Transactions of the ASME®



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

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

# Journal of Engineering for Gas Turbines and Power

**Published Quarterly by The American Society of Mechanical Engineers**

**VOLUME 124 • NUMBER 2 • APRIL 2002**

## TECHNICAL PAPERS



(**Contents continued on inside back cover**)

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

Journal of Engineering for Gas Turbines and Power APRIL 2002

Volume 124, Number 2



## ANNOUNCEMENTS AND SPECIAL NOTES

# Influence of Processing on Microstructure and Performance of Electron Beam Physical Vapor Deposition (EB-PVD) Thermal Barrier Coatings

**U. Schulz**
**K. Fritscher**
**C. Leyens**
**M. Peters**

DLR–German Aerospace Center,
Institute of Materials Research,
D-51170 Cologne, Germany

*The paper addresses the effect of processing parameters on microstructure and lifetime of electron beam physical vapor deposition, partially yttria-stabilized zirconia (EB-PVD PYSZ) coatings deposited onto NiCoCrAlY-coated Ni-base superalloys. In particular, the formation of a thermally grown oxide layer, an equi-axed zone, and various columnar arrangements of the highly textured PYSZ layers are discussed with respect to processing conditions. Three different microstructures were cyclically tested at 1100°C. The intermediate columnar structure was superior with respect to cyclic life times to a fine and to a coarse columnar structure which was mainly attributed to differences in the elastic properties. The effect of PYSZ microstructure on hot corrosion behavior of the thermal barrier coating (TBC) system at 950°C is briefly discussed.* [DOI: 10.1115/1.1447238]

## Introduction

Efficiency and performance of advanced gas turbines for aircraft and land-based applications can be increased considerably if the gas temperatures in the high-pressure turbine is raised. Application of thermal barrier coatings (TBCs) on superalloy turbine blades allows such an increase of gas temperature owing to their thermal insulation properties and correspondingly lowered temperature of the metal components.

Partially yttria-stabilized zirconia (PYSZ) coatings have demonstrated their capabilities as the most durable TBCs on rotating parts in the turbine. Currently, TBCs applied by electron beam physical vapor deposition (EB-PVD) are more and more employed because the inherent columnar microstructure offers the advantage of a superior tolerance against straining and thermoshock. Although the EB-PVD process is meanwhile prevalent for application of TBCs on turbine blades, there is still potential for further improvements of the properties of the ceramic coating by varying their microstructure. For example, TBCs of a modified microstructure consisting of bulged club-like columns in combination with adjacent S-shaped lean columns having a uniform diameter represent a very low in-plane Young's modulus microstructure (Schulz et al. [1]).

Integration of TBCs into the design of modern aero engine parts—so-called ''designed-in TBCs''—requires a quantitative understanding of the correlation between processing, microstructure, and the resulting lifetime of the coatings. Although oxidation of the bond coat and the associated formation of a thermally grown oxide (TGO) between the bond coat and TBC is essential for the lifetime of TBCs, the microstructure of the ceramic top coat plays another important role for the properties of the coating. The present paper describes the interplay between processing, microstructure, and cyclic lifetime of EB-PVD TBCs having different PYSZ microstructures. The role of the mechanical properties of the ceramic top layer for the lifetime is emphasized, and the effect of PYSZ microstructure on the hot corrosion behavior of the TBC system is briefly discussed.

## Experimental

Cast cylindrical samples of DS IN 100 (6 mm diameter $\times$100 mm) and coupons of CMSX-4[1] (15 mm diameter $\times$1.5 mm thick) were uniformly coated with a 100 to 120-$\mu$m-thick bond coat by EB-PVD. Compositions are given in Table 1. For analysis and measurement of the Young's modulus, 4-mm thick plates of IN 617 were bond coated on one side. All samples were bead peened and subsequently vacuum annealed for four hours at 1080°C to establish a dense and uniform alumina scale on top of the bond coat required for good bonding of the ceramic top layer.

