Transactions of the ASME®

Technical Editor
H. L. JULIEN (1998)

Associate Technical Editors
Advanced Energy Systems
M. J. MORAN (1999)
Gas Turbine
D. COOKE (1999)
H. NELSON (1999)
J. PETERS (1999)
J. N. SHINN (1996)
Internal Combustion Engines
D. ASSANIS (1999)
Power
D. LOU (1998)

BOARD ON COMMUNICATIONS
Chairman and Vice President
R. MATES

OFFICERS OF THE ASME
President, **R. J. GOLDSTEIN**
Executive Director, **D. L. BELDEN**
Treasurer, **J. A. MASON**

PUBLISHING STAFF
Managing Director, Engineering
CHARLES W. BEARDSLEY
Director, Technical Publishing
PHILIP DI VIETRO
Managing Editor, Technical Publishing
CYNTHIA B. CLARK
Managing Editor, Transactions
CORNELIA MONAHAN
Production Coordinator
VALERIE WINTERS
Production Assistant
MARISOL ANDINO

Transactions of the ASME, Journal of Engineering for Gas Turbines and Power (ISSN 0742-4795) is published quarterly (Jan., April, July, Oct.) for $185.00 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017. Periodicals postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Engineering for Gas Turbines and Power, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.
CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.
PRICES: To members, $40.00, annually: to nonmembers, $185.00. Add $30.00 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 (B7.1, par. 3).

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

Journal of Engineering for Gas Turbines and Power

Published Quarterly by The American Society of Mechanical Engineers

VOLUME 118 • NUMBER 4 • OCTOBER 1996

TECHNICAL PAPERS

(Contents continued on p. 896)

ANNOUNCEMENT

A Second-Law-Based Optimization: Part 1—Methodology

Y. M. El-Sayed
President,
Advanced Energy Systems Analysis,
41658 Higgins Way,
Fremont, CA 94539

This paper deals with the optimization of complex energy systems given a cost objective function. The optimization uses a decomposition strategy based on the second law of thermodynamics and a concept for costing the components of a system. A large number of nonlinear decision variables can be optimized with enhanced convergence to an optimum. The paper is in two parts. In this part, the methodology is described. In Part 2, the methodology is applied to a simple energy system of 10 components and 19 manipulated decision parameters. The system is treated once as a single purpose combined cycle and once as a power-heat cogenerating system. The results of the application are summarized and evaluated. Further development is encouraged.

Introduction

Optimal energy systems having a cost objective function have been the target of many investigators in the last decade. Gaggioli and El-Sayed (1989), Evans (1980), Melli and Sciubba (1992), Valero et al. (1993), Tsatsaronis and Lin (1990), and Valero et al. (1992) are among the major contributors. The last contribution is an inspiring combined effort by Valero, Tsatsaronis, von Spakovsky, Frangopoulos, and Lozano.

In spite of the advancements achieved, the costing of the components of a system lacks rationality and optimization techniques are still slow.

In this paper two implementations are introduced:

- Rationally derived costing equations from the design models of the system components as the communicating channels between system analysis and component design analysis, or in other words, as a decomposition strategy of the two analyses.
- Enhanced optimization by partly decomposing a system into dissipation–dissipator pairs using the second law of thermodynamics.

Both implementations are natural extensions of the current state of the art.

In this part, the implementations are addressed and the methodology is described. In Part 2, the methodology is demonstrated and tested on two application examples. The two examples were treated in more detail separately by the author (1995a, b).

Derivation of Costing Equations

The most convenient way of costing many manufactured devices is in terms of a unit price of a characterizing extensive size parameter of the device such as mass, volume, area, or length:

$$Z = c_a{}^*A \qquad (1a)$$

c_a is the unit price and A is the extensive parameter. Strength and efficiency parameters as well as parameters of other relevant features are implicit in c_a and are usually stated separately.

A system model that computes the performance of a system in terms of the performance characteristics of its components is essentially a thermodynamic model. Power, mass rate, heat rate, exergy rates, pressure, temperature, composition, enthalpy, entropy, specific volume, pressure loss, heat loss, temperature difference, adiabatic efficiency, heat exchange effectiveness, extent of reaction, and stoichiometry excess are the major system variables $\{V_s\}$. For an optimization that minimizes or maximizes a cost objective function, it is desirable that the system model compute costs as well. Therefore it is desirable to express the characterizing parameter A of a device as a function of the same system variables. The system variables $\{V_s\}$ may be grouped into three groups: extensive V_x, intensive V_i, and performance V_p. Thus the desirable costing equation becomes

$$Z = c_a{}^*A(\{V_s\}) = c_a{}^*A(\{V_x\}, \{V_i\}, \{V_p\}) \qquad (1b)$$

The variables of Eq. (1*b*) may be dependent or independent system variables. Including $\{Z\}$ as expressed by Eq. (1*b*), the system model becomes a *performance-cost* model.

Whether Eq. (1*a*) or Eq. (1*b*) is used, a device cost Z is determined in terms of the variables of a design process $\{V_d\}$ and the variables of a manufacturing process $\{V_m\}$, i.e.:

$$Z = Z(\{V_d\}, \{V_m\}) \qquad (1c)$$

Let Eq. (1*c*) be expressed by the unit price c_a and the characteristic shape parameter A. Let A be a function of design variables only and c_a be a function of both design and manufacturing variables, i.e.,

$$Z = c_a(\{V_d\}, \{V_m\})^*A(\{V_d\}) \qquad (1d)$$

The unit cost c_a can be expressed for families of devices from commercially available price information in terms of materials and the intensity of operation (levels of pressure, temperature, corrosion, ...). The connection between Eqs. (1*b*) in terms of thermodynamic variables, and (1*d*) in terms of design and manufacturing variables for a device, is *a design process* through which $A\{V_s\}$, and $A(\{V_d\})$ are related by design equations. The variables of $\{V_x\}$, $\{V_i\}$, $\{V_p\}$ of Eqs. (1*b*) are a subset of the component design variables $\{V_d\}$. The design relations and the design variables constitute a design model.

The number of the design variables $\{V_d\}$ minus the number of design equations represents the degrees of design freedom for a device. The number of variables of the costing Eq. (1*b*) may be less, equal, or more than the design degrees of freedom. If less, the difference may be partly used to capture the best design experiences and partly to minimize A. If equal, the problem reduces to the exchange of equal numbers of design and

Contributed by the Advanced Energy Systems Division for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Advanced Energy Systems Division September 28, 1995; revision received April 19, 1996. Associate Technical Editor: M. J. Moran.

thermodynamic variables. If more, the difference represents system variables forced on the system, thus reducing the system degrees of freedom. In most situations, however, the number of variables of the costing Eq. (1*b*) is appreciably less than the design degrees of freedom.

A design model treats some of the variables of the costing Eq. (1*b*) as independent variables (inputs to the model) and the rest as dependent variables (outputs of the model). The costing Eq. (1*b*) is generated by running the design model with various sets of inputs such that the variables of the costing equation, whether inputs or outputs of the design model, cover the range of interest in which the costing equation is intended to be applicable for the system at hand. It is *important* to note that design models are not only useful for deriving and verifying costing equations; they can also be used to predict off-design performance by fixing a designed geometry and varying input parameters.

Many selections are possible for what should represent the extensive parameter A of a device and for the form of its function as expressed by Eq. (1*b*). In the current study the *surface area* is selected for A. This may represent a heat exchange surface in heat exchangers or momentum exchange surface in turbomachines. For the form of the function, the following equation is selected:

$$A = k_a * \Pi\{V_x^{nx}\}\{V_i^{ni}\}\{V_p^{np}\} \tag{2}$$

This form is favored because it is adequate for correlating equipment cost and also because it enhances optimization, as will become clear later. The thermodynamic variables needed to express adequately the characterizing surface of a component are not unique but their number seldom exceeds four. The costing equation becomes:

$$Z = c_a * k_a * \Pi\{V_x^{nx}\}\{V_i^{ni}\}\{V_p^{np}\} \tag{3}$$

Implicit in c_a, are other relevant features not expressed by $\{V\}$ such as materials, complexity of surface, and intensity of operation.

The essential *communication* between a thermodynamic model of a system and the design models of its components is shown in Fig. 1. The communication through the two dotted lines with arrows is active only during the generation of costing equations. Once the costing equations of the components are generated, the system model operates on its own as a performance-cost model. Costing equations in the form of Eq. (3) are given in Table 2 of Part 2. The coefficients and the exponents (c_a, k_a, $\{n\}$) are deduced using specific design models described in the appendix of Part 2. (It is hoped that generating more costing equations based on different design assumptions and dealing with different design variants of components will help establish the statistical nature of their coefficients and exponents and will increase the confidence in their applicability.)

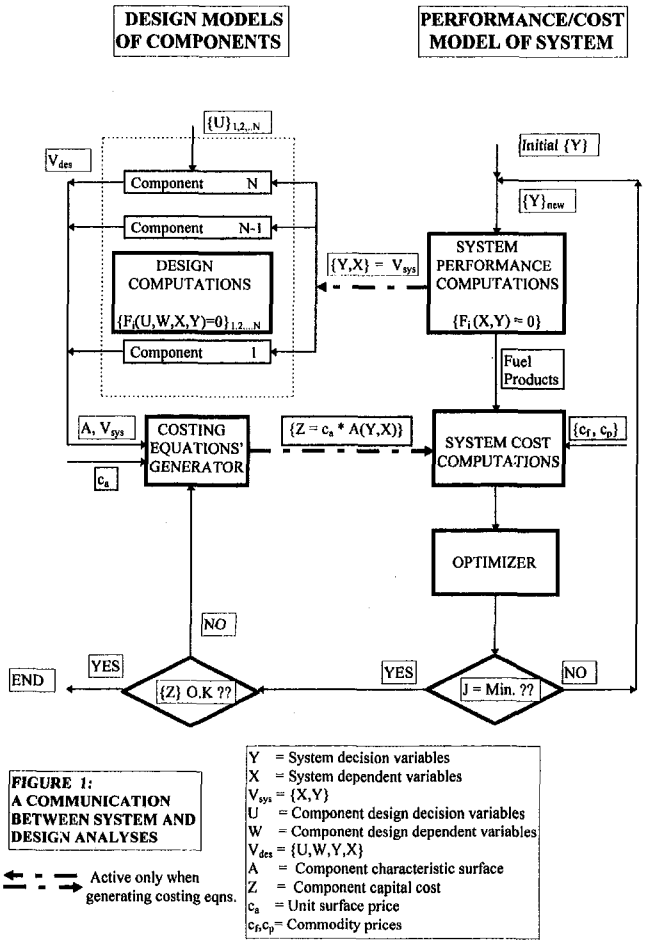


FIGURE 1: A COMMUNICATION BETWEEN SYSTEM AND DESIGN ANALYSES

Fig. 1 A communication between system and design analyses

A Second-Law-Based Decomposition

Strategies to decompose a system into subsystems for the purpose of optimization may be achieved by adding priced terms to the objective function, each of which in effect is zero (a constraint); but when combined with the objective function, the features of a decomposition emerge. The Lagrangian is the augmented objective function by the added terms and the prices

Nomenclature

A = surface area, size characterizing parameter of a device
c = price per specified unit; c_F of feed, c_P of product; c_f, c_p exergy-based of feed and product, respectively; c_d of dissipation, $c_{d\lambda}$ Lagrangian, c_{da} average (Eq. (21)), c_{dd} marginal (Eq. (19)), c_{d0} special case (Eq. (20)); c_z of capital
D = exergy destruction (dissipation); D_P, D_T, D_C mechanical, thermal, and chemical dissipations; d dissipation per unit mass or unit reference mass
E = exergy rate; E_p product, E_f feed
ΔE = change in exergy
e = efficiency parameter for adiabatic efficiency = $\eta/(1-\eta)$
F = feed; function
J = objective function; J_2 second-law objective, J_R remainder = $J - J_2$, J_i local of component i
k = coefficients; k_d and k_e of exergy destruction; k_z of capital; and k_a of surface or a volume function of a device
n = exponent; n_e of exergy destruction, n_z of capital
P = product; pressure
U = decision design variable of a component
V = variable; V_s a system variable (thermodynamic), V_x, V_i, V_p system variables as extensive, intensive, and performance; V_m manufacturing process variable; V_d a design variable of a device
X = system-dependent variable (thermodynamic variable)
Y = system-decision variable; Y_L local, Y_G global (thermodynamic variables)
Z = equipment cost
η = isentropic efficiency
λ = Lagrange multiplier
{ } = set of variables

are undetermined Lagrange multipliers. The decomposition features depends on the selected constraints augmenting the objective function. One early feature of second-law-based decomposition is given by El-Sayed and Evans (1970). Other features were used by Tsatsaronis and Lin (1990) and by Valero et al. (1992, 1994). The present second-law-based decomposition uses another different feature. One constraint at the system's boundaries is added to the objective function, which is the overall exergy balance of the system multiplied by a price yet to be determined, and what emerges is the cost of dissipators and the cost of dissipations in pairs. The advantage of this feature is that *one* price is used while a possibility of decomposing the system into its components is created.

A suitable cost objective function for an energy system of single feed and single product (e.g., a power plant) is the minimization of production cost minus product revenues under fixed unit prices $\{c\}$:

$$J = c_F * F + \Sigma\, c_z * Z - c_P * P \tag{4}$$

Profit $= -J$ and is maximized.

Focusing on the system boundaries, overall mass, energy, and entropy balances should be satisfied as well as the sizing parameter of the system. If exergy balance is based on these three balance equations, two multipliers would adequately describe the conditions at the system's boundaries: one associated with the exergy balance equation, and one with the sizing parameter, which in this simple case is usually related to the product rate P or the feed rate F. If the system's computational algorithm permits the selection of the sizing parameter, then one boundary multiplier would take care of system decomposition.

Equation (4) can be transformed to a second-law objective function by rating fuel and product prices per unit exergy and adding the overall exergy balance equation as a priced constraint:

$$c_d * (E_p + \Sigma\, D + \Sigma\, E_j - E_f) = 0 \tag{5}$$

The following Lagrangian is obtained:

$$L = \Sigma\, (c_d * D + c_z * Z) + \Sigma\, c_d * E_j + (c_f - c_d) * E_p - (c_p - c_d) * E_p \tag{6}$$

where

$$J_2 = \Sigma\, (c_d * D + c_z * Z) \tag{6a}$$

and

$$J_R = \Sigma\, c_d * E_j + (c_f - c_d) * E_f - (c_p - c_d) * E_p \tag{6b}$$

J_2 becomes a second-law transformation of the objective function J if $\{D\}$ and $\{Z\}$ are grouped as independently interacting pairs and J_R is a constant K. J_R becomes a constant when the cost of the wasted exergy $\Sigma c_d * E_j$ is negligible and when either the feed is constant or the product is constant. The price of dissipation c_d becomes the unit exergy price of feed c_f when given the product rate E_p and becomes the unit exergy price of product c_p when given the feed rate E_f. Associating c_d with c_p and c_f is helpful in guessing an initial value for c_d as a Lagrange multiplier in the cases not tied to the specific assumptions just mentioned.

Note that the dissipator–dissipation terms of Eq. (6*a*) include any waste treatment before dumping the wasted exergy ΣE_j in the environment.

Enhanced Optimization

In general, Eq. (6*a*) presents dissipations and dissipators as interacting pairs and not as independent pairs. A system decision variable, in principle, influences all items of Eq. (6*a*) to varying degrees. Considering the interactions among pairs as secondary interactions helps in gaining insight to the complexity of the interactions of the decision variables. Therefore, the following two idealizations are introduced:

- The localized influence of certain decision variables.
- The form $k * Y^n$ for the local influence.

The first *rationalizes* a second-law decomposition of the system. The second *enhances* optimization.

The first idealization states that when the influence of a decision variable is primarily localized, the decision variable is considered local. Otherwise it is global. The global effects of the local decision variables are ignored and hence, Eq. (6*a*) represents a decomposed system of components having local objective functions and the local decisions can then be optimized one by one. Performance decisions of components (efficiency parameters) can be considered local decisions in *most,* but *not* all, situations. A suitable two-step optimization, repeatable if necessary, is to minimize J with respect to $\{Y_L\}$ given $\{Y_G\}$, then minimize J with respect to $\{Y_G\}$ given $\{Y_{L\text{opt}}\}$.

Optimization of Local Decision Variables

For an exergy destruction price c_d, the local optimization problem for a component i becomes:

$$\underset{Y_i}{\text{Minimize}}\ J_i = c_d * D_i + c_z * Z_i \tag{7}$$

where Y_i is the component performance decision variable assumed local. D_i and Z_i have opposing trends that are best expressed by a generalized version of the equation $J = a * Y + b/Y$, known as Kelvin equation:

$$J_i = k_e * Y_i^{ne} + k_z * Y_i^{nz} \tag{7a}$$

where n_e and n_z are constant exponents of opposite signs and k_e and k_z are constant coefficients.

If a component has more than one local decision, decomposing its dissipation into thermal, mechanical and chemical helps preserving the same form of Eqs. (7) and (7*a*) for each decision:

$$\underset{Y_P,\, Y_T,\, Y_C}{\text{Minimize}}\ Ji = J_P + J_T + J_C = c_d * (D_P + D_T + D_C) + c_z * Z_i \tag{7b}$$

where $c_z * Z_i$ has the form of Eq. (3) and

$$J_P = k_{eP} * Y_P^{nep} + k_{zP} * Y_P^{nzp}$$

$$J_T = k_{eT} * Y_T^{net} + k_{zT} * Y_T^{nzt}$$

$$J_C = k_{eC} * Y_C^{nec} + k_{zC} * Y_C^{nzc}$$

The exponent n_z (or the set n_{zp}, n_{zt}, n_{zc}) is already available from the costing equations having the form of Eq. (3). The dissipation cost $c_d * D_i$ of a component can be generally expressed in the same way as the capital cost Z. Dissipation in the form of Eq. (2) will have V_x a mass rate with $n_x = 1$, V_i a pressure or a temperature input parameter with $n_i \cong 0$ and V_p the decision Y_i with $n_p = n_e$. The assumption $n_i \cong 0$ is used because D_i is weakly affected by any input pressure or temperature, and hence the dissipation takes the form:

$$c_d * D_i = c_d * M_i * d_i = c_d * M_i * k_d * Y_i^{ne} = k_e * Y_i^{ne} \tag{8}$$

d_i is a dissipation per unit mass and k_d is a constant depending on Y_i and $k_e = c_d * M_i * k_d$. Equation (8) should include pressure and/or temperature terms when $n_i \neq 0$. The thermodynamic process of a component determines the exponent n_e. Note that the thermodynamic process of a component is already available as an element of the system thermodynamic model. Pressure drop decisions have $n_e = 1$. Efficiencies $e = \eta/(1 - \eta)$ have n_e close to -1. For some performance parameters such as an

approach temperature difference or an extent of reaction n_e may deviate appreciably from 1 (or -1).

If Eqs. (7) and (7a) are identical, then k_e and k_z, given n_e and n_z, become:

$$k_e = c_d * D_i / Y_i^{ne} \tag{9}$$

$$k_z = c_z * Z_i / Y_i^{nz} \tag{10}$$

Equation (7a) (or any form of Eq. (7b)) has an analytical solution for the unconstrained optimum decision variable by setting the derivative of J_i with respect to Y_i equal to zero.

$$Y_{i\text{opt}} = [-(k_z * n_z)/(k_e * n_e)]^{(1/(ne-nz))} \tag{11}$$

Substituting Eqs. (9) and (10) into Eq. (11):

$$Y_{i\text{opt}} = Y_i * [(-n_z/n_e) * (c_z * Z_i)/(c_d * D_i)]^{1/(ne-nz)} \tag{12}$$

If k_e and k_z are constants then Eq. (12) gives optimum Y_i in one step from a feasible solution. Since D_i and Z_i involve variables (basically mass rates), which may depend partly on Y_i (a global effect of Y_i), k_e and k_z deviate from constant values and the derivative of J_i of Eq. (7a) with respect to Y_i gives

$$k_e * (dY_i^{ne}/dY_i) + Y_i^{ne} * (dk_e/dY_i) + k_z * (dY_i^{nz}/dY_i) + Y_i^{nz} * (dk_z/dY_i) = 0 \tag{13}$$

If the sum of $Y_i^{ne} * (dk_e/dY_i)$ plus $Y_i^{nz} * (dk_z/dY_i)$ is close to zero (mild global effect), or if k_z and k_e change such that the change in their ratio is insignificant (converging global effect), then k_e and k_z will adjust to constant values in few steps from a feasible solution and Y_i settles to its optimum value. In this case Eq. (12) acts as an updating equation of the decision variable Y_i. Otherwise Y_i may not qualify as local and updating Y_i causes its optimum to fluctuate and transmit fluctuations to the systems objective function (non-converging global effect). Fast convergence to optimum by equation (12) for the local decision variables implies that the global influence is *mild* or *converging*. The application examples of Part 2 have no problem with convergence when all performance decision variables are assumed local. No more than four system computations were needed in most cases.

Thus, once the local decisions are identified and their $\{n_e, n_z\}$ are known, their optimization is fast irrespective of their number. However, it is *important* to notice certain *limitations:* The exponents $\{n_e, n_z\}$ are not absolutely constants but only around a reference system solution. The decisions $\{Y_i\}$ are not absolutely local but their influence is primarily localized. Local optimizations are expected to be unconstrained, but imperfections in modeling may occasionally force the lower or the upper bound of a local decision to be active. Overly low and overly high exergy destruction prices activate bounds as well. It should also be expected that occasionally, though not in the present study, a decomposed dissipation (e.g., thermal) may be influenced by two local decisions that should be computed simultaneously.

Optimization of Global Decision Variables

Not all the system decision variables can be assumed local. For those variables left as global or treated as local while showing nonconverging global effect, the objective function is Eq. (4) for the system undecomposed, but not Eqs. (7), which represent the decomposed system. A suitable nonlinear programming algorithm must be invoked. A simplified gradient-based approach that ignores cross second derivatives is used (El-Sayed, 1989) having the following updating equation:

$$Y_{\text{new}} = Y_{\text{old}} \pm \Delta Y \tag{14}$$

$$\Delta Y = ABS[0.5 * (Y_2 - Y_1)/(g_2 - g_1) * (-g_1)] \tag{14a}$$

$$g_1 = (J_0 - J_1)/(Y_0 - Y_1) \tag{14b}$$

$$g_2 = (J_2 - J_0)/(Y_2 - Y_0) \tag{14c}$$

$$Y_2 > Y_0 > Y_1 \tag{14d}$$

The updating Eq. (14) requires three system computations to obtain three neighboring values of the objective function, assuming for example Y_0, 1.05 Y_0, and 0.95 Y_0. The $\pm$ sign of Eq. (14) is selected to direct the change in the favored direction, because zero gradient represents both maximum and minimum.

Note that for each global decision, two system computations are needed (Y_1 and Y_2) besides the current Y_0 in each update, and the number of updates to reach an optimum is usually large. For local decisions, only the current system computation is needed and the number of updates is relatively much smaller.

Extending Local Optimization to Multipurpose Systems

The objective function for a multiple-feed multiple-product plant (multiple commodities) becomes:

$$J = \Sigma\, c_F * F + \Sigma\, c_z * Z - \Sigma\, c_P * P \tag{15}$$

Basing prices per unit exergy of the commodities and adding the overall exergy balance as a constraint priced at an exergy destruction price c_d yet to be determined, the following Lagrangian is obtained:

$$L = \Sigma\, c_f * E_f + \Sigma\, c_z * Z - \Sigma\, c_p * E_p + c_d * (\Sigma\, D + \Sigma\, E_j + \Sigma\, E_p - \Sigma\, E_f) = \Sigma\, (c_z * Z + c_d * D) + \Sigma\, c_d * E_j + \Sigma\, (c_f - c_d) * E_f - \Sigma\, (c_p - c_d) * E_p \tag{16}$$

Let

$$J_2 = \Sigma\, (c_z * Z + c_d * D) \tag{17}$$

$$J_R = \Sigma\, c_d * E_j + \Sigma\, (c_f - c_d) * E_f - \Sigma\, (c_p - c_d) * E_p \tag{18}$$

where J_2 may be defined as a second-law objective function and J_R as a remainder cost function.

Any positive value of c_d minimizes J_2 but not necessarily J_R. To determine the Lagrangian dissipation price $c_{d\lambda}$ that minimizes the actual objective function $J = J_2 + J_R$ through minimizing J_2, is simply to minimize J_2 with different values of the dissipation price c_d. $\{dJ_2/dY\} = 0$ gives provisional $\{Y\}_{\text{opt}}$ for the local decision variables by which a solution is obtained and a value of J is computed. The decision set $\{Y\}$, corresponding to the least computed value of J, is considered the optimum set of local decision variables and the corresponding dissipation price is the lagrangian price $c_{d\lambda}$. An automated way to find $c_{d\lambda}$ is to use Eq. (14) of global optimization with c_d as the only decision variable.

Two average exergy destruction prices are found to improve J and to give a head start to the minimization of J through the minimization of J_2 in most cases. One is a marginal average c_{dd} (a ΔE price) and the other is simply a mean average c_{da}.

Marginal Average Price. If a set of local decision variables $\{Y\}$ produces the same improvement in the Lagrangian ΔL (or ΔJ) as in the second-law objective ΔJ_2, then ΔJ_R must equal zero (J_R is a constant K or zero). For J_R = constant, the condition of $\Delta J_R = 0$ gives the following exergy destruction price:

$$c_{dd} = (\Sigma\, c_f * \Delta E_f - \Sigma\, c_p * \Delta E_p)/(\Sigma\, \Delta E_f - \Sigma\, \Delta E_p - \Sigma\, \Delta E_j) = (\Sigma\, c_f * \Delta E_f - \Sigma\, c_p * \Delta E_p)/\Sigma\, \Delta D \tag{19}$$

Equation (19) states that using c_{dd} to minimize J_2, then the change of J_2 is the same as the change of J, but does not assure the best change. Also Eq. (19) does not always guarantee a

positive c_d, as will be noticed in Part 2, because global effects whether mild or converging are ignored.

In general, all commodities (feeds and products) are dependent variables on the sizing parameters of the plant. A sizing parameter is usually an independent nominal or actual commodity rate. The number of the independent sizing parameters is often one and selom exceeds two. If actual rates determine the size of the plant, then their ΔE's = 0 and Eq. (19) includes *only* the dependent commodities. The changes of these dependent commodities during optimization affects slightly the price of dissipation c_{dd}.

The condition of $J_R = 0$ is a special case of the constant K and this case gives the following exergy destruction price:

$$c_{d0} = (\Sigma\ c_f{*}E_f - \Sigma\ c_p{*}E_p)/(\Sigma\ E_f - \Sigma\ E_p - \Sigma\ E_j) \quad (20)$$

Equation (20) is more convenient to use than Eq. (19) since it does not deal with differences. Equation (20) like Eq. (19) does not guarantee a positive dissipation price.

Average Price. Because of the possibility of a negative exergy dissipation price from Eqs. (19) and (20), a simple mean average price as given by the following equation is considered:

$$c_{da} = (\Sigma\ c_f{*}E_f + \Sigma\ c_p{*}E_p)/(\Sigma\ E_f + \Sigma\ E_p) \quad (21)$$

c_{da} may be the average of all commodities or the average of the dependent commodities only. In this study the average of the dependent commodities only is used. Equations (20) and (21) give the same result when the dependent commodities are all products or all feeds.

The marginal price c_{dd} from Eq. (19), the average price c_{da} from Eq. (21), and the searched Lagrangian price $c_{d\lambda}$ are investigated for local optimization and their results are compared. The small changes in c_{dd} and in c_{da} as the exergies of the dependent commodities vary during optimization are accounted for automatically. Optimizing the system with different values of c_d proves that $c_{d\lambda}$ does exist. Minimum J flattens over a small range around $c_{d\lambda}$. Above this range, the increase in efficiency does not pay off, and below it, the decrease of efficiency also does not pay off. The Lagrangian price $c_{d\lambda}$ depends on the unit prices involved in the objective function and the sizing commodity. The minimum value of J using $c_{d\lambda}$ is compared with those obtained using c_{dd} and c_{da}. The agreement is satisfactory in some cases but not in others.

However, thus far, the complexity of predicting $c_{d\lambda}$, particularly for systems with multiple feeds and products, is solved by computing the local optimum for more than one dissipation price. Four to eight dissipation prices are usually sufficient due to the relatively flat relation between J and c_d when J_2 is minimized.

Concluding Remarks

- The proposed decomposition strategy extends second-law analysis to the minimization of the cost of energy intensive systems.
- The proposed concept of costing equations provides a rational basis for costing the components of a system in terms of their design methodologies.
- As shown in Part 2, the foregoing offer a *sense of direction* for the search for optimal decision variables of power and cogeneration systems.
- The communication established between performance and design to generate costing equations is also useful in predicting off-design performance of the components of a system and hence that of the system.

Acknowledgments

The recommendations and the constructive criticism of Professor Richard A. Gaggioli of Marquette University are highly appreciated. Thanks are also due to Professor Michael J. Moran of Ohio State University for his support and helpful remarks and to Dr. Michael R. von Spakovsky of EPFL, Lausanne, Switzerland, for his valuable comments.

References

El-Sayed, Y. M., and Evans, R. B., 1970, "Thermoeconomics and the Design of Heat Systems," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 92, pp. 27–35.

El-Sayed, Y. M., 1989, "A Decomposition Strategy for Thermoeconomic Optimization," ASME *Journal of Energy Resources Technology,* Vol. 111, pp. 111–120.

El-Sayed, Y. M., 1992, "Targeting an Optimal Design of Heat Exchange in a Power System," *Thermodynamics and the Design Analysis and Improvement of Energy Systems,* ASME AES-Vol. 27/HTD-Vol. 228, pp. 347–353.

El-Sayed, Y. M., 1995a, "The Integration of Thermodynamic Analysis and Design Analysis for Optimal Thermal Design," *Proc. International Conference ECOS'95,* Istanbul, Turkey, July 11–14, Vol. 2, pp. 375–386.

El-Sayed, Y. M., 1995b, "Repowering Second-Law Based Optimization," presented at ASME International Mechanical Congress and Exposition, Nov. 12–17, San Francisco, CA.

Evans, R. B., 1980, "Thermoeconomic Isolation and Essergy Analysis," *Energy,* Vol. 5, p. 805.

Gaggioli, R. A., and El-Sayed, Y. M., 1989, "A Critical Review of Second-Law Costing Methods," ASME *Journal of Energy Resources Technology,* Vol. 111, pp. 1–15.

Melli, R., and Sciubba, E., 1992, "Second-Law Based Synthesis and Optimization: A Third Generation AI Code," presented at ECOS'92, Zaragoza, Spain.

Tsatsaronis, G., and Lin, L., 1990, "On Exergy Costing in Exergoeconomics," ASME AES-Vol. 21, p. 1.

Valero, A., Tsatsaronis, G., von Spakovsky, C., Frangopoulos, X., and Lozano, M., 1992, "Thermoeconomic Optimization of the CGAM Problem," presented at ECOS'92, Zaragoza, Spain.

Valero, A., et al., 1993, "Structural Theory of Thermoeconomics," ASME AES-Vol. 30, pp. 189–198.

Valero, A., Tsatsaronis, G., von Spakovsky, C., Frangopoulos, Lozano, M., Serra, L., and Pisa, J., 1994, "CGAM Problem: Definitions and Conventional Solution," *The International Journal of Energy,* Vol. 19, No. 3.

A Second-Law-Based Optimization: Part 2—Application

Y. M. El-Sayed
President,
Advanced Energy Systems Analysis,
41658 Higgins Way,
Fremont, CA 94539

This paper deals with an application of the second-law-based optimization methodology described in Part 1 of this paper. The methodology is applied to a simple energy system of 10 components and 19 manipulated decision parameters. The system is treated once as a single-purpose combined cycle and once as a power-heat cogenerating system. The results of the application are summarized and evaluated. The required computer programs are described. Further investigations are suggested.

The Application Examples

The system considered for the two application examples is shown in Fig. 1. The system has 10 components and the system performance-cost model has 35 parameters that can be independently specified (decision variables). For the current study, 19 of the independent parameters are manipulated for system optimization; two are global and 17 are local (performance decision variables). The rest of the decisions are not manipulated for optimization. These are: 2 global, 10 boundary parameters, and 4 conditions. The two global decisions are the top cycle temperature T_3 = 870°C (1600°F) (implying uncooled gas turbine blades) and condensate temperature T_8 = 38°C (100°F) for the single-product case and 120°C (250°F) for the two-product case.

The 10 parameters determined by boundary conditions are a sizing parameter taken as the natural gas feed to the system of rate 4 kg/s (8.8 lb/s), and a pressure, a temperature, and/or a composition at inlet to or exit from system. Air of specified composition enters at 1.013 atm (14.7 psia) and 27°C (80°F). Natural gas of specified composition enters at 15°C (59°F). Exhaust gases leave at 1.013 atm (14.7 psia). For the single-product case, the cooling water enters at 1.03 atm (15 psia) and 21°C (70°F) and leaves at 1.03 atm. For the two-product case, water enters at 3.8 atm (55 psia) and 77°C (170°F) and is delivered at 4.5 atm (65 psia). The four conditions are fuel pressure $P_5 = P_2$, saturated liquid at locations 8, 10, and saturated vapor at 11.

The manipulated global decisions are the pressure exiting the compressor P_2 and the entering steam turbine pressure P_6. The local decisions are the adiabatic efficiencies of the power machines 1 to 5, the combustor pressure loss, the shell-side and tube-side pressure losses of exchangers 7 to 10, and terminal temperature differences $T_4 - T_6$, $T_{13} - T_{10}$, and $T_7 - T_{17}$. The effectiveness or logarithmic mean temperature difference could have been used instead of a terminal difference. Temperature differences have a lower bound of 2.8°C (5°F). Pressure losses/inlet pressures have an upper bound of 0.5. The temperature of the delivered hot water in the two-product case has a lower bound of 110°C (230°F).

Unit surface area prices $\{c_a\}$ are evaluated with the help of available price information (ASPEN Project, 1988; Hu, 1985; Belding, 1982) and are listed in the second column of Table 2. Unit exergy prices are assumed to be 0.01 and 0.04 $/kWh for fuel and product electricity, respectively. For heat a price per unit heat of 0.015 $/k Wh is assumed (0.067 $/kWh exergy). The capital recovery factor c_z is assumed to be 0.09 $/year. 8000 operating hours per year is assumed.

The system is first treated as a single-feed single-product system, then modified to a cogeneration system of single feed and power and heat products. With the first system, local optimization is performed once with fixed fuel, and once with fixed product. With the second system, local optimization is performed with fixed fuel, with fixed power, and with fixed heat. In both systems three types of dissipation prices are used and compared. These are the marginal and the average prices c_{dd} and c_{da} (Eqs. (19) and (21) of Part 1) and the Lagrangian price $c_{d\lambda}$. The price c_{d0} of the special case $K = 0$ is computed but not used. Table 1 summarizes the cases considered in local optimization.

One set of costing equations (generated in the British system of units, IP) is used in all optimizations. Six design models covering the 10 system components are used to generate the costing equations and to verify them whenever a verification is needed. The system thermodynamic model and the six design models are summarized in the appendix. The derived costing equations and the local objectives as function of local decisions are summarized in Table 2. $\{c_a, k\}$ are given in both SI and IP units. The ranges of applicability of the costing equations in IP units are also given.

For each costing equation, between 10 and 20 design computations are made while only 4 or 5 design computations are needed to derive the constant k and the exponents $\{n_z\}$ of any costing equation. Combinations from the design computations are used to compute k and $\{n_z\}$. The group giving the best fit for all design computation is used. A fit with scatter ±10 percent is considered acceptable.

For each component the correlation of its dissipation as a function of its performance parameters by the form $D_i = k_e {}^*Y_i^{ne}$ is obtained using the component's thermodynamic model (an element in the system thermodynamic model). Fewer thermodynamic computations are needed to obtain n_e.

The number of independent performance parameters of a component depends on the thermodynamic process of the component and the computational algorithm of the system as a whole. Some components may have more than one independent performance parameter and others may have none. For example, a heat exchanger has three parameters describing its performance (two pressure losses and a temperature difference). Decomposing the exergy destruction into thermal and mechanical, three local objectives are generated. However, for the superheater-boiler-economizer combination only two temperature differences (or two effectiveness values) are free to decide but not the third, leaving the combination with five tradeoffs and not six. Also for the combustor, there is nothing to trade the chemical part of the exergy destruction against. Only the me-

Contributed by the Advanced Energy Systems Division for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Advanced Energy Systems Division September 28, 1995; revision received April 19, 1996. Associate Technical Editor: M. J. Moran.

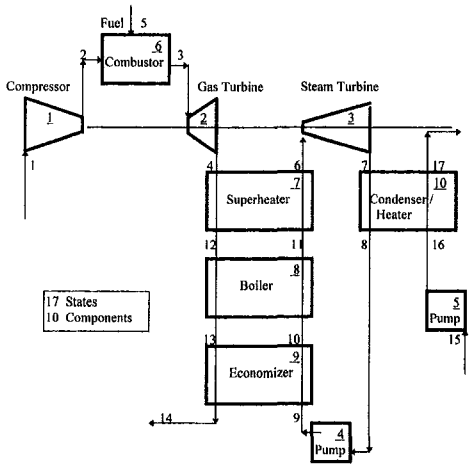


Fig. 1 Energy system of the application examples

Table 1 The cases of local optimization considered

Case and Objective $J = J_2 + J_R$	Dissipation price c_d
Single Product	
$J = c_f*E_f + \sum c_z*Z - c_p*E_p$	$c_{dd} = (c_f*\Delta E_f - c_p*\Delta E_p)/(\Delta E_f - \Delta E_p - \Delta E_j)$
$J_2 = \sum(c_z*Z + c_d*D)$	
$J_R = (c_f - c_d)*E_f - (c_p - c_d)*E_p + c_d*E_j$	
1) fixed feed $J = \sum c_z*Z - c_p*E_p$	$c_{da} = c_p$
c_f*E_f = const.	$c_d = c_{da}$, c_{dd}, $c_{d\lambda f}$
2) fixed Product $J = c_f*E_f + \sum c_z*Z$	$c_{da} = c_f$
c_p*E_p = const.	$c_d = c_{da}$, c_{dd}, $c_{d\lambda p}$
Two Products	
$J = c_f*E_f + \sum c_z*Z - c_p*E_p - c_q*E_q$	$c_{dd} = (c_f*\Delta E_f - c_p*\Delta E_p - c_q*\Delta E_q)/(\Delta E_f - \Delta E_p - \Delta E_q - \Delta E_j)$
$J_2 = \sum(c_z*Z + c_d*D)$	
$J_T = (c_f - c_d)*E_f - (c_p - c_d)*E_p - (c_q - c_d)*E_q + c_d*E_j$	
3) fixed feed $\sum c_z*Z - c_p*E_p - c_q*E_q$	$c_{da} = (c_p*E_p + c_q*E_q)/(E_p + E_q)$
c_f*E_f = const	$c_d = c_{da}$, c_{dd}, $c_{d\lambda f}$
4) fixed power $c_f*E_f + \sum c_z*Z - c_q*E_q$	$c_{da} = (c_f*E_f + c_q*E_q)/(E_f + E_q)$
c_p*E_p=const	$c_d = c_{da}$, c_{dd}, $c_{d\lambda p}$
5) fixed heat $c_f*E_f + \sum c_z*Z - c_p*E_p$	$c_{da} = (c_f*E_f + c_p*E_p)/(E_f + E_p)$
c_q*E_q=const	$c_d = c_{da}$, c_{dd}, $c_{d\lambda q}$

chanical part of the exergy destruction is tradable against the cost of the combustor since the flame temperature is fixed. The case of more than one decision per one local objective was not met in this study.

Table 3 lists the decision variables and their values. Table 3(a) lists the fixed ones and Table 3(b) lists the manipulated

Table 2 Costing equations and the local objective functions

Costing Equation $Z = ca*A$; $A = k*x_1^{n1}*x_2^{n2}*x_3^{n3}*x_4^{n4}$				Local Objectives $J(Y)$; $c_d*D + c_z*Z = k_e*Y^{ne} + k_z*Y^{nz}$			
component	c_a units IP/SI	k	$x_1^{n1}\ x_2^{n2}\ x_3^{n3}\ x_4^{n4}$ ranges of {x}, IP*	Y	ne	nz	D
1 Compressor	25	.15	$M^1\ Pr^{.45}\ e^{.65}$	e	-.95	.65	D_{PT}
	269	.0307	50-1000,5-15, 2.3-11.5				
2 G turbine	25	.32	$M^1\ Pr^{-.5}\ e^{.85}$	e	-.8	.85	D_{PT}
	269	.0654	50-1000 ,5-15, 4-19				
3 St turbine	25	.225	$M^1(T_i/P_i)^{-.05}P_e^{.75}\ e^{.9}$	e	-.8	.90	D_{PT}
	269	2.395	25-100, 1.5-30, 1-150, 4-19				
4 Feed pump	3	.0025	$M^{.55}\ \Delta P^{.55}\ e^{1.05}$	e	-1	1.05	D_{PT}
	32	.0001245	5-70, 14-900, 1.8-9				
5 C.W pump	3	.0063	$M^1\ \Delta P^1\ e^{.7}$	e	-1	.7	D_{PT}
	32	.00106	100-500, 2-25, 4-14				
6 Combustor	.2	5.85	$M^{.5}\ P^{.24}\ \delta p^{-.75}$	δp	1	-.75	D_P
	2.15	1.9352	400-900, 50-200, .01-3				
7 Superhtr	.015	340	$Q^1\ \Delta T_m^{-1}\ \delta P_t^{-.15}\ \delta P_s^{-.14}$	ΔT_m	1	-1	D_T
	.16	30.71	10-15, 100-200, 6-13, .06-.44	δP_t	1	-.15	D_{Pt}
				δP_s	1	-.14	D_{Ps}
8 Boiler	.015	340	$Q^1\ \Delta T_m^{-1}\ \delta P_t^{-.33}\ \delta P_s^{-.26}$	ΔT_m	.45	-1	D_T
	.16	54.80	25-55, 75-200, 6-13, .06-.44	δP_t	1	-.33	D_{Pt}
				δP_s	1	-.26	D_{Ps}
9 Econmizr	.015	310	$Q^1\ \Delta T_m^{-1}\ \delta P_t^{-.16}\ \delta P_s^{-.125}$	ΔT_m	.45	-1	D_T
	.16	27.73	15-40, 70-105, 6-26, .07-.56	δP_t	1	-.16	D_{Pt}
				δP_s	1	-.125	D_{Ps}
10 Condnsr	.02	15	$Q^1\ \Delta T_m^{-1}\ \delta P_t^{-.15}\ \delta P_s^{-.04}$	ΔT_m	.9	-1	D_T
or Heater	.22	1.117	26-70, 15-33, .16-20, .005-.22	δP_t	1	-.15	D_{Pt}
				δP_s	1	-.04	D_{Ps}

* IP Units: c_a k$/ft^2, A ft^2, M lb/s, Q kW (range MW), P_i, P_e psia, T_i R, ΔT F, ΔP, δP psi, D kW.
SI Units: c_a k$/m^2, A m^2, M kg/s, Q kW (range MW), P_i, P_e MPa, T_i K, ΔT C, ΔP, δP kPa, D kW.
Pr = pressure ratio, $e = \eta/(1-\eta)$, $D = D_P + D_T + D_C$, $D_{PT} = D_P + D_T$

ones for the reference systems along with four examples of their optimized values. Note from Table 3(b) that high dissipation prices in local optimization call for more efficient components and conversely. Note also that the hot-side ΔT's of the economizer and the condenser are constrained by their lower bounds and that their cold side pressure loss ratios are sometimes constrained by their upper bounds.

Local Optimization. Table 4 shows the results of the local optimization runs for five cases of the sizing commodity using three types of the dissipation price. Table 4(a) gives the dissipation prices and the overall system parameters. The intention is to reveal the influence of sizing the system by a particular commodity as well as the response of the system to different dissipation prices. All optimization runs were performed at the

Nomenclature

A = surface area
ata = pressure in atmospheres absolute
c = unit price per specified unit; c_F feed, c_P power, c_Q heat; c_f feed, c_p product, c_q exergy-based; c_d dissipation price, $c_{d\lambda}$ Lagrangian price, c_{da} mean average, c_{dd} marginal average, c_{de0} special average; c_z of capital
D = exergy destruction (dissipation); D_P, D_T, D_C mechanical, thermal, and chemical dissipations, D_{Ps}, D_{Pt} mechanical due to shell and tube sides pressure losses
E = exergy rate, E_p product, E_f feed
e = efficiency parameter for adiabatic efficiency $= \eta/(1-\eta)$
eff = efficiency as defined in text; eff_1 first-law efficiency, eff_2 second-law efficiency
F = feed; function
IP = inch-pound (British system of units)
J = objective function; J_2 exergy-based, J_R remainder ($= J - J_2$), J_i local of component i
k = coefficient; k_d, k_e of exergy destruction, k_z of capital
M = a mass rate
MW = megawatt; MWt thermal, MWe electrical or exergical
n = exponent; n_e of exergy destruction, n_z of capital
P = product, power, pressure
Q = rate of heat exchange
SI = international system of units
sh = specific humidity
T = temperature
X = system-dependent variable
Y = system-decision variable; Y_L local, Y_G global
Z = equipment cost
ΔP = pressure rise in a pump
δP = pressure loss by friction; δP_t tube side, δP_s shell side
ΔT_m = logarithmic mean temperature difference
ΔT = temperature difference
η = adiabatic efficiency
λ = Lagrange multiplier
Ψ = loading or head coefficient
Φ = flow coefficient
{ } = set of variables

Table 3 Decision variables

(3a) Boundary and Fixed Decisions

Variable Boundary Parameter	Location	Values*	Variable Fixed Decisions and Conditions	Location	Values*
Compressor inlet P,T,{X}	1	1.013, 27, sh .0026	Top Cycle T (global)	3	871
Fuel inlet T, {X}, M	5	15 natural gas	Teond (global) 1 prod	8	38
Exhaust exit P	14	1.013	2-prod	8	121
Water inlet P, T 1 prod	15	1.03 27	Fuel pressure condition	5	= P2
2 prod	15	3.8 77	Sat. liquid condition	8,10	
Water exit P 1 prod	17	1.03	Sat. vapor condition	11	
2 prod	17	4.5			
Sizing parameter, Fuel	5	4 kg/s			

(3b) Manipulated Decisions

Manipulated Decisions		Reference	Examples of optimized decisions** (1)	(2)	(3)	(4)
Global Decisions	location					
Press. exiting compressor ata	2	9.1	9.1	9.1	9.1	7.7
Steam turbine inlet press ata	6	41.4	41.4	41.4	41.4	12.1
Local Decisions	component					
Compressor efficiency	1	.82	.90	.80	.89	.91
Gas Turbine efficiency	2	.87	.93	.86	.92	.92
Stm turbine efficiency	3	.85	.94	.89	.94	.94
Feed pump efficiency	4	.75	.95	.92	.94	.94
Circulating pump efficiency	5	.75	.71	.61	.67	.75
Combustor loss dP/P in	6	.001	.001	.003	.001	.001
Superhtr loss dP/P in hot side	7	.005	.001	.002	.001	.001
cold side	7	.01	.010	.016	.012	.01
Boiler loss dP/P in hot side	8	.005	.003	.008	.004	.005
cold side	8	.01	.082	.156	.093	.088
Econmzr loss dP/P in hot side	9	.005	.001	.004	.001	.001
cold side	9	.01	.303	.500	.356	.500
Condnsr loss dP/P in hot side	10	.005	.008	.028	.01	.001
cold side	10	.5	.336	.500	.43	.141
Superhtr ΔT C hot side	7	28	2.8	2.8	2.8	2.8
Econmzr ΔT C hot side	9	11	2.8	2.8	2.8	2.8
Condnsr ΔT C hot side	10	5.5	2.8	2.8	2.8	2.8

* P ata, T C

** (1) Local optimization, single product, fixed fuel rate = reference fuel rate, $c_d = c_f$
(2) Local optimization, single product, fixed product rate = reference product rate, $c_d = c_f$
(3) Local optimization, single product, fixed product rate = reference product rate, $c_d = c_{d\lambda}$
(4) Local/Global optimization, two products, fixed fuel rate= reference fuel rate, $c_d = c_{dd}$

reference values of the global decisions. In all the runs the dissipation prices c_{dd}, c_{d0}, and c_{da} (Eqs. (19), (20), (21), Part 1) are computed, but only c_{dd} and c_{da} are used in optimization along with the searched Lagrangian price $c_{d\lambda}$.

The price c_{d0} is negative in two cases. The price c_{dd} is negative in one case and is a very low positive in another. The price $c_{d\lambda}$ is close to c_{dd} for the cases sizing by fuel rates and is close to c_{da} for the cases sizing by power but nowhere near any c_d for the case of sizing by heat.

The search for the Lagrangian price $c_{d\lambda}$ is done systematically by minimizing J_2 using dissipation prices that are increased in small steps. The price $c_{d\lambda}$ corresponds to the most improved value of the objective J.

Table 4(b) shows an example of the convergence to optimum in four system computations. Treating performance decision parameters as local decisions of converging global effects is justified by the fact that convergence occurs in less than four system computations in most cases. Performance parameters of nonconverging global effects are expected to be associated with persisting significant fluctuations in their optimal values and hence in the values of the objective function, which contradicts the observed convergence.

The search for $c_{d\lambda}$ shows that as c_d is increased, the efficiency of the system keeps increasing while the objective function J decreases then increases. The most improved value of J flattens over a range of $c_d \pm 15$ percent. Below this range, the reduced efficiency does not pay, and above it, the increased efficiency does not pay either.

Local-Global Optimization. Table 5 shows the results of local–global optimization runs. The optimization is carried out for the single-product and two-product systems. The fuel rate is assumed fixed and the dissipation price c_{dd} (Eq. (19), Part 1) is used. c_{dd} is very close to $c_{d\lambda f}$. As shown in Table 4, c_{dd} is 0.0355 and $c_{d\lambda f}$ is 0.038.

Table 4 Local optimizations investigated

Case	Prices		c_d $/kWh	Profit $/h	Fuel MWt	Power MWe	Heat MWt	eff1%	eff2%
a) Comparison of Cases									
Single Product System									
Reference	--	--	--	940	212.2	82.97	----	39.10	42.47
1) Fixed Fuel	c_{d0}	.04							
	c_{da}	.04							
	c_{dd}	.04		1284.3	212.2	97.02	---	45.71	49.66
	$c_{d\lambda f}$	.0475		1286.4	212.2	97.66	----	46.02	49.99
2)Fixed Power	c_{d0}	.01							
	c_{da}	.01		933.2	215.2	82.97	---	38.48	41.80
	c_{dd}	.0107		957.5	212.9	82.97	----	38.96	42.33
	$c_{d\lambda p}$	.0317		1106.2	186.2	82.97	----	44.62	48.48
Two-Product System									
Reference	--	--	--	1410.0	212.2	72.25	59.00	61.66	45.08
3) Fixed Fuel	c_{d0}	.0418,0419*							
	c_{da}	.0418		1626.7	212.2	87.75	45.47	62.68	51.21
	c_{dd}	.0355		1629.3	212.2	86.52	46.56	62.61	50.73
	$c_{d\lambda f}$	.0380		1629.8	212.2	86.46	46.58	62.58	50.69
4)Fixed Power	c_{d0}	.0058,0061							
	c_{da}	.0133		1443.5	202.6	72.25	54.00	62.13	46.50
	c_{dd} (use .005)	.0018**		1397.8	281.2	72.25	101.4	61.38	38.33
	$c_{d\lambda p}$	.0108		1455.2	215.4	72.25	62.18	62.2	44.84+
5)Fixed Heat	c_{d0}	-.0014,-.0108							
	c_{da}	.0187		1786.3	239.5	90.58	59.00	62.31	48.28
	c_{dd} (use .005)	-.053		814.1	163.3	42.00	59.00	61.42	38.33
	$c_{d\lambda q}$	.0600		2161	284.7	119.9	59.00	62.76	51.85++

Case	Runs	c_d $/kWh			Profit $/h	Fuel MWt	Power MWe	Heat MWt	eff1%	eff2%
4b : Convergence Example										
Two-Product System										
	run	c_{d0}	c_{da}	c_{dd}						
Fixed Fuel	1	.0419	.0419	.0358	1628.4					
$c_d = c_{dd}$	2	.0419	.0419	.0354	1628.8					
	3	.0419	.0419	.0355	1629.7					
	4	.0419	.0419	.0355	1629.3	212.23	86.52	46.56	62.61	50.73

* c_{d0} when c_{dd} is used and when c_{da} is used (different commodity values in each case).
** a set minimum price below which unexpected violations occur such that compressor work larger than gas turbine work.
\+ second law efficiency lower than reference while objective improves
++ Most bounds are active.

Two starting points were used in the optimization of the two global variables, amounting to 64 and 128 system computations for the single and two-product systems, respectively, while no more than four computations were needed to optimize the 17 local decision variables. The system decision variables are selected such that the computation of a feasible solution is done with a small number of iteration loops. On IBM PC computers of today (e.g., 80,486 CPU chip, 66 MHz), one system computation takes about seven seconds. Thus, while 17 local decisions variables are optimized in less than half a minute using a given dissipation price, the two global decision variables require 15

Table 5 Results of the local/global optimization runs

Case	P2,P6 ata	optzn runs*	Profit $/h	Fuel $/h	Equip $/h	Power $/h	Heat $/h	c_{dd} $/kWh	Power MWe	Heat MWt	eff1 %	eff2# %
Single Product, Given fuel rate=reference												
Ref.	9.1, 41.4	--	940	1953.6	424.7	2378.3	---	---	83.3	---	39.10	42.47
Local	9.1, 41.4	4	1284	1953.6	642.7	2596.4	---	.04	97.0	---	45.71	49.66
Global	9.7, 17.2	16	1324	1953.6	686.6	2640.2	---	.04	99.1	---	46.70	50.73
Local	9.7, 17.2	4	1331	1953.6	677.3	2630.9	---	.04	99.0	---	46.67	50.70
Two Products, Given fuel rate=reference												
Ref.	9.1, 41.4	--	1410	1953.6	395.2	2873.7	885.2	-----	72.25	59.00	61.66	45.08
Local	9.1, 41.4	4	1629	1953.6	449,5	3149.9	1047.2	.0355	86.52	46.56	62.61	50.73
Global	7.7, 12.1	32	1795	1953.6	562.8	3283.3	1028.7	.0418	82.21	68.58	70.99	51.85
Local	7.7, 12.1	4	1796	1953.6	574.2	3297.5	1026.5	.0417	82.57	68.44	70.09	52.01

* Local : a run = 1 system computation
Global: a run = 4 system computations
Eff1: First Law eff = (Power+Heat) / Fuel HHV rate
Eff2: Second law eff = (Power+Heat exergy) / Fuel exergy rate

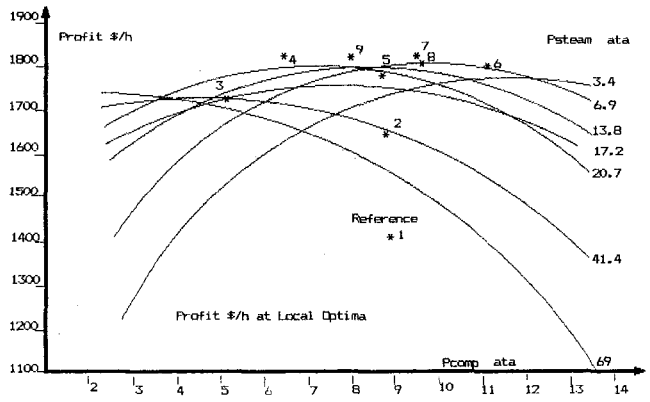


Fig. 2 Profit at local optima as function of global decision variables

minutes. Optimizing local decision variables by searching for the Lagrangian price $c_{d\lambda}$ would take less than 2 minutes assuming eight trial dissipation prices are sufficient to idendify $c_{d\lambda}$ and two computations are allocated for each trial price.

The optimal global decisions for the single-product and the two-product systems are distinctly different but the optimal local decisions are not significantly different.

Figure 2 shows the journey of the profit $(-J)$ during optimization for the two-product system. The optimization journey of the single product system has been given by the author (1995).

A map for profit, power, heat, and first-law efficiency is constructed by computing them at the local optimal decisions for various values of compressor pressure given one value for the steam pressure and the process is repeated for different steam pressures. Only the profit is shown in Fig. 2.

Figure 2 shows a flat zone for maximum profit. The power/heat ratio increases continuously with the increase of compressor and the steam pressures while efficiency shows an opposite trend. Both show a narrow variation in the optimization runs. The power/heat ratio varies from 1.96 to 1.26 and first-law efficiency varies from 0.6 to 0.7. Low-pressure steam seems to be an optimum when heat is cogenerated and sold at a relatively high price.

Point 1 represents the profit of the system before optimization (reference). Point 1 has compressor and steam pressures $P2$ and $P6$ = 9.1, 41.4 ata (132, 600 psia), profit = 1410 $/h, and first-law efficiency = 61.7 percent. The location is correct with respect to profit and compressor pressure only since the local decisions were not yet optimized. Point 2 represents the locally optimized system at the same $P2$ and $P6$. The profit increased from 1410 to 1629 $/h. Points 3 to 8 are all by global optimization. From point 2 to point 3, the profit increased further to 1743 $/h. From point 3 to point 8 the profit fluctuates in a band between 1764 and 1795 $/h.

Proposed Methodology in Steps

In addition to the thermodynamic property equations and conventional mass, energy, and entropy balances, the proposed methodology uses Eqs. (2), (3), (7), (8), (12), (14), (16), (19), and (21) of Part 1. Since the computations require computer use, the following description of the methodology focuses on the creation of the required computer programs:

1 Create a thermodynamic model for the system at hand that describes the system by its connectivity through numbers for the states and numbers for components. Select the inputs (decision variables) such that the performance parameters of the components of the system, whenever possible, are input variables, and such that the computations of the system's performance go with least iteration loops. The sequence of the computations does not need to follow the sequence of the components. A suitable computational hierarchy is to compute pressures that can be determined through allowed pressure losses; then states of known pressure, temperature, and composition; then processes of the essential boundary parameters known; then masses relative to a reference mass (usually assumed 1). As the computational sequence proceeds, pressures give rise to more states to be computed, states give rise to more processes to be computed, and processes give rise to more relative masses to be computed. Finally the sizing parameters determine all mass rates and hence powers, heat loads, exergy, and exergy destruction rates. Although a thermodynamic model that can handle a family of systems is feasible as the one used in this study, it is often more convenient to create an individual thermodynamic model for each system.

2 Prepare two sets of equations for each component: a costing equation for its capital cost as given by Eq. (3); an equation for its dissipation (or for each of the thermal, mechanical and chemical components of its dissipation) as given by Eq. (8) in terms of the selected decision variables describing the performance of the component. The costing equation requires a design model. Dissipation equation(s) are generated by the system's thermodynamic model. The costing equations of Table 2 may offer partial help in filling in this step. The prepared equations, provided that the ranges of their validity are observed, can be used by families of systems and not only by the system at hand. In this case, the equations should be entered in a matrix, with rows representing equipment and columns representing the parameters: unit costs $\{c_a\}$, coefficients $\{k\}$, and exponents $\{n\}$. A system uses any of these parameters by referring to their locations in the matrix. This also permits expanding the matrix in the future, updating its unit costs in time, and updating its coefficients and indices with improved correlations.

3 Implement the thermodynamic model by a costing subroutine that includes the costing equations. This creates the performance-cost model of the system. Allow the model to access the parameters of the costing matrix for modification to permit the investigation of the sensitivity of the objective function to changes in any of the parameters.

4 Create a local optimization routine with Eq. (7) as the local objective function and a global optimization routine with Eq. (16) as the global objective function. The dissipation prices given by Eq. (19) and (21) are used in local optimization as provisional prices. The routines call the performance-cost model whenever a system computation is required and update decisions whenever required. For local decisions, the updating equation is Eq. (12) and the decisions are updated in each system computation. For global decisions, Eq. (14) is the updating equation. Two system computations beside a current one are needed to update each global decision. Optimization routines can be adapted to families of systems.

5 Create a routine to display whatever results needed to be displayed and to print whatever results needed to be printed. This routine can be adapted to almost all systems to be investigated.

Discussion

The proposed methodology is based on the concept of costing equations and the local-influence idealization for certain decision variables in order to manage the large number of decision variables involved in the design of complex energy systems. Expanding the base of costing equations by more types of devices and wider ranges of applicability is beneficial to the methodology as well as to most second-law-related methodologies. A few questions regarding the optimization of local decision variables are left unanswered:

All performance decision variables are treated here as local decisions with no convergence problems. When convergence problems occur, an automated way of isolating the decision variables failing the local-influence idealization is needed.

Equation (19) for the marginal average price c_{dd} does not always predict the Lagrangian dissipation price $c_{d\lambda}$. On the basis that the set $\{Y_i\}$ is local, Eq. (19) ignores the interactions $\Sigma(\partial Z_j/\partial Y_i)$ and $\Sigma(\partial D_j/\partial Y_i)$ where i does not equal j. These interactions occur due to changes in mass rates. For the case of fixed fuel rate, the changes in mass rates from reference to optimum are relatively small. Therefore, Eq. (19) gives a satisfactory prediction of the Lagrangian price $c_{d\lambda}$. In all the other cases, larger changes in mass rates occur and hence the unsatisfactory prediction. A better prediction of the Lagrangian dissipation price is needed.

It is also worth considering whether augmenting the objective function with the exergy balances of the components of a system to obtain different dissipation prices rather than one price would lead to improved minimization. On the other hand, local optimization may lose its speed and improvement may not be significant. Optimum local decision variables do not change drastically with small changes in dissipation prices, particularly when the comparison shows that one dissipation price is in fact a weighed-average of all system dissipations, i.e., $c_d = \Sigma\, c_{di}^* D_i / \Sigma\, D_i$.

The proposed methodology does not prove that the improvement achieved is a global minimum for the objective function. Probably no methodology can guarantee global minimum for multivariable nonlinear problems and the question of possible further improvements of the objective function remains to be answered.

Concluding Remarks

- Considering performance decision parameters as local decisions proves to be an acceptable idealization. The idealization enhances the optimization of performance decisions significantly irrespective of their number. The direction of changing each performance decision to minimize cost is revealed. However, the accuracy of the value of any change depends on the quality of the costing equations and on the quality of the design models used to derive them.
- For local optimizations, there exists an exergy destruction price $c_{d\lambda}$, which minimizes the cost objective function J by a trade-off between the cost of the system's components as dissipators and the cost of their corresponding dissipations. Two average prices c_{dd} and c_{da} sometimes guide to this particular price $c_{d\lambda}$, and if used by themselves, often produce improvement. The marginal average price c_{dd} is a fair prediction of $c_{d\lambda}$ when the system is optimized for fixed fuel rate.
- Global optimization is still slow but the number of decision variables to handle as global is dramatically reduced.

Acknowledgments

The recommendations and the constructive criticism by Professor Richard A. Gaggioli of Marquette University are highly appreciated. Thanks are also due to Professor Michael J. Moran of Ohio State University for his support and helpful remarks and to Dr. Michael R. von Spakovsky of EPFL, Lausanne, Switzerland, for his valuable comments.

References

ASPEN Project, 1988, MIT, Third Annual Report, June 1, 1978–May 31, 1988.

Belding, J., 1982, *Industrial Energy Conservation Manual,* MIT Press, Manual 15.

Cohen, H., Rogers, G., and Saravanamuttoo, H., 1987, *Gas Turbine Theory,* 3rd ed., Wiley, New York.

El-Sayed, Y. M., 1989, "A Decomposition Strategy for Thermoeconomic Optimization," ASME *Journal of Energy Resources Technology,* Vol. 111, pp. 111–120.

El-Sayed, Y. M., 1995, "The Integration of Thermodynamic Analysis and Design Analysis for Optimal Thermal Design," *Proc. International Conference ECOS'95,* Istanbul, Turkey, July 11–14, Vol. 2, pp. 375–386.

Hegetschweiler, H., and Bartlet, R., 1957, "Predicting Performance of Large Steam Turbine-Generator Units," *Transactions ASME,* Vol. 79, July, pp. 1085–1114.

Hu, S., 1985, *Cogeneration,* Prentice Hall, New York.

Rohsenow, W., Hartnett, J., and Ganic, E., 1985, *Handbook of Heat Transfer Applications,* 2nd ed., McGraw-Hill, New York.

Sabersky, R., Acosta, A., and Hauptmann, E., 1989, *Fluid Flow, A First Course in Fluid Mechanics,* 3rd ed,, Macmillan Publishing Co.

Sehra, A., Bettner, J., and Cohen, A., 1992, "Design of a High Performance Axial Compressor for Utility Gas Turbines," ASME *Journal of Turbomachinery,* Vol. 114, pp. 277–286.

APPENDIX

Performance-Cost Modeling

The following is a brief description of the models used in generating the costing equations as well as the system performance-cost model. These models are one thermodynamic model and six design models.

The Thermodynamic Model (El-Sayed, 1989): This is a general program capable of computing the performance of a good number of power generation and cogeneration systems in terms of the overall performance parameters of its components. A system is described by its decision parameters and connectivity with the help of clear instructions. Saved descriptions can be run for analysis any time or modified to derive alternative systems. Consistency of the results is always checked by overall energy and exergy balances. The program can handle up to 70 stations and 40 components. However, to save time and memory, smaller versions of the general program can be derived to handle specific configurations. The general model is used in this study.

The Design Models: *All* design models target the evaluation of a *dominating flow passage surface* for which a unit price gives a fair prediction of the expected component cost in relation to its performance parameters.

The Compressor (Cohen et al., 1987; Sehra et al., 1992): Basic features are axial, two-dimensional analysis at the mean radius, subsonic, 50 percent reaction, diffusion factor < 0.45, and ideal gas properties. Blade geometry is kept constant. All stages have the same temperature rise except the last. Tip blade speed, axial velocity, root/tip radius ratio, and work factors are kept constant at 350 m/s, 150 m/s, 0.5, and 0.98–0.83 (0.83 after the third stage), respectively. A polytropic efficiency is assumed, velocity triangles computed, and the stage efficiency is evaluated from cascade tests corresponding to the blade geometry. Computations are iterated until polytropic and stage efficiencies are matched. Mass rate, pressure ratio, and temperature rise per stage are varied and number of stages, total surface of fixed and moving blades, adiabatic efficiency, speed, and recommended solidity are computed. An arbitrary value of solidity can also be entered as input.

The Gas Turbine (Cohen et al., 1987): Basic features are axial, two-dimensional analysis at mean radius, subsonic, 50 percent reaction, and ideal gas properties. Blade geometry is kept constant. Loading and flow coefficients Ψ, Φ, mean blade speed, and inlet temperature are kept constant at 1.4, 0.8, 340 m/s, and 871°C respectively. The first two values happen to minimize blades. Stage efficiencies of nozzle and rotor blades are assumed, velocity triangles are computed, stage and tip clearance losses are evaluated from cascade tests corresponding to the blade geometry. Computations are iterated until the assumed efficiencies and the losses are matched. Mass rate, pressure ratio, and speed are varied and number of stages, total surface of nozzle and rotor blades, adiabatic efficiency, and recommended solidity are computed. An arbitrary value of solidity can be entered as input. The model does not guarantee that the speed matches that of the compressor.

The Steam Turbine (Cohen et al., 1987; Hegetschweiler and Bartlet, 1957): Same as the gas turbine but with few differences. Actual steam properties were used to compute the specific heat and the isentropic index instead of the constant values assumed in the case of air and combustion gases. Inlet temperature and pressure and exit pressure became inputs instead of the pressure ratio. Exit pressure was changed to cover both condensing and back pressure turbines. In some cases the blade heights were too short and high rotational speeds were entered to reduce mean diameter and increase blade height. The total surface of blades did not change with the change of speed.

Pumps (Sabersky et al., 1989): Basic features are centrifugal, axial flow at inlet and radial at exit with velocity head recovery. Loading (head) and flow coefficients Ψ, Φ, number of impeller blades, root/eye radius ratio, velocity exiting casing, specific volume, and maximum head per stage are kept constant at 1.4, 0.8, 7, 0.4, 1.8 m/s, 0.001 m^3/kg, and 150 m, respectively. Velocity triangles and flow passages are computed given specific speed. Mass rate, head, and specific speed are varied. Speed, surface of impeller, diffuser surface, and efficiency are computed. Specific speed is changed such that surface is minimized. One costing equation did not fit all cases. One equation was used for low flow rates and high-pressure heads (feed pumps) and one for large flow rates and low pressure heads (circulating pumps). Extending flow rates to higher than 225 kg/s need to be implemented.

Combustor (Cohen et al., 1987): Basic features are tube-annular, burning natural gas. Inlet and exit temperatures, air/fuel mass ratio, and number of cans are kept constant at 260°C, 871°C, 75, and 7, respectively. Air mass rate, pressure and pressure loss are varied and the combustor surface computed. 60 m/s was set as a limit on velocity. Combustion intensity varied around 2880 kW/m^3 atm.

Heat Exchangers (El-Sayed, 1989; Rohsenow et al., 1985): Basic features are forced convection heat exchange, single and two-phase fluids, three generic types of exchangers (double tube, fin-plate, and shell-and-tube). For the shell-and-tube type, flow may be counter or cross-counter, tubes may be plain or finned on the outside, and shell may be cylindrical or duct-type. In two-phase cases, more than one equation is used for film coefficients and friction factor multipliers. Pressure losses were based on the worst case multiplier. The plate-fin type is layers of plates with straight parallel fins on each side of each plate. The fins on one side are perpendicular to those of other side. Two sets of layers may be connected in series to allow mixing. A surface geometry is selected. For shell and tube geometry, tube length, diameter, and pitches, and shell diameter or width and depth are entered. For the plate fin geometry, number of plates, fins/inch on either side, their heights and thicknesses are entered. Two groups of boundary parameters can be entered: mass rates and temperature and pressures at all inlets and exits or mass rates and inlet pressures and temperatures and effectiveness. With both entries, film coefficients of heat transfer and pressure drops are computed. The heat exchange surface is computed two ways: surface by geometry and surface = $Q/U\Delta T$. With the first entry, computations are iterated until the two areas are matched and the two pressure drops are met. With the second, only the surface iteration is needed. The pressure drops are output parameters. The iterations are both manual and automated to help minimizing the heat exchange surface. The superheater, the boiler, and the economizer assume duct-type shell and tubes with outside circular fins. The condenser assumes plain tubes and a cylindrical shell. Fin geometry on the outside of the steam generator tubes and fouling factors in heat exchange are kept constant. Inlet parameters are varied. The temperature profile is first computed and checked for crossings and pinch point. Rate of heat exchange, effective temperature difference of heat transfer, and heat exchange surface are computed among other detailed heat transfer outputs. Pressure drops, if outputs, are computed.

Long-Term Testing of Advanced Ceramics: Concerns, Insights, and Recommendations

M. G. Jenkins
Department of Mechanical Engineering,
University of Washington,
Seattle, WA 98195

Advanced ceramics have reached a level of material development to warrant serious consideration for use in advanced heat engine designs. Typically, design requirements based on service conditions may include 25,000 h lifetimes under stresses and temperatures of up to 250 MPa and 1370°C, respectively, with probabilities of failure of <0.1 percent. To assure that materials meet these stringent requirements requires long-term testing under the service conditions. Tensile tests at 1370°C in ambient air have been conducted on silicon nitride alloys to 5000 h with reports of 10,000 h tests for silicon carbide. To provide useful data, such long-term tests must incorporate such meticulous attention to detail as: strict temperature control (±5°C); accurate temperature measurement (1 percent of the nominal temperature); close control of grip cooling (±0.1°C) and ambient environment (±0.25°C); stable, high-resolution extensometry (±0.5 μm); reliable heating (MTBF > 10,000 h) and load control (gravity-controlled, dead load), and responsive data acquisition systems (12-bit, digital collection). Data thus obtained can be used as input into design codes such as NASA CARES/LIFE to predict and confirm reliability/durability.

Introduction

Energy and environmental concerns have placed increasingly greater demands on materials used in advanced engineering designs. For example, as greater efficiencies are sought and achieved in the design of gas turbine engines, so too have the temperature, strength, and weight requirements of their components changed to "push the limit" of the mechanical properties of the various materials (generally metallic alloys) now used in those applications. Much time and effort have been devoted to the search for structural materials for a variety of high-temperature applications that will keep pace with these engineering demands; only recently have ceramic materials such as silicon nitride and silicon carbide alloys reached the developmental stage to receive focused attention as plausible successors to the more traditional metallic alloys. A number of high melting point ceramics exhibit, at elevated temperatures, greater resistance to oxidation and chemical change, lower density, higher strength and elastic modulus than even the best nickel-based super alloys.

However, although ceramics as a class of materials have found their way into a wide variety of applications, including microelectronics, sensors and transducers, superconductors, bioceramics, and cutting tools, their implementation in elevated-temperature applications is still evolving. The reasons for this may be the engineering community's perception of the lack of a credible data base, reluctance to change, and other factors, such as cost. Nonetheless, it is in the high-temperature world of turbine engines (where turbine inlet temperatures are projected at >1370°C) and other related aerospace applications that advanced ceramics have their greatest potential, and simultaneously experience their greatest skepticism. Therefore, perhaps the greatest challenge to the ceramic engineering community is in turbojet engine technology and beyond (i.e., ramjet, scramjet, etc.).

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 27, 1995. Paper No. 95-GT-237. Associate Technical Editor: C. J. Russo.

Although the perception of ceramics as brittle and susceptible to thermal shock might indicate their limited, if not undesirable use in high-stress, high-temperature environments, their high-temperature strengths and oxidation resistances make ceramics ideal materials for sustained use at high temperatures. Even so, there are a number of factors contributing to the slow rise of detailed investigation into the creep of ceramic materials. A primary factor is that long-term creep testing and characterization of creep fracture properties of materials is an expensive and time-consuming endeavor, although recent studies have reported successful long-term tests of ceramics to 5000–10,000 h [1, 2]. An additional factor is the limited availability of these newly developed and still-emerging materials, which restrict mechanical characterization to those ceramics that produce the more immediate results.

In design, a suggested general rule is that the test data for extrapolating long-term creep data to "real-world" applications should extend to ~10–20 percent of the desired service life [3]. Thus, if an engine component design or service life is 25,000–50,000 h, over 2500–10,000 h of testing would be required to estimate service performance. However, even this general rule is suspect when one considers the probabilistic nature of the failure of ceramics. Currently available probabilistically based reliability design codes for ceramics based on weakest link theory [4] assume linear-elastic stress conditions for fast fracture. At present it is not certain what type of design methodology should be implemented to account for such cumulative damage mechanisms as creep [5–7]. Additionally, turbine engine components may have complex three-dimensional designs, experience severe stress and temperature gradients, and must operate under time-varying stresses and temperature. All these add to the reluctance of engineers to replace their time-proven metallic alloys with brittle, somewhat unproven materials. Nonetheless, long-term testing must be implemented to allay these fears and establish the database necessary for designers to incorporate advanced ceramics into new turbine engines.

In this paper, aspects of long-term testing of ceramic materials are discussed in the following order. First, temperature control is addressed, followed by discussion of temperature measurement. Next, grip cooling and ambient environment are discussed. Stable, high-resolution extensometry is reviewed, after

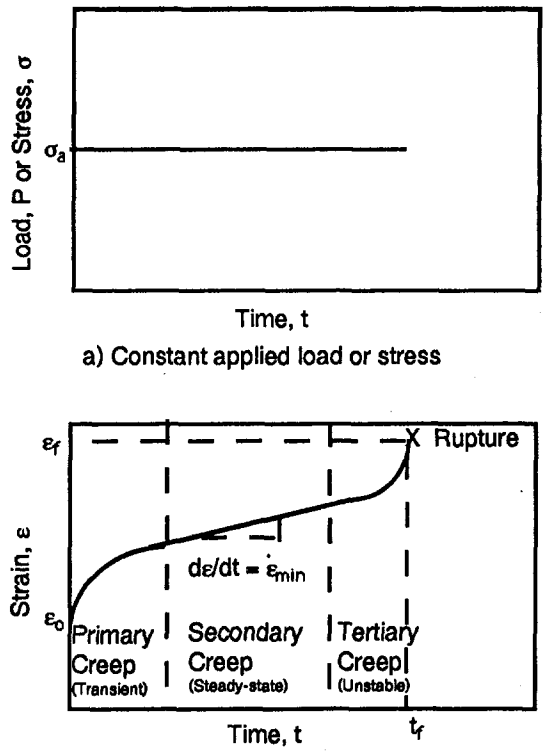


Fig. 1 Creep deformation results

which reliable heating and load control are presented. Finally, responsive data acquisition systems are detailed before summary and conclusions are provided that give specific recommendations and guidelines for long-term testing of ceramics.

Review: Creep and Creep Rupture

Results of long-term testing of many materials follow the schematic illustrations of Figs. 1(a–c). Under a constant load or stress (Fig. 1(a)), the material response is the time-dependent deformation (creep strain) shown in Fig. 1(b). Some advanced ceramics, such as a silicon nitride with 6 wt% yttria [8] (as well as aluminas and other ceramics with predominantly vitreous secondary phase) under varying stress and temperature conditions, have shown all three regions of creep behavior as illustrated in Fig. 1(b). However, other types of silicon nitride with 4 wt% yttria [2, 9] show a dominant steady-state regime with no tertiary regime.

Analysis of these creep data is typically based upon the assumption that the minimum (or steady-state) creep strain rate ($d\epsilon_{min}/dt = \dot{\epsilon}_{min}$) (Fig. 1($c$)) dominates the overall behavior. For this case, the applied tensile stress, σ_a, and temperature dependencies of $\dot{\epsilon}_{min}$ are given by the expression [8, 9]:

$$\dot{\epsilon}_{min} = A\sigma_a^n \exp(-Q_c/RT) \quad (1)$$

where A is a pre-exponential factor, n is the creep (or stress) exponent, Q_c is the activation energy for creep, R is the universal gas constant, and T is the absolute temperature.

The creep rupture life or time to failure, t_f, can be described by a similar equation [8, 9]:

$$t_f = B(\sigma_o)^{-N} \exp(Q_f/RT) \quad (2)$$

where B is a pre-exponential factor, Q_f is the activation energy for creep rupture (static fatigue), and N is the stress (or creep) rupture exponent. As shown in Fig. 2, this creep rupture relation, while also a function of temperature, may not always be a simple relation. Indeed, so-called knees in the curves may indicate changes in failure mechanisms (e.g., from dominant slow crack growth of intrinsic flaws at high values of stress to cumulative creep damage at low values of stress) [8, 9].

When failure is controlled by the accumulation of creep damage, t_f is often found to be a unique function of $\dot{\epsilon}_{min}$ independent of applied stress and temperature. This function is adequately described by the Monkman–Grant relation [8, 9]:

$$t_f = C(\dot{\epsilon}_{min})^m \quad (3)$$

where C and m are constants. A comparison of Eq. (3) (which is expected to be valid when failure is controlled by the accumulation of creep damage, i.e., activation energies for creep and creep rupture are approximately equal) to Eqs. (1) and (2) indicates that $-N = n \cdot m$.

As Eqs. (1) and (2) and Figs. 1 and 2 indicate, the creep deformation and creep rupture behavior are strongly dependent on both applied stress (and hence load) and test temperature. Stress is raised to a power (typically n = 2–6 in tension) and temperature is the independent variable of the exponential function.

Tensile Testing Setup

In the discussion and results that follow, references to the testing experience of the author will deal primarily with uniaxial tensile tests [2, 8, 9, 11]. The primary motivations for long-term testing in tension are (1) uniform, uniaxial stress state and (2) unambiguous interpretation of the applied stress and measured strain results. It should be noted that ceramics, unlike most metals, show directional anisotropy during high-temperature deformation. Therefore, creep test results for tension will differ from creep test results in compression. Creep tests in flexure will be complicated by this difference leading to involved data reduction, which may lead to ambiguity in interpreting the test results [10, 11].

Although details of the elevated-temperature, tensile testing setup are given elsewhere [8, 9]; a brief description is provided

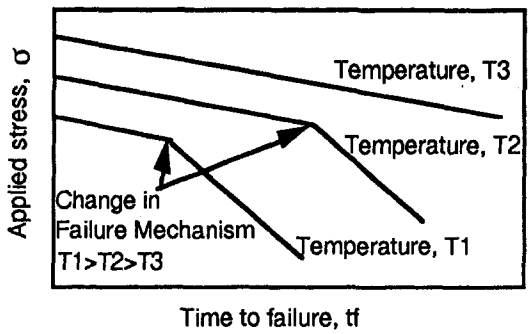


Fig. 2 Creep rupture results

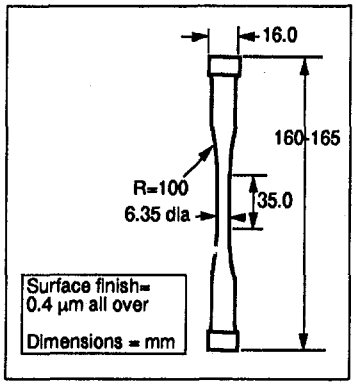


Fig. 3 Typical button-head tensile specimen

here for clarity. Button-head tensile specimens [8, 9] (see Fig. 3) (6.35 mm gage section diameter and 35 mm gage section length) were fabricated using diamond-grit grinding wheels from as-processed rods (typically 160–175 mm long and 20–25 mm in diameter).

Tensile creep/creep rupture tests [8, 9] were conducted in ambient air under electronic load control on commercial electromechanical load machines having both load- and strain-control capabilities. The specimen grips were located outside the compact two-zone resistance-heated furnaces capable of generating maximum temperatures of ~1600°C. The grips, which utilized copper collets as interfaces with the button heads of the specimens, were attached to the load frame with self-aligning, self-contained, hydraulic couplers to minimize bending moments. In some cases specimens were strain gaged and the percent bending measured at room temperature to ensure that minimum bending was imposed at the lower loads and stresses applied during the creep tests. In all these cases bending at the anticipated creep stresses was on the order of 3 percent of the uniaxial tensile stress [8].

Specimen displacements were measured over 25-mm gage lengths using direct-contact extensometers employing remote capacitance sensors (see Fig. 4). By carefully controlling the temperatures of both the measurement hardware and grip, cooling water [8] resolutions of ~0.5 μm could be achieved with these arrangements. Personal computers and digital data acquisition systems were used to monitor and store various output signals including time, displacement, load, load error, and test temperature.

Temperature Control

Temperature, as any other engineering independent variable, must be controlled during testing so as to provide valid information about the material behavior. Generally a feedback control system such as a PID controller is used (regardless of the temperature transducer) to provide self-correcting temperature about a desired setpoint.

Based on current ASTM recommendations [12], for the duration of the test the difference between the indicated temperature and the nominal test temperature should not exceed the following limits:

$\leq$1000°C ±3°C

>1000°C ±6°C

In addition, temperature within the uniformly heated gage section should not exceed the following per 25 mm of gage section length:

$\leq$500°C ±5°C

>500°C ±1 percent of the test temperature

The term "indicated temperature" means temperature that is indicated by the temperature measuring device using good quality pyrometric practice. It is recognized that true temperature may vary more than the indicated temperature. The permissible indicated temperature variations should not be construed to minimize the importance of good pyrometric practice and precise temperature control. Both indicated and true temperature variations should be kept as small as practicable. It is well recognized, in view of the extreme dependency of long-term behavior of materials on temperature, that close temperature measurement is necessary. The limits prescribed here represent ranges that are common practice, based on experience and practicality.

Temperature overshoots during heating should not exceed the limits stated above. The heating characteristics of the furnace and the temperature control system should be studied to determine the power input, temperature set point, proportioning control adjustment, and control-thermocouple placement to limit transient temperature overshoots. It is sometimes desirable to stabilize the furnace at a temperature 10 to 25°C less than the nominal test temperature before making the final adjustments.

The rate at which temperature can be increased from the ambient to final test temperature depends on many factors, such as: heating system, temperature controller, test material, and test environment. In some materials, so-called oxidation due to low-temperature chemical instabilities occurs at intermediate temperatures (800–1000°C) much less than the upper limit elevated temperatures (1370–1400°C) of long-term tests. In these materials, the temperature ramp should be as rapid as possible to minimize the exposure time to these intermediate temperatures. Generally, good results have been obtained for heating rates in which the specimen temperature is ramped from ambient to the test temperature in ~30 minutes. The hold time at temperature prior to the start of the test should be governed by the time necessary to ensure that the specimen has reached a steady-state temperature and that the temperature can be maintained within the limits specified above. In practice this may require up to 2 h to ensure that the grips, the extensometer, and other heated parts of the system have also reached steady-state temperatures.

It has also been found, that if possible, the empty furnace should be maintained at the test temperature for 12–24 h prior to testing to create a stable-temperature "thermal mass" before the initiation of the long term. This stable thermal mass provides more consistent data during the initial hours of the long-term

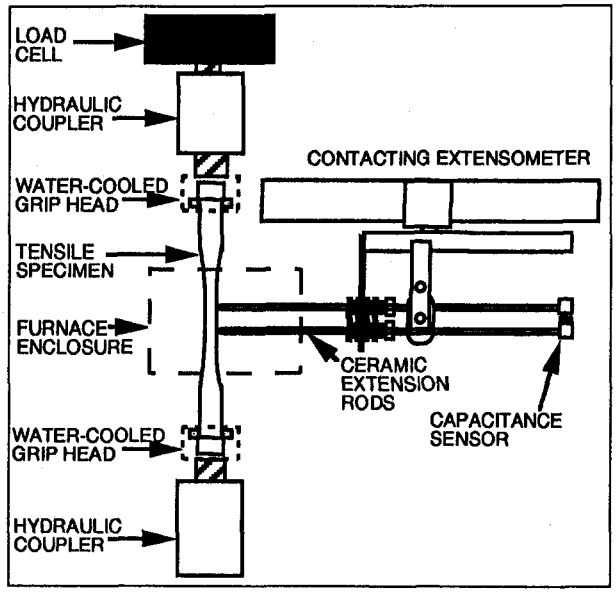


Fig. 4 Schematic of tensile test setup

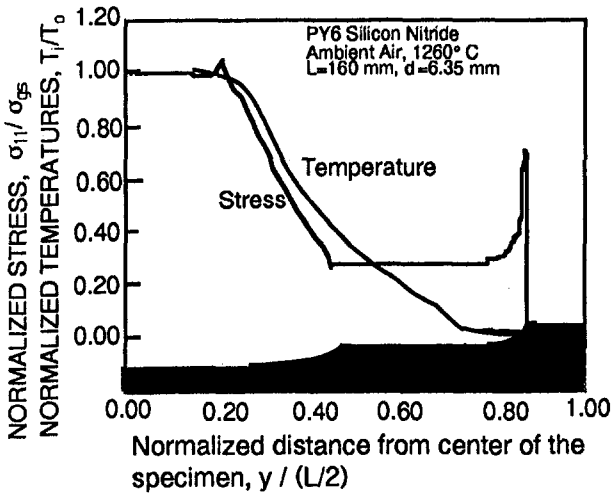


Fig. 5 Stress and temperature distributions for a cold grip

test, although consequences of this procedure at the several thousand hour level are insignificant.

Temperature Measurement

The method of temperature measurement must be sufficiently sensitive and reliable to ensure that the temperature of the specimen is within the limits specified in the previous section [13]. Although optical methods (i.e., infrared or emissivity) or pyrometric substances are viable methods of measuring gross temperatures, for the temperature resolution required in long-term materials testing the primary temperature measurement should be with thermocouples in conjunction with electronic temperature controllers and/or readout units. Although potentiometers and millivoltmeters have been recommended as suitable temperature measurement devices in the past [13], more modern instruments eliminate unnecessary complication and ambiguity.

Such measurements are subject to two types of error. Thermocouple calibration and instrument measuring errors initially produce uncertainty as to the exact temperature. Secondly both thermocouples and measuring instruments may be subject to variations over time. Common errors encountered in the use of thermocouples to measure temperatures include: calibration error, drift in calibration due to contamination or deterioration with use, lead-wire error, error arising from method of attachment to the specimen, direct radiation of heat to the bead, heat conduction along thermocouple wires, etc.

Temperature measurements should be made with thermocouples of known calibration. Representative thermocouples should be calibrated from each lot of wires used for making noble (e.g., Pt or Rh) metal thermocouples. Except for relatively low temperatures of exposure, noble metal thermocouples are eventually subject to error upon reuse, unless the depth of immersion and temperature gradients of the initial exposure are reproduced. Consequently noble-metal thermocouples should be calibrated using representative thermocouples. Oxidized noble-metal thermocouples also should not be reused without clipping back to remove wire exposed to the hot zone, rewelding, and annealing. Any reuse of noble metal thermocouples after relatively low-temperature use without this precaution should be accompanied by recalibration data demonstrating that calibration was not unduly affected by the conditions of exposure [12, 13].

Measurement of the drift in calibration of thermocouples during use is difficult. When drift is a problem during tests, a method should be devised to check the readings of the thermocouples on the specimen during the test. For reliable calibration of thermocouples after use, the temperature gradient of the test furnace must be reproduced during the recalibration. Note that Eq. (1) indicates that minimum creep strain rate is an exponential function of temperature and that creep rupture time as shown in Eq. (3) is a function of the minimum creep strain rate raised to a power. Thus, errors in temperature measurement can have significant consequences on not only the outcome of long-term tests but also subsequent interpretation and use of the results in long-term predictions.

Generally, noble-metal thermocouples should not be attached directly to ceramic materials due to chemical incompatibility. The thermocouple junction may be brought close to the specimen (3–6 mm) and shielded from thermal radiation. Shielding may be omitted if, for a particular furnace, the difference in indicated temperature from an unshielded bead and a bead inserted in a hole in the specimen has been shown to be less than one half the variation listed above. The bead should be as small as possible and there should be no shorting of the circuit (such as could occur from twisted wired behind the bead). Ceramic insulators should usually be used on the thermocouples in the hot zone. If some other electrical insulation material is used in the hot zone, it should be carefully checked to determine whether the electrical insulating properties are maintained at higher temperatures.

When the length of the specimen gage section is 25–50 mm, at least two thermocouples are employed, one near each end of the gage section. For lengths of >50 mm, a third thermocouple is added near the center of the gage section length [12].

Grip Cooling and Ambient Temperatures

A primary requirement for long-term tests at elevated temperatures is that the gage section must be exposed uniformly to the test temperature. This requirement poses an interesting problem. If a conventional hot wall furnace is used in conjunction with cold grips to test a straight-sided test specimen, the specimen will be exposed to a nonuniform temperature profile but uniform stress profile. If the material displays little dependence of strength on temperature, a failure outside the heated gage section would be considered unacceptable. Therefore, it is often necessary to define a gage section where the material is not only exposed uniformly to the test temperature but where it is also subjected to the maximum stress (see Fig. 5). This is accomplished by reducing the cross section of the specimen, thereby "forcing" the failures to occur in the uniformly heated and stressed region.

For tests at elevated temperatures the grips can be located outside the furnace and either liquid-cooled (cold grips), or noncooled (or heated) (warm grips). Alternatively, the grips can be exposed to the test temperature (hot grips). Thermal analyses such as those shown in Fig. 6 have indicated substantial

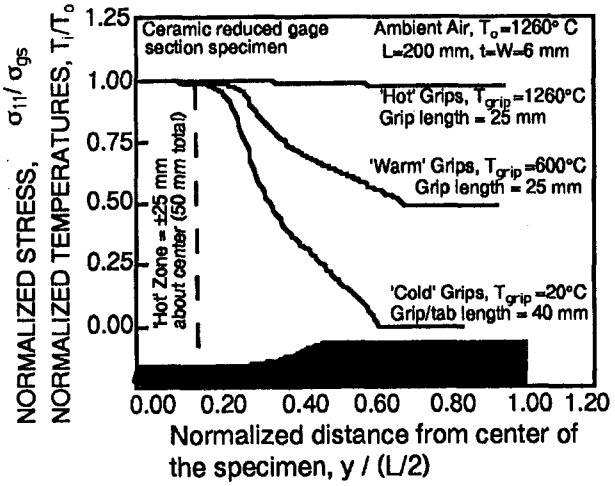


Fig. 6 Temperature distributions for cold, warm, and hot grips

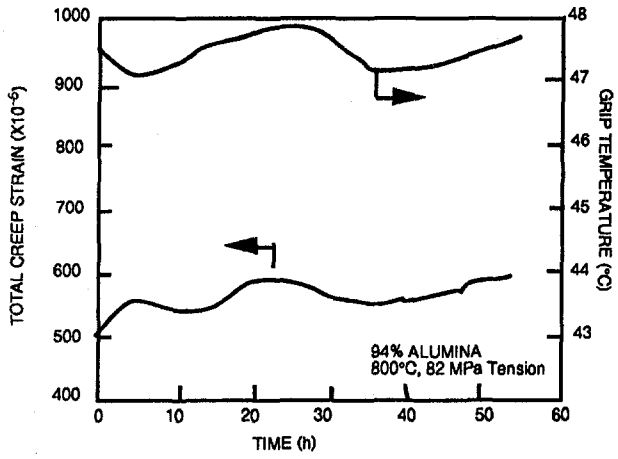


Fig. 7 Strain fluctuations with grip temperature fluctuations

temperature gradients when using cold grips ($\Delta T \approx 1200$°C in $\Delta L \approx 50$ mm). Such steep temperature gradients may introduce thermal stresses and thereby promote non-gage section failures. Progressively less steep temperature gradients exist for warm and hot grips.

However, in practice cold grips have been found to produce the most economical testing setups since they can be fabricated conventionally from easily obtained engineering materials (e.g., stainless steels). In addition, because the grips are not exposed directly to the high-temperature oxidizing environment of the test chamber, the grips can be used repeatedly with no change in performance.

A major limitation of cold grips is that fluctuations in the grip temperature can produce large fluctuations in the temperature gradient in the specimen, thus introducing fluctuations in the strain readings (see Fig. 7). Similarly, changes in the temperature of the ambient environment will affect the temperature at the grips and measured strain fluctuations will result. Additionally, temperature changes of the ambient environment can affect the strain readouts from the extensometer sensor and conditioner as well as the load readout from the load cell and conditioner.

To minimize these strain fluctuations, a series of corrective actions were taken for the test setup of this study. These included installation of a closed-loop laboratory chiller with water-cooled condenser and a hot gas bypass circuit, which produced grip cooling water temperatures within ±0.1°C of a setpoint (versus cooling tower water temperature fluctuations of up to ±2°C). A well-controlled HVAC system (closed room with no windows and ~10 percent make up air) was placed into operation, which reduced ambient temperature fluctuations from ±1°C to ±0.25°C.

These modifications reduced strain fluctuations due to grip and ambient environment from about 75×10^{-6} m/m to $<10 \times 10^{-6}$ m/m over a 24 h period, thus greatly reducing the variability of the data and increasing the reliability of the creep strain rate analyses.

Stable High-Resolution Extensometry

Typical strains at creep rupture of advanced ceramics (~1.5–3.5 percent) are a fraction of those of metals or polymers. Thus, the creep strain evaluation of ceramics requires the use of extensometers with fine resolutions. In addition, ceramics are nonductile and very sensitive to stress concentrations, which may complicate the use of contact-type strain measurement devices.

Strain measuring devices for structural ceramics may be classified as optical and mechanical. Optical extensometers (for example, a constant mercury light source, which creates a shadow of the flags attached to the specimen [1] or a scanning laser beam interrupted by the flags [1, 14]) have been used successfully to determine strains during the tensile creep evaluation of monolithic ceramics in tests lasting from several thousand hours up to 10,000 h [1, 14].

However, these devices may require ceramic flags attached to the specimen to monitor displacements. These flags may be difficult to retain on the specimen and may, over time, creep and deform thus altering the strain measurement. Finally, contaminants or oxidation products on the surface of the flags may disrupt the strain measurements [14], necessitating shifting of the light beam to intersect a "clean" portion of the flags.

An additional complication is the need for a clear optical path into the hot zone of the furnace. Silica-glass windows are typically used to contain the heat of the furnace. However, a major problem with this type of system is the air density fluctuations that disturb the index of refraction of the beam of light. The source of density fluctuations can be divided into air seepage around the pull rods and the thermal gradient in the air adjacent to the silica glass window.

The former source can be suppressed by insulating the pull rods and averaging the displacement measurements. For the latter problem, the window is mounted about 100 mm outside the inner refractory wall of the furnace, thus avoiding the most severe thermal gradients. However, surface degradation and contaminants reduce the clarity of the silica glass over time, forcing replacement every 200–300 h [14]. Nonetheless, optical extensometers have been shown to provide resolutions of ±2 μm over 25 mm gage lengths [14].

The most successful mechanical extensometers involve long, ceramic, contact rods and remote sensing elements (typically capacitance sensors or strain-gaged elements). The introduction of bending moments by mechanical extensometers may be a major concern and unless strains are measured using similar extensometers on opposite sides of the specimen, the use of a light force extensometer is favored (on the order of 50–100 g contact force per rod). The use of knife edges on the ends of the contact rods is favored over the V-notch because it minimizes the chance of slipping on the surface of the specimen. It has been found that initially, light contact with the specimen will minimize slippage until oxidation products fuse the extensometer tips to the specimen, thus eliminating any chance of slippage.

Two major concerns are (1) chemical and thermal compatibility of the extensometer contact rods and the specimen and (2) thermal effects due to temperature gradients and ambient environments. The first concern has been addressed by using α-SiC extension rods. This material is chemically compatible with most advanced ceramics and displays almost no creep susceptibility and little degradation due to oxidation in the temperature ranges of interest (800–1400°C).

The second concern has been addressed by first blocking the entrance ports in the furnace for the extension rods with ceramic felt (zirconia) disks to minimize temperature fluctuations due to convective air currents. In addition, the thermal effect due to conduction from the hot zone to the sensing elements has been minimized by either air or water cooling the contact rods at a point near the transition from the rods to the pivot points of the extensometer. Finally, containment of the extensometer assembly in a sealed enclosure has helped create a stable, thermal mass of air around the extensometer while isolating from air currents due to the HVAC system in the ambient environment.

Having addressed the two major concerns of contacting extensometry, the extensometer shown schematically in Fig. 4 has been shown to give resolutions of ±0.5 μm over 25 mm gage lengths [2, 8, 9], a major advantage over optical systems. An additional advantage is that since no performance-inhibiting degradation occurs over the course of a test, no components need be replaced until test completion.

Reliable Heating and Load Control

Heating can be by indirect electrical resistance (heating elements), direct induction, indirect induction through a susceptor, or radiant lamp with the specimen in ambient air at atmospheric pressure unless other environments are specifically applied and reported. Note that direct resistance heating is not recommended due to possible differences of the electrical resistances of the constituent materials, which may produce nonuniform heating of the specimen.

Two-zone compact furnaces employed in these studies [2, 8, 9, 11] utilize six resistance heating elements equally spaced around the specimen for each zone (total of twelve elements). The molybdenum disilicide (Super Kanthal™) elements are U-shaped and introduce a reduction from 12 mm diameter at the point of the electrical connections to 6 mm at the resistively heated U. These elements, while electrically connected in series, have proven quite robust and have provided a minimum of 10,000 h of life under normal use. Generally, elements fail due to physical damage or arcing at the electrical connection and not because of degradation at temperature. Elements that do fail, fail either upon furnace restart after a cooldown or due to excessive physical handling during replacement of other failed elements. During normal service, elements have not been known to fail, and thus continuous tests of over 10,000 h should be possible without element replacement. In case of element failure during the course of the test, elements can be electrically isolated such that failure of one element will not result in the loss of the entire test. Such systems have been employed for tests of up to 10,000 h in duration in which elements have been changed while the test was ongoing [1].

As discussed in previous sections, recommended temperature control and measurement involve thermocouples. For tests at temperatures <1500°C, Type S (Pt–Pt10Rh) unshielded, thermocouples are preferred for their responsiveness and accuracy. However, degradation of the unshielded thermocouples in the temperature range of 1300–1400°C limits their useful (reliable accuracy) life to ~1000 h. Thus, replacement of thermocouples can be anticipated and replacement performed before failure of a thermocouple produces large temperature fluctuations. As an alternative, type B (Pt6Rh–Pt30Rh) thermocouples have been shown [13] to provide superior long-term stability (compared to Type S) with some sacrifice in responsiveness.

The PID temperature controllers used in these studies have proven reliable and trouble free in the automatic control mode. However, during thermocouple replacement, the manual mode can be used to hold the output of the silicon-controlled rectifier (SCR) at the level called for in the automatic control mode. Thus, even though the temperature controller has no input (with the thermocouple removed), the elements will still receive a level of power required to maintain the furnace near (within ±10°C) the required test temperature. Once the thermocouple has been replaced (after ~5 min) the automatic control mode can again be selected to return the temperature to that which is required.

Load must be controlled within a reasonable level since creep strain rate is a function of stress raised to a power (2–6). In previous studies [2, 8, 9] commercial electromechanical load machines having both load- and strain-control capabilities were used in electronic load control for creep testing. Typically, desired load (strain)-time profiles were controlled by analog, trapezoidal function generators. While electronic load cells and associated conditioners provided load resolutions on the order of 1 N, power glitches or stray electrical signals over the course of long-term tests caused the test machine to "lose control." During isothermal creep tests, any change in the applied stress (load) can have significant effect on the creep damage evolution and subsequent creep rupture.

Therefore for maximum reliability, dead weight loading (preferably through the mechanical advantage of a lever arm as shown in Fig. 8) with an electronic load cell for digital recording of the applied load is recommended. Since dead weight loading relies upon gravity to maintain the load, loss of load control (short of a change in the gravitational constant of the universe) is nearly impossible. While use of an electronic load cell provides high resolution of the applied load, the finite masses used to produce the dead weight, limit the applied loads to those produced by combination of various discrete weights. However, the increased reliability of the dead-weight, gravity-controlled system more than compensates for the loss of the infinite range of loads.

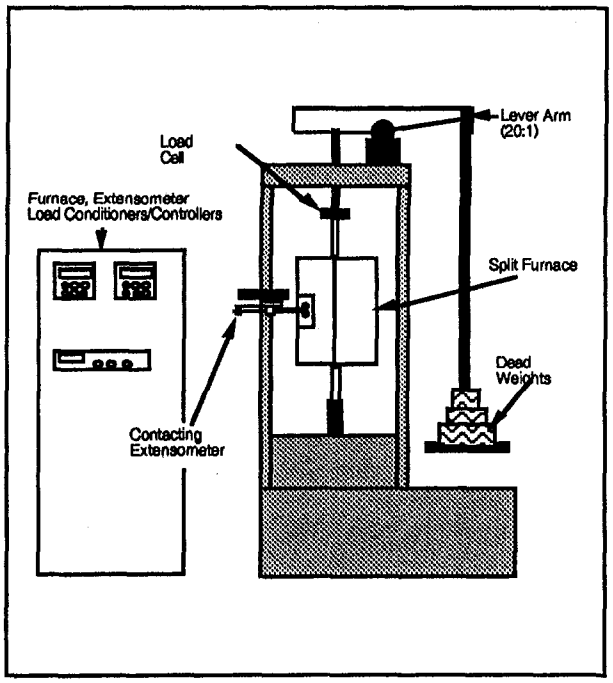


Fig. 8 Dead weight load frame

It should be noted that although the implication of a dead-weight creep test is a constant load, not constant stress test. Ordinarily (for metals) this would be of concern since large longitudinal tensile strains (and subsequent necking) can produce significant reductions in cross-sectional areas of tensile specimens with subsequent and significant changes in the applied stress. However, given the relatively small creep strains to failure for advanced ceramics (1–4 percent) and the somewhat smaller Poisson's ratios, so-called stress errors would be on the order of 1–3 percent, nearly on the order of the variability of dimensional measurements of the cross section.

Data Acquisition System

A computer-based data acquisition system is required for a number of reasons, not the least of which is the ease of data analysis. In addition, a 12-bit word provides a resolution of ~2 mV for a 10 V signal typical of most transducer circuits. Finally, infinitely adjustable data acquisition rates provide rapid data collection rates necessary at initial loading and just prior to fracture and lower data collection rates during steady-state creep.

At a minimum, strain, time, load, and temperature should be monitored. In addition, a signal such as load error or a break detector should be monitored to provide a positive indication of specimen fracture.

To minimize the effects of noise, it has been found useful to sample signals every 50 ms but to record the signals as a running average at discrete intervals somewhat greater than the sampling rate. To this end, two simultaneous recording modes have been

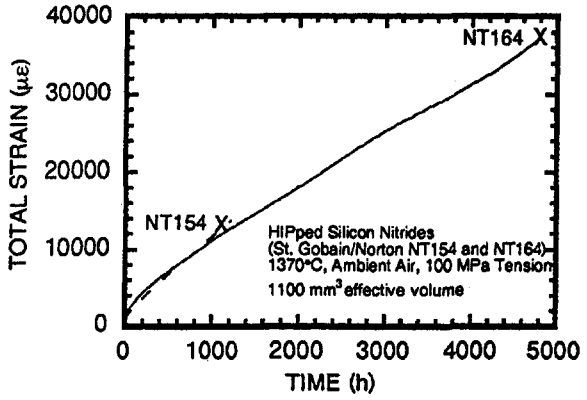


Fig. 9 Comparisons of creep tests for two silicon nitrides

used. The first mode is continuous and records a datum every 30–60 s. The second mode is a loop, which records 100–1000 data points every 60 s. In this way, the basic creep data are recorded continuously at a rate sufficient to give a reasonable number of data points (1000–2000). Meanwhile, events such as the final fracture are recorded only around the instant in which they occur.

Storage of the data as either binary or text files simplifies data reduction by allowing direct importation of file contents into data analysis programs.

Summary and Conclusions

To provide useful engineering data, long-term tests of ceramics must incorporate such meticulous attention to detail as: strict temperature control (±5°C); accurate temperature measurement (1 percent of the nominal temperature); close control of grip cooling (±0.1°C) and ambient environment (±0.25°C); stable, high-resolution extensometry (±0.5 μm); reliable heating (MTBF > 10,000 h) and load control (gravity control, dead load), and responsive data acquisition systems (12-bit, digital collection).

When these aspects are closely controlled and monitored, creep tests can be reliably conducted to nearly 5000 h with measured minimum creep strain rates as low as 10^{-10}/s at 1370°C as shown in Fig. 9. Such low strain rates are in the range of those that are physically measurable.

It is worth noting that replicate tensile creep tests of a silicon nitride at 1370°C for two stresses (145 and 180 MPa) were conducted in a recent study [9] using the methods, techniques, and equipment discussed here. In these tests, despite meticulous attention to experimental detail, the materials showed enough variability that the shapes of the resulting strain–time curves varied markedly over the 16 tests conducted at 145 MPa. While the steady-state strain rates varied only by at most a factor of 3, the times to failure varied by an order of magnitude (45–450 h). Thus, the tester must be aware that strict attention to detail and controlled testing will eliminate the test methods, techniques, or equipment as sources of this variability, thus allowing him to focus on the inherent material variability.

Acknowledgments

Research was sponsored by the U.S. Department of Energy, Assistant Secretary for Conservation and Renewable Energy, Office of Transportation Technologies, as part of the Ceramic Technology Project of the Materials Development Program, under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

In addition, empirical work was conducted at the High Temperature Materials Laboratory sponsored by the U.S. Department of Energy, Assistant Secretary for Conservation and Renewable Energy, Office of Transportation Technologies, High Temperature Materials Laboratory User Program, under contract DE-AC05-84OR21400 with Martin Marietta Energy Systems, Inc.

References

1 Ohji, T., and Yamauchi, Y., "Long-Term Tensile Creep Testing for Advanced Ceramics," *J. Am. Ceram. Soc.*, Vol. 75, 1992, pp. 2304–2307.

2 Ferber, M. K., Jenkins, M. G., Nolan, T. A., and Yeckley, R., "Comparison of the Creep and Creep Rupture Performance of Two HIPed Silicon Nitride Ceramics," *J. Am. Ceram. Soc.*, Vol. 77, 1994, pp. 657–665.

3 Dowling, N. E., *Mechanical Behaviour of Materials: Engineering Methods for Deformation, Fracture and Fatigue*, Prentice Hall, Englewood, NJ, 1993.

4 Nemeth, N. N., Manderscheid, J. M., and Gyekenyesi, J. P., "Designing Ceramic Components for Reliability," *Am. Ceram. Bull.*, Vol. 68, 1989, pp. 2064–2072.

5 Nemeth, N. N., Powers, L., Janosik, L., and Gyekenyesi, J. P., "Designing Ceramic Components for Durability," *Am. Ceram. Soc. Bull.*, Vol. 72, 1993, pp. 59–69.

6 Snedden, J. D., and Sinclair, C. D., "Volume Effect on Creep Strength of Ceramics: Fact or Fiction," *Br. Ceramic Transactions*, Vol. 91, 1992.

7 Snedden, J. D., "Prediction of Creep Reliability of Ceramics: Some Approximate Engineering Techniques," *Br. Ceramic Transactions*, Vol. 92, 1993.

8 Ferber, M. K., and Jenkins, M. G., "Evaluation of the Strength and Creep-Fatigue Behaviour of a HIPed Silicon Nitride," *J. Am. Ceram. Soc.*, Vol. 75, 1992, pp. 2453–2462.

9 Menon, M. N., Fang, H. T., Wu, D. C., Jenkins, M. G., Ferber, M. K., More, K. L., Hubbard, C. M., Nolan, T. A., "Creep and Stress Rupture Behavior of an Advanced Silicon Nitride, Part I—Experimental Observations," *J. Am. Ceram. Soc.*, Vol. 77, 1994, pp. 1217–1227.

10 Chuang, T. J., "Estimation of Power-Law Creep Parameters From Bend Test Data," *J. Mater. Scien.*, Vol. 21, 1986, pp. 165–175.

11 Ferber, M. K., Jenkins, M. G., and Tennery, V. J., "Comparison of Tension, Compression, and Flexure Creep for Alumina and Silicon Nitride Ceramics," *Ceram. Eng. Sci. Proc.*, Vol. 11, 1990, pp. 1028–1045.

12 American Society for Testing and Materials, "Standard Test Method for Tensile Strength of Monolithic Advanced Ceramics at Elevated Temperatures," Committee C28 "Advanced Ceramics," ASTM, Philadelphia, PA, 1995.

13 ASTM STP 470A "Manual on the Use of Thermocouples in Temperature Measurement," ASTM, Philadelphia, PA, 1960.

14 Carroll, D. F., Wiederhorn, S. M., and Roberts, D. E., "Technique for Tensile Creep Testing of Ceramics," *J. Am. Ceram. Soc.*, Vol. 72, 1989, pp. 1610–1614.

CT Multiscan: Using Small Area Detectors to Image Large Components

E. A. Sivers

W. A. Ellingson

S. A. Snyder

D. A. Holloway

Argonne National Laboratory,
Argonne, IL 60439

The small size and dynamic range of the best two-dimensional X-ray detectors are impediments to the use of three-dimensional X-ray computed tomography (3D-XRCT) for 100 percent inspection of large ceramic components. The most common industrial 3D-XRCT systems use a "rotate-only" geometry in which the X-ray source and the area detector remain stationary while the component placed between them is rotated through 360 deg. This configuration offers the highest inspection speed and the best utilization of X-ray dose, but requires that the component be small enough to fit within the X-ray/detector "cone." Also, if the object is very dense, the ratio of an unattenuated X-ray signal to that through the longest path in the component may exceed the dynamic range of the detector. To some extent, both of these disadvantages can be overcome by using "Multiscan CT," i.e., scanning small overlapping regions of a large component separately while maximizing the X-ray dose to each. The overlapping scans can then be combined seamlessly into a single scan with optimal contrast.

Introduction

Three-dimensional X-ray computerized tomography (3D-XRCT) is one of the best nondestructive inspection methods for dimensional analysis, density measurement, and flaw detection in complex objects. Of the possible implementations of 3D-XRCT, the rotate-only geometry is most favored because of its efficient dose utilization and speed. Its only moving part is the turntable, which rotates the object through 360 deg during data collection.

In the conventional method, the scanned object must fit within the cone formed by the X-ray source and the area detector, and contrast discrimination within the object is limited by the dynamic range of the area detector/data collection system. The best area detectors for use with 3D-XRCT are less than 25 cm in width (Bueno et al., 1992; Sivers et al., 1994) and the diameter of an object that can fit within the source/detector cone is much less than this. However, with a simple modification of the usual geometry, it is possible to scan overlapping regions of a large object separately and merge them into a single data set. This process is called "multiscan CT." In many cases, it is also possible to isolate regions having a narrow range of thickness, which permits the maximization of X-ray flux in these sections and stretches the effective dynamic range of the system.

For multiscan CT, the origin of the turntable must be rotated at a constant radius around the X-ray source in the plane perpendicular to the area detector between data acquisition periods. However, it can be shown that one rotation of the turntable about the X-ray source is equivalent to two orthogonal translations in the perpendicular plane plus a rotational increment of the turntable about its own axis. Since most 3D-XRCT systems have two orthogonal translation stages to position the turntable, multiscan CT can usually be performed without modifying the hardware of a conventional 3D-XRCT system.

Multiscan CT image quality is limited primarily by the precision with which the imaging system is calibrated. It is necessary to measure the distance between the X-ray source and the turntable very accurately, and the overlapping regions of multiple scans must be scaled to equalize incident X-ray flux. Errors in these determinations cause concentric rings to appear at the boundaries of adjacent scans. However, these artifacts can be reduced to nearly the level of the image noise in a well-calibrated system. Because of the ability to stretch the dynamic range of the detector, signal-to-noise ratios in a multiscan reconstruction will be higher than those on a single-scan reconstruction made by the same scanner, and contrast discrimination will be correspondingly superior. It is also possible to increase spatial resolution over that obtainable in a single scan by the use of geometric magnification, but multiscan samples are often averaged to reduce the data load to manageable proportions.

X-ray source strength and scan time impose operational limits on the size of an object that can be scanned on a given system using multiscan CT. Although the distance between the source and the detector is usually adjustable, X-ray flux must be sufficiently great that a statistically meaningful signal can be collected in a reasonable time. In addition, the total fan angle cannot exceed $\pi/2$ because of algorithmic concerns.

The steps needed to perform multiscan CT on a conventional 3D-XRCT system are outlined here. Several experimental images are also presented, and a qualitative comparison is made between a multiscan reconstruction taken on the system shown in Fig. 1 and a single-scan reconstruction of the same object taken on a two-dimensional (2D) medical system. We have found multiscan CT to be a useful tool for increasing the effective capacity and contrast discrimination of our scanner.

Scanner Design

The 3D-XRCT system diagrammed in Fig. 1 was described earlier by Ellingson et al. (1991). Its primary components are a 10-μm microfocus X-ray source, a 9-in. dia image intensifier (II), and a CCD camera coupled to the II output screen via a lens. The object to be imaged is placed on a turntable that is rotated through 360 deg about its vertical axis during a conventional scan. The turntable is mounted on a translational (X) stage that moves perpendicular to the area detector, and the X stage is itself mounted on a translational (Y) stage that moves parallel to the area detector.

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute March 2, 1995. Paper No. 95-GT-297. Associate Technical Editor: C. J. Russo.

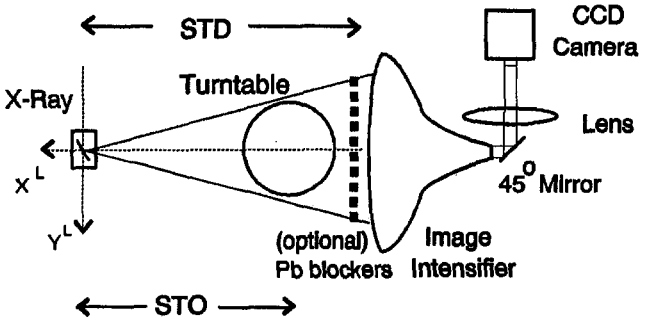


Fig. 1 Diagram of experimental microfocus 3D-XRCT system using an image intensifier area detector coupled to a CCD camera by a lens

Also shown in Fig. 1 are optional lead "blockers" that can be interposed between the X-ray source and the detector. These are lead strips that are ≈0.5 cm wide and thick enough to stop virtually all of the incident X-ray flux. Any signal recorded by the detector in the path of these blockers will be spurious. In practice, each scan is taken twice: once without the blockers and once with them. Then, a low-order curve is fit to the data taken through lead paths to approximate any low-spatial-frequency offset, which is subtracted from the scan without blockers.

Typical factors contributing to the detector signal offset are the background signal of the detector circuit (dark current), X rays scattered from other paths into the detector, and light scattered within the output phosphor of the II (veiling glare). Of these, veiling glare makes the largest contribution to an II detector offset, and can be several hundred percent higher than the true signal in paths through a dense object. Dark current can be measured independently by turning off the X-ray source. X-ray scatter is usually only 10–20 percent of the true signal if the object is not too close to the detector. Thus, lead-blocker measurements can be omitted if an area detector that does not suffer from veiling glare is used. Alternatively, it is also possible to correct for II veiling glare with the 2D deconvolution method of Siebert et al. (1985) if a large vertical portion of the screen is used and pixel averaging is not performed in the data acquisition circuitry of the camera.

Ordinarily the X and Y stages are used to position an object within the cone made by the X-ray source and the II. However, they can also be used to position the turntable for multiscan data acquisition as shown in Fig. 2. Here, R is the radius of the object to be scanned, STO is the source-to-object distance, STD is the source-to-area-detector distance, ΔX is the perpendicular increment from dead center, and ΔY is the parallel increment from dead center. Dead center is defined to coincide with the perpendicular from the source to the detector plane. ΔX and ΔY are constrained by the requirement that the turntable axis must remain a constant distance STO from the X-ray source. In addition, when the turntable is repositioned from dead center, its initial orientation must be rotated by an amount $\Delta\theta = \tan^{-1}(-\Delta Y/\mathrm{STO})$.

It is shown in the appendix that the combination of two translations plus a rotation of the turntable is equivalent to rotating the turntable about the X-ray source. As is evident from Fig. 2, rotating the origin of the turntable about the perpendicular axis of the X-ray source in one direction is equivalent to rotating the source and detector rigidly about the same axis in the opposite direction, but the former is much easier to implement. It is also seen that the samples obtained by a rotated source and detector must be interpolated to provide the same sampling as a "virtual" extension of the detector in its original orientation. Analogous resampling is required to standardize the sample increments in data obtained by rotating the turntable.

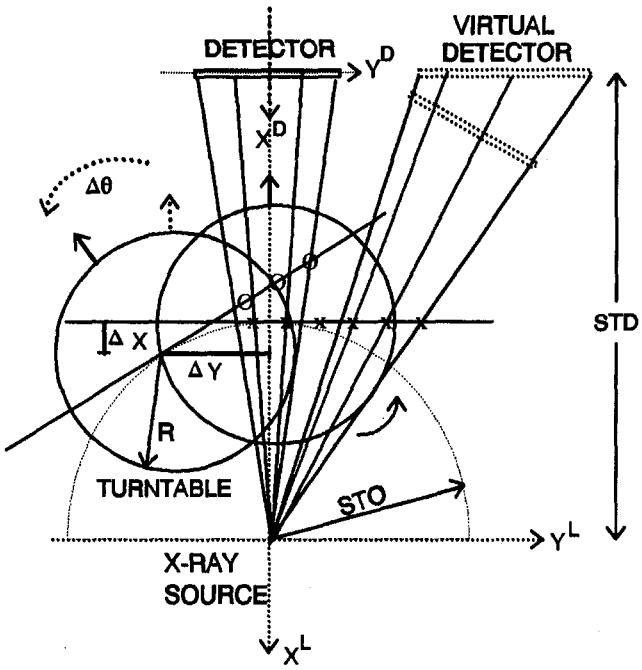


Fig. 2 Diagram of 3D-XRCT data acquisition system

Multiscan CT

Figure 3 is an outline of the procedures used to implement multiscan CT on a conventional rotate-only 3D-XRCT system.

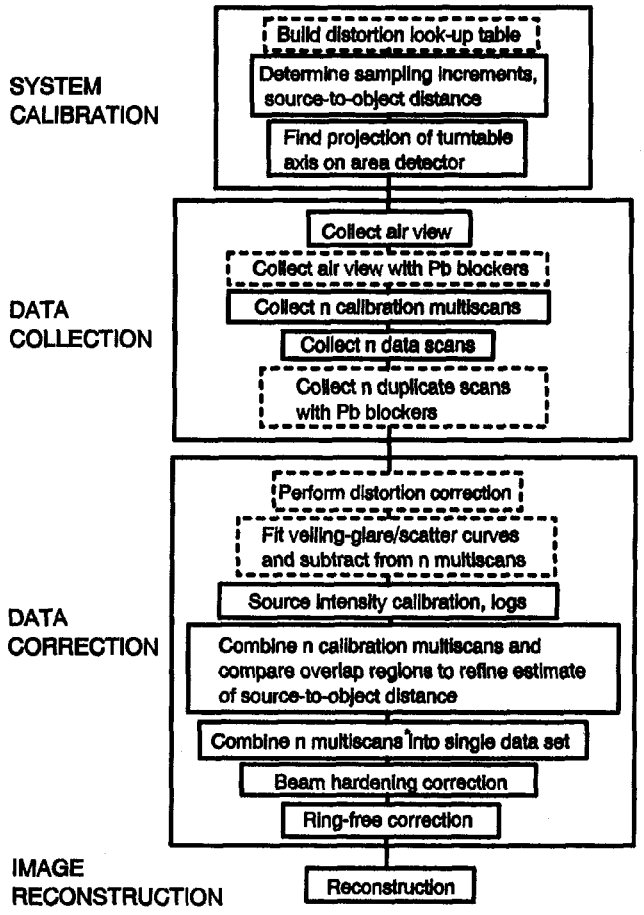


Fig. 3 Block diagram of multiscan CT data acquisition and processing

The processes divide naturally into system calibration, data collection, data correction, and reconstruction. Steps in dashed blocks relate to veiling glare/scatter correction and can be omitted if the area detector used has very low veiling glare and X-ray scatter is not significant.

System Calibration

Each time the dead-center source-to-object distance or the source-to-detector distance is changed to accommodate a new object, the 3D-XRCT system must be recalibrated. If only the dead-center position of the turntable has changed, the source-to-object distance and the sample spacing at the turntable axis must be determined. Usually, this is done by comparing two digital radiographs (views) of a precise grid in two known positions. It is also necessary to find the projection of the turntable axis on the area detector by scanning a vertical wire placed near the center of the turntable.

If the X-ray source or the detector is moved and the area detector is an II, it is also advisable to remake the spatial distortion look-up table. The output image from an II is distorted by the curvature of its glass envelope and by nonuniformities in the field accelerating the electron beam. The distortion look-up table is derived from a view of a uniform matrix that is placed on the face of the II. Note that all views taken on an II should be distortion-corrected before further processing.

Data Collection

Because of the spatial nonuniformity of the X-ray source intensity, it is necessary to take a view in air for data calibration. In an idealized form, the signal $I(y, z)$ detected by an X-ray detector when an X-ray of initial intensity $I_o(y, z)$ traverses a pathlength x of a material having linear attenuation coefficient μ can be written as

$$I(y, z) = I_o(y, z)e^{(-\mu x)}. \tag{1}$$

Taking the natural logarithm of both sides of Eq. (1) gives

$$\mu x = \ln I_o(y, z) - \ln I(y, z), \tag{2}$$

which is used to reconstruct a 2D distribution of μ. The digital radiograph of air provides the values $I_o(y, z)$. Although the air view contains no scatter, it will be corrupted with veiling glare if the detector is an II and the exposure should be repeated with intervening lead blockers.

The number n of multiscans to be collected for a particular object is a function of the width of the X-ray cone at the center of the turntable at dead center and the width of the object. The object should be positioned so that there is a generous overlap between adjacent scans, and regions of air should be available on either side. In addition, it may be desirable to adjust geometric magnification so that regions of the object having similar pathlengths appear within one scan. For instance, if the object is a solid cylinder, magnification should be adjusted so that center regions of the cylinder fill the detector during the dead-center scan. This will allow higher flux to be used for this scan without saturating the detector.

Because of the requirement for a precise measurement of the source-to-object distance, it is useful to take a multiscan data set of some high-resolution object. By cross-correlating overlapping regions of these scans, it is possible to determine any sample misalignment and adjust the estimate of the source-to-object distance to remove it.

If the area detector is an II, or if X-ray scatter is expected to be high, $2n$ scans of the object should be taken; n with lead blockers and n without (or a software veiling-glare correction should be performed). Note that the veiling-glare/scatter offset is more harmful to readings taken through long pathlengths or highly attenuating objects because it will be a higher percentage of the true transmitted signal. The process of taking natural logarithms of these corrupted readings will introduce large nonlinearities. These are manifested in a reconstructed image as decreased density in the center of uniform objects and as "streaking" between high-density objects.

Data Correction

Data taken with an II should be corrected for distortion and veiling-glare/scatter before further processing. The distortion correction is a simple 2D interpolation made from the look-up table created in the system calibration. The veiling-glare/scatter correction involves fitting low-order curves to the data values recorded under the lead blockers in each view of each scan and subtracting the resulting offset curve.

Before multiscan data combination, the spatial variation of the X-ray intensity must be removed by subtracting the natural logarithm of each view from the natural logarithm of the air view as shown in Eq. (2). Because X-ray flux is tailored to each scan and may fluctuate during a single scan, it is also necessary to normalize source intensity before combining data. We have chosen to scale the intensity of one exterior scan to the air view and to chain the scaling of adjacent scans by comparing overlapping regions.

Before the n scans can be combined, it is necessary that the horizontal and vertical sample increments in each scan be the same. It is evident from Fig. 2 that the data collected when the turntable is rotated from dead center has different sample spacing. It is shown in the appendix (see also Sivers, 1995) that equilinear samples (y, z) on the virtual detector plane can be interpolated from equilinear samples (y'', z'') in off-centered multiscans by the formulas

$$y'' = \frac{\text{STD}(\text{STD} \sin \Delta\theta + y \cos \Delta\theta)}{(\text{STD} \cos \Delta\theta - y \sin \Delta\theta)} \tag{3}$$

and

$$z'' = \frac{(\text{STD}\ z)}{(\text{STD} \cos \Delta\theta - y \sin \Delta\theta)}, \tag{4}$$

where STD is the perpendicular distance between the detector and the X-ray source. Sample spacing is adjusted by means of 2D interpolation.

To minimize artifacts in multiscan data sets, it is useful to examine the n scans of the high-resolution calibration object to estimate any alignment mismatch. This is done by cross-correlating overlapping regions from adjacent scans, determining the amount of misalignment, and computing the value of the source-to-object distance that will eliminate it. Once the appropriate value has been determined, it can be used for all multiscans taken with the same geometric configuration.

It is possible to obtain extremely high spatial resolution by using geometric magnification and multiscan data acquisition. However, the processing time increases as the square of the number of samples, and it is usually preferable to average samples by the number of multiscans to reduce the data load.

The expression for transmitted radiation given in Eq. (1) is a simplification that is valid only for mono-energetic X-ray sources. Conventional bremsstrahlung X-ray sources have a normalized probability distribution of energies $P(E)$ in the range $(E_{\min}, E_{\max})$, and linear attenuation coefficients also have an energy dependence $\mu(E)$. With these modifications, Eq. (1) becomes

$$I(y, z) = I_o(y, z) \int_{E_{\min}}^{E_{\max}} P(E)e^{-\mu(E)x}dE \tag{5}$$

Subtracting the natural logarithm of Eq. (5) from the natural logarithm of an air path does not give a simple product of linear attenuation coefficient and pathlength. The nonlinearities

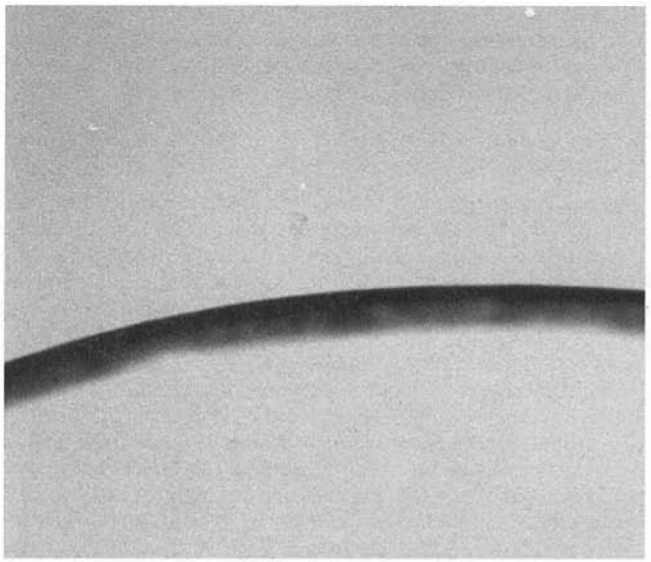


Fig. 4(*a*) Relatively homogeneous segment from 2-mm-thick multiscan 3D-XRCT slice of Al_2O_3/Al_2O_3 combustor liner

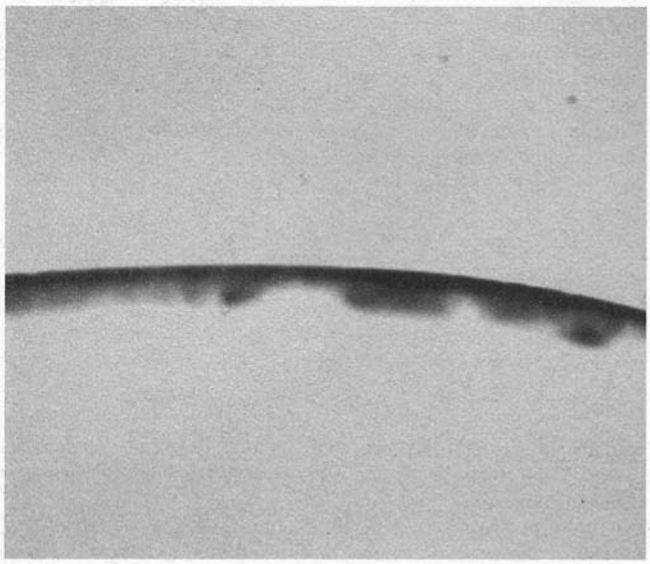


Fig. 4(*b*) Segment from 2-mm-thick multiscan 3D-XRCT slice of Al_2O_3/Al_2O_3 combustor liner showing voids and inclusions

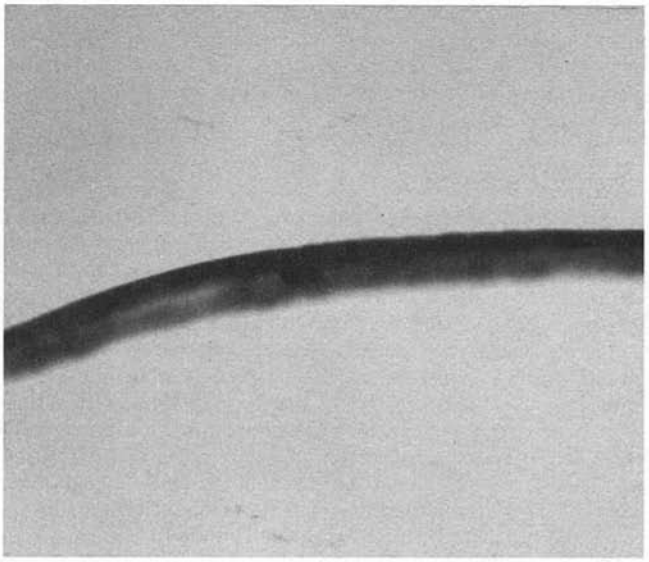


Fig. 4(*c*) Segment from 2-mm-thick multiscan 3D-XRCT slice of Al_2O_3/Al_2O_3 combustor liner showing delamination (left)

introduced by a polyenergetic X-ray source produce "beam-hardening" artifacts similar to those caused by veiling-glare/scatter offset. If only one type of material is present, it is possible to build a beam-hardening calibration table by recording the signals transmitted through known thicknesses of the material and linearizing them.

Scanners of the design shown in Fig. 1 have the disadvantage that a sample at a fixed distance from dead center is always recorded by the same detector element. Because detector elements are not precisely identical, their sensitivity will differ, resulting in concentric-ring artifacts in reconstructed slices near the plane perpendicular to the area detector. These rings can be mitigated somewhat by computing scan-wide averages for a given sample and generating equalizing factors.

Multiscan Reconstructions

Combined multiscan data sets are reconstructed in the same manner as single-scan data sets. We have used the 3D algorithm of Feldkamp et al. (1984). Objects that we have scanned successfully with a 20-cm-dia II include a 20-cm-dia Al_2O_3/Al_2O_3 combustor liner, a 20-cm-dia SiC/SiC combustor liner, and a 17-cm-dia Si_3N_4 turbine rotor.

Combustor liners are hollow, right-circular, ceramic-composite cylinders with wall thicknesses of 0.3–0.4 cm and densities of 3–4 g/cm^3. Figure 4 shows three selected arcs from a reconstructed 2-mm-thick slice of an Al_2O_3/Al_2O_3 combustor liner. In all of the images presented here, the dark areas represent high density and the light areas represent air. Five scans were used to make the composite data set, and this segment of the reconstructed image has 803 × 803 0.00044-cm^2 pixels. Although the segment in Fig. 4(*a*) shows a relatively homogeneous region, some density variation can be seen. The segment in Fig. 4(*b*) contains obvious voids and inclusions, and the segment in Fig. 4(*c*) shows a clear delamination.

Figure 5 shows three selected arcs from a reconstructed 2-mm-thick slice of a SiC/SiC combustor liner. Five scans were used to make the composite data set, and this segment of the reconstructed image has 715 × 715 0.0003-cm^2 pixels. The segment shown in Fig. 5(*a*) is a relatively "good" section with uniform structure and density. The lighter sections of the segments shown in Figs. 5(*b*) and 5(*c*) have mean densities of only about half that of the good section.

For comparison, the SiC/SiC combustor liner was also scanned on a 2D medical CT scanner. Figure 6 shows the same slice and segment shown in Fig. 5(*c*). This reconstructed image is also 2 mm thick and has 512 × 512 0.0019-cm^2 pixels. Obviously, more than one image is required to make a definitive comparison of two CT scanners, but the superior spatial resolution of Fig. 5(*c*) makes it the clear choice for this particular application.

Figure 7 shows one 2-mm slice from the multiscan reconstruction of a 17-cm-dia Si_3N_4 turbine rotor (the inner core is ≈10 cm in diameter). Three scans were used to make the composite data set, and this image has 757 × 757 0.0006-cm^2 pixels. The image is comparatively free from image artifacts, but suffers from a rather high level of statistical noise. In this case, noise would be improved by taking five or seven multiscans and increasing the flux through the central regions substantially.

Conclusions

We have found multiscan CT to be effective for increasing the capacity and dynamic range of our 3D-XRCT system. The ability to increase the flux through the thicker regions of an

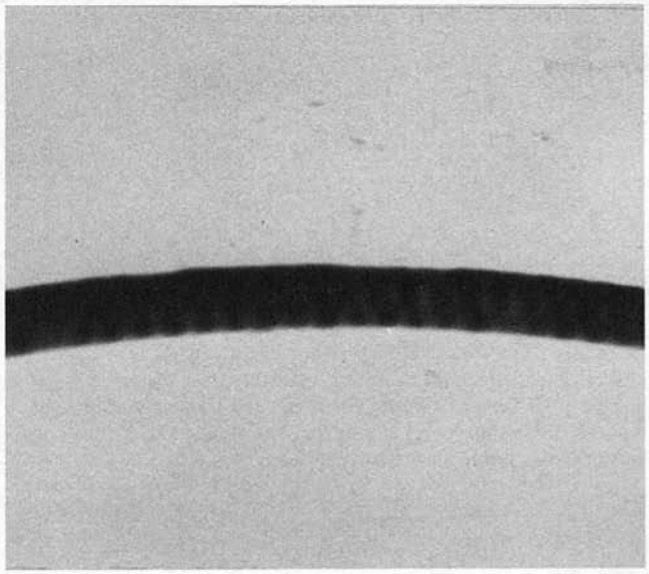


Fig. 5(*a*) Relatively homogeneous segment from 2-mm-thick multiscan 3D-XRCT slice of SiC/SiC combustor liner

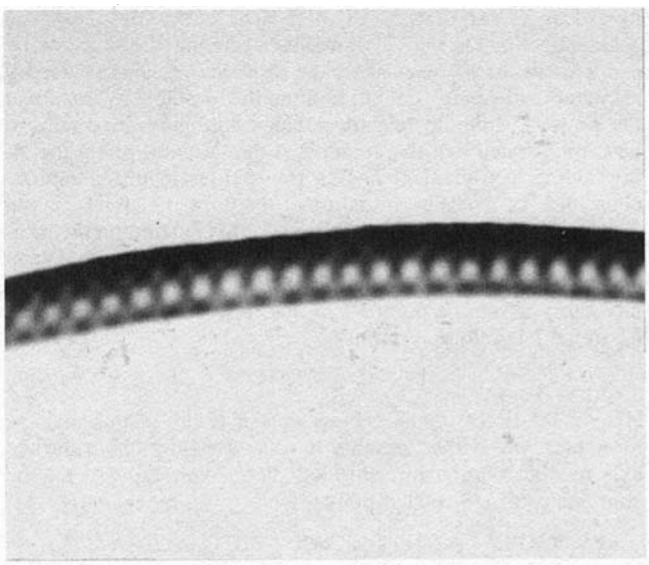


Fig. 5(*b*) Segment from 2-mm-thick multiscan 3D-XRCT slice of SiC/SiC combustor liner showing low-density (50 percent) regions

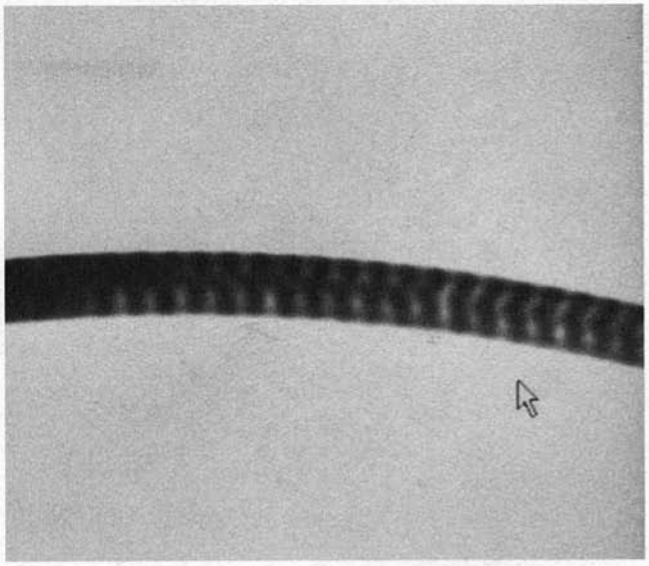


Fig. 5(*c*) Segment from 2-mm-thick multiscan 3D-XRCT slice of SiC/SiC combustor liner showing weave debond

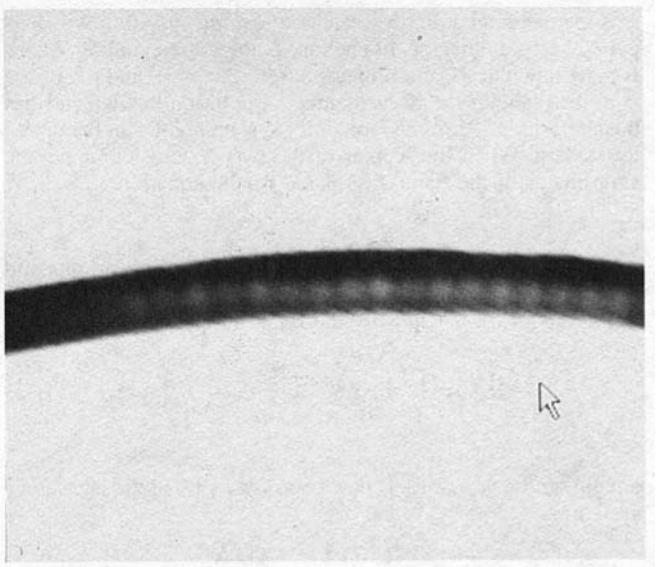


Fig. 6 Same 2-mm segment of SiC/SiC combustor liner shown in Fig. 5(*c*), reconstructed from single 2D-XRCT scan

object is so useful that multiscan data collection has been used even on objects that fit within the beam cone. It also makes possible the use of other excellent area detectors, such as fiber-optic scintillators, which are even smaller than an II (Bueno et al., 1992). Its major drawback is an increase in data-collection and preprocessing time, but reconstruction time is somewhat less than that needed for a single scan if samples are averaged by the number of constituent scans. There can also be ring artifacts at the interfaces between adjacent data sets, but these can be reduced to nearly the level of statistical noise in a well-calibrated system.

Acknowledgments

This work was supported by the US Department of Energy, Office of Fossil Energy, Advanced Research and Applied Technology, Materials Program, under Contract No. W-31-109-ENG-38.

References

Bueno, D., Barker, M. D., Condon, P. E., and Betz, R. A., 1992, "Solid-State X-Ray Imaging Methodology," WL-TR-92-4003, Materials Directorate, Wright Laboratory, Wright-Patterson Air Force Base, OH.

Ellingson, W. A., Vannier, M. W., and Stintor, D. P., 1991, in: *Characterization of Advanced Materials,* W. Altergath and E. Henneke, eds., Plenum Press, New York, pp. 2–25.

Feldkamp, L., Davis, L., and Kress, J., 1984, "Practical Cone Beam Algorithm," *J. Optical Society of America, A,* Vol. 1(6), pp. 612–619.

Siebert, J. A., Nalcioglu, O., and Roeck, W. W., 1985, "Removal of Image Intensifier Veiling Glare by Mathematical Deconvolution Techniques," *Medical Physics,* Vol. 12(3), pp. 280–288.

Sivers, E. A., Holloway, D. A., and Ellingson, W. A., 1994, "Predicting the Performance of 3-D X-ray Computerized Tomography Systems," *Nuclear Instruments and Methods in Physics Research, A,* Vol. 345, pp. 179–197.

Sivers, E. A., 1995, "Use of Multiple CT Scans to Accommodate Large Objects and Stretch Dynamic Range of Detectability," *Proc. Thirteenth International Accelerator Conference,* Denton, TX, Nov 7–10, 1994; accepted for publication in *Nuclear Instruments and Methods in Physics Research, B.*

APPENDIX

To derive Eqs. (3) and (4), consider the two coordinate systems illustrated in Fig. 2: The laboratory (L) coordinate system is centered on the X-ray source; the detector (D) coordinate system is centered on the detector with the positive $\mathbf{X}^D$

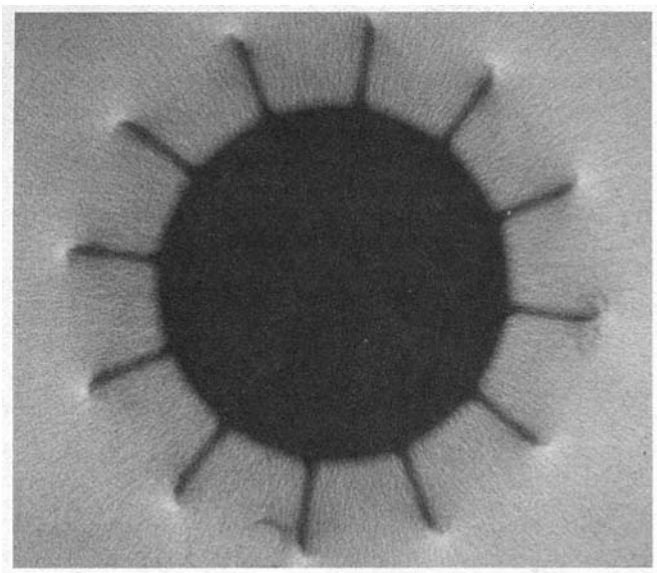


Fig. 7 2-mm-thick multiscan 3D-XRCT slice of 17-cm-dia Si_3N_4 turbine rotor

axis perpendicular to the detector and the positive $\mathbf{Y}^D$ axis parallel to the detector. In each case, the corresponding $\mathbf{Z}$ axis is perpendicular to, and projects out of, the page, and the origin is in the plane of the X-ray source. A point on the detector has the coordinates $\mathbf{P}^L$ in the laboratory system and $\mathbf{P}^D$ in the detector system. When the $\mathbf{X}$ axes of the two systems coincide, the coordinates of the point are related by the equation

$$\mathbf{P}^L = \mathbf{P}^D - \text{STD}\ \mathbf{X}^L, \tag{A-1}$$

where STD is the distance between the coordinate origins and $\mathbf{X}^L$ is the unit vector in the X direction. If the detector is rotated rigidly about the $\mathbf{Z}^L$ axis by the rotation operator

$$\mathbf{R}_z^{\Delta\theta} = \begin{pmatrix} \cos\Delta\theta & -\sin\Delta\theta & 0 \\ \sin\Delta\theta & \cos\Delta\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{A-2}$$

a point on the detector has new coordinates $\mathbf{P}_R^L$ in the laboratory system,

$$\mathbf{P}_R^L = \mathbf{R}_z^{\Delta\theta}(\mathbf{P}^D - \text{STD}\ \mathbf{X}^L). \tag{A-3}$$

Equation (A-3) can also be inverted to express the coordinates of a point in the laboratory system in the coordinates of the rotated detector system as follows:

$$\mathbf{P}^D = (\mathbf{R}_z^{\Delta\theta})^{-1}\mathbf{P}_R^L + \text{STD}\ \mathbf{X}^L. \tag{A-4}$$

An X-ray in the laboratory system can be expressed as the line $\mathbf{L}$ drawn between the origin $\mathbf{O}$ and a point on the virtual detector $\mathbf{V}$, where

$$\mathbf{O} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}, \quad \mathbf{V} = \begin{pmatrix} -\text{STD} \\ y \\ z \end{pmatrix},$$

$$\text{and} \quad \mathbf{L} = \begin{pmatrix} \tau\cos\alpha \\ \tau\cos\beta \\ \tau\cos\gamma \end{pmatrix}. \tag{A-5}$$

In Eq. (A-5), τ is a distance scalar and $\cos\alpha$, $\cos\beta$, and $\cos\gamma$ are the direction cosines of the line in the $\mathbf{X}^L$, $\mathbf{Y}^L$, and $\mathbf{Z}^L$ directions, respectively. If $\mathbf{L}$ is substituted for $\mathbf{P}_R^L$ in Eq. (A-4), the point $\mathbf{P}^D$ in the rotated detector coordinated system that is intersected by the X-ray can be expressed in terms of the coordinates of the point on the virtual detector:

$$\mathbf{P}^D = \begin{pmatrix} 0 \\ y'' \\ z'' \end{pmatrix} \tag{A-6}$$

Solving for τ and substituting in the expressions for y'' and z'' gives Eqs. (3) and (4).

