrightarrow{\mathrm{OH,\,H}}\mathrm{NH}\xrightarrow{\mathrm{NO}}\mathrm{N_2} \qquad (6)$$

In the 1.03 SR_R reburn condition, little HCN is formed in the reburn zone so little N_2 formation occurs. The only significant reduction of NO_{PRI} that is achieved under this condition is via dilution of the primary flue gases by the low NO_x reburning flame gases.

Thus, in-flame measurements of intermediate nitrogen species confirm bench-scale experiments that suggest a reburning mechanism. The qualitative argument presented here is based on typical data trends; however, the boiler volume is too large, and turbulent mixing too great, to provide a more quantitative analysis of the process.

Optimization Tests. The North American boiler was used to show how fuel oil reburning might be optimized in actual application. Baseline tests were performed to demonstrate that the North American boiler and PBS had similar thermal and aerodynamic environments in the radiant section. Reburning characterization tests were then performed to verify PBS test results. Finally, two optimized examples were performed and are presented here: distillate fuel oil reburning of a distillate fuel oil primary flame, and distillate/residual fuel oil reburning of a distillate/residual fuel oil flame. Constraints on these applications were to achieve good fuel burnout by minimizing fuel-rich zones and to compare conditions at full load, defined to be 650 kW (2.2×10^6 Btu/hr), and 15 percent excess air. These applications are shown in Figs. 10 and 11 in a stepwise fashion to demonstrate how the overall NO_x reduction via reburning is achieved.

Initial NO levels of 125 ppm (dry, at zero percent O_2) and 265 ppm were observed for distillate and distillate/residual fuel oil firing, respectively. Step 1 consists of reducing primary flame excess air from 15 to 5 percent at full load. This resulted in reduced NO emissions of 21 percent for distillate oil firing and 26 percent for distillate/residual fuel oil firing. In step 2 the primary burner load is dropped by 20 percent, maintaining constant excess air. Some reduction in NO was achieved here due to the reduced mixing intensities in the primary flame. In step 3 reburning is applied, bringing the boiler back to a full load condition. Here, primary NO is reduced by staged hydrocarbon injection. Finally, the overall excess is increased to 15 percent resulting in slightly increased NO emissions. It is significant to note that the effect of increasing excess air in the burnout zone is less severe than the effect of decreasing excess air in the primary zone, due to lower flame temperatures in the staged combustion.

The overall NO_x reduction for the distillate fuel oil example was 46 percent; the overall NO_x reduction for the distillate/residual fuel oil example was 51 percent. Thus, approximately 50 percent NO_x reduction can be achieved via fuel oil reburning in the firetube package boiler.

Conclusions

The results from the firetube package boiler fuel oil reburning application tests demonstrate that 50 percent NO_x reduction is achievable with minimal retrofit modification requirement. The parametric tests indicate that, as primary NO_x is decreased, such as by low NO_x burner design, NO_x destruction via reburning becomes increasingly difficult, perhaps requiring the use of low-nitrogen reburning fuels. There exists an optimum fuel-rich stoichiometry for reburning application of 0.7 to 1.0, dependent on system and fuel parameters. Distillate fuel oil reburning was shown to be slightly more effective than natural gas in reducing NO_x. However, fuel oil reburning requires longer reburn zone residence times than natural gas reburning due to droplet evaporation and mixing requirements. A minimum gas phase residence time of 50 ms is required to complete reducing zone reactions. In the firetube package boiler employed in this study, in which the peak combustion gas velocity in the boiler radiant section was 15 m/s, a reburn zone residence length of 1.0 m or greater was required to optimize fuel oil reburning.

In-flame measurements of intermediate nitrogen species were consistent with proposed reburning NO reduction mechanisms. In the low temperature environment of this application (1300-1600 K), actual N_2 formation occurred in the burnout zone; in the reburn zone, total fixed nitrogen speciation was shifted from all NO_x in the primary zone to NO, HCN, and NH_3. Unpublished test results indicate that, although higher temperature fuel oil reburning applications result in increased reburn zone N_2 formation, higher temperatures in the burnout zone result in less N_2 formation. Thus, the net effect of temperature on the overall reburning process is not as dramatic as was first expected.

In optimizing reburning application to a firetube package boiler, while maintaining fuel-lean conditions in the primary zone to avoid boiler wall corrosion and/or sooting, reductions of about 50 percent were achieved for 20 percent distillate and distillate/residual fuel oil reburning. The overall NO_x emission for distillate fuel oil firing was reduced from 125 to 68 ppm (dry, at zero percent O_2); for distillate/residual fuel oil firing the overall NO_x emission was reduced from 265 to 131 ppm (dry, at zero percent O_2).

Acknowledgments

This work was funded by EPA's Air and Energy Engineering Research Laboratory at Research Triangle Park, NC. The authors wish to express their appreciation for the support provided by E. E. Stephenson and C. Pendergraph of Acurex Corporation.

References

1 Patry, M., and Engel, G., *Compte Rendu Hebd. des Seances de l'Academie des Sciences*, Vol. 31, 1950, p. 1302.

2 Meyerson, A. L., et al., *Sixth Symposium (International) on Combustion*, The Combustion Institute, 1957, p. 154.

3 Reed, R. D., "Process for the Disposal of Nitrogen Oxide," John Zink Company, U. S. Patent No. 1274637, 1969.

4 Wendt, J. O. L., et al., *Fourteenth Symposium (International) on Combustion*, The Combustion Institute, 1973, p. 897.

5 Takahashi, Y., et al., "Development of Mitsubishi 'MACT' In-Furnace NO_x Removal Process," Presented at the U.S.-Japan NO_x Information Exchange, Tokyo, Japan, May 25-30, 1981; published in *Mitsubishi Heavy Industries, Ltd., Technical Review*, Vol. 18, No. 2.

6 Chen, S. L., et al., "NO_x Reduction by Reburning With Gas and Coal – Bench Scale Studies." *Proceedings of the 1982 Joint Symposium on Stationary Combustion NO_x Control*, Vol. 1, 1983, Paper No. 16.

7 Kelly, J. T., et al., "Fuel Staging for Pulverized Coal Furnace NO_x Control," *Proceedings of the 1982 Joint Symposium on Stationary Combustion NO_x Control*, Vol. 1, 1983, Paper No. 17.

8 Okigami, N., et al., "Three-Stage Pulverized Coal Combustion System for In-Furnace NO_x Reduction," *Proceedings of the 1985 Joint Symposium on Stationary Combustion NO_x Control*, May 7, 1985, Paper No. 4-5.

9 Miyamae, S., et al., "Evaluation of In-Furnace NO_x Reduction," *Proceedings of the 1985 Joint Symposium on Stationary Combustion NO_x Control*, May 7, 1985, Paper No. 4-4.

10 Murakami, N., "Application of the 'MACT' In-Furnace NO_x Removal Process Coupled With a Low-NO_x SGR Burner," *Proceedings of the 1985 Joint Symposium on Stationary Combustion NO_x Control*, May 8, 1985, Paper No. 5-6.

11 Seeker, R., et al. "Pilot-Scale Evaluation of Reburning for In-Furnace NO_x Reduction," *Proceedings of the 1985 Joint Symposium on Stationary Combustion NO_x Control*, May 7, 1985, Paper No. 4-1.

12 Brown, R., "Application of Reburning for NO_x Control in Cogeneration," *Proceedings of the 1985 Joint Symposium on Stationary Combustion NO_x Control*, May 7, 1985, Paper No. 4-6.

13 Yang, R. J., et al., "Laboratory Investigation of Advance Biased Firing for Industrial Processes," *Proceedings of the 1985 Joint Symposium on Stationary Combustion NO_x Control*, May 7, 1985, Paper No. 4-3.

14 Mulholland, J. A., and Lanier, W. S., "Application of Reburning for NO_x Control to a Firetube Package Boiler," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 107, 1985, pp. 739-743.

15 Lanier, W. S., *An Investigation of Chemical and Mixing Phenomena Associated With Reburning Applied to Firetube Package Boilers*, Ph.D. Dissertation, University of Virginia, 1984.

16 Mulholland, J. A., and Hall, R. E., "The Effect of Fuel Nitrogen in Reburning Application to a Firetube Package Boiler," *Proceedings of the 1985 Joint Symposium on Stationary Combustion NO_x Control*, May 7, 1985, Paper No. 4-2.

17 England, G., et al., "The Development of Design Criteria for Low NO_x Oil Burners Suitable for Application to TEOR Steamers," Air Pollution Control Association, 1982.

D. W. Anderson
Assoc. Mem. ASME

R. Viskanta
Fellow ASME

F. P. Incropera
Fellow ASME

Heat Transfer Laboratory,
School of Mechanical Engineering,
Purdue University,
West Lafayette, IN 47907

Effective Thermal Conductivity of Coal Ash Deposits at Moderate to High Temperatures

The effective thermal conductivity of coal ash deposits strongly influences heat transfer in pulverized coal-fired boilers. In this study thermal conductivity measurements were performed over a wide range of temperatures for fly ash, slagging deposits, and fouling deposits. The effects of ash particle size, thermal history, and physical structure of the deposit are discussed. Thermal history and deposit structure were observed to have the greatest influence on the local thermal conductivty, which increased by an order of magnitude with particle melting. Conductivities for solid-porous deposits were twice those of the same sample in particulate form.

Introduction

The thermal performance of coal-fired boilers is affected greatly by the deposition of coal ash particles on furnace wall steam tubes, superheater tubes, and other heat transfer surfaces. The reduction in heat absorption and the accumulation of low-thermal-conductivity ash deposits are known to influence temperature and heat flux distributions and to drastically reduce heat transfer rates [1-5]. Consequently, coal ash deposits have a major impact on the control (operation) and efficiency of a coal-fired unit.

Heat transfer through a porous coal ash deposit occurs by both conduction and thermal radiation. While the latter consists mainly of radiation transfer across voids, thermal radiation absorption, emission, and scattering can also be important at certain wavelengths for semitransparent slags. Depending on the temperature, physical state, and boundary conditions at the wall and deposit surface, the ratio of conductive to radiative heat fluxes will change with position in the deposit.

Since conductive and radiative heat transfer in the ash deposit occur both in parallel and in series, the concept of an "effective" thermal conductivity is used. This quantity k_e is defined such that the following form of the heat equation for one-dimensional, steady-state conduction is satisfied

$$\frac{dq}{dy} = -\frac{d}{dy}\left[k_e(T)\,\frac{dT}{dy}\right] = 0 \tag{1}$$

Since the heat flux is constant, the local effective thermal conductivity can be defined as

$$k_e(T) = -q/(dT/dy) \tag{2}$$

An overall effective thermal conductivity $k_{e,o}$, calculated for the entire ash deposit thickness, can be defined in a similar manner.

Heat transfer rates depend primarily on the effective thermal conductivity, surface absorptivity, and thickness of the deposit [6]. The deposit thickness, however, depends indirectly on the temperature distribution in the deposit, which in turn is governed by the effective thermal conductivity. The low thermal conductivities of ash deposits (typically less than 1.0 W/m K for sintered deposits) [7, 8] can produce temperature gradients of 100°C/mm or larger in the deposit. As a result, variations in effective thermal conductivity with temperature can be expected to be significant. While measurements of the overall effective thermal conductivity (or conductance) of ash deposits have been made [7, 9-12], measurements of local conductivities of coal ash at temperatures well above the sintering range are lacking.

An experimental facility has been constructed to study steady-state, high-temperature heat transfer through coal ash deposits obtained from utility boilers. The primary objectives of the study were: (1) to measure local thermal conductivities accurately over a large temperature range for three types of coal deposit (fly ash, slagging deposits, and fouling deposits), (2) to correlate experimental results and identify the parameters which significantly influence thermal conductivities, and (3) to compare results for solid-porous samples with those obtained from crushed deposits.

Experimental Facility and Procedures

The facility consists of a large-cavity kiln capable of continuous operation at 1500°C. Six silicon carbide heating electrodes irradiate a uniform and relatively thin ash deposit sample, as shown in Fig. 1. A water-cooled heat sink at the bottom of the test cell assembly produces a one-dimensional heat flux through the sample. Surface irradiation across the sample is measured with a water-cooled, circular-foil transducer. Irradiation levels were typical of those existing in coal-fired boilers.

Figure 2 shows details of the test cell assembly. Heat is con-

---

Contributed by the Power Division for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Power Division August 1985.

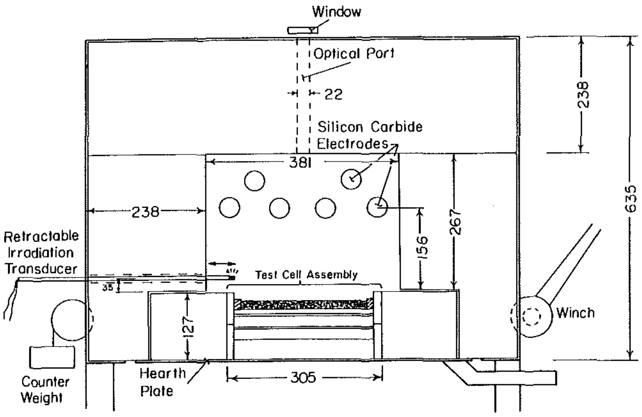


Fig. 1 Schematic of kiln interior (dimensions are in mm)

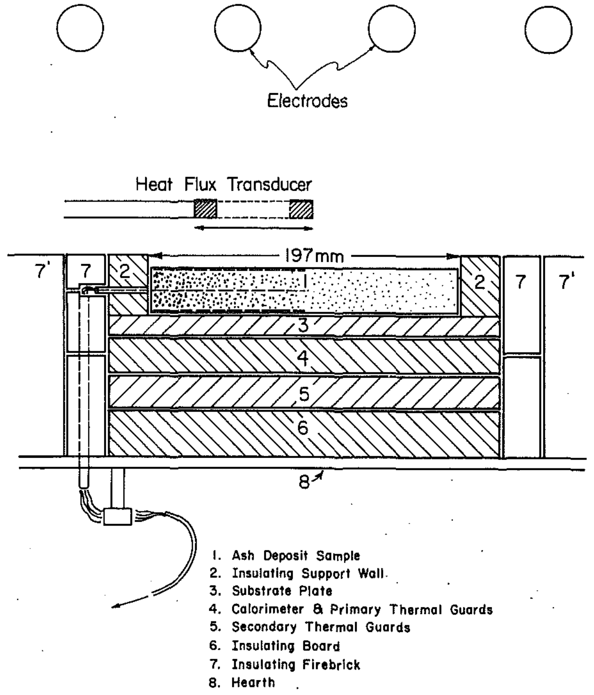


Fig. 2 Details of test cell assembly

ducted and radiated through the porous ash sample (1) and conducted through the thin, high-conductivity substrate plate (3). The heat is then transferred by radiation across an air gap which separates the substrate plate from the blackened top surface of the heat-flux calorimeter and primary thermal guards (4). The flux calorimeter (100 × 100 mm) and the surrounding thermal guards are water-cooled assemblies constructed of 25 mm copper plate with internal flow channels arranged in a serpentine fashion. The calorimeter is located directly below the center of the substrate plate and occupies 16 percent of the total primary layer area. Temperatures across the calorimeter surface were uniform to within 0.5°C. A set of secondary thermal guards (5), located below the calorimeter and primary guards, is used to isolate the calorimeter completely from any extraneous heat gains or losses.

Internal ash temperatures were measured with 0.13 mm (0.005 in.) "S" type platinum thermocouples in 1.2 mm (3/64 in.) o.d. alumina tubing. Accurate measurement of ash surface temperatures would have been difficult, especially for samples of particulate coal ash which have no well-defined surface. To circumvent this problem, the top and bottom thermocouple probes were positioned slightly below the top surface and slightly above the substrate plate, respectively. Using these and two other interior thermocouples, temperature profiles were determined and the surface temperatures were estimated by extrapolation.

Each experiment consisted of a heating cycle and a cooling cycle. The ash sample was first heated from 25°C to approximately 300°C (at the irradiated surface). After steady-state conditions were reached, data were collected and effective thermal conductivities were determined. The kiln temperature was then increased in increments of 100°C up to the desired maximum temperature, with data collected at each kiln temperature. This constituted a heating cycle. In the cooling cycle, thermal conductivities were determined in a similar fashion. The cooling cycle in each experiment lasted much longer than the heating cycle, since the kiln could cool only by natural convection from the kiln exterior and from the residual heat removed by the primary and secondary guards.

At each steady-state condition, the ash surface irradiation, the internal ash temperatures, and the heat flux through the ash sample were determined. Local thermal conductivities were calculated from the Fourier-Biot law, equation (2), where the gradient dT/dy was estimated from the temperature drop between two adjacent thermocouples. The heat flux was measured calorimetrically. Overall thermal conductivities, which represent an average conductivity for the entire sample thickness, were calculated from the following equation

$$k_{e,o} \equiv \int k(T)dT/\int dT = -qH/\Delta T \tag{3}$$

where ΔT is the difference between extrapolated surface temperatures. Alternatively, $k_{e,o}$ can be calculated using the top and bottom thermocouple temperatures and the corresponding separation distance. Results from these two methods agreed with one another to within 5 percent. The overall uncertainty in determining k_e was estimated to be less than 10 percent. More detailed descriptions of the facility, the experimental procedures, and the methods of data reduction are presented elsewhere [13].

Materials Studied

Thermal conductivities were measured for fly ash, crushed boiler-wall slag and fouling superheater deposits, and solid-porous fouling deposits. Hygroscopic moisture, specific gravity, and characteristic particle sizes were determined using American Society for Testing and Materials (ASTM) methods [14, 15]. Properties of the samples and relevant operating parameters are listed in Table 1.

With the exception of the fly ash samples, there was a

Nomenclature

A = area, m^2
D_p = particle diameter, m
H = ash sample thickness, m
k_e = local effective thermal conductivity, W/mK
$k_{e,o}$ = overall effective thermal conductivity, W/mK
q = heat flux, W/m^2
T = temperature, K
y = distance from substrate plate, m
ξ = dimensionless height = y/H

Subscripts

fc = heat flux calorimeter
h = hybrid
p = particulate
sp = solid porous
t = total

Table 1 Ash characteristics and experimental conditions

Experiment	Sample thickness,[a] mm	Particle diameter, μm	Specific gravity[b]	Bulk density,[a] g•cm^{-3}	Porosity,[a] percent	Silica ratio[c]	Maximum surface temperature, °C
A	19.0	25.5[d,e]	2.45	0.80	67.4	59.6	1150
B	20.3	13.5[d]	2.38	0.99	58.4	86.3	~1050[f]
C	20.3	715	2.33	0.67	71.5	58.5	1380
D	20.5	420	2.37	0.67	71.6	58.5	1155
E	20.5	420	2.37	0.67	71.6	58.6	1394
F[g]	19.7	—	2.58	0.80	69.1	80.0	1255
G[g]	19.7	—	2.58	0.80	69.1	80.0	1365
H	20.6	160	2.58	1.04	59.6	80.0	1235[f]
I	21.8	1015	2.58	0.95	63.2	58.5	1420

[a] Values prior to melting, where applicable.

[b] ASTM C-188 (1982).

[c] Silica ratio = $[(SiO_2)/(SiO_2 + Fe_2O_3 + CaO + MgO)]*100$ percent.

[d] ASTM D-422 (1982).

[e] Value for particles passing #200 mesh (84 w/o of sample).

[f] Useful data limited due to exposed thermocouple probes.

[g] Approximately 1/3 of sample was in particulate form with median particle diameter of ~200 μm. Where relevant, table entries are for solid pieces.

significant spatial variation in the chemical composition of the deposits. It was believed that a conventional mineral analysis (using x-ray fluorescence, for example) would not yield representative compositions for the samples. For this reason each sample was characterized by its average silica ratio. The silica ratios in Table 1 were provided by the utilities supplying the deposits. These are average values in the sense that they were obtained from the results of numerous mineral analyses performed over a period of several weeks.

The fly ash samples (Experiments A and B) were prepared by loading as-received fly ash into the test cell tray (Fig. 2) to a depth of approximately 20 mm. Care was taken to insure that the samples were not compacted during the loading process. Nonhomogeneities in the boiler-wall slag samples, such as their chemical composition and the size, shape, and distribution of voids, as well as machining difficulties, precluded studying these samples in their natural form. In order to obtain uniform samples, crushing the boiler-wall slag prior to testing was required. In order to investigate the influence of particle size on thermal conductivity, experiments were performed on three different particle sizes of the crushed samples. The particle sizes listed in Table 1 for Experiments C through I represented the mean characteristic particle size D_p.

While the thermal resistance of a collection of discrete slag particles is different from that of the original slag deposit, this difference diminishes considerably when the particles experience sintering and melting. Thus, at high temperatures the thermal conductivity of sintered or fused particles can be considered to be representative of that for the corresponding slag deposit, bonded to a steam-tube wall.

To assess objectively the effect of crushing deposits prior to heating, Experiments F and G were performed on a sample of solid-porous superheater deposit. Due to an insufficient amount of suitable material, preparation of a hybrid sample was necessary. Two thirds of this sample consisted of a solid-porous deposit 19.7 mm thick, with a 19.7 mm layer of crushed particles, approximately 160 μm in diameter, constituting the remaining third of the particulate sample. The heat flux calorimeter was located below part of the solid-porous deposit and part of the particulate sample. Since heat transfer through the solid and crushed regions was in parallel, the effective thermal conductivity of the solid-porous specimen could be determined. For comparative purposes, measurements were made on the same fouling deposit after it had been crushed into particles of approximately 160 μm diameter (Experiment H).