Finally, samples were coated with a 200 to 250-$\mu$m-thick thermal barrier coating of 6.5 to 7.2 wt% $Y_2O_3$–$ZrO_2$. While the coupons were coated in a 150 kW EB-PVD coater, a smaller 60 kW coater was used to coat cylinders and plates. After preheating by radiation in a separate chamber, cylinders and coupons were rotated on a planetary drive in both coaters to achieve a uniform thickness distribution around the circumference of the samples whereas the plates were rotated in standard mode. Unlike the coupons, plates and cylinders were rotated on an axis that was tilted 27 deg towards the evaporation plane. To ensure continuous deposition of the ceramic layers ingots of 63.5-mm diameter (150 kW coater) and 50 mm (60 kW coater), respectively, were bottom fed into the crucible during evaporation. To achieve stoichiometric zirconia a controlled amount of oxygen was bled into the deposition chambers. Rotational speed of the individual samples was adjusted between 4 and 30 $min^{-1}$, condensation rates were between 5 and 7 $\mu$m/min for all samples. In order to study the influence of substrate temperature during deposition, this parameter kept on a constant level, varying from 920°C to 1080°C for different batches by controlling the heat input from the pool during deposition. Several batches were produced with distinct combinations of rotational speed and substrate temperature to generate different PYSZ microstructures.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-579. Manuscript received by IGTI, November 1999; final revision received by ASME Headquarters, February 2000. Associate Editor: D. Wisler.

[1]CMSX-4 is a registered trademark of Cannon-Muskegon Corporation.

**Table 1 Chemical composition (wt%) of substrate alloys and EB-PVD bond coat**

| | Ni | Co | Cr | Al | Mo | Ti | Others |
|---|---|---|---|---|---|---|---|
| IN 100 | Bal. | 15 | 10 | 5.5 | 3 | 4.7 | C, B, Zr |
| CMSX-4 | Bal. | 9 | 6.5 | 5.6 | 0.6 | 1 | 6W, 6.5 Ta, 3 Re, Hf |
| IN 617 | bal. | 12.5 | 22 | 1.2 | 9 | 1.5 Fe | |
| Bond coat | Bal. | 21–22 | 18–20 | 12–13 | | 0.11–0.15 Y | |

Thermal cyclic tests were performed on the CMSX-4 coupons at Oak Ridge National Laboratory (ORNL) and on the IN 100 cylinders at DLR. Samples were exposed to oxygen for 1 hour (ORNL) and for 50 minutes in air (DLR), respectively, at 1100°C and cooled down for 10 minutes between the cycles. At ORNL the specimens were weighed after fixed intervals in an analytical balance. At DLR an optical inspection was employed after each 24 cycles to determine cracks and the location and time of TBC spallation. The testing procedures at ORNL and DLR are described in detail by Leyens et al. [2] and by Fritscher et al. [3], respectively. Hot corrosion tests were performed at ORNL depositing sodium sulfate onto the specimens prior to thermal cycling at 950°C (Leyens et al. [4]).

Samples were analyzed by a field emission gun SEM equipped with an energy dispersive spectrometer (EDS) after TBCs deposition and after testing. TEM studies were performed on a Philips EM 430 electron microscope using standard sandwich-type preparation methods.

The Young's moduli of the TBCs were measured using dynamical mechanical analysis (DMA) on both free-standing TBCs and composite layers attached to the substrate. Free-standing EB-PVD TBCs were prepared by chemical dissolution of the bond coat. The load in the three-point-bending arrangement was 6 MPa (static) $\pm$ 4 MPa (dynamic) for free-standing samples and $15 \pm 10$ MPa for composite samples. The method is described in detail by Szücs [5].

## Results and Discussion

**Processing Related Microstructure of the TBCs.** Manufacturing of an EB-PVD TBC comprises a number of successive steps. After preheating of the parts the ceramic coating is formed during the deposition phase while the cooling step concludes the physically relevant part of the process.

In the *preheating* phase, gaseous and liquid surface contaminants like water are removed, and a thermally grown oxide (TGO) starts or continues to grow. In Fig. 1 such a dense and approximately 1-$\mu$m-thick alumina scale is shown at the interface between the NiCoCrAlY bond coat and the PYSZ TBC in the as-coated condition.