Examination of Fig. 2 shows that rotating the detector by $-\Delta\theta$ about the $\mathbf{Z}^L$ axis while the turntable remains stationary is operationally equivalent to rotating the turntable by $\Delta\theta$ about the $\mathbf{Z}^L$ axis while the detector remains stationary. It remains to be demonstrated that this rotation of the turntable about the X-ray source is equivalent to two translations in the laboratory coordinate system plus a rotation of the turntable about its own Z axis. Consider a third coordinate system centered on the turntable (T) with its origin in the plane of the X-ray source. When the $\mathbf{X}^T$ axis of the turntable coincides with the $\mathbf{X}^L$ axis of the laboratory system, a point $\mathbf{P}^T$ on the turntable has coordinates $\mathbf{P}^L$ in the laboratory system.

$$\mathbf{P}^L = \mathbf{P}^T - \text{STO}\ \mathbf{X}^L \tag{A-7}$$

Here, STO is the distance between the X-ray source and the turntable axis. If the turntable is rotated rigidly about the $\mathbf{Z}^L$ axis by the rotation operator $\mathbf{R}_z^{\Delta\theta}$ defined in Eq. (A-2), the coordinates of the rotated point $\mathbf{P}_R^L$ are given by

$$\mathbf{P}_R^L = \mathbf{R}_z^{\Delta\theta}\mathbf{P}^T - \text{STO}\cos\Delta\theta\mathbf{X}^L - \text{STO}\sin\Delta\theta\mathbf{Y}^L \tag{A-8}$$

The difference between Eqs. (A-8) and (A-7) is

$$(\mathbf{R}_z^{\Delta\theta} - \mathbf{I})\mathbf{P}^L = (\mathbf{R}_z^{\Delta\theta} - \mathbf{I})\mathbf{P}^T + \text{STO}(1 - \cos\Delta\theta)\mathbf{X}^L - \text{STO}\sin\Delta\theta\mathbf{Y}^L. \tag{A-9}$$

Thus, the rotation of the turntable about the laboratory $\mathbf{Z}^L$ axis is equivalent to a rotation of the turntable about its own $\mathbf{Z}^T$ axis, followed by two orthogonal translations in the laboratory system.

T. Tsuchiya

Y. Furuse

S. Yoshino

Engineering R&D Center,
Tokyo Electric Power Company,
Yokohama, Kanagawa, Japan

R. Chikami

Y. Tsukuda

Takasago Machinery Works,
Mitsubishi Heavy Industries, Ltd.,
Takasago, Hyogo, Japan

M. Mori

Takasago R&D Center,
Mitsubishi Heavy Industries, Ltd.,
Takasago, Hyogo, Japan

Development of Air-Cooled Ceramic Nozzles for a Power-Generating Gas Turbine

The development of air-cooled ceramic nozzle vanes for a power-generating gas turbine has been reported. To make up the limited temperature resistance of present ceramic materials, the utilization of a small amount of cooling air has been studied for the first-stage nozzle vanes of a 1500°C class gas turbine. A series of cascade tests were carried out for the designed air-cooled Si_3N_4 nozzle vanes under 6 atm and 1500°C conditions. It was confirmed that the maximum ceramic temperature can be maintained below 1300°C by a small amount of cooling air. In spite of the increased thermal stresses from local cooling, all Si_3N_4 nozzle vanes survived the cascade tests, including both steady-state and transients of emergency shutdown. The potential for an air-cooled ceramic nozzle was demonstrated for a 1500°C class gas turbine application.

Introduction

Considering the future energy situation, great expectation has recently been placed on combined cycle power plants as an effective power-generating method for the utilization of fossil fuels. It is becoming the mainstream of thermal power generation. In order to improve the efficiency of the plant, it is necessary to develop a highly efficient gas turbine, which can be operated at higher temperature and with less cooling air than the present level. This can be realized by introducing ceramics, of which the major attraction is their potential capability to operate at high temperatures and in corrosive environments that far exceed the capability of any conventional superalloy systems.

Since 1984, Tokyo Electric Power Company (TEPCO) has been conducting a cooperative research program for an application of ceramics to a power-generating gas turbine with three Japanese gas turbine manufacturers (Toshiba Corporation, Mitsubishi Heavy Industries, Ltd., and Hitachi Ltd.). The research program has been conducted to apply ceramics for hot parts of a power-generating gas turbine to improve the thermal efficiencies of combined cycle power-generating plants. The first objective of this program was to verify adaptability of Si-based monolithic ceramics to the combustor, the first and second-stage nozzles, and the first-stage rotor of a 20 MW class gas turbine with turbine inlet temperature of 1300°C. Combustion tests on the combustor and cascade tests on the nozzles were conducted under full-pressure (15 atm) and full-temperature (1300°C) conditions. Hot spin tests were conducted on the rotor after confirming the validity of the design by cold spin tests and thermal loading cascade tests in a static test rig. The soundness of ceramic components was verified by these tests simulating the actual 1300°C class gas turbine conditions and a positive prospect for an application of ceramics to gas turbine components was reported in our previous papers [1–14].

Further efforts have been conducted in a cooperative program between TEPCO and Mitsubishi Heavy Industries (MHI) to apply ceramics for a gas turbine with a TIT higher than 1300°C.

In the case of applying ceramics to a 1500°C class gas turbine, materials should withstand high-temperature gas flow above 1600°C due to the temperature distribution of combustion gas. Although some B-C-doped SiC ceramics have demonstrated excellent high-temperature strength, high creep resistance, and high oxidation resistance under a static environment, it seems difficult to apply these materials for gas turbine components due to their low toughness and poor thermal shock resistance.

To make up the limited temperature resistance of present ceramic materials, the utilization of a small amount of cooling air has been studied for turbine nozzles. The inner face of the nozzle is cooled by air impingement and this air is discharged and mixed into the main gas flow through cooling slits located at trailing edge of the nozzle: An air-cooled nozzle vane was designed so that the maximum surface temperature can be maintained below 1300°C by a small amount of cooling air. Cascade tests were carried out for the air-cooled Si_3N_4 nozzles to evaluate the validity of the design.

A silicon nitride nozzle with a small amount of cooling air may be useful in the first-stage ceramic nozzle with TIT of 1500°C. This paper will review the design of the air-cooled ceramic nozzle developed for a power-generating gas turbine from structural and material standpoints, and will present the results obtained from tests conducted under various operational conditions.

Ceramic Materials

In recent years a wide variety of silicon-based ceramics has emerged with potential as structural components in gas turbines. Silicon nitride (Si_3N_4) and silicon carbide (SiC) are currently regarded as the most promising candidates for gas turbine application. The available materials represent a large family with wide property variations and different responses to the gas tur-

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 13, 1995. Paper No. 95-GT-105. Associate Technical Editor: C. J. Russo.

Table 1 Properties of ceramics

Material		Si_3N_4 (SN-88)	SiC (C-600)
Manufacturer		NGK	AGC
Density at RT		3.5	3.2
Young's modulus at RT (GPa)		260	412
Thermal expansion coefficient at RT. ~ 800°C ($\times 10^{-6}$/°C)		3.5	4.3
Thermal conductivity at RT (W/mk)		63	45
Bending strength at 4pts. (MPa)	RT	790	577
	HT (1400°C)	760	423
Fracture toughness ($MPam^{1/2}$)		7	5.1

bine environment. Silicon carbide is one of the leading candidates for gas turbine application because of its high strength, good oxidation, and wear resistance at elevated temperatures. SiC also has extremely good creep strength and microstructural stability, and higher thermal conductivity than Si_3N_4. The major disadvantages of SiC when directly compared with Si_3N_4 are its lower fracture toughness and lower thermal shock resistance. The low toughness of SiC is due to its low critical stress intensity factor and low fracture surface energy. The low thermal shock resistance of SiC is due to the combination of its higher thermal expansion and higher elastic modulus in comparison with Si_3N_4.

Si_3N_4 ceramics have excellent strength, toughness, and thermal shock resistance at temperatures below 1300°C, although they tend to degrade at temperatures above 1300°C. However, high-performance Si_3N_4 ceramics have been developed recently that demonstrate little degradation of strength and excellent oxidation resistance up to around 1400°C.

Although corrosion resistance in the combustion gas environment and long-term durability at high temperature are not yet established, the reliability of ceramic components is expected to improve if Si_3N_4 can be adopted instead of SiC. It also appears that advances in Si_3N_4 technology are being made much more rapidly than advances in SiC technology. Therefore, Si_3N_4 ceramics are considered as the major candidate material for the present case.

These materials range from commercially available to highly developmental. Even though some are termed commercially available, they are in a continual state of development and change. Thus, the intended application and the processing maturity must be considered in selecting materials. In order to select appropriate materials, material evaluation tests have been carried out for many ceramics that are commercial and under development. Tests were conducted at room temperature to establish a baseline, and at the appropriate elevated temperature ranges where the properties are expected to begin to change rapidly.

The material and mechanical properties of the ceramics that were used for the cascade testings are presented in Table 1.

Structural Design

In designing the first-stage nozzle vane for a 1500°C class gas turbine, the radial temperature distributions were taken into consideration by assuming the pattern factor of the combustor is 0.10 at an average gas temperature of 1500°C. Then the maximum gas temperature at the inlet of the first-stage nozzle is about 1600°C.

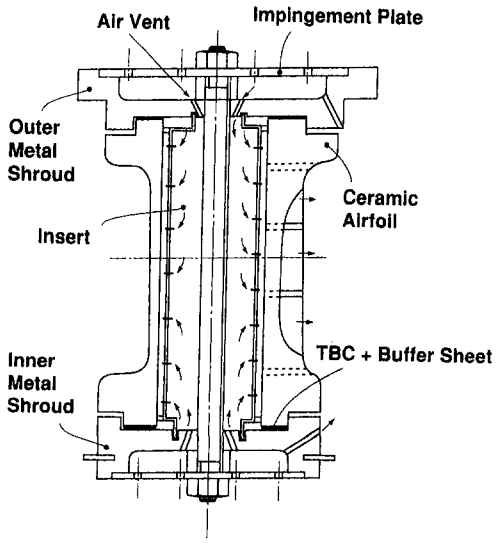


Fig. 1 Assembly construction of cooled ceramic nozzle

Judging from the limited temperature resistance of present ceramic materials, it is necessary to find a way for these materials to withstand high-temperature gas flow of 1600°C. In order to cope with the difficulty, an air-cooled ceramic nozzle vane system has been developed that avoids exposure of ceramic material to the high temperatures exceeding its tolerance level. By avoiding high temperatures in the nozzle vanes, commercially available monolithic materials can be used.

In the preliminary design and stress analysis, it was predicted that the maximum surface temperature of ceramics can be maintained below 1300°C by utilizing a small amount of cooling air.

The assembly construction of the designed air-cooled ceramic nozzle vane is shown in Fig. 1. The details of the cooling slits are presented in Fig. 2. The components of the nozzle assembly are presented in Fig. 3. A one-piece solid ceramic construction was adopted for the airfoil and inner/outer shroud configuration.

Although the three-piece construction for an airfoil and outer/inner shrouds seems to be effective to reduce thermal stress, one-piece construction was chosen in the present case. This is to avoid the unknown factors at the contact surface of ceramics and problems associated with gas leakage between ceramic

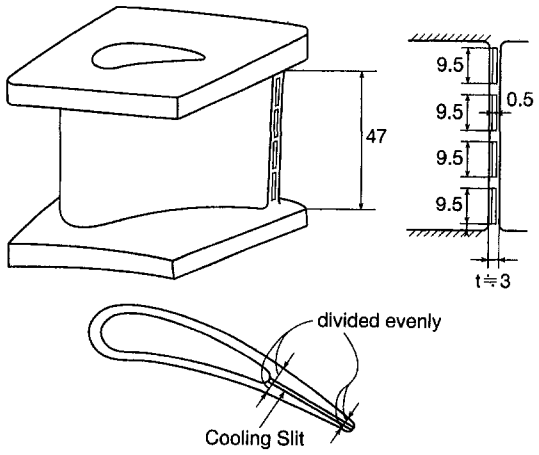


Fig. 2 Configuration of cooled ceramic nozzle vane

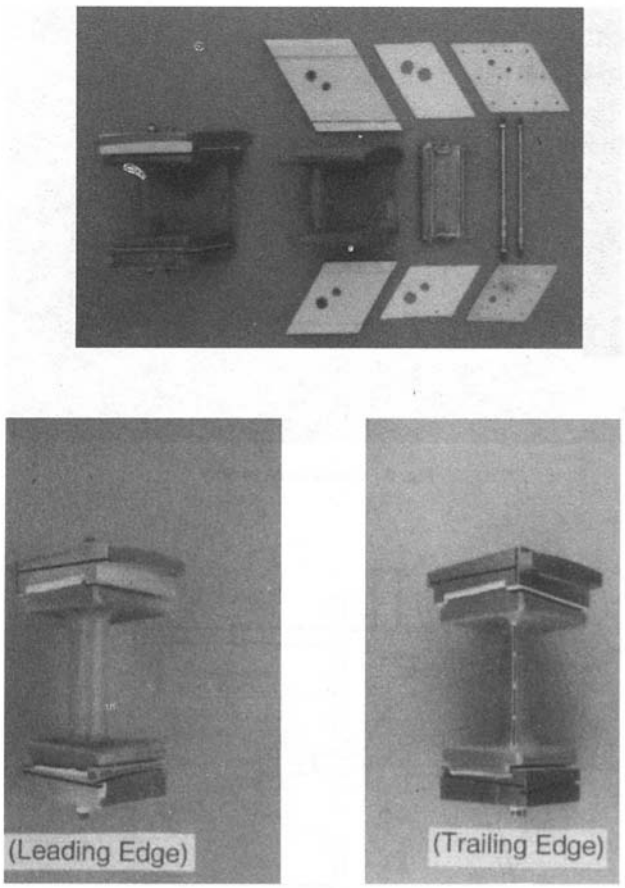


Fig. 3 Component of cooled ceramic nozzle vane

parts, which might be a cause of unexpected failure of ceramics. Moreover, delicate handling is required in assembling the divided construction.

The cooling air is introduced into the nozzle vane through impingement plates, which are located at the outside of inner and outer metal shrouds, and enters into the inside of insert after cooling down the inner and outer metal shrouds. The inner surface of the ceramic nozzle vane is cooled by impingement and convection. The cooling air is discharged and mixed into the main gas flow through cooling slits located at the trailing edge of the ceramic nozzle vane.

A hybrid construction was adopted with metal shrouds and a metal insert along with the ceramic part. The ceramic airfoil and the inner and outer ceramic shrouds were one-piece construction and this ceramic part was tightened and fixed from both sides with the metal shrouds and the metal insert. There were a number of small holes on the surface of the insert to cool the inner surface of nozzle vane by air impingement. Buffer materials were inserted between metal shrouds and ceramic shrouds to make the tightening loads uniform and to insulate heat transfer between them. In addition, to reduce the thermal expansion differences between these metal and ceramic parts, a metal insert of low thermal expansion Ni-base alloy was used. Furthermore, a thermal barrier coating was applied to the inner surface of the metal shrouds. To facilitate manufacture of the three-dimensional configurations of the vanes, the difference between each adjacent airfoil cross section was minimized, resulting in a vane configuration with little twisting. In addition, for the leading edge of the airfoil where the heat flux tends to be quite high, a blunt nose configuration was adopted to keep its heat transfer coefficient as low as possible.

The cooling slits located at a trailing edge of the nozzles are machined using the ultrasonic wave.

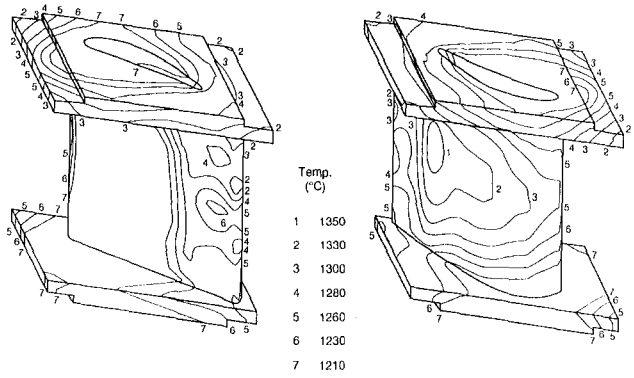


Fig. 4 Temperature distribution (steady state at 1500°C)

In determining the amount of cooling air, two levels were considered. One was the standard cooling in which the amount of cooling air was determined so as to maintain the ceramic temperature at 1300°C. The other was the mild cooling in which the amount of cooling air was reduced so as to maintain 1400°C of ceramic temperature. The amount of cooling air was 1.4 percent of total cycle core flow for the standard cooling and 1.2 percent for the mild cooling.

Stress Analysis

Nozzle vanes experience thermal stresses due to uneven temperature distributions. An FEM stress analysis was carried out for the Si_3N_4 nozzle vane with standard cooling under both steady-state and transient conditions. According to the results of stress analysis, it was found that the most critical stress would be the transient thermal stress generated during emergency shutdown (tripping).

Figure 4 shows the calculated temperature distribution for a steady state condition, and the corresponding stress distribution is presented in Fig. 5. It is observed that the maximum surface temperature of ceramic nozzle vane can be maintained below 1300°C as intended in design. The maximum tensile stress is about 210 MPa, which is generated at either leading edge or trailing edge portion on inner surface of outer shroud.

During shutdown transients, the shroud remains relatively hotter than airfoil section due to the volume effect. The temperature difference between shroud section and airfoil section results in the generation of thermal stresses that tend to be maximized at either leading edge or trailing edge on inner surface of outer shroud. It was found that reducing the shroud thickness was effective to reduce thermal stresses generated during emergency

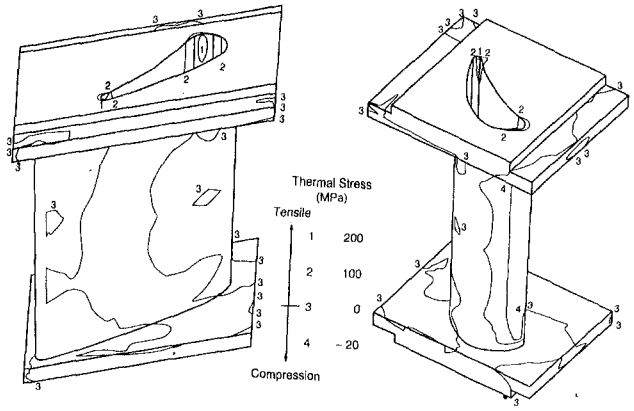


Fig. 5 Thermal stress distribution (steady state at 1500°C)

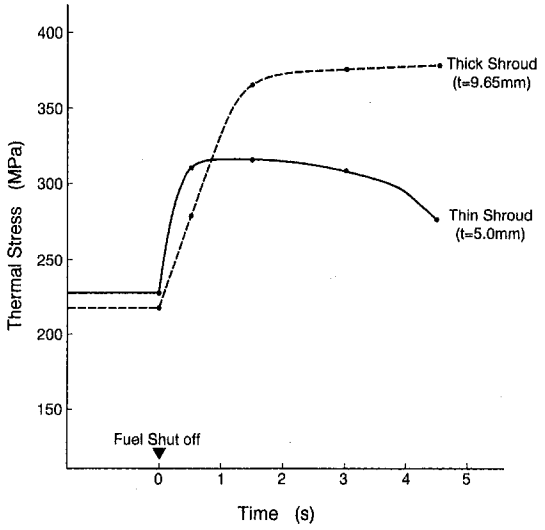


Fig. 6 Transient thermal stresses after emergency shutdown

shutdown as shown in Fig. 6. It may be necessary to consider a trade-off between stress levels and structural integrity.

Cascade Tests

In order to evaluate the design concept of the air-cooled ceramic nozzle vane, a series of intermediate pressure tests were carried out at 6 atm pressure condition. Although the full-pressure condition for the designed first-stage nozzle is 14.9 atm, the lower pressure tests such as 6 atm cascade testings allow an assessment of the validity of the air-cooled hybrid construction and the soundness of ceramics against thermal stresses that are induced by steady/transient state conditions.

Test Facilities. Figure 7 shows the schematic diagram of 6 atm cascade testing equipment, which has been used to evaluate the air-cooled ceramic nozzle vane. The overall test facility is presented in Fig. 8 and a cross section of the cascade testing section is shown in Fig. 9. The cascade testing equipment consists of the housing unit as well as the combustion air and cooling air systems, fuel, exhaust, and cooling waterlines. The testing housing unit consists of the combustor basket, transition piece, inlet duct, ceramic vane cascade, and a casing in which these components are contained. Mounted on the rear end of the casing is a window for an infrared radiation thermometer and a sight glass to observe the ceramic vane under testing.

As shown in Fig. 10, the cascade consists of 4 ceramic vanes and 2 metal dummy vanes at both ends. The view of the tested cascade for the first-stage nozzle vanes is shown in Fig. 11.

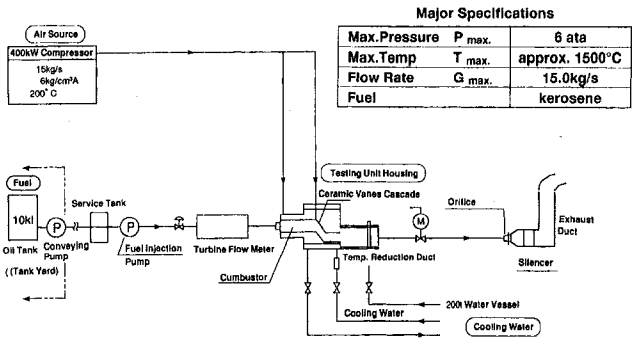


Fig. 7 Schematic diagram of cascade test facility

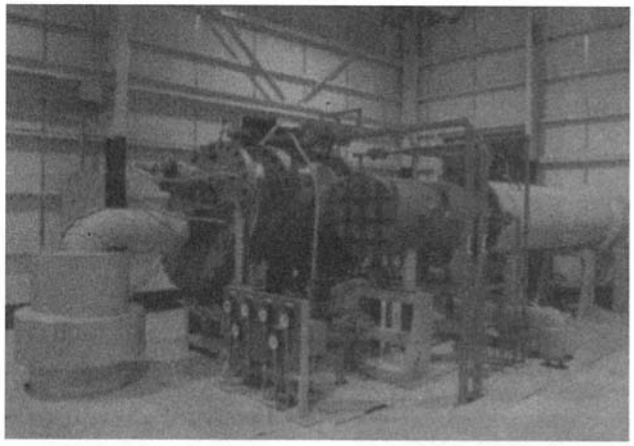


Fig. 8 Overall test facility

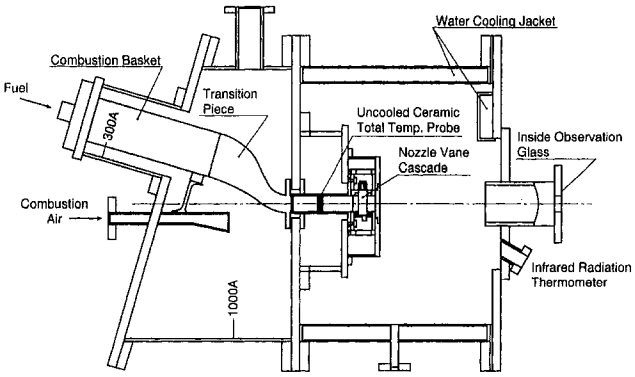


Fig. 9 Cross section of cascade test facility

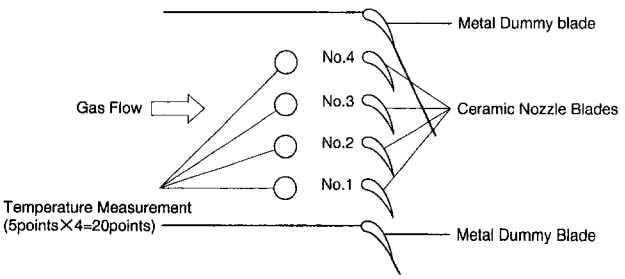


Fig. 10 Configuration of nozzle cascade test

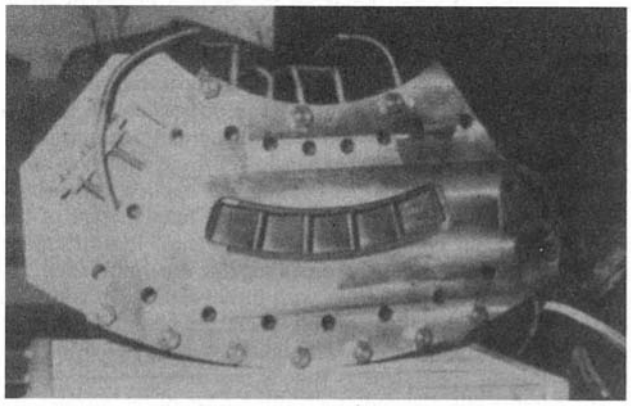


Fig. 11 View of nozzle vane cascade

Table 2 Major measuring items

Major Measuring Items	Measuring Method
1. Main Gas Flow (combustion air flow)	Orifice
2. Fuel Amount (kerosine)	Flow Meter
3. Main Flow Pressure (inlet & outlet)	Pressure Tap
4. Main Gas Flow Temp. Distribution	Ceramic Total Temp. Probe
5. Ceramic Stator Vane Temp. Distribution	Thermocouple & radiation thermometer
6. Metal Core & Metal Shrouds Temperature	Thermocouple
7. Metal Core Cooling Air Flow	Orifice
8. Metal Core Cooling Air Pressure	Pressure Tap
9. Metal Core Cooling Air Temperature	Thermocouple
10. Ceramic Stator Vane Observation	Video Observation from vane suction side

Measurement. The major measuring items and methods are presented in Table 2. The gas temperature was measured by uncooled ceramic total temperature probes, and main air flow and cooling air flow were measured with orifices. The airfoil and shrouds were provided with bores through which thermocouples (0.5 ϕ inconel sheath thermocouple; K type) were inserted to measure temperature distribution. The temperature distribution in the metal core and metal shrouds was measured by means of thermocouples inserted in the metal. Suction side surface temperature of the ceramic vane was measured with infrared radiation thermometer mounted on casing at rear end. The ceramic conditions under testing were observed through the sight glass. The temperature fluctuations at ignition or emergency shutdown were recorded by means of a chart recorder.

Test Conditions and Procedures. Cascade testing conditions are presented in Table 3. The combustion gas temperature and gas flow velocity simulated the full-load conditions of the designed ceramic nozzle vane. The tests were conducted in two steps, i.e., the steady-state test with normal shutdown and emergency shutdown test. The most rigorous was the emergency shutdown test. Due to immediate cutoff of fuel, the gas temperature dropped at once from 1500°C down to air temperature of nearly 400°C. Under emergency shutdown conditions, the ceramic vanes are suddenly cooled down and are subjected to severe thermal stresses. The tested nozzles were disassembled

Table 3 Cascade test conditions

		Design (full load condition)	Test Condition
Turbine Inlet Temp.	Tg	1500 °C	1500 °C
Inlet Total Press.	Pgin	14.9 ata	6 ata
Static Exit Press.	Pgex	8.9 ata	3.6 ata
Press. Ratio	X	1.67	1.67
Velocity (IN/EX)	Vin/Vex	117/635/m/s	117/635/m/s
Gas Flow	Gg	46.4 kg/s*1	2.1 kg/s*2

*1 44 gas flow passages
*2 5 gas flow passages

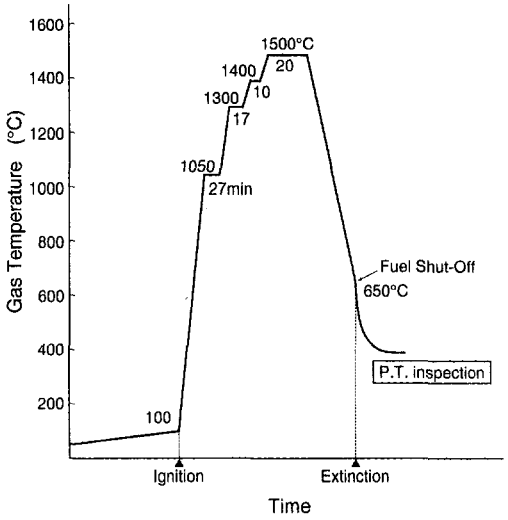


Fig. 12 Test pattern for test 1

and each part was inspected after a series of tests. Visual inspection and fluorescence penetrant inspection were carried out carefully for each part.

Results and Discussion. A series of cascade tests (Tests 1–3) were carried out to evaluate the designed air-cooled nozzle vanes. All tests were conducted at pressure of 6 atm and temperatures up to 1500°C. After confirming the validity of the design concept by Test 1, two emergency shutdown tests (Tests 2 and 3) were conducted to evaluate the durability of the designed ceramic nozzle vanes against severe thermal shock generated by abrupt fuel cutoff.

Test 1: Steady-State Operation Without Emergency Shutdown. The objective of Test 1 was to confirm the validity of the design concept for the air-cooled ceramic nozzle vane. Four air-cooled ceramic nozzle vanes made of Si_3N_4 were attached to the cascade test rig. All tested ceramic vanes were solid one-piece configuration for an airfoil and outer/inner shrouds. Two of them (No. 2 and No. 4 vanes) were the standard cooling type and the rest (No. 1 and No. 3 vanes) were the mild cooling type. The test pattern is presented in Fig. 12. The test was conducted only at steady-state conditions with normal shutdown and no emergency shutdown test was performed. After the test, each vane was disassembled and inspected carefully by the visual appearance and fluorescent penetrant inspections. There was no evidence of erosion on the surface of the tested air-cooled ceramic nozzle vanes. As for the two different cooling levels, there was little difference in the nozzle surface conditions. No cracks were found through the inspection and it was confirmed that the tested air-cooled ceramic nozzle vanes were sound after the steady-state test under 1500°C and 6 atm conditions.

Test 2: Emergency Shutdown Test. In order to evaluate the structural reliability of the developed air-cooled ceramic nozzle vanes against transient thermal stresses generated by emergency shutdown, Test 2 was conducted. Three air-cooled one-piece ceramic nozzle vanes made of Si_3N_4 (No. 1, No. 2, and No. 4 vanes) and one uncooled three-piece ceramic nozzle vane made of Si_3N_4 (No. 3 vane) were assembled and tested under 6 atm pressure and up to 1500°C temperature conditions. Among three air-cooled vanes, two of them (No. 1 and No. 2 vanes) were the standard cooling type and the rest (No. 4 vane) was the mild cooling type. An uncooled ceramic vane, which was developed mainly for a 1300°C application in our previous efforts, was included just for a comparison. The test pattern is presented in

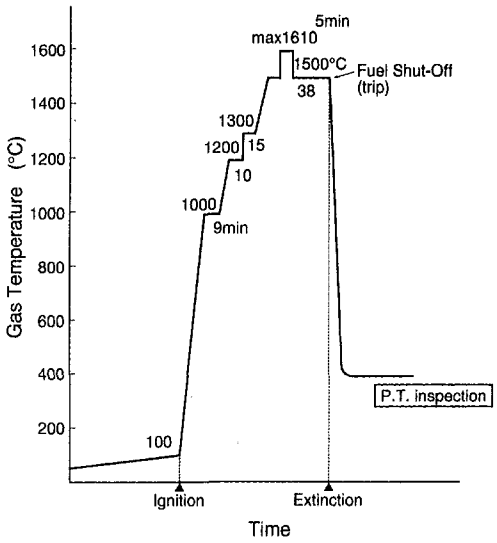


Fig. 13 Test pattern for test 2

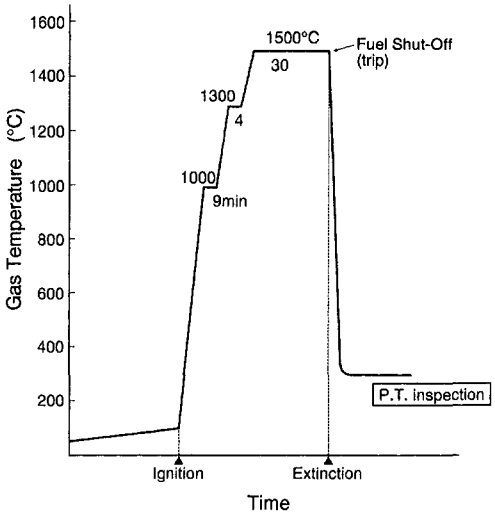


Fig. 14 Test pattern for test 3

Fig. 13. During the test, the gas temperature was quite uniform at the test section and the maximum gas temperature was below the expected one. Therefore, the gas temperature was intentionally increased to 1610°C for a short period to expose ceramics to the design condition. The emergency shutdown test was carried out by abrupt cutoff of fuel at the average gas temperature of 1500°C and pressure of 6 atm. The tested vanes were disassembled for the post-test inspections. The soundness of each vane was confirmed by the visual appearance inspection and fluorescent penetrant inspection.

The erosion of ceramics was observed at trailing edge portion of uncooled Si_3N_4 vanes due to the exposure of ceramics to high temperatures exceeding its tolerance levels. On the other hand, there was no evidence of such erosion on the surface of air-cooled Si_3N_4 vanes. It is considered that this proves the effectiveness of the cooling by a small amount of air. In considering severe thermal stresses generated during emergency shutdown, the degradation due to the corrosion may lead to a fatal damage on ceramic parts. Therefore, it is very important to maintain ceramic temperatures below tolerable levels.

Test 3: Ceramic Temperature Measurement. The detailed temperature measurements were performed in Test 3 for the ceramic nozzle vanes. The tested vanes were two air-cooled one-piece ceramic vanes made of Si_3N_4 (No. 1 and No. 2 vanes) and two uncooled one-piece ceramic vanes made of SiC (No. 3 and No. 4 vanes). The uncooled SiC nozzles were included mainly to conduct the detailed temperature measurement. The test pattern is presented in Fig. 14. The main objective of this test was to conduct the detailed temperature measurement at the design condition. Therefore, the test was carried out in short dwell time and fewer temperature steps at off-design conditions.

After performing the steady-state tests, the emergency shutdown test was conducted at 6 atm pressure and 1500°C temperature conditions.

Figure 15 shows the results of temperature measurement of the air-cooled ceramic vane (No. 2 vane) under 6 atm and 1500°C conditions. The ceramic temperatures were measured at outer shroud and two different radial locations (50 and 95 percent vane heights) at the leading edge portion of the airfoil section. As seen in this figure, the ceramic temperatures are maintained below 1300°C as intended in the design.

Figure 16 shows the comparison of measured temperatures between air-cooled type (No. 2 vane) and uncooled type (No. 3 vane). The effect of the cooling by a small amount of air is clearly demonstrated in this figure. Thus, it is confirmed that the ceramic material temperature can be maintained below 1300°C even if the gas temperature is 1500°C by utilizing a small amount of cooling air.

Figure 17 presents the results of measurement for ceramic internal temperatures at the time of emergency shutdown

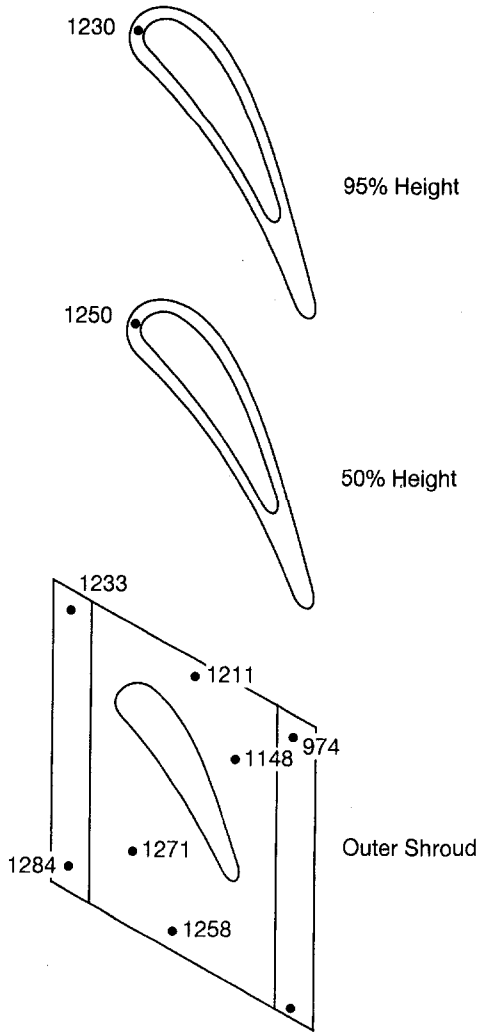


Fig. 15 Measured ceramic temperature (T_g = 1500°C)

(abrupt cutoff of fuel). The gas temperature at 50 percent vane height was abruptly dropped from 1550°C to nearly 700°C at 2 seconds after emergency shutdown and to air temperature of 400°C afterward. The ceramic airfoil was cooled down most quickly at midspan section of leading edge portion. The thick portions of the outer shroud were cooled down moderately. Thus, a large temperature difference was observed between the leading edge portion of airfoil midspan section and the outer shroud, which resulted in the generation of thermal stresses. The maximum temperature difference in measured points is about 400°C. This phenomenon coincides with the prediction by thermal stress analysis described in the stress analysis section.

After the tests, each vane was disassembled and inspected by the fluorescent penetrant inspection. Although no cracks were found for the air-cooled Si_3N_4 ceramic nozzle vanes, the uncooled SiC ceramic nozzle vanes experienced cracks. This is partly because the thermal stress generated on SiC vane is higher by 30–40 percent compared with the Si_3N_4 vane due to the difference in material properties, such as Young's modulus, thermal expansion, and thermal conductivity. In order to apply SiC for the nozzle vanes, it seems to need further efforts to understand the actual operational conditions and to improve the design and application technologies.

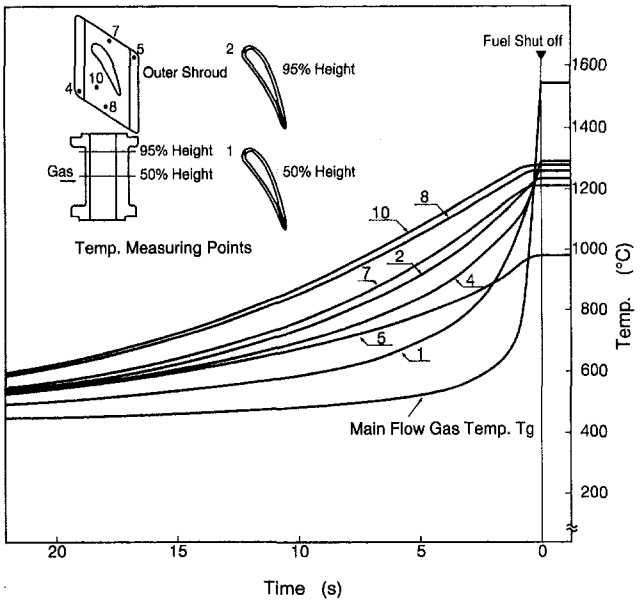


Fig. 17 Temperature change after emergency shutdown

Summary

The development of air-cooled ceramic nozzle vanes for a power generating gas turbine has been reported. To make up the limited temperature resistance of present ceramic materials, the utilization of a small amount of cooling air has been studied for the first-stage nozzle vanes of a 1500°C class gas turbine. The inner face of the ceramic vane was cooled by air impingement and convection, and this cooling air was discharged and mixed into the main gas flow through cooling slits located at trailing edge of the nozzle.

A series of cascade tests were carried out for the designed air-cooled Si_3N_4 nozzle vanes under 6 atm and 1500°C conditions. It was confirmed that the maximum ceramic temperature can be maintained below 1300°C by a small amount of cooling air. In spite of the increased thermal stresses by local cooling, all Si_3N_4 nozzle vanes survived in the cascade tests including both steady-state and transients of emergency shutdowns and the possibility of air cooled ceramic nozzle was demonstrated for 1500°C class gas turbine application.

It is planned to verify the validity of the air-cooled ceramic nozzle vanes in a forthcoming test under full-pressure (15 atm) and full-temperature (1500°C) conditions. It is also necessary to confirm the long-term durability and reliability of ceramics under actual operational conditions.

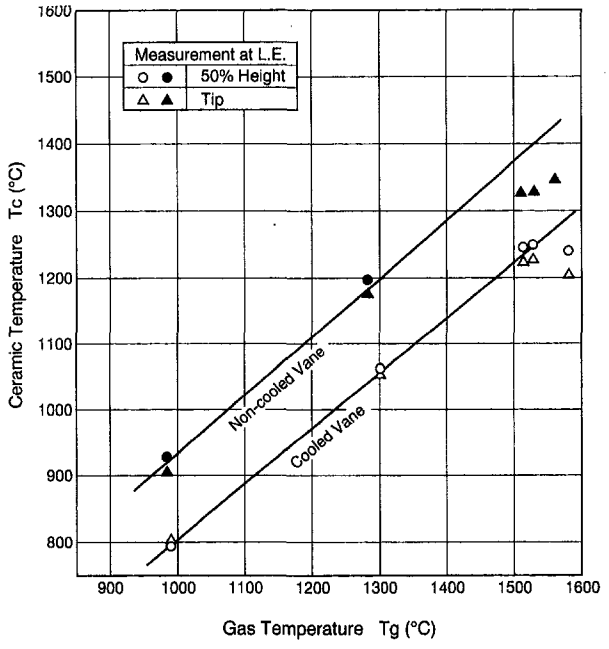


Fig. 16 Comparison of ceramic temperature for cooled and uncooled vanes

Acknowledgments

The authors wish to thank heartily the persons concerned from NGK Insulators, Ltd. for their kind assistance in manufacturing the ceramic vanes for this study.

References

1 Hara, Y., et al., *JSME Reports,* No. 870–10, Nov. 1987.
2 Tsuchiya, T., *Journal of the Gas Turbine Society of Japan,* Vol. 18, No. 69, June 1990.
3 Hara, Y., FC Reports, Vol. 8 (10), 1990.
4 Hara, Y., et al., *JSME Reports,* No. 870–60, Nov. 1989.
5 Tsuchiya, T., et al., *Proceedings of the 18th Regular Meeting for the Gas Turbine Society of Japan,* June 1990.
6 Ohkoshi, A., et al., *Proceedings of the 1987 TOKYO International Gas Turbine Congress,* 87-TOKYO-IGTC-121, Oct. 1987.
7 Hara, Y., et al., *Proceedings of the 18th Regular Meeting for the Gas Turbine Society of Japan,* June 1990.
8 Furuse, Y., et al., *Proceedings of the 18th Regular Meeting for the Gas Turbine Society of Japan,* June 1989.
9 Tsuji, I., et al., ASME Paper No. 91-GT-311, 1991.
10 Hara, Y., et al., ASME Paper No. 91-GT-319, 1991.
11 Maeda, F., et al., *Proceedings of the 1991 YOKOHAMA International Gas Turbine Congress,* 91-YOKOHAMA-IGTC-44, 1991.
12 Miyazaki, S., et al., *Proceedings of the 1991 YOKOHAMA International Gas Turbine Congress,* 91-YOKOHAMA-IGTC-101, 1991.
13 Machida, T., et al., *Proceedings of the 1991 YOKOHAMA International Gas Turbine Congress,* 91-YOKOHAMA-IGTC-103, 1991.
14 Furuse, Y., IUMRS-ICAM-93, Sept. 1993.

Operation of Gas Turbine Engines in Volcanic Ash Clouds

M. G. Dunn

A. J. Baran

Calspan Corporation,
Buffalo, NY 14225

J. Miatech

Headquarters, Defense Nuclear Agency,
Alexandria, VA

Results are reported for a technology program designed to determine the behavior of gas turbine engines when operating in particle-laden clouds. There are several ways that such clouds may be created, i.e., explosive volcanic eruption, sand storm, military conflict, etc. The response of several different engines, among them the Pratt & Whitney JT3D turbofan, the Pratt & Whitney J57 turbojet, a Pratt & Whitney engine of the JT9 vintage, and an engine of the General Electric CF6 vintage has been determined. The particular damage mode that will be dominant when an engine experiences a dust cloud depends upon the particular engine (the turbine inlet temperature at which the engine is operating when it encounters the dust cloud), the concentration of foreign material in the cloud, and the constituents of the foreign material (the respective melting temperature of the various constituents). Further, the rate at which engine damage will occur depends upon all of the factors given above, and the damage is cumulative with continued exposure. An important part of the Calspan effort has been to identify environmental warning signs and to determine which of the engine parameters available for monitoring by the flight crew can provide an early indication of impending difficulty. On the basis of current knowledge, if one knows the location of a particle-laden cloud, then that region should be avoided. However, if the cloud location is unknown, which is generally the case, then it is important to know how to recognize when an encounter has occurred and to understand how to operate safely, which is another part of the Calspan effort.

Introduction

The reality of the danger posed by volcanic clouds to commercial aircraft operation has been dramatically illustrated during the past ten years. For the past four or five years, the airframe and engine companies have become involved in trying to understand the nature of the problems associated with aircraft interactions with volcanic clouds. Campbell (1990) presented an in-depth review of the problem along with suggested operational procedures at a Boeing 747 Flight Operations Symposium. As part of this review, he described four 747 aircraft encounters with volcanic ash clouds, beginning with the June, 1982, encounter of a British Airways aircraft near Bandung, Indonesia, and concluding with an encounter in December, 1989, of a KLM aircraft with the Mt. Redoubt cloud near Anchorage, Alaska. More recently, in June of 1991, there were numerous (14 or more) encounters between commercial aircraft and various Mt. Pinatubo volcanic clouds. In these encounters, the aircraft and the propulsion system both can sustain significant damage. However, the discussion presented herein will be confined to the power plant.

The adverse influence of the dust cloud has been observed to occur in most modern power plants, e.g., the Rolls Royce RB211, the Pratt/Whitney JT9D, and the General Electric CF6. Measurement programs designed to ascertain the response of gas turbine engines when operated in dust-laden environments provide the fundamental data necessary to develop a predictive capability in order to recommend operating procedures. The specific type of dust of interest in the experiments described herein is material that becomes airborne as a result of an explosive volcanic eruption.

Gas turbine engines are routinely tested for the effect of ingestion of solid particles according to the procedures of Military Specification MIL-E-5007D, commonly known as the "Arizona road-dust test." During the last several years the material used in this test was changed from a mixed silicate mineralogy to crushed quartz. Encounters with volcanic clouds are felt to be much different than those experienced with MIL-E-5007D. The principal differences are material chemical composition and particle concentration in the cloud.

A technology program ongoing at Calspan Advanced Technology Center for many years in the area of gas turbine propulsion research utilizes a unique facility and experimental technique. In these experiments, operating gas turbine engines are exposed to an adverse environment like that associated with a volcanic ash cloud without endangering either an airplane or a flight crew. The response of several different engines, including the Pratt & Whitney JT3D turbofan, the Pratt & Whitney J57 turbojet, an engine of the Pratt & Whitney JT9 vintage, and a General Electric engine of the CF6 vintage to typical volcanic ash clouds has been determined.

The behavior of the individual components that make up a volcanic cloud are studied separately using one of two combustor rigs: (1) an Allison T56 combustor can and a row of T56 high-pressure turbine inlet vanes or (2) a Pratt & Whitney F100 annular combustor and sector of high-pressure turbine vanes (Kim et al., 1993). The combustors just described are equipped with a dust injection system, an external air compression system, and a heat exchanger that is used to increase the temperature of the compressed air prior to introduction into the combustor. Data obtained using these test sections provide guidance to the researcher prior to initiation of full-scale engine measurements. A reasonable estimate of the behavior of individual cloud materials and the threshold temperature at which some components of the material might melt and deposit on the vanes can be determined with these test systems.

The damage mode that will be dominant for a particular encounter depends upon the particular engine (the turbine inlet temperature at which the engine is operating at the time of encounter), the constituents of the cloud, and the respective melting temperature of the cloud constituents. Further, the rate at which engine damage will occur depends upon the items just

Contributed by the International Gas Turbine Institute and presented at the 39th International Gas Turbine and Aeroengine Congress and Exposition, The Hague, The Netherlands, June 13–16, 1994. Manuscript received by the International Gas Turbine Institute February 15, 1994. Paper No. 94-GT-170. Associate Technical Editor: E. M. Greitzer.

mentioned and upon particle density (or concentration) of the cloud.

Several different damage modes may be responsible for altering engine performance during or after exposure to a particle-laden cloud and the potential of each mode should be recognized. The dust cloud represents an unanticipated engine degradation and may be manifested in one or more of the following ways: (*a*) deposition of material on hot-section components, (*b*) erosion of compressor blading and rotor path components, (*c*) partial or total blockage of cooling passages, (*d*) engine control system degradation, and (*e*) oil system, bleed air supply, or fuel contamination. In addition to the power plant difficulties just noted, there are several aircraft problems that are also of concern, e.g., (*f*) fogging of the windscreen and landing lights, (*g*) contamination of the electronics because of dust ingested through the Environmental Control System (ECS), (*h*) erosion of antenna surfaces, and (*i*) plugging of the pitot-static system. Most of these phenomena have been experienced in the ash cloud encounters reported by Mitchell and Gilmore (1982), Chambers (1985), and Smith (1985). Compressor erosion takes the form of severe tip and blade leading edge damage and causes a loss of surge margin even if deposition does not occur in the hot section. Tabakoff and Hussein (1971a), Hussein and Tabakoff (1973), Grant and Tabakoff (1975), Tabakoff and Balan (1981), and Tabakoff and Hussein (1971b) are representative of laboratory studies of erosion conducted for an isolated compressor blade row. Fogging of the windscreen as a result of flying through particle-laden air has made landing difficult and dust contamination of the ECS has added to the millions of dollars worth of damage to aircraft.

The intent of the work described in this paper is to: (1) ascertain how a flight crew might recognize a particle-laden environment if natural light or cloud conditions preclude direct visual observation, (2) to describe the type of engine damage that should be anticipated, (3) to describe which of the diagnostics provide an indication of engine degradation and how one might operate to generate significant thrust after the engines have been severely damaged. Early results of the Calspan research were reported by Dunn et al. (1987a), Batcho et al. (1987), and Dunn et al. (1987b).

Experimental Apparatus

This type of research requires a unique facility, one that allows large engines to be subjected to realistic environmental conditions under controlled laboratory conditions. The engine test cell at Calspan has been built within a large building, permitting tests unhampered by weather. The test setup includes the engine test stand, operating equipment, and instrumentation. Two major features of the experimental apparatus are a noise suppression system and a dust injection system.

The noise suppression system is an air-cooled unit consisting of a primary intake silencer, a secondary intake silencer, an engine test stand enclosure, and an exhaust silencer system (made up of an augmented tube, an exhaust plenum, and an exhaust silencer) (Fig. 1). This test cell has the capability of handling engines with an air weight flow as large as 300 lbm/s. A blanket made of triple woven $\frac{5}{8}$-in. diameter steel cable is hung around the engine to provide protection in the event of a thrown component. A Halon fire suppression system is available for use if an engine fire were to occur.

The Dust Injection System (DIS) is an important part of the test apparatus and was designed to accurately regulate the dust environment. The dust injection system must provide a controlled range of particle sizes and concentrations, all evenly dispersed in the incoming air. The DIS, shown schematically in Fig. 2, consists of a dust tube, dust injection nozzles, weigh-belt feeder and dust reservoir, and a bellmouth. Prior to initiation of the measurement program, the specific components of the dust are specified, the fraction of each component is specified by particle size, these components are mixed in an electric mixer, and a known amount of the mixture is loaded into the dust reservoir. The weigh-belt feeder that delivers the dust from the reservoir to the ejector includes process monitoring instrumentation to allow precise control of dust being fed to the test engine. The dust is delivered to the injection nozzles by compressed air. At the completion of an experiment, the material remaining in the reservoir is removed and weighed. A more detailed description of the experimental apparatus is given in Dunn et al. (1987a).

Cloud Preparation and Analysis

The engines discussed here were all operated in air laden with particulate material composed of a mixture of basically three types of soils (sand, feldspar, and clay), and a volcanic ash. The soil materials were either commercially available or were obtained from southern California. The sandy soil is composed of decomposed granite and the clayey soil is weathered material from surrounding rock outcrop units. During the measurement program, two different volcanic ash materials have been routinely used: Mt. St. Helens ash and Twin Mountain black scoria. The Mt. St. Helens ash was obtained from a location near Portland, Oregon, and the black scoria was obtained from Twin Mountain quarry near Des Moines, New Mexico.

An analysis of the scoria indicated that the basaltic glass fraction was approximately 85 percent. The remaining 15 percent of the material was metrosphere, plagioclases, and pyroxene. A similar analysis of the Mt. St. Helens ash indicated a basaltic glass fraction of approximately 25 percent. The remainder was similar to that described above.

A description of one of the cloud compositions that was passed through the engine is given in Dunn et al. (1987b). Figure 3(*a*) is a scanning electron micrograph and Fig. 3(*b*) shows the elemental composition spectrum of the black scoria. Figures 4(*a*) and 4(*b*) are a micrograph and an elemental composition, respectively, for Mt. St. Helens ash. The elemental composition of the two volcanic materials is different as can be seen by comparing the composition spectrums of Figs. 3(*b*) and 4(*b*). However, aluminum, silicon, calcium, and iron are present in each material. The SEM of the two materials, Figs. 3(*a*) and 4(*a*), illustrate the very different physical characteristics of the two materials. The scoria has the appearance of a plate whereas the Mt. St. Helens material has the more rounded appearance.

Test Procedures and Teardown Observations

Test Procedures. As noted earlier, many different engines have been utilized during the course of this testing program. The detailed test procedure differs depending upon the particular engine because the available engine instrumentation and the control systems are different. However, for all engines, a set of baseline measurements is obtained at many different throttle or power level angle (PLA) settings prior to exposing the engine to a dusty environment. All of the available engine parameters are recorded. Depending upon the particular engine, these will include: (*a*) some form of exhaust gas temperature (EGT), (*b*) fuel flow rate, (*c*) inlet total pressure (P_{T_2}), (*d*) fan exit total pressure ($P_{T_{2.5}}$), (*e*) engine pressure ratio (EPR), (*f*) interstage pressure, (*g*) compressor discharge pressure (P_{S_3}), (*h*) engine thrust, (*i*) burner pressure (P_B), (*j*) core speed (N_2), and (*k*) fan speed (N_1). During this baseline measurement series, the engine also undergoes several rapid (six second) accelerations and decelerations to obtain pre-dust exposure response data.

Observations During Testing. For all engines tested, immediately upon introduction of the dust, a striking glow is observed at the fan face (or compressor face if a turbojet). The glow is the well-known phenomenon St. Elmo's glow (or fire)

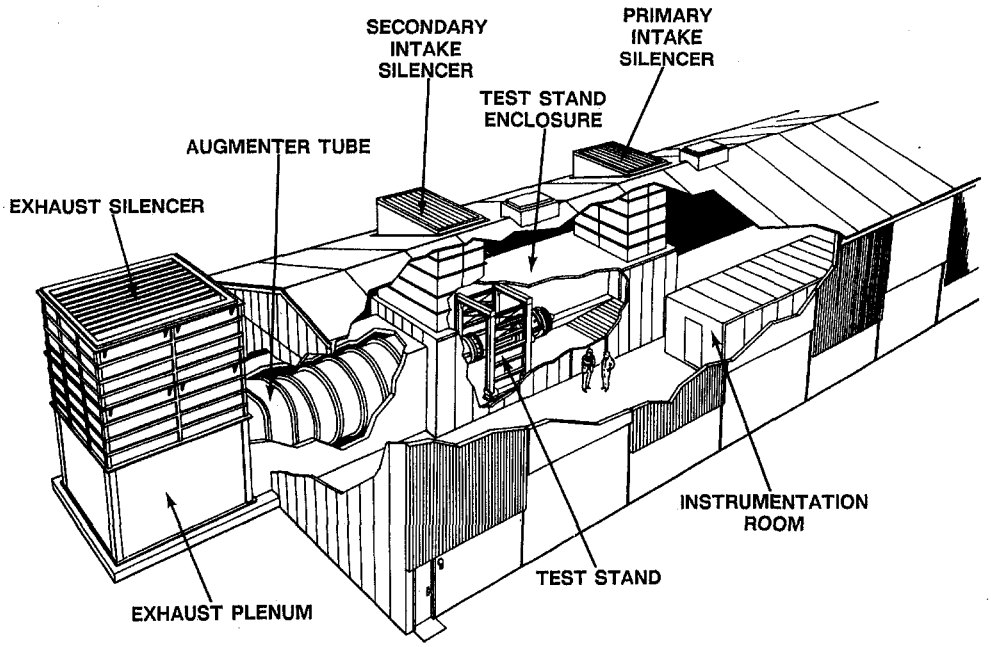


Fig. 1 Large engine research cell

that is the result of a discharge of electrostatic charge that has built up because of dust particles striking the moving metallic surfaces. St. Elmo's glow was observed by the British Airways crew during their encounter with the Galunggung volcanic cloud (Chambers, 1985). A video camera viewing the engine inlet records this glow for each of the measurement programs. Figure 5 is a frame of the video tape taken during a dust ingestion experiment showing the glow along with the bellmouth, the dust tube, and the dust injection nozzles. A very bright ring of light appears at the first rotor tip/shroud region. The illumination in the tip region may be brighter for several reasons, e.g., (*a*) higher dust particle concentrations caused by centrifuging, (*b*) greater optical depth and unobstructed view of the illumination, and (*c*) scintillation at the rotor tip. It is, however, felt that the primary reason for the increased brightness is possibility (*b*).

It has also been observed that selected engine parameters respond very quickly to the presence of a dust cloud if the engine is operating at a turbine inlet temperature (TIT) in excess of 2000°F. For example, the burner static pressure (PB) and the high compressor static discharge pressure (PS3) respond in almost identical fashion as illustrated in Fig. 6 for an engine operating at a TIT of approximately 2345°F. The pressure increase can be very rapid (recall that the CF6 engines described by Campbell, 1990, spooled down after about 1 minute of high thrust operation in the Mt. Redoubt cloud). The rate of pressure increase shown on Fig. 6 is almost directly proportional to the dust concentration. For the modest concentration used in the

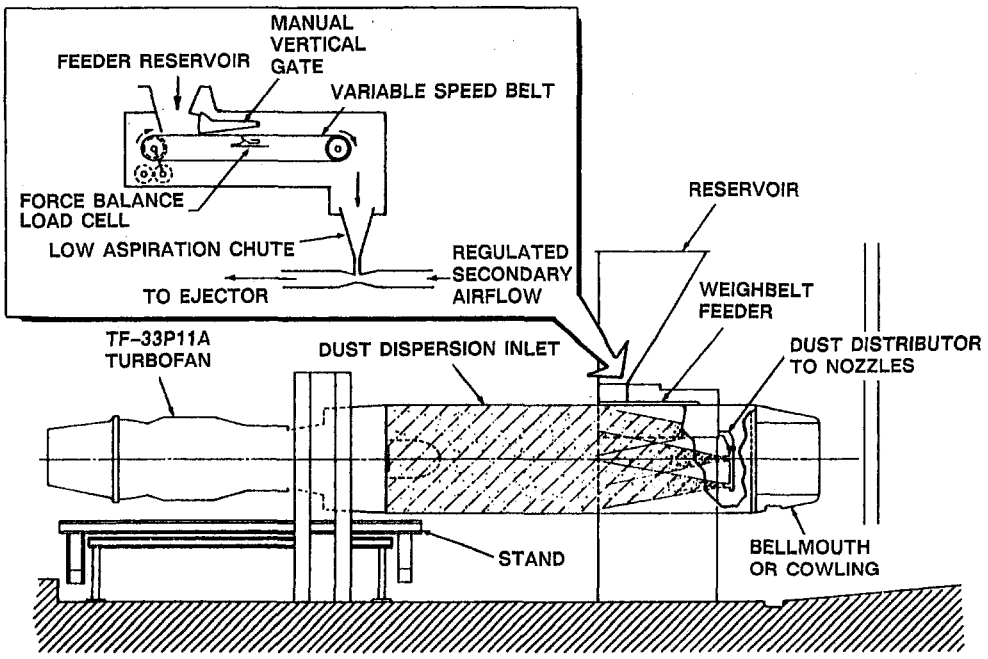
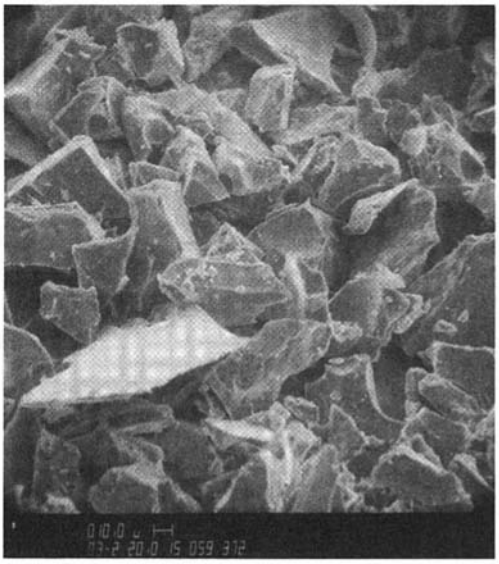
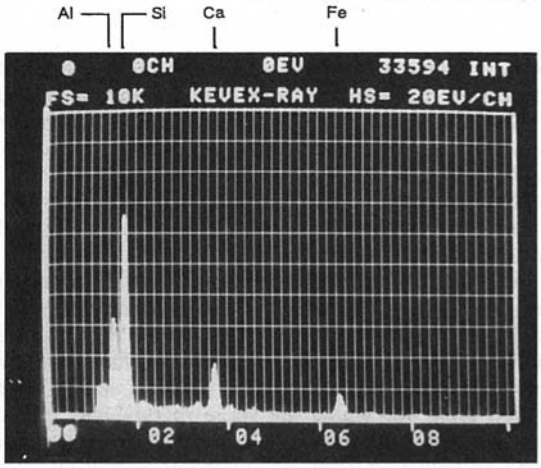


Fig. 2 Schematic of dust injection system

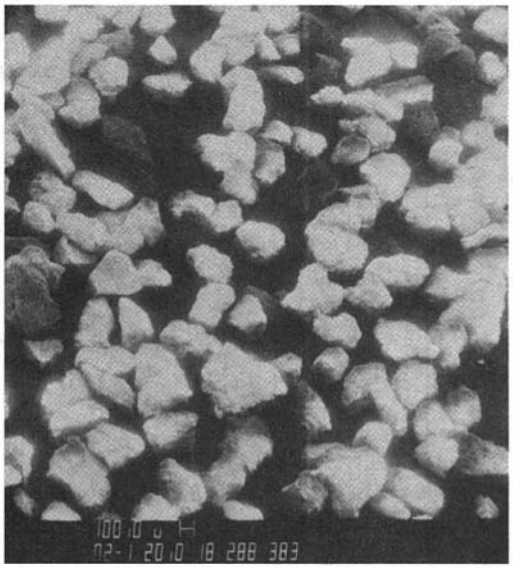


(a) Scanning Electron Micrograph of Black Scoria Sample

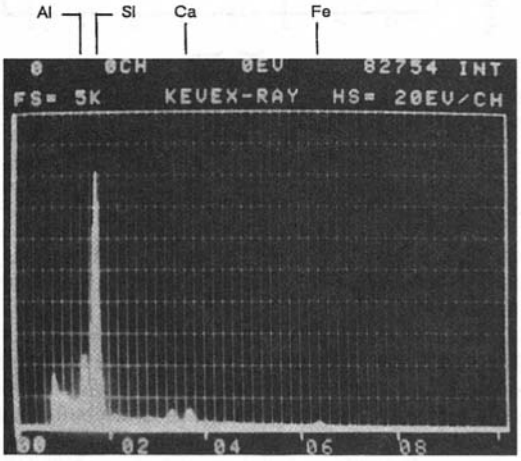


(b) Elemental Composition Spectrum

Fig. 3 Scanning electron microscope micrograph and composition spectrum of black scoria



(a) Scanning Electron Micrograph of Mt. St. Helens Ash



(b) Elemental Composition Spectrum

Fig. 4 Scanning electron microscope micrograph and composition spectrum of Mt. St. Helens ash sample (unprocessed ash)

experiments of Fig. 6, the pressure increase amounted to about 4 psi/min. It is possible to consume all of the available surge margin very quickly even in this relatively modest dust cloud. As will be demonstrated later, this pressure build-up is the result of material from the dust cloud being deposited on the high-pressure turbine first vane. Also included on Fig. 6 are the fuel flow rate and the fan total pressure ratio, both of which are influenced by the dust cloud, but not as dramatically as the burner and compressor discharge pressures. Figure 7 illustrates that the exhaust gas temperature (EGT) also responds rapidly to the presence of the dust cloud, again if the TIT is in excess of 2000°F. As was the case for pressure, the rate of increase of EGT is proportional to the cloud dust density. The engine pressure ratio (EPR) and core speed ($N2$) are also shown on Fig. 7. Changes in the core speed ($N2$) and the fan speed ($N1$) are much more subtle indicators of the material deposition or erosion and are detected through the response of the control system. An example of control system deterioration will be shown a little later in this section.

The importance of the TIT can be demonstrated by reducing the throttle so that the TIT falls below 2000°F while maintaining the dust-laden environment. Figure 8 illustrates the time history of the fuel flow, the fan total pressure ratio, the compressor

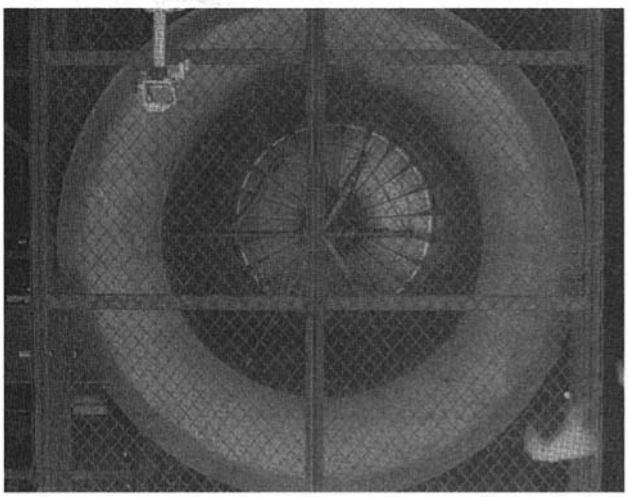


Fig. 5 Photograph of St. Elmo's glow at fan face

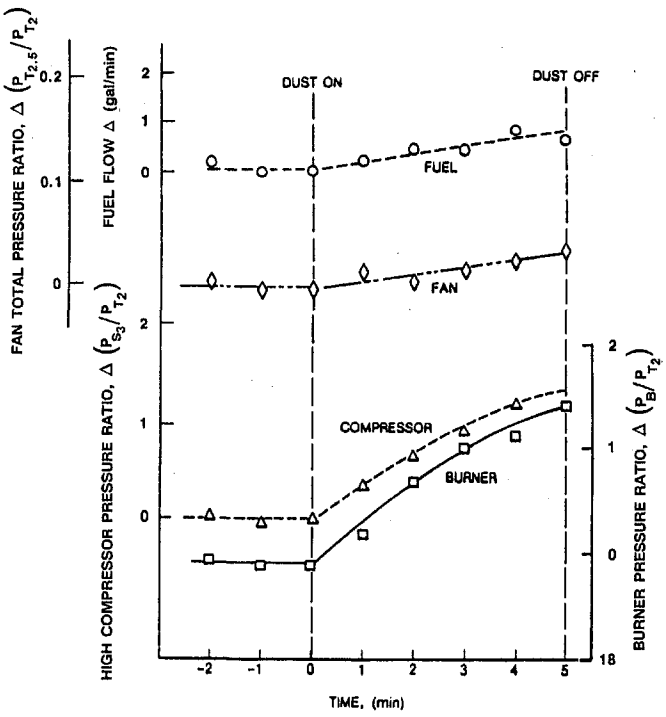


Fig. 6 Time history of engine parameters during dust experiments for relatively high turbine inlet temperature

discharge pressure and the burner pressure for a TIT of about 1850°F and the same dust concentration as used in the experiments of Figs. 6 and 7. The EGT, though not shown, also remained constant during dust exposure at this lower temperature condition.

It is obvious that if the burner pressure is allowed to increase unchecked at a rate as illustrated in Fig. 6, then the engine is going to surge and continued operation will become difficult. Perhaps the best way to prevent further pressure build-up is to retard the throttle (to maintain TIT below 2000°F), and exit the dust cloud using the technique described in Campbell (1990).

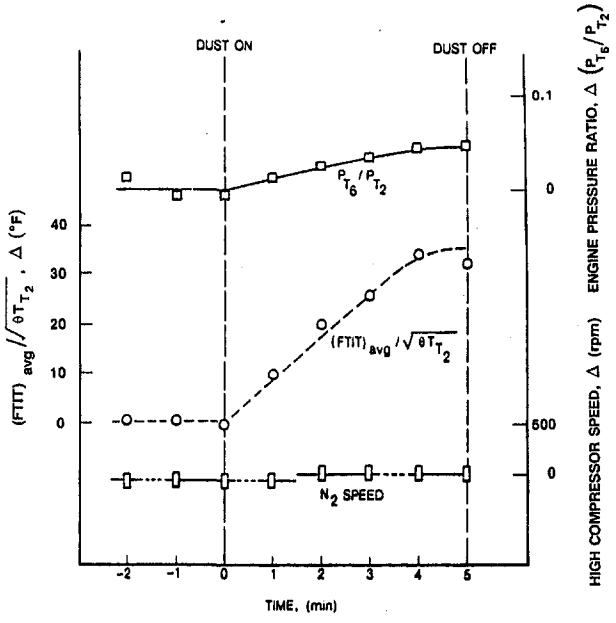


Fig. 7 Time history of engine parameters during dust experiments for relatively high turbine inlet temperature

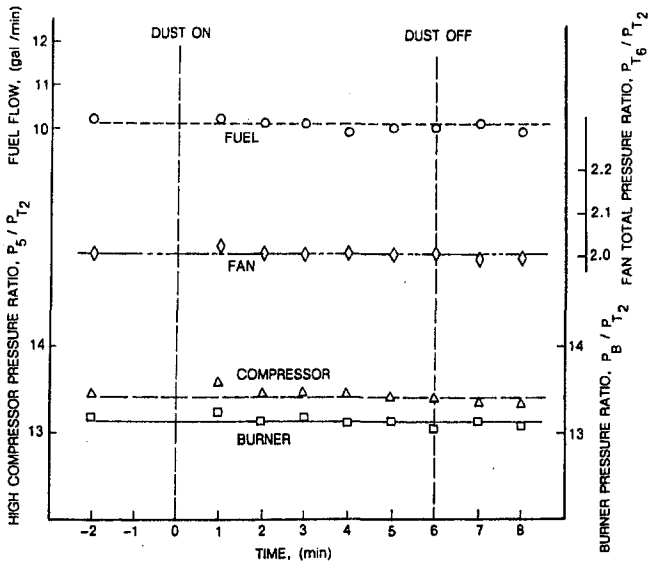


Fig. 8 Time history of engine parameters during dust experiments for relatively low turbine inlet temperature

However, if the crew is not in a position to exercise this option, but rather must maintain significant thrust, then what should be done? One alternative is to clear the material that has been deposited on the HPT vane. This can be done by slowly retarding the throttle to idle, dwelling for thirty seconds, and advancing the throttle to full thrust. This process may have to be repeated two or three times. For a multi-engine aircraft, it should be done one engine at a time.

Figure 9 is a time history of the burner pressure for an engine operating in a particle-laden environment at a relatively high turbine inlet temperature. Very soon after the dust was turned on, the burner pressure began to increase significantly. When the pressure increase reached approximately three atmospheres, a situation arose for which the surge margin was felt to be almost consumed (earlier it was noted that this pressure increase is caused by flow area blockage created as a result of material deposition on the high-pressure turbine first vane). It, therefore, became prudent to purge the engine (as described above) in an attempt to remove some of this material and thus move the burner pressure back towards normal. As can be seen on Fig. 9, the purge does allow the burner pressure to return to a pressure level that is near the predust value. The dust flow was not terminated during the purge sequence. Upon completion of the purge, the throttle was returned to the previous setting. The pressure returned to a level closer to the predust exposure value for a relatively brief period of time and then began to increase again. When the pressure increase again reached about three

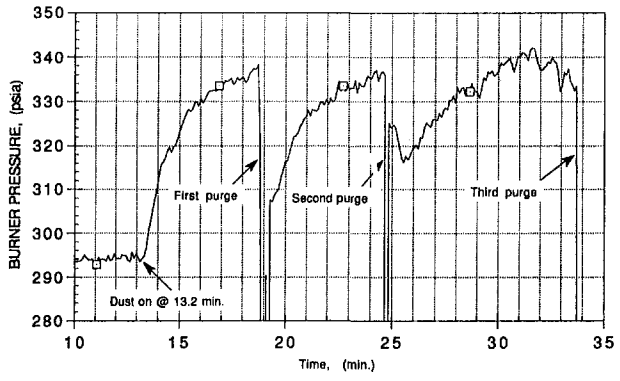


Fig. 9 History of burner pressure

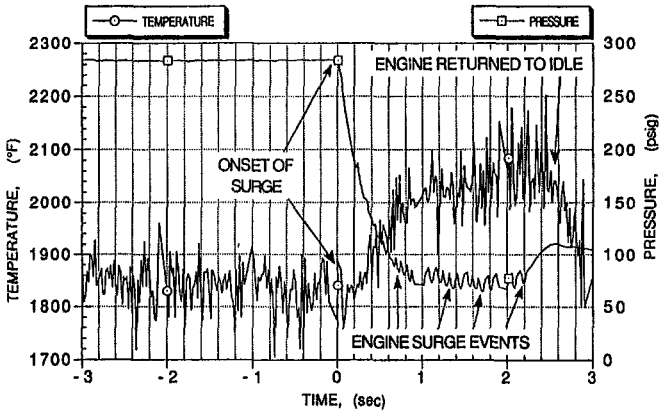


Fig. 10 Burner pressure and temperature response

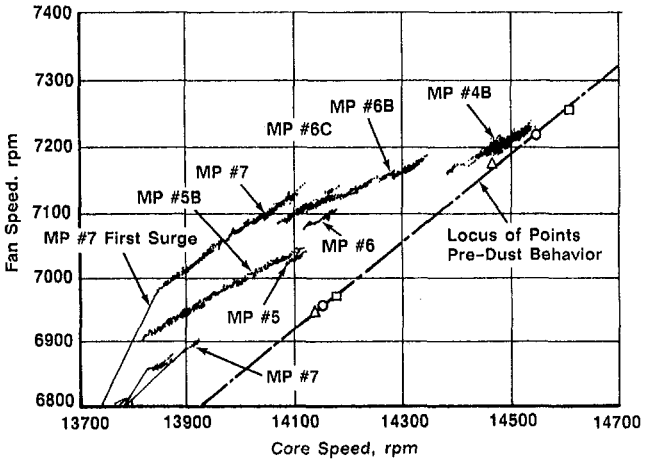


Fig. 12 Fan and core speed map during dust exposure

atmospheres, the purge sequence was repeated. As can be seen from the results presented in Fig. 9, purging the engine does help to clear the engine and extend the operational time at high thrust. However, the postpurge pressure does not decrease as much with succeeding purges. This purge process can't be continued indefinitely because the integrated damage becomes so severe that the engine eventually begins to surge in an uncontrolled manner.

The particular engine being used to obtain the results shown in Fig. 9 provides fan turbine inlet temperature (FTIT) instead of exhaust gas temperature and a time history of that parameter was also obtained. The FTIT is measured inside the vane of the fan turbine first vane for several different vanes. The qualitative behavior of the FTIT is very similar to the burner pressure from the time of dust on until shortly after the second purge. Between the second and third purges, the FTIT rate of increase is much greater than between dust-on and first purge and between first purge and second purge. This rapid increase is indicative of impending difficulty. After completion of the third purge, the engine began to surge rapidly. This surging was captured with video cameras looking at the tailpipe and at the engine face. Pictures were recorded at 26 frames per second.

Figure 10 represents the corresponding time history of the burner pressure and FTIT during the surging event described above. Time 0-seconds on Fig. 10 is the initiation of the rapid pressure decrease that is associated with the surge event. During the time interval shown on Fig. 10, multiple surges have occurred. The FTIT begins to increase rapidly during the surging sequence as might be expected. Figure 11 is a portion of the data shown in Fig. 10 that has been used to determine the surging frequency. For this particular engine, that frequency was approximately 10 Hz. It is also interesting to note that upon initiation of the surge, the pressure drops from about 285 psig to about 74 psig and remains at this lower pressure until the engine can be recovered. This recovery is achieved by returning the throttle to idle as noted on Fig. 10.

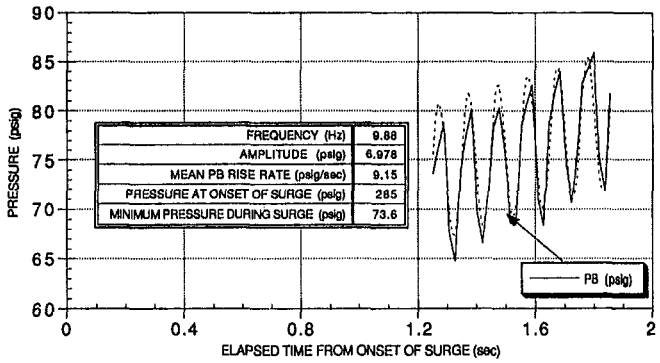


Fig. 11 Modeling of burner pressure during engine surge

Figure 12 is a fan versus core speed map for one of the engines undergoing dust exposure. This plot reflects the integrated influence of the dust environment on the engine control system. The dashed line and the points along it represent the predust exposure behavior of the engine. The data denoted by MP4B (matrix point 4B), MP5, MP5B, etc., represent data taken during dust exposure. Data from matrix point 1–3 are not included on this figure. The initial engine surge occurred during MP7 as illustrated. By the time this surge occurred, a significant mismatch between fan and core speed had been experienced. However, this mismatch would not be as obvious to the flight crew as the increases in burner pressure, CDP, and FTIT shown earlier. Unfortunately, neither burner pressure nor CDP are generally available to the flight crew. These parameters are available in the engine control box and could be made available to the crew.

Engine Teardown Results. All of the engines used for the measurement program are disassembled for inspection when they can no longer be operated. Both in-flight experience and laboratory experience have demonstrated that if one continues to operate a modern gas turbine engine at high thrust levels in a particle-laden environment, then it is likely that the engine will surge and possibly flame out if at high altitude. The time over which one can operate before surge depends upon many factors. Among them are: (*a*) the particular engine involved, (*b*) the concentration of the dust cloud, and (*c*) the constituents of the dust cloud. As noted earlier, the five most dominant factors of immediate concern to the crew are material deposition occurring on the high-pressure turbine vanes, blocking of turbine vane and/or blade cooling holes, erosion of the fan and compressor blades, degradation of the engine control system, and deposition of carbonlike material on the fuel nozzles. Fogging of the windscreen and landing lights, contamination of the environmental control system (ECS), and contamination of the oil supply can also pose serious problems to the flight crew. The particulates gouge and scratch the windscreen. One can easily imagine how difficult it might be to land an aircraft while looking through a windscreen that has the appearance of a shower door. The material ingested in the ECS and the oil supply does not pose an immediate threat to safe operation for commercial aircraft engines.

Accumulation of material on the high-pressure turbine vanes is responsible for the very rapid increase in burner pressure and in compressor discharge static pressure shown in Fig. 6. Figure 13 is a photograph of a portion of the high-pressure turbine vane row taken from an engine that was subjected to a dust

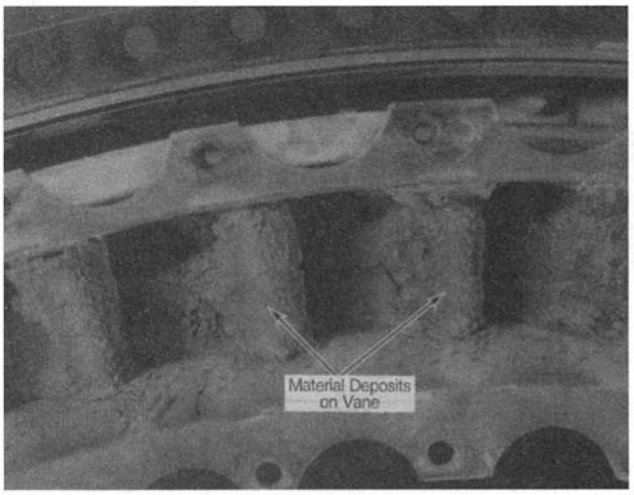


Fig. 13 Material deposits on HP turbine vanes

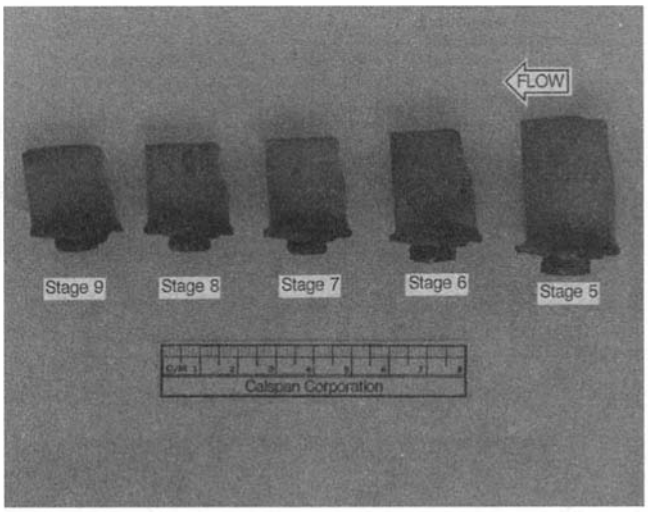


Fig. 15 Compressor erosion from dust exposure

environment. The engine shown in Fig. 13 was exposed to an environment containing the volcanic material black scoria. A similar engine was exposed to a dust environment that did not contain a volcanic material and the results were very similar. The heaviest deposits are generally found on the leading edge and on the pressure surface of the vane. Figure 13 illustrates that the cooling holes over the entire shower head region are blocked, but that those near the trailing edge on the pressure surfaces are still open.

It has been observed that the deposited material is relatively brittle. The thermal properties of the material appear to be very different from those of the metal airfoils. It is felt that one of the reasons that it has been possible to restart engines that have shut down at altitude after encountering a volcanic cloud is that a significant portion of the deposited material can be blown off the surface by ram air after the engine cools down. The same reasoning is felt to explain why the engine purge technique described earlier can remove substantial material.

Turbine damage is not limited to the HP vane. The ingested dust also gets into the turbine cooling passages and can clog the cooling holes of the HPT blade from the inside. Because the temperature is lower, the local velocity is higher, and the local flow disturbances are greater, deposition does not generally occur on the HPT blades to nearly the extent that it does on the vanes. However, this is not meant to imply that the blades are free of damage. Figure 14 is a photograph of several HPT blades illustrating that cooling hole blockage can result in overheating and subsequent burn out at the leading edge. This damage shows up as an increase in the aft turbine vibration level and as a loss in turbine performance.

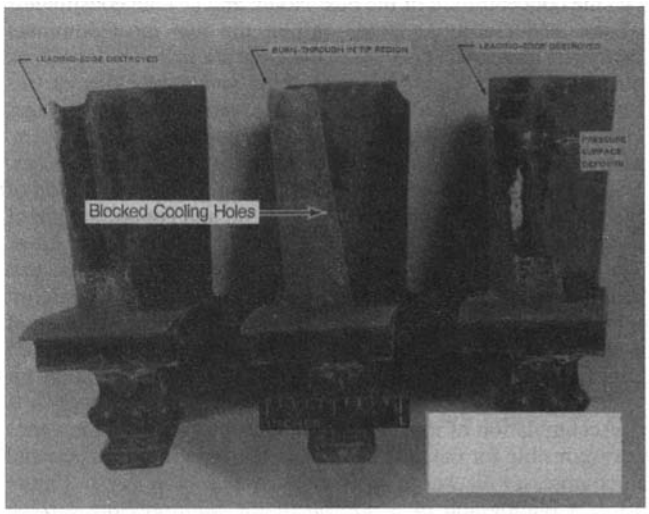


Fig. 14 Post-test photograph of HP turbine blades

Depending upon the physical characteristics of a particular fuel nozzle, it has been observed that the entire surface of the nozzle, with the exception of the fuel hole, can be covered by a thick carbon layer. It is felt that these heavy carbon deposits on the fuel nozzle swirl vanes are caused by a rich fuel-to-oxygen ratio, which, in turn, is brought about because of the severe compression system erosion to be discussed later. Carbon deposits occurring in this manner block the swirl vanes, thus inhibiting atomization of the fuel. If carbon deposits of this type do occur, then it is extremely difficult (if not impossible) to restart the engine. This carbon material cannot be blown off by ram air and there is no way to clean the fuel nozzles while in flight.

In addition to the difficulties caused in the hot section of the engine as a result of exposure to the volcanic cloud, the compression system can also sustain severe damage. Figure 15 is a photograph of typical blades taken from the HP compressor, stage 5–9 of one of the machines described above. The damage to the blade from about 50 percent span to the tip and in the tip region is very obvious. The damage becomes more severe with increasing compressor stage as might be expected. In general, the blades are badly damaged in the leading edge and tip region. In some cases, the blades break near the root. Several blades were broken in this manner and resulted in severe damage to other blades not shown in this figure. Compressor wear is an irreversible process and once the engine has operated in a particle-laden environment for any significant time period the blades must be replaced.

The engine damage illustrated in Figs. 13–15 is severe and can make operation difficult. However, laboratory experiments make it possible to determine ways in which damaged engines can be operated successfully long after severe damage has been incurred. For example, the engine from which the photographs shown in Figs. 13–15 were taken was still operational and able to generate some thrust when the test was terminated.

The engine response data described herein has been formatted into algorithm format for each of the important engine parameters. These algorithms can then be combined with an environment predictive code in order to predict what might happen to the propulsion system during specified in-route events.

Conclusions

As a result of information obtained during the course of measurement programs utilizing many different gas turbine engines, the following conclusions can be drawn:

- The presence of St. Elmo's glow at the engine face is indicative of a particle-laden environment.
- The manner in which the engine will behave in a particle-laden environment is dependent upon the turbine inlet temperature, the particulate concentration, the particulate chemical constituents, and the engine control system.
- The turbine inlet temperature required to cause material deposition on the hot section components is on the order of 2000°F (1094°C). Many of the older engines that operate at lower temperature will not experience deposition, but will experience compression system erosion. The newer engines will experience both deposition and erosion. If the thrust level is reduced in order to lower the turbine inlet temperature, then the material deposition rate can be significantly reduced.
- It must be emphasized that engine operation time in a particle-laden environment is limited and if possible entry into the environment should be avoided. If entrance is unavoidable, then the air crew would be advised to reduce thrust, exit the cloud, carefully monitor the EPR, EGT, fuel flow, core speed, and fan speed. This combination of parameters will help determine the degree of engine damage.
- Prolonged operation in the dust-laden environment may result in permanent engine and aircraft damage. If the aircraft is operating at high altitude, the surging associated with this type of damage may cause flame out. If the fuel nozzle swirl vanes have not become covered with carbon, then it should be possible to restart the engine at lower altitude. Some of the deposited material will have blown off during the period of time that the engine was inoperable. If the swirl vanes have been covered with carbon, then restart is unlikely.

References

Batcho, P. F., Moller, J. C., Padova, C., and Dunn, M. G., 1987, "Interpretation of Gas Turbine Response Due to Dust Ingestion," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 109, pp. 344–352.

Campbell, E. E., 1990, "Volcanic Ash," *Proc. 1990 Flight Operations Symposium,* pp. 2.3.1–2.3.34.

Chambers, J. C., 1985, "The 1982 Encounter of British Airways 747 With the Mt. Galuggung Eruption Cloud," AIAA Paper No. 85-0097.

Dunn, M. G., Padova, C., Moller, J. E., and Adams, R. E., 1987a, "Performance Deterioration of a Turbofan and a Turbojet Engine Upon Exposure to a Dust Environment," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 109, pp. 336–343.

Dunn, M. G., Padova, C., and Adams, R. M., 1987b, "Operation of Gas Turbine Engines in Dust-Laden Environments," AGARD-Advanced Technology of Aero Engine Components, Paris, France.

Grant, G., and Tabakoff, W., 1975, "Erosion Prediction of Turbomachinery Resulting From Environmental Solid Particles," *Journal of Aircraft,* pp. 471–478.

Hussein, M. F., and Tabakoff, W., 1973, "Dynamic Behavior of Solid Particles Suspended by Polluted Flow in a Turbine Stage," *Journal of Aircraft,* pp. 434–440.

Kim, J., Dunn, M. G., Baran, A. J., Wade, D. P., and Tremba, E. L., 1993, "Deposition of Volcanic Materials in the Hot Sections of Two Gas Turbine Engines," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 115, pp. 641–651.

Mitchell, H. J., and Gilmore, F. R., 1982, "Dust-Cloud Effects on Aircraft Engines—Emerging Issues and New Damage Mechanisms," RDA-TR-120012-001.

Smith, W. S., 1985, "International Efforts to Avoid Volcanic Ash Clouds," AIAA Paper No. 85-0101.

Tabakoff, W., and Balan, C., 1981, "Effects of Solid Particles Suspended in Fluid Flow Through an Axial Flow Compressor Stage," presented at the 5th International Symposium on Airbreathing Engines, Bangalore, India.

Tabakoff, W., and Hussein, M. F., 1971a, "Effects of Suspended Solid Particles on the Properties in Cascade Flow," *AIAA Journal,* pp. 1514–1519.

Tabakoff, W., and Hussein, M. F., 1971b, "Pressure Distribution on Blades in Cascade Nozzle for Particulate Flow," *Journal of Aircraft,* pp. 736–738.

New Technology Trends for Improved IGCC System Performance

A. K. Anand

C. S. Cook

J. C. Corman

GE Power Generation,
Schenectady, NY 12345

A. R. Smith
Air Products and Chemicals, Inc.,
Allentown, PA 18195

The application of gas turbine technology to IGCC systems requires careful consideration of the degree and type of integration used during the system design phase. Although gas turbines provide the primary output and efficiency gains for IGCC systems, as compared with conventional coal-fired power generation systems, they are commercially available only in specific size ranges. Therefore, it is up to the IGCC system designer to optimize the IGCC power plant within the required output, efficiency, and site conditions by selecting the system configuration carefully, particularly for air separation unit (ASU) integration incorporated with oxygen blown gasification systems. An IGCC system, based on a generic, entrained flow, oxygen blown gasification system and a GE STAG 109FA combined cycle has been evaluated with varying degrees of ASU integration, two fuel equivalent heating values and two gas turbine firing temperatures to provide net plant output and efficiency results. The data presented illustrate the system flexibility afforded by variation of ASU integration and the potential performance gains available through the continued use of gas turbine advances. Emphasis is placed on system design choices that favor either low initial investment cost or low operating cost for a given IGCC system output.

Introduction

The commercialization of Integrated Gasification Combined Cycle (IGCC) power generation systems depends equally on component technology advances and on site and fuel-specific system integration. The gas turbine dictates the primary efficiency and size of the IGCC system. Benefits from improvements in gas turbine performance translate directly to the system when an appropriate IGCC system optimization has been performed. As IGCC power generation systems move to large, commercial-scale demonstration plants, many system and component technology developments are being scheduled for introduction. The basis for these developments is a continuing systems analysis and product improvement program. Significant insight has been gained for oxygen blown IGCC systems with respect to air separation unit integration with the gas turbine's compressor in terms of net plant output, system flexibility and overall IGCC system efficiency. Advances in gas cleaning systems, both wet and dry, have allowed increased use of low-temperature energy and preservation of syngas sensible heat in the fuel gas delivered to the gas turbine. The gas turbine firing temperature, which ultimately governs the system's efficiency potential, has been steadily increased to the current 1300°C class and can be expected to increase further, to as high as 1450°C or higher. In addition, the trend in gas turbine design is toward larger sizes and this will lead to lower capital cost per unit of electrical output due to economy of scale.

The performance analyses presented here have been based on the generic IGCC system configuration shown in Fig. 1. The results have been based on a system utilizing a GE model MS9001FA gas turbine with appropriate steam turbine generator configurations, and an Air Products and Chemicals, Inc. elevated pressure (EP) cryogenic air separation unit developed specifically for IGCC applications. The intent is to demonstrate the effect on output and efficiency of some of the primary system variables such as air separation unit integration, fuel equivalent heating value, ambient temperature range of operation required, and gas turbine firing temperature. The analysis has been performed for a range of these parameters for specific system design points for each configuration. However, a fixed air separation unit design, including the auxiliary air compressor, is required for a site-specific IGCC power plant. The final optimization procedure for the overall IGCC system is an iterative process, which factors in the projected operating ambient temperature range, as well as the ambient temperature design point, in the selection of the degree of air separation unit integration and the resulting auxiliary compressor size.

Design Point IGCC System Performance Evaluations

The IGCC system configuration shown in Fig. 1 has been evaluated for two syngas equivalent heating values [120 and 150 Btu/scf (4731 and 5913 kJ/nm^3)], three levels of air separation unit air supply integration [0, 50, and 100 percent (percentage of air separation unit air requirement supplied by the gas turbine's compressor)], and a wide range of ambient temperatures, picked to include the maximum operating ambient temperature and maximum ambient temperature at which the maximum gas turbine output can be obtained. The latter ambient temperature is termed the "break point" for gas turbine and IGCC system output and is specific to each fuel composition and degree of air extraction for a given gas turbine. In addition, system performance was evaluated for the case of zero air extraction at a gas turbine firing temperature of 1400°C, for both fuel equivalent heating values, to examine the impact of future gas turbine development. The specific performance calculations were performed utilizing the GE MS9001FA gas turbine in its single shaft, combined cycle configuration (STAG 109FA) shown in Fig. 2 and at maximum firing temperature and flow to maximize overall IGCC system efficiency. The methodology for the system analysis is shown in Fig. 3 and has been described previously by Cook et al. (1995). The dry, clean syngas composition used is shown in Table 1. The final equivalent heating values used have been obtained by first adding nitrogen from the air separation process to the combustor to provide the 150 Btu/scf (5913 kJ/nm^3) equivalent heating value and then add-

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute March 2, 1995. Paper No. 95-GT-227. Associate Technical Editor: C. J. Russo.

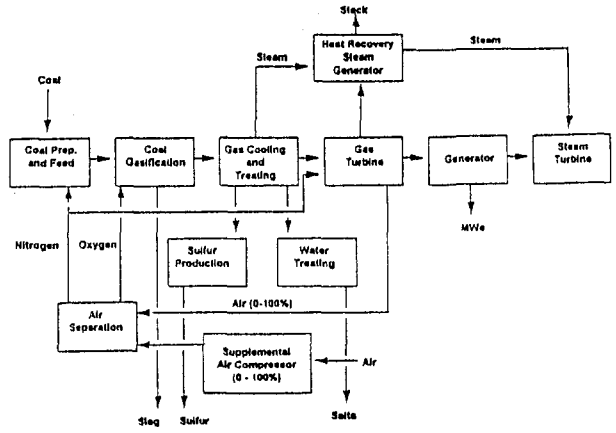


Fig. 1 Overall block diagram showing variable integrated air separation IGCC system

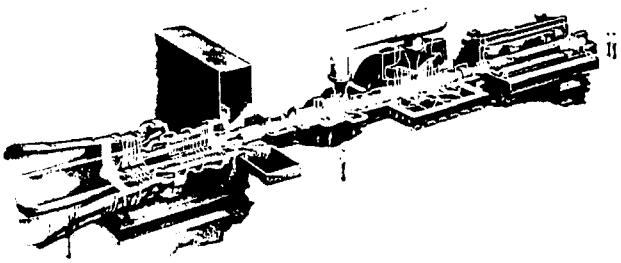


Fig. 2 General electric stag 109FA single shaft combined cycle power generation system

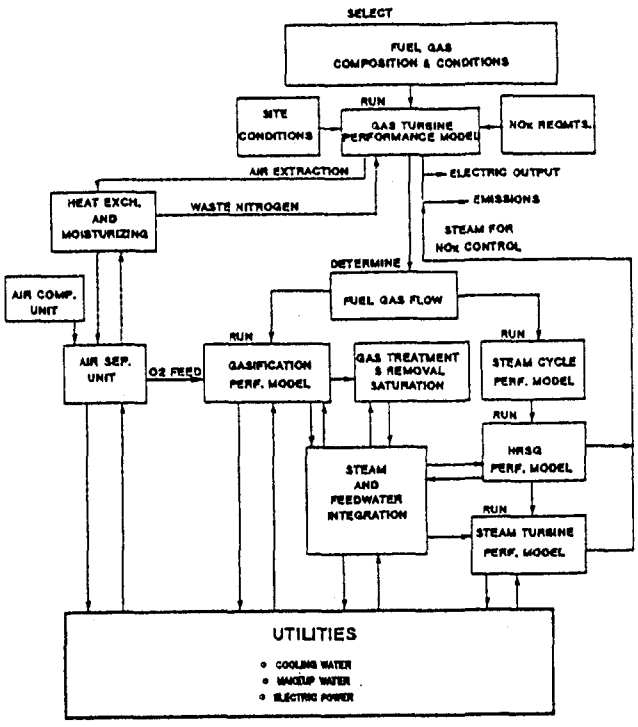


Fig. 3 GE IGCC system study flow diagram

ing moisture to the nitrogen to lower the equivalent heating value to 120 Btu/scf (4731 kJ/nm^3). Nitrogen, or moisturized nitrogen, is injected separately at a pressure somewhat lower than required for the syngas, similar to standard steam injection practice for gas turbines. The nitrogen requirement was within

Table 1 Clean syngas composition

Constituent	Volume %
CO	63.0000
CO2	1.0000
H2	30.0000
CH4	0.0500
A	0.9480
H2S	0.0007
COS	0.0013

the nitrogen flow rate that would be available from the air separation unit and the moisturized nitrogen was sufficiently superheated to prevent condensation prior to admission to the combustor.

Gas Turbine Performance. In order to illustrate the effects on IGCC system design of variation of fuel equivalent heating value and air separation unit integration, gas turbine performance was calculated for a 1260°C firing temperature for both fuel equivalent heating values and the three levels of ASU air supply integration. The results of these calculations are shown as the percent of maximum power-limited gas turbine output versus ambient temperature in Fig. 4. The 100 percent output points correspond to the maximum output from the gas turbine being considered, this maximum output being mechanically limited. The ambient temperature location of these points increases with decreasing gas turbine compressor air extraction and with decreasing fuel equivalent heating value. This feature offers the system designer wide flexibility in optimizing system output and efficiency for a given syngas fuel and desired ambient temperature operating range. An additional data point (at 35°C) is shown for the case of 120 Btu/scf equivalent heating value and 0 percent air extraction since, for this case, the gas turbine output "break point" occurs at greater than 35°C. In general, all of the individual performance curves shown represent fully open inlet guide vane operation and can be extended to lower ambient temperature at the 100 percent output level. The low ambient temperature limit to the range for each curve is determined by full closure of inlet guide vanes, full utilization of compressor discharge inlet air bleed heating, and ultimate reduction in firing temperature.

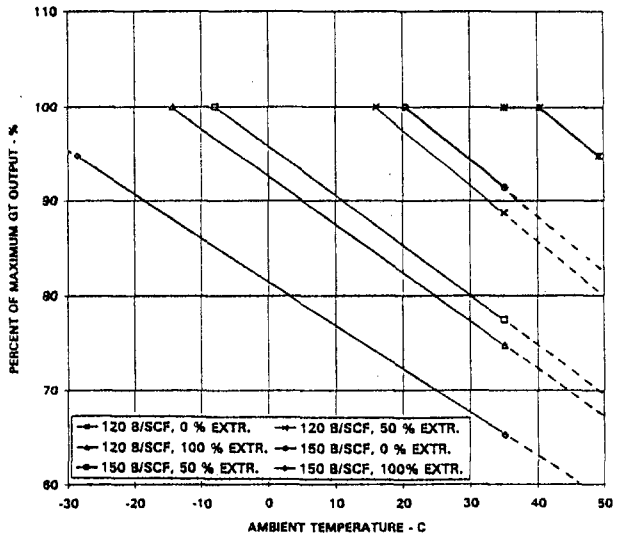


Fig. 4 Gas turbine performance as a function of ambient temperature for two fuel equivalent heating values and three levels of ASU integration

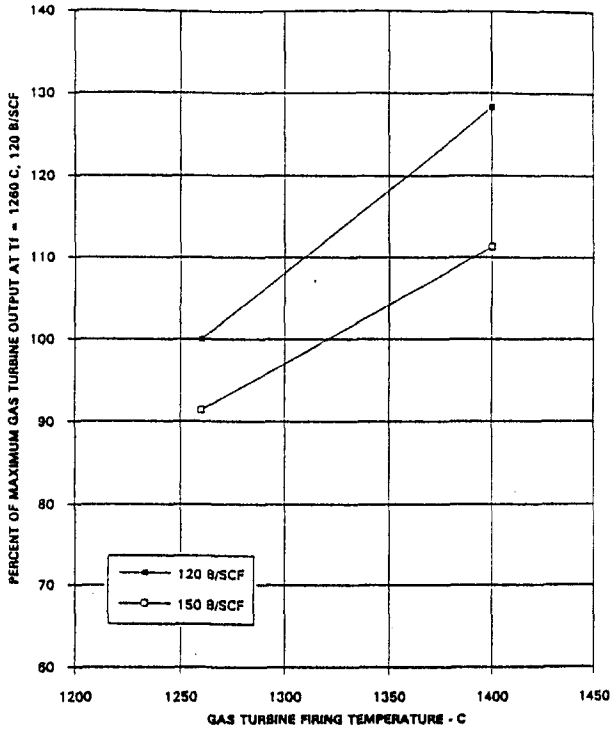


Fig. 5 Gas turbine output as a function of firing temperature for two fuel equivalent heating values and zero ASU integration

Although the performance comparisons shown in Fig. 4 have been made at a gas turbine firing temperature of 1260°C, gas turbine technology currently allows significant increases in firing temperature. Accordingly, the performance data shown in Fig. 5 have been included to show the general dependence of gas turbine output on firing temperature for the two fuel equivalent heating values and gas compositions considered. This assessment has been made using the performance parameters of the GE MS9001FA gas turbine and does not reflect the full advantage of increased firing temperature that could be achieved with the proper pressure ratio matching in the design. The data have been normalized to 100 percent for the case of 120 Btu/scf fuel equivalent heating value and 1260°C firing temperature and are meant to provide a conservative illustration of the importance of firing temperature on gas turbine and overall IGCC system performance.

Figure 6 shows the adiabatic flame temperature predicted by a thermochemical equilibrium calculation for the two fuels considered in the IGCC system analysis. The equivalent fuel/air ratios of interest for the gas turbine are in the range of 0.26 to 0.48 depending on the fuel equivalent heating value and firing temperature assumed. In general, current 1260°C firing temperature class gas turbines can be expected to have about a 100°C temperature difference between the actual combustion temperature and the gas temperature at the entry to the first rotating stage of the turbine, the latter being defined here as the firing temperature. Examination of Fig. 6 reveals that significant increases in firing temperature may still be obtained, even with the lower fuel equivalent heating value, although other limitations such as thermal NO_x formation will have to be addressed.

IGCC System Performance. Once the gas turbine performance has been established, this determines the required fuel flow from the gasification system as well as the requirements for energy and mass transfer to and from the air separation unit. In addition, as the gasification process flow information is developed, the heat and mass interchange with the steam bottoming cycle and the low level heat integration with the cleaned syngas and nitrogen can be developed on an iterative basis until all heat and mass balances for the overall system are closed. Referring again to Fig. 1, the IGCC system configuration used for the normalized data presented below has been based on an elevated pressure air separation system with utilization of the waste nitrogen through adiabatic compression and admission to the gas turbine's combustor. The nitrogen has been moisturized in the case of the 120 Btu/scf fuel equivalent heating value and heated in heat exchange with the air extracted from the gas turbine's compressor discharge, where applicable. The gasification system has been modeled as a full heat recovery system with high-pressure saturated steam supplied to the bottoming cycle from the gas cooling sections. The steam bottoming cycle employs a three pressure level heat recovery steam generator with reheat steam turbine inlet conditions of 101 bara and 538°C/538°C. The gas cleaning system utilizes water scrubbing, conventional acid gas removal, and conversion to elemental sulfur and reheating with boiler feedwater to 288°C prior to admission to the gas turbine.

The net electric output of the various system configurations evaluated is shown as a function of ambient temperature in Fig. 7. This plot shows two additional data points, which incorporate performance estimates at 1400°C gas turbine firing temperature. The data have been normalized with respect to the maximum IGCC system net electric output for 1260°C firing temperature for the range of ambient temperature considered. The system designer must then overlay such a plot with site constraints on operating temperature range, design ambient temperature, and whether the primary economic drivers are capital or operating costs. In some cases, relatively low or negative fuel cost, coupled with the requirement for output for a specific power sale agreement, can make output the primary system discriminant. In particular, if high system output at high ambient temperature is required, the lower fuel equivalent heating value, zero air extraction case is preferred. However, at lower ambient temperature, this system reaches a gas turbine limited output at a lower IGCC system net electric output level than other configurations having partial to full air extraction for either of the fuel equivalent heating values considered.

Each of the cases shown in Fig. 7 has been treated as a "design" case with the auxiliary power requirements, low level heat exchange, and steam cycle integration being specific to the

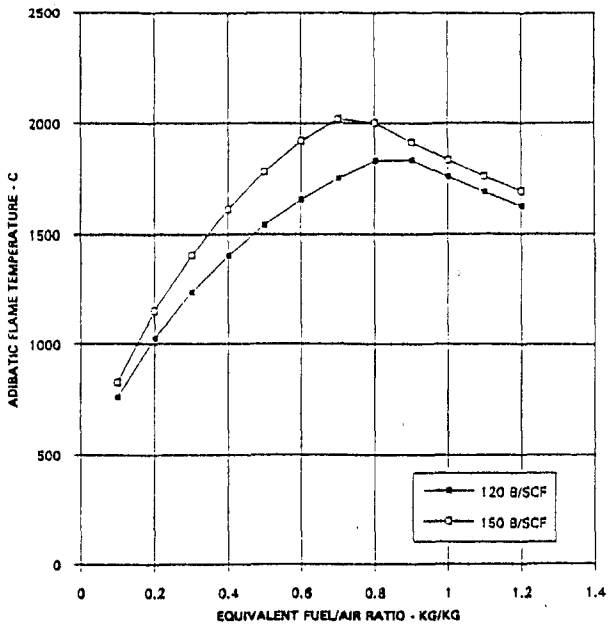


Fig. 6 Adiabatic flame temperature versus equivalent fuel/air ratio for two fuel equivalent heating values: T_{fuel} = 233°C, T_{air} = 416°C

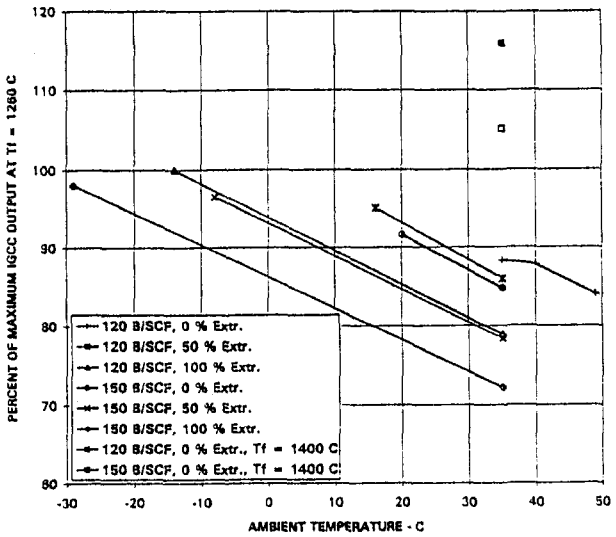


Fig. 7 IGCC system performance as a function of ambient temperature for two fuel equivalent heating values and three levels of ASU integration; relative system output is also shown for elevated gas turbine firing temperature

particular case. Once a preliminary output evaluation such as this has been performed, and the design operating ambient temperature and temperature range have been selected, it is necessary to fix the equipment specifications and performance and then re-evaluate the output over the operating temperature range. In the case of air separation integration, this requires careful selection of an auxiliary air compression system that will provide sufficient flexibility to meet the highest temperature output requirement while meeting the degree of extraction required for design point operation.

If low operating cost, rather than low capital cost, is the primary design requirement, then data similar to those plotted in Fig. 8 must be examined. The two curves shown represent the overall IGCC system net efficiency (LHV) for the two fuel equivalent heating values and different degrees of integration of the air separation unit. The curves are specific to 35°C ambient temperature, a typical upper limit specified for summer peaking IGCC power generation systems. Although higher system efficiencies (up to 46.5 percent, LHV) can be calculated for lower temperature operation, these data for 35°C ambient temperature operation illustrate the high ambient temperature efficiency consequences of selecting a particular ambient temperature (15°C) design point for both fuel equivalent heating values. A data point for 16°C ambient temperature operation with the lower fuel equivalent heating value is included to illustrate the higher efficiency obtainable.

In the case of zero air extraction, the entire air separation unit air supply must come from an auxiliary, intercooled compressor, which, in general, will have a lower component efficiency than the gas turbine's multistage, axial flow compressor. In addition, the intercooled auxiliary compressor must reject a large portion of the heat of compression to cooling water at low temperature while the air extracted from the gas turbine compressor can exchange heat at high temperature with the nitrogen being returned to the gas turbine. The nonintegrated system's efficiency also suffers from the additional electrical inefficiencies related to motor-driven compressors as well as the effect of scale of the additional compressor relative to the large gas turbine compressor. As pointed out earlier under the discussion of Fig. 4, the gas turbine's output "break point" for the lower fuel equivalent heating value and zero air extraction occurs at an ambient temperature greater than 35°C. Therefore, full, mechanically limited gas turbine output is available at 35°C for the zero extraction case. In going from 0 to 50 percent extraction, loss of gas turbine output is more than compensated by the thermal integration and increased compression efficiency resulting from ASU integration and results in an increase in overall system efficiency. Further integration with the ASU reduces gas turbine output sufficiently to offset further compression efficiency and regenerative thermal gains for the lower fuel equivalent heating value system with the result that full cycle efficiency for full extraction equals that for zero extraction.

In the case of higher fuel equivalent heating value system performance, the gas turbine output "break point" has previously been shown to be 21°C (Fig. 4) with gas turbine output significantly reduced at 35°C (to approximately 92 percent of maximum). In this case, further reductions in gas turbine output due to air extraction more than offset any thermal integration and compression efficiency gains for all values of air extraction percentage and both overall system output and efficiency decrease. It can be concluded that if maximum output at high ambient temperature is desired, the lowest tolerable fuel equivalent heating value will allow the greatest IGCC system efficiency and that partial air extraction may provide increased efficiency at reduced output (Fig. 7). Both output and efficiency for the 16°C design point benefit substantially from partial integration of the ASU.

Performance calculations have been made for the MS9001FA based IGCC system with higher gas turbine firing temperatures than the current 1260°C limit. The results of these analyses are shown in Fig. 9. As noted earlier, the full thermodynamic benefit of the increased (1400°C) firing temperature on cycle efficiency has not been reached for this system configuration as the gas turbine pressure ratio is lower than optimum. However, it can be seen that significant increases in efficiency are possible for both fuel equivalent heating values, providing ample incentive to utilize the continuing advances in gas turbine development.

Figure 10 shows a key result of the system analyses performed. Once the choice of a particular gasification system and system output has been made, the primary independent variable left to the system designer is the percentage of the ASU air

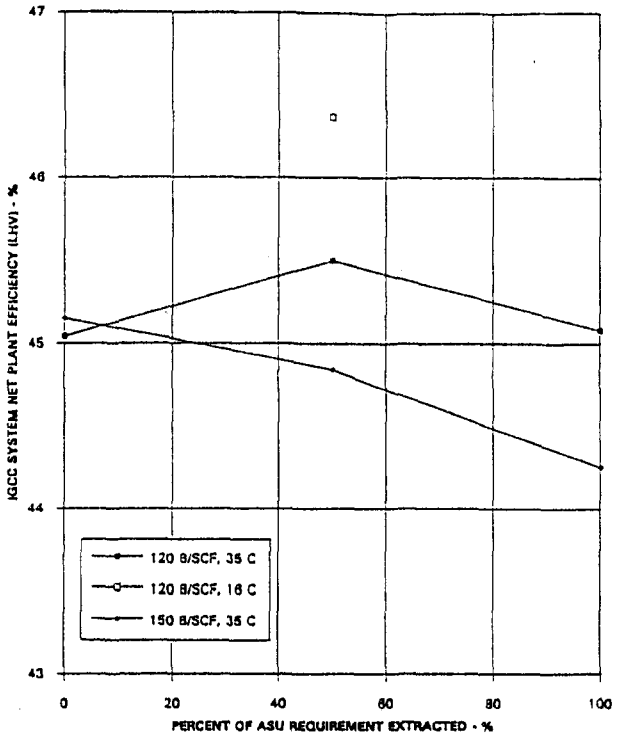


Fig. 8 Overall IGCC system efficiency (LHV) as a function of level of ASU integration for two fuel equivalent heating values at 16°C and 35°C ambient temperature

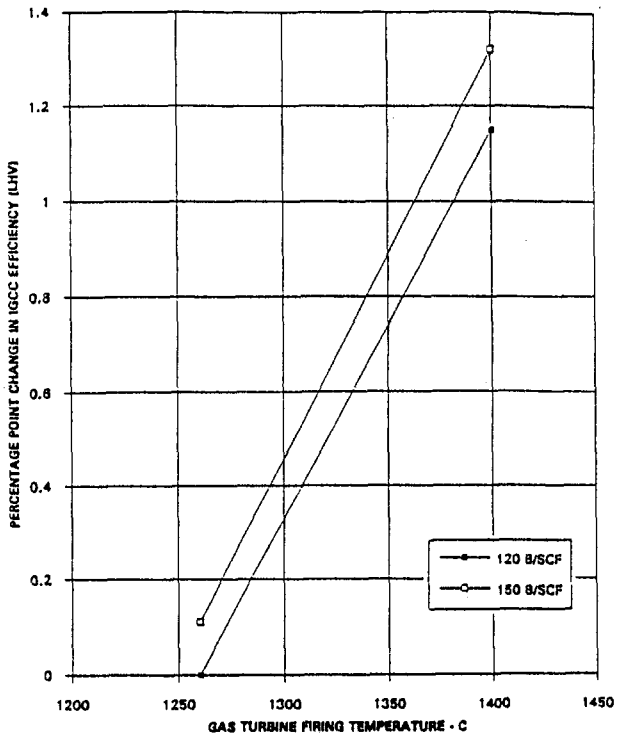


Fig. 9 Percentage point change in overall IGCC system net efficiency for two fuel equivalent heating values as a function of gas turbine firing temperature

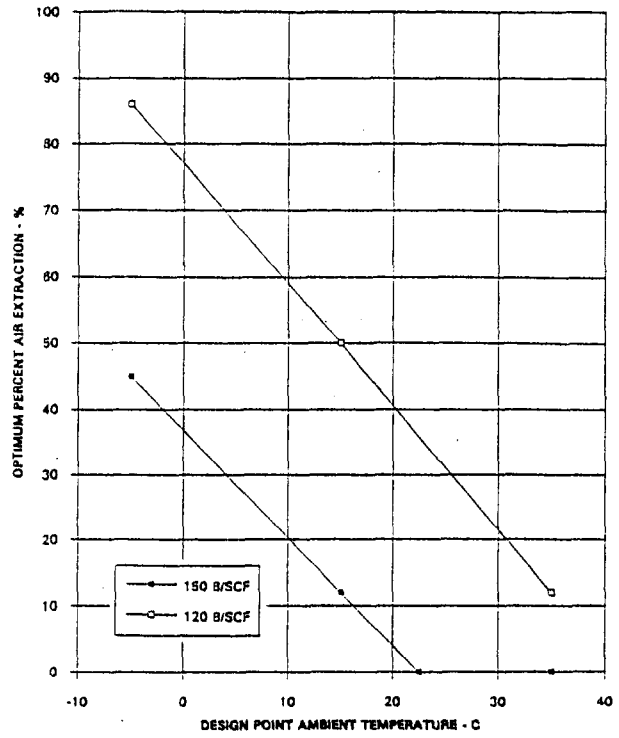


Fig. 10 Optimum gas turbine air extraction for design point operation of an IGCC system versus design point ambient temperature for two fuel equivalent heating values

supply to be extracted from the gas turbine compressor. The points shown in Fig. 10 have been estimated from the data shown in Fig. 7 utilizing a linear relationship between the change in ambient temperature between gas turbine "break points" and the change in percent air extraction. This results in a set of data points (Fig. 10) that represent the optimum air extraction percentage for each fuel equivalent heating value and ambient temperature for the specific IGCC system being considered. In practice, an auxiliary compressor must be sized for one specific operating point with the air extraction percentage then dictated by the particular compressor's operating map. Typically, centrifugal compressors designed for this application can allow a range of volumetric flow from 60 to 110 percent of their nominal design value. The system designer must then consider the operating ambient temperature range required (summer peaking, winter peaking) and specify the compressor's fractional capacity at the design point ambient temperature.

Conclusions

It is essential that IGCC power generation systems be as carefully integrated at the design stage as possible. However, carefully integrated does not necessarily mean fully integrated. Start-up, part-load, and other off-design operational requirements must be considered in the design. In particular, air supply integrated ASUs must be designed for variable air pressure operation based on gas turbine compressor discharge pressure variation with load. The primary energy converter, and the component that determines the output of the IGCC system, is the gas turbine and advances in IGCC efficiency and output will derive primarily from this technology. However, within the set of available IGCC-compatible gas turbines, the system designer can substantially adjust the output and efficiency to suit the requirements of a particular project by varying the amount of air extraction from the gas turbine compressor. This is particularly true for the oxygen blown gasification system considered here.

Advances in gas turbine technology leading to higher firing temperature cycles can be readily adapted to IGCC systems with the range of fuel equivalent heating values in use today. Injection of waste nitrogen from the air separation process, along with heat exchange with the extracted air flow for the case of air extraction, contributes to power augmentation, NO_x suppression, and cycle efficiency and is a part of several major IGCC demonstration projects in construction. These project applications range from zero to full ASU integration and will provide a basis for future system design decisions utilizing the type of analyses presented. The system designer must clearly identify the roles of capital and operating cost as well as the need for part or base-load operation prior to selecting the gas turbine size and extent of ASU integration in order to provide the lowest life cycle cost IGCC power generation system.

Acknowledgments

The authors wish to acknowledge the contributions of M. Jandrisevits, B. Sandhu, J. Sholes, B. Gardener, and L. O. Tomlinson to the general development of IGCC system analysis at GE Power Generation, on which the results of this paper are based.

Reference

Cook, C. S., Corman, J. C., and Todd, D. M., 1995, "System Evaluation and LBTU Fuel Combustion Studies for IGCC Power Generation," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 117, pp. 673–677.

Combined-Cycle Power Stations Using "Clean-Coal Technologies": Thermodynamic Analysis of Full Gasification Versus Fluidized Bed Combustion With Partial Gasification

G. Lozza

P. Chiesa

Department of Energetics,
Politecnico di Milano,
Milan, Italy

L. DeVita
Enricerche,
Milan, Italy

A novel class of power plants for clean conversion of coal into power has been recently proposed, based on the concept of partial coal gasification and fluidized-bed combustion of unconverted char from gasification. This paper focuses on the thermodynamic aspects of these plants, in comparison with full gasification cycles, assessing their performance on the basis of a common advanced power plant technology level. Several plant configurations are considered, including pressurized or atmospheric fluidized-bed, air- or steam-cooled, with different carbon conversion in the gasifier. The calculation method, used for reproducing plant energy balances and for performance prediction, is described in the paper. A complete second-law analysis is carried out, pointing out the efficiency loss breakdown for both technologies. Results show that partial gasification plants can achieve efficiencies consistently higher than IGCC, depending on plant configuration and carbon conversion, making this solution a viable and attractive option for efficient coal utilization.

1 Introduction

Environmental protection and rationalization of primary energy sources consumption call for new technologies for utilization of coal and residual fuels in the power industry. The marriage between coal and gas turbine-based combined cycle (Webb et al., 1993) has been proposed as a winning strategy to achieve better conversion efficiencies and lower pollutant emissions than those of pulverized-coal fired-boiler steam power stations. Two main basic technologies have been widely studied in recent years: full gasification (IGCC) and pressurized fluidized bed combustion (PFBC).

The gasification concept involves the addition of large pieces of hardware to the power plant, including, beyond the gasifier itself, an air separation process, syngas cooling and clean-up, acid gas treatment, etc. Coal conversion efficiency is high (42–45 percent HHV values are now quoted in recent projects) but is limited by the inefficiencies of the several components and by the necessity of gas cooling. NO_x control is another sensitive matter, requiring steam injection or fuel humidification in addition to the best burner technology.

On another side, the PFBC technology is much simpler in terms of plant configuration and emission control (obtained by sorbent addition to control SO_2 emissions and by the low combustion temperature to control NO_x); however, it imposes a severe limitation on the highest cycle temperature to ≈850°C of the combustion bed, thus not allowing the use of high-temperature high-efficiency gas turbines.

Recently, fluidized-bed technology has been re-evaluated by a novel concept, named by various authors as "hybrid cycle" or "topping cycle" or "second-generation PFBC." The concept embodies a partial gasification process providing synthetic fuel to a topping combustor, which enables the utilization of advanced gas turbines; the unconverted carbon-rich residuals (char) from partial gasification are then processed through a FBC, providing heat to combustion air or to steam. We will refer to these plants with the acronym PGFBC (Partial Gasification Fluidized Bed Combustion). The gasifier is air-blown and sulfur is captured by sorbent addition in both the gasifier and the FBC; therefore typical IGCC substations like ASU and acid gas removal/treatment are not included in PGFBC plants.

Due to the potential advancement in conversion efficiency and in capital cost reduction, PGFBC is now the subject of several recent studies (Robertson et al., 1993; Robertson and Bonk, 1994; Maude, 1993; Minchener et al., 1993; Soothill and Brown, 1993; Foster-Pegg and Fraas, 1993; Bohn et al., 1993), with reference to different plant concepts. In these papers, there is a general agreement about a significant improvement of conversion efficiency with respect to IGCCs, but a comparison among the various possible PGBFC solutions cannot be found, since most authors "marry" a specific proprietary technology. In addition, a detailed analysis of thermodynamic processes and second-law considerations cannot be found in the literature.

The present paper is devoted to the thermodynamic analysis of PGFBC cycles, with different plant schemes including both pressurized and atmospheric fluidized bed combustion. These plants will be compared to a representative IGCC scheme based on an entrained-bed slurry gasifier, including the latest technology of pressurized ASU and hot-gas-clean-up. As far as power cycle components are concerned, all cases are based on large heavy-duty gas turbines, having TIT = 1280°C and β = 15, and

Contributed by the International Gas Turbine Institute and presented at ASME Cogen Turbo Power '94, Portland, Oregon, October 25–27, 1994. Manuscript received by the International Gas Turbine Institute January 15, 1994. Associate Technical Editor: E. M. Greitzer.

a three-pressure reheat steam cycle: This provides a common set of assumptions in the performance calculation, enabling a direct comparison among the various solutions.

The performance analysis will be carried out by using the second-law approach, allowing a deeper understanding of the loss generation.

2 Calculation Method

The calculation model used to generate the results described here has been specifically developed to predict the performance of complex gas-steam cycles, e.g., gas cycles with intercooling, reheat, recuperation, bottoming steam cycle, and mixed gas–water cycles like steam-injected, water-injected, and HAT (humid air turbine), as described by Macchi et al. (1994). The structure and the capabilities of the model have been extensively described in previous papers (Consonni et al., 1991; Consonni, 1992; Lozza, 1990, 1993; Chiesa et al., 1993); we recall here only the most significant features and the modifications introduced to enable the calculation of gasification processes.

2.1 Basic Outline. The system to be calculated is defined modularly as an ensemble of interconnected components, which can be of twelve basic types: compressor, gas turbine expander, splitter, mixer, heat exchanger, combustor, pump, saturator, steam cycle (including all its components), shaft (accounts for turbomachine spool interconnections, as well as electric losses), chemical converter, air separation plant. Operating characteristics and mass/energy balance of each component are calculated sequentially until the conditions (pressure, temperature, mass flow, etc.) at all interconnections converge toward a stable value. Aside from the algorithm handling the component network (it is virtually possible to analyze any cycle configuration), the most distinctive features of the model lie in the calculation of the key cycle components: turbomachines and heat recovery steam generator.

The cooled gas turbine expansion is calculated as a sequence of small steps, each consisting of an expansion followed by gas-coolant mixing. At each step, the coolant flow required to maintain the blade temperature within an assigned value is found by the heat flux balance across the blade wall. The coolant is bled at the minimum pressure required to overcome coolant circuit pressure drops, and then discharged into the main flow. The polytropic efficiency of both the cooled expansion steps and the uncooled turbine varies with a similarity size parameter to account for scale effects; exit kinetic energy is partly recovered in the diffuser. The compressor polytropic efficiency is also predicted by means of a similar size parameter.

The calculation of the steam bottoming cycle and the evaluation of the steam turbine expansion have been addressed in previous papers by Lozza (1990, 1993). We can recall here that: (*i*) the model can handle steam cycles with a virtually unlimited number of evaporating pressures (a maximum of four has been set), the higher can be supercritical; (*ii*) at any pressure, steam generation, reheating, or both can occur; (*iii*) steam turbine efficiency is predicted by built-in correlations, using a stage-by-stage calculation on the basis of the stage specific speed; (*iv*) steam and/or saturated water can enter or exit the cycle from several points (drums, deaerator, superheaters outlet, turbine); (*v*) steam pressures can be optimized to maximize efficiency, by a numerical optimization routine.

The saturator is another peculiar component, which has been extensively discussed by Macchi et al. (1994), in the framework of HAT cycles. This component is extensively used in gasification cycles to moisturize the synthetic fuel, lowering NO_x emissions and recovering low-grade heat.

To render the calculation method capable of reproducing gasification processes, some new features were added: (*i*) a novel component, the ASU; (*ii*) more powerful tools to handle chemical conversions; (*iii*) the "equivalent coal" model. These features are described in the next paragraph.

2.2 New Features

2.2.1 Air Separation Unit. The air separation process is rather complex and its detailed calculation requires several features, which are beyond the capabilities of the computational method. In fact, our analysis is devoted to the thermodynamic assessment of the overall plant capabilities and not to the design of subsystems: therefore, our interest toward the ASU is here limited to the power (or compressed air) consumptions and to the interaction with the remainder of the plant. Consequently, the ASU component is rather simple: The pressure of air entering the "cold-box" (i.e., cryogenic heat exchangers, separation column, etc.) is an input datum, as well as the pressure of the streams (O_2 with assigned purity, pure N_2, wastes) exiting the cold-box. The program calculates the power requirement to: (*i*) compress the incoming air from the entrance condition to the cold-box inlet pressure; (*ii*) compress O_2 and N_2 from cold-box exit to the pressure at which those streams are used in the plant (gasifier or others). In these compressions, the polytropic efficiency, the number and the effectiveness of intercoolers are input data to calculate the power requirement. This schematization can manage both cases of conventional ambient air supply and of pressurized ASU, with air supplied by the gas turbine compressor.

2.2.2 Chemical Conversions. Among the above quoted components, chemical reactions can take place in four components:

- Combustor: All chemical species containing C, H, and S are fully oxidized to CO_2, H_2O, and SO_2; the heat balance does not require the concept of LHV (or HHV), since the enthalpy of all chemical species is referred to their heat of formation.
- Mixer: The chemical species mixed together can or cannot react, depending on the user's choice. If they can react, the final composition and temperature is calculated by the chemical equilibrium hypothesis. The method is the

Nomenclature

G = mass flow, kg/s
G_a = air mass flow, kg/s
Δi_{is} = isentropic enthalpy drop, J/kg
Δp = pressure drop, Pa
ΔT = temperature drop, K
η_y = polytropic efficiency

Acronyms

AFBC = Atmospheric Fluidized Bed Combustion
ASU = Air Separation Unit
BFW = Boiler Feed Water
CGCU = Cold Gas Clean-Up
FBHE = Fluidized Bed Heat Exchanger
FBC = Fluidized Bed Combustion
GCCR = Gasifier Carbon Conversion Ratio
GT = Gas Turbine
HGCU = Hot Gas Clean-Up
HHV = Fuel Higher Heating Value
HRSG = Heat Recovery Steam Generator
IGCC = Integrated Gasification Combined Cycle
LHV = Fuel Lower Heating Value
PFBC = Pressurized Fluidized Bed Combustion
PGFBC = Partial Gasification Fluidized Bed Combustion
ST = Steam Turbine
TIT = Turbine first rotor Inlet total Temperature

Table 1 Consequences of assuming a syngas composition derived from chemical equilibrium at gasifier outlet, rather than using the actual syngas composition (as respectively estimated from Matchak et al., 1984; Dawkins et al., 1985; delaMora et al., 1985). All data are relative to the raw syngas exiting the gasifier; specific heat is the mean value between the gasifier outlet temperature and 100°C; the quoted efficiencies refer to a power cycle with state-of-the-art gas turbines and improved steam cycle conditions.

Gasifier type	Entrained bed		Fluidized bed		Fixed bed	
	litera-ture	chem. equil.	litera-ture	chem. equil.	litera-ture	chem. equil.
Temperature °C	1295	1307	1003	928	377	698
Specific heat J/kg	1769	1764	1810	1805	1681	1789
HHV kJ/kg	10380	10440	12830	13070	13640	13090
LHV kJ/kg	9321	9381	11620	11870	12350	11790
CO mols%	39.01	40.84	38.08	40.10	43.58	30.63
CO_2 mols%	10.62	8.78	11.43	9.28	2.73	12.59
H_2 mols%	29.76	28.54	29.30	32.70	25.22	29.82
H_2O mols%	17.72	19.19	12.10	11.29	19.81	13.94
LHV efficiency %	45.66	45.65	44.88	45.08	46.96	44.60

minimization of the Gibbs free energy, as developed by Reynolds (1986); the chemical species present in the output stream have to be assigned. For instance, a combustor can be represented by a mixer: The difference is that CO formation can be predicted by the mixer (at chemical equilibrium), but not by the combustor. This option will be used for modeling gasifiers, provided that chemical equilibrium is a sufficiently accurate hypothesis.

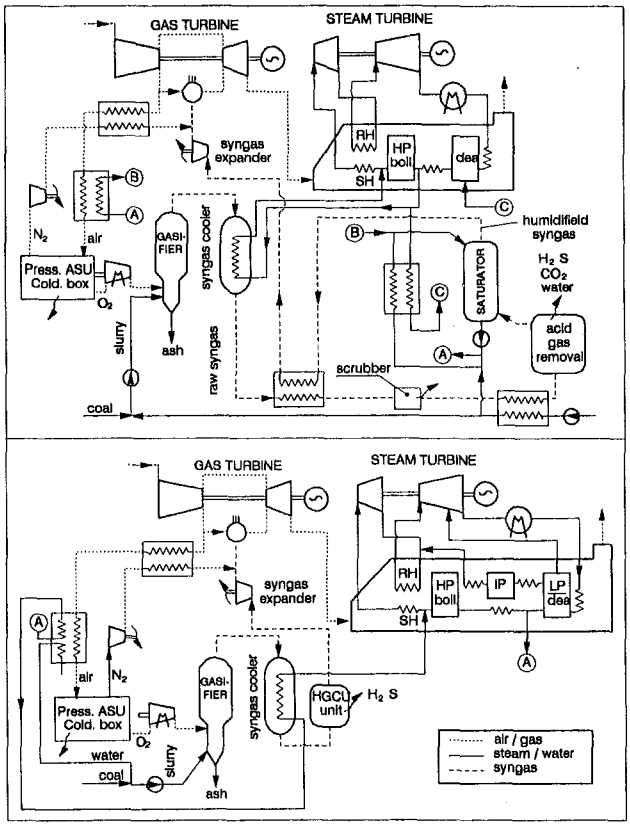


Fig. 1 Plant schemes of IGCC cycles considered in the paper, based on entrained-bed slurry gasifier with pressurized ASU. The upper configuration refers to cold-gas-clean-up, with syngas saturation; the lower one includes hot-gas-clean-up.

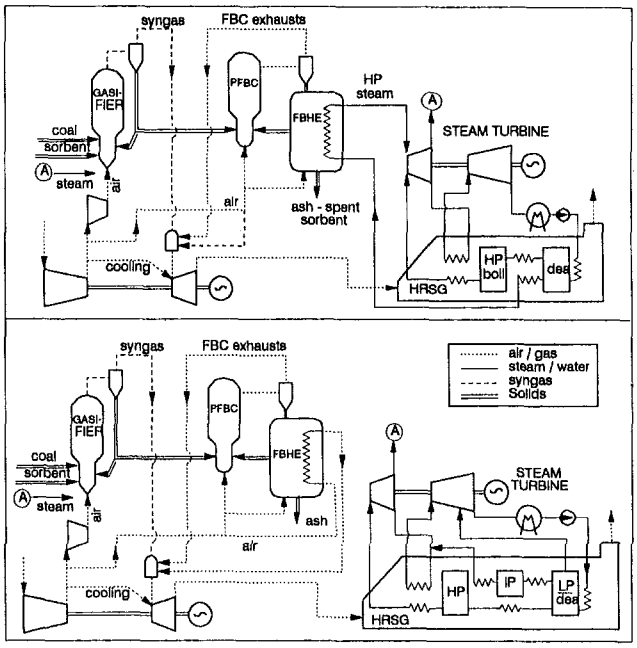


Fig. 2 Partial-gasification plants with pressurized FBC. The FBHE can be steam-cooled (upper part) or air-cooled (lower part).

- Heat exchanger: One of the streams in a heat exchanger can chemically react as in the mixer (the reaction here accounts for the heat transferred to the other stream, while the mixer is adiabatic). It is useful, for instance, for modelizing FBCs.
- Chemical converter: It is essentially a mixer, whereby the composition of the outlet stream is assigned by means of molal fractions of the representative components. The element simply completes the atomic balance and provides the outlet temperature and enthalpy. This can be useful to reproduce gasifiers providing syngas with a known composition, far from the one at equilibrium.

All the elements above can include heat and/or pressure losses. The question of whether chemical equilibrium can appropriately represent the reactions in the plant schemes here considered requires some comments. In a previous analysis concerning IGCCs (Lozza et al., 1993), three oxygen-blown gasification plants, with fixed-bed (delaMora et al., 1985), fluidized-bed (Dawkins et al., 1985) and entrained-bed (Matchak et al., 1984) gasifiers, were calculated by both options of chemical equilibrium and of syngas composition as quoted by respective authors. The results, summarized in Table 1, show that equilibrium is very accurate for entrained-bed, high-temperature gasifiers, but provides large errors for fixed-bed gasifiers. For fluidized-bed, there are some discrepancies about syngas composition (within ±4 percentage points in molal composition), providing heating values 2 percent higher and a lower gas outlet temperature: The error on the cycle thermodynamics is small (0.2 points), since the larger gas turbine output, made possible by higher gas HHV, is partly balanced by the lower steam production from the syngas coolers. In the present paper, it was decided to resort to chemical equilibrium in all calculations, mainly because of the much higher flexibility to handle different situations (coal ranks, carbon conversion, steam and/or air quantities); however, possible errors on syngas composition will provide uncertainties on the cycle performance much lower than the ones introduced by the several assumptions common to all cycle calculations (i.e., turbomachine efficiency, heat exchanger effectiveness, and so on). In addition, PGFBC performance will be less dependent on the gasifier behavior than IGCC, since a relevant fraction of the fuel input will be handled by the combustion process.

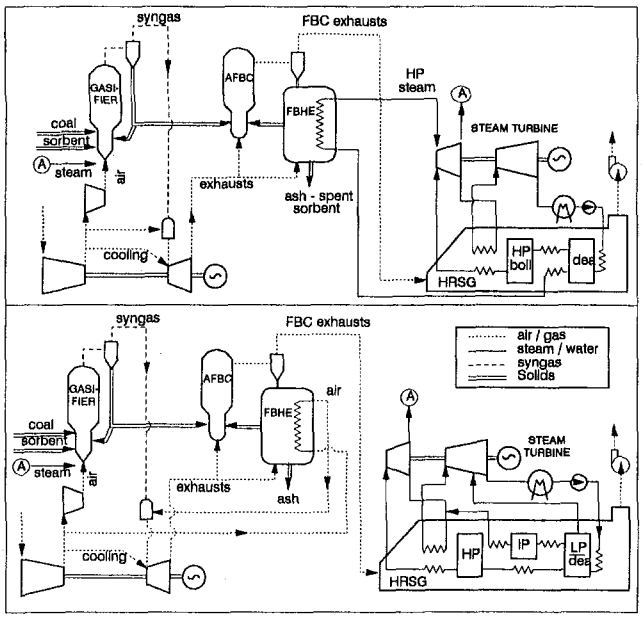


Fig. 3 Partial-gasification plants with atmospheric FBC; the FBHE can be steam-cooled (upper part) or air-cooled (lower part)

2.2.3 Equivalent Coal. Coal is normally defined by its "ultimate analysis" providing an atomic composition (C, H, O, N, S), moisture and ash contents, and heating value. The actual molecular composition is generally unknown. The present calculation method does not make use of heating values and it is therefore necessary to create a "virtual" coal composition providing the same atom balance and the same heating values. This task is operated by a routine that determines the best combination among the species considered (including C, CO, CO_2, CH_4, C_2H_2, C_2H_6, C_3H_8, H_2, H_2S, COS, SO_2, NH_3, H_2, N_2, O_2, S), minimizing the quadratic error on atom balance and HHV. Water and ash (considered as an inert) are eventually added. This provides a fuel perfectly equivalent under all points of view, except for entropy (stipulating the entropy of coal is a very arduous job!): entropy productions in gasifiers and combustors will therefore be referred to that of the equivalent fuel.

2.2.4 Gasifier Carbon Conversion Ratio (GCCR). It is the ratio—by weight—between the carbon converted in the gasifier to syngas components—mainly CO and CO_2—and the total carbon input: It can be considered 1 for IGCCs (apart from incomplete conversion) and 0 for PFBC plants, assuming intermediate values for PGFBCs. It is the main parameter governing the thermodynamics of cycles considered in this paper. GCCR must remain inside an "operating window," for any configuration, depending on constraints discussed later. It is important to notice that a low GCCR can be obtained also bypassing part of the coal directly to the FBC, to improve combustion characteristics: This will not substantially change the results of the following analysis.

Table 2 Coal and sorbent composition

Coal Composition (by weight)	
Carbon	64.44%
Hydrogen	3.95%
Oxygen	7.40%
Sulphur	0.85%
Nitrogen	1.49%
Ash	12.67%
Moisture	9.2%
Coal HHV, MJ/kg	25.71
Coal LHV, MJ/kg	24.66
Sorbent: dolomite ($CaCO_3 \cdot MgCO_3$)	
Ca/S ratio (0.098 $kg_{dolomite}/kg_{coal}$)	2

Table 3 Assumptions for calculation of gasifiers and fluidized bed combustors

Coal input (600 MW_t LHV):	24.33 kg/s
Auxiliary power consumption:	6 MW_e
Entrained Bed Gasifier (for IGCC)	
Water/coal ratio in slurry:	0.323
O_2/coal ratio [1]:	0.852
Gasifier pressure:	40 bar
Overall pressure loss[2]:	14%
Heat loss (% of input LHV):	0.5%
Other components of IGCC	
ASU cold box pressures:	15/4.2 bar
N_2/O_2 compressors η_y:	0.78
HGCU inlet temperature:	550°C
CGCU inlet temperature:	37°C
Syngas recup. effectiveness:	0.87
Syngas cooler heat loss:	1.5%
Syngas recup. heat loss:	1.5%
Scrubber heat loss:	50°C
HGCU unit heat loss:	67°C
Fluidized Bed Gasifier	
Bed temperature:	870°C
Carbon conversion (GCCR):	0.3÷0.7
Steam/coal ratio:	GCCR/2
Pressure loss:	6%
Heat loss (% of input LHV):	0.5% (1%[3])
Fluidized Bed Combustor	
Bed temperature:	870°C
Pressure loss (pressurized):	6%
Pressure loss (atmospheric):	10%
Heat loss (% of input LHV):	0.5% (0%[3])
Minimum O_2 in exhausts: 6%	

[1] Such to obtain a raw syngas temperature of 1370°C
[2] From gasifier to syngas expander inlet
[3] After correction for sorbent reactions

3 Plant Schemes and Assumptions

3.1 Reference IGCC. The IGCC plant used as a reference in this paper is based on an entrained-bed oxygen-blown gasifier fed by a coal–water slurry, one of the most popular gasification techniques. Two versions are considered with cold (CGCU) or hot (HGCU) gas clean-up. The first version, schematically shown in Fig. 1 (upper part), requires an efficient heat recovery system from gas cooling to the near-ambient temperature required by conventional acid gas removal processes. With more detail, the raw syngas exiting the syngas cooler (which raises saturated steam from BFW preheated in the HRSG) is processed through a syngas recuperator, increasing the temperature of clean syngas, a scrubber, and a heat exchanger warming water for saturator make-up, slurry formation, and steam cycle make-up. Acid gases, together with CO_2 and condensate, are removed from syngas. Cleaned syngas is satu-

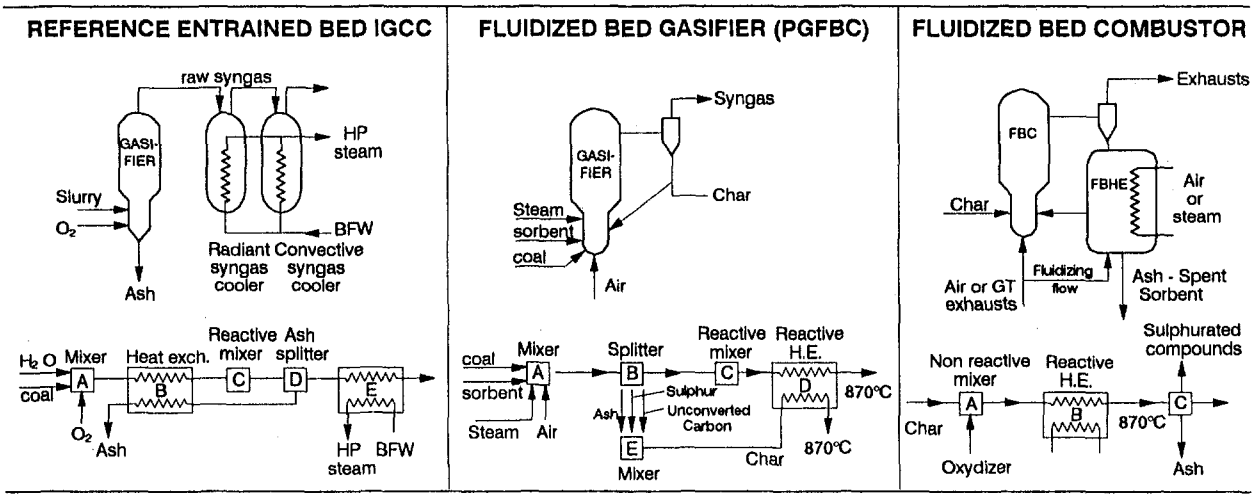


Fig. 4 Modelization of gasifiers and FBC by means of the components made available by the computational method, with real (upper drawings) and simulated (lower drawings) streams entering and exiting the systems

rated with water coming from HRSG economizer and from ASU heat recovery. Saturator is useful for two purposes: NO_x control, due to addition of large amounts of water into fuel, and efficient heat recovery from low-temperature streams. After saturation, clean syngas is heated by the recuperator and, before being ducted to the gas turbine combustor, its excess pressure is converted into mechanical power by an expander. The HGCU scheme (Fig. 1, lower part) is simpler: The clean-up unit operates at high temperature and provides cleaned syngas directly to the combustor, after expansion.

Both schemes include a pressurized ASU, using air entirely extracted from gas turbine compressor outlet. This option is often proposed in IGCC projects (Klosek et al., 1993; Rao et al., 1993), allowing capital cost savings and a slight efficiency improvement. In our schemes, the oxygen-rich stream exiting the cold box is compressed by an intercooled compressor to the gasifier pressure, while the nitrogen stream is completely recycled into the gas turbine, after adiabatic compression and mixing with fuel, for NO_x suppression. To render the pressurized ASU attractive on the efficiency point of view, the heat recovery from incoming air should be very carefully studied: For instance, in our case about 22 percent of compressor air flow is ducted to the ASU at about 400°C, with a large energy content. Part of heat is used to increase the temperature of nitrogen reinjected into the gas turbine; the remainder is used to warm up the saturator water in the CGCU scheme, or to preheat water to be evaporated in the syngas cooler in the HGCU scheme, where BFW is extracted from the deaerator (rather than after the HP economizer, as in the CGCU scheme).

The power cycle includes, beyond the gas turbine, a three-pressure reheat steam cycle. However, calculations show that the multipressure arrangement is not necessary (nor possible) for the CGCU scheme, since a complete heat recovery is accomplished by means of BFW preheating, providing near-saturated liquid to the syngas cooler boiler and to the saturator, in addition to HP drum: In this case, a single-pressure cycle was adopted (see Fig. 1).

3.2 Pressurized FBC. The reference plant schemes are shown in Fig. 2, for two configurations with steam-cooled FBHE (upper part) and air-cooled FBHE (lower part). The steam-cooled scheme reproduces the layout proposed by Robertson et al. (1993) and Robertson and Bonk (1994). The plant includes a gasifier, fed by coal, sorbent, air, and steam. Coal and sorbent treatment are simply considered by an auxiliary power requirement; air comes from the gas turbine compressor outlet and is further compressed by a booster compressor, in order to provide the head necessary to move the syngas through the fluidized bed, cyclone, filters, piping, and gas turbine combustor nozzles; steam is bled from a proper steam turbine location. Two streams exit the gasifier at the imposed bed temperature of 870°C: syngas and carbon-rich solids, including char, ash and partially reacted sorbent. Solids are ducted to PFBC, where combustion takes place with air coming from the main compressor. The FBHE raises HP superheated steam, from preheated feedwater coming from the HRSG. The oxygen-rich PFBC exhausts (at 870°C) serve as oxidizer to the gas turbine combustor, achieving the assumed TIT of 1280°C. After the expansion, exhaust gases are used for steam production, reheating, and feedwater heating through the HRSG. The steam turbine, collecting steam from FBHE and HRSG units, completes the plant scheme.

An alternate plant scheme is possible, by introducing an air flow bypassing the PFBC and coming directly to the gas turbine combustor, represented by the bold dotted line in Fig. 2 (upper part), thus quenching the oxidizer and reducing the air flow at the PFBC to the amount strictly necessary for combustion. This reduces size and cost of the PFBC (including the filtering system), especially when the gas turbine air flow becomes very large due to high GCCR (i.e., large syngas production requires a large combustion air flow, to maintain the imposed turbine inlet temperature).

The plant scheme with an air-cooled bed is shown in the lower part of Fig. 2. Air from compressor outlet is split into two streams (in addition to the ones feeding the gasifier and cooling the turbine): The first one is ducted to the PFBC as oxidizer, the second one as coolant in the FBHE, instead of steam. The two streams are then mixed together, as oxidizer to the gas turbine combustor.

Therefore, three plant schemes are depicted: two of them with steam-cooled FBHE, for which part of heat input serves for steam raising (we will call them PFBC-A, with full air feed to the FBC, and PFBC-B, with air bypass to the GT combustor); the third with air-cooled FBHE, in which all the heat input (through gasifier or fluidized-bed) remains inside the gas turbine cycle.