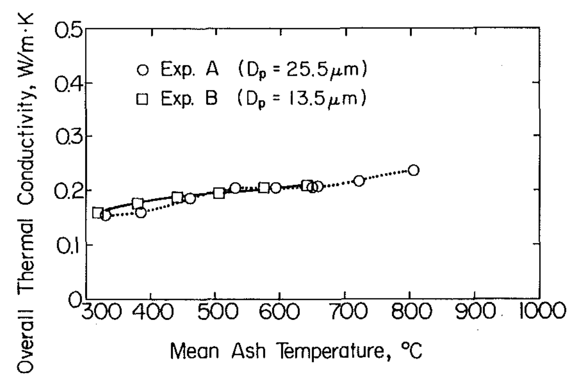


Fig. 3 Overall thermal conductivities for fly ash

Results and Discussion

Fly Ash. The overall effective thermal conductivities, based on the temperature drop across the entire sample thickness, are shown in Fig. 3 as a function of temperature. The magnitude of $k_{e,o}$ is quite small, indicating that the fly ash will create a large thermal resistance if it is allowed to accumulate on heat transfer surfaces.

A very light sintering of fly ash particles near the heated surface was produced at the higher temperatures. Local conductivities in this region were only a few percent greater than those in cooler regions, where sintering did not occur. That the overall effective thermal conductivity curve in Fig. 3 is reversible (Experiment A) suggests that the thin layer of lightly sintered ash particles did not significantly alter the effective conductivity of the ash sample.

Within experimental error, the overall effective conductivities for both types of fly ash are the same up to a mean ash temperature of approximately 650°C (see Fig. 3), despite the disparity in particle size, porosity, and the silica ratios listed in Table 1. (Experimental difficulties limited the useful data from Experiment B to the six points shown.) This agreement suggests that either the foregoing parameter effects cancel each other, or that the effects are negligible under these conditions. The likelihood of the first possibility is small, lending credance to the second alternative. Indeed, thermal radiation models for packed beds [16, 17] predict a negligible radiation contribution for 13.5 μm and 25.5 μm particles. Silica ratio is important only at higher temperatures [9], where strong sintering and fusion can occur. While porosity is important,

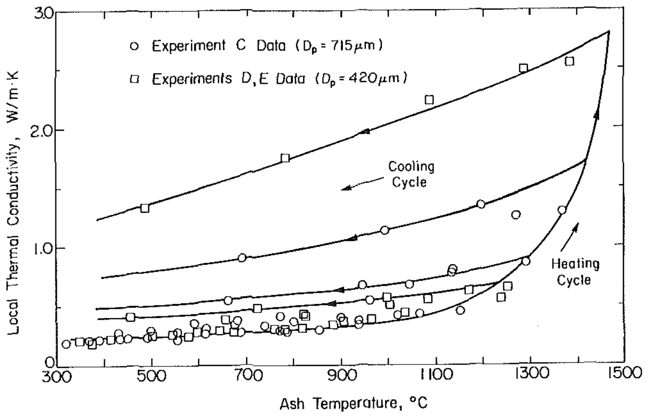


Fig. 4 Local thermal conductivities for crushed boiler-wall slag deposits with $D_p \le 715\ \mu$m

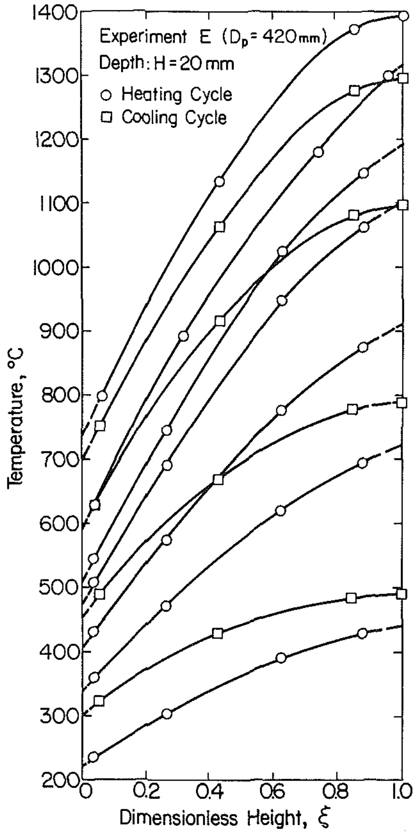


Fig. 5 Steady-state temperature profiles for crushed boiler-wall slag deposits showing effects of thermal history

regardless of particle size [18, 19], the particle size distribution can play a much larger role than porosity alone. The fly ash sample used in Experiment A consisted of particles 1.5 μm to 50 μm in diameter; fly ash particles in Experiment B ranged from less than 1 μm to 30 μm in diameter.

Crushed Slag Deposits. Experiments were performed on samples of crushed boiler wall slag for three different particle sizes: 420 ± 170 μm, 715 ± 125 μm, and 1015 ± 175 μm. Temperatures were sufficiently high to produce extensive sintering and melting, with corresponding decreases in sample thickness ranging from 23 to 33 percent. The maximum surface temperature of the melt was 1420°C.

Experiments D and E were performed on the same sample of boiler-wall slag (see Table 1) to check the reproducibility of the experimental results for a lightly sintered sample. Maximum ash temperatures in Experiment D were limited to 1155°C in order to preserve the sample for Experiment E. The sample was allowed to cool to 25°C prior to Experiment E. The effective thermal conductivities measured during the heating cycle in Experiment E agreed to within 10 percent with those measured during the cooling cycle of Experiment D.

Results from Experiments C, D, and E (Table 1) were used to generate the composite curve in Fig. 4. The local thermal conductivities k_e measured in the heating cycle are seen to be fairly independent of temperature below 1100°C, and increasing particle size from 420 μm (Experiments D and E) to 715 μm (Experiment C) did not significantly influence k_e. At higher temperatures, particle sintering and fusion become important. The increase in k_e with temperature is seen to be greater than tenfold (0.25 to 2.5 W/m•K).

The following empirical correlations for the local thermal conductivity in Fig. 4 (heating cycle) were obtained using least-squares curve-fitting methods. With k_e and T having units of W/mK and °C, respectively,

$T \le 1170°$C:

$$k_e = 0.105 + 2.99 \times 10^{-4} T \tag{4}$$

$T > 1170°$C:

$$k_e = 24.3 - 0.0416T + 1.81 \times 10^{-5} T^2 \tag{5}$$

These correlations have been used in modeling boiler-wall heat transfer for this particular slag [13]. However, in practice equation (4) should be modified to account for the loose particles which are predominant in the laboratory slag samples at relatively low temperatures. Since the actual boiler-wall deposit will at least be sintered, the minimum thermal conductivity will be closer to that of sintered material (~0.4 W/m-K in Fig. 4).

The most obvious trend in Fig. 4 is the irreversibility in k_e. Four heat transfer mechanisms govern overall heat transfer in a bed of loose particles in stagnant air [20]: (1) conduction through the solid particles, (2) radiation transfer between discrete particles not in contact, (3) radiation transfer between neighboring voids, and (4) conduction through the air near the contact points or surfaces. When individual particles become sintered or fused at high temperatures, the thermal resistance due to the fourth mechanism vanishes. In addition, heat conduction through the resulting solid matrix increases. These two effects produce a decrease in the total thermal resistance offered by the particles, i.e., an increase in the local thermal conductivity.

Another factor responsible for the increase in the effective thermal conductivity in Fig. 4 is that thermal radiation heat transfer is enhanced at the higher temperatures. Since radiation transfer depends on T^4, the curve shows a nonlinear temperature dependence at high temperatures. Packed-bed models [16, 20, 21] predict that the radiative contribution to k_e depends on T^3, but that the contribution is negligible for particles smaller than 100 μm. For spherical particles 1.0 mm in diameter with an emissivity of 0.8, radiation can become significant at temperatures of approximately 1000 to 1200 K. For smaller particles, radiation becomes significant at higher temperatures (Fig. 4). While the slag particles are certainly not spherical, packed bed heat transfer models are useful in explaining the observed thermal conductivity trends.

Figure 5 shows the effect of increased thermal conductivity on steady-state temperature distributions in the sample. The dimensionless height ξ is defined as $\xi = y/H$, where y is the distance from the substrate plate and H is the thickness of the ash sample. Each curve contains four ash temperatures measured at a single steady-state condition. The dashed line segments are the surface temperature extrapolations.

Near the cooled substrate plate (small ξ) the profiles are nearly linear, with temperature gradients of ~85 °C/mm. In

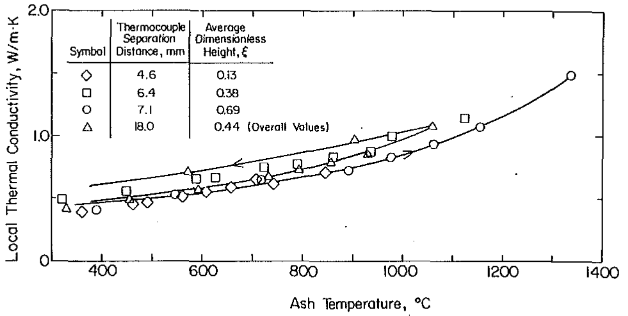


Fig. 6 Local thermal conductivities for slag deposits with D_p = 1015 μm (Experiment I)

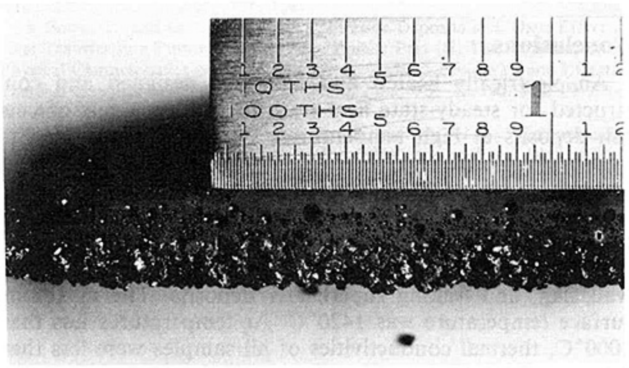


Fig. 7 Cross section of slag melt layer from Experiment I (1 in. scale)

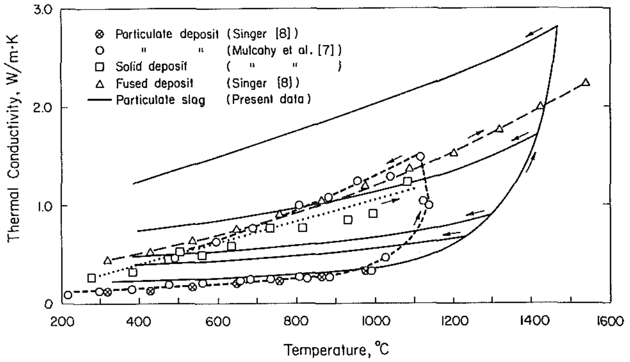


Fig. 8 Comparison of present data with those of previous experiments

contrast, the gradient is reduced near the heated surface, where k_e is largest. Temperature profiles are also greatly influenced by the thermal history of the sample itself, as clearly evidenced in the cool-down curves. The shift in dimensionless height of the thermocouple positions to larger ξ at elevated temperatures results from the decrease in sample thickness during melting.

Local effective thermal conductivities for the sample of 1015-μm-dia crushed slag particles (Experiment I) are shown in Fig. 6. Conductivities were approximately twice as large as those for particles with diameters smaller than 715 μm (see Fig. 4). The increase results from increased solid particle conduction and longer path lengths available for thermal radiation transfer. Quantitative predictions of k_e for various particle sizes using packed-bed heat transfer models are difficult due to the randomness of particle shape, the number of contact points, and the particle size distribution existing for the crushed slag samples. Comparing Fig. 6 (D_p = 1015 μm) with Fig. 4 (D_p < 715 μm), however, shows that crushing coal ash deposits into particles of arbitrary size can have a significant influence on effective thermal conductivity.

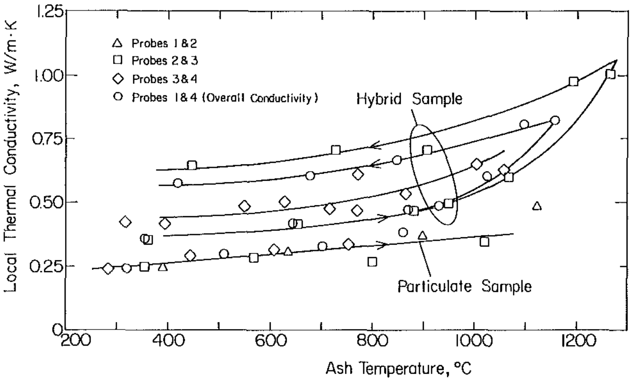


Fig. 9 Local thermal conductivities for hybrid and particulate fouling deposits

A cross section of the melt layer produced in Experiment I is shown in Fig. 7. Spherical voids, most of which ranged in size from 0.5 to 1.5 mm, are present throughout most of the hard melt region. Very few voids, however, are present in the top millimeter of the melt. Evidently the slag viscosity in this region was sufficiently low for the trapped voids to migrate to the surface and escape.

Results from the crushed slag experiments are compared with the data of Mulcahy et al. [7] and Singer [8] in Fig. 8. The present data are seen to agree quite well with previous data at temperatures below 1000°C. The deviation above 1000°C is believed to result from a slightly higher silica content in the present slag samples (46 percent) than that in Mulcahy's samples (41 percent). In addition, Mulcahy's data are for particles approximately 200 μm in diameter, whereas the particles used in Experiments C, D, and E had diameters of 420 and 715 μm. The smaller particles can be expected to sinter and melt sooner than larger particles for a given sample of material.

As expected, the curves for Singer's fused deposit and Mulcahy's solid deposit in Fig. 8 lie above the particulate-sample data. The thermal resistance of the particulate sample is greater than that of the solid sample due to the added contact resistance offered by the discrete particles. The magnitude of k_e for the solid and fused deposits in Fig. 8 depends primarily on the extent of sintering. The locations of the four cooling cycle curves (solid lines) reflect the relative degree of particle sintering and fusing.

Considering that the data in Fig. 8 are for different slag samples with different-sized particles, the close agreement for particulate samples at temperatures below 1000°C is encouraging. However, thermal conductivities at higher temperatures (above the sintering point) strongly influence heat transfer to the boiler steam tubes. The large discrepancy between the present data and those of Mulcahy et al. [7] at temperatures above 1000°C clearly indicates the importance of extending the measurement of thermal conductivities to temperatures above the sintering point.

Fouling Deposits. Local effective thermal conductivities measured for the hybrid sample of fouling superheater deposit (2/3 solid-porous, 1/3 particulate) in Experiments F and G and for the sample of crushed deposit (Experiment H) are shown in Fig. 9. Results for the hybrid sample exceed those for the particulate example by as much as a factor of two, indicating that the practice of routinely crushing samples of porous coal ash deposits can significantly bias effective thermal conductivities.

Results from Experiments F and G are not strictly valid for either a solid-porous deposit or for a particulate sample, since the heat flux calorimeter absorbed heat which was transferred through both materials. Results from Experiments F and G are therefore bounded by the two limiting cases. An estimate

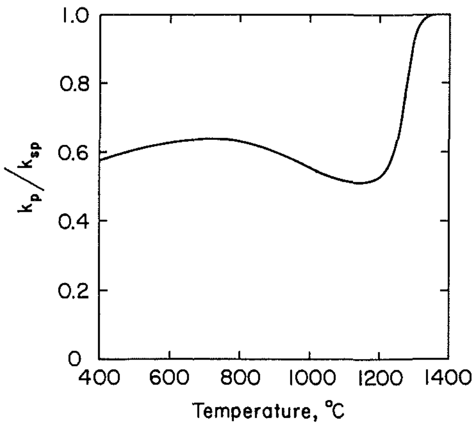


Fig. 10 Ratio of local thermal conductivities for crushed k_p and solid-porous k_{sp} fouling deposits

of the solid deposit conductivity can be made, however, since the heat transfer paths are in parallel. The thermal conductance for the hybrid sample is the sum of the two values

$$\frac{k_h A_{fc}}{H} = \frac{k_{sp} A_{sp}}{H} + \frac{k_p A_p}{H} \tag{6}$$

or

$$k_{sp} = \frac{A_{fc}}{A_{sp}} k_h - \frac{A_p}{A_{sp}} k_p \tag{7}$$

where the subscript *sp* denotes solid-porous, *h* the hybrid sample (Experiments F and G), and *p* the particulate sample (Experiment H). The areas A_{sp} and A_p represent the surface areas of the flux calorimeter below the solid-porous specimen (80 percent) and the particulate sample (20 percent), respectively. By applying equation (7) to the data of Fig. 9, estimates for k_{sp} as a function of temperature were made. Results are shown in Fig. 10, where the ratio of particulate sample conductivity k_p to solid-porous conductivity k_{sp} is plotted. This ratio clearly shows that crushing the fouling deposit reduces the effective thermal conductivity at moderate temperatures ($T < 1200°\text{C}$). (The curve in Fig. 10 was drawn assuming that the thermal conductivities of both the particulate sample and the solid-porous samples would be equal at 1350°C, i.e., when partial melt conditions exist.) The error in assuming that thermal conductivities measured for the crushed sample are representative of those for the actual solid-porous fouling deposit is approximately 100 percent at moderate temperatures ($k_p \approx 0.5\ k_{sp}$ at 1100°C).

While the curve in Fig. 10 was obtained for one specific fouling deposit, it is expected that similar trends will hold in general. The common practice of measuring thermal conductivities at moderate temperatures for crushed coal ash deposits in loose particle form appears to be of questionable value. Efforts should be directed toward measuring thermophysical properties of deposits in their natural form when possible. The difference between solid and particulate sample effective thermal conductivity will depend on the chemical composition, degree of sintering, and the experimental procedures used.

Packed-Bed Heat Transfer Models

Attempts to correlate the effective thermal conductivities shown in the previous figures using packed-bed heat transfer models were not successful. The main difficulties in applying packed-bed models to coal ash deposits include satisfying the model assumptions and obtaining a spatially uniform ash deposit sample with known properties.

Of the many correlations appearing in the literature, most are restricted to constant-property materials having uniform particle size, uniform packing, and porosities which are well below those of the ash samples used in this study. In addition, unknown ash properties such as the solid-phase thermal conductivity and surface emissivity, as well as local variations in porosity, particle shape, particle size, particle-to-particle spacing, and the number of contact points per particle are difficult to account for.

For these reasons, a detailed comparison of the experimental results with predictions from packed-bed correlations was not considered to be practical. Nonetheless, the packed-bed heat transfer models surveyed [16–21] were useful in explaining, in a qualitative sense, the observed trends in k_e for the fly ash, fouling, and slagging deposits.

Conclusions

An electrically heated kiln has been designed and constructed for steady-state heat transfer measurements of coal ash deposits at high temperature. The facility was used to measure the "effective" thermal conductivities of several ash samples under conditions similar to those existing in a large-scale utility boiler. Both overall and local thermal conductivities were measured.

Experiments were performed on samples of fly ash, boiler-wall slag, and fouling superheater deposits. The maximum surface temperature was 1420°C. At temperatures less than 1000°C, thermal conductivities of all samples were less than 0.5 W/m-K and were relatively independent of temperature. Temperature gradients in the unsintered regions of the samples approached 100°C/mm. At higher temperatures, conductivities increased significantly and irreversibly where particle sintering and fusion occurred. Local thermal conductivities for crushed slag samples increased tenfold upon melting.

Particle size of crushed deposits was not an important parameter for $D_p \leq 715\ \mu\text{m}$. For particles larger than 715 μm in diameter, the effective thermal conductivities increased with particle size. Conductivities at temperatures less than 1200°C for 1015 μm particles were approximately 100 percent higher than those for particles smaller than 715 μm.

Both the thermal history and structural form of the deposits (solid-porous versus crushed samples) were important parameters. Thermal conductivities in sintered or fused regions were considerably larger than those in unsintered regions at a given temperature. Crushing a sample of fouling superheater deposit reduced k_e by 50 percent at moderate temperatures.

Results for crushed slag deposits compared favorably with previously existing data at temperatures below 1000°C. Discrepancies above 1000°C clearly indicated the need to extend the measurement of thermal conductivities to temperatures well above the sintering range and into the melting range.

Acknowledgments

This work was supported, in part, by the U.S. Department of Energy, Pittsburgh Energy Technology Center, under Grant DE-FG22-81PC40789. One of the authors (D.W.A.) is grateful for the financial support provided through the Fluor Foundation Fellowship and the Donnan Scholarship.

References

1 McDonald, C. R., Spink, C. D., Barratt, D. J., and Cowley, L. T., "A New Pilot-Scale Combustor for the Evaluation of Pulverized Coal Combustion Performance," *Proceedings, Third Engineering Foundation Conference on Slagging and Fouling Due to Impurities in Combustion Gases*, Copper Mountain, CO, July 29–August 3, 1984, in press.

2 Reid, W. T., *External Corrosion and Deposits*, Elsevier, New York, 1971.

3 Wall, T. F., Lowe, A., Wibberley, L. J., and Stewart, I. M., "Mineral Matter in Coal and the Thermal Performance of Large Boilers," *Progress in Energy and Combustion Science*, Vol. 5, 1979, pp. 1-29.

4 Abraham, K. U., and Rajaram, S., "Measurements of Furnance Heat Transfer Performance in a Corner-Fired, Pulverized Coal Boiler," *Journal Institute of Energy*, Vol. LVI, No. 429, 1983, p. 217-225.