The formation of the TGO during processing depends on various factors, namely:

- material related parameters such as
- substrate alloy,
- bond coat chemistry and microstructure (phase composition, grain size, etc.), and
- TBC chemistry and microstructure (in part);
- processing parameters such as
- bond coat surface treatment,
- preheating and deposition conditions (temperature, duration, atmosphere), and
- post deposition treatments.

In order to keep the TGO formation as reproducible as possible the same procedure was applied to all samples in the present study, the substrate temperature during deposition excepted. Because of the formation of a ''starting'' TGO during the vacuum annealing prior to the deposition and the comparatively small temperature variation during deposition, nearly identical TGO formation can be assumed for all samples. Immediate formation of thermodynamically stable alpha-alumina is considered extremely important for the adherence and lifetime of TBCs (Clarke [6]), but mainly due to kinetic reasons, sometimes harmful transient aluminas or spinels can form instead.

Immediately after the transfer of the samples from the heating to the deposition chamber the growth of the TBC starts. The initial phase of *deposition* plays a crucial role for the formation of microstructure and crystallographic orientation of the TBC. In Fig. 2 a typical TEM micrograph of a TBC close to the substrate is shown.

The first thin layer adjacent to the substrate (denoted A in Fig.

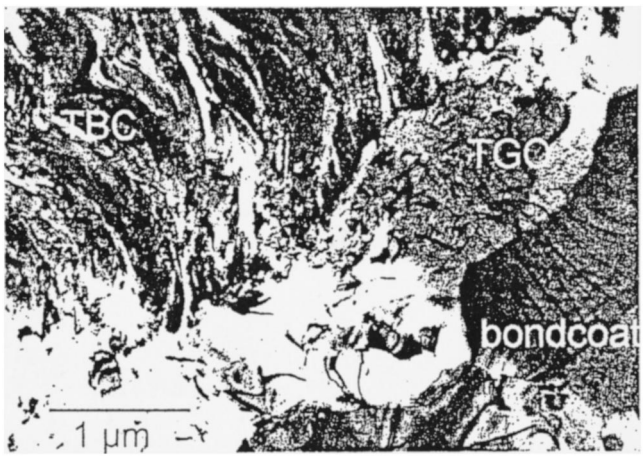



**Fig. 1 TEM picure of interface between NiCoCRAlY bond coat and YPSZ TBC in as-coated condition**

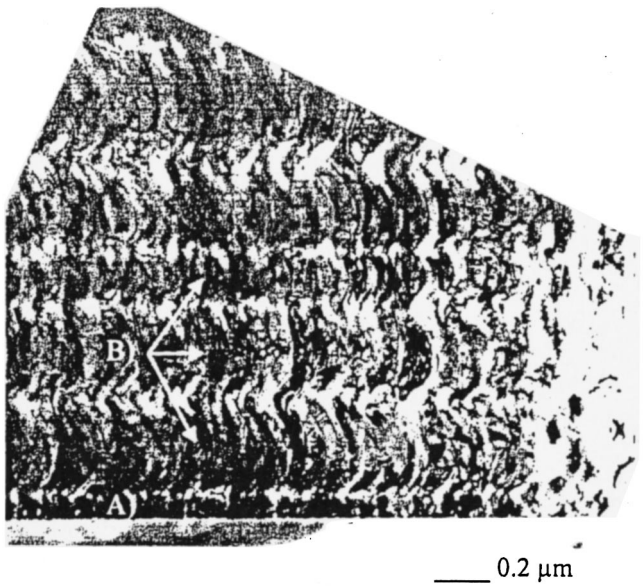



**Fig. 2 TEM of an EB-PVD PYSZ TBC close to the substrate**

2) consists of equiaxed grains of about 30 nm in diameter, forming the so-called equiaxed zone (Unal et al. [7]). The grains are randomly oriented and the thickness of the equi-axed zone varies somewhat from sample to sample. A geometrical model was proposed to explain these thickness differences, taking into account the actual position of the sample during the first rotation (Schulz and Schmuecker [8]). From the model it can be concluded that the equi-axed zone is formed exclusively during the very first rotation of the sample.