3.3 Atmospheric FBC. The difference between plant schemes with AFBC and PFBC is basically due to the fact that FBC is fed by gas turbine exhausts (provided that O_2 content is sufficient for combustion) rather than by air bled from compressor outlet. The AFBC fuel consists of char coming from the partial gasification (conceptually similar to the one of the previous case) via a pressure-let-down lockhopper. Exhausts

Table 4 Assumptions adopted for power cycle calculations

Gas turbine compressor

Δi_{is} =27 kJ/kg for all stages
Leakage 0.8% of inlet flow, at HP exit
$\eta_p = 0.895 \cdot [1 - 0.07108 \cdot \log_{10}^2(SP)]$ for $SP < 1$
$\eta_p = 0.895$ for $SP \geq 1$
Inlet Δp (filter)=1 kPa

Combustors

Oxidizer Δp/p=3%, fuel Δp/p=15%
Heat losses=0.4% of combustion heat (LHV)

Gas turbine

Δi_{is}: 300 kJ/kg (cooled stages),
100 kJ/kg (uncooled stages)
$\eta_p = \eta_{p,\infty} \cdot [1 - 0.02688 \cdot \log_{10}^2(SP)]$ for $SP < 1$
$\eta_p = \eta_{p,\infty}$ for $SP \geq 1$; $\eta_{p,nozzle} = 0.95$
$\eta_{p,\infty}$: 0.89 (cooled stages), 0.925 (uncooled stages);
diffuser recovery =50% of exit kinetic head
Maximum blade temperature:
830°C (1st nozzle), 800°C (cooled turbine)

HRSG and steam cycle

Live steam pressure: 110 bar (A) or 260 bar (B)
Live steam temperature (SH and RH):
538°C (A) or 565°C (B)
Reheat and IP drum pressure: 18 bar (A) or 35 bar (B)
LP pressure: 3.1 bar
Condensing pressure: 0.05 bar
Approach ΔT=25°C, Pinch point ΔT=10°C
Gas side Δp 3 kPa, Δp/p superheaters 8%,
economizers 10%, heat losses 0.7%
Minumum gas stack temperature: 100°C
Steam turbine efficiency: see Lozza (1990)
Auxiliary power consumption: 0.5% of condenser heat

Others

Electric generators: see Lozza (1990)
Organic losses 0.03% of turbomachine work

The size parameter SP used to evaluate turbomachinery efficiencies is defined as $V^{0.5}/\Delta i_{is}^{0.25}$, where V is the stage exit volumetric flow for turbines, and the average volumetric flow for the compressor.

from the AFBC at the bed temperature of 870°C enter the HRSG to produce steam. Heat developed by char combustion, besides raising the exhausts temperature, can be used to produce superheated steam (Fig. 3, upper part) or to preheat combustion air coming from the compressor (Fig. 3, lower part).

These two situations will be analyzed in the paper, but plant schemes with both air/steam heating (as stipulated by Foster-Pegg and Fraas (1993) with an air-cooled FBHE and a steam-cooled AFBC) will not be considered here (even if interesting) to limit the extent of the analysis. For the same reason, we will not consider plants for which the AFBC oxidizer is ambient air (Minchener et al., 1993) rather than gas turbine exhausts.

3.4 Assumptions. This paragraph will discuss the assumptions necessary to evaluate the performance of plants described in Sections 3.1–3.3 and to clarify the schematization used for simulating the various components by means of our computational model.

3.4.1 Coal and Sorbent. A low sulfur content (<1 percent) coal, imported from South Africa and widely used in Italian power stations, has been considered as fuel in this paper.

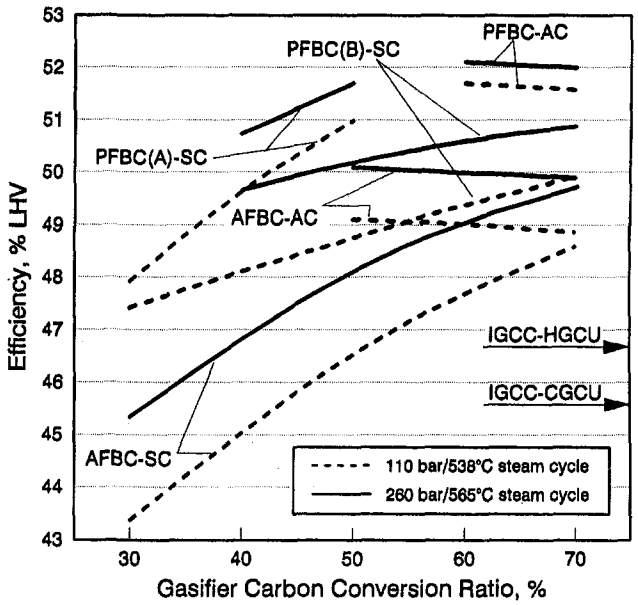


Fig. 5 LHV net efficiencies of plant configurations considered in the paper, as a function of the gasifier carbon conversion ratio

Its composition and heating values are shown in Table 2. The sulfur content of coal definitely affects plant performance, imposing the amount of sorbent necessary to control SO_2 emissions and the related heat of reaction and heat losses, due to the discharge of hot spent sorbent. In a IGCC, it affects the steam needed for H_2S absorption regeneration and the loss related to sulfur production (sulfur is a "fuel", contributing to the coal HHV). The assumption of a low-sulfur coal will increase plant efficiency of about 1–1.5 percentage points, with respect to a 3–4 percent S content coal.

In the present paper, a simplified approach, accounting for the reactions of Ca and Mg in the sorbent cycle, has been introduced, due to the inaccuracies of handling this matter by chemical equilibrium. As a matter of fact, in a PFB gasifier the following reactions can be assumed to take place: (*i*) decomposition of dolomite into $CaCO_3$ and $MgCO_3$; (*ii*) full $MgCO_3$ conversion to MgO and CO_2; (*iii*) formation of CaS from $CaCO_3$ and part of sulfur. All these reactions are endothermic: for 1 percent S in coal, with HHV = 25.7 MJ/kg, they require a 0.63 percent of coal HHV, assuming that 75 percent of S reacts into CaS. This heat requirement can be simulated by an additional "equivalent" heat loss of such amount.

Table 5 Performance and main characteristics of entrained-bed IGCC cycles. "Syngas fraction" is the ratio of thermal power obtainable by syngas combustion and the coal thermal input (both LHV). "Steam fraction" is the ratio between the steam flow produced in syngas coolers and the total steam flowing in the HP turbine.

IGCC plant		HGCU	CGCU
Net LHV efficiency	%	46.68	45.58
Gas turbine power	MW	167.1	164.3
Steam turbine power	MW	136.5	133.6
Compressor air flow	kg/s	448.9	434.7
Air flow to ASU	kg/s	85.0	85.0
Clean syngas flow	kg/s	49.13	49.12
LHV clean syngas	kJ/kg	9427	9430
LHV syngas+nitrogen	kJ/kg	4244	4245
Syngas fraction	%	77.19	77.20
Steam fraction	%	61.26	73.02

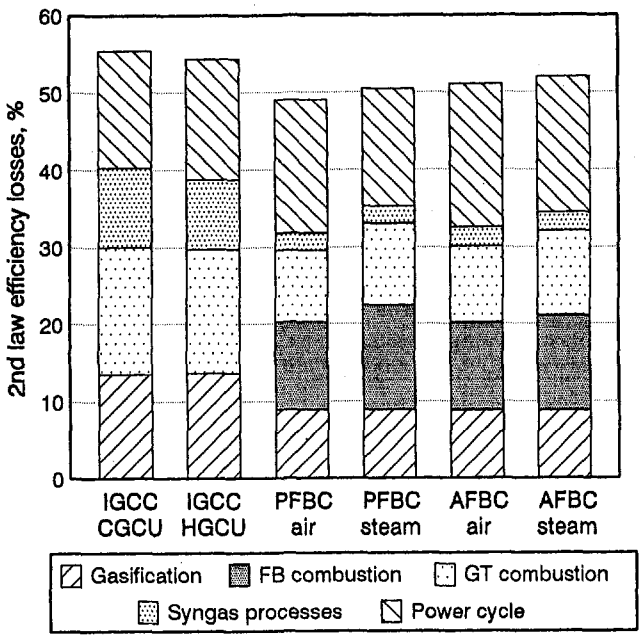


Fig. 6 Second-law losses of IGCC and PGFBC plants, at a gasifier carbon conversion ratio of 60 percent with supercritical steam cycle. The group "Syngas processes" includes the fourth to seventh rows of Table 6, while "Power cycle" includes the eighth to twelfth rows of the same table.

In a combustor, sulfur present in sorbent and char coming from the gasifier (thus containing CaS and unconverted S (25 percent)) is fully oxidized to $CaSO_4$, by exothermic reactions, producing a 1.05 percent of the HHV input by coal, under the same assumptions. Since chemical equilibrium fully converts S into SO_2 (producing 0.36 percent of input HHV) and since this reaction does not occur in FBC (considering an ideal sorbent activity), the heat balance can be corrected by introducing a "virtual" heat addition, for each percent of S, of 1.05–0.36 = 0.69 percent of input HHV.

In practice, we will simply correct the heat losses of both gasifier and FBC: Starting from an assumed "actual loss" of 0.5 percent, it becomes, with some approximations for our 0.85 percent S coal, 1 percent of input HHV for gasifier and zero for FBC (see Table 3). In this way, we will not detect the real sorbent cycle, but we will simulate a reasonably correct heat balance by simply considering the sorbent as an inert in both components.

As shown in Table 3, all calculations will refer to a coal input of 600 MW_t LHV. The coal and sorbent treatment plant (handling, grinding, lock-hoppers) has not been considered in detail, but a power consumption of 6 MW_e (1 percent of input) was introduced in all plants, to account for solid treatment and all other auxiliaries, not shown in Figs. 1–3, of the gasifier sections.

3.4.2 Gasifier and FBC. The main assumptions concerning the cycle performance are listed in Table 3. For PGFBC, it has to be noticed that the steam/coal ratio depends on gasifier type and technology: Since our work does not refer to any particular technology, the assumed value represents an average among various possible processes.

For the reference IGCC, the slurry composition and the gas temperature (imposing the oxidant flow) are taken from Matchak et al. (1984). The ASU operating pressures are derived from Rao et al. (1993). HGCU is supposed feasible at 550°C, an average value among various possible processes described by Newby and Bannister (1994); the heat loss of 67°C (120°F) is also quoted by the same authors. For CGCU, it was supposed that steam produced by the Claus furnace is sufficient for absorber regeneration and for SCOT process: This is quite realistic, even if a modest quantity of LP steam is generally supplied from the steam cycle.

All calculations were performed by assuming that chemical equilibrium well represents the outlet condition from gasifiers (see Section 2.2.2). The schematization used for reproducing gasifiers and FBCs within our calculation method is shown in Fig. 4. For the entrained bed gasifier used in the reference IGCC, the actual gasification is simulated by the reactive mixer C; the heat exchanger B simply lowers the temperature of discharged ashes, split by D, since in real plants ashes leave the gasifier at a lower temperature than raw syngas; the radiant and convective syngas coolers have been incorporated in a single heat exchanger E. The fluidized bed gasifier in PGFBC plants should have the capability of considering the partial gasification: Therefore, an imposed amount of carbon (here considered the only component of char) must be split before the reacting air–coal stream enters the mixer E, where the reactions producing syngas take place (otherwise the entire carbon content would be converted). Together with carbon, the splitter B separates ash and sorbent (which is an inert, as previously discussed): These streams are then heated by the heat exchanger D to the bed temperature. This arrangement does not reproduce the real sequence of operations, but provides the correct heat balances: In fact, considering all components A–E, they convert the entering flows into two exit flows at the bed temperature (syngas and char + ash + sorbent) as in the real plant. For a given coal input, the air flow entering the system is calculated to obtain the imposed outlet temperature of 870°C. The fluidized bed combustor is represented by a mixer, collecting char and air (or gas turbine exhausts, Fig. 3), and by a reactive heat exchanger, producing steam or hot air. Ash and spent sorbent are then separated (splitter C), as in cyclone and filter of the real plant. Heat released to coolant is calculated by imposing a bed temperature of 870°C.

3.4.3 Power Cycle. In this paper we will consider a heavy-duty gas turbine with main cycle parameters, β and TIT, representative of 200 MW_e class last generation units. The selected values of β = 15 and TIT = 1280°C have been kept constant in the analysis, since the possibility of using the same machines of gas-fired combined cycles (with some combustor modifications) is one of the key points of coal-fueled plants using gasification processes. An extended analysis, considering various pressure ratios, would be an interesting exercise, but would probably remain rather academic. On the contrary, the air flow will be changed according to the cycle requirements: In fact, a fixed coal input has been stipulated (Table 3) in order to avoid

Table 6 Second-law losses of IGCC and PGFBC considered in the paper. All PGFBC have GCCR = 60 percent and supercritical steam cycle. The group "other, gasification plant" includes thermal losses and auxiliaries consumptions (coal treatment, booster compressor, etc.). The group "other, power cycle" includes thermal, mechanical, electric losses and auxiliaries consumptions (as condenser cooling circuit pumps, etc.).

Plant configuration / Second-law loss sources	IGCC-CGCU	IGCC-HGCU	PFBC air-cooled	PFBC steam-cooled	AFBC air-cooled	AFBC steam-cooled
Gasification	13.590	13.721	8.957	8.957	8.846	8.846
Fluidized-bed combustion	-	-	11.345	13.511	11.341	12.284
Gas turbine combustion	16.496	16.075	9.270	10.596	9.918	11.004
Air separation unit	2.521	2.261	-	-	-	-
Syngas cooling	3.057	2.715	-	-	-	-
Syngas clean-up	2.590	1.885	0.490	0.490	0.483	0.483
Other, gasification plant	2.055	2.047	1.782	1.782	1.893	1.877
GT compression	2.096	2.166	2.694	1.919	2.015	1.674
GT expansion	3.626	3.651	4.338	3.269	3.412	2.934
Steam cycle and HRSG	6.250	6.792	6.953	7.017	9.782	9.706
Exhaust gas	1.280	1.152	1.302	1.109	1.420	1.300
Other, power cycle	1.917	1.937	1.953	1.831	1.962	1.883
Second-law efficiency	44.522	45.598	50.916	49.519	48.868	48.009

Table 7(*a*) Performance and main characteristics of partial-gasification cycles with pressurized FBC. "Syngas fraction" is the ratio of thermal power obtainable by syngas combustion and the coal thermal input (both LHV). "Steam fraction" is the ratio between the steam flow produced in FBHE and the total steam flowing in the HP turbine.

GCCR, %	30		40		50		60		70	
Steam pressure, bar	110	260	110	260	110	260	110	260	110	260
Results common to all PFBC schemes										
Gasifier air flow, kg/s	34.2	-	40.0	39.4	45.7	45.0	51.45	50.57	57.20	56.20
Syngas flow, kg/s	47.96	-	56.48	55.89	65.00	64.26	73.52	72.63	82.07	81.02
LHV syngas, kJ/kg	3479	-	3783	3861	4006	4094	4179	4273	4315	4416
Syngas fraction %	27.81	-	35.61	35.96	43.40	43.84	51.21	51.72	59.02	59.63
Air cooled PFBC										
Net LHV efficiency, %	-	-	-	-			51.69	52.09	51.57	51.97
GT gross power, MW	-	-	-	-	Air exiting		200.3	201.2	200.7	202.1
ST gross power, MW	-	-	-	-	FBHE exceeds		119.6	122.7	118.6	121.6
Compressor air flow, kg/s	-	-	-	-	the bed temper-		555.8	558.7	553.6	556.9
Air flow to PFBC, kg/s	-	-	-	-	ature of 870°C		104.4	104.4	78.9	78.9
Air flow to FBHE, kg/s	-	-	-	-			319.7	323	335.9	339.8
T_g exiting FBHE, °C	-	-	-	-			827	823	709	705
Steam cooled PFBC with full air feed to PFBC (PFBC-A)										
Net LHV efficiency, %	47.92		49.70	50.72	50.99	51.70	Steam produc-		-	-
GT gross power, MW	113.9	HRSG	146.5	148.1	179.0	180.7	tion in FBHE		-	-
ST gross power, MW	184.2	cannot	162.0	169.2	136.9	141.5	becomes zero.		-	-
Compressor air flow, kg/s	305.5	provide	400.1	404.7	494.8	500.2	Bed temperature		-	-
Air flow to PFBC, kg/s	223.1	BFW	300.1	304.8	377.0	382.5	of 870°C cannot		-	-
% O_2 PFBC exit	8.75	and RH	13.05	13.16	15.60	15.70	be sustained,		-	-
Steam fraction %	85.8		65.5	74.6	33.5	36.5	due to large		-	-
% O_2 in exhausts	6.14		9.40	9.52	11.44	11.53	excess air		-	-
Steam cooled PFBC with 6% O_2 at PFBC exit (PFBC-B)										
Net LHV efficiency, %	47.56		48.11	49.66	48.75	50.16	49.40	50.66	49.92	50.88
GT gross power, MW	106.2	HRSG	119.7	120.3	133.2	134.0	146.8	147.7	160.4	161.5
ST gross power, MW	189.4	cannot	179.6	191.7	169.9	180.7	160.2	169.7	149.8	156.9
Compressor air flow, kg/s	282.1	provide	320.2	321.6	357.9	359.6	395.5	397.5	432.8	435.7
Air flow to PFBC, kg/s	181.1	BFW	155.5	155.5	130.0	130.0	104.4	104.4	78.9	78.9
Air to GT combustor,kg/s	21.3	and RH	73.01	75.3	124.6	127.4	176.2	179.5	227.8	231.8
Steam fraction %	89.6		82.7	88.6	74.7	80.8	65.4	71.5	54.0	60.2
% O_2 in exhausts	5.01		6.74	6.80	8.09	8.15	9.19	9.24	10.09	10.15

Table 7(*b*) Performance and main characteristics of partial gasification cycles with atmospheric FBC. Syngas and steam fractions were defined in the previous table.

GCCR, %	30		40		50		60		70	
Steam pressure, bar	110	260	110	260	110	260	110	260	110	260
Results common to all AFBC schemes										
Gasifier air flow, kg/s	-	-	-	-	45.4	44.6	51.1	50.1	56.8	55.7
Syngas flow, kg/s	-	-	-	-	64.67	63.91	73.14	72.22	81.62	80.54
LHV syngas, kJ/kg	-	-	-	-	4041	4131	4214	4312	4352	4457
Syngas fraction %	-	-	-	-	43.55	44.00	51.37	51.90	59.20	59.83
Air cooled AFBC										
Net LHV efficiency, %	-	-			49.10	50.10	49.03	50.00	48.87	49.89
GT gross power, MW	-	-	Air exiting FBHE		159.6	160.2	160.6	161.0	161.0	161.8
ST gross power, MW	-	-	exceeds the bed		153.3	161.3	152.1	160.2	151.2	159.1
Compressor air flow, kg/s	-	-	temperature of		417.2	419.0	414.5	417.1	412.6	415.6
Air flow to FBHE, kg/s	-	-	870°C		308.1	310.5	299.1	301.1	288.5	291.7
T_g exiting FBHE, °C	-	-			745	741	609	603	454	451
Steam cooled AFBC										
Net LHV efficiency, %	43.43	45.34	45.00	46.79	46.57	48.26	47.72	49.11	48.60	49.76
GT gross power, MW	77.6	77.7	97.1	97.2	116.6	116.7	135.1	136.1	145.5	155.7
ST gross power, MW	200.1	215.3	190.1	204.4	180.2	193.6	169.1	179.6	155.8	164.3
Compressor air flow, kg/s	190.8	191.1	242.5	242.7	293.8	294.2	342.8	345.7	394.0	397.4
Air to GT combustor, kg/s	122.9	123.2	161.2	161.5	199.3	199.7	234.4	237.9	273.1	276.1
Steam fraction %	57.9	65.2	54.1	61.4	49.8	56.4	35.6	38.4	10.3	10.9
% O_2 in exhausts	6	6	6	6	6	6	7.53	7.63	9.10	9.19
Air added in FBC, kg/s	111.6	111.4	61.2	60.9	10.90	10.77	0	0	0	0

significant differences in plant size (otherwise yielding different performance of components with size-related efficiency).

The thermodynamic calculation of the gas turbine is based on the main assumptions reported in Table 4. Blade cooling is calculated by the method described by Consonni (1992), with the same hypotheses used by Macchi et al. (1994) for heavy-duty machines.

The steam cycle will assume different configurations, depending on the plant scheme and on the carbon conversion ratio. For instance, in the air-cooled scheme of Fig. 2, it is essentially

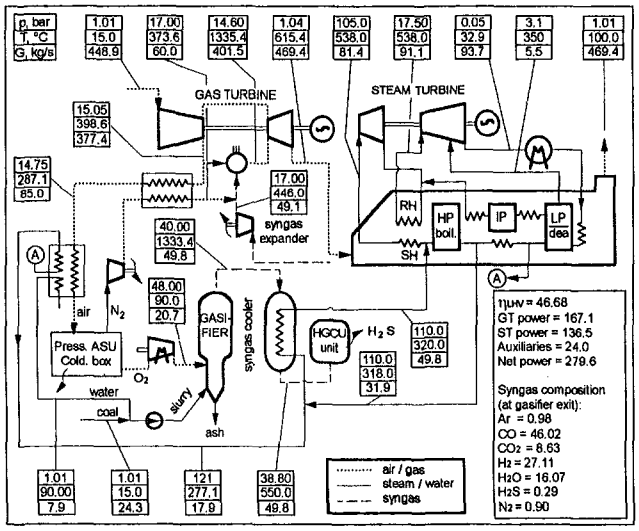


Fig. 7 Main operating parameters of the IGCC scheme with hot-gas-clean-up

a heat recovery cycle, not different from the one of gas-fired combined cycles (multipressure). As the direct heat input from coal to steam increases (i.e., steam-cooled plants at low GCCR), the cycle much more resembles a conventional fired-boiler cycle (single pressure with feedwater heaters). However, a general strategy has been stipulated: In FBC high-pressure steam is generated and superheated from saturated feedwater provided by the HRSG. In the HRSG, another fraction of superheated HP steam is generated, depending on the exhaust gas heat capacity, and feedwater heating and reheating is provided to the entire steam flowing in the turbine; if the exhaust heat capacity exceeds feedwater heating requirements, IP and LP steam is generated (as in multipressure schemes), otherwise BFW heating, for the part that cannot be supplied by exhaust heat recovery, is accomplished by regenerative bleedings from the steam turbine, as in fired plants. The general method of calculation of the steam section is exhaustively treated in Lozza (1993). The main assumptions are listed in Table 4. Two steam conditions were assumed in this paper:

(A) a subcritical single-reheat cycle (110/18 bar, 538/538°C), representing the most common conditions of advanced combined cycles,

(B) a supercritical single-reheat cycle (260/35 bar, 565/565°C), representing an advanced but widely diffused technology of conventional steam power stations.

The use of double-RH, high-supercritical, high-temperature (600–650°C) cycles, undoubtedly beneficial to efficiency (Lozza, 1993), was not considered here, to avoid solutions including further technological risks.

4 Discussion of Results

The predicted performances of IGCC and PGFBC plants described in Section 3 are shown in Fig. 5. As a first general consideration, one can say that all partial gasification schemes, even if including a conservative steam cycle or a poor GCCR, can surpass the predicted efficiency of IGCCs (in the range of 46 percent, Table 5), in some cases by several percentage points, and can reach values up to 52 percent. Efficiency losses can be seen in Fig. 6, showing the second-law analysis of IGCCs and PGFBCs at a GCCR of 60 percent, with supercritical steam cycle. On a general basis, Fig. 6 demonstrates that the main reason for the superiority of PGFBCs schemes is the abatement of losses grouped as "Syngas processes," i.e., (*i*) cooling of syngas, directly used by PGFBCs in the power cycle without highly irreversible heat transfer processes, (*ii*) low-temperature treatment and acid gas separation, (*iii*) air separation (responsible for more than 2 points loss). On the contrary, the combustion-gasification processes, even if taking place within different devices with different modes provide similar second-law losses; these losses are somewhat higher in steam-cooled FBC where combustion heat is transferred to low temperature steam.

Let us now discuss with detail the results for each cycle, by referring to Fig. 6 (showing the second-law loss breakdown), Table 6 (including the same information with higher detail) and Table 7 (reporting the main cycles characteristics, outlining the main variations due to different GCCR):

- **IGCC:** The efficiency of the HGCU scheme is 46.64 percent LHV (44.67 percent HHV). One percentage point in efficiency is gained compared to the more conventional CGCU (Table 5), because of (*i*) lower syngas cooler losses, due to its smaller duty (raw syngas exits coolers at 550°C instead of 368°C), (*ii*) elimination of all low temperature heat exchangers with related irreversibilities. However, a 1.9 percent efficiency loss is attributed to HGCU, due to thermal losses and to the loss of the heating value of removed acid gas. For both schemes, the chemical conversion of coal into syngas accomplishes about 13.6 points of efficiency loss; the gas turbine combustor is responsible for a loss of more than 16 points, mainly due to the low oxidizer temperature. It is worth noting that the sum of these two losses is very similar to the one taking place in the combustor of a natural gas-fired combined cycle (about 28 points). Figure 7 shows mass flow, pressure, and temperature of most relevant points of the HGCU cycle. From Table 5, one can notice that syngas flow and LHV are unexpectedly the same in the two cases: Water reintroduced in the saturator is almost equal to that originally enclosed in the raw syngas, not removed by hot-gas-clean-up.
- **Air-Cooled PFBC:** This is the most efficient configuration. As all other PGFBC schemes, it accomplishes a reduction of gasification losses (about 9 points versus the 13.6 of IGCC, due to the partial chemical conversion and to the lower operating temperatures) and of gas cleaning losses (only related to the discharge of hot ashes and spent sorbent), in addition to the elimination of ASU and syngas cooler losses. However, the second-law loss in the PFBC combustion is as large as 11.3 points, even if high-temperature air is raised in the process; on the other end,

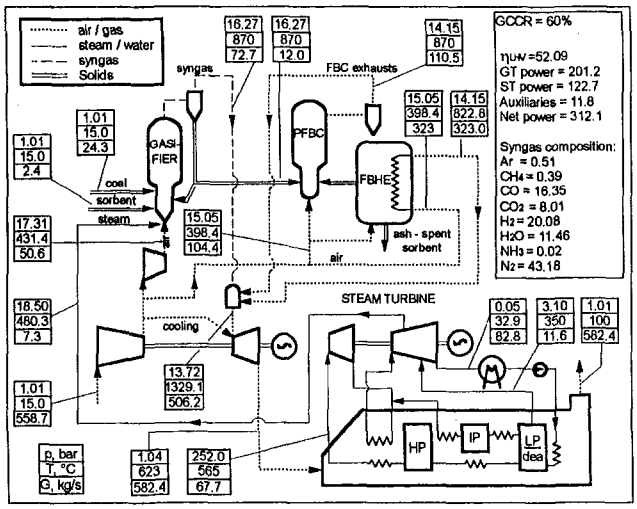


Fig. 8 Main operating parameters of the partial-gasification scheme with air-cooled pressurized FBC, at 60 percent gasifier carbon conversion ratio and supercritical steam cycle

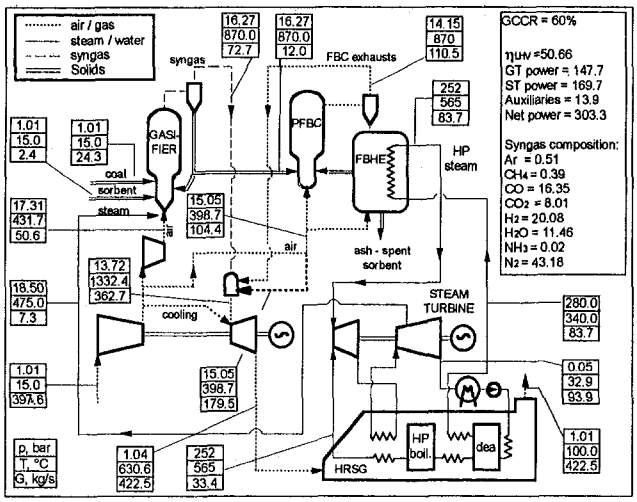


Fig. 9 Main operating parameters of the partial-gasification scheme with steam-cooled pressurized FBC (with 6 percent O_2 at FBC exit), at 60 percent carbon conversion ratio and supercritical steam cycle

high-temperature oxidizer (835°C) contributes to a dramatic reduction of the combustion loss to 9.3 points. Looking at Table 7(*a*), it can be noticed that GCCR lower than 60 percent are not viable, since air flow can no longer moderate the bed temperature, unless a reduction of the TIT (not considered here) is accepted: In this case, with a given syngas flow, with a larger air flow it becomes possible to cool the bed, but efficiency drops rapidly. A GCCR of 70 percent or higher is recommended to keep the air temperature within reasonable limits: Efficiency slightly decreases, only due to the larger steam flow required by the gasifier (from assumption of Table 3). The use of supercritical steam cycle improves efficiency by 0.4 percentage point only. The switching between the two considered steam cycles also modifies the gas turbine power output by a very small amount (Table 7*a*). This is simply due to the fact that steam ducted to the gasifier is at different temperatures, since it is bled at 18 bar before being reheated in subcritical cycle, and after reheat (and therefore at higher temperature) in the other one: This modifies the gasifier energy balance and the resulting syngas composition. A more detailed picture of the plant operating parameters at GCCR = 60 percent is given in Fig. 8.

- **Steam-Cooled PFBC With Full Air Feed to PFBC:** In this plant scheme, combustion air is fully ducted to the PFBC and heated up to 870°C. The excess heat is used for steam production. By increasing the GCCR, the air flow increases to make use of the available syngas; therefore the oxygen content of the PFBC increases, while steam production decreases, until it becomes zero at about 67 percent carbon conversion: In this condition the plant is conceptually similar to the air-cooled one, obtaining the same high efficiency. On the contrary, at low GCCR it becomes a steam-cooled one, having a large steam production and a decreasing efficiency. At 30 percent GCCR the supercritical steam cycle is not possible, since the gas turbine exhaust heat content is not sufficient for reheating (the cycle would be feasible by performing RH in the PFBC): This does not happen for the subcritical cycle, requiring a larger vaporization heat. In general, this plant arrangement, if adopted at GCCR of 50–60 percent, involves a very large air flow at the PFBC, increasing its size and cost: This drawback is removed in the following configuration, with a slight loss of efficiency.
- **Steam-Cooled PBFC With 6 percent O_2 at PFBC Exit:** This configuration presents a lower efficiency than the air-cooled one because of a larger entropy production in the PFBC, due to generation of low-temperature steam rather than air, and in the GT combustor, due to low-temperature oxidizer. At the same GCCR (50 percent of Table 6) this yields a larger loss of 3.6 percent, partly compensated by lower losses in the gas turbine, having a lower power output. It is important to notice that the adoption of a more efficient steam cycle yields efficiency improvements much larger than in previous cases: In fact, steam generation from the high-temperature environment of the PFBC, producing large irreversibilities, concerns a large part of the heat input, especially at low GCCR. The plant operating parameters at GCCR = 60 percent are given in Fig. 9. For this plant configuration, the second-law analysis was performed at different carbon conversion ratios (Fig. 10 and Table 8): By decreasing GCCR, the reduction of GT combustion and gasification losses (due to the decreasing duty of these components) is more than counterbalanced by a large improvement of combustion loss in the fluidized-bed, since heat is transferred at low-temperature steam; the power cycle losses remains approximately the same, in spite of the different gas/steam turbine output (Table 7*a*).
- **Air-Cooled AFBC:** The use of an atmospheric bed yields a loss breakdown very similar to the one with pressurized bed. The only significant difference, providing an efficiency loss of 2 points, is an higher heat transfer loss in the HRSG, in which gas enters at 870°C (bed exhausts) rather than at about 600°C (gas turbine exhausts). Despite this similarity, the plant operating parameters (Table 7*b*) are different from the ones of PFBC: The gas turbine power is lower, because the high-temperature exhausts from the fluidized-bed do not enter the gas turbine expander, but are devoted to steam production, enhancing the steam turbine output. A somewhat higher efficiency (at high GCCR) can be obtained by using the high-temperature AFBC exhausts to preheat combustion air further, rather than raising steam, but this would require a

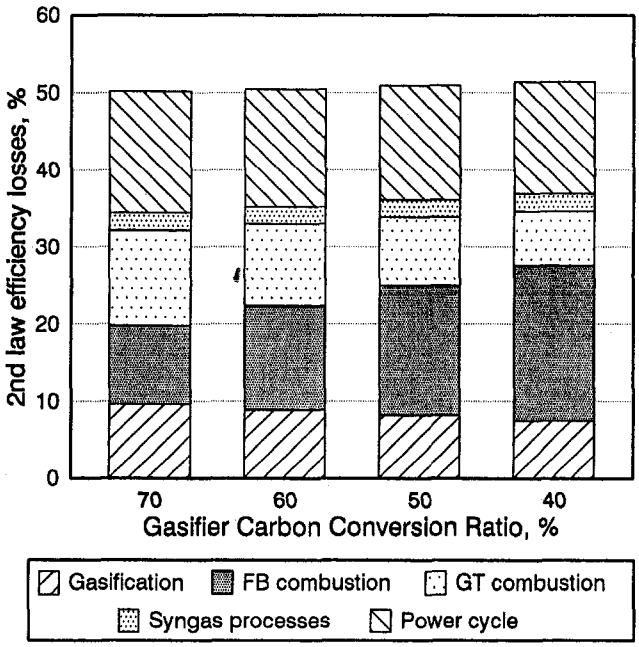


Fig. 10 Second-law losses of partial-gasification plants with steam-cooled PFBC at various carbon conversion ratios. The group "Syngas processes" represents the fourth and fifth rows of Table 8, while "Power cycle" includes the sixth to tenth rows of the same table.

Table 8 Second-law losses of PGFBC cycles with steam-cooled pressurized FBC and supercritical steam cycle, at various carbon conversion ratios. Losses are grouped as in Table 6.

Steam-cooled PFBC	Gasifier carbon conversion ratio			
2nd law loss sources	40%	50%	60%	70%
Gasification	7.526	8.243	8.957	9.673
FB combustion	20.174	16.844	13.511	10.175
GT combustion	7.011	8.804	10.596	12.389
Syngas clean-up	0.490	0.490	0.490	0.490
Other, gasification	1.772	1.776	1.782	1.789
GT compression	1.559	1.739	1.919	2.100
GT expansion	2.660	2.964	3.269	3.575
Steam cycle, HRSG	7.646	7.334	7.017	6.980
Exhaust gas	0.939	1.024	1.109	1.195
Other, power cycle	1.681	1.756	1.831	1.901
2nd law efficiency	48.542	49.026	49.519	49.733

high-temperature air-to-air heat exchanger. As in the PFBC case, the lower limit of GCCR is set by the temperature of air exiting the FBHE, approaching the bed temperature.

- **Steam-Cooled AFBC:** This is the less efficient configuration, including the drawbacks of the atmospheric solution and of a large steam production (see Table 7*b*), yielding a lower contribution of the gas turbine to the power output. However, an efficiency as high as 49 percent can be obtained at GCCR = 60 percent (Fig. 11), using technical solutions less innovative than previous schemes. Gas turbine exhausts can provide the full oxygen content required by char combustion in EBC only for GCCR exceeding 53 percent; for lower values, an addition of external air (Table 7*b*) is necessary to maintain the minimum oxygen content of 6 percent in exhausts: This rapidly decreases cycle efficiency and the gas turbine provides a very limited fraction of the total output, as in conventional PFBC cycles without gasification. In this situation, the thermodynamic convenience of using a supercritical steam cycle is enhanced.

5 Technology and Emission Problems

With the exception of steam-cooled AFBC at low GCCR, all PGFBC cycles present an efficiency consistently higher than the one of IGCC. As stated in the introduction, the plant configuration is simpler and the potential exists for capital cost reduction, compared to IGCC. However, it should be clear that these plants are at present subject to several technological hazards, while IGCC with cold-gas-clean-up is now "proven technology" (for instance, the 250 MW_e Buggenum plant in the Netherlands is in commercial operation) and HGCU is close to the demonstration phase in utility-size plants. The most challenging issue is the filtration of syngas and PFBC exhausts (including dust and alkali removal) at the bed temperature of about 870°C, by means of ceramic filters. In addition, the gas turbine combustor should be completely redesigned, as well as a fuel control system able to withstand extremely high temperatures and large mass flows. Many of the filtering problems have been encountered (and certainly solved) by ABB Carbon, which now proposes, as a commercial package, a PFBC plant (first generation, i.e., without gasification) of 800 MW_e. However, this plant makes use of a low-temperature uncooled gas turbine, rather than an advanced unit. On another side, the realization of a PFBC remains a sensitive issue, especially if air-cooled; however, commercial steam-cooled units are available (ABB Carbon, Alhstrom). The AFBC solution is simpler, because gas filtration can be operated at low temperature (i.e., before discharge to the atmosphere). The fact that important manufacturers and engineering companies (like Foster Wheeler) are devoting relevant R&D resources to PGFBC plants is probably the best evidence that quoted technological problems can be solved within a near-medium term.

As far as emissions are concerned, a 90 percent sulfur capture is a realistic figure obtainable by sorbent addition within a fluidized bed: By considering the limit on sulfur dioxide emission imposed by EEC for solid fuel plants (400 mg/Nm^3, 6 percent O_2 dry, for thermal power exceeding 500 MW), and the fact that our 0.85 percent S coal (Table 2) would produce as much as 1720 mg/Nm^3 of SO_2 without any sulfur removal, an abatement of 90 percent would enable to employ a coal with about 2 percent S content. The use of largely diffused coals with 3–4 percent sulfur content calls for sulfur retention close to 95 percent, a figure hardly obtainable by fluidized-bed combustors in real operating conditions, but largely within the capabilities of IGCC with low-temperature adsorption of H_2S. Considering also that the sorbent needs (and therefore the solid transports to and from the plant) are proportional to the coal sulfur content, it is the authors' opinion that PGFBC are a very attractive option, due to their efficiency and potential cost reduction, for low-sulfur coals (as the one considered in this paper), but IGCCs can provide unsurpassed environmental performance with high-sulfur coal (see also Moore, 1993).

The situation is of course reversed as far as CO_2 emissions are concerned, due to the higher efficiency of PGFBCs, therefore providing less CO_2 per kWh.

NO_x emissions cannot be predicted within thermodynamic analyses like the present one. However, an important parameter is the adiabatic flame temperature, resulting from stoichiometric fuel combustion, which can be correlated to the NO_2 production in a diffusion burner. Limiting the analysis to PFBC schemes (where, contrarily to AFBC, NO_x production can be reasonably concentrated in the gas turbine combustor), Table 9 shows the flame temperature of some schemes proposed in the paper, as calculated by the computer model for the gas turbine described in Fig. 3, in the hypothesis of full oxidation.

The flame temperature of a IGCC without any NO_x suppression system is consistently higher than the one achieved with natural gas, but the massive addition of nitrogen and water saturation, shown in Fig. 1, strongly reduces the flame tempera-

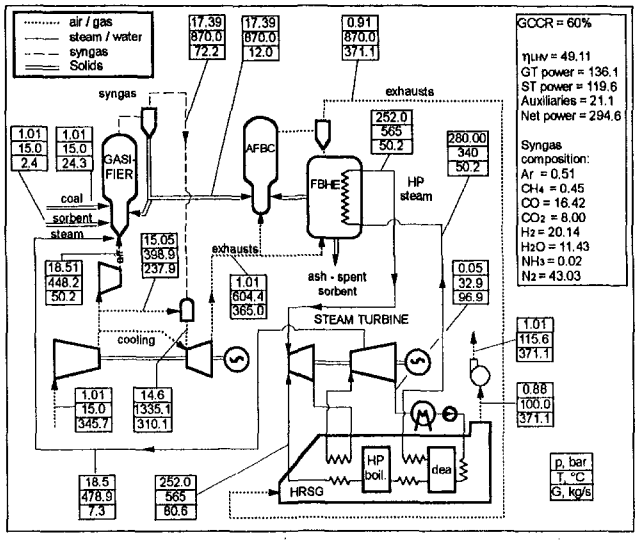


Fig. 11 Main operating parameters of the partial gasification scheme with steam-cooled atmospheric FBC: 60 percent gasifier carbon conversion ratio and supercritical steam cycle

Table 9 Flame temperature of plant configuration considered in the paper

Plant configuration	Flame temperature, K
Natural gas fired gas turbine	2526
IGCC (without water and N_2 addition)	2725
IGCC HGCU (with N_2 addition)	2185
IGCC CGCU (with water and N_2 addition)	2129
Air-cooled PFBC (GCCR=60%)	2373
Steam-cooled PFBC (free O_2, GCCR=40%)	2163
Steam-cooled PFBC (6% O_2, GCCR=40%)	1972
Steam-cooled PFBC (6% O_2, GCCR=60%)	2178

ture and thermal NO_x. For different reasons (mainly related to air-blown gasification), the PFBC plants here considered have a flame temperature consistently lower (150–400 K, depending on configuration) than natural-gas-fired gas turbines. The difference among various PFBC schemes of Table 9 is due to the oxidizer temperature (higher for the air-cooled scheme) and to the syngas heating value (Table 7*a*). The values quoted in Table 9 make feasible the realization of burners with acceptable NO_x emissions, without syngas dilution as necessary in IGCC plants.

6 Conclusions

Partial coal gasification, together with fluidized bed combustion, can achieve conversion efficiency significantly higher (up to 5–6 percentage points) than IGCC, which in turn are more efficient than systems involving low-temperature cycles, like pulverized coal steam stations and PFBC plants without gasification. Reasons for this superiority are the reduction or elimination of loss sources typical of IGCCs (air separation, syngas cooler heat transfer). Several plant configurations are possible; in the most efficient schemes, the heat of combustion (via gasification or directly in fluidized-bed combustion) must remain, as much as possible, inside the gas turbine cycle. This can be obtained by relatively high-carbon conversion ratio or by air-cooled fluidized bed.

Together with plant simplification, with respect to IGCC, and potential capital cost reduction, PGFBC plants can be considered as a very interesting solution for coal utilization, justifying considerable efforts in R&D activities on those components which are not commercially available today. This option is particularly attractive for low-sulfur coals (like the ones imported in Italy), due to acceptable emissions and sorbent consumption; however, IGCC will remain one of the most attractive technologies for very low-grade fuels, also due to superior sulfur retention capabilities.

References

Bohn, G. H., et al., 1993, "Optimizing a PFBC Combined Cycle With Gas Turbine Topping Cycle," ASME Paper No. 93-GT-390.

Chiesa, P., Consonni, S., Lozza, G., and Macchi, E., 1993, "Predicting the Ultimate Performance of Advanced Power Cycles Based on Very High Temperature Gas Turbine Engines," ASME Paper No. 93-GT-223.

Consonni, S., 1992, "Performance Prediction of Gas/Steam Cycles for Power Generation," MAE Dept. Ph.D. Thesis No. 1893-T, Princeton University, Princeton, NJ.

Consonni, S., et al., 1991, "Gas-Turbine-Based Advanced Cycles for Power Generation. Part A: Calculation Model," *Proc. 1991 Yokohama Int'l Gas Turbine Congress,* pp. III-201–210.

Dawkins, R. P., et al., 1985, "Cost and Performance of Kellogg Rust Westinghouse-Based Gasification Combined Cycle Plants," Final Report EPRI AP-4018, Project 2029-4, Electric Power Research Institute, Palo Alto, CA.

delaMora, J. A., et al., 1985, "Evaluation of the British Gas Corporation/Lurgi Slagging Gasifier in Gasification-Combined-Cycle Power Generation," Final report EPRI AP-3980, Electric Power Research Institute, Palo Alto, CA.

Foster-Pegg, R. W., and Fraas, A. P., 1993, "A Coal Fired Combined Cycle With High Efficiency," *Proc. 7th ASME Cogen Turbo Power,* ASME IGTI-Vol. 8, pp. 119–126.

Klosek, J., Sorenson, J. C., and Wong, M., 1993, "Air vs Oxygen for Gasification Combined Cycle Power," in: *Buggenum IGCC Demonstration Plant,* VGB Lecture TB216, edited by VGB, Essen, Germany.

Lozza, G., 1990, "Bottoming Steam Cycles for Combined Gas–Steam Power Plants: a Theoretical Estimation of Steam Turbine Performance and Cycle Analysis," *Proc. 4th ASME Cogen-Turbo,* pp. 83–92.

Lozza, G., 1993, "Steam Cycles for Large-Size High-Gas-Temperature Combined Cycles," *Proc. 7th ASME Cogen-Turbo Power,* pp. 435–444.

Lozza, G., Chiesa, P., and Consonni, S., 1993, "Theoretical Modelization of Integrated Coal Gasification Combined Cycles," *Proc. of the VII National Congress Combined Cycles: Technical and Economic Perspectives,* pp. 287–312 [in Italian].

Macchi, E., et al., 1991, "Gas-Turbine-Based Advanced Cycles for Power Generation. Part A: Calculation Model; Part B: Performance Analysis of Selected Configurations," *Proc. 1991 Yokohama International Gas Turbine Congress,* pp. III-201–219.

Macchi, E., et al., 1994, "An Assessment of the Thermodynamic Performance of Mixed Gas–Steam Cycles. Part A: Intercooled and Steam-Injected Cycles; Part B: Water-Injected and HAT Cycles," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 117, pp. 489–508.

Matchak, T. A., et al., 1984, "Cost and Performance for Commercial Applications of Texaco-Based Gasification-Combined-Cycle Plants," Final report EPRI AP-3486, Electric Power Research Institute, Palo Alto, CA.

Maude, C., 1993, "Advanced Power Generation—A Comparative Study of Design Options for Coal," published by IEA Coal Research, IEACR/55, London, United Kingdom.

Minchener, A. J., et al., 1993, "Development of the CRE Gasifier for Use in Integrated Partial Gasification Combined Cycle Systems," *Proc. of the Twelfth EPRI Conference on Gasification Power Plants,* San Francisco, CA.

Moore, T., 1993, "Brighter Future for PFBC," *EPRI Journal,* Dec.

Newby, R. A., and Bannister, R. L., 1994, "Advanced Hot Gas Cleaning System for Coal Gasification Processes," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 116, pp. 338–344.

Oganowski, G., 1987, "LM5000 and LM2500 Steam Injection Gas Turbines," *Proc. 2nd Tokyo Int'l Gas Turbine Congress,* pp. III-393–397.

Rao, A. D., et al., 1993, "HP Integrated vs. Non-integrated ASU in Destec IGCC," *Proc. of the Twelfth EPRI Conference on Gasification Power Plants,* San Francisco, CA.

Reynolds, W. C., 1986, "The Element Potential Method for Chemical Equilibrium Analysis: Implementation in the Interactive Program STANJAN—version 3," Dept. of Mech. Eng., Stanford University, Stanford, CA.

Robertson, A., et al., 1993, "The Advanced PFB Process: Pilot Plant Results and Design Studies," *Proc. of the Twelfth EPRI Conference on Gasification Power Plants,* San Francisco, CA.

Robertson, A., and Bonk, D., 1994, "Effect of Pressure on Second-Generation Pressurized Fluidized Bed Combustion Plants," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 116, pp. 345–351.

Soothill, C., and Brown, D., 1993, "The Department of Trade and Industry's Topping Cycle Working Party—Final Report," Contractor Report No. Coal R011 to the British Dept. of Trade and Industry.

Webb, H. A., Parsons, E. L., and Bajura, R. A., 1993, "Advanced Turbine Systems Program and Coal Application," ASME Paper No. 93-GT-356.

Feasibility Study on Oil Droplet Flow Investigations Inside Aero Engine Bearing Chambers—PDPA Techniques in Combination With Numerical Approaches

A. Glahn

M. Kurreck

M. Willmann

S. Wittig

Lehrstuhl und Institut für Thermische Strömungsmaschinen, Universität Karlsruhe (T.H.), 76128 Karlsruhe, Federal Republic of Germany

The present paper deals with oil droplet flow phenomena in aero engine bearing chambers. An experimental investigation of droplet sizes and velocities utilizing a Phase Doppler Particle Analyzer (PDPA) has been performed for the first time in bearing chamber atmospheres under real engine conditions. Influences of high rotational speeds are discussed for individual droplet size classes. Although this is an important contribution to a better understanding of the droplet flow impact on secondary air/oil system performance, an analysis of the droplet flow behavior requires an incorporation of numerical methods because detailed measurements as performed here suffer from both strong spatial limitations with respect to the optical accessibility in real engine applications and constraints due to the extremely time-consuming nature of an experimental flow field analysis. Therefore, further analysis is based on numerical methods. Droplets characterized within the experiments are exposed to the flow field of the gaseous phase predicted by use of our well-known CFD code EPOS. The droplet trajectories and velocities are calculated within a Lagrangian frame of reference by forward numerical integration of the particle momentum equation. This paper has been initiated rather to show a successful method of bearing chamber droplet flow analysis by a combination of droplet sizing techniques and numerical approaches than to present field values as a function of all operating parameters. However, a first insight into the complex droplet flow phenomena is given and specific problems in bearing chamber heat transfer are related to the droplet flow.

Introduction

Secondary air and lubrication oil systems have a decisive influence on performance characteristics and capabilities of modern jet engines. Therefore, to reduce performance losses, air has to be taken as early as possible from the compressor and has to be expelled back into the mainstream at the highest possible pressure. To meet this requirement, sufficient knowledge on discharge characteristics and heat transfer phenomena of all secondary air system components, i.e., rotating orifices, labyrinth seals, and annular gaps, are necessities within the design process. In contrast to satisfying results that have been elaborated for single-phase air flows in rotating machinery (e.g., Waschka et al., 1991, 1992a, b; Wittig et al., 1995; Jakoby et al., 1994), little knowledge is available on the flow phenomena at the interface between the secondary air and the lubrication oil system, namely the bearing chamber.

Based on the fact that rolling element bearings cannot be substituted in aero engines within the near future, special effort has to be directed to the problem of heat generation due to bearing friction and, therefore, these bearings have to be lubricated by oil. As a consequence, chambers have to be built separating the lubricant from the hot zones of the engine in order to prevent oil fires. Bearing chamber seals (e.g., labyrinth or brush seals) have to be pressurized by air, which is supplied by the secondary air system described above. Entering the chambers, the air mixes with the lubrication oil generating extremely complex two-phase flow conditions. Finally, the air is discharged through vent lines, carrying a certain amount of oil with it.

Designing the flow elements upstream and downstream of bearing chambers is a highly iterative procedure due to several two-phase flow phenomena involved in this process. Problems arise due to uncertainties in the theoretical description as well as due to a lack of data on these flow conditions. Besides a consideration of interactions between the oil film at the housing walls and the turbulent gas phase in the core of bearing chambers (Glahn and Wittig, 1995), special emphasis has to be directed to a better insight into effects of droplet flow. Evidently, droplet flow can cause a significant enhancement of local heat transfer in bearing chambers as reported by Wittig et al. (1994) and, therefore, a detailed analysis of heat transfer phenomena has to include investigations of droplet size, velocity, and flow rate. In addition, these quantities are necessities for the calculation of equivalent two-phase densities and separation mechanisms required for the proper matching of all air and oil flows linked to bearing chambers (Zimmermann et al., 1991).

Unfortunately, obstacles to improvements are given by the severe spatial limitations, which make any application of mea-

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 16, 1995. Paper No. 95-GT-100. Associate Technical Editor: C. J. Russo.

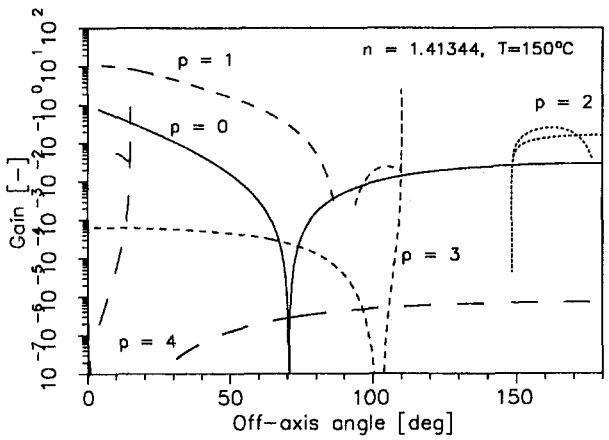


Fig. 1 Modes of reflection ($P = 0$) and refraction ($P = 1$–4)

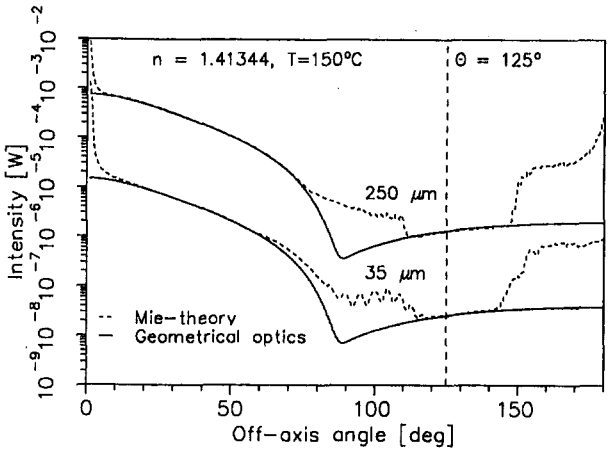


Fig. 2 Scattered light intensities for two diameters

suring techniques for a determination of the above-mentioned quantities nearly impossible.

In the present study, considerable experience of opto-electronic diagnostic has been used to develop a new PDPA setup, which allows for the first time investigations of droplet size and velocity in bearing chamber atmospheres. Measurements have been performed for real engine conditions simulated at our test rig up to high rotational speeds of n_W = 16,000 rpm. Although this is an important contribution to a better understanding of the flow phenomena described above, a complete flow field analysis cannot be performed by this approach due to restrictions with respect to the accessibility to the bearing chamber. However, measured data can be used as initial conditions for a computational investigation of droplet/gas interactions. The present paper has been initiated as a feasibility study on such an approach. Therefore, the gaseous phase of the two-phase flow has been predicted by our Finite-Volume code EPOS (Kurreck and Wittig, 1994). In addition, droplet trajectories calculated within a Lagrangian frame of reference are visualized in order to intensify the knowledge of the physical phenomena in bearing chambers.

Experimental Setup and Measuring Techniques

PDPA Setup. Simultaneous measurements of droplet size and velocity have been performed using a Phase Doppler Particle Analyzer (PDPA) as one of the few measuring devices capable of this task. This technique, described in detail by Bachalo and Houser (1984) and Saffmann (1989) and discussed with respect to its resolution and accuracy by Bachalo and Sankar (1994), measures the distance of scattered fringe patterns, which are generated by single droplets passing the measuring volume in the intersection of two laser beams. By use of two or more photomultipliers separated by fixed spacings, the fringe patterns of droplets moving through the beam intersection region are analyzed by use of the phase shift between the signals of the detectors. This phase shift shows for each scattering mode (i.e., reflection, first and higher orders of refraction) a linear relationship to the droplet diameter. However, for successful adaptation it has to be checked whether one of the modes is dominant at the scattering angle chosen for the measurements. Otherwise, the relationship between phase shift and the droplet diameter cannot be used for the sizing of particles due to superimposition effects. Therefore, scattered light intensities of a droplet due to reflection and four orders of refraction have been calculated as a function of the scattering or off-axis angle, respectively, using geometric optics (van der Hulst, 1981). Results are presented in Fig. 1 in terms of normalized values, which have been based on the intensities of an isotropical scattering particle of the same size. For the selected scattering angle, only the refractive index n for the liquid has to be taken into account. It can be concluded that two regions are suitable for the measurements. In the forward scatter region of 30 deg $\leq \theta \leq$ 80 deg first-order refraction ($p = 1$) is the dominant mode. The intensity of reflected light ($p = 0$) is two to three orders of magnitude lower. This is well known and nearly independent of the refractive index. Unfortunately, due to the geometric constraints given by the rig, this scatter configuration cannot be used here. However, a second region in the backscatter range of 120 deg $\leq \theta \leq$ 140 deg showing a dominant mode of reflection seems to be suitable too. This observation is confirmed by additional calculations of the scattered light intensity considering the superimposition of all modes described above. Additional scattering modes have been included by a comparison with calculations using Mie theory (van der Hulst, 1981). Examples are given in Fig. 2 for two droplet diameters. Without going into all details, some points should be highlighted. It can be seen that the backscatter region identified above is not significantly affected by other modes than reflection. A proper off-axis angle suitable for the bearing chamber rig adaption is given where geometric optics and Mie theory show the best fit. Therefore, an angle of θ = 130 deg should be selected. However, due to the constraints given by the rig, this angle has to be slightly shifted to an off-axis angle of θ = 125 deg. It can also be concluded that measurements in the backscatter region suffer from low intensities three orders of magnitude lower than in the forward scatter region, which has to be considered in the discussion of measured diameters with respect to the lower limit of the distributions.

Bearing Chamber Rig Adaptation. As mentioned above, the test rig developed at our institute for the investigation of air/oil flow and heat transfer phenomena in bearing chambers

Nomenclature

d = diameter, m
$\dot{m}_G$ = mass flow, air, kg/s
n = refractive index
n_W = rotational speed, min^{-1}
p = mode of reflection/refraction
r, ϕ, z = cylindrical coordinates, m, deg, m
R_0 = housing inner radius, m
T = temperature, K
U = circumferential velocity, m/s
U_s = shaft surface velocity, m/s
$\dot{V}_L$ = volume flow oil, L/h
Θ = off-axis angle, deg
Θ^* = backscatter angle, deg

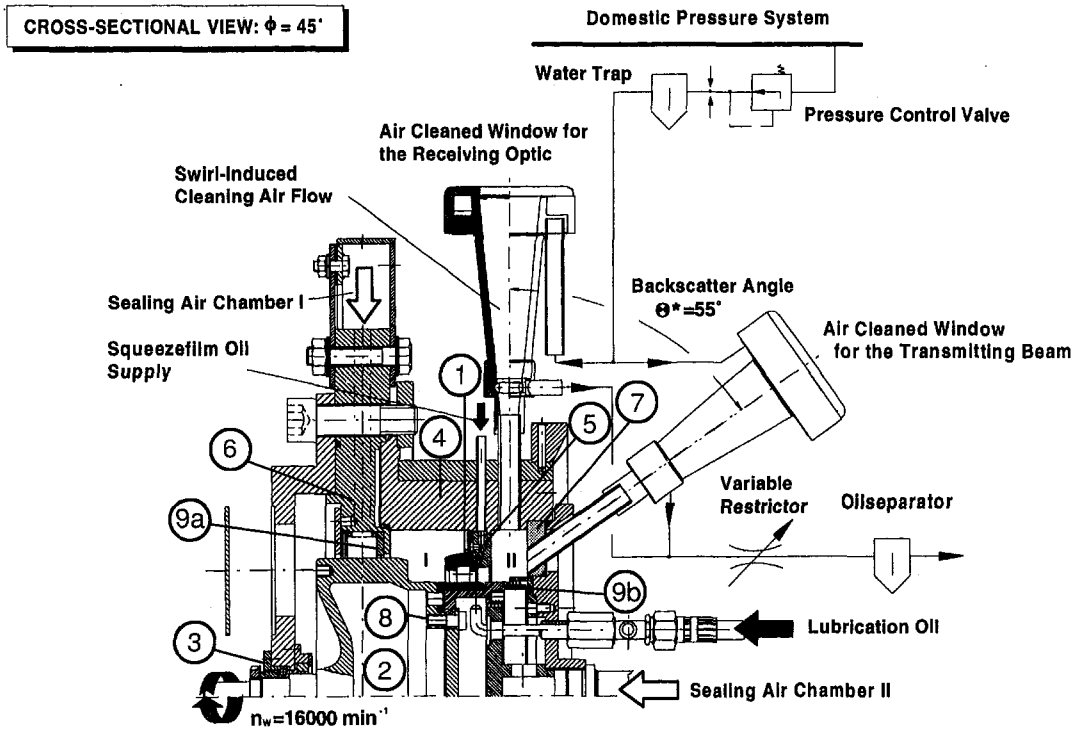


Fig. 3 High-speed bearing chamber test rig

has been adapted to the present investigation. Design considerations and a description of all test facility components are given by Wittig et al. (1994). However, for a better understanding of air and oil flows inside the bearing chamber rig, a short description of the general arrangement is presented below. A sectional view of the test rig is shown in Fig. 3 for a co-axial plane located at a circumferential position of $\phi = 45$ deg relative to the vertical.

At each side of the squeeze-film-damped roller bearing (1, diameter $d_1 = 124$ mm), separate chambers (I, II) are formed. The axial location of the rotor (2) is realized by use of a conventional groove-ball bearing (3). The chambers are bounded by a thick-walled housing (4), the roller bearing support (5), the rotor, a flange (6) realizing the sealing air supply of chamber I and the support of the housing, and a transparent cover (7) for chamber II representing the optical access for a qualitative flow analysis. Although the bearing chamber geometries of our test rig have been abstracted, air and oil flows have been arranged in the same way as in the real engine. An under-race lubrication jet (8) supplies the roller bearing with preheated oil. To prevent oil leakage, the chambers are sealed using three-fin labyrinth seals (9*a*, *b*). Both labyrinth seals are pressurized by air, which is preheated to the same temperature level as the lubrication oil. Air/oil mixtures are discharged through the vent lines at the top of each chamber, while the oil collected at the bottom is dropped out via radial scavenges utilizing the pressure difference from the bearing chamber atmosphere to the environment. Thus, another simplification is given to a real engine application, where scavenge pumps are used. However, droplet flow investigations have been performed in the upper part of the chamber at a position close to the bearing and, therefore, the results gained with the setup will be independent of specific scavenge port arrangements and transferable to other applications using the same bearing. The adaptation of the measuring technique is reflected by the optical access for the transmitting and the receiving beams, respectively, mounted at an angle of $\theta^* = 55$ deg (referring to an off-axis angle of $\theta = 125$ deg) to each other. The setup is provided to chamber II at a circumferential position of $\theta = 45$ deg (see Fig. 9) and the measurements described in the next section have been conducted at a radial position of $r = 65.25$ mm and at an axial distance of $\Delta z = 4.1$ mm from the bearing support.

Experimental Results

Measurements have been performed for real engine conditions as described by sealing air flows and lubrication oil flows of $\dot{m}_{G,II} = 10$ g/s and $\dot{V}_{L,II} = 100$ L/h, respectively, and rotational speeds up to $n_W = 16{,}000$ min^{-1}. An advanced data reduction software developed at our institute for the correlation of a discrete droplet class with its individual velocities has been used to extract specific data from the measured distributions. As a result, mean values for the droplet velocity components in the radial and tangential directions according to the PDPA setup described in the section above are shown for three discrete droplet diameter classes in Figs. 4 and 5 as a function of rotational speeds. It can be readily seen from the figures that the droplet flow is strongly influenced by the circumferential veloc-

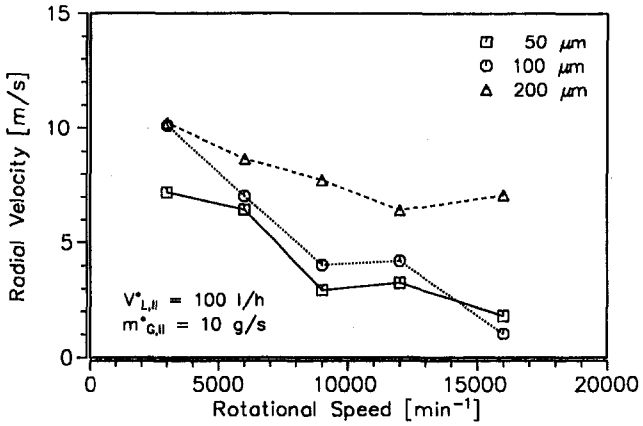


Fig. 4 Mean radial velocity versus rotational speed

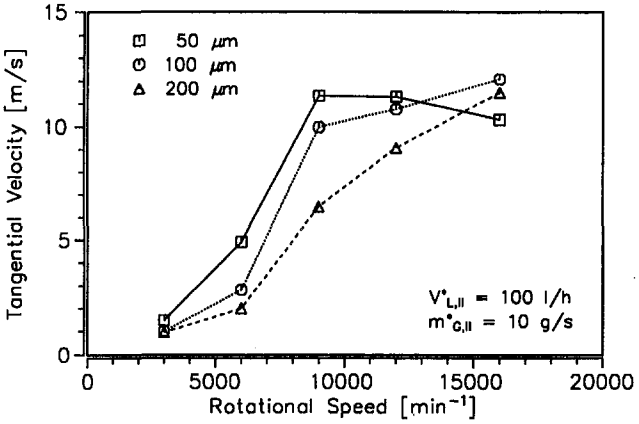


Fig. 5 Mean tangential velocity versus rotational speed

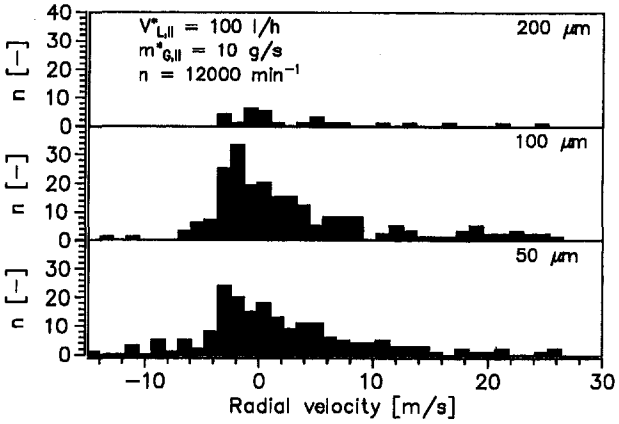


Fig. 7 Radial velocity distribution for 50 μm, 100 μm, and 200 μm droplets

ity of the rotor. In general, the radial velocity of the droplet is decreasing, whereas the tangential velocity component strongly increases with an enhancement of the shaft speed. Thus, it can be concluded that the droplet flow angle is changing from a radial to a tangential direction, reflecting a dominant effect of air flows to the vent at low speed and dominant influences of rotating core flows inside the chamber at high speed conditions, respectively. With the exception of the highest speed of n_W = 16,000 min^{-1} this is more significant for small droplets. Larger droplets are accelerated by the rotation, but keep a radial velocity component even for high-speed conditions due to their inertia.

An analysis of droplet flow phenomena in combination with numerical methods requires information on each individual droplet characterized by its diameter and velocities. In the present paper we have investigated in detail the operating condition shown in Fig. 6 where the droplet diameter histogram for $\dot{m}_{G,II}$ = 10 g/s and $\dot{V}_{L,II}$ = 100 L/h at n_W = 12,000 min^{-1} is given. It represents a total number of 3000 valid samples and ranges from approximately 14 μm to 500 μm in diameter. The number distribution shows a maximum at diameters from 70 μm to 80 μm. However, it can also be seen that larger droplets, i.e., ≥ 200 μm, must not be neglected. From the measured diameter distribution, droplet classes referring to those discussed above (dark bins) have been separated and the velocity distributions for both, the radial (Fig. 7) and the tangential (Fig. 8) components, have been extracted from the raw data. Although the data are presented in two-dimensional histograms for the sake of clarity, it has to be pointed out that the initially given correlation of diameter and two velocity components is still available for each individual droplet.

By use of these figures, it can be concluded that the mean values discussed above result from a comparatively broad distribution and, therefore, the necessity of a detailed analysis is pointed out again. In addition, the problem of statistical approaches to complex two-phase flow patterns is demonstrated in the light of low number densities for large droplets. Individual data for the 200 μm droplets given in the upper part of Fig. 7 and Fig. 8, respectively, show clearly that mean values can only be used for a first impression of the flow behavior. As an example, the radial velocity component in Fig. 7 is similarly distributed for all droplet sizes, but the mean value for the 200 μm droplets is shifted to a higher value due to a few droplets with high speeds and the low number of samples in this size class.

It will be demonstrated in the next section that any further analysis has to be based on the detailed characterization of each individual droplet because mean values would not represent the flow conditions correctly. Thus, numerical methods have to be included into an investigation of specific bearing chamber droplet flow phenomena.

Numerical Investigation

Since the experimental investigation of bearing chamber flows requires tremendous efforts, the numerical prediction of the flow pattern and its visualization offers the opportunity to gain a better understanding of the fundamental physical phenomena taking place. The aim of this numerical study is rather the qualitative visualization of the flow regime than the prediction of quantitative results. The gaseous phase of the two-phase flow in the bearing chamber is described by an Eulerian ap-

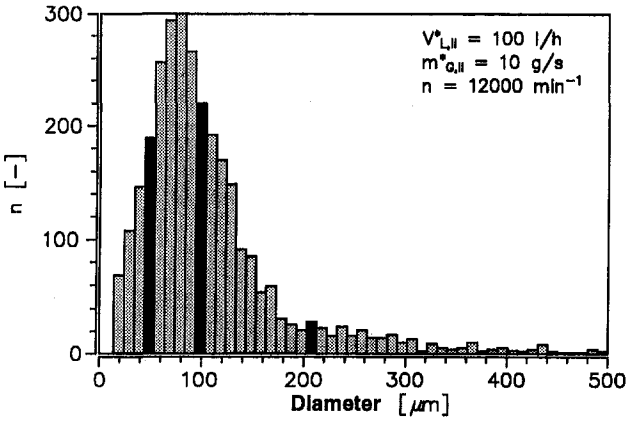


Fig. 6 Diameter histogram at n_W = 12,000 rpm

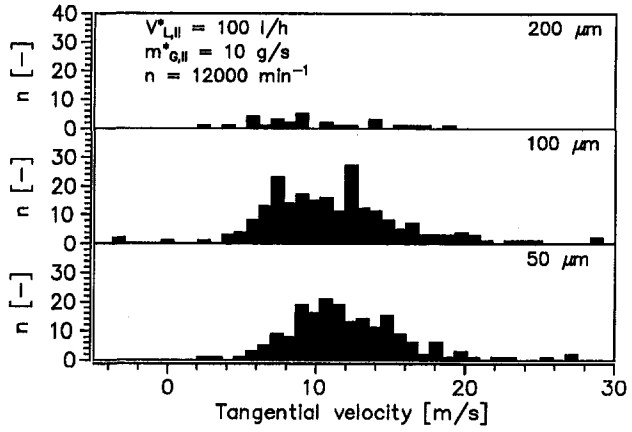


Fig. 8 Tangential velocity distribution for 50 μm, 100 μm, and 200 μm droplets

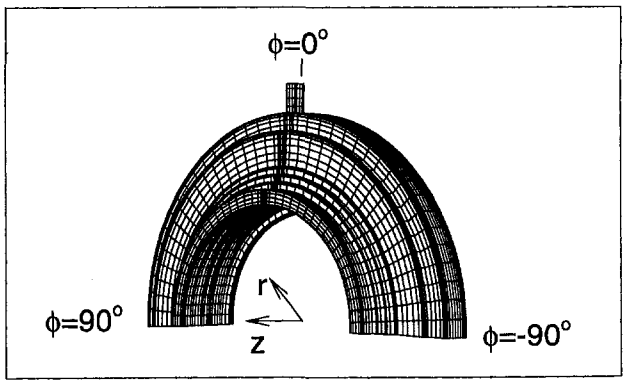


Fig. 9 Three-dimensional grid (31 × 44 × 23 ($r/\phi/z$) grid points)

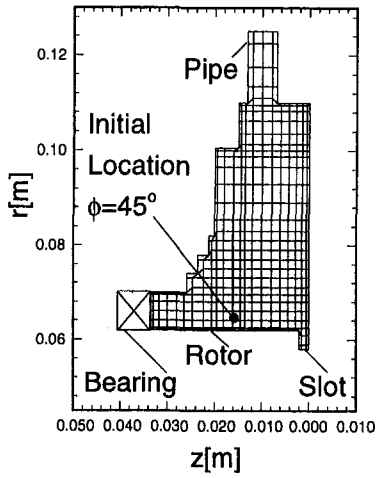


Fig. 10 Cross section of the computational grid (ϕ = 0 deg)

proach, whereas the dispersed phase is calculated within a Lagrangian frame of reference. The repercussion on the gaseous phase caused by the liquid phase is neglected. The same holds for the distortion of droplets, i.e., the droplets are considered as spherical nonevaporating particles and wall films are not taken into account. Because of these assumptions, this numerical study represents a first step for the detailed theoretical description of the two-phase flow in bearing chambers.

Gaseous Phase. The gaseous phase is predicted by applying the CFD code EPOS developed at the Institut für Thermische Strömungsmaschinen. The governing equations describing conservation of mass, momentum, and enthalpy are discretized by a Finite-Volume method on nonstaggered grids in Cartesian, cylindrical, or body-fitted coordinates. The SIMPLEC algorithm for the calculation of the pressure field in combination with the $k-\epsilon$ turbulence model and wall functions is applied. A high spatial accuracy and boundedness of the solution is ensured by the discretization of the convective fluxes with the MLU scheme developed by Noll (1992). All the calculations have been performed with the highly vectorized version of the code using the implicit iterative ILU-CG algorithm proposed by Noll and Wittig (1991).

The three-dimensional grid representing a model of the bearing chamber is shown in Fig. 9. The problem is described in cylindrical coordinates (31 × 44 × 23 ($r/\phi/z$) grid points), which is a compromise in terms of effort for the grid generation and accuracy for this first study of the flow. Due to cyclic boundary conditions at the planes $\phi = -90$ deg and $\phi = 90$ deg only the upper half of the chamber is discretized. The assumption of cyclic boundaries is feasible since wall films are neglected and therefore no oil collection will be found in the lower half of the chamber. Figure 10 gives an insight into the chamber representing a cross section of the computational grid at $\phi = 0$ deg. The air enters the chamber through a slot with a radial velocity of 7.74 m/s at an inlet pressure of 2 bar. A pipe at the top of the chamber ($\phi = 0$ deg) ensures parabolic outflow conditions. The bearing is replaced by a fixed wall, neglecting any influence on the flow.

The velocity field of the gaseous phase within the chamber predicted for three different rotational speeds (n_W = 7000, 12,000, 16,000 rpm) is shown in Fig. 11 at the plane ϕ = 45 deg. Due to the momentum transfer from the rotating shaft to the flow, a secondary motion is established within the chamber. The strength of the secondary flow increases as the rotational speed is increased and caused by the complex shape of the chamber several vortices are induced. At n_W = 7000 rpm neither the inlet air nor the secondary flow is dominant and two vortices of comparable size and strength are generated. As the rotation is increased to n_W = 12,000 rpm the vortex due to the secondary motion is enlarged and occupies a large portion of the chamber's cross-sectional area. Merely a small vortex can be found near the air intake. This small recirculation zone is reduced as the shaft speed is further enhanced. At n_W = 16,000 rpm the secondary flow within the chamber is intensified again.

In Fig. 12, nondimensional profiles of the circumferential velocity U versus the chamber radius at z = 10 mm and ϕ =

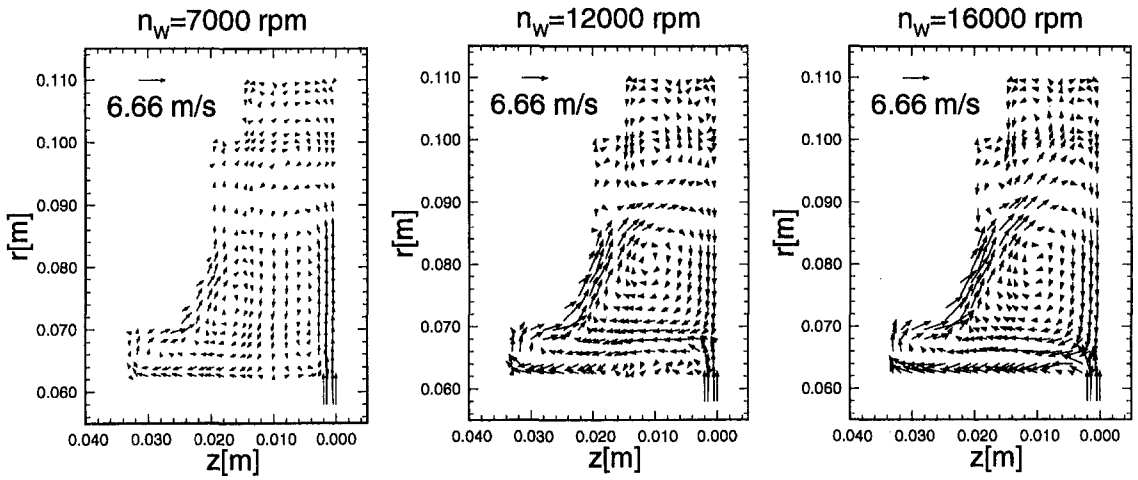


Fig. 11 Velocity distributions (ϕ = 45 deg)

45 deg are depicted. The dimensionless circumferential velocity is based on the shaft surface velocity U_s, whereas the coordinate is divided by the outer radius R_0 = 110 mm of the chamber. The profiles are quite similar for the different rotational speeds and the circumferential velocity within the chamber varies in the range of 8 to 16 percent of the shaft's surface velocity.

Liquid Phase. The liquid phase is predicted within a Lagrangian frame of reference based on the predescribed turbulent flow of the gaseous phase. A detailed description of the numerical procedure is given by Hallmann et al. (1995) and Hallmann (1994). The droplet trajectories and velocities are calculated by forward numerical integration of the particle momentum equation. Turbulence is taken into account by a stochastic approach (Gosman and Ioannides, 1983), splitting the instantaneous velocity into a mean and a fluctuating part. The fluctuating velocities are chosen randomly from an isotropic Gaussian distribution of the gas phase velocity. Evaporation of the oil is negligible for the temperatures considered in this study.

For the numerical simulation of the droplet motion a number of trajectories given by the experimental data is calculated. Each of the three drop-size classes (50, 100, and 200 μm in diameter) considered in this numerical study consists of a number of velocity classes represented by a radial and a tangential velocity component and a corresponding number of particles being measured (see Figs. 7 and 8). The Lagrangian calculations are performed taking the experimental data into account. Therefore, for the same initial condition such as velocity, location, and drop size, different trajectories are predicted due to the stochastic turbulence approach for the liquid phase. Since the axial velocity component has not been measured, this component is assumed to be zero.

The radial and axial initial location of the droplets within the chamber is shown in Fig. 10. The droplets start at a position corresponding to the location of the PDPA measuring volume described above. Predicted trajectories of the three classes of droplets are given in Fig. 13. Due to the initial conditions and the rotating flow within the chamber, the paths of the droplets are directed toward the outer radial wall of the chamber. Since droplets with diameters of 50, 100, and 200 μm are considered, the influence of the turbulence of the gaseous phase is relatively small as can be seen from the straight trajectories. The spreading of the trajectories is mainly caused by the initial conditions of the droplet velocities. Small droplets with 50 μm in diameter are more influenced by the gaseous phase than the larger ones. Therefore, the curvature of the trajectories increases as the droplet diameter decreases.

The volume fraction of the dispersed phase, represented by the dark gray areas (see Fig. 13), is taken as an indicator for the droplet distribution at the radial plane at r = 100 mm. Some of the particles are reaching the radial plane (r = 100 mm),

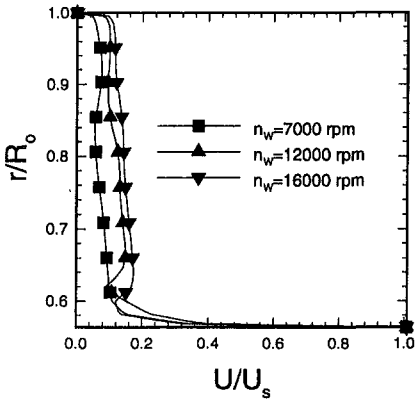


Fig. 12 Nondimensional profiles of the circumferential velocity component U (ϕ = 45 deg, z = 10 mm)

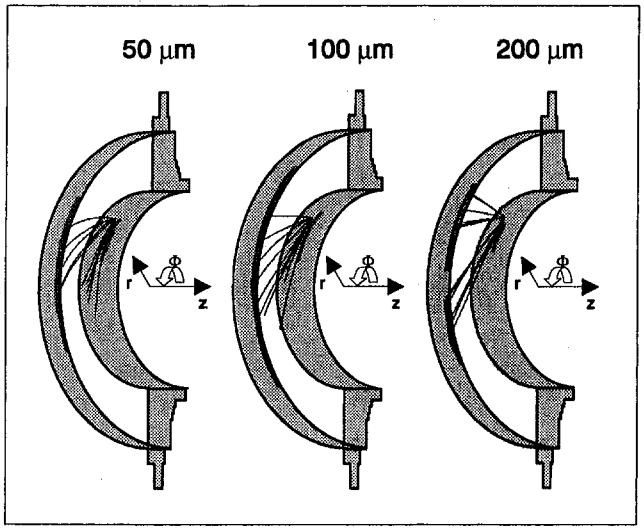


Fig. 13 Trajectories and volume fraction of the liquid phase (planes r = 62.5 mm and r = 100 mm)

whereas especially the small droplets are driven toward the vertical wall of the housing (z = 20 mm, Fig. 10). This leads to a wall oil film, which strongly influences the heat flux through the chamber walls. The process of transferring thermal energy generated within the bearing with the droplets to the outer wall of the chamber which has been reported by Wittig et al. (1994) is visualized by our "numerical test rig."

Conclusion

An innovative method of bearing chamber droplet flow analysis has been presented utilizing advanced optoelectronic techniques in combination with numerical approaches on two-phase flow phenomena. Results of the numerical simulation of two-phase flow in aero engine bearing chambers are presented for the first time. Although several simplifications have been made, reasonable results have been obtained. Therefore, a first insight into the complex two-phase flow can be given and the methods presented here are thought to be a valuable contribution to overcoming uncertainties in the design of bearing chamber and vent pipe systems due to the impact of the droplet flow behavior.

Acknowledgments

We wish to thank Rolls-Royce plc, Bristol and Motoren- und Turbinen-Union, München for funding the experiments documented in this paper. Special thanks are due to R. M. Coleman and H. Zimmermann for helpful discussions.

References

Bachalo, W. D., and Houser, M. J., 1984, "Phase Doppler Spray Analyzer for Simultaneous Measurements of Drop Size and Velocity Distributions," *Optical Engineering,* Vol. 23, pp. 583–590.

Bachalo, W. D., and Sankar, S. V., 1994, "Factors Affecting the Measurement Resolution and Accuracy of the Phase Doppler Particle Analyzer," presented at The Second International Conference on Fluid Dynamics Measurement and Its Applications, Beijing, China.

Glahn, A., and Wittig, S., 1995, "Two-Phase Air/Oil Flow in Aero Engine Bearing Chambers—Characterization of Oil Film Flows," ASME Paper No. 95-GT-114; ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 118, 1996, pp. 578–583.

Gosman, A. D., and Ioannides, E., 1983, "Aspects of Computer Simulation of Liquid-Fueled Combustors," *Journal of Energy,* Vol. 7.

Hallmann, M., 1994, "Numerische Untersuchung der Gemischbildung in Verbrennungskraftmaschinen," Dissertation, Institut für Thermische Strömungsmaschinen, Universität Karlsruhe, Germany.

Hallmann, M., Scheurlen, M., and Wittig, S., 1995, "Computation of Turbulent Evaporating Sprays: Eulerian Versus Lagrangian Approach," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 117, pp. 112–119.

Jakoby, R., Benz, E., Willmann, M., and Wittig, S., 1994, "Phase Resolved LDV-Measurements in High-Speed Rotating Disk Systems With Orifices," *Proc. Fifth International Symposium on Transport Phenomena and Dynamics of Rotating Machinery,* Honolulu, HI, May 8–11, Vol. A, pp. 573–591.

Kurreck, M., and Wittig, S., 1994, "Numerical Simulation of Combustor Flows on Parallel Computers—Potential, Limitations and Practical Experience," ASME Paper No. 94-GT-404.

Noll, B., and Wittig, S., 1991, "Generalized Conjugate Gradient Method for the Efficient Solution of Three-Dimensional Fluid Flow Problems," *Numerical Heat Transfer,* Vol. 20, pp. 207–221.

Noll, B., 1992, "Evaluation of a Bounded High-Resolution Scheme for Combustor Flow Computations," *AIAA Journal,* Vol. 30, pp. 64–69.

Saffmann, M., 1989, "Phasen Doppler Methode zur optischen Partikelgrößenmessung," *Technisches Messen,* Vol. 56, pp. 298–303.

van der Hulst, H. C., 1981, *Light Scattering by Small Particles,* Dover, New York.

Waschka, W., Wittig, S., Scherer, Th., and Kim, S., 1991, "Leakage Loss and Heat Transfer in High-Speed Rotating Labyrinth Seals: An Experimental Verification of Numerical Codes," *Proc. 1991 Yokohama International Gas Turbine Congress,* Vol. II, pp. 239–247.

Waschka, W., Scherer, Th., Kim, S., and Wittig, S., 1992a, "Study of Heat Transfer and Leakage in High-Rotating Stepped Labyrinth Seals," presented at the 4th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery, Honolulu, HI.

Waschka, W., Wittig, S., and Kim, S., 1992b, "Influence of High Rotational Speeds on the Heat Transfer and Discharge Coefficients in Labyrinth Seals," ASME *Journal of Turbomachinery,* Vol. 114, pp. 462–468.

Wittig, S., Glahn, A., and Himmelbach, J., 1994, "Influence of High Rotational Speeds on Heat Transfer and Oil Film Thickness in Aero Engine Bearing Chambers," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 116, pp. 395–401.

Wittig, S., Kim, S., Jakoby, R., and Weißert, I., 1995, "Experimental and Numerical Study of Orifice Discharge Coefficients in High-Speed Rotating Disks," ASME *Journal of Turbomachinery,* Vol. 118, pp. 400–407.

Zimmermann, H., Kammerer, A., Fischer, R., and Rebhahn, D., 1991, "Two-Phase Flow Correlations in Air/Oil Systems of Aero Engines," ASME Paper No. 91-GT-54.

An Experimental Investigation of the Conversion of NO to NO_2 at High Pressure

J. W. Hunderup
Department of Mechanical Engineering,
Virginia Polytechnic Institute and
State University,
Blacksburg, VA 24061-0238

R. J. Roby
Hughes Associates, Inc.,
Fire Science & Engineering,
Columbia, MD 21045

Unexpectedly high concentrations of NO_2 have been noted in stack emissions from industrial combustors. NO_2 formation has been reported to occur through the so-called "HO_2 mechanism" in which NO combines with HO_2 to produce NO_2 and OH. In this study, the formation of NO_2 was investigated at superatmospheric pressures through experiments and computer modeling. Computer modeling utilized the CHEMKIN chemical kinetics program and a subset of a previously published C–H–O–N system mechanism. Experimental work was conducted using a high-pressure flow reactor designed and built in the course of the study. The effects of pressure, temperature, and the presence of a NO_2-promoting hydrocarbon, methane, were investigated. It was discovered that as pressure increased from 1 atm to 8.5 atm, the rate and amount of NO converted to NO_2 also increased. The results also show a temperature "window" between approximately 800 K and 1000 K in which NO to NO_2 conversion readily occurred. The presence of methane was seen to enhance NO conversion to NO_2, and a ratio of $[CH_4]/[NO]$ was found to be a useful parameter in predicting NO_2 formation. Significant NO conversion to NO_2 was noted for $[CH_4]/[NO] > 1$ at the hydrocarbon injection point. Experimental results validated those trends obtained from modeling with a modified C–H–O–N mechanism.

Background and Motivation

Since the passage of the Clean Air Act of 1970, the production of major air-borne pollutants—NO_x, SO_x, CO, unburned hydrocarbons (UHCs), and particulate matter—from combustion sources has decreased significantly. However, the recent Amendments to the Clean Air Act, enacted in 1990, stipulate further reduction in combustion source emissions. The formation of oxides of nitrogen is of particular interest due to the role it plays in the formation of photochemical smog and acid rain. Of the two main constituents of NO_x (NO + NO_2), NO_2 is more toxic. It is also a gas that becomes visible (brown color) at relatively low concentrations. NO is colorless. In many practical combustion systems, NO_2 remains a small fraction of NO_x. However, it has been noticed in some systems that the ratio of NO_2/NO_x increases to a significant level, particularly as total NO_x emissions decrease. Since NO_2 becomes visible at relatively low concentrations, a small increase can lead to visible concentrations.

In nearly all investigations of NO_2 formation, concentrations of NO_2 have been determined through analyses of combustion gas samples. Hot combustion gases are typically sampled via a quartz or stainless steel probe. During the 70s, combustion gas samples were generally obtained through water-cooled probes at relatively high pressures. Hot gases were rapidly quenched as they passed through the probe. The samples were usually found to contain high levels of NO_2. In 1975, J. D. Allen questioned the reliability of "conventional probe sampling techniques." He suggested that NO_2 is formed in the probe during sampling, and, therefore, postsampling measurements would not reflect actual combustion system concentrations. Johnson et al. (1978) showed through laser-induced fluorescence that NO_2 levels measured after probe sampling do not, in fact, represent levels found in the actual combustion system. The samples contained a much higher level of NO_2. These important research findings make suspect much of the work done in the 70s in which probe effects were neglected.

Probe effects remain an unavoidable obstacle in combustion gas sampling, but they can be minimized. Conditions in the probe that have been found conducive to NO_2 formation are rapid quenching of the sample, high pressure, and the presence of unburned fuel species (Amin, 1977; Johnson et al., 1978; Kramlich and Malte, 1978; Chen et al., 1979; Hori, 1980). It is believed that much of the NO_2 is produced by the catalytic conversion of NO at the probe walls where the temperature gradient is greatest. Keeping these points in mind, probe effects can be minimized by a slow sampling rate at low pressure in an uncooled quartz probe.

Conditions conducive to NO_2 formation in the probe can also be extended to the combustion system. It has been observed that NO_2 is produced in areas of large temperature gradients in the flame. Hargraves et al. (1981) found that little or no NO_2 is formed in the main combustion region but that the NO_2/NO_x ratio is greatest along the periphery of the flame where the hot combustion gases mix with cool surrounding air. Sano (1984) confirmed this through a numerical study of NO_2 production in mixing regions. Temperature is therefore an important variable in NO_2 formation. Bromly et al. (1992), Hori et al. (1992), and Marinov et al. (1993) have all recently noted very distinct temperature limits above and/or below which little NO is converted to NO_2. Nearly complete NO conversion only appears to occur in the relatively low-temperature region below 1200 K.

The presence of unburned fuel and/or other oxidizable species, similar to the probe case, also increases the amount of NO_2 produced (Jaasma and Borman, 1980; Hargraves et al., 1981; Sano, 1985; Hori, 1988; Bromly et al., 1992; Hori et al., 1992; Marinov et al., 1993). In separate research projects, Hori et al. (1992), Marinov et al. (1993), and Bromly et al. (1992) have investigated the effect of the presence of hydrocarbons on the promotion of NO–NO_2 conversion. Hori et al. (1992) and Marinov et al. (1993) observed that higher order hydrocarbons are more effective in promoting NO–NO_2 conversion. They

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute March 16, 1995. Paper No. 95-GT-306. Associate Technical Editor: C. J. Russo.

also noted that the type of hydrocarbon has an effect on the temperature region in which $NO-NO_2$ conversion is nearly complete. Bromly et al. (1992) observed that trace amounts of n-butane greatly enhanced NO_2 formation in a similar temperature region.

It is widely believed that the majority of NO_2 is produced from the conversion of NO by the HO_2 radical:

$$NO + HO_2 \rightarrow NO_2 + OH.$$

Increased concentrations of HO_2 would then result in increased conversion of NO to NO_2. One important pathway that has been suggested for HO_2 formation (Chen et al., 1979; Hargraves et al., 1981; Hori, 1986; Miller and Bowman, 1989) is

$$H + O_2 + M \rightarrow HO_2 + M.$$

Because this is a three-body reaction, it would be particularly sensitive to pressure. It is also a reaction that is favored at relatively low temperatures. In turn, the formation of NO_2 would then also be pressure and temperature sensitive. This has been seen in probe and combustion system studies. In probes, it has been observed that relatively high sampling pressures and rapid quenching have a substantial effect on NO_2 formation. Although research on combustion systems has shown that temperature and the presence of unburned hydrocarbons has an important effect on NO_2 formation, no studies prior to this one have investigated the effect of superatmospheric pressure on NO_2 formation.

The goal of this study was to determine under what conditions NO may be converted to NO_2 in a high-pressure environment similar to that of an industrial gas turbine combustor. In achieving this, the effect of pressure on the $NO-NO_2$ conversion was of particular interest. The objective was achieved through computer modeling and experimental work. The modeling offered a basis of comparison for those results achieved through experimental study. Modeling may also prove to be a useful tool in the "field" for predicting incidences of high $NO-NO_2$ conversion.

Experimental Setup

Experimental testing took place in a high-pressure flow reactor. A schematic of the reactor facility is shown in Fig. 1. Figure 2 shows a system schematic of the experimental facility. The high-pressure flow reactor design can be divided into several areas: (1) the pressure vessel, (2) the burner and related components, and (3) instrumentation and control. These will each be discussed below. Further details are given by Hunderup (1993).

The Pressure Vessel. The vessel is based on a similar design developed by Carter et al. (1989) at Purdue University. In order to simulate the environment of a high-pressure combustor, the pressure vessel was designed to withstand working pressures up to 10 atm (1.013 MPa gage). Blind flanges were machined for access to the interior of the vessel. The bottom flange contains all inlet ports as well as a water trap assembly and electrical/instrumentation feed-throughs. Inside the vessel, polypropylene tubing conducts the air/fuel mixture to the burner. Tygon tubing provides passage to the injector inlet. The injector does not extend below the bottom flange. Extending off the water trap tube are a relief valve and pressure transducer. The top flange is of somewhat simpler design than the lower flange. It houses the sampling probe feed-through and a Velmex translation stage, which provides for accurate positioning of the probe tip within the sealed vessel. Offset slightly from center is the exhaust port. The top flange also houses a Fike rupture disk fitting.

The pressure vessel, which was designed to ASME 300 psia pressure standards, was hydrostatically tested to 18.3 atm (1.86 MPa). The vessel integrity remained intact. The vessel is very similar to the design of Carter et al., which was successfully tested hydrostatically to 60 atm. (6.08 MPa) (Carter et al., 1989). Although this vessel was only tested to 18.3 atm (1.86 MPa), it is believed that it could withstand much greater pressures without compromising integrity of the welds and seals.

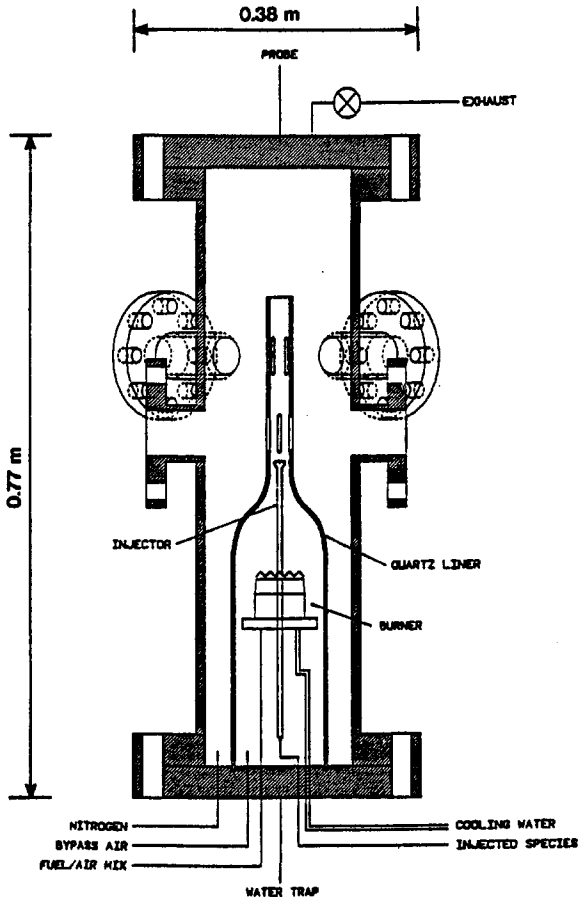


Fig. 1 High-pressure flow reactor schematic

The Burner and Related Components. The phenomenon of $NO-NO_2$ conversion appears to be a postflame event due to the relatively low temperatures required for it to occur. For the scope of this study, the burner was solely a source of postflame gases. The burner used for this purpose was a McKenna sintered bronze, flat flame burner that had been specially modified to allow injection of a hydrocarbon promoter into the post flame gases. A 6.35 mm hole had been bored through the center of the burner. This allowed passage of a quartz injector. Flame ignition was initiated by throwing an arc between a nichrome wire electrode and the burner head.

The injector and liner, both constructed of quartz, were designed to enhance mixing. Modeling suggested that $NO-NO_2$ conversion is an extremely fast process. It was therefore desirable that mixing occur as thoroughly and quickly as possible. This was achieved with the injector-liner design by converging the liner and creating a throat between the liner and the injector. Convergence of the liner past the burner promoted shear layer mixing of the flame products and bypass air by accelerating the flow through a 94 percent reduction in flow area. Methane was injected into the flow at the throat created between the bulb of the injector and the wall of the liner. Injection occurred through six, 0.508 mm orifices. The injection jets were spaced 60 deg apart and flowed normal to the main flow.

The liner created an annulus between its outside wall and the inside wall of the pressure vessel. Nitrogen was passed through this annulus in order to cool the walls of the pressure vessel. After some experimental testing, it was found that significant

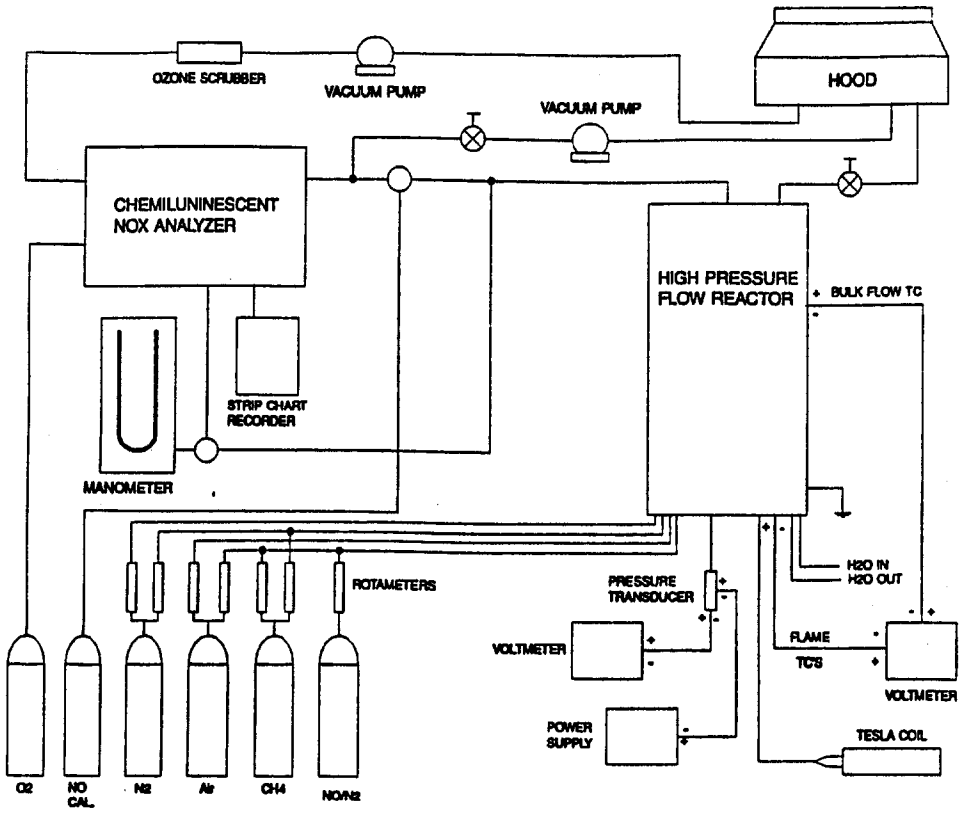


Fig. 2 Experimental system schematic

cooling nitrogen was being entrained into the main flow through the slots in the liner. This caused a significant temperature gradient in the direction of the flow and made it impossible to determine mass flow inside the liner. To eliminate entrainment, the slots were filled with Sauereisen (Electrotemp Cement) No. 8. A small opening was left in one of the top slots for thermocouple access. Any entrainment through this small space was neglected.

Instrumentation and Control. Pressure inside the vessel was controlled with a Whitey stainless steel forged body regulating valve located downstream from the exhaust port. The pressure inside the vessel was monitored via an Ashcroft Model K1 pressure transmitter attached to the bottom flange water trap. To prevent overpressurization, a Nupro relief valve was installed off the water trap. This valve was backed up with a Fike rupture disk mounted on the top flange.

Temperature inside the vessel was monitored with Type S Pt/Pt-10%Rh thermocouples. A double thermocouple probe fed through an NPT port in the bottom flange. One thermocouple sat in the flame front. The other was positioned approximately 2.5 cm above the flame. They were used to monitor flame extinction. The third thermocouple fed through a port in a top side flange. This thermocouple probe extended through an opening in the liner to the postflame gas flow. It was used to monitor bulk temperature of the postflame gases. It was assumed that temperature of the postflame gases was relatively isothermal as it passed through the upper section of the liner at steady state. Experimental testing showed this to be a valid assumption. Temperature of the postflame gases was controlled by the amount of air allowed to bypass the burner. Thermocouple junctions were kept at room temperature. The thermocouple readings were then corrected for the junction temperature. Thermocouple radiation effects were neglected.

Flow into the facility was metered with various Matheson rotameters. Working pressure for the rotameters was 1.38 MPa gage (200 psig). Figure 2 contains a schematic of the flow system. Methane was used both as the burner fuel and as the injected hydrocarbon. Methane was selected as a fuel because it is the major constituent of natural gas, a common industrial gas turbine fuel. Methane was mixed upstream of the burner with both air and a NO/N_2 mixture. A mixture of 0.99 percent NO in a balance of nitrogen was used to seed the flame.

All NO_x concentrations were determined using a Thermo Environmental Corporation Model 10 chemiluminescent analyzer (CLA). As past research has shown, probe conditions are extremely important in avoiding the formation of NO_2 in the probe. In order to minimize probe-formed NO_2, gases were sampled at low pressure in an uncooled quartz probe. Probe sampling pressures varied from 49 to 117 torr. In order to sample at such low pressures, the CLA was modified in a manner similar to the modifications made by Kimball-Linne (1984) in his NH_i/NO reaction kinetics research.

The CLA was calibrated using 240 ppm NO in nitrogen and room air as a zero gas. NO_2 concentrations were determined by recording NO and NO_x concentrations at a point and then calculating the NO_2 present. NO and NO_x levels were recorded using a strip chart recorder. No corrections for quenching effectiveness were made for other gases besides nitrogen in the NO_x samples.

Experimental Procedure. From the results of computer modeling and past research, it was seen that several variables are very important in the formation of NO_2. It was decided that experimental work should focus on these variables and their effect on the rate and extent of NO conversion to NO_2. Those variables examined experimentally were temperature, pressure, and the presence of unburned hydrocarbons. Twenty-six sets of data were collected using the high-pressure flow reactor. These runs encompassed data taken at temperatures in the following ranges: below 800 K, 800 K–1000 K, 1000 K–1200 K, and above 1200 K. Pressure was varied between atmospheric and 9.85 atm. Various data sets were taken both with and without

methane and nitric oxide injection. All flames were run fuel lean at or around an equivalence ratio of 0.75.

For a typical experimental run, the burner was ignited then sealed. The facility was allowed to come to a quasi-steady state. Data acquisition was begun at the zero point, defined as the point at which the sample probe met the injection bulb. Species concentrations at this point will be denoted with a subscripted "o." NO and NO_x measurements were obtained from the center of the flow at 5 mm to 20 mm intervals beginning at the zero point. The values of NO and NO_x concentrations were recorded at each spatial point in the flow by determining a time average of the analyzer output. Reactor flow conditions were also monitored during the time of data acquisition. There was usually some slight change in pressure and bulk flow temperature during the period of data acquisition. Again, a time average was estimated for these properties.

Computer Modeling. As a preliminary step to experimental work, computer modeling of the system was undertaken. This modeling served various purposes. It offered insight into the problem of NO–NO_2 conversion. It showed which parameters appeared to be important in NO_2 formation and gave a general idea of the kinetics involved in the process. This affected the design of the high-pressure flow reactor used in experimental work. It also offered support for much previous research as well as indicating which areas should be explored in experimental work.

Modeling was conducted using the CHEMKIN chemical kinetics code (Kee et al., 1989) and a plug flow reactor model of the combustor. The combustor was divided into two sections: a flame zone and a postflame zone. The flame zone was considered to be a constant-temperature reaction zone at the adiabatic flame temperature. The reactants entering the flame zone were methane, air, and nitric oxide. The significant species (those with a concentration greater than or equal to 0.1 ppm) at the end of the flame zone were then used as reactants (input) for the post flame zone. At this point, relatively cool air (580 K) was added to simulate the dilution air in an actual combustor. Thermodynamically, this was modeled as an instantaneous bulk quench of the postflame gases. At the end of the flame zone, CH_4 concentration was boosted to simulate a hydrocarbon promoter. The beginning of the postflame zone is analogous to the injection point of the experimental flow reactor. Although the injection point is not physically the same as the zero point, they shall be assumed to be the same in evaluating the data. The postflame gases were allowed to react isothermally, and the production of NO_2 was studied as a function of time.

The reaction mechanism that was used in the computer model is a subset of the Miller and Bowman (1989) mechanism. The mechanism includes 46 different species and 219 reactions (reactions involving HCCOH, H_2CN, HCNO, and HOCN species were omitted). The rate constants for the reactions are of the form: $AT^n e^{-Ea/RT}$, where A is the pre-exponential factor, T is temperature in K, n is a constant between 2 and -2, Ea is the activation energy, and R is the universal gas costant.

The reactions thought to be the most important in the NO–NO_2 conversion mechanism are:

$$NO + HO_2 \rightarrow NO_2 + OH \quad (1)$$

$$H + O_2 + M \rightarrow HO_2 + M \quad (2)$$

The Miller and Bowman mechanism showed that the conversion through these reactions is an extremely fast process. Under some conditions their mechanism predicted nearly complete NO_2 conversion within a millisecond. Based on the conditions under which the reaction constants were determined and the extraordinarily fast process they seemed to indicate, it was hypothesized that one or both were in reality too high. To investigate this, a good portion of the modeling was run with the "A" constant for reaction (1) lowered an order of magnitude. Results

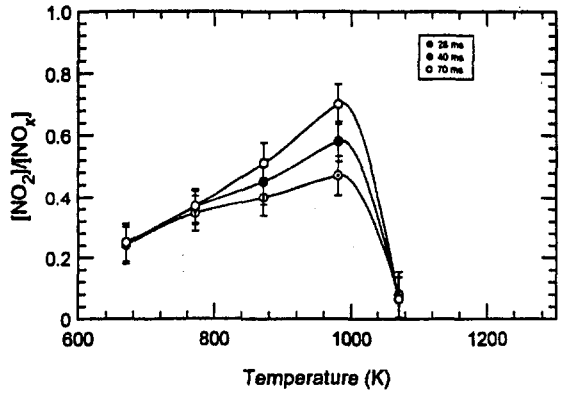


Fig. 3(*a*) Temperature profiles for data at P = 5.0 atm and $[CH_4]_{inj}$ = 1000 ppm

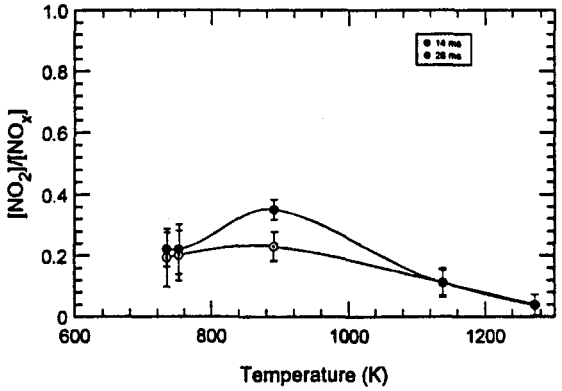


Fig. 3(*b*) Temperature profiles for data at P = 1.2 atm and $[CH_4]_{inj}$ = 1000 ppm

presented are those arrived at with the Miller and Bowman mechanism unless otherwise noted. The mechanism containing the reduced "A" constant for reaction (1) will be referred to as the modified Miller and Bowman mechanism.

Experimental Results

This section will examine the effects of varying temperature, pressure, and concentration of injected methane on conversion of NO to NO_2. All NO_x concentrations are reported on a volumetric basis, wet.

Experimental data were collected at temperatures between 664 K and 1272 K. The results show that temperature does indeed have a profound effect on NO_2 formation. There was a definite temperature range in which significant NO_2 production occurred. For a given pressure, peak levels of NO_2 were noted around 1000 K. However, above 1000 K, NO_2 formation diminished, as shown in Fig. 3. This figure shows NO_2/NO_x plotted as a function of temperature at given residence times for pressures of approximately one and five atmospheres. It should be noted that as temperature decreases below 1000 K, NO_2 formation is seen to fall off. Slight differences in pressure were neglected in comparing results. Although more NO is converted to NO_2 at higher pressure, it can be seen that there is a similar temperature range which is conducive to the conversion process. A temperature "window" becomes apparent.

The majority of experimental data was taken at near-atmospheric conditions and at five atmospheres. All near-atmospheric data were taken at pressures between 0.94 and 1.41 atm. Pressure differences in the near-atmospheric tests were neglected. In general, increasing pressure resulted in increased NO_2/NO_x. Without hydrocarbon injection, NO_2/NO_x profiles are seen to be relatively constant, and there is an increase in

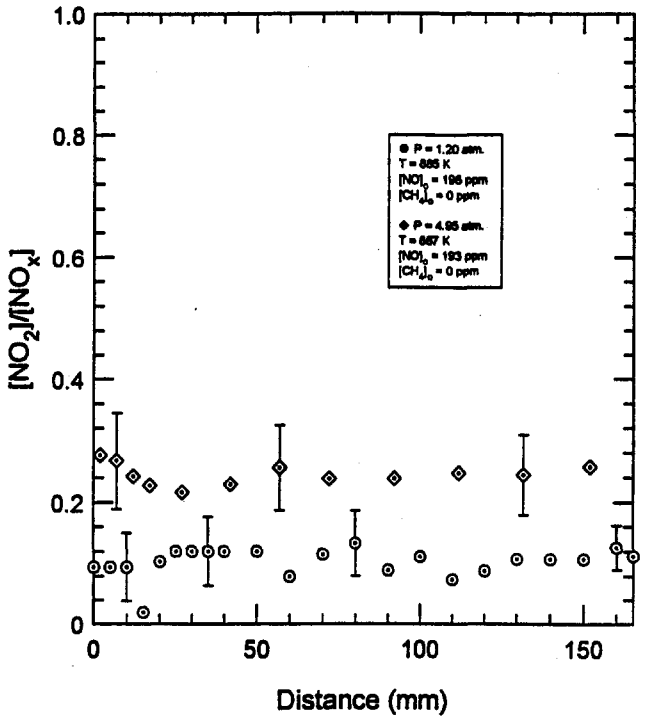


Fig. 4 The effect of pressure on NO_2/NO_x: no methane injection cases

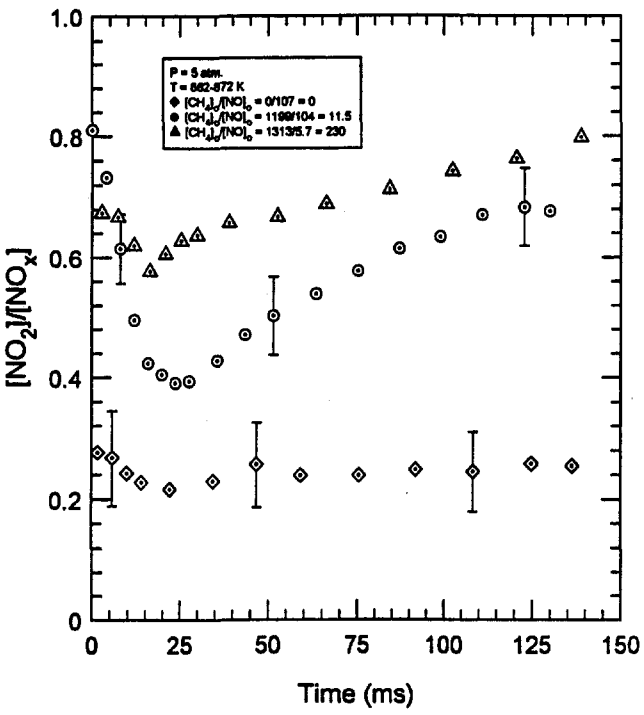


Fig. 6 The effect of hydrocarbon promotion on the conversion of NO to NO_2

NO_2/NO_x as pressure increases (Fig. 4). Between 1.2 and 5.0 atm, NO_2 formation approximately doubles. Figure 5 shows NO_2/NO_x profiles as a function of pressure in the temperature range 847 K − 891 K with hydrocarbon injection. The greatest NO–NO_2 conversion occurred at 8.54 atm. In this case, NO_2/NO_x approached 0.90 within the 150 ms residence time. It should be kept in mind that these data points were collected from various runs, which had slightly different mixing characteristics and temperatures. This may explain some of the data spread.

Experimental results support claims of past research that unburned hydrocarbons have a significant promotional effect on the conversion of NO to NO_2. Since initial NO concentration also appears to affect NO_2 formation, an important ratio may in fact be $[CH_4]_o/[NO]_o$. This is similar to a ratio of [NO] to [*n*-butane] defined by Bromly et al. (1992) in his research. Figure 6 shows a comparison of varying levels of $[CH_4]_o$ and $[NO]_o$ at 5 atm in a temperature range of 862 K − 872 K. It was difficult to measure $[NO]_o$ at the injection point because of the mixing characteristics of the system. To offer a more consistent basis of comparison, at a given pressure $[NO]_o$ is assumed to be equal to the value of $(1 - NO_2/NO_x)$ with no CH_4 injection multiplied by $[NO_x]$ at the end of the sample range. As $[CH_4]_o/[NO]_o$ increases, NO_2 formation also appears to increase. The increase is very dramatic with relatively small additions of hydrocarbon. However, continuing to increase the amount of injected hydrocarbon with $[NO]_o$ held constant has little further effect on NO_2 formation. Although results are only presented at 5 atm, the same trend is seen at other pressures.

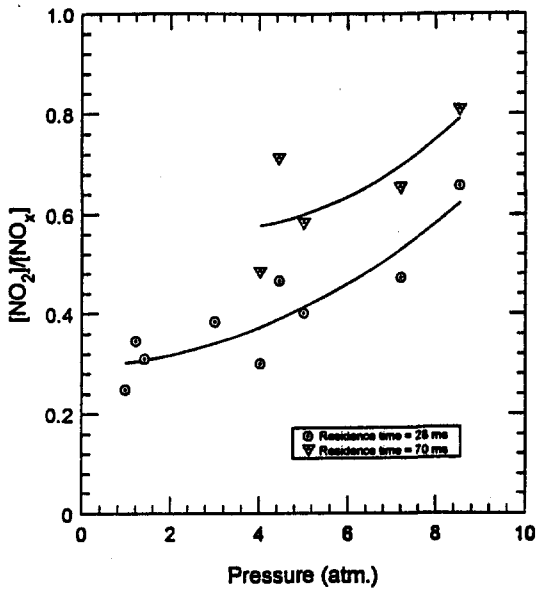


Fig. 5 The effect of pressure on NO_2 formation: T = 847–891 K, $[CH_4]_o$ = 1000 ppm

Discussion

In evaluating experimental results and comparing them to computer modeling results, an area of importance is the mixing characteristics of the flow reactor facility. Since NO–NO_2 conversion appears to be a relatively rapid process, mixing in the flow reactor is an important issue. A result of past studies and computer modeling not readily seen in the experimental work is a conservation of total NO_x. For near-atmospheric tests, NO_x concentration was at a maximum at $t = 0$ (i.e., that point where the probe meets the injector bulb). It then decreases relatively quickly to a minimum, at which point it begins to increase again slowly or remains relatively constant. This could be the result of mixing characteristics or a chemical effect of the methane injection. It was concluded that this was not a chemical effect because identical NO_x profiles were obtained under the same flow conditions both with and without methane injection. It follows that the inflated NO_x concentration around the injection point is a mixing effect. The fact that NO_x profiles level off shortly after the injection point suggest that NO_x is actually conserved in the NO–NO_2 conversion process.

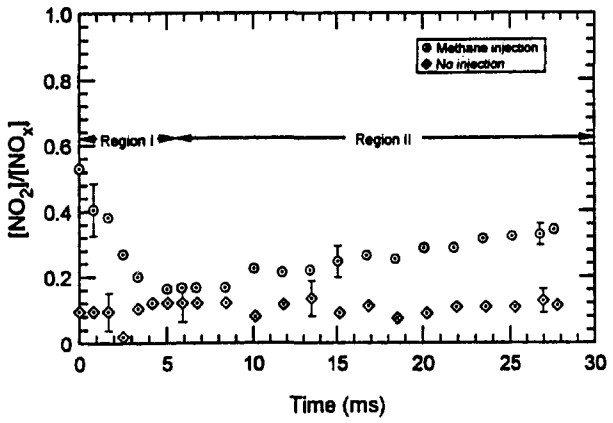


Fig. 7(*a*) NO_2/NO_x profiles at 1.2 atm and 890 K with $[NO]_o$ = 195 ppm

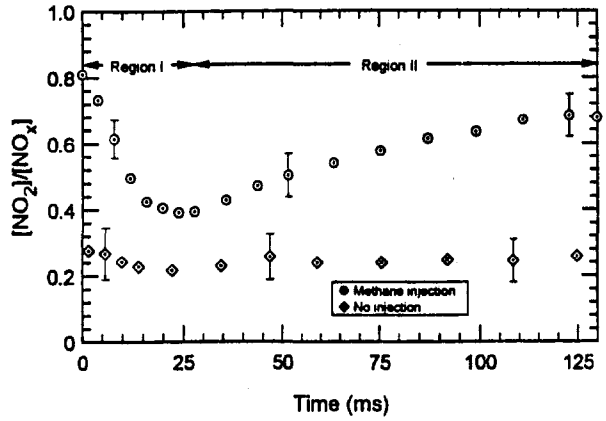


Fig. 7(*b*) NO_2/NO_x profiles at 5 atm and 870 K with $[NO]_o$ = 185 ppm

Fig. 7 Comparison of NO_2/NO_x profiles with and without methane injection

In comparing NO_2 formation for cases in which CH_4 is injected and cases in which it is not, it becomes apparent that there was NO_2 being formed at the injection point (Fig. 7). NO_2/NO_x then decreased as the flow mixed. In the region past the NO_2/NO_x minimum, there is an increase in NO_2/NO_x. This behavior of the system can be explained by mixing effects. The region between $t = 0$ and the NO_2/NO_x minimum shall be referred to as region I. The region past the NO_2/NO_x minimum shall be called region II. At the injection point, there were regions of locally high methane concentration. NO was quickly converted to NO_2 here. As the gases flowed downstream and mixed in region I, both CH_4 and NO_2 concentrations decreased as the flow became diluted with regard to these species at the center. At the minimum in NO_2/NO_x, mixing catches up with kinetics, and NO is converted to NO_2 at a slower rate. Thus, it appears that in region I, NO_2 formation is mixing controlled, and in region II it is kinetics controlled.

It has been shown that mixing is an important factor in assessing the experimental data of this study. Keeping this consideration in mind, the effects of several important variables—temperature, pressure, and unburned hydrocarbons—on the conversion of NO to NO_2 can be examined more closely.

Temperature. Modeling, experimental data, and past research indicate that the formation of NO_2 from NO is an extremely temperature sensitive process. Bromly et al. (1992) reported in their recent study that "at a given residence time, the system goes from being inert (no NO conversion) to completely reacted (100 percent NO → NO_2) over just a few degrees of temperature." Experimental data from this study support these observations.

Initial computer modeling with the Miller and Bowman mechanism revealed significant NO_2 conversion only at relatively low temperatures (below approximately 1200 K). NO–NO_2 conversion would therefore appear to be a postflame phenomenon. Temperatures in the flame zone would be much too high to support the conversion. The modified Miller and Bowman mechanism predicted an upper temperature limit similar to the original mechanism; however, unlike the original Miller and Bowman mechanism, a lower temperature limit also became evident (Fig. 8). This indicated that there may be a temperature window in which NO is readily converted to NO_2.

Experimental data supported the predictions of the modified mechanism. In the experimental work a clearly defined region was apparent in which significant NO_2 formation took place. This temperature "window" is more easily visualized when plotted as $[NO_2]/[NO_x]$ at a point in time as a function of temperature. The computer results for both mechanisms along with an experimentally obtained curve are plotted in Fig. 8. Points for these curves are taken at 150 ms in the case of the computer model. By this point, NO_2/NO_x is approaching steady state or has reached steady state. The reaction does not appear to have proceeded to this extent for the experimental case; therefore, different residence times were chosen. A similar temperature window is noted at the experimental residence times. It should be noted that the Miller and Bowman mechanism does not predict a lower limit to the temperature window that the experimental data and the modified mechanism indicate.

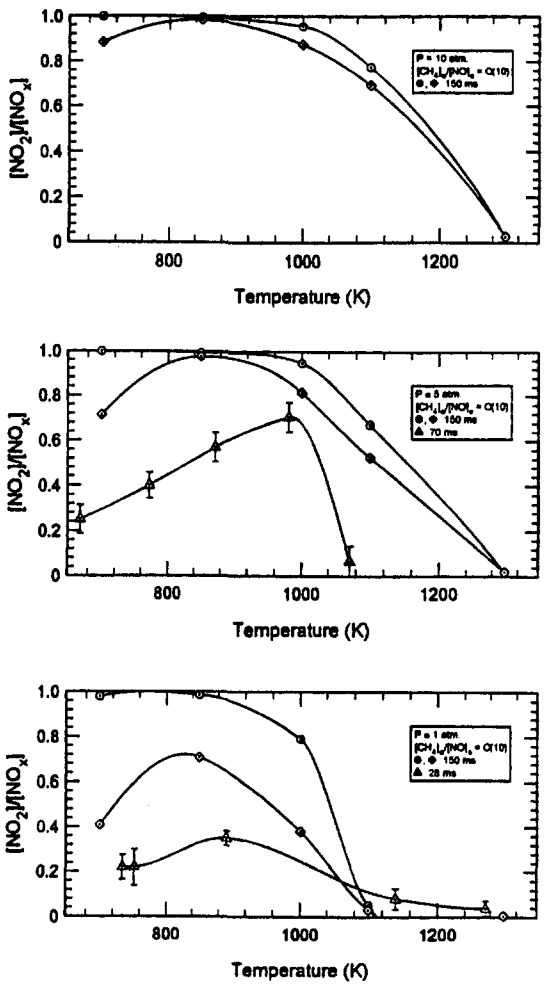


Fig. 8 Temperature effect on NO_2 formation: ⊙ Miller and Bowman mechanism, ◇ modified Miller and Bowman mechanism, △ experimental data

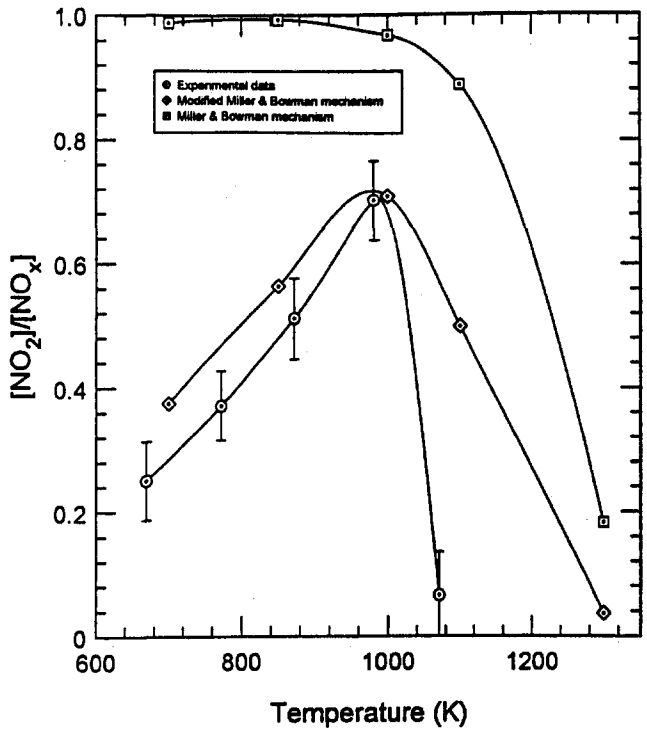


Fig. 9 Comparison of NO_2/NO_x temperature profile for modeling and experimental results; experimental points taken at 28 ms; modeling points taken at 10 ms; P = 5 atm, $[CH_4]_o$ = 1000 ppm

As a better basis of comparison, the maximum value of the experimental data curve at 5 atm was matched with the modified Miller and Bowman model. Points were then taken at this residence time from the model. The resulting curves are shown in Fig. 9. In comparing the modified mechanism curve with that obtained experimentally, the curves are similar. The modified mechanism, however, predicts a slightly higher upper temperature limit that does not drop off as quickly as experimental data indicates.

In examining these temperature curves, it should be noted that the lower boundary observed in experimental testing and predicted by the modified Miller and Bowman mechanism is of a different nature than the upper boundary. Above approximately 1000 K, the formation of NO_2 quickly decreases as temperature increases. Some NO_2 formed in this temperature region is destroyed relatively quickly. This is because at higher temperatures, the formation of HO_2 via the three-body reaction

$$H + O_2 + M \rightarrow HO_2 + M \qquad (2)$$

is limited by competition with the reaction

$$H + O_2 \rightarrow OH + O.$$

This results in less HO_2 formed and more OH and O formed, which, in turn, attack HO_2 and NO_2. Destruction of NO_2 and HO_2 then becomes significant and NO is favored. In the area of the lower "boundary," NO continues to be converted to NO_2, but at a much slower rate. The very rapid conversion seen in the ideal temperature range does not occur at lower temperatures. The "window" then is a temperature range in which NO conversion to NO_2 not only is favored, but occurs rapidly.

Pressure. Pressure does appear to be another variable that is important in the conversion of NO to NO_2. If the reaction $H + O_2 + M \rightarrow HO_2 + M$ is indeed an important reaction, increasing pressure would increase the probability of a three-body reaction occurring as the mean free path of the molecules decreases. If HO_2 production is increased, NO_2 production should also increase due to the reaction of HO_2 with NO (reaction (1)).

At first glance, the experimental results with regard to pressure appear somewhat ambiguous. Different residence times and mixing conditions make it difficult to compare various runs. In general, both experimental data and modeling show an increase in NO_2 production as pressure increases. If NO_2/NO_x is plotted at a given residence time for various pressures (with hydrocarbon promotion), an increase in NO_2 production is seen at higher pressures (Fig. 5). The spread of the data is most likely affected by the varying mixing conditions of different runs. By eliminating hydrocarbon promotion, the effect of pressure is isolated, and it is clearly seen that increased pressure leads to increased NO_2/NO_x levels (Fig. 4).

Data indicate that the temperature region favoring rapid and significant NO conversion changes as pressure changes. The upper temperature limit of NO conversion increases as pressure increases (Fig. 8). The lower limit appears to decrease with a pressure increase. In short, as pressure increases results suggest that the NO_2 temperature window grows larger with the greatest change occurring between 1 and 5 atm.

Unburned Hydrocarbons. It has been well documented in past research that the presence of unburned hydrocarbons greatly increases NO–NO_2 conversion. Recent studies have confirmed reports from previous studies that trace amounts of hydrocarbons greatly promote NO conversion to NO_2. These studies have investigated the effectiveness of various hydrocarbons in the promotion of the conversion of NO to NO_2 (Bromly et al., 1992; Hori et al., 1992; Marinov et al., 1993).

In this study, CH_4 was selected to be injected into the post-flame gas flow. The current results agree with past research in that the presence of UHCs is indeed an important factor in NO–NO_2 conversion. When no CH_4 was injected, the conversion was dramatically reduced. However, when a small amount of CH_4 was injected, NO_2 production significantly increased. As a baseline injection for computer modeling, approximately 1000 ppm was injected. Doubling this had little effect on NO_2 formation. Similarly, halving it also had little effect. However, when the injection concentration was decreased an order of magnitude, a significant reduction in NO_2 formation was noted. The question naturally arises as to how much injected hydrocarbon is required for optimum NO_2 formation.

The results indicated that $[NO]_o$ is also a variable in NO_2 production. In general, as $[NO]_o$ decreases with injected CH_4 held constant, NO_2/NO_x increases. Taking this into account, as well as the relationship with $[CH_4]_o$ described above, a ratio $[CH_4]_o/[NO]_o$, similar to that of Bromly et al. (1992), can be defined. This ratio is sensitive to changes in both injected CH_4 and NO, and, therefore, may be a useful quantity in predicting those conditions conducive to NO–NO_2 conversion. Computer modeling showed that the value of this ratio is indeed useful; however, the individual values of $[CH_4]_o$ and $[NO]_o$ must also be taken into account in predicting the extent of NO_2 formation. This is because the two variables appear to have a slightly disproportionate effect on NO_2 production. Bromly et al. (1992) report that $[NO]/[n\text{-butane}] = 1$ is an important quantity. Similarly, results of this study found $[CH_4]_o/[NO]_o = 1$ to be important. At $[CH_4]_o/[NO]_o$ ratios greater than one, there is an increase in NO_2/NO_x. For $[CH_4]_o/[NO]_o > 10$, the maximum effect of the hydrocarbon has apparently already been reached and there is little further increase in NO_2/NO_x as $[CH_4]_o/[NO]_o$ is further increased. As $[CH_4]_o/[NO]_o$ approaches zero, there is significantly less NO–NO_2 conversion (Fig. 10). This was noted in both experimental and modeling results. Similar experimental results at both 1 and 5 atm indicate that pressure probably does not significantly affect the $[CH_4]_o/[NO]_o$ ratio necessary to enhance NO–NO_2 conversion.

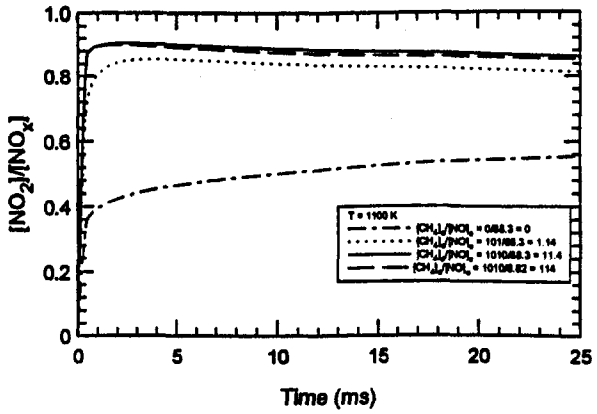


Fig. 10(*a*) Hydrocarbon promotion at 10 atm (computer model)

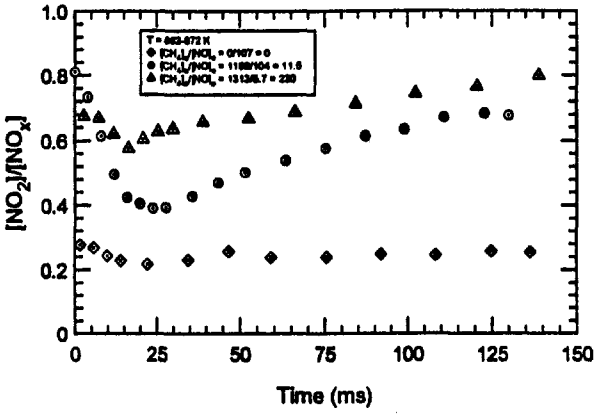


Fig. 10(*b*) Hydrocarbon promotion at 5 atm (experimental data)

Fig. 10 The effect of hydrocarbon promotion on the conversion of NO to NO_2

Comparison of Promoting Factors. In assessing the relative importance of the three promoting factors on $NO-NO_2$ conversion, it would appear that the order of decreasing importance is the presence of unburned hydrocarbons, temperature, and pressure.

This study as well as past studies indicate that hydrocarbons play an essential role in $NO-NO_2$ conversion. It has been shown that in their absence, NO is not converted to NO_2 even if the system is within the ideal temperature range. This study only examined the effects of one hydrocarbon—methane—on $NO-NO_2$ conversion. Recent studies have compared the promotional effects of different hydrocarbons on $NO-NO_2$ conversion and have shown methane to be the poorest of the hydrocarbon promoters (Bromley et al., 1992; Hori et al., 1992; Marinov et al., 1993). Even if methane is the least effective NO_2 promoting hydrocarbon agent, its effect on the conversion of NO to NO_2 is still significant.

Hydrocarbons must be considered in terms of their ratio with [NO] to be of use in explaining possible sources of high NO_2 formation. For a given system, NO may be decreased; however, if UHCs remain constant or increase, the decrease in NO may push the system into an $NO-NO_2$ conversion regime ($[CH_4]/[NO] > 1$). In this regime, the majority of NO may be converted to NO_2. If NO_2 levels increase enough, visible emissions may result. Therefore, although total NO_x is decreased, the resulting increase in NO_2/NO_x may result in visible emissions.

The conversion mechanism has been shown to be extremely temperature sensitive. Experimental data as well as some of the modeling indicate that there is a temperature region or window, outside of which NO conversion does not occur or occurs at a much slower rate. Research by Hori et al. (1992) also supports this idea of a temperature window.

A hydrocarbon promoter and the temperature are two essential factors in $NO-NO_2$ conversion. Outside of parameters involved with these variables, NO_2 formation does not occur or is drastically decreased. Pressure is not as essential as hydrocarbons and temperature, but as pressure increases, NO_2 is seen to increase. The results indicate that pressure may affect the peak values of NO_2/NO_x as well as the temperature range of optimal conversion. Pressure, however, is not as important as the first two factors because $NO-NO_2$ formation cannot be eliminated by changing the pressure above 1 atm. Although NO_2 formation may not be as fast or great at 1 atm as it is at 10 atm, this study as well as past studies have shown that NO conversion still does occur to significant levels, even at 1 atm.

Summary and Conclusion

The conversion of NO to NO_2 continues to be a problem in industry. In order to address the issue, a study of this conversion process was undertaken. Computer modeling of the system was undertaken. Based on the modeling results and past research, a high-pressure flow reactor was designed and built. This facility was used to collect experimental data on $NO-NO_2$ conversion. Results point to three variables that are of particular importance to the conversion process. These are temperature, the presence of unburned hydrocarbons, and pressure.

It was found that rapid, significant conversion only took place over a temperature range of several hundred degrees. Experimentally this temperature window was determined to be between approximately 800 K and 1000 K. The upper limit of this temperature window appears to be quite abrupt. At 70 ms and under similar conditions, NO_2/NO_x increases from approximately 0.07 to 0.70 as temperature changes from 981 K to 1071 K. Within less than 100 K, the system becomes "inert" with respect to $NO-NO_2$ conversion (i.e., no conversion of NO to NO_2). The lower limit is the result of a relatively gradual decrease in the rate at which NO is converted to NO_2. These temperature constraints on the mechanism limit NO_2 production to occurring in the relatively cool postflame region of a combustion system.

A conversion promoter appears to be another essential element in significant NO conversion. Hydrocarbons have been shown to be excellent species for promoting NO to NO_2 conversion. The relative concentration of methane with respect to NO concentration has been shown to be an important consideration. For $[CH_4]/[NO] > 1$, significant conversion was seen to occur. It was seen experimentally that at 5 atm and approximately 870 K, NO_2/NO_x was approximately 0.25 for the no-injection case, 0.70 for $[CH_4]_o/[NO]_o = 11.5$, and 0.80 for $[CH_4]_o/[NO]_o = 230$. Above the upper limit of the temperature window (approximately 1000 K), the presence of hydrocarbons had little effect on $NO-NO_2$ conversion.

Although pressure does not seem to have as dramatic an impact on $NO-NO_2$ conversion as temperature and hydrocarbon promotion, it is a contributing factor in the process. At higher pressure, increased levels of NO_2/NO_x were noted, both with and without hydrocarbon promotion. In exploring the effect of pressure alone (no hydrocarbon injection), as pressure of the reactor was increased from 1.20 to 4.95 atm, NO_2/NO_x was seen to increase from 0.12 to 0.25. Pressure, therefore, is an important variable in consideration of pressurized combustion systems.

The model does indeed appear to be a good tool in predicting incidences of significant NO to NO_2 conversion. Experimental data suggested that the rate of NO_2 formation is not as rapid as the model indicated. The important consideration, however, is that the model, particularly when incorporating the modified Miller and Bowman mechanism, predicted those trends seen in the experimental tests.

The results of this study yield more evidence that the major pathway for NO_2 formation is, in fact, the HO_2 mechanism,

$$NO + HO_2 \rightarrow NO_2 + OH.$$

The existence of an upper temperature limit to NO_2 formation and the increase in NO to NO_2 conversion noted at superatmospheric conditions suggest that the three-body reaction,

$$H + O_2 + M \rightarrow HO_2 + M,$$

which forms HO_2, is indeed important in the mechanism.

It has been shown that hydrocarbon promotion, temperature, and, to a lessor extent, pressure are important variables in the conversion of NO to NO_2. In a practical sense, as NO_x emissions are decreased past a certain point, they must be accompanied by a decrease in UHCs. This may help to explain why NO_2 has not been much of a concern until recently. NO_x reductions have finally reached the point where they have surpassed a limit with respect to UHC emissions. It is likely that further reductions in UHCs will be necessary before further NO_x reductions can be achieved while avoiding significant NO_2 production.

Acknowledgments

The authors wish to thank the Turbine Technology Department of the General Electric Corp. for their support of this research effort. They also wish to acknowledge Dr. Norm Laurendeau and his colleagues from Purdue University and Dr. Bill Flower from Sandia National Laboratories for their suggestions and influence in the design of the high-pressure flow reactor utilized in this study.

References

Allen, J. D., 1975, "Probe Sampling of Oxides of Nitrogen From Flames," *Combust. Flame,* Vol. 24, p. 133.

Amin, H., 1977, "Effect of Heterogeneous Removal of Oxygen Atoms on Measurement of Nitrogen Dioxide in Combustion Gas Sampling Probes," *Combust. Sci. Tech.,* Vol. 15, p. 31.

Bromly, J. H., Barnes, F. J., Mandyczewsky, R., Edwards, T. J., and Haynes, B. S., 1992, "An Experimental Investigation of the Mutually Sensitised Oxidation of Nitric Oxide and *n*-Butane," *Twenty-Fourth (International) Symposium on Combustion,* The Combustion Institute.

Carter, C. D., King, G. B., and Laurendeau, N. M., 1989, "A Combustion Facility for High-Pressure Flame Studies by Spectroscopic Methods," *Rev. Sci. Instr.,* Vol. 60, p. 8.

Chen, J. Y., McLean, W. J., and Gouldin, F. C., 1979, "The Oxidation of NO to NO_2 During Combustion Quenching Processes," WSS Paper No. 79-17, The Combustion Institute.

Hargraves, K. J. A., Harvey, R., Roper, F. G., and Smith, D. B., 1981, "Formation of NO_2 by Laminar Flames," *Eighteenth (International) Symposium on Combustion,* The Combustion Institute, p. 133.

Hori, M., 1980, "Effects of Probing Conditions on NO_2/NO_x Ratios," *Combust. Sci. Tech.,* Vol. 23, p. 131.

Hori, M., 1986, "Experimental Study of Nitrogen Dioxide Formation in Combustion Systems," *Twenty-First (International) Symposium on Combustion,* The Combustion Institute, p. 1181.

Hori, M., 1988, "Nitrogen Dioxide Formation by the Mixing of Hot Combustion Gas With Cold Air," *Twenty-Second (International) Symposium on Combustion.* The Combustion Institute, p. 1175.

Hori, M., Matsunaga, N., and Malte, P. C., 1992, "The Effect of Hydrocarbons on the Conversion of Nitric Oxide to Nitrogen Dioxide," WSS/CI Paper No. 92-36, The Combustion Institute.

Hunderup, J. W., 1993, "An Experimental Investigation of the Conversion of NO to NO_2 in a Simulated Gas Turbine Environment," M. S. Thesis, Virginia Polytechnic Institute & State University, Blacksburg, VA.

Jaasma, D., and Borman, G., 1980, "Peculiarities Associated With the Measurement of Oxides of Nitrogen Produced by Diffusion Flames," *Combust. Sci. and Tech.,* Vol. 23, p. 83.

Johnson, G. M., Smith, M. Y., and Mulcahy, M. F. R., 1978, "The Presence of NO_2 in Premixed Flames," *Seventeenth (International) Symposium on Combustion,* The Combustion Institute, p. 647.

Kee, R. J., Rupley, F. M., and Miller, J. A., 1989, "CHEMKIN II: A FORTRAN Chemical Kinetics Package for the Analysis of Gas Phase Chemical Kinetics," Sandia Report SAND89-8009, UC-401, Sandia National Laboratories.

Kimball-Linne, M. A., 1984, "Flow Reactor Studies of NHi/NO Reaction Kinetics, From 1050 K to 1450 K," Ph.D. Dissertation, Stanford University, CA.

Kramlich, J. C., and Malte, P. C., 1978, "Modeling and Measurement of Sample Probe Effects on Pollutant Gases Drawn From Flame Zones," *Combust. Sci. Tech.,* Vol. 18, p. 91.

Marinov, N. M., Steele, R. C., Hori, M., and Malte, P. C., 1993, "The Effect of Hydrocarbons on the Promotion of Nitric Oxide to Nitrogen Dioxide in a Well-Stirred Reactor," WSS/CI Paper, The Combustion Institute.

Miller, J. A., and Bowman, C. T., 1989, "Mechanism and Modeling of Nitrogen Chemistry in Combustion," *Prog. Energy Combust. Sci.,* Vol. 15, p. 287.

Sano, T., 1984, "NO_2 Formation in the Mixing Region of Hot Burned Gas with Cool Air," *Combust. Sci. Tech.,* Vol. 38, p. 129.

Sano, T., 1985, "NO_2 Formation in the Mixing Region of Hot Burned Gas With Cool Air—Effect of Surrounding Air," *Combust. Sci. Tech.,* Vol. 43, p. 259.

A NO_x Prediction Scheme for Lean-Premixed Gas Turbine Combustion Based on Detailed Chemical Kinetics

W. Polifke

K. Döbbeling

T. Sattelmayer

ABB Research Center,
Baden-Dättwil,
Switzerland

D. G. Nicol

P. C. Malte

Dept. of Mechanical Engineering,
University of Washington,
Seattle, WA 98195

The lean-premixed technique has proven very efficient in reducing the emissions of oxides of nitrogen (NO_x) from gas turbine combustors. The numerical prediction of NO_x levels in such combustors with multidimensional CFD codes has only met with limited success so far. This is to some extent due to the complexity of the NO_x formation chemistry in lean-premixed combustion, i.e., all three known NO_x formation routes (Zeldovich, nitrous, and prompt) can contribute significantly. Furthermore, NO_x formation occurs almost exclusively in the flame zone, where radical concentrations significantly above equilibrium values are observed. A relatively large chemical mechanism is therefore required to predict radical concentrations and NO_x formation rates under such conditions. These difficulties have prompted the development of a NO_x postprocessing scheme, where rate and concentration information necessary to predict NO_x formation is taken from one-dimensional combustion models with detailed chemistry and provided—via look-up tables—to the multidimensional CFD code. The look-up tables are prepared beforehand in accordance with the operating conditions and are based on CO concentrations, which are indicative of free radical chemistry. Once the reacting flow field has been computed with the main CFD code, the chemical source terms of the NO transport equation, i.e., local NO formation rates, are determined from the reacting flow field and the tabulated chemical data. Then the main code is turned on again to compute the NO concentration field. This NO_x submodel has no adjustable parameters and converges very quickly. Good agreement with experiment has been observed and interesting conclusions concerning superequilibrium O-atom concentrations and fluctuations of temperature could be drawn.

Introduction

For the combustion of fuels without fuel-bound nitrogen at gas turbine conditions, the lean premixed technique has already proven to be an enormously successful primary measure to significantly reduce emissions of oxides of nitrogen (NO_x) (Sattelmayer et al., 1992). Further reductions in NO_x from gas turbine combustors should be possible with the lean-premixed technique, and there is great interest in the development of such combustors. A numerical NO_x prediction scheme adapted to these conditions would obviously be a tool of great practical value for the designer. The concept, development, and implementation of such a model are presented here.

Experimental Setup and Operating Conditions

As a test case for the model of NO_x formation in lean-premixed combustion, experiments were conducted with an ABB Double Cone Burner with 0.1 m nominal diameter mounted in an atmospheric test rig. A Double Cone Burner comprises two halves of a cone that are shifted with respect to each other in the radial direction such that two inlet slits of constant width are formed. A strong tangential component is imparted to the gas entering the burner through these slits. The degree of swirl is chosen such that near the burner outlet the flow undergoes a vortex breakdown, resulting in a zone of recirculation, which acts as an aerodynamic flame holder. This type of flame stabilization in free space is the characteristic property of the double cone burner. More details have been given by Sattelmayer et al. (1992).

The burner used in the experiments is mounted in a movable tube and fires into a SiC-ceramic combustion chamber of square cross section with 0.14 m side length, which in turn is enclosed by a water-cooled steel liner; see Fig. 1. With an electrical heater, the inlet temperature of the combustible mixture can be increased up to 670 K. Typical operating conditions for the atmospheric test rig are inlet temperature $T_i \approx 630$ K, and combustible mixture volume flow $\dot{V} = 200\ m_N^3/h$.

In these experiments, fuel is injected far upstream of the burner in order to achieve perfect premixing of fuel and air. As a drawback, flashback occurs for excess air ratios λ below 1.9 (lean blow-out occurs at $\lambda \approx 2.2$). This is different from gas turbine applications, where gaseous fuel is injected along the Double Cone Burner's inlet slits to eliminate the risk of flashback into the premixing section, and consequently save operation at richer conditions is possible.

Sample gas is taken from the combustor with a movable water-cooled stainless steel probe. The concentrations of O_2, CO, CO_2, and NO_x are determined with the paramagnetic, the NDIR, and the chemiluminiscence principle, respectively. An uncoated Pt–PtRh thermocouple, mounted on an uncooled ceramic suction probe, is used to measure temperature. No corrections for radiation loss are made. In addition to the temperature and species measurements in the reacting flow, LDA measurements of velocity were conducted in a water model of the burner and combustion chamber in order to determine appropriate inlet

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 10, 1995. Paper No. 95-GT-108. Associate Technical Editor: C. J. Russo.

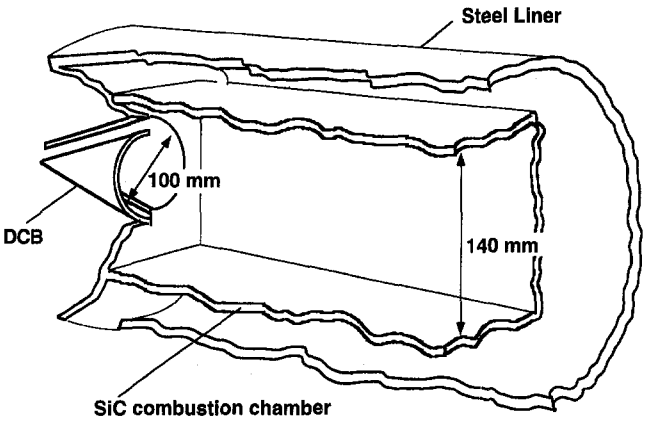


Fig. 1 Sketch of the test rig with Double Cone Burner (not to scale)

conditions for the two-dimensional numerical model, and to assess the quality of the turbulent flow modeling; this is discussed in more detail by Polifke et al. (1995).

Turbulent Flow and Heat Release Modeling

A commercial CFD code—finite volume based with "standard" SIMPLE(C) pressure correction algorithm (Patankar, 1980)—was used as the basis for the implementation of the turbulent combustion model and the NO_x postprocessor. The geometry of the burner and the combustion chamber were mapped to a two-dimensional cylindrically symmetric Cartesian grid with 6000 cells. Although both the burner and the combustion chamber are not exactly cylindrically symmetric, this simplification seems appropriate to keep memory and CPU time demands at an acceptable level.

Flow conditions at the inlet section of the two-dimensional model are not uniform and must be based on experimental measurements or three-dimensional numerical predictions of the flow in the double cone burner's inlet slits. Experiments were conducted in a water model of the burner at the appropriate Reynolds number. Axial, tangential, and radial velocity and turbulent intensity along the inlet slits and downstream of the burner were measured with standard LDA equipment. Additionally, a three-dimensional, boundary-fitted grid of the double cone burner was generated, which includes also the test rig section upstream of the burner. It has been found that the numerical predictions of the flow through the burner's inlet slits are rather insensitive to the details of the boundary conditions prescribed at the inlet of the three-dimensional model and are in good agreement with experiment.