5 Bryers, R. W., ed., *Ash Deposits and Corrosion Due to Impurities in Combustion Gases*, Hemisphere, Washington, DC, 1977.

6 Smith, J. D., Smith, P. J., Smoot, L. D., and Rees, D. P., "Part I. Prediction of the Effects of Coal Quality on Utility Furnace Performance," *Proceedings, Third Engineering Foundation Conference on Slagging and Fouling Due to Impurities in Combustion Gases*, Copper Mountain, CO, July 29-August, 3, 1984, in press.

7 Mulcahy, M. F. R., Boow, J.,and Goard, P. R. C., "Fireside Deposits and Their Effect on Heat Transfer in a Pulverized-Fuel-Fired Boiler: Part I. The Radiant Emittance and Effective Thermal Conductance of the Deposits," *Journal Institute of Fuel*, Vol. 39, No. 308, 1966, pp. 385-394.

8 Singer, J. G., ed., *Combustion Fossil Power Systems*, Combustion Engineering, Inc., Windsor, CT, 1981.

9 Boow, J., and Goard, P. R. C., "Fireside Deposits and Their Effect on Heat Transfer in a Pulverized-Fuel-Fired Boiler: Part III. The Influence of the Physical Characteristics of the Deposit on Its Radiant Emissivity and Effective Thermal Conductivity," *Journal Institute of Fuel*, Vol. 42, No. 346, 1969, pp. 412-419.

10 Karasina, E. S., Abryutin, A. A., and Efimenko, A. N., "Determination of Heat Transfer Resistance of Slag and Ash Deposits on Panel Walls With a Water Cooled Heat Flow Meter," *Thermal Engineering*, Vol. 26, No. 1, 1979, pp. 27-30.

11 Fetters, G. D., Viskanta, R., and Incropera, F. P., "Experimental Study of Heat Transfer Through Coal-Ash Deposits," ASME Paper No. 82-WA/HT-30, 1982.

12 Abryutin, A. A., and Karasina, E. S., "Thermal Conductivity and Thermal Resistance of Ash Deposits in Boiler Furnaces," *Thermal Engineering*, Vol. 12, 1970, pp. 46-50.

13 Anderson, D. W., "Heat Transfer Through Coal Ash Deposits," Ph.D. Thesis, Purdue University, West Lafayette, IN, 1985.

14 American Society for Testing and Materials, Designation C188-78, "Density of Hydraulic Cement," Annual Book of ASTM Standards, Part 13, 1982, pp. 194-196.

15 American Society for Testing and Materials, Designation D422-62, "Particle-Size Analysis of Solids," Annual Book of ASTM Standards, Part 19, 1982, pp. 112-122.

16 Kuzay, T. M., "Effective Thermal Conductivity of Porous Solid-Gas Mixtures," ASME Paper No. 80-WA/HT-63, 1980.

17 Balakrishnan, A., and Pei, D., "Heat Transfer in Gas-Solid Packed Bed Systems, Part 2. The Conduction Mode," *Industrial Engineering and Chemistry Process Design and Development*, Vol. 18, 1979, pp. 40-46.

18 Imura, S., and Takegoshi, E., "Effect of Gas Pressure on the Effective Thermal Conductivity of Packed Beds," *Heat Transfer—Japanese Research*, Vol. 3, 1974, pp. 13-26.

19 Vortmeyer, D., "Radiation in Packed Solids,'" *Heat Transfer—1978*, Hemisphere, Washington, DC, 1978, Vol. 8, pp. 525-539.

20 Yagi, S., and Kunii, D., "Studies on Effective Thermal Conductivities in Packed Beds," *AIChE Journal*, Vol. 3, 1957, pp. 373-381.

21 Schotte, W., "Thermal Conductivity of Packed Beds," *AIChE Journal*, Vol. 6, 1960, pp. 63-67.

C. W. Somerton
Assistant Professor,
Department of Mechanical Engineering,
Michigan State University,
East Lansing, MI 48824

T. Brouillette[1]

C. Pourciau

D. Strawn

L. Whitehouse

Undergraduate Research Assistants,
Mechanical Engineering Department,
Louisiana State University,
Baton Rouge, LA 70803

RANKINE: a Computer Software Package for the Analysis and Design of Steam Power Generating Units

A software package has been developed for the analysis of steam power systems. Twenty-eight configurations are considered, all based upon the simple Rankine cycle with various additional components such as feedwater heaters and reheat legs. The package is demonstrated by two examples. In the first, the optimum operating conditions for a simple reheat cycle are determined by using the program. The second example involves calculating the exergetic efficiency of an actual steam power system.

Introduction

Nearly all major electric-generating power systems run through a series of processes which is based upon the theoretical thermodynamic cycle called the Rankine cycle. Because the Rankine cycle is the basis for steam power cycles, its analysis is vital to the design of power systems and to ensuring their proper operation. By introducing additional components, such as feedwater heaters or reheat legs, or adjusting the operating temperatures and pressures, a Rankine cycle analysis can be used to maximize the thermal efficiency and exergetic efficiency (second law effectiveness) within the constraints of turbine exit quality and pump capacity. The process of analyzing these Rankine cycles and attempts at optimizing the operating conditions can be very complex as well as time consuming. Such an analysis dictates the use of the computer. A software package (RANKINE) has been designed and tested which will solve a wide range of Rankine cycle problems very quickly. This software package is the subject of this paper.

Another use of the software package is in the instruction of undergraduate mechanical engineering students. Because the Rankine cycle is the basis for steam power cycles, it is a primary topic covered in an intermediate thermodynamics course. In order to provide the students with a feeling for how the ideal Rankine cycle may be modified in order to model any actual steam power cycle, a large number of examples need to be worked. With the RANKINE software package, the students will be able to gain a feeling for actual steam power cycles in a reasonable time period.

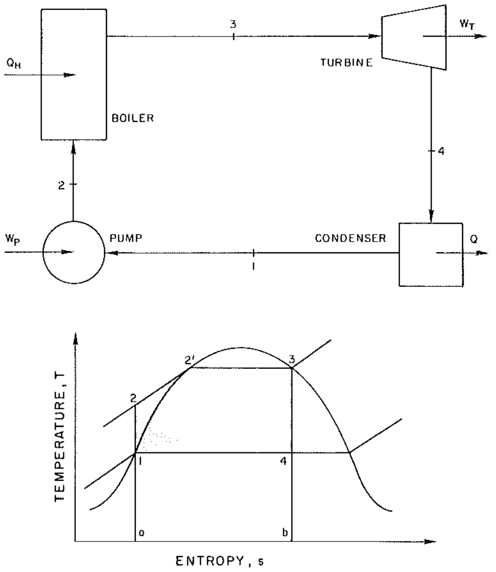


Fig. 1 The simple Rankine cycle

Previous software packages dealing with Rankine or modified Rankine cycles fall into one of three groups. There have been some very simple programs developed which can be run on hand-held calculators (Whitehouse, 1984). These are normally restricted to considering only the simple Rankine cycle and hence do not provide a realistic analysis of the power plant. The second group involves software packages that run on micro- or minicomputers. Preston et al. (1973) have written a program for the PDP-8 computer which analyzes the operation of the Rankine cycle for classroom use. Unfortunately, this second group is also restricted to only the simplest steam power cycles such as the Rankine or reheat cycle since the primary use of these programs is for classroom instruction. Finally, there are some massive computer programs available for the analysis of actual steam power plants, but these are

---

[1] Current address: Dow Chemical, Plaquemine, LA.

Contributed by the Power Division for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Power Division October 1, 1985.

often very costly to run and are designed for a specific power plant. Among these codes are Pepsi, SNYTHA, and Therm (Babin, 1985). Recently, Rosen and Scott (1985) have modified the Aspen Plus code for energy–exergy analysis. A modular technique was employed by Sonnenschein (1982) to perform calculations involving power station energy balances and plant efficiency. The RANKINE program is sufficiently flexible to analyze a number of different power plants and is also sufficiently simple to be used for instructional purposes. The Rankine software package was developed by the authors from scratch using a cycle solution methodology employed by one of the authors in the classroom.

This paper continues by providing a short theoretical background on the Rankine cycle, followed by a detailed description of the designed computer program. This will include an outline of the overall structure of the program, details of each subroutine, the operating structure of the program (with appropriate flow charts), and a discussion on the acquisition of the physical properties of the working fluid (steam). The paper concludes by employing the software package to optimize the thermal efficiency of a simple reheat cycle subject to a constraint on the exit quality of the turbines. Finally, the software package is demonstrated by conducting a second law analysis on a more complicated steam power system.

Theoretical Background

The theoretical Rankine cycle is shown in Fig. 1 in a schematic diagram as well as a temperature–entropy diagram. In the ideal Rankine cycle the process from 1–2 represents a reversible isentropic pumping process, from 2–3 is a constant pressure transfer of heat in the boiler, from 3–4 is a reversible isentropic expansion in the turbine, and from 4–1 is a constant-pressure transfer of heat in the condenser. The temperature–entropy diagram shows the possible states at each node but due to friction or heat losses these states may not be as the diagram indicates. The area denoted by a-2-2′-3-b-a indicates the heat transferred to the working fluid Q_H (steam in this case) and the area denoted by a-1-4-b-a indicates the heat transferred from the fluid Q_L. The net work is represented by the difference in the heat added and the heat rejected, or the portion of the diagram which is shaded.

The major components of the Rankine cycle are:

(*i*) *Boiler*—converts the working fluid of the system from the liquid state into a vapor state through heat addition.

(*ii*) *Turbine*—converts the available energy of the steam into mechanical energy by an expansion process.

(*iii*) *Condenser*—converts the working fluid of the system from the vapor state into a liquid state with rejection of heat.

(*iv*) *Pump*—raises the pressure of working fluid in liquid state.

(*v*) *Feedwater Heater*—mixes extracted stream from the turbine with the liquid from the condenser.

In modifying the simple theoretical Rankine cycle for steam power use, a reheat leg is often employed. This uses steam extracted off the turbine, passed back through the boiler and finally expanded through a second turbine. The reheat cycle is used to increase efficiency with higher pressures but avoids high moisture in the turbine.

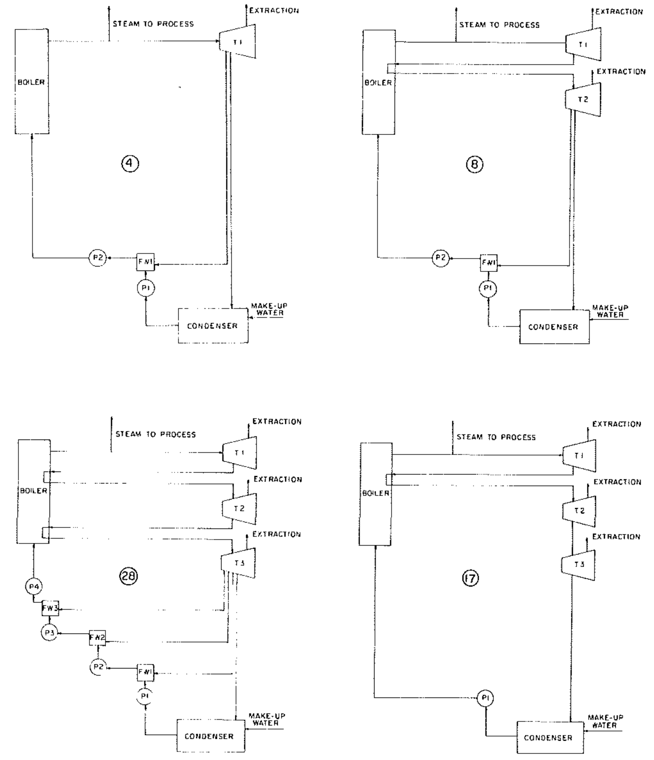


Fig. 2 Examples of cycle configurations

When a modified Rankine system is used to model an actual steam power cycle a number of deviations from ideal operation must be incorporated. The principal losses considered in the software package are:

(*i*) turbine losses
(*ii*) pump losses
(*iii*) condenser losses
(*iv*) piping losses, both pressure losses and heat losses

These deviations from ideal performance are all incorporated into the program using the method outlined by Van Wylen and Sonntag (1974). Shaft and valve leakages are not accounted for in the program.

Description of the Software Package

The main program of the software package has been designed to analyze 28 different cycle configurations. Some of these configurations are shown in Fig. 2 and range in complexity from the simple Rankine cycle to a steam power cycle that includes three turbines, two reheat legs, and three feedwater heaters. Some generality of the program is lost by allowing a maximum of only three feedwater heaters, since the standard in today's large control station power cycles is seven feedwater heaters. For all the configurations extraction of process steam has been accounted for with makeup water being added at the condenser. For each cycle, the exit conditions of the boiler and the exit pressure of the turbine(s) are required inputs.

Once the cycle to be analyzed has been selected, a noding of the cycle is carried out. Beginning with the condenser, a node is identified and numbered after every device. The piping between components is considered a device so that node 1 is at

Nomenclature

E = exergy
h = enthalpy
L = fluid state index
P = pressure
P_I^* = dimensionless intermediate pressure
Q = quality
Q_H = heat added
Q_L = heat rejected
s = entropy
T = temperature
W = work
ϵ = exergetic efficiency

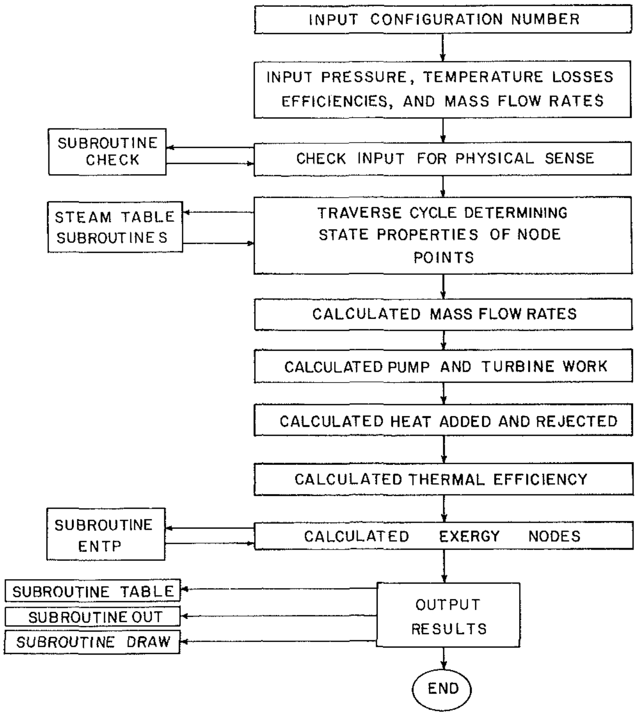


Fig. 3 Flow chart of RANKINE

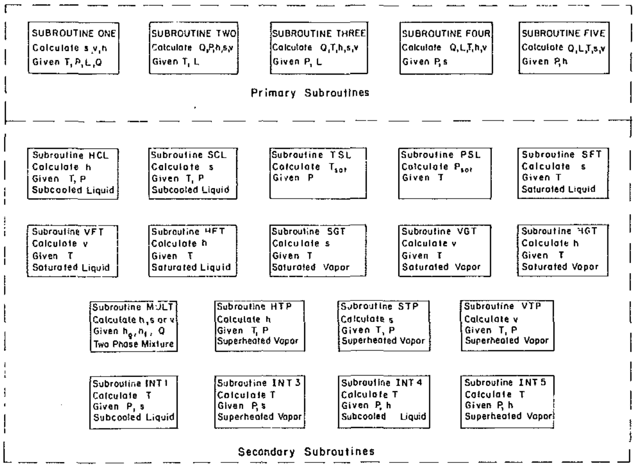


Fig. 4 Steam table subroutines

SUBROUTINE FIVE

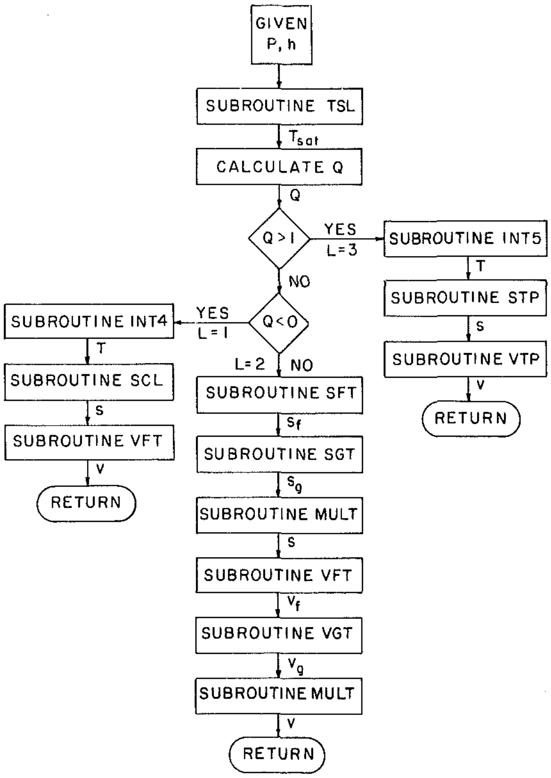


Fig. 5 Flow chart of subroutine FIVE

the exit of the condenser, but node 2 is at the entrance to the first pump. This is, of course, the exit of the "pipe device." After the nodes have been numbered, the cycle is traversed, using the appropriate thermodynamic relations to obtain the state variables at the next node. At this time the required inputs of temperature T, pressure P, quality Q, and fluid state L are made. The fluid state index L takes on the values of 1 for subcooled liquid, 2 for two-phase mixture, and 3 for superheat vapor. Once all of the state variables are specified at each node the calculations for net work, thermal efficiency, and exergy can be performed. The results are all printed out in tabular form and the corresponding temperature–entropy diagram is drawn. Once all of the state variables are specified at each node the calculations for net work, thermal efficiency, and exergy can be performed. The results are all printed out in tabular form and the corresponding temperature–entropy diagram is drawn.

A key element in determining the state variables at each node is employing the appropriate equation of state or fundamental relation for the working fluid. The equation of state for steam is represented by the steam tables. The steam table routines of RANKINE were programmed from the equation derived in the Keenan and Keyes edition of the *Steam Tables* (1947).

A flow chart of the overall operation of the software package is shown in Fig. 3. In Fig. 4 a description of the Steam Table Subroutines is provided. The interaction of the Steam Table Subroutines is somewhat unique and an example is given in the flow chart of Fig. 5.

Although this software package is multipurpose, it does have certain limitations. The numerical methods used to calculate the steam table properties introduce considerable errors whenever the pressure is greater than the critical pressure. Therefore, the maximum allowable input pressure is 22.09 MPa. This prohibits use of the program to simulate systems with supercritical boilers. The maximum allowable temperature input is 1300°C. The software package also fixes the state at the exit of the condenser and exit of a feedwater heater to be saturated liquid. The input values for pressure and heat losses can be either positive or negative, but the pressure loss cannot exceed the pressure at the entrance of the pipe and the heat loss cannot exceed the enthalpy at the entrance of the pipe. Finally only open feedwater heaters can be incorporated, as closed feedwater heaters which require iterative solution schemes cannot be solved with RANKINE.