Adjacent to the equi-axed zone a layered structure is apparent which is caused by the rotation of the substrate (denoted B). The thickness of the layers shows periodically modulated striations which arise due to changing distances between the evaporation pool and each single sample during evaporation in a planetary drive. A small distance corresponds to a high vapor density and hence to a thicker layer and vice versa. While formation of c-shaped columns is a result of rotation alone, the modulation of layer thickness is directly affected by the sample distance from the melting pool and therefore by the vapor density.

During the first 15 to 20 rotations an evolutionary selection of crystals during growth is evident. The number of columns continually decreases while their diameter increases. A selection of the crystallographic orientation with the most favored growth rate takes place resulting in a strong texture of the TBCs with a {100} plane parallel to the interface. The final morphology of the coatings is mainly ruled by substrate temperature, rotational speed, the vapor incidence angle, surface roughness of the substrate, energy of the vapor particles, and also by chamber pressure, curvature of the substrate, and deposition rate.

Cooling after deposition finalizes the manufacturing of an EB-PVD TBC. Here, only slightly different stresses in the TGO and in the TBC may occur because of different cooling rates at different locations. Large components may cool differently at various positions because of the unbalanced mass distribution which can cause detrimental stress peaks.

In the present study three different microstructures of the columnar ceramic layers were chosen for further evaluation:

- a fine-columnar structure, obtained at a comparably low substrate temperature and low rotational speed,
- an intermediate-columnar structure, obtained at intermediate temperature of approximately 1000°C and intermediate rotational speed of 12 $min^{-1}$, and
- a coarse-columnar structure, obtained at high temperature and high rotational speed.

These microstructures that are exemplified in Fig. 3 can be characterized by the following features.

The columns of the fine-columnar structure often vary in diameter from root to top and from one column to the other. Nucleation of new columns at the middle of the height and columns that do not span from top to bottom across the TBC were frequently observed. The coarse structure is characterized by 10 to 15 $\mu$m thick, regular, and parallel columns. Although the coating density is strongly related to substrate temperature with a higher density for high-temperature deposited TBCs (Fritscher and Bunk [9]) the coarse structure seems to exhibit a more open structure. The intermediate structure shows features in between fine and coarse. The differences of the microstructure were less pronounced for the CMSX-4 coupons than for plates and cylinders because of the process condition chosen for the individual deposition runs. An additional difference was that TBCs on plates and cylinders possessed tilted columns while on coupons columns were oriented rectangular to the surface. A model was proposed (Schulz et al. [10,11]) that describes the microstructure of EB-PVD PYSZ coatings as a function of both temperature and rotation of the samples, thereby extending the classical structural zone model given by Movchan and Demchishin [12].

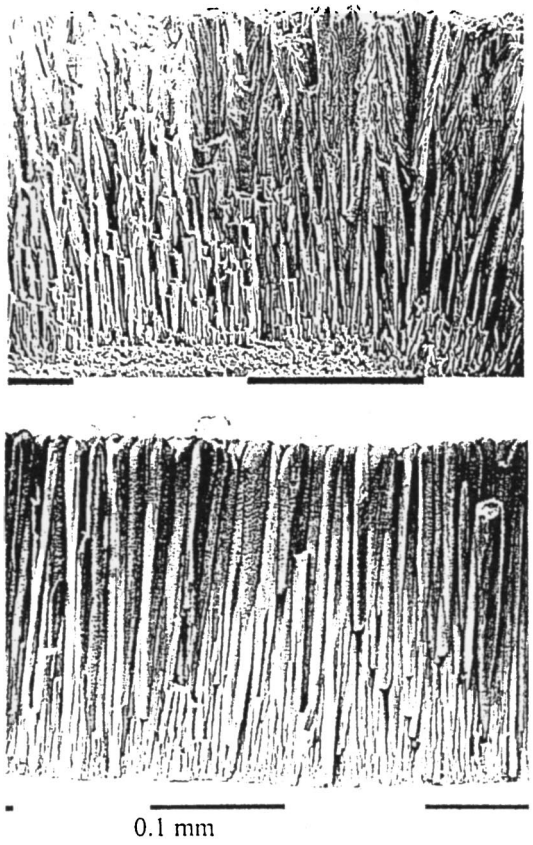



**Fig. 3 Fractured cross sections of EB-PVD TSCs with fine (top) and coarse (bottom) microstructure**

**Cyclic Testing.** Weight change versus number of 1 h cycles data of the CMSX-4 coupons in thermal cyclic tests at 1100°C are summarized in Fig. 4.