Inlet flow conditions for the two-dimensional geometry were then set to reproduce along the burner's inlet the fluxes of mass, and axial and azimuthal momentum observed in the three-dimensional flows. The boundary values for turbulent kinetic energy k and dissipation ϵ were determined similarly from the three-dimensional data by appropriate averages. The inflow distributions resulting from this prescription appear entirely reasonable, and characteristic features of the flow inside the combustor—e.g., the size and location of the recirculation bubble—obtained with these boundary conditions are in good agreement with experiment and three-dimensional computation.

Table 1 Reaction rate constants in SI units (m/s/K/J/kmol)

Reaction	A	n	a	b	E_A
(1)	4.56×10^{11}	0	0.2	1.3	2.03×10^8
(2)	3.16×10^{12}	0	1.5	0.25	1.67×10^8

It is well known, see, e.g., Sloan et al. (1986), that the standard $k-\epsilon$ model is not well suited for modeling swirling flows, and use of a full differential Reynolds stress model (RSM) is usually recommended for such cases. However, the use of an RSM turbulence model entails significantly longer compute times and often severe convergence difficulties. The swirl number of the Double Cone Burner, defined as the ratio of azimuthal to axial momentum fluxes divided by a characteristic radial length scale is less than unity. Such a level of swirl would commonly be classified as "moderate" swirl. For moderately swirling flows, a $k-\epsilon$ model with a gradient Richardson number correction (see Sloan et al., 1986, or Phillip et al., 1991) seems to be an acceptable compromise between accuracy and computational requirements.

As discussed below, CO concentrations are required to predict NO_x formation rates with our model. This requires at least the use of a two-step global reaction mechanism with CO as an intermediate:

$$CH_4 + \tfrac{3}{2}O_2 \rightarrow CO + 2H_2O, \quad (1)$$

$$CO + \tfrac{1}{2}O_2 \leftrightarrow CO_2. \quad (2)$$

Rate expressions for the forward reactions are of a generalized Arrhenius form, based on reactant concentrations $[R_i]$ and temperature T:

$$w_f = AT^n[R_a]^a[R_b]^b \exp\left(-\frac{E_A}{RT}\right), \quad (3)$$

with pre-exponential factor A, temperature and species exponents n, a, b, and activation energy E_A based on the work of Westbrook and Dryer (1980). Heeding the warning given by Coffee (1985), negative species concentration exponents have not been used in this study, as they may adversely affect both the convergence behavior and the accuracy of a computation. Table 1 lists the reaction rate constants used in this study.[1]

[1] Westbrook and Dryer (1980) do provide a set of rate constants for a one-step global mechanism that uses only positive exponents ("Set 3" in Table II of their paper). Note that the exponential prefactor A_3 of this set differs significantly from the one for "Set 2," i.e., the set with a fuel concentration exponent $a = -0.3$. Unfortunately, for the two-step mechanism Westbrook and Dryer provide only a set of rate constants with $a = -0.3$. The product of this prefactor with A_3/A_2 is used in the set of rate constants for reaction (1) in this study; see Table 1.

Nomenclature

A = Arrhenius pre-exponential factor
CO, O_2, NO, ... = chemical species
Da, Ka = Damköhler, Karlovitz numbers
E_A = activation energy
Re = turbulent Reynolds number
S_L, δ_L = laminar flame speed and thickness
T = temperature
$\dot{V}$ = volumetric flow rate
Y = species mass fraction
a, b, n = species, temperature exponents
k = turbulent kinetic energy
k_f, k_r, k_1, k_2, ... = reaction rate constants
l_T, t_T = turbulent macro length and time
p = pressure
t_c = chemical heat release time
w = reaction rate
ϵ = turbulent dissipation
η = unmixedness factor
λ = excess air ratio
ν = dynamic viscosity
τ = reactor residence time

Rate constants for the reverse of reaction (2) have also been provided by Westbrook and Dryer (1980), i.e., $k_r = 5 \times 10^8 \exp(-1.67 \times 10^8/RT)$. However, the value of the activation energy E_A given for the reverse reaction is equal to that of the forward reaction. The ratio of [CO] to [CO_2] at equilibrium resulting from such a prescription will obviously be independent of temperature, which is not plausible. Indeed, at lean conditions, totally unrealistic equilibrium levels of [CO] are obtained with the rate constants suggested by Westbrook and Dryer (1980). Therefore, the reverse rate of reaction (2) was determined from the equilibrium condition $k_f/k_r = K_c$ for the rate constants.[2]

Admittedly, the rate constants have been manipulated in a very ad hoc manner to suit our needs. The justification for doing so is, first, that this work deals with the description of the concept and implementation of a NO_x postprocessor, and not with the development or detailed testing of a turbulent combustion model. The results obtained with the set of rate constants listed in Table 1 are qualitatively correct, and this shall suffice to test the validity of our NO_x postprocessor concept. Second, the reaction progress limitations imposed by chemical kinetics are only one aspect of turbulent combustion. The limitations imposed by turbulent mixing must—depending on the relative magnitudes of relevant chemical and turbulent time scales—also be considered.

Indeed, it is generally recognized that different types or "regimes" of turbulent combustion can be distinguished (Borghi, 1988; see also Libby and Williams, 1994, and references therein). The regimes are usually characterized by three dimensionless numbers: the turbulent Reynolds number $\mathrm{Re} \equiv u' l_T/\nu$, the turbulent Damköhler number $\mathrm{Da} \equiv t_T/t_c$, and the Karlovitz number $\mathrm{Ka} \equiv u' \delta_L/l_t S_L$. Here l_T is the turbulent macro scale, while l_t denotes the Taylor micro scale, $t_T \sim l_T/u'$ a mixing time typical of the macro scale turbulent eddies. The laminar flame is characterized by its flame speed S_L and thickness δ_L; a typical chemical heat release time t_c may be expressed as $t_c \sim \delta_L/S_L$. For the Double Cone Burner atmospheric test stand at typical operating conditions, we estimate $\mathrm{Re} \approx 250$, while the turbulent Damköhler number is of order unity, i.e., turbulent mixing time scales t_T are of the same order of magnitude as chemical time scales t_c. Estimates for the Karlovitz number yield $\mathrm{Ka} > 2$, which indicates that thin laminar reaction layers can be quenched by the smallest turbulent eddies. Following Borghi (1988), the turbulent flame front in the Double Cone Burner should be categorized as a (wrinkled) thickened flame. See also Hoffmann et al. (1994) for a more detailed discussion and experimental results concerning the structure of highly turbulent lean-premixed flames.

This suggests that eddy break-up, eddy dissipation, and flamelet models (see, e.g., Libby and Williams, 1994, and references therein), all of which assume that the Damköhler number is large and the Karlovitz number not much greater than one, are not appropriate for modeling this type of flame. Instead, a turbulent combustion model has been used in this study that incorporates both turbulent mixing and chemical kinetic rate limitations. The overall reaction rate R is expressed as:

$$\frac{1}{w} = \frac{\eta}{w_{\mathrm{Mix}}} + \frac{1-\eta}{w_{\mathrm{Kin}}}. \tag{4}$$

See Philip et al. (1991) for a similar formulation. Here we introduce an "unmixedness factor" η, based on a nondimensionalized and normalized gradient of temperature:

[2] Note that for reasons of numerical convenience and robustness, the [H_2O] dependence of the CO oxidation given by Westbrook and Dryer (1980) has been lumped into the prefactor for reaction (2) by assuming that $[H_2O] \approx 2[CO]$, which holds as long as CO_2 concentrations are small. Note that this procedure does not affect the equilibrium concentrations of CO and CO_2 if the reverse rate is computed from $k_f/k_r = K_c$.

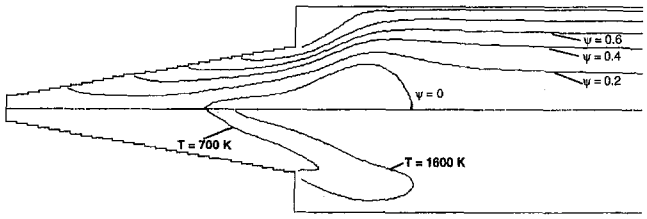


Fig. 2 Contours of streamfunction and temperature

$$\eta \equiv \tanh\left(C_\eta |\nabla T| \frac{l_T}{T}\right). \tag{5}$$

The tanh function has been chosen for mere convenience to normalize the gradient to the interval [0, 1]; other formulations should be possible. C_η is an adjustable constant of order 10. It has been observed that computed species and temperature distributions are rather insensitive to the exact value of C_η as long as the resulting unmixedness η is near unity in the vicinity of the flame front and small elsewhere.

What is the rationale behind this formulation? In homogeneous regions of the computational domain, where $|\nabla T|$ is small, $\eta \approx 0$, and $w \approx w_{\mathrm{Kin}}$, i.e., chemical kinetics is rate limiting—as it should be, because mixing processes cannot impede the reaction progress if the temperature and species distributions are already uniform. Furthermore, it is often assumed in turbulence models that temporal or statistical fluctuations are small in regions where spatial gradients are small. From this assumption it follows that in regions where temperature gradients are small, the mean values of temperature (and species concentrations) may be used to determine the mean chemical rates. By the same token fluctuations will be large near the flame zone, where hot products are in proximity with the fresh fuel/air mixture. Due to the strong nonlinearity of the exponential term in Eq. (3), rates based on mean values will then underpredict the mean reaction progress. Therefore, turbulent mixing is set to be rate determining in these regions. In intermediate regions, both turbulent mixing and kinetics can influence the overall rate. Again, see Philip et al. (1991) for a similar approach.

Figure 2 shows distributions of temperature and stream function from a computation with a total mass flux of 0.089 kg/s, mixture inlet temperature $T_i = 630$ K, methane fuel at an air excess ratio $\lambda \approx 2$, and zero heat flux boundary conditions. Presented are contour plots of stream function ψ (top half) and temperature T (bottom half) in an axial cross section of the burner and combustion chamber. Near the burner exit, the central recirculation bubble ($\psi = 0$) is visible. The two contour lines of temperature at $T = 700$ K and $T = 1600$ K show the location of the flame front.

A detailed comparison (see Polifke et al., 1995) of numerical and experimental temperature and CO distributions reveals that although the overall flow and temperature distributions agree well with burner design expectations and experimental observations, there is significant disagreement in the vicinity of the flame front. In particular, it seems that the turbulent combustion model employed underpredicts the turbulent intensity and flame speed and overpredicts the equilibration or burn-out rate.

In this work, however, we are mainly concerned with the implementation of the NO_x postprocessing model. Considering that overall flow and temperature patterns are qualitatively correct, and that the predicted peak and equilibrium CO levels are of the correct order of magnitude, one may expect that at least the feasibility of our NO_x prediction concept can be tested on the basis of the employed heat release model.

NO_x Chemistry in Lean-Premixed Combustion

The formation of NO in lean (λ = 1.5 to 2.5), premixed combustors fired on methane or natural gas occurs by the nitrous oxide, the Zeldovich, and—to a lesser extent—the prompt mechanisms. This is quite different from the situation in diffusion flames, where the Zeldovich and prompt mechanisms are dominant, and most certainly the reason why the application of a NO_x prediction model (Missaghi et al., 1990; Dupont et al., 1993) suitable for richer conditions and/or diffusion flames to the Double Cone Burner gave totally unacceptable results.

The rate of NO production by the Zeldovich mechanism under such conditions is given with good accuracy by

$$d[NO]/dt = 2k_1[N_2][O], \quad (6)$$

where the rate constant k_1 in the Eq. (6) is that of the initial Zeldovich reaction:

$$N_2 + O \rightarrow NO + N. \quad (7)$$

A value of $k_1 = 1.8 \times 10^{11} \exp(-38370/T)$ m^3 kmol^{-1} s^{-1} has been used, as suggested by Hanson and Salimian (1984).

As λ and p increase and the inlet air temperature decreases, the ratio of NO formed by the nitrous oxide mechanism to that by the Zeldovich mechanism increases. Indeed, under the leanest conditions and at high pressures, the nitrous oxide mechanism is predominant and can account for essentially all of the NO formed (Nicol et al., 1993). The nitrous oxide is formed by the following reaction:

$$N_2 + O + M \leftrightarrow N_2O + M. \quad (8)$$

For high-pressure combustion, the N_2O is destroyed mainly by the reverse reaction. Destruction of the N_2O by free radical attack plays a secondary role. This is different from the situation at atmospheric pressure, in which case destruction by H-atom attack (i.e., $N_2O + H \rightarrow N_2 + OH$) is important.

Nitric oxide is formed from N_2O and O or H:

$$N_2O + O \rightarrow NO + NO, \quad (9)$$

$$N_2O + H \rightarrow NO + NH. \quad (10)$$

For high-pressure, lean-premixed combustion, reaction (9) is predominant, whereas for 1 atm combustion, reaction (10) also is important.

Under lean conditions, the NH formed by reaction (10) is converted to NO. This occurs mainly through reaction of the NH to N and HNO, followed by reaction of these intermediates to NO. If the steady-state assumption is valid for the NH, N, and HNO, the rate of NO production by the nitrous oxide mechanism is:

$$d[NO]/dt = 2k_2[N_2O][O] + 2k_3[N_2O][H], \quad (11)$$

where k_2 and k_3 are the rate constants of reactions (9) and (10).

The third mechanism of interest is the so-called prompt[3] or Fenimore mechanism, i.e., fixation of N_2 by hydrocarbon attack leading to NO. This mechanism, based on available rate data, exerts a negligible-to-weak effect in high-pressure, lean, premixed combustion. However, for the flame zone of 1 atm, lean, premixed combustors, the mechanism can be significant and should be considered. The initiating reaction is taken to be:

$$N_2 + CH \rightarrow HCN + N \quad (12)$$

The HCN forms in the flame zone, followed rapidly by complete conversion to NO under lean conditions. A second effect is the reduction of NO by hydrocarbon attack, leading to additional HCN and other cyano species (e.g., HCNO). These molecules form in the flame zone and then are reconverted to NO. With respect to NO production, the net effect of the attack of hydrocarbons on the nitrogen system is the formation of NO in the flame zone at a rate less than two times the rate of reaction (12), because some of the nitrogen is tied up as cyano molecules, followed by a yield of NO as the cyano molecules are converted to NO.

By comparing PSR/PFR models of turbulent combustion of methane for λ = 2 (Nicol et al., 1994), it is possible to determine the contributions of the three NO_x mechanisms for a range of pressure levels. For lean premixed combustion at 1 atm, it is found that all of the mechanisms contribute in the flame zone (which is represented by the PSR). Due to the influence of the radicals, the nitrogen oxides are formed mainly in the flame zone (and immediate post-flame zone), rather than in the burnout zone. In particular, the maximum rate of NO_x production occurs throughout the PSR and initial part of the PFR, until a time of about 3τ is reached, where τ is the PSR residence time. This is a relatively long time period, which arises because of the slow relaxation of the O atom toward its equilibrium level. Thus, the nitrous oxide mechanism (which is sensitive to O atom) persists into the PFR, and has a contribution of NO_x competitive with that of the Zeldovich mechanism until a time of about 15τ is reached. Downstream, in the burnout zone, the Zeldovich mechanism is predominant, though by this point the rate of production of the NO_x is very low compared to that in the flame zone and immediate post-flame zone. Since the CH radical is very short-lived, the prompt NO_x mechanism only contributes in the flame zone.

For high-pressure combustion (e.g., 14.3 atm from Nicol et al., 1995) the NO_x behavior is markedly different. The nitrous oxide mechanism is the predominant source of the NO_x—again the NO_x forms mainly in the flame zone. The prompt NO_x mechanism contributes weakly, and the Zeldovich mechanism does not become significant (relative to the nitrous oxide mechanism) until well into the burnout zone. Because of the fast relaxation of the O atom toward equilibrium in high-pressure combustion, the period of maximum NO_x production is very short. This behavior may explain the lack of a positive pressure dependency of the NO_x, as observed in some experiments (e.g., see Joshi et al., 1994; Steele et al., 1995). For 1 atm combustion, because of the persistence of the maximum NO_x production rate, the radical dominated NO_x formation is spread over a significant range of eddy sizes, up to times of several milliseconds. However, at high pressure the region of super-equilibrium O atoms can be very small, although the maximum rate of NO_x production is very high (e.g., about 10-fold higher at 14.3 atm than at 1 atm; Nicol et al., 1995). The result of these competing effects is that the NO_x emission does not increase with pressure, and in some situations actually falls off moderately with increasing pressure (Joshi et al., 1994; Steele et al., 1995).

The Postprocessor Concept

The total amount of nitrogen oxides formed in lean, premixed combustion is typically (much) less than 100 ppmv, so that the influence of NO_x formation on the heat release process and the concentrations of stable species and the more important radicals may safely be neglected. It is therefore possible to decouple the prediction of nitrogen oxide concentrations from the "main" heat release computation. This approach is only feasible, if the chemical source terms for NO_x may be derived from the main solution field with sufficient accuracy. If this is the case, the CFD solver can be used in a "postprocessing" step to solve the transport equations of the nitrogen oxides and compute their two- or three-dimensional concentration fields, while keeping the values of the "main" variables frozen.

However, the discussion in the previous section has shown that a sizable chemical mechanism is required to compute the contributions of the three NO_x formation routes to the source

[3] Obviously, in lean-premixed combustion a substantial portion of the NO_x stemming from the nitrous oxide and Zeldovich mechanisms is also produced "promptly" in the flame zone.

term, especially at atmospheric pressures. It is not always possible to reduce the number of reactions to a value that can be comfortably handled by a CFD code not designed for stiff problems. Empirically fitted global reaction mechanisms have therefore been introduced to predict, e.g., the formation of prompt NO_x, see Missaghi et al. (1990) and Dupont et al. (1993).

Furthermore, Eqs. (6)–(12) show that NO_x formation rates depend directly on the concentrations of a number of radicals, which are typically not predicted by the turbulent combustion models that are currently in use for practical applications. It is often suggested to approximate radical concentrations by their respective (partial) equilibrium values, see e.g., Warnatz and Maas (1993) and Missaghi et al. (1990). However, the temperature levels achieved in lean-premixed low-NO_x combustion systems are too low to allow the use of this approximation. For example, Zeldovich NO_x formation depends linearly on the O atom concentration, while at $p = 1$ bar the O/H/OH radical pool is in general not in partial equilibrium at temperatures of interest here.

In order to circumvent these difficulties, the postprocessor developed in this study extracts radical concentrations from numerically evaluated dependencies, i.e., look-up tables that are based on local CO concentration. Among the variables predicted by the CFD code, CO has been chosen as the basis of the tables as it is most indicative of free radical chemistry and less sensitive to numerical inaccuracies than temperature. The look-up table is generated (before the postprocessor's run) by adiabatic one-dimensional laminar flame or reactor modeling with detailed chemistry. Note that the operating conditions (equivalence ratio, inlet temperature, pressure) for the laminar flame or reactor models with detailed chemistry are always set in accordance with the main simulation. We emphasize that it is not assumed that there is a "universal" correlation between CO and radical concentration and NO_x formation that is valid over a wide range of operating conditions. It has nevertheless been found that the CO correlations are surprisingly universal inasmuch as they do not depend on the details of the reactor modeling, provided that the combustion process is (nearly) adiabatic and that the CO oxidation process is not quenched (see below and Nicol et al., 1994).

With the O atom mass fraction Y_O taken from the appropriate look-up table, the prediction of Zeldovich NO_x formation rate from Eq. (6) is straightforward. How are the other formation paths treated? In the high-pressure case, Fenimore prompt NO_x may safely be neglected (Nicol et al., 1995), and the nitrous-oxide route is well described by Eq. (11) with the H atom contribution dropped. Providing CO correlations for Y_{N_2O} and Y_O from the detailed chemical model allows then a prediction of the NO_x formation rate via the nitrous oxide route. At atmospheric pressure, which is considered in this study, the situation is more complicated: Fenimore NO_x should not be neglected, and the destruction of N_2O by hydrogen atom attack is important. It has therefore been decided to provide the formation rate of NO via the nitrous and prompt route directly to the postprocessor.

The NO_x postprocessor operates then in two steps: (1) The one-dimensional adiabatic laminar flame or chemical reactor model is used as a preprocessor to generate a look-up table of Y_O and $d[Y_{NO}]/dt$ by the nitrous oxide and prompt mechanisms versus Y_{CO} for the given λ, p, and inlet temperature. The chemical mechanism used is that of Miller and Bowman (1989) with Zeldovich NO_x deactivated. (2) Calls to the look-up table provide the CFD code with local values for the O atom concentration and the rates of NO production by the nitrous oxide and prompt mechanisms. Zeldovich NO formation is then computed according to Eq. (6), using also the CFD generated local temperature.

Adopting this special treatment for Zeldovich NO, it appears principally feasible to consider the effect of heat losses and/or

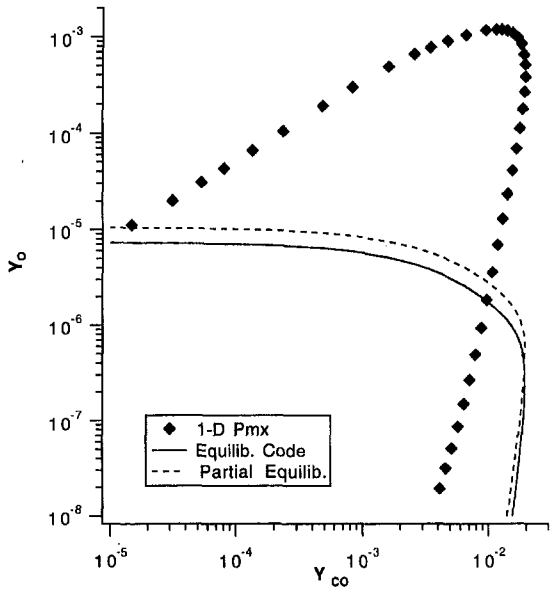


Fig. 3 O atom mass fraction plotted versus CO mass fraction. Shown are one-dimensional premix flame results (T_i = 630 K, λ = 2.1), equilibrium computations based on the local one-dimensional premix flame concentrations, and the partial equilibrium expression by Warnatz.

temperature fluctuations at least on Zeldovich NO_x formation—which shows the strongest temperature dependence of all NO_x formation paths considered due to the high activation energies involved.[4] Of course, this would require that the [O] – [CO] correlation be fairly insensitive to heat losses. However, since this has not yet been investigated, the present work deals only with adiabatic cases and the effects of temperature fluctuations are not included. See below for a further discussion of the latter aspect.

As an important prerequisite for the NO_x prediction it is useful to investigate how strongly the preprocessor results, i.e., the CO correlations, depend on the type of chemical model (one-dimensional laminar premix flame or PSR/PFR).

Kinetic Modeling Results

As discussed above, our NO_x postprocessor does not perform detailed chemical kinetic calculations, it merely links—via look-up tables—reaction kinetic data derived from one-dimensional premix flame or PSR/PFR computations into the two-dimensional or three-dimensional solution field obtained from the turbulent combustion modeling. Specifically, O atom concentrations and prompt/nitrous NO_x formation rates for a given value of CO mass fraction and temperature are extracted from the detailed chemical kinetics computations. Throughout this paper, the detailed mechanism used for combustion of CH_4 is that of Miller and Bowman (1989).

This is illustrated in Fig. 3, which shows O atom versus CO mass fraction taken from a one-dimensional premix flame computation at atmospheric pressure with inlet temperature T_i = 630 K, air excess ratio λ = 2.1, and methane fuel. Also plotted are equilibrium O atom mass fractions obtained with the Chemkin equilibrium code (Kee et al., 1985, 1989) and with the partial equilibrium expression $[O] = k_p[O_2]^{1/2}$, where $k_p = 34.6T^{0.5} \exp\{-27123/T\}$ as suggested by Warnatz and Maas (1993). Please note that traversing the flame front from

[4] The effects of heat losses on nitrous/prompt NO_x are less important because the relevant activation energies are generally lower and because heat losses in the flame zone, where the nitrous and prompt routes dominate, may be expected to be small as radiative heat transfer may be neglected in lean-premixed flames and cooling air (if any) is usually not injected into the flame zone.

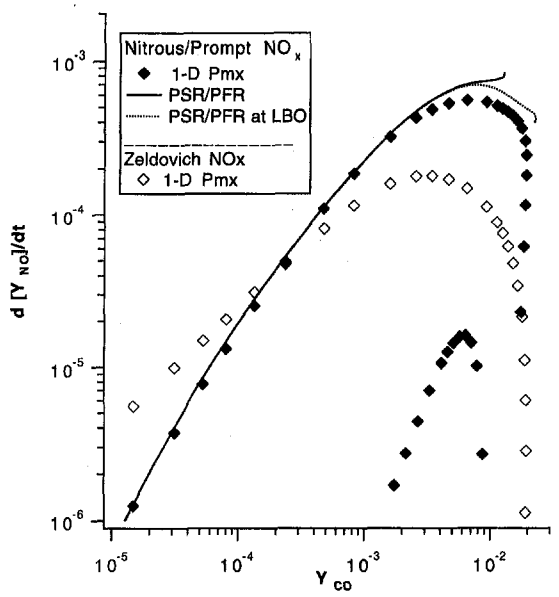


Fig. 4 NO production rates (in mass fraction per second) from one-dimensional premix flame computations and PSR/PFR models with residence times $\tau \approx 0.2$ ms (lean blow out) and $\tau \approx 0.4$ ms

the unburned to the burned side corresponds to traversing the curves in the diagram in the counterclockwise direction. Equilibrium conditions are near the upper left end of the curves. The portion of the curves with O atom mass fractions below 10^{-8}—corresponding to the cold part of the preheat zone—have been omitted from the graph.

It is found that in the flame zone, i.e., near the CO peak, O atom mass fractions are up to two orders of magnitude above (partial) equilibrium values, i.e., superequilibrium radical overshoot is enormous at the rather moderate temperatures that are found in this lean premixed flame. Only with CO nearing its equilibrium value are the O atom concentrations reasonably well approximated by the partial equilibrium assumption. Obviously, at the prevailing conditions NO_x models based upon partial equilibrium O atom concentrations would severely underpredict the production of Zeldovich NO_x in the flame front. The NO_x prediction scheme introduced in this study should be significantly more precise than any such model, even if predictions of CO concentrations are with the available combustion model only qualitatively correct.

Production rates of prompt/nitrous NO_x—which are also input to the NO_x postprocessor in the form of look-up tables—are plotted versus CO mass fraction in Fig. 4. Please note that all concentrations are on a wet, actual O_2 basis, except where noted otherwise. The prompt/nitrous NO_x production rate of the one-dimensional premix flame (filled diamonds) behaves essentially[5] as expected, i.e., there is a strong production rate peak in the flame zone (upper right corner of the diagram), the production diminishes rapidly in the burn-out zone, and is negligibly small near equilibrium conditions (left side of the diagram).

The production of Zeldovich NO_x (open diamonds) also peaks in the CO-rich flame zone with a value of about one fifth of the maximum prompt/nitrous NO_x rate. Zeldovich NO_x production diminishes less rapidly in the burn-out zone and—in this particular case—forms near equilibrium at approximately 1 ppm/s, which is more than five times faster than prompt/nitrous production.

[5] The sudden dip near the maximum CO concentration in the lower branch of the nitrous/prompt curve, which corresponds to the beginning of the reaction zone of the flame, is due to a combination of strong molecular diffusion and NO $\leftrightarrow$ NO_2 conversion.

Also shown in Fig. 4 are production rates from PSR/PFR reactor models at lean blowout (PSR residence time $\tau \approx 0.2$ ms), which provides correlations over the greatest range of CO space, and with a larger PSR with $\tau \approx 0.4$ ms. The larger reactor's volume was set in such a way that the Double Cone Burner's lean blowout equivalence ratio $\phi \approx 0.4$ was reproduced by the reactor model. It can be seen that the production rates taken from the two chemical reactor models are quite close to each other and the one-dimensional premix flame values, except near the CO peak, where differences up to a factor of two are observed. The CO–O correlations produced by PSR/PFR models differ even less from their one-dimensional premix counterparts (not shown). This is a remarkable observation, considering that the PSR/PFR and laminar flame models correspond to turbulent Damköhler numbers equal to zero and infinity, respectively.

This apparent universality has been further examined using several chemical reactor zone arrangements, including cases with multiple PSRs and PFRs, and cases with recycle zones included. These results, which are described by Nicol et al. (1994), show (for given λ, p, and inlet temperature) that the dependence on the CO concentration of the O atom and N_2O concentrations, and the rate of NO production, is essentially independent of the chemical reactor zone arrangement selected. The provisions are that the combustion process is adiabatic (or nearly adiabatic), and that the combustion process is not quenched.

Postprocessor Results

Two-dimensional contour plots of total NO_x production rate and concentration obtained with the postprocessor are shown in Fig. 5. (Operating conditions are air excess ratio $\lambda = 2.1$, inlet temperature $T_i = 630$ K, pressure $p = 1$ bar.) The production rate rises very steeply on the cold side of the flame front, and a finer grid would be required to resolve this region adequately. Peak production rates in the flame are about 1000 ppm/s; in the burn-out zone the production rate decreases rapidly to very small values around 20 ppm/s. Figure 5 confirms our expectations that the formation of oxides of nitrogen is essentially restricted to the flame zone, where significant concentrations of radicals are found in lean premixed combustion. Only near the burner axis, where residence times in the vortex breakdown and its wake are longer than elsewhere, is postflame NO_x formation noticeable.

A comparison of measured and predicted NO_x concentrations (dry, 15 percent O_2) at the centerline, 400 mm downstream of the burner exit, is presented in Fig. 6. Mixture inlet temperature was 630 K for both experiment and computation with natural gas and methane used as fuels, respectively. Maximum experimental flame temperatures are near 1750 K due to the risk of flashback at richer conditions (see above). The agreement between experiment and postprocessor predictions is remarkable, especially if one considers that the postprocessor has no adjustable parameters. The disagreement at higher temperatures can be attributed to heat losses, which increase with temperature in the experiment, while they are not considered in the (adiabatic) computations. However, recalling the shortcomings of

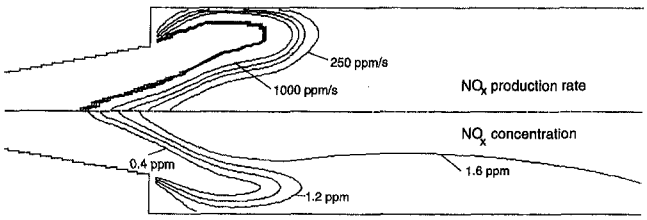


Fig. 5 Contour lines of NO_x production rates (top) and NO_x concentration (bottom, contour line spacing equals 0.4 ppm)

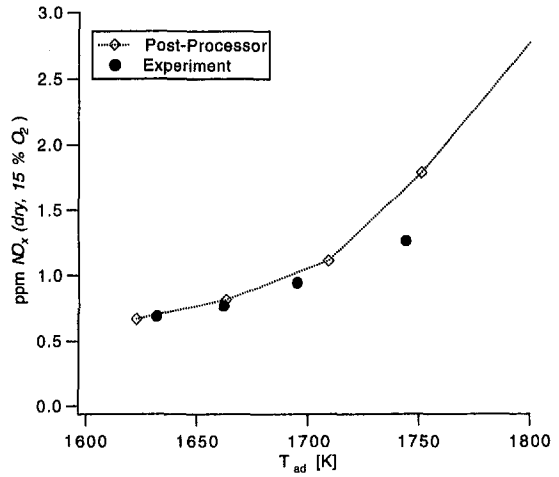


Fig. 6 Measured (●) and computed (– –) NO_x concentrations versus adiabatic flame temperature

the heat release and CO modeling discussed in Section 3, and the neglect of turbulent fluctuations in the present implementation of the postprocessor, we must concede that the level of agreement shown in Fig. 6 is most likely at least to some extent coincidental. Nevertheless, the postprocessor is conceptually sound, and reproduces qualitative aspects of NO_x formation in lean premixed flames correctly. A more reliable turbulent combustion model would certainly increase our confidence in the NO_x model's quantitative predictions.

A few comments concerning the influence of temperature fluctuations on NO_x formation in lean premixed combustion shall conclude this section. The formation of Zeldovich NO_x is extremely sensitive to temperature, due to the high activation energy of the rate limiting reaction (7). A Zeldovich NO_x formation rate based on mean temperature can therefore be very different from the actual mean formation rate. This is particularly so in diffusion flames, where significant NO_x production occurs only in near stoichiometric regions where high peak temperatures and strong fluctuations are realized. For such flames, the turbulent fluctuations of temperature must be taken into account in a meaningful way. In the context of a NO_x postprocessor, this has been attempted—see, e.g., Missaghi et al. (1990) and Dupont et al. (1993)—by introducing a presumed probability distribution function (pdf) for temperature, which is based on the first two moments, i.e., the mean and the variance. If the variance is not provided by the main combustion model, it may be determined approximately from the mean gradient. Species concentrations enter the rate expressions linearly or with small exponents, therefore their mean values may be used under certain circumstances. It is also possible to use a presumed pdf for species, with species and temperature fluctuating independently from each other. This methodology has been applied by Missaghi et al. (1990) and Dupont et al. (1993) to turbulent diffusion and (partially) premixed flames with some success.

The situation is quite different in (perfectly) premixed combustion. First, the nitrous and prompt NO_x formation paths are significantly less sensitive to temperature than the Zeldovich path. Second, the momentary temperature is always restricted to values below the adiabatic flame temperature: less than 1800 K at conditions of interest. Finally, the highest radical concentrations and the highest NO_x formation rates are found at intermediate temperatures (compare Fig. 4). These observations suggest that the error incurred by neglecting temperature fluctuations altogether is insignificant. This supposition is supported by Fig. 7, which shows the computed NO_x formation rate and the temperature gradient sampled from about 500 computational cells in the vicinity of the flame front. The temperature gradient ∇T has been nondimensionalized by the turbulent length scale l_T and the temperature T and then normalized to the interval [0, 1]. It is remarkable that the NO_x formation predicted by the postprocessor is considerable only in regions where gradients of temperature, and thereby also the fluctuations of temperature, are relatively small. This behavior is consistent with our conjecture.

Summary

A NO_x prediction scheme suitable for lean-premixed combustion of methane or natural gas at atmospheric and gas turbine conditions has been developed. The model is implemented as a postprocessor, i.e., the NO transport equation is solved after the "main" turbulent combustion simulation has been completed. Source terms for NO are derived from the "main" solution via correlations based on mean CO concentration and temperature. These correlations are computed with one-dimensional laminar flame or PSR/PFR models with detailed chemistry and provided to the NO_x module in the form of a look-up table. The correlations are *not* universal; for every operating condition (inlet temperature, pressure, air excess, etc.) of interest an appropriate look-up table must be prepared. However, the correlations are remarkably universal inasmuch as unstrained laminar flame or PSR/PFR models, corresponding in some sense to infinite and zero turbulent Damköhler number, produce very similar correlations.

At atmospheric conditions, which this study is restricted to, correlations for nitrous/prompt NO_x formation rate and O atom mass fraction, necessary to predict Zeldovich NO_x, are supplied. At high pressures, prompt NO_x may be neglected, and a global expression for the nitrous NO_x rate may be used, which is based on N_2O and O. The effect of turbulent fluctuations on NO_x formation rates is not considered in the present implementation. It has been argued, however, that these effects are possibly quite small in lean-premixed combustion.

As a test case, combustion in a swirl-stabilized gas turbine burner at atmospheric pressure has been simulated. Dimensional estimates suggest that this burner operates in a regime where both turbulent mixing and kinetic rate limitations must be considered ($Da \approx 1$, $Ka > 1$). An extended eddy break-up turbulent combustion model has been developed and implemented, which allows one to consider kinetic limitations in a physically reason-

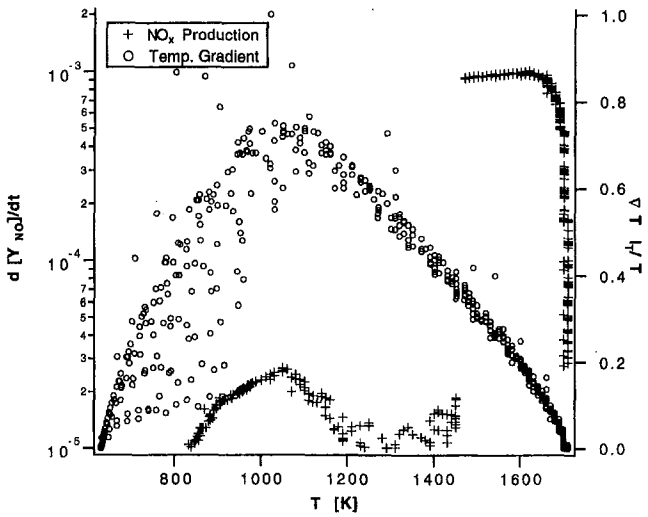


Fig. 7 Computed NO_x production rate and normalized temperature gradient versus temperature T from the vicinity of the flame front. The sudden jump in production rate corresponds to the transition from the low-T to the high-T branch of Fig. 4 and occurs because the CO concentration reached in the CFD model is lower than the maximum CO concentration from the detailed model.

able manner. Qualitatively correct distributions of flow, temperature and CO mass fraction have been obtained.

For the leanest conditions considered, where effects of nonadiabaticity are smallest, the NO_x postprocessor predicts concentrations of oxides of nitrogen with excellent quantitative agreement with experiment. Also, the production of NO_x occurs mainly in the flame zone, which is in agreement with our present understanding of NO_x formation in lean-premixed combustion. The model shall be applied to high-pressure test cases in the near future; the possibility of including effects of turbulent fluctuations and imperfect premixing is being studied.

References

Borghi, R., 1988, "Turbulent Combustion Modelling," *Prog. Energy Combustion Sci.,* Vol. 14, p. 1.

Coffee, T. P., 1985, "On Simplified Reaction Mechanisms by Oxidation of Hydrocarbon Fuels in Flames by C. K. Westbrook and F. T. Dryer," *Combustion & Flame,* Vol. 43, pp. 333–339.

Dupont, V., Pourkashanian, M., and Williams, A., 1993, "Modelling of Process Heaters Fired by Natural Gas," *J. Inst. Energy,* Vol. 66, pp. 20–28.

Hanson, R. K., and Salimian, S., 1984, "Survey of Rate Constants in H/N/O System," in: *Combustion Chemistry,* W. C. Gardiner, ed.

Hoffmann, D., Kampmann, S., Leipertz, A., Polifke, W., Döbbeling, K., and Sattelmayer, Th., 1994, "An Experimental Investigation of the Spatial Structure of Highly Turbulent Lean-Premixed Flames," submitted to *Combustion & Flames.*

Joshi, N. D., Epstein, M. J., Durlak, S., Marakovits, S., and Sabla, P. E., "Development of a Fuel Air Premixer for Aeroderivative Dry Low Emissions Combustors," ASME Paper No. 94-GT-253.

Kee, R. J., Grcar, J. F., Smooke, M. D., and Miller, J. A., 1985, "A Fortran Program for Modeling Steady Laminar One-Dimensional Premixed Flames," Sandia Report No. SAND85-8240 UC-401.

Kee, R. J., Rupley, F. M., and Miller, J. A., 1989, "Chemkin-II: A Fortran Chemical Kinetics Package for the Analysis of Gas Phase Chemical Kinetics," Sandia Report No. SAND89-8009B UC-706.

Libby, P. A., and Williams, F. A., 1994, *Turbulent Reacting Flow,* Academic Press, London.

Miller, J. D., and Bowman, C. T., 1989, "Mechanism and Modelling of Nitrogen Chemistry in Combustion," *Prog. Energy Combustion Sci.,* Vol. 15, pp. 287–338.

Missaghi, M., Pourkashanian, M., Williams, A., and Yap, L., 1990, *Proceedings of the American Flame Days Conference,* USA.

Nicol, D. G., Steele, R. C., Marinov, N. M., and Malte, P. C., 1995, "The Importance of the Nitrous Oxide Pathway to NO_x in Lean-Premixed Combustion," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 117, pp. 100–111.

Nicol, D. G., Malte, P. C., and Steele, R. C., 1994, "Simplified Models for NO_x Production Rates in Lean-Premixed Combustion," ASME Paper No. 94-GT-432.

Patankar, S. V., 1980, *Numerical Heat Transfer and Fluid Flow,* Hemisphere Publishing, New York.

Phillip, M., Hoffmann, S., Habisreuter, P., Lenze, B., and Eickhoff, H., 1992, "Experimental and Numerical Study Concerning Stabilization of Strongly Swirling Premixed and Nonpremixed Flames," *24. Symp. (Int.) on Combustion,* The Comb. Inst., Pittsburgh, p. 361.

Polifke, W., Döbbeling, K., and Sattelmayer, Th., 1995, "A Computational Model for Lean-Premixed, Highly Turbulent Combustion," presented at CIMAC 21st International Congress on Combustion Engines, Interlaken, Switzerland.

Sattelmayer, Th., Felchlin, M. P., Haumann, J., Hellat, J., and Styner, D., 1992, "Second-Generation Low-Emission Combustors for ABB Gas Turbines: Burner Development and Tests at Atmospheric Pressure," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 114, pp. 118–125.

Sloan, D. G.,. Smith, P. J., and Smooth, L. D., 1986, "Modeling of Swirl in Turbulent Flow Systems," *Prog. Energy Combustion Sci.,* Vol. 12, pp. 163–250.

Steele, R. C., Jarrett, A. C., Malte, P. C., Tonouchi, J. H., and Nicol, D. G., 1995, "Variables Affecting NO_x Formation in Lean-Premixed Combustion," ASME Paper No. 95-GT-107; ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 119, Jan. 1997.

Warnatz, J., and Maas, U., 1993, *Technische Verbrennung,* Springer Verlag, Berlin.

Westbrook, C. K., and Dryer, F. L., 1980, "Simple Reaction Mechanisms for the Oxidation of Hydrocarbon Fuels in Flames," *Com. Sci. Technology,* Vol. 27, pp. 31–43.

G. J. Sturgess[1]
Pratt & Whitney,
East Hartford, CT 06108

S. P. Heneghan

M. D. Vangsness

D. R. Ballal

University of Dayton,
Dayton, OH 44444

A. L. Lesmerises
WRDC,
Wright-Patterson Air Force Base, OH 44135

Lean Blowout in a Research Combustor at Simulated Low Pressures

A propane-fueled research combustor has been designed to represent the essential features of primary zones of combustors for aircraft gas turbine engines in an investigation of lean blowouts. The atmospheric pressure test facility being used for the investigation made it difficult to approach the maximum heat release condition of the research combustor directly. High combustor loadings were achieved through simulating the effects on chemical reaction rates of subatmospheric pressures by means of a nitrogen diluent technique. A calibration procedure is described, and correlated experimental lean blowout results are compared with well-stirred reactor calculations for the research combustor to confirm the efficacy of the calibration.

Introduction

A propane-fueled research combustor, Fig. 1, has been designed (Sturgess et al., 1992) and developed (Heneghan et al., 1990) to investigate lean blowouts in simulated primary zones of the combustors for aircraft gas turbine engines. The fundamental flow features of a gas turbine combustor primary zone, within which the flame is held, are generated by the geometrically simple design of the research combustor.

The research combustor consists of co-axial jets with a 29.7 mm ID central fuel jet surrounded by a 40 mm diameter annular air jet. The fuel is gaseous propane. The jets are located centrally in a 150 mm nominal diameter duct. A backward-facing step at the jet discharge plane completes the sudden expansion, giving a step height of 55 mm. The step provides a recirculation region that can stabilize the flame. The combustor test section incorporates flat quartz windows to provide optical access. Curved metal fillets, located in the corners of the test section, reduce and distribute the vorticity generation due to wall secondary flows, and so eliminate their impact on the bulk flowfield in the combustor. An orifice plate with a 45 percent blockage ratio forms the exit from the combustor.

When operated in a fuel-rich mode, the flame in the combustor is very stable and is anchored in the jet shear layers by a pilot flame attached to the step near the outer edge of the air supply tube. Fuel for this flame is recirculated from downstream by the step recirculation zone; ignition is by hot gases also recirculated. As the bulk equivalence ratio is reduced, the flame becomes less stable, and eventually reaches a point where the pilot flame becomes detached from the base region (lifts), and the entire flame structure becomes stabilized downstream, Fig. 2. Thus, there are two distinct operating modes for the combustor: a fully anchored flame, and a lifted flame.

As a consequence of the flame-lift, a significant degree of partial premixing of the reactants takes place upstream of the lifted flame position. When the equivalence ratio is further reduced, the flame becomes progressively less stable (Heneghan et al., 1990) in its lifted condition, and eventually flows out. This lean blowout limit at atmospheric pressure has been found to correspond to a fuel equivalence ratio (around 0.5) that is very close to the lean flammability limit of propane in air (Lewis and von Elbe, 1987). Due to the partial premixing, the combustor lean blowout performance characteristic has been observed to be consistent with that of well-stirred reactor data (Sturgess et al., 1992).

Lean blowout is strongly affected by chemical reaction rates. Application of chemical reaction rate theory to well-stirred reactors has resulted in the formulation of the gas loading parameter $\dot{m}/(V P^n F)$, where n is the effective order of the reaction, F is a temperature correction factor, V is the reactor volume, $\dot{m}$ is the oxidant mass flow rate, and P the operating pressure, against which the equivalence ratio at blowout is frequently plotted. The concept of the well-stirred reactor can provide a simplified description of practical combustors. Bragg (1953) originated the view that any reasonably efficient practical combustor should consist of an initially well-stirred reactor section for ignition and flame holding, followed by a plug-flow reactor section for burnout. Since the research combustor was designed specifically to reproduce the major features of an aircraft gas turbine combustor primary zone (Sturgess et al., 1992), it might be expected to conform to well-stirred reactor behavior. The lean blowout behavior supports this expectation. Therefore, the loading parameter can be used to define lean blowout performance.

The research combustor is being tested in a facility that operates at atmospheric pressure and that was limited to maximum propane flow rates of about 20 kg/h. Under these circumstances, sufficient loading to approach the maximum heat release condition of the research combustor cannot be achieved. However, to investigate the lean blowout phenomenon thoroughly, it is desirable to probe the flow fields for flames that are anchored and lifted at both lightly strained and heavily strained flame conditions. Heavily strained flames occur at high loadings, and in practice, result from combustor operation at a given airflow with low air inlet temperatures and low (subatmospheric) pressures. Reducing air inlet temperatures alone to achieve high loadings without choking the combustor requires very low temperatures, and this demands drying of the air to avoid severe icing problems. Given these difficulties, alternative methods of achieving heavily strained flames were needed.

The quantities involved in the loading parameter are those that influence the speed at which chemical activity converts reactants to products. When the combustor inlet conditions cannot be manipulated to achieve a desired loading, the chemical reaction can be influenced by some other means to yield a

[1] Current address: Innovative Scientific Solutions, Inc., Beavercreek, OH 45432.

Contributed by the International Gas Turbine Institute and presented at the 36th International Gas Turbine and Aeroengine Congress and Exposition, Orlando, Florida, June 3–6, 1991. Manuscript received at ASME Headquarters March 1991. Paper No. 91-GT-359. Associate Technical Editor: L. A. Riekert.

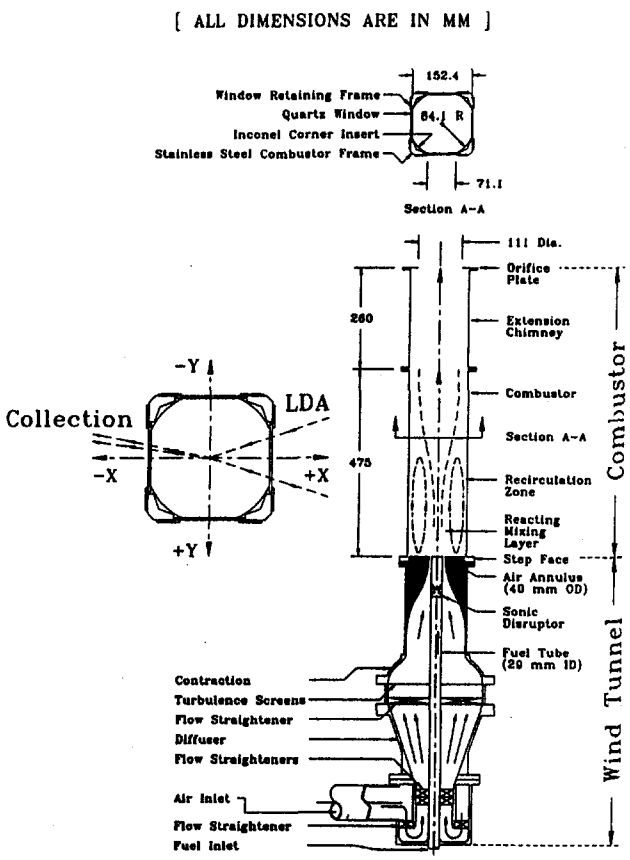


Fig. 1 Cross section of research combustor

similar rate of reactants conversion. If pressure is made the variable of influence, then a suitably altered chemical reaction would reflect the "simulated pressure."

Low pressures can be simulated by the introduction of an inert diluent that slows down the chemical reaction in approximately the same way that low pressures do. The addition of the diluent has two effects: first, it lowers the concentration of reactants, and second, it lowers the reaction temperature by virtue of its heat capacity. Thus, tests may be made at atmospheric pressure and high loading still achieved.

Such a simulation technique for achieving high loadings while operating at room temperature and atmospheric pressure was developed for the research combustor.

Simulation Technique

A number of techniques for simulating low combustion pressures are available, e.g., Lefebvre and Halls (1959), Lefebvre (1961), and Greenhough and Lefebvre (1956), using water as the diluent. However, water is inconvenient for the present purposes. It requires preheating of the inlet air to ensure complete vaporization of the droplets. Gaseous nitrogen has also been used with success (Norster, 1968), and does not require preheating. Although a direct calibration for nitrogen was not available, nitrogen as a diluent is extremely convenient for the present purposes, and was therefore selected.

Theoretical Background

Consider the combustion of the general hydrocarbon C_xH_y in air, and let the equivalence ratio ϕ be such that there is an excess of air, i.e., $\phi < 1.0$. Also, let the combustion efficiency be such that B is the fraction of fuel that is actually burned. The unburned fuel will not pass through the combustor unchanged, but will, in general, appear in the exhaust as carbon monoxide, hydrogen, and lower hydrocarbons. In modern combustors the lower hydrocarbons are always small in quantity under normal circumstances; further, the fast reaction rate of hydrogen suggests that it too, will not be a significant product. Therefore, a simple bimolecular global reaction can be proposed,

$$\phi C_xH_y + (x + y/4)O_2 + 79/21(x + y/4)N_2 = \alpha CO + \beta CO_2 + y/2\phi H_2O + 79/21(x + y/4)N_2 + (1 - \phi B)(x + y/4)O_2 \quad (1)$$

where all of the hydrogen in the fuel is burned to water vapor, and the inefficiency is reflected solely in the CO/CO_2 balance. This is reasonable because of the relatively slow conversion of CO to CO_2. Nitrogen is assumed inert. The molar fractions α and β are found from C/O balances.

For propane this reaction becomes,

$$\phi C_3H_8 + 5O_2 + 18.8095N_2 = 10\phi(1 - B)CO + \phi(10B - 7)CO_2 + 4\phi H_2O + 18.8095N_2 + 5(1 - \phi B)O_2 \quad (2)$$

Note that with this simple reaction mechanism CO_2 disappears from the products for a fractional conversion less than or equal to 0.7.

Let K be the mass ratio of diluent nitrogen to fuel. The reaction expression becomes

$$\phi C_3H_8 + 5O_2 + 18.8095\gamma N_2 = 10\phi(1 - B)CO + \phi(10B - 7)CO_2 + 4\phi H_2O + 18.8095\gamma N_2 + 5(1 - \phi B)O_2 \quad (3)$$

so that the additional mass of N_2 is $18.8095 \times 28\ (\gamma - 1)$.

Nomenclature

A = pre-exponential rate constant (may also include convenient unit conversion factors)
B = combustion efficiency
$C = E/\mathcal{R}$
E = activation energy
F = temperature correction factor
K = mass ratio of excess nitrogen to fuel
m = volumetric fractional concentration
$\dot{m}$ = mass flow rate
$\dot{m}'''$ = volumetric rate of oxidant mass consumption
n = effective order of global chemical reaction
P = pressure
R = global reaction rate expression
$\mathcal{R}$ = universal gas constant
T = reaction temperature
V = volume associated with combustion
α = molar fraction of CO in products of reaction
β = molar fraction of CO_2 in products of reaction
γ = defined by Eq. (4)
ϕ = fuel/air equivalence ratio

Subscripts

air = air
f = carbon monoxide in products of reaction, fuel
LBO = lean blowout
N_2 = excess nitrogen as diluent
o = oxygen in products of reaction
Tot = inlet sum of air, propane, and excess nitrogen

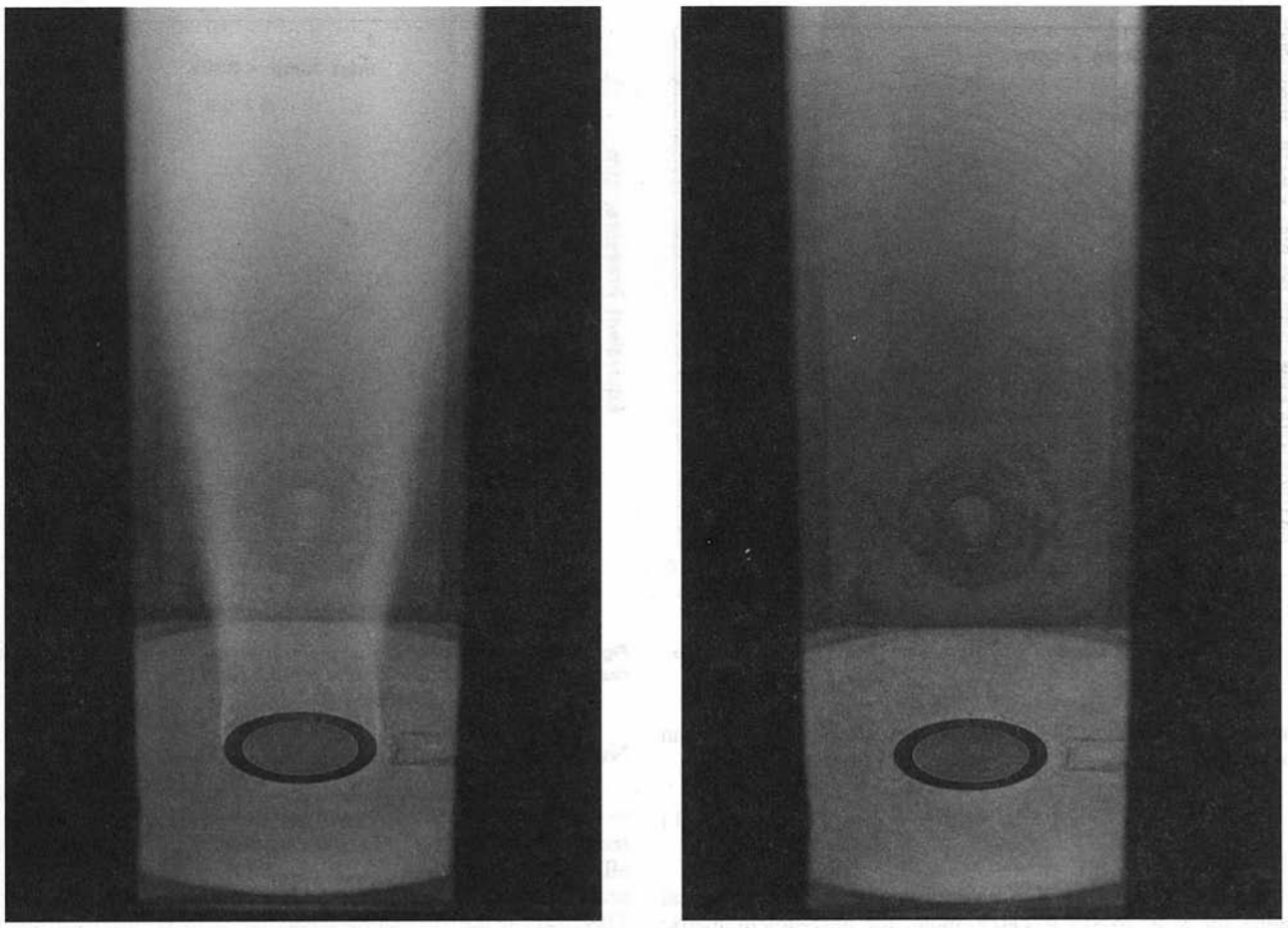


Fig. 2 Photographs of anchored and lifted flame operation

Hence,

$$\gamma = (\phi K/11.9697 + 1) \tag{4}$$

Equations (3) and (4) together provide the volumetric fractional concentrations of oxidant and fuel present in the exhaust, i.e.,

$$m_f = \frac{10\phi(1 - B)}{G} \tag{5}$$

where

$$G = 10\phi(1 - B) + \phi(10B - 7) + 4\phi + 18.8095\gamma + 5(1 - \phi B)$$

and

$$m_o = \frac{5(1 - \phi B)}{G} \tag{6}$$

From the kinetic theory of gases a reaction rate expression in Arrhenius form can be derived for bimolecular, single-step reactions. Longwell et al. (1953) have proposed that for lean mixtures this takes the form,

$$\phi B\dot{m}''' = Am_o m_f P^2 T^{-3/2} \exp(-E/\mathcal{R}T) \tag{7}$$

where $\dot{m}'''$ is the volumetric rate of oxidant mass consumption.

Although Longwell and others have proposed fractional indices in Eq. (7) to accord with the results of various experiments, for the present purposes a pure second-order reaction in pressure and first order in concentrations is taken. In Eq. (7), E is a reaction rate constant (activation energy), and A can incorporate a pre-exponential rate constant as well as various unit conversion factors and other constant terms as may be convenient. T is the reaction temperature, and $\mathcal{R}$ the universal gas constant.

Rearranging Eq. (7) into a form incorporating the loading parameter yields,

$$\frac{\dot{m}_{\text{air}}}{VP^2} = \frac{Am_o m_f}{\phi B} \frac{1}{T^{3/2} \exp(E/\mathcal{R}T)} \tag{8}$$

Combining Eqs. (5), (6), and (8) gives,

$$\frac{\dot{m}_{\text{air}}}{VP^2} = \frac{A(1 - B)}{B} \times \frac{2(1 - \phi B)}{(\phi(1.4 - B) + 79/21\gamma + 1)^2} \frac{1}{T^{3/2} \exp(E/\mathcal{R}T)} \tag{9}$$

Equation (9) represents the appropriate global kinetic expression for the reaction rate of propane in air with additional nitrogen. With K equal to zero, it also represents a similar expression for propane in pure air.

A reaction rate simulation can be obtained if

$$\left(\frac{\dot{m}_{\text{air}}}{VP^2R}\right)_{\text{simulation}} = \left(\frac{\dot{m}_{\text{air}}}{VP^2R}\right)_{\text{desired}} \tag{10}$$

where R is the appropriate reaction rate expression from the right-hand side of Eq. (9), i.e., with $K > 0$ and $P = 1$ atm for simulation and $K = 0$ and $P < 1.0$ atm for desired reaction rate. If it is taken that the presence of the diluent has no effect

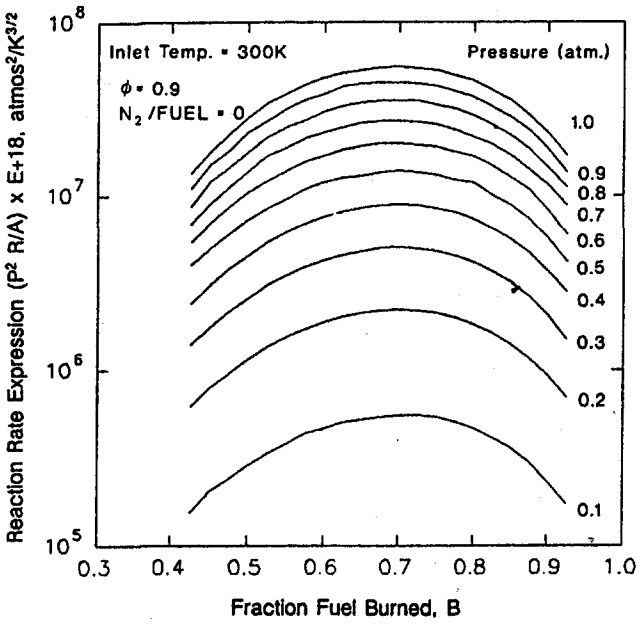


Fig. 3 Variation of reaction rate expression (Eq. (11)) with pressure

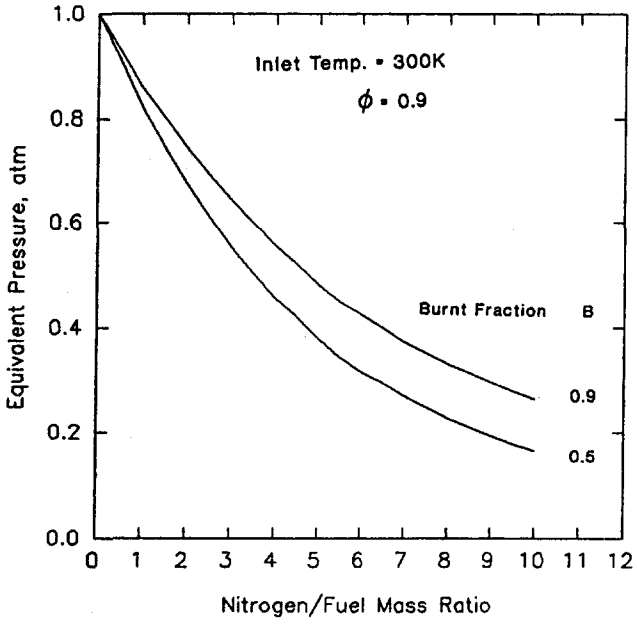


Fig. 5 Dependency of simulated pressure on diluent addition for a global reaction rate based on indices for ϕ = 1.0 (Eq. (7))

on the rate constants, the identity represented by Eq. (10) can be written,

$$\left(\frac{P^2R}{A}\right)_{\text{simulation}} = \left(\frac{P^2R}{A}\right)_{\text{desired}} \tag{11}$$

Thus, A does not have to be described explicitly. A typical value for E is 26,613 g·cal/g·mole, for near-stoichiometric mixtures (Clarke et al., 1960).

Figures 3 and 4 may be used to compare the left- and right-hand sides of the identity for an equivalence ratio of 0.9 and inlet temperature of 300 K. Figure 3 shows the effect of pressure on propane/air combustion, while Fig. 4 shows the effect of nitrogen addition on the combustion at one atmosphere pressure. The similarities of the two curves are obvious.

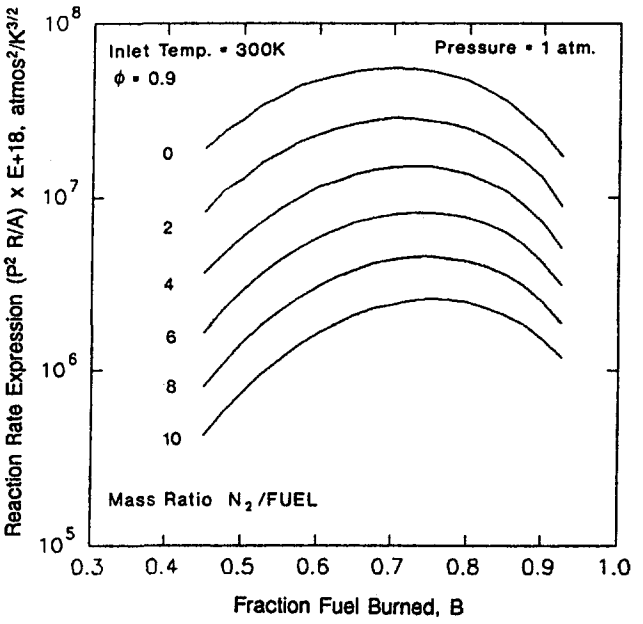


Fig. 4 Variation of reaction rate expression (Eq. (11)) with diluent addition

Need for Calibration

The simplified reaction mechanism and rate expression used to derive the simulation are only a convenient representation of reality. A given rate expression cannot be expected to represent all the actual reactions taking place adequately so that the expression is appropriate over wide ranges of operating conditions. Therefore, if the simulation technique is to be used for other than merely comparative purposes, some form of calibration is required.

Various global Arrhenius rate expressions have been published in the literature to represent propane/air combustion at different conditions, e.g., Greenhough and Lefebvre (1956), Clarke et al. (1960), and Herbert (1960). One that is more appropriate for low equivalence ratios (around 0.5) is given by

$$\dot{m}''' = Am_o^{0.5}m_f^{0.5}PT^{-0.5}\exp(-56{,}600/\mathcal{R}T) \tag{12}$$

This has the form of Eq. (7), but different fractional indices and rate constants. When the simulation is based on Eq. (12) rather than Eq. (7), considerable differences result.

As is to be described, comparison of curves like Figs. 3 and 4 can be used to obtain the relationships between nitrogen to fuel mass ratio and the equivalent reaction pressure. This has been done in Figs. 5 and 6 for an equivalence ratio of 0.9 and a mixture initial temperature of 300 K, using simulations based on Eqs. (7) and (12), respectively. For given combustion efficiency and nitrogen/fuel mass ratio, the differences in equivalent pressure from the two curves are marked. This illustrates vividly the need for calibration of the technique.

Calibration

An ideal calibration would be based on a comparison of the simulated pressure technique against true low-pressure tests in the same reactor. Unfortunately, such a direct approach was not immediately available for the research combustor. Therefore, an indirect approach is used. This maintains the bimolecular reaction mechanism Eq. (3), unchanged, but calibrates the Arrhenius reaction rate expression that goes with it.

Fortunately, a calibration of the Arrhenius rate expression for propane/air/excess nitrogen systems has already been carried out by Kretschmer and Odgers (1972), and their work has been utilized here.

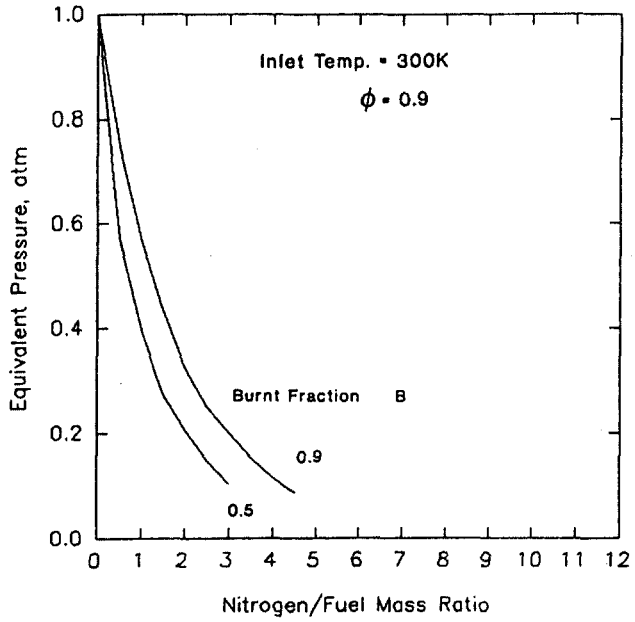


Fig. 6 Dependency of simulated pressure on diluent addition for a global reaction rate based on indices for ϕ = 0.5 (Eq. (12))

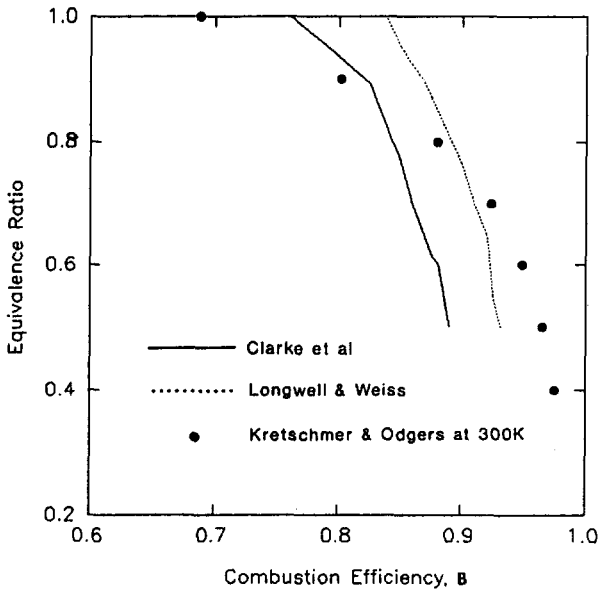


Fig. 7 Variation of combustion efficiency at blowout with equivalence ratio, from several sources

Kretschmer and Odgers used for their calibration experimental results from spherical stirred reactors and gas turbine combustors, burning fuels from propane to aviation kerosene over a wide range of operating conditions including pressures from 0.1 to 5.4 atm and inlet temperatures from 200 K to 900 K; equivalence ratios equal to and less than unity were included. Analysis of these data resulted in a rate expression giving the calibrated form for Eq. (8) as

$$\frac{\dot{m}_{\text{air}}}{VP^2\phi} = \frac{A(m_o m_f)^{\phi}}{\phi B} \frac{1}{T^{2\phi-0.5}\exp(C/T)} \qquad (13)$$

where C is $E/\mathcal{R}$. For an equivalence ratio of unity, Eq. (13) is identical to Eq. (8); while for an equivalence ratio of one-half, it is identical with the form of Eq. (12). Now, however, the exponential rate constant contained in C is a complex function of equivalence ratio and reactant inlet temperature. This variation arises because of the empirical character given to the rate equation.

Figure 7 gives the dependency on ϕ of B at blowout for 300 K inlet temperature as produced by Kretschmer and Odgers. This dependency is compared with earlier recommendations also given by Kretschmer and Odgers (1972).

Unfortunately, Kretschmer and Odgers assumed a reaction mechanism that was different from that given in Eq. (2). Hence, their expression for m_f was different from that given in Eq. (5). Therefore, it was necessary to derive a variation of C, containing the exponential rate constant, with ϕ that was appropriate to the present reaction mechanism.

The derivation of an appropriate C was done through Eq. (13) so that,

$$C = T\,[\ln\,(m_{f,new}/m_{f,\text{Kretschmer}})\,\exp(C/T)_{\text{Kretschmer}}] \qquad (14)$$

A comparison between the variations of two C's with equivalence ratio for an inlet temperature of 300 K is shown in Fig. 8. Although the differences are small, it should be remembered that C is eventually to be used in an exponential relationship. The variation of C for the present reaction mechanism is conveniently represented by the polynomial

$$C = 14{,}926.5 + 73{,}774.7\phi - 109{,}826\phi^2 + 35{,}933.3\phi^3 \qquad (15)$$

which has a correlation coefficient of 0.99998 and a maximum error of $-54.3K$ in the range $0.4 < \phi < = 1.0$.

Equations (13) and (15) establish the form of the global reaction rate, and relate it to the selected reaction mechanism.

The calibration of simulated pressure may now be constructed for a range of equivalence ratios. The appropriate reaction rate expressions, one with variation in pressure and the other with variation in diluent mass flow rate, can be compared as functions of pressure and nitrogen/fuel mass ratio, respectively. The intersection of these expressions represents a common value of reaction rate expression indicating equivalency of pressure and nitrogen/fuel mass ratio. This enables the relationship between diluent mass flow rate and simulated pressure to be established. Figure 9 gives an example of this technique for a ϕ of 0.8; Fig. 10 shows the resulting calibration curves for an inlet temperature of 300 K.

Figure 10 can be considered valid for well-stirred and partially stirred reactors. Although the empirical adjustment of the

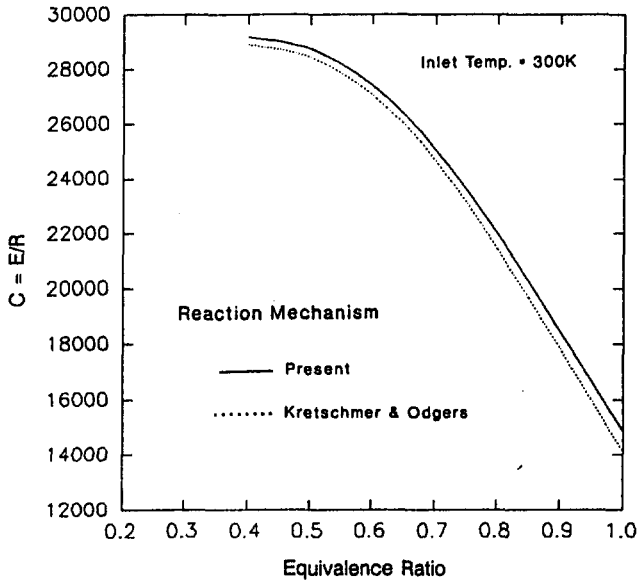


Fig. 8 Comparison of dependencies of activation energy on equivalence ratio for Kretschmer and Odgers reaction mechanism and present reaction mechanism

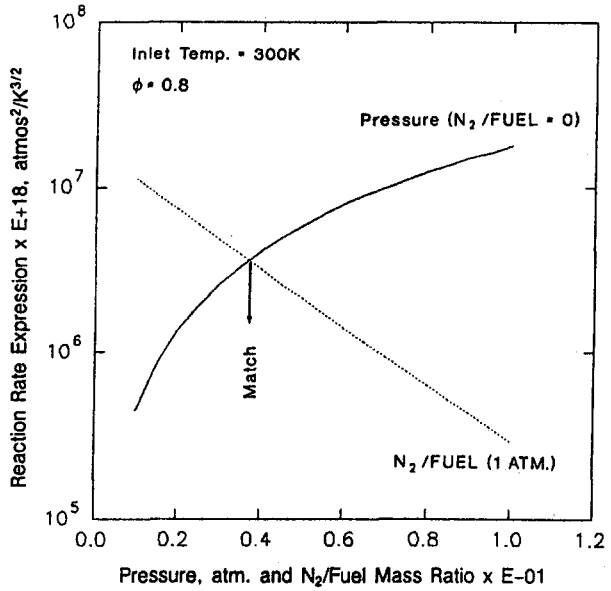


Fig. 9 Relationship between low pressure and simulated reaction rates for $\phi = 0.8$

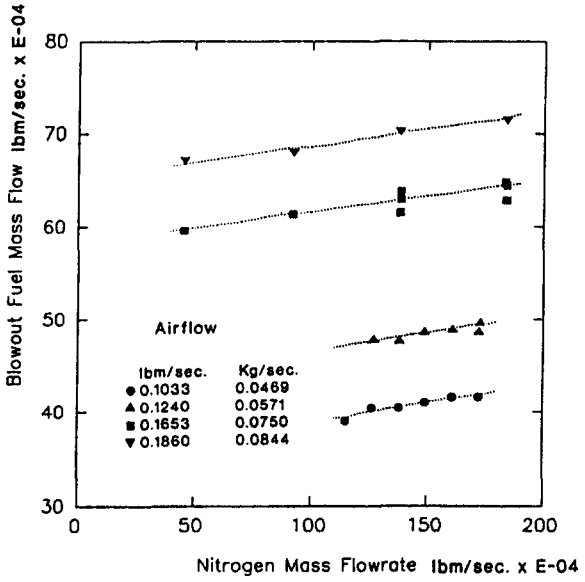


Fig. 11 Relationship between fuel flow at blowout and nitrogen flow at several airflows in the research combustor

global Arrhenius reaction rate expression may be considered as being valid for a wide variety of hydrocarbons burning in air, the simulation calibration is for propane only. Further, it is limited to fuel-lean mixtures, as the figure indicates. The calibration shown is for 300 K inlet temperatures; Swithenbank (1974) contains sufficient information to construct curves for inlet temperatures from 200 K to 900 K.

Experimental Results

After the ignition sequence was completed and the desired airflow established with the anchored flame condition (Sturgess et al., 1992), gaseous nitrogen was introduced into the air supply line to the research combustor. This was done far upstream so that the excess nitrogen was uniformly mixed with the air entering via the annular air jet surrounding the fuel jet. The nitrogen flow rate was set at a desired level, and a blowout sequence obtained by reducing the fuel flow rate until the flame was extinguished. Ignition was re-established, the airflow reset at the previous value in conjunction with the fuel flow, a new nitrogen flow rate was selected, and blowout was again obtained.

This procedure was repeated until a maximum nitrogen flow condition was reached. The combustor has been operated with nitrogen mass flow rates up to 55 percent of the air mass flow rate. No difficulties were encountered at this extreme condition. Repeatability and hysteresis were checked through obtaining blowouts by turning down the nitrogen flow rate range, and also by holding nitrogen and fuel flow rates constant while the airflow rate was increased. In addition, test points were repeated on different days with a different observer.

Examples of the basic results are displayed in Fig. 11, and show at a given airflow that increasing the nitrogen flow rate linearly increases the fuel flow rate at which a blowout takes place. In the figure, the multiple tests points shown at a given nitrogen flow rate for the 0.653 lbm/sec (0.075 kg/s) airflow represent the day-to-day repeatability of the blowout. Increasing the combustor airflow increases the fuel flow at which blowout occurs, but does not change its dependency on the nitrogen flow rate.

Figure 12 gives an indication of the range of equivalent pressures represented by the nitrogen simulation, for various airflows. It ranges from just less than one atmosphere down to one-tenth of an atmosphere for these tests. The equivalent pressures were obtained from Fig. 10 for the blowout conditions given in Fig. 11. When the equivalent pressures are displayed against the mass ratio of nitrogen to fuel at blowout in this form, the data for the different airflows collapse onto a single curve representing the blowout characteristic for the research combustor. Note that for nitrogen to fuel mass ratios greater than about 5, there is a loss of effectiveness of the technique.

When excess nitrogen is introduced into the air supply, the visual appearance of the flame and its behavior does not change from that observed for zero excess nitrogen (Sturgess et al., 1992; Heneghan et al., 1990). Figure 13 gives a description of the flame behavior at constant airflow as the nitrogen flow rate is increased. Visually, there is no change in the equivalence ratio at which flame lift takes place. The equivalence ratio for lean blowout does increase, in accord with the behavior of Fig. 11. Blowout with excess nitrogen follows the same sequence

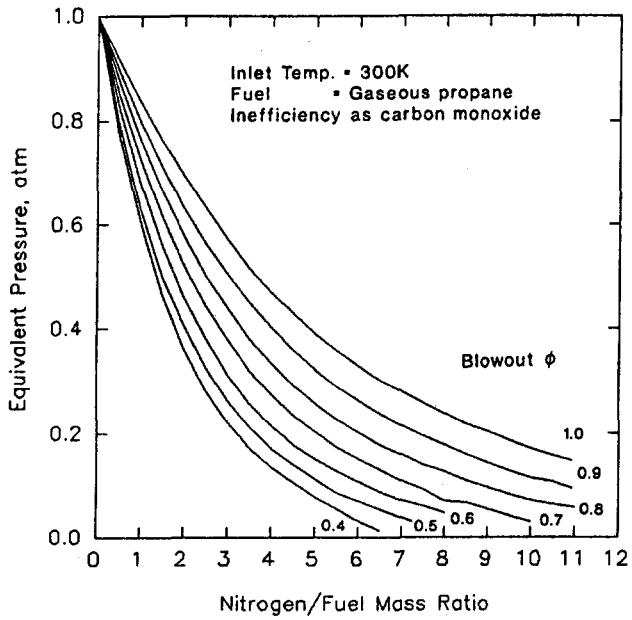


Fig. 10 Nitrogen-simulation calibration curve at 30 K

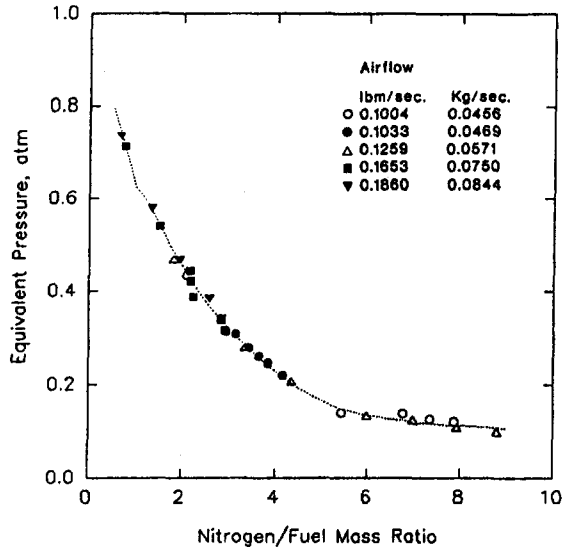


Fig. 12 Relationship between equivalent pressure and nitrogen to fuel mass ratio at blowout in the research combustor

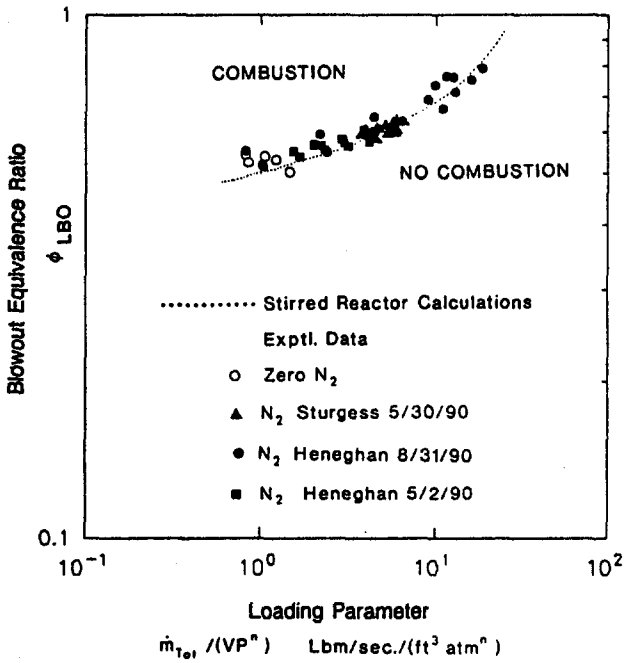


Fig. 14 Correlated blowout data for the research combustor, and comparison with the calculated well-stirred reactor blowout characteristic for the research combustor

of events as was observed for zero excess nitrogen (Sturgess et al., 1992; Heneghan et al., 1990).

Correlation of Experimental Data

As described above, the equivalence ratio at blowout, ϕ_{LBO}, of a well-stirred reactor can be related to the gas loading parameter, $\dot{m}/VP^nF$, for the combustor. Here, the mass flow rate is interpreted as $\dot{m}_{Tot}$, which is the sum of the air, excess nitrogen, and fuel mass flow rates. The fuel is included in this sum since its volume is not insignificant; the nitrogen flow rate is generally several times that of the fuel. Thus, the residence time in the reactor is materially affected by these flows.

For the present experiments the temperatures of the reactants were not varied, and the values for air, excess nitrogen and propane were sensibly equal. Therefore, the temperature correction factor F, is taken as unity as a matter of convenience.

The pressure is interpreted as the effective pressure, and depends on the quantity of excess nitrogen flowing, as given by Fig. 12.

The (constant) volume is taken as the total combustor volume minus the volume of the combustor associated with flame lift.

The apparent order of the reaction, n, is taken as equal to ϕ for lean mixtures. However, this is for propane/air systems. When a diluent is present, the order is reduced. The expression,

$$n = \frac{2\phi_{LBO}}{(1 + \dot{m}_{N2}/\dot{m}_{air})} \tag{16}$$

is used to account for this. When no excess nitrogen is present, Eq. (16) reverts to the pure air form.

In Fig. 14 the data in terms of ϕ_{LBO} and $\dot{m}_{Tot}/VP^n$ are plotted in logarithmic form. The measured data define most of the lean portion of the stability loop for the research combustor. The data cover points at atmospheric pressure with zero excess nitrogen, as well as those at atmospheric pressure with excess nitrogen; a variety of airflows were used.

The figure shows that the loading parameter correlates the blowout equivalence ratios very well. This can be appreciated when considering the repeatability shown in Fig. 11. The blowout data extend over three orders of magnitude of loading parameter, and range from a blowout equivalence ratio of around 0.5 to around 0.8.

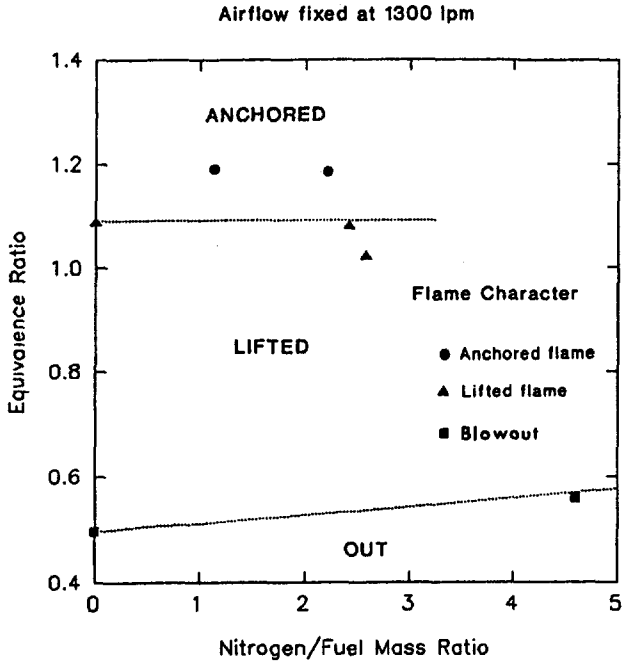


Fig. 13 Flame behavior in the research combustor at fixed airflow as nitrogen to fuel mass ratio is varied

Verification of Calibration

Use of the simulation calibration curves for effective pressure (Fig. 12) resulted in a very satisfactory correlation of experimental lean blowout data obtained at constant true pressure for wide ranges of air and excess nitrogen flows. Fortunately, since the research combustor with zero excess nitrogen appeared to behave like a well-stirred reactor (Sturgess et al., 1992; Heneghan et al., 1990), an independent check of the calibration can be made through the use of well-stirred reactor theory.

Swithenbank (1974) presented a dissipation gradient approach for defining perfectly stirred regions of combustors. This

approach was applied to the research combustor, using a computational fluid dynamics (CFD) calculation to provide the turbulence characteristics.

Based on the dissipation gradient analysis, the rapid-stirring region of the research combustor was identified as the volume enclosing the highly turbulent portion of the flow. It was found that the surface contour enclosing 96 percent of the turbulence kinetic energy and 99 percent of its rate of dissipation corresponded to a surface over which the total dissipation gradient had a value of 10.0, which is 100 times the value recommended by Swithenbank. This well-stirred region extended over most of the length of the combustor, with exceptions of portions of the step recirculation zone and the stagnation region on the upstream face of the orifice plate placed at the combustor exit.

While the turbulence characteristics of a flow can be used to define a region of rapid mixing, this alone is not a sufficient condition for a well-stirred reactor. For stable reaction to take place within the rapid mixing region, the fuel and air mixture must locally fall inside the flammability limits, a source of continuous ignition (mixture at exit temperature) must be present, and local flow velocities must not exceed the local turbulent burning velocity.

CFD analysis was again used to provide necessary information to determine places within the defined rapid mixing region where stable combustion could take place for conditions close to blowout. On the basis of these calculations and the observations of the lifted flame position, which was found to remain roughly constant as equivalence ratio was reduced, a suitable well-stirred reactor volume was defined. This was 44 percent of the total combustor volume, and corresponded to a lifted flame mean position of about 40 cm (16 in.) from the central portion of the step, assuming that the flame completely filled the cross section. Visual observations of the flame indicated that on the centerline of the combustor the closest position to the step-plane of the lifted flame was about 16 cm (6 in.) downstream. Of course, the real flame never completely filled the cross section.

With a well-stirred reactor defined, the stirred reactor network code MARK2I by David Pratt and Brian Pratt was used to make calculations for the research combustor. The MARK2I code uses the CREK chemical kinetics code (Pratt and Wormeck, 1976) and a hydrocarbon reaction mechanism by Roberts et al. (1972). The research combustor near blowout was modeled in MARK2I as a single, perfectly stirred reactor with separate air and fuel inlets, and a single discharge of products; there was no recirculation of products. It was operated with gaseous propane as fuel and air as oxidant, with equal inlet temperature of 293 K. The volume was that defined above, although two other volumes were used for spot checks to see the sensitivity to volume. This was done because of the uncertainty associated with the flame position. (The results for the range of volumes were subsequently correlated successfully by the loading parameter, via the volume term incorporated in it.) The model was operated at a number of air mass flow rates, with several pressure levels of atmospheric and less, for each airflow. No cases with excess nitrogen were run.

Shown on the stability plot of Fig. 14 is a line representing the blowout characteristic calculated by the well-stirred reactor code. The line runs through the correlated experimental blowout data very nicely, thereby confirming that the research combustor is behaving as a well-stirred reactor. Furthermore, since the loading variation in the well-stirred reactor calculations was achieved through variations in mass flow rate and true operating pressure while the experimental data were correlated on the basis of mass flow rate and equivalent pressure, the agreement verifies the calibration (Fig. 12) of the nitrogen dilution technique.

Discussion

A technique has been presented whereby the lean reaction of gaseous propane in air at subatmospheric pressures can be simulated by approximating the subatmospheric pressure reaction rate with a reaction rate at atmospheric pressure in the presence of excess gaseous nitrogen as a diluent.

The technique has been calibrated through the assumption of a one-step global reaction mechanism for propane/air/excess nitrogen systems in which nitrogen remains inert, and where combustion inefficiency is accounted for solely through the CO/CO_2 balance in the products. A global reaction rate expression in Arrhenius form for a bimolecular, single-step reaction was calibrated using experimental data from the literature. Finally, the relationship between excess nitrogen and the equivalent pressure was established by comparing appropriate reaction rate expressions with separate variation in pressure and diluent mass flow rate. This relationship forms a calibration for the simulation technique at blowout conditions. The calibration is valid for well-stirred and partially stirred reactors using propane as a fuel in a lean mixture with air. The specific calibration curve presented is for 300 K inlet temperature of reactants, although similar curves for other temperatures are readily constructed.

Lean blowout data in a research combustor, obtained using the simulation technique, were successfully correlated using the conventional loading parameter in which the pressure term was taken as the equivalent pressure obtained from the calibration. The range of equivalence ratios covered by the variation in loading parameter was from near the flammability limit to near the maximum heat release rate; the range of loading parameter obtained extended over three orders of magnitude. These were achieved at atmospheric pressure and with a limited range of mass flow rates. The lean stability of the research combustor was completely defined by these data.

The combustor was operated without difficulty with an excess nitrogen mass flow rate up to 60 percent of the air mass flow rate. At this condition the equivalent pressure corresponded to 0.1 atmosphere. However, it is not recommended that the maximum excess nitrogen flow rate be greater than about 45 percent of the air mass flow rate. This is because of inaccuracies in the calibration curve (lower right-hand corner of Fig. 10), at operating conditions with very high levels of excess nitrogen.

The effective order of the reaction when excess nitrogen is used becomes a function of the mass ratio of excess nitrogen to air, as well as equivalence ratio at blowout. There was no visual evidence that the excess nitrogen changed the flame characteristics in any way from that observed without excess nitrogen.

A final check on the accuracy of the calibration was made by comparing the correlated experimental results using the simulation against well-stirred reactor calculations based on propane/air systems for true subatmospheric pressures. This comparison was justified since it was known from the initial test results (Sturgess et al., 1992; Heneghan et al., 1990) that the research combustor had well-stirred reactor lean blowout performance when operated at atmospheric pressure without any excess nitrogen. The subsequent agreement in the stirred reactor calculation comparison confirmed the accuracy of the calibration.

The demonstration of the successful calibration for the nitrogen dilution simulation of low operating pressures opens an important possibility for small-scale facility combustion experiments. Their operating range may be extended for certain low pressure effects in an easy, inexpensive, and safe fashion while the convenience of operating at atmospheric pressure is maintained.

Experimental lean blowout data have been obtained in the research combustor over a wide range of equivalence ratios. These data will provide a good data base for lean blowout modeling. The fact that the research combustor has been definitively shown to behave as a well-stirred reactor when operated in the lifted flame condition is an important help in understanding, and hence modeling, the lean blowout process.

Conclusions

1 The nitrogen dilution technique for simulating the effects on chemical reaction rates of low pressures has been successfully calibrated. The calibration removes the simulation from being a qualitative technique, to being a quantitative one. It permits certain low-pressure combustion effects to be examined while maintaining the convenience of operating at atmospheric pressure.
2 The lean blowout data obtained in a research combustor via the low-pressure simulation technique were correlated by the familiar loading parameter when the effective pressure was used for actual pressure and the reaction order was made a function of the excess nitrogen.
3 Nitrogen dilution enables the loading parameter for the research combustor to be extended over three orders of magnitude, and the blowout equivalence ratio to be increased from near the flammability limit to close to the maximum heat release condition. This enables the lean stability of the combustor to be adequately defined.
4 The research combustor at blowout conditions behaves like a classical well-stirred reactor.

Acknowledgments

This work was supported by the U.S. Air Force Wright Research and Development Center, Aero Propulsion and Power Laboratory, under Contract No. F33615-87-C-2822 to Pratt & Whitney, East Hartford, CT (Contract Monitor: Lt. A. L. Lesmerises) and Contract No. F33615-87-C-2767 to University of Dayton, Dayton, OH (Contract Monitor: Dr. W. M. Roquemore).

References

Bragg, S. L., 1953, Aeronautical Research Council, Paper No. 16170, CF 272, U.K., Sept.

Clarke, A. E., Odgers, J., and Ryan, P., 1960, "Further Studies of Combustion Phenomena in a Spherical Reactor," *Proc. 8th Symposium (International) on Combustion,* The Combustion Institute, pp. 983–994.

Greenhough, V. W., and Lefebvre, A. H., 1956, "Some Applications of Combustion Theory to Gas Turbine Development," *Proc. Sixth Symposium (International) on Combustion,* Reinhold Publishing Co., pp. 858–869.

Heneghan, S. P., Vangsness, M. D., Ballal, D. R., Lesmerises, A. L., and Sturgess, G. J., 1990, "Acoustic Characteristics of a Research Step Combustor," AIAA Paper No. AIAA-90-1851.

Herbert, M. V., 1960, "A Theoretical Analysis of Reaction Rate Controlled Systems; Part II," *Proc. 8th Symposium (International) on Combustion,* The Combustion Institute, pp. 970–982.

Kretschmer, D., and Odgers, J., 1972, "Modeling of Gas Turbine Combustors—A Convenient Reaction Rate Equation," ASME JOURNAL OF ENGINEERING FOR POWER, pp. 173–180.

Lefebvre, A. H., and Halls, G. A., 1959, "Simulation of Low Combustion Pressures by Water Injection," *Proc. Seventh Symposium (International) on Combustion,* Butterworths Scientific Publications, London, pp. 654–658.

Lefebvre, A. H., 1961, "Some Simple Techniques for the Performance Evaluation of Gas Turbine Combustion Systems," *Experimental Methods in Combustion Research,* J. Surugue, ed., AGARD, Pergamon Press.

Lewis, B., and von Elbe, G., 1987, *Combustion, Flames and Explosion of Gases,* 3rd ed., Academic Press.

Longwell, J. P., Frost, E. E., and Weiss, M. A., 1953, "Flame Stability in Bluff-Body Recirculation Zones," *Ind. Eng. Chem.,* Vol. 45, No. 8, pp. 1629–1633.

Norster, E. R., 1968, "Subsonic Flow Flameholder Studies Using a Low Pressure Simulation Technique," *Proc. Intl. Prop. Symp. Combustion in Advanced Gas Turbine Systems,* I. E. Smith, ed., Cranfield Intl. Symp. Series, Vol. 10, Pergamon Press, pp. 79–94.

Pratt, D. T., and Wormeck, J. J., 1976, "CREK—A Computer Program for Calculation of Combustion Reaction Equilibrium and Kinetics in Laminar or Turbulent Flow," Rpt. No. WSU-ME-TEL-76, Dept. Mech. Engrg., Washington State University.

Roberts, R., Aceto, L. D., Kollrack, R., Teixeira, D. P., and Bonnell, J. M., 1972, "An Analytical Model for Nitric Oxide Formation in a Gas Turbine Combustor," *AIAA J.,* Vol. 10, No. 6, pp. 820–826.

Sturgess, G. J., Sloan, D. G., Lesmerises, A. L., Heneghan, S. P., and Ballal, D. R., 1992, "Design and Development of a Research Combustor for Lean Blow-Out Studies," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 114, pp. 13–19.

Swithenbank, J., 1974, "Flame Stabilization in High Velocity Flow," *Combustion Technology—Some Modern Developments,* H. B. Palmer and J. M. Beer, eds., Academic Press, pp. 91–125.

Si-Doek Oh
New Project Development Team,
Hyosung Industries Co. Ltd.,
Seoul, Korea

Hyo-Sun Pang
Korea Gas Corporation R&D Center,
Ansan, Kyunggi-Do, Korea

Si-Moon Kim

Ho-Young Kwak[1]

Mechanical Engineering Department,
Chung-Ang University,
Seoul, Korea

Exergy Analysis for a Gas Turbine Cogeneration System

A general exergy balance equation that is applicable to any component of thermal systems has been formulated in this study. One of distinct features of this formulation is that the exergy involved in the component of any thermal system can be decomposed into exergy flows, entropy production flows, and the appropriate exergy rate terms such as fuel and available work. The exergy analysis based on this equation permits one to predict the thermal efficiency of the system, the exergy destruction in each component as well as the mass flow rate, the composition, and the temperature of the exhaust gases. We have examined the performance of a 1000 kW gas turbine cogeneration system when it is operated at part and full-load conditions through this analysis. We have also tested the effect of the inlet air temperature and the relative humidity of the inlet air on the performance of the system. The predicted values of the performances for the system have been compared with the actual performance data provided by the gas turbine manufacturer. It has been found that the measured data of net power and the properties of exhaust gases are in good agreement with calculation ones, differing by less than 3 percent. The exergy balance equation may be utilized in the exergoeconomic analysis to estimate the production costs depending on various input costs in a gas turbine cogeneration system.

1 Introduction

Cogeneration may be defined as a thermal system that produces electrical and heat energy simultaneously from a single source of fuel (Baughn and Kerwin, 1987; Baughn and Bagheri, 1987). For industrial and domestic applications where both kinds of energy are demanded, this turns out to be a very effective energy-saving system (Lundberg, 1991). In this study, a general exergy balance equation that is applicable to any component of thermal systems has been formulated and the energy and exergy analysis based on this equation permit one to predict the thermodynamic performance of the system as well as to provide appropriate design data for the auxiliary components such as absorption refrigerators and cooling towers, which are needed for gas turbine cogeneration plants (Kim, 1995). This analysis also yields the thermal efficiency of the system, the exergy destruction in each component, the mass flow rate, the compositions and temperature of the exhaust gases at turbine outlet, and NO concentration with appropriate input data. In this analysis the mass, energy, and exergy balance in each component are strictly applied, which has never been done before. Previous exergy analysis (Dunbar et al., 1993; El-Masri, 1985) is a rather ideal and simplified approach in this sense.

We have examined the performance of the gas turbine cogeneration system as well as the exergy destruction in each component in the system when it is operated at part and full-load conditions through this analysis. We have also tested the effect of the inlet air temperature, the relative humidity of the inlet air, and water or steam injection on the performance of the system. The predicted values of the performances for a 1000 kW gas turbine system have been compared with the actual performance data provided by the gas turbine manufacturer.

One distinct feature of the exergy balance equation is that the exergy involved in the component of any thermal system can be decomposed into exergy flows, entropy production flows, and the appropriate exergy rate terms such as fuel and available work. The exergy balance equation may be utilized in the exergoeconomic analysis to estimate the production costs depending on various input costs in a gas turbine cogeneration system (Kim et al., 1995).

2 Governing Laws for Energy and Exergy

For steady flow processes, the law of thermodynamics, or the energy conservation law for an open system that interacts with two thermal reservoirs including atmosphere (Howell, 1987; Bejan, 1988) could be extended to the case where there is heat exchange between streams in the system (Kim, 1995). That is, including mass exchange:

$$(\dot{Q}_o + \dot{Q}_s + \dot{Q}_f) + \sum_{\text{inlet}} \dot{m}_i\left(h_i + \tfrac{1}{2}v_i^2 + gz_i\right) = \sum_{\text{outlet}} \dot{m}_i\left(h_i + \tfrac{1}{2}v_i^2 + gz_i\right) + \dot{W}_{c.v} + \sum h_{o,i}(\dot{m}_{i,\text{in}} - \dot{m}_{i,\text{out}}) \quad (1)$$

where $\dot{Q}_s$ is the heat exchange rate between two material streams in the system. The entropy generation rate of the system described by Eq. (1) is given by

$$\dot{S}_{\text{gen}} = -\left(\frac{\dot{Q}_o}{T_o} + \frac{\dot{Q}_s}{T_s}\right) - \frac{\dot{Q}_f}{T_f} - \sum_{\text{inlet}} \dot{m}_i s_i + \sum_{\text{outlet}} \dot{m}_i s_i + \sum s_{o,i}(\dot{m}_{i,\text{in}} - \dot{m}_{i,\text{out}}) \quad (2)$$

Eliminating the rate of heat transfer with atmosphere, $\dot{Q}_o$, from Eqs. (1) and (2), one may obtain the rate of available work of the system, $\dot{E}_w$, or the exergy balance law. It is given by

$$\dot{E}_w = \dot{W}_{c.v} = \dot{E}_{Qf} + \dot{E}_{Qs} + \sum_{\text{inlet}} \dot{E}_{x,i} - \sum_{\text{outlet}} \dot{E}_{x,i} - T_o(\dot{S}_{\text{gen}}) \quad (3)$$

The first term of the right-hand side in Eq. (3) is the rate of maximum available energy to perform work for the thermal reservoir at T_f. That is

[1] To whom correspondence should be addressed.

Contributed by the International Gas Turbine Institute and presented at the International Mechanical Engineering Congress and Exposition, San Francisco, California, November 12–17, 1995. Manuscript received by the International Gas Turbine Institute March 22, 1995; revision received March 6, 1996. Associate Technical Editor: L. A. Riekert.

$$\dot{E}_{Qf} = \left(1 - \frac{T_o}{T_f}\right)\dot{Q}_f \quad (4)$$

Similarly, the second term of the right-hand side in Eq. (3) is the rate of exergy, which is related heat exchange between material streams at T_s. It is given by

$$\dot{E}_{Qs} = \left(1 - \frac{T_o}{T_s}\right)\dot{Q}_s \quad (5)$$

The rate of flow exergy, $\dot{E}_x$, in Eq. (3) can be expressed in terms of the mass flow rate and the enthalpy and entropy at given state and at restricted dead state:

$$\dot{E}_x = \dot{m}e_x = \dot{m}[(h + \tfrac{1}{2}v^2 + gz - h_o^*) - T_o(s - s_o^*) + (\mu_o^* - \mu_o)] \quad (6)$$

where the notation ( )* indicates properties associated with the restricted dead state and $\mu_o^* - \mu_o$ is the chemical exergy term (Moran, 1982). The last term in Eq. (3) represents the rate of lost exergy (work), which is always positive.

In addition to these governing equations for the energy and exergy of the system, the following mass conservation equation should be employed for thermal analysis of open system:

$$\sum_{\text{inlet}} \dot{m}_i = \sum_{\text{outlet}} \dot{m}_i \quad (7)$$

In this analysis all noninteracting gases are assigned zero enthalpy and entropy at 273.15 K and the kinetic and potential energies in exergy flow are neglected. For entropy calculation, mixing of different gaseous components and pressure elevation effects is taken into account. Thus, the enthalpy and entropy of each nonreacting component can be calculated by using the appropriate polynomials (Gordon and McBride, 1971) fitted into thermophysical data in the *JANAF Tables* (1971). Careful attention has been paid to the quantitative balance for the governing laws at each plant component in this analysis.

3 Thermodynamic Analysis for the Components of the Gas Turbine System

3.1 System Description. A schematic diagram for the typical 1000 kW class gas turbine cogeneration system, which shows the main work and exergy flows and state points, is given in Fig. 1. The mass flow rate of air at T_1 = 15°C is about 5.0 kg/s. The mass flow rate of air is adjusted depending on the inlet air temperature so that the air temperature does not affect the isentropic efficiency of the compressor very much. The compression ratios at the first and second stage of the compressor for the gas turbine system concerned in this study are about 3.18 and 3.34, respectively. The air fuel ratio on mass basis at design condition is about 59.5. The air temperature at the inlet to the second stage of the compressor (T_3) is fixed to the predetermined value depending on the inlet air temperature (T_1) by controlling the mass flow rate ($\dot{m}_{15}$) and the temperature of water (T_{15}) of one of heat exchangers in the intercooler.

The temperature (T_5) and the pressure (P_f) of the gaseous fuel from the gas compressor are about 58°C and 1500 kPa, respectively. The pressure (P_w) of the injected water or saturated steam to the combustor is assumed to be 1500 kPa. The temperature of the injected water is assumed to be 55°C. The mass flow rate of the injected water or steam is almost same amount as that of fuel. The pressure of the exhaust gases at turbine outlet is about 102 kPa. The isentropic efficiencies of these components are estimated by using a set of experimental

Nomenclature

a, b, c, d = mole fraction of fuel species
$\bar{a}_{ch}$ = availability of exergy fuel
C_p = constant pressure specific heat
$\bar{E}_{che}$ = exergy of fuel
E_{Qf} = available energy to perform work at T_f with chemical exergy of fuel
$\dot{E}_x^m$ = rate of thermomechanical exergy
$\dot{E}_x$ = rate of total exergy
e_x = exergy per unit mass flow rate
$\bar{e}_x$ = average value of exergy of mixture
$\dot{E}_w$ = rate of available work
$\bar{h}$ = average value of enthalpy of mixture
$\bar{h}_f^o$ = enthalpy of formation
h_{fg} = latent heat
h_i = enthalpy of i-species gas in mixture
h_o = pure substance enthalpy at standard state
$\dot{m}$ = mass flow rate
n = number of moles of gas
$n_{\alpha t}$ = total number of C atoms in fuel
$n_{\beta t}$ = total number of H atoms in fuel
P_o = ambient pressure
P = pressure
Q_f = lower heating value of fuel
$\dot{Q}_o$ = heat transfer rate to the environment
$\dot{Q}_s$ = heat exchange rate between two material streams
R = gas constant
$\bar{s}$ = average value of entropy of mixture
s_i = entropy of i-species gas in mixture
s_o = pure substance entropy at standard state
$\dot{S}_{\text{gen}}$ = entropy generation rate
T = temperature
T_f = temperature of high-temperature thermal reservoir; adiabatic flame temperature
T_o = ambient temperature
T_s = temperature in which heat exchange between streams occurs
v = velocity
$\dot{V}$ = flow rate
$\dot{W}$ = work transfer rate
z = elevation
α = number of carbon atoms in fuel
β = number of hydrogen atoms in fuel
η = efficiency
η_e = exergy efficiency of the gas turbine system
η_{th} = thermal efficiency of gas turbine system
$\mu = h - T_o s$-chemical potential
ω = humidity ratio

Subscripts

a = air
c = compressor
cb = combustor
ch = chemical exergy
co = correct oxygen
$c.v$ = control volume
e = exergetic
da = dry air
eo = excess oxygen
f = fuel
j, k, l, m = index of fuel species
p = products (exhaust gases)
rev = reversible
s = isentropic
t = turbine
ta = total air
to = total oxygen
w = water or steam
1–8 = index of stations in gas turbine system
2′, 4′, 8′ = index of stations achieved by isentropic process

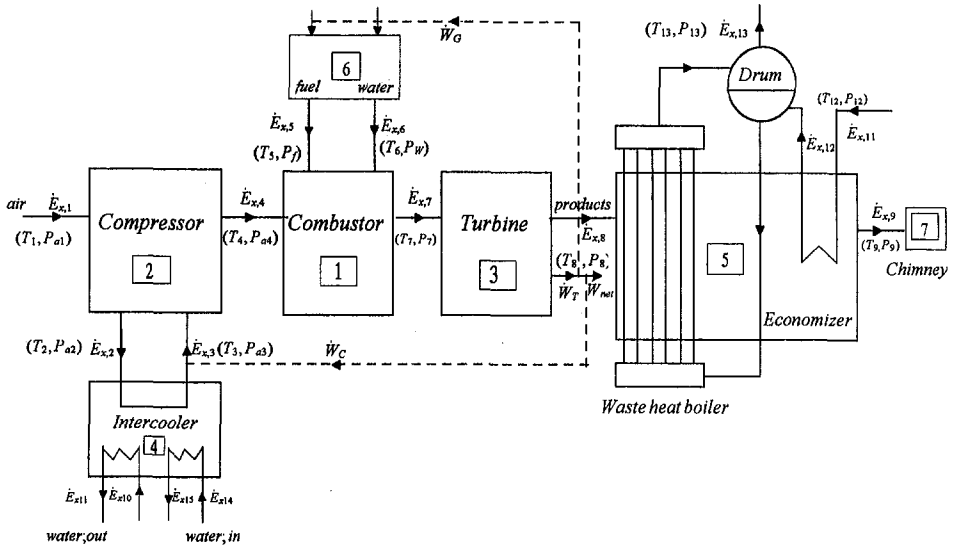


Fig. 1 Work and exergy flow in gas turbine cogeneration system

data of inlet and outlet temperatures at full-load condition. The values obtained are employed in the other cases. The isentropic efficiencies for the compressor and turbine turn out to be 0.814 and 0.850, respectively. However, for part-load condition, where the temperature of the product is sufficiently low, the turbine isentropic efficiency 0.860 has been used in this analysis. The part-load condition is accomplished by reducing the fuel flow rate. Turbine blade cooling is achieved by compressed air drawn from the second stage of the compressor outlet in this system.

3.2 Compressor. A simple version of the gas turbine cycle neglecting the pressure losses in the intercooler and combustor is shown in Fig. 2. The work needed in the first stage of the compressor can be expressed in terms of the isentropic efficiency of the compressor and the pressure ratio in that stage. Or, on the basis of ideal gas behavior,

$$-\frac{\dot{W}_{c,1}}{\dot{m}_a} = (\bar{h}_2 - \bar{h}_1) = \frac{C_{p,1}T_{2'}}{\eta_{cs}}\left[1 - \left(\frac{P_{a,2}}{P_{a,1}}\right)^{-(R_1/C_{p,1})}\right] \quad (8)$$

where $\bar{h}$ is the average enthalpy per unit mass of the mixture of air, which consists of N_2, O_2, Ar, and water vapor. The temperature dependence on the specific heat is taken into account in this calculation:

$$\bar{h} = \frac{\Sigma\, \dot{m}_i h_i}{\Sigma\, \dot{m}_i}. \quad (9)$$

The other average properties such as entropy, $\bar{s}$, and exergy, $\bar{e}_x$, per unit mass flow rate are defined as in Eq. (9). Neglecting the heat transfer to the environment, the air enthalpy change experienced in the first stage of the compressor may be written as

$$\bar{h}_2 - \bar{h}_1 = \left(\int_{2'}^{2} T d\bar{s}\right)_{P=P_{a2}} + (\bar{h}_{2'} - \bar{h}_1) \quad (10)$$

Equation (10) provides the temperature of the air at the outlet of the first stage of the compressor (T_2) on the basis of ideal gas behavior,

$$T_2 = T_{2'}\left\{1 + \left(\frac{1}{\eta_{c,s}} - 1\right)\left[1 - \left(\frac{P_{a,2}}{P_{a,1}}\right)^{-(R_1/C_{p,1})}\right]\right\} \quad (11)$$

Temperature dependence on the heat capacity of gas was taken into account to evaluate T_2. Similarly, one may obtain the work needed in the second stage of the compressor and the air temperature at the compressor outlet. These are on the basis of ideal gas behavior,

$$-\frac{\dot{W}_{c,2}}{\dot{m}_a} = \bar{h}_4 - \bar{h}_3 = \frac{C_{p,3}T_{4'}}{\eta_{c,s}}\left[1 - \left(\frac{P_{a,4}}{P_{a,3}}\right)^{-(R_1/C_{p,3})}\right] \quad (12)$$

and

$$T_4 = T_{4'}\left\{1 + \left(\frac{1}{\eta_{c,s}} - 1\right)\left[1 - \left(\frac{P_{a,4}}{P_{a,3}}\right)^{-(R_1/C_{p,3})}\right]\right\} \quad (13)$$

The constant-pressure specific heats, $C_{p,1}$ and $C_{p,3}$, which appear

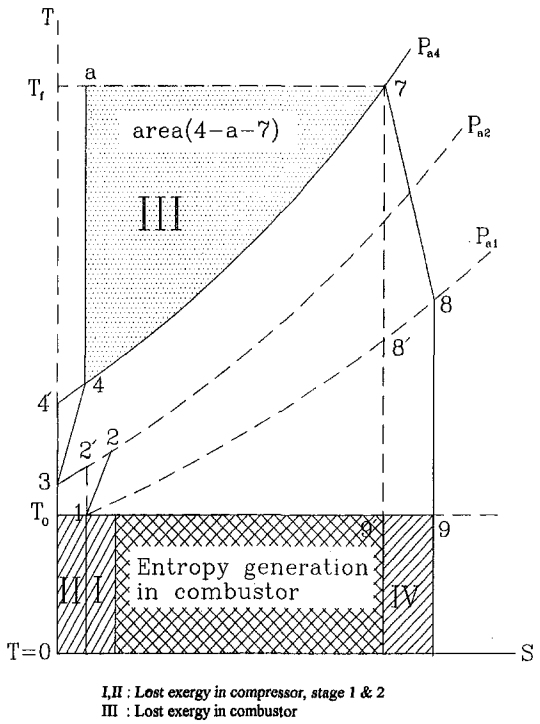


Fig. 2 The process and internal irreversibilities of a gas turbine system in *T–S* diagram

in Eqs. (8), (11), (12), and (13), are the average taken between the values at states 1 and 2, and the values at states 3 and 4 respectively.

The exergy balance for the compressor can be obtained from Eqs. (3), (8), and (12). That is

$$\frac{(\dot{E}_w)_c}{\dot{m}_a} = \frac{1}{\dot{m}_a}(\dot{W}_{c,1} + \dot{W}_{c,2}) = (\bar{e}_{x,1} - \bar{e}_{x,2}) + (\bar{e}_{x,3} - \bar{e}_{x,4}) - \frac{T_o(\dot{S}_{\text{gen}})_c}{\dot{m}_a} \quad (14)$$

The lost work in the compressor, given in Eq. (14), can be obtained from either Eq. (2) or Eq. (15):

$$\frac{T_o(\dot{S}_{\text{gen}})_c}{\dot{m}_a} = T_o[(\bar{s}_4 - \bar{s}_3) + (\bar{s}_2 - \bar{s}_1)] \quad (15)$$

where $\bar{s}$ denotes $\bar{s} - \bar{s}_o^*$. This lost exergy per unit mass flow rate is represented by the shaded area in the lower-left corner of Fig. 2. The isentropic ($\eta_{c,s}$) and the exergy ($\eta_{c,e}$) efficiency of the compressor (Bejan, 1988), which are evaluation tools for the component, are given by

$$\eta_{c,s} = \frac{(\bar{h}_{2'} - \bar{h}_1) + (\bar{h}_{4'} - \bar{h}_3)}{(\bar{h}_2 - \bar{h}_1) + (\bar{h}_4 - \bar{h}_3)} \quad (16)$$

and

$$\eta_{c,e} = 1 + \frac{T_0(\dot{S}_{\text{gen}})_c}{(\dot{E}_w)_{c,\text{rev}}} \quad (17)$$

3.3 Combustor. On a one kmole of fuel basis, the overall combustion reaction for typical gas turbine operation may be written as on the basis of complete combustion,

$$(aC_{\alpha j}H_{\beta j} + bC_{\alpha k}H_{\beta k} + cC_{\alpha l}H_{\beta l} + dC_{\alpha m}H_{\beta m}) + n_{to}O_2 + 3.728\,n_{to}N_2 + 0.044n_{to}\text{Ar} + (n_{wv} + n_{wi})H_2O \rightarrow n_{\alpha t}CO_2 + (0.5n_{\beta t} + n_{wv} + n_{wi})H_2O + n_{eo}O_2 + 3.728n_{to}N_2 + 0.044n_{to}\text{Ar} \quad (18)$$

The combustion process is assumed to be complete because the flame temperature in the combustor is around 1000°C. The gaseous fuel is known to be composed of various hydrocarbons such as CH_4, C_2H_6, C_3H_8, and C_4H_{10}. The mole numbers in the reaction given in Eq. (18) are given as follows:

$$n_{\alpha t} = a\alpha_j + b\alpha_k + c\alpha_l + d\alpha_m \quad (19\text{-}1)$$

$$n_{\beta t} = a\beta_j + b\beta_k + c\beta_l + d\beta_m \quad (19\text{-}2)$$

$$n_{co} = n_{\alpha t} + 0.25n_{\beta t} \quad (19\text{-}3)$$

$$n_{to} = xn_{co} \quad (19\text{-}4)$$

$$n_{eo} = (x - 1)n_{co} \quad (19\text{-}5)$$

$$n_{wv} = n_{ta}\frac{1.608\omega}{1 + 1.608\omega} \quad (19\text{-}6)$$

$$n_{da} = 4.772xn_{co} \quad (19\text{-}7)$$

$$n_{ta} = n_{da} + d_{wv} \quad (19\text{-}8)$$

where

$$a + b + c + d = 1.0 \quad (19\text{-}9)$$

The lower heating value of the fuel can be obtained from the chemical reaction given in Eq. (18) by assuming that the reaction occurs at standard state, or

$$Q_f = a(\bar{h}_f^o)_j + b(\bar{h}_f^o)_k + c(\bar{h}_f^o)_l + d(\bar{h}_f^o)_m - n_{\alpha t}(\bar{h}_f^o)_{CO_2} - 0.5n_{\beta t}(\bar{h}_f^o)_{H_2O(g)} \quad (20)$$

Also the temperature in the combustor, T_f (or T_7) can be calculated from the following energy conservation equation by assuming that the heat transfer to the environment is negligible:

$$Q_f + [a(\Delta\bar{h})_j + b(\Delta\bar{h})_k + c(\Delta\bar{h})_l + d(\Delta\bar{h})_m]_{T_5} - n_{wi}(\bar{h}_{fg})_{T_6} + n_{wi}(\Delta\bar{h})_{T_6} + [n_{to}(\Delta\bar{h})_{O_2} + 3.728n_{to}(\Delta\bar{h})_{N_2} + 0.044n_{to}(\Delta\bar{h})_{\text{Ar}} + n_{wv}(\Delta\bar{h})_{H_2O}]_{T_4} = [n_{\alpha t}(\Delta\bar{h})_{CO_2} + (0.5n_{\beta t} + n_{wv} + n_{wi})(\Delta\bar{h})_{H_2O} + n_{eo}(\Delta\bar{h})_{O_2} + 3.728n_{to}(\Delta\bar{h})_{N_2} + 0.044n_{to}(\Delta\bar{h})_{\text{Ar}}]_{T_7} \quad (21)$$

The bisection method (Chapra and Canale, 1988) was employed to calculate the adiabatic flame temperature, T_7. Once the flame temperature is known, one can calculate the equilibrium concentration of NO from the following chemical reaction with the help of the corresponding equilibrium constant and the equilibrium composition of O_2 and N_2, which can be obtained from Eq. (18) approximately:

$$\tfrac{1}{2}O_2 + \tfrac{1}{2}N_2 \rightleftarrows NO \quad (22)$$

With consideration of the resident time of products in a combustor, one can calculate the transient NO concentration by using the Zeldovich mechanism (Heywood 1988), which turns out to be much less than that of the equilibrium concentration of NO in the system considered.

The exergy balance for the combustor may be written as

$$\dot{E}_{Qf} + \dot{E}_{x,4} + \dot{E}_{x,5} + \dot{E}_{x,6} - \dot{E}_{x,7} - T_o(\dot{S}'_{\text{gen}})_{cb} = 0 \quad (23)$$

The lost exergy in the combustor can be obtained from Eqs. (2) and (23). That is,

$$\frac{T_o(\dot{S}'_{\text{gen}})_{cb}}{\dot{m}_p} = \frac{\dot{m}_6}{\dot{m}_p}(h_{fg})_{T_6} - \frac{\dot{Q}_f T_o}{\dot{m}_p T_f} + T_o\left(\bar{s}_7 - \frac{\dot{m}_4}{\dot{m}_p}\bar{s}_4 - \frac{\dot{m}_5}{\dot{m}_p}\bar{s}_5 - \frac{\dot{m}_6}{\dot{m}_p}\bar{s}_6\right) \quad (24)$$

With the help of the energy conservation for combustor, the equation above can be written as approximately, since $\dot{m}_p \cong \dot{m}_a$,

$$\frac{T_o(\dot{S}'_{\text{gen}})_{cb}}{\dot{m}_p} = \frac{T_o}{T_f}\left[T_f(\bar{s}_7 - \bar{s}_4) - \left(\int_4^7 Tds\right)_{p=P_{a,4}}\right] \quad (24')$$

The lost exergy in the combustor per unit mass flow rate is equal to the area (4-a-7) with a reduction factor of T_o/T_f, as shown in Fig. 2, which shows that the lost exergy occurring in the combustor may be considered as external (Bejan, 1988). However, the exergy balance equation for the combustor can be rewritten with the net entropy flow rate for the component. It is provided that exergy stream includes thermomechanical exergy only.

$$\dot{E}_{che} + \dot{E}^m_{x,4} + \dot{E}^m_{x,5} + \dot{E}^m_{x,6} - \dot{E}^m_{x,7} - T_o(\dot{S}_{\text{gen}})_{cb} = 0 \quad (23')$$

where

$$\frac{T_o(\dot{S}_{\text{gen}})_{cb}}{\dot{m}_P} = \frac{\dot{m}_6}{\dot{m}_p}(h_{fg})_{T_6} + T_o\left(\bar{s}_7 - \frac{\dot{m}_4}{\dot{m}_p}\bar{s}_4 - \frac{\dot{m}_5}{\dot{m}_p}\bar{s}_5 - \frac{\dot{m}_6}{\dot{m}_p}\bar{s}_6\right) \quad (25)$$

and

$$E_{\text{che}} = a(\bar{a}_{ch})_j + b(\bar{a}_{ch})_k + c(\bar{a}_{ch})_l + d(\bar{a}_{ch})_m \quad (26)$$

With this net entropy flow rate term, the exergy balance equation for any component of thermal system can be written generally as

$$\dot{E}_{\text{che}} + \dot{E}_{Qs} + \Big(\sum_{\text{inlet}} \dot{E}^m_{x,i} - \sum_{\text{outlet}} \dot{E}^m_{x,i}\Big) + T_o\Big(\sum_{\text{inlet}} \dot{S}_i - \sum_{\text{outlet}} \dot{S}_i\Big) = \dot{E}_w \quad (27)$$

The fourth term in Eq. (27) represents the negentropy flow, which is essentially the entropy production rate due to internal irreversibilities. Equation (27) can be regarded as a general exergy balance equation in the sense that the exergy of material stream can be decomposed into exergy flows and entropy production flows for any component in thermal system.

As can be seen in Eqs. (24) and (25), quite a different interpretation of the entropy generation in combustor is possible by choosing the exergy input due to chemical reaction as $\dot{E}_{che}$ rather than $\dot{E}_{che}\,(1 - T_o/T_f)$. The entropy generation given in Eq. (25) could be regarded as the internal one, which can be seen in Fig. 2. Equation (23′) is more appropriate for the exergy balance equation of the combustor, since the heat transfer to the environment is negligible, so that all chemical exergy is spent to raise the thermal and mechanical exergy of the material stream in combustor. It should be noted that the water injection to the combustor ($\dot{m}_6$) induces the lost work in the system due to heat loss by evaporation of water, which is $-\dot{m}_6(h_{fg})_{T_6}$, as can be seen in Eq. (25).

The exergy efficiency of the combustor may be defined as

$$\eta_{cb,e} = 1 - \frac{T_o \dot{S}_{\text{gen},cb}}{\dot{E}_{\text{che}}} \quad (28)$$

3.4 Turbine. Using the same procedure as done for the analysis of the compressor, one could obtain the work done in the turbine and the temperature at the turbine outlet (T_8). These are on the basis of ideal gas behavior,

$$\frac{\dot{W}_t}{\dot{m}_p} = \bar{h}_7 - \bar{h}_8 = \eta_{t,s} C_{p,7} T_{8'}\left[\left(\frac{P_{a,4}}{P_{a,1}}\right)^{R_p/C_{p7}} - 1\right] \quad (29)$$

and

$$T_8 = T_7\left\{1 - \eta_t\left[1 - \left(\frac{P_{a,4}}{P_{a,1}}\right)^{R_p/C_{p7}}\right]\right\} \quad (30)$$

The exergy balance for the gas turbine is given by

$$\frac{\dot{W}_t}{\dot{m}_p} = \bar{e}_{x,7} - \bar{e}_{x,8} - \frac{T_o \dot{S}_{\text{gen},t}}{\dot{m}_p} \quad (31)$$

which provides the lost exergy per unit mass flow rate in the turbine such as

$$\frac{T_o(\dot{S}_{\text{gen}})_t}{\dot{m}_p} = T_o(\bar{s}_8 - \bar{s}_7) \quad (32)$$

This lost exergy is represented by the shaded area in the lower-right corner of Fig. 2. The isentropic efficiency ($\eta_{t,s}$) and the exergy efficiency ($\eta_{t,e}$) for the turbine may be defined as

$$\eta_{t,s} = \frac{\bar{h}_8 - \bar{h}_7}{\bar{h}_{8'} - \bar{h}_7} \quad (33)$$

and

$$\eta_{t,e} = 1 - \frac{T_o(\dot{S}_{\text{gen}})_t}{(\dot{E}_W)_{t,\text{rev}}} \quad (34)$$

Table 1 Specification of waste heat boiler and intercooler

Waster Heat Boiler			
		Evaporator	Economizer
Type of tube		Finned	Finned
Type of tube arrangement		staggered	staggered
number of tube rows in flow		24	12
number of tubes normal to the		12	22
Transverse tube pitch [m]		0.106	0.065
Parallel tube pitch [m]		0.09	0.055
Width of heat exchanger [m]		1.325	1.4625
Height of heat exchanger [m]		1.8	1.08
Length of heat exchanger [m]		2.25	0.65
Total inside heat area [m^2]		73.29	17.56
Total outside heat area [m^2]		82.73	22.75
Fouling factor [m^2 K/kW]	gas side	0.14394	0.12992
	water side	0.14394	0.12992
Overall heat transfer		0.262282	0.284782
Intercooler			
		Hot side	Cold side
number of tubes		120	120
Length of heat exchanger [m]		1.1	1.1
Total inside heat area [m^2]		107.8	107.8
Fouling factor [m^{2}K/kW]		0.17697	0.17697
Overall heat transfer coefficient [kW/m^{2}K]		0.0708	0.0741

3.5 Intercooler and Waste Heat Boiler. Detailed analysis for the intercooler and waste heat boiler will not be discussed here. Essentially the intercooler employed in this system is made of two identical multipass and counterflow heat exchangers. The waste heat boiler, on the other hand, consists of evaporator and economizer section as shown in Fig. 1. Specification of the waste heat boiler and intercooler utilized in this cogeneration system is shown in Table 1. The equation suggested by IFC (International Formulation Committee) have been employed to calculate the water and steam properties in the intercooler and waste heat boiler (*JSME Steam Tables,* 1968).

Heat exchange between air and flue gas and water takes place in the intercooler and economizer. In this case, the outlet temperature of air or flue gas may be obtained by using the following equation (Stoecker, 1989) provided that the overall heat transfer U, the heat transfer area A for the heat exchangers, and the properties at inlet streams are known:

$$T_{g,o} = \frac{(1 - R)T_{g,i} + \{1 - \exp[\text{NTU}(R - 1)]\}RT_{w,i}}{1 - R\exp[\text{NTU}(R - 1)]} \quad (35)$$

where NTU is the number of heat transfer units and R is the ratio of the fluid capacity rates for cold and hot streams, which are defined as

$$\text{NTU} = \frac{UA}{C_c} \quad (36)$$

and

$$R = \frac{C_c}{C_h} = \frac{\dot{m}_w C_{p,w}}{\dot{m}_g C_{p,g}} \quad (37)$$

The outlet temperature of water in the exchanger can also be obtained by

$$T_{w,o} = T_{w,i} + \frac{T_{g,o} - T_{g,i}}{R} \tag{38}$$

Equation (38) has been employed to calculate the outlet temperature of air at the first and second stage of the intercooler as well as the flue gas temperature at the exit of the waste heat boiler. The overall heat transfer coefficient and the other informations on the intercooler and waste heat boiler are given in Table 1 (Kwon et al., 1994).

The amount of heat recovered from the waste heat boiler as steam may be obtained from

$$\dot{Q}_{BQ} = \dot{m}_{11}(h_{13} - h_{11}) \tag{39}$$

The corresponding exergy recovered from the boiler is given by

$$\dot{E}_{BQ} = \dot{m}_{11}(e_{x,13} - e_{x,11}) \tag{40}$$

The exergy balance equation for the intercooler can be obtained with the help of Eq. (25):

$$\dot{E}_{Qs} - T_0(\dot{S}_3 - \dot{S}_2 + \dot{S}_{11} - \dot{S}_{10} + \dot{S}_{15} - \dot{S}_{14}) = \dot{E}_{x,3} - \dot{E}_{x,2} \tag{41}$$

where $\dot{E}_{Qs} = \dot{E}_{x,10} - \dot{E}_{x,11} + \dot{E}_{x,14} - \dot{E}_{x,15}$ is the exergy gain by water streams due to the heat exchange between gas and water in the intercooler, which is given in Eq. (5).

The exergy balance for the waste heat boiler is given by

$$\dot{E}_{x,8} - \dot{E}_{x,9} - T_0(\dot{S}_9 - \dot{S}_8 + \dot{S}_{13} - \dot{S}_{11}) = \dot{E}_{BQ} \tag{42}$$

where $\dot{E}_{BQ}$ is the rate of exergy recovered as steam in the waste heat boiler.

3.5 Gas Compressor and Water Injection Equipment. Gas compressor and water injection equipment are auxiliary components to compress gas and water. The exergy balance for these components altogether is given by

$$|\dot{W}_G| - T_0(\dot{S}_5 + \dot{S}_6) = \dot{E}_{x,5} + \dot{E}_{x,6} \tag{43}$$

where $|\dot{W}_G|$ is the work needed to compress the gas and water streams. Using the results for each component of the cogeneration system, one can obtain the overall thermal (η_{th}) and exergy (η_e) efficiency of the system as

$$\eta_{th} = \frac{\dot{W}_t + \dot{W}_c + \dot{W}_G + \dot{Q}_{BQ}}{\dot{Q}_f} \tag{44}$$

and

$$\eta_e = \frac{\dot{W}_t + \dot{W}_c + \dot{W}_G + \dot{E}_{BQ}}{\dot{E}_{che}} \tag{45}$$

4 Computer Program

A computer program for the energy and exergy analysis has been written to simulate the thermodynamic performance of the gas turbine cogeneration system. The computational procedure is outlined in the flow chart of the program shown in Fig. 3. To perform its simulation, the program uses the following input data sets:

- (a) Standard pressure (P_o) and temperature (T_o).
- (b) Fuel compositions.
- (c) Isentropic efficiency of the compressor ($\eta_{c,s}$) and turbine ($\eta_{t,s}$).
- (d) Mass flow rate ($\dot{m}_a$) and the relative humidity of the inlet air to the first stage of compressor.

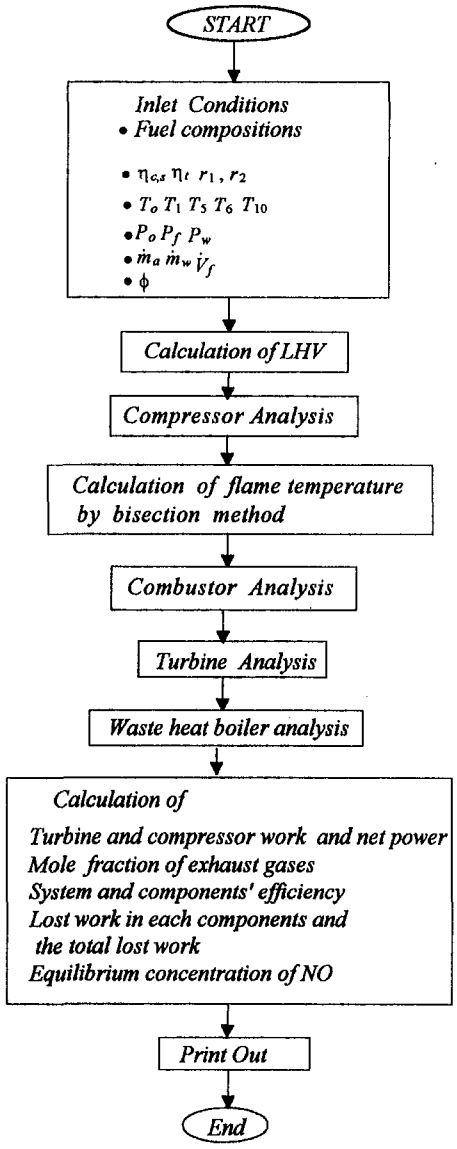


Fig. 3 Flow chart of the performance prediction program for the gas turbine cogeneration system

- (e) The compression ratio of the first and second stage of compressor.
- (f) The volume flow rate ($\dot{V}_f$), temperature (T_5) and pressure (P_f) of the fuel to the combustor.
- (g) The mass flow rate ($\dot{m}_w$), temperature (T_6), and pressure (P_w) of the injection water or steam to the combustor.
- (h) The overall heat transfer coefficient (U) and the heat transfer area (A) of the intercooler and waste heat boiler.
- (i) The mass flow rate and the temperature of water to intercooler.

Using these input data, and the thermodynamic analysis for the components of the gas system as discussed in the previous section, one can get the properties of exhaust gases such as temperature (°C), entropy (kW/K), enthalpy (kW), and exergy (kW) at each station, as shown in Fig. 1. The efficiencies of the components and of the system and the lost exergy in each component are then calculated by using these properties values.

Table 2 Calculated system and components' efficiency and lost work in each component (case 2); *data provided from gas turbine manufacturer

Suction air pressure*(kPa)	103.32
Suction air temp.(°C)	15
Fuel*(mole fraction in %)	CH_4:89.35, C_2H_6:8.5 C_3H_8:1.57, C_4H_{10}:0.53
Fuel flow rate*(Nm^3/h)	378
Air mass flow rate*(kg/s)	5.01
Water injection rate*(kg/h)	0
Temperature(°C)/exergy(kW) at LPC outlet (T_2)	153.5/602.7
Temperature(°C)/exergy(kW) at HPC outlet (T_4)	207.7/1202.4
Temperature(°C)/exergy(kW) at turbine inlet (T_7)	928/3951.7
Temperature(°C)/exergy(kW) at turbine outlet (T_8)	484.4/1063.4
Temperature(°C) of flue gas at stack (T_9)	158.2
Thermal /exergy efficiency(%) of the system	68.40/41.2
Lost work(kW)/exergy efficiency(%) of compressor	207.5/84.44
Lost work(kW)/exergy efficiency(%) of combustor	657/66.83
Lost work(kW)/exergy efficiency(%) of turbine	180.4/93.8

The computer output includes:

(*a*) Temperature, enthalpy, entropy, and exergy of the gases at each station.
(*b*) Mole fraction of the exhaust gases.
(*c*) Compressor and turbine work and net power.
(*d*) The lost work in each component and the total lost work.
(*e*) Equilibrium concentration of NO and the mass flow rate of the exhaust gases.

5 Calculation Results and Discussions

The computer program for the exergy analysis has been used primarily to test the effect of the inlet air temperature (T_1) on the gas turbine performance. The operation cases tested are the full-load case without water and steam injection, and with water or steam injection, in which the input data and the corresponding experimental results for the gas turbine performance are available. We have also examined the performance of the system when it is operated at part-load conditions and the performance depending on humidity of air. The calculation results have been compared with the experimental results or correction curves provided by gas turbine manufacturer. Calculation results for the system and components' efficiencies and the lost work in each component for the case without water or steam injection are shown in Table 2, which shows that the thermal efficiency of the system is about 68.4 percent. As expected, the exergy efficiency of the system, which is about 41.2 percent, is much less than that obtained from the first law standpoint. This large deviation is mainly due to the difference in the evaluation method of the heat recovered from waste heat boiler.

In Table 3, the measured values of the net power at the shaft end of gas turbine and the temperature, mass flow rate and mole fraction of exhaust gases at turbine outlet along with calculation results are shown for full load case without water or steam injection. The calculation and experimental results obtained with different inlet air temperatures of 0°C, 15°C, and 30°C are given in this table and correspond to cases 1, 2, and 3. As mentioned before, the mass flow rate of air and correspondingly the fuel flow rate are adjusted when the inlet air temperature increases so that the air temperature does not affect the compressor efficiency. As can be seen in this table, the agreement between experimental results and predicted ones is exceptionally good for the net power and the temperature and mole fraction of exhaust gases, differing by less than 3 percent. Certainly the isentropic efficiency of either compressor or turbine varies as the mass flow rate, if pressure and temperature of working gases are changed (Phillips, 1986). One would be able to predict the net power and the exhaust gas temperature at turbine outlet more accurately if the efficiencies as function of the appropriate parameters are provided.

In Figs. 4 and 5, the change of the thermal efficiency without waste heat boiler and the net power at the shaft end of gas turbine depending on the inlet air temperature are shown graphically. Calculation results show that the net power decreases more rapidly than the thermal efficiency as the inlet air temperature increases, which shows the same behavior as one obtained by experimental results. Note that the power output is the value measured at the generator for the steam injection case.

In Tables 4 and 5, the measured data along with predicted data for the gas turbine performance depending on inlet air temperature are shown for the full-load case with steam and water injection, respectively. The agreement between experi-

Table 3 Measured and predicted gas turbine performance and exhaust gas data for the cases of different inlet air temperatures without water injection (ISO 1155 kW basis, ISO condition; no duct losses but including gear box loss at 1 atm and 15°C); *measured at the shaft end of gas turbine; **measured NO_x includes NO as well as NO_2

Suc. air press, P_{a1}	kPa	Case 1	103.32	Case 2	103.32	Case 3	103.32
Suc. air temp, T_1	°C		0		15		30
Suc. air humidity, ϕ	%RH		60		60		60
HPC inlet air temp, T_3	°C		36.5		45.8		52.5
Fuel (mole fraction in %)	-	CH_4 : 89.3 , C_2H_6 : 5.4 , C_3H_8 : 3.5 , C_4H_{10} : 1.8					
Fuel flow rate,	Nm^3/h		395		378		351
Air mass flow rate,	kg/sec		5.37		5.01		4.63
Water injection rate,	kg/h		0		0		0
		Cal.	Exp.	Cal.	Exp.	Cal.	Exp.
Power*	kW	1200	245	1117	1155	997	1017
Exhaust gas temp, T_8	°C	467	457	484.4	481	489	495
Exhaust gas mass flow rate,	kg/sec	5.46	5.46	5.1	5.09	4.71	4.71
Constituent,exh. gas							
CO_2	Vol.%	3.00	3.1	3.07	3.2	3.06	3.2
H_2O	Vol.%	6.00	6	6.76	6.7	8.22	8.2
N_2	Vol.%	75.6	75.6	75.1	75.1	73.95	73.9
O_2	Vol.%	14.47	14.4	14.19	14.1	13.09	13.8
Ar	Vol.%	0.89	0.9	0.89	0.9	0.87	0.9
NOx emission**	ppmv	127.3	123	142.8	109	137.6	84

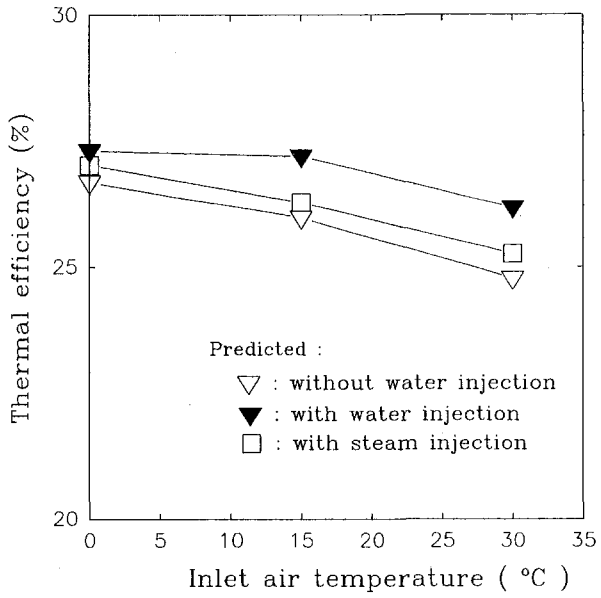


Fig. 4 System thermal efficiency without waste heat boiler for the cases without water injection, with water injection, and with steam injection depending on inlet air temperatures (ISO 1155 kW basis)

mental results and predicted ones for these cases is also good. The tendency of the thermal efficiency and the net power depending on the inlet air temperature are the same as for the cases without water or steam injection.

In Fig. 6, the fuel flow rate and the temperature and the mass flow rate of exhaust gases at various part-load conditions are shown. The load characteristic curves calculated behave as those by experimental results provided by gas turbine manufacturer.

The system's thermal efficiency and the net power, depending on the humidity of inlet air, are shown in Fig. 7. Both the thermal efficiency and the net power decrease slightly in the same manner as the relative humidity increases. On the other hand, the temperature of exhaust gases experiences 3°C changes in full variation of the relative humidity. As can be seen in this figure, the calculated values are close to the experimental results.

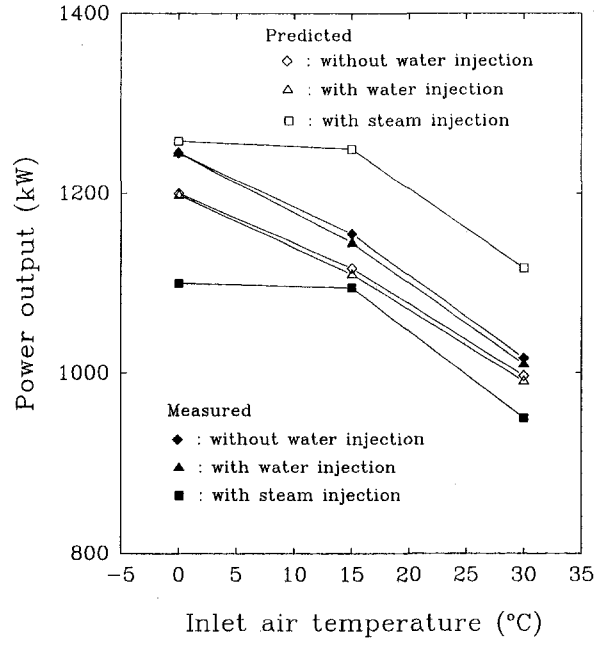


Fig. 5 Predicted and measured system power output depending on inlet air temperature for the cases without water injection, with water injection, and with steam injection (ISO 1155 kW basis); for the steam injection case, power output is measured at generator

Table 4 Measured and predicted gas turbine performance and exhaust gas data for the cases of different inlet air temperatures with steam injection (ISO 1090 kW basis); *measured at the shaft end of gas turbine; **measured NO_x includes NO as well as NO_2

		Case 4		Case5		Case6	
Suc. air press.	kPa	103.32		103.32		103.32	
Suc. air temp.	°C	0		15		30	
Suc. air humidity	%RH	60		60		60	
HPC inlet air temp	°C	38		48		58	
Fuel (mole fraction in %)	-	CH_4 : 89.3 , C_2H_6 : 5.4 , C_3H_8 : 3.5 , C_4H_{10} : 1.8					
Fuel flow rate	Nm^3/h	390		371		345	
Air mass flow rate	kg/sec	5.34		4.98		4.6	
Steam injection rate	kg/h	328		312		290	
		Cal.	Exp.	Cal.	Exp.	Cal.	Exp.
Power*	kW	1198	1245	1110	1,145	991	1010
Exhaust gas temp	°C	451	441	466	462	471.4	476
Exhaust gas mass flow rate	kg/sec	5.51	5.52	5.14	5.15	4.75	4.76
Constituent,exh. gas							
CO_2	Vol.%	2.9	3	2.94	3.1	2.95	3.1
H_2O	Vol.%	8.38	8.4	9.1	9.1	10.56	10.5
N_2	Vol.%	73.71	73.6	73.16	73.1	72.04	72
O_2	Vol.%	14.15	14.1	13.91	13.8	13.6	13.5
Ar	Vol.%	0.87	0.9	0.86	0.9	0.85	0.9
NOx emission**	ppmv	97.2	51	106.7	46	104.6	36

The exergy study reveals that the majority of system irreversibility occurs within the combustor, wherein 14.8 percent of the exergy of the fuel is destroyed, as shown in Fig. 8. The exergy destroyed in the remainder of the system such as compressor and turbine is about 8.8 percent of the fuel exergy. About 4.5 percent of the fuel exergy is destroyed by the processes of energy conversion and heat transfer, which were provided by gas turbine manufacturer. It should be noted that the exergy escaping through the chimney is about 25 percent of the fuel exergy. Consequently, a component to recover the exergy expelled from the system should be developed for more cost-saving system.

The water injection to the combustor to control NO emission increases the exergy destruction in combustor from 14.8 to 15.8 percent of the fuel exergy because of heat absorption due to evaporation of injected water in combustor, which suggests that steam injection to the combustor is a better method to control NO. NO concentrations calculated with equilibrium assumption at oxygen 15 percent dry basis were considerably larger than those obtained by experiments as shown in Tables 3, 4, and 5. Note that the measured NO_x consists of NO and NO_2. More detailed analysis should be done on the formation of NO_x in this system.

Finally it is noted that the pressure drop in the combustor is about 4 percent of the gas pressure at the outlet of the compressor. However, calculation has revealed that such pressure drop

Table 5 Measured and predicted gas turbine performance and exhaust gas data for case of different inlet air temperatures with water injection (ISO 1155 kW basis); *measured at generator; **measured NO_x includes NO as well as NO_2

		Case 7		Case8		Case9	
Suc. air press.	kPa	103.32		103.32		103.32	
Suc. air temp.	°C	0		15		30	
Suc. air humidity	%RH	60		60		60	
HPC inlet air temp	°C	40.2		46.6		50	
Fuel	-	CH_4 : 89.3 , C_2H_6 : 5.4 , C_3H_8 : 3.5 , C_4H_{10} : 1.8					
Fuel flow rate	Nm^3/h	405		404		375	
Air mass flow rate	kg/sec	5.305		4.915		4.537	
Water injection rate	kg/h	255		255		236	
		Cal.	Exp.	Cal.	Exp.	Cal.	Exp.
Power*	kW	1258	1100	1,250	1095	1117	950
Exhaust gas temp	°C	474	451	509.6	497	511.5	511
Exhaust gas mass flow rate	kg/sec	5.47	5.47	5.08	5.08	4.68	4.69
Constituent,exh. gas							
CO_2	Vol.%	3.04	3.1	3.25	3.4	3.25	3.4
H_2O	Vol.%	8.1	8.1	9.26	9.3	10.69	10.7
N_2	Vol.%	74.03	74	73.28	73.2	72.17	72.1
O_2	Vol.%	13.95	13.9	13.34	14.7	13.04	13
Ar	Vol.%	0.87	0.9	0.86	0.9	0.85	0.9
NOx emission**	ppmv	119.6		150.4		142.9	

does not affect the outcome such as the temperature at turbine outlet as well as the efficiency of the turbine.

6 Conclusions

The exergy balance equation, which is applicable to any component of thermal system, has been formulated in this analysis.