Demonstration of the Software Package

To demonstrate the usefulness of RANKINE for power plant design, we first consider the simple reheat cycle specified as configuration #7 and shown in Fig. 6. The design problem faced in considering the reheat cycle is to select the intermediate pressure, that between the turbines, so that the thermal efficiency η_{th} is maximized while the exit quality of

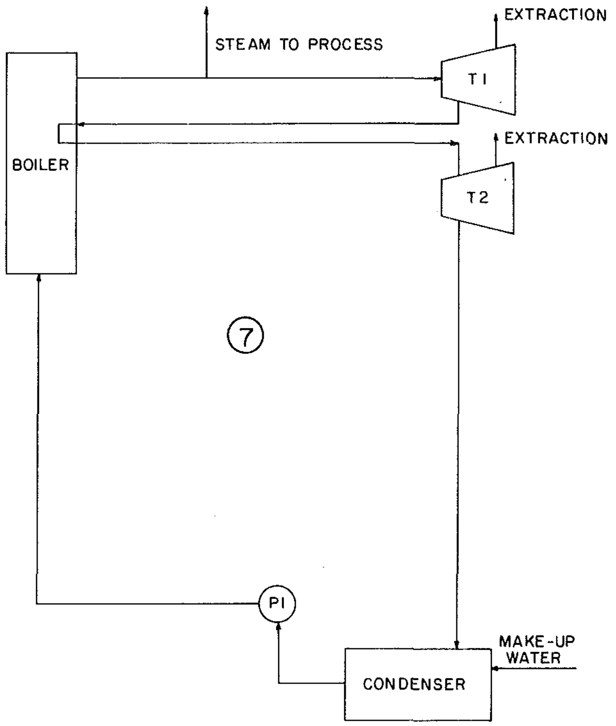


Fig. 6 Simple reheat cycle

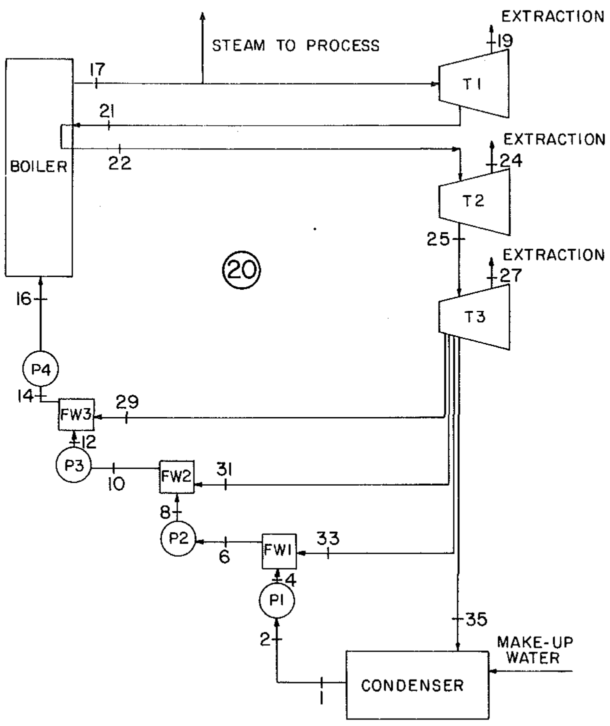


Fig. 8 Actual steam power system

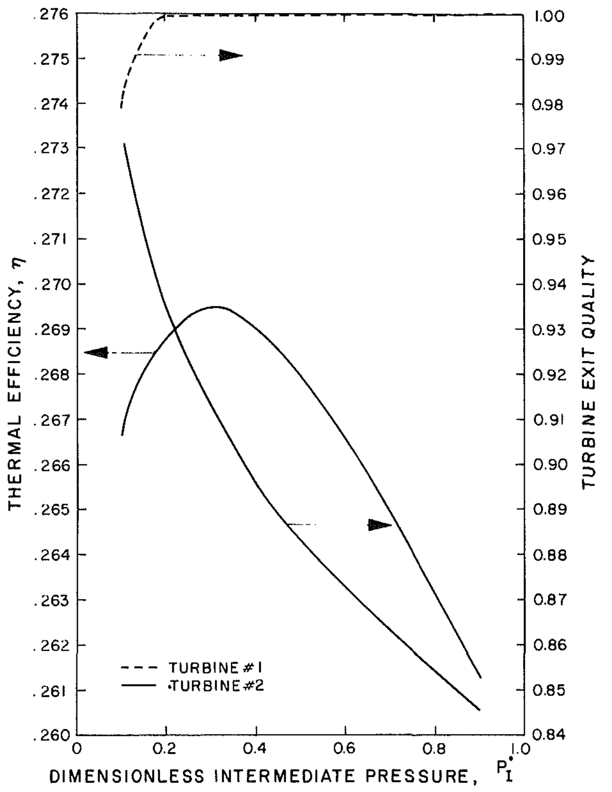


Fig. 7 Thermal efficiency and turbine exit qualities for reheat cycles

the turbine is maintained above 0.9. This optimization can be done very easily and quickly using RANKINE. To simulate actual conditions, pump and turbine efficiencies are taken to be 85 percent. Further, it is assumed that the temperature at the exit of the reheat leg is 80 percent of that at the exit of boiler. The boiler exit conditions are set at 2 MPa and 100°C superheat and the condenser pressure is 0.01 MPa. A dimensionless intermediate pressure is defined as

$$P_I^* = \frac{P_{\text{Intermediate}} - P_{\text{Condenser}}}{P_{\text{Boiler}} - P_{\text{Condenser}}} \tag{1}$$

and the computer program is run over a range of values of P_I^*. The results of these computer runs are shown in Fig. 7. The computer time required for these runs is very small with the program residing on a main-frame computer such as an IBM3033 or a VAX machine.

The first use of Fig. 7 is to determine the range of P_I^* for which the constraint of turbine exit quality greater than 0.9 is satisfied for both turbines. From the graph this range is seen to be 0.1 to 0.37. Now restricting our attention to this range of P_I^*, we seek to maximize the thermal efficiency. Fortunately the global maximum of η_{th} occurs within this range of $P_I^* = 0.315$. This is in fair agreement with the rule of thumb used for optimum reheat pressure being 20–25 percent of throttle for large single reheat power plant cycles. For the given pressures of the boiler and condenser this translates to an actual pressure of 0.637 MPa. By operating the reheat cycle at this intermediate pressure its maximum efficiency will be achieved.

Another question which may arise involving the reheat cycle concerns the reheat leg exit temperature. Normally, this temperature is some fraction of the boiler exit temperature (80–85 percent), the standard being 1000°F maximum reheat temperature. In the example just presented it was taken as 80 percent. A question one may well ask is whether the improvement in the thermal efficiency is worth the effort to raise this percentage to 85 or 90 percent. To increase this percentage usually involves a capital expenditure, such as an additional heat exchanger pass. The RANKINE package can be used for decision-making purposes via cost-worth analyses.

Finally, the RANKINE software package can be used to conduct a second law analysis of a power system. To demonstrate this analysis, configuration #20, shown in Fig. 8, is employed, which does a reasonable job of modeling the power plant schematic given by Babcock and Wilcox in their handbook (1978).

```
                              ***** CONFIGURATION 20 *****
NODE        T(C)      P(MPA)    L        Q      S(KJ/KG K) H(KJ/KG) V(KG/M**3) EXERGY(KJ/KG)
  1         32.9      0.0050    1      *****     0.4763     137.88   0.0010        0.6
  2         32.9      0.0050    1      *****     0.4763     137.88   0.0010        0.6
  3         32.9      0.0170    1      *****     0.4763     137.89   0.0010        0.6
  4         32.9      0.0170    1      *****     0.4763     137.89   0.0010        0.6
  5         56.6      0.0170    1      *****     0.7876     236.78   0.0010        6.7
  6         56.6      0.0170    1      *****     0.7876     236.78   0.0010        6.7
  7         56.6      0.0660    1      *****     0.7876     236.84   0.0010        6.8
  8         56.6      0.0660    1      *****     0.7876     236.84   0.0010        6.8
  9         88.4      0.0660    1      *****     1.1732     370.07   0.0010       25.1
 10         88.4      0.0660    1      *****     1.1732     370.07   0.0010       25.1
 11         88.4      0.1830    1      *****     1.1733     370.22   0.0010       25.2
 12         88.4      0.1830    1      *****     1.1733     370.22   0.0010       25.2
 13        117.4      0.1830    1      *****     1.4991     492.65   0.0011       50.6
 14        117.4      0.1830    1      *****     1.4991     492.65   0.0011       50.6
 15        120.2     22.0900    1      *****     1.5026     519.92   0.0011       76.8
 16        120.2     22.0900    1      *****     1.5026     519.92   0.0011       76.8
 17        538.0     22.0900    3      *****     6.2290    3338.04   0.0143     1486.4
 18        538.0     22.0900    3      *****     6.2290    3338.04   0.0143     1486.4
 19        538.0     22.0900    3      *****     6.2290    3338.04   0.0143     1486.4
 20        289.4      3.7200    3      *****     6.3564    2940.93   0.0621     1051.4
 21        289.4      3.7200    3      *****     6.3564    2940.93   0.0621     1051.4
 22        538.0      3.7200    3      *****     7.2327    3532.45   0.0982     1381.8
 23        538.0      3.7200    3      *****     7.2327    3532.45   0.0982     1381.8
 24        538.0      3.7200    3      *****     7.2327    3532.45   0.0982     1381.8
 25        284.4      0.5050    3      *****     7.3977    3031.52   0.5027      831.7
 26        284.4      0.5050    3      *****     7.3977    3031.52   0.5027      831.7
 27        538.0      0.5050    3      *****     8.1818    3562.75   0.7389     1129.2
 28        182.6      0.1830    3      *****     7.4746    2836.65   1.1369      613.9
 29        182.6      0.1830    3      *****     7.4746    2836.65   1.1369      613.9
 30         98.0      0.0660    3      *****     7.5521    2676.70   2.5686      430.8
 31         98.0      0.0660    3      *****     7.5521    2676.70   2.5686      430.8

 32         56.6      0.0170    2      0.956     7.6472    2498.98   8.5152      224.8
 33         56.6      0.0170    2      0.956     7.6472    2498.98   8.5152      224.8
 34         32.9      0.0050    2      0.916     7.7287    2357.65  25.8099       59.2
 35         32.9      0.0050    2      0.916     7.7287    2357.65  25.8099       59.2
THE WORK OF PUMP          1IS -1.2215804E-02(KW)
THE WORK OF PUMP          2IS -5.2589539E-02(KW)
THE WORK OF PUMP          3IS -0.1353554    (KW)
THE WORK OF PUMP          4IS  -27.26770    (KW)
THE WORK OF TURBINE          1IS   397.1057    (KW)
THE WORK OF TURBINE          2IS   500.9348    (KW)
THE WORK OF TURBINE          3IS   628.2141    (KW)
THE NET WORK IS    1498.787     (KW)
THE HEAT ADDED IS    3409.638     (KW)
THE HEAT REJECTED IS  -1910.851     (KW)
THE THERMAL EFICIENCY IS  0.4395736
```

Fig. 9 Computer results for Configuration 20

The results of the computer run are shown in Fig. 9. The authors define an exergetic efficiency as

$$\epsilon = \frac{\text{exergy and work produced}}{\text{exergy and work used}} \tag{2}$$

This is a somewhat different efficiency than the second law effectiveness defined by Bejan (1982) and Paolino and Burghardt (1982). The advantage of using the exergetic efficiency as defined by the authors in steam power cycle analyses is that the exergy used in delivering process or extracted steam can be easily incorporated. Exergy and work produced connote the outputs of the cycle such as process and extracted steam and turbine work. Exergy and work used include the exergy supplied from the boiler and reheat legs. For the problem being considered we have

$$\epsilon = \frac{W_{T1} + W_{T2} + W_{T3}}{(\Delta E)_{\text{boiler}} + W_{p1} + W_{p2} + W_{p3} + W_{p4} + (\Delta E)_{\text{reheat}}} \tag{3}$$

The terms $(\Delta E)_{\text{boiler}}$ and $(\Delta E)_{\text{reheat}}$ represent the exergy change of the steam as it passes through the boiler or reheat leg. That is,

$$(\Delta E) = h_{\text{out}} - h_{\text{in}} - T_0(s_{\text{out}} - s_{\text{in}}) \tag{4}$$

where T_0 is taken at 298 K. For the calculation of exergy at a node the definition is used

$$E = h - h_0 - T_0(s - s_0) \tag{5}$$

where h_0 and s_0 are taken for saturated liquid at 298 K. From Fig. 9 we can thus determine the exergetic efficiency as

$$\epsilon = \frac{397.11 + 500.93 + 628.21}{1409.6 + 0.0122 + 0.0526 + 0.135 + 27.268 + 330.4} = 0.864 \tag{6}$$

Conclusions

A software package called RANKINE has been developed to analyze 28 different steam power cycle configurations. The package allows for inefficiencies in the pumps and turbines and accounts for pressure losses and heat losses in the piping.

Results are provided in terms of the thermodynamic state properties before and after each device, the work of the pumps and turbines, heat added to and rejected by the cycle, and the overall thermal efficiency of the cycle. Along with these results, drawings of the cycle and the temperature-entropy diagram are also provided as output.

The software package has been shown to be useful in power plant design and analysis through calculations performed for two cases. The package may also be used for classroom demonstration.

Authors' Note

For educational purposes, the authors are willing to make the RANKINE software package available to interested parties at no charge. The program is currently running on a VAX/VMS system in Fortran 77. The KERMIT file transfer program will be used to transfer the package to a diskette for either IBM or DEC compatible machines.

References

Babcock and Wilcox, 1978, *Steam/Its Generation and Use*, 39th ed., Babcock and Wilcox, New York.

Babin, T., 1985, Personal Communication.

Bejan, A., 1982, *Entropy Generation Through Heat and Fluid Flow*, Wiley, New York.

Keenan, J. H., and Keyes, F. G., 1947, *Thermodynamic Properties of Steam*, Wiley, New York.

Paolino, M. A., and Burghardt, M. D., 1982, "Energy Conservation and Second Law Efficiency," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 104, pp. 241-246.

Preston, R., Harmer, D. S., Carlson, R. W., and Wrege, D. E., 1973, "Minicomputer Program for Evaluating Rankine Cycles Using Steam Tables," *DECUS Proceedings*, San Francisco, pp. 279-282.

Rosen, M. A., and Scott, D. S., 1985, "The Enhancement of a Process Simulator for Complete Energy-Exergy Analysis," AES-Vol. 1, ASME Winter Annual Meeting, Miami Beach, FL.

Sonnenschein, H., 1982, "A Modular Optimizing Calculation of Power Station Energy Balance and Plant Efficiency," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 104, pp. 255-259.

Van Wylen, G. J., and Sonntag, R. E., 1978, *Fundamentals of Classical Thermodynamics*, 2nd ed., Wiley, New York.

Whitehouse, G. D., 1984, Personal Communication.

Exergy Analysis of Combined Cycles: Part 1—Air-Cooled Brayton-Cycle Gas Turbines

M. A. El-Masri
Associate Professor,
Department of Mechanical Engineering,
Massachusetts Institute of Technology,
Cambridge, MA 02139

Quantitative analytical tools based on the second law of thermodynamics provide insight into the complex optimization tradeoffs encountered in the design of a combined cycle. These tools are especially valuable when considering approaches beyond the existing body of experience, whether in cycle configuration or in gas turbine cooling technology. A framework for such analysis was provided by the author [1-3] using simplified, constant-property models. In this paper, this theme is developed to include actual chemical and thermodynamic properties as well as relevant practical design details reflecting current engineering practice. The second-law model is applied to calculate and provide a detailed breakdown of the sources of inefficiency of a combined cycle. Stage-by-stage turbine cooling flow and loss analysis calculations are performed using the GASCAN program and examples of the resulting loss breakdowns presented. It is shown that the dominant interaction governing the variation of cycle efficiency with turbine inlet temperature is that between combustion irreversibility and turbine cooling losses. Compressor and pressure-drop losses are shown to be relatively small. A detailed analysis and loss breakdown of the steam bottoming cycle is presented in Part 2 of this paper.

Introduction

Combined cycle power plants are gaining increasing acceptance as alternatives to conventional or nuclear steam systems. This is leading to the development of gas turbines dedicated to combined cycle applications, such as the AGTJ-100 "Moonlight" project in Japan [4-6]. Traditionally, the power-generation gas turbine has been built to operate on the Brayton cycle with thermodynamic design constraints originating from its airborne counterpart. Thus the cycle has consisted of one adiabatic compression followed by combustion and expansion through an air-cooled turbine. While heavy-duty utility gas turbines differ from aircraft engines and their derivatives in many design features, the basic thermodynamic processes outlined above are identical. Those processes are not necessarily optimum for power-generation gas turbine systems for a given set of system performance requirements and component performance characteristics. An optimum system for a given power generation duty may involve alternate cycle configurations such as compressor intercooling, turbine reheat [4-9], steam injection into the gas turbine [10], steam cooling of turbine components [11, 12], use of gas turbine regenerators, integration with alternative bottoming cycles [13], or a variety of other options and combinations thereof.

Assessment and optimization of such complex thermal systems is a complicated task. The traditional first-law cycle analysis based upon component performance characteristics coupled with energy balances can invariably lead to the correct final result. However, such analysis cannot locate and quantify the sources of loss which lead to that result. This is because the first law embodies no distinction between work and heat, no provision for quantifying the quality of heat, and no accounting for the work lost in adiabatically throttling a pressurized gas. These inherent limitations of energy-balance analysis are not a serious drawback in the hands of experienced engineers dealing with familiar systems, since an intuitive understanding of the different parametric influences on system performance and a second-law qualitative appreciation of "grade of heat" and effect of pressure loss are developed. When analyzing novel and complex thermal systems, however, such experience needs to be supplemented by more rigorous quantitative analytical tools. Second-law analysis provides those tools.

One reason second-law analysis has not gained widespread engineering acceptance is thought to lie in the fact that it has not been readily integrable into traditional first-law techniques. The time-honored, practical definition of power plant efficiency as the ratio of work output to heat input

$$\eta_1 = \frac{W_{\text{out}}}{Q} = \frac{W_{\text{out}}}{m_f \Delta H_r} \tag{1}$$

differs from the second-law definition

$$\eta_2 = \frac{W_{\text{out}}}{\Omega_{\text{in}}} = \frac{W_{\text{out}}}{m_f \Delta G_r} \tag{2}$$

Contributed by the Gas Turbine Division and presented at the Joint ASME/IEEE Power Generation Conference, Portland, Oregon, October 19-23, 1986. Manuscript received by the Gas Turbine Division July 31, 1986. Paper No. 86-JPGC-GT-9.

since for most fuels the enthalpy of reaction (or heating value) differs slightly from the availability[1] (or Gibbs free energy of reaction). In a previous paper [1], the author has circumvented this inconsistency by a hybrid methodology where a work/exergy balance of air-standard constant-property uncooled gas turbine cycles was drawn up with exergy losses expressed as fractions of the heat addition. By proposing a simplified model for the combustion irreversibility, wherein the exergy lost as a fraction of the heat added in combustion was expressed as the complement to the Carnot efficiency based upon the logarithmic-mean temperature of heat addition, the work/exergy-loss balances summed up to the heat addition. Thus the efficiency definition of equation (1) was preserved with the balance of the heat unconverted to work being quantitatively assigned to the irreversibility mechanisms responsible for its rejection. This is in contrast to the energy balance that assigns unconverted heat to the points of its rejection rather than to the causes therefor.

In second-law analysis of gas turbine systems, models for the turbine cooling irreversibilities are required. The author has presented a simplified constant-property general model for the thermodynamics of cooled turbines in [2] and applied it to a wide-ranging comparative assessment of alternative cycles and cooling technologies in [3]. In the present work, the models presented in [1-3] are developed to account for actual variable chemical/thermodynamic properties, rather than constant-property air-standard cycles, and to incorporate stage-by-stage turbine cooling and thermodynamic analysis to a level of detail suitable for absolute performance calculations rather than the comparative results presented in [1-3]. Those models are applied to a Brayton cycle air-cooled gas turbine in simple and combined cycle modes, with detailed analysis of the bottoming steam cycle presented in Part 2 of this paper. The exergy balances for the cycle and its components are presented and compared to energy balances. The loss mechanisms in combustion and gas turbine cooling are identified, quantified, and broken down into their sources by component and by thermodynamic process. Emphasis is placed throughout on realistic component modeling based upon current technological constraints.

Analysis

Figure 1(*a*) shows a flow diagram for the cycle under consideration and its thermodynamic state points are illustrated on temperature-entropy coordinates in Fig. 1(*b*). Ambient air is assumed to be at 1.013 bar, 289 K, 50 percent humidity. The fuel is assumed to be methane. All gaseous mixtures appearing in the calculations can be treated as mixtures of varying proportions of three components. Those three components are: air (77.44 percent N_2, 20.76 percent O_2, 0.92 percent Ar, 0.85 percent H_2O, 0.03 percent CO_2); water (100 percent H_2O); and stoichiometric gases[2] (70.16 percent N_2, 19.58 percent H_2O, 9.43 percent CO_2, 0.83 percent Ar). Polynomial fits for the specific heats of each of those three components as a function of temperature are used in the calculations and plotted in Fig. A1 of Appendix I. The specific enthalpy, entropy, and availability of all substances are treated as the sum of a chemical component and a thermomechanical component. All nonreacting gases[3] are arbitrarily assigned zero thermomechanical enthalpy, entropy, and availability at the ambient pressure and temperature regardless of chemical composition, thereby neglecting the entropy of mixing different gaseous components. Thus the enthalpy, entropy, and exergy of each nonreacting component can be calculated from the specific heat polynomials[4]

$$h=\int_{T_a}^{T} C_p(T)dT \tag{3}$$

$$\phi=\int_{T_a}^{T} C_p \frac{dT}{T} \tag{4}$$

$$s=\phi-R\ln(p/p_a) \tag{5}$$

$$\omega=h-T_a s=h-T_a\phi+RT_a\ln(p/p_a) \tag{6}$$

It is convenient to think of the first two terms on the right-hand side of equation (6) as the sensible component of the exergy and the last term as the pressure component. From the above relations and the polynomial fits for $C_p(T)$, the enthalpy and exergy of all streams can be calculated as a function of their pressure, temperature, and chemical composition. The latter state variables can be calculated from the cycle parameters and component performance characteristics as discussed below. An interactive computer program, GASCAN (GAS turbine Cycle ANalysis), has been written to set up the gas turbine cycle, select the compressor bleed and turbine-coolant source configuration, perform mass, energy, and exergy balances on each component, compute turbine cooling

---

[1] The terms "exergy," "availability," and "available work" are synonymous and shall be used interchangeably.

[2] It is assumed that a fraction η_Q of the fuel is completely burned and that the small unburned portion has the same specific heat as stoichiometric gases.

[3] The only reacting component is taken to be CH_4, but unburned methane resulting from incomplete combustion is treated as nonreacting.

[4] In the current terminology, T and p refer to stagnation values. The property polynomials are based upon static temperatures. In calculations where the Mach number is sufficiently large to introduce appreciable property errors, correction to the static state is made before the polynomials are applied.