Initially, all samples suffered from weight loss that can be attributed during the initial cycles to TBC chipping at the specimen edges. However, this was not considered as a failure criterion. Samples with a fine-columnar structure failed after 750 cycles by spallation of more than 20 percent of the PYSZ from the surface, followed by the coarse structure after 850 cycles. Although losing continuously mass by spallation of the PYSZ, the intermediate structure survived the 1200-cycles test duration without reaching the 20 percent failure criterion. The fine and coarse-columnar

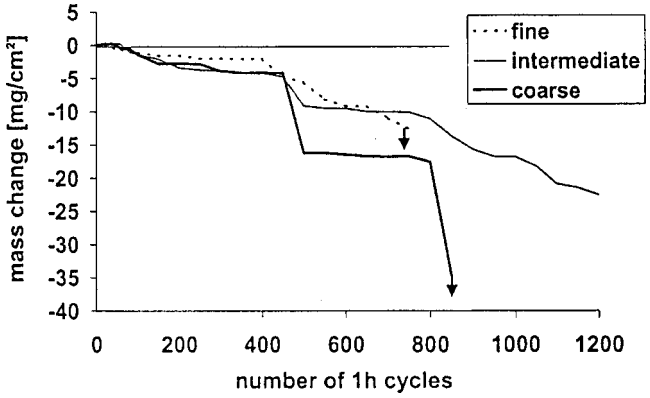



**Fig. 4 Weight change of EB-PVD TBCs on CMSX-4 coupons at 1100°C cycles. Arrows indicate spallation of the TBC.**

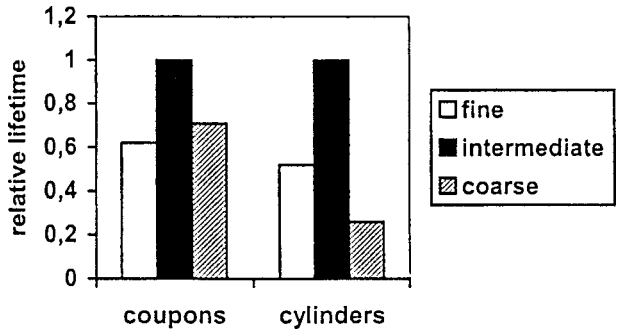



**Fig. 5 Normalized cyclic lifetimes obtained in two alternative furnace tests at 1100°C**

structures failed after cooling down from test temperature at the interface between the bond coat and the TGO with the TGO still adherent to the PYSZ.

A similar behavior was observed for the IN 100 cylinders in the second cyclic test with a slightly changed ranking with regard to TBC microstructures: again, the intermediate structure was the best followed by the fine and the coarse structures, the latter with a cyclic lifetime less than half of that of the intermediate. The results are summarized in Fig. 5. Note that for the intermediate structure on the coupons a basic lifetime of 1200 cycles was chosen although samples did not fail until the test was stopped.

Although the microstructure of the TBC is only one factor among others that influences the lifetime, a distinct influence on the spallation life must be considered. Additionally, the different specimen geometry has also an influence on cyclic life since on curved samples it is easier to get a tensile stress component perpendicular to the surface while on flat samples the free edges are more prone to spallation. The present results suggest there is an optimum microstructure. The larger the variance from the optimum the lower is the cyclic lifetime. As shown by Johnson et al. [13], the mechanical behavior of the ceramic layers with various microstructures is different, especially the elastic modulus and the stress state, and they are further influenced extensively by sintering.

**Elastic Modulus.** A wide range of in plane Young's moduli of EB-PVD TBCs has been published. Table 2 summarizes several room temperature values for YPSZ EB-PVD TBCs. The data measured by Bartsch [14], Kirchhoff [15], Morell [16], and Szücs [5] originate all from TBCs produced in the same coaters at DLR as the samples of the present study. Therefore, the morphologies were similar to the intermediate structure and hence differences of the values in Table 2 must be attributed to other effects than the origin of the samples.