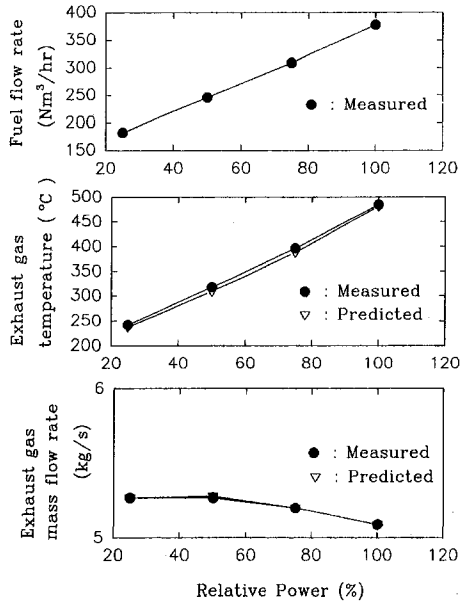


Fig. 6 Predicted exhaust gas temperature at turbine outlet and fuel and exhaust mass flow rate at full and part-load condition without water injection (ISO 1155 kW) together with the correction curves from the gas turbine manufacturer

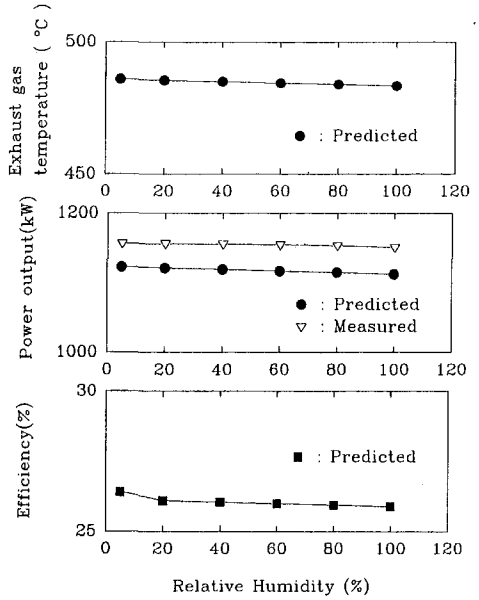


Fig. 7 Predicted gas turbine performance depending on the relative humidity (reference case is No. 2) together with corrected curve from gas turbine manufacturer

One of distinct features of this formulation is that the exergy involved in the component of any thermal system can be decomposed into exergy flows, entropy production flows, and the appropriate exergy rate terms such as fuel and available work. It turns out that the chemical exergy of fuel is the appropriate form of the exergy term for the combustor provided that the entropy generation due to internal irreversibilities is only our concern. Computerized energy and exergy analysis for the gas turbine cogeneration system permit to estimate the thermal efficiency of the system, the exergy destruction in each components as well as the mass flow rate, composition, and temperature of the exhaust gases at the outlet of the waste heat boiler. It has been found that the measured data of net power and the properties of exhaust gases are in good agreement with predicted ones. It has also been found that the exergy destruction due to water injection to the combustor is about 1 percent of the fuel exergy. The exergy balance equation may be utilized in exergoeconomic analysis (Kim et al., 1995) to estimate the production costs depending on various input costs in a gas turbine cogeneration system.

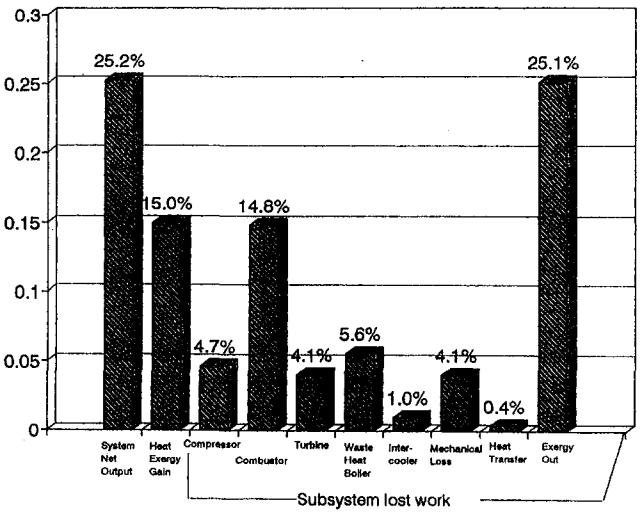


Fig. 8 Exergy losses in the gas turbine cogeneration system

References

Baughn, J. W., and Bagheri, N., 1987, "The Effect of Thermal Matching on the Thermodynamic Performance of Gas Turbine and IC Engine Cogeneration System," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 109, pp. 39–45.

Baughn, J. W., and Kerwin, R. A., 1987, "Comparison of the Predicted and Measured Thermodynamic Performance of a Gas Turbine Cogeneration System," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 109, pp. 32–38.

Bejan, A., 1988, *Advanced Engineering Thermodynamics,* Wiley, New York.

Chapra, S. C., and Canale, R. P., 1988, *Numerical Methods for Engineers,* McGraw-Hill, New York.

Dunbar, W. R., Mood, S. D., and Lior, N., 1993, "Exergy Analysis of the Lasalle Country Nuclear Power Station," *Thermodynamics and the Design, Analysis and Improvement of Energy Systems,* H. J. Richter, ed., ASME HTD-Vol. 266, pp. 377–387.

El-Masri, M. A., 1985, "On Thermodynamics of Gas-Turbine Cycle: Part 1—Second Law Analysis of Combined Cycles," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 107, pp. 880–889.

Gordon, S., and McBride, B., 1971, "Computer Program for Calculation of Complex Chemical Equilibrium Compositions, Rocket Performance, Incident and Reflected Shocks and Chapman–Jouguet Detonations," NASA SP-273.

Heywood, J. R., 1988, *Internal Combustion Engine Fundamentals,* McGraw-Hill, New York.

Howell, J. R., 1987, *Fundamentals of Engineering Thermodynamics,* McGraw-Hill, New York.

JANAF Thermochemical Tables, 1971, National Bureau of Standard Publications, NSRDS-N3537, Washington, DC.

JSME Steam Tables, 1968, JSME.

Kim, S., 1995, "Study on Exergy, Exergoeconomic and Optimization Analysis and Operational Strategy for a Gas Turbine Cogeneration Systems," Master's Thesis, Chung-Ang University, Seoul, Korea.

Kim, S., Oh, S. D., and Kwak, H. Y., 1995, "A New Approach of Exergoeconomic Analysis of Thermal Systems," *Proc. KSME Spring Annual Meeting,* pp. 504–510.

Kwon, Y. H., Park, S. H., Cha, J. H., Oh, S. D., Kim, C. T., Han, D. H., Kim, S. W., Sohn, W. S., Pang, H. S., Kim, S. M., Kim, T. K., and Kwak, H. Y., 1995, "Research and Development of 1000 kW Class Packaged Type Gas Turbine Cogeneration System," *Proc. KSME Thermal & Fluids Engineering Division,* pp. 34–41.

Lundberg, M., 1991, "Latent Heat Utilization in Steam-Injected Gas Turbine Application," *Proc. ASME COGEN-TURBO,* IGTI-Vol. 6, pp. 9–18.

Moran, M. J., 1982, *Availability Analysis,* Prentice-Hall, New York.

Phillips, J. N., 1986, "A Study of Off-Design Performance of Integrated Coal Gasification Combined-Cycle Power Plant," Ph.D. Dissertation, Stanford University, CA.

Stoecker, W. F., 1989, *Design of Thermal Systems,* McGraw-Hill, New York.

Thermal Cycle Analysis and Component Aerodesign for Gas Turbine Concept in Low-Range Cogenerating Systems

F. Bozza

A. Senatore

R. Tuccillo

Dipartimento di Ingegneria Meccanica per l'Energetica (D.I.M.E.), Università di Napoli Federico II, Napoli, Italy

The authors link together their previous experiences in gas turbine plant analysis and aerodynamic design of radial flow compressors. In recent papers they have introduced a method for the performance estimation of gas turbine engines, based on the prediction of the matching conditions among the several components in the whole operating range. On the other hand they have expressly paid attention to the problem of optimal design of radial flow compressors for satisfactory operation within an assigned operating range. In this paper, the authors present an integrated method, which aims to define the optimal characteristics of a low-power gas turbine engine (i.e., in the range 500–2000 kW). In this case, the radial compressor performance plays an important role as regards gas turbine operation for both power generation and cogeneration applications. The analysis proceeds with the optimization of rotating components (i.e., radial compressor and axial flow turbine) for given thermal cycle parameters. The prescribed objectives of the optimizing procedure are related to performance levels not only at the reference design conditions but also throughout the operating field. A particular emphasis is given to the extension of the field of satisfactory performance for cogeneration applications, with best fitting of mechanical and thermal power requirements. The aerodynamic design of radial flow compressor utilizes a method based on genetic algorithms.

Introduction

The increasing interest in low-range gas turbines for self-production of electricity in cogenerating plants solicits the researchers in this field to develop methods for their reliable design and analysis. As is well known, a particular requirement for this kind of plants is to meet a time-varying thermal and mechanical load. For this reason, both design criteria and thermodynamic analyses must take into account the off-design performance of both the gas turbine and the whole cogenerating plant.

In their recent papers, the authors have dealt with the analysis of gas turbine based power plants, by introducing methods for an accurate prediction of its behavior under several operating conditions and plant configuration (Tuccillo and Bozza, 1993, 1994). The analysis essentially consists of a thermodynamic model, which accounts for gas composition, combustion process, and blade cooling, linked to fluid-dynamic models of both compressor and turbines. In particular, the compressor performance characteristics are derived from a preliminary fluid-dynamic simulation and are introduced in terms of operating maps. The off-design behavior of the turbine is, instead, directly estimated within the gas turbine model. In this case, in fact, the turbine performance prediction is strictly dependent on the inlet conditions, in terms of gas thermal state and composition, which affect the flow characteristics (i.e., choking conditions, Reynolds and Mach numbers, etc.). In their experience, the authors have demonstrated the ability of their proposed methodology in determining the gas turbine operating field, both in the case of changes to its initial configuration and in the study of off-design operation. The latter is of special interest for the subject of the present paper, which essentially deals with the determination of the operating field of a gas turbine based cogenerating set.

A further field of interest for the authors has been the fluid-dynamic optimization of turbomachines such as axial and radial flow compressors and turbines (Tuccillo, 1990; Tuccillo and Senatore, 1991, 1994). Their most recent experiences (1994) have, in particular, shown the possibility of performing an effective design of radial flow impellers for a wide variety of design specifications and even when an extension of its operating range is expressly required. This authors' methodology consists of a "quasi-three-dimensional" relative flow model, which includes the loss estimation resulting from calculation of boundary layer growth along end-wall and blade surfaces. Flow separation conditions are detected when a limit value for the blade load coefficient is reached, and correspond to a limitation of the compressor operating range. The optimal impeller configuration is found by applying an algorithm derived from a *genetic analysis* (Goldberg, 1989), which allows a fast search of the most suitable geometry for meeting the prescribed objective within a number of imposed constraints (i.e., best adiabatic efficiency for an assigned pressure ratio and within limits on blade loading, peripheral velocity, Mach numbers, etc.).

The radial flow compressor optimization is of particular interest, since this compressor type is usually employed in low-range gas turbine generating sets. The capability of the method to find the compressor characteristics for a better fitting of an assigned operating range, corresponds to extended possibilities for a cogenerating plant to operate with satisfactory levels of both overall efficiency and energy saving index.

This will be demonstrated in the next sections, which proceeds with a brief description of the procedure followed for the selection of the component characteristics and of the thermal cycle parameters and, then, with the analysis of the plant performance that result from different choices of the compressor con-

Contributed by the International Gas Turbine Institute and presented at Cogen-Turbo Power, Vienna, Austria, August 23–25, 1995. Manuscript received by the International Gas Turbine Institute June 2, 1995. ASME Paper No. 95-CTP-1. Associate Technical Editor: C. J. Russo.

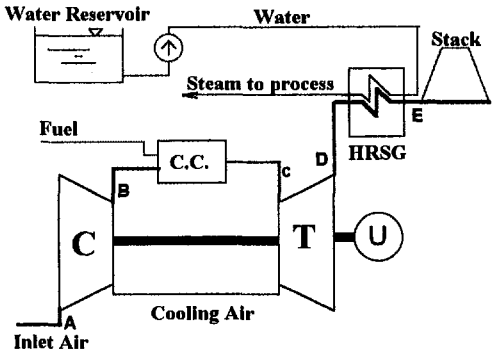


Fig. 1 Gas turbine power plant with heat recovery device

Table 1 Thermal cycle parameters

Gas Turbine Parameters			
Inlet Conditions	ISO std.	Max. Blade Temp.	1100 K
Pressure Ratio	9.0	Inlet Press. Loss	0.01
T.I.T.	1450 K	Exh. Press. Loss	0.01
Air Flow Rate	4.870 kg/s	Comb. Press. Loss	0.02
Comp. Ad. Eff.	0.855	Mechanical Eff.	0.98
Turbine Ad. Eff.	0.880	Combustion Eff.	0.98
Cogeneration Parameters			
(Saturated Steam)		ΔT_{min} approach	30 K
Steam Temp.	443.6 K	ΔT_{min} pinch point	10 K
Steam Pressure	8 bars	min. Stack Temp.	403 K

Table 2 Thermodynamic model results

Gas Turbine Performance		Cogeneration Performance	
Mechanical Output	1592 kW	Steam Flow Rate	1.188 kg/s
Thermal Eff.	0.316	FEUF	0.889
Exhaust Temp.	927.6 K	FESR	37.64 %
Exh. Flow Rate	4.973 kg/s	Thermal Output	2891 kW
Fuel Flow Rate	0.103 kg/s	Ref. overall Eff., η_{or}	0.34
Second Law Eff.	0.592	Ref. burner Eff., η_{br}	0.85

figuration. More comprehensive information, and details on the thermodynamic and fluid-dynamic models and on the optimizing procedure as well, may be found in the authors' papers referenced above.

Gas Turbine Thermal Cycle

In order to develop the examples presented in this paper, a 1600 kW level has been established for the gas turbine rated mechanical output. A single-shaft, simple cycle, gas turbine (Fig. 1) was considered with the choice of thermodynamic cycle parameters as indicated in Table 1.

The thermodynamic model results in Table 2 allow a first estimation of the performance level both for mechanical power generation and for cogenerating activity.

For evaluating the cogeneration plant performance, the authors have assumed some classical energy indices (Horlock, 1987), since a more comprehensive thermo-economic analysis is not the purpose of this paper. The performance criteria which were considered are, therefore:

Fuel Energy Utilization Factor (FEUF):

$$\text{FEUF} = \frac{W + Q_u}{F} \tag{1}$$

Fuel Energy Saving Ratio (FESR):

$$\text{FESR} = \frac{\Delta F}{F} = 1 - \frac{W/\eta_o}{W/\eta_{or} + Q_u/\eta_{br}} \tag{2}$$

In these relationships, W and Q_u represent the mechanical and thermal plant output, respectively, and η_{or} and η_{br} are the reference values for overall efficiency and burner (or thermal device) efficiency. The actual plant efficiency is given by η_o, which implies the actual fuel rate consumption, F.

With regard to the rotating components, a 9:1 pressure ratio radial flow compressor was chosen, as an effective device for gas turbine size and weight reduction. The details on compressor characteristics and results of its optimized design are discussed in the following section. An axial flow turbine was considered for gas expansion and its features were determined by a different optimizing process (Rizzo and Tuccillo, 1987). A fundamental choice was represented, in this case, by the number of stages. A preliminary comparison was made between a two-stage and a three-stage turbine. This allowed the two-stage solution to be discarded, since highly loaded stages would involve poor flexibility of the gas turbine under different operating conditions. An increased number of stages allows the gas turbine to meet wider variations in external load requirements, which usually produce changes in reduced mass flow rate and rotating speed. The three-stage turbine represents, therefore, a good compromise between a reduced length of the whole generating set and a satisfactory extension of the operating field.

Radial Flow Compressor Optimization

In a previous paper (Tuccillo and Senatore, 1994), the authors have already applied a *genetic algorithm optimization procedure* (Goldberg, 1989) to the design of a radial flow impeller. By means of this new optimizing technique, it is possible to overcome some typical limits of the numerical optimization methods. These are mainly related to the presence of local minima in the objective function, which make it difficult to move toward the absolute optimal solution. In addition, methods that operate with continuous derivatives encounter difficulties when

Nomenclature

$c(\bar{x})$ = equality constraint
C_L = load coefficient
c_h, c_s, d_h, d_s = form factors for blade meridional profiles
$f(\bar{x})$ = objective function
F = fuel flow rate
FEUF, FESR = fuel energy utilization and saving factors
F_{dim} = dimensional factor
k_1, k_2 = blade angle law constants
$\dot{m}$ = mass flow rate
M = Mach number
n = rotational speed
Q_u = thermal output
$\mathcal{R}$ = reaction degree
u = peripheral velocity
V_x, V_r = axial or radial coordinate variation
W = mechanical output
Z = number of blades
β_b = blade angle
β_{tt} = total-to-total pressure ratio
η_{tt} = total-to-total efficiency
η_{or} = reference overall efficiency
η_{br} = reference burner efficiency

Subscripts

A, B, C, D = end points of impeller profile
b = burning
c = absolute velocity
h, s = hub, shroud
W = relative velocity
* = design point

dealing with discrete impeller geometric variables like the number of blade and the presence or absence of splitter blades. The objective function itself is characterized by discontinuities due to both the effect of constraint violation and unreliable flow conditions induced by the compressor geometry.

The *genetic algorithm* allows generation of a *population* of impellers by means of the variation of the *genes* associated with the variables of the optimization problem. The algorithm applies satisfactorily to both continuous and discrete independent variables. So it is possible to discard more easily those impeller geometries characterized by unsatisfactory performance level or constraint violations. This results in an extended field of search, because no zero-derivative condition has to be found.

The method starts from an actual impeller geometry and allows us, therefore, to obtain geometries strongly different from the original one. So, the search of the optimum condition can be enlarged within a wide range of design specifications. Moreover, it is possible to obtain an impeller able to produce optimum flow conditions in an assigned operating range. This is of particular interest in the case of applications to combined cycle installations and in off design conditions. An acceptable compressor operation is considered to occur within the typical limits for blade loading (Eckardt, 1976) and choked flow.

The impeller design method is based on an initial geometry configuration whose performance level is well established and experimentally tested and on the preliminary identification of some geometric parameters to define the impeller configuration completely. These parameters are used as independent variables of the optimizing procedure (Senatore and Tuccillo, 1990, 1991). So the initial geometry generates a wide variety of new impeller configurations through changes in these parameters or *genes.* Each of these represents the actual position of an independent variable within an equally spaced interval of search. The interval between the minimum and maximum value is subdivided into N subintervals by equally spaced points. Therefore, each independent variable for the impeller definition may be identified by $N + 1$ integer values of the corresponding *gene.* The whole configuration is, in this way, identified by a *genetic code,* which consists of the string of the gene values.

The search of the optimum genetic code is organized in two steps. In the first one, each impeller genetic code is obtained with a random generation so that an impeller population is generated and a statistical investigation is carried on, to the aim of calculating a preliminary optimum level of the objective function. This leads to a first choice for the optimal genetic code.

In the second step, the same statistical investigation is repeated within modified search intervals. In particular, the initial interval:

- will be *restricted* around the previous optimal value found for each gene of the impeller geometry, if this value has been found inside the interval investigated;
- will be *shifted* if the optimal genetic code suggests variable values close to the end of the corresponding interval;
- each variable domain is again *enlarged* if the first step of the procedure has not produced improvements with respect to the optimal conditions previously found.

The whole procedure continues with the reiteration of the first and second steps until the investigation no longer provides improvements in the objective function and variations in the optimal variable values. It is to be highlighted that in this procedure (Tuccillo and Senatore, 1994) are present a number of significant modifications with respect to the classical genetic algorithms. In particular:

- the adaptive variation of the investigated domains allows an optimization refinement;
- the search of the optimum condition may be extended outside the initial interval established for each variable.

In this way the typical problems related to the existence of a variety of local minima and to discontinuities in the objective function may be overcome.

The geometric variables are chosen based on the different influence that they exert on compressor characteristics:

- variables mainly affecting the whole operating range, like a dimensional factor, F_{dim}, and the rotational speed n;
- variables mainly influencing the blade shape, like the coordinates of the end-points of the blade meridional profile (Fig. 2) and the coefficients they define the elliptic curve used in the meridional plane to link the inlet and outlet straight lines at leading and trailing edges and to define the three-dimensional blade development;
- those mainly affecting the flow conditions and the blade loading, like the number of blade, Z, and the presence of the splitter blades.

The main purpose of the optimizing procedure is to meet a prescribed value of the total-to-total pressure ratio, β_{tt}^*, with a maximum efficiency level, so that the primary definition of the objective function is:

$$f(\bar{x}) = 1 - \eta_{tt} \tag{3}$$

with the equality constraint:

$$c(\bar{x}) = 1 - \frac{\beta_{tt}^*}{\beta_{tt}} = 0 \tag{4}$$

The following inequality constraints are also established:

Relative inlet Mach number:	$M_{w1} \leq 1.10$
Peripheral velocity:	$u_2 \leq 700$ m/s
Blade load coefficient:	$C_L \leq 2$
Absolute outlet Mach number:	$M_{c2} \leq 1.05$
Degree of reaction:	$\mathcal{R} \geq 0.5$

The objective function of the optimizing procedure is therefore built with the method of the augmented Lagrangian function (Gill et al., 1984).

The initial configuration is that of a compressor for aircraft engines, characterized by high rotational speed and blade loading. Variations in the initial characteristics are required to meet the condition established for the gas turbine:

$$\dot{m}^* = 4.9 \text{ kg/s} \quad \beta_{tt}^* = 9 \tag{5}$$

Two optimized designs were carried out. The first one establishes, as design specification, those defined above and assumes as primary objective the best adiabatic efficiency at the design point. The rotational speed is included in the independent variable set. This optimizing procedure starts from the original impeller configuration, whose geometric and operating characteristics are rather different from those to be found for the new compressor. The process for the second optimized design starts from the results of the first one but assumes a more complex objective function. In this case, the primary objective is the maximum in an average adiabatic efficiency level evaluated in a number of off-design points (i.e., for a ±20 percent variation in air flow at the design speed and for two reduced rotating speeds). The reference rotating speed is the same as that found by the first optimization. The purpose of this second design step is to ensure an acceptable compressor operation in a wider range than that resulting from the first step. Really, it should be considered that off-design flow conditions must produce blade loading levels not exceeding the constrained value, so allowing an extension in compressor stall limits.

Table 3 Results of the compressor optimizations

Var. Name	Search Interval	Init. Config.		Optim. n.1		Optim. n.2	
$\dot{m}$*	--	2.8		4.9		4.9	
β_{tt}	--	5.83		9.01		9.00	
η_{tt}	--	0.862		0.854		0.851	
		value	gene	value	gene	value	gene
n	i.v.± 12000	38095	5	26238	1	26238	1
F_{dim}	0.25 ÷ 4	1.	5	1.49	6	1.46	6
V_{xC}	i. v. ± 15%	0.	5	-7.93	3	-15.13	1
V_{xD}	i. v. ± 15%	0.	5	-12.72	2	-19.78	1
V_{rA}	i. v. ± 15%	0.	5	4.50	6	6.37	7
V_{rB}	i. v. ± 15%	0.	5	-3.40	4	4.48	6
V_{rC}	i. v. ± 15%	0.	5	7.70	7	7.74	7
c_h	i. v. ± 0.1	-0.1	5	-0.005	9	-0.091	5
d_h	i. v. ± 0.1	0.35	5	0.316	4	0.300	3
c_s	i. v. ± 0.1	0.42	5	0.455	6	0.423	5
β_{1bs}	i. v. ± 10	-62.9	5	-50.54	1	-54.52	2
β_{2b}	i. v. ± 20	-34.78	5	-40.20	6	-31.30	4
k_1	i. v. ± 2	2.6	5	2.78	5	2.80	5
k_2	i. v. ± 4	1.77	5	1.99	5	2.98	6
Z	6÷30	13	13	9	9	17	17
split	0÷1	yes	1	yes	1	yes	1

Table 3 summarizes the results of the two optimizing processes, in terms of optimal values found for independent variables, compared with the initial ones (i.v.), and of performance level at the design point. This confirms the ability of the genetic algorithm to identify impeller characteristics strongly different from the initial ones for the best fitting of design data. Figure 2 compares the new impeller meridional profiles with the original one.

Figures 3–5 show in higher detail the conditions resulting from the two different design steps. In particular, Fig. 3 allows comparison of the blade load coefficients at the design point, while Figs. 4–5 refer to off-design operation with both a reduced mass flow rate (Fig. 4) and a lower rotational speed (Fig. 5). These latter flow conditions refer to the minimum allowable flow rates for the first optimized impellers (i.e., with a maximum C_L value close to the limit value of 2), while the second one presents a significantly larger stall margin. This situation is summarized by Fig. 6, which compares the maximum values found for C_L within the operating fields of the two different impellers.

This explains the difference found when predicting the behavior of the operating curves for the two optimized compressors (Figs. 7–8). The first one presents a slightly higher efficiency all over its range but the latter is strongly reduced because of the blade loading conditions. The second compressor presents extended stall margins, which would produce benefits in gas turbine and cogenerating plant operation.

Cogenerating Plant Performance Analysis

A brief outline of the plant model is given, since a more comprehensive description can be found in some previous papers (Tuccillo and Bozza, 1993, 1994).

Thermodynamic Model. The model is based on the accurate evaluation of thermodynamic state in each characteristic point of the plant in Fig. 1; thermal cycle calculation can be performed only once the operating characteristics of turbomachinery, in terms of both pressure ratio and efficiency, have been determined. The calculation procedure can be hence divided into the following subsequent steps:

1 Values are chosen for the only two degrees of freedom of gas turbine: *shaft speed* and *fuel flow rate.*
2 An initial trial value is assigned to the air flow rate entering the compressor. The air flow rate value is of fundamental interest and will be modified according to the following steps.
3 Compressor pressure ratio and adiabatic efficiency can be thus evaluated from the performance map (Figs. 7 and 8) resulting from the previous optimization.
4 Mass and energy balance equations are applied to the evaluation of thermodynamic conditions at turbine inlet. Air bleeding for turbine blade cooling is also calculated with an "inverse" procedure, based on the well-known El-Masri (1986) approach, which determines the actual stage-by-stage blade temperature.
5 Turbine expansion ratio and efficiency are evaluated by a direct off-design calculation. This solution was preferred to the employment of given turbine performance maps, in order to take into account variations in chemical composition of the expanding gas. JANAF Tables (Svhela and McBride, 1974) were employed for the calculation of working fluid thermodynamic properties and composition.
6 A correction to the initial estimate of the air flow rate is based on the value calculated for the turbine discharge pressure, which should meet a given level (i.e., atmospheric pressure plus exhaust pressure losses). Steps 3–6 are hence repeated, to find the air flow through the compressor that ensures the achievement of the prescribed discharge pressure.
7 In correspondence to the actual exhaust gas temperature, the maximum steam mass flow rate through the heat recovery device is calculated as a function of imposed values of steam pressure and temperature and gas/steam differential temperature at the pinch point, according to data in Table 1.
8 Steps 2–7 are repeated for a different (shaft speed, fuel flow rate) pair.

Gas Turbine Operating Domain. A typical operating domain of the gas turbine, in a (shaft speed, fuel flow rate) plane, is reported in Fig. 9(*a*). In the following, shaft speed, fuel flow rate, mechanical output, and thermal output will be expressed as a fraction of their reference values in Tables 2 and 3. The operating region is bounded, on one side, by turbine choking and/or compressor stall and, on the other one, by compressor choking. Each point inside such a region represents a feasible plant operating point. Values of gas turbine power or efficiency are hence known in each point inside the operating domain, together to the cogenerating plant performance.

A better representation of the same domain can be drawn in a (shaft speed, mechanical output) plane (Fig. 9(*b*)). In this

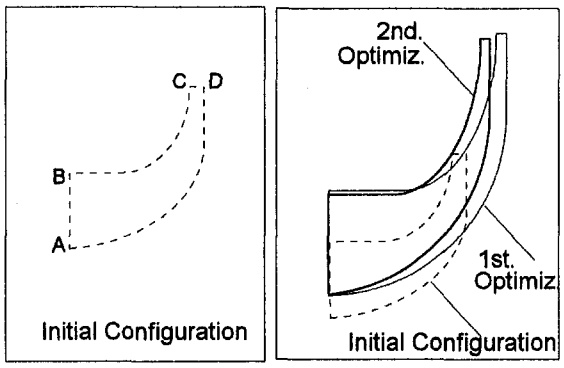


Fig. 2 Meridional profiles of the reference and optimized impellers

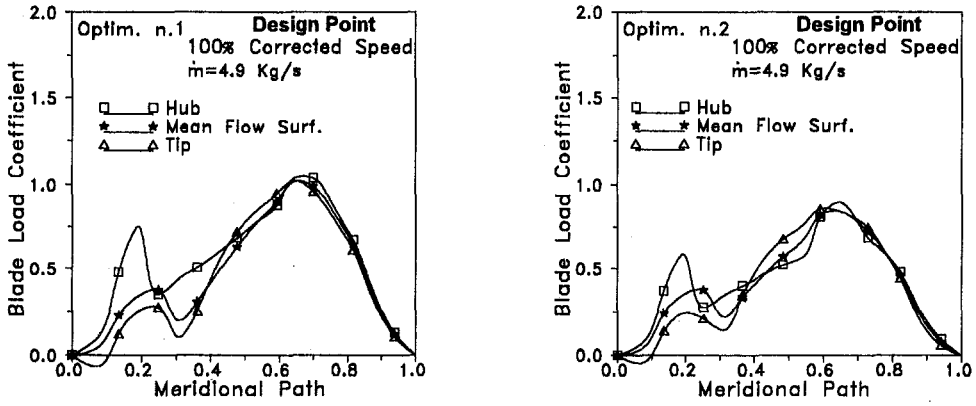


Fig. 3 Blade load coefficient pattern at reference conditions

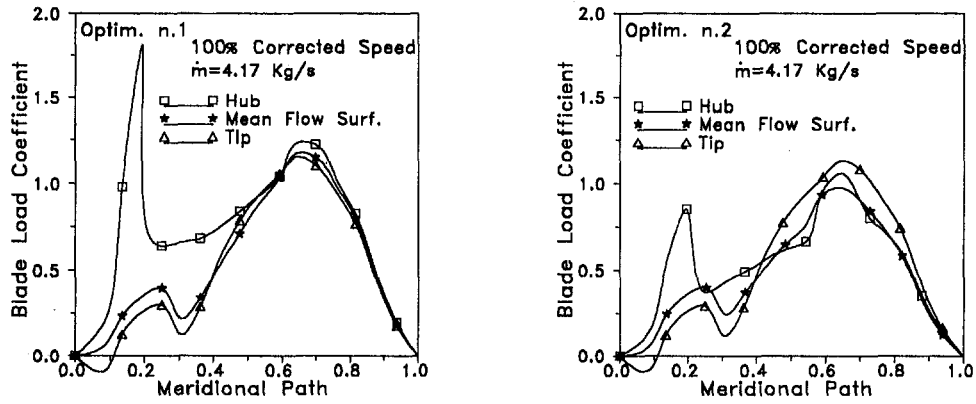


Fig. 4 Blade load coefficient pattern at reduced air flow

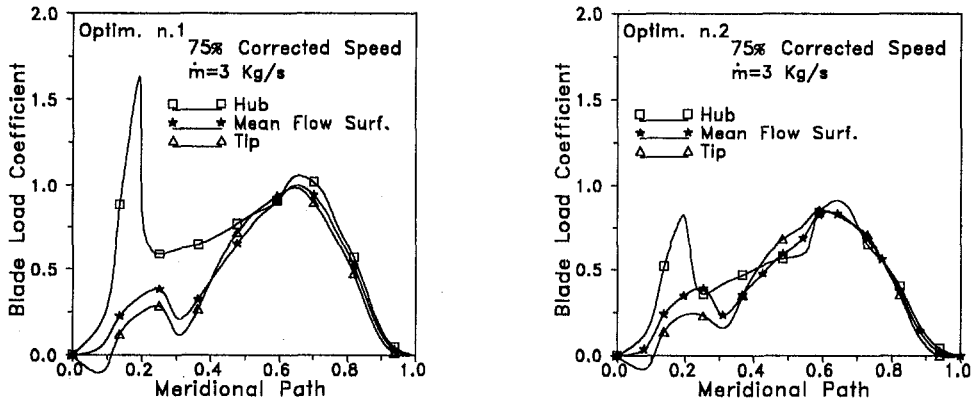


Fig. 5 Blade load coefficient pattern at reduced air flow and rotating speed

way the actual external load lines can be superimposed to check their matching with the gas turbine. In the same figure, examples are presented of both a constant-speed load and a variable-speed load, which are typical of electric generator sets and mechanical-drive applications, respectively.

The characteristics of the domain in Fig. 9 highlight the importance of a proper choice of the rotating components for a satisfactory extent of the gas turbine operating range. In particular, the left-upper side of this domain in Fig. 9(*a*) is bounded by either turbine choking or compressor stall conditions. This explains the more favorable situation that results from the three-stage turbine in terms of problems related to choked expansion. On the other hand, the importance of a proper compressor design for a wider gas turbine field of operation is confirmed. Really, the enlargement of stall limits, produced by the second compressor optimization, is of interest for the extent of the gas turbine operating domain.

Employment of the Optimized Compressors. In Fig. 10, the superimposition of the operating domain to the compressor characteristic maps demonstrates that the most significant benefits from the adoption of the second optimized impeller is expected at reduced shaft speed and mechanical output. Really, the second compressor stall would affect the operating range only at very low air flow rates, while the influence of the first compressor would be already significant starting from 90 percent corrected speed. Above these speeds, the main influence is exerted by turbine choking.

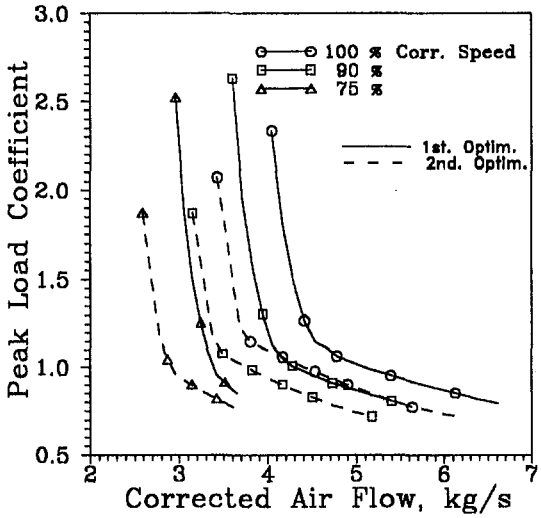


Fig. 6 Comparison of blade loading estimation for the optimized impellers

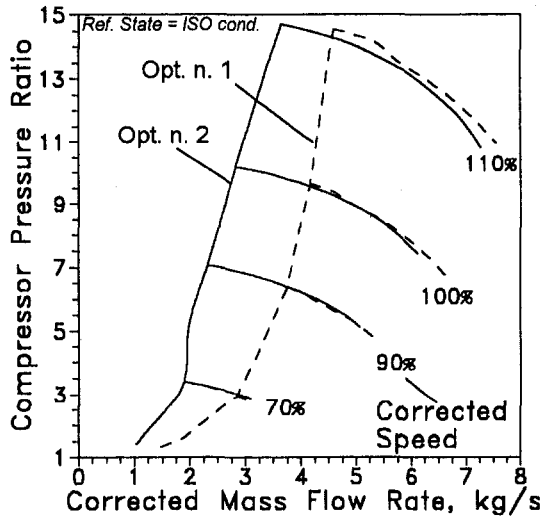


Fig. 7 Predicted compressor characteristic curves

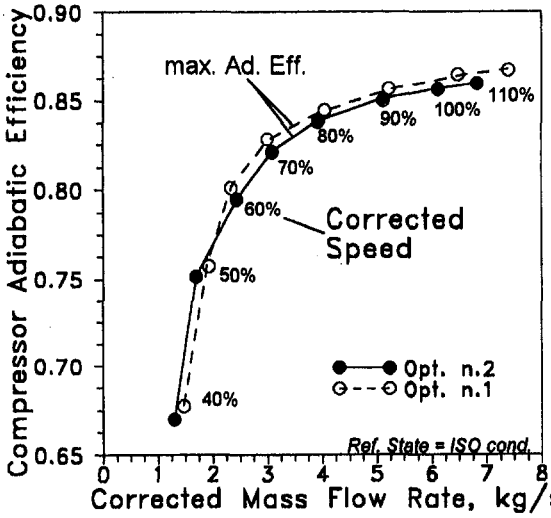
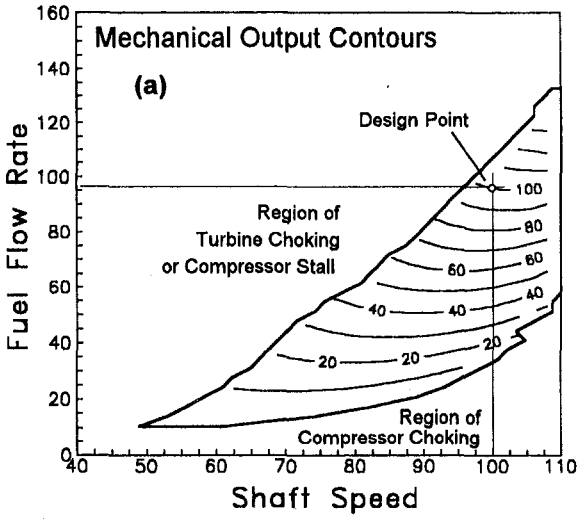


Fig. 8 Estimated efficiency levels for optimized compressors



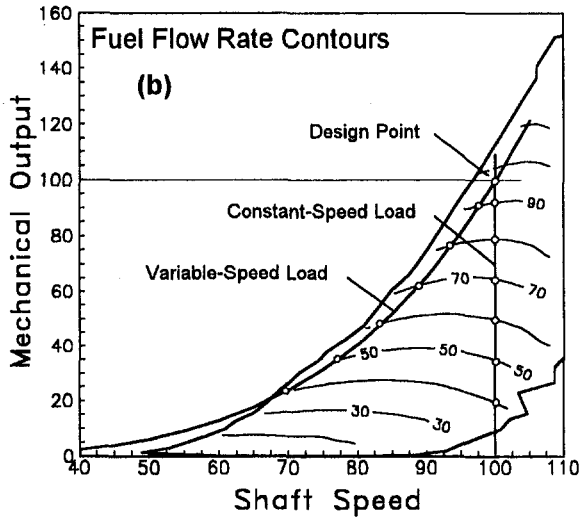


Fig. 9 Regions of gas turbine operation

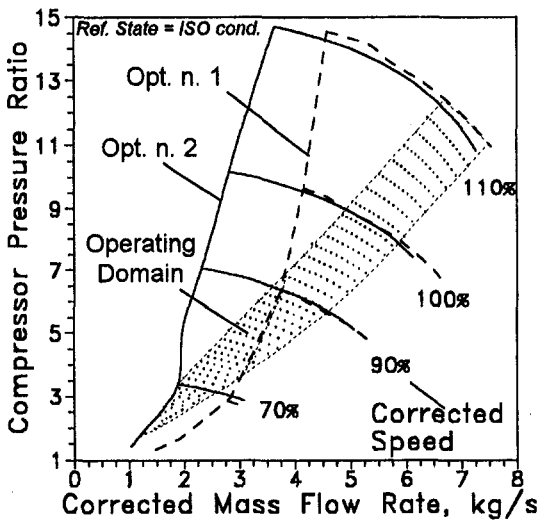


Fig. 10 Operating domain in the compressor plane

In Fig. 11, the performance parameter contours are drawn within the operating domains that result from the adoption of the two different compressors. This allows comparison of gas turbine performance when either the first or the second optimized compressor operates. The previous considerations are

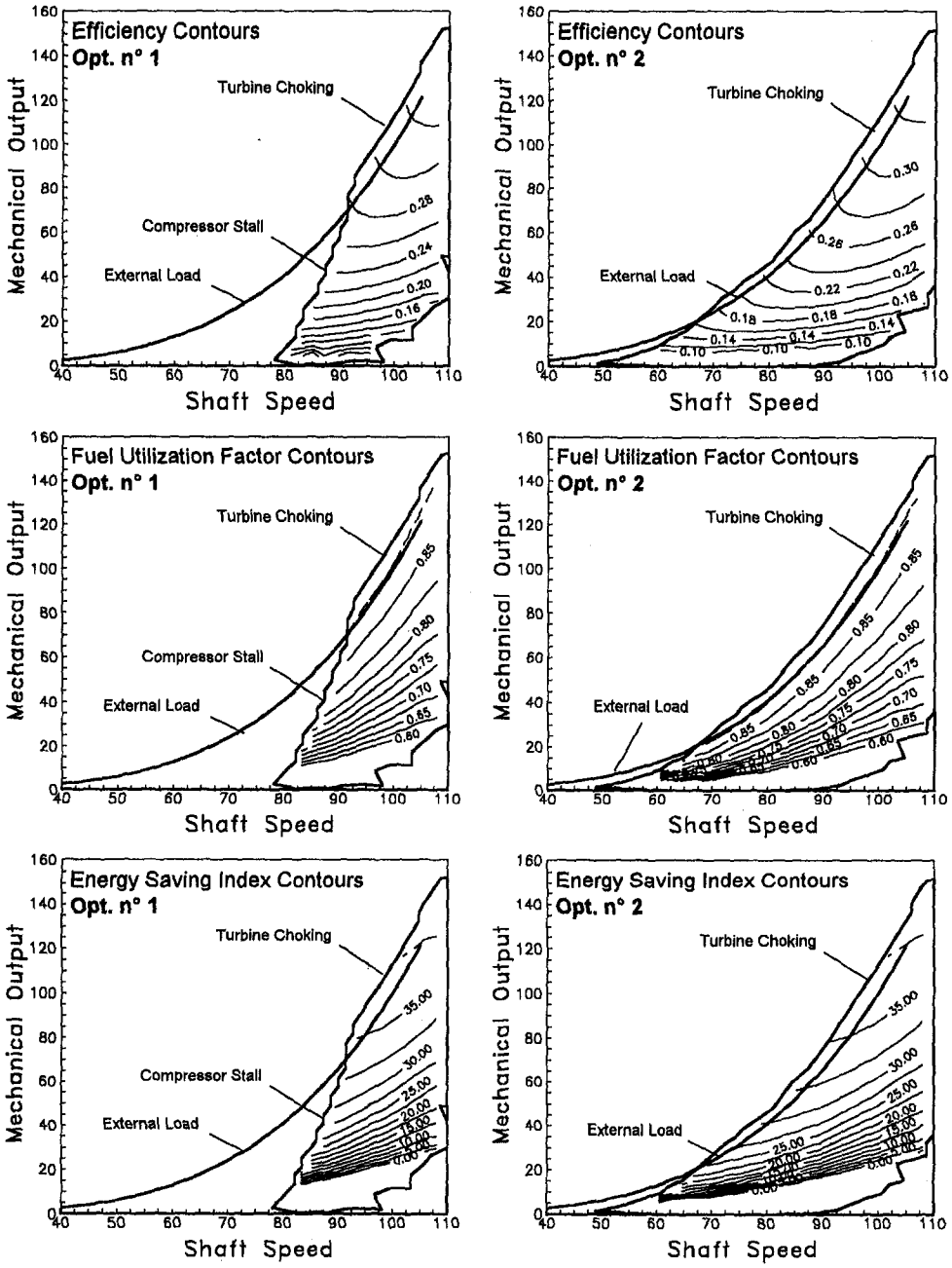


Fig. 11 Performance parameters contours for gas turbines adopting the two optimized compressors

confirmed, since more satisfactory performance levels are expected when the second compressor is introduced. The main advantages do not refer to the peak values of performance parameters, but more specifically, to the extended possibilities of matching with the external load with satisfactory values of cogeneration indices.

The general behavior of the cogenerating domain of the gas turbine may be then represented in terms of (mechanical versus thermal output, Fig. 12) or (mechanical output versus FESR index, Fig. 13). The light shaded region of these domains pertains to both the optimized compressors, while the dark-shaded one would be obtained only by adopting the compressor proposed by the second optimization. This demonstrates that the latter choice would allow one to meet a larger variety of simultaneous requirements of thermal and mechanical load and to extend the region of cogeneration with a satisfactory level of fuel energy saving. In fact, besides the matching with the typical variable-speed external load running line in Figs. 12 and 13, the gas turbine could operate in a reduced load range, with enhanced possibilities when the second optimized compressor is considered.

For a more realistic investigation of the different possibilities offered by the two compressor choices, some typical operating curves were extracted from the general operating domains in Figs. 9–13. In particular, further comparisons were made by referring to:

(*a*) Constant-speed load, for electric generator sets (Figs. 14 and 15).
(*b*) Variable speed load, for mechanical drive applications (Fig. 16).
(*c*) Constant thermal output (Fig. 17).

The constant-speed operation curves (Figs. 14 and 15) confirm that no significant variation in gas turbine efficiency may

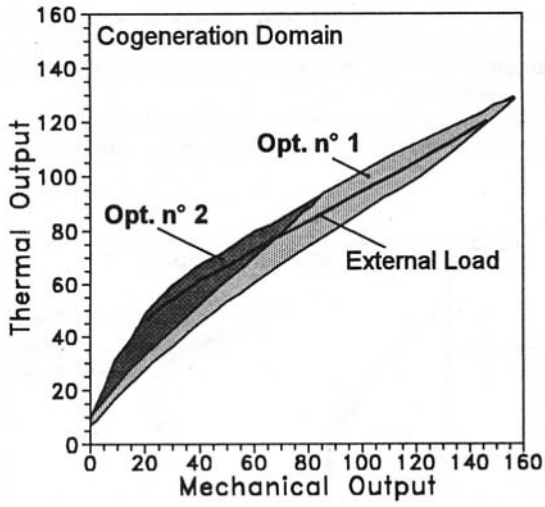


Fig. 12 Thermal and mechanical output ranges for cogenerating plants

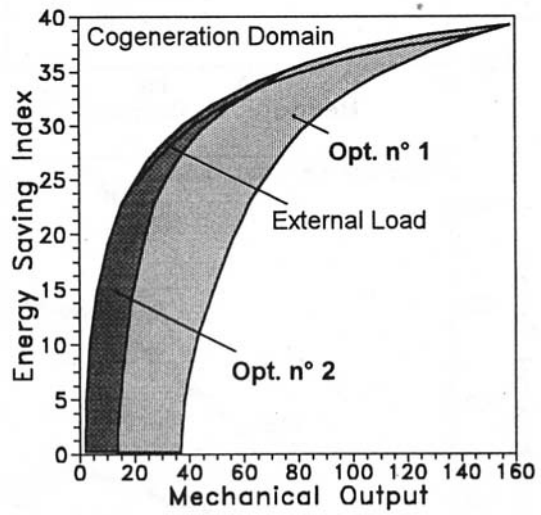


Fig. 13 Energy saving levels expected for cogenerating plants

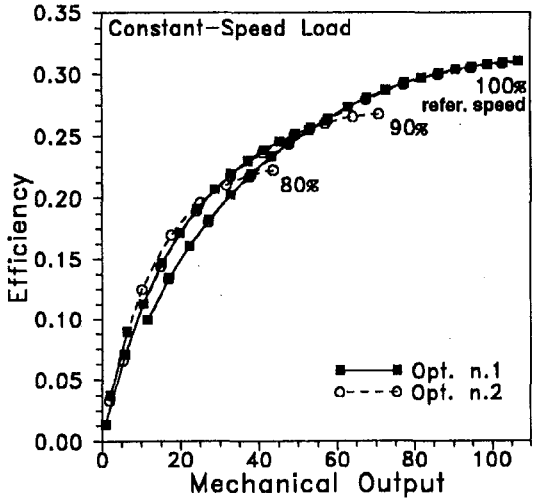


Fig. 14 Gas turbine efficiency for constant speed operation

be expected by the adoption of either the first or the second compressor. The most relevant information refers to the possibility that exists of operating at a constant-reduced speed (for instance, at 80 percent of nominal speed) only when adopting the second optimized impeller. Really, the introduction into the objective function of specific requirements for off-design operation has allowed an extension of the stall limits, thus im-

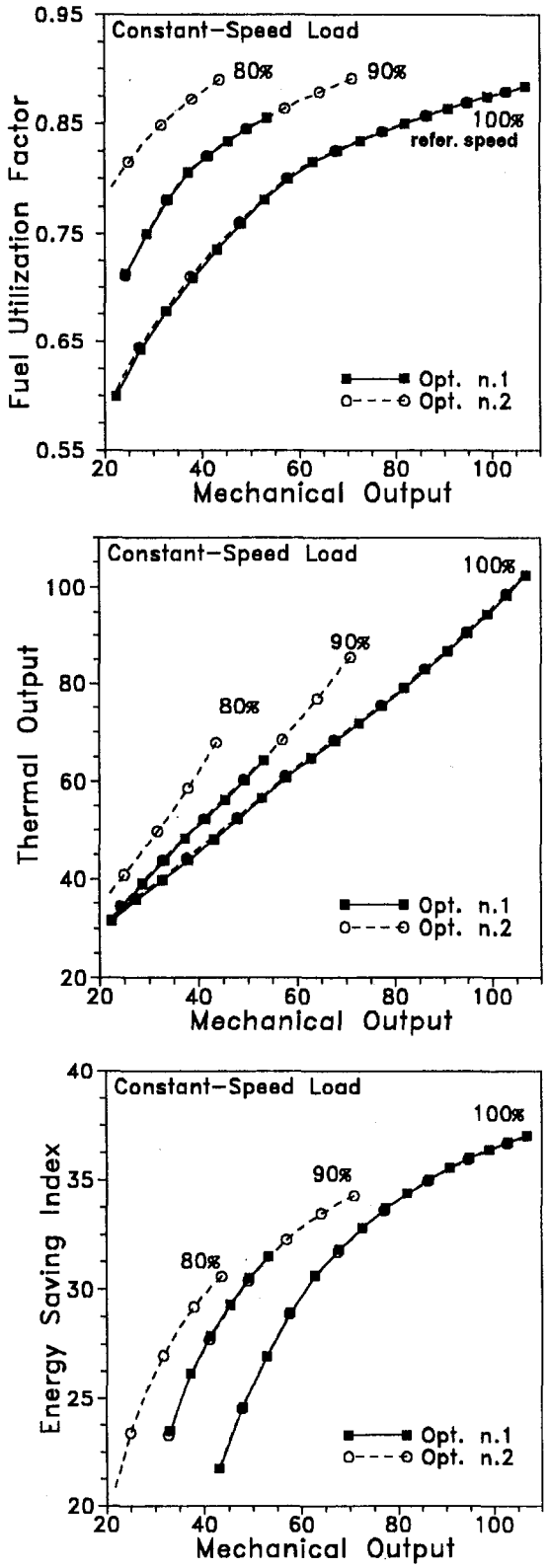
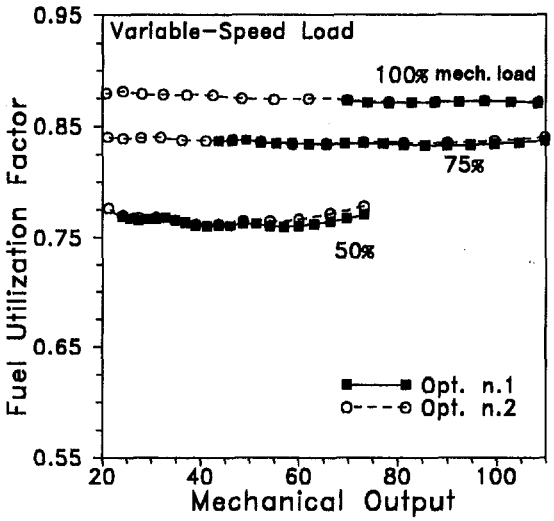
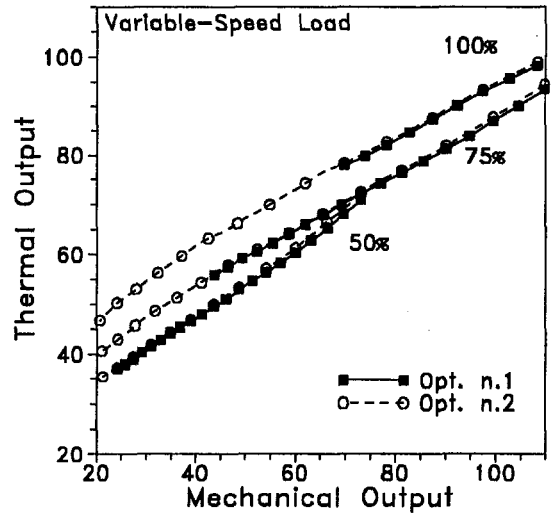


Fig. 15 Cogeneration performance for gas turbine constant speed





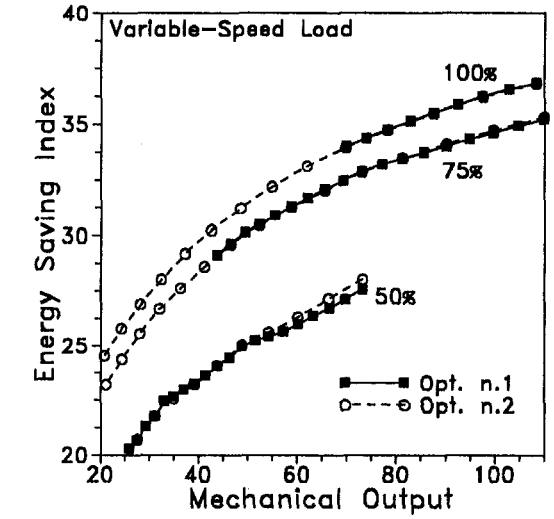
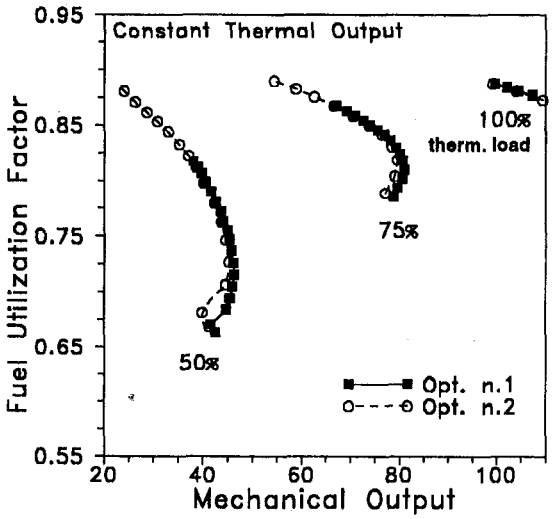


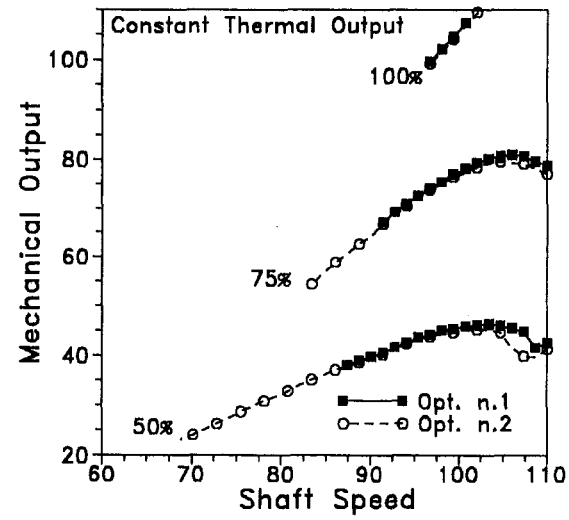
Fig. 16 Cogeneration performance for gas turbine variable speed

proving both compressor and plant operation at reduced air flow rates.

More important differences arise in terms of cogeneration performance parameters (Fig. 15), since satisfactory levels of fuel utilization factors (FEUF) and energy saving index (FESR) can be obtained when adapting the gas turbine generating set to a situation with a reduced electrical load. The curves in Fig. 15 confirm that some interesting conditions could be obtained only when adopting the second compressor. Of course, these conditions could only be obtained through a different choice of the gear-box ratio.

In the authors' opinion, more attractive advantages are shown in the case of mechanical drive application within a combined





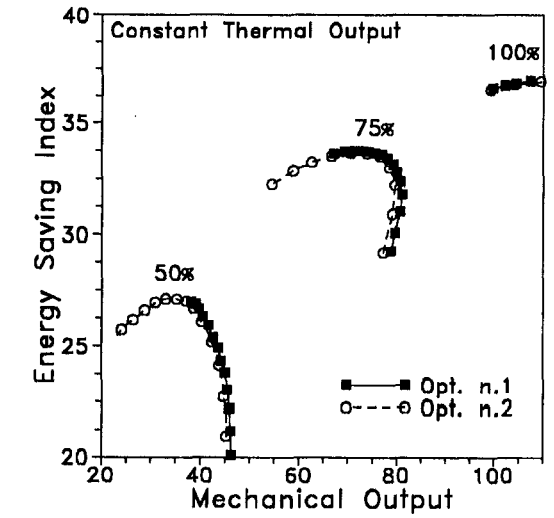


Fig. 17 Cogeneration performance for constant thermal load

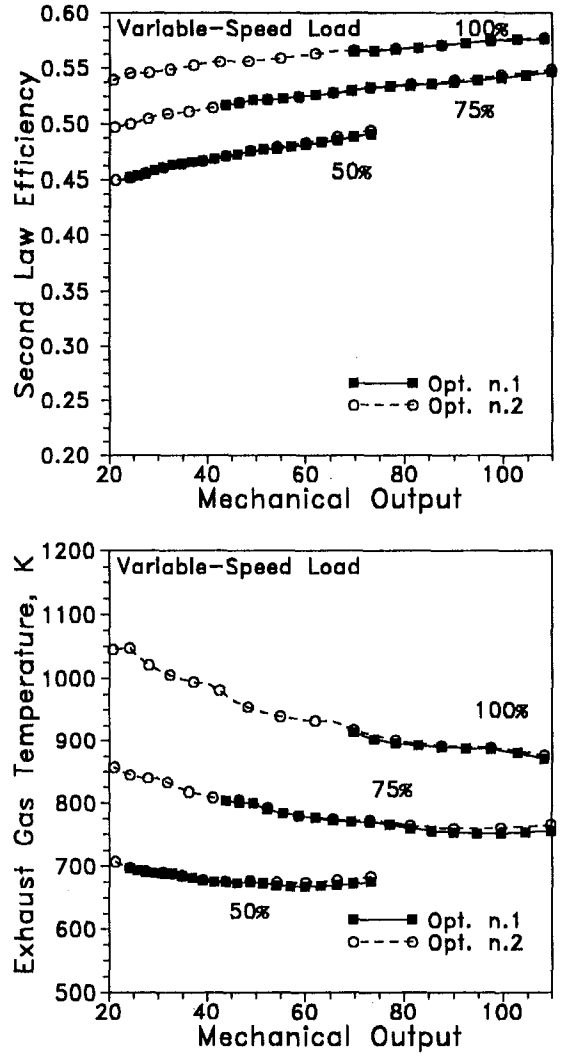


Fig. 18 Gas turbine second law efficiency and exhaust temperature

heat and power plant (Fig. 16). The 100–50 percent curves in this figure refer to different laws imposed by the external device for full-load or reduced-load operation with variable rotational speed. It should be pointed out that curves in Fig. 16 correspond to those drawn in Figs. 9–13, representing actual matching conditions of the power plant with the external load. The best benefits result from the employment of the second optimized compressor and for full-load conditions, which allow a satisfactory value of both FEUF and FESR to be maintained all over the gas turbine operating range. The utilization of the first compressor would dramatically reduce this possibility.

The third kind of operating curve at 100–50 percent of reference thermal output (Fig. 17) confirms that the advantages from the second compressor are mainly expected when a reduced level of thermal and mechanical output is required. This is a situation that may frequently occur within an industrial productive cycle and demonstrates that a proper choice of the gas turbine components leads to its better utilization for cogenerating applications. The (shaft speed, mechanical output) curves also suggest an optimal control strategy for taking the maximum advantage in the case of a constant thermal load.

A final consideration refers to the trend of second law efficiency and exhaust gas temperature shown in Fig. 18. This confirms that the choice of the second optimized impeller is favorable in terms of a significant extension of the operating range with a high thermodynamic quality of the recoverable heat.

Conclusion

The authors have proposed, in this paper, a method of analysis of gas turbine based cogenerating plants, which pays particular attention to the behavior of its single components. In particular, they have demonstrated the effectiveness of the procedure for radial compressor optimized design, in successfully modifying the gas turbine off-design operation.

The main advantages from this proper choice of the compressor result in a wider range of possibilities of part-load operation with satisfactory levels of the energy-saving criteria.

The resulting discussion has proposed that the most interesting benefits would arise in the case of mechanical drive applications of the gas turbine power plant. In this case, the cogenerating system would take full advantage of the extension of the compressor operating range.

Future developments of this work will take into account a more comprehensive optimization of the whole cogenerating system for its best operation in assigned conditions for thermal and mechanical load.

Acknowledgments

The research is developed with the financial support of the Italian Government (M.U.R.S.T. 40 percent funds) and of the National Research Council (C.N.R. contract No. 94.1937/CT07).

References

Al-Zubaidy, S. N. J., 1990, "Towards Optimizing the Design of Centrifugal Impeller," *Proc. of ASME COGEN-TURBO IV,* pp. 227–234.

Bejan, A., 1988, *Advanced Engineering Thermodynamics,* Wiley, New York.

Benvenuti, E., Bettocchi, R., Cantore, G., Negri di Montenegro, G., and Spina, R., "Experimental Validation of a Gas Turbine Cycle Model Based on a Simultaneous Solution Method," *ASME COGEN TURBO POWER '94 Proc.,* pp. 245–256.

Beveridge, G. S. G., and Sheckter, R. S., 1970, *Optimization in Theory and Practice,* McGraw-Hill, New York.

Bozza, F., and Tuccillo, R., 1992, "Performance Upgrading of Aero-Derivative Gas Turbines," *Proc. of FLOWERS '92 World Energy Research Symposium,* Firenze, June, pp. 271–288; NOVA SCIENCE, New York.

Bozza, F., Fontana, G., and Tuccillo, R., 1994, "Performance and Emission Levels in Gas Turbine Plants," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 116, pp. 53–62.

Colantuoni, S., and Colella, A., 1993, "Aerodesign and Performance Analysis of a Radial Transonic Impeller for a 9:1 Pressure Ratio Compressor," ASME *Journal of Turbomachinery,* Vol. 115, pp. 573–581.

Consonni, S., and Macchi, E., 1988, "Gas Turbine Cycles Performance Evaluation," *Proc. of ASME Cogen-Turbo III,* pp. 67–77.

Craig, H. R. M., and Cox, H. J. A., 1971, "Performance Estimation of Axial Flow Turbines," *Proc. of Instn. of Mech. Engn.,* Vol. XX.

Eckardt, D., 1976, "Detailed Flow Investigations Within a High-Speed Centrifugal Compressor Impeller," ASME *Journal of Fluids Engineering,* Vol. 98, pp. 390–402.

El Masri, M. A., 1986, "On Thermodynamics of Gas Turbine Cycles: Part 2—A Model for Expansion in Cooled Turbines," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 108, pp. 151–159.

Gill, P. H., Murray, W., and Wright, M. H., 1984, *Practical Optimization,* Academic Press, London.

Goldberg, D. E., 1989, *Genetic Algorithms in Search, Optimization, and Learning,* Addison-Wesley, New York.

Horlock, J. H., 1987, *Cogeneration: Combined Heat and Power Thermodynamics and Economics,* Pergamon Press, New York.

Horlock, J. H., 1992, *Combined Power Plants,* Pergamon Press, New York.

Katsanis, T., 1966, "Use of Arbitrary Quasi-Orthogonals for Calculating Flow Distribution in a Turbomachine," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 88, pp. 197–202.

Manfrida, G., and Stecco, S. S., 1982, "Exergy Loss Modeling of Power Plants and Industrial Processes," *Proc. of 17th IECEC Conference,* IEEE Paper No. 829004.

Perdichizzi, A., and Savini, M., 1983, "Aerodynamic and Geometric Optimization for the Design of Centrifugal Compressors," *International Journal of Heat and Fluid Flow,* Vol. 6, No. 1, pp. 49–55.

Rao, S. S., and Gupta, R. S., 1980, "Optimum Design of Axial Flow Turbine Stage (Part I and II)," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 102, pp. 782–797.

Rizzo, G., and Tuccillo, R., 1987, "Optimal Design Techniques for Multistage Turbomachinery," [in Italian], *Proc. of 42nd 'Congresso A.T.I.', Genova,* pp. III-209, 224; CLEUP, Padova.

Russo, C., Nicklaus, D., and Tong, S., 1987, "Initial User Experience With an Artificial Intelligence Program for the Preliminary Design of Centrifugal Compressors," ASME Paper No. 87-GT-217.

Senatore, A., and Tuccillo, R., 1990, "A Study on the Blade Shape Influence on Flow Distribution and Performance of Centrifugal Compressor Rotors," ASME Paper No. 90-GT-99.

Svhela, S. A., and McBride, B. J., 1974, "Fortran IV Computer Program for Calculation of Thermodynamic and Transport Properties of Complex Chemical Systems," NASA TND-7056.

Tong, S. S., and Gregory, B. A., 1992, "Turbine Preliminary Design Using Artificial Intelligence and Numerical Optimization Techniques," ASME *Journal of Turbomachinery,* Vol. 114, pp. 1–7.

Tuccillo, R., 1990, "Optimized Design of Axial Flow Compressors With Respect to the Field of Operation," *Proc. of 1990 ASME COGEN-TURBO IV,* pp. 241–248.

Tuccillo, R., and Senatore, A., 1991, "Optimized Design of Radial Flow Compressor Rotors," *Proc. of ASME COGEN-TURBO V,* pp. 225–234.

Tuccillo, R., and Bozza, F., 1993, "The Influence of Innovative Solutions on Gas Turbine Plant Operation," *Proc. of ASME-IGTI COGEN TURBO POWER '93,* pp. 367–378.

Tuccillo, R., and Bozza, F., 1994, "An Integrated Study of Advanced Cycles and Component Operation in Gas Turbine Power Plants," *Proc. of FLOWERS '94 Int. Symp.,* pp. 381–394.

Tuccillo, R., and Senatore, A., 1994, "Optimum Design of Radial Flow Impellers for Variable Load Operation," *Proc. of ASME COGEN TURBO POWER '94,* pp. 151–162.

R. Yokoyama
Associate Professor.
Mem. ASME

K. Ito
Professor.
Mem. ASME

Department of Energy Systems Engineering,
Osaka Prefecture University,
1-1, Gakuen-cho, Sakai, Osaka, 593 Japan

Y. Matsumoto
Manager,
Technical Research Center,
The Kansai Electric Power Co., Inc.,
3-11-20, Nakoji, Amagasaki,
Hyogo, 661 Japan

Optimal Multistage Expansion Planning of a Gas Turbine Cogeneration Plant

A multistage expansion planning problem is discussed concerning a gas turbine cogeneration plant for district heating and cooling using an optimization approach. An optimal sizing method for single-stage planning proposed by the authors is extended to this case. Equipment capacities and utility maximum demands at each expansion stage are determined so as to minimize the levelized annual total cost subject to increasing energy demands. A numerical study on a simple-cycle gas turbine cogeneration plant to be installed in a district development project clarifies the relationship between optimal expansion planning and energy demand trend, and shows the effectiveness of the proposed method.

Introduction

Recently, district heating and cooling systems have been installed increasingly in urban areas with their renewal projects. As dispersed and small-scale plants for district heating and cooling, cogeneration plants with gas turbines are expected to compensate for energy supply by a large-scale electrical grid because of their high economical and energy-saving potentials. In order to utilize these potentials, planners are urged to propose improved plans for the fundamental design of cogeneration plants. It is considered an important challenge to determine rational equipment capacities and utility maximum demands as well as operational strategies.

First, the authors proposed an optimal operational planning method of determining plants' operational strategies for the variations of both electricity and thermal energy demands, and applied it to an evaluation of several types of gas turbine cogeneration plants (Ito et al., 1987, 1988, 1990, 1995; Yokoyama and Ito, 1991). Second, they presented an optimal sizing method to determine equipment capacities and utility maximum demands from a long-term viewpoint in consideration of plants' operational strategies (Yokoyama et al., 1994). In these papers, optimal operational strategies and optimal sizes were discussed only for energy demands estimated at the final state of district heating and cooling projects.

However, in many projects, buildings are constructed stepwise for several years, while energy demands increase with the progress of building construction. In such cases, only one installation of equipment at the initial state of projects causes a decrease in availability of equipment and an increase in investment. Therefore, multistage expansion planning of cogeneration plants should be considered corresponding to increasing energy demands. However, many alternatives exist concerning the installation timing, capacities, and numbers of equipment to be installed, which makes it very difficult for planners to find good expansion plans. The simultaneous consideration of plants' operational strategies makes it even more difficult.

The purpose of this paper is to present an optimization method for such expansion planning of gas turbine cogeneration plants, and to investigate the relationship between optimal expansion planning and energy demand trend. First, the optimal expansion planning problem is defined by introducing several fundamental assumptions. Second, an optimization method is described to solve the problem efficiently. Finally, a numerical study is carried out on a simple-cycle gas turbine cogeneration plant to be installed in a district heating and cooling project.

Multistage Expansion Planning

Definition of Multistage Expansion Planning Problem. A multistage expansion planning problem shown in Fig. 1 is discussed in this paper. Energy demands are assumed to increase stepwise at the beginning of a year and M times with the progress of building construction. Equipment is assumed to be installed simultaneously together with an increase in energy demands. The period between the mth and $(m + 1)$th installations of equipment is called the mth expansion stage, and the number of years of the period is denoted by $T(m)$ $(m = 1, 2, \ldots, M - 1)$. The Mth expansion stage corresponds to the period between the Mth installation of equipment and the evaluation of levelized annual total cost (after L years since first installation of equipment). Thus the number of years of the period between the first and mth installations of equipment $Y(m)$ is expressed by

$$Y(m) = \begin{cases} 0 & (m = 1) \\ \sum_{j=1}^{m-1} T(j) & (m = 2, 3, \ldots, M) \end{cases} \quad (1)$$

The life of equipment is assumed to be the same for all pieces of equipment installed simultaneously and its value for the equipment installed at the beginning of the mth expansion stage is denoted by $\kappa(m)$.

A cash flow diagram is drawn at the lower part of Fig. 1. The initial capital cost, designated by $P(m)$, is invested with the mth installation of equipment. The salvage value of equipment, designated by $V(m)$, is assumed to be left after the life of equipment $\kappa(m)$. The operational cost paid per year during the mth expansion stage, which is denoted by $C(m)$, increases with energy demands.

Economic Evaluation. The levelized annual total cost is adopted as an objective function of the optimal multistage expansion planning problem, and it is evaluated by the annualized costs method (Witte et al., 1988). Here, the balance of cash flow is considered after L years since the first installation of equipment, on the assumption that L is greater than

Contributed by the International Gas Turbine Institute for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received at ASME Headquarters February 25, 1996. Associate Technical Editor: L. A. Riekert.

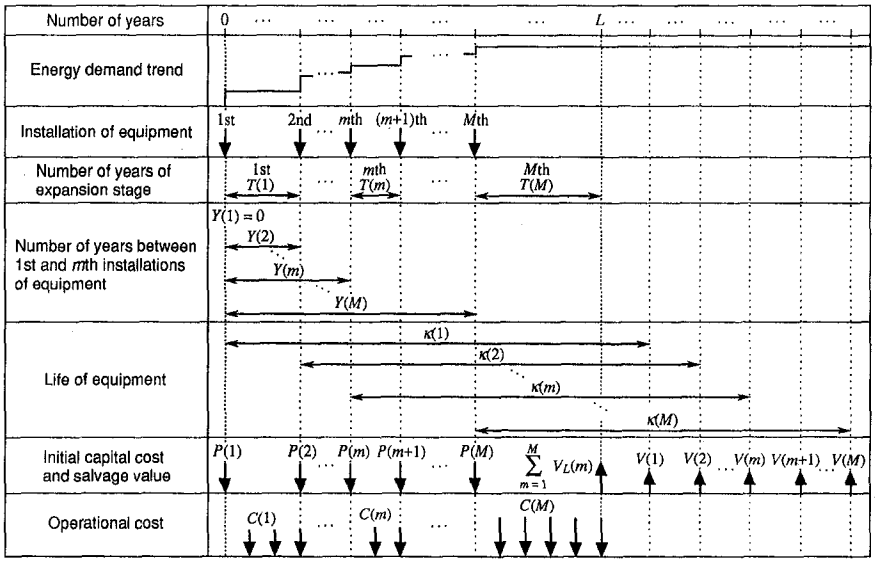


Fig. 1 Multistage expansion planning

$Y(M)$ and not greater than $Y(m) + \kappa(m)$ for any m ($m = 1, 2, \ldots, M$).

First, the salvage value of equipment $V(m)$ is transformed to the value after L years since the first installation of equipment $V_L(m)$ by

$$V_L(m) = \frac{P(m)\{(1+i)^{\kappa(m)} - (1+i)^{L-Y(m)}\}}{(1+i)^{\kappa(m)} - 1} + \frac{V(m)\{(1+i)^{L-Y(m)} - 1\}}{(1+i)^{\kappa(m)} - 1} \quad (m = 1, 2, \ldots, M), \quad (2)$$

where i is the interest rate. Equation (2) is obtained by equalizing the levelized annual capital cost of equipment with an initial capital cost of $P(m)$, a life of $\kappa(m)$, and a salvage value of $V(m)$ to that of equipment with an initial capital cost of $P(m)$, a life of $L - Y(m)$, and a salvage value of $V_L(m)$.

Using the salvage value $V_L(m)$ of Eq. (2), the annual capital cost levelized over L years, C_c, is evaluated by

$$C_c = \sum_{m=1}^{M} \frac{i\{P(m)(1+i)^{L-Y(m)} - V_L(m)\}}{(1+i)^L - 1}. \quad (3)$$

On the other hand, the annual operational cost levelized over L years, C_o, is expressed by

$$C_o = \sum_{m=1}^{M} \sum_{j=1}^{T(m)} \frac{iC(m)(1+i)^{L-Y(m)-j}}{(1+i)^L - 1}. \quad (4)$$

As a result, the annual total cost levelized over L years, C_t, is given by

$$C_t = C_c + C_o. \quad (5)$$

Optimization Method

An optimal sizing method for single-stage planning proposed by the authors is extended to the case of multistage expansion planning (Yokoyama et al., 1994). Equipment capacities and utility maximum demands at each expansion stage are deter-

Nomenclature

C_c = levelized annual capital cost, yen/y
$C(m)$ = operational cost per year at mth expansion stage, yen/y
C_o = levelized annual operational cost, yen/y
C_t = levelized annual total cost, yen/y
E_{buy} = electric power purchased from outside electric power company, MWh/h
E_{GT} = electric power generated by gas turbine generator, MWh/h
$\bar{E}_{GT}$ = capacity (electrical output at rated load status) of gas turbine generator, MW
i = interest rate
L = number of years of period between first installation of equipment and evaluation of levelized annual total cost, y
M = number of expansion stages
m = index for mth expansion stage
$P(m)$ = initial capital cost of equipment installed at mth expansion stage, yen
$T(m)$ = number of years of mth expansion stage, y
$V(m)$ = salvage value of equipment installed at mth expansion stage, yen
$V_L(m)$ = salvage value of equipment installed at mth expansion stage, which is evaluated after L years since first installation of equipment, yen
$Y(m)$ = number of years between first and mth installations of equipment, y
η_e = generating efficiency at rated load status of gas turbine generator
η_h = waste heat recovery efficiency at rated load status of gas turbine generator plus waste heat recovery boiler
$\kappa(m)$ = life of equipment installed at mth expansion stage, y

Equipment Symbols

BA = gas-fired auxiliary boiler
BW = waste heat recovery boiler
EP = equipment for purchasing electricity
GT = gas turbine generator
PC = pump for space cooling
RE = electric compression refrigerator
RS = steam absorption refrigerator

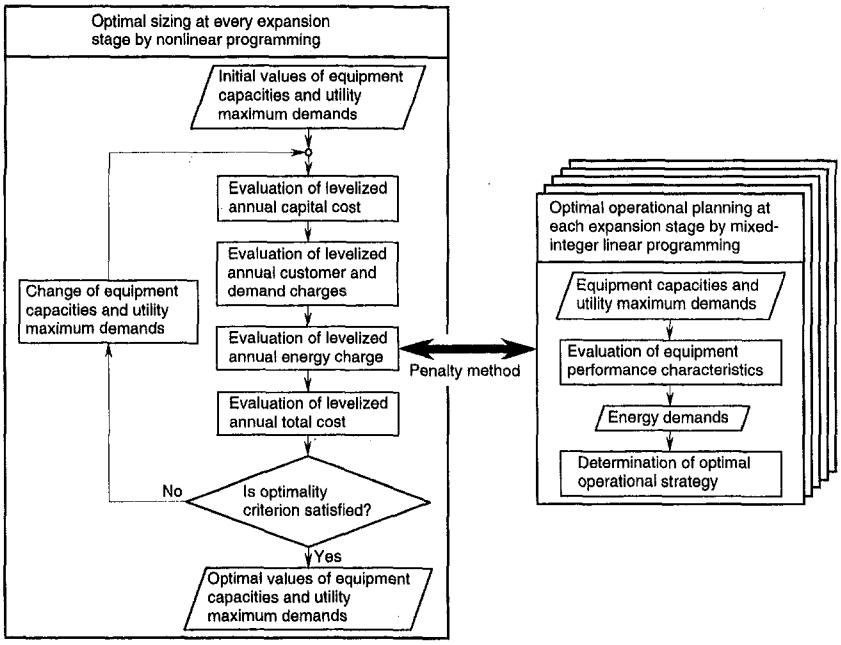


Fig. 2 Hierarchical optimization method of determining plants' sizes and operational strategies

mined in consideration of the plant's operational strategy for increasing energy demands. Only a summary of this optimization method is described here.

Basic Concept of Optimal Planning Method. As stated above, the levelized annual total cost is minimized from the viewpoint of long-term economics. The initial capital cost $P(m)$ and the salvage value $V(m)$ of equipment are considered functions of their capacities. The operational cost $C(m)$ is the sum of the customer, demand, and energy charges of utilities such as purchased electricity and natural gas used for the plant's operation. The demand charge is considered a function of utility maximum demands, and the energy charge is calculated from the plant's operational strategy.

As constraints at each expansion stage, it is necessary to consider performance characteristics of each piece of equipment and energy balance relationships of each energy flow for average hourly energy demands estimated on several representative days in one year. In addition to the average energy demands, peak energy demands in summer and winter must be considered as well. Here, since performance characteristics of equipment change with the capacities, they should be considered as functions of the capacities.

Design variables are composed of equipment capacities and utility maximum demands at each expansion stage in the sizing problem, and the variables expressing the operational strategy at each expansion stage in the operational planning problem. Although the capacity of each piece of equipment is selected from a set of discrete values in the practical design, it is regarded as a continuous variable in this paper. The operational strategy is expressed by the binary and continuous variables that correspond to the on/off status of operation and energy flow rates, respectively (Ito et al., 1990).

Hierarchical Optimization Method. The algorithm of the optimization method is shown in Fig. 2. The whole is divided into two parts related hierarchically to each other by a penalty method. In the optimal sizing problem at the upper level, optimal values of equipment capacities and utility maximum demands at every expansion stage are searched simultaneously so as to minimize the levelized annual total cost. In the optimal operational planning problem at the lower level, for a plant with the equipment capacities and utility maximum demands given at each searching step, the operational strategy is assessed individually at each expansion stage and the levelized annual energy charge is evaluated. Here, the operational strategy is assessed using an optimal operational planning method proposed by the authors (Ito et al., 1990). This strategy minimizes the levelized annual energy charge, which is consistent with minimization of the levelized annual total cost.

The optimal sizing problem is formulated as a nonlinear programming problem, and it is solved by the sequential linear programming method (Cheney and Goldstein, 1959). The optimal operational planning problem is formulated as a mixed-integer linear programming problem, and it is solved by the branch and bound method (Garfinkel and Nemhauser, 1972).

Table 1 Total floor areas and energy demands of buildings

Type of building	Symbol	Total floor area m^2	Energy demand					
			Hourly maximum value			Annual total value		
			Electricity MWh/h	Space cooling MWh/h	Steam MWh/h	Electricity GWh/y	Space cooling GWh/y	Steam GWh/y
Office building	A_1	66.5×10^3	3.77	6.81	5.07	12.56	5.99	4.33
	A_2	108.9×10^3	6.17	11.15	7.66	20.56	9.82	7.08
	A_3	12.8×10^3	0.72	1.31	0.98	2.43	1.16	0.84
	A_4	80.6×10^3	4.57	8.25	6.14	15.23	7.26	5.25
Hotel	H_1	72.6×10^3	3.25	4.34	5.93	16.78	5.06	10.75
	H_2	42.3×10^3	1.90	2.53	3.46	9.79	2.95	6.26
Total	—	383.7×10^3	20.36	31.57	24.84	77.35	32.24	34.51

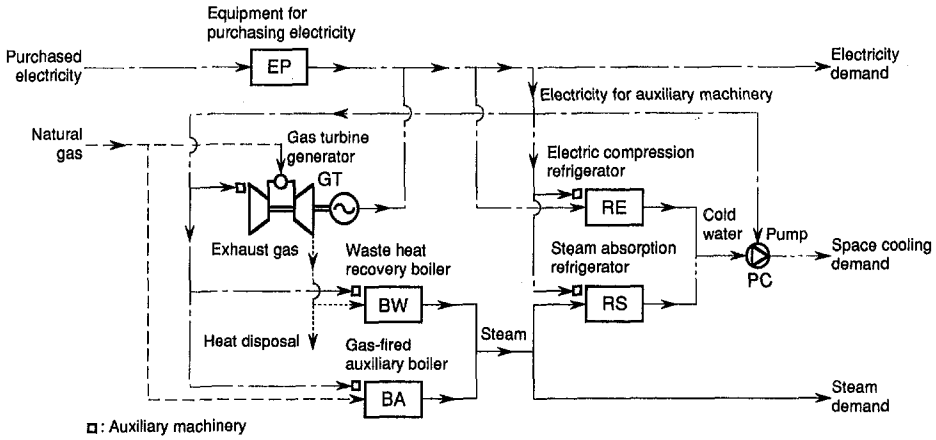


Fig. 3 Structure of simple-cycle gas turbine cogeneration plant

The concrete algorithm adopted here is based on the Land–Doig method (Land and Doig, 1960).

Numerical Study

A numerical study is carried out on the multistage expansion planning of a simple-cycle gas turbine cogeneration plant to be installed in a district heating and cooling project.

District Heating and Cooling Project. The district studied here is a area to be renewed in the near future. Four office buildings, A_1 to A_4, and two hotels, H_1 and H_2, are to be constructed stepwise over a period of several years. Table 1 shows the total floor area of each building and the hourly maximum and annual total values of energy demands estimated. Here, one representative day is considered for each month; i.e., the operational strategy is to be determined on 12 representative days at each expansion stage. Hourly energy demands are given as input data for each representative day. In addition to these average energy demands, peak energy demands in summer and winter are also given by multiplying the average energy demands on the representative day in August and January by 1.2, respectively.

Plant Structure. Figure 3 shows schematically the structure of a simple-cycle gas turbine cogeneration plant investigated in this study. Although only one unit is illustrated for each kind of equipment, there may be several units. Solid lines, dot-dash lines, two dots-dash lines, dotted lines, and broken lines denote the flows of steam, electricity, cold water, exhaust gas, and natural gas, respectively.

Electricity is supplied to users by operating gas turbine generators and by purchasing electricity from an outside electric power company. Electricity is used to drive electric compression refrigerators, pumps, and other auxiliary machinery in the plant. Exhaust heat generated by the gas turbines is recovered by waste heat recovery boilers, and it is reused for several kinds of thermal energy in a cascade fashion. Surplus exhaust heat is disposed of through exhaust gas dampers. Shortage of steam is supplemented by gas-fired auxiliary boilers. Cold water for space cooling is supplied by electric compression and steam absorption refrigerators. Steam is used for space heating and other miscellaneous purposes.

The capacities of each kind of equipment except the pumps are determined optimally at each expansion stage, together with the maximum demands of purchased electricity and natural gas. It is assumed that the capacity of equipment for purchasing electricity is equal to the maximum demand of purchased electricity. The size of the gas turbine generator will determine the capacity of the waste heat recovery boiler.

Input Data. Performance characteristics and initial capital costs of equipment are given as functions of their capacities. Nonlinear curve fitting is adopted for the performance characteristics of gas turbine generators plus waste heat recovery boilers. For example, the generating efficiency at the rated load η_e tends to increase with the capacity, and conversely the waste heat recovery efficiency at the rated load η_h tends to decrease. These values can be approximated by

$$\eta_e = 0.211 \bar{E}_{GT}^{0.15} \tag{6}$$

and

$$\eta_h = 0.555 \bar{E}_{GT}^{-0.04}, \tag{7}$$

where $\bar{E}_{GT}$ is the capacity (electrical output at the rated load status) of a gas turbine generator. Variations in these efficiencies at part-load status are also taken into account. Linear curve fitting is adopted for performance characteristics of other kinds of equipment. Linear functions are used to express the relationship between initial capital cost and capacity for all kinds of equipment. The representative values of performance characteristics and initial capital costs are given in Table 2.

According to a numerical study on the single-stage planning carried out by the authors, gas turbine generators of a large capacity were selected optimally from the viewpoint of long-term economics (Yokoyama et al., 1994). However, this is not satisfactory from the energy-saving viewpoint because a large amount of exhaust heat must be disposed of. In this paper, the numerical study is carried out by assuming that exhaust heat less than 50 percent of the capacity of gas turbines is permitted to be disposed of.

The rates for purchased electricity and natural gas adopted in this study are standard ones used in Japan, and they are given in Table 3. In evaluating the levelized annual total cost, it is assumed that the life of equipment $\kappa(m) = 15$ y and the salvage value of equipment $V(m) = 0$ for any m ($m = 1, 2, \ldots, M$), and that the interest rate $i = 0.1$. The levelized annual total cost is evaluated after 15 years since the first installation of equipment; i.e., $L = 15$ y.

Table 2 Performance characteristic values and initial capital costs of equipment

Equipment	Performance characteristic value		Initial capital cost
Gas turbine generator plus waste heat recovery boiler	Generating efficiency Waste heat recovery efficiency	Eq. (6) Eq. (7)	220.0×10^6 yen/MW
Gas-fired auxiliary boiler	Thermal efficiency	0.90	6.6×10^6 yen/MW
Electric compression refrigerator	COP	3.60	21.5×10^6 yen/MW
Steam absorption refrigerator	COP	1.20	51.6×10^6 yen/MW
Equipment for purchasing electricity	—		56.3×10^6 yen/MW
Pump	Electricity consumption rate	0.025 MW/MW	—

Table 3 Rates for purchased electricity and natural gas

Utility	Unit cost		
	Customer charge	Demand charge	Energy charge
Purchased electricity	—	1 740 yen/(kW·month)	10.77 (Jul.–Sep.) yen/kWh 9.79 (Other months)
Natural gas	36 000 yen/month	1 200 (Apr.–Oct.) yen/(m³/h·month) 3 200 (Other months)	30.88 yen/m³

Table 4 Scenarios of building construction

Case			1	2	3	4	5	6
1st expansion stage	Number of years $T(1)$		3	3	3	9	9	9
	Buildings constructed	A_1	○			○		
		A_2		○	○		○	○
		A_3		○	○		○	○
		A_4			○			○
		H_1	○	○		○	○	
		H_2			○			○
2nd expansion stage	Number of years $T(2)$		12	12	12	6	6	6
	Buildings constructed	A_1		○	○		○	○
		A_2	○			○		
		A_3	○			○		
		A_4	○	○		○	○	
		H_1			○			○
		H_2	○	○		○	○	

Scenarios of Building Construction. Optimization calculation is carried out for several scenarios of building construction, and the relationship is investigated between optimal expansion planning and energy demand trend. These scenarios are given in Table 4. Two-stage expansion planning ($M = 2$) is investigated by changing the ratio of energy demands at the first expansion stage to those at the second one as well as the years of expansion stages. In cases 1–3, the period of the first expansion stage is shorter than that of the second one. In cases 4–6, the period of the former is longer than that of the latter. The ratio of energy demands at the first expansion stage to those at the second one increases with the case number for each group.

It is assumed that two units of gas turbine generator plus waste heat recovery boiler are installed at each expansion stage, and that one unit is installed only at the first expansion stage concerning the gas-fired auxiliary boiler and electric compression refrigerator. It is also assumed that one unit is installed at each expansion stage concerning the other kinds of equipment.

Results and Discussion. The results obtained by the optimization calculation are summarized in Table 5, Figs. 4 and 5. Table 5 gives the optimal values of equipment capacities per unit and utility maximum demands at each expansion stage. Figures 4(a)–(c) show the load duration curves for electricity supply at each expansion stage for cases 1–3, respectively. In these figures, E_{GT} and E_{buy} denote the electric power generated by a gas turbine generator and purchased from an outside electric power company, respectively. The subscripts 1 and 2 show the expansion stages when gas turbine generators are installed. The subscripts A and B are the indices for two gas turbine generators installed simultaneously. Figure 5 shows the levelized annual total cost and its items for each case. Here, optimal sizes and optimal operational strategies are discussed only for gas turbine generators.

According to the results shown in Table 5, a correlation can be found between the amount of energy demands increased at each expansion stage and the optimal values of equipment capacities and utility maximum demands at the corresponding stage. For example, the capacity of gas turbine generators tends to increase with the amount of energy demands. However, the correlation is not simple quantitatively. The ratio of the capacity of gas turbine generators to the hourly maximum value of electricity demand changes with cases.

In cases 1–3, the period of the first expansion stage is shorter than that of the second one, and consequently the levelized annual operational cost at the second stage dominates that at the first one as shown in Fig. 5. Thus the amount of energy demand at the second stage has a significant effect on the sizing at the first stage. In case 1, since the possibility of installing gas turbine generators of a large capacity at the second stage remains, gas turbine generators of a relatively small capacity are installed at the first stage. In case 3, gas turbine generators of a relatively large capacity are installed at the first stage in spite of their low availability, in consideration of a small increase in energy demands at the second stage. These are due to the feature that the generating efficiency of gas turbine generators tends to increase with capacity as stated above.

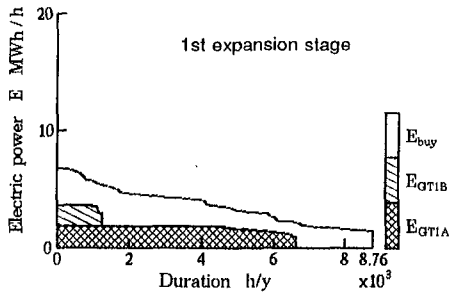
The above-described tendency is relaxed in cases 4 and 6. Namely, the capacity of gas turbine generators installed at the first stage in case 4 is larger than that in case 1, and conversely the capacity in case 6 is smaller than that in case 3. This is because the period of the first stage is longer than that of the second one, and the amount of energy demands at the second stage does not have a significant effect on the sizing at the first stage.

Concerning the operational strategy shown in Fig. 4, it can be seen in case 1 that the operating time of both gas turbine generators installed at the first stage becomes shorter at the second stage, and that one of them is operated only for peak periods of energy demands. In case 2, however, no remarkable change occurs in the operating time. In case 3, the operating time becomes longer. This effect is caused by the difference in capacity.

In cases 1–3, the operational strategy of gas turbine generators differs from the one following electricity demand. The reason for this strategy is as follows: The strategy determined by minimizing the energy charge depends significantly on the prices of electricity and natural gas purchased as well as the characteristics of electricity and thermal energy demands, as analyzed previously (Ito et al., 1990). Under the prices and demand characteristics in this study, it is advantageous economically that gas turbine generators be stopped or operated following thermal energy demand in the case of low thermal energy

Table 5 Optimal values of equipment capacities per unit and utility maximum demands

Case			1	2	3	4	5	6
1st expansion stage	Gas turbine generator	MW	1.79	2.40	5.42	2.09	2.53	4.46
	Waste heat recovery boiler	MW	4.21	5.34	10.34	4.78	5.58	8.83
	Gas-fired auxiliary boiler	MW	0.92	3.16	3.04	0.49	2.01	1.73
	Electric compression refrigerator	MW	0.86	1.35	3.57	1.17	0.83	3.47
	Steam absorption refrigerator	MW	8.77	14.36	21.67	8.19	14.78	21.71
	Equipment for purchasing electricity	MW	4.59	7.07	5.20	4.19	6.72	6.98
2nd expansion stage	Gas turbine generator	MW	4.50	3.01	0.15	4.26	3.34	1.06
	Waste heat recovery boiler	MW	8.89	6.42	0.57	8.50	6.98	2.76
	Gas-fired auxiliary boiler	MW	—	—	—	—	—	—
	Electric compression refrigerator	MW	—	—	—	—	—	—
	Steam absorption refrigerator	MW	21.26	15.89	6.34	22.45	15.96	6.38
	Equipment for purchasing electricity	MW	6.84	6.14	8.32	7.17	5.52	6.60



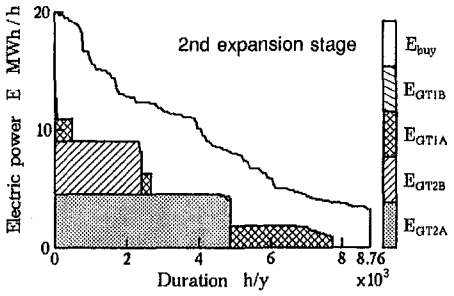
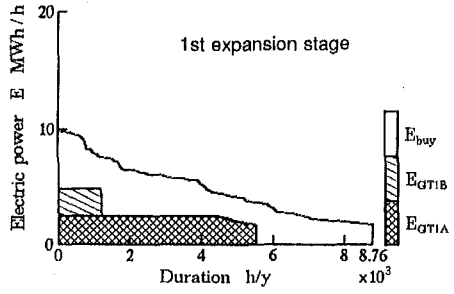


Fig. 4(*a*) Case 1



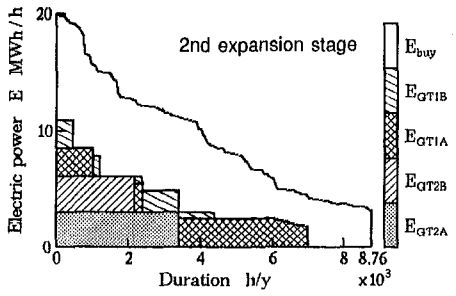
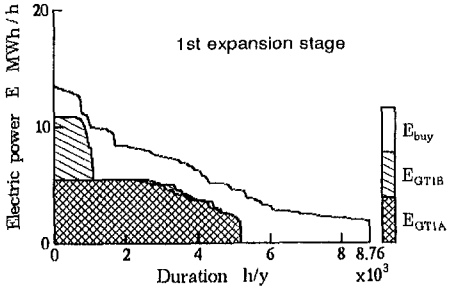


Fig. 4(*b*) Case 2



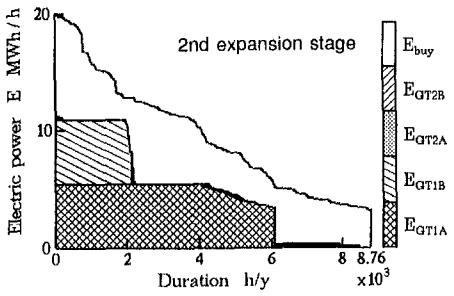


Fig. 4(*c*) Case 3

Fig. 4 Load duration curves for electricity supply

demand during the nighttime. As an alternative plan, it is possible to reduce the capacity of gas turbine generators and to improve their load factors. However, this plan decreases the power-generating efficiency and increases the exhaust heat recovery efficiency, which tends to increase the mismatch between supply and demand characteristics of electricity and thermal energy. The high price of electricity used in the study is also a reason for a large capacity and low load factors of gas turbine generators. The results obtained here are similar to those obtained previously for single-stage planning (Yokoyama et al., 1994).

Conclusions

An optimal multistage expansion planning problem has been considered concerning a gas turbine cogeneration plant for district heating and cooling. An optimization method has been proposed to solve the problem efficiently. Optimal values of equipment capacities and utility maximum demands at each expansion stage have been determined so as to minimize the levelized annual total cost subject to increasing energy demands. A numerical study has been carried out on a simple-cycle gas turbine cogeneration plant to be installed in a district renewal project. The study shows that equipment capacities and utility maximum demands can be determined rationally and easily by the proposed method, and the relationship between optimal expansion planning and energy demand trend has been made transparent by investigating several scenarios of building construction.

Although a numerical study has been carried out on a restricted number of cases for scenarios of building construction, this method is applicable to other cases.

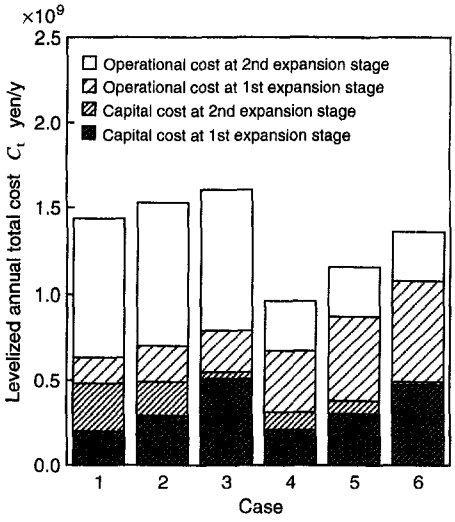


Fig. 5 Levelized annual total cost and its items

References

Cheney, E. W., and Goldstein, A. A., 1959, "Newton's Method for Convex Programming on Tchebycheff Approximation," *Numerical Mathematics,* Vol. 1, pp. 253–268.

Garfinkel, R. S., and Nemhauser, G. L., 1972, *Integer Programming,* Wiley, New York, pp. 5–7.

Ito, K., Yokoyama, R., Horii, S., Matsumoto, Y., and Akagi, S., 1987, "An Optimal Operational Planning Model for Gas Turbine Cogeneration Plants," *Proc. 1987 Tokyo International Gas Turbine Congress,* Vol. I, pp. 97–104.

Ito, K., Yokoyama, R., Yamaguchi, T., Matsumoto, Y., and Akagi, S., 1988, "Optimal Operational Planning of a Gas Turbine Combined Heat and Power Plant Based on the Mixed-Integer Programming," *Proc. IFAC International Symposium on Power Systems Modelling and Control Applications,* pp. 13.4.1–13.4.7.

Ito, K., Yokoyama, R., Akagi, S., and Matsumoto, Y., 1990, "Influence of Fuel Cost on the Operation of a Gas Turbine-Waste Heat Boiler Cogeneration Plant," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 112, pp. 122–128.

Ito, K., Yokoyama, R., and Matsumoto, Y., 1995, "Optimal Operation of Cogeneration Plants With Steam-Injected Gas Turbines," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 117, pp. 60–66.

Land, A. H., and Doig, A. G., 1960, "An Automatic Method of Solving Discrete Programming Problems," *Econometrica,* Vol. 28, pp. 497–520.

Witte, L. C., Schmidt, P. S., and Brown, D. R., 1988, *Industrial Energy Management and Utilization,* Hemisphere Publishing, Washington, pp. 68–69.

Yokoyama, R., and Ito, K., 1991, "An Evaluation of Gas Turbine Cogeneration Plants for a District Energy Supply Center," *Proc. 1991 Yokohama International Gas Turbine Congress,* Vol. I, pp. 121–128.

Yokoyama, R., Ito, K., and Matsumoto, Y., 1994, "Optimal Sizing of a Gas Turbine Cogeneration Plant in Consideration of Its Operational Strategy," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 116, pp. 32–38.

Theoretical and Experimental Comparisons for Damping Coefficients of a Short-Length Open-End Squeeze Film Damper

L. A. San Andres
Associate Professor,
Mechanical Engineering Department,
Texas A&M University,
College Station, TX 77843

Squeeze film dampers (SFD) provide load isolation and attenuate rotor vibrations in high speed turbomachinery. Operating parameters such as whirl frequency, amplitude of journal motion, and value of external pressure supply determine the SFD dynamic force response and its dissipation of mechanical energy. Measurements of pressure fields and fluid film forces in a fully submerged open-end squeeze film damper are presented for tests with rotor speeds to 5000 cpm and low supply pressures. The damper has a clearance of 381 μm (0.015 in.) and the journal describes circular centered orbits of amplitudes ranging from 30 to 50 percent of the bearing clearance. Experimental film pressures depict a vapor cavitation (close to zero absolute pressure) zone increasing in extent as the whirl frequency increases. Estimated fluid film forces from the measured pressure profiles are found to be proportional to whirl speed and lubricant viscosity. Test cross-coupled damping coefficients (C_{rr}) are smaller than predicted values based on the short-length bearing model with a π film cavitation assumption. The direct damping coefficients (C_{tt}) are larger than theoretical values, especially at low frequencies where the dynamic cavitation region has not grown to half the circumferential flow extent. The experiments demonstrate the viscous character of the fluid film forces in a SFD test apparatus where fluid inertia effects are minimal (squeeze film Reynolds number less than one). On the other hand, the extent of the cavitation zone appears to be dominant on the generation of fluid film forces.

Introduction

Light-weight, high-performance engines exhibit a trend toward increased flexibility, leading to a high sensitivity to unbalance with large vibration levels and reduced reliability. Squeeze film dampers (SFDs) are essential components of high-speed turbomachinery and are used in aircraft jet engines to provide viscous damping to rolling element bearings, which otherwise would contribute no appreciable amount of damping to the rotor-bearing system. One other important application is related to high-performance compressor units where dampers are installed in series with tilting pad bearings to reduce bearing support stiffness while providing additional damping as a safety mechanism to avoid rotordynamic instabilities. Despite the many benefits a squeeze film damper could offer, experience has demonstrated that the force response of SFDs is complex and difficult to predict analytically for actual hardware applications. Engineering design of SFDs is based on overly simplified theoretical models, which either fail to incorporate or simply neglect unique features (structural and fluidic) affecting the damper dynamic force performance. Squeeze film dampers in certain regimes of operation can also bring strong nonlinearities to the performance of rotor-bearing systems. The optimum damping required by the rotor-bearing system may conflict with the damping available in a squeeze film damper. If damping is too large the SFD acts as a rigid constraint to the rotor-bearing system with large forces transmitted to the supporting structure. If damping is too light, the damper is ineffective and likely to permit large amplitude vibratory motion with possible subharmonic resonances.

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 13, 1995. Paper No. 95-GT-98. Associate Technical Editor: C. J. Russo.

Squeeze film dampers in practice operate with levels of external pressurization not large enough to prevent *dynamic cavitation* of the lubricant in the thin film lands. Cavitation in squeeze film flows is a broad term, which means either liquid vaporization, or release of gaseous media, or allowance of air entrainment, or formation of bubbly mixtures, or a combination of all these conditions. Dynamic cavitation is the single factor most affecting the performance of a squeeze film damper (Feng and Hahn, 1986, 1987). This complex phenomenon is yet little understood, although some recent advances point to the mechanics of bubble evolution and two-phase flow considerations (Sun and Brewe, 1992; Sun et al., 1993; Al-Sharif et al., 1993). By extension from hydrodynamic journal bearing theory, current cavitation models for squeeze film dampers are based on Elrods' algorithm (1981). However, this model is of little practical value when considering the complexity of the observed pressures and flow in bearings undergoing dynamic loading or pure squeeze film flow. Hibner and Bansal (1979) noted the effect of liquid compressibility and lubricant cavitation on a small clearance damper ($c/R = 0.0032$). Walton et al. (1987) found that fluid cavitation onset and extent depend on the damper operating parameters like whirl frequency, journal orbit size and offset position, level of supply pressure and end seal restrictions. Zeidan and Vance (1989a, b, 1990) experimentally identified five regimes of dynamic cavitation depending on the damper configuration tested and the operating conditions. The experiments showed persistent gas bubbles in the fluid film and leading eventually to a well-mixed two-phase fluid at large whirl frequencies. The significance of these experimental investigations cannot be overlooked since they have already prompted significant changes in current damper design. On the other hand, Zhao et al. (1994) have recently suggested that large negative absolute pressures (fluid in tension) are likely to occur in SFDs. These observations are not supported by exhaustive pressure

measurements on a SFD, but try to justify poor correlation between the test data and numerical predictions for a rotor supported on short-length SFDs.

In many practical circumstances, dampers are designed with a feeding central groove to insure a continuous flow of lubricant through the squeeze film lands. The groove, usually of large volume, is conceived as a flow source with constant pressure around the journal surface. In practice, however, grooved dampers have shown more damping than theoretical predictions (Holmes and Sykes, 1990). Zeidan and Vance (1989b) reported dynamic pressures at a groove feeding a squeeze film land. Roberts et al. (1986), Roberts and Ellis (1990), Ramli et al. (1987), and Zhang et al. (1994) report experimental force coefficients an order of magnitude larger than predictions from the short bearing model for all dynamic eccentricities tested. Systematic experiments and analysis by Arauz and San Andres (1994) have shown experimentally that dampers with groove depths up to ten times the radial clearance can generate more damping forces than predictions based on a classical lubrication model. Feeding grooves then do not isolate the adjacent film lands, but rather interact with the squeeze film regions and are capable of producing an appreciable dynamic pressure, which adds to the damper force response.

Large-clearance, light-viscosity squeeze film dampers operating at high frequencies are affected by fluid inertia forces, which could introduce significant added masses and dramatically change the dynamic response of a rotor–SFD system (San Andres and Vance, 1988). The relevance of fluid inertia is related to the squeeze film Reynolds number ($\mathrm{Re}_s = \rho\omega c^2/\mu$), which ranges from 1 to 50 in many applications. It is worth noting that some modern damper applications have film radial clearances as large as 0.508 mm (0.020 in.) radial to allow large unbalance levels and withstand high maneuvering loads (Kocur, 1992). Ramli et al. (1987), Roberts et al. (1986), Roberts and Ellis (1990), Kinsale and Tichy (1989), and San Andres et al. (1993) provide experimental evidence of fluid inertia forces in test damper configurations. Correlations of experimental measurements with analytical predictions have ranged from poor to adequate. In general, measured results seem to be highly dependent on actual operating conditions including the level of external pressurization, test hardware configuration, coupling to the dynamics of the structural system, and most importantly on the early onset or avoidance of fluid dynamic cavitation in one of its many forms.

The need to carry out planned experiments to validate theoretical analysis and to enhance our understanding of the flow mechanics in squeeze film dampers is always present. This paper reports experimental measurements of absolute pressures in a short-length ($L/D = 0.24$), fully *submerged* open-end squeeze film damper with a clearance ratio of $c/R = 0.0059$. The damper journal describes circular centered orbits with orbit radii ranging from 30 to 50 percent of the radial clearance. The experiments were carried out with a low level of external pressurization (approximately 35 kPa), enough to warrant a continuous flow of lubricant across the damper. Controlled tests were performed for a range of whirl frequencies to 83 Hz and lubricant inlet temperatures. The major objectives were to determine the type and extent of the dynamic cavitation region for the different operating conditions, and to quantify its effect on the damper fluid film forces. The tests detail absolute pressure measurements and verify that the lubricant used is unable to sustain negative pressures (fluid in tension). The present test results differ greatly from previously published experiments for the same damper but with a larger radial clearance (1.55 mm), and where fluid inertia brought a large added mass effect into the SFD dynamic force response (San Andres et al., 1993). The tests performed with a small clearance show squeeze film Reynolds numbers (Re_s) less than unity, and this warrants that fluid inertia effects have a negligible influence on the measured pressure field and fluid film forces.

Theoretical Model

Most dampers in practice are of short axial length ($L/D \leq 0.25$) and accommodate some type of end seals to increase their damping capability. Additional features of this mechanical element may include high resistance orifices for pressure delivery and discharge, and deep grooves acting as flow sources or sinks of uniform pressure. The literature provides a wealth of numerical and theoretical analysis for the prediction of the fluid film pressure generated under the squeeze film action and computation of the dynamic forces (see, for example, Marmol and Vance, 1978; Szeri et al., 1983; Kinsale and Tichy, 1989).

The short-length open-end SFD model is amenable to provide a closed-form solution to the flow equations including fluid inertia and accommodating a rather simplistic cavitation model. For circular centered journal motions of amplitude (e) and whirl frequency (ω), and in the absence of geometric distortions in the circumferential direction, a damper produces a constant fluid film force in a reference frame rotating with angular frequency ω (see Fig. 1). The radial and tangential components of the fluid film force are equal to:

$$F_r = -C_{rr}V_t, \quad F_t = -C_{tt}V_t, \tag{1}$$

where $V_t = e\omega$ is the tangential speed of the whirling journal center, and (C_{rr}, C_{tt}) are known as the cross and direct damping force coefficients. Note that (C_{tt}, C_{rr}) are not strictly rotordynamic coefficients, as their classical definition implies small-amplitude motions (perturbations) about a journal equilibrium position. The formula given in Eq. (1) does not account for fluid inertia effects as these are dominant for large squeeze film Reynolds number applications ($\mathrm{Re}_s = \rho\omega c^2/\mu > 10$). The interested reader may find the work of Tichy and Modest (1978) and San Andres and Vance (1987) to be of relevance for a

Nomenclature

c = damper radial clearance = 381 μm
C_{tt}, C_{rt} = direct and cross damping coefficients, N · s/m
D = damper journal diameter = 0.129 m
e = journal center orbit radius = 190 μm typically
$F_r = -C_{rr}V_t$ = radial force, N
$F_t = -C_{tt}V_t$ = tangential force, N
L = damper journal axial length = 0.0307 m
n = number of cycles of test data, typically 4 to 6
P = fluid pressure, Pa
$\mathrm{Re}_s = \rho\omega c^2/\mu$ = squeeze film Reynolds number
T = fluid temperature, °C
$T_\omega = 2\pi/\omega$ = period of orbital motion, s
$V_t = e \cdot \omega$ = journal center tangential velocity, m/s
z = axial coordinate, m
Z = axial location of measurement, m
$\epsilon = e/c$ = dimensionless orbit radius = 0.3–0.5
θ = circumferential coordinate
μ = fluid viscosity, Pa-s
ρ = fluid density, kg/m^3
ω = damper whirl frequency, rad/s

Subscripts

r, t = radial and tangential coordinate system
1, 2 = axial locations of measurement of pressure and temperature
3 = location of measurement of journal motion
π = related to π film—short-length SFD model

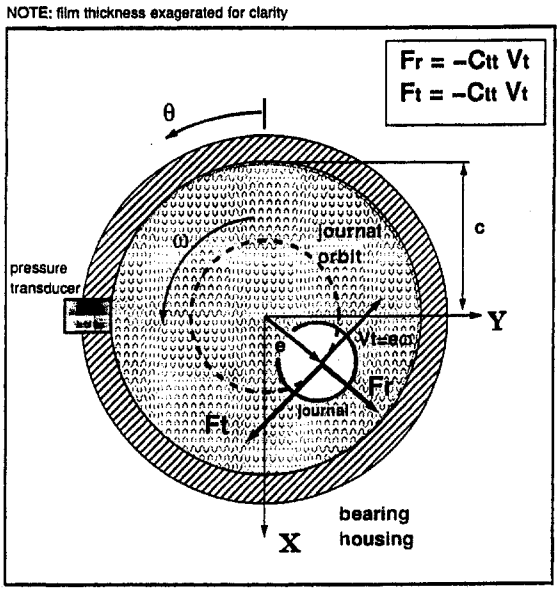
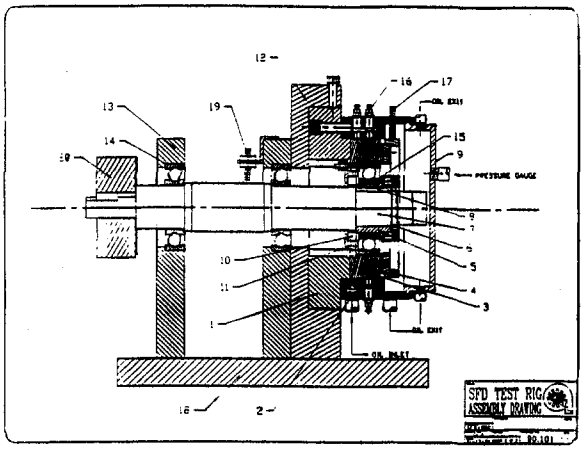


Fig. 1 Schematic view of journal center motion and fluid film forces for circular centered orbits



REFERENCES

1.- HOUSING END PLATE
2.- HOUSING
3.- JOURNAL
4.- END PLATE
5.- ECCENTRIC
6.- ECCENTRIC HOLDER
7.- SHAFT
8.- KEY
9.- HOUSING CAP
10.- LIP SEAL
11.- LEFT SEAL PLATE
12.- HOUSING SUPPORT
13.- SHAFT SUPPORT
14.- SHAFT BEARINGS
15.- JOURNAL BEARING
16.- PRESSURE NUTS
17.- PROXIMITY PROBE
18.- BASE PLATE
19.- OPTICAL SPEED SENSOR
20.- TRANSMISSION WHEEL

Fig. 2 TRC low clearance squeeze film damper test rig

detailed exposition on fluid inertia forces in squeeze film dampers.

The damping coefficients for the simple π film cavitation model applicable to the test configuration described below are (Vance, 1986):

$$C_{tt(\pi)} = \frac{\pi\mu DL^3}{(1-\epsilon^2)^{1.5}}; \quad C_{rr(\pi)} = \frac{4\mu\epsilon DL^3}{(1-\epsilon^2)^2} \qquad (2)$$

For the full film model (no cavitation), the direct damping coefficient is twice the value given above and there is no cross-coupled damping, i.e., $C_{rt} = 0$. The formulae given assume the fluid to be isoviscous and incompressible undergoing an isothermal flow process within a damper journal and bearing completely submerged in a lubricant bath to prevent air entrainment (Pan, 1980).

Experimental Facility and Procedure

Figure 2 depicts the squeeze film damper section of the test apparatus. The damper journal outer diameter (D) is equal to 129.41 mm (5.095 in.) and its axial length (L) is equal to 30.73 mm (1.21 in.). The nominal radial clearance (C) is 0.381 mm (0.015 in.). Thus, the damper length-to-diameter (L/D) and clearance-to-radius (C/R) ratios are 0.237 and 0.0059, respectively. San Andres et al. (1993), and Arauz and San Andres (1994) describe in detail the test apparatus and experimental procedure. The damper left end is sealed with an O-ring located on the axial face of the journal and in contact with the damper housing. This O-ring is preloaded on assembly of the test journal and dynamic tests with no lubricant and pressurized air demonstrate that this seal does not allow any measurable leakage.

In the experiments, the supply pressure is kept constant at a low value of 35 kPa (5 psig) above ambient pressure. The lubricant flows from a closed reservoir, through the SFD test section and into a fully flooded plenum. A small secondary pump removes the oil from the discharge plenum kept at ambient pressure. The flow rate across the damper varied between 0.10 to 0.25 lt/min as temperature in the squeeze film damper increased. The lubrication circulation system and test rig configuration do not allow for air entrainment into the squeeze film lands.

The journal motion is constrained mechanically to a centered circular orbit with an eccentric sleeve of radius (e) equal to 0.19 mm (0.0075 in). The nominal dimensionless orbit radius ($\epsilon = e/c$) is then equal to 0.50. Figure 3 shows a schematic view of the damper configuration tested, denoting the relevant dimensions and location of instrumentation sensors. The journal orbital motion is measured by means of two orthogonally positioned displacement sensors (PP) at Z_3. Five piezoelectric pressure transducers (PT) are used to measure the dynamic pressure generated in the fluid film. Four transducers are located 90 deg apart at axial location Z_2 (16.7 mm from the left end). These transducers are used to verify the synchronicity of the dynamic pressure signal for circular centered orbits. The remaining pressure transducer is located at Z_1 (5.6 mm from the left end). An additional strain-gage pressure transducer is installed at location Z_2, and 60 deg from the top. This last transducer measures absolute pressure and serves as a reference to verify the actual magnitude of measured pressures at the other (Z_2) locations.

For a rotationally symmetric damper and circular centered orbits, the time variations in circumferential pressure measured

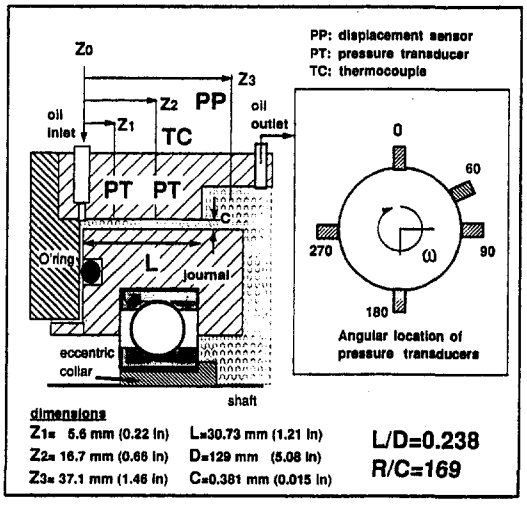


Fig. 3 Schematic axial view of damper test section and location of instrumentation

at any location on the damper annulus are identical for any one cycle of journal orbiting. In the experimental procedure a long and tedious process of calibration insures the orbit to be centered; and it is verified that pressure signals measured at four circumferential locations 90 deg apart are identical but for a phase shift corresponding to the relative position of the sensors on the circumference of the damper. Fluid film forces are estimated by integration of the pressure field acting on the journal surface. For circular centered orbits well-known kinematic relationships allow the evaluation of film forces from measured dynamic pressures at a fixed circumferential location (θ'). The numerical integration is performed on the circumferential direction, and the axial variation is approximated with a simple arithmetic average from the forces obtained at the two axial locations of measurement. The formula used to evaluate the forces numerically is:

$$\begin{bmatrix} F_r \\ F_t \end{bmatrix} = \int_0^{2\pi} \int^{L} P_i(\theta) \begin{pmatrix} \cos\theta \\ \sin\theta \end{pmatrix} dz R d\theta = -\frac{\omega}{n} \sum_i^2 \left[\int_0^{T_\omega \cdot n} P_i(\theta' - \omega t) \begin{pmatrix} \cos(\theta' - \omega t) \\ \sin(\theta' - \omega t) \end{pmatrix} dt \right] R \Delta z_i \quad (3)$$

where T_ω corresponds to the period of motion at whirl frequency ω, and n denotes the number of full cycles of data taken at each test condition, typically four to six. Each test measures the journal motion and pressure fields with at least 70 data points per orbital motion. The error introduced by the numerical integration of the pressure signals is about 13.8 percent since measurements are taken only at two axial locations along the damper. The numerical uncertainty is thus larger than the uncertainty of the pressure transducers and displacement probes, which are 2 and 1 percent of the full scale, respectively.

Experimental Results and Discussion

Measurements of the pressure field and estimation of the fluid film forces at fixed whirl frequencies (20, 33, 50, 67, 83 Hz) while the inlet oil temperature increased from 30°C (87°F) to 38°C (100°F) were performed first. Then, tests with decreasing whirl frequencies from 83 Hz to 20 Hz at a constant inlet fluid temperature were subsequently made. For these last tests, the calculated squeeze film Reynolds number (Re_s) based on the experimental parameters (fluid viscosity, clearance, and whirl frequency) are rather small and varied from 0.2 to 0.80. It was then anticipated that the effect of fluid inertia on the damper pressure field and force response would be negligible.

The lubricant used in the experiments is SAE 10 engine oil. Oil viscosity and density values at different temperatures and at ambient pressure are measured prior to the SFD tests. The fluid viscosity used for presentation of the SFD experimental measurements corresponds to that determined from the average of temperatures measured at the damper inlet and discharge ports and at one location within the squeeze film lands (Z_2). Preliminary dynamic tests without lubricant on the damper film lands confirmed that the journal orbit was circular of radius equal to 0.19 mm ($\epsilon = 0.50$). However, in all subsequent experiments with oil, the measured journal orbit size decreased proportionately with both whirl frequency and oil viscosity as shown in Fig. 4. The orbit radii noted on the figure *key* correspond to average values at the rated test frequency. The measured orbit radius decreases due to the large SFD forces generated and acting to center the rotor by producing a deflection in the elastic shaft.

Figures 5 present the measured pressure profiles at axial locations Z_1 and Z_2 versus time for increasing values of the whirl frequency. Note that for each test frequency the actual orbit radius (ϵ) and fluid temperature (and viscosity) are different while the level of absolute supply pressure corresponds to 138 kPa (20 psia). Figure 5(b) also depicts the qualitative variation of the film thickness at the location of measurement. For test times ($\omega \cdot t$) $\leq \pi$, the gap is increasing and the film pressures drop, while for times ($\omega \cdot t$) $\geq \pi$, the film gap is decreasing (positive squeeze action) and generating squeeze film pressures well above the supply values. All the pressure signals at Z_1 (closest to inlet port and O–ring seal) show a region of uniform pressure close to zero absolute pressure with the cavitation zone extending as the whirl speed increases. The pressures at Z_2, closest to the damper open end, show subambient pressures but none below zero absolute. The measurements indicate that the dynamic cavitation observed is of the vapor type and demonstrate the lubricant used is not able to sustain tension.

Figure 6 shows the radial and tangential fluid film forces estimated from the measured pressures as the whirl frequency decreases in a fast coast down from 5000 to 1000 rpm of shaft speed. The theoretical predictions shown in the figure are computed with the formulae given in Eqs. (1) and (2) for the π film model and use as input the measured values of orbit radii (ϵ), whirl frequency (ω), and fluid viscosity estimated at the damper average temperature. The test results show the forces to be proportional to the whirl frequency and with a large data scatter at the highest frequencies due to the unsteady nature of the dynamic cavitation zone. The theoretical damping (tangential) force is well below the test values, since the circumferential extent of the cavitation zone is less than π radians at low frequencies. On the other hand, predictions for the radial force agree better with the measurements, especially in the mid-frequency range (200 to 400 rad/s).

Figure 7 shows the damping coefficients (C_{tt}, C_{rt}) versus the whirl frequency estimated from the pressure measurements and the corresponding theoretical predictions for the π film model. The experimental damping values are the direct ratio between the forces shown in Fig. 6 and the journal center tangential velocity (V_t). Please note that at each whirl frequency, the test orbit radii and viscosity do vary, and henceforth, the theoretical damping values depicted are not constant. The direct damping coefficient C_{tt} approaches the analytical value at the largest frequency while it is 1.8 times the theoretical value at 100 rad/s. The discrepancy is fundamentally due to the extent of the cavitation zone, which increases as the whirl speed increases. Note also that the direct damping coefficient for uncavitated conditions is twice the value of the π film coefficient, and thus, the test data is bracketed between the two theoretical models. The test cross-coupled damping coefficient (C_{rt}) is smaller than

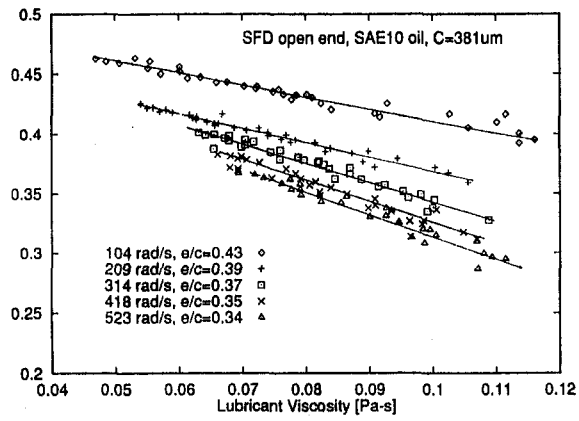


Fig. 4 Journal orbit radii versus mean lubricant viscosity: experimental results at various whirl frequencies

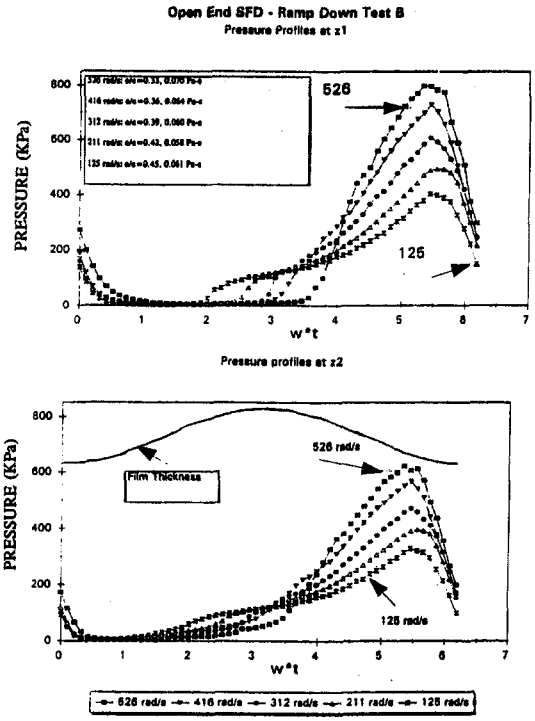


Fig. 5 Absolute pressures versus time for increasing values of whirl frequency: (*a*) measurements at location Z1; (*b*) measurements at location Z2

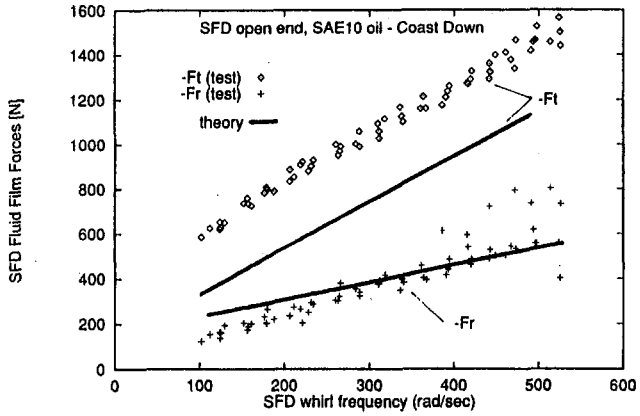


Fig. 6 Open end SFD radial and tangential forces versus whirl frequency: measurements and predictions from shorth-length film model

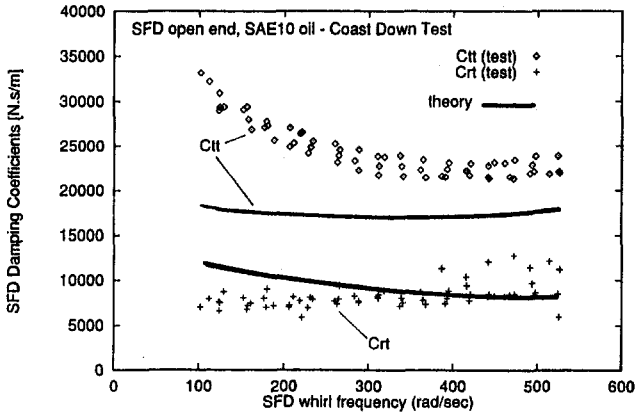


Fig. 7 Open end SFD damping coefficients versus whirl frequency: measurements and predictions from shorth-length film model

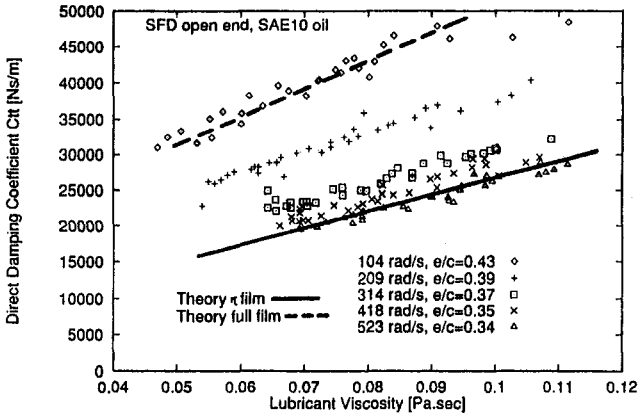


Fig. 8 Open end SFD direct damping coefficient C_{tt} versus lubricant viscosity: measurements and predictions from shorth length film model

the π film theoretical value at the lowest frequencies and approaches the analytical value at the largest test frequencies.

Figures 8 and 9 show the direct and cross-coupled damping coefficients for test conditions in which the whirl speed is kept constant and the inlet fluid temperature steadily increases due to viscous heating at the feeding pump and the damper itself. The solid lines in the figures correspond to theoretical predictions for the π film model and based on the measured values of frequency and fluid viscosity calculated at the average temperature. The orbit radius used for the calculations is equal to $\epsilon = 0.37$. The test direct damping coefficients (C_{tt}) are directly proportional to fluid viscosity and agree well with the π film model values for the largest test frequencies (418 and 523 rad/s). At the lowest frequency (104 rad/s) the damping coefficient C_{tt} is virtually identical to theoretical values for the full film model, while at intermediate frequencies (209 and 314 rad/s), the experimental coefficient is in between the two theoretical models.

The test cross-coupled damping coefficients (C_{rt}) shown in Fig. 9 are also proportional to fluid viscosity and lower than the π film model predictions except at large fluid viscosities. Note that for the full film model, the coefficient C_{rt} is identically equal to zero. The measurements then show that the cross-damping coefficient, customarily also conceived as a frequency dependent stiffness and largely responsible for the nonlinear characteristics associated to SFDs, is not as large as the theoretical predictions and depends greatly on the extent of the cavitation zone.

The experimental damping force coefficients (C_{tt}, C_{rt}) are situated in a well-defined region bounded by the analytical pre-

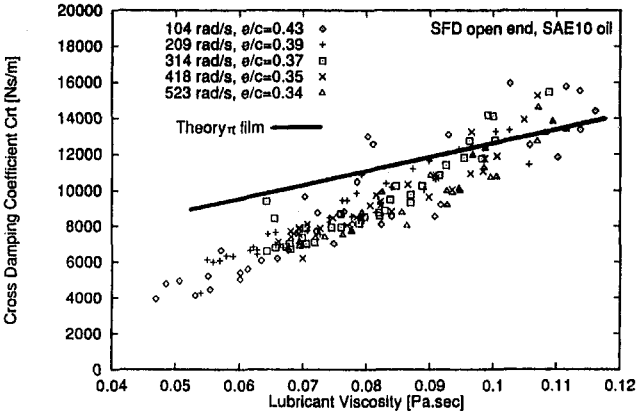


Fig. 9 Open end SFD cross damping coefficient C_{rt} versus lubricant viscosity: measurements and predictions from shorth-length film model

dictions for the π film and the full film short-length SFD models. The correlation is reasonably good mainly due to the fully submerged character of the damper test rig (no air entrainment possible), and the minimal effect of fluid inertia forces. The squeeze film Reynolds number, $Re_s = \rho\omega c^2/\mu$, ranged from 0.1125 at the lowest whirl frequency and largest lubricant viscosity, to 1.35 at the highest frequency (520 rad/s) and lowest viscosity (0.05 Pa).

Conclusions

Measurements of pressure fields and fluid film forces in a fully submerged open-end squeeze film damper are presented for tests with whirl frequencies to 524 rad/s and low supply pressures. The experiments demonstrate the radial and tangential fluid film forces to be directly proportional to whirl frequency and lubricant viscosity. The test results are compared to predictions based on the short-length squeeze film damper with the simple π film cavitation model. The test direct damping coefficients (C_{tt}) increase with fluid viscosity and decrease as the whirl frequency increases. C_{tt} coefficients correlate well with π film model predictions at the largest test frequency (523 rad/s), while at low frequencies (104 rad/s), C_{tt} values agree better with a full film model. The experimental cross-coupled damping coefficients (C_{rt}) are lower than analytical predictions, especially at low fluid viscosities and test whirl frequencies. The experiments also show vapor cavitation at zero absolute pressure with no signs of negative pressures (fluid in tension) as the whirl frequency increases. The experiments demonstrate the viscous character of the fluid film forces in a SFD test apparatus where fluid inertia effects are minimal. On the other hand, the extent of the lubricant cavitation zone has a great influence on the squeeze film forces.

Acknowledgments

The help of my wife, Gloria, in sorting the data and preparing the figures, is gratefully acknowledged. Also thanks to student worker Mr. Kevin Bivens who performed most of the measurements under my direction.

References

Al-Sharif, A., Chamiprasat, K., Rajagopal, K. R., and Seri, A. Z., 1993, "Lubrication With Binary Mixtures," ASME *Journal of Tribology,* Vol. 115, pp. 46–55.

Arauz, G., and San Andres, L., 1994, "Effect of a Circumferential Feeding Groove on the Force Response of a Short Squeeze Film Damper," ASME *Journal of Tribology,* Vol. 116, pp. 369–377.

Elrod, H. G., 1981, "A Cavitation Algorithm," ASME *Journal of Lubrication Technology,* Vol. 103, pp. 350–354.

Feng, N. S., and J. Hahn, 1986, "Cavitation Effects on Squeeze-Film Damper Performance," *ASLE Transactions,* Vol. 29, pp. 353–360.

Feng, N. S., and Hahn, J., 1987, "Effect of Gas Entrainment of Squeeze Film Damper Performance," ASME *Journal of Tribology,* Vol. 109, pp. 149–154.

Hibner, D., and Bansal, P., 1979, "Effect of Fluid Compressibility on Viscous Damper Characteristics," *Proc. Conference on the Stability and Dynamic Response of Rotors With Squeeze Film Dampers,* Charlottesville, VA, pp. 116–132.

Holmes, R., and Sykes, J., 1990, "The Effects of Manufacturing Tolerances on the Vibration of Aero-Engine Rotor–Damper Assemblies," 6th Workshop on Rotor-Dynamic Instability Problems in High-Performance Turbomachinery, Texas A&M University, College Station, TX, May 21–23.

Kinsale, I., and Tichy, J., 1989, "Numerical and Experimental Study of a Finite Submerged Squeeze Film Damper," *Proc. 1989 ASME Vibrations Conference, Machinery Dynamics: Applications and Vibration Control Problems,* ASME DE-Vol. 18-2.

Kocur, J., 1992, personal communication, Pratt & Whitney Co., May.

Marmol, R. A., and Vance, J. M., 1978, "Squeeze Film Damper Characteristics for Gas Turbine Engines," ASME *Journal of Mechanical Design,* Vol. 100, pp. 139–146.

Pan, C. H. T., 1980, "An Improved Short Journal Bearing Analysis for the Submerged Operation of Plain Journal Bearings and Squeeze Film Dampers," ASME *Journal of Lubrication Technology,* Vol. 102, pp. 320–332.

Ramli, M. D., Roberts, J. B., and Ellis, J., 1987, "The Determination of Squeeze Film Dynamic Coefficients From Experimental Transient Data," ASME *Journal of Tribology,* Vol. 109, pp. 155–163.

Roberts, J. B., Holmes, R., and Mason, P. J., 1986, "Estimation of Squeeze Film Damping and Inertial Coefficients From Experimental Free-Decay Data," *Proc. Instn. of Mechanical Engineers,* Vol. 200, No. 2C, pp. 123–133.

Roberts, J. B., and Ellis, J., 1990, "The Determination of Squeeze Film Dynamic Coefficients From Transient Two Dimensional Experimental Data," ASME *Journal of Tribology,* Vol. 112, pp. 288–298.

San Andres, L., and Vance, J. M., 1987, "Effect of Fluid Inertia on Finite Length Sealed Squeeze Film Dampers," *ASLE Transactions,* Vol. 30, No. 3, pp. 384–393.

San Andres, L., and Vance, J. M., 1988, "Effect of Fluid Inertia on the Performance of Squeeze Film Damper Supported Rotors," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 110, pp. 51–57.

San Andres, L., Meng, G., and Yoon, S., 1993, "Dynamic Force Response of an Open Ended Squeeze Film Damper," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 115, pp. 341–343.

Sun, D. C., and Brewe, D., 1992, "Two Reference Time Scales for Studying The Dynamic Cavitation of Liquid Films," ASME *Journal of Tribology,* Vol. 114, pp. 612–615.

Sun, D. C., Brewe, D., and Abel, P., 1993, "Simultaneous Pressure Measurements and High Speed Photography Study of Cavitation in Dynamic Loaded Journal Bearings," ASME *Journal of Tribology,* Vol. 115, pp. 88–95.

Szeri, A. Z., Raimondi, A. A., and Giron-Duarte, A., 1983, "Linear Force Coefficients for Squeeze Film Dampers," ASME *Journal of Lubrication Technology,* Vol. 105, pp. 326–334.

Tichy, J., and M. Modest, 1978, "Squeeze Film in Arbitrary Shaped Journal Bearings Subject to Oscillations," ASME *Journal of Lubrication Technology,* Vol. 100, pp. 323–340.

Vance, J., 1988, *Rotordynamics of Turbomachinery,* Wiley, New York.

Walton, J., Walowit, F., Zorzi, J., and Schtrand, E., 1987, "Experimental Observation of Cavitating Squeeze Film Dampers," ASME *Journal of Tribology,* Vol. 109, pp. 290–295.

Zeidan, F. Y., and Vance, J. M., 1989a, "Cavitation Leading to a Two Phase Fluid in a Squeeze Film Damper," *STLE Tribology Transactions,* Vol. 32, pp. 100–104.

Zeidan, F. Y., and Vance, J. M., 1989b, "Experimental Investigation of Cavitation Effects on the Squeeze Film Force Coefficients," presented at the 1989 ASME Design Technology Conference, Rotating Machinery Dynamics, Montreal, Canada.

Zeidan, F. Y., and Vance, J. M., 1990, "Cavitation Regimes in Squeeze Film Dampers and Their Effect on the Pressure Distribution," *STLE Tribology Transactions,* Vol. 33, pp. 447–453.

Zhang, J., Roberts, J. B., and Ellis, J., 1994, "Experimental Behavior of Short Cylindrical Squeeze Film Damper Executing Circular Centered Orbits," ASME *Journal of Tribology,* Vol. 116, pp. 528–535.

Zhao, J. Y., Linnett, I. W., and McLean, L. J., 1994, "Unbalanced Response of a Flexible Rotor Supported by a Squeeze Film Damper," ASME *Journal of Vibration and Acoustics,* Vol. 116, Paper # JVA-93-016.

The Synchronous Instability of a Compressor Rotor Due to Bearing Journal Differential Heating

F. M. de Jongh
Research and Development,
Demag Delaval Turbomachinery,
Delaval Stork V.o.f.,
Hengelo, The Netherlands

P. G. Morton
Consulting Mechanical Engineer,
Wolverhampton, United Kingdom

The paper describes a synchronous vibration instability problem encountered on a centrifugal compressor with oil-lubricated bearings. The problem was solved by modification of the compressor rotor; however, the root cause was not completely understood at that time. A possible explanation was based on a theory that suggested differential heating of the bearing journals. It was decided to verify this theory by experiments. Therefore a test rotor was designed with identical rotor dynamic characteristics to those of the compressor rotor. To fill a gap in the published research on bearing thermohydrodynamics, an experimental technique was devised to measure the surface temperature variations around one of the journals of this rotor. The dependence of significant temperature differentials across the journal upon its orbit was confirmed.

Introduction

Turbomachinery for offshore applications is extensively tested before it is shipped to the site. Compressor testing for these applications usually involves aerodynamic performance measurements according to the ASME/PTC-10 code [1] and API-617 mechanical running tests [2]. The subject of this paper is a vibration problem encountered during the mechanical running test on a LP/IP centrifugal compressor, for an offshore gas-lift application. The rotor of this drive-through compressor weighs around 450 kg and is supported by two oil-lubricated tilting pad bearings, each bearing having five pads arranged in a load-on-pad configuration. A longitudinal section of the compressor is shown in Fig. 1. During testing in the manufacturer's workshop, the rated speed could not be attained due to high rotor vibration.

Problem Description

After the aerodynamic performance test at reduced speed according PTC-10/Class III, the compressor was subjected to a no-load API mechanical running test. The machine was run as a stand-alone unit driven by a shop motor and gear. When the compressor was run up in speed it was impossible to achieve the rated maximum continuous speed of 11,947 rpm, because of high synchronous vibration levels of the rotor, although the casing vibration levels remained low. Figure 2 gives an example of the variation of vibration level and phase as a function of running speed, measured adjacent to the non-drive end (NDE) bearing. At speeds below 9000 rpm the shaft vibration level was low. From 9000 rpm the rotor speed was gradually increased in steps of 500 rpm up to 11,000 rpm. At this speed neither the vibration level nor the phase stabilized, but fluctuated over a certain range. A further speed increase to 11,400 rpm resulted in continuously fast growing shaft vibrations; at a level of around 50 percent of the bearing clearance the run was interrupted by reducing the speed to 10,000 rpm. After a dwell of about three minutes at this speed the vibrations returned to their original run-up values. Additional vibration measurements showed that the high vibration levels were associated with a second bending mode, i.e., a mode with three nodes, two of which were located near the bearings.

The complete rotor was rebalanced in the manufacturer's high-speed balancing facility. Although it proved to be very sensitive to small unbalances, it was found possible to run up to 13,142 rpm (10 percent overspeed) without any problems. On refitting in the compressor, however, the original unstable behavior was reproduced exactly. The rotor of a duplicate machine was then built into the compressor with identical results. Moreover, the instability persisted when, in a further run, the mechanical oil seals were removed. Also, testing showed that by reducing the NDE overhang weight, the speed at which instability occurred was increased. By removing the half coupling spacer weight as well as the coupling hub (total 13.7 kg), it was possible to achieve 10 percent overspeed. Further experimentation resulted in a final overhang weight reduction of 12.5 kg at the DE and 8.1 kg at the NDE, obtained mainly by a lighter coupling hub and by a change of coupling spacer material from steel to titanium. With this configuration a successful API rotor sensitivity test was carried out. Since additional investigation and testing on the compressor were not possible due to practical delivery time constraints, the compressor unit was shipped. No further vibration problems have recurred since commissioning at site.

It was decided that a full explanation of the unstable synchronous vibrations was required for two reasons: first, to establish the margin of safety of the machine in service, and second, to enable design rules to be formulated so that this particular problem could be avoided in the future.

Diagnosis/Theory

Whenever unstable vibrations occurred, they could be suppressed by reducing the rotor speed to 10,000 rpm. "Hysteresis" occurred, in that after reducing the speed, it took some time before the original status was restored. This ruled out the simplest explanation, namely that the machine was merely running too near to a low damped critical speed. Rotor response calculations showed a critical speed at 14,500 rpm (second bending mode), which according to the API [2] was far enough separated from the operating speed. In any case, the fact that

Contributed by the International Gas Turbine Institute and presented at the 39th International Gas Turbine and Aeroengine Congress and Exposition, The Hague, The Netherlands, June 13–16, 1994. Manuscript received by the International Gas Turbine Institute February 4, 1994. Paper No. 94-GT-35. Associate Technical Editor: E. M. Greitzer.

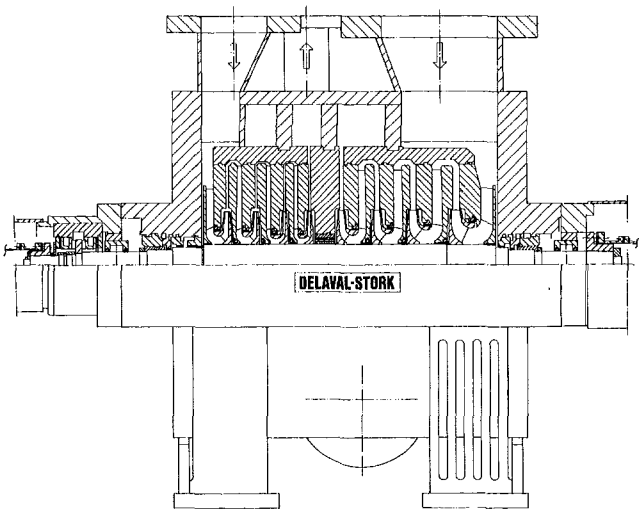


Fig. 1 Longitudinal section of LP/IP centrifugal compressor

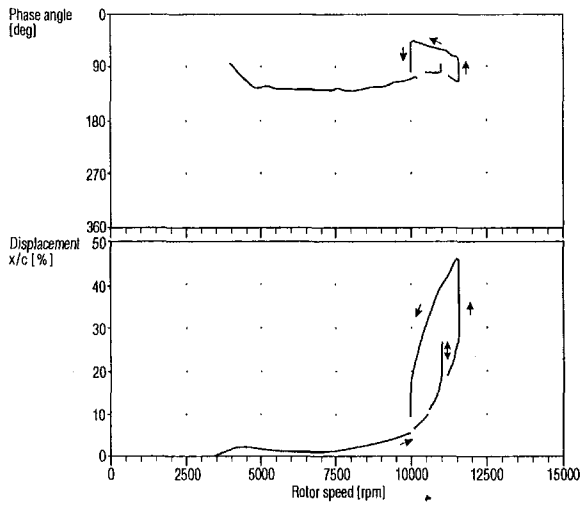


Fig. 2 Fundamental component (1X) of relative shaft vibration, measured at the N.D.E. bearing (LP/IP compressor)

the vibration level was increasing at constant speed suggested instability rather than response to unbalance. The time constants of both the growth and decay of vibrations were those normally associated with a thermal phenomenon. The indications were similar in fact to those of a seal rub in which the thermal "Newkirk effect" causes synchronous vibrations to increase, until the rotor is damaged or the rub clears due to seal abrasion [3, 4]. However, the fact that the rotor also exhibited instability with the mechanical seals and labyrinths removed precluded this rubbing possibility.

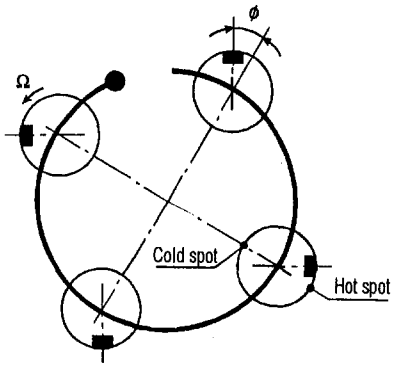


Fig. 3 Differential heating at bearing journal for synchronous forward rotor whirl

The Newkirk effect results from the contact of a stationary element with the rotor, producing a hot spot on the rotor, which in time develops a thermal bend. Any journal orbiting in a bearing can also produce a temperature difference across a diameter due to differential shearing in the oil-film. This causes the rotor to bend. It has been shown that this thermal bend, if powerful enough, in conjunction with dynamic amplification due to running in the vicinity of a critical speed, can under certain circumstances generate unstable vibrations [5–7].

Explanation of Mechanism

Since no rotor is perfectly balanced, every journal will always execute some orbit within the bearing. The orbit will in general be an ellipse whose center is eccentric to the geometric center of the bearing. In a linear system the elliptical orbit can be decomposed into a forward and a backward circular orbit, thus:

$$x = X \cos(\Omega t + \alpha) = \frac{X}{2}\{e^{i(\Omega t+\alpha)} + e^{-i(\Omega t+\alpha)}\} \quad (1)$$

$$iy = iY \sin(\Omega t + \beta) = \frac{Y}{2}\{e^{i(\Omega t+\beta)} - e^{-i(\Omega t+\beta)}\} \quad (2)$$

whence:

$$x + iy = \left(\frac{Xe^{i\alpha} + Ye^{i\beta}}{2}\right) e^{i\Omega t} + \left(\frac{Xe^{-i\alpha} - Ye^{-i\beta}}{2}\right) e^{-i\Omega t} \quad (3)$$

Considering first the forward orbit, it can be seen from Fig. 3 that in the case of synchronous vibration, one point on the journal will always be nearest to the bearing wall and the point opposite, farthest from it. A study of the thermohydrodynamics of the bearing is a complex matter, but it can be stated that the heat input to the journal comprises a circumferentially constant value and a sinusoidal term distributed around the rotor.