Nomenclature

C_p = specific heat at constant pressure, kJ/kgK
E = effectiveness or second-law efficiency
h = specific stagnation enthalpy, kJ/kg
m = mass flow rate, kg/s
p = stagnation pressure, bar
Q = heat flow rate, kW
R = gas constant, kJ/kgK
s = specific entropy, kJ/kgK
T = stagnation temperature, K
W = work rate, kW
ΔH_r = enthalpy of reaction (lower heating value)
ΔG_r = Gibbs function of reaction, kJ/kg
η = efficiency
η_1 = first-law efficiency
η_2 = second-law efficiency
ϕ = Keenan function, kJ/kgK, defined in equation (4)
Ω = availability (or exergy) rate, kW
Ω^+ = increase in sensible exergy through combustion, kW
ω = specific stagnation exergy, kJ/kg

Subscripts

a = ambient air
c = compressor
f = fuel
i = index indicating a compressor bleed stream
j = index indicating a cooling stream entering a turbine stage
L = lost
LM = logarithmic mean
Q = combustor
TS = turbine stage
1, 2, 3, 4 = cycle state points, Fig. 1

Superscripts

$+$ = added

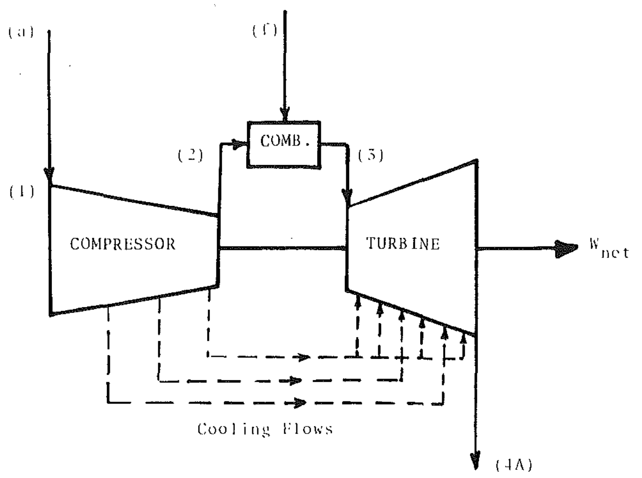


Fig. 1(*a*)

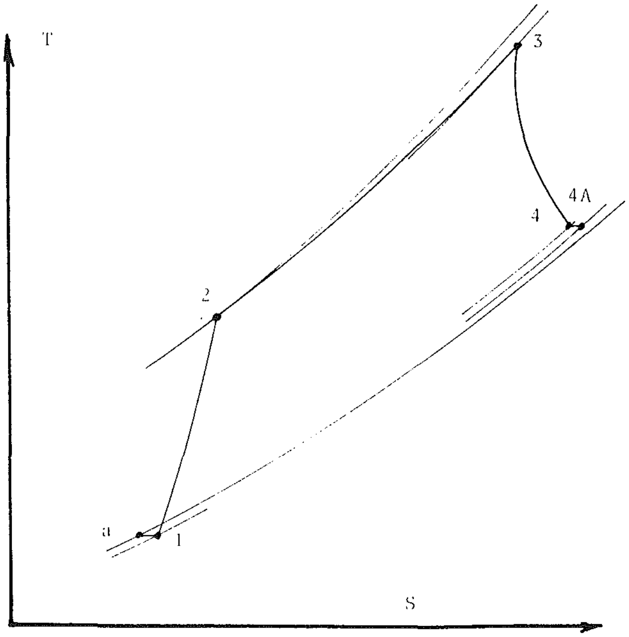


Fig. 1(*b*)

flow rates and performance penalties for each component, and conduct a turbine pitchline thermodynamic analysis. The program methodology is given in [14]. The calculations presented in this paper are based upon the inputs described in the paper or given in Appendix I. Appendix II shows mass, energy, and exergy streams in an example cycle with a pressure ratio of 14 and combustor discharge temperature (T.I.T.) of 1550 K. Reference to this table should facilitate the discussions of each component given below. All tabulated results refer to this sample cycle.

Compressor. For a given polytropic efficiency, pressure ratio, and bleed point pressures, the temperature and thus other state variables for each stream can be calculated. Mass and energy balances yield the compressor work, and the availability balance for the compressor gives the exergy lost

$$m_1 = m_2 + \Sigma m_i \tag{7}$$

$$W_c = m_2 h_2 + \Sigma m_i h_i \tag{8}$$

$$\Omega_{L,c} = W_c + m_1\omega_1 - m_2\omega_2 - \Sigma m_i \omega_i \tag{9}$$

It should be noted that state 1 has zero enthalpy but negative exergy due to the total pressure drop across the inlet filter/ductwork. The compressor effectiveness defined as the ratio of increased availability of the working substance to work input

$$E_c = (m_2\omega_2 + \Sigma m_i\omega_i - m_1\omega_1)/W_c \tag{10}$$

has been related to the polytropic efficiency in [1].

Combustor. For a given discharge temperature, pressure drop, and combustion efficiency, the fuel flow is found from the mass and energy balances, and the availability lost from the exergy balance. Those are, respectively,

$$m_3 = m_2 + m_f \tag{11}$$

$$\eta_Q m_f \Delta H_r = m_3 h_3 - m_2 h_2 \tag{12}$$

$$\Omega_{L,Q} = m_f[\Delta G_r + R_f T_a \ln (p_f/p_a)] - (m_3\omega_3 - m_2\omega_2) \tag{13}$$

In equations (11) and (12), the fuel line is assumed to be at a pressure p_f and ambient temperature. The first term on the right of equation (12) includes the thermomechanical availability of the pressurized fuel. Some rearrangement of equation (13), together with equations (6), (11), and (12), allows the availability lost in the combustor to be expressed as

$$\Omega_{L,Q} = \Omega_{L,Q,FT} + \Omega_{L,Q,IC} + \Omega_{L,Q,PD} + \Omega_{L,Q,TD} \tag{13a}$$

where the terms on the right-hand side respectively represent the availability lost due to throttling the fuel from p_f to p_2, incomplete combustion, total flow combustor pressure drop in the combustor from p_2 to p_3, and thermal degradation of the heat released by combustion to a finite temperature. Those quantities are

$$\Omega_{L,Q,FT} = m_f R_f T_a \ln (P_f/P_2) \tag{13b}$$

$$\Omega_{L,Q,IC} = m_f(1-\eta_Q)\Delta G_r \tag{13c}$$

$$\Omega_{L,Q,PD} = m_3 R_3 T_a \ln (P_2/P_3) \tag{13d}$$

$$\Omega_{L,Q,TD} = m_f \eta_Q (\Delta G_r - \Delta H_r) + T_a(m_3\phi_3 - m_2\phi_2) \tag{13e}$$

It is particularly interesting to form the ratio of (13*e*) to the heat-added quantity of equation (12). Such a manipulation gives

$$\frac{\Omega_{L,Q,TD}}{\eta_Q m_f \Delta H_r} = \frac{T_a(\phi_3-\phi_2)}{h_3-h_2} + \left(\frac{\Delta G_r}{\Delta H_r} - 1\right) + 0\left(\frac{m_f}{m_3}\right) \tag{14}$$

where the last term represents miscellaneous terms of order (m_f/m_3), a small number in practical cycles. The next to last term is very small for most fuels. Thus the ratio of availability lost to heat added in the combustor can be approximated by the first term on the right. Neglecting property variations, this term can be written, in view of equations (3) and (4), as

$$\frac{T_a(\phi_3-\phi_2)}{h_3-h_2} = \frac{\ln (T_3/T_2)}{T_3-T_2} = \frac{T_a}{T_{LM}} \tag{15}$$

where T_{LM} is the logarithmic mean temperature in the combustor. Thus equation (15) can be recognized as the complement of the Carnot efficiency based on T_{LM}. The heat-addition effectiveness of a combustor defined as the increase of sensible thermal availability of the working fluid divided by the heat supplied is therefore reasonably approximated by the Carnot efficiency based upon T_{LM}

$$E_Q^{+} = 1 - \frac{\Omega_{L,Q,TD}}{\eta_Q m_f \Delta H_r} \cong 1 - \frac{T_a}{T_{LM}} \tag{16}$$

This statement is validated by the results discussed below and justifies the simplified model presented in [1]. The rigorous second-law definition of combustor effectiveness, which accounts for all losses including incomplete combustion and pressure drops, is

$$E_Q = 1 - \Omega_{L,Q}/m_f \Delta G_r \tag{17}$$

Turbine. For a cooled turbine stage, the mass, energy, and exergy balances may be written as

$$m_{\text{out}} = m_{\text{in}} + \Sigma m_j \tag{18}$$

$$W_{TS} = m_{\text{in}} h_{\text{in}} + \Sigma m_j h_j - m_{\text{out}} h_{\text{out}} \tag{19}$$

$$\Omega_{L,TS} = m_{\text{in}} \omega_{\text{in}} + \Sigma m_j \omega_j - m_{\text{out}} \omega_{\text{out}} - W_{TS}\,. \tag{20}$$

where Σm_j represents the sum of all cooling flows to the stage: stator, rotor, and wheelspaces. Those flows may originate from different bleed points and have different properties. In contrast to the compressor[5] and combustor, the mass and energy balances are in themselves insufficient to specify the thermodynamic state leaving a cooled turbine stage of a given pressure ratio since both the exit temperature and the work are unknown. Thus the availability losses cannot be determined a posteriori, and the second law must form an integal part of the solution. This may be accomplished by specifying stage cooling-loss efficiency penalties or through fundamental entropy-generation models as presented by the author in [2]. To further complicate matters, unlike the compressor and combustor, the mass balance is coupled nonlinearly to the energy and irreversibility equations through the cooling flow calculation and the cooling thermodynamic loss penalty. The calculation procedure used is described in [14]. Its key assumptions and results are mentioned here. In all the results presented, the turbine consists of three equal-pressure-ratio stages. This implies either a single shaft machine or a two-spool engine without a free-power turbine, which would further constrain the stage pressure ratios.

The first step is to relate the cooling flows to the cycle parameters. In the illustrative examples given here, the wheelspace purge flows are taken as 0.6 percent by volume of the local gas flow. The gas path component flowrates are based upon the curves shown in Fig. 2 where cooling effectiveness

$$\Phi = \frac{T_{\text{in}} - T_{\text{blade}}}{T_{\text{in}} - T_{\text{coolant}}} \tag{21}$$

is plotted against the ratio $(m_c C_{p_c} / m_{\text{in}} C_{p_{\text{in}}})$ for the blade row. Those curves are based upon the analysis of [15]. An additional coolant leakage flow is assumed as 0.4 percent of the local gas flow by volume per blade row. The curves of Fig. 2 include endwalls and platforms for each blade row. The coolant for all rotors and the first stator is assumed to originate at compressor discharge. The second and third stators are supplied from two intermediate compressor bleeds at pressures 30 percent higher than those of gas inlet to the cooled stator. All wheelspace purge flows are taken from the same source as the adjacent stator row.

The gas temperature T_{in} for determining the rotor cooling flowrate through equation (21) and Fig. 2 is taken as the rotor-relative stagnation temperature.[6] Calculation of this temperature requires specification of rotor pitchline velocity relative to the speed of sound at ambient as well as absolute backswirl velocity component at rotor exit relative to pitchline velocity. Those quantities also enable the pitchline degree of reaction and stage loading to be determined. In the results presented, rotor pitchline speed is taken as 1.05, 1.1, and 1.15 times the speed of sound at ambient and rotor exit backswirl at 14, 7, and 0 percent of pitchline speed for stages 1, 2, and 3, respectively. Temperature pattern factors, representing the ratio of peak gas temperature above mixed mean to temperature rise through the combustor are taken as 0.12, 0.06, and 0.03 for the first three blade rows. Metal temperatures T_b used in the examples are listed in Appendix I.

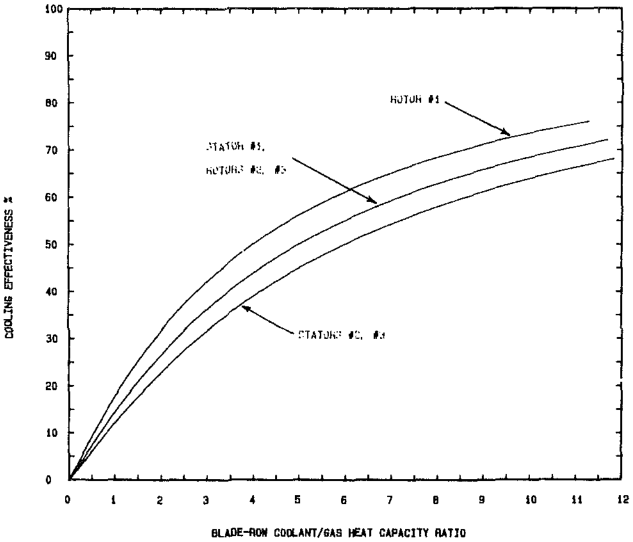


Fig. 2

Table 1 Cycle state points

POINT	PRESSURE BAR	TEMPERATURE K Absolute	Rotor-Relative	MASS FLOW % al	MOLE COMPOSITION % air	water	st.gas
Air Inlet (a)	1.013	289.0		100.00	100.00	0.00	0.00
Comp. Inlet (1)	1.005	289.0		100.00	100.00	0.00	0.00
Comp. Disch.(2)	14.076	652.7		75.49	100.00	0.00	0.00
Comb. Disch.(3)	13.513	1550.0		77.27	57.60	0.00	42.40
Inlet STATOR 1	13.513	1550.0		77.27	57.60	0.00	42.40
Inlet ROTOR 1	12.774	1469.3	1335.6	85.92	61.80	0.00	38.20
Inlet STATOR 2	5.873	1235.0		91.70	64.17	0.00	35.83
Inlet ROTOR 2	5.716	1202.9	1082.9	96.69	65.99	0.00	34.01
Inlet STATOR 3	2.552	1003.8		100.13	67.15	0.00	32.85
Inlet ROTOR 3	2.534	999.1	893.8	101.03	67.44	0.00	32.56
Turb. Exit (4)	1.109	825.7		101.78	67.67	0.00	32.33
Diffuser Ex(4A)	1.048	825.7		101.78	67.67	0.00	32.33

Table 2 Turbine analysis

BLADE ROW	COOLING EFFECT%	COOLANT % OF ROW INLET Gaspath Compnts	Gaspath w/Lekge	Dnstrm Whspce	STAGE EFF.%	STAGE WORK KJ/KG R.I. KJ/KG al	CHARACTERISTICS React-ion	Loading
STATOR 1	0.620	9.490	10.089	1.105	72.9	253.160		
ROTOR 1	0.529	5.229	5.729	1.004	88.0	217.514	0.156	1.969
STATOR 2	0.343	4.003	4.518	0.916	82.6	227.338		
ROTOR 2	0.262	2.220	2.627	0.937	91.3	219.807	0.264	1.612
STATOR 3	0.000	0.000	0.000	0.900	90.3	201.077		
ROTOR 3	0.000	0.000	0.000	0.736	95.9	203.158	0.348	1.305

The cooling thermodynamic loss mechanisms and calculation procedure adopted in GASCAN are physically similar to those given in [2]. However, greater differentiation between the gas momentum loss penalties arising from stator, rotor, foreward and aft wheelspace coolant mixing is provided. Also, unlike [2], a closed-form algebraic solution is not possible due to the property variations and the discrete rather than distributed model for staging.

Tables 1 and 2 for the sample cycle illustrate the results, with Table 2 showing some details of the turbine cooling flows and stage characteristics. Several different definitions exist for the efficiency of a cooled turbine stage. The column on stage efficiency in Table 2 shows two values for each stage. The top value represents the actual temperature drop across the stage due to work extraction only (had it been "cooled" by coolant

---

[5] Note that the second law is applied implicitly in the compressor analysis since T_2 depends on the polytropic efficiency which is a second-law concept.

[6] The recovery factor is assumed to be unity.

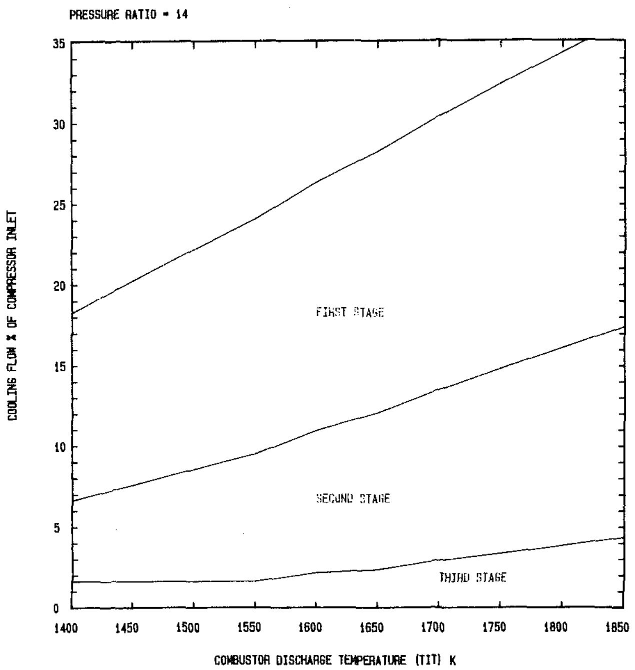


Fig. 3

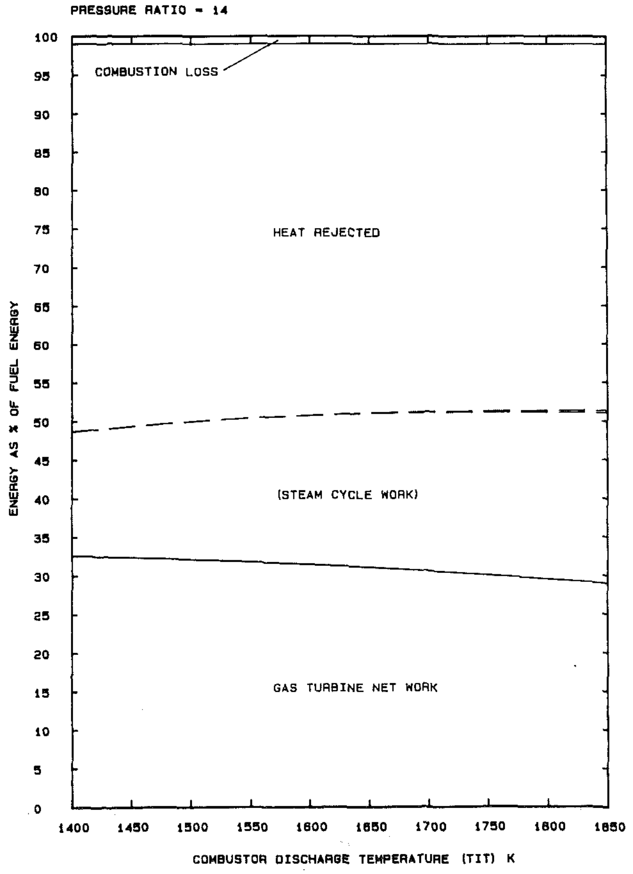


Fig. 4

at the local gas temperature) to the isentropic temperature drop based upon the total stage pressure ratio. This definition effectively accounts for the total pressure losses due to coolant injection as well as fluid friction. The lower value is the definition preferred by the author, the ratio of stage work to availability consumed

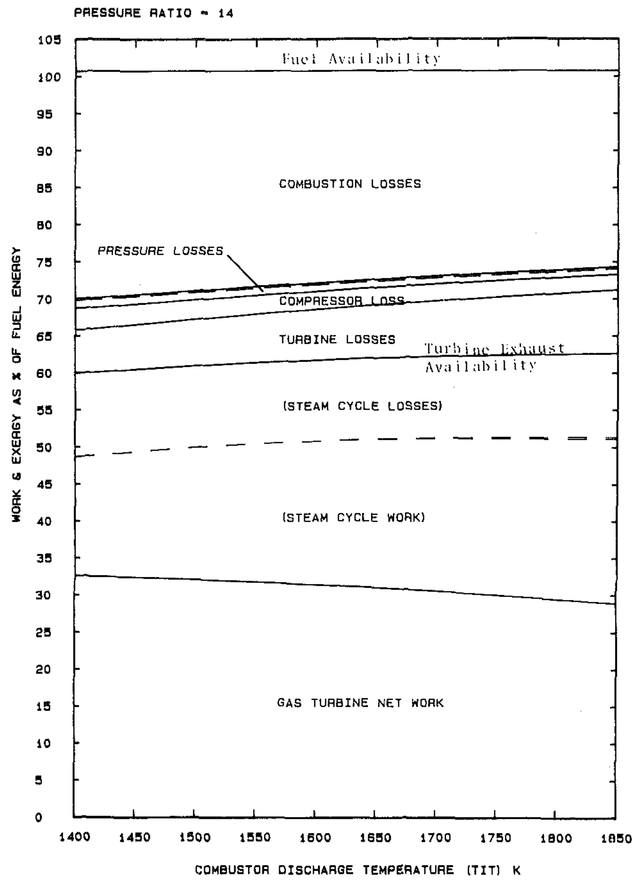


Fig. 5

$$E_{TS} = W_{TS}/(m_{\text{in}}\omega_{\text{in}} + \Sigma m_j\omega_j - m_{\text{out}}\omega_{\text{out}}) \qquad (22)$$

The next columns in Table 2 show the stage specific works per unit rotor inlet flow as well as per unit compressor inlet. The last two columns give the degree of reaction and stage loading parameters. Figure 3 shows cooling flow variation against T.I.T. for each stage at a cycle pressure ratio of 14.

Cycle Results and Discussion

Table 3 shows energy and exergy balances for the sample cycle with a pressure ratio of 14 and T.I.T. of 1550 K. This table is complemented by Figs. 4–8, which illustrate how those balances vary with T.I.T. and pressure ratio. The energy and exergy balances are presented as percentages of fuel heating value in the second and fifth columns of Table 3 and Figs. 4–6. Thus net work represents traditional first-law efficiency in those cases. The balances are presented as specific quantities in kJ/kg compressor inlet in the first and third columns of Table 3 and in Figs. 7 and 8, wherein the net work represents cycle specific power in kW/kg/s. Work quantities are identical in energy and exergy terms since work is entropy-free energy. The fourth column in Table 3 shows the exergy balance referred to fuel availability. This gives true second-law efficiency as defined by equation (2).