**Table 2 Room temperature Young's modulus of YPSZ EB-PVD TBCs**

| $E$ in GPa | Method | Reference |
|---|---|---|
| 210 | Bulk | |
| | TBC composite | |
| 35 | tensile test | Bartsch [14] |
| 42 | tensile test | Meier [21] |
| 50–55 | laser acoustic test | Kirchhoff [15] |
| 100 | laser ultrasonic (compression) | Morell [16] |
| 15–77[1] | resonant flexural frequency | Johnson [13] |
| 10–30[2] | resonant flexural frequency | Johnson [13] |
| 27 | 3-point-bending (DMA) | Szücs [5] |
| | Free-Standing TBC | |
| 13–15[1] | 3-point-bending (DMA) | Szücs [5] |
| 6[2] | 3-point-bending (DMA) | Szücs [5] |

[1]∥ to rotation axis,
[2]⊥ to rotation axis

A number of factors must be considered that influence the elastic moduli of those TBCs. We believe that the most important factors are

- microstructure of the TBC.
- stress state of the TBC ruled by:
  - substrate material,
  - type of sample (free-standing or composite), and
  - loading mode (tensile, compression, bending, etc).
- orientation of the measured modulus (parallel or perpendicular to rotational axis, in-plane or parallel to columns).
- crystallographic orientation of the stabilized zirconia TBC.
- methodology of the measurement.

Of course, the modulus increases significantly by thermal exposure or in service. Taking one special method into account, namely the DMA, the large difference between Young's moduli of free-standing TBCs and of TBC composites (TBC attached to a bondcoated substrate) is noticeable. Obviously, the stress state is different for both conditions since a composite TBC will be under compression at room temperature because the thick and stiff substrate dictates the strain during cooling from processing temperature according to its higher thermal contraction while a free-standing TBC is nearly stress free. Interestingly, free-standing TBCs exhibit a concave shape if the upper side of the PYSZ coating is considered. In addition to the differences in the microstructure between top and bottom of the TBC as discussed above, a stress gradient through the thickness of the ceramic layer is evident. After annealing above 1000°C the extent of curvature decreases, indicating stress relaxation in the coating. For real TBC systems attached to a metallic blade or vane these findings imply the presence of a stress gradient in the PYSZ. The gradient includes compressive stress components near the bond coat/TBC interface and much lower compressive or even tensile stresses near the free surface of the PYSZ top coating, which are superimposed by inherent thermal mismatch stresses. After annealing, stress relaxation by plastic flow and sintering results in a lower radius of curvature of the free-standing TBCs.

The anisotropy of EB-PVD processed PYSZ coatings is caused by several phenomena. Firstly, substrate rotation during deposition causes a bent columnar microstructure in only one direction which in turn leads to different in-plane moduli. Data reported by Johnson [13] and Szücs [5] clearly demonstrate this anisotropy (Table 2). Parallel to the rotational axis significantly higher Young's moduli were observed than perpendicular. Secondly, because of the strong texture of EB-PVD TBCs (Schulz et al. [17]) and the crystallographic anisotropy of tetragonal stabilized zirconia, the mechanical properties differ both in the plane parallel to the interface and perpendicular to it. Thirdly, in growth direction of the columns perpendicular to the metallic substrate variations of the microstructure from bottom to top cause an anisotropic stress field as stated before. It is difficult to separate these interacting factors from each other but they all contribute to the strong anisotropic properties of EB-PVD TBCs.

Notably, the columnar structure shows a different elastic behavior under compression and under tension. While under tension the columns can widely accommodate strain, the degree of mechanical interconnection between single columns is much higher under compression. This leads to a strain-dependent Young's modulus of the TBC. Figure 6 shows schematically the trends obtained during the first heating of composite samples (after Szücs [5]). Because the uncertainties caused by inaccurate input values into the calculation were quite high only the trends and not the absolute values are plotted. At least