Apart from oil-film convection, there will be conduction along the rotor from the mean heat input and across the journal

Nomenclature

G = gain factor
I = influence coefficient
M_c = concentrated overhang mass
T = thermal gain
c = bearing clearance
i = imaginary unit
l = overhang length
m = mass
q = orbit vector
t = time
x, y = displacement
z = distance
α = phase angle
β = phase angle
γ = shaft slope compressor rotor
ζ = angle between the hot spot and the point of thinnest oil film on the journal
η = shaft slope test rotor
θ = change in slope at bearing location
ϕ = angle between unbalance vector and rotor displacement vector
Ω = rotational speed

Subscripts

B = bearing position
i = input
O = overhang position
o = output

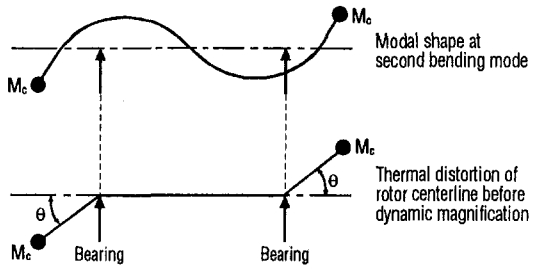


Fig. 4 Comparison of thermal bend and second bending mode

from the differential input. The latter produces a temperature differential locally across the journal, which results in a bend. Similar arguments apply to the backward whirl case when the journal is eccentric, although the heat input will always be smaller. Since linearity is assumed, the forward and backward whirl effects can be superimposed.

For the purpose of illustration, further argument will be restricted to the forward circular whirl case. Referring to Fig. 4, the thermal distortion of the rotor when stationary is shown to affect only the overhangs. Imagine for simplicity that the overhang mass is concentrated at the ends and the thermal bend θ localized at the bearings. Also consider θ arising thermally from an orbit at the bearing, denoted by vector q, then:

$$\theta = T(t, \Omega)q \tag{4}$$

where θ defines both magnitude and direction relative to rotor axes. T is a complex number varying with time and speed. Consider now a small bend θ_i imposed on this rotor at the bearing location. This results in an unbalance at the overhang of $M_c l\theta_i$. According to the influence coefficient balancing theory [8], the orbit produced at the bearing will be:

$$q = M_c l\theta_i I_{OB} \tag{5}$$

where I_{OB} is the influence coefficient between the overhang and bearing locations. Applying q to the bearing, then from Eq. (4) the value of θ resulting from thermal action will be:

$$\theta_o = TM_c l\theta_i I_{OB} \tag{6}$$

According to the control theory [9] the simple control loop of Fig. 5 is obtained, where the open loop gain is given by:

$$G = \theta_o/\theta_i = M_c l I_{OB} T \tag{7}$$

I_{OB} can be expressed as a complex number implicitly dependent on speed, mode shape, system damping, and proximity to a critical speed. A high value indicates a sensitive dynamic system. T, which is time dependent due to its thermal nature, can also be expressed as a complex number in the steady-state open-loop case, i.e., after a time elapse associated with the thermal inertia of the journal. The full theory for the time-dependent solutions is quite complex [9] but a great deal of information about stability can be found from the simple steady-state representation.

Given information either from theory or test, the values of I_{OB} and T can be obtained at any given running speed and G can be evaluated. In most practical systems there will be at least one running speed at which G becomes a real number. If this number is greater than unity, the system is unstable at that speed. If G has a negative real part, then clearly the criterion for instability is not satisfied. T will be largely negative since the orbit vector q will produce a bend whose direction results in an unbalance vector on the overhang roughly antiphase to q. The sign of the real part of G is thus strongly influenced by the sign of the real part of I_{OB}. When the latter value becomes positive, the system must be stable.

Experimental Verification

To verify the foregoing theory, it was decided to build a special test rotor, equipped with temperature sensors in one of the bearing journals in order to measure a possible temperature difference across the journal.

Design of Test Rotor. The test rotor was required to run on rotor supports, provided by an existing R&D test compressor, which would accommodate a test rotor with a bearing span of only 1041 mm instead of the original span of 1700 mm. The drive system, that is the electric motor, step-up gear and the necessary oil system, was also available. The problem then was to design a shorter test rotor with dynamic characteristics identical in all respects to the full-size machine rotor and using the same bearings.

The initial modeling proces is described in detail in the appendix, but basically the diameter of the rotor was reduced to counter the stiffening arising from shortening all rotor elements. The impellers and other additional rotor weights were simulated by simple disks. In order to maintain the same mass matrix, these disks were increased and additional disks were added. Bearings of size identical to the originals were retained. Although this process gave a dynamically identical system, the shortened overhang gave an incorrect response to the thermal bending. The overhang length of the model rotor was therefore increased and the change in dynamics was compensated for by increasing the overall stiffness of the rotor, while retaining the same mass distribution. In this way, a model for the test rotor that gave the correct influence coefficients between journal bend and journal response over the appropriate speed range was obtained. Figure 6 shows the design of the test rotor and the original compressor rotor and Fig. 7 gives a comparison of the critical speeds of both rotors. Figure 8 gives the most relevant comparison concerning the rotor response at the NDE journal for a 100 grmm unbalance at the NDE overhang. To be strictly accurate the correct response value would be that which arises from distributed unbalance of $m \cdot z$, where $m(z)$ is the incremental mass on the overhang at a distance z from the bearing. The error in concentrating the overhang mass at the coupling position is not great, however.

Journal Temperature Measurement. For verification of the theory, measurement of journal differential temperature was essential. Since the mass overhang moment at the NDE was greater than at the DE and because of the better accessibility, it was decided to measure at the NDE journal only. To develop the measuring system, preliminary experiments were conducted on a simple shaft running in the balancing facility. From these experiments small Pt-100 resistance temperature detectors (RTDs) were selected for the test rotor. Assuming that according to the theory the temperature distribution around the journal varied sinusoidally, a minimum of four sensors was required to establish the direction and magnitude of any differential temperature in the journal (Fig. 9). It was important to make as little change as possible to the heat conduction paths in the journal, so that the number and size of the sensors, as well as the guide holes for the wiring, had to be kept to a minimum. The temperature sensors were first glued into a small insert and then the assembly was mounted both with glue and

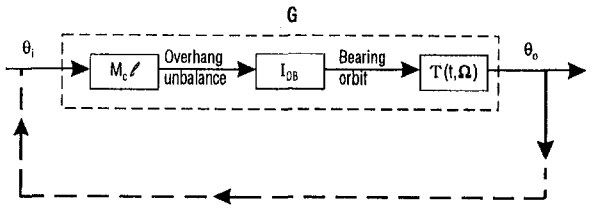


Fig. 5 Scheme of instability phenomenon

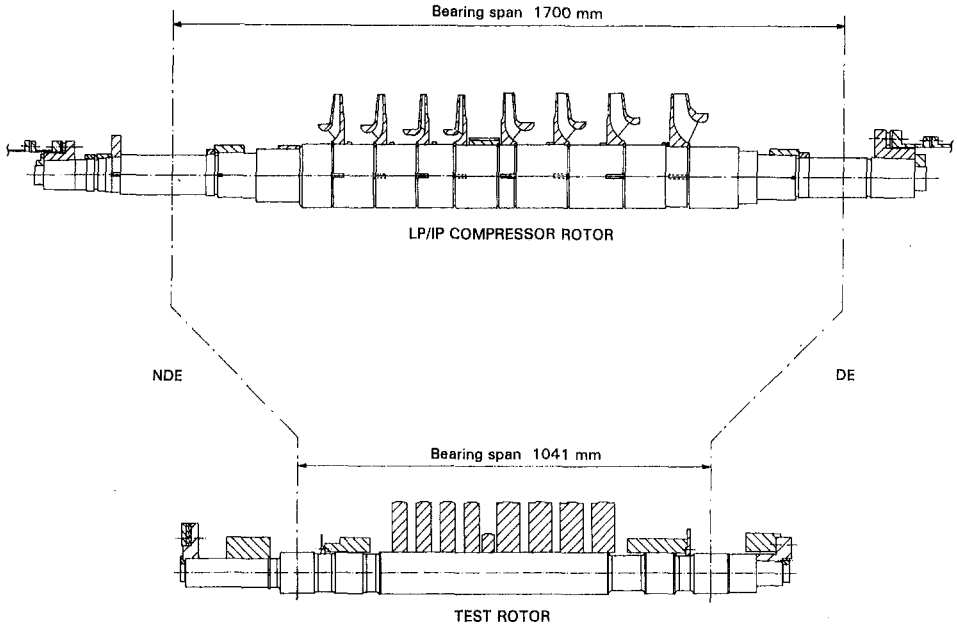


Fig. 6 Comparison of LP/IP compressor rotor and test rotor

a light interference fit into the journal, the latter to ensure a good thermal contact between sensor and journal. The sensing tip of each temperature sensor was located 1.3 mm below the journal surface. All sensors were electrically insulated from the shaft. To transfer the electrical signals from the rotating shaft to stationary measuring equipment, a special slipringless transmitter has been used. This instrument was mechanically connected to the shaft at the NDE with a flexible coupling. The temperature sensors were electrically wired to the transmitter to establish both absolute and differential temperature of the journal.

Test Setup. After the complete test rotor had been balanced at operating speed, it was built into the R&D test compressor (Fig. 10). A simple oil-lubricated axial bearing was fitted to keep the rotor in position. Since the rotor would run under atmospheric conditions, air was purged through the suction nozzle of the compressor to remove heat generated by the disk friction of the rotor. To prevent any drain oil from the bearings entering the compressor casing, labyrinth seals with an air purge were installed. The rotor was driven through a test stand planetary gear by a speed-controlled electric motor.

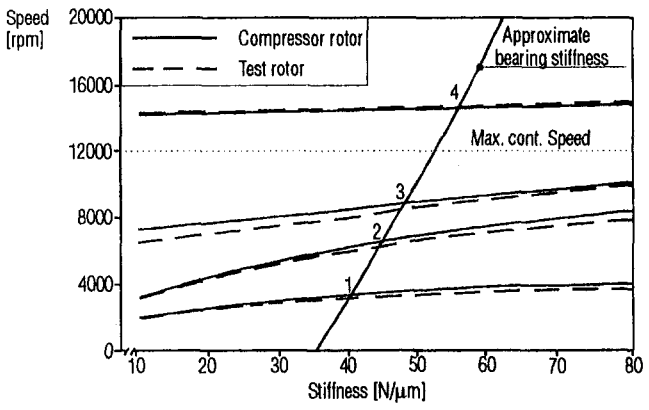


Fig. 7 Undamped critical speed map of compressor rotor and test rotor

Instrumentation. Radial shaft vibration relative to the casing was measured with noncontacting eddy-current displacement probes placed at six different locations along the test rotor, each location containing two probes spaced 90 deg apart. To correlate the journal differential temperature with the orbit size in the bearing, three additional vibration probes of the same type were mounted at the center of the NDE bearing between the pads and as a result had to be spaced 72 deg. Due to the nonorthogonality, corrections were made to obtain the correct orbit representation at this location.

Piezo-electric accelerometers were attached to the bearing pedestals both in horizontal and vertical directions to measure absolute casing vibration. A once-per-turn pulse signal was generated for phase reference and rotor speed indication. Vibration signals could be connected to oscilloscopes for time and orbit

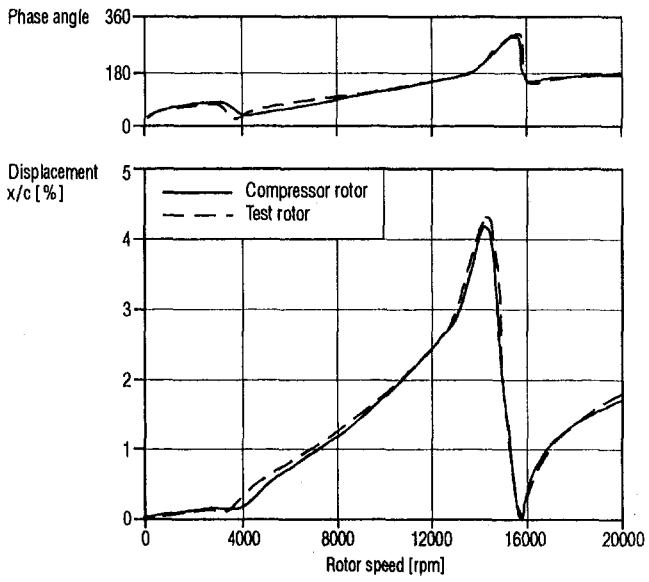


Fig. 8 Rotor response computed at the center of the N.D.E. bearing due to 100 grmm unbalance at the N.D.E. overhang

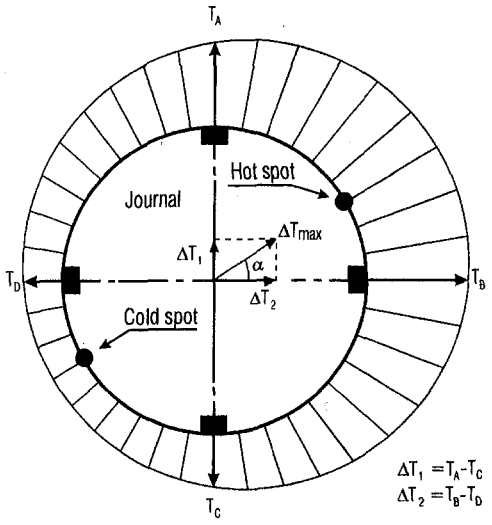


Fig. 9 Cross section of bearing journal with sinusoidal temperature distribution. Direction and magnitude of maximum temperature difference derived from four sensors.

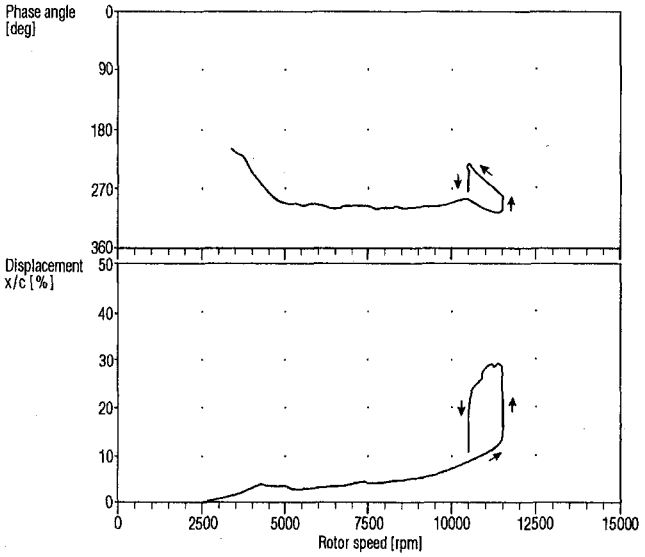


Fig. 11 Fundamental component (1*X*) of relative shaft vibration, measured at the N.D.E. bearing (test rotor)

presentation and to a two-channel FFT analyzer for frequency spectra of the individual time signals. All vibration signals were connected to an automated data acquisition system for the reduction of dynamic vibration data.

The most heavily loaded pads of the bearings were provided with temperature sensors. Thermocouples were used to measure the oil supply and drain temperatures. An orifice flow measuring device was inserted in the oil supply piping to the journal bearings.

Bearing journal differential temperatures were measured with a two-channel carrier amplifier, each channel being connected to two RTDs, which were spaced 180 deg, in a half-bridge configuration.

The vibration signals and the two journal differential temperatures were recorded on a digital tape recorder for later analysis. All other parameters were measured with an automatic data acquisition system and recorded on disk. A control monitor was capable of generating machine alarms and trips when relative shaft vibrations or bearing pad temperatures exceeded predefined set-points.

Results

Under specified bearing operating conditions, the test rotor was gradually increased in speed up to 10,500 rpm (Fig. 11). At this speed the vibration level was around 8 percent of the bearing clearance with an observed temperature difference across the bearing journal of about 3°C (Fig. 12(*a*)), but the rotor-bearing system was completely stable. When the rotor was accelerated up to 11,500 rpm, the system became unstable. Vibration was increasing at constant speed with a growth rate of about 15 percent of the bearing clearance per minute. A phase decrease was also observed. The vibration vector was spiraling in the same direction as the running speed. The differential temperatures measured at the NDE journal were continuously growing. When the vibration level reached about 30 percent of the bearing clearance, the rotor speed was reduced to 10,500 rpm and held constant. After around two minutes, the initial conditions with respect to vibration amplitude, phase,

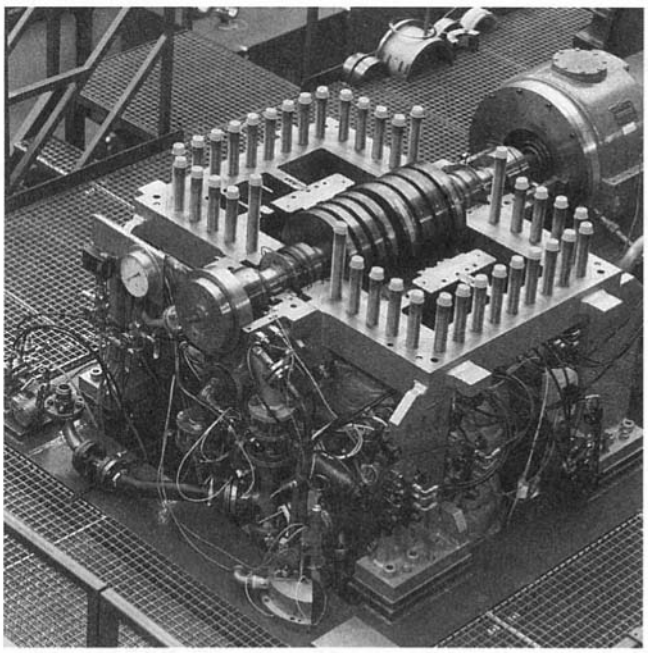


Fig. 10 Test rotor in R&D test compressor

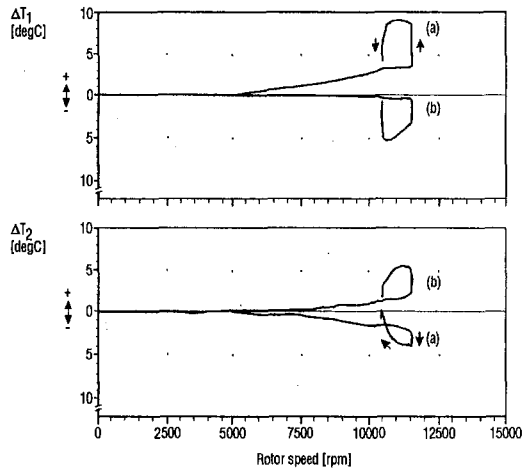


Fig. 12 Measured temperature difference at N.D.E. bearing journal: (*a*) unbalance weight at the N.D.E. overhang; (*b*) unbalance weight rotated over 180 deg

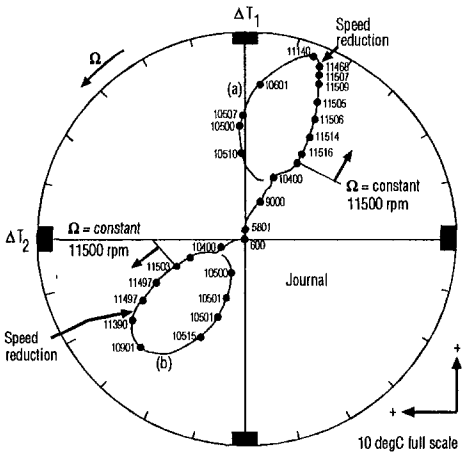


Fig. 13 Direction and magnitude of maximum temperature difference versus rotor speed: (*a*) unbalance weight at the N.D.E. overhang; (*b*) unbalance weight rotated over 180 deg

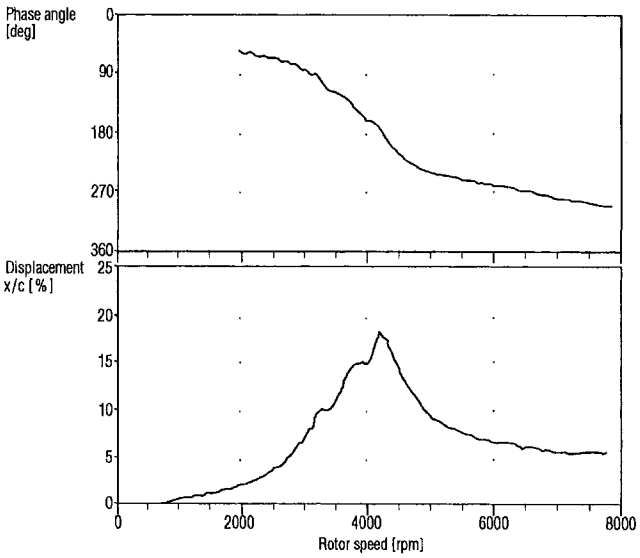


Fig. 14 Rotor response (1*X*) due to unbalance weight at midspan (passing first critical speed)

and journal differential temperature were reached again. In an attempt to establish the instability boundary more precisely, the rotor was run at 11,200 rpm. At this speed the vibration level and phase fluctuated over a range. The rotor mode shape throughout this exercise was that of the second bending mode, with large overhang deflections.

The magnitude and direction of the maximum journal differential temperature (hot spot) are shown in Fig. 13(*a*), and have been derived from the two measured differential temperatures (Fig. 12(*a*)).

The above-mentioned rotor behavior was obtained on a balanced test rotor to which a small unbalance weight was attached at the NDE overhang. This was done to force a defined point on the NDE journal circumference to be at the outside of the orbit, being the point of the thinnest oil film and so influencing the location of the hot spot. Similar results, however, were obtained without an additional unbalance weight, where the location of the hot spot was determined by the position of the residual unbalance of the rotor. Rotating the attached unbalance weight over an angle of 180 deg resulted in a change of the location of the hot spot of about the same angle (Figs. 12(*b*) and 13(*b*)). At 11,500 rpm the experiments showed a significant journal differential temperature proportional to the size of the orbit.

The journal orbit was very nearly circular and the vibration measurements showed that the journal was moving predominantly in forward whirl. The point on the journal circumference exhibiting the thinnest oil film therefore coincided with the high spot of the journal, measured with vibration probes at the center of the bearing. It was found that a phase lag existed between the point of the thinnest oil film and the hot spot on the journal circumference. At a speed of 10,500 rpm the hot spot is lagging the point of the thinnest oil film by an angle ζ of about 20 deg.

By quickly raising the rotor speed, it was found possible to run through the unstable range and reach stability at 13,600 rpm. It was thus established that the band of instability lay within the range 10,500 rpm to 13,600 rpm. The precise location of instability boundaries was clearly an impossible task because by definition the vibration growth at these speeds is around zero. Also slight variations in system parameters resulted in fluctuations in the vibration level. Somewhere in this range the rate of unstable vibration growth was at a maximum and roughly at this same region the spiraling of the journal vibration vector changed from the same direction as the rotation to the opposite direction. For reasons of safety it was judged inadvisable to dwell too long in this region, so that the exact speed of change-over in spiraling was not established.

Differential heating could also be generated by the oil seal retaining rings, which were situated on both sides of each bearing. These rings slide in grooves and should float radially to follow the rotor displacement. In the event of these rings sticking, differential journal temperatures could be generated. A test was repeated without the rings and similar results were obtained. It was concluded that the seals were functioning properly and were not contributing to the instability problem.

An additional test was carried out to observe the journal temperature behavior at lower rotor speeds. A distinct orbit could be generated at the bearing location when passing the first critical speed (4250 rpm). An excitation unbalance weight was attached at midspan of the test rotor, being the antinode of the first mode shape. A Bode plot of the run-up is shown in Fig. 14 and the measured differential temperatures are given in Fig. 15. The rotor was accelerated in speed slowly with a rate of around 150 rpm/min. When passing the critical speed, the location of the hot spot was changing continuously in one direction as a function of the phase shift between the unbalance vector and the rotor displacement vector (Fig. 16). Although additional unbalance was created due to the thermal effect, the rotor system was completely stable since in this case the open-loop gain G is smaller than unity. The measured journal differ-

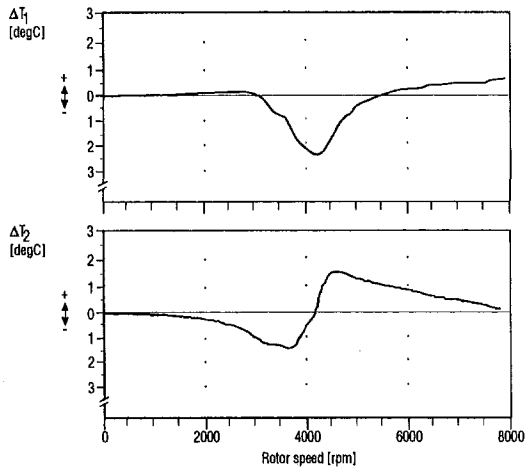


Fig. 15 Measured temperature difference at N.D.E. bearing journal (passing first critical speed)

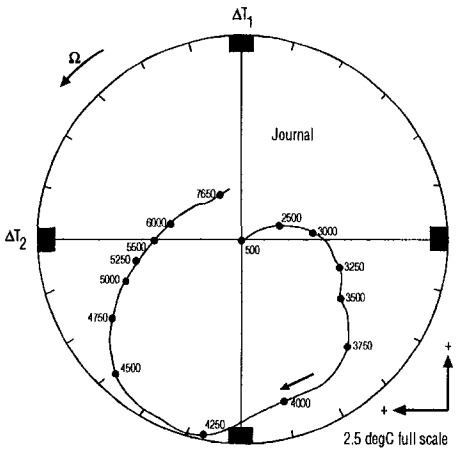


Fig. 16 Direction and magnitude of maximum temperature difference versus rotor speed (passing first critical speed)

ential temperature at a speed of 4000 rpm was about one third of that observed at 11,500 rpm. At this speed the hot spot was lagging the point of the thinnest oil film by an angle ζ of about 15 deg.

Finally, it was decided to investigate the effect on the onset of instability, of changing the overhang mass. With regard to the difficulties mentioned earlier regarding the precise location of instability boundaries, it appeared that the range was shifted upward by around 700 and 1800 rpm by reductions in the NDE coupling mass of 2 kg and 5 kg, respectively. Increases of 2 kg and 5 kg reduced the range by 800 rpm and 1300 rpm, respectively.

Discussion

At the outset, it is necessary to establish whether the test rotor correctly represented the original compressor rotor in all essential aspects. Although the test rotor was not physically identical, computed comparisons with the original machine of essential dynamic characteristics were remarkably close in the following respects:

- The critical speeds, particularly the most relevant second bending mode, were close over a wide support stiffness range.
- The response of the test rotor at the NDE journal to unbalance at the NDE overhang was very similar to that of the original compressor rotor.

Furthermore, the bearings, usually the most indeterminate components in any rotating system, were identical in terms of geometry, loading, and operating conditions.

It could be argued that the diagnosis put forward earlier in this paper would be correct, provided the same pattern of synchronous instability as in the production machine was reproduced during the test rotor experiments. In this respect the agreement was good. Although it was impossible to establish exactly the onset speed of instability, both systems were in a transitional stage at around 11,000 rpm. Also both systems exhibited very similar hysteresis effect at this boundary.

A reduction in coupling masses both at DE and NDE locations had been shown to eliminate instability within the running range on the production compressor. The experiments on the test rotor involved changing the coupling mass at the NDE only. It was clearly confirmed that such changes had a powerful effect on the onset of instability. A reduction of 5 kg in the test rotor coupling mass (around 60 percent of the reduction in the production machine), raised the onset speed to 13,000 rpm, which is roughly the 10 percent overspeed level in service. A similar increase in coupling mass reduced the onset speed to 9900 rpm. Figure 5 shows that the overhang moment directly influences the open-loop gain G. Reductions in this moment reduce G with a consequential increase in stability, provided G is positive and has only a small imaginary part.

The test rotor exhibited a feature that did not occur during the production machine tests, merely because in the latter case the maximum speed was limited. The test rotor exhibited an upper limit above which stability was restored. The reason for this limit has been discussed in the theory; a sufficient but not exclusive condition for stability is that the real part of I_{OB} becomes positive, i.e., that the I_{OB} vector has a phase angle that increases through 270 deg with increasing speed. Figure 8 shows that such a condition exists at around 14,500 rpm, thus confirming the existence of an upper stability boundary. The exact speed at which this boundary occurs cannot be predicted from Fig. 8, but it must be less than 14,500 rpm. There is every reason to believe that the original compressor rotor too would have restabilized, if it had been possible to run at a high enough speed.

Conclusion

A vibration problem that could not directly be explained was encountered on a centrifugal compressor. A test rotor with dynamic characteristics identical to those of the compressor rotor has been built to verify a theory that could possibly explain the phenomenon. The experimental results obtained confirmed that differential heating of the bearing journals was the root cause of the unstable vibration behaviour. Even for a small orbit a significant temperature difference across the bearing journal was measured.

In particular, bearing thermal characteristics play an important role in thermal bending and a comprehensive program of theoretical and experimental research is being carried out in this field. The objective is to arrive at a bearing design, in which journal differential heating is minimized, while still retaining acceptable operating and dynamic characteristics.

Acknowledgments

The authors wish to thank Demag Delaval Turbomachinery for their permission to publish this paper.

References

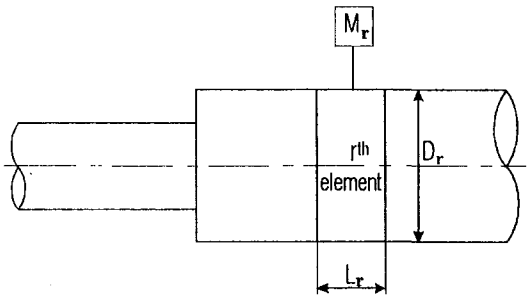
1 ASME Power Test Codes "Compressors and Exhausters," PTC 10, 1965.
2 American Petroleum Institute standard "Centrifugal Compressors for General Refinery Service," API 617, 1979.
3 Newkirk, B. L., "Shaft Rubbing," *Mechanical Engineering,* Vol. 48, 1926, pp. 830–832.
4 Dimarogonas, A. D., "Newkirk Effect. Thermally Induced Dynamic Instability of High-Speed Rotors," ASME Paper No. 73-GT-26, 1973.
5 Kellenberger, W., "Spiral Vibrations Due to Seal Rings in Turbogenerators. Thermally Induced Interaction Between Rotor and Stator," ASME Paper No. 79-DET-61, 1979.
6 Schmied, J., "Spiral Vibrations of Rotors," *Rotating Machinery Dynamics,* Vol. 2, ASME, 1987.
7 Keogh, P. S., and Morton, P. G., "Journal Bearing Differential Heating Evaluation With Influence on Rotor Dynamic Behaviour," *Proceedings of the Royal Society, London,* Vol. A441, 1993, pp. 527–548.
8 Goodman, T. P., "A Least Squares Method for Computing Balance Corrections," ASME *Journal of Engineering for Industry,* Vol. 68, 1964, pp. 273–279.
9 Keogh, P. S., and Morton P. G., "The Dynamic Nature of Rotor Thermal Bending Due to Unsteady Lubricant Shearing Within a Bearing," *Proceedings of the Royal Society, London,* Vol. A445, 1994, pp. 273–290.

APPENDIX

Modeling Technique for Test Rotor

This appendix deals with a mathematical justification of the modeling technique for a reduced rotor length.

Any shaft can be modeled from a series of elements of diameter D_r, length L_r, and mass M_r. The bearing system can be



modeled as one set of notational stiffness and damping coefficients per bearing. Any "real" bearing needs four stiffness and four damping coefficients, and in the general case the shaft has to have degrees of freedom in both X and Y stationary axes.

The modeling technique can, however, be demonstrated more simply, but equally rigorously, by considering only vibrations in a single plane, with one stiffness and one damping coefficient per bearing. The homogeneous equations of motion for the system are expressed by the following matrix equation:

$$\left[\begin{array}{c|c} M & \\ \hline & k^2M \end{array}\right]\left[\begin{array}{c} \ddot{x} \\ \hline \ddot{\gamma} \end{array}\right] + \left[\begin{array}{c|c} C & \\ \hline & \end{array}\right]\left[\begin{array}{c} \dot{x} \\ \hline \dot{\gamma} \end{array}\right] + \left[\begin{array}{c|c} \dfrac{EI}{L^3}+K & \dfrac{EI}{L^2} \\ \hline \dfrac{EI}{L^2} & \dfrac{EI}{L} \end{array}\right]\left[\begin{array}{c} x \\ \hline \gamma \end{array}\right] = 0 \quad (8)$$

where:

- M is a diagonal submatrix with typical elements of mass $M + \rho\pi LD^2/4$
- k is the radius of gyration of each mass element
- C is a submatrix with two diagonal elements of bearing damping coefficients
- K comprises two diagonal elements of bearing stiffness coefficients
- E is the modulus of elasticity
- I is the moment of inertia
- EI/L^3 is the following submatrix:

$$\begin{bmatrix} \dfrac{12EI_1}{L_1^3} & \dfrac{-12EI_1}{L_1^3} & & \\ \dfrac{-12EI_1}{L_1^3} & \left(\dfrac{12EI_1}{L_1^3}+\dfrac{12EI_2}{L_2^3}\right) & \dfrac{-12EI_2}{L_2^3} & \\ & \dfrac{-12EI_2}{L_2^3} & \left(\dfrac{12EI_2}{L_2^3}+\dfrac{12EI_3}{L_3^3}\right) & \\ & & \ddots & \text{etc.} \end{bmatrix} \quad (9)$$

- EI/L^2 is the following submatrix:

$$\begin{bmatrix} \dfrac{6EI_1}{L_1^2} & \dfrac{6EI_1}{L_1^2} & & \\ \dfrac{6EI_1}{L_1^2} & \left(\dfrac{6EI_1}{L_1^2}+\dfrac{6EI_2}{L_2^2}\right) & \dfrac{6EI_2}{L_2^2} & \\ & \dfrac{6EI_2}{L_2^2} & \left(\dfrac{6EI_2}{L_2^2}+\dfrac{6EI_3}{L_3^2}\right) & \\ & & \ddots & \text{etc.} \end{bmatrix} \quad (10)$$

- EI/L is the following submatrix:

$$\begin{bmatrix} \dfrac{4EI_1}{L_1} & \dfrac{2EI_1}{L_1} & & \\ \dfrac{2EI_1}{L_1} & \left(\dfrac{4EI_1}{L_1}+\dfrac{4EI_2}{L_2}\right) & \dfrac{2EI_2}{L_2} & \\ & \dfrac{2EI_2}{L_2} & \left(\dfrac{4EI_2}{L_2}+\dfrac{4EI_3}{L_3}\right) & \\ & & \ddots & \text{etc.} \end{bmatrix} \quad (11)$$

We will now reduce the length of each rotor element by n (<1) and I by n^3. We will also reduce the radius of gyration k by $n^{1/2}$.

Consider the original shorthand form of the dynamic Eq. (8) for the full-scale rotor-bearing system. Now write a new set of equations for the model thus:

$$\left[\begin{array}{c|c} M & \\ \hline & Mk'^2 \end{array}\right]\left[\begin{array}{c} \ddot{x} \\ \hline \ddot{\eta} \end{array}\right] + \left[\begin{array}{c|c} C & \\ \hline & \end{array}\right]\left[\begin{array}{c} \dot{x} \\ \hline \dot{\eta} \end{array}\right] + \left[\begin{array}{c|c} \dfrac{EI'}{L'^3}+K & \dfrac{EI'}{L'^2} \\ \hline \dfrac{EI'}{L'^2} & \dfrac{EI'}{L'} \end{array}\right]\left[\begin{array}{c} x \\ \hline \eta \end{array}\right] = 0 \quad (12)$$

Substitute:

$$\eta = \gamma/n$$

$$k' = n^{1/2}k$$

$$L' = nL$$

$$I' = n^3I \quad \text{or} \quad D' = n^{3/4}D$$

The new equation now becomes:

$$\left[\begin{array}{c|c} M & \\ \hline & nk^2M \end{array}\right]\left[\begin{array}{c} \ddot{x} \\ \hline \ddot{\gamma} \end{array}\right] + \left[\begin{array}{c|c} C & \\ \hline & \end{array}\right]\left[\begin{array}{c} \dot{x} \\ \hline \dot{\gamma} \end{array}\right] + \left[\begin{array}{c|c} \dfrac{EI}{L^3}+K & \dfrac{EI}{L^2} \\ \hline \dfrac{EIn}{L^2} & \dfrac{EIn}{L} \end{array}\right]\left[\begin{array}{c} x \\ \hline \gamma \end{array}\right] = 0 \quad (13)$$

If we divide the lower equations throughout by n then the original form of Eq. (8) is obtained. This proves that the model is valid. However, it is inevitable that all shaft slopes will be increased by a factor $1/n$ because of the shortening of the rotor. This is not consequential.

In summary the diameter of shaft sections has to be reduced by the factor $n^{3/4}$ and the length by n. In order to achieve the same mass, the existing disk masses have to be increased to allow for the reduced shaft sections. Where there are no existing disks, these have to be added by shrinking onto the shaft. The shrink length of the added mass disks has to be kept small to avoid appreciable shaft stiffening. The diameter of the disks has to be adjusted to retain the same mass moment of inertia. It is necessary to maintain the geometry of the bearing journals.

This model neglects the effect of shear deflection in the scaling process, in the interests of geometric simplicity. To obtain the correct shear stiffness with a simple cylindrical element requires a diametral reduction of $n^{1/2}$, which is inconsistent with the scaling strategy adopted for bending. This is not an important consideration in a slender rotor.

DISCUSSION

R. Holmes[1]

The authors describe a very interesting rotor dynamic phenomenon. The results cast a new light on certain aspects of the design of high-speed rotating machines. I would be grateful for the author's comments on the following two questions:

1 Under "Results" it is mentioned that a significant journal differential temperature was measured and that a phase lag existed between the positions of thinnest oil film and the maximum differential temperature (hot spot). What is the reason for this phase lag?
2 Although instability occurred during the no-load API mechanical running test, under "Problem Description" it was mentioned that when the same rotor was run in the high-speed balancing facility, there was apparently no indication of instability. Could the authors advance an explanation for this difference in behavior?

Authors' Closure

We are greatful to Professor Holmes for his perceptive comments. Regarding his first point, the heat transfer from oil to bearing and journal is both conductive and convective. In the absence of convection the steady-state peak temperature of the journal would coincide with the position of highest heat flux in the forward whirl case, which position is necessarily at the thinnest oil film. Oil is, however, convected counter to the journal rotation relative to the journal axes. It follows that when forward whirl is dominant, the journal hot spot must lag this position with respect to the direction of rotation. Indeed, the fact that the maximum journal temperature did not occur at the position of closest proximity to the bearing was one factor that led us to believe that the phenomenon was due to thermohydrodynamic action rather than solid friction.

Regarding Professor Holmes' second point, in the balancing facility, the bearing pedestals were "soft," whereas in the actual machine the bearing supports were of very high impedance. The two systems were therefore dynamically different. Although the absolute journal vibrations were similar at rated speed, measurements showed that the amplitudes of the journal orbits relative to the bearing were much lower in the balancing facility than in the machine for similar unbalance at the overhangs. Thus, using the simplified theory presented in the paper according to the definition of Eq. (5), the influence coefficient I_{OB} (journal orbit per unit overhang unbalance) was lower in the balancing facility. We have shown that system stability depends on the real part of the open loop gain G as defined in Eq. (7) of the paper. This is directly influenced by I_{OB}, a reduction of whose magnitude reduces the real part of G and therefore increases stability.

[1] Professor of Mechanical Engineering, University of Southampton, Highfield, Southampton, S09 3NII, United Kingdom.

Adaptive Vibration Control of a Rigid Rotor Supported by Active Magnetic Bearings

O. Lang

J. Wassermann

H. Springer

Department of Mechanical Engineering,
Technical University of Vienna,
Vienna, Austria

In this paper a digital numerical simulation is carried out in which active magnetic bearings (AMB) are applied to control nonlinear and nonsynchronous vibrations of a rigid rotor excited by nonconservative cross-coupling mechanisms. Through an adaptive control algorithm of the AMB controller, unknown cross-coupling parameters of a rotor are estimated on-line by a standard least-square estimator along with a time-varying so-called forgetting factor. The parameters of the AMB controller are adapted in order to compensate for the cross-coupling effects and to stabilize the rotor system. Transient cross-coupling response characteristics are calculated by numerical simulation of a rigid rotor supported by two active magnetic bearings. With this kind of control strategy, the stability of the system can be guaranteed for much higher values of cross-coupling coefficients than with common nonadaptive feedback controllers, designed with pole placement or least square algorithms.

Introduction

One of the advantages of using magnetic bearings in rotating machinery is the possibility of actively controlling the forces exerted on the rotor. In this way, machine vibrations caused by internal or external disturbing forces can be compensated for by properly controlling the bearing forces. Excitation forces in the rotor bearing system may be subdivided into two classes: first, time-dependent excitation forces as for example, synchronous unbalance; second, so-called self-exciting forces, as described by nonlinear mechanisms, such as fluid film bearings, annular seals, steam-to-blade excitation mechanisms, internal damping in the shaft, loose parts of the rotor, centrifuges partially filled with liquid, etc. In many applications the forces of the second class may drive the system unstable depending on the running speed and the load of rotating machinery.

In the case of time-dependent excitation forces, the dynamic model of the rotor bearing system does not change its parameters, while in the second case of self-excitation the parameters of the system (plant) may change and an adaptive control algorithm for the magnetic bearings has to be applied in order to keep the system in stable conditions.

In this paper, nonconservative cross-coupling stiffness forces are investigated. They are usually represented in a linearized form of a highly complicated nonlinear self-exciting mechanism, such as seal forces, for example.

Nordmann (1994) measured and calculated stiffness, damping, and mass coefficients of short and long annular seals. He showed that the direct stiffness decreases quadratically with the running speed, while cross-coupling stiffness and damping terms increase linearly with the running speed, and the direct damping coefficients are almost independent of the speed.

It can be stated that some theoretical models are currently available to calculate nonconservative cross-coupling stiffness coefficients approximately for known operating conditions in rotating machinery. However, when the operating conditions of a machine for some reason change in an unforeseen way, a common rotor bearing system may become unstable due to increasing cross-coupling stiffness terms. With active magnetic bearings, this problem can be solved by on-line parameter estimation of the plant (including developing destabilizing mechanisms) and adaptation of the AMB controller parameters.

There are several works that deal with various forms of adaptive control. Burrows (1991) uses adaptive control to minimize any vibrations of a general rotor bearing system. This method needs FFT, which is time-consuming because of the necessary time window needed to analyze subharmonic effects and therefore cannot be applied to quickly changing parameters, and it seems to work best for periodic disturbances.

A paper dealing with cross-coupling stiffness was written by Matsushita et al. (1990). They assume that the self-excited rotor vibration appears in a forward whirl motion and the frequency of the self-excited vibration must be known in advance.

In this paper an estimation control algorithm is discussed that needs a priori information neither about the direction of the unstable whirl motion nor about its frequency. Further, because of the shortness and simplicity of the algorithm, it can be applied to processes with quickly changing parameters of destabilizing mechanisms in a rotor bearing system.

Rotor Bearing Plant and Controller Design

Rigid Rotor Model. The shaft displacements may be measured either in the bearing or in the sensor planes. In the first case the displacement vector is defined as

$$\mathbf{x}_B = [x_A, x_B, y_A, y_B]^T \tag{1}$$

with the corresponding equations of motion for the rigid rotor

$$\mathbf{M}_B\ddot{\mathbf{x}}_B + \mathbf{G}_B\Omega\dot{\mathbf{x}}_B = \mathbf{f}_B \tag{2}$$

$\mathbf{M}_B$ and $\mathbf{G}_B$, respectively, are mass and gyroscopic matrices of the rotor, Ω is its angular speed, and $\mathbf{f}_B$ is a corresponding force vector including bearing forces, unbalance, and other excitation forces.

Linear Modeling of Nonconservative Cross-Coupling Stiffness Effects. Consider a rigid shaft with a nonconservative and nonuniformly distributed cross-coupling stiffness density $\eta(\xi)$; see Fig. 2.

The cross-coupling force density can then be described as

$$\frac{d\mathbf{f}_{cc}}{d\xi} = \begin{bmatrix} 0 & \eta(\xi) \\ -\eta(\xi) & 0 \end{bmatrix} \begin{bmatrix} x(\xi) \\ y(\xi) \end{bmatrix} = \mathcal{H}(\xi)\mathbf{x}(\xi) \tag{3}$$

For small displacements of a rigid rotor, the linear transfor-

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 13, 1995. Paper No. 95-GT-181. Associate Technical Editor: C. J. Russo.

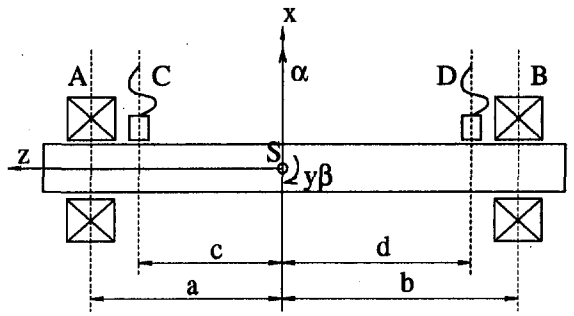


Fig. 1 Model of the rigid rotor with magnetic bearings (*A*, *B*) and sensors (*C*, *D*)

mation $\mathbf{x}(\xi) = \mathbf{R}(\xi)\mathbf{x}_B$ holds, and the nonconservative cross-coupling stiffness matrix corresponding to bearing coordinates finally becomes

$$\mathbf{N}_B = l \int_0^1 \eta(u) \begin{bmatrix} 0 & 0 & v^2 & uv \\ & 0 & uv & u^2 \\ & & 0 & 0 \\ \text{skew sym.} & & & 0 \end{bmatrix} du \tag{4}$$

with $u = \xi/l$, $v = 1 - u$. Equation (4) shows that any distribution $\eta(\xi)$ of self-exciting forces along the rotor axis can be described by not more than three different coefficients of the matrix $\mathbf{N}_B$. If the self-exciting mechanism is discrete and located at $\xi = \xi_0$ with a discrete stiffness n of unit N/m, then

$$\mathbf{N}_B = n \begin{bmatrix} 0 & 0 & v_0^2 & u_0 v_0 \\ & 0 & u_0 v_0 & u_0^2 \\ & & 0 & 0 \\ \text{skew sym.} & & & 0 \end{bmatrix} = n\mathcal{N}_B. \tag{5}$$

State Space Equations and AMB Control. Introducing self-exciting forces represented by $\mathbf{N}_B$, position stiffness forces $\mathbf{K}_{SB}\mathbf{x}_B$, and current control forces $\mathbf{K}_{iB}\mathbf{i}_C$ of the magnetic actuators, the plant equations can be established in the form

$$\underbrace{\begin{bmatrix} \dot{x}_B \\ \ddot{x}_B \end{bmatrix}}_{\dot{\mathbf{z}}} = \underbrace{\begin{bmatrix} 0 & \mathbf{I} \\ \mathbf{M}_B^{-1}(\mathbf{K}_{SB} - \mathbf{N}_B) & -\mathbf{M}_B^{-1}\mathbf{G}_B\Omega \end{bmatrix}}_{\mathbf{A}} \underbrace{\begin{bmatrix} x_B \\ \dot{x}_B \end{bmatrix}}_{\mathbf{z}} + \underbrace{\begin{bmatrix} 0 \\ \mathbf{M}_B^{-1}\mathbf{K}_{iB} \end{bmatrix}}_{\mathbf{B}} \mathbf{i}_c + \mathbf{f}(t) \tag{6}$$

with $\mathbf{K}_{SB}$ and $\mathbf{K}_{iB}$ being position stiffness and current coefficient matrices, respectively. An extended state feedback control law including integral feedback is established in the form

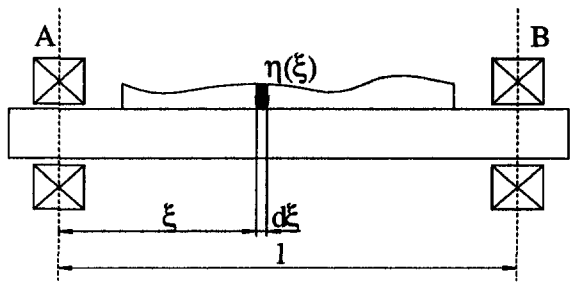


Fig. 2 Rigid rotor excited by a distributed cross-coupling stiffness density $\eta(\xi)$ with units N/m²

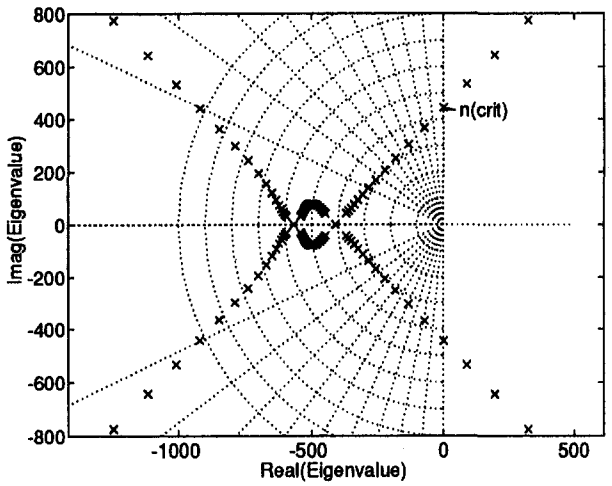


Fig. 3 Stability map of the closed-loop system for different values of the cross-coupling stiffness, n_{crit} = 7.8 MN/m

$$\mathbf{i}_c = [\mathbf{K}_P\mathbf{K}_D\mathbf{K}_I] \begin{bmatrix} \mathbf{x} \\ \dot{\mathbf{x}} \\ \int \mathbf{x}\,dt \end{bmatrix} \tag{7}$$

Stability of the LQR-Controlled Rotor Bearing System. In order to design the basic controller feedback matrix in Eq. (7), the common cost functional

$$J = \int (\mathbf{z}^T\mathbf{Q}\mathbf{z} + \mathbf{i}_c^T\mathbf{R}\mathbf{i}_c)dt \tag{8}$$

is minimized for the system of Eq. (6) with N_B and G_B set to zero. In this investigation, the weighting matrices are chosen as $\mathbf{Q} = 10^{-4}\,\mathbf{I}\ \text{m}^{-2}\,\text{s}^{-1}$ and $\mathbf{R} = \mathbf{I}\,A^{-2}\,\text{s}^{-1}$, i.e., the unstable poles of the open loop system are approximately mirror-located to the stable poles of the closed-loop system with respect to the imaginary axis of the complex plane.

If cross-coupling self-excitation $N_B \neq 0$ is added to the closed-loop rotor bearing system, one pair of poles crosses the imaginary axis and the system becomes unstable at a critical value of n; see Fig. 3. As long as the actual cross-coupling excitation n_{act} is below n_{crit}, the system is stable. In a real machine, depending on the operating conditions, the critical value n_{crit} might be exceeded. In order to keep the system stable, this has to be recognized by identification and to be compensated by adaptive control of the magnetic bearings.

Parameter Estimation and Adaptive Control

The method of parameter estimation as used in this paper is described by Slotine and Li (1991). This method is generally based on a prediction error, which is defined as the difference between a predicted output $\hat{\mathbf{y}}$ and a real output $\mathbf{y}$, in the form

$$\mathbf{e}(t) = \hat{\mathbf{y}}(t) - \mathbf{y}(t). \tag{9}$$

First, a linear parametrized form is introduced so that

$$\mathbf{y}(t) = \mathbf{W}(t)\mathbf{n} \tag{10}$$

where the k-dimensional vector $\mathbf{y}(t)$ contains the "output" of the system and the m-dimensional vector $\mathbf{n}$ contains the unknown parameters to be estimated ($m < k$). The $(k \times m)$-dimensional matrix $\mathbf{W}(t)$ is the so-called signal matrix. The output $\mathbf{y}$ and the signal matrix $\mathbf{W}$ have to be calculated from measurement data.

A standard least-square estimator can be used to minimize the total prediction error so that

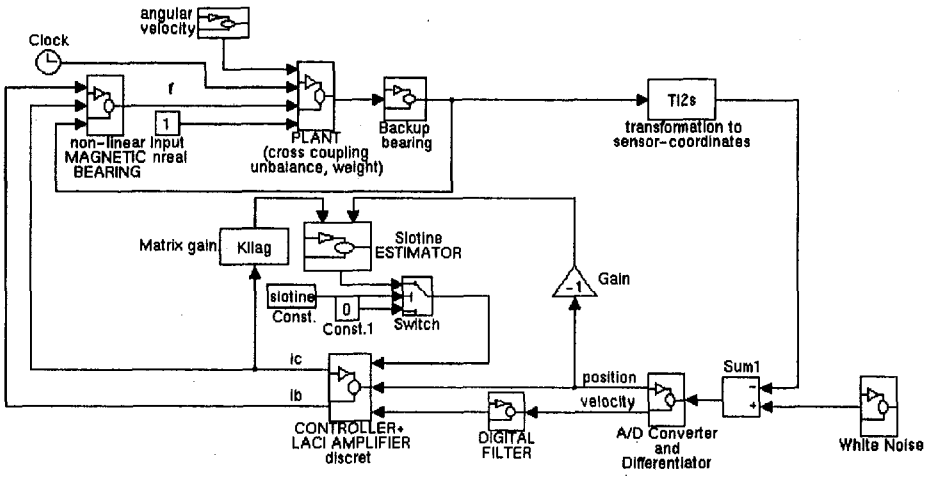


Fig. 4 Block diagram of the rotor-AMB discrete simulation model with adaptive control

$$\mathbf{J} = \int_0^t \|\mathbf{y}(\tau) - \mathbf{W}(\tau)\hat{\mathbf{n}}(t)\|^2 d\tau \rightarrow \min \tag{11}$$

with respect to an estimated parameter vector $\hat{\mathbf{n}}(t)$. If a gain matrix $\mathbf{P}$ defined by

$$\mathbf{P}(t) = \left[\int_0^t \mathbf{W}^T(\tau)\mathbf{W}(\tau)d\tau\right]^{-1} \tag{12}$$

is introduced, a parameter update law

$$\dot{\hat{\mathbf{n}}} = \mathbf{P}\mathbf{W}^T\mathbf{e} \tag{13}$$

and a gain matrix update law

$$\dot{\mathbf{P}} = -\mathbf{P}\mathbf{W}^T\mathbf{W}\mathbf{P} \tag{14}$$

can be established for on-line estimation (see Slotine and Li, 1991).

A disadvantage of the standard least-square estimator is that the gain matrix $\mathbf{P}$ is always decreasing, and therefore this algorithm can only be used for estimating time-independent parameters. However, in the present case time-varying cross-coupling terms of a rotor have to be estimated and the standard least-square estimator has to be modified. In order to estimate the current parameters properly, it is important to discount past measurement data. This can be realized by minimizing the modified functional

$$J = \int_0^t e^{-\int_\tau^t \lambda(r)dr}\|\mathbf{y}(\tau) - \mathbf{W}(\tau)\hat{\mathbf{n}}(\tau)\|^2 d\tau \tag{15}$$

where λ is a time-varying "forgetting factor."

The parameter update law is again given by

$$\dot{\hat{\mathbf{n}}} = -\mathbf{P}\mathbf{W}^T e \tag{16}$$

while the modified gain update law has the form

$$\dot{\mathbf{P}}^{-1} = -\lambda\mathbf{P}^{-1} + \mathbf{W}^T\mathbf{W}. \tag{17}$$

To keep the benefits of data forgetting (parameter tracking ability) while avoiding the possibility of gain unboundness, it is desirable to tune the forgetting factor so that forgetting is activated when $\mathbf{W}$ is persistently exciting and suspended when it is not. Since the magnitude (norm) of the gain matrix $\mathbf{P}$ is an indicator of the excitation level, it is reasonable to correlate the forgetting factor with the variation of the norm $\|\mathbf{P}\|$. So

$$\lambda(t) = \lambda_0\left(1 - \frac{\|\mathbf{P}\|}{k_0}\right) \tag{18}$$

is chosen with λ_0 being the maximum forgetting rate and k_0 the prescribed bound for the gain matrix magnitude. The factor $\lambda(t)$ implies forgetting with the maximum rate λ_0 when $\mathbf{W}$ is persistently exciting.

In a real rotor bearing system with cross-coupling effects, the output $\mathbf{y}(t)$ in Eq. (9) is supposed to be the cross-coupling self-excitation force:

$$\mathbf{y} = (\mathcal{N}_B\mathbf{x}_B)n \tag{19}$$

However, this force cannot be measured separately in the real system. Simulation studies have shown that good results can be obtained when the total bearing control force

$$\mathbf{y} = \mathbf{K}_{iB}\mathbf{i}_c \tag{20}$$

is taken as the output with $\mathbf{i}_c$ and $\mathbf{K}_{iB}$, respectively, being the control current and the current amplification matrix. The estimation error is then given by

$$e = \hat{y} - y = (\mathcal{N}_B\mathbf{x}_B)\hat{n} - \mathbf{K}_{iB}\mathbf{i}_c \tag{21}$$

and the corresponding gain matrix reduces to the scalar function

$$P(t) = \left[\int_0^t \mathbf{x}_B^T(\tau)\mathcal{N}_B^T\mathcal{N}_B\mathbf{x}_B(\tau)d\tau\right]^{-1} \tag{22}$$

Discrete Numerical Simulation

Simulation Model. The model is composed of five major components as follows (see Fig. 4):

(*a*) *Controller and Amplifier.* From the A/D converter and the differentiator the controller receives the digital values of the position and velocity signals of the rotor in sensor coordinates. The controller, as mentioned in the section above, is provided with an additional integrator. To compensate the estimated cross-coupling forces exerted on the rotor, a current is calculated that generates counteracting cross-coupling forces with no phase deviation. For this model a linear power amplifier with current injection (LACI) is applied (Wassermann and Springer, 1994).

(*b*) *Nonlinear Electromagnetic Actuators.* With a differential current control scheme (i.e., the control current is increased in one actuator while decreased by the same amount in the opposite actuator) the air gap magnetic force can be calculated as a function of the bias current i_b, the control current i_c, the nominal air gap x_0, the actual gap x, the pole angle α, the number of windings w, and the pole surface A in the form

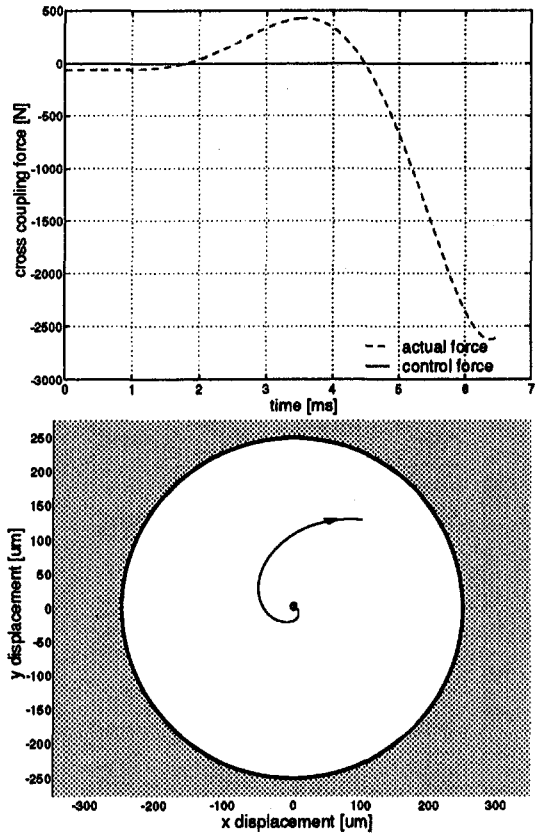


Fig. 5 Simulation of the rotor characteristics without adaptive control (white noise variance 1 μm, cross-coupling stiffness jump 35 MN/m, unbalance eccentricity 1 μm, running speed 20,000 rpm)

$$f = \frac{\mu_0 w^2 A}{4} \left\{ \frac{(i_b + i_c)^2}{(x_0 - x)^2} - \frac{(i_b - i_c)^2}{(x_0 + x)^2} \right\} \cos \alpha \quad (23)$$

(*c*) *Rigid Rotor Model (Plant) Represented by Eq. (6).* Further, a rotor unbalance eccentricity of 1 μm is introduced. It turned out that any external excitation as unbalance, for example, helps the estimator (*d*) to identify the destabilizing cross-coupling forces more quickly. The cross-coupling stiffness excitation is simulated by a step function in the time domain.

(*d*) *Estimator Designed to Negate Destabilizing Cross-Coupling Forces Present in the Rotor System (Including a Switch to Turn It on and off).* The estimator is designed as a least-square estimator with exponential forgetting as described in the section above. If the estimator is turned off the model reduces to a rigid rotor in two common active magnetic bearings providing stiffness and damping to the rotor.

(*e*) *Digital Filter.* A white noise with a Gaussian-distributed data sequence of zero average and a variance of 1 μm or 0.2 μm is added to the sensor signal in order to simulate a disturbance in the eddy current transducer signals. The digital differentiation of the sensor signals leads to high amplitudes at high frequencies. Therefore it was decided to filter the differentiated sensor signals with a Chebyshev-type I digital filter.

The simulation input variables include the applied cross-coupling stiffness n, the white noise amplitude, the unbalance eccentricity, and the running speed of the rotor.

Numerical Results. Two simulation runs of a rotor supported by two active magnetic bearings are compared in Figs. 5 and 6. In both cases the rotor is exposed to an abrupt onset of a nonconservative cross-coupling stiffness of 35 MN/m. In the first case no estimation and no adaptive control are applied. It can be seen in Fig. 5 that within 6 ms the destabilizing cross-coupling force reaches a value of 2500 N and almost immediately drives the rotor into instability. In the bottom figure the corresponding shaft orbit at a bearing station is shown, where the shaded area represents the boundary of the backup bearing.

Figure 6 shows the cross-coupling stiffness response characteristics for the same rotor bearing system with the same initial conditions after employing adaptive control. In the top of Fig. 6 the step function of the applied (actual) cross-coupling stiffness n is compared with the value $\hat{n}$ as estimated by the adaptive control algorithm. Although significant deviations between n and $\hat{n}$ can be observed, the corresponding actual and counteracting control forces are very close. The reason for this lies in Eq. (21). For higher cross-coupling forces the total bearing forces, see Eq. (20), are close to the real cross-coupling forces,

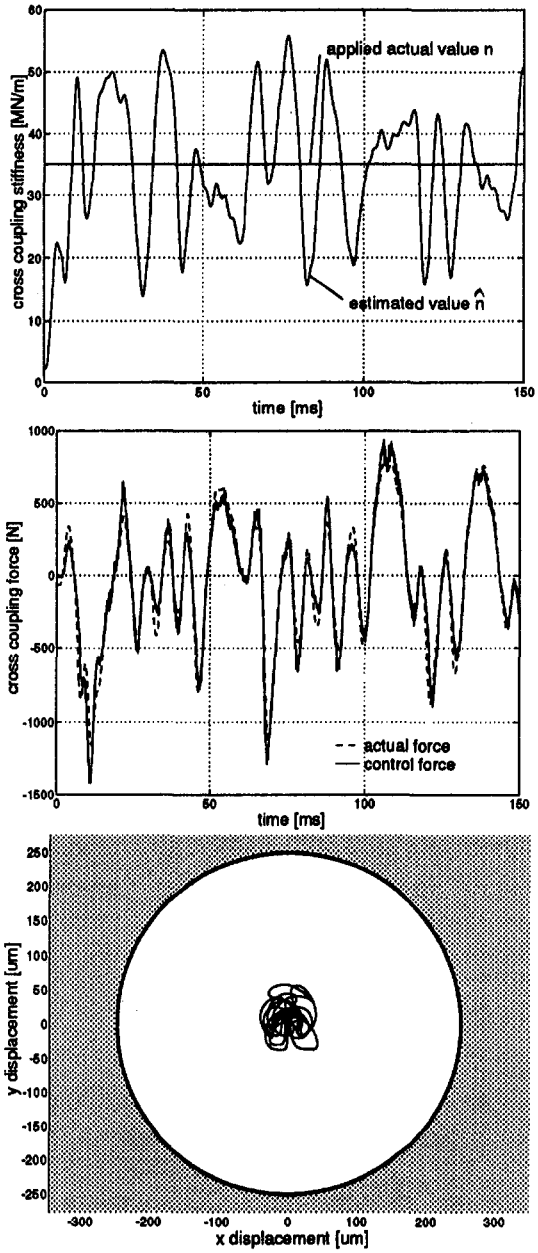


Fig. 6 Simulation of the rotor characteristics employing adaptive control (system data see Fig. 5)

Table 1 Comparison of different controller designs with respect to cross-coupling stability

Cross-coupling stiffness stability limits		
LQR-design Equation (8) **	without estimator	with estimator
$\mathbf{Q}=10^{-4}\mathbf{I}m^{-2}s^{-1}$ $\mathbf{R}=\mathbf{I}A^{-2}s^{-1}$ $\mathbf{K}_I=10^5 A/sm$	±8 MN/m	±35 MN/m
Pole placement design **	without estimator	with estimator
4 poles at $-550s^{-1}$ 4 poles at $-600s^{-1}$ $\mathbf{K}_I=10^5 A/sm$	±30 MN/m	±50 MN/m
** The controller was designed for a feedback matrix without integral feedback		

and a good estimation of $\hat{n}$ is achieved. For a small force level the difference between $\mathbf{N}_b\mathbf{x}_b n$ and $\mathbf{K}_{iB}\mathbf{i}_c$ is higher, and the estimation of $\hat{n}$ is less accurate but also less important. In the bottom of Fig. 6 the shaft orbit at a bearing station is kept within a diameter limit of about 100 μm, which is quite satisfactory for the applied cross-coupling stiffness step magnitude of 35 MN/m. It is worthwhile to note that the counteracting cross-coupling control forces are limited by the maximum current available in the actuator coil and by the magnetic saturation of the actuator core material. In the present example of Fig. 6 the dynamic control forces reach their maximum values.

One reason the estimator works better than a common control algorithm is that when the damping is increased, there is also an increase in the phase lag. This is, apart from the high sensitivity to noise of a *D*-feedback, a reason why it is not always possible to increase the damping as much as necessary to stabilize the rotor. On the other hand, the counteracting force determined from the estimated parameter $\hat{n}$ has no phase deviation to the exciting cross-coupling force.

Table 1 compares stability limits of applied cross-coupling stiffness values for different controller design strategies with and without adaptive control capabilities. It turned out that, compared with a common LQR design, a basic AMB controller design with pole placement is capable of stabilizing higher self-exciting cross-coupling forces. However, in both cases the stability performance is substantially improved by an additional adaptive control algorithm.

The proposed control algorithm might work quite well even in the presence of "auto balancing," because any persistent excitation helps the estimation algorithm to find the cross-coupling stiffness parameter more quickly.

Conclusion

A major problem in turbomachinery is self-excitation and stability loss due to nonconservative cross-coupling forces applied to the shaft. With magnetic bearings, a tool is available to compensate these forces in an active way. It turned out that a nonadaptive controller design for magnetic bearings may stabilize a rigid rotor up to a certain level of cross-coupling stiffness values. However, for transient operating conditions of a machine these values might be temporarily and significantly exceeded, and an adaptive control is necessary to guarantee stable running conditions. The applied algorithm is quite robust with respect to noise. In a future investigation this control concept will be applied to flexible rotors. Additionally, a test rig is built where realistic cross-coupling excitation will be applied and the estimation algorithm will be tested experimentally.

Acknowledgments

The authors gratefully acknowledge the support for this project by the Austrian science foundation "Fonds zur Förderung der wissenschaftlichen Forschung (FWF)."

References

Burrows, C. R., 1991, "An Evaluation of Some Strategies for Vibration Control of Flexible Rotors," in: *Int. Symp. on Magnetic Susp. Techn.*, NASA Conf. Publ., No. 3152, Part 2, pp. 691–701.

Matsushita, O., Takagi, M., Yoneyama, M., Yoshida, T., and Saitoh, I., 1990, "Control of Rotor Vibration Due to Cross Stiffness Effect of Active Magnetic Bearings," in: *Proc. 3rd Int. Conf. on Rotordynamics (IFToMM)*, Paris, pp. 515–520.

Nordmann, R., 1994, "Rotordynamische Kennzahlen von berührungslosen Dichtungen und Laufrädern; Grundsatzreferat," *VDI Handbuch 1994.*

Slotine, J. E., and Li, W., 1991, *Applied Nonlinear Control,* Prentice-Hall, Englewood Cliffs, NJ 07632.

Wassermann, J., and Springer, H., 1994, "A Linear Power Amplifier With Current Injection (LACI) for Magnetic Bearings," in: *Proceedings of the Fourth International Symposium on Magnetic Bearings,* Hochschulverlag AG, ETH Zürich, pp. 371–376.

Machine Health Monitoring and Life Management Using Finite-Element-Based Neural Networks

M. J. Roemer

C. Hong

S. H. Hesler

Stress Technology Inc.,
1800 Brighton-Henrietta Town Line Rd.,
Rochester, NY 14623

This paper demonstrates a novel approach to condition-based health monitoring for rotating machinery using recent advances in neural network technology and rotordynamic, finite-element modeling. A desktop rotor demonstration rig was used as a proof of concept tool. The approach integrates machinery sensor measurements with detailed, rotordynamic, finite-element models through a neural network that is specifically trained to respond to the machine being monitored. The advantage of this approach over current methods lies in the use of an advanced neural network. The neural network is trained to contain the knowledge of a detailed finite-element model whose results are integrated with system measurements to produce accurate machine fault diagnostics and component stress predictions. This technique takes advantage of recent advances in neural network technology that enable real-time machinery diagnostics and component stress prediction to be performed on a PC with the accuracy of finite-element analysis. The availability of the real-time, finite-element-based knowledge on rotating elements allows for real-time component life prediction as well as accurate and fast fault diagnosis.

Introduction

Maximizing operating life and availability of all critical components on rotating machinery, while minimizing unplanned maintenance downtime and the risk of catastrophic failure, is a common goal within all industry. This paper demonstrates a finite element based neural system for improving the present state of the art in machinery health monitoring by increasing the effectiveness of structural component diagnostics and monitoring. In particular, neural network classifiers were developed that operate as a "hub" for information gathering and served to make informed decisions on a rotor system's health using experimental and analytical data. The network observes the behavior of the rotor system being monitored to diagnose structural faults and predict component stresses from a variety of potential failure sources.

A desktop demonstration rotor was used as a proof concept tool. Sensors on the demonstration rig measure vibration amplitude and phase at appropriate locations throughout the rotor system. From these measurements, the neural system will diagnose faults and predict rotating member component stresses by way of a neural network trained extensively from a detailed, rotordynamics, finite-element model (FEM). Currently, commercially available expert systems used for condition monitoring use only measured system data, with no knowledge of rotating component stresses. Without these stress data, calculating remaining component life directly would be very difficult.

The miniaturized rotor rig demonstrates how the neural system can be used to obtain both real-time, finite-element model results and machine fault diagnostics. The finite-element model capability is demonstrated by estimating the dynamic stresses on the rotating shaft and reaction forces on the bearings. The diagnosis ability of the network is illustrated by predicting the location, magnitude, and phase of disk unbalances, amount of misalignment, degree of rotor rub or mechanical looseness, and bearing clearance problems. The dynamic stress estimation and structural diagnoses are both performed from the vibration measurements taken from the bearing locations. This paper also shows the ability of the network to predict the *nonlinear* dynamic stresses in the shaft, while simultaneously predicting mechanical faults.

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 27, 1995. Paper No. 95-GT-243. Associate Technical Editor: C. J. Russo.

Rotor Demonstration Rig and Measurement Processing

Rotor System Configuration. A desktop rotor rig was constructed to demonstrate the concepts proposed in this paper on actual hardware. The demonstration rig was designed to be versatile enough to duplicate various vibration-producing phenomena found in all types of rotating systems. Many different types of vibration-related characteristics were created and measured by changing rotor speed, degree of unbalance, degree of misalignment, shaft bow, shaft rub, and rotor bearing clearances. The resulting dynamic characteristics are measured with proximity probes and/or accelerometers and are processed with a multichannel dynamic signal analyzer.

The rotor configuration studied in this paper is shown in Fig. 1. The rotor setup consists of the following components:

1 $\frac{1}{10}$ HP electric motor.
2 Flexible rubber coupling.
3 Rigid steel coupling (user-controlled source of shaft misalignment)
4 3 ball bearings and 3 journal bearings.
5 2 rotating disks with balance weight holes.
6 $\frac{3}{8}$ in. diameter and 25 in.-long steel shaft.
7 Motor speed controller with closed-loop feedback.
8 Various proximity probes and accelerometers.
9 Fixturing to provide rotor preloads, rotor rub, and mechanical looseness conditions.

Two roller bearings support the motor armature, while four oil-impregnated bronze sleeve bearings are positioned between the various couplings and disks. A solid 36 in. aluminum base with adjustable bearing pedestal locations and rubber isolation feet provide sufficient rigidity to the rotor configuration. Motor speed control is maintained with a proportional speed feedback

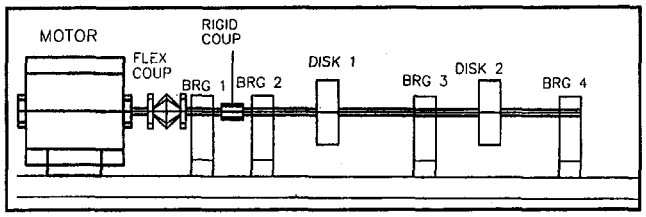


Fig. 1 Rotor demonstration rig

algorithm, with speed sensed by a dedicated proximity probes and toothed wheel.

The rotor was initially balanced within 0.05 mils in two planes before any measurements were taken. A speed run-up test was performed to experimentally determine the rotor's critical speeds. The measured rotor response from 0 to 100 Hz is given in Fig. 2. The first resonant rotor mode was identified at approximately 80 Hz or 4800 rpm. The rotor was run continuously at 40 Hz in a balanced condition to determine the sensitivity of the rotor to changing conditions.

Data Acquisition and Database Development. Vibration measurements obtained from proximity probes and accelerometers were signal conditioned and then processed by an Ono-Sokki CF6400, four-channel, digital signal analyzer. The measured frequency responses were then transferred to a personal computer where the pertinent, per-rev magnitude and phase readings were determined. Note, the input parameters to the neural network classifiers were magnitude (mils) and phase (degrees) of the 1× per rev rotor speed at all transducer locations.

Seeded faults were introduced into the rotor demonstration system by applying mass unbalances to the disks, misalignment across the rigid coupling, loosening the bearing pedestals, and installing preworn bearings. Under each of these conditions, measurements were obtained from each of four proximity probes to determine the magnitude and phase of each transducer with respect to the reference key phaser. The specific magnitude and phase measurements were then logged into a database with specific input–output pairs that are used in the neural network training procedure. A list of the input–output pairs that are included in the database is given below.

Rotordynamics Finite Element Model

A detailed model of the rotor demonstration system was developed using a dedicated finite element rotor program developed at STI called RDA (Rotor Dynamics Analysis). This computer program was used to simulate rotor operation and to train the neural network classifiers. RDA is finite-element based, and contains an array of preprocessor routines to facilitate grid

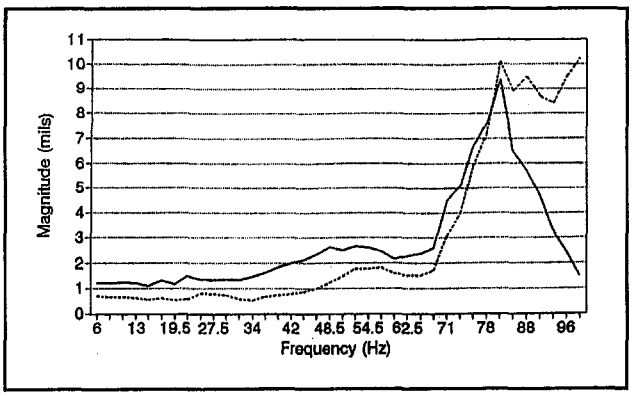


Fig. 2 Rotor response 0–100 Hz

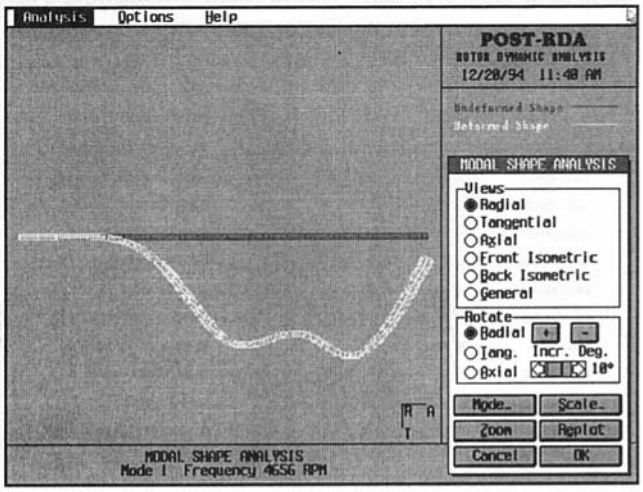


Fig. 3 Calculated first critical rotor mode

generation. The finite-element based model predicts overall rotor vibratory characteristics as well as local vibratory stress levels. The general geometry of the rotor is prescribed to the code at the outset, to allow selection of the preprocessor (and input instructions) to be made.

The added value of having a finite-element-model-based diagnostic system is that it provides a very accurate picture of the rotor stress distribution and reaction forces. These stresses and forces are the causes of many of the component failures in the rotor, bearings, seals, etc. With the rotating shaft component stresses predicted, an automated life analysis algorithm will be able to determine what the expected component life will be with any damage condition.

The finite element model of the demonstration rotor configuration was developed and correlated to the experimental results. The model was used as an additional source of information for enhanced training of the neural network. In particular, the network was trained from the model to determine dynamic stresses and forces in critical mechanical components so that it would be able to calculate remaining component life as a diagnostic output. Figure 3 illustrates the first critical mode associated with the finite element model. Note the close agreement between the measured and calculated first critical modes. This model was used for calculating dynamic stresses in the shaft and bearing reaction forces under various operating conditions including unbalances and misalignment.

Neural Network Description and Development

The neural network architectures developed in this paper served as a hub for information gathering/processing and resulted in informed diagnoses of the condition of the demo rotor rig using a combination of experimental and analytical data.

The internal interconnections of the proposed neural network architectures were developed based on the amount of data to be processed by the neural net. This is analogous to modeling the number of neurons in the system's "brain" to be utilized for a particular network. The more neurons used in the entire network, the larger the solution space will be for generalizing a system's behavior. Several multilayer, feed forward networks were developed for this project, utilizing the back propagation algorithm for minimizing the error signals. Two principal neural network architectures were developed in order to examine the sensitivity and accuracy of different network design philosophies.

Single Network Architecture. The single network configuration developed first utilized four bearing vibration input measurements (including magnitude and phase) and functional

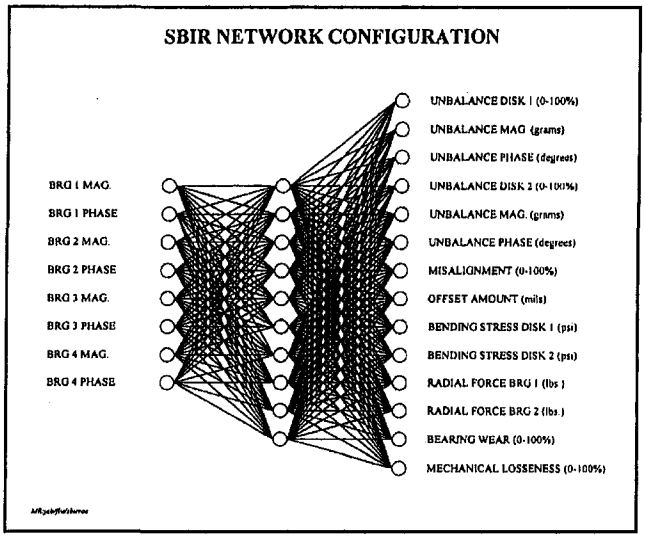


Fig. 4 Single neural network architecture

enhancements of these four sensor inputs to yield 24 input nodes to the network. A discussion on the practice of using functional enhancements to improve training accuracies and timing is given later. One hidden layer, consisting of 24 nodes, is used to increase the "flexibility" of the network. Hidden layers, when used properly, can provide more accurate correlation between complex, linear, and nonlinear training patterns. The output layer of the network consists of 14 nodes. Figure 4 is a representation of this type of single neural network architecture with its corresponding input/output parameters. Note, due to the space limitation associated with the figure, the 24 input and hidden layer nodes were reduced to fit on the page.

The first six nodes of the output layer are dedicated to determining: (1) the probability that an unbalance may exist, (2) the magnitude of the identified unbalance, and (3) the phase location of the unbalance on the out-of-balance disk. The next two output nodes determine if a misalignment exists across the rigid coupling. The probability of having a misalignment is determined along with the magnitude of the offset in mils.

Four output nodes of the network are dedicated to "virtual" sensing. Virtual sensing refers to indirectly measuring a parameter such as shaft stress or bearing forces by matching patterns of directly sensed data (such as bearing displacement) with a finite element model to yield an accurate measurement of the unmeasured parameter. For the demonstration rotor system, the shaft bending stresses and bearing forces are calculated using a detailed finite-element model of the rotor for particular rotor conditions. The neural network is then trained to recognize the sensed patterns and relate them to the values calculated from the model. The result is a neural network (trained from measurements and FE model) that is capable of "virtually" sensing stresses and reaction forces on particular components in real time without actually having installed strain gages or force transducers onboard. The last two nodes of the output layer diagnose the probability of rotor rub or bearing clearance problems and structural support looseness.

Divided Network Architecture. A divided, multilayer network architecture was developed that used the same four bearing vibration input measurements (including magnitude and phase) as the previous architecture. However, in this case, the new network configuration was broken up into smaller, more specialized classifiers. An illustration of this network architecture is given in Fig. 5.

The first section of this new network configuration diagnoses the gross fault condition as either: (1) an unbalance on disk No. 1, (2) an unbalance on disk No. 2, (3) a misalignment across the rigid coupling, (4) a bearing wear or clearance problem, or (5) a structural/mechanical looseness problem. The second layer utilizes the same bearing vibration inputs to determine specific levels of unbalance and/or misalignment about the particularly identified fault as well as give important "virtual sensing" information about shaft stresses and bearing reaction forces.

The top network architecture in the second layer determines the fault specifics with respect to a disk No. 1 unbalance. The severity of the unbalance is diagnosed in the first output node. The severity output values range from 0 of 1, with 1 representing the most severe condition. The second and third output nodes determine the magnitude and phase of the unbalance condition so that corrective action can be taken at any time. The severity of the unbalance diagnosis is continuously monitored and tracked to identify a worsening condition. The network architecture in the second layer diagnoses an unbalance condition on disk No. 2. The output node specifics are identical to the diagnosis network associated with disk No. 1.

A third network in the second layer is used to determine the severity and magnitude of any misalignment across the coupling. Severity values range between 0 and 1, as in the previous cases, while the misalignment offset amount is reported in mils. The final network in the second layer is dedicated to "virtually" sensing maximum shaft stresses and bearing reaction forces from the vibration patterns recognized at the sensor locations.

Neural Network Training and Consulting

Training of a neural network involves the process of evaluating the weights and thresholds of the numerous interconnections between the input and output layers. The training of the neural networks was conducted utilizing both unsupervised and supervised procedures. The unsupervised training was used to group similar input patterns to facilitate processing of the large number of training patterns used. The supervised training technique is

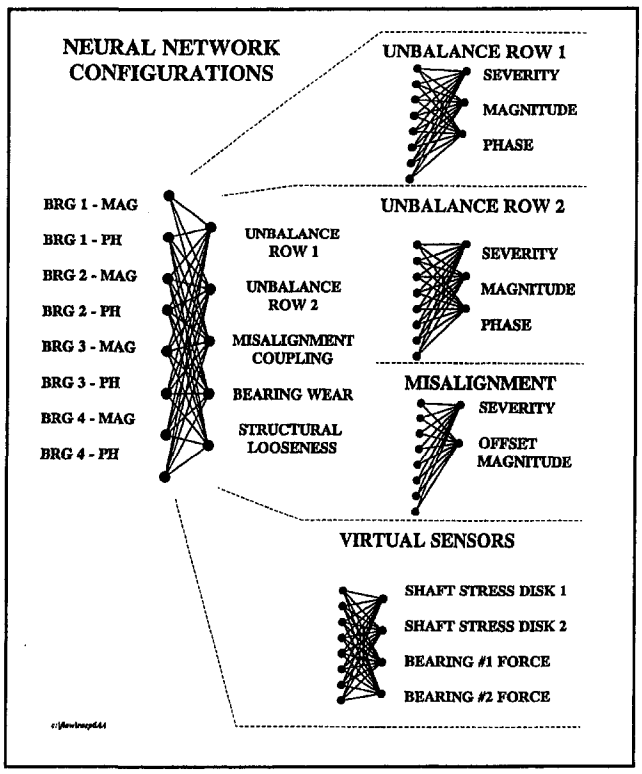


Fig. 5 Divided neural network architecture

used for specifying what target outputs should result from an input pattern. The neural network variables (weights and thresholds) are then self-adjusted to generate that target output. The combination of these two training procedures was utilized during this project in order to achieve the desirable network accuracy.

Once the internal structures of the networks were constructed, they were trained based on experimental case histories and analytically derived input/output pairs derived from the rotor dynamics computer model. Development of this database containing the neural network input/output training patterns represented a major portion of this paper's effort.

Unsupervised Training. Given a set of training patterns, an unsupervised learning algorithm will self-organize the input patterns into groups of patterns called clusters. Based on a Euclidean distance similarity measure, a large number of patterns can be separated into several clusters. During the training process, network weights and thresholds are modified and cluster centers are determined. The number of clusters formed is controlled by adjusting the cluster center radius value. After the training process is finished, the network can be consulted with either known or unknown input patterns.

Supervised Training. Supervised learning, as opposed to unsupervised learning, utilizes pairs of associated input/output patterns. This technique is commonly implemented using a Generalized Delta Rule network architecture with back propagation of error. During this procedure, the network architecture is specified in terms of the number of input and output nodes, as well as hidden layer nodes. The training set is then used to specify what target outputs should result from an input pattern, and the network automatically learns the set of parameters (weights and thresholds) that will generate this desired output. In this learning procedure, the network learns a single set of network parameters that satisfies *all* the training input/output pairs. The learning is not perfect, but is optimum on the basis of the least mean square error. In the consulting mode, the network is able to generalize and generate appropriate output patterns for any input pattern applied to the network. This attribute is the principal advantage to utilizing neural networks in condition monitoring applications.

An additional mathematical enhancement used in Phase I that helps the network architecture reduce the error associated with the numerous input/output pairs is called the "Functional Link." In this approach, the input patterns are expanded to include higher order terms associated with the original input values. Although this enhancement isn't always necessary, it often reduces the need for hidden layers and results in dramatically reduced training times.

Specific Network Training and Consulting. Both network architectures were trained with the same 232 input/output training patterns devised from both experimental measurements and the finite element model analysis. The training patterns of the network database focused on diagnosing unbalance conditions, misalignment, bearing reaction forces, and shaft stresses. As an example, experimental data were collected from the rig to train the neural network to distinguish the differences between misalignment and an unbalance condition. Both of these conditions exhibit similar one/rev vibration characteristics. Phase angle measurements were obviously very important for the network to make this distinction.

A major portion of the training sets were derived in order to recognize the differences between small changes in magnitude and phase of the applied (seeded) unbalance forces. Due to the fact that the key phaser signal was only accurate to within ±10 deg, changes in unbalance forces applied every 22.5 deg were used as the base resolution from which to identify the locations of the unbalance. Due to the fact that unbalance magnitude changes of 1.2 g-in. (0.0425 oz-in.) only produced a minimal vibration amplitude change of 0.2 mils, this value was used as the best resolution possible within the practical constraints imposed by the rotor system.

The rotordynamics finite-element model was exercised extensively with numerous unbalance force and shaft misalignment conditions. The results from each run of the finite element model (taking approximately $\frac{1}{2}$ hour each) yielded steady-state shaft bending stresses and bearing reaction forces for each of these forcing conditions. These results were then used in conjunction with the measured data to build the training pattern database.

Component Life Accumulation

A fatigue life algorithm was developed that utilized the virtually sensed shaft stresses and bearing reaction forces as a basis for computing fatigue initiation life. The algorithm estimates the amount of time to crack initiation, with crack propagation not being considered. Neuber's rule is used to compute the true stress and strain in the crack initiation region. Morrow's method is used to incorporate the mean stress effects in the life calculations, which are based on strain amplitude and the number of reversals. Miner's law computes the cumulative fatigue damage.

Strain-Life Equation. The local strain approach was used to calculate the total strain, ϵ, including elastic and plastic components, from the given stress state and the fatigue properties of material:

$$\epsilon = (\sigma_f' - \sigma_o)\,(2N_f)^b/E + \epsilon_f'(2N_f)^c$$

where E is the elastic modulus, σ_o is the true mean stress or the true steady stress, and N_f is the number of cycles required for crack initiation. In the right-hand side of the equation, the first term represents an elastic strain and the second term represents a plastic strain. This equation is the foundation for the cyclic strain-based approach to fatigue prediction and is usually called the strain-life equation.

Cycle Counting and Cumulative Damage. Under spectral loading, the dynamic strain conditions at critical locations of a component may have very complex waveforms. Several procedures exist to deal with this situation, of which the Rainflow cycle counting procedure is well known. Simply stated, this procedure consists of dividing the complex waveform into a sequence of simple cycles, and then counting the number of strain cycles within a given strain range. The resulting number is then compared with the tested fatigue life of the material at this strain level to determine the degree of incremental damage. The best-known cumulative damage assessment procedure is Miner's law, which states that the cumulative damage is equal to the sum of the incremental damage at the various strain ranges:

$$\sum \frac{n_i}{N_i} = 1;$$

This procedure is utilized in this fatigue life algorithm. The number of cycles n_i occurring at a given strain level is first computed from the Rainflow cycle counting procedure. The number of cycles to failure at each strain level, N_i, is based on test sample data and adjusted for mean stress effects. This is obtained from the strain-life equation. The portion of damage at this strain level becomes n_i/N_i. The summation sign in the Miner's law equation indicates that the cumulative damage is the sum of damage portion due to all existing strain levels. Accordingly, the crack initiation is expected to occur when the cumulative damage is equal to or greater than unity.

Shaft and Bearing Life Results. An example output from the fatigue life algorithm used to assess the shaft and bearing life is given in Fig. 6. The shaft fatigue life summary is given at the top of the page, while the bearing life results are given

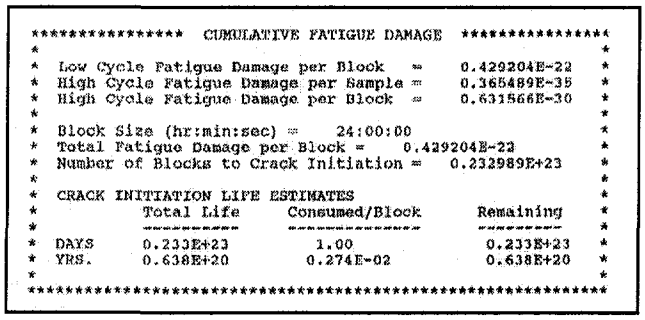


Fig. 6 Fatigue lifing algorithm output for shaft

at the bottom. One can see from the results of the shaft life calculations that shaft fatigue under an unbalance condition is primarily dominated by the low cycle fatigue associated with rotor start/stop cycles. The bearing life, on the other hand, is dominated by the high cycle fatigue associated with the rotor speed. This phenomenon is due to the fact that the shaft moves with the unbalance force (no cycling of the shaft takes place with respect to shaft speed), while the sleeve bearings are fixed with respect to the unbalance vibratory forces. In this example, the unbalance forces are too low to create any significant damage. However, the concept of assessing component life on a PC-based health monitor is demonstrated.

System Verification and Results

Several parametric tests were conducted involving seeded fault conditions applied to the laboratory rotor system. Comparison between the neural network diagnostic output and the seeded faults was performed using the output percentages from the neural network inquiry, as discussed previously. For example, if a mass unbalance was imposed on disk No. 1, then the network diagnostic output should ideally respond "100 percent chance of unbalance on disk No. 1." However, this percentage is never exactly 100 percent. Therefore, an uncertainty measure can be calculated on this basis. In this paper, specific faults diagnosed by the networks were compared to the actual, seeded fault condition. Also, comparisons were generated between the virtually measured shaft stresses and those predicted by the rotordynamics finite element model.

Table 1 gives the results from a dozen randomly selected test cases for both network configurations. Each of the test cases was selected from a group of rotor balance/alignment conditions that were not used in the training process. The first column in Table 1 describes the seeded or actual fault condition applied to the demo rotor system. The second column gives the results produced from the single network architecture. The third column gives the results produced from the divided network architecture when presented with the same rotor measurements as applied to the single network architecture.

The divided network architecture was generally more accurate than the single, all-encompassing network. The gross fault (i.e., unbalance, misalignment, bearing, mechanical looseness) was always diagnosed properly with either network configuration. However, the diagnosed fault specifics (i.e., magnitude and phase of unbalance or degree of misalignment) were identified more accurately with the smaller, individual networks. This trend is understandable based on the reduced network complexity and number of training patterns needed for identifying the fault specifics.

Conclusions

A neural network-based health monitor that performs both mechanical diagnostics and component life prediction was demonstrated with the use of a miniaturized rotor system. The rotor system was subjected to several different mechanical fault scenarios including: mass unbalances, misalignment, rotor rub, and structural looseness in order to examine the benefits of utilizing neural network technologies for monitoring rotating machinery. The neural network classifiers, developed under this Phase I project, were able to diagnose mechanical faults accurately based on the associated vibration signatures measured from the desktop system. In addition, an algorithm for component life estimation was demonstrated based on the diagnostic outputs provided by the neural networks. In particular, the shaft stresses and bearing forces were predicted by the diagnostic network after being trained extensively using results of the rotor dynamic finite element model and system measurements.

The concept of training diagnostic neural network classifiers with both rotor system measurements and detailed finite element models is highlighted as a significant advancement in condition monitoring applications. The rotordynamics finite-element model was used to train the diagnostic network to recognize fault patterns and their resulting effect on shaft and bearing stresses. This real-time, "virtual" sensing of shaft stresses and bearing radial forces allows for component life monitoring to be achieved in real time. In other words, the ability of the neural networks to recognize particular vibration signatures and correlate them with associated shaft stresses is of particular significance. Using sensory data with calculated rotordynamics results as demonstrated in this project represents the most effective use of each data source.

Finally, the two primary diagnostic network classifiers developed under this program were examined for architecture-based sensitivities and accuracies. Not surprisingly, the divided network architecture was consistently more accurate than the all-encompassing architecture for diagnosing fault specifics. In other words, the more dedicated and smaller network architectures of the divided network provided a better "tuned" classifier for diagnosing fault specifics like unbalance phase angle and particular amount of offset in a misalignment condition. The

Table 1 Network consulting results

#	True Machinery Fault Condition	Diagnosed Fault Condition (Single Network)	Diagnosed Fault Condition (Divided Network)
1	Unbalance Disk 1 2.0 grams @ 1.2" 180 degrees phase	96% Chance Disk 1 1.94 grams @ 1.2" 208 degrees phase	104% Chance Disk 1 1.94 grams @ 1.2" 185 degrees phase
2	Unbalance Disk 1 1.0 grams @ 1.2" 157 degrees phase	97% Chance Disk 1 0.93 grams @ 1.2" 172 degrees phase	101% Chance Disk 1 0.97 grams @ 1.2" 152 degrees phase
3	Unbalance Disk 2 2.0 grams @ 1.2" 225 degrees phase	99% Chance Disk 2 1.98 grams @ 1.2" 204 degrees phase	100% Chance Disk 2 1.97 grams @ 1.2" 213 degrees phase
4	Unbalance Disk 2 2.0 grams @ 1.2" 157 degrees	98% Chance Disk 2 1.80 grams @ 1.2" 112 degrees phase	99% Chance Disk 2 1.82 grams @ 1.2" 178 degrees phase
5	6 Mils Misalignment at Coupling	78% Misalignment 3.1 Mils	101% Misalignment 6.3 Mils
6	Unbalance Disk 2 2.0 grams @ 1.2" 180 degrees	99% Chance Disk 2 1.99 grams @ 1.2" 159 degrees	102% Chance Disk 2 2.01 grams @ 1.2" 185 degrees
7	Unbalance Disk 1 2.0 grams @ 1.2" 90 degrees	101% Chance Disk 1 1.99 grams @ 1.2" 90 degrees	100% Chance Disk 1 2.16 grams @ 1.2" 94 degrees
8	Unbalance Disk 1 3.0 grams @ 1.2" 112 degrees	94% Chance Disk 1 1.93 grams @ 1.2" 171 degrees	101% Chance Disk 1 2.96 grams @ 1.2" 128 degrees
9	Unbalance Disk 1 1.0 grams @ 1.2" 270 degrees	99% Chance Disk 1 0.87 grams @ 1.2" 293 degrees	102% Chance Disk 1 0.96 grams @ 1.2" 291 degrees
10	Unbalance Disk 2 3.0 grams @ 1.2" 225 degrees	95% Chance Disk 2 1.91 grams @ 1.2" 177 degrees	99% Chance Disk 2 3.01 grams @ 1.2" 262 degrees
11	2 Mils Misalignment at Coupling	75% Misalignment 1.2 Mils	95% Misalignment 2.5 Mils
12	4 Mils Misalignment at Coupling	72% Misalignment 2.4 Mils	98% Misalignment 4.5 Mils

overall gross diagnostics for strictly identifying unbalances, misalignment, etc., were diagnosed accurately using either network configuration. Current knowledge-based expert systems used on rotating machinery are only capable of diagnosing these gross fault conditions. Important specifics related to these particular faults are not considered.

References

Aleksander, I., and Morton, H., 1990, *An Introduction to Neural Computing,* Chapman and Hall Publishing, New York.

Dietz, W. E., Kiech, E. L., and Ali, M., 1989, "Jet and Rocket Engine Fault Diagnosis in Real Time," *Journal of Neural Network Computing,* Vol. XX, pp. XX–00.

Hong, C. A., 1991, "Analysis and Replacement of the 15th Stage Blading of Martins Creek Unit," STI Technical Paper.

Pomfret, C., 1993, "Engine Life Monitoring, Performance Trending, and Diagnosis in the USAF: A Concept for the Future," Discussion paper, WPAFB.

STI, 1993, "Development of Advanced Rotordynamics Software for Root Cause Investigations," Report PA791.

Widrow, B., 1989, "DARPA Neural Network Study," AFCEA International Press.

Aerodynamic Loading and Magnetic Bearing Controller Robustness Using a Gain-Scheduled Kalman Filter

R. D. Smith
Center for Electromechanics.

W. F. Weldon
Department of Mechanical Engineering and Department of Electrical and Computer Engineering.

A. E. Traver
Department of Mechanical Engineering.

The University of Texas at Austin, Austin, TX 78712

Modeling or predicting aerodynamic loading effects on rotating equipment has been a source of concern to those who wish to examine stability or response of critical components. The rotordynamic model of the system employed for such examination assumes greater importance for active bearings than for passive ones, if only because of the additional potential for instability introduced by the controller. For many systems, aerodynamic loading may vary widely over the range of operation of the bearings, and may depend on extended system variables. Thus, potential controllers for active magnetic bearings require sufficient robustness or adaptation to changes in critical aerodynamic loading parameters, as might be embodied in cross-coupled stiffness terms for compressor impellers. Furthermore, the presence of plant or measurement noise provides additional sources of complication. Here, the previous development of a nonlinear controller for a hypothetical single-stage centrifugal gas compressor is extended by comparing the compensator performance using a multivariable Luenberger observer against that of a stationary Kalman filter, both gain-scheduled for rotational speed. For the postulated system, it was found that the slower poles of the Kalman filter did not observably detract from controller convergence and stability, while predictably smoothing out the simulated sensor noise.

Introduction

Reducing bearing frictional losses, eliminating lubrication systems, and modifying vibrational behavior through active control have been some of the reasons motivating the use of active magnetic bearings. Such sophisticated systems as hybrid and fully electromagnetic bearings offer a challenging multivariable control problem whose approximate solution is most often found in the application of PD or PID feedback (Allaire et al., 1983; Lewis et al., 1988; Humphris et al., 1986). The merits of local PD or PID control are most pronounced with regard to implementational ease, as the structural simplicity of the control laws evades the necessity of additionally cumbersome state estimation. Of course, when applying SISO techniques to the multivariable problem, one sacrifices the ability to effect greater structural control as one may when operating on the entire state space. Thus, the potential exists to affect rotordynamic instabilities—to alter the closed-loop eigenstructure favorably through the mechanism of state feedback—and to increase the stable operating range of a given device. Also, the state space approach has been a conduit to certain optimal control methods. Matsumura and Yoshimoto (1986) provided a theoretical and experimental treatment of an optimal controller for a rigid rotor, which relied on full-state feedback, while others (e.g., Anton and Ulbrich, 1985) have dealt with flexible rotors controlled via output feedback. One of the advantages of advancing to modern control solutions for the active bearing problem is the increased capacity for load accommodation, or load modeling, which is aimed at improving the imbalance response characteristics through frequency shaped filters, noise driven load models, etc. Mizuno and Higuchi (1984) offer an early work with this approach.

Contributed by the Aeromechanical Interaction Division for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Aeromechanical Interaction Division September 1995; revision received April 1996. Associate Technical Editor: R. E. Kielb.

For rigid rotors, a variety of control strategies is possible as full-state measurement is not necessarily prohibitive. For high-order flexible rotordynamic systems, however, the implementation of usefully stable on-line state estimation further complicates the modeling problem and may restrict the performance of the linear state space methods employing full-state feedback, or other techniques in nonlinear control, which often fall short of requiring complete state feedback.

It is not the intention here to deliberate the respective merits of possible controllers; rather, the interest is briefly and specifically concerned with a robust nonlinear compensator designed for a hypothetical single-stage centrifugal gas compressor in the face of wide variation in the aerodynamic loading, as popularly embodied in a cross-coupled stiffness term acting on the impeller. Measurement noise will be assumed, which will necessitate recourse to the Kalman filter for state estimates in the feedback law. Considerations of rotordynamic loading from various sources, such as liquid or gas annular seals, fluid interaction forces, etc., is often left unclear in contemporary discussions of active magnetic bearing control. Unfortunately, these self-exciting mechanisms are highly correlated with other system variables (e.g., speed, displacement, flow rate, pressure), are often time dependent, and merit close scrutiny. For example, the examination of magnetic bearings for submarine pump noise control might require careful attention to fluid loading. While bulk flow models and perturbation techniques for liquid and gas annular seals have produced accessible and semi-successful results (see Childs, 1993; Childs and Scharrer, 1986; or Thieleke and Stetter, 1990), the same cannot be said for aerodynamic loading on compressor impellers. These forces have been known to cause instabilities, and will motivate an examination of controller robustness and state estimation. Lang et al. (1994) offered a brief discussion of the treatment of the potentially destabilizing effects of cross-coupled seal stiffness by making use of an on-line parameter estimation routine. Although their focus was restricted to that of a rigid rotor, they did emphasize

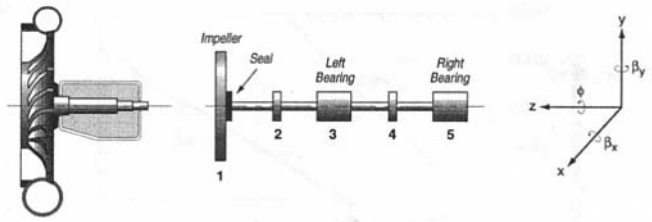


Fig. 1 Hypothetical centrifugal gas compressor

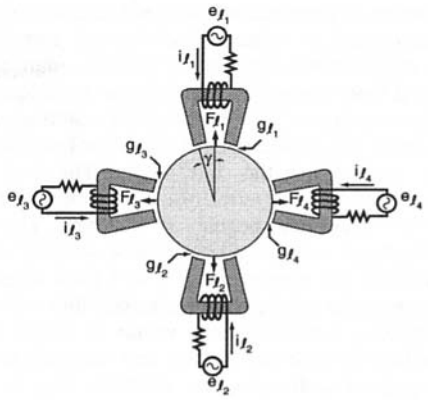


Fig. 2 Representative four-axis magnetic bearing

the paucity of reliable analytical calculations for key rotordynamic terms, which lend to instability.

Dynamic Model

The hypothetical centrifugal gas compressor of Fig. 1 is intended to serve as the physical system of interest whose single-stage overhung impeller and annular gas seal will provide the requisite model uncertainty that is the basis for the objective compensator design, namely a nonlinear sliding mode controller implemented via the stationary Kalman filter, gain-scheduled for rotational speed. This compensator will be compared to a previous design, which made use of a gain-scheduled multivariable Luenberger observer whose poles were placed through straightforward eigenstructure assignment techniques, although not according to any specific algorithm (e.g., Gopinath, 1971). From Fig. 1, the system will be simplified by neglecting the eye packing labyrinth seal at the impeller shroud interface, as well as the balance drum, but the mathematical ideas behind the compensator design will not be greatly enhanced by considering additional levels of uncertainty.

Using real coordinates, a five-node rotor displacement vector in the global X, Y, Z frame may be written as $\mathbf{z} = [\mathbf{z}_{X-Z}\ \mathbf{z}_{Y-Z}]^T = [R_{1X}\ \beta_{1Y} \ldots R_{5X}\ \beta_{5Y}\quad R_{1Y}\ \beta_{1X} \ldots R_{5Y}\ \beta_{5X}]^T$, where $R_{iX,Y}$ symbolizes the displacement of the subscripted inertial node along the subscripted axis direction, while $\beta_{iY,X}$ similarly implies an associated rotation about an orthogonal axis, producing the horizontal and vertical planar orientations of the flexible rotor as indicated in the partitioning. These global axes are indicated in Fig. 1. In the mathematical development that follows, the finite element representation of the rotordynamic matrices may be assumed, although a five-node discrete model was taken here for initial simplicity in the vector of forcing functions; that is, the mass imbalance forces were taken as "lumped" at the five nodes: the impeller, the two bearing journals, and the midspans of the intervening shafts. Gravitational loading was similarly concentrated. The net horizontal and vertical bearing forces were assumed to be acting at discrete points on their journals, which implies several contemporary assumptions about journal and electromagnet pole face curvature, fringing, etc. One-dimensional field treatments of these forces have been presented by many authors (e.g., Hebbale, 1985; Sortore, 1990) and others (Schweitzer, 1988; Traxler, 1985) have presented empirical results to justify/evaluate some of these popular simplifying assumptions. Finally, decoupling the rotational dynamics from the lateral dynamics results in:

$$M\ddot{\mathbf{z}} + [C_D + C_G]\dot{\mathbf{z}} + K\mathbf{z} = \mathbf{F} \tag{1}$$

$$\bar{J}\dot{\phi} = M_z \tag{2}$$

where $\bar{J}$ is the net effective polar moment of inertia and the seal, impeller, and bearing forces are included with the harmonic imbalance forces and gravitational terms in $\mathbf{F}$. A representative four-axis magnetic bearing is shown in Fig. 2, and the relevant electrical dynamics and the net, left and right, vertical and horizontal bearing force vector is given as:

$$\frac{d}{dt}\tilde{\mathbf{i}} = -\text{diag}[(R/L)_{i=1,4}]\tilde{\mathbf{i}} + \text{diag}[(1/L)_{i=1,4}]\tilde{\mathbf{e}} \tag{3}$$

$$\mathbf{F}_{\text{Brg}} = \mathbf{F}' + K_p\tilde{\mathbf{g}} + K_i\tilde{\mathbf{i}}$$

Nomenclature

0 = null matrix
I = identity matrix
κ = empirical constant
f = operating frequency
Diam = impeller diameter
$\mathbf{D}$ = "disturbance" vector
$\mathbf{z}$ = rotor displacement vector
$\mathcal{P}$ = compressor stage power
M_w = molecular weight of gas
$\mathbf{X}_r$ = reduced order state vector
G = process noise input matrix
$R_{iX,Y}$ = translation of ith inertial node along X or Y direction
A, B = system matrix, input matrix
$\mathbf{y}$, H = output vector, output matrix
L = stationary Kalman gain matrix
$\mathbf{w}$, $\mathbf{v}$ = process, measurement noise
$\Psi(\cdot)$ = general, nonlinear functional component of feedback signal
$\beta_{iY,X}$ = associated rotation of ith inertial node about orthogonal axis to trans.
$\hbar$ = restrictive dimension in flow path
υ = tuning parameter
Δ = nominal (regulated) air gap length
ρ_d/ρ_s = discharge/suction density ratio
$(R/L)_i$ = resistance/inductance ratio of electromagnet pair
K_D, K_i = bearing displacement, current stiffness matrices
P = positive definite solution to Riccati equation
M_z = total applied moment about rotational axis
$(1/L)_i$ = inductance ratio of electromagnet pair
$\bar{J}$ = rotor net effective polar moment of inertia
$\mathbf{F}$ = rotordynamic vector of excitation sources
$\mathbf{F}_{\text{Brg}}$ = magnetic bearing force component of $\mathbf{F}$
ξ = bound on model uncertainty magnitude
$s(\cdot)$ = sliding surface in nonlinear state space
$C_{D,G}$ = rotor damping, gyroscopic matrices
λ = pole of sliding surface dynamics
M, K = rotor mass, stiffness matrices
Q_f, R_f = process, measurement noise spectral density matrices
$\tilde{\mathbf{i}}$, $\tilde{\mathbf{e}}$ = vector of electromagnet perturbation currents, voltages

where the electrical parameters have been indexed to accommodate any intentional or unintentional bearing asymmetry, and the bearing displacement and current stiffness matrices may be easily derived from first-order perturbation techniques such as found in Matsumura and Yoshimoto (1986), among other references. Equation (3) will augment the rotordynamics of Eqs. (1) and (2) to lengthen the state vector. The four net bearing forces are expressed in terms of the gap length error associated with the respective degree of freedom. That is, $\tilde{\mathbf{g}} = -[R_{3Y}R_{5Y}R_{3X}R_{5X}]^T$, and the displacement stiffness and current stiffness matrices are appropriately 4×4 (and diagonal), yet this ordering remains arbitrary. For compatibility with the rotordynamic matrices, a bearing force vector of length 20 will be assumed, where the relevant vectors and matrices are now appropriately padded with zeros. For example, $\hat{\mathbf{F}}_{\text{Brg}} = \hat{\mathbf{F}}' - \hat{K}_D\mathbf{z} + \hat{K}_i\tilde{\mathbf{i}}$, where dim $(\hat{K}_i) = 20 \times 4$.

For plain annular gas seals, perturbation treatments have been used to produce numerical algorithms for calculating rotordynamic coefficients. The approach of Thieleke and Stetter (1990) follows that of Childs and Scharrer (1986), yet reduces the number of variables to be calculated, thus simplifying the algorithm. It was the method adopted for calculating the seal coefficients used here.

For the aerodynamic loading on the single impeller, no generalized analytical approaches have been satisfactory. This is in contrast to pump impellers, where substantial data have been collected to support reasonable numbers for rotordynamic models. Wachel and von Nimitz (1981) introduced an empirical formula for estimating the cross-coupled stiffness coefficient based on extensive experimental work with several unstable units. It has a wide acceptance and appears to yield reasonable estimates of actual stability points. Thus, with variables defined in the Nomenclature section,

$$-\begin{bmatrix} F_X \\ F_Y \end{bmatrix}_{\text{Imp}} = \begin{bmatrix} & K_{XY} \\ -K_{YX} & \end{bmatrix}\begin{bmatrix} R_{1X} \\ R_{1Y} \end{bmatrix} \quad \text{where}$$

$$K_{XY} = \frac{\kappa M_w{}^{\varphi}}{\hbar f \,\text{Diam}}\left(\frac{\rho_d}{\rho_s}\right) \quad (4)$$

The direct stiffness, which remains obscure, would tend to affect the natural frequencies (and, hence, critical speeds) while the cross-coupled term has been the focus of stability. In this simple system, the single seal will develop forces, which are modeled as acting at the same node as those of the impeller. In a more realistic system, this would not be the case.

Controller Development

Modeling and controlling continuous physical systems with a finite number of modes (structural uncertainty) has been insightfully shown by Balas (1978) to offer the opportunity for control and observation spillover, the effects of which may produce instability in the closed-loop system. In that paper, Balas recommended comb-filtering the measured output signals to pass only the frequencies in the modes to be controlled as a means of "eliminating" the observation spillover, thus assuring stability. Unfortunately, he provides no conditions on the filter that would *guarantee* the stability. Doyle and Stein (1979) discuss the closed-loop instability potential for systems whose dynamic models are merely parametrically uncertain (i.e., the system matrices are in error to within a constant). To date, there seems to be no simple, general approach to eliminating these interactions. These ideas, which they discuss in the context of LTI systems, would appear to be even more complicated with time-varying or nonlinear systems. Despite the invalidity of the separation principle with respect to the anticipated model uncertainty and nonlinearity, we will proceed conventionally and design the controller and the observer separately, and evaluate the resulting compensator via simulation, postponing further comment.

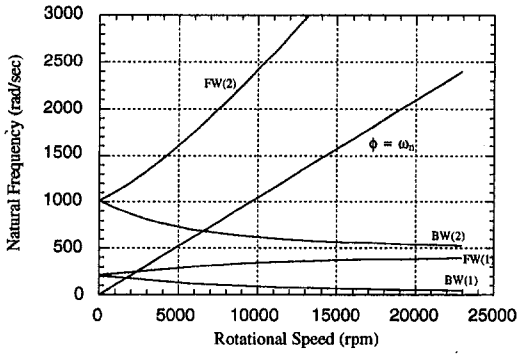


Fig. 3 Campbell diagram

With the state vector defined as $\mathbf{x} = [\mathbf{z}\ \dot{\mathbf{z}}\ \tilde{\mathbf{i}}]^T$, the controlled input is now $\tilde{\mathbf{e}}$. Following Slotine and Li (1990) or Fernandez and Hedrick (1987), one may derive a sliding mode controller for the system. Choosing a vertical and horizontal gap length error at each bearing as the system outputs to be regulated, the four independent control laws desired are easily produced in straightforward fashion. Combining Eqs. (1)–(4) with the seal loading forces not detailed here:

$$\dot{\mathbf{x}} = A(\dot{\phi})\mathbf{x} + B\mathbf{u} + \mathbf{D} \quad \mathbf{x} \in R^n, \quad \mathbf{u} \in R^p \qquad (5)$$

$$\bar{J}\ddot{\phi} = M_z \qquad (6)$$

where, for the matrices appropriately partitioned according to the defined state vector,

$$A(\dot{\phi})\mathbf{x} = \left[\begin{array}{cc|c} 0 & 1 & 0 \\ -M'^{-1}K' & -M'^{-1}C' & M'^{-1}\hat{K}_i \\ \hline 0 & & -\text{diag}[(R/L)_{i=1,4}] \end{array}\right]\begin{bmatrix} \mathbf{z} \\ \dot{\mathbf{z}} \\ \tilde{\mathbf{i}} \end{bmatrix}$$

$$B\mathbf{u} = \left[\begin{array}{c|c} 0 & 0 \\ \hline 0 & \text{diag}[(1/L)_{i=1,4}] \end{array}\right]\begin{bmatrix} \mathbf{0} \\ \tilde{\mathbf{e}} \end{bmatrix} \quad D = \begin{bmatrix} 0 \\ M'^{-1}(\hat{\mathbf{F}} + \hat{\mathbf{F}}') \\ 0 \end{bmatrix}$$

In the equations above, the primes denote the original rotordynamic matrices as modified by the aerodynamic and seal loading terms, and the elements of the bearing stiffness matrix K_D (or $\hat{K}_D$). $\hat{\mathbf{F}}$ is simply that from Eq. (1) minus the bearing, seal, and impeller forces. Since the intention is to invoke a linear speed of rotation law, and because the angular degree of freedom is decoupled from the remaining system, $\dot{\phi}$ has not been included in the state vector definition, although the dependence of the system matrix on $\dot{\phi}$ has been noted to remind the reader of the gyroscopic effects; of course, $\mathbf{D} = \mathbf{D}(\phi, \dot{\phi}, \ddot{\phi})$ and is also time-varying. Although real, effective process noise could easily be modeled and included in Eq. (5), this has presently been excluded for the generic problem at hand. Proceeding, then, with the control law(s), one may take a representative air gap length (left bearing, top) as $g_{\ell 1} = \Delta - R_{3Y}$. With an associated error of $\tilde{g}_{\ell 1} = -R_{3Y}$, one begins with the 15th row of Eq. (5), as dictated by the rotordynamic vector originally assumed and dimensioned for the reader.

$$\dot{\tilde{g}}_{\ell 1} = -\dot{R}_{3Y}$$

$$\ddot{\tilde{g}}_{\ell 1} = -\ddot{R}_{3Y} = M'^{-1}K'_{(15,:)}\mathbf{z} + M'^{-1}C'_{(15,:)}\dot{\mathbf{z}} - M'^{-1}\hat{K}_{i(15,:)}\tilde{\mathbf{i}} - M'^{-1}\hat{\mathbf{F}}_{(15)}$$

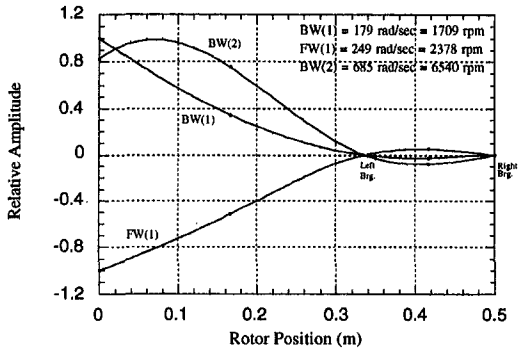


Fig. 4 Synchronous critical speed mode shapes

$$\ddot{\tilde{g}}_{\ell 1} = M'^{-1}K'_{(15,:)}\dot{\mathbf{z}} + M'^{-1}C'_{(15,:)}\ddot{\mathbf{z}} - M'^{-1}\hat{K}_{i(15,:)}\{-(R/L)_1\tilde{\mathbf{i}}_{\ell 1} + (1/L)_1\tilde{e}_{\ell 1}\} - M'^{-1}\hat{\mathbf{F}}_{(15)}$$
$$= \Psi_{\ell 1}(\dot{\mathbf{z}}, \ddot{\mathbf{z}}, \tilde{\mathbf{i}}, \dot{\hat{\mathbf{F}}}) - M'^{-1}\hat{K}_{i(15,:)}(1/L)_1\tilde{\mathbf{e}}_{\ell 1}$$
$$= \Psi_{\ell 1}(\mathbf{x}, \phi, \dot{\phi}, \ddot{\phi}) - M'^{-1}\hat{K}_{i(15,:)}(1/L)_1\tilde{e}_{\ell 1}$$

where, for example, $M'^{-1}K'_{(15,:)}$ indicates 15th row, all columns of $M'^{-1}K'$, and the substitution of $\ddot{\mathbf{z}} = f(\mathbf{z}, \dot{\mathbf{z}})$ from the system dynamics has been left implicit above. Since the input voltage may control the third derivative of the respective gap length error, the relative degree is $3 - 1 = 2$, and one may define the time-varying sliding surface in the state space as:

$$s_{\ell 1}(\mathbf{x}, t) = \left(\frac{d}{dt} + \lambda\right)^{3-1}\tilde{g}_{\ell 1} = \ddot{\tilde{g}}_{\ell 1} + 2\lambda\dot{\tilde{g}}_{\ell 1} + \lambda^2\tilde{g}_{\ell 1} \quad (7)$$

Differentiating the surface once recovers the control input,

$$\dot{s}_{\ell 1} = \dddot{\tilde{g}}_{\ell 1} + 2\lambda\ddot{\tilde{g}}_{\ell 1} + \lambda^2\dot{\tilde{g}}_{\ell 1}$$
$$= \Psi_{\ell 1}(\mathbf{x}, \phi, \dot{\phi}, \ddot{\phi}) - M'^{-1}\hat{K}_{i(15,:)}(1/L)_1 e_{\ell 1} + 2\lambda(-\ddot{R}_{3Y}) + \lambda^2(-\dot{R}_{3Y})$$

Therefore one may choose:

$$\tilde{e}_{\ell 1} = (L_1/M'^{-1}\hat{K}_{i(15,:)})\{\Psi_{\ell 1}(\mathbf{x}, \phi, \dot{\phi}, \ddot{\phi}) + 2\lambda(-\ddot{R}_{3Y}) + \lambda^2(-\dot{R}_{3Y}) + \xi_{\ell 1}\tanh(\upsilon s_{\ell 1})\} \quad (8)$$

as the control law to provide an attractive surface. That is, $\dot{s}_{\ell 1} = -\xi_{\ell 1}\tanh(\upsilon s_{\ell 1})$ (where υ represents a tuning parameter for the tanh $(\cdot)$ function) insures stable attraction for uncertain disturbances bounded by ξ, which may be constant, or perhaps a function of $\dot{\phi}$. The three remaining gap length error control laws are similarly derived by substituting the appropriate row index into the indicated matrix algebraic operations. In this procedure, a nondiagonal inertia matrix, as with FEM techniques, will distribute the control input $\tilde{e}$'s among all four gap length error equations, and a simultaneous solution of the terms must be sought.

Eigenanalysis

Substitution of the four (regulating) control laws, as with Eq. (8), will uncouple the respective bearing translational degree of freedom (e.g., R_{3Y}) from the remaining system by virtue of the feedback cancellation term under the restriction of no model uncertainty. Of course, with model imprecision one is left with a robust nonlinear regulator, which intentionally maintains the bearing journals near their nominally zero displacement (rigid bearing) positions. Thus, taking four independent gap length errors as $\tilde{\mathbf{g}} = -[R_{3Y}\ R_{5Y}\ R_{3X}\ R_{5X}]^T$, one removes the 5th, 9th, 15th, and 19th rows and columns of submatrix A_{11} in Eq. (5), in accordance with the system matrix as assembled here, and considers the homogeneous system. Taking $\boldsymbol{\chi}_r = [\mathbf{z}_r\ \dot{\mathbf{z}}_r]^T$ one has the first-order eigenvalue form:

$$\dot{\boldsymbol{\chi}}_r - \begin{bmatrix} 0_r & I_r \\ -M_r'^{-1}K_r' & -M_r'^{-1}C_r' \end{bmatrix}\boldsymbol{\chi}_r = \mathbf{0} \quad (9)$$

The Campbell diagram is illustrated in Fig. 3, while the first three synchronous whirl critical speeds and their respective mode shapes are provided in Fig. 4.

Kalman Filter

Expanding Eq. (5), one may now express the output relationship:

$$\dot{\mathbf{x}} = A_{(\phi)}\mathbf{x} + B\mathbf{u} + \mathbf{D} + G\mathbf{w}$$
$$\mathbf{x} \in R^n, \quad \mathbf{u} \in R^p, \quad \mathbf{w} \in R^n \quad (10)$$
$$\mathbf{y} = H\mathbf{x} + \mathbf{v}, \quad \mathbf{y}, \mathbf{v} \in R^m \quad (11)$$
$$E[\mathbf{w}\mathbf{w}^T] = Q_f, \quad E[\mathbf{v}\mathbf{v}^T] = R_f$$

Cross-correlation of the noise vectors is assumed to be zero, and the reader is reminded that the process noise indicated in Eq. (10) is artificial, and its spectral density matrix is, of course, taken for the sake of calculation accuracy and filter performance. As was earlier indicated, no actual process noise was implemented with the dynamic model of Eq. (5). R_f was chosen

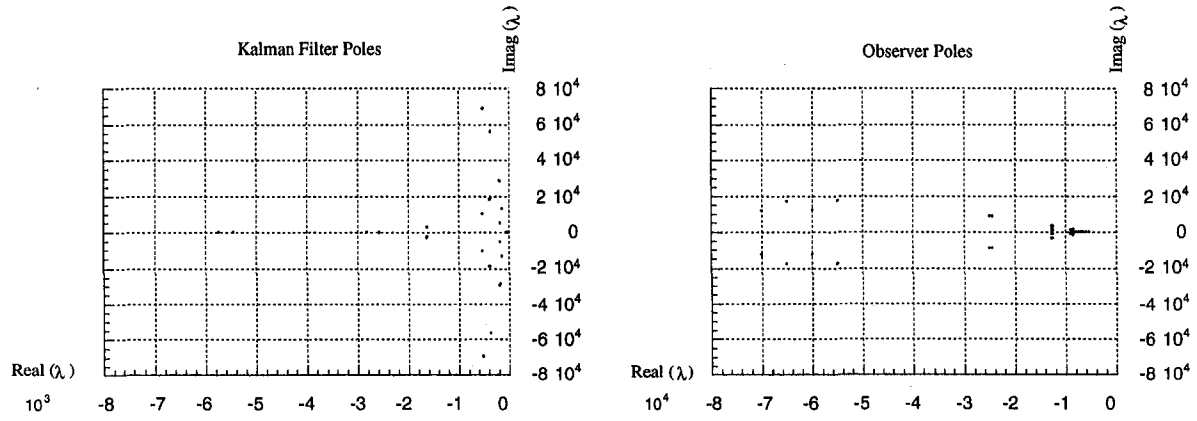


Fig. 5 Estimator pole locations

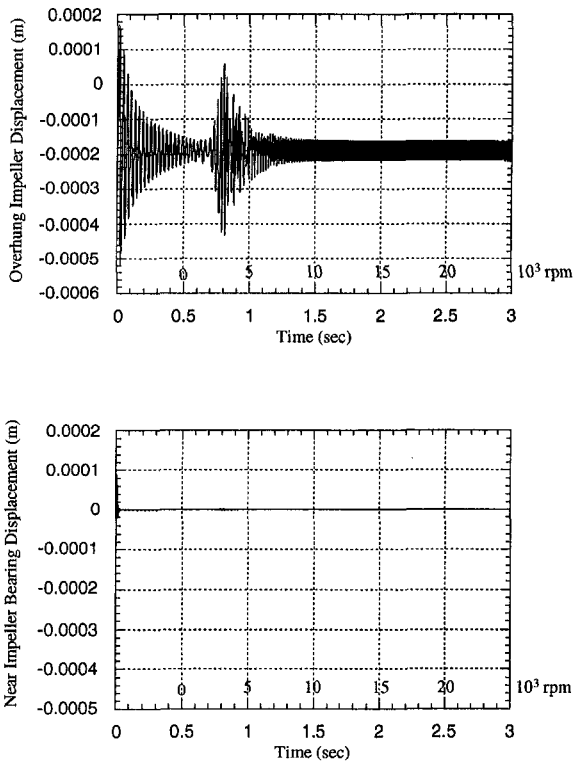


Fig. 6 Vibration behavior: Kalman filter

as the $(m \times m)$ diagonal matrix of the appropriate position and velocity noise statistics generated in the model; that is the noise introduced in the feedback signals of the four independent bearing journal positions and velocities stemming from the nonlinear controller. Q_f was taken to be a banded, nondiagonal matrix as was found to be appropriate for best filter performance. The stationary, or algebraic Ricatti equation (ARE) was solved:

$$\dot{P} = 0 = AP + PA^T + Q_f - PH^T R_f^{-1} HP \qquad (12)$$

and in straightforward manner, the state estimates using the stationary Kalman gain matrix L were retrieved from:

$$\dot{\hat{\mathbf{x}}} = A_{(\phi)}\hat{\mathbf{x}} + B\mathbf{u} + \mathbf{D} + L(\mathbf{y} - H\hat{\mathbf{x}}) \qquad (13)$$

where, of course, $L = PH^T R_f^{-1}$, and the argument of the system matrix A in Eq. (13) has been reinserted to emphasize the rotational speed dependence.

Guided by Shamma and Athans (1992) and Shahruz and Behtash (1992), who present recent and insightful treatments of gain scheduling, the Kalman gain matrix was calculated over the relevant speed range and scheduled to a smooth function of the measured rotational speed signal. Of course, a variety of methods can be used actually to schedule the feedback matrices between the operating points at which they are calculated in the linear parameter varying (LPV) paradigm. Here, curve-fitting to a cubic polynomial of rotational speed was favored over linear interpolation. While Shahruz and Behtash (1992) provide explicit upper bounds on the speed of variation tolerable in the scheduled parameter, it can also be shown that the rate of variation of the rotational speed (the rotational acceleration) obeys the Lyapunov stability criterion. With

$$\dot{\mathbf{x}} = A_{c(\phi)}\mathbf{x} \qquad (14)$$

where $A_{c(\phi)}$ implies the closed-loop estimator matrix, one finds the positive definite solution $\Xi_{(\phi)}$ of the matrix Lyapunov equation, $\Xi_{(\phi)}A_{c(\phi)} + A_{c(\phi)}^T\Xi_{(\phi)} = -I$, where the specific case of the identity matrix has been assumed as a suitable choice of an arbitrary p.d. symmetric matrix on the right-hand side. The Lyapunov function is taken as $V = \mathbf{x}^T\Xi_{(\phi)}\mathbf{x}$, whose negative definite time derivative may be verified from:

$$\dot{V} = \dot{\mathbf{x}}^T\Xi_{(\phi)}\mathbf{x} + \mathbf{x}^T\Xi_{(\phi)}\dot{\mathbf{x}} + \dot{\phi}\mathbf{x}^T\Xi'_{(\phi)}\mathbf{x} \overset{?}{<} 0 \qquad (15)$$

Or, by substituting Eq. (14) into Eq. (15) appropriately (and where $\Xi'_{(\phi)} = \partial\Xi_{(\phi)}/\partial\phi$),

$$\mathbf{x}^T(-I + \ddot{\phi}\Xi'_{(\phi)})\mathbf{x} \overset{?}{<} 0 \qquad (16)$$

Thus, the stability of the time-varying closed-loop estimator is addressed by evaluating, numerically, the negative definiteness of $(-I + \ddot{\phi}\Xi'_{(\phi)})$ for the cubic polynomial scheduling of the Kalman filter matrices subject to the known time rate of change of rotational speed.

In the realm of linear systems, and linear multivariable observers in particular, the heuristic approach still seems to hold that the slowest observer poles should be anywhere from 5–10 times the speed of the fastest closed-loop pole (see Lewis, 1992, among others). Still, Luenberger's original position was that there was little need for observer poles much faster than the controller poles. With regard to nonlinear, sliding mode controllers, Bondarev et al. (1985) detail the positive benefits and sometime necessity of using observers in sliding mode systems, but note that for asymptotic observers, the rates of motion in the observer should be an order higher than the rates of motion of the state vector. Of course, this appears intuitive and a previous compensator whose gain-scheduled observer poles were placed via eigenstructure assignment techniques to meet just such an objective has been considered (Smith, 1995). Its performance will be compared with that of the present design, relying as it does on the Kalman filter for measurement noise accommodation, but at the potential performance expense of comparatively slow estimator poles. The pole locations of the respective estimators are shown in Fig. 5. LOG controllers have previously been considered for magnetic bearings (Matsumura and Yoshi-

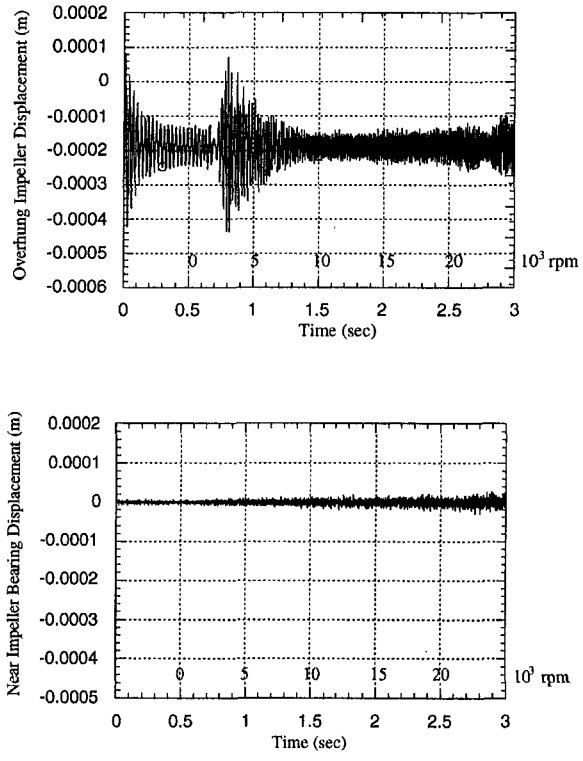


Fig. 7 Vibration behavior: multivariable observer

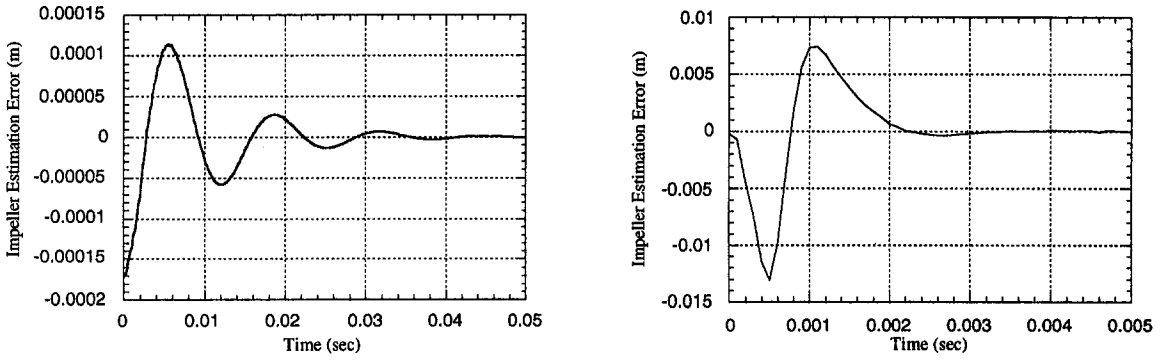


Fig. 8 Comparative error dynamics

moto, 1986; Maslen, 1991, etc.), but the convergence of the estimator is never a theoretical problem for perfect linear models obeying the separation principle.

Numerical Simulation

The exhorbitant data for the system description have not been included for the sake of brevity, but Figs. 6 and 7 indicate the performance comparison for the same controller implemented with the Kalman filter and a multivariable Luenberger observer for identical levels of position variance and velocity variance in the ideally assumed measurements. The details of the observer may be found in Smith (1995). The overhung impeller and the near-impeller bearing displacement are shown. For the Kalman filter implementation, +40 percent was the maximum tolerable increase in the aerodynamic cross-coupling term over its nominal design value. For the multivariable Luenberger observer, this term was taken to 55 percent above its nominal value before serious performance degradation occurred. The respective (convergent) errors in the impeller estimate for both the filter and the observer indicate the "fast dynamics" design criterion that the observer was based on, yet the comparatively slow poles of the Kalman filter have not apparently hindered the sliding mode action that the implemented cross-coupled stiffness variation would have relied upon, and that Bondarev et al. (1985) recommend.

Acknowledgments

This work was supported by the University of Texas Center for Electromechanics (UT-CEM), and funded through an ARO grant under the AASERT program.

Conclusions

Rotordynamic loading effects from seals, fluid interaction forces, etc., are not often directly addressed in contemporary discussions of magnetic bearing control system design. The recent paper by Lang et al. (1995) strikes at the heart of fundamental issues of potentially unstable uncertainties. Although they dealt with a rigid rotor in their parameter estimation algorithm, the more difficult issue to tackle in adaptive control might be the genesis and reliability of the signal matrix itself. In this paper, the attempt was merely to focus attention on the cross-coupled stiffness terms arising from both aerodynamic loading of compressor impellers and annular gas seal forces, and this has been done by recourse to a hypothetical single-stage centrifugal gas compressor of the overhung design. A nonlinear controller was successfully implemented through the stationary Kalman filter, gain-scheduled for rotational speed, and theoretically examined for sensitivity to the aerodynamic loading in the presence of measurement noise. Compared to its counterpart, the multivariable Luenberger observer, the Kalman filter performs well, although admits of a slightly lower toleration for modeling error in the cross-coupled stiffness term, while predictably smoothing out the noise. Of course, a comprehensive evaluation of both compensators in the face of the competing needs of robustness and low noise sensitivity to an array of possible variations requires a more thorough and deliberate effort.

References

Allaire, P. E., Lewis, D. W., and Knight, J. D., 1983, "Active Vibration Control of a Single Mass Rotor on Flexible Supports," *Journal of the Franklin Institute,* Vol. 315, No. 3, pp. 211–222.

Anton, E., and Ulbrich, H., 1985, "Active Control of Vibrations in the Case of Asymmetrical High-Speed Rotors by Using Magnetic Bearings," ASME *Journal of Vibration, Acoustics, Stress, and Reliability in Design,* Paper No. 85-DET-28.

Balas, M. J., 1978, "Feedback Control of Flexible Systems," *IEEE Transactions on Automatic Control,* Vol. AC-23, No. 4, pp. 673–679.

Bondarev, A. G., Bondarev, S. A., Kostylyova, N. Y., and Utkin, V. I., 1985, "Sliding Modes in Systems With Asymptotic State Observers," *Automatic and Remote Control,* Vol. 6.

Childs, D. W., and Scharrer, J. K., 1986, "An Iwatsubo-Based Solution for Labyrinth Seals: Comparison to Experimental Results," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 108, pp. 325–331.

Childs, D. W., 1993, *Turbomachinery Rotordynamics: Phenomena, Modeling, and Analysis,* Wiley Interscience, New York.

Doyle, J. C., and Stein, G., 1979, "Robustness With Observers," *IEEE Transactions on Automatic Control,* Vol. AC-24, No. 4, pp. 607–611.

Fernandez, Benito R., and Hedrick, J. Karl, 1987, "Control of Multivariable Non-linear Systems by the Sliding Mode Method," *International Journal of Control,* Vol. 46, No. 3, pp. 1019–1040.

Gopinath, G., 1971, "On the Control of Linear Multiple Input-Output Systems," *Bell Systems Technical Journal,* Vol. 50, pp. 1063–1081.

Hebbale, K. V., 1985, "A Theoretical Model for the Study of Nonlinear Dynamics of Magnetic Bearings," Ph.D. Dissertation, Cornell Univ., Jan.

Humphris, R. R., Kelm, R. D., Lewis, D. W., and Allaire, P. E., 1986, "Effect of Control Algorithms on Magnetic Journal Bearing Properties," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 108, pp. 624–632.

Lang, O., Wassermann, J., and Springer, H., 1995, "Adaptive Vibration Control of a Rigid Rotor by Active Magnetic Bearings," ASME Paper No. 95-GT-181.

Lewis, D. W., Maslen, E. H., and Humphris, R. R., 1988, "Dynamic Response of a Magnetically Supported Rotor," *Magnetic Bearings: Proc. of the 1st International Symposium,* June, pp. 263–272.

Lewis, Frank L., 1992, *Applied Optimal Control and Estimation,* Prentice-Hall, New York.

Maslen, E. H., 1991, "Magnetic Bearing Synthesis for Rotating Machinery," Ph.D. Dissertation, Univ. of Virginia, Charlottesville, VA, Jan.

Matsumura, F., and Yoshimoto, T., 1986, "System Modeling and Control Design of a Horizontal-Shaft Magnetic-Bearing System," *IEEE Trans. on Magnetics,* Vol. MAG-22, May.

Mizuno, T., and Higuchi, T., 1984, "Compensation for Unbalance in Magnetic Bearing Systems," *Trans. Soc. Instrum. and Control Engineering (Japan),* Vol. 20, No. 12, pp. 1095–1101.

Schweitzer, G., 1988, "Magnetic Bearings," in: *Rotordynamics 2: Problems in Turbomachinery,* N. F. Rieger, ed., Springer-Verlag, New York.

Shahruz, S. M., and Behtash, S., 1992, "Design of Controllers for Linear Parameter Varying Systems by the Gain Scheduling Technique," *Journal of Mathematical Analysis and Applications,* Vol. 168, No. 1, pp. 195–217.

Shamma, Jeff S., and Athans, M., "Gain Scheduling: Potential Hazards and Possible Remedies," *IEEE Control Systems Magazine,* June, pp. 101–107.

Slotine, J. J., and Weiping Li, 1990, *Applied Nonlinear Control,* Prentice-Hall, New York.

Smith, R. D., 1995, "Robust Compensation of Actively Controlled Bearings: A Performance Comparison of Regulation Methods," Ph.D. Dissertation, University of Texas (Austin), Dec.

Sortore, C. K., 1990, "Design of Permanent Magnet Biased Magnetic Bearings for a High Speed Rotor," M.S. Thesis, Univ. of Virginia, Jan.

Thieleke, G., and Stetter, H., 1990, "Experimental Investigations of Exciting Forces Caused by Flow in Labyrinth Seals," presented at the 1990 Rotordynamic Instability Workshop, Texas A&M University.

Thieleke, G., and Stetter, H., 1992, "Identification of Friction Factors for Modelling the Exciting Forces Caused by Flows in Labyrinth Seals," *Rotordynamics '92, Proceedings of the International Conference on Rotating Machine Dynamics,* Venice, Apr. 28–30.

Traxler, A., 1985, "Eigenschaften und Auslegung berührungsfreier elektromagnetischer Lager," Diss., ETH Zurich, No. 7851.

Wachel, J. C., and von Nimitz, W. W., 1981, "Ensuring the Reliability of Offshore Gas Compression Systems," *Journal of Petroleum Technology,* Nov. pp. 2252–2260.

Test Results of a New Damper Seal for Vibration Reduction in Turbomachinery

J. M. Vance

J. Li

Department of Mechanical Engineering,
Texas A&M University,
College Station, TX 77843

A new type of labyrinth gas seal for damping vibration and whirl, called the TAMSEAL, has been evaluated in both nonrotating and rotating tests at Texas A&M University. Test results of the prototype, along with comparison tests of a conventional labyrinth seal, show up to 100 times more direct damping than the conventional bladed seal. The new design also has a feature that blocks swirl of the working fluid, which is known to be rotordynamically destabilizing in machines with conventional seals. Coastdown tests of the new seal were conducted at various pressures on a rotordynamic test apparatus with a critical speed at 4000 rpm and compared with identical testing of a conventional labyrinth seal. Rap tests of both seals were also conducted to measure the logarithmic decrement of free vibration, and the leakage of both seals was measured. Test results show large reductions in peak vibration at the critical speed in all cases, with the critical speed being completely eliminated by the TAMSEAL at some pressure drop conditions. The leakage rate of the tested TAMSEAL is higher than the conventional seal at the same clearance, but the large reductions in vibration and whirl amplitudes suggest that the TAMSEAL could be operated with smaller clearances than conventional labyrinth seals.

Description of the TAMSEAL and Test Rig

The most common type of pressure seal in turbomachinery is the labyrinth seal. A labyrinth seal consists of a series of circular blades and annular grooves that present a tortuous path for leakage of the working fluid. The TAMSEAL is a revolutionary development of the teeth on stator labyrinth seal. There are two distinct features of the TAMSEAL that work together to produce high damping. The first distinct feature is the radial rotor-to-blade clearance that increases along the axial flow path of the gas. The second distinct feature is placement of segregating partitions in the annular grooves of the seal. The combination of these two features is unique and has never been incorporated in any other seal design. A computer code predicts that the optimal blade number for damping is 2; increasing blade number will decrease both leakage and damping. Figure 1 is a cross-sectional drawing of a two-bladed configuration with four circumferential pockets.

Test results from a nonrotating apparatus were reported in 1993 (Vance and Schultz). This paper reports rotating tests. Figure 2 is a cross-sectional sketch of the rotordynamic test apparatus. The seal journal is attached to the end of the shaft. The journal is enclosed inside a sealed housing. Actually there are two identical TAMSEALs installed back-to-back around the overhung journal disk. Pressured air is supplied to a plenum chamber between the two seals. There are two leakage flows going from inside out in opposite directions and exhausting to the atmosphere. This cancels the axial thrust load and precludes the necessity for any ancillary seals that could pollute the test results. The seal journal diameter is 101.6 mm (4.000 in.). The upstream and downstream radial clearances of the tested damper seal are 0.102 mm (0.004 in) and 0.203 mm (0.008 in.), respectively, while the configuration of the conventional seal is straight with 0.102 mm (0.004 in.) radial clearance. The first critical speed of the rotor-ball bearing support assembly is 3960 rpm (66 Hz) with no seal installed.

Instrumentation consists of two X-Y Bently proximity probes mounted on the bearing housing near the overhung end of the rotor. An air pressure transducer is installed into the inlet plenum between the seals to measure the inlet pressure. There is a flowmeter installed in the air supply line to measure the total mass flow rate (leakage). Rotor speed is measured by an optical pickup. Coastdown curve measurements are made by acquiring vibration displacement data from a Bently-Nevada synchronous tracking filter. An HP 3561A signal analyzer is used for logarithmic decrement measurements when conducting rap tests.

Theory of Operation

Alford (1965) published a theory to predict direct stiffness and damping coefficients of a labyrinth seal with two blades and choked flow. The most notable prediction of Alford's analysis was that seals with clearances converging in the direction of flow would have negative damping, and seals with diverging clearances would have positive damping. The predicted coefficients are quite large. This theory modeled only the axial flow and neglected circumferential flow effects. Ten years of testing labyrinth seals at TAMU by Childs and Scharrer (1988), and by Vance and Schulz (1993), has shown that the direct damping coefficients of conventional labyrinth seals are very small, even with diverging clearances. Instead, the dominant rotordynamic coefficient of conventional seals is cross-coupled stiffness, which reduces the effective damping and can be destabilizing to rotor whirl. The reason for this failure of Alford's theory is that conventional labyrinth seals have continuous and unobstructed annular grooves so that pressure variations across the seal diameter cannot exist without large circumferential flow rates. Unfortunately, circumferential flow induces the destabilizing force represented by cross-coupled stiffness. The TAMSEAL design eliminates or greatly reduces the circumferential flow.

Iwatsubo (1980) made photographs through a transparent labyrinth seal wall that reveals a very complex flow pattern, containing vortices around the annulus. The TAMSEAL design makes this complex flow disappear, or at least makes it of

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 11, 1995. Paper No. 95-GT-36. Associate Technical Editor: C. J. Russo.

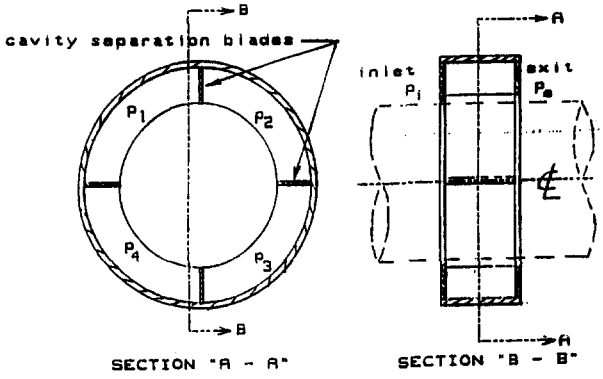


Fig. 1 Two-bladed damper seal

secondary importance relative to the axial flow in and out the cavities. The current theory for the new damper seal design is therefore based on an assumption that the average pressure in the seal cavities can be calculated based on gross mass flow rates into and out of the cavities. Local flow fields within the cavities are ignored. This theory is mathematically the same as that developed by Sundararajan and Vance (1993) for a gas-operated bearing damper also invented at Texas A&M University. It is important to note that the new damper seal is unlike all other damping devices commonly used, in that it does not rely on viscosity of a fluid to dissipate energy.