The rows showing steam cycle and combined cycle work in Table 3 and the dotted lines on Figs. 4–8 are based upon the optimized dual-pressure bottoming cycle discussed in Part 2 of this paper. The figures actually show two dotted lines, barely distinguishable from each other, representing optimized dual pressure steam cycles with maximum allowable throttle

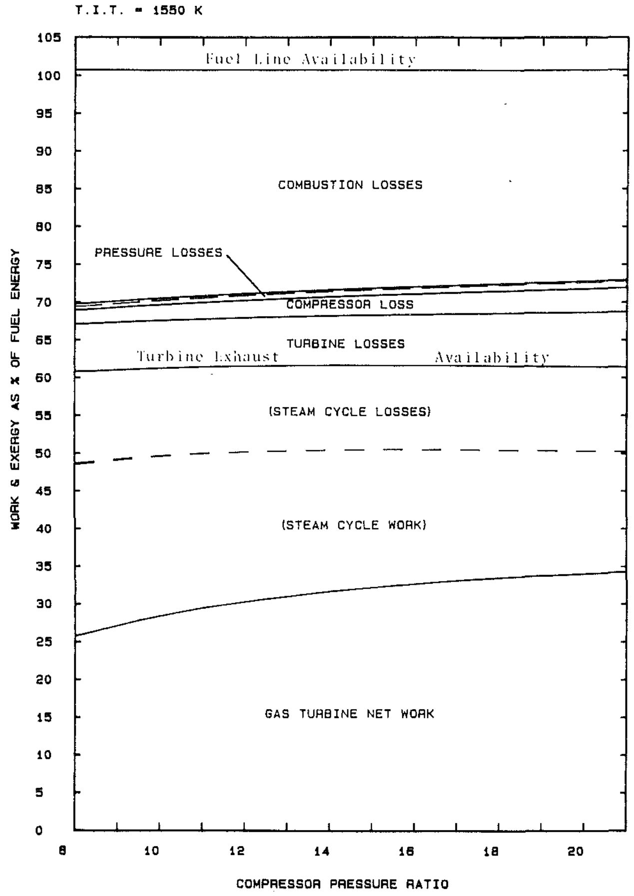


Fig. 6

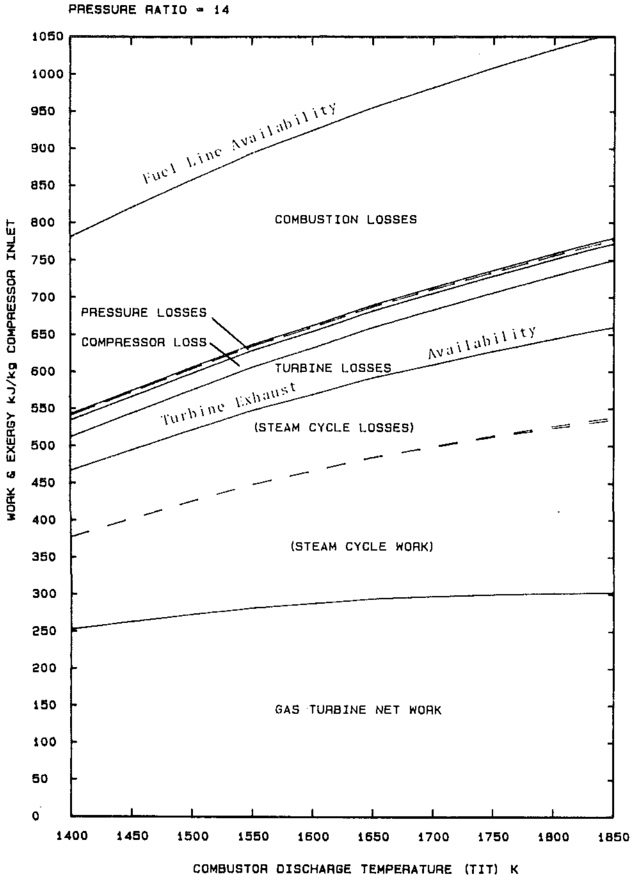


Fig. 7

Table 3 Energy and exergy balances

	ENERGY		EXERGY		
	KJ/KG al	% of Fuel Enthalpy	KJ/KG al	% of Fuel Exergy	% of Fuel Enthalpy
FUEL SUPPLY LINE	882.513	100.000	888.692	100.000	100.700
G.T. CYCLE NET WORK	276.941	31.381	276.941	31.163	31.381
LOSSES :-					
Inlet pressure loss	0.000	0.000	0.627	0.071	0.071
Compressor	0.000	0.000	22.424	2.523	2.541
Combustor	8.825	1.000	258.826	29.124	29.328
Turbine	0.000	0.000	59.493	6.694	6.741
Exit diffuser loss	0.000	0.000	4.839	0.545	0.548
(Exhaust Gas Stream)	596.747	67.619	265.543	29.880	30.089
STEAM CYCLE NET WORK	166.392	18.854	166.392	18.723	18.854
COMBINED CYCLE NET WORK	443.333	50.235	443.333	49.886	50.235
NET PLANT ELECTRIC WORK	428.941	48.605	428.941	48.267	48.605

temperatures of 1000°F and 1100°F. The bottom row in Table 3 shows net combined cycle electric output based upon electric generator efficiency of 98 percent after subtracting 0.6 percent shaft losses for each of the gas and steam turbines. Motor driven auxiliaries are assumed at 0.5 percent and balance-of-plant at 0.6 percent of gross output.

The fuel exergy in those balances includes fuel line pressure, taken as 20.67 bar (300 psi), in addition to fuel chemical availability ΔG_r, as shown in equation (13).

It is interesting to contrast the second and fifth columns of Table 3 and Figs. 4 and 5. The "bottom lines" are identical in both cases; the information presented, however, is different. The energy balance indicates that roughly 2/3 of the heat not converted to work by the gas turbine is rejected at its exhaust,

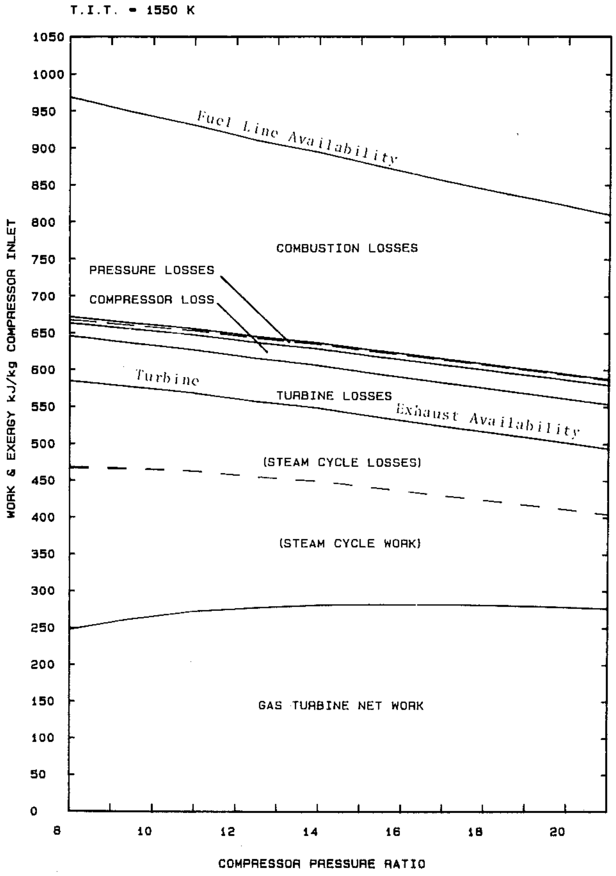


Fig. 8

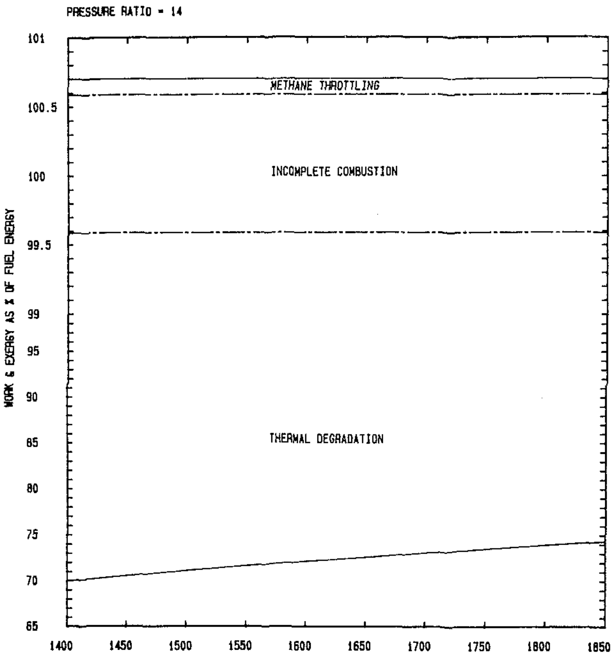


Fig. 9

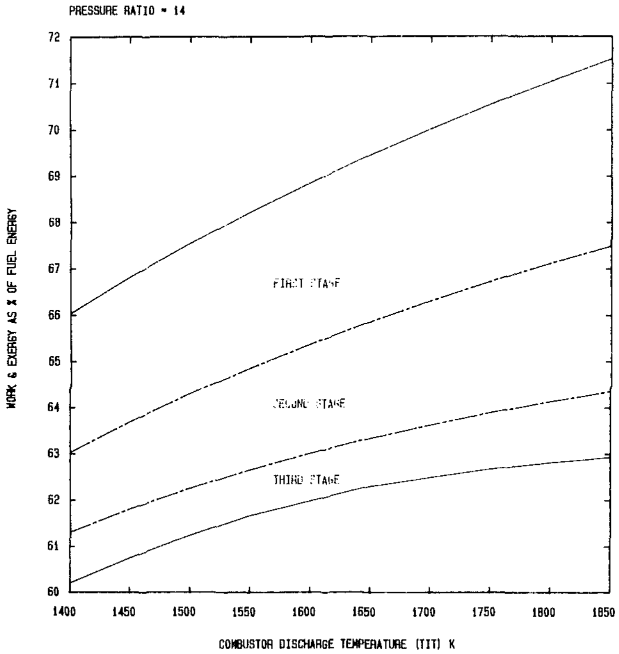


Fig. 11

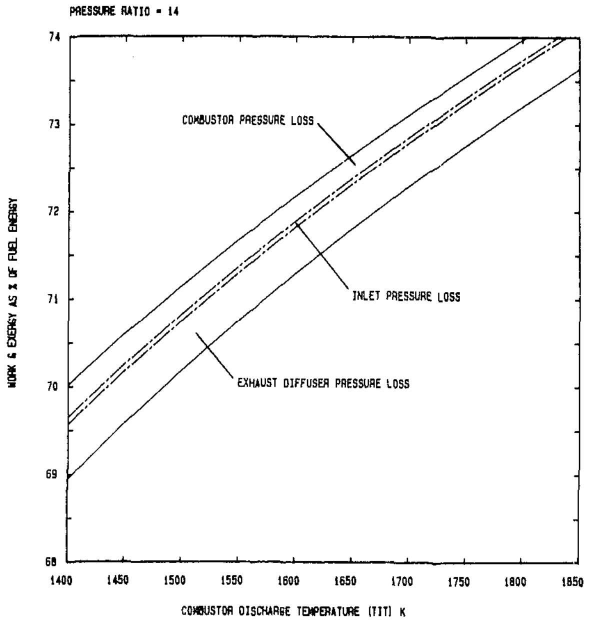


Fig. 10

and that the steam cycle recovers less than 1/3 of that heat, the balance being rejected mostly at the condenser. It does not show why such a large amount of energy has to be rejected. Moreover, it indicates no losses in the system pressure drops and in compressor friction, turbine friction, and coolant mixing, and just the 1 percent loss arising from incomplete combustion in the combustor. All those processes, which produce obvious physical losses, conserve energy and do not appear significant as sources of loss on the energy balance. The availability balance, on the other hand, provides a detailed quantitative accounting for the sources of those losses.

Table 4 Turbine irreversibilities breakdown

SOURCE OF IRREVERSIBILITY	EXERGY LOSS KJ/KG al AND MECHANISM					
	coolant throttling	mainstream pressure losses	thermal degradation	cooling totals	friction &aerodynamic	STAGE TOTAL
STATOR 1 cooling	0.577	3.154	6.876	10.607		
ROTOR 1 cooling	3.554	4.728	2.749	11.031		
Fwd.WS 1 cooling	0.069	0.576	0.711	1.355		
Aft WS 1 cooling	0.188	0.508	0.714	1.410		
Totals	4.389	8.965	11.050	24.404	5.321	
TOTAL STAGE # 1						29.725
STATOR 2 cooling	0.974	1.555	3.328	5.857		
ROTOR 2 cooling	3.596	2.367	0.674	6.637		
Fwd.WS 2 cooling	0.202	0.565	0.655	1.422		
Aft WS 2 cooling	0.198	0.533	0.805	1.536		
Totals	4.970	5.020	5.462	15.453	5.595	
TOTAL STAGE # 2						21.048
STATOR 3 cooling	0.000	0.000	0.000	0.000		
ROTOR 3 cooling	0.000	0.000	0.000	0.000		
Fwd.WS 3 cooling	0.202	0.607	0.793	1.602		
Aft WS 3 cooling	0.679	0.438	0.417	1.533		
Totals	0.881	1.044	1.210	3.135	5.585	
TOTAL STAGE # 3						8.720

It should be mentioned that the combustion-loss band in Figs. 5-8 includes the losses given by equations (13*b*), (13*c*), and (13*e*). The loss due to combustor pressure drop of equation (13*d*) is included in the pressure loss bands of Figs. 5-8.

The dotted line appearing at the bottom of the combustion-loss band in Figs. 5-8 represents the simplified, air-standard cycle model of sensible exergy addition to the cycle working fluid, i.e.,

$$E_Q^+ = \left(1 - \frac{T_a}{T_{LM}}\right)\eta_Q \Delta H_r \tag{23a}$$

in Figs. 5 and 6, and

$$\frac{\Omega^+}{\Delta H_r} = \left(1 - \frac{T_a}{T_{LM}}\right)\eta_Q \tag{23b}$$

in Figs. 7 and 8, T_{LM} being the logarithmic mean temperature of the working fluid during the heat addition prcess. This line

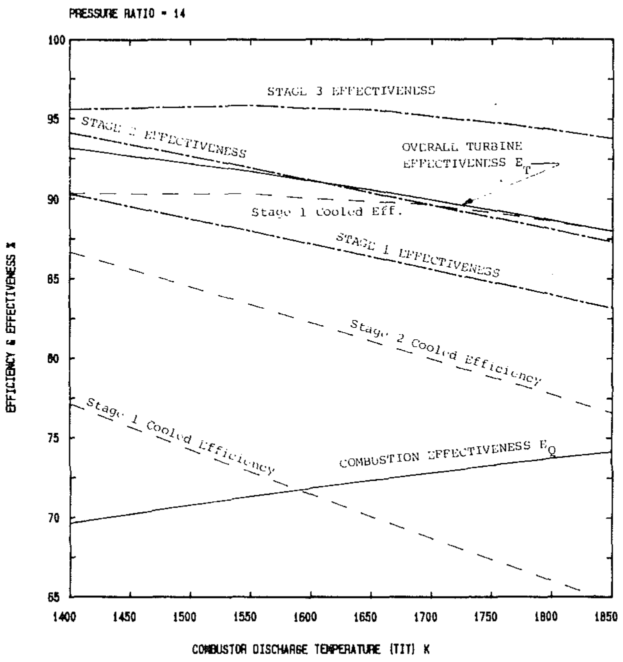


Fig. 12

lies very close to the solid line representing the actual computed sensible exergy addition, validating the approximate model in the case of methane-fired gas turbines.

To illustrate further the power of availability-balance analysis, Figs. 9-11 and Table 4 provide detailed magnification of the exergy-loss bands shown in Fig. 5. Figure 9 shows the combustion-loss details, Fig. 10 the breakdown of the pressure drop irreversibilities, and Fig. 11 a breakdown of the turbine loss band by stage. Table 4 provides a further breakdown of the turbine exergy losses by stage and by irreversibility mechanism for the sample cycle. The heading on the top of each row indicates each of the four irreversibility mechanism discussed in [2]. The contribution due to the cooling of each component is shown. The total column to the right shows the sum of cooling irreversibilities by component. The totals to the bottom show the sum by loss mechanism. Such a "spreadsheet" analysis can rapidly provide quantitative insight into the potential thermodynamic impact of any projected design modification.

Figure 5 shows that as T.I.T. increases the compressor and pressure drop exergy-loss bands are essentially constant. The combined cycle efficiency is governed mostly by the interaction between the decreasing combustion irreversibility, brought about by raising the logarithmic-mean combustion temperature, and the increasing turbine irreversibilities resulting from the necessity of larger cooling flows and their associated losses. This tradeoff is illustrated in Fig. 12, which shows the variation of combustion effectiveness as well as turbine effectiveness, for each stage and as a whole, with T.I.T. The graph also shows how the cooled-stage efficiency described earlier varies with T.I.T. Thus cycle efficiency changes very slightly with T.I.T. Means of decreasing the cooling-loss penalties are essential if cycle efficiencies are to increase with T.I.T. Higher material working temperatures or improved cooling techniques can accomplish that.

One way of decreasing combustion irreversibility without increasing T.I.T. is to employ reheat gas turbine cycles, where the logarithmic mean combustion temperature for the cycle can be increased without increasing T.I.T. The increased cooling losses due to subjecting more turbine stages to elevated temperatures, however, must be traded off against those benefits.

Conclusions

1 The method of exergy-balance analysis presented enables the time-honored first-law definition of efficiency to be preserved, while detailing and quantifying the sources of inefficiency. This quantification, which cannot be performed by traditional energy-balance analysis, can be a valuable tool for system optimization.

2 The simplified model for the combustion irreversibility, based on the Carnot efficiency at the logarithmic-mean combustion temperature, is in good agreement with the detailed calculation.

3 The dominant influence on cycle efficiency as turbine inlet temperature is raised is the tradeoff between decreased combustion exergy losses and increased turbine cooling losses.

References

1 El-Masri, M. A., "On Thermodynamics of Gas Turbine Cycles—Part I: Second Law Analysis of Combined Cycles," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 107, 1985, pp. 880-889.

2 El-Masri, M. A., "On Thermodynamics of Gas Turbine Cycles—Part II: Model for Expansion in Cooled Turbines," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 108, 1986, pp. 151-159.

3 El-Masri, M. A., "On Thermodynamics of Gas Turbine Cycles—Part III: Potential and Limitations of Reheat Turbine Combined Cycles," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 108, 1986, pp. 160-170.

4 Teshima, K., et al., "A Report for the Engineering Status of Some Confirmation Tests in the Development of AGTJ-100A," ASME Paper No. 84-GT-53.

5 Tekeya, K., Oteki, Y., and Yasui, H., "Current Status of Advanced Reheat Gas Turbine AGTJ-100A (Part 3) Experimental Results of Shop Tests," ASME Paper No. 84-GT-57.

6 Arai, M., et al., "Results From High Temperature Turbine Test on the HPT of the AGTJ-100A," ASME Paper No. 84-GT-235.

7 Rice, I. G., "The Combined Reheat Gas Turbine/Steam Turbine Cycle; Part 1—A Critical Analysis of the Combined Reheat Gas Turbine/Steam Turbine Cycle," ASME Paper No. 79-GT-7.

8 Rice, I. G., "The Combined Reheat Gas Turbine/Steam Turbine Cycle: Part 2—The LM5000 Gas Generator Applied to the Combined Reheat Gas Turbine/Steam Turbine Cycle," ASME Paper No. 79-GT-00.

9 Rice, I. G., and Jenkins, P. E., "Comparison of the HTTT Reheat Gas-Turbine Combined Cycle With the HTTT Nonreheat Gas Turbine Combined Cycle," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 104, 1982, pp. 131-142.

10 Messerlie, R. L., and Strother, J. R., "Integration of the Brayton and Rankine Cycles to Maximize Gas Turbine Performance—A Cogeneration Option," ASME Paper No. 84-GT-52.

11 Rice, I. G., "The Reheat Gas Turbine With Steam-Blade Cooling —A Means of Increasing Reheat Pressure, Output, and Combined Cycle Efficiency," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 104, 1982, pp. 9-22.

12 Rice, I. G., "Steam Cooled Blading in a Combined Reheat Gas Turbine/Reheat Steam Turbine Cycle: Part I—Performance Evaluation," ASME Paper No. 79-JPGC-GT-2.

13 Kalina, A. I., "Combined Cycle System With Novel Bottoming Cycle," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 106, 1984, pp. 737-742.

14 El-Masri, M. A., "GASCAN—An Interactive Code for Thermal Analysis for Gas Turbine Systems," presented at the ASME Winter Annual Meeting, Anaheim, CA, Dec. 7-12, 1986.

15 El-Masri, M. A., and Pourkey, F., "Prediction of Cooling Flow Requirements for Advanced Utility Gas Turbines," Parts 1 and 2, ASME Paper No. 86-WA/HT-43 and 86-WA/HT-44.