A physical description of TAMSEAL phenomenon proceeds as follows:

Refer to Fig. 1 and assume counterclockwise whirling of a rotor within the clearance circle of section "A-A." Consider the seal journal in the three o'clock position and moving up. The blade clearances adjacent to the p_3 and p_4 cavities are opening up. The clearances adjacent to p_1 and p_2 are closing. The percent rate of change of inlet area (at Pi on section "*B-B*") is greater than the percent rate of change of exit areas (at Pe on section "*B-B*"). The pressures in all cavities will therefore change with time, at the same frequency as the shaft whirl speed. It turns out that the dynamic pressure in each cavity leads the whirl displacement vector by a phase angle. If this phase angle is 90 deg, then the pressure force on the rotor will continuously oppose the rotor velocity. This is damping, by definition. The actual phase angle is frequency dependent and deviates considerably from 90 deg at high frequencies. The deviation from 90 deg reduces the damping and produces a negative stiffness coefficient that is small compared to typical shaft and bearing support stiffness in turbomachinery.

Test Results and Comparison

The inlet air pressure to the seals was varied from 1 bar (14.7 psia) to 3.4 bar (50 psia) for the TAMSEAL, and from 1 bar (14.7 psia) to 4.4 bar (65 psia) for the conventional seal exhausting to the atmosphere. The inlet temperature was 23°C (74°F). The synchronous response to imbalance was measured by running the rotor up to 6000 rpm, then decreasing the speed at a uniform rate, and recording the *X* and *Y* probe amplitudes as a function of the running speed. Figures 3 and 4 show how the maximum peak-to-peak amplitudes at the critical speed vary with inlet pressures. The coastdown results for a number of different inlet pressures are shown in Figs. 5–10, respectively. The coastdown results show that the TAMSEAL has much more effective damping than the conventional seal. The TAMSEAL always has positive effective damping for imbalance response while the conventional seal has negative effective damping in the vertical direction. When inlet air pressure is above 3.1 bar (45 psia), the maximum amplitude of imbalance response for the TAMSEAL is decreased greatly, so that the coastdown curves are flat (no peak) in the horizontal direction. The effects of the TAMSEAL and the conventional seal on the critical speed are different from each other. The critical speed is raised with inlet air pressure for the conventional labyrinth seal. For the TAMSEAL, however, the critical speed first is lowered with inlet air pressure and then is raised with inlet pressure above 40 psia.

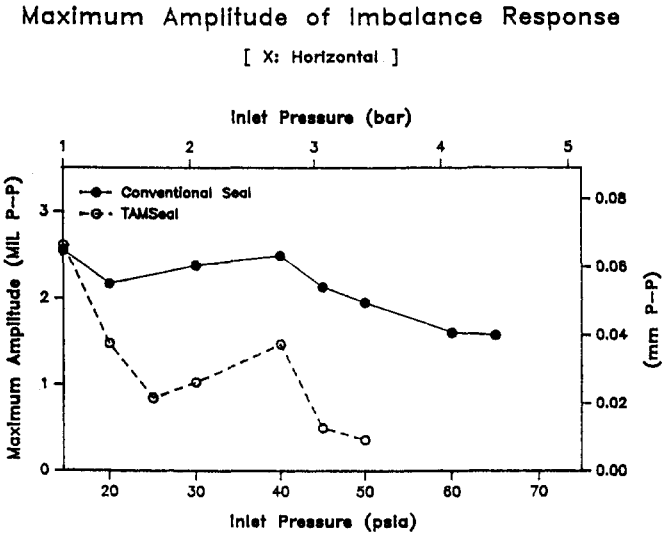


Fig. 3 Critical speed amplitude versus inlet pressure

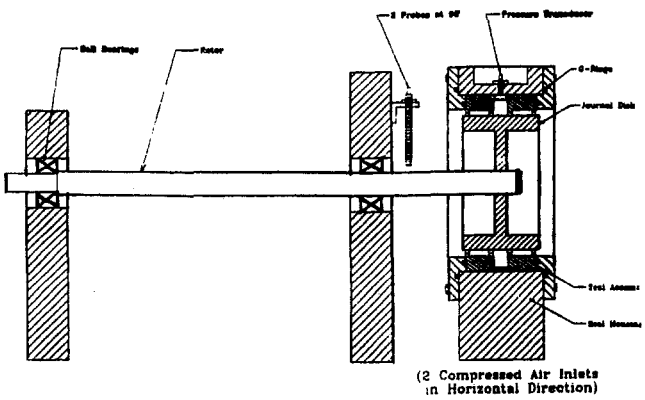


Fig. 2 Cross section of damper seal test rig

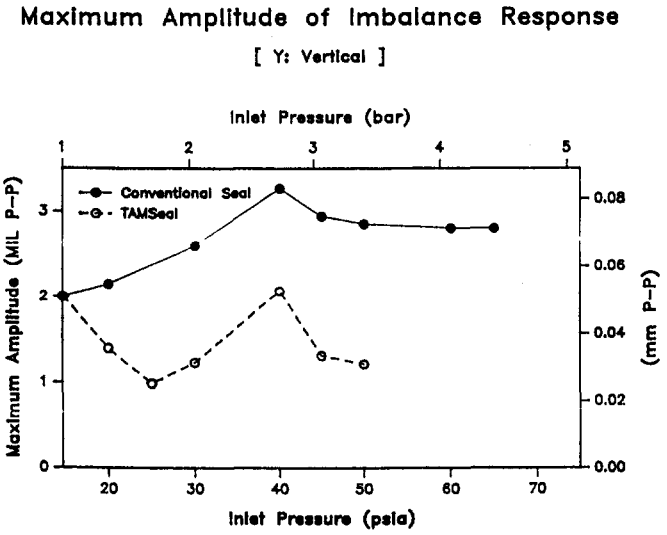


Fig. 4 Critical speed amplitude versus inlet pressure

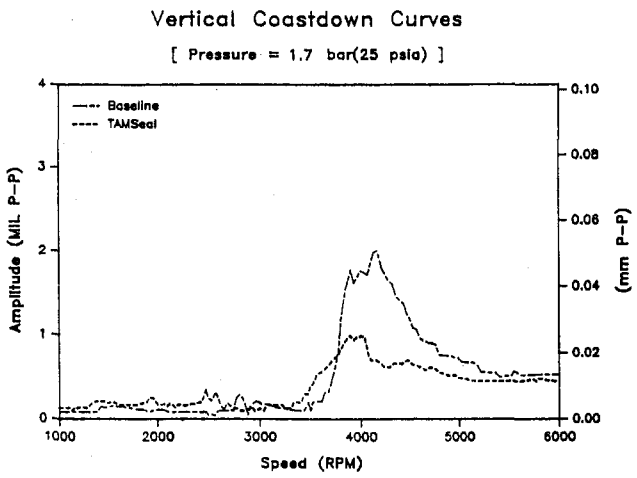


Fig. 5 Vertical imbalance response (1.7 bar)

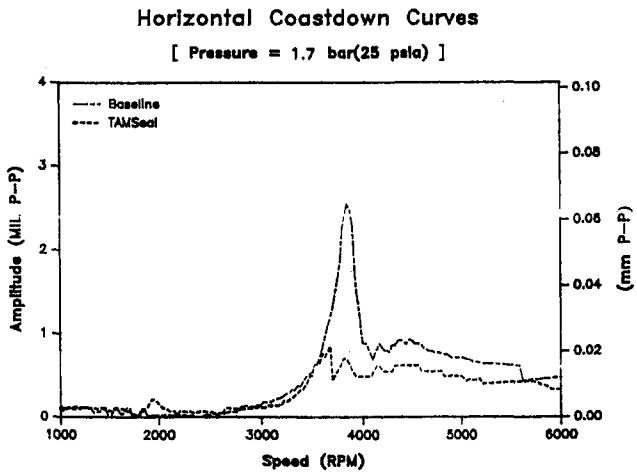


Fig. 6 Horizontal imbalance response (1.7 bar)

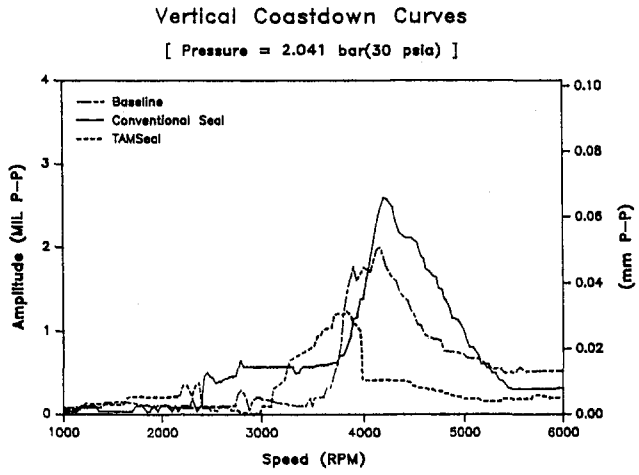


Fig. 7 Vertical imbalance response (2.041 bar)

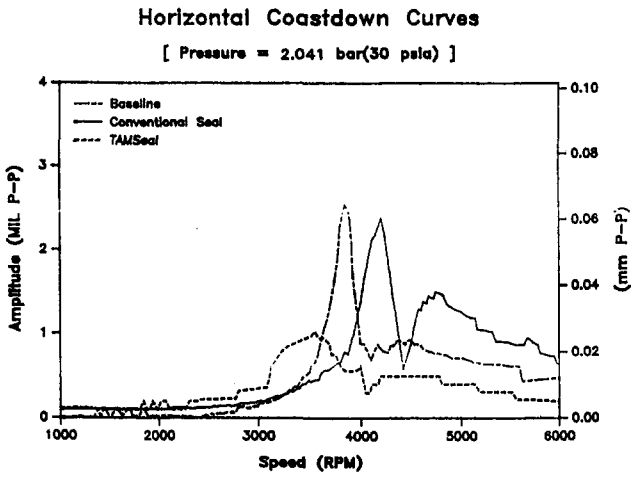


Fig. 8 Horizontal imbalance response (2.041 bar)

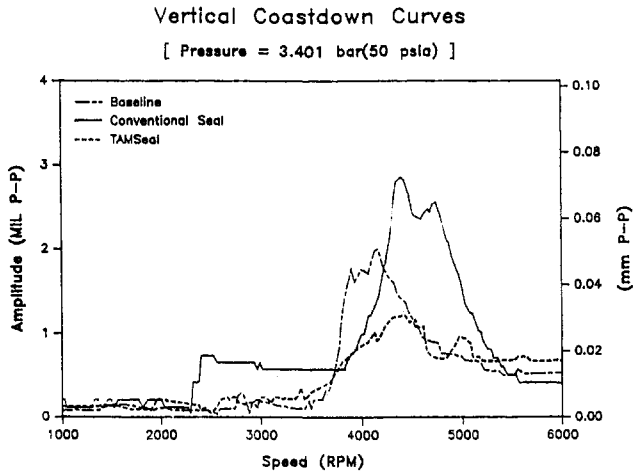


Fig. 9 Vertical imbalance response (3.401 bar)

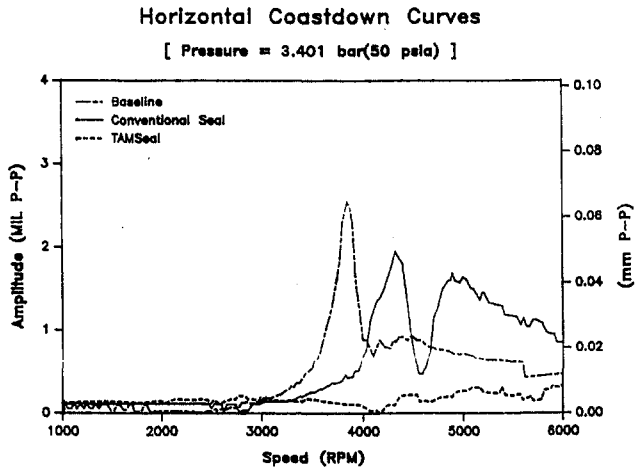


Fig. 10 Horizontal imbalance response (3.401 bar)

Rap tests of both seals were conducted with the rotor at rest for several different inlet pressures. Examples of the decaying vibration of rap tests of the TAMSEAL compared with the conventional labyrinth seal are shown in Figs. 11 and 12. The rapidly decaying vibrations are with the TAMSEAL. Figure 13 shows logarithmic decrements extracted from these wave forms. When increasing the inlet air pressure to the seal, the damper seal was found to increase the logarithmic decrement sharply, while the conventional seal decreased the logarithmic decrement.

The mass leakage rates of the TAMSEAL and the conventional seal were also measured under identical conditions. The

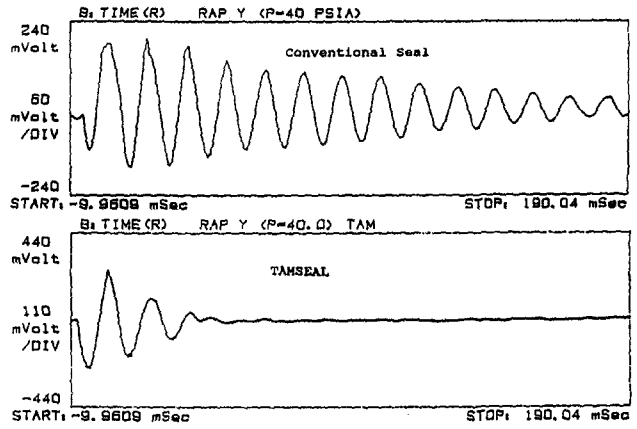


Fig. 11 Rap test response (2.721 bar, vertical)

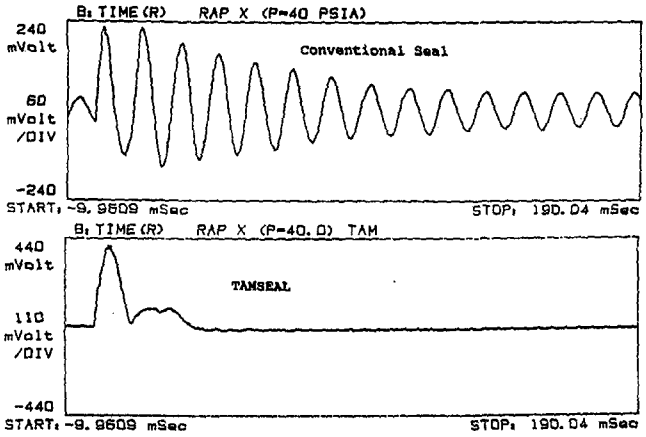


Fig. 12 Rap test response (2.721 bar, horizontal)

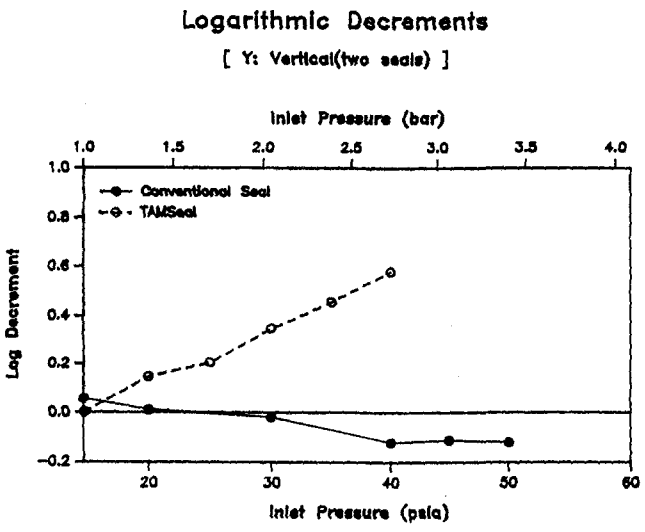


Fig. 13 Logarithmic decrement (vertical direction)

comparison of the leakage of both seals is shown in Fig. 14. The leakage of the TAMSEAL is 1.30 times that of the conventional seal. The relationship between leakage and inlet air pressure is approximately linear for both seals, but the leakage

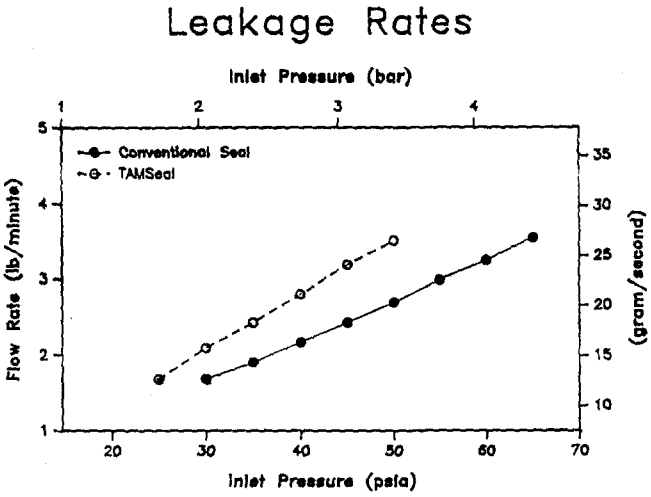


Fig. 14 Leakage rates versus inlet pressure

rate of the TAMSEAL is increased more quickly with inlet air pressure than that of the conventional seal.

Conclusion

1 The TAMSEAL has much more equivalent damping for imbalance response than a conventional labyrinth seal with identical dimension, including whatever cross-coupled stiffness results from shaft rotation. Equivalent damping is indicated by reduction of the vibration amplitude at the critical speed.

2 For low inlet air pressure (2.7 bar and less), the TAMSEAL lowers the first critical speed frequency. When inlet air pressure is above 2.7 bar (40 psia), the TAMSEAL increases the first critical speed frequency. This is in contrast to the conventional labyrinth seal, which was found to always raise the critical speed.

3 For free vibration, the TAMSEAL has positive effective damping and increases the logarithmic decrement rapidly with inlet air pressure, while the effective damping of the conventional seal is not only very small, but becomes negative for inlet air pressure exceeding 2 bar (30 psia). When inlet air pressure is above 2.7 bar (40 psia), the effective damping of the TAMSEAL approaches the critical damping.

4 The leakage rate of the TAMSEAL is about 30 percent larger than that of a same dimension conventional labyrinth seal under identical conditions. The large damping effect of the TAMSEAL suggests that smaller clearances could be used in many applications, which would reduce the leakage. (Additional testing has shown that smaller clearances also produce higher damping.)

References

Alford, J. S., 1965, "Protecting Turbomachinery From Self-Excited Whirl," ASME JOURNAL OF ENGINEERING FOR POWER, Vol. 87, pp. 333–344.

Childs, D. W., and Scharrer, J. K., 1988, "Theory Versus Experiment for the Rotordynamic Coefficients of Labyrinth Gas Seals, Part II—A Comparison to Experiment," *Journal of Vibration, Acoustics, Stress, and Reliability in Design,* Vol. 110, pp. 281–287.

Iwatsubo, T., 1980, "Evaluation of the Instability Forces of Labyrinth Seals in Turbines or Compressors," NASA CP 2133, Workshop on Rotordynamic Instability Problems in High Performance Turbomachinery, College Station, TX, May 12–14.

Sundararajan, P., and Vance, J. M., 1993, "A Theoretical and Experimental Investigation of a Gas-Operated Bearing Damper for Turbomachinery—Part I: Theoretical Model and Predictions," *Proc. 1993 ASME Vibration and Noise Conference,* Albuquerque, NM.

Vance, J. M., and Schultz, R. R., 1993, "A New Damper Seal for Turbomachinery," *Vibration of Rotating Systems,* ASME DE-Vol. 60.

The Development of Life Prediction Techniques for Structural Ceramics

P. K. Khandelwal

N. J. Provenzano

W. E. Schneider

Allison Engine Company,
Indianapolis, IN 46206

One of the major challenges involved in the use of ceramic materials in advanced vehicular heat engines is ensuring adequate strength and durability. This Department of Energy supported activity has developed methodologies to predict the structural behavior of ceramic components. The effort involved the characterization of injection-molded and hot isostatic pressed PY6 silicon nitride and the development of analytical life prediction techniques. Three failure modes are addressed: fast fracture, slow crack growth, and creep rupture. The technique deals with surface as well as internal component failures. The life prediction methodologies for fast fracture and slow crack growth have been verified using two types of confirmatory specimens: (1) flat circular disks subjected to bending stresses, and (2) high-speed rotating spin disks. Correlation was achieved for a variety of test conditions and failure mechanisms. The predictions associated with surface failures proved to be optimistic, requiring re-evaluation of the components' initial fast fracture strength. Correlation was achieved for the spin disks that failed in fast fracture from internal flaws. Time-dependent, elevated-temperature spin disk failures were also successfully predicted.

Introduction

The goals associated with this effort are to develop and demonstrate:

- design methodology, which can predict the useful life of ceramic components
- a ceramic material database appropriate for use in the design of engine components

To accomplish the goals, three major tasks were involved:

- material characterization and database generation
- development of analytical life prediction models
- verification

The characterization of the PY6 material involves room and elevated temperature testing of a variety of specimens. The fast fracture and slow crack growth testing used modulus-of-rupture (MOR), button-head, and dogbone tensile specimens. The creep/stress rupture testing employed the button-head and dogbone tensile specimens. Fractography of representative samples from each of the tests was conducted to ensure adequate knowledge of the failure modes and mechanisms.

The development of the analytical life prediction models builds on the knowledge and information gained during the materials characterization phase. A series of statistical models has addressed each of the failure modes. Both uniaxial and biaxial models have been developed and correlated with the material test data. The fast fracture model is based on a two-parameter Weibull distribution, which is the current industry standard, except the proposed model utilizes a different distribution depending on the specimen/component failure origin. Separate durability models have been developed for slow crack growth and creep. Both models have been enhanced and time-dependent probability of survival is predicted such that the probability of failure increases with time. In addition to the life prediction models, a creep deformation model has been developed that couples the creep rate to the stress rupture information. The life prediction models developed by this program have been incorporated into the NASA software, CARES [1].

The final phase involves additional tests to confirm the adequacy of the durability models and establishes suitable safety factors for use in component design. The verification tests used flat circular disks subjected to bending loads and high-speed spin tests. The confirmatory tests were conducted at room and elevated temperatures up to a maximum of 1400°C. Pretest inspection of the PY6 spin disks revealed that the material quality was unacceptable. The material supplier discontinued all efforts to develop and supply ceramic materials. Under these circumstances, an alternate verification method had to be used. This involved the use of the material database and confirmatory test results for the NT154 silicon nitride material. This information was obtained from the associate program conducted by AlliedSignal Aerospace Company, Garrett Auxiliary Division. The methodology and correlations developed by this program have also been applied to the NT154 material.

Material and Specimens

The principal requirement of the Ceramic Life Prediction program was to select a material system that was a candidate material for the gasifier rotor application in the DOE-funded Advanced Turbine Technology Applications (ATTAP) program at Allison. Allison selected the GTE Laboratories PY6 injection-molded and hot isostatic pressed (HIP) silicon nitride, which was a leading candidate material and processing method when the program was first envisioned and initiated. The selection of the ceramic material was approved by the Oak Ridge National Laboratory (ORNL) program management as required by the contract.

A mix of coarse and fine silicon nitride powders was ball milled with 6 wt% yttria. The milled powders were mixed with a binder, injection molded, and HIP processed to fabricate test specimens of PY6 silicon nitride material. Four types of specimen were tested during this program: Type-B modulus-of-rupture (MOR) bar, button-head and dogbone tensile specimens, and flat circular disk specimens. All the specimens were subjected to density and dimensional measurements, and character-

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 8, 1995, Paper No. 95-GT-30. Associate Technical Editor: C. J. Russo.

ized by several nondestructive evaluation (NDE) flaw detection methods including visual (10X to 40X), fluorescent penetrant inspection (FPI), microfocus X-ray, and acoustic microscopy inspection prior to mechanical testing.

Analytical Methodology Development. Three major failure mechanisms were identified and addressed for two high-strength silicon nitride materials, GTE PY6 and Norton NT154. The life prediction models were developed for the following damage mechanisms:

- fast fracture (FF)
- subcritical (slow) crack growth (SCG)
- creep deformation and rupture

Fast Fracture

Corner Failure Mode. A total of 320 Type-B MOR specimens were characterized in four-point bending from room temperature to 1400°C at a constant cross-head speed of 0.5 mm/min using a servohydraulic test system. Failure analysis of the specimen fracture surfaces revealed that in addition to volume and surface failure populations, a third type of failure originating at the specimen corners was present. In fact, 80 percent of the room temperature specimens failed from the corners. The results suggest that each type of failure plays a separate role in the assessment of the fast fracture (FF) strength and Weibull properties.

The CARES program was modified [2] to account for the corner flaw population. As part of this process MOR failures are identified as surface, volume, or corner failures. The MLE method is used with the other two failure modes treated as suspended items for the third. This approach is used to determine the Weibull properties of the corner failures, which are applied during the FF analysis of the component.

Consolidation of the Weibull Exponent (m). One of the problems incurred in the development of a ceramic design system is the large amount of testing necessary to obtain the fast fracture Weibull properties at operational temperatures.

A possible approach in reducing the data requirements involves consolidating the elevated temperature data using a single Weibull modulus. It can be analytically demonstrated [3] that if the fast fracture failure mode is dominated by a Mode I failure, then the Weibull exponent is independent of temperature. This is true for high-strength PY6 and NT154 silicon nitride materials.

The surface Weibull modulus (m) for the PY6 material was calculated to be 8.6 at room temperature. At elevated temperatures, the mean Weibull exponent was between 11 and 14 from 1000°C to 1400°C except for the value of 18 at 1300°C. The consistency of these data suggests it is reasonable to consolidate the elevated temperature Weibull modulus into a single value. Individual specimen strengths may be normalized to a single base temperature by multiplying by the ratio of average strength at the base temperature to the strength at the test temperature. If competing failure modes exist, it is first necessary to find the mean strength after accounting for the suspended items, calculate the characteristic strength and m at each temperature, and subsequently ascertain the consolidated characteristic strength. As described in detail in [2], the methodology worked extremely well for the PY6 material. The consolidated Weibull modulus was calculated to be 11.97 for FF surface failures. The capability of determining the consolidated Weibull modulus and the characteristic strengths has been added to the CARES code.

Slow Crack Growth Methodology. Two experimental methods are commonly used to characterize the SCG behavior:

- the constant load test (static fatigue)
- the constant loading rate test (dynamic fatigue)

In the constant load test, the specimen rupture time is determined at various applied stress levels. In dynamic fatigue, the specimen failure strength is measured at various stressing rates. The equations for the two types of test are derived from empirical crack velocity relationships. When an activation energy is used to consolidate temperature data, the equations are as follows [4]:

Constant load:

$$Tf = B*Si^{N-2}*S^{-N}*\mathrm{EXP}(Q/RT) \quad (1)$$

Constant rate:

$$Sf = (B*(N+1)*Si^{N-2}*\dot{S})^{1/(N+1)} \times *\mathrm{EXP}(Q/RT*(1+N)) \quad (2)$$

where

Tf = time to failure (constant load test), s
Sf = fracture strength (rate test), MPa
Si = initial fast fracture strength, MPa
S = applied stress, MPa
$\dot{S}$ = applied stress rate, MPa/s
T = temperature (absolute), K
Q = activation energy (material constant), KJ/mole
N = crack velocity exponent (material constant)
B = crack velocity coefficient (material constant) MPa^2-s
R = universal gas constant = 0.0083 kJ/(mole- K)

Dynamic and static fatigue (stress rupture) testing of the button-head tensile specimens was conducted at the Southern Research Institute (SoRI). An analysis of the static data, using Eq. (1), resulted in slow crack growth constants somewhat different from those obtained dynamically. As shown in Table 1, the static fatigue crack velocity exponent is 50.9 versus 31.1 from dynamic testing.

When the static test data were predicted by using dynamic material constants at 1200°C (Fig. 1) the dynamic fatigue model predicted strengths about 25 percent higher than observed. This may indicate a higher level of environmental damage associated with the static load tests. The activation energy of the static test was 587 kJ/mole compared to 207 kJ/mole in dynamic testing. This also indicates that the material strength, in static test, is

Nomenclature

CARES = Ceramic Analysis and Reliability Evaluation of Structures
FF = fast fracture
Gamma = gamma function
HIP = hot isostatic pressing
K_{ic} = Mode I fracture toughness
kJ = kilo-Joules
Type B = standard ceramic four-point bending specimen, 3 × 4 × 50 mm
MLE = maximum likelihood estimator
MOR = modulus of rupture
mole = molecular weight
MPa = megapascal
NIST = National Institute of Standards
NT154 = Norton silicon nitride material
PIA = principle of independent action
POS = probability of survival
PY6 = GTE silicon nitride material
SCG = slow crack growth
SoRI = Southern Research Institute
Weibull = a type of statistical methodology

Table 1 Comparison of static and dynamic slow crack material constants using button-head tensile specimen tests

Type of test	Number of specimens	Crack velocity exponent—N	Activation energy Q—KJ/mole	Coefficient "B"— MPa^2-sec
Dynamic fatigue	16	31.1	207	2.8E-4
Static fatigue	31	50.9	587	1.88E-25

more sensitive to a change in temperature and associated environmental effects. Thus SCG information obtained from a static test tends to be more conservative.

Creep Analysis Methodology. At high temperatures, particularly at low stresses and high lives, the creep rupture damage mechanism dominates the failures of specimens. For the high-strength silicon nitride materials such as PY6, the onset of creep occurs between 1250 and 1300°C. While there are a number of equations in the literature that adequately describe the creep rupture phenomenon, the "classical" Norton equation was used to describe the relationship between creep rate and applied stress [4]. The "classical" Monkman–Grant equation was used to relate rupture time to creep rate. Finally, by combining the equations, a relationship between rupture time and applied load is obtained. The equations follow:

$$CR = B*S^N *\text{EXP}(-Q/RT) \quad \text{(Norton equation)} \quad (3)$$

$$TF = A*CR^a \quad \text{(Monkman–Grant)} \quad (4)$$

$$TF = A*B^a*S^{a*N} *\text{EXP}(-a*Q/R*T) \quad \text{(combined)} \quad (5)$$

where:

CR = creep rate, 1/s
S = static applied stress, strength, MPa
T = absolute temperature, K
TF = rupture time, s
N = Norton constant (material constant)
a = Monkman–Grant material constant (generally near −1)
A = material coefficient, s^{a+1}
B = material coefficient, MPa^{-N}/s
Q = activation energy (material constant), kJ/mole
R = universal gas constant 0.0083 kJ/(mole- K)

When both creep rate and rupture data are available, it is desirable to use all of the available data to obtain an optimum estimate of the material constants. This is done by first writing Eqs. (3) and (5) with the applied stress on the left side of the equation. Both equations are linearized by taking the natural logarithms. Finally, a least-squares regression technique is used to minimize scatter in the stress term.

Figures 2 and 3 illustrate the correlation between the predicted line and the actual data for both the time-to-failure and creep rate as a function of stress for PY6 material.

Creep Statistical Failure Model. The large amount of scatter observed in the time-to-failure in the tensile stress rupture data, Fig. 2, suggests that a prudent component design practice should incorporate a statistical "weakest link" concept. Such an approach applicable to stress rupture was developed and incorporated into the Allison CARES program.

Rearranging Eq. (5), the mean strength of the specimen at any time for a given temperature is given by

$$Sm = (A*B^a)^{-1/aN}*TF^{1/aN} *\text{EXP}(Q/NRT) \quad (6)$$

where:

Sm = mean strength at time TF

An assumption is made that at all fixed time "slices" the distribution of strength (but not the mean strength) is constant and described by a Weibull distribution. Thus, the Weibull modulus related to creep, mcr, is constant with time. This is similar to the standard practice of characterizing the time-dependent SCG strength distribution with the fast fracture Weibull modulus when developing the SCG statistical theory. After a coefficient of variation is determined from the mean creep strength and the data scatter, it is then possible to estimate the value of mcr.

The coefficient of variation may be written as:

$$Cv = (\sum (Smi - Si)/Smi)^2/(Npts - 3))^{0.5} \quad (7)$$

where:

Si = observed strength of the ith specimen, MPa
Smi = mean strength of the ith specimen, MPa (Eq. (6))
$Npts$ = total number of test specimens

Based on Weibull theory, a Weibull modulus related to the variation of creep strength is approximately [5]:

$$mcr = 1.2/Cv \quad (8)$$

Weibull theory [5] also relates a unit characteristic strength to specimen geometry, mean strength, and Weibull modulus. The characteristic strength is also a function of time and temperature. Assuming creep is primarily a volume related phenomenon, the following equation applies:

$$Soc = Sm*Vspec^{1/mcr}/\text{Gamma}(1 + 1/mcr) \quad (9)$$

where:

$Vspec$ = tensile specimen volume, m^3
Gamma = gamma function
Soc = unit characteristic strength in creep

Finally, for each of the three principal stresses (Weibull PIA) acting on a finite element of volume Vi, the survival probability may be written as [6]:

$$\text{POS}ij = \text{EXP}(-Vi*(Sij/Soc)^{mcr}) \quad (10)$$

where:

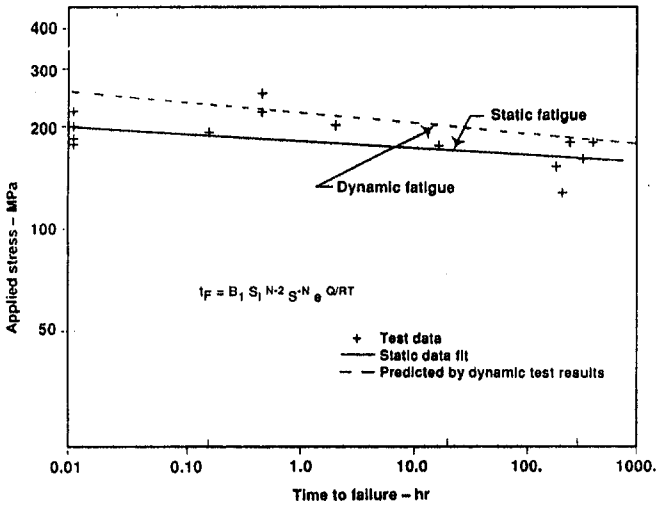


Fig. 1 Comparison of prediction from dynamic test with static test results at 1200°C

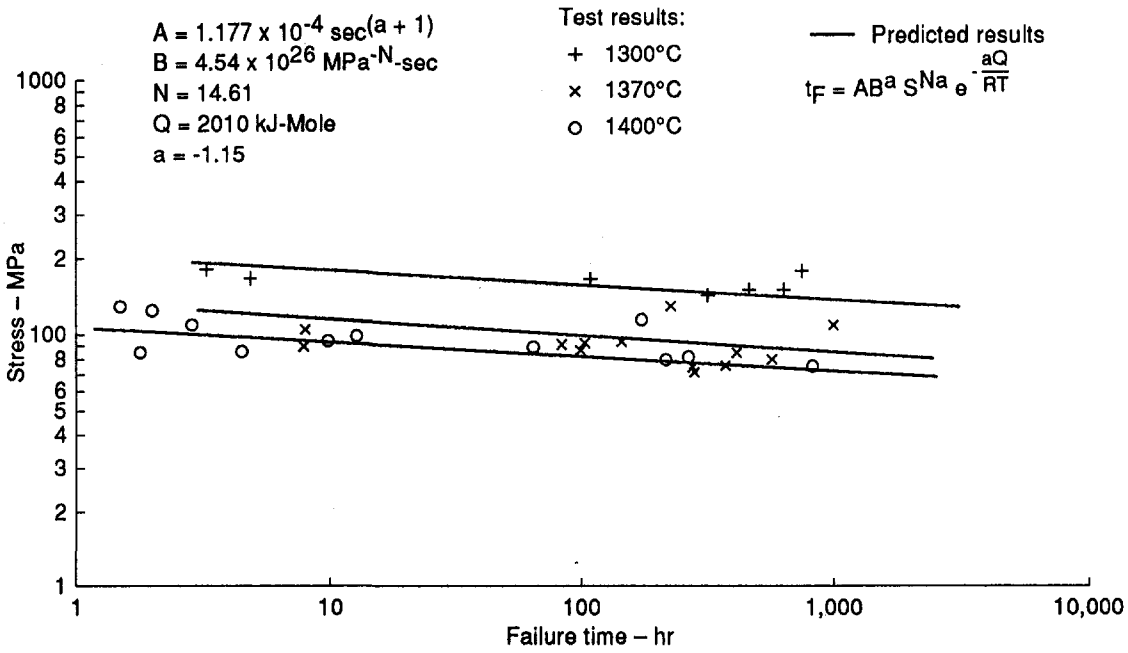


Fig. 2 Predicted versus observed stress rupture failure time for PY-6 button-head tensile bars that failed at elevated temperatures

Sij = the jth (j = 1, 3) principal stress acting on the ith element

POSij = the probability of survival of the ith element resulting from the jth principal stress

The component survival probability is the product of the individual POSij. Thus, at a given time, the survival probability of a component may be computed. Since *Soc* also depends on the local element temperature, this value will vary with location as well as time.

Confirmatory Specimen Correlation. An important element of the program was the verification of the life prediction models. Confirmatory specimens subjected to multiaxial loading were designed and tested to failure in both fast fracture and time dependent modes. Table 2 summarizes the verification tests involved. While the 1371°C spin disk was expected to fail from creep, analysis showed that SCG caused the failure. Thus SCG was the only time-dependent failure mode correlated by test.

Fast Fracture. Three types of multiaxial specimen using two high-strength silicon nitride materials were tested in fast fracture at room temperature. The two flat circular disks failed on the surface only while the spin disk failed both on the surface and internally.

PY6 Silicon Nitride Ball-on-Ring Confirmatory Tests. Battelle Research Laboratories, Columbus, Ohio, conducted biaxial room temperature tests using flat circular disk specimens of 31.75 mm nominal diameter and 2.5 mm thickness. The disk surface was machined using a protocol developed at Battelle. The specimens were tested using a ball-on-ring arrangement and loaded at the center with a 6.35 mm diameter ball and supported on a 25.4 mm diameter ring. The maximum stress on the disk occurs on the tensile surface at the center of the disk where both radial and tangential stresses are equal and given by the equations in [7].

After several iterations, a promising approach was found that allowed the prediction of the biaxial results from uniaxial MOR test specimens. Initially, Weibull shape and scale parameters obtained from room temperature testing of longitudinally ground Type-B MOR bars were used to predict the biaxial strengths. Results were not encouraging. The room temperature Weibull modulus of 8.6 differed greatly from the value of 17 for the disk specimens. Fractographic examination of the ball-

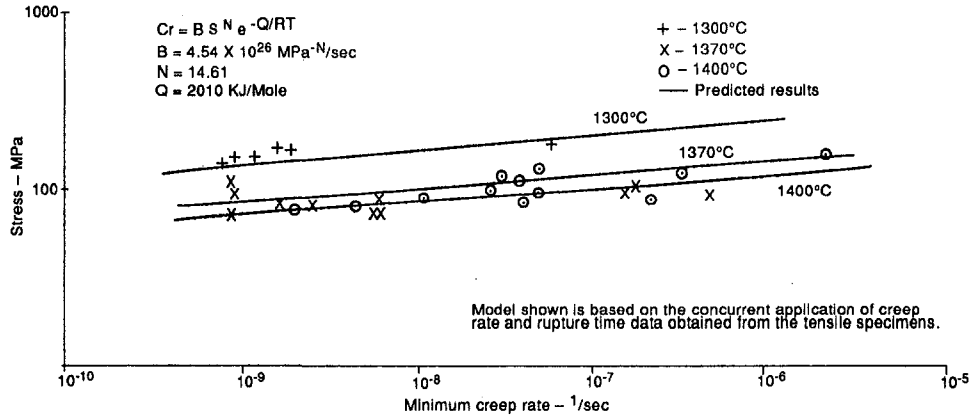


Fig. 3 Predicted versus observed creep rates for PY-6 button-head tensile bars that failed from stress rupture at elevated temperatures

Table 2 Summary of multiaxial slow crack growth and fast fracture verification tests

Component	Type of loading	Material	Temp—°C	Failure mode analyzed	Failure location
Flat circular disk	Center Ball	PY6	Room temp	Multiaxial fast fracture	Surface
Flat circular disk	Center ball	PY6	1200	Multiaxial SCG	Surface
Flat circular disk	Center ball	PY6	1300	Multiaxial SCG	Surface
Flat circular plate	Uniform pressure	NT154	Room temp	Multiaxial fast fracture	Surface
Spin disk	Rotation	NT154	Room temp	Multiaxial fast fracture	Surface/ Internal
Spin disk	Rotation	NT154	1204	Multiaxial SCG	Surface
Spin disk	Rotation	NT154	1371	Multiaxial SCG	Surface

on-ring specimens showed that failure occurred on the tensile surface normal to the machining marks. To approximate the observed biaxial results better, a set of Type-B MOR bars was manufactured with 320 grit transverse machining and evaluated at room temperature. When these data were analyzed, the average strength more accurately simulated that of the disk, but the Weibull modulus was approximately equal to 8, again a poor correlation.

Finally, 20 MOR specimens were machined from ball-on-ring disk specimens. The specimen surface was first ground and polished identical to the biaxial disk specimens. MOR bars were then dissected from the disks such that machining marks were transverse to the specimen length. The MOR bars represented the surface condition of the biaxial disk specimen. The measured Weibull modulus of 20 agreed well with that of the disks.

Material properties from the MOR specimens cut from the disks were used to compute the disk characteristic strength using effective areas from two differing multiaxial models. In the first model, only the stress normal to machining marks was assumed to influence the fast fracture failure probability. In the second model, the effective area calculation was based on the Weibull principle of independent action (PIA) [6]. Characteristic strength of the ball-on-ring specimens was predicted using the aforementioned models and Weibull properties of the transversely ground MOR specimens. The first model predicted a characteristic strength of 785 MPa versus an experimental value of 776 MPa, a difference of only 1 percent. The PIA value of 759 MPa was within 2 percent of the measured value and is slightly conservative.

In Fig. 4, the first model is used to predict a 90 percent scatter-band for the ball-on-ring disks. This was done by first converting the MOR measured strengths to that of the ball-on-ring specimens, which have a smaller effective area. The CARES program was used to predict the scatterband. The correlation appears reasonable. All of the disk data lie within the 90 percent confidence band and the Weibull slope is similar to the predicted value.

NT 154 Silicon Nitride Pressurized Flat Plates Confirmatory Tests. A total of 37 simply supported flat circular plates were loaded with a uniform pressure at room temperature until FF occurred. The failure origin was near the center of the plate on the tensile surface. The plate nominal diameter was 50.46 mm, the nominal thickness 3.191 mm, with a support diameter of 48.26 mm. From calculations of the peak surface stress [8] and measurements of the center strain, the elastic modulus was calculated to be about 322 GPa with a Poisson's ratio of 0.27. A maximum likelihood estimate (MLE) analysis of the FF strengths calculated the Weibull modulus to be 10.4 with a characteristic strength of 808.6 MPa.

Room temperature NT154 FF data sets included Type *B* and *E* type MOR bars and button-head tensile specimens, all of which exhibit surface failures. Mixed mode FF testing of NT154 has shown this material to be shear insensitive. Therefore, the PIA method was used to predict the strength of the NT154 pressurized plates. Table 3 shows that the predicted characteristic strength using three different types of specimen data was within 3 percent of the average failure strength of 808.6 MPa.

Room Temperature Verification Using NT154 Spin Disks. A total of nine disks were spun to failure at room temperature. Three of them failed from internal defects whereas six failed from surface flaws.

The analysis of spin disks with internal failures was conducted using test data from two types of NT154 MOR bars, which were tested at room temperature and demonstrated volume failures. These were the *E*-Bars and 4-in.-long bars cut from actual disks. Unlike the MIL-B bar, the thickness of these specimens exceeded their width, resulting in relatively high effective volumes.

Machining direction has little or no effect on the internal strength. Both Type-*E* and 4-in.-long bar data were reduced using corner and surface failures as suspended items to generate material properties. The CARES MLE methodology was used to determine the Weibull modulus and the characteristic strength. In addition, both data sets were pooled, after normalizing the 4-in. bar strength to that of Type-*E* bar, which resulted in the consolidated Weibull properties.

Once fast fracture material data were available, the spin disk stresses were calculated using a NASTRAN finite element analysis (FEA). The finite element information as well as the Weibull FF properties allowed CARES to calculate the disk failure probability. The experimental survival probabilities for the three disks that experienced volume failure were obtained from median ranks with surface failures treated as suspended items. Table 4 compares the predicted failure probabilities obtained from *E*-Bar, 4-in. bar, and consolidated data with the experimental results.

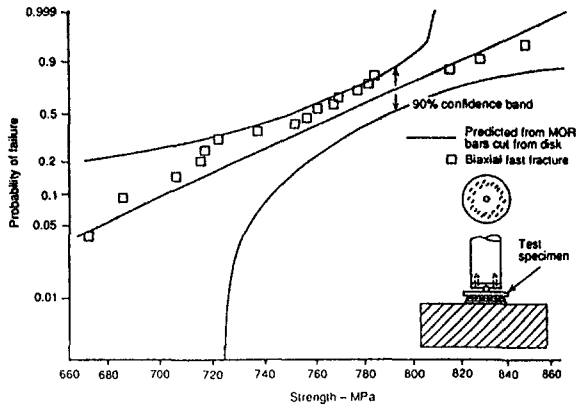


Fig. 4 PY-6 biaxial fast fracture room temperature ball-on-ring confirmatory test prediction of surface failure probability versus experimental results

Table 3 Room temperature NT154 pressurized flat plate surface characteristic strength predicted from uniaxial fast fracture test specimens and the Weibull PIA methodology

Specimen type	Weibull modulus—ms	Unit characteristic strength—MPa-mm$^{2/ms}$	Disk effect area—mm^2	Predicted disk characteristic strength—MPa	Measured disk characteristic strength	Percent difference predicted vs. measured
Tensile	6.219	2313	705.4	805.6	808.6	-0.4%
E-bar	8.105	1803	559.3	826	808.6	+2.1%
Mil-B	8.070	1808	561.4	830	808.6	+2.6%
Flat plate	10.4	—	—	—	808.6	—

The experimental results are bracketed by predictions made from *E*-Bar and 4-in. bar data. The predictions made from the 4-in. bars are conservative, while the *E*-bar predictions are optimistic. When all data are combined, the observed and predicted results agree well. Figure 5 illustrates the predicted mean and 90 percent scatter band based on the combined Weibull properties. After the disk failure data are corrected for the effective volume, only disk 3 lies outside the scatter band. This disk is approximately 10 percent stronger than predicted.

The analysis of the spin disks that failed at the surface used the NT154 uniaxial tensile data since test data that simulate the surface finish of the spin disk were not available. The analysis yielded Weibull modulus (ms) of 6.22 and unit characteristic strength (So) = 250.83 MPa-m$^{2/ms}$.

The failure probability of the six burst tested disks were predicted using the finite element analysis and the CARES PIA method. A comparison of the median rank disk failure probability to that computed based on the above-described methodology, Table 5, indicates the mean strength was overestimated by 15 percent. It is imperative that material data be obtained from the specimens with identically finished surfaces. Machining has a major impact on the surface properties of high-strength silicon nitride.

Slow Crack Growth

SCG Verification of PY6 Ball-on-Ring Specimens. A total of 25 flat circular plates using the ball-on-ring arrangement were tested to correlate the static fatigue methodology at 1200°C and 1300°C. All specimens had a 31.75 mm nominal diameter and were supported by balls at a diameter of 15.9 mm. The 1200°C plates were 1.78 mm thick, while the 1300°C plates were 2.00 mm thick. At elevated temperatures, a surface oxide layer is formed that decreases the surface sensitivity to the initial surface preparation. As a result, both FF and SCG constants obtained from statically loaded (stress rupture) button-head tensile test specimens were used to predict the lives of the flat circular plates. The SCG material constants were calculated using the 1200°C, 1260°C, and 1300°C database. Only specimens that failed in less than 3.3 h at 1300°C were used.

The predicted initial biaxial strength and the SCG constants were used in Eq. (1) to predict the effect of time on the biaxial strength of the circular disks at 1200°C and 1300°C. Figure 6 shows that the experimental and predicted results are in good agreement. The predictions pass through the mean of the short

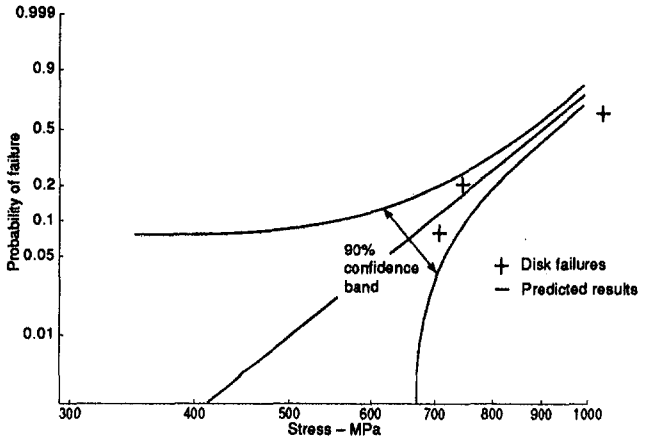


Fig. 5 Comparison of predicted fast fracture behavior versus experimental results for NT154 room temperature spin disks with volume failures

Table 4 Comparison of NT154 room temperature spin disk internal fast fracture experimental and predicted failure probabilities

Disk ID	Internal stress—MPa	Experimental failure probability	Predicted failure probability 4 in.	E-bar	Combined bar
3	506.0	0.618	0.98	0.23	0.80
31	358.9	0.204	0.27	0.017	0.13
34	345.5	0.086	0.21	0.013	0.10

Table 5 Comparing experimental and predicted failure probability for disks with surface failures at room temperature

Disk ID	Stress—MPa	Experimental failure probability	Predicted failure probability from tensile bar data
9	424.5	0.758	0.444
17	405.3	0.622	0.349
18	378.1	0.485	0.243
36	359.3	0.348	0.184
4	331.6	0.211	0.116
21	263.9	0.074	0.029

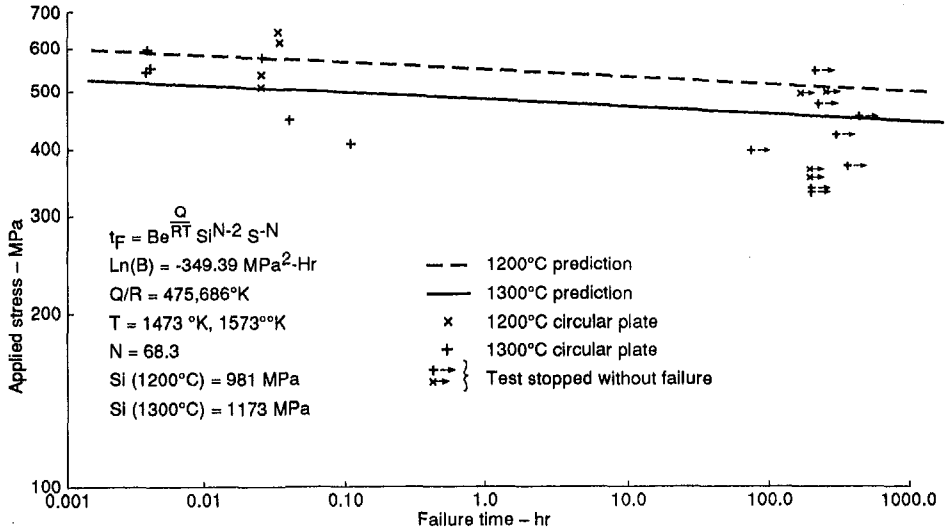


Fig. 6 Comparison of predicted and observed slow crack growth behavior for PY-6 ball-on-ring flat circular disks tested at 1200°C and 1300°C

time failure data and lie above the bulk of the longer duration tests that were not tested to failure.

It should be noted that the use of an elastically calculated stress is inherently conservative at 1300°C where creep relaxation will reduce the surface bending stresses. This may explain why two of the run-outs at 1300°C lie above the predicted line.

NT154 Spin Disk Tests at 1204°C (2200°F). A total of 13 NT154 spin disks were subjected to various rotational speeds at 1204°C (2200°F). Nine disks failed in fast fracture while four failed on the surface in a time-dependent failure mode. Stress rupture tensile testing has shown that slow crack growth was the dominant time-dependent damage mechanism at this temperature.

SCG analysis of the disk requires material constants to define both FF and SCG. This information is used with the predicted stress distribution by the CARES code to compute the failure probability at a given component life. A Weibull analysis of the disk fast fracture strength showed that the disks were about 15 percent weaker than the strength predicted from the button-head tensile specimens. As a result, fast fracture material constants were obtained directly from FF disk testing and used in the SCG analysis, Table 6. A CARES MLE analysis, using the volume failures as suspended items, was used to find the disk surface Weibull modulus and characteristic strength. The FF constants were used with the finite element stress analysis and the CARES code to determine the unit characteristic strength. The unit characteristic strength was determined to be 155.23 MPa-$m^{2/ms}$; the Weibull modulus was 6.0. The slow crack growth material constants were obtained by applying Eq. (1) to the analysis of static failures of button-head tensile specimens. The coefficient, B, was 77,060 MPa^2-h while the crack velocity exponent, N, was 14.1.

Table 6 NT154 spin disks failed in fast fracture at 1204°C

Disk ID	Failure speed—rpm	Peak surface stress—MPa	Type
32	76,000	227.1	Surf.
33	76,000	227.1	Surf.
13	83,000	269.1	Surf.
23	83,000	269.1	Surf.
37	74,000	215.8	Vol.
29	80,400	253.1	Vol.
6	93,000	335.6	Vol.

The Allison CARES program computed the time-dependent probability of failure of a statically loaded component by applying equations of the following form to each finite element and finding the product of the survival probabilities.

$$Si, o = (S^{N-2} + S^N * t/B)^{1/(N-2)} \quad (12)$$

$$\text{POS}i = \text{Exp}(-Ai(Si, o/So)^m) \quad (13)$$

where:

Si, o = effective stress at time = 0
S = element stress
Ai = element surface area
t = time
POSi = probability of survival of ith element

Table 7 defines the rpm levels and the hold time for the four disks that failed on the surface as a result of SCG. Since only disk 22 was held at a constant rpm until final failure, the previous equation for Si, o must be modified for the complex missions. If the time spent moving from one rpm to another is neglected and if the loads are constantly increasing, equations of the following more general forms were applied [3]:

$$Sc, o = (Sf^{N-2} + S \text{ max } 1^N * t1/Bc1 + \ldots . S \text{ max } n^N * tn/Bcn)^{1/(N-2)} \quad (14)$$

$$\text{POS} = \text{Exp} - (Sc, o/Sch)^m \quad (15)$$

where:

Sc, o = effective component stress at time = 0
Sch = disk characteristic strength
Sf = final mission peak stress
$t1 \ldots n$ = time intervals at each stress
$S \text{ max} \ldots S \text{ max } n$ = peak surface stress at each rpm
$Bc1 \ldots Bcn$ = SCG coefficient for the disk at each rpm

The values of Bc are found by first using the CARES analysis to determine the failure probability of the component after time, ti, at a constant rpm, at which a peak surface stress, S max i, is produced. Equations (14) and (15) are then rearranged to find Bc.

Table 7 Applied load spectrum and predicted failure probability for four NT154 spin disks that failed on the surface in slow crack growth at 1204°C

Disk ID	rpm x 10^3	Load time —hr	Predicted mean load time —hr	Peak surface stress— MPa	Effective T. zero stress— MPa	Predicted POS
41	67/75	150/27	150/16	178.5/220.5	285.3	0.4657
22	71	27.33	116	199.4	249.1	0.7259
7	71/75/82	100/10/*	100/2.7	199.4/220.5/262.9	291.7	0.4174
19	75/84.5	50/*	23.5	220.5/278.6	303.8	0.328

* Burst

Table 7 summarizes the applied load spectrum, the peak stresses, the predicted probabilities of survival, and the predicted mean mission lives. The predicted mean lives were slightly low for three of four disks tested. For instance, disk 41 spent 150 h at 67,000 rpm and 27 h at 75,000 rpm prior to failure. The predicted result was 150 h followed by 16 h at 75,000 rpm. Only disk 22 demonstrated a life below the predicted mean life, 27 h compared with a prediction of 116 h.

The CARES program was used with the 1204°C disk fast fracture failures to predict the 90 percent confidence band for the 1204°C disk fast fracture failures. Using median ranks along with the effective time zero stress, it was found that the SCG test results lie within the confidence band and do not differ greatly from the mean Weibull line (Fig. 7). This tends to show that the SCG methodology provided a reasonable prediction.

NT154 Spin Disk Testing at 1371°C (2500°F). Five NT154 disks were burst tested at 1371°C (2500°F) to failure. While creep damage mechanisms are important at this temperature, an SCG analysis predicted the observed test results. The creep analysis was not particularly accurate. All the disk failures initiated from a surface flaw. Fast fracture surface failure data obtained from identical disks tested at 1204°C were scaled to the 1371°C temperature. In scaling the Weibull modulus was kept the same, 6.0, while the characteristic strength was ratioed by the values of K_{ic} and found to be 146.38 MPa-$m^{2/m}$. Stress rupture (static fatigue) data for tensile specimens tested at 1295°C, 1315°C, 1371°C, and 1400°C were pooled and regressed using Eq. (1) to ascertain the SCG constants. The value of the crack growth exponent (N) was 17, while the coefficient B was 0.040864 MPa-h^2 at 1371°C for surface failures.

In determining the spin disk life at a designated failure probability, a NASTRAN finite element analysis is used to find both the stress distribution and the corresponding areas and volumes.

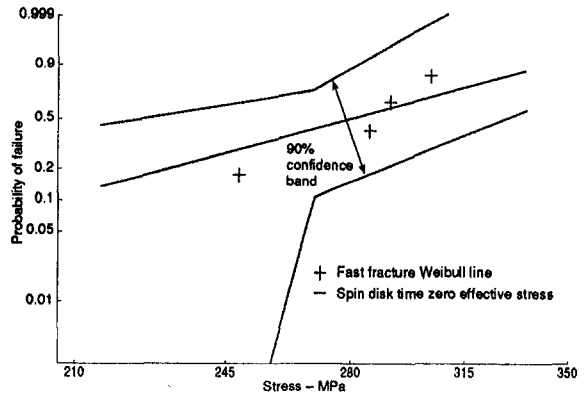


Fig. 7 NT154 1204°C spin disk confirmatory test: comparison of predicted failure probability versus experimental surface SCG failure results

CARES is then used to determine the life using the Weibull PIA methodology.

Table 8 summarizes the loadings and failure time of disks tested at 1371°C. The results indicate a large scatter in observed lives even at the same loading. At 50,000 rpm disk 26 lasted 11 h, slightly longer than the predicted mean life of 10.2 h. At the same rpm, disk 24 failed after only 0.017 h, slightly less than the expected 5 percent failure rate life. The lives demonstrated by the other three disks were roughly bounded by the predicted mean life and the 5 percent failure life, Table 8. Thus, the life prediction at 1371°C using SCG was slightly optimistic. The large amount of scatter demonstrated in the test lives is not that unusual for a high-strength silicon nitride ceramic material since the fracture controlling defects may have differing flaw characteristics. The low Weibull modulus of NT154 is also partially responsible for the large scatter. Figure 8 further illustrates that although the time zero effective stress lies within the 90 percent scatter band, the prediction of failure probability is slightly low.

Conclusions

Multiaxial Fast Fracture

1 Correlation was obtained between the experimental and predicted biaxial fast fracture strength of both PY6 and NT154 silicon nitride materials.

Multiaxial Slow Crack Growth

1 A difference in slow crack growth material properties was observed between the test results obtained from static and dynamic tests. The static results are conservative compared to the dynamic test results.
2 The SCG methodology resulted in a reasonable prediction for both the PY6 and the NT154 verification specimens.
3 The SCG mechanism appears to dominate the failure of PY6 biaxial specimens and NT154 spin disks even at test conditions where creep deformation was observed in uniaxial tensile testing. The SCG models accurately predicted the verification test results even though the uniaxial specimens showed creep damage at comparable temperatures.

Recommendations

Multiaxial Fast Fracture

1 Shear insensitive fast fracture methodology should be used to analyze high-strength silicon nitrides.
2 The reliable prediction of component failure probability requires material data base from specimens with similar surface finish. If this type of information is not available,

Table 8 Comparison of observed versus mean and 5 percent failure SCG lives of NT154 spin disks failed from surface at 1371°C

Disk ID	Failure—rpm	Predicted 5% failure life—hr	Observed life—hr	Predicted mean life—hr	Peak surface stress—MPa
8	41,800	7.85	6.966	4,300	79.35
26	50,000	0.0186	11.08	10.16	107.66
24	50,000	0.0186	0.0166	10.16	107.66
14	55,000	0.0007	0.95	0.40	127.4
15	45,000/50,000	0.22/0.009	10.0/0.40	121/4.9	89.81/107.66

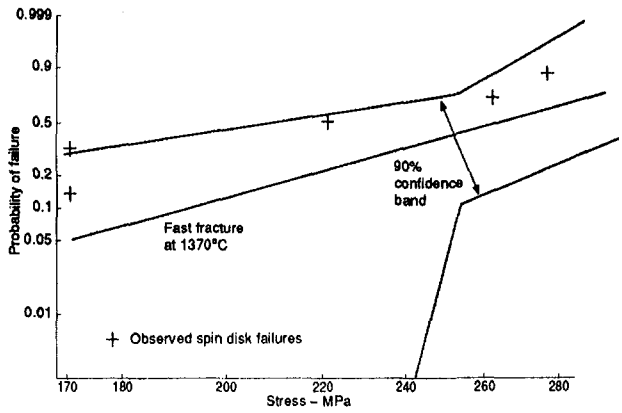


Fig. 8 Comparison of NT154 1371°C spin disk SCG time zero effective strength versus predicted fast fracture strength. Test results lie within the predicted 90 percent confidence band.

the lowest characteristic strength data base should be used.

Multiaxial Slow Crack Growth

1 To insure the long-term durability of ceramic components, the time-dependent failure mechanisms should be considered. For PY6 and NT154 types of silicon nitride, the SCG methodology should be applied below 1300°C.
2 Above 1300°C, both SCG and creep should be considered to determine the dominant failure mechanism. The failure mode resulting in the minimum component life should be used to ensure a robust design.

Acknowledgments

Research was sponsored by the U.S. Department of Energy, Assistant Secretary for Energy Efficiency and Renewable Energy, Office of Transportation Technologies, as part of the Ceramic Technology Project of the Propulsion System Materials Program, under contract DE-AC05-840R21400 with Martin Marietta Energy Systems, Inc.

The authors would like to thank Dr. D. R. Johnson, Project Manager, Ceramic Technology, and Dr. C. R. Brinkman, Ceramic Life Prediction Program Manager, Oak Ridge National Laboratory, for their patience and continued support. Thanks also are due to Dr. A. R. Rosenfield and Dr. B. S. Majumdar of Battelle Memorial Institute for conducting biaxial testing studies, to Dr. S. M. Wiederhorn, National Institute of Science and Technology, for dog-bone creep studies, and to Mr. Terry Barnett at Southern Research Institute for conducting button-head fast fracture and creep testing. The authors would also like to thank Dr. J. Chang for his contribution to the failure analysis activity. The authors sincerely appreciate the cooperation of AlliedSignal Aerospace Company for providing the NT154 silicon nitride data.

References

1 Gyekenyesi, J. P., and Nemeth, N. N., "Surface Flaw Reliability Analysis of Ceramic Components With the SCARE Finite Element Postprocessor Program," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 109, 1987, pp. 274–281.

2 Khandelwal, P. K., Provenzano, N. J., and Schneider, W. E., "Life Prediction Methodology for Ceramic Components of Advanced Vehicular Heat Engines," *Proc. Annual Automotive Technology Development Contractors' Coordination Meeting,* 1993, pp. 199–211.

3 Khandelwal, P. K., Provenzano, N. J., and Schneider, W. E., "Life Prediction Methodology for Ceramic Components of Advanced Vehicular Heat Engines—Final Report," to be published as a part of Ceramic Technology Project of the Materials Development Program, DOE contract DE-AC05-840R21400 with Martin Marietta Energy Systems, Inc.

4 Wiederhorn, S. M., Quinn, G. C., and Krause, R., "Fracture Mechanisms Maps: Their Applicability to Silicon Nitride," Life Prediction Methodologies and Data for Ceramic Materials, ASTM STP 1201, C. R. Brinkman and S. F. Duffy, eds., American Society for Testing and Materials, Philadelphia, 1994, pp. 36–61.

5 DeSalvo, G. J., "Theory and Structural Design Applications of Weibull Statistics," Westinghouse Astronuclear Laboratory Publication, WANL-TME-2688, May 1970.

6 Freudenthal, A. M., "Statistical Approach to Brittle Fracture," *Fracture, Vol. 2—Mathematical Fundamentals,* H. Liebowitz, ed., Academic Press, New York, 1968, pp. 591–619.

7 Shetty, D. K., Rosenfield, A. R., Mcquire, P., Bansal G. K., and Duckworth W. H., "Biaxial Flexural Tests for Ceramics," *Journal of Ceramic Society,* Vol. 59, No. 12, 1980, pp. 1193–1197.

8 Shetty, D. K., Chao, Luen-Yuan, "Reliability Analysis of Structural Ceramics Subjected to Biaxial Flexure," *Journal of Am. Ceramic Society,* Vol. 74, No. 2, 1991, pp. 333–344.

Strength Prediction of Ceramic Components Under Complex Stress States

A. D. Peralta

D. C. Wu

P. J. Brehm

J. C. Cuccio

M. N. Menon

AlliedSignal Engines,
Phoenix, AZ 85034

The capability to perform accurate fast-fracture strength predictions for ceramic components under complex stress states must be available in order to transition the use of advanced, high-strength ceramic materials from the laboratory to the high-strength/high-temperature applications they are intended for. Multiaxial strength prediction theories have provided the prediction capabilities, but only limited testing of these theories under complex states of stress and stress gradient conditions has been performed previously. Presented here are comprehensive test results and strength predictions for ceramic components subjected to complex states of stress and stress gradient conditions. The results show excellent agreement of the predictions from the multiaxial theories with test results for volumetrically distributed flaws. An important finding of this work is the problem that arises in performing component surface strength predictions from database-type specimens. Database-type specimens and component surface properties are not necessary correlated, and in many cases it may be completely inaccurate to use database-type specimen surface properties for component surface strength predictions.

Introduction

As part of the Life Prediction Methodology for Ceramic Components of Advanced Heat Engines program, funded by the Department of Energy/Oak Ridge National Laboratories (DOE/ORNL) under Contract No. DE-AC05-840R21400, AlliedSignal Engines has been developing methodology to perform life predictions for high-strength ceramic materials. The selected approach included the development of a comprehensive materials database, nondestructive evaluation methods, material behavior models, statistical tools for data analysis, and probabilistic tools for component design. The accuracy of the developed methods was confirmed by comparison of predictions with the results of subelement tests. This paper focuses on the progress made and the lessons learned in confirming life prediction methods for the fast fracture mode of ceramic components.

Acceptance of advanced ceramics for use in heat engine applications is based to a large extent on the fast fracture strength capabilities of the materials. Time-dependent failure modes also play a major role in the successful implementation of ceramics in heat engines, because of the associated long life requirements of these applications. To those familiar with heat engine component design, the complex stress fields found in engine components are a fact of life. But laboratory testing of specimens for database generation is restricted by specimen costs and the complexity of testing. Most testing is currently performed on four-point bend test specimens and, to a lesser extent, on tensile specimens, both providing simple uniaxial stress fields. Weibull (1939) developed the uniaxial theory to predict the behavior of uniaxially loaded ceramic components. The multiaxial theories developed by Batdorf and Crose (1974), Batdorf and Heinisch (1978), and Evans (1978) provide the tools that allow the data generated from laboratory specimens to be used for heat engine components with complex stress fields.

An important consideration in prediction of the fast fracture strength of ceramic components from specimen data is in the similitude of the specimen to the component properties. Considerations have to be made as to the similitude of the volume properties, the surface conditions, and the presence or nonpresence of chamfers in the component. The volume considerations can be satisfied by using the same powder and processing method for both the test specimens and the components. The surface considerations are much more complex. One must make sure that any difficulties associated with the fabrication of the component are also reflected in the specimen fabrication. The machining operations must be the same for all machined surfaces, and the mold material and mold removal operations must be the same for all as-processed surfaces.

A second important consideration is the failure modes associated with the specimen and the component. One must make sure that the failures expected in the component are also present in the specimen. For example, the strength of a component of volume V_o, with machined surface area A_o and chamfer length L_o, cannot be accurately predicted from tensile specimen data alone, because the tensile specimen does not provide information on the chamfer failure mode. Also, the pertinent machined-surface failure mode of the component may be the transverse machining direction, whereas in a longitudinally machined tensile specimen such surface property data cannot be captured.

Testing

A large database was generated during the AlliedSignal program to investigate fast fracture and time-dependent failure modes. In total, 1400 specimens were tested, in the various configurations shown in Fig. 1. The specimens were made from cold isostatically pressed plus hot isostatically pressed NT154 silicon nitride ceramic material supplied by Norton Advanced Ceramics of Worcester, MA.

Four specimen types for fast fracture strength were tested: MIL-A, MIL-B, MIL-E, and ORNL tensile buttonhead specimens. Strength data from these simple, uniaxially loaded specimens were then used to predict the fast fracture strength of confirmatory specimens having significant stress gradients and complex multiaxial stress fields. Successful strength predictions of these components with complex stress states would validate the multiaxial theories.

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 22, 1995. Paper No. 95-GT-394. Associate Technical Editor: C. J. Russo.

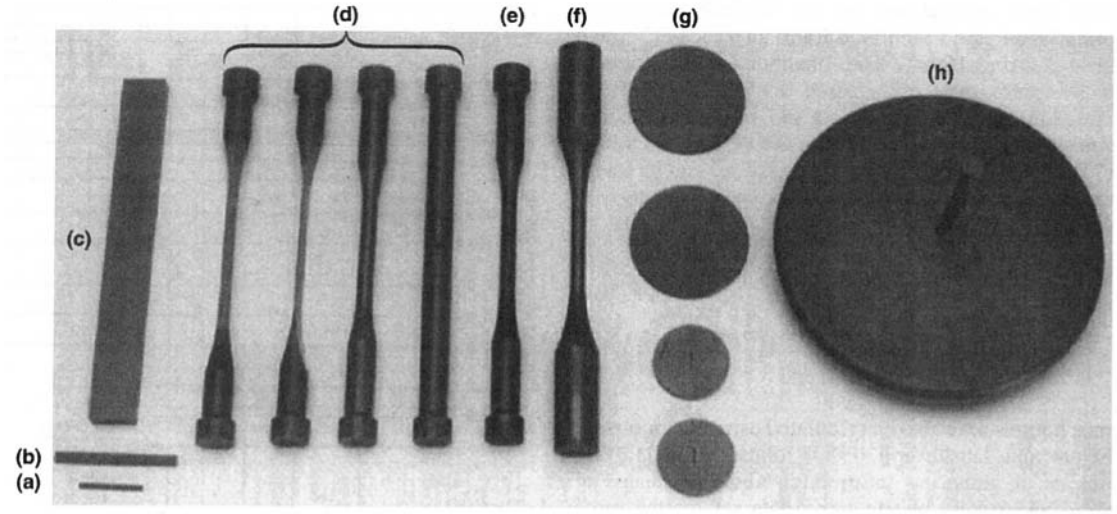


Fig. 1 Ceramic test specimens used in this program: (a) MIL-A; (b) MIL-B; (c) MIL-E; (d) flaw growth and notched tensile; (e) buttonhead tensile; (f) tension-torsion; (g) plate bending and diametral compression; (h) spin disk

Five confirmatory specimen geometries were designed to provide for a variety of stress gradients, multiaxial stress conditions, and of sizes significantly different than the database specimens: a spin test disk, a tension-torsion specimen, a plate bending specimen, and two notched specimens having stress concentration factors (K_t) = 1.6 and 2.1, respectively. Care was taken in the processing of these specimens to ensure that the material was the same for all the different specimen types and the confirmatory specimens. Specific instructions were given to the machining vendors to control machining process factors such as the grinding wheel, grit size, and down feed rate. The notched specimens were designed with the same notch root radius, but different gross cross-sectional areas; this was done with the intent of producing similar surface machining conditions and finishes.

Background

The data analysis was performed using statistical tools developed under this program (Cuccio et al., 1994). The tools developed consisted of a reformulation of the classical likelihood function to account for the specimen effective sizes, as shown in Eq. (1), which is the reliability equation using Evans' notation for the size scale term:

$$P_S = \exp - \left\{ I_V V \left(\frac{\sigma_{e,\max}}{\sigma_o} \right)^m \right\} \tag{1}$$

The likelihood solution, using censored data analysis, for m and σ_o using the new formulation is given by Eqs. (2) and (3) (Peralta, 1991):

$$\frac{1}{\hat{m}} = \frac{\sum_{k=1}^{n} V_k \sigma_k^{\hat{m}} [\hat{I}'_k + \hat{I}_k \ln(\sigma_k)]}{\sum_{k=1}^{n} \hat{I}_k V_k \sigma_k^{\hat{m}}} - \frac{1}{r} \sum_{i=1}^{r} \left[\frac{\hat{I}'_i}{\hat{I}_i} + \ln(\sigma_i) \right] \tag{2}$$

$$\hat{\sigma}_o = \left[\frac{1}{r} \sum_{k=1}^{n} \hat{I}_k V_k \sigma_k^{\hat{m}} \right]^{1/\hat{m}} \tag{3}$$

The Weibull parameters calculated from this approach are material constants, independent of specimen size and loading conditions. Note that the second Weibull parameter, σ_o, can be thought of as the characteristic strength of a specimen of unit effective size. The multiaxial stress gradient factors are defined as follows, in Eqs. (4)–(6):

$$I_V = \iiint_V \int_{-\pi/2}^{\pi/2} \int_0^{2\pi} \frac{\sigma_e^m(x, y, z, \phi, \psi)}{4\pi V_o \sigma_{e,\max}^m} \cos\phi \, d\phi \, d\psi \, dV \tag{4}$$

$$I_A = \iint_A \int_0^{2\pi} \frac{\sigma_e^m(x, y, \theta)}{2\pi A_o \sigma_{e,\max}^m} \, d\theta \, dA \tag{5}$$

$$I_L = \int_L \frac{\sigma_e^m(L)}{L_o \sigma_{e,\max}^m} \, dL \tag{6}$$

The multiaxial stress gradient factors are defined such that $0 \leq I_V < 1$. The volume, I_V, is unity for a uniform equitriaxial state of stress; the surface I_A is unity for a uniform equibiaxial state of stress; the chamfer, I_L is unity for a uniform uniaxial state of stress. The effective stress can be defined according to any of the many failure theories available in the literature. A good summary can be found in the CARES Users Manual (Nemet et al., 1990). The normal stress criterion was used in this analysis.

Nomenclature

A_o = component physical size (area)
I_V = multiaxial stress gradient factor
$I_V V$ = effective size
K_t = stress concentration factor
L_o = component physical size (chamfer length)
$l(\cdot, \cdot)$ = likelihood function
m = first Weibull parameter, Weibull modulus or Weibull slope
$\hat{m}_{\sigma_o}$ = Weibull slope, given σ_o
P_f = probability of failure
P_f = probability of success
r = total number of failures
V_o = component physical size (volume)
$W(\cdot)$ = likelihood ratio statistic
σ_e = equivalent stress
σ_o = second Weibull parameter
$\hat{\sigma}_{o,m}$ = second Weibull parameter, given m

Confidence limit calculations for the Weibull parameters were performed using the Likelihood Ratio approach (Cox and Oakes, 1984; Tucker, 1990). The likelihood ratio statistic, $W(\cdot)$, has an approximate chi-squared distribution with one degree of freedom, $\chi^2_{1,\alpha}$. To determine the $1 - \alpha$ confidence limits, one must compute the values of m and σ_o for which Eqs. (7) and (8) hold. Two solutions to these equations can be found that give the upper and lower confidence bounds:

$$\left\{ \begin{array}{c} W(m) = 2[l(\hat{m}, \hat{\sigma}_o) - l(m, \hat{\sigma}_{o,m})] \\ W(m) = \chi^2_{1,\alpha} \end{array} \right\} \quad (7)$$

$$\left\{ \begin{array}{c} W(\sigma_o) = 2[l(\hat{m}, \hat{\sigma}_o) - l(\hat{m}_{\sigma_o}, \sigma_o)] \\ W(\sigma_o) = \chi^2_{1,\alpha} \end{array} \right\} \quad (8)$$

Confidence bounds may also be calculated using the bootstrap method (Efrom and Tibshiriami, 1986; Johnson, 1990). This method consists of gathering information about estimates of variability and uncertainty by obtaining data sets of the same number of samples as the original data set via a Monte Carlo simulation method. The confidence limits are calculated by generating new data sets and their respective Weibull parameters a sufficient number of times to define the required confidence limits accurately.

Data Analysis

Fractographic examination of the specimens was performed to determine the failure modes present. Three failure modes were found in the four-point bend specimens: volume inclusion failures, surface failures, and chamfer failures. Surface and volume inclusion failures were found in the tensile specimens. One set of specimens of MIL-A, MIL-B, MIL-E and tensile specimens was machined longitudinally and one set of MIL-B specimens was machined transversely. The longitudinally machined specimens were analyzed separately from the transverse specimens, since these produce two entirely different flaw populations. Censored data analysis was performed on all these specimens to extract the most information from the data.

Volume Failures. The MIL-A and MIL-B specimens were machined from square tiles, whereas the buttonhead tensile and the MIL-E specimens were machined from cylindrical billets. Since the same processing steps were followed in making all the specimens, it is reasonable to assume that all of the specimens may be pooled together to calculate the volume Weibull parameters for the material. The data analysis using the methods described above yield the results plotted in Fig. 2. Note that the tensile specimens and the MIL-E specimen data do pool together nicely and that the MIL-A and MIL-B specimens tend to show a lower strength than would otherwise be expected, as observed from their location to the left of the Weibull line. In order to understand this behavior, an analysis of the individual data sets was performed and the results show no significant discrepancies in the parameters, as shown in Table 1. Hence, it was concluded that all of these specimens belong to the same population and therefore can be pooled.

It is important to state that the data are plotted using adjusted ranking methods for censored data (Johnson, 1990), and that the specimen data is scaled and plotted to a prescribed specimen size. The scaling is done on the reliability of the specimen as shown in Eq. (9). This means that specimen size scales vertically in a reliability plot. The Weibull line plotted is obtained from the maximum likelihood Weibull parameters:

$$\frac{\ln(1 - P_f)_i}{\ln(1 - P_f)_C} = \frac{(I_V V)_i}{(I_V V)_C} \quad (9)$$

Surface Failures. In order to obtain the best surface properties on all the specimens, the flexure-type specimens were machined by BOMAS Machine Specialties Inc. of Somerville, MA, and the tensile-type specimens were machined by Allied Manufacturing of Worcester, MA. Since the grinding wheel geometry, feed rate, and grit size were specified to be the same, it was initially assumed that the surface properties of all of the longitudinally machined specimens would be the same; the data were pooled, analyzed, and plotted (Fig. 3). However, the data showed that the sets did not pool well. As a result, an analysis of each of the sets was performed individually. This analysis showed that there are significant differences in the Weibull parameters calculated from the individual data sets, as shown in Table 1.

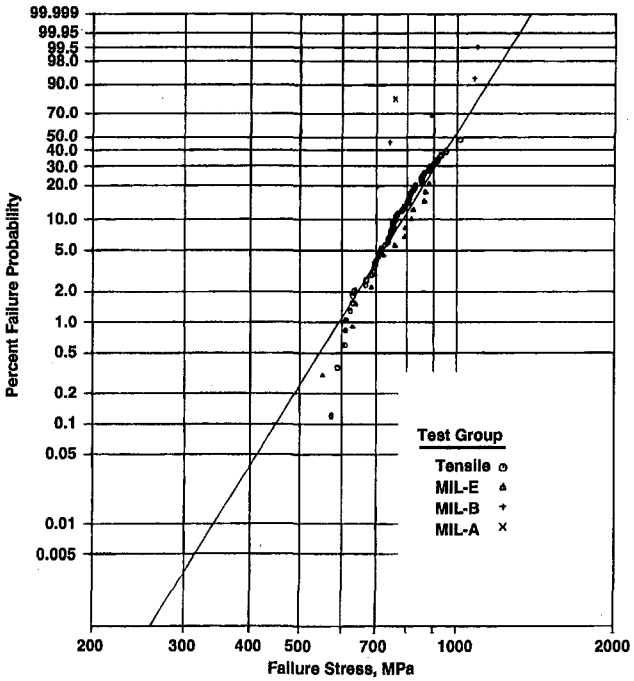


Fig. 2 Weibull plots of the volume distributed failure data. Note that the MIL-A and MIL-B specimens show some deviation from the best estimate of the Weibull curve.

Upon examination of the data, it can be concluded that even though extreme care was taken in the machining of these specimens, there were enough differences in the Weibull parameters that it cannot be safely assumed that the surface flaw distribution is the same between the different specimen sizes.

A problem arises as to which one of the surface properties will be used to predict the component strength. Regardless of which specimen is chosen to predict the component strength, one can not be assured that this was the correct choice. An accurate strength prediction may only mean that the component machining flaw population distribution is similar, but not the same.

An analysis of the transversely machined specimens was also performed to determine the Weibull parameters for this machining direction. Since only one specimen type was machined in this orientation, there is no need to plot the data to show how well the data fit the calculated distribution. The results are summarized in Table 1.

Chamfer Failures. The analysis of chamfer failures is typically not presented in the literature, but it must be recognized that chamfers may represent a real failure mode in components and therefore must be taken into account in any component life prediction. With this in mind, the analysis of the chamfer failures for the three types of flexure specimen is evident as for the surface failure modes; hence, individual specimen analysis was also performed for this failure mode. The results are shown in Table 1. Again, the data indicate that the chamfer failures of these specimens may not belong to the same population.

Table 1 Statistical data analysis results for the various failure modes. Individual and pooled set results are presented.

SPECIMEN	WEIBULL SLOPE (m)			2ND WEIBULL PARAMETER (σ_o), MPa-mm$^{1/m}$		
	Lower C.L.	Best Estimate	Upper C.L.	Lower C.L.	Best Estimate	Upper C.L.
VOLUME FAILURES						
Tensile	6.63	7.97	9.42	1325	1471	1703
MIL-E	5.23	7.98	11.3	1230	1568	2576
MIL-B	2.77	6.62	12.5	1121	1473	4560
MIL-A	1.6	4.1	15.7	980	2020	5120
Pooled	7.18	8.33	9.56	1319	1440	1615
LONGITUDINAL SURFACE FAILURES						
Tensile	5.52	7.60	9.98	1550	1911	2791
MIL-E	6.47	8.06	9.82	1398	1593	1924
MIL-B	8.23	12.16	16.98	1280	1399	1656
MIL-A	16.0	25.73	38.98	1086	1118	1196
Pooled	7.84	8.95	10.1	1451	1560	1712
TRANSVERSE SURFACE FAILURES						
MIL-B	4.76	13.59	26.98	953	1135	2408
CHAMFER FAILURES						
MIL-E	3.2	4.34	5.67	2209	3059	5064
MIL-B	4.42	5.89	7.60	1713	2011	2565
MIL-A	6.60	8.46	10.6	1318	1423	1591
Pooled	5.82	6.64	7.5	1659	1797	1987

C.L.= Confidence Limit

Component Predictions

The capabilities of the multiaxial theories to predict component strength will be presented in this section. The capabilities and accuracy of the theories by Batdorf and Crose and by Evans are put to test, since the components used in this exercise contain very complex stress fields, with large stress gradients and significant multiaxial stress conditions. Component predictions were performed only for the volume and surface failure modes. Analysis of the chamfer failure mode was not performed on the confirmatory specimens because the stresses operating at these locations were either zero or very low or because the specimens did not have any chamfers.

Volume Failure Prediction. Out of the five confirmatory specimen types used in this analysis, three exhibited internal failures: the tension-torsion, plate bending, and spin disk specimens. The tension-torsion specimens address the capabilities of the multiaxial theories in predicting failure for a tension-compression state of stress; the plate bending specimens address the capabilities to predict failure for a tension-tension state of stress; and the spin disk addresses the capabilities of the theories to predict failure for much larger components with large stress gradients and significant tension-tension states of stress.

The results of the predictions are shown in Figs. 4–6. The figures show that volume-distributed failure modes are well predicted using the Weibull theory and multiaxial theories. Limited agreement is shown for the pure-torsion specimens (Fig. 4); this may be an indication of the sheer sensitivity of the material. Tension-torsion specimens loaded under an axial-to-torsion stress ratio of 0.5 showed a much better agreement than the pure-torsion specimens (test data are not shown). Figure 5, for the plate bending specimens, shows excellent agreement even when only a single volume failure was obtained out of 44 specimens. The successful prediction of the volume failure modes in most of these different specimens provides the evidence required to prove the validity of the multiaxial failure theories.

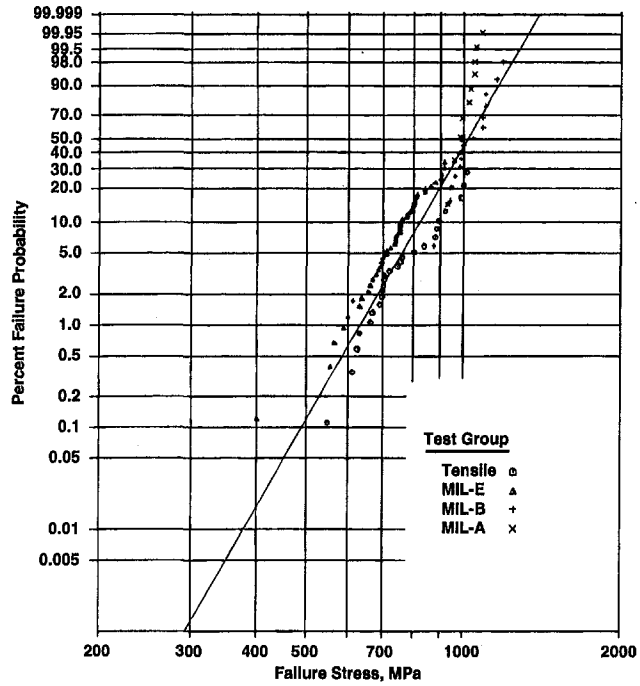


Fig. 3 Weibull plot of the pooled surface distributed failure data

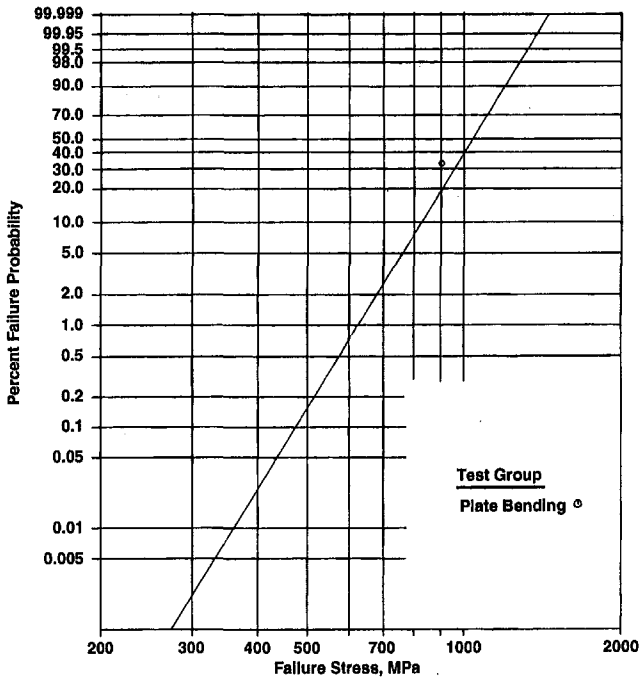


Fig. 5 Excellent agreement is observed in prediction of volume-distributed failure mode for plate bending specimens, even when only a single volume failure occurred in the specimen set

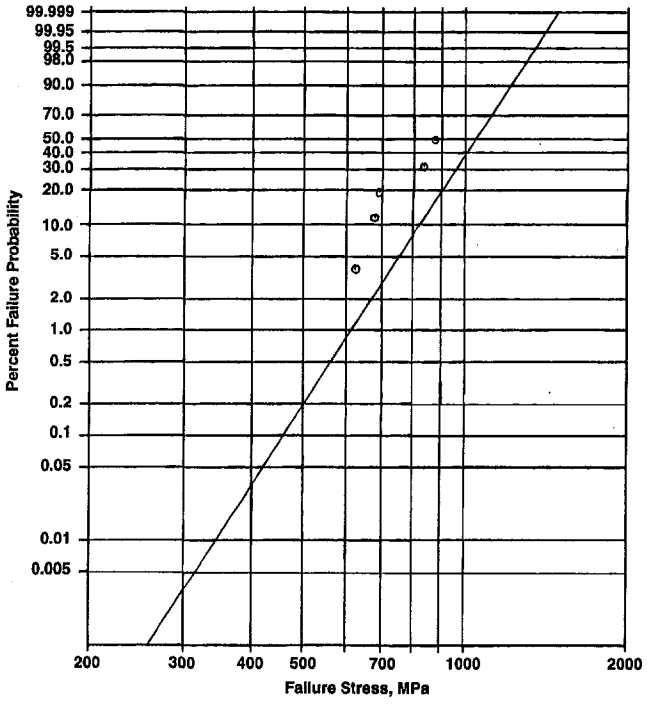


Fig. 4 Prediction of volume distributed failure mode for the pure torsion specimens

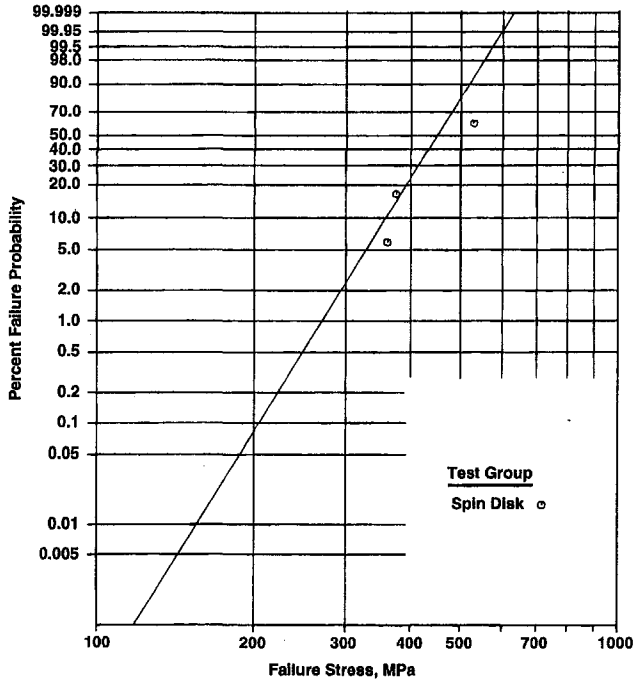


Fig. 6 Good agreement is observed in prediction of volume-distributed failure mode for spin disk specimens

Surface Failure Predictions. All of the confirmatory specimens tested exhibited surface failures. The tension-torsion, plate bending, and spin disk specimens were analyzed assuming the surface properties relate to the longitudinally machined surface. The two notched specimens were assumed to relate to the transversely machined specimens.

Since the database specimens did not show agreement in the longitudinal surface strength properties, it was decided to use the tensile specimen properties for the tension-torsion specimen predictions and the MIL-E bar properties for the plate bending specimen predictions. Since none of the database specimens were machined exactly the same as the spin disks, the tensile specimen data was used for these predictions. The results in Fig. 7 show good agreement in the pure-torsion specimen surface failure mode prediction, which should be expected since the tensile and the tension-torsion specimens were machined exactly the same by Allied Manufacturing. The surface life predictions for the tension-torsion specimen with stress ratio $\sigma/\tau = 0.5$ are as good as those for the pure-torsion specimens,

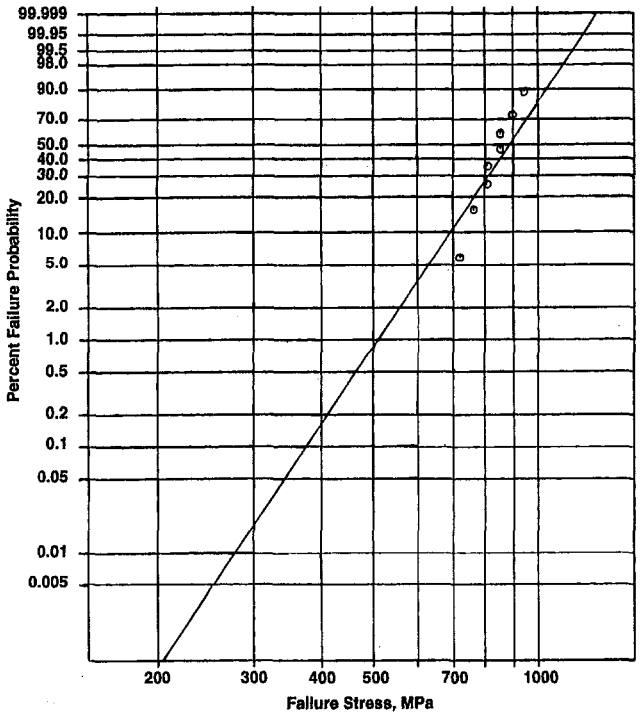


Fig. 7 Good agreement is observed in prediction of surface distributed failure mode for pure torsion specimens

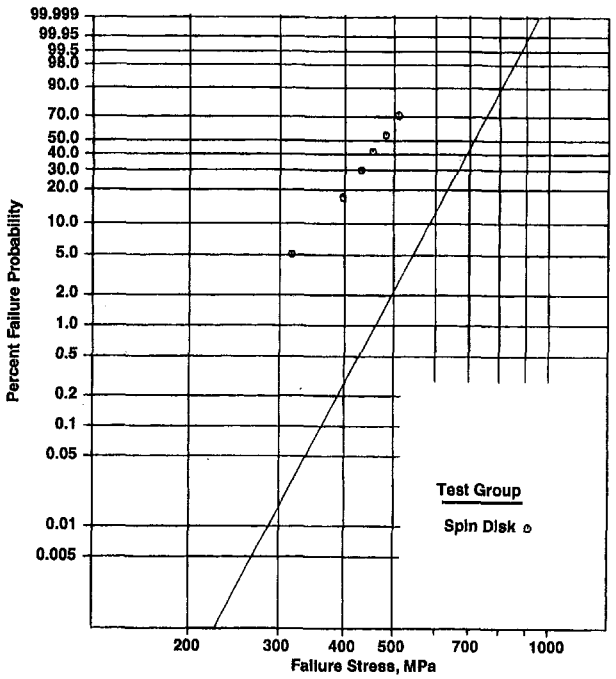


Fig. 9 Spin disk surface properties exemplify the differences that may exist between database-type specimens and actual component(s)

but are not presented here. The surface failure mode prediction for the plate bending specimens is shown in Fig. 8; good correlation is observed between the MIL-E specimen properties and the plate bending specimen data. Here again, an indication of the validity of the multiaxial failure theories is provided.

The spin disk predictions shown in Fig. 9 highlight the most important lesson learned in this program: that the machined surface strength is very sensitive to the machining process variables. Figure 9 serves to prove that the variability in the database specimens (Fig. 2) is real and strongly cautions that evaluation of machined surface properties of components must be carefully considered. One must make sure that the specimen properties accurately reflect the component properties. This suggests that if possible, the database specimens should be machined from actual components.

The two different notched specimens were machined with a shaped grinding wheel, in an attempt to make the notch diameters the same for the two specimens. The *original* machining

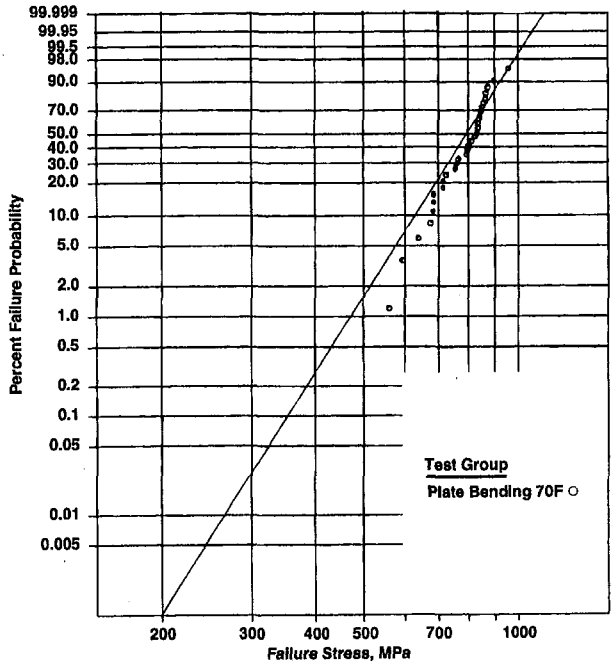


Fig. 8 Good agreement is observed in prediction of surface distributed failure mode for plate bending specimens (assuming correlation to the MIL-E specimens)

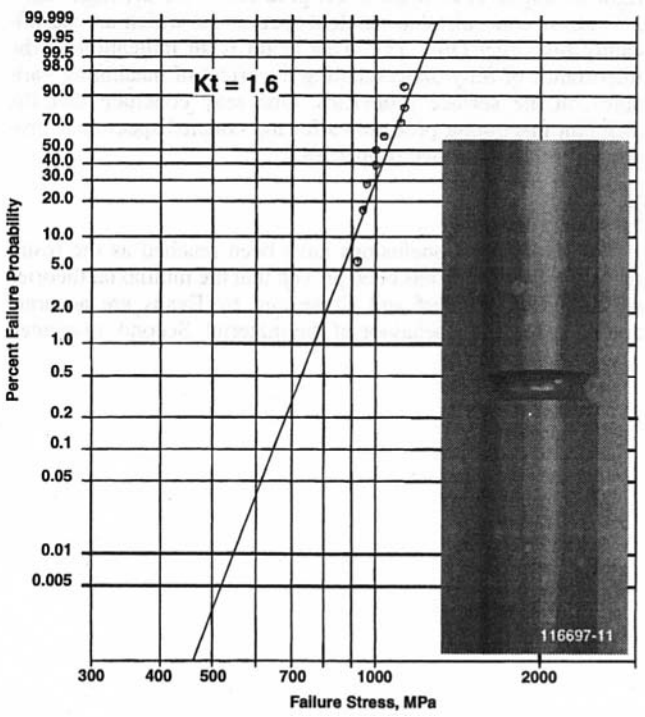


Fig. 10 Transverse surface properties of $KT = 1.6$ notched specimens correlate well with transversely machined MIL-B specimens

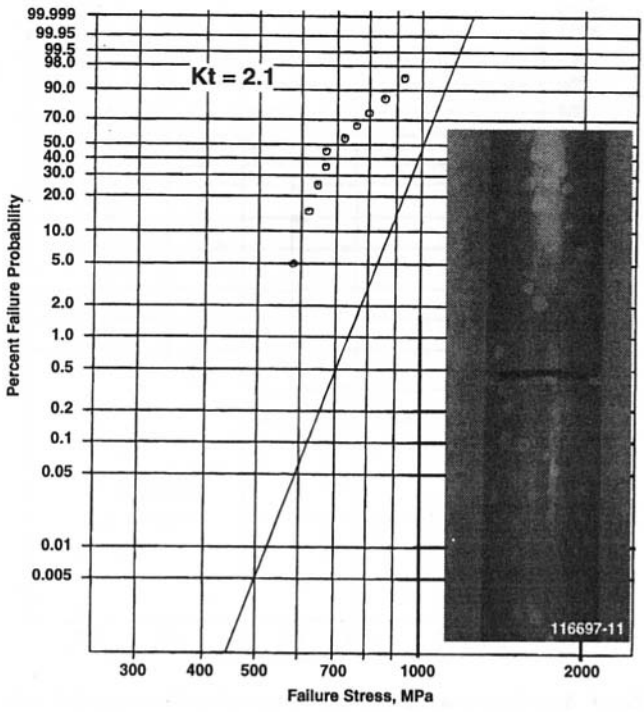


Fig. 11 Transverse surface properties of K_t = 2.1 notched specimens deviate from MIL-B specimen properties

plan for these specimens called for initial machining of *all* the specimens to the stress concentration factor (K_t) = 2.1 configuration, followed by further machining of the K_t = 1.6 notches. But the *actual* machining of the K_t = 1.6 specimens was performed by machining the notch directly into tensile specimen blanks. This minor change in the machining process steps clearly showed up as a differing material strength value for those specimens. As shown in Fig. 10, the failure prediction for the low (K_t = 1.6) stress concentration notch specimens is right on target, in contrast to the predictions for the high (K_t = 2.1) stress concentration notched specimens, which are significantly off target (Fig. 11). This again is an indication of the importance of fully understanding the effect of machining variables on the surface properties. One may conclude that the different machining procedures for the notched specimens produced different surface properties.

Conclusions

Two important conclusions have been reached as the result of this work. First, it has been proven that the multiaxial theories developed by Batdorf and Crose and by Evans are accurate and do capture the behavior of the material. Second, machined surface properties are extremely sensitive to the machining operations, and care must be exercised when component predictions are made based on database-type specimen data alone. Component life predictions must be based on data that truly represent the properties of the *actual* components. Many times the only way to generate *accurate* baseline properties is by cutting specimens from actual components, with the understanding that these specimens must conserve the component surface conditions.

Acknowledgments

We would like to thank the U.S. Department of Energy for financial support, and Dr. C. Brinkman of the Oak Ridge National Laboratories for program direction; Mr. H. T. Fang for his valuable fractographic examination of the specimens; Dr. D. Shetty of The University of Utah for testing the plate bending specimens; and Dr. S. Suresh of Brown University for the testing of the tension-torsion specimens. We would also like to thank Dr. C. Johnson and Dr. W. Tucker of the General Electric Corporate Research Center for their valuable cooperation in this project.

References

Batdorf, S. B., and Crose, J. G., 1974, "A Statistical Theory for the Fracture of Brittle Structures Subjected to Nonuniform Polyaxial Stresses," ASME *Journal of Applied Mechanics,* Vol. 41, pp. 459–464.

Batdorf, S. B., and Heinisch, H. L., 1978, "Weakest Link Theory Reformulated for Arbitrary Failure Criterion," *Journal of the American Ceramic Society,* Vol. 61, Nos. 7–8, pp. 355–358.

Cox, D. R., and Oakes, D., 1984, *Analysis of Survival Data,* Chapman and Hall, NYC, Chapter 3.3.

Cuccio, J. C., Peralta, A. D., Wu, D. C., Fang, H. T., Brehm, P. J., Menon, M. N., Strangman, T., Meade, W., and Wade, J., 1994, *Life Prediction Methodology for Ceramic Components of Advanced Heat Engines Phase I Final Report,* Dept. of Energy/Oak Ridge Natl. Laboratory Report No. ORNL/Sub/89-SC674/1/V1 and V2, March 1995; Report No. 31-11591 (two volumes), AlliedSignal Engines, Phoenix, Arizona, Mar. 1.

Efrom, B., and Tibshiriami, R., 1986, "Bootstrap Methods for Standard Errors, Confidence Intervals, and Other Measures of Statistical Accuracy," *Statistical Science,* Vol. 1, pp. 54–57.

Evans, A. G., 1978, "A General Approach for the Statistical Analysis of Multiaxial Fracture," *Journal of the American Ceramic Society,* Vol. 61, No. 7–8, pp. 302–308.

Johnson, C. A., 1990, "Bootstrap Techniques for Confidence and Tolerance Bounds," General Electric Corporate Research and Development, Oct. 10; in Cuccio et al., 1994 (op. cit.), AlliedSignal Engines Phase I Final Report No. 31-11591 (Vol. 2), Appx. VI, Sect. 2, pp. 41–46.

Nemet, N. N., Manderscheid, J. M., and Gyekenyesi, J. P., 1990, *(CARES): Users and Programmers Manual, Ceramic Analysis and Reliability Evaluation of Structures,* NASA Technical Paper 2916, Aug.

Peralta, A. D., 1991, "Derivation of Statistical Models: Mid-Program Summary Report," Report No. 32-3722, AlliedSignal Auxiliary Power, Phoenix, AZ, Aug. 22; reprinted in Cuccio et al., 1994 (op. cit.), AlliedSignal Engines Phase I Final Report No. 31-11591 (Vol. 2), Appx. VI, Sect. I, pp. v, 1–40.

Tucker, W. T., 1990, "Likelihood Ratio Methods for Ceramic Life Prediction," General Electric Corporate Research and Development, Oct. 10; in Cuccio et al., 1994 (op. cit.), AlliedSignal Engines Phase I Final Report No. 31-11591 (Vol. 2), Appx. VI, Sect. 3 and 4, pp. 47–65.

Weibull, W., 1939, "A Statistical Theory of the Strength of Materials," *Royal Swedish Academy of Engineering Science and Proceedings,* Vol. 151, pp. 1–45.

J. A. Salem

N. N. Nemeth

NASA Lewis Research Center,
Cleveland, OH 44135

L. M. Powers

S. R. Choi

Department of Civil Engineering,
Cleveland State University,
Cleveland, OH 44135

Reliability Analysis of Uniaxially Ground Brittle Materials

The fast fracture strength distribution of uniaxially ground, alpha silicon carbide was investigated as a function of grinding angle relative to the principal stress direction in flexure. Both as-ground and ground/annealed surfaces were investigated. The resulting flexural strength distributions were used to verify reliability models and predict the strength distribution of larger plate specimens tested in biaxial flexure. Complete fractography was done on the specimens. Failures occurred from agglomerates, machining cracks, or hybrid flaws that consisted of a machining crack located at a processing agglomerate. Annealing eliminated failures due to machining damage. Reliability analyses were performed using two and three-parameter Weibull and Batdorf methodologies. The Weibull size effect was demonstrated for machining flaws. Mixed mode reliability models reasonably predicted the strength distributions of uniaxial flexure and biaxial plate specimens.

Introduction

A basic requirement for reliability analysis of components made from brittle materials is a knowledge of the strength distribution of the flaws that will induce failure. Two basic types of strength-limiting flaw are encountered: surface defects and volumetric defects. Volumetric defects include large grains, pores, agglomerates, coarse precipitates, and inclusions, while surface defects could include exposed volume defects (e.g., a pore machined open) and machining or handling damage occurring during fabrication [1].

In some cases the failure-inducing flaws are distributed randomly and thus the strength distribution is independent of orientation and simple, uniaxial test data can be used in conjunction with an isotropic model. However, for components made by processes such as extrusion, which induce texture, a bias in the distribution of processing flaws will exist. Also, components finished by surface grinding will also contain machining damage in the form of cracks that are oriented parallel and transverse to the grinding direction. Fortunately for the designer, processes typically used to make larger, three-dimensional components do not induce strong textures. However, uniaxial grinding is used for finishing of components and to make test specimens for strength measurements.

The strength of a ceramic material is typically measured in accordance with ASTM C1161 [1], which specifies that machined specimens be ground uniaxially in the longitudinal direction and tested so that the maximum principal stress is longitudinal. Such a grinding process typically induces minimal damage in the transverse direction, but significant damage in the longitudinal direction, resulting in an anisotropic flaw distribution on the surface of the specimen. Since the beam is stressed longitudinally, such a preparation is typically sufficient to avoid failures from machining damage, and result in measurements that are representative of strength limiting defects associated with the materials processing (e.g., agglomerates, inclusions, pores, coarse grains, etc.). However, structural components and multiaxial test specimens are subjected to multiaxial stresses and thus are sensitive to planar flaws with the crack plane oriented in any direction.

Thus, if uniaxial grinding that imparts damage in one direction is used or if anisotropic damage occurs, component design and life prediction methods must account for such anisotropies. The objective of this work is to measure the effect of unidirectional grinding on the strength distribution of a ceramic material under various loading conditions and determine whether adequate component reliability models could be formulated.

The fast-fracture strength of a sintered alpha silicon carbide was measured in four-point flexure with the principal stress oriented at angles between 0 and 90 deg relative to the grinding direction. Also, uniaxially ground plate specimens were loaded in biaxial flexure. Finally, flexure bar specimens were tested in an annealed condition to determine whether the machining damage could be healed. Modeling of the strength distributions was done with two and three-parameter Weibull models and shear sensitive and insensitive cracks. Alpha silicon carbide was chosen because it exhibits a very low fracture toughness, no crack growth resistance, high elastic modulus, and a very low susceptibility to slow crack growth. Such properties should make silicon carbide an ideal ceramic for the verification of fast fracture reliability models and codes.

Experimental Method

The material used in this study was a commercially available sintered alpha silicon carbide (Carborundum's Hexoloy) processed in the form of 25 by 25 by 42 mm billets. Several sets of plates were ground from the billets and finished by 320 grit diamond grinding one face of the plates at angles ranging from 0 to 90 deg relative to the plate edge. Flexure beams were then cut from the plates. Beams with a given grinding angle were cut from a random selection of plates in order to block the effects of billet and location on the test results. The specimen edges along the tensile surface were beveled by hand to eliminate spurious edge failures. The specimens nominally measured 2 by 3 by 25 mm in height, width, and length, and were tested in four-point flexure with inner and outer spans of 8 and 20 mm, respectively, at room temperature with a stroke rate of 0.05 mm/min. The rollers of the test fixture were free to roll and the upper span was free to articulate relative to the lower span. A minimum of 30 specimens were tested per condition.

In order to determine whether the deleterious effects of grinding damage could be negated, a group of 0 and 90 deg specimens were annealed at 1200°C for two hours in air prior to testing.

Results and Discussion

Strength Distributions. Specimen strength as a function of grinding angle relative to the longitudinal direction is shown

Contributed by the International Gas Turbine Institute and presented at the 40th International Gas Turbine and Aeroengine Congress and Exhibition, Houston, Texas, June 5–8, 1995. Manuscript received by the International Gas Turbine Institute February 8, 1995. Paper No. 95-GT-31. Associate Technical Editor: C. J. Russo.

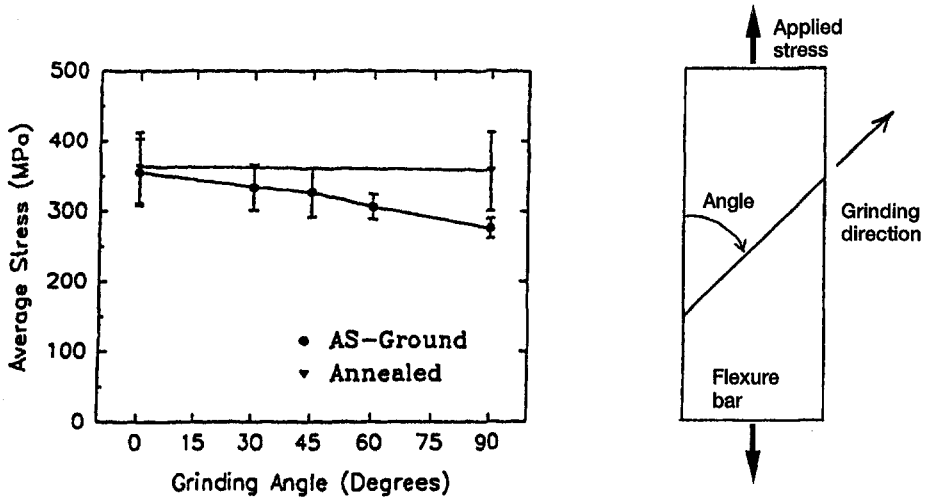


Fig. 1 Average four-point flexure strength as a function of grinding angle. Error bars are ± one standard deviation.

in Fig. 1 and summarized in Table 1. The average strength and standard deviation decrease continually as the angle increases, with exception of the 30 and 45 deg data, which are quite similar. Annealing does not significantly change the average strength of specimens ground at 0 deg (i.e., longitudinally). However, specimens ground at 90 deg (i.e., transversely) and annealed exhibited strengths not significantly different from those of the 0 deg annealed and 0 deg as-ground specimens. This implies that significant, strength-limiting damage relative to the processing flaws, for this material, only exists parallel to the grinding direction (such that specimen strength is affected when loading is transverse to the grinding direction), and that annealing eliminates the grinding damage but does not significantly enhance the strength of 0 deg ground specimens.

Weibull plots of the annealed and as-ground specimens are shown in Figs. 2(*a*)–2(*d*). The Weibull modulus continuously increases with increasing grinding angle while the characteristic strength decreases. Annealing does not significantly change the distribution parameters of 0 deg ground specimens, and annealing appears to totally heal machining damage associated with the 90 deg ground specimens. As all these data sets lay will within each others 90 deg confidence intervals, the annealed specimen data and the 0 deg as-ground data were pooled.

Fractography. Fractographic analysis is a necessary aspect of reliability analysis in order to determine whether surface, volume, or combined (surface and volume) flaw reliability analysis should be performed, and if flaws of different processing sources are present. Fractography to determine the sources and locations of failure was done in accordance with military handbook procedures [2]. The results are summarized in Table 2. Three types of flaw were encountered: processing flaws, machining cracks, and hybrid flaws. The processing flaws that induced failure were agglomerates located at the tensile surface (i.e., cut open during machining) or located slightly below (within one flaw diameter). Occasionally (< 10 percent of the time) failure occurred from agglomerates located at a distance larger than the flaw diameter below the surface. These failures

Nomenclature

A = surface area
$\bar{C}$ = Shetty's constant in mixed-mode fracture criterion
H = step function
h = thickness
K = stress intensity factor
k = crack density coefficient
k_{BS} = Batdorf crack density coefficient for surface distributed flaws
l = direction cosine = cos α
m = Weibull modulus (shape parameter), or direction cosine = sin α
m_S = Weibull modulus for surface flaws
P = load
P_f = cumulative failure probability
R_s = diagonal half length
r = radius
x, y, z = Cartesian coordinate directions
Y = crack geometry factor
α = angle between σ_n and the stress σ_1
η = crack density function
π = 3.1416
ν = Poisson ratio
σ = applied stress distribution
σ_{cr} = critical mode I strength of a crack
σ_e = effective (equivalent) mode I stress on a crack
σ_o = Weibull scale parameter
σ_u = threshold strength
$\sigma_1, \sigma_2, \sigma_3$ = tensor stress components; principal stresses ($\sigma_1 \geq \sigma_2 \geq \sigma_3$)
τ = shear stress acting on oblique plane whose normal is determined by angle α
Ψ = spatial location (x, y, z) and orientation (α) in a component
ω = angle in two-dimensional principal stress space for which $\sigma_e \geq \sigma_{cr}$

Subscripts

B = Batdorf
cr = critical
e, ef = effective
eq = equivalent
f = failure; fracture
I = crack opening mode
II = crack sliding mode
III = crack tearing mode
i = inside
max = maximum
n = normal; normal stress averaging
o = outside
S = surface
u = threshold
θ = characteristic

Superscripts

$-$ = normalized quantity

Table 1 Summary of strength data

Grinding Angle	Number Tested	Range (MPa)	Average[1] (MPa)	Characteristic Strength (MPa)	Weibull[2] Modulus
As-Ground					
0°	36	266-458	356±47	377	8.2
30	34	271-400	334±32	348	11.6
45	32	267-392	327±35	341	10.5
60	31	273-331	306±17	314	21.3
90	35	246-308	276±14	282	19.8
Annealed					
0°	36	265-465	362±51	384	8.3
90	36	222-481	357±56	380	6.8

[1]Average ± one standard deviation.
[2]Maximum likelihood estimator.

Table 2 Summary of failure origins

Grinding Angle	Surface and Near Surface Agglomerates	Volume Agglomerates	Machining Damage	Hybrid Flaws	Not Identifiable
As-Ground					
0°	31	1	0	3	0
30	30	1	0	4	0
45	21	0	9	2	0
60	11	0	17	3	0
90	0	0	33	2	0
Annealed					
0°	32	3	0	0	1
90	31	3	2	0	0

are classified as "volume" in Table 2. Machining cracks were semi-elliptical and cartoidal cracks typically oriented parallel to the grinding direction. Hybrid flaws were defined as a machining crack connected with an agglomerate. It is not clear whether hybrids constituted a distinct failure population and their overall numbers are relatively small. The various types of origins are shown in Figs. 3 and 4. Although both the processing and machining flaws are complex shapes, the processing flaws are typically larger in overall size, indicating a bluntness relative to the machining flaws.

Annealed specimens and 0 deg as-ground specimens appear to fail from surface and near surface processing agglomerates, while 30, 45, 60, and 90 deg as-ground specimens fail from either surface and near-surface agglomerates, machining cracks, or hybrid flaws. At lower grinding angles, final fracture of hybrids occurred from the agglomerate and at higher angles fracture occurred from the machining crack (i.e., the river markings

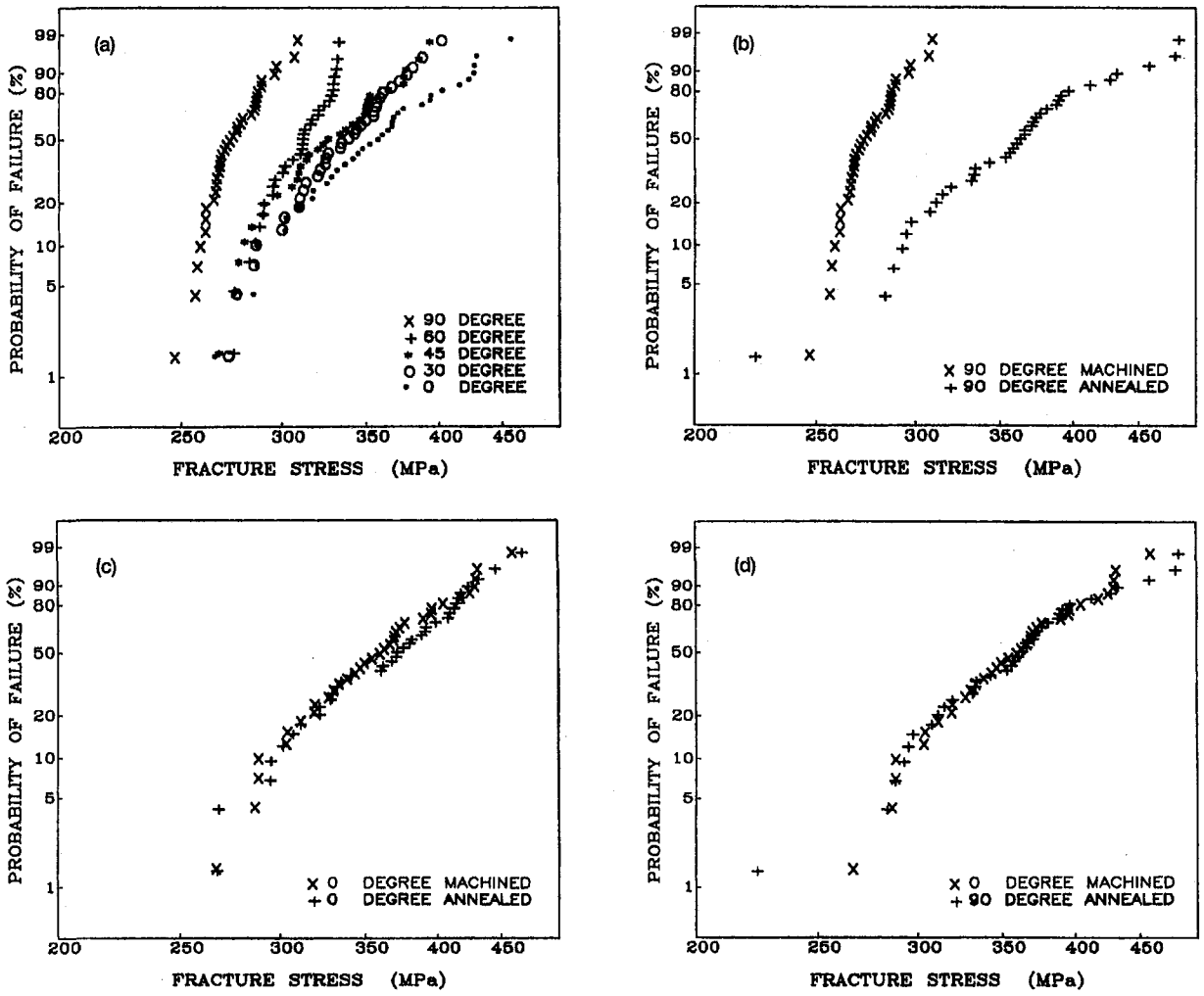
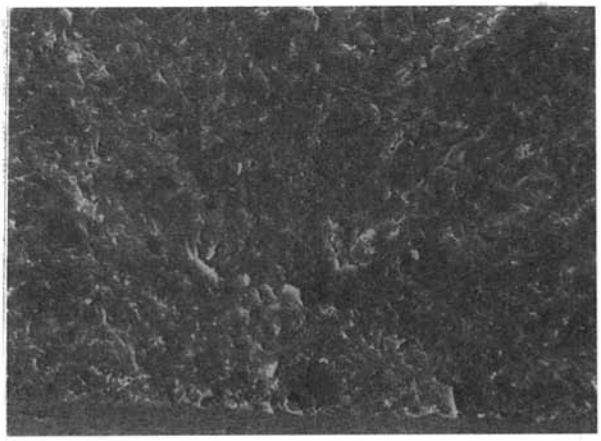
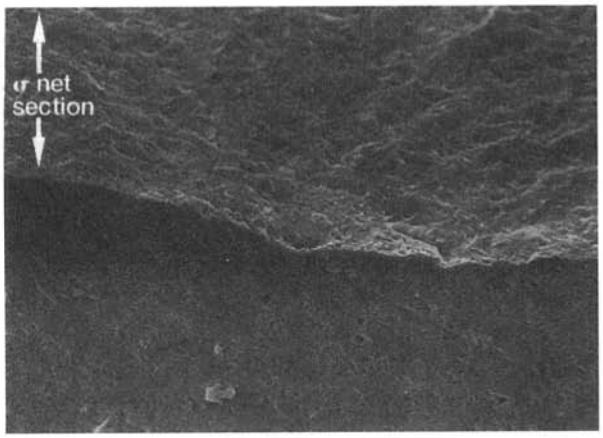


Fig. 2 Weibull plots of the fracture stresses for (*a*) all as-ground specimens, (*b*) 90 deg annealed and as-ground specimens, (*c*) 0 deg annealed and as-ground specimens, and (*d*) 90 deg annealed and 0 deg as-ground specimens



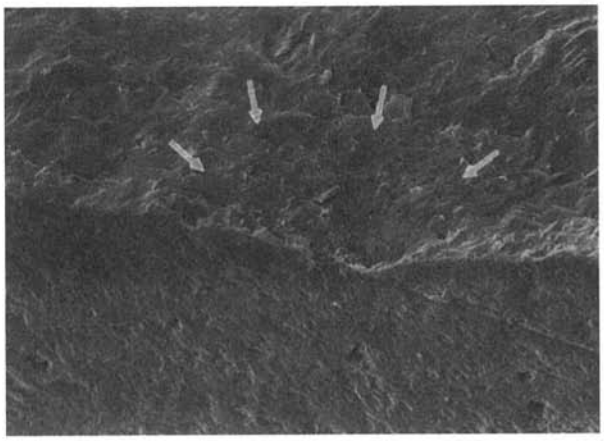
(a)

50 μm



(b)

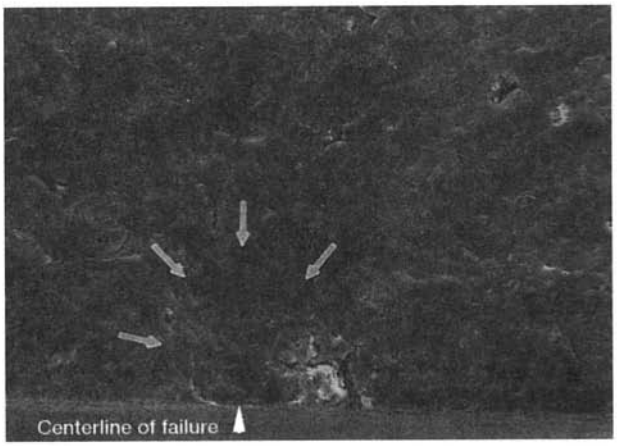
50 μm



(c)

20 μm

Fig. 3 Failure origins showing (*a*) surface connected processing agglomerate, (*b*) 30 deg tilt, machining crack, mirror and machining scratch, (*c*) 30 deg tilt, machining crack detail

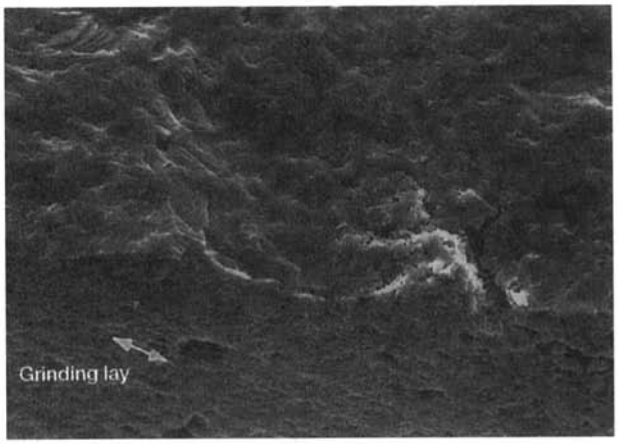


(a)

50 μm

(b)

20 μm



(c)

50 μm

Fig. 4 Hybrid flaw and detail: (*a*) river marks, machining crack and agglomerate, (*b*) machining crack; and (*c*) grinding lay

point to either the agglomerate or the machining crack, respectively). As the grinding angle increased, the number of agglomerates associated with failure decreased.

Note (Fig. 3(*b*)) how the crack path from a 30 deg machining flaw is initially parallel to the grinding direction and gradually turns normal to the principal stress direction on one side of the flaw, but not on the other. The river marks indicated that fracture initiated from the deepest point of the crack and followed the crack plane (coplanar extension) to the surface where it gradually turned normal to the applied stress. Evidently, mixed stress intensity modes exist during such failures, but the amount of

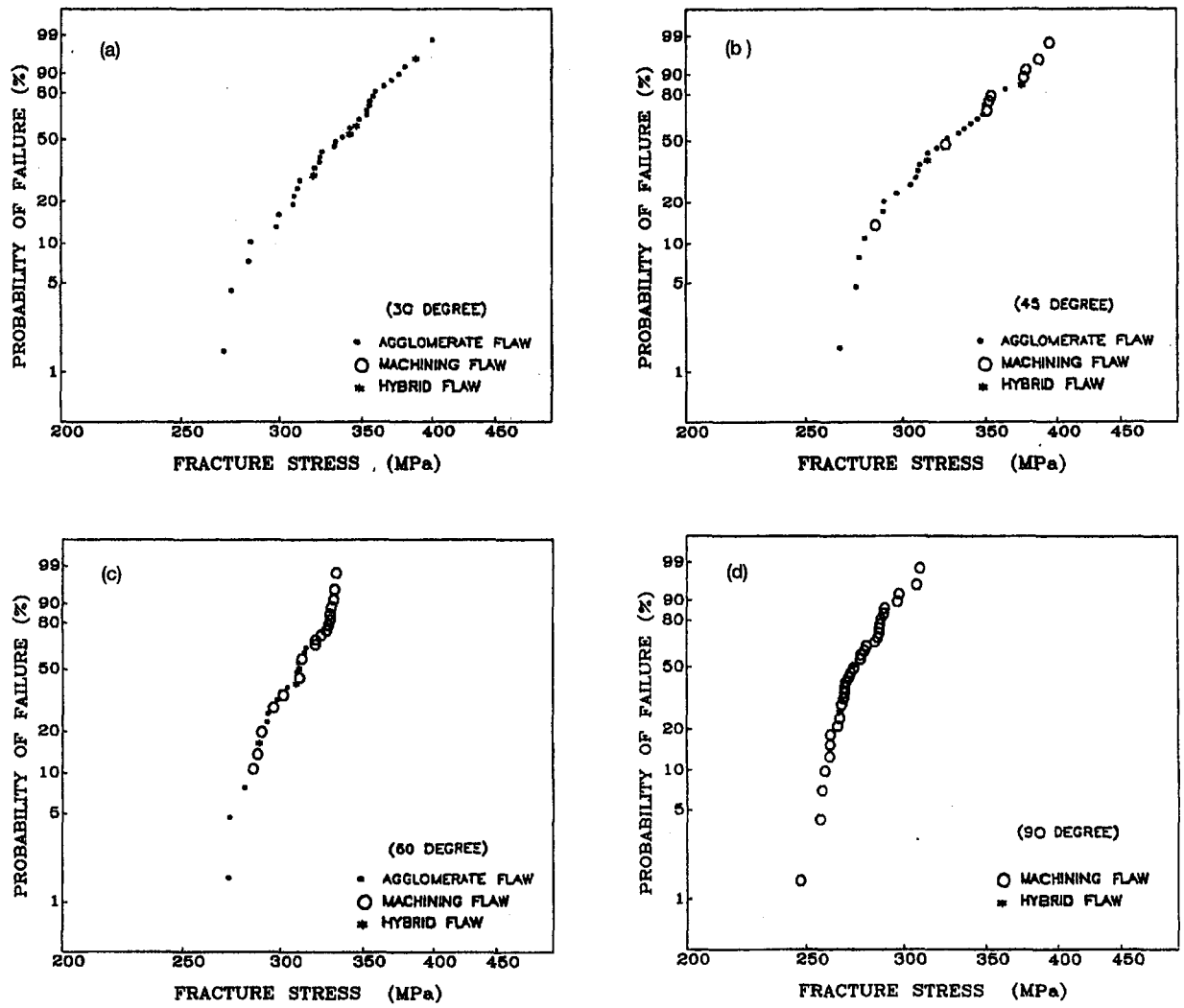


Fig. 5 Weibull plots of the fracture stresses for 30, 45, 60, and 90 deg as-ground specimens showing the different failure sources

extension under mixed modes is unclear. Also, the macroscopic fracture plane can occasionally be observed to jog parallel to the grinding direction in regions well away from the origin. A machining crack can typically be observed at such jogs.

Note also (Fig. 4) that the mirror of the hybrid flaw centers on the machining crack and that it dominated failure. Coplanar extension is not apparent at the surface; however, it is apparent in the interior.

Modeling of the Strength Distribution. As a result of different flaw populations (agglomerate and machining flaw) being simultaneously present in any given specimen, the contribution of each flaw population to the cumulative failure probability of the specimen must be considered. The survival probability of the specimen from these simultaneously present (concurrent) flaw populations is the product of the survival probabilities of the specimen from each flaw population considered separately. Weibull plots of the concurrent populations for 30, 45, 60, and 90 deg grinding angles are shown in Fig. 5. Note that as the grinding angle increases, the frequency of failure at machining flaws increases and that they populate the upper end of the overall distributions. No distinct pattern for the distribution of hybrid flaws is apparent in Fig. 5.

Two and three-parameter Weibull formulations were used to represent the strength distribution. Response of the flaw population to multiaxial stresses was predicted with shear sensitive and insensitive flaw (Batdorf) models. These models were modified such that an anisotropic flaw orientation or a threshold/truncation stress was possible. The overall strength distribution was modeled as a bimodal distribution of agglomerate flaws and machining flaws. The effects of hybrid and volume flaws were not considered, as these populations were small.

Agglomerate Flaws. The Batdorf [3–5] theory was used to describe the strength response of the ground surface. The agglomerate flaws were modeled as randomly distributed and randomly oriented microcracks with the exception that the crack plane was perpendicular to the surface. The probability of failure for a ceramic component using the Batdorf model for such surface flaws is [4]

$$P_{fs} = 1 - \exp\left\{ - \int_A \left[\int_{\sigma_u}^{{}^{\sigma}e_{\max}} \frac{\omega}{2\pi} \frac{d\eta_s(\sigma_{cr})}{d\sigma_{cr}} d\sigma_{cr} \right] dA \right\} \quad (1)$$

where A is the surface area, η_S is the crack density function, ${}^{\sigma}e_{\max}$ is the maximum value of effective stress, σ_e, for all values of spatial angle Ψ, and ω is the arc length of an angle α projected onto a unit radius circle in stress space containing all of the crack orientations for which the effective stress is greater than or equal to the critical mode I stress, σ_{cr}. The integration limit σ_u represents a threshold value below which no failures exist. This effectively truncates the statistical

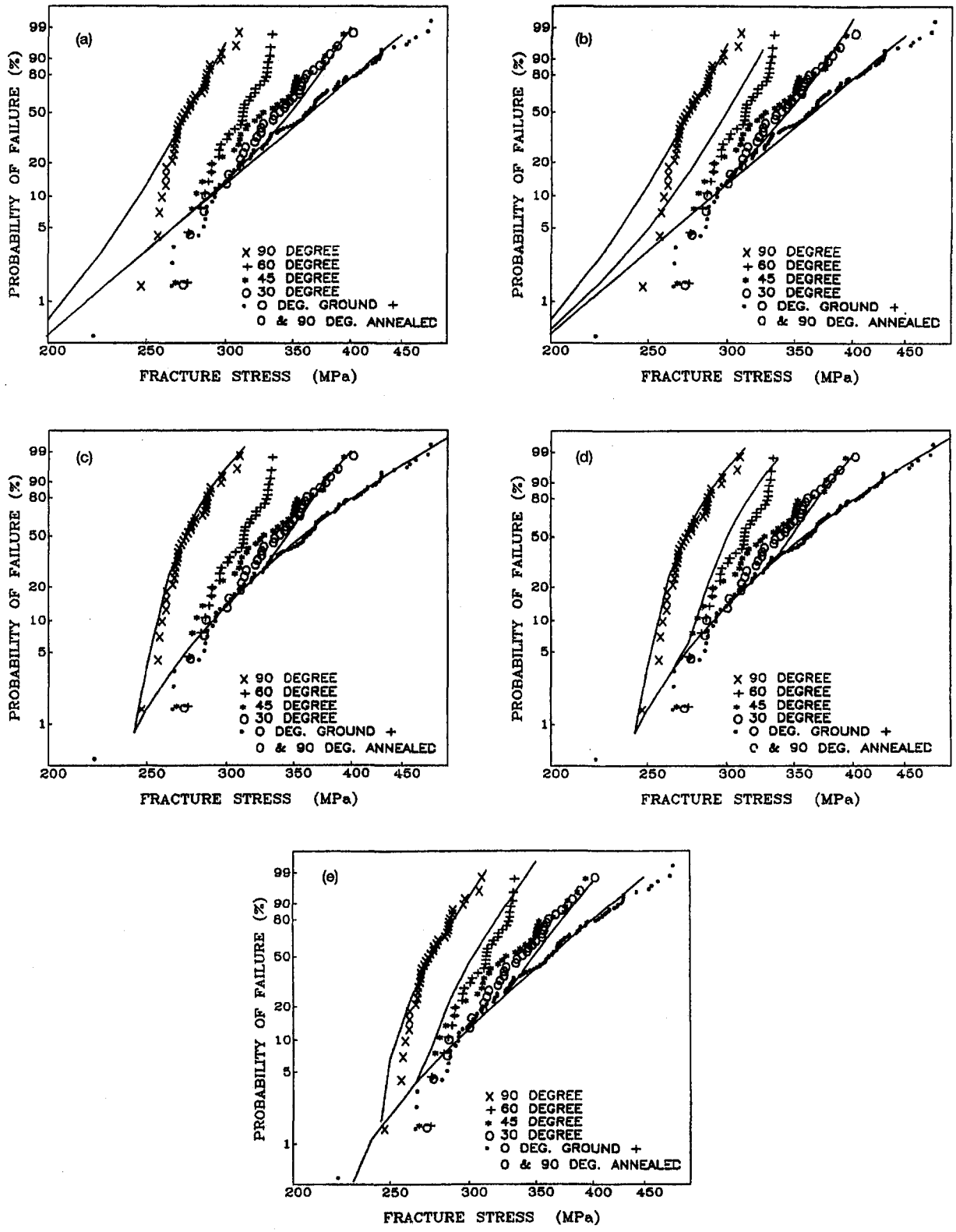


Fig. 6 Experimental data and (*a*) two-parameter shear-insensitive model, (*b*) two-parameter shear-sensitive model, (*c*) three-parameter shear-insensitive model, (*d*) three-parameter shear-sensitive model, and; (*e*) two-parameter shear-sensitive model with a truncation load. In these plots the value of $\bar{C}$ ranges between 1.0 and 1.05 (or approximately 1.0).

strength distribution and is similar to a proof test. However, in this case the truncation may occur naturally or result from processing controls on the maximum flaw size. The crack density distribution is a function of the critical stress distribution. For surface flaw analysis, the crack density function is expressed as

$$\eta_s(\sigma_{cr}(\Psi)) = k_{BS}\sigma_{cr}^{m_s} \tag{2}$$

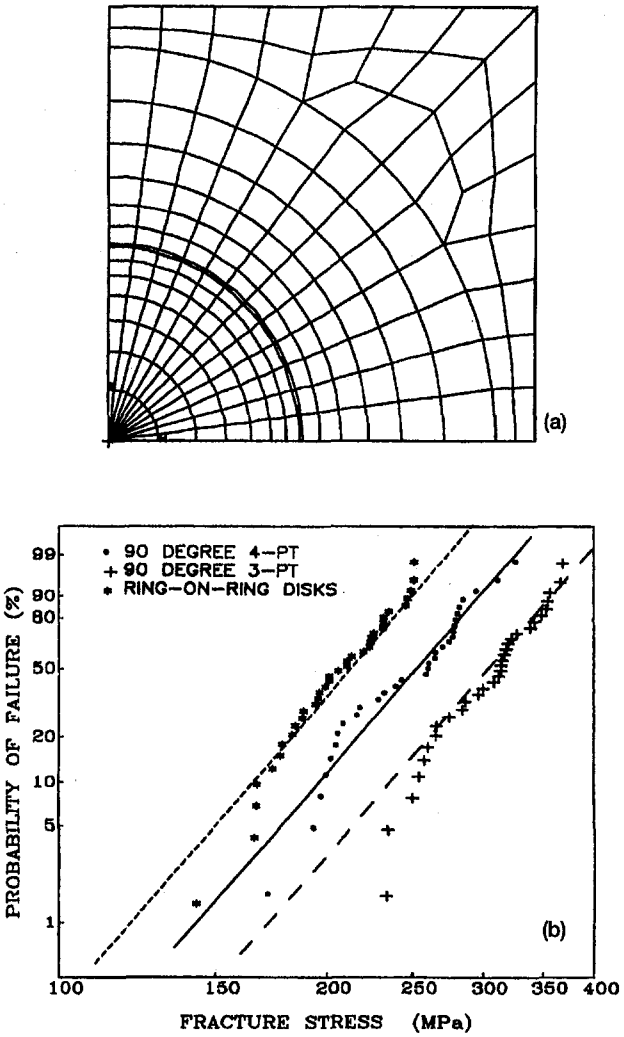


Fig. 7 (*a*) FEM mesh and (*b*) experimental data with MLE fit (solid line) and predictions based on a two-parameter model (dashed lines)

where k_{BS} and m_S are material constants. The flaw orientation function is expressed as

$$\frac{\omega}{2\pi} = \frac{\int_0^{2\pi} H(\sigma_e, \sigma_{cr})d\alpha}{\int_0^{2\pi} d\alpha} \tag{3}$$

where

$$H(\sigma_e, \sigma_{cr}) = \begin{cases} 1 & \sigma_e \geq \sigma_{cr} \\ 0 & \sigma_e < \sigma_{cr} \end{cases}$$

Equation (3) represents the fraction of critically oriented flaws over all possible flaw orientations. The equivalent stress σ_e is dependent on the appropriate fracture criterion and crack shape. Equation (1) can be simplified by performing the integration of σ_{cr} [5], yielding the probability of failure for surface flaw analysis as

$$P_{fs} = 1 - \exp\left[- \frac{k_{BS}}{2\pi} \int_A \int_0^{2\pi} [\sigma_e^{m_s}(\Psi) - \sigma_u^{m_s}] H(\sigma_e, \sigma_u) d\alpha dA \right] \tag{4}$$

When the threshold stress, σ_u, is zero, then Eq. (4) reduces to a two-parameter Weibull type distribution for a specified stress state.

Assuming that small cracklike imperfections control the failure, the material strength in multiaxial stress states can be correlated to the effects of mixed-mode loading on the individual cracks [4, 5] Shetty [6] developed a simple equation describing the ability of a crack to extend under the combined actions of a normal and shear load on the crack face using an empirically determined parameter, $\bar{C}$. For a semicircular crack, the equation for the effective stress is [7]

$$\sigma_e = \frac{1}{2}\left[\sigma_n + \sqrt{\sigma_n^2 + 3.301\left(\frac{\tau}{\bar{C}}\right)^2} \right] \tag{5}$$

where σ_n is the normal stress on the flaw and τ is the shear stress on the flaw. The Batdorf methodology used herein is normalized to uniaxial specimen rupture data [3, 4]. Therefore, the choice of $\bar{C}$ does not affect the failure probability predictions for the flexure bar specimens. For the agglomerates, a value of $\bar{C}$ of 1.0 is assumed (to be consistent with the results from grinding damage). A reference frame relative to machining direction is used to define the normal and shear stress components. The normal and shear stress components are computed from

$$\sigma_n = l^2\sigma_x + m^2\sigma_y + 2lm\tau_{xy} \tag{6}$$

and

$$\tau^2 = (l\sigma_x + m\tau_{xy})^2 + (l\tau_{xy} + m\sigma_y)^2 - \sigma_n^2 \tag{7}$$

where l and m are the direction cosines:

$$l = \cos\alpha$$
$$m = \sin\alpha \tag{8}$$

Comparison of the Batdorf to the following form of the three-parameter Weibull distribution can be performed:

$$P_{fs} = 1 - \exp\left\{ - \int_A \left[\left(\frac{\sigma_1 - \sigma_u}{\sigma_{\theta S}} \right)^{m_S} + \left(\frac{\sigma_2 - \sigma_u}{\sigma_{\theta S}} \right)^{m_S} \right] dA \right\} \tag{9}$$

where σ_1 and σ_2 are the principal stresses, and $\sigma_{\theta S}$ is the characteristic strength.

Machining Flaws. Richerson [8] has described machining flaws as consisting of populations of median (longitudinal) cracks, lateral cracks, and radial cracks (transverse). The longitudinal cracks are parallel to the grinding groove and perpendic-

Table 3 Summary of strength data for model verification

Grinding Angle	Number Tested	Range (MPa)	Average[1] (MPa)	Characteristic Strength[2] (MPa)	Weibull[2] Modulus
As-Ground					
90° (3-point)	31	231-364	303±37	319	9.6 (7.3-12.6)
90 (4-point)	31	172-327	249±39	265 (254-277)	7.4 (5.6-9.0)
Plates	36	142-250	206±28	218	8.5 (6.6-11.0)

[1]Average ± one standard deviation.

[2]Maximum likelihood estimator. Values in parentheses are the 95% confidence interval.

ular to the surface of the material. The radial cracks are perpendicular to the longitudinal cracks and lateral cracks run parallel to, but branch off at some angle from the grinding groove. As the only strength-limiting flaws induced by unidirectional grinding were observed to be oriented parallel to the grinding direction, the flaw distribution, for modeling purposes, can be treated as an anisotropic distribution. The fact that all the strength-limiting machining cracks are parallel to each other eliminates the need for the orientation function $\omega/2\pi$ in Eq. [3] and Eq. (1) is then rewritten as

$$P_{fS} = 1 - \exp\left\{ - \int_A \left[\int_{\sigma_u}^{\sigma_{e_{\max}}} H(\sigma_e, \sigma_{cr}) \frac{d\eta_S(\sigma_{cr})}{d\sigma_{cr}} d\sigma_{cr} \right] dA \right\} \tag{10}$$

Analogous to the formulation shown in Eq. (4), Eq. (10) can be simplified, yielding

$$P_{fS} = 1 - \exp\left[-k_{BS} \int_A (\sigma_e^{m_s} - \sigma_u^{m_s}) dA \right] \tag{11}$$

where σ_e represents the effective stress on the longitudinal machining flaw. The threshold strength σ_u in this case may possibly be related to the largest grit size of the grinding wheel. Since grinding particle sizes are screened, this may translate to a maximum flaw size that can be induced in the material surface.

As the machining cracks tended to be semielliptical, a semicircular flaw geometry was assumed here and Eq. (5) applied. A reference frame relative to grinding direction was used to define the normal and shear stress components.

Since the machining cracks are oriented in one direction, comparison to a three-parameter Weibull distribution is feasible:

$$P_{fS} = 1 - \exp\left\{ - \int_A \left(\frac{\sigma_e - \sigma_u}{\sigma_{\theta S}} \right)^{m_S} dA \right\} \tag{12}$$

where σ_e is the effective stress in Eq. (5), and $\sigma_{\theta S}$ is the characteristic strength.

Other Modeling Considerations. As a result of the agglomerates and machining flaws being simultaneously active, a multimodal distribution function was used to describe the overall probability of failure

$$P_{fS_{\text{total}}} = 1 - (1 - P_{fS_{\text{agglomerate}}})(1 - P_{fS_{\text{machining}}}) \tag{13}$$

The finite element method or a simple numerical procedure can be used for stress analysis as they enable discretization of a component or specimen into incremental volume (and surface) elements from which the probability of failure (i.e., Eq. (13)) can be evaluated. Either the Gaussian integration points of the elements or, optionally, the element centroids can be used to evaluate the stress-area integrals [9] numerically in Eqs. (4), (9), (11), and (12). Assuming that the probability of survival for each element is a mutually independent event, the overall component reliability is then the product of all the calculated element (or subelement) survival probabilities.

A simple numerical model of one half of the ground tensile surface of the flexure bar was prepared for the strength predictions. The model consisted of a single subarea between the inner loading points (where a constant stress exists) and 600 equispaced subareas between the inner and outer loading point to capture the linearly varying stress along the specimen length. Each subarea was assumed to have a constant stress state. This model is analogous in effect to a finite element model of the tensile surface of the specimen where each subarea would correspond to a surface element and the element centroidal stress is used in the reliability analysis.

The bar surface stress distribution and preceding equations were used to predict the behavior of the 30, 45, and 60 deg orientations based on the best-fit distributions (maximum likelihood estimator) of the 0 and 90 deg data. The resulting distributions are shown in Fig. 6 compared to the actual data for several cases. As shown in the figure, the two-parameter model appears to fit the data poorly at the lower probabilities of failure.

Both the shear-sensitive three-parameter and truncated distribution models are better fits to the data; however, no specific justification for the lower strength boundary can be given. Although truncation of the distribution may be due to largest grit size, this was not proven here. The shear-insensitive models did not correlate well to the data. The shear-sensitive models use a Shetty shear sensitivity coefficient that ranged in value between 1.00 and 1.05.

Prediction of Strength Distributions Resulting From Multiaxial Stresses. The previous results were derived from loading conditions resulting in a uniaxial stress state; however, components are frequently subjected to multiaxial stresses. Thus, a second set of plates was ground and tested in biaxial flexure via ring-on-ring loading. As both the specific grinding damage and agglomerates within these plates were expected to control strength, uniaxial (90 deg) specimens were also cut from these plates and tested in either four-point flexure as described above or in three-point flexure with a 20 mm support. In order to test the capability of the models to account for area changes, multiaxial stresses, and machining damage along one axis, a quarter symmetry, finite element model (Fig. 7(*a*)) of the biaxial specimen was made and the output interfaced with a version of the CARES [7] code containing the previously described models. The finite element model was prepared using MSC/NASTRAN and consisted of 200 solid elements (a single element spanned the plate thickness) and 200 shell elements. The shell elements were attached to the tensile surface of the plate model and had negligible thickness and membrane properties only. The reliability of the plate was determined from the stress and surface area output of the shell elements. Ring-on-ring load induces an equibiaxial stress state within the inner ring. The fracture stress, σ_f, at the specimen center was computed as a function of the fracture load, P, and is given by [10]

$$\sigma_f = \frac{3P}{4\pi h^2} \left[2(1 + \nu) \ln \left(\frac{r_o}{r_i} \right) + \frac{(1 - \nu)(r_o^2 - r_i^2)}{R_S^2} \right] \tag{14}$$

where R_S is one-half the average of the diagonal and edge lengths, r_i is the inner radius, r_0 is the outer radius, h is the thickness, and ν is Poisson ratio.

The results are summarized in Table 3 and shown in Fig. 7(*b*). A distinct effect of scale can be seen when machining damage controls the uniaxial specimen failure. Note that the Weibull parameters of the second set of 90 deg specimens were significantly different (exceeds 90 percent confidence bounds) from those of the first set (Table 1), even though the same grinding specification was used. A more stringent control, other than just the grit size and removal depth specified in ASTM C1161, may be required to control damage adequately parallel to the longitudinal direction. Prediction of the plate failure distribution and the three-point bend specimens distribution based on the four-point data is shown by the dashed lines in Fig. 7(*b*) for a two-parameter, shear-sensitive model. Use of a shear-insensitive model did not substantially change the results, as the flaws experience little mixed mode loading for this specimen configuration. Goodness-of-fit tests indicated that the two-parameter model gave a slightly better fit than a three-parameter model. This fact does not rule out the existence of a threshold strength, but it is not as clearly manifested as before (and therefore not investigated). From a design point of view, a two-parameter model will give a more conservative failure probability than a three-parameter model. A threshold strength must be clearly and consistently demonstrated for the purpose of a safe design. These results firmly establish that a statistical approach

incorporating multiaxial failure criterion must be used as a basis for designing ceramic components. Variations between sample sets must either be resolved by improved control on the machining process, or accounted for by sampling multiple specimen sets. Ultimately, more research is needed in this area.

Conclusions

The strength distribution of silicon carbide was found to be a function of grinding orientation for a typical uniaxial grinding procedure (e.g., ASTM C1161). However, annealing greatly reduced or eliminated the effect of grinding orientation on strength. Annealing did not increase the strength of longitudinally ground flexure specimens. The Weibull size effect was exhibited by the as-ground specimens. Annealed and longitudinally ground specimens typically fail from surface and near-surface agglomerates while transversely ground specimens predominantly fail from machining cracks.

Truncated distribution and three-parameter models appear to approximate best the first experimental data generated in this work (beams ground at various angles). The effect of mixed-mode fracture of machining cracks was adequately predicted by the models for the larger angles of orientation (>30 deg). Strength data for low grinding angles showed some deviation from predicted results, possibly reflecting damage to the agglomerate flaws during the machining process. However, fractography did not detect this damage. For the second data set (plates and beams cut from plates), the uniaxial data were reasonably fit with a two-parameter distribution and a two-parameter model accurately predicted the biaxial strength distribution.

Acknowledgments

The authors would like to thank Ben Willard for help with the experimental work. This research was sponsored in part by the Ceramic Technology Project, DOE Office of Transportation Technologies, under contract DE-AC05-840R21400 with Martin Marietta Energy Systems.

References

1 ASTM C 1161-90, "Standard Test Method for Flexural Strength of Advanced Ceramics at Ambient Temperature," Annual Book of ASTM Standards, Vol. 15.01, 1990, pp. 333–339.

2 "Fractography and Characterization of Fracture Origins in Advanced Structural Ceramics," MIL-HDBK-790, July, 1992.

3 Batdorf, S. B., and Crose, J. G., "A Statistical Theory for the Fracture of Brittle Structures Subjected to Nonuniform Polyaxial Stresses," ASME *Journal of Applied Mechanics,* Vol. 41, 1974, pp. 459–464.

4 Batdorf, S. B., and Heinisch, H. L., Jr., "Weakest Link Theory Reformulated for Arbitrary Fracture Criterion," *Journal of the American Ceramic Society,* Vol. 61, No. 7–8, 1978, pp. 355–358.

5 Batdorf, S. B., "Fundamentals of the Statistical Theory of Fracture," *Fracture Mechanics of Ceramics,* Vol. 3, R. C. Bradt, D. P. H. Hasselman, and F. F. Lange, eds., Plenum Press, New York, pp. 1–30, (1978).

6 Shetty, D. K., "Mixed-Mode Fracture Criteria for Reliability Analysis and Design With Structural Ceramics," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 109, 1987, pp. 282–289.

7 Nemeth, N. N., Manderscheid, J. M., and Gyekenyesi, J. P., "Ceramic Analysis and Reliability Evaluation of Structures (CARES)," NASA TP-2916, Aug. 1990.

8 Richerson, D. W., *Modern Ceramic Engineering,* Marcel Dekker, New York–Basel, 1982, pp. 260–73.

9 Powers, L. M., Starlinger, A., and Gyekenyesi, J. P., "Ceramic Component Reliability With the Restructured NASA/CARES Computer Program," NASA TM-105856, Sept. 1992.

10 Ritter, J. E., Jr., Jakus, K., Batakis, A., and Bandyopadhyay, "Appraisal of Biaxial Strength Testing," *Journal of Non-crystalline Solids,* Vols. 38 and 39, 1980, pp. 419–424.

The Application of a Thermal Efficiency Maximizing Control Strategy for Ignition Timing and Equivalence Ratio on a Natural Gas-Fueled Hercules G1600

M. L. Franklin

D. B. Kittelson

R. H. Leuer

Department of Mechanical Engineering,
University of Minnesota,
Minneapolis, MN 55455

A two-dimensional optimization process, which simultaneously adjusts the spark timing and equivalence ratio of a lean-burn, natural gas, Hercules G1600 engine, has been demonstrated. First, the three-dimensional surface of thermal efficiency was mapped versus spark timing and equivalence ratio at a single speed and load combination. Then the ability of the control system to find and hold the combination of timing and equivalence ratio that gives the highest thermal efficiency was explored. NO_x, CO, and HC maps were also constructed from our experimental data to determine the tradeoffs between efficiency and emissions. The optimization process adds small synchronous disturbances to the spark timing and air flow while the fuel injected per cycle is held constant for four cycles. The engine speed response to these disturbances is used to determine the corrections for spark timing and equivalence ratio. The control process, in effect, uses the engine itself as the primary sensor. The control system can adapt to changes in fuel composition, operating conditions, engine wear, or other factors that may not be easily measured. Although this strategy was previously demonstrated in a Volkswagen 1.7 liter light duty engine (Franklin et al., 1994b), until now it has not been demonstrated in a heavy-duty engine. This paper covers the application of the approach to a Hercules G1600 engine.

Introduction

Objectives. The objectives of this research were as follows:

1 To demonstrate an engine control strategy that allows simultaneous optimization of thermal efficiency with respect to spark advance and equivalence ratio in a heavy-duty commercial natural gas engine.
2 To map the thermal efficiency and emissions surfaces against ignition timing and equivalence ratio for the engine operating at a constant speed and torque.
3 To explore the effect of biasing either the spark advance or the equivalence ratio away from the optimum for emissions control or other reasons.
4 To explore the effect on response and stability of changing correction gains and the interaction between correction gains and maximum correction limits.

Background. It is common knowledge that the torque, efficiency, and emissions of a spark ignition engine are influenced by its equivalence ratio and ignition timing (Obert, 1973; Heywood, 1988). In a previous paper, Leuer (1991) described the effect of air fuel ratio and spark advance on torque at constant fuel flow. In more recent papers, we have described a PC-based control system (Franklin et al., 1994a) and some optimization results on a light duty gasoline engine modified to operate on a gaseous fuel (Franklin et al., 1994b). This paper extends the work to a heavy-duty commercial natural gas engine.

The Optimization Strategy. When the amount of fuel injected per engine cycle is held constant and torque versus spark advance and equivalence ratio plots are made, it is easy to see the effect of small, systematic disturbances in the independent variables. For example, if the spark timing is not at its optimum (Maximum Brake Torque (MBT)) value and a synchronous timing disturbance is added, it results in a fluctuating torque that causes synchronous speed variations. The magnitude and phase of these variations may be related to the slope of the surface with respect to spark advance and the dynamic response of the system. This information may be used to move in a systematic manner to the MBT spark advance. If the air flow is systematically disturbed at constant fuel flow per engine cycle, the same approach may be used to find the equivalence ratio which gives the maximum thermal efficiency for that condition (Franklin et al., 1994a, 1994b).

In many stationary applications, the engine is equipped with a speed governor and powers a load with a well-defined torque versus speed characteristic. At any nonoptimum point on the map, if the spark advance (or air flow) is disturbed, a tiny, momentary change in the torque occurs, which results in a change in speed. These speed changes are too small and too fast to be detected by the speed governor. In fact, the changes are typically smaller than those associated with individual cylinder firing pulses, but their synchronous nature makes it possible for the control system to detect them, much like a lock-in amplifier (Horowitz and Hill, 1984). Changes that result in greater torque at constant fuel flow per cycle indicate higher efficiency. The control system described in this paper systematically disturbs both ignition timing and equivalence ratio and monitors the corresponding speed fluctuations. Based on these speed fluctuations, the strategy moves the parameters in the direction of increasing engine efficiency until the Equivalence Ratio and Spark Advance for Best Efficiency (ERSABE) is reached.

Details of the Dither Cycle. Adaptive control of timing alone has been demonstrated on an assortment of engines, both compression and spark ignition, running on a variety of fuels

Contributed by the Internal Combustion Engine Division and presented at the 16th Annual Fall Technical Conference of the ASME Internal Combustion Engine Division, Lafayette, Indiana, October 2–6, 1994. Manuscript received at ASME Headquarters June 1994. Associate Technical Editor: W. K. Cheng.

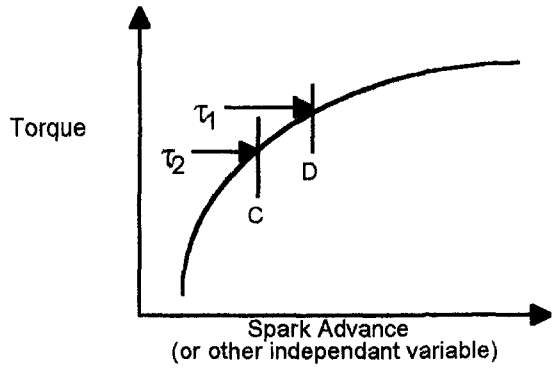


Fig. 1 Idealized torque (τ) versus spark advance, showing the spark timing dithered between levels C and D and the resulting torque change

(Pipho and Kittelson, 1983; Kittelson and Pipho, 1983; Kittelson et al., 1989; Pipho et al., 1991). In a compression ignition engine, the fuel injection timing is controlled, while in a spark ignition engine, the spark timing is controlled. The strategy is based on introducing small disturbances into the timing (dithering the timing) and detecting the corresponding changes in engine speed. The basic control logic is designed to find the timing that corresponds to maximum brake torque. However, a technique called biasing can be used to maintain the timing retarded (or advanced) of that for MBT if necessary for control of emissions, knock, or rate of pressure rise.

The equivalence ratio (ϕ) is defined as the ratio of the chemically correct to the actual air–fuel ratio. It can also be controlled in a manner similar to the timing. To achieve the equivalence ratio that gives maximum fuel economy, the air flow must be dithered with a constant fuel flow. On the other hand, the maximum power equivalence ratio can be determined by dithering the fuel flow at constant air flow. The dithering is superimposed on top of changes in equivalence ratio associated with normal engine operation. It does not interfere with normal engine operation. A fuel and air management system that allows independent control of the fuel and air flow rates without significant interaction between the two (aside from the normal proportionality between fuel flow and air flow) is needed in order to implement this type of equivalence ratio control.

The optimum values of timing and equivalence ratio are interdependent. Therefore, it is desirable to optimize timing and equivalence ratio simultaneously. To do this, a control strategy that varies both parameters during each dither cycle and determines appropriate corrections for each must be used. In this manner, the spark timing and equivalence ratio will advance to their optimum values. Ideally, the system will follow the steepest path to the summit of the three-dimensional efficiency surface in spark timing, equivalence ratio, and efficiency space.

Although this work deals with steady-state engine operation, modifications to the strategy could allow for transients. In an application that requires rapid response to transients, the strategy could be used effectively in conjunction with open-loop maps to set approximate values of timing and equivalence ratio. The maps would provide rapid transient response while the optimizing strategy would fine tune values in the maps. The adaptive part of the system would use the engine itself as a sensor to correct for changes in fuel composition, engine wear, fuel metering, or other variables that may not be easily measured. With such a system, ignition timing and equivalence ratio could be kept near their optimum values even during transients.

Dither Cycles. The two-dimensional dither cycle of spark advance and equivalence ratio requires eight revolutions. That dither cycle is fully detailed in a previous paper (Franklin et al., 1994b). To give sufficient background for the understanding of this paper, a one-dimensional example is described here. The relation between torque and spark timing is shown in Fig. 1. The structure of the disturbance in the case where only the spark advance is optimized is shown in Fig. 2. Here the spark is retarded from optimum so that the "high" portion of the dither cycle causes the torque to increase, increasing the speed. A simple calculation based on revolution times (T_1) for the four revolutions of the simple dither cycle points in the direction of increased efficiency. The result of this calculation is referred to as either the ERROR or the slope estimate:

$$\text{ERROR} = T_1 - T_2 - T_3 + T_4 \qquad (1)$$

The ERROR is multiplied by the correction gain to obtain the correction. The correction is added to the current spark advance and the cycle repeats. When the efficiency reaches a maximum with respect to spark advance, the ERROR approaches zero, causing the correction also to approach zero. The final stable operating point can be shifted from the optimum for reasons of emissions, knock, or stability, by subtracting a bias term:

$$\text{ERROR} = T_1 - T_2 - T_3 + T_4 - \text{BIAS} \qquad (2)$$

A bias term causes the system to seek a specific slope of the response curve rather than its maximum.

Experimental Apparatus and Procedure

While the control strategy does not require an engine map, the maps allow us to evaluate its effectiveness. Data sets necessary for mapping include torque, engine speed, manifold pressure, air flow, fuel flow, and emissions information. These data sets were then used to construct the three-dimensional surface plots.

Apparatus. Apparatus for the experiments includes engine, controller, emissions rack, and dynamometer.

Engine. The specifications of the engine are listed in Table 1.

Engine Controller. Engine control was provided by the same PC-based system described in a previous paper (Franklin et al., 1994a).

Speed Measurement. A relative encoder installed on the engine crankshaft delivers a pulse for every one-sixteenth degree of crank angle rotation and a once-per-revolution pulse, which corresponds to top-dead-center of cylinder number one. A data acquisition card installed in the engine control computer continuously measures the period between these pulses to within 0.1 μs. The engine speed and the ERROR estimations are then

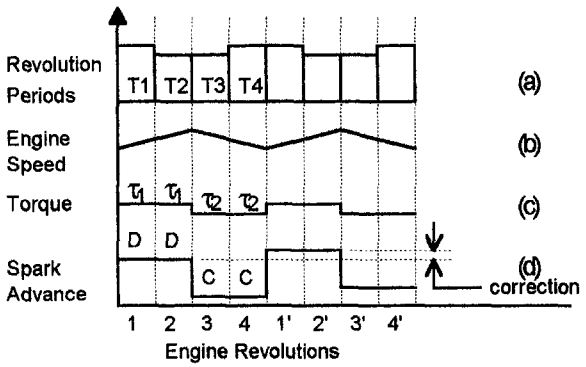


Fig. 2 Response curves (speed and revolution period) to show effect on speed when spark advance is dithered for spark advance-only optimization: (*a*) period (Ti) for each revolution versus time (revolutions); (*b*) engine speed response versus time (revolutions); (*c*) torque versus time (revolutions); (*d*) spark advance versus time (revolutions) with correction to spark advance.

Table 1 Engine specifications

manufacturer	Hercules
cylinders	4
aspiration	normal
displacement	2.7 liters
configuration	inline, overhead valve
cooling	liquid
bore	101.6 mm
stroke	82.55 mm
compression ratio	10

calculated from this period. This particular detail of the system is fully described in a previous paper (Franklin et al., 1994a).

Manifold Pressure Dither Valve. A schematic of the valve used to produce the disturbance in the manifold absolute pressure (and thus the air flow) and its placement on the intake manifold are shown in Fig. 3. The valve is an OEM part originally intended to control idle air bypass flow in a fuel-injected 1986 Ford 5.0 liter V-8. It has fast time response and a flow rate of 110 lpm at a pressure drop of about 2.5 kPa, which make it an acceptable dither valve for this study. For the operating condition used in this study, the valve gave a dither amplitude of one-half to one percent of the manifold pressure, which we feel is a reasonable value. Under lower manifold pressures, the amplitude of the equivalence ratio disturbance is as high as 2 percent, which is larger than desired. An adjustable valve in series with the dither valve could be varied as a function of manifold pressure to ensure the one-half percent dither amplitude.

Throttle Control Servo. The throttle position was controlled with a Futaba model S3302 servo motor.

Ignition System. The ignition system was an Altronic, Inc. PM1 configured to run in the waste spark mode.

Fuel Injectors. The fuel injectors were supplied by BOSCH and are similar to production part number 0-280-150-837.

Emissions Instruments. An emissions rack includes the following instruments:

1 Beckman 400A, hydrocarbon analyzer, using a flame ionization detector.
2 Horiba PIR-2000, CO_2 analyzer.
3 Horiba PIR-2000, CO analyzer.
4 Teledyne Analytical Instruments, percent oxygen analyzer.
5 Thermo Environmental Instruments, Inc., model 10, chemiluminescent $NO-NO_2-NO_x$ analyzer.

Load Control. A Go Power water brake dynamometer controlled by a personal computer provided the load.

Governor. The engine speed for these experiments was governed by a fuel-based governor. This governor works by sensing the difference between the speed set point and the actual speed and changing the pulse width of the fuel injectors. It is an integral-only controller. It was adjusted to provide a slow response to avoid any interaction with the optimization strategy.

Procedure

The experimental work was broken down into two major parts: mapping and optimization.

Mapping. The mapping was done under conditions typical of stationary engine application. To simulate the load, the water control valve of the water brake dynamometer was held approximately constant by the PC-based dynamometer controller. This produced a well-defined load versus speed characteristic curve. The fuel-based engine speed governor simulated typical speed control for a large stationary lean-burn engine. The governor was tuned to be very slow to keep the engine stable during data acquisition. For most of the experiments reported upon here, the engine load and speed were 0.459 MPa brake mean effective pressure (bmep) and 2000 rpm, respectively. This was chosen to represent a medium speed and load condition. The map was created by setting the spark advance and manifold absolute pressure and allowing the governor to set the fuel flow. The equivalence ratio was calculated from air flow, fuel flow, and emissions data. The spark advance was varied between 30 and 70 deg BTDC. The equivalence ratio ranged from approximately 0.65 to greater than 1.0. When the operating condition stabilized, fuel flow, air flow, torque, speed, and emissions data were taken. Then the spark advance and manifold pressure were adjusted to new settings, and the governor was allowed to set the new fuel flow. Spark advance and manifold pressure adjustments were randomized to eliminate hidden trends.

Optimization. The optimization runs were performed to see where the control algorithm would settle. To compare changes in optimization parameters fairly, each run was started at approximately the same point. Modeling has shown that the controller can reach the same steady-state condition, regardless of where the system is started. As shown in previous work (Franklin et al., 1994b), the small speed changes associated with impending misfire were not fatally disruptive to the control scheme. However, if the actual system were started in a region of severe misfire, the cyclic variation might dominate the ER-ROR signal, rendering the current strategy ineffective. The initial conditions were an equivalence ratio of approximately 1.0 and a spark advance of 30 deg BTDC. The system was then allowed to optimize. The spark advance correction bias and the equivalence ratio correction bias were varied to determine the effects each had on the performance of the control system and on the final operating condition. In addition, the gains for both the spark timing and air flow corrections were varied. Finally, interactions between spark timing correction gains, equivalence ratio correction gains, their respective correction magnitudes, and also their respective maximum correction limits were studied.

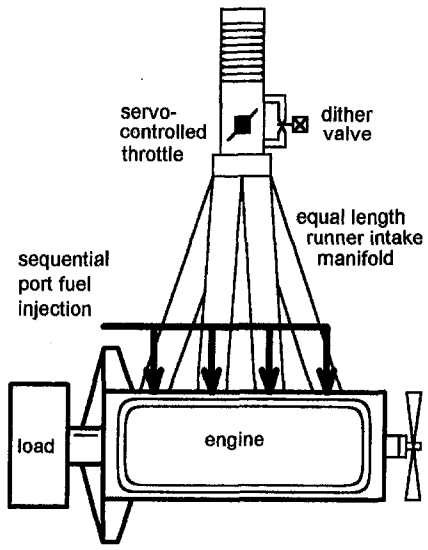


Fig. 3 Schematic of engine showing intake manifold, throttle, and dither valve

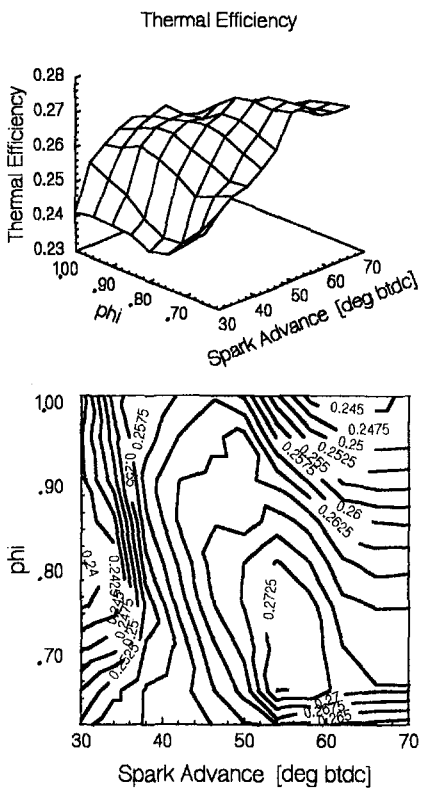


Fig. 4 Thermal efficiency surface versus equivalence ratio and spark advance: (*a*) surface plot; (*b*) contour plot

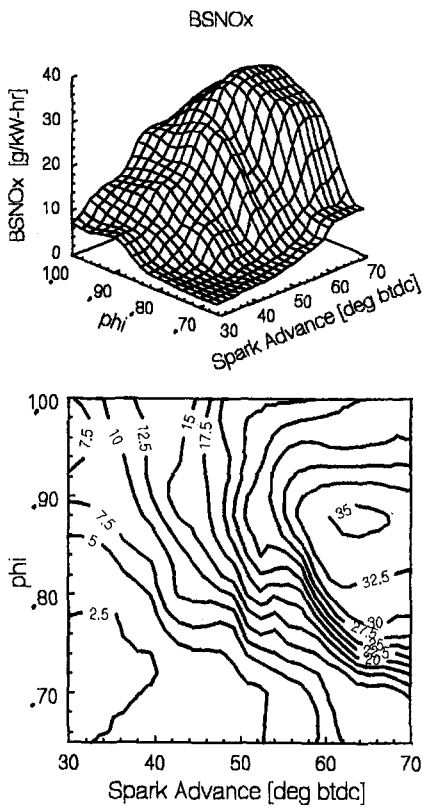


Fig. 5 Brake specific NO_x surface versus equivalence ratio and spark advance: (*a*) surface plot; (*b*) contour plot

Results and Discussion

Mapping. In Figs. 4(*a*) and 4(*b*) the thermal efficiency is plotted against spark advance and equivalence ratio. The contour map, shown in Fig. 4(*b*), indicates a maximum (ERSABE) centered about a spark advance of 56 deg BTDC and an equivalence ratio of 0.72. Operating here gives a thermal efficiency of about 0.27. The thermal efficiency is within one percent (or 0.27 percentage points) of the maximum over a range of spark advances from 50 to 60 deg BTDC and a range of equivalence ratios from 0.63 to 0.84.

Many gas-fired engines are currently run at an equivalence ratio of 1.0 (stoichiometric) to allow the use of a three-way catalyst. Thus, it is interesting to compare performance and emissions at this operating condition with the ERSABE condition, which is the basic set-point of our optimizing strategy. For the speed and load conditions of this experiment, thermal efficiency improves by about 4 percent when adjusted from stoichiometric and MBT timing to ERSABE. This change was smaller than the typically expected 10 to 12 percent improvement (Franklin et al., 1994b). This may be caused by a slower burning combustion chamber, not designed to run very lean. Better results could be expected from a combustion chamber redesigned for lean operation.

As shown in Figs. 5(*a*) and 5(*b*), operating the engine at ERSABE would produce relatively low NO_x emissions (about 10 g/kW-h). This is approximately a 36 percent reduction compared to the corresponding NO_x emissions if the engine were operated at the stoichiometric condition. As expected, the NO_x emissions depend strongly on both equivalence ratio and spark advance.

Operating with such a low equivalence ratio increases the hydrocarbon emissions by approximately 75 percent over the stoichiometric condition as shown in Figs. 6(*a*) and 6(*b*). Spark advance also influenced the hydrocarbon emissions.

Figures 7(*a*) and 7(*b*) show the relative standard deviation of revolution times plotted against spark advance and equiva-

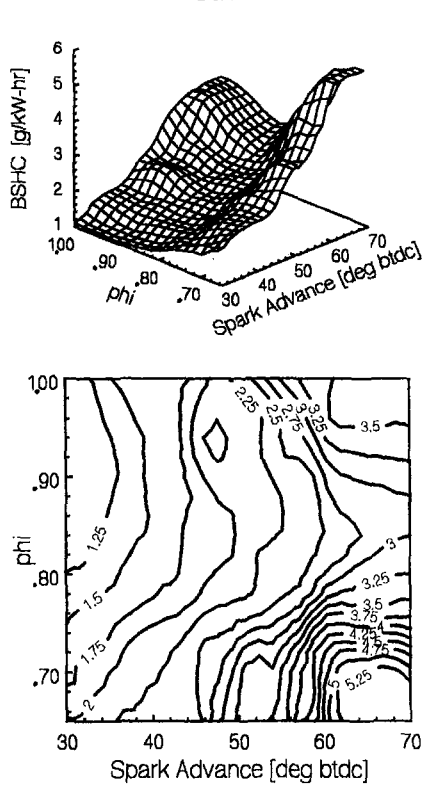


Fig. 6 Brake specific hydrocarbon surface versus equivalence ratio and spark advance: (*a*) surface plot; (*b*) contour plot

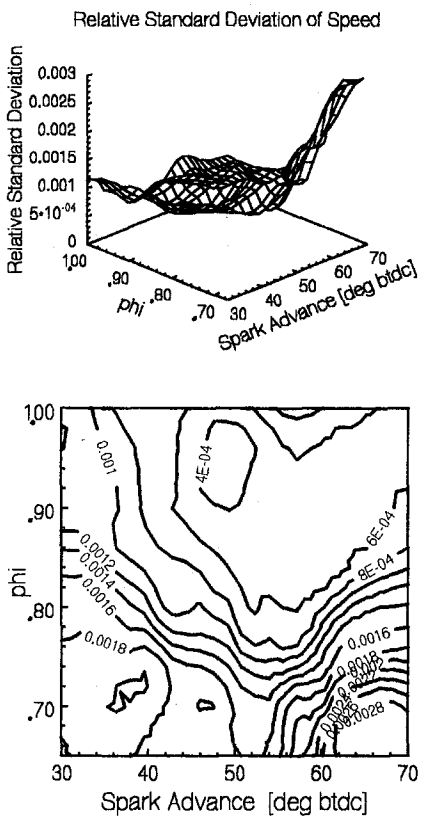


Fig. 7 Relative standard deviation of speed surface versus equivalence ratio and spark advance: (*a*) surface plot; (*b*) contour plot

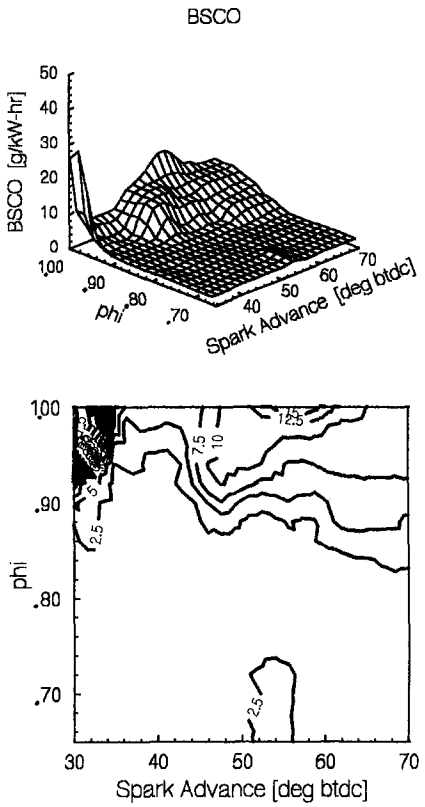


Fig. 8 Brake specific carbon monoxide surface versus equivalence ratio and spark advance: (*a*) surface plot; (*b*) contour plot.

lence ratio. These were calculated by taking four pairs of adjacent revolution times, calculating their standard deviation, and calculating a running average. As expected, the data show that the engine speed variation increases with decreasing equivalence ratio. There is also an increase in speed variation with any departure from MBT timing for a given equivalence ratio.

In general, CO emissions from a lean-burn engine are very low. Figures 8(*a*) and 8(*b*) show the CO emissions versus spark advance and equivalence ratio. The CO emissions are less than 3 g/kW-h for equivalence ratios less than 0.86.

These maps show that if only nonmethane hydrocarbons are regulated, and small fluctuations in speed are not important, the equivalence ratio and spark advance for best efficiency may be the most desirable operating point. If total hydrocarbons are regulated, a slightly richer operating condition would be a better compromise. On the other hand, NO_x emissions could be reduced even further by operating the engine slightly lean of the ERSABE condition. The tradeoffs among efficiency, stability, and regulated emissions must be considered in adjusting the design point with respect to the ERSABE point. This adjustment is implemented by introducing appropriate bias values.

At a typical equivalence ratio of about 0.80 the uncertainty in the equivalence ratio is about ±0.03. The uncertainty in the spark timing is less than one half of one degree. The uncertainties in the calibrations of the torque and flow meters and emissions instruments combine for a total uncertainty in thermal efficiency of about 2 percent. Sparse data and the gridding process introduced an uncertainty in knowing the ERSABE point. The equivalence ratio at ERSABE is known to within ±0.03 and the spark timing is known to within ±3 deg. The conclusions of this work depend only on the relative uncertainties in knowing the equivalence ratio and spark advance at the optimum (ERSABE).

Optimization. The optimization study showed that the system can move to the ERSABE point in a relatively short time. Figure 9 shows a typical optimization trajectory. The spark advance correction gain was 0.0025, with a maximum correction amplitude of 0.25 deg per dither cycle. The equivalence ratio correction gain was 0.002, with a maximum throttle correction of approximately 0.4 deg, which is about a 1.2 percent change in airflow at typical conditions. The system reached the ERSABE point in about 2000 dither cycles. At the tested speed of 2000 rpm, this requires 8.0 minutes. Although an eight minute response seems long, note how far from ERSABE the system started and compare that to how slowly gas composition or engine characteristics typically change. Also recall that the only information the controller needs is crank shaft speed change. Once the controller reached the optimum, there was some wander, but the system remained stable and close to the ERSABE. The ignition timing wandered between 52 and 65 deg BTDC. The equivalence ratio wandered between 0.70 and 0.77. Using the current strategy, the response can be quickened,

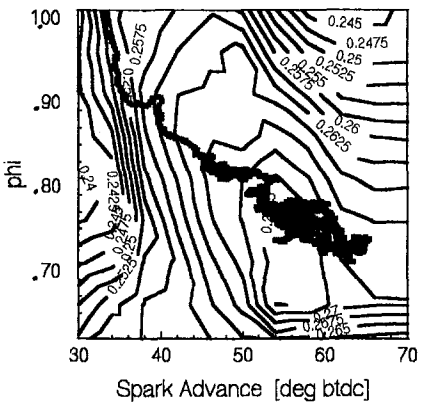


Fig. 9 Typical optimization path taking about 2000 dither cycles (about 16,000 revolutions) to approach ERSABE: 2000 rpm, BMEP = 0.459 MPa

Table 2 A comparison between different combinations of spark advance bias and equivalence ratio bias. Spark advance bias on left axis.

	ϕ bias			
SA bias	-50	0	100	200
-25	A(0.73, 65°)	B(0.75, 60°)		
0	C(0.73, 60°)	D(0.73, 62°)	E(0.78, 50°)	F(0.83, 48°)
50	G(0.76, 50°)			
100		H(0.78, 45°)	I(0.84, 41°)	
200		J(0.82, 38°)		K(0.87, 34°)

but only at the expense of wander. Increasing the correction gains speeds up the response, but then wander increases.

Bias. Bias was demonstrated to effectively move the steady-state operating point away from the ERSABE point. This is shown in Table 2 and in Fig. 10. Each alphabetical label (e.g., "A") in the table corresponds directly with a label on the graph. Increasing the equivalence ratio bias causes the final resting point to be richer. Increasing the spark advance bias causes the final resting point to be retarded from ERSABE.

Another example of the influence of bias is shown in Fig. 11, which shows optimization paths on a $BSNO_x$ surface. These results are for a slightly higher speed and load than the other results presented here, 2200 rpm and 0.511 MPa bmep. The two paths show bias levels of 0, 0 and 50, −50 for spark advance and equivalence ratio, respectively. Here the stoichiometric, MBT condition gives a $BSNO_x$ of 20 [g/kW-h] while the bias condition gives a $BSNO_x$ level of 3 [g/kW-h], an 85 percent reduction while the fuel efficiency is actually 1.5 percent higher. Thus, biasing may be used for effective control of emissions with little or no fuel consumption penalty.

Correction Gains. Changing the correction gains affected the response and stability of the control system. Figure 12 shows the effect of spark advance correction gain on speed of response and stability for three spark advance correction gains: 0.0025 (base case), 0.00025, and 0.000125. As expected, the system responded more slowly with smaller gains. Each curve shown is an average of four data runs. To show the effect of spark advance correction gain on wander, Figs. 13 and 14 show individual curves for two spark advance correction gains: 0.0025 and 0.000125. The wander is much more pronounced in the high gain case.

In this system the equivalence ratio depends on both spark advance and manifold pressure. As the efficiency improves due to a more advanced spark, the fuel is decreased by the governor to keep the speed constant, because the engine is equipped with a fuel-controlled governor. This naturally decreases the equivalence ratio. So, the equivalence ratio shows the effect of both a more efficient spark timing and air-flow. Figure 15 shows equivalence ratio plotted against time for three different equivalence ratio correction gains: 0.002 (base case), 0.001, and 0.0005. The effect of equivalence ratio correction gains on wander is similar to that for the spark advance in that the smaller gains respond more slowly and have less wander. The larger gains respond more quickly but wander more.

Maximum Correction Limits. To prevent a large errant correction in either the spark advance or the throttle setting from an occasional atypical firing, the maximum correction per dither cycle was limited. Changing the maximum correction limits affected the response and stability of the control system. Experience has shown that correction limits of 0.25 deg of spark advance and 0.4 deg of throttle angle work well. One interesting effect of maximum correction limit is that it makes all gains above a certain level have about the same response because the correction is almost always at the maximum level and the only thing the ERROR equation determines is the sign of the correction.

This effect was observed for the spark timing correction when the spark advance correction gain was at or above 0.0025. At a gain of 0.0025 the correction amplitude was at the maximum limit over 85 percent of the time: either +0.25 or −0.25 deg of spark advance. When the gain was cut by a factor of ten, the portion of the cycles hitting the maximum limits was decreased to about 25 percent. This combination of spark advance correction gain and maximum correction limit, while slow in response, was very stable. This effect was observed to a lesser degree for the equivalence ratio correction. When the equivalence ratio correction gain was at 0.004, the correction amplitude was at the maximum limit over 15 percent of the time: either +0.4 or −0.4 deg of throttle angle correction. When the equivalence ratio correction gain was at 0.002, the correction amplitude was at the maximum limit only 2 percent of the time.

Optimization Strategies. It can be seen from the results described above that the optimizing control strategy, combined with bias, allows considerable flexibility in selecting the mean timing and equivalence ratio at which the engine operates. Thus, appropriate tradeoffs may be made among efficiency, emissions, and knock control. What makes this approach different from other strategies currently in use is that it sets these parameters relative to the Equivalence Ratio and Spark Advance for Best Efficiency (ERSABE), rather than to some type of fixed map. Furthermore, the control system does not require expensive and potentially unreliable sensors in order to define operating condi-

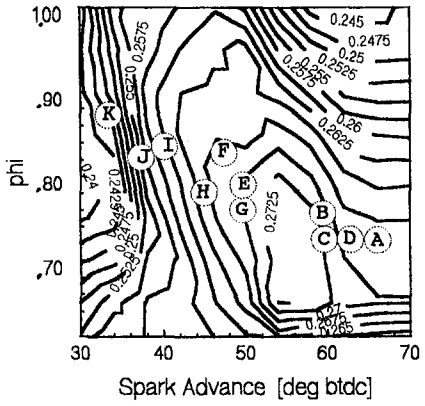


Fig. 10 A comparison between different combinations of spark advance bias and air-fuel ratio bias. A legend is shown in Table 2.

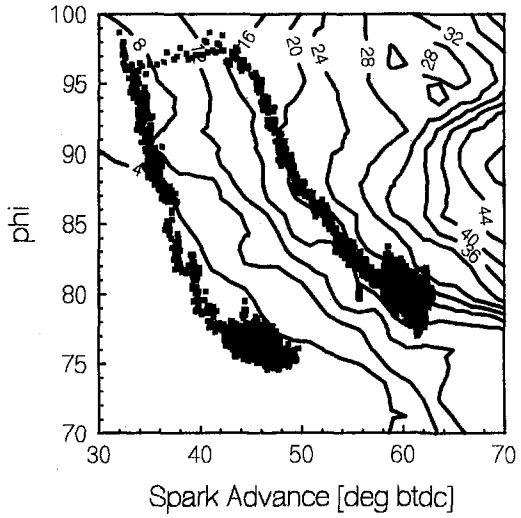


Fig. 11 Optimization paths with and without bias plotted on $BSNO_x$ surface. Speed and load kept at 2200 rpm and 0.511 MPa bmep, respectively. Left curve is for a spark bias of 50 and an equivalence ratio bias of −50.

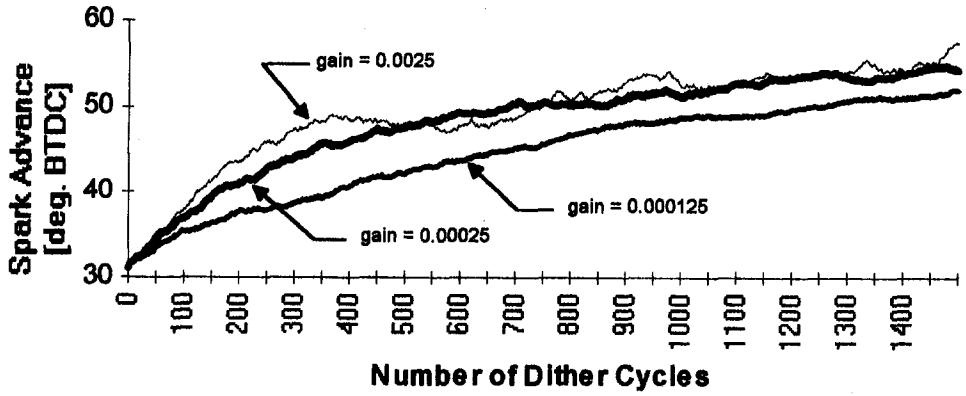


Fig. 12 Spark advance versus number of dither cycles for different spark advance correction gains: 0.0025 (base case), 0.00025, and 0.000125

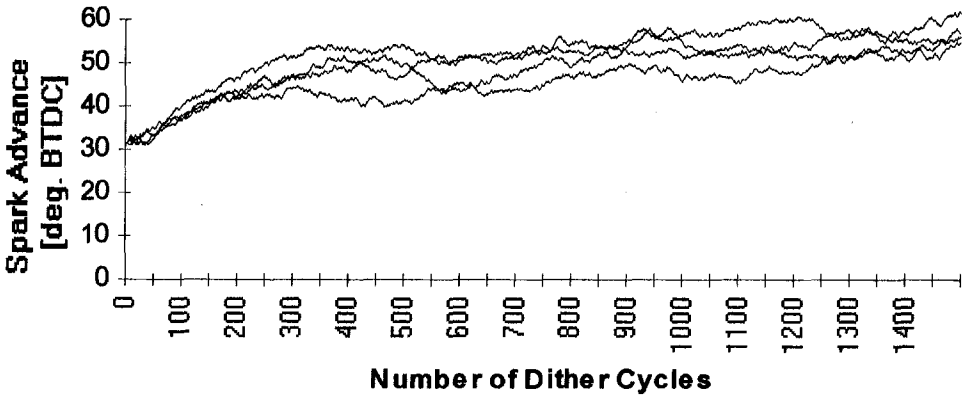


Fig. 13 Spark advance versus number of dither cycles for base case gain. Notice wander and spread of trajectories indicating that spark advance correction gain is probably too high for the noise present.

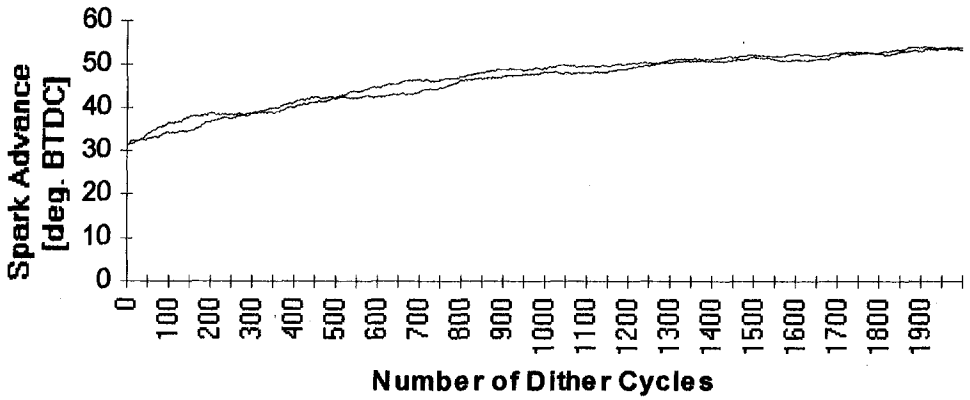


Fig. 14 Spark advance versus number of dither cycles for one-twentieth base case gain. Notice less wander and better repeatability.

tions. The only required sensors are a crankshaft rotation sensor and the engine itself.

Applications. This approach may be particularly attractive in applications where fuel properties are subject to significant variation. Natural gas engines like the one used in this study are one such application. Another application would be engines that run on biogasses like landfill gas or sewage sludge gas. These gasses are even more variable in properties than natural gas. An obvious mobile application would be multifuel engines. This last application would, however, require the use of a combination of the optimizing strategy with an open-loop system to handle transients.

Conclusions

1 The Equivalence Ratio and Spark Advance for Best Efficiency (ERSABE) for the speed-load condition tested was 56 deg of spark advance and an equivalence ratio of 0.72.
2 The thermal efficiency at ERSABE is approximately 4 percent better than at the stoichiometric condition.

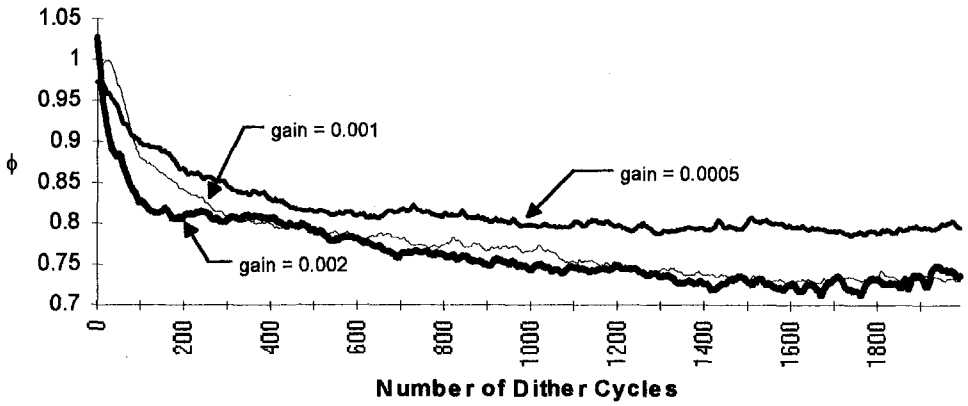


Fig. 15 Equivalence ratio (ϕ) versus number of dither cycles for different equivalence ratio correction gains: 0.002 (base case), 0.001, and 0.0005

3 Compared to stoichiometric, operation at ERSABE gives engine-out $BSNO_x$ emissions about 36 percent lower while total hydrocarbons are increased by about 75 percent.
4 The cyclic combustion variations near ERSABE cause no major problems for the optimization strategy if the correction gains are kept small.
5 The optimization strategy, when run at zero spark advance bias and zero equivalence ratio bias, settles at 58 deg of spark advance and an equivalence ratio of 0.76, which is close to the ERSABE point.
6 Bias effectively moves both equivalence ratio and spark advance away from the ERSABE point. This allows for tradeoffs among efficiency, emissions, and other constraints.
7 At 2200 rpm, 0.511 MPa bmep, NO_x emissions were reduced 85 percent compared to stoichiometric through the use of bias.
8 Spark advance corrections were limited by the maximum correction factor 25 percent of the time at what was considered to be a slowly responding but stable gain setting.
9 Equivalence ratio corrections were limited by the maximum correction factor 15 percent of the time at what was considered to be a slowly responding but stable gain setting.

Future Work

Work remains to be done in the following areas:

1 Developing means of applying this strategy to supercharged engines.
2 Adding an open-loop control system with learning to handle transients. It would use look-up tables that are updated by optimization information.
3 Testing the control system on an engine fueled with natural gas diluted with CO_2 to simulate operation on landfill gas. This will demonstrate its adaptive capability for varying fuel compositions.
4 Testing the control system on an engine fueled with natural gas enhanced with H_2 to explore extension of the lean operating limit.
5 Exploring methods of speeding up response without making the system unstable.

Acknowledgments

The funding for this work comes from the Gas Research Institute, with additional support from Minnegasco and Optimizer Control, Corp. Equipment donations and technical assistance have been provided by Altronic, Inc., Robert Bosch Corp., Hercules Engine, Co., Tescom Corporation, and Thermo King Corp.

This work was also supported in part by a stipend from the Murphy–Robertson Fellowship.

The authors wish to acknowledge the generous technical assistance of Mr. John Kubesh of Southwest Research Institute, and the helpful suggestions and support of Allen D. Wells of Gas Research Institute.

The authors also wish to acknowledge the awe-inspiring laboratory assistance of Mr. Evran Ener, and the brilliant design and machine work of Mr. Dale Lindbeck.

References

Franklin, M. L., Kittelson, D. B., Leuer, R. H., and Pipho, M. J., 1994a, "A PC-Based Fuel and Ignition Control System Used to Map the 3-D Surfaces of Torque and Emissions Versus Equivalence Ratio and Ignition Timing," SAE Paper No. 940546.

Franklin, M. L., Kittelson, D. B., Leuer, R. H., and Pipho, M. J., 1994b, "Synchronous, Simultaneous Optimization of Ignition Timing and Air–Fuel Ratio in a Gas-Fueled Spark Ignition Engine," SAE Paper No. 940547.

Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals,* McGraw-Hill, New York.

Horowitz, P., and Hill, W., 1984, *The Art of Electronics,* Cambridge University Press, Cambridge.

King, S. R., 1992, "The Impact of Natural Gas Composition on Fuel Metering and Engine Operational Characteristics," SAE Paper No. 920593.

Kittelson, D. B., and Pipho, M. J., 1983, "Optimizing Control of Diesel Injection Timing," *ISATA 83 Proceedings,* Vol. 1, pp. 245–271.

Kittelson, D. B., 1986, "System for Optimizing the Timing of Diesel or Spark Ignition Engines," U.S. Patent No. 4,575,800.

Kittelson, D. B., Pipho, M. J., and Franklin, M. L., 1989, "Dynamic Optimization of Spark Advance and Equivalence Ratio for a Natural Gas Engine," SAE Paper No. 892142.

Kittelson, D. B., Pipho, M. J., and Franklin, M. L., 1992, "System for Optimizing the Performance of Internal Combustion Engines," U. S. Patent No. 5,168,853.

Leuer, R. H., 1991, "Mapping of a Lean-Burn Natural Gas Engine," NSF/UROP/REU Project Report.

Meyer, R. C., Cole, J. J., Kienzle, E., and Wells, A. D., 1991, "Development of a CNG Engine," SAE Paper No. 910881.

Obert, E. F., 1973, *Internal Combustion Engines and Air Pollution,* Harper and Row, New York.

Pipho, M. J., and Kittelson, D. B., 1983, "Closed-Loop Digital Electronic Control of Diesel Engine Timing," SAE Paper No. 830579.

Pipho, M., Kittelson, D., and Zarling, D., 1991, "Application of the Optimizer to a Natural Gas Fueled Ford 429," Project Report for Minnegasco.

Vinyard, S., 1991, "Natural Gas Engine Combustion and Emissions," presented at the Continental Congress on Natural Gas Vehicles, Norman, OK, June 5.

Computer-Aided Simulation of Piston and Piston Ring Dynamics

G. Knoll

H. Peeken

R. Lechtape-Grüter

J. Lang

Institut für Maschinenelemente und Maschinengestaltung, Rheinisch-Westfälische Technische Hochschule, Aachen, Federal Republic of Germany

A numerical computer simulation program was developed, aiding in finding optimum design parameters in the multibody-system piston, piston-rings, and cylinder with respect to optimum sealing, minimal friction, and minimum noise stimulation (impact impulse). In the simulation of piston secondary movement and piston ring motion, forces arising from the combustion process, subsonic/supersonic gas flow between the combustion chamber and the crank case, inertial forces and forces resulting from the hydrodynamic lubrication between cylinder liner and piston shaft and piston rings and between piston ring flanks and piston grooves are considered. In addition it is possible to account for effects of global, three-dimensional ring deformation as well as local piston deformation, roughness effects in lubricated contacts, and variable viscosity and variable oil supply. The governing differential equations for the pressure as well as the deformation are solved via finite element techniques, while initial value problems are solved by efficient implicit time integration schemes. The application of the developed computer code is presented in examples.

Introduction

The computer-aided simulation of technical processes is an important tool for decreasing time-consuming and expensive experiments. Calculation of complex system motions can be realized with a small amount of time and costs. Moreover extensive variations of system parameters are possible. The mutual dependencies between the influence variables can be analyzed, which is required for a specific optimization of design parameters as, e.g., piston skirt contour, lateral pin offset, piston ring contour, prestressing of the ring, and clearance between ring and groove. Furthermore, computer-aided simulation can be used during the product development process, when experimental verifications are not yet possible.

Kinetics of Piston and Rings

Piston, piston rings, and cylinder bore represent a coupled system of elastic bodies; its properties are governed by the interaction of the gas and inertial forces as well as by the hydrodynamic lubrication film reactions.

For the design of piston and piston rings of a combustion engine, the following simulation results are of interest:

- friction
  - between piston skirt and cylinder bore
  - between piston rings and cylinder bore
- noise emission
- oil consumption

These mutual dependencies can only be determined by detailed knowledge of the motion behavior of piston and rings.

The analysis of motion of dynamic systems is based on Newton's equation of motion. The local inertial forces can be determined by the given outer forces, the local deformations and its velocities. The velocities and positions of the next time step can be calculated by integration of the local acceleration. In doing so the time-dependent system motion can be calculated iteratively.

Contributed by the Internal Combustion Engine Division and presented at the 16th Annual Fall Technical Conference of the ASME Internal Combustion Engine Division, Lafayette, Indiana, October 2–6, 1994. Manuscript received at ASME Headquarters December 13, 1995. Associate Technical Editor: W. K. Cheng.

The finite element method (FEM) formulation of Newton's equation is

$$\mathbf{M}\cdot\ddot{\mathbf{s}} + \mathbf{D}\cdot\dot{\mathbf{s}} + \mathbf{K}\cdot\mathbf{s} = \sum \mathbf{F}_a \qquad (1)$$

In addition to the gas forces over and between the piston rings, the outer forces of the system consist of the hydrodynamic reaction forces in the gap between piston, piston rings, and cylinder wall. The latter are decisively influenced by the gap function between the sliding surfaces. This makes it necessary to consider the elastic deformations of the surface of piston and cylinder wall as well as the global ring deformation (constant cross section). The calculated piston secondary movements during a duty cycle serve as input data for the computation of the piston ring dynamics, as they determine the ring motion. It can be assumed that the ring motion doesn't influence the piston motion due to small radial forces between piston and rings.

For an exact computation of the motion of the crank mechanism the geometric boundary conditions have to be defined correctly. This is in particular the clearance and the contour of the piston skirt, which arises from a superposition of the cold contour and the radial thermal deformations. The distorted assembled geometry of the cylinder, the local shape of the bearing surface, and the prestress are of further interest.

In order to consider such influences as inertia, hydrodynamic effects, and elasticity in a computer program, various computation methods have to be developed and coupled. The mass and inertia effects are provided by the method of multibody systems (MBS) under consideration of the dynamics of the connecting rod. The lubricant film reactions at the piston rings and the piston skirt are calculated by solving the average Reynolds equation for rough surfaces (Patir and Cheng, 1978) with the finite element method. The governing differential equations for the elastic deformations are also solved via FEM.

Outer Forces

Hydrodynamic Forces. The hydrodynamic pressure function in the lubricated gaps is determined under consideration of surface roughness by the average Reynolds equation developed by Patir and Cheng (1978)

$$\frac{\partial}{\partial x}\left[\phi_x \frac{h^3}{12\eta}\frac{\partial \bar{p}}{\partial x}\right] + \frac{\partial}{\partial z}\left[\phi_z \frac{h^3}{12\eta}\frac{\partial \bar{p}}{\partial z}\right]$$

$$= -\frac{1}{2} W \frac{\partial \bar{h}_T}{\partial z} + V - \frac{1}{2} W\sigma \frac{\partial \phi_s}{\partial z} \quad (2)$$

The variation of the gap due to a crowned and oval piston skirt as well as a thermal and elastic distortion and, if provided, a wear profile have to be considered. This complete Reynolds equation is solved at each time step via FEM to calculate the piston skirt and ring hydrodynamics.

For cavitation effects at the piston skirt and the rings, Reynolds boundary condition is assumed via Murty's Algorithm as implemented by Goenka (1984). For mixed lubrication ($h/\sigma < 4$) a solid-to-solid contact is additionally considered by an algorithm developed by Greenwood and Tripp (1970). Herein the contact pressure is a function of Young's modulus, the nominal gap, and roughness parameters of the surface.

Motivated by the particular ring geometry, the hydrodynamic pressure on the bearing surface, the flanks and the backs of the piston rings can be calculated by a simplified form of Reynolds equation: Neglect of the circumferential fluid flow yields a one-dimensional Reynolds equation, which is solved at a number of circumferential locations:

$$\frac{\partial}{\partial z}\left[\phi_z \frac{h^3}{12\eta}\frac{\partial \bar{p}}{\partial z}\right] = -\frac{1}{2} W \frac{\partial \bar{h}_T}{\partial z} + V - \frac{1}{2} W\sigma \frac{\partial \phi_s}{\partial z} \quad (3)$$

In comparison to the full two-dimensional solution, the computing time can be reduced to about $\frac{1}{50}$.

Different radial fluid film thicknesses in the gap between piston skirt or the rings and cylinder liner can be considered by using the cavitation algorithm mentioned above. Additionally, hydrodynamic friction and Coulomb solid friction on base of the pressure gradients and the shear velocities are taken into account.

Gas Flow Simulation. The measured pressure distribution for the whole duty cycle is required for the computation of the piston secondary movement. An eccentric loadpoint of the gas force can be accounted for.

A detailed analysis of the gas flow in the ring pack is essential for the computation of the piston ring movements, because the pressures over, behind, and under the piston rings change rapidly over the duty cycle and local interring pressures serve as boundary conditions for the lubrication analysis. Interring pressures significantly influence piston ring motion. During phases of compression and expansion, for instance, the axial movement of the first compression ring is almost exclusively determined by the gas forces.

The pressure distributions in two places as a function of the duty cycle must be known in advance: the combustion pressure, which is identical with the pressure over the first compression ring, and the pressure in the crankcase, which in most cases will be identical with ambient pressure.

Normally the pressures between and behind the rings are not known from measurements, so that interring pressures are derived from numerical gas flow simulations.

The common theory for the determination of the gas flow and interring pressures bases on Eweis' transient labyrinth model (1935) and was extended in the present analysis to get more accurate results with respect to interring pressures and global blow by rates.

In Eweis' model discrete volumes above, behind, and under the rings are considered to be connected by orifices whose flow properties are defined by flow coefficients. Under the assumption of axisymmetry, the thermodynamic equations for quasi-static and isentropic mass flow $\dot{m}_{12}$ out of a big reservoir

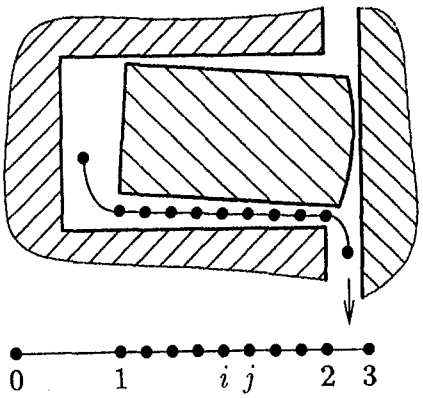


Fig. 1 Discretized gas flow simulation model

$$\dot{m}_{12} = \varphi_{12} \cdot A_{12} \frac{p_1\sqrt{2}}{\sqrt{RT_1}} \cdot \sqrt{\frac{\kappa}{\kappa-1}\left[\left(\frac{p_2}{p_1}\right)^{2/\kappa} - \left(\frac{p_2}{p_1}\right)^{(\kappa+1/\kappa)}\right]} \quad (4)$$

can be solved for each time step, yielding the thermodynamic quantities pressure, temperature and density at each time step in each volume.

Pressure changes between time increments can easily be determined from the equation of state for ideal gases and integration of mass flow quantities. The problem in the model proposed by Eweis lies in the determination of flow coefficients $0 < \varphi_{ij} \leq 1$, which must be chosen in tedious work to ensure realistic interring pressures and global blow-by rates at the same time.

In the present analysis, the orifice model was replaced by a discrete one-dimensional finite difference model in which the equations of mass conservation, momentum, conservation of energy, and equation of state for ideal gases are locally satisfied in each time step. Figure 1 shows the discretized flow simulation model for a ring flank.

The governing equations for flow, momentum, energy, and state can be combined and form a system of two independant nonlinear Eqs. (5) for the unknown velocity u_j and pressure p_j with known velocity and pressure at point i:

$$\left.\begin{aligned} f_1 &= p_j A_j u_j - p_i A_i U_i + \frac{\kappa-1}{2\kappa}\dot{m}(u_i + u_j)(u_j - u_i) = 0 \\ f_2 &= u_j - u_i + \frac{A_i + A_j}{2\dot{m}}(p_j - p_i) \\ &\quad + \frac{\lambda}{2}(u_i + u_j)\frac{x_j - x_i}{d_{H_i} + d_{H_j}} = 0 \end{aligned}\right\} \quad (5)$$

The coefficient of friction λ depends on local flow velocities u and is determined from

$$\lambda = \frac{64}{\mathrm{Re}} \quad 0 \leq \mathrm{Re} \leq 2300 \quad \text{(laminar flow)} \quad (6)$$

and

$$\lambda = \frac{0.316}{\mathrm{Re}^{0.25}} \quad 2300 \leq \mathrm{Re} \leq 10^5 \quad \text{(turbulent flow)} \quad (7)$$

While this dependency leads to a fully nonlinear system of two equations if the coefficient of friction λ is taken to be the spatial average between points i and j, it has been proven to be sufficiently accurate to solve the quadratic problem with λ determined from known quantities at point i. Under the assumption that the flow between the reservoir (point 0) and

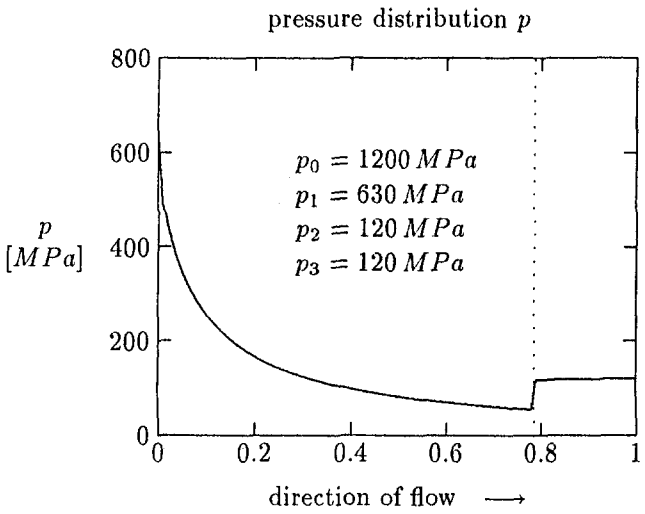


Fig. 2 Pressure distribution at lower ring flank

gap entrance (point 1) is isentropic, the mass flow $\dot{m}$ is iterated until the pressure at the end of the gap (point 2) conforms to the static pressure in the next reservoir (point 3). This model is well suited to account for sub- and supersonic flow. Figure 2 shows the local pressure in a supersonic case in which the location of a compression shock must be found in an additional iteration.

Deformations

In order to obtain precise results for the hydrodynamic pressure between piston and cylinder bore, it is essential to know the deformations of the elastic bodies for each time step.

The deformation $\mathbf{u}$ of an elastic body in local coordinates can be calculated as the difference of its global position $\mathbf{s}$ and the rigid body movement $\mathbf{q}$. Thus

$$\mathbf{u} = \mathbf{s} - \mathbf{q} \tag{8}$$

Equation (1) can then be written as

$$\underbrace{\mathbf{M}\ddot{\mathbf{u}} + \mathbf{D}\dot{\mathbf{u}} + 2\mathbf{M}\dot{\varphi} \times \dot{\mathbf{u}} + \mathbf{K}\mathbf{u}}_{\text{deformation}} = \sum \mathbf{F}_a - \underbrace{\mathbf{M}[\ddot{\mathbf{q}} - \ddot{\varphi} \times \mathbf{r} - \dot{\varphi} \times (\dot{\varphi} \times \mathbf{r})]}_{\text{rigid body movement}} \tag{9}$$

For large structures with a large number of degrees of freedom, solving Eq. (9) becomes difficult: On one hand computation time increases significantly, on the other hand numerical problems do arise due to the huge dimension of the problem. Consequently the number of degrees of freedom has to be reduced, maintaining, however, significant properties of the system.

Modal Reduction. By reducing the number of unknowns of a system of equations the original vector of degrees of freedom $\mathbf{u}$ with length n is transformed into a vector $\mathbf{v}$ with length l. This is realized by multiplying the vector $\mathbf{v}$ with an time-invariant transformation matrix $\mathbf{T}$ of dimension $n \times l$:

$$\mathbf{u} = \mathbf{T}\mathbf{v} \quad \dot{\mathbf{u}} = \mathbf{T}\dot{\mathbf{v}} \quad \ddot{\mathbf{u}} = \mathbf{T}\ddot{\mathbf{v}} \tag{10}$$

In the modal reduction technique, the transformation matrix is composed by the first l structure's eigenvectors. The higher the number of used eigenvectors, the better the deformation behavior is reproduced. Using Eq. (10) with Eq. (9) and premultiplying with $\mathbf{T}^T$ from the left leads to

$$\underbrace{\mathbf{T}^T\mathbf{M}\mathbf{T}}_{\mathbf{M}_{red}}\ddot{\mathbf{v}} + \underbrace{\mathbf{T}^T\mathbf{D}\mathbf{T}}_{\mathbf{D}_{red}}\dot{\mathbf{v}} + \underbrace{\mathbf{T}^T\mathbf{K}\mathbf{T}}_{\mathbf{K}_{red}}\mathbf{v} = \mathbf{T}^T \cdot (\mathbf{RS}) \tag{11}$$

The reduced matrices are now of dimension $l \times l$ so that the equation system has been considerably reduced. The modal degrees of freedom $\mathbf{v}$ obtained by the reduction can, however, not be assigned to the nodes of the FEM structure any more. Therefore the right side ($\mathbf{RS}$) of Eq. (11) has to be reduced at each time step. This disadvantage is compensated by the fact that mass-stiffness- and damping matrices (in the case of Rayleigh damping) have diagonal structure. In consequence, the equations of motion are independent and have not to be solved by a numerical solution procedure.

If outer forces—gas forces, piston pin forces, hydrodynamic piston skirt forces, and friction forces—and the rigid body movement are known, the deformations of the piston structure can be calculated with Eq. (11), which again do have influence on the outer forces and the rigid body movement.

Rigid Body Movement. The rigid body movement of the piston depends on the outer forces of the piston and the kinematics of the crank mechanism. The link between the connecting rod and the crank shaft is regarded as an ideal rigid joint. The equation of motion of the rigid body accelerations $\ddot{\mathbf{q}}$ are formulated on basis of the multibody algorithms described by Haug (1989).

The link between piston, connection rod, and crankshaft with ideal rigid revolute joints can be described by kinematic constraint equations:

$$\Phi^K(\mathbf{q}, t) = [\Phi_1^K(\mathbf{q}, t), \ldots, \Phi_n^K(\mathbf{q}, t)] = 0 \tag{12}$$

Here, n is the number of degrees of freedom that are removed by the revolute joints in the system. The velocities of the rigid bodies' center of mass are obtained by differentiation of the constraint equations

$$\Phi_{\mathbf{q}}\dot{\mathbf{q}} = -\Phi_t \equiv v \tag{13}$$

where Φ_q is the Jacobian matrix. A second differentiation leads to the accelerations of the center of masses of the crank mechanism

$$\Phi_{\mathbf{q}}\ddot{\mathbf{q}} = -(\Phi_{\mathbf{q}}\dot{\mathbf{q}})_{\mathbf{q}}\dot{\mathbf{q}} \equiv \gamma \tag{14}$$

In addition to the kinematic aspects of the crank mechanism, kinetic aspects due to the outer forces have to be considered in order to describe the piston motion fully.

For each body of the dynamic system, a virtual displacement $\delta\mathbf{q}$ can be defined where the virtual work can be written in the form

$$\delta\mathbf{q}^T[\mathbf{M}\ddot{\mathbf{q}} - \mathbf{Q}_a] = \mathbf{0} \tag{15}$$

Here, $\mathbf{M}$ contains the masses and moments of inertia of each rigid body, $\mathbf{Q}_a$ the outer forces and moments. The virtual work of the inner forces and moments between the bodies of the system (in particular for the joints) are respectively eliminated and therefore do not appear in Eq. (15). The virtual displacement $\delta\mathbf{q}$ has to be consistent with the system constraints and has to satisfy the differential of Eq. (12) with time held fixed (Haug, 1989):

$$\Phi_{\mathbf{q}}\delta\mathbf{q} = \mathbf{0} \tag{16}$$

By introduction of Lagrange multipliers, Eqs. (15) and (16) can be combined to yield

$$[\mathbf{M}\ddot{\mathbf{q}} + \Phi_{\mathbf{q}}^T\boldsymbol{\lambda} - \mathbf{Q}_a]^T\,\delta\mathbf{q} = 0 \tag{17}$$

which yields a mixed differential–algebraic equation,

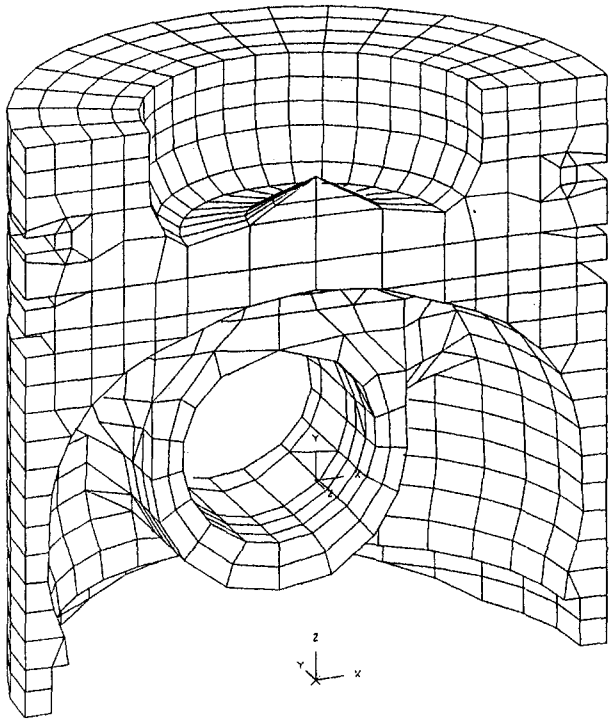


Fig. 3 Finite element structure of the piston

$$\begin{bmatrix} \mathbf{M} & \Phi_q^T \\ \Phi_q & 0 \end{bmatrix} \begin{bmatrix} \ddot{\mathbf{q}} \\ \lambda \end{bmatrix} = \begin{bmatrix} \mathbf{Q}_a \\ \gamma \end{bmatrix} \tag{18}$$

since no derivatives of the Lagrange multipliers λ appear. The accelerations of the rigid bodies are thus determined by a kinematic and kinetic analysis of the system, which, after integration, leads to the velocities and positions of each component at each time step.

The calculation of the piston ring movement by Newton's equation of motion is less complex than the piston movement analysis, because no multibody system occurs. Moreover, the small number of degrees of freedom does not require any reduction.

Results

Piston Secondary Movement. The computation of the piston secondary movement under consideration of the piston deformation is carried out for the crank mechanism of a turbocharged three-cylinder in-line diesel engine with wet-liner-type cylinder, engine type Deutz MWM TD 226 B-3. The basic input data are listed below:

Young's modulus of piston	72,000 N/mm^2
Cylinder bore	105 mm
Radial clearance	40 μm
Center of mass of piston (from bottom)	64.78 mm
Moment of inertia of piston	1922 kg · mm^2
Mass of piston	1.282 kg
C. o. mass of connecting-rod (from big end bearing)	68.7 mm
Moment of inertia of connecting rod	14,000 kg · mm^2
Mass of connecting-rod	1.522 kg
Stroke	120 mm
Viscosity	0.012 Pa · s
Speed	1900 rpm
Coulomb friction coefficient	0.05
Thickness of oil film	30 μm

The piston's contour is found by superposition of the cold contour and thermal deformation of the skirt. Hydrodynamic effects appear at the piston skirt, solid-to-solid contact and friction, and over the whole surface of the piston. Presently, the load is assumed to be evenly distributed over the piston pin bearing; the cylinder liner is modeled as a rigid body. The program is subject to further extensions in the near future.

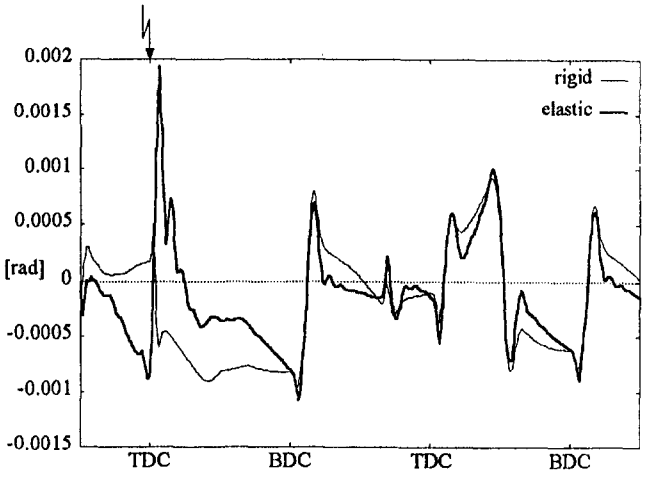


Fig. 5 Tilt angle of the piston

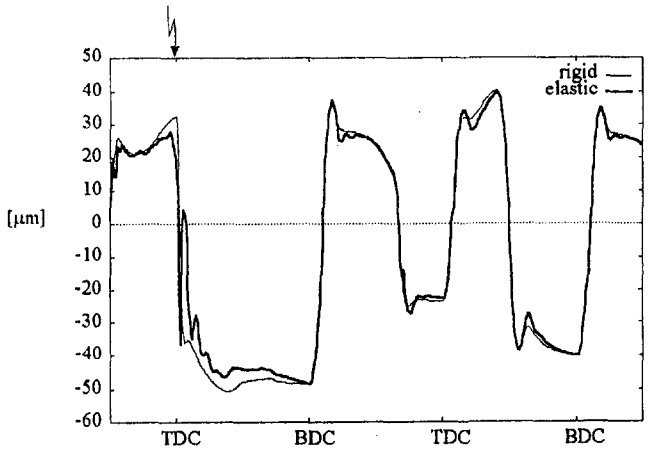


Fig. 4 Radial center of mass motion of the piston

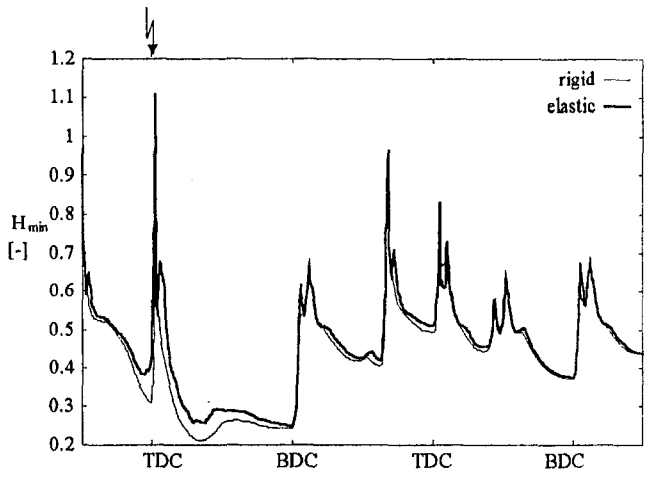


Fig. 6 Minimal relative gaps at the piston skirt

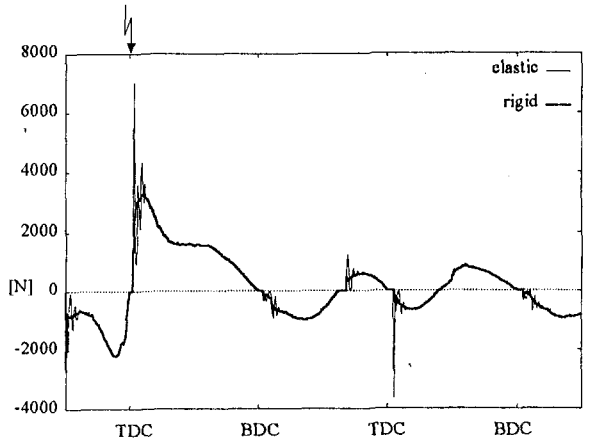


Fig. 7 Lateral skirt force

Figure 3 shows the FEM structure of the piston. Due to symmetry only half of the piston's geometry need be considered. The finite element model consists of 1720 nodes and 1324 elements and is reduced to 17 modal degrees of freedom.

In order to reveal the effect of elasticity, calculations of the piston secondary movement have been carried out for a rigid and an elastic piston. Figure 4 shows the radial motion of the center of mass of the piston during the first duty cycle.

The secondary movements do not differ significantly as the center of mass is located in a region of high stiffness. In the top dead center, elastic vibrations occur due to the impact with the rigid cylinder liner. The tilt angle of the piston is of higher interest (Fig. 5).

Particularly in a range of 90 deg before and 180 deg crank angle after ignition the piston positions differ considerably due to the gas force induced deformation effects of the piston skirt.

To characterize wear and frictional behavior the minimal gaps are of great importance. Figure 6 shows the minimal gap between piston skirt and cylinder liner. As one can expect, the gaps of the rigid piston are smaller than the gaps of the elastic piston.

Figure 7 shows the comparison of the lateral skirt force of the rigid and the elastic piston. One can see that the vibrations of the piston skirt, in particular close to the firing point and in phases of large translational movements, have a significant impact on the skirt hydrodynamics and as a result also on noise emissions.

Piston Ring Motion. Each piston ring is modeled via finite element beam elements, which meet the demands of high accuracy and ease of model generation. In contrast to modal approaches for the piston, a modal reduction of the number of unknowns for the piston ring is not necessary because already ten beam elements per ring will yield sufficiently accurate results. The number of unknowns can be halved, however, if symmetry regarding the plane in which the connecting rod moves can be assumed (Fig. 8).

In a preprocessing step, the cross-sectional properties area A and moments of inertia (bending I_y, I_z, torsion I_t) are determined. For the computation of I_t boundary element techniques are applied for the solution of Poisson's equation over the ring's

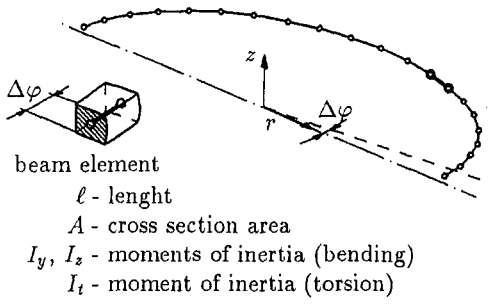


Fig. 8 FE ring model

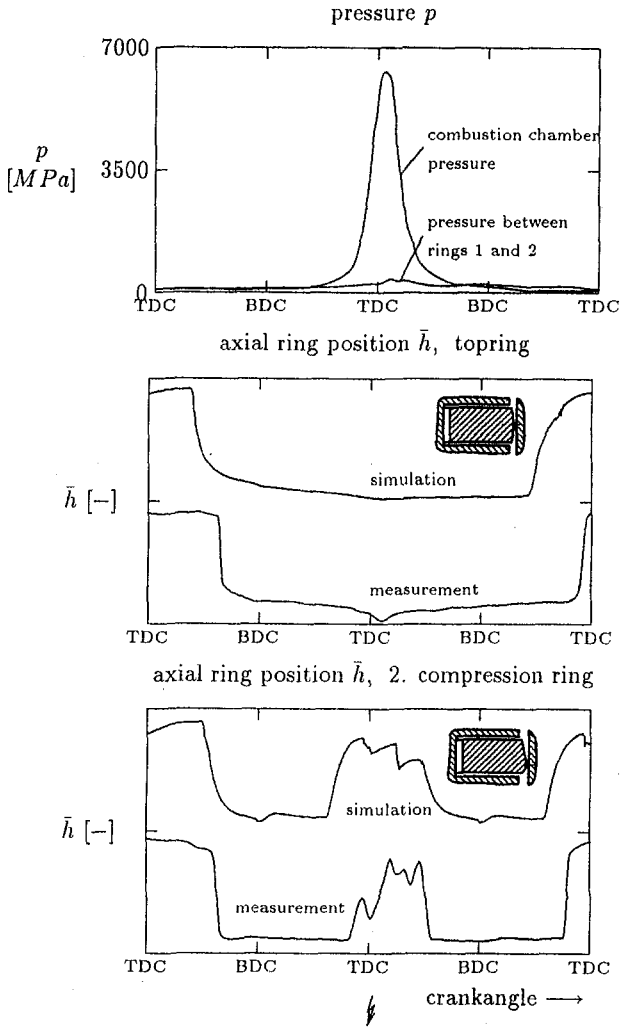


Fig. 9 Axial ring movement

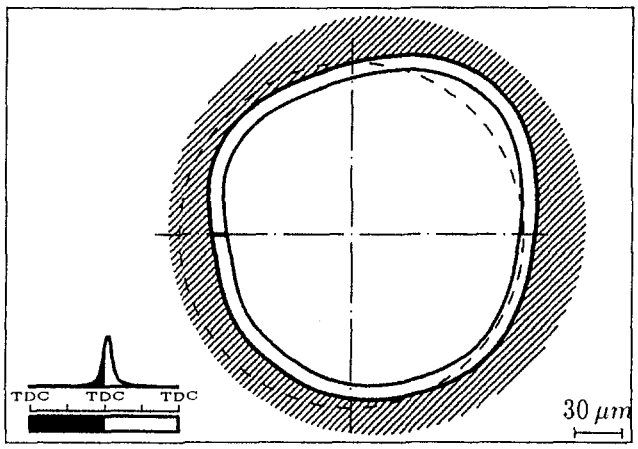


Fig. 10 Ring confirmability with an unaligned cylinder liner (TDC)

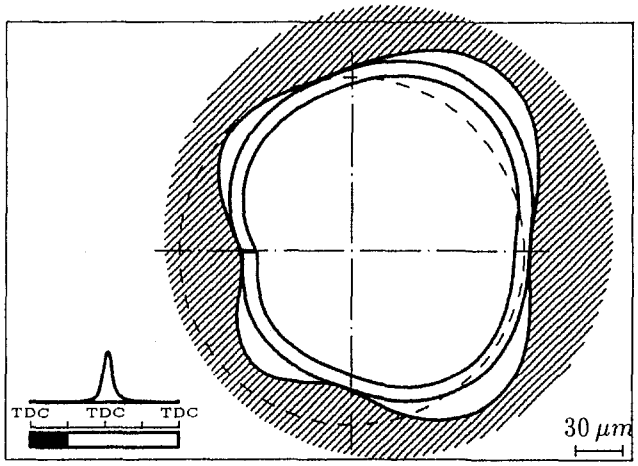


Fig. 11 Ring confirmability with an unaligned cylinder liner (BDC)

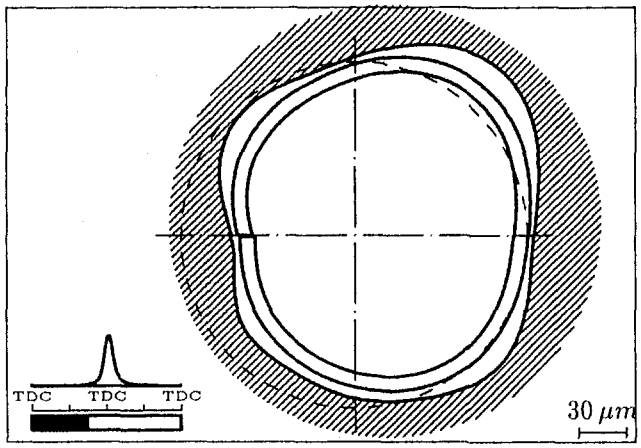


Fig. 12 Ring confirmability with an unaligned cylinder liner (90 deg before ignition (TDC)

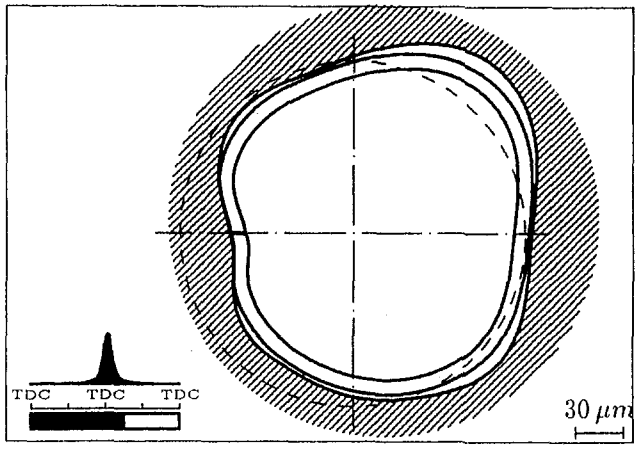


Fig. 13 Ring confirmability with an unaligned cylinder liner (90 deg after ignition TDC)

cross section. This way only the boundary of the cross section, directly derived from CAD data, needs to be discretized.

Figure 9 shows measured and calculated piston ring movements for a 1.6 *l* automotive diesel engine with bore of 76.5 mm under 3600 rpm.

The measurements were taken at the facilities of Kolbenschmidt AG, Neckarsulm. The motion of the oil control ring was not measured. The analysis shows that a lift-off of the first compression ring is prevented by combustion pressures.

Very often the cylinder liner geometry shows a clover-shaped misalignment caused by the cylinder head bolts. This misalignment is very crucial with regard to the sealing properties of the piston rings that must adapt to the distorted geometry. Therefore examinations of the conformability of piston rings are of great interest.

Figure 10 shows simulation results for a 4 cylinder otto engine with bore of 80 mm at a speed of 4000 rpm.

The cylinder distortion, determined with the measuring system Incometer (ARE Goetze GmbH, Burscheid), and expressed in Fourier series expansions, was available for nine planes perpendicular to the cylinder axis.

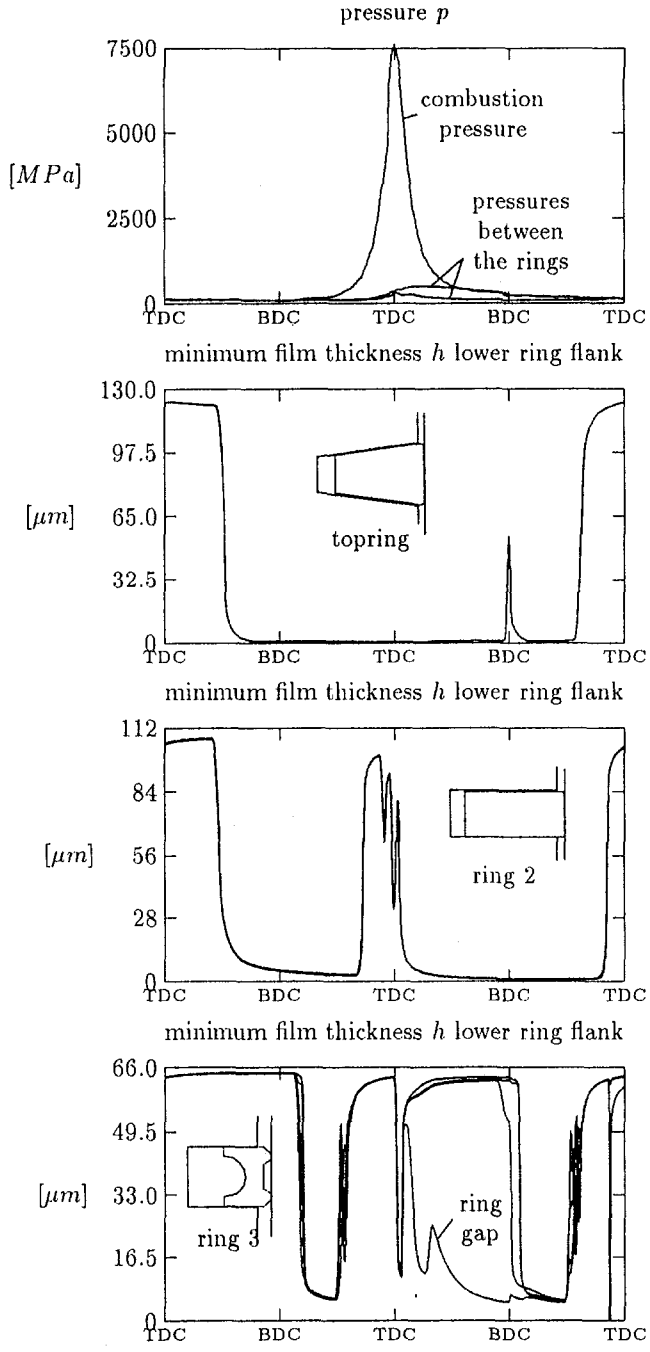


Fig. 14 Axial ring movement relative to a fixed piston

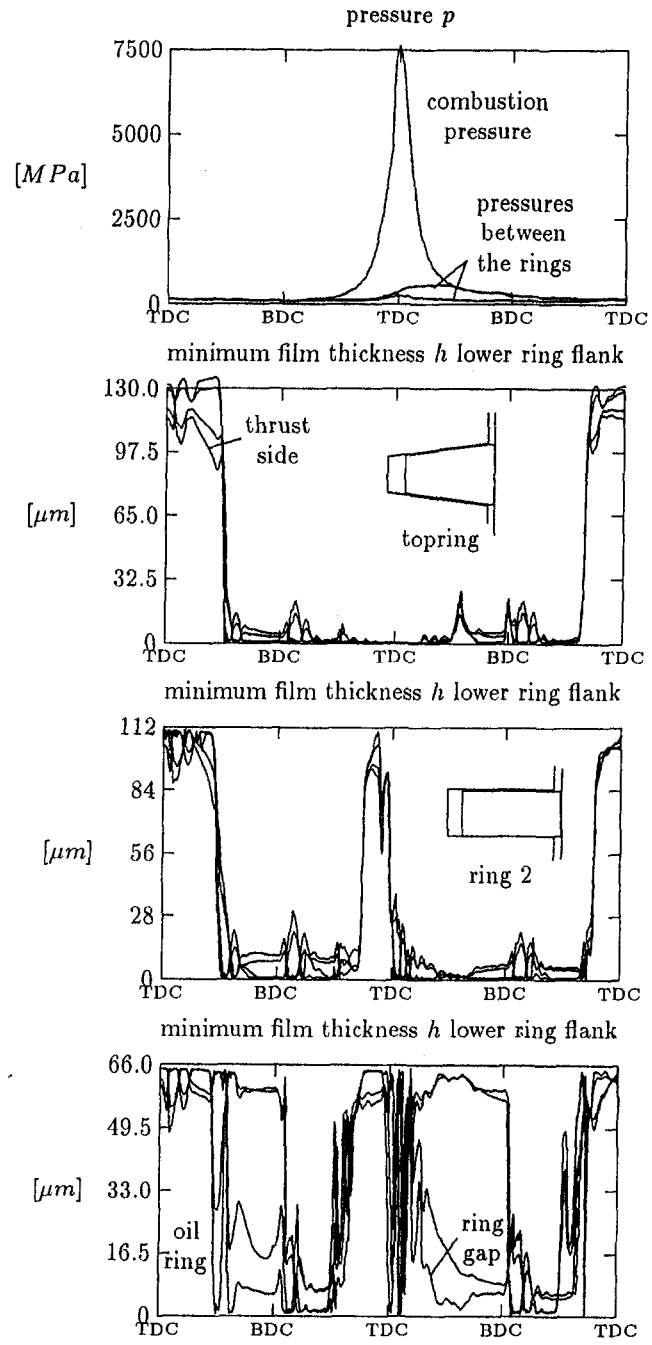


Fig. 15 Axial ring movement with piston secondary movement

Figures 11–13 show, enlarged by a factor of 350, calculated radial positions of the top ring for crank angles of BDC, 90 deg before TDC (ignition) and 90 deg after ignition TDC.

It is apparent that combustion pressures, acting on top of and from behind the ring, smoothly fit the ring to the cylinder liner. Radial gaps increase between dead centers, when sliding velocities allow for a hydrodynamic pressure build up, and vary circumferentially, since gas pressures behind the ring are low. Metal-to-metal contact at four circumferential positions occurs at BDC, where sliding velocities are zero and hydrodynamic squeeze actions do not suffice for a separation of the surfaces.

The calculation of the axial ring movement relative to a fixed piston is shown in Fig. 14; the axial ring movement under consideration of the secondary movement is presented in Fig. 15.

The calculation is based on the following types of piston ring:

- top ring: 15 deg keystone type
- second ring: rectangular compression ring with taper faced running surface
- third ring: coil spring-loaded level edged oil control ring

Each piston ring is modeled by eight beam elements in the circumferential direction.

The gap function in the circumferential direction differs substantially due to the radial piston secondary movement. The "jagged" ring movement in Fig. 15 is caused by the piston motion. In principal the motions of the compression rings are very similar in both calculations. The oil control ring changes the side of contact between the TDC's at the pressure side, while during the same time the piston doesn't. From the calculation with consideration of the piston movement it can be observed that the oil control ring's open end changes the side of contact on intake between the dead centers, while without consideration of the piston movement it does not. The calculated blow by rates (8.5 L/min per cylinder with piston secondary movement, 7.0 L/min with a fixed piston) indicate that the sealing properties are better when the piston secondary movement does not take place.

Summary

A numerical computer simulation program was developed, calculating the piston secondary movement by considering the full three-dimensional elasticity of piston and rings. The piston ring motion is computed for a given secondary movement.

The average Reynolds equation for rough surfaces provides the hydrodynamic reaction forces in the lubricated gaps, where solid-to-solid contact and friction is considered by an elastic contact pressure model.

The local gas flows between rings and grooves are calculated with a discrete capillary model. The interring pressures, which significantly influence the axial ring movement at each time step, as well as such global parameters as the blow-by rate are determined via this model.

The solution Newton's equation of motion for the piston is carried out in a modal reduced system. However, the physical boundary conditions are considered correctly. Finite element method (FEM) and multibody systems are used for solving the governing equations.

Acknowledgments

The authors are indebted to the Bundesministerium für Wirtschaft (BMWi, Bonn) and the Arbeitsgemeinschaft industrieller Forschungsvereinigungen (AiF, Köln, AiF-Nr. 8894, 7269 and 8488) for the financial support that made this research project possible. The investigations have been accompanied by a council under the chairmanship of Dr.-Ing. P. Reipert, Kolbenschmidt AG and Dr.-Ing. R. Jakobs, AE Goetze GmbH. Their support is also acknowledged with thanks. The complete reports are available at the FVV, Frankfurt.

References

Eweis, M., 1935, "Reibungs- und Undichtigkeitsverluste an Kolbenringen," *VDI-Forschungsheft 371*, Berlin.

Goenka, P. K., 1984, "Dynamically Loaded Bearings: Finite Element Method Analysis," ASME *Journal of Tribology*, Vol. 106, pp. 429–439.

Greenwood, J. A., and Tripp, J. H., 1970, "The Contact of Two Nominally Flat Rough Surfaces," *Proc. Instn. Engrs.*, Vol. 185, pp. 696–706.

Haug, E. J., 1989, *Computer Aided Kinematics and Dynamics of Mechanical Systems, Vol. I: Basic Methods*, Allyn and Bacon, Boston.

Patir, N., and Cheng, H. S., 1978, "An Average Flow Model for Determining Effects of Three-Dimensional Roughness on Partial Hydrodynamic Lubrication," ASME *Journal of Lubrication Technology*, Vol. 100, pp. 12–17.

Evaluation of a Three-Dimensional Mathematical Model of a Power Station Boiler

P. J. Coelho
Assistant Professor.

M. G. Carvalho
Professor.

Mechanical Engineering Department,
Instituto Superior Técnico,
Technical University of Lisbon,
Av. Rovisco Pais,
1096 Lisboa Codex,
Portugal

A three-dimensional mathematical model based on the numerical solution of the governing equations of mass, momentum, and energy, and transport equations of scalar quantities is presented. The model is applied to a power station boiler of the Portuguese Electricity Utility for which experimental data are available. Predictions of the gas temperature and chemical species concentrations are compared with the measurements. Three different grids were employed to ensure that conclusions are not affected by the grid refinement. It was found that using a combustion model based on the chemical equilibrium assumption, the predicted gas species concentrations are in good agreement with the data and the temperature profiles display the correct trend, although the temperature was underpredicted, especially in the burner region. The influence of the air/fuel ratio and power load were also successfully simulated. The present results reveal the predictive capabilities of the model and its usefulness for design or optimization studies.

1 Introduction

In power station boilers, a significant amount of fossil fuels is consumed. The limited sources of these fuels and their uneven distribution over the world will tend to increase the price of the fuels with time. This trend is accentuated by increased energy demand. Therefore, there is a need to improve combustion efficiency and to reduce fuel consumption. This motivates the search for new fuel resources and the burning of fuels of lower quality. But these fuels generally yield more pollutants. To face the increasingly more restrictive emission regulations, improved combustion strategies and burner designs are required. In addition, the foregoing requirements should not adversely affect the investment and operation costs and the lifetime of the equipment. Experimental techniques and computational methods are two complementary tools that may help engineers meet all these requirements.

Only a few experimental studies have been performed in full-scale utility boilers. This is due to the difficulties involved, namely the limited physical access to data and the harsh environment. Consequently, the experiments require skill and time and become very expensive. Moreover, they are subject to inaccuracy. Most of them have been performed in boilers with a power output below 80 MWe (see Butler and Webb, 1991). Recently, Cassiano et al. (1994) carried out experimental work in a 250 MWe power station boiler of the Portuguese Electricity Utility. The results are used in the present numerical study for evaluation of the mathematical model.

The mathematical modeling of utility boilers based on the solution of the three-dimensional governing equations for mass, momentum, energy, and other scalar variables has only recently been attempted. Although most of the numerical techniques and physical models were developed about two decades ago, the first numerical study of this kind applied to utility boilers was reported by Robinson (1985). He applied the model to a tangentially fired gaseous furnace and compared the predicted gas temperatures and heat fluxes with experimental data acquired on two large furnaces. Thereafter, several studies have been published in the literature.

Contributed by the International Gas Turbine Institute for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the International Gas Turbine Institute June 23, 1993; revision received April 4, 1996. Associate Technical Editor: H. L. Julien.

Abbas and Lockwood (1986) and Lockwood et al. (1988) studied two power station boilers, one front-wall fired and the other corner-fired. Comparison with measurements was restricted to isothermal velocity data collected in a scale model. A similar study was reported by Sargianos et al. (1990), who simulated the isothermal air flow in a laboratory-scale boiler model and methane combustion in a large utility boiler. The isothermal air flow predictions in the scale model showed qualitative trends aligned with certain preliminary measurements. Carvalho and Coelho (1990) applied a three-dimensional model to an oil-fired water tube boiler where a few heat flux measurements were available. The calculated fluxes were in good agreement with the observed ones. Bonvini et al. (1991) applied a similar model to the prediction of NO_x emissions from a large utility boiler furnace. Comparison with experimental measurements was limited to average concentration at the furnace outlet. All these works report applications to gas or oil-fired boilers, but other studies in coal-fired boilers have also been performed (Boyd and Kent, 1986; Görner and Zinser, 1988; Fiveland and Wessel, 1988; Fiveland and Latham, 1991; Luo et al., 1991; Epple and Schnell, 1992; Coimbra et al., 1994). The evaluation of the mathematical models used in the studies mentioned above is limited by the scarceness of available data.

The present paper describes the application of a mathematical model to a power station boiler of the Portuguese Electricity Utility. Preliminary calculations for standard operating conditions were presented previously (Carvalho et al., 1991) as was a comparative study of the performance of the boiler for fuel oil and natural gas firing (Carvalho et al., 1990). An analysis of the heat transfer phenomena with emphasis on the influence of the uncertainties in the wall temperature and emissivity distributions on the predicted heat transfer was also performed (Coelho and Carvalho, 1992). In these previous works only predicted results were presented, since experimental measurements were not available. Recently, an experimental study was carried out in this boiler. The main objective of the present paper is to report the evaluation of the model by means of comparison of the predicted results with the measurements obtained. These include operation at two different air/fuel ratios and power loads. Due to the limited data available in utility boilers, this study represents an important contribution to the evaluation of current three-dimensional mathematical models.

The mathematical model is presented in the next section. In section 3 the modeled boiler is described, and in section 4 some

numerical details of the computations are provided. The results are presented and discussed in section 5. The main conclusions are drawn in the last section.

2 Mathematical Model

The mathematical model is based on the numerical solution of the conservation equations for mass, momentum, and energy, and transport equations for scalar variables. All these equations may be cast in the following general form:

$$\frac{\partial}{\partial x_j}(\bar{\rho}\tilde{u}_j\tilde{\phi}) = \frac{\partial}{\partial x_j}\left(\Gamma_\phi \frac{\partial \tilde{\phi}}{\partial x_j} - \bar{\rho}\widetilde{u_j''\phi''}\right) + \bar{\rho}\tilde{S}_\phi \qquad (1)$$

where ϕ is the dependent variable; u_j, the Cartesian velocity component along direction x_j; ρ, the density; Γ_ϕ, the diffusion coefficient; and S_ϕ, the source term.

Equation (1) stands for the mass conservation equation when $\phi = 1$; the momentum conservation equation when ϕ is a velocity component; the energy conservation equation when ϕ is the stagnation enthalpy; or the transport equation of a scalar when ϕ is a scalar variable, such as mixture fraction.

To close Eq. (1) the Reynolds stresses $\widetilde{u_i''u_j''}$ and the turbulent scalar fluxes $\widetilde{u_j''\phi''}$ should be expressed in terms of known quantities. The $k-\epsilon$ turbulent viscosity/diffusivity model has generally been employed for that purpose in the mathematical modeling of utility boilers and it was also used in the present work. It assumes a linear relationship between the Reynolds stresses and the rate of strain:

$$-\bar{\rho}\widetilde{u_i''u_j''} = \mu_t\left(\frac{\partial \tilde{u}_i}{\partial x_j} + \frac{\partial \tilde{u}_j}{\partial x_i}\right) - \frac{2}{3}\left(\bar{\rho}k + \mu_t\frac{\partial \tilde{u}_k}{\partial x_k}\right)\delta_{ij} \qquad (2)$$

and a gradient diffusion hypothesis for the calculation of the turbulent scalar fluxes:

$$-\bar{\rho}\widetilde{u_j''\phi''} = \frac{\mu_t}{\sigma_\phi}\frac{\partial \tilde{\phi}}{\partial x_j} \qquad (3)$$

In these expressions it is assumed that the approximations used for constant density flows remain valid without explicitly accounting for density fluctuations (see Jones and Whitelaw, 1982). The solution of transport equations for the turbulent kinetic energy, k, and its dissipation rate, ϵ, allow the calculation of the turbulent viscosity μ_t ($\mu_t = c_\mu\rho k^2/\epsilon$). Standard values were assigned to all the constants of the model (Launder and Spalding, 1974).

The density is calculated from the pressure, temperature, and chemical gas composition using the equation of state for a perfect gas. The temperature is related to the enthalpy by a well-known thermodynamic relationship. The calculation of the chemical gas composition requires a combustion model.

The combustion model assumes that the reaction rates are high enough that the chemical reactions take place instantaneously as soon as the reactants are brought together. Combustion is controlled by mixing rather than kinetic phenomena. Therefore, the instantaneous thermochemical state of the gaseous mixture depends only upon strictly conserved scalar variables. All these variables are linearly related when it is further assumed that the system is adiabatic and the mass diffusion coefficients of all the chemical species and the thermal diffusion coefficient are equal. In such a case, the mixture fraction is often taken as the dependent variable. The instantaneous values of the mass fractions of all chemical species, temperature, and density may be found from the instantaneous value of the mixture fraction.

The relation between the gas composition and the mixture fraction depends on the model employed. In this work two different models were used. In the chemical equilibrium model, which was used in most of the cases, it is assumed that chemical equilibrium prevails and the instantaneous gas composition is calculated from the minimization of the free Gibbs energy, following Gordon and McBride (1971). Complete chemical equilibrium was assumed for values of mixture fraction, f, smaller than a threshold value, f_c, and frozen mixing (adiabatic mixing of mixtures at $f = f_c$ and $f = 1$) was assumed for $f > f_c$. The threshold value f_c was selected to match the peak measured CO concentration. The values of the instantaneous mass fractions of the chemical species as a function of the mixture fraction are stored in a look-up table. In the simple chemically reacting system (SCRS) model it is supposed that combustion takes place as a single-step irreversible reaction, where the fuel and the oxidant react instantaneously, yielding combustion products. This assumption yields piecewise linear relationships between the instantaneous mass fractions of oxygen, fuel, nitrogen, products, and the mixture fraction (see, e.g., Lockwood and Naguib, 1975).

In utility boilers a significant proportion of the heat released in combustion is transferred to the walls and it is not correct to consider the combustion chamber as adiabatic. Hence, the enthalpy is not linearly related to the mixture fraction and an additional assumption on how these variables are related is needed. Here, a piecewise linear relationship as suggested by Abou-Ellail et al. (1978) was used. The temperature can be determined from the enthalpy using well-known thermodynamic relations and the density is obtained from the ideal gas equation of state.

Due to the turbulent nature of the flow, mixture fraction fluctuations should be taken into account to calculate the mean values of the mass fractions, density and temperature. This was accomplished by prescribing a clipped Gaussian shape for the mixture fraction probability density function (Lockwood and Naguib, 1975). Hence, density weighted and nonweighted mean values are calculated as:

$$\tilde{\phi} = \int_0^1 \phi(f)p(f)df \qquad (4)$$

$$\bar{\phi} = \bar{\rho}\int_0^1 \frac{\phi(f)}{\rho(f)}p(f)df \qquad (5)$$

Nomenclature

C_μ = constant of $k-\epsilon$ turbulence model
k = turbulent kinetic energy
p = pressure
$p(f)$ = probability density function
S_ϕ = source term of the discretized ϕ equation
u_i = velocity component along i direction
x_i = Cartesian coordinate in i direction
Γ_ϕ = diffusion coefficient of variable ϕ
δ_{ij} = Kronecker tensor
ϵ = dissipation rate of turbulent kinetic energy
ϕ = dependent variable in a transport equation
μ = dynamic viscosity
ρ = density
σ_ϕ = Prandtl/Schmidt number of variable ϕ

Subscripts

i = spatial direction
t = turbulent
ϕ = dependent variable

Superscripts

$\bar{}$ = Reynolds-averaged value
$\tilde{}$ = Favre-averaged value
$''$ = fluctuation quantity in Favre-averaging

The probability density function is specified by the mean value and variance of the mixture fraction. Transport equations are solved for these two quantities.

The radiative heat transfer plays a dominant role in utility boilers and enters in the energy conservation equation as a source term. Here, it was calculated by means of the discrete transfer model (Lockwood and Shah, 1981), which is based on the solution of the radiative heat transfer equation along specified directions. In this model, the temperature and the absorption coefficient of the medium are assumed to be constant in each control volume resulting from the discretization of the physical domain. For all the control volumes adjacent to the boundary, the central points of the faces of the control volumes coincident to the boundary are determined. Let P be one such point. The hemisphere centered in P is discretized into a given number of solid angles. Each solid angle defines a direction along which radiation beams are fired and the radiative heat transfer equation is solved. Hence, given a point P at the center of a cell face on the boundary, a radiation beam is fired from P for each one of the directions selected above. The path of a radiation beam is followed until it hits another boundary. Let Q_i be the impingement point. Although, in general, Q_i is not the central point of a boundary cell, it is assumed that the radiation intensities at Q_i and at the central point of the cell that contains Q_i are equal. The radiation intensity leaving that cell is either known or guessed from a previous iteration. Then, starting from Q_i, the path of the beam is followed back to the origin (point P), and the radiative transfer equation is integrated analytically along this path, allowing the calculation of the radiation intensity arriving at point P. The incident radiative heat flux at point P is calculated by adding the contributions due to all the radiation beams that reach point P, i.e., from all the directions resulting from the discretization of the hemisphere centered in P. This procedure is repeated for all the cell faces on the boundary. Then, the boundary conditions (prescribed temperature in the present study) are applied. This allows for an improved guess of the radiosity at the boundary. The iterative procedure continues until the convergence criterion is satisfied.

The source term of the energy equation in each control volume is obtained by the summation of the change of radiation intensity within the control volume for all the rays that cross it. Additional details of the method are given by Lockwood and Shah (1981).

Scattering of the participating medium was neglected. This is a good assumption for gas or oil-fired boilers. The absorption coefficient of the medium is calculated from the emissivity of the gas/soot mixture. The emissivity is calculated using the two grey plus one clear gas approach of Hottel and Sarofim (1967) extended to account for soot. The coefficients determined by Truelove (1976) were used. Soot concentration was predicted using the soot formation model of Khan and Greeves (1974) and the soot oxidation model of Magnussen and Hjertager (1977).

The governing equations are integrated over each control volume and discretized using finite differences. The central differences scheme is employed for all but convective terms, which are discretized using the hybrid central differences/upwind scheme. The solution algorithm is based on the SIMPLE method. The sets of discretized equations are solved using the Gauss–Seidel line-by-line iteration procedure.

3 The Modelled Boiler

The mathematical and physical models described in the previous section were applied to a boiler of the Portuguese Electricity Utility, schematically shown in Fig. 1(a). It is a natural circulation drum boiler with a pressurized combustion chamber, two parallel passages by the convection zone and superheating. The walls of the combustion chamber, the vaporization panels, are vertical welded tubes. Water circulates in these tubes and partly vaporizes due to the heat released in the combustion chamber, which is transferred to the walls. The boiler is fired from three levels with four oil burners in each level, as shown in Fig. 1(b). A detail of the burners is presented in Fig. 1(c).

At maximum capacity the boiler produces 771 ton/h of steam at 167 bar and 545°C. The operating conditions are those for which experimental data are available, as summarized in Table 1. Case 1 corresponds to the standard operating conditions at full load (250 MWe). In case 2 the air/fuel ratio was reduced but the power load is the same, whereas in case 3 the boiler is operating at partial load. The fuel oil properties are given in Table 2.

The present model was applied to the combustion chamber whose overall dimensions are $8.57 \times 11.43 \times 19.96$ m^3 along the x, y, and z directions, respectively. The simplified combustion chamber geometry is presented in Fig. 1(b). Due to the symmetry, only one half was modeled. Figure 1(b) also shows the inspection ports, which provide access for the probes.

Experimental work was recently carried out in this boiler (Cassiano et al., 1994). Gas temperatures were measured using a 3.5-m-long, double-shielded, water-cooled suction pyrometer, which consists of a Pt/Pt-13 percent Rh thermocouple of 350 μm diameter protected by radiation shields from losses to the combustion chamber walls. Major gas species concentrations were obtained by a gas sampling probe connected to gas analyzers. The probe consists of a sampling tube of 2 mm diameter mounted in a water-cooled jacket of 50 mm diameter and 3.5 m length. According to Cassiano et al. (1994) it is expected that the measured gas species concentrations are close to the density-weighted averages within 10 percent of the maximum measured value.

4 Numerical Details

The fuel oil is atomized in each burner and enters the combustion chamber as a spray. In the present model it is assumed that the droplets vaporize instantaneously and only the gas-phase flow is considered. The wall temperature and emissivity were taken as 700 K and 0.65, respectively, which is expected to be a good assumption. In fact, a previous sensitivity study (Coelho and Carvalho, 1992), where a coupling algorithm between the physical phenomena occurring in the combustion chamber and in the water wall tubes was proposed, showed that the influence of the wall temperature distribution on the predictions is negligible. Moreover, it was shown that the heat transfer phenomena are influenced by the wall emissivity but the gas temperature and composition are only marginally affected, provided that the wall emissivity is not too different from the actual one.

Calculations were performed using a mesh with $16 \times 34 \times 60$ grid nodes along the x, y, and z directions, respectively (see Fig. 2(a)). Hereafter, this grid will be called grid 1. The grid is very fine in the burner region where neighboring grid lines are undistinguishable. Therefore, a detail of the grid is shown in Fig. 2(b). In this figure the dashed lines are the faces of the control volumes and the solid lines identify the boundaries of the fuel inlet, atomizing air inlet, divider wall, and combustion air inlet. A coarser grid was employed for the calculation of radiative heat transfer comprising $10 \times 9 \times 20$ grid nodes. This allows savings in CPU time without adverse effects on the solution accuracy (see, e.g., Abbas and Lockwood, 1986). Additional savings were achieved by calling the radiation subroutines only every 10 iterations. Hence, the converged solution, achieved when the sum of the normalized absolute residuals of each dependent variable decreased to values below 2×10^{-3}, is obtained in about 1000 iterations and 6 hours CPU time (VAX 9000-440).

Two other meshes were employed to evaluate the influence of grid refinement on numerical accuracy. It would be desirable to duplicate the number of grid nodes in each direction since this would allow the use of Richardson extrapolation to estimate

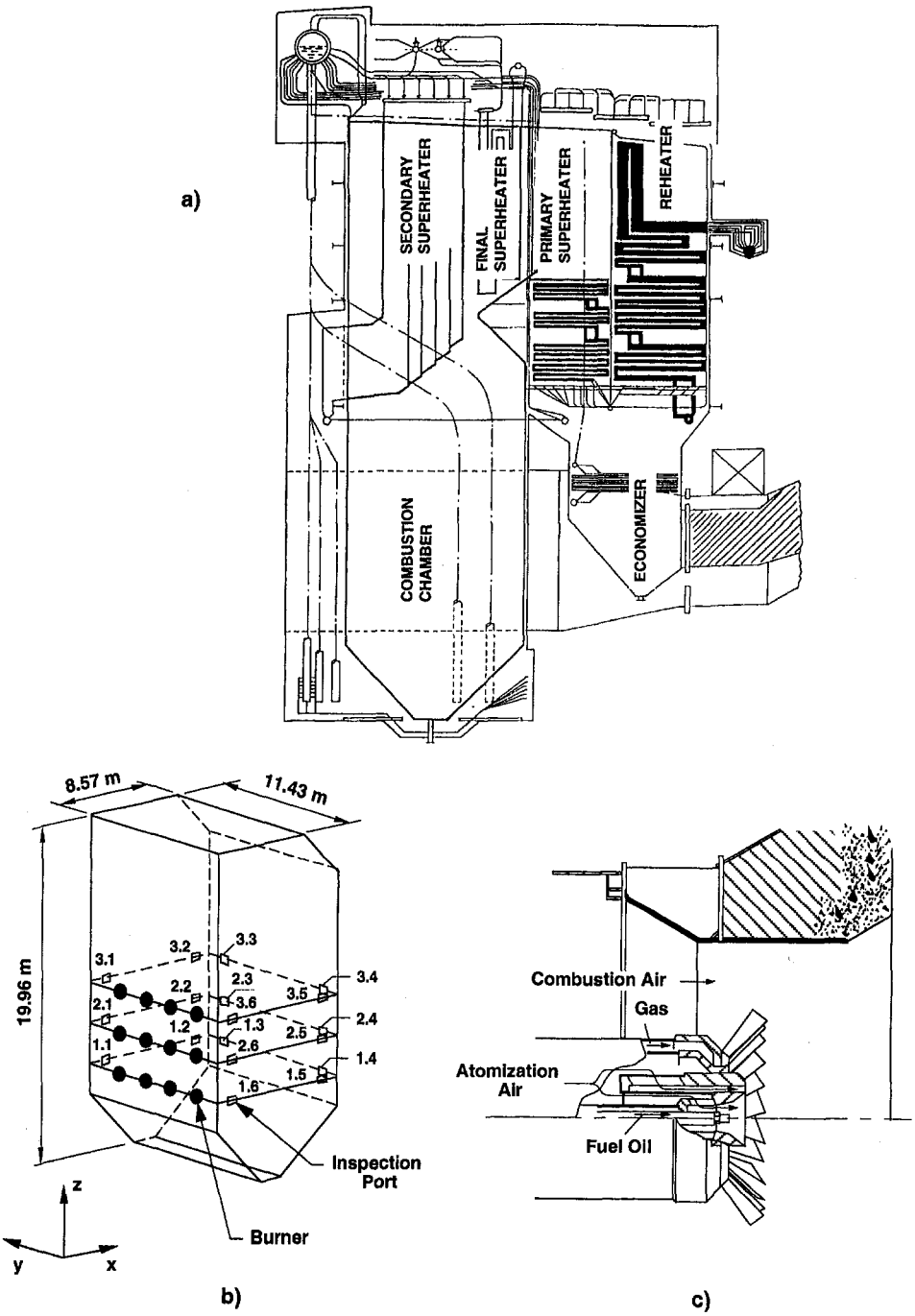


Fig. 1 (*a*) Sketch of the power station boiler; (*b*) simplified geometry of the combustion chamber with identification of the burners and inspection ports; (*c*) detail of a burner

the solution error. However, this would require many grid nodes and several days of CPU time. Therefore, a compromise solution was sought in which the grid was refined only either along the y and z directions or along the x direction. Hence, additional computations were carried out using grid 2 ($16 \times 54 \times 98$ grid nodes) and grid 3 ($32 \times 34 \times 60$ grid nodes).

5 Results and Discussion

5.1 Standard Operating Conditions. Figures 3(a), 3(b), and 3(c) show the predicted velocity component along the x direction, temperature, and mixture fraction profiles along the axis of the top level burner, close to the side wall. The profiles obtained using the three grids mentioned above are plotted. These profiles were selected to illustrate the grid dependency. It can be seen that the three grid schemes yield qualitatively similar predicted profiles. However, significant differences may be found between the numerical solutions at certain locations. For example, the predicted u velocity at $x = 4.3$ m is about 10 m/s higher for grid 3 than for grids 1 and 2. The temperature and mixture fraction profiles are similar for the three grids and exhibit steep gradients near the burner. The calculations performed using grid 3, which has twice as many grid nodes along the x direction as grids 1 and 2, reveal a higher fuel consumption rate for x values up to $x \approx 1.3$ m and a smaller consumption rate further downstream up to $x \approx 4.0$ m.

Therefore, the predicted results cannot be regarded as completely grid independent. This conclusion is consistent with the

Table 1 Operating conditions

Case	1	2	3
Load (MWe)	250	250	125
Air mass flow rate (ton/h)	805	769	459.3
Fuel mass flow rate (ton/h)	52.7	52.7	28.8
O_2 volumetric concentration in flue gases (%)	1.5	0.75	3.0
Inlet air temperature (°C)	335	343	290

Table 2 Fuel-oil properties

Composition (average)	Carbon: 86.52% Hydrogen: 11.06% Nitrogen: 0.72% Sulfur: 2.08%	(ASTM D129)
Heating Value	10076 kcal/kg	(ASTM D129)

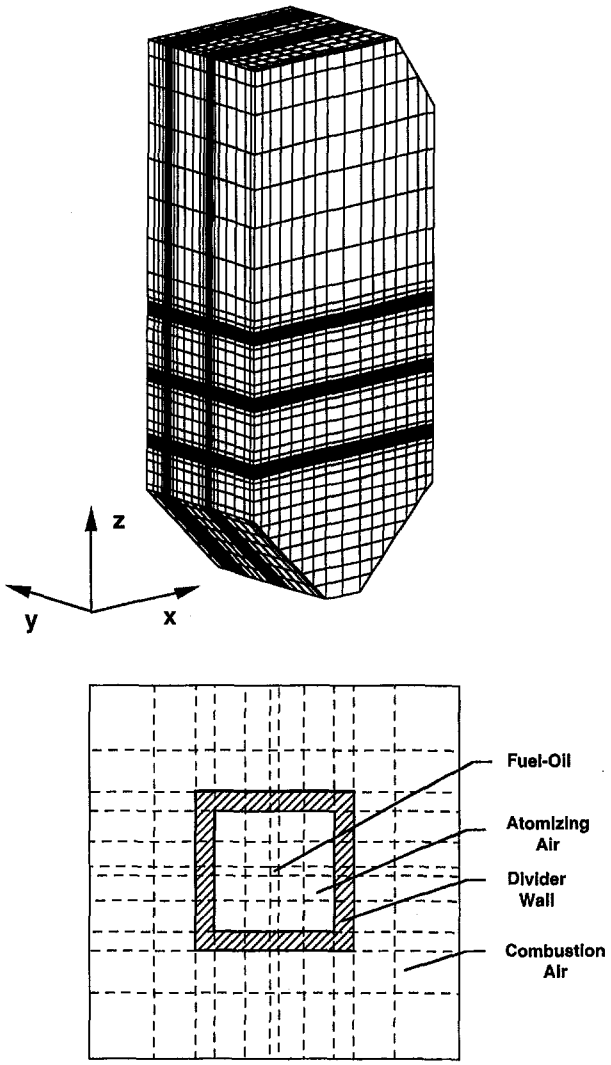


Fig. 2 (*a*) Computational grid (16 × 34 × 60 grid nodes); (*b*) detail of the grid in the burners region

one of Gillis and Smith (1990), who have found that the calculation of an isothermal flow in a pilot scale wall-fired furnace requires more than 250,000 grid nodes. So, the comparison presented below between the predictions and the data for the

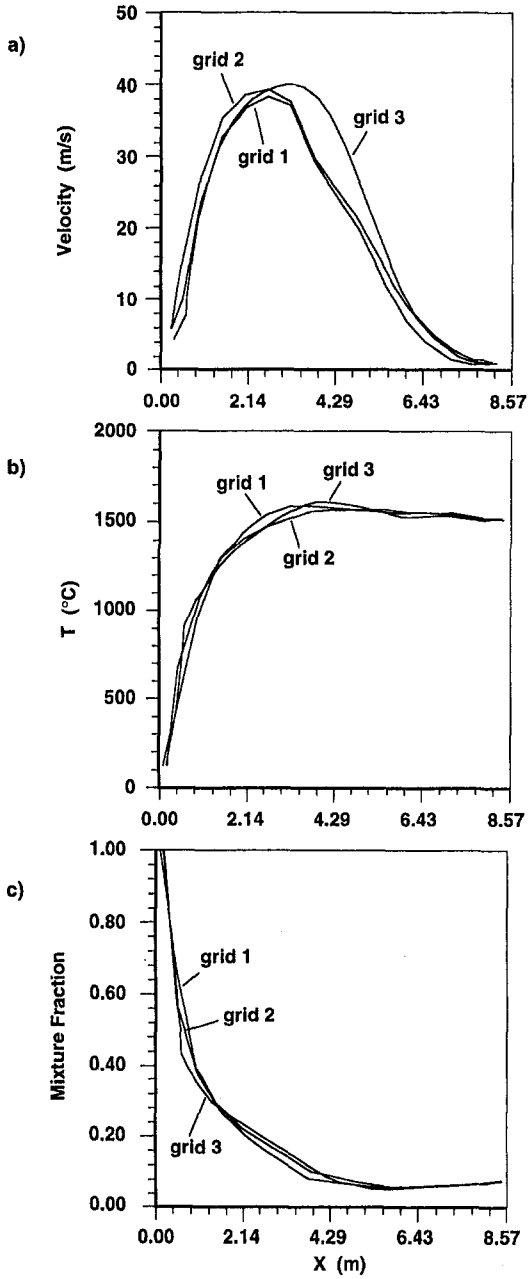


Fig. 3 Predicted velocity, temperature, and mixture fraction profiles along the axis of the top level burner close to the side walls (grid 1: 16 × 34 × 60; grid 2: 16 × 54 × 98; grid 3: 32 × 34 × 60)

boiler analyzed here includes consideration of the numerical solutions obtained for the three grids. This enables one to find out whether discrepancies may be attributed to grid refinement or not.

The comparison between predicted and measured temperature profiles along the axis of inspection ports 1.5 and 3.6 is illustrated in Fig. 4. Close to the back wall and just above the ash pit (inspection port 1.5; see Fig. 1(*b*)) the predicted temperature profile is relatively flat and nearly grid independent. However, the temperature is underpredicted by 50° to 150°C. The agreement between the three different numerical solutions suggests that this discrepancy is due to the modeling assumptions rather than numerical reasons. Since this profile is in a region where the fuel has already been consumed, the mixture fraction fluctuations are negligible. Therefore, the underprediction of temperature reveals an overprediction of the radiation loss. This suggests that the absorption coefficient was overesti-

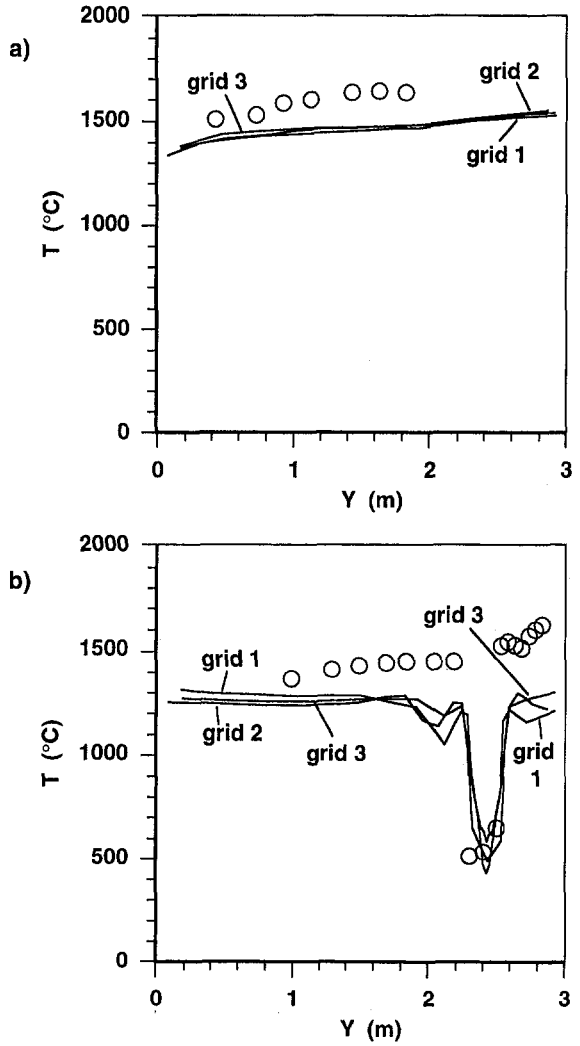


Fig. 4 Comparison between predicted (solid lines) and measured (symbols) temperature profiles: (*a*) profile along the axis of inspection port 1.5; (*b*) profile along the axis of inspection port 3.6

mated. A possible explanation is the overestimation of soot concentration since, as seen below, the gas species concentrations are in good agreement with the data.

Inspection port 3.6 is located near the front wall and at the top burners level (see Fig. 1(*b*)). Along the axis of this port the temperature profiles computed using the three grids mentioned above are similar, as shown in Fig. 4(*b*). The temperature is underpredicted, except in front of the combustion air inlet ($y \approx 2.4$ m), where a temperature drop is observed due to the stream of relatively cold combustion air. The temperature rises for larger values of y due to the heat released in the combustion of the fuel injected through the burner closest to port 3.6. There may be several reasons for the temperature underestimation, but first it should be mentioned that this profile is located in a region of steep temperature gradients (see Fig. 3 for small x values) and very close to the burner inlet. It would be very useful to have temperature measurements along the burner axis, but this is not possible because there is no physical access for the probes.

The burner geometry is rather complex and its details are not taken into account. Due to these limitations, it is not surprising that there is a discrepancy between the data and the predictions for the profile through port 3.6. Apart from the geometry simplifications and the assumptions underlying the combustion model, namely the instantaneous fuel vaporization, other possible explanations for the temperature underprediction may be devised: an overprediction of the heat radiation loss; errors in the assumed relationship between instantaneous values of enthalpy and mixture fraction; and an overestimation of mixture fraction fluctuations. However, much more detailed data would be needed to draw conclusions about the actual reason of the temperature underprediction.

Figure 5 shows the comparison between predicted and measured O_2, CO_2, and CO mole fractions on a dry basis through inspection ports 1.5 and 2.6. At inspection port 1.5 the predicted CO_2 profiles are grid independent and in good agreement with the data, although there is a slight overprediction of CO_2 mole fraction. The influence of grid refinement can be seen in the predicted O_2 and CO profiles. Nevertheless, there appears to be a tendency for a slight overprediction of CO and a slight underprediction of O_2. The general agreement, however, is rather good.

The mole fractions of O_2, CO_2, and CO are approximately constant along the axis of inspection port 2.6 up to $y \approx 1.8$ m, and their values are typical of combustion products that recirculate in this region (see Figs. 5(*c*) and 5(*d*)). The presence of CO with levels between 1 and 2 percent suggests that combustion is not complete. For larger values of y, a sharp rise of O_2 along with a sudden decrease of CO_2 mole fractions occurs as the stream of combustion air, injected through the burner closest to the injection port 2.6, is approached. As y increases, the O_2 mole fraction drops and the CO and CO_2 mole fractions increase, as a result of the combustion phenomena. The CO_2 mole fraction decreases again, as y increases, up to the axis of the burner, because in front of the burner there is a fuel-rich mixture, and the oxygen mole fraction is zero. The presence of the streams of atomizing air is recognized only if grid 2 is used. This is the grid with the largest number of grid nodes in the burners region. There are no data in the burner region where combustion takes place because droplets of unburned liquid fuel block the probes and prevent measurements from being taken closer to the burner axis.

As far as the species mole fraction profiles though inspection port 2.6 are concerned, it can be seen that the influence of the

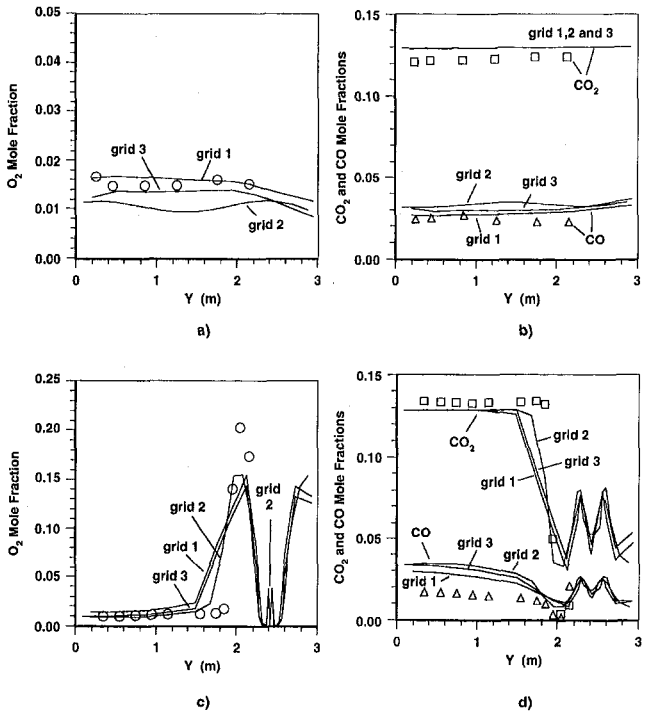


Fig. 5 Comparison between predicted (solid lines) and measured (symbols) mole fraction profiles: (*a*) O_2 profile along the axis of inspection port 1.5; (*b*) CO_2 and CO profiles along the axis of inspection port 1.5; (*c*) O_2 profile along the axis of inspection port 2.6; (*d*) CO_2 and CO profiles along the axis of inspection port 2.6

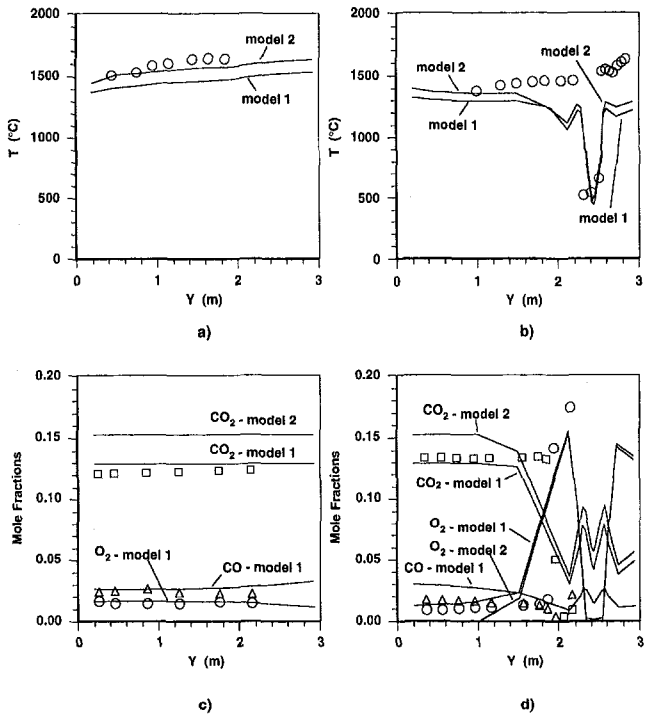


Fig. 6 Comparison between predictions obtained using two different combustion models (model 1: chemical equilibrium; model 2: SCRS; symbols: data): (*a*) temperature profile through port 1.5; (*b*) temperature profile through port 3.6; (*c*) mole fraction profiles through port 1.5. (□: CO_2; △: CO; ○: O_2); (*d*) mole fraction profiles through port 2.6. (□: CO_2; △: CO; ○: O_2)

grid on the predictions is negligible. All the predicted profiles show a good agreement with the data, at least qualitatively. The CO_2 mole fraction is slightly underpredicted up to the burner position. In front of the burner the measured mole fraction decreases rapidly with distance from the burner, whereas the numerical solution exhibits a smoother behavior. However, grid 2 shows a steeper gradient than grids 1 and 3, suggesting that a finer grid along the y direction might improve the results in this region. The predicted O_2 mole fraction is in close agreement with the data up to $y = 1.5$ m. Farther downstream the rise in O_2 mole fraction is smoother than the data suggest and, once again, the steeper gradient was calculated by grid 2. The peak value of O_2 is underpredicted and this may be attributed to a need for grid refinement. The CO mole fraction is overpredicted but the shape of the profile is in good agreement with the measurements.

The results presented so far show that although it cannot be argued that the results are completely grid independent, it can be concluded that the discrepancies between the data and the predictions are in most cases only explainable by physical reasons or experimental uncertainties. The range of the errors attributable to grid refinement was identified by performing the calculations using three different grids. Therefore, in the following discussion only the results obtained using grid 1 shall be considered.

The influence of the combustion model on the predicted results was investigated by repeating the calculations using the SCRS model instead of the chemical equilibrium model. The results of the two models and the measured values are compared in Fig. 6.

It can be seen that the SCRS yields higher temperature levels. This is an expected result since this model does not take dissociation into account. The SCRS yielded an improvement of the predicted temperature profile through port 1.5 but only marginal improvements were achieved for the profile through port 3.6. However, the gas composition is not adequately predicted by the SCRS model. In fact, Fig. 6(*c*) shows that, according to this model, no oxygen is present along the axis of inspection port 1.5. This is not corroborated either by the chemical equilibrium model or by the experimental data. Consequently, the CO_2 mole fraction is significantly overpredicted by the SCRS model. A similar behavior can be seen through port 2.6. Out of the burner region, the SCRS model overpredicted the CO_2 mole fraction and predicted no oxygen. On the other hand, the chemical equilibrium model is in good agreement with the data. Moreover, the SCRS does not take the presence of CO into account.

Hence, it can be concluded that the chemical equilibrium model is preferable to the SCRS model. The reason the SCRS model yielded a better agreement with the measurements for the temperature profile through port 1.5 should not be attributed to its superiority but rather to an overestimated heat loss. This was compensated by the higher temperature calculated when the SCRS model is employed.

5.2 Influence of the Air/Fuel Ratio. The effect of decreasing the air/fuel ratio is illustrated in Fig. 7, where the comparison between predicted and measured temperature pro-

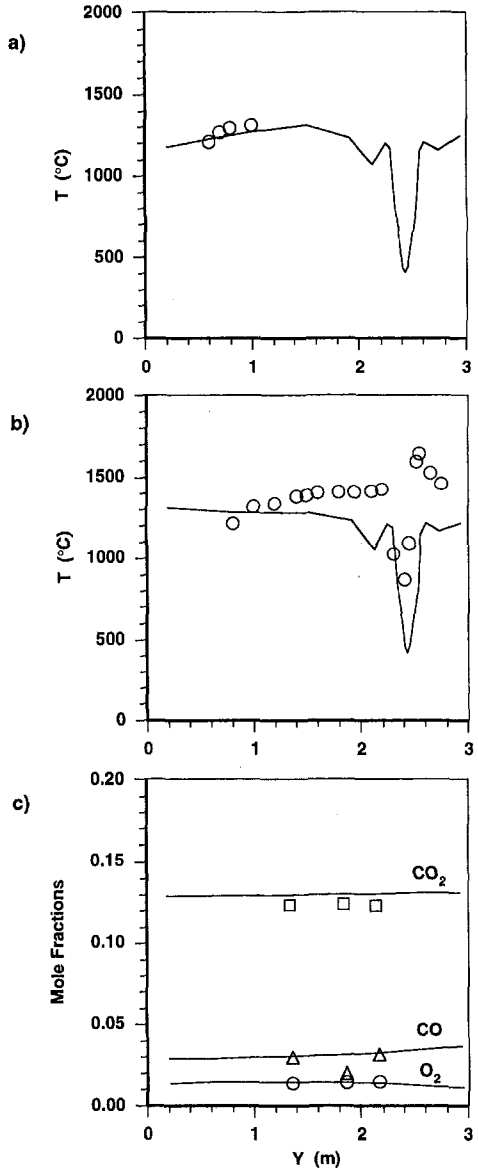


Fig. 7 Comparison between predicted (solid lines) and measured (symbols) for 0.75 percent O_2 in flue gases: (*a*) temperature profile through port 2.6; (*b*) temperature profile through port 3.6; (*c*) mole fraction profiles through port 1.5

files through ports 2.6 and 3.6 and mole fraction profiles through port 1.5 are presented. These results (test case 2) were obtained for a volumetric concentration of O_2 in flue gases of 0.75 percent, whereas at standard operating conditions (test case 1) this concentration is 1.5 percent.

Along the axis of inspection port 2.6, the calculated temperatures are in close agreement with the data, but this is restricted to a small zone out of the flame region. The predicted and measured temperature profiles through port 3.6 are very similar to those presented in Fig. 4(*b*) for standard operating conditions. The variation in gas temperatures between the two test cases is very small. Nevertheless, the measured gas temperature at the outer side of the flame region is slightly higher for case 1 and this is corroborated by the predictions. The temperature drop observed for $y \approx 2.4$ m is caused by the stream of combustion air.

The computed mole fractions of CO_2, CO, and O_2 through port 1.5 are in good agreement with the data. The CO_2 mole fraction is about the same for test cases 1 and 2. However, the reduction of the air/fuel ratio yielded a small increase of CO and a small decrease of O_2 mole fractions. These variations are correctly predicted by the model.

5.3 Influence of the Boiler Load. The influence of decreasing the boiler load from 250 MWe (test case 1) to 125 MWe (test case 3) is shown in Fig. 8. The reduction of the boiler load is achieved by turning off the burners of the top level and blocking air access to these burners. However, the blockage is not completely effective and some air escapes through the burners.

The gas temperature through inspection port 1.5 decreases with the power load as both experiments and calculations reveal. The predicted temperature profile for test case 3 is in good agreement with the data, suggesting that, in this case, the radiative heat loss is well predicted. The gas temperature profile along the axis of port 2.6 shows the same kind of discrepancy observed for temperature profiles close to the front wall in other test cases: The temperature is underpredicted except close to the side wall and in front of the combustion air inlet where the temperature falls down. The temperature profile through port 3.6 is rather different because the burners at the third level are out of service. The calculated temperature is constant during the first 2 m and then drops due to the mixture of combustion gases with cold air that escapes through the top level burners. This temperature profile is correctly predicted despite a slight temperature underprediction in the flat temperature region.

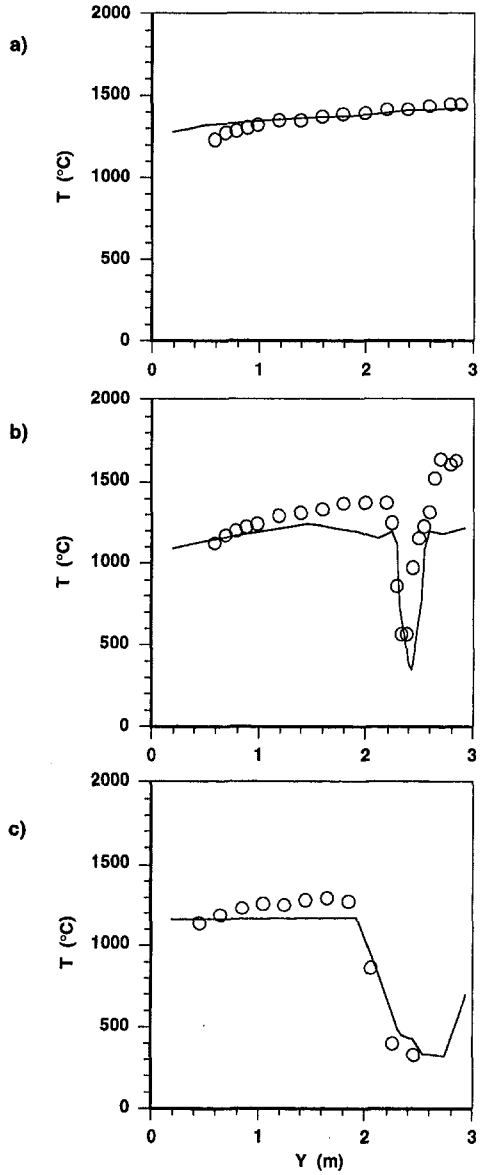


Fig. 8 Comparison between predicted (solid lines) and measured (symbols) temperature profiles at partial load (125 MWe): (*a*) profile through port 1.5; (*b*) profile through port 2.6; (*c*) profile through port 3.6

6 Conclusions

This paper reports the application of a mathematical model to a power station boiler of the Portuguese Electricity Utility. The model is based on the numerical solution of the partial differential equations governing conservation of mass, momentum, energy, and chemical species. The results are compared with measurements recently obtained for several operating conditions.

Calculations were performed using three different grids and a careful analysis was conducted to conclude if the solution was dependent on the grid refinement. The results obtained for standard operating conditions are in good agreement with the available data, except near the burner exit where the temperature is underpredicted. The influence of grid refinement on the predictions is generally negligible, compared with the differences between the numerical solutions and the data, although this is not always the case. Situations where grid refinement might influence the predictions were identified.

The simple chemically reacting system combustion model yields concentrations of the chemical species that disagree with the data. Contrary to the data, according to this model there would be no oxygen close to the side wall. In addition, it did not take CO into account. Therefore, the chemical equilibrium model was preferred since it does not suffer from these shortcomings.

The influence of the air/fuel ratio and power load were correctly predicted by the model, demonstrating its ability to perform parametric studies.

The data are restricted to a small number of temperature and chemical species profiles. Moreover, these profiles are located very close to the front and back walls. This is a consequence of the position of inspection ports that provide physical access to the probes. Despite these limitations, this is one of the most complete sets of data for full-scale utility boilers ever reported, and the evaluation reported here has proved the predictive capabilities of the mathematical model, although there is room for further improvements.

Acknowledgments

This work has been performed within the ESPRIT II European Collaborative Research Programme (Project No. 2192-AIMBURN).

References

Abbas, A. S., and Lockwood, F. C., 1986, "Prediction of Power Station Combustors," *21st Symposium (Int.) on Combustion,* The Combustion Institute, pp. 285–292.

Abou-Ellail, M. M. M., Gosman, A. D., Lockwood, F. C., and Megahed, I. E. A., 1978, "Description and Validation of a Three-Dimensional Procedure for Combustion Chamber Flows," *Journal of Energy,* Vol. 2, No. 2, pp. 71–80.

Bonvini, M., Piana, C., and Vigevano, L., 1991, "Predictions of Thermal NO_x Emissions From Utility Boiler Furnaces," *Proc. 1st International Conference on Combustion Technologies for a Clean Environment,* Vilamoura, Portugal, Sept. 3–6, Vol. 2, paper 23.3.

Boyd, R. K., and Kent, J. H., 1986, "Three-Dimensional Furnace Computer Modelling," *21st Symposium (Int.) on Combustion,* The Combustion Institute, pp. 265–274.

Butler, B. W., and Webb, B. W., 1991, "Local Temperature and Wall Radiant Heat Flux Measurements in an Industrial Scale Coal Fired Boiler," *Fuel,* Vol. 70, pp. 1457–1464.

Carvalho, M. G., and Coelho, P. J., 1990, "Numerical Prediction of an Oil-Fired Water-Tube Boiler," *Engineering Computations—Int. J. Computer Aided Engineering and Software,* Vol. 7, No. 3, pp. 227–234.

Carvalho, M. G., Coelho, P. J., and Costa, F., 1990, "A Comparative Study of the Performance of a Large Industrial Boiler for Fuel Oil and Natural Gas Firing," *Proc. 6th ESPRIT CIM—Europe Annual Conference,* Lisbon, May 15–17.

Carvalho, M. G., Coelho, P. J., and Costa, F., 1991, "Mathematical Modelling of a Power Station Boiler of the Portuguese Electricity Utility," *Proc. 2nd European Conference on Industrial Furnaces and Boilers,* Vilamoura, Portugal, Apr. 2–5, Vol. 2.

Cassiano, J., Heitor, M. V., Moreira, A. L. N., and Silva, T. F., 1994, "Temperature, Species and Heat Transfer Characteristics of a 250 MWe Utility Boiler," *Combust. Sci. and Tech.,* Vol. 98, pp. 199–215.

Coelho, P. J., and Carvalho, M. G., 1992, "Heat Transfer in Power Station Boilers," ASME AES-Vol. 27, R. F. Boehm, ed., New York, pp. 365–372.

Coimbra, C. F. M., Azevedo, J. L. T., and Carvalho, M. G., 1994, "3-D Numerical Model for Predicting NO_x Emissions From an Industrial Pulverized Coal Combustor," *Fuel,* Vol. 73, pp. 1128–1134.

Epple, B., and Schnell, U., 1992, "Modelling and Simulation of Coal Combustion Process in Utility Boiler Furnaces," *Proc. 2nd International Forum on Expert Systems and Computer Simulation in Energy Engineering,* Erlangen, Germany, Mar. 17–20, paper 12-L.

Fiveland, W. A., and Wessel, R. A., 1988, "Numerical Model for Predicting Performance of Three-Dimensional Pulverized-Fuel Fired Furnaces," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 110, pp. 117–126.

Fiveland, W. A., and Latham, C. E., 1991, "Use of Numerical Modelling in the Design of a Low-NO_x Burner for Utility Boilers," *Proc. 1st International Conference on Combustion Technologies for a Clean Environment,* Vilamoura, Portugal, Sept. 3–6, Vol. 1, paper 15.1.

Gillis, P. A., and Smith, P. J., 1990, "An Evaluation of Three-Dimensional Computational Combustion and Fluid-Dynamics for Industrial Furnaces Geometries," *23rd Symposium (Int.) on Combustion,* The Combustion Institute, pp. 981–991.

Gordon, S., and McBride, B. J., 1971, "Computer Program for Calculation of Complex Equilibrium Composition, Rocket Performance, Incident and Reflected Shocks and Chapman–Jouguet Detonations," NASA SP-273.

Görner, K., and Zinser, W., 1988, "Prediction of Three-Dimensional Flows in Utility Boiler Furnaces and Comparison With Experiments," *Combust. Sci. and Tech.,* Vol. 58, pp. 43–58.

Hottel, H. C., and Sarofim, A. F., 1967, *Radiative Transfer,* McGraw-Hill, New York.

Jones, W. P., and Whitelaw, J. H., 1982, "Calculation Methods for Reacting Turbulent Flows: a Review," *Combustion and Flame,* Vol. 48, pp. 1–26.

Khan, I. M., and Greeves, G. A., 1974, "A Method for Calculating the Formation and Combustion of Soot in Diesel Engines," in: *Heat Transfer in Flames,* Afgan, and Beer, eds., pp. 391–402.

Launder, B. E., and Spalding, D. B., 1974, "The Numerical Computation of Turbulent Flows," *Comput. Methods Appl. Mech. Eng.,* Vol. 3, pp. 269–289.

Lockwood, F. C., and Naguib, A. S., 1975, "The Prediction of the Fluctuations in the Properties of Free, Round-Jet, Turbulent, Diffusion Flames," *Combustion and Flame,* Vol. 24, pp. 109–124.

Lockwood, F. C., and Shah, N. G., 1981, "A New Radiation Solution Method for Incorporation in General Combustion Prediction Procedures," *18th Symposium (Int.) on Combustion,* The Combustion Institute, pp. 1405–1414.

Lockwood, F. C., Papadopoulos, C., and Abbas, A. S., 1988, "Prediction of a Corner-Fired Station Combustor," *Combust. Sci. and Tech.,* Vol. 58, pp. 5–24.

Luo, X.-L., Boyd, R. K., and Kent, J. H., 1991, "Computational Investigation of Burnout in a Furnace Firing Pulverized Coal," *J. Institute of Energy,* Vol. 64, Dec., pp. 230–238.

Magnussen, B. F., and Hjertager, B. H., 1977, "On Mathematical Modelling of Turbulent Combustion With Special Emphasis on Soot Formation and Combustion," *16th Symposium (Int.) on Combustion,* The Combustion Institute, pp. 719–728.

Robinson, G. F., 1985, "A Three-Dimensional Analytical Model of a Large Tangentially Fired Furnace," *J. Institute of Energy,* Sept., pp. 116–150.

Sargianos, N. P., Anagnostopoulos, J., and Bergeles, G., 1990, "A Numerical Algorithm for Gas Combustion in Utility Boilers," in: *Advanced Computational Methods in Heat Transfer,* L. C. Wrobel, C. A. Brebbia, and A. J. Novak, eds., Springer-Verlag, Vol. 3, pp. 155–170.

Truelove, J. S., 1976, "A Mixed Grey Gas Model for Flame Radiation," AERE Harwell, Report HL76/3448/KE.