APPENDIX I

Inputs not mentioned in the body of the paper but used in the examples cited are:

Fractional total pressure drops:
Atmosphere to compressor inlet = 0.0075
Across combustor = 0.04
Turbine exit to diffuser exit = 0.06
Diffuser exit to stack (including waste heat boiler) = 0.035
Compressor polytropic efficiency = 90.5 percent
Combustion efficiency = 99 percent

Maximum metal temperatures:

Stator #1: 1035 K	Rotor #1: 1000 K
Stator #2: 1015 K	Rotor #2: 970 K
Stator #3: 1015 K	Rotor #3: 950 K

The polynomials used to express $C_p(T)$ for each of the three gaseous components defined in the paper are plotted on dimensionless coordinates in Fig. A1.

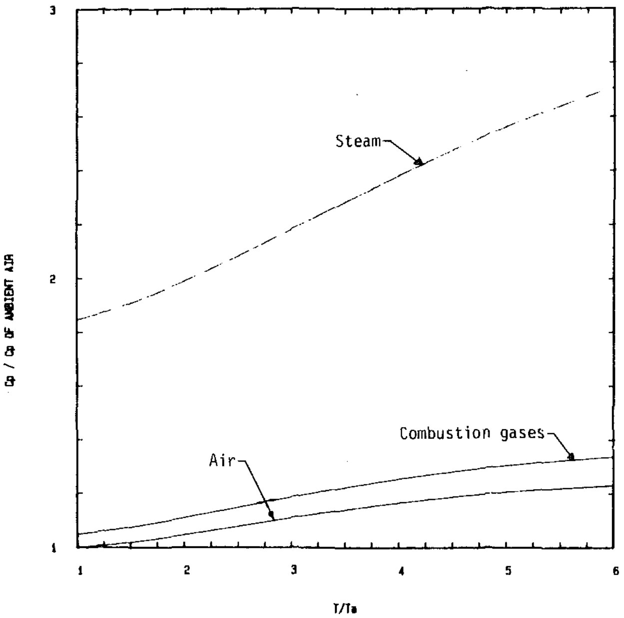


Fig. A1

APPENDIX II

CYCLE ENERGY & EXERGY STREAMS

STREAM	THERMODYNAMIC STATE				STREAM/KG COMPR. INLET		
	P BAR	T K	h KJ/KG	omega KJ/KG	Mass KG	Enthalpy KJ	Exergy KJ
Atmosphere inlet	1.013	289.00	0.00	0.00	1.00	0.00	0.00
Compr. Inlet (1)	1.005	289.00	0.00	-0.63	1.00	0.00	-0.63
Bleed port # 1	14.076	652.75	376.53	352.85	0.16	60.67	56.85
Bleed port # 2	7.634	543.10	260.41	242.07	0.06	15.22	14.15
Bleed port # 3	3.318	420.36	133.29	122.23	0.03	3.40	3.12
Compr. Disch.(2)	14.076	652.75	376.53	352.85	0.75	284.25	266.37
Fuel supply line	20.665	289.00	50187.34	50538.73	0.02	882.51	888.69
Comb. Disch. (3)	13.513	1550.00	1498.56	1159.88	0.77	1157.94	896.24
STAGE # 1							
Main Gasflow in	13.513	1550.00	1498.56	1159.88	0.77	1157.94	896.24
STATOR coolant	14.076	652.75	376.53	352.85	0.08	29.35	27.51
ROTOR coolant	14.076	652.75	376.53	352.85	0.05	18.53	17.37
Fwd.WS coolant	14.076	652.75	376.53	352.85	0.01	3.22	3.01
Aft WS coolant	7.634	543.10	260.41	242.07	0.01	2.25	2.09
Stage shaft work						217.51	217.51
STAGE # 2							
Main Gasflow in	5.873	1234.99	1083.67	762.20	0.92	993.77	698.97
STATOR coolant	7.634	543.10	260.41	242.07	0.04	10.79	10.03
ROTOR coolant	14.076	652.75	376.53	352.85	0.03	9.56	8.96
Fwd.WS coolant	7.634	543.10	260.41	242.07	0.01	2.19	2.03
Aft WS coolant	3.318	420.36	133.29	122.23	0.01	1.21	1.11
Stage shaft work						219.81	219.81
STAGE # 3							
Main Gasflow in	2.552	1003.80	796.65	479.61	1.00	797.71	480.25
STATOR coolant	3.318	420.36	133.29	122.23	0.00	0.00	0.00
ROTOR coolant	14.076	652.75	376.53	352.85	0.00	0.00	0.00
Fwd.WS coolant	3.318	420.36	133.29	122.23	0.01	1.20	1.10
Aft WS coolant	3.318	420.36	133.29	122.23	0.01	0.99	0.91
Stage shaft work						203.16	203.16
Turbine exit (4)	1.109	825.70	586.32	265.66	1.02	596.75	270.38
Diffuser exit 4A	1.048	825.70	586.32	260.90	1.02	596.75	265.54

Compressor effectiveness = 93.83185 %
Combustor effectiveness = 71.3719 % Fuel enthalpy 70.87566 % Fuel exergy
Turbine Stage 1 effectiveness = 87.97719 %
Turbine Stage 2 effectiveness = 91.26108 %
Turbine Stage 3 effectiveness = 95.88434 %
Turbine effectiveness as a whole= 91.50059 %

W. W. Chin
Research Assistant.

M. A. El-Masri
Associate Professor.

Department of Mechanical Engineering,
Massachusetts Institute of Technology,
Cambridge, MA 02139

Exergy Analysis of Combined Cycles: Part 2—Analysis and Optimization of Two-Pressure Steam Bottoming Cycles

Results of a study for selecting the optimum parameters of a dual-pressure bottoming cycle as a function of the gas turbine exhaust temperature are presented. Realistic constraints reflecting current technological practice are assumed. Exergy analysis is applied to quantify all loss sources in each cycle. Compared to a single pressure at typical exhaust gas temperatures the optimized dual-pressure configuration is found to increase steam cycle work output on the order of 3 percent, principally through the reduction of the heat transfer irreversibility from about 15 to 8 percent of the exhaust gas energy. Measures to further reduce the heat transfer irreversibility such as three-pressure systems or use of multicomponent mixtures can therefore only result in modest additional gains. The results for the efficiency of optimized dual-pressure bottoming cycles are correlated against turbine exit temperature by simple polynomial fits. Sensitivity of the results to variations in the constraint envelope are presented.

Introduction

Thermal optimization of steam bottoming cycles operating on gas turbine exhaust heat differs from that for directly fired boilers in that the heat supply is of relatively low grade. Table 1 of Part 1 of this paper illustrates and quantifies this fact by displaying the relatively low exergy content of typical turbine exhaust gas relative to its energy content. The decreasing temperature of the heating gas as energy is extracted therefrom necessitates careful matching of the steam-side heating curve to maximize the availability transferred between the streams and minimize the exergy lost in the heat transfer process. This matching can be improved by increasing the number of evaporator pressures on the steam side from one to two or more, or by application of developments such as the Kalina cycle [1], currently under development, wherein a water-ammonia mixture boils at a temperature which varies with composition. Either of those means to reduce the heat transfer exergy loss introduces additional cost and complexity, which have to be weighed against the performance improvement.

One purpose of this paper is to provide a basis for quantifying the improvements to cycle performance that can be brought about by adopting measures which reduce heat transfer irreversibility as well as variation of the design parameters or constraints. This basis is provided by the presentation of detailed exergy balances of steam cycles based upon practical performance constraints. Thus the extent of improvement made possible by reducing (or in the theoretical limit, eliminating) any of those sources of availability loss can be assessed.

The second purpose of this paper is to provide results of an optimization study on a two-pressure evaporator cycle configuration. This is selected since it is thought to provide a reasonable tradeoff between performance improvement and complexity, and the model includes as a subset the performance of a single-pressure system. The results show the optimum selection of evaporator pressures as a function of gas turbine exhaust temperature. The optimization is carried out within a constraint envelope representative of current technological limitations. Sensitivity of the final results to departures from the optimum evaporator pressures and to variations in the constraint envelope are also presented.

Cycle Description and Analysis

The cycle model is a two-pressure steam Rankine cycle with reheat running off of a gas turbine exhaust at a specified temperature T_{4A}. Due to energy losses in the connecting ductwork, this gas reaches the boiler at a temperature T_{4B} and no additional firing in the exhaust gas is assumed. A single-pressure system is a subset of the model and results are presented for this case also. At each pressure the steam streams are superheated to the same temperature, as shown in the temperature-entropy cycle diagram of Fig. 1. In that figure, the high-pressure stream mixes with the low-pressure stream at a temperature below $T_{\mathrm{sat},H}$, the saturation

---

Contributed by the Gas Turbine Division and presented at the Joint ASME/IEEE Power Generation Conference, Portland, Oregon, October 19-23, 1986. Manuscript received by the Gas Turbine Division July 31, 1986. Paper No. 86-JPGC-GT-10.

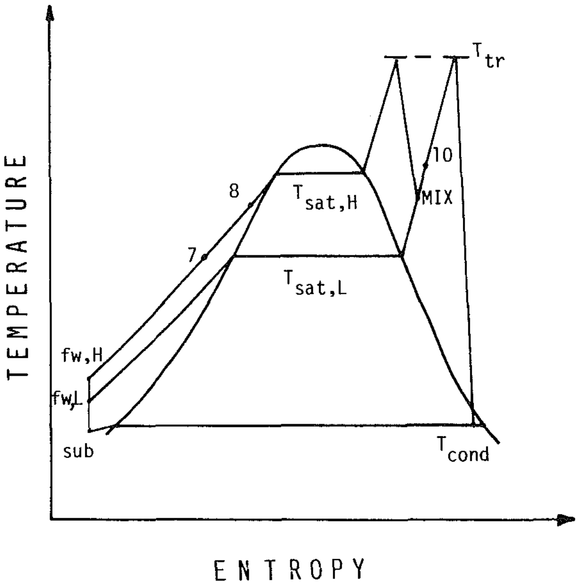


Fig. 1 Temperature-entropy diagram of the steam cycle for the case $T_{sat,H}$ T_{mix}

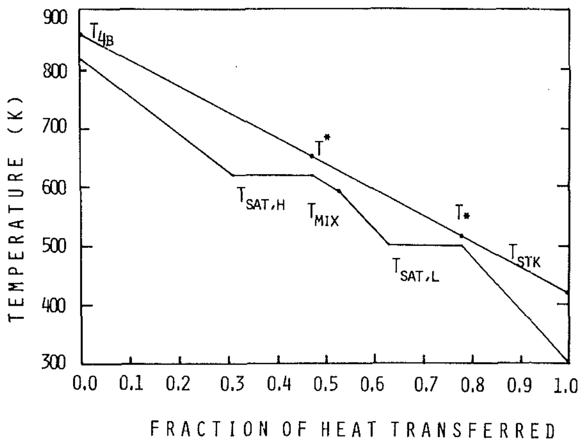


Fig. 2 Heat transfer diagram for the case $T_{sat,H}$ T_{mix}

temperature corresponding to the high pressure. A corresponding heat transfer diagram (Fig. 2) shows the fraction of energy transferred from the gas stream to the steam side above any temperature. This arrangement implies that part of the high-pressure stream economizer surface is in parallel with the low-pressure superheater tube bank. Note that the change in slope that occurs at the mixing temperature T_{mix} is below $T_{\text{sat},H}$. For a fixed high pressure, if the low pressure is high enough the two streams will mix at a temperature above $T_{\text{sat},H}$, resulting in a heat transfer diagram with T_{mix} slope change above $T_{\text{sat},H}$, in contrast to the case depicted in Figs. 1 and 2. In each of those two cases, a different algorithm is used to perform an energy balance.

For given values of T_{4A}, T_{tr}, both pressures and the set of constraints given in the Appendix and discussed below, it is possible to find the properties at all the necessary steam cycle state points. The only unknowns on the steam side then are the low pressure mass flow fraction m_L and the total steam mass flow per unit of gas flow m. On the gas side, the only unknown is the stack temperature. Introducing the definitions

$$T^* = T_{\text{sat},H} + \Delta T_{pp,\,H}$$

$$T_* = T_{\text{sat},L} + \Delta T_{pp,\,L}$$

A = fraction of energy used to heat steam at or above $T_{\text{sat},L}$

$$= \frac{q_{st} - [m_H(h_7 - h_{fw,H}) + m_L(h_{f,L} - h_{fw,L})]}{q_{st}}$$

where

$$q_{st} = m_L(h_{\text{mix}} - h_{fw,L}) + m_H(h_{sh,H} - h_{fw,H}) + (h_{sh,L} - h_{\text{mix}}) + \text{blowdown}$$

$$\text{blowdown} = BD[m_L(h_{f,L} - h_{fw,L}) + m_H(h_{f,H} - h_{fw,H})]$$

and

B = fraction of energy used to heat steam at or above $T_{\text{sat},h}$

$$= \frac{m_H(h_{sh,H} - h_{f,H}) + (h_{sh,L} - h_{10})^{1}}{q_{st}}$$

---

[1] This expression is for the case where $T_{\text{mix}} < T_{\text{sat},H}$. For the other case another heat balance is used.

Nomenclature

BD = blowdown factor = 2 percent
ΔP = pressure loss
ΔT_{apr} = exhaust gas approach temperature difference to boiler
ΔT_{pp} = pinch point temperature difference
η_i = internal cycle efficiency
η_1 = first-law cycle efficiency
η_2 = second-law cycle efficiency
h = enthalpy
h^* = enthalpy of gas at T^*
h_* = enthalpy of gas at T_*
m = dimensionless mass fraction of steam per unit
m_L = dimensionless fraction of steam mass flow in low-pressure stream
m_H = $1 - m_L$
Q = heat
q = heat per unit mass of flow
T = absolute temperature
Γ = nondimensional temperature = T/T_a
W = work
Ω = availability
ω = availability per unit mass

Subscripts

$4A$ = gas state at turbine exhaust
$4B$ = gas state at boiler inlet
a = ambient
cond = condenser
duct = ducting between gas turbine exit and boiler inlet
f = saturated liquid
fw = feedwater
H = high pressure
HT = heat transfer
L = low pressure, loss
min = minimum
mix = where high-pressure stream mixes with low-pressure stream
pipe = piping between superheater exit and throttle
sat = saturation
sh = superheater
st = steam
stk = stack
sub = subcooled
T = turbine
tr = throttle

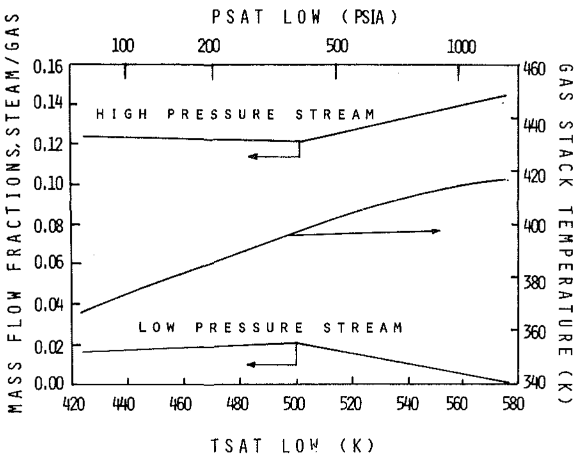


Fig. 3 Gas stack temperature and mass flow fractions as a function of $T_{sat,L}$

The steam enthalpies h_7 and h_{10} are for states 7 and 10, respectively, as defined in Fig. 1. Referring to Fig. 2, it can be seen that the fraction of heat on the gas side above T_* must be equal to A, giving

$$A = \frac{h_{4B} - h_*}{h_{4B} - h_{\text{stk}}} \tag{1}$$

Similarly,

$$B = \frac{h_{4B} - h^*}{h_{4B} - h_{\text{stk}}} \tag{2}$$

Finally by performing a heat balance for the entire cycle one obtains

$$m = \frac{m_{st}}{m_{\text{gas}}} = \frac{h_{4B} - h_{\text{stk}}}{q_{st}} \tag{3}$$

If we assume $\Delta T_{pp,\,L}$ is equal to the minimum pinch point temperature difference, then for a given m_L equation (1) can be solved for h_{stk}, the enthalpy of the gas at the stack. Using equation (2), T^* and thus $\Delta T_{pp,\,H}$ is found.

A computer program was constructed to perform the cycle analysis according to the above methodology. The steam properties were determined according to the equations given in Irvine and Liley [2]. The exhaust gas properties were assumed to be those for CH_4 burned with 250 percent theoretical air.

It was found that the best cycle efficiency was achieved when both pinch point temperature differences were at the minimum allowed by the specified constraints. So for a given set of inputs, the program iterates through m_L until both pinch point temperature differences are equal to the minimum allowable ΔT_{pp}. Then an overall energy balance is performed using equation (3) and the mass fraction ratio m is found. The results of the iterations are shown in Fig. 3. For the reference case example defined below, the mass flow fractions are found for each low pressure, then the stack temperature, work of each turbine and pump, availabilities, losses, and all other quantities of interest can be calculated directly for the cycle.

In an optimizing loop, given T_{4A} and the high pressure, the program will iterate through low pressures and save the $T_{\text{sat},L}$ that gives the highest net cycle work. If a cycle constraint is reached during an iteration, the program goes to the next logical loop. With the fixed inputs the program lowers $T_{\text{sat},L}$ by a decrement of 2 K for each iteration. So if the quality in the condenser falls below the minimum allowable quality, the program stops further calculations and goes to the next $T_{\text{sat},L}$. If the quality of the mix state is below the minimum or if the stack constraint is reached, the program stops the $T_{\text{sat},L}$ loop.

Cycle Constraints and Reference-Case Parameters

Table 1 Cycle constraints and parameters

T_a	289 K	(60°F)
P_{cond}	0.0677 bar	(2.0 in. HgA)
T_{sub}	303 K	(86°F)
$T_{\text{stk,min}}$	367 K	(200°F)
$\Delta T_{pp,\text{min}}$	14 K	(25°F)
ΔT_{apr}	42 K	(75°F)
$P_{tr,\text{max}}$	166.5 bar	(2400 psig)
$T_{tr,\text{max}}$	867 K	(1100°F)
ΔT_{duct}	−2.2 K	(−4°F)
ΔP_{duct}	3.5 percent	
ΔP_{pipe}	6 percent	
ΔP_{sh}	5 percent	
ΔP_{reheat}	5 percent	
Boiler radiant loss	1.5 percent of total	
Blowdown loss	2 percent of $(h_f - h_{fw})$ for each pressure	
Exhaust gas supply pressure above P_a	3.5 percent of P_a	
η_{fw}	70 percent	

The operating constraint and parameters are summarized in Table 1. The maximum throttle temperature of 867 K (1100°F) is not a set cycle parameter. The throttle temperature will be at this maximum only if the gas temperature T_{4B} is sufficiently high to satisfy the minimum gas approach temperature difference. Otherwise, the throttle temperature is taken as

$$T_{tr} = T_{4B} - \Delta T_{\text{apr}}$$

The effect of decreasing the maximum throttle temperature constraint to 811 K (1000°F) will be examined later.

It was found that the cycle had the highest efficiency at the highest throttle pressure. Therefore, for all program runs, the maximum allowable throttle pressure of 166.5 bars (2400 psig) was assumed for the high-pressure stream. Again, the maximum is later decreased to quantify its effect on the cycle efficiency.

The minimum stack temperature of 367 K (200°F) is based on the assumption that clean natural gas is the fuel for the gas turbine.

The low-pressure turbine efficiency is assumed to be 87.0 percent. The high-pressure turbine efficiency is 79.18 percent for the reference-case cycle. A correlation for the change in steam turbine efficiency as a function of inlet temperature and pressure was presented by Foster-Pegg [3] based upon the results of Spencer et al. [4]. A subroutine was generated to adjust the turbine efficiency according to the curves given in the former reference.

Results and Discussion

The internal steam cycle efficiency is defined as

$$\eta_i = \frac{W_{\text{net}}}{Q_{\text{gas}}} = \frac{W_{\text{net}}}{m_{\text{gas}}(h_{4A} - h_{\text{stk}})} \tag{4}$$

However, in this study we use the definition of the overall bottoming cycle first-law efficiency as

$$\eta_1 = \frac{W_{\text{net}}}{\text{heat available}} = \frac{W_{\text{net}}}{m_{\text{gas}}(h_{4A} - h_a)} \tag{5}$$

Therefore the cycle efficiency is penalized for the available heat that the cycle does not recover. Thus, the latter definition incorporates the conventional efficiency with the efficiency with which the heat recovery boiler recoups the gas turbine exhaust energy.

The second-law bottoming cycle effectiveness is given by

$$\eta_2 = \frac{W_{\text{net}}}{\text{available work}} = \frac{W_{\text{net}}}{m_{\text{gas}}(\omega_{4A} - \omega_a)} \tag{6}$$

Effect of Low Pressure Selection. The efficiency of the bottoming cycle is plotted versus $T_{\text{sat},L}$ for several combinations of maximum throttle pressures and throttle temperatures in Fig. 4. The lowest low pressure ($T_{\text{sat},L}$) at which each curve

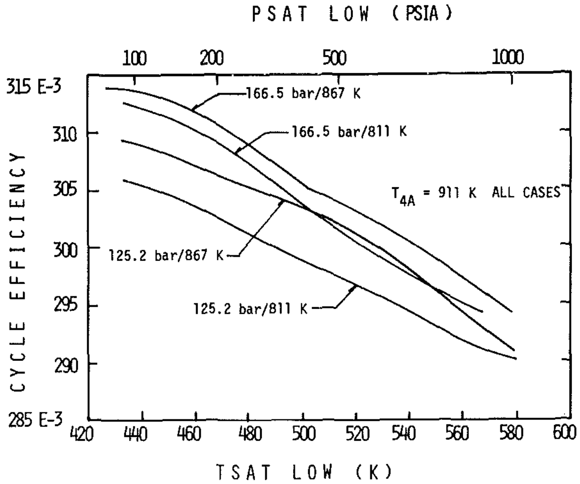


Fig. 4 First-law cycle efficiencies for two-pressure cycles with various throttle constraints

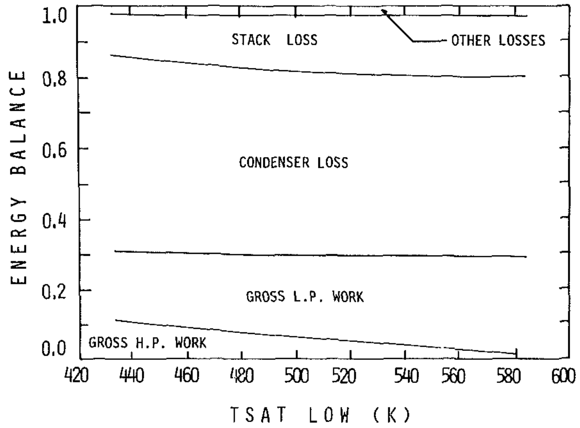


Fig. 5 An energy balance for the cycle versus $T_{\text{sat},L}$. Throttle constraints are 166.5 bar/867 K, and T_{4A} = 911 K.

ends corresponds to the pressure at which the stack temperature constraint is reached.

In Fig. 5, an energy balance is given for the case corresponding to the highest efficiency curve in Fig. 4, namely the case with maximum throttle pressure of 166.5 bar (2400 psig), throttle temperature of 867 K (1100°F) and gas turbine exit temperature T_{4A} of 911 K. This energy balance illustrates the tradeoff between energy losses in the condenser and unrecovered heat rejected through the stack. However, the figure does not show to which components these losses can be attributed. An availability flow chart for the steam cycle is shown in Fig. 6. The Ω for each steam stream are total exergies of that stream per unit of gas mass flow. By performing an exergy balance for the control volume we get

$$\Omega_{4A} = W_{\text{net}} + \Omega_{\text{stk}} + \Omega_{L,HT} + \Omega_{L,\text{cond}} + \Omega_{L,T} + \Omega_{L,\text{etc.}} \qquad (7)$$

where

$W_{\text{net}} = W_{T,H} + W_{T,L} - W_{\text{pump}}$

Ω_{stk} = exergy of stack gas

Ω_{HT} = exergy lost in heat transfer

$\Omega_{L,\text{cond}} = \Omega_{\text{cond}} - \Omega_{\text{sub}}$

$\Omega_{L,T}$ = exergy lost in turbines

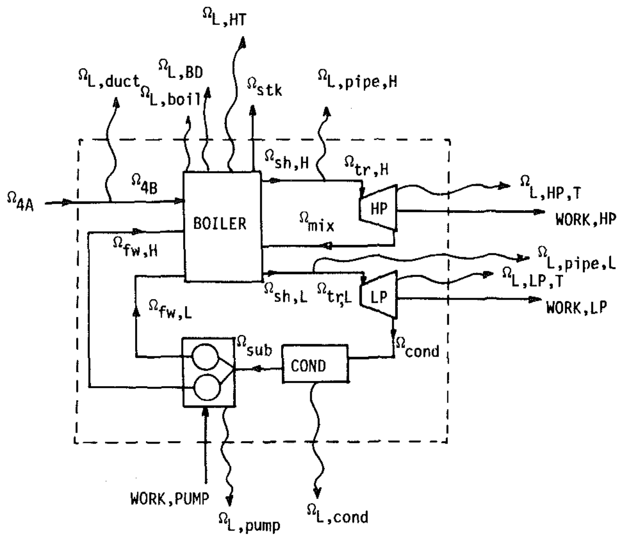


Fig. 6 An availability flow diagram for the two-pressure bottoming cycle

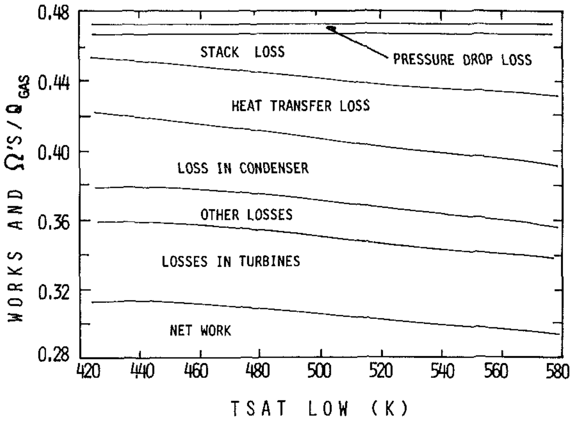


Fig. 7 An availability balance for the 166.5 bar/867 K throttle constraint two-pressure cycle with T_{4A} = 911 K, with exergy and work terms divided by thermal energy of gas at state 4*A*

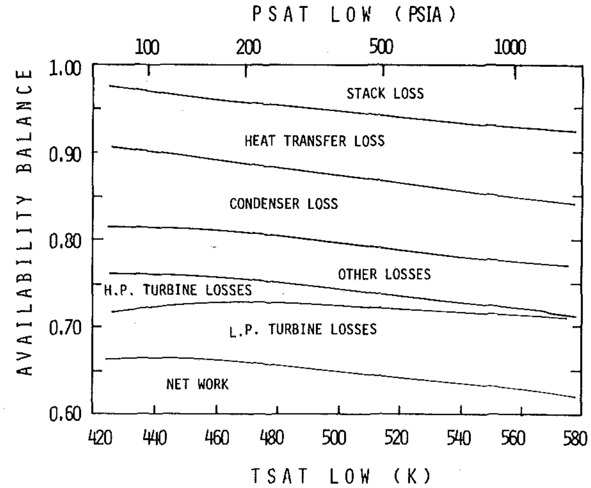


Fig. 8 An availability balance for the 166.5 bar/867 K throttle two-pressure cycle with T_{4A} = 911 K, plotted as a function of the low pressure

$\Omega_{L,\text{etc.}}$ = exergy losses in blowdown, radiant loss ducting, piping, and irreversibility in feedwater pump

The heat transfer exergy loss term occurs because the

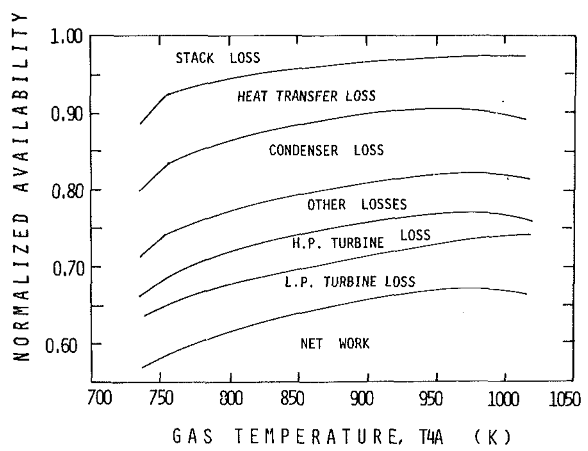


Fig. 9 An availability balance for the two-pressure cycle, with 166.5 bar/867 K throttle constraint

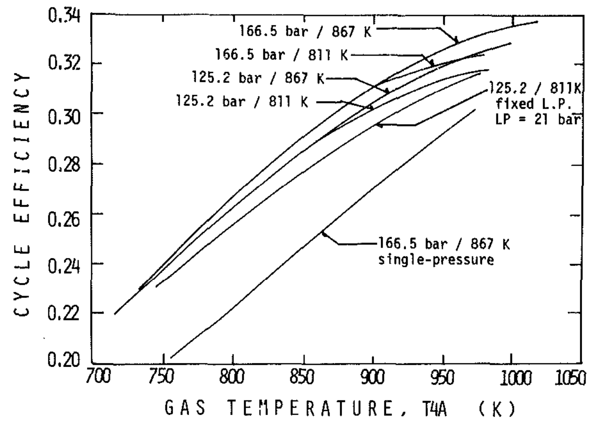


Fig. 12 First-law efficiencies for various throttle constraints as a function of exhaust temperature T_{4A}

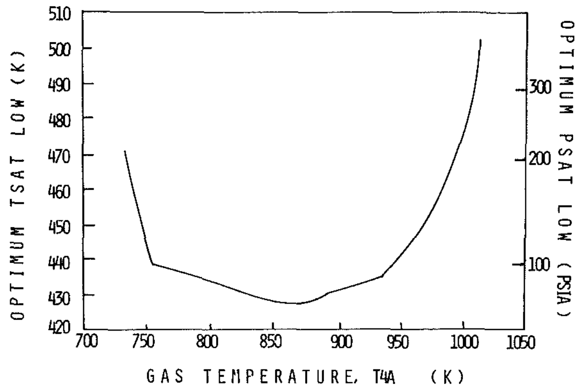


Fig. 10 The optimum choice of the low pressure for the two-pressure cycle with throttle constraints of 166.5 bar/867 K

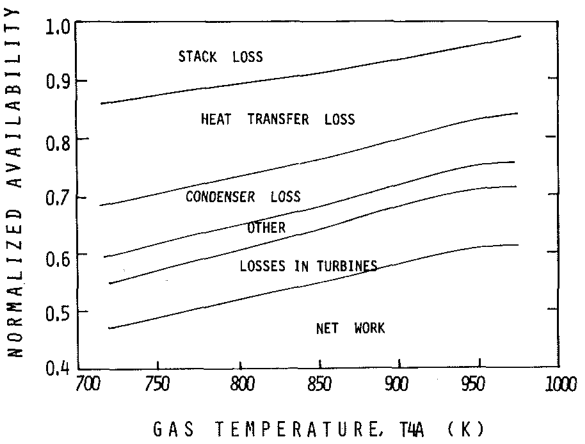


Fig. 11 An availability balance for a single-pressure cycle, with 166.5 bar/867 K throttle constraint

"quality" of energy at the gas temperature is degraded during the process of transferring it to a lower temperature stream. Figures 7 and 8 present graphic representations of the availability balance of equation (7). In Fig. 7 the availabilities are divided by $Q_{\rm gas}$, such that the cycle net work corresponds to the first-law efficiency definition in equation (1) and is identical to the efficiency curve for that cycle shown in Fig. 4. The narrow band at the top of the figure labeled "Pressure Drop Loss" is the fraction of availability lost through the pressure drop in the duct between the gas turbine exit and the boiler inlet as well as in gas flow through the boiler to the stack. All the other bands represent exergy losses within the steam cycle and sum up to the sensible thermal exergy contained in the gas stream. Figure 8 presents the same availability balance, availabilities in this case being normalized by dividing them by the availability of the exhaust gas. Thus cycle net work corresponds to the second-law efficiency definition of equation (6). Here the small pressure drop loss is incorporated in the "Duct loss" term with the loss in the duct due to the temperature drop.

Correlation of Optimized Cycles Versus Exhaust Gas Temperature. In Fig. 9 an availability balance is given for a class of cycles having the high throttle pressure constraint of 166.5 bars (2400 psig) and maximum throttle temperature of 867 K (1100°F). Each cycle has been optimized by the selection of the best low pressure at the given gas exhaust temperature, which is plotted as the abscissa. As an example, the optimum choice of the low pressures for the 166.5 bar/867 K (2400 psig/1100°F) case is given in Fig. 10. For each gas temperature the highest allowable throttle temperature consistent with the minimum approach temperature-difference constraint is used. For comparison, the availability balance for the single pressure cycle with the same cycle constraints is shown in Fig. 11. Note that the heat transfer exergy loss in the two-pressure case is nearly one-half the heat transfer loss in the single pressure.

The steam cycle efficiency as defined in equation (1) is plotted against the gas exhaust temperature for various combinations of throttle conditions in Fig. 12. Note that for a set of curves with the same maximum throttle pressure, the cycle with the higher maximum throttle temperature constraint gives better efficiencies unless the gas temperature T_{4B} is too low to satisfy the minimum gas approach temperature difference. Under such conditions the two efficiency curves coincide. Also given in the figure is a comparison of efficiencies between the optimized low pressure and fixed low pressure for the 125.2 bar/811 K (1800 psig/1000°F) case. The difference of approximately 0.8 percent is attributable to reduced stack loss, as the optimized cycle brings the low pressure down to the stack constraint. This reduction in stack loss is large enough to overcome the increased condenser loss associated with the lower low pressure. An efficiency curve for a single-pressure cycle with the same constraints is also included for comparison.

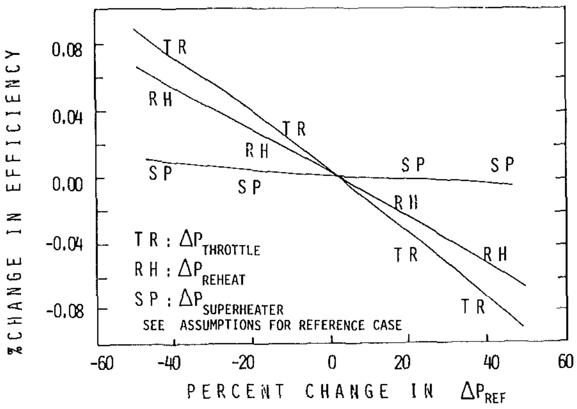


Fig. 13(*a*) The change in cycle efficiency as a function of the various pressure drops given in the assumptions

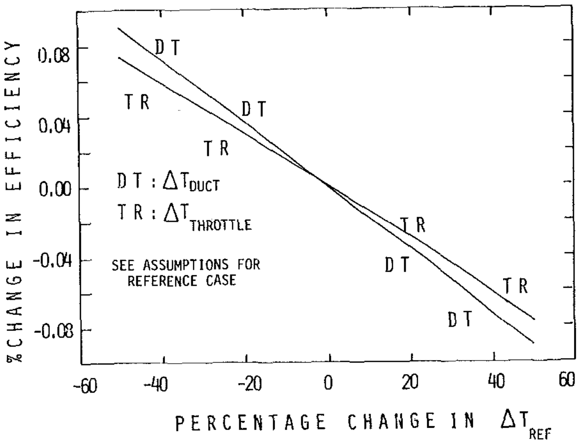


Fig. 13(*b*) The change in efficiency of the baseline case as the assumed *T* are varied

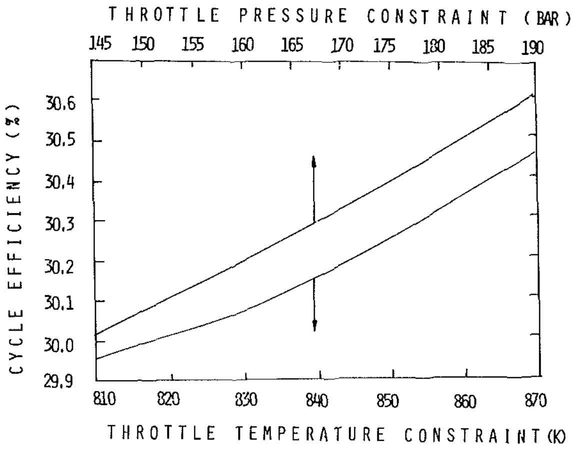


Fig. 13(*c*) The effect of the throttle constraints on the first-law cycle efficiency

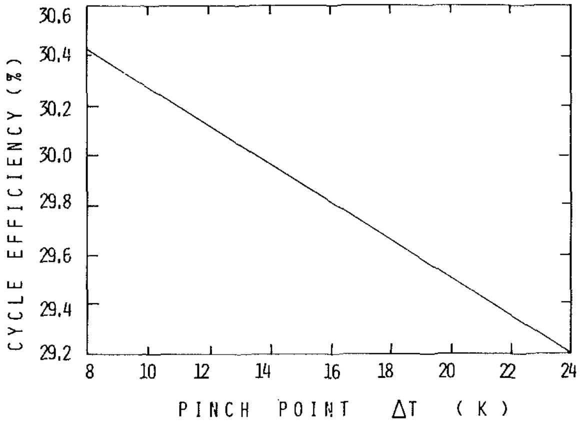


Fig. 13(*d*) The efficiency of the cycle as the minimum pinch point temperature difference is increased

The 166.5 bar maximum throttle pressure represents the typical constraint based upon current technological practice. A polynomial curve fit for the first-law efficiency of the two-pressure cycle with this throttle pressure constraint and maximum throttle temperature of 867 K (1100°F) as a function of nondimensional gas temperature gives

$$\eta_1 = -0.5915 + 0.4801\,\Gamma - 0.06064\,\Gamma^2 \tag{8}$$

$$2.337 < \Gamma < 3.513$$

For the same cycle, except with a throttle temperature constraint of 811 K (1000°F), the equation is

$$\eta_1 = -0.7365 + 0.5925\,\Gamma - 0.08229\,\Gamma^2 \tag{9}$$

$$2.337 < \Gamma < 3.306$$

The lower limits of Γ represent the lowest T_{4A} that would be practical to be considered for a bottoming cycle. However, the high limits are not the highest temperatures encountered in practice. Instead, these limits represent the gas temperatures at which the dual-pressure cycle effectively becomes a single-pressure cycle with reheat, as the mass fraction in the low-pressure stream goes to zero in the cycle optimization. In order to find the cycle efficiencies for the entire temperature range of interest, the single-pressure with reheat cycle must be analyzed for the higher temperature range [5]. Combining the results of single-pressure reheat cycles and the dual-pressure cycles, the optimized cycle efficiencies for the entire temperature range are found. These efficiencies for the two throttle temperature constraints are as follows:

$$\eta_1 = -0.6327 + 0.5094\,\Gamma - 0.06578\,\Gamma^2 \tag{10}$$

$$2.337 < \Gamma < 3.652$$

$$\eta_1 = -0.7285 + 0.5855\,\Gamma - 0.08081\,\Gamma^2 \tag{11}$$

$$2.337 < \Gamma < 3.652$$

These polynomials can be used to approximate the optimum bottoming-cycle efficiency, given a gas turbine exhaust temperature, thus aiding combined cycle analyses.

Finally, the sensitivity of the cycle efficiency to various cycle parameters is revealed in Fig. 13(*a-d*). The base case for the sensitivity curves is the 125.2 bar/811 K (1800 psig/1000°F) cycle with T_{4A} of 911 K. In Fig. 13(*a*) the reference pressure drops are 6, 5, and 5 percent for the throttle pressure drop, reheat pressure drop, and superheater pressure drop, respectively. The base efficiency for the reference case is 29.97 percent and the figure shows the change in efficiency as the pressure drops are changed. The reference temperature drops for Fig. 13(*b*) are 2.2 K (4°F) for the exhaust gas temperature loss between gas turbine flange and boiler inlet, and 2.8 K (5°F) for the steam temperature drop in the piping between the superheater exit to the throttle.

The sensitivity of cycle efficiency to the throttle state is shown in Fig. 13(*c*). It is seen that the cycle efficiency increases nearly linearly with increasing throttle pressure and throttle temperature.

The effect of the minimum pinch point temperature difference constraint on the cycle efficiency is illustrated in Fig. 13(*d*). Clearly, a smaller difference results in a higher efficiency. In the limit with infinite heat transfer surface area, the difference becomes zero and all of the exergy from the gas side is transferred to the steam side. However, a reasonable difference must be chosen for economic reasons.

Conclusions

The energy balance approach to cycle analysis makes no distinction between heat and work and thus does not quantify the "quality" of the heat. Therefore, this method of analysis does not show which system components the cycle losses should be attributed to. The exergy balance method of analysis enables all loss sources to be quantified.

An exergy balance analysis of steam bottoming cycles shows that the selection of a two-pressure evaporator cycle results in an efficiency improvement principally through the reduction of the heat transfer irreversibility from that for a single-pressure system. For typical turbine exhaust gas temperatures of 800–850 K, using two evaporator pressures instead of one reduces the exergy loss in heat transfer from approximately 15 percent of the exhaust gas availability to 8 percent. This corresponds to a gain of about 3 percentage points in first-law efficiency. Further reductions of heat transfer irreversibilities are possible through increasing the number of evaporator pressures or using binary fluid mixtures, but the improvements over an optimized two-pressure system at typical exhaust temperatures are unlikely to exceed 3–4 percent of the exhaust gas availability, which corresponds to about one percentage point gain in first-law efficiency.

With cycle parameter constraints and assumptions consistent with current technology, the first-law efficiency of optimized two-pressure steam bottoming cycles are correlated by equations (8) and (9) given above.

Acknowledgments

The authors wish to acknowledge the support of the Gas Turbine Division of the General Electric Company and of the Electric Power Research Institute. The discussions held with Messrs. G. A. Cincotta, A. Perugi, and Dr. A. Cohn were invaluable during the course of the work.

References

1 Kalina, A. I., "Combined Cycle and Waste-Heat Recovery Power Systems Based on a Novel Thermodynamic Energy Cycle Utilizing Low-Temperature Heat for Power Generation," ASME Paper No. 83-JPGC-GT-3.

2 Irvine, T. F., Jr., and Liley, P. E., *Steam and Gas Tables With Computer Equations*, Academic Press, New York, 1984.

3 Foster-Pegg, R. W., "Steam Bottoming Plants for Combined Cycles," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 100, 1978, pp. 203–211.

4 Spencer, R. W., Cotton, K. C., and Cannon, C. N., "A Method for Predicting the Performance of Steam Turbine-Generators 165000 kW and Larger," ASME Paper No. 62-WA-209.

5 Chin, W. W., "Computer Modelling of Steam Bottoming Cycles," S. M. Thesis, Department of Mechanical Engineering, M. I. T., Cambridge, MA, Jan. 1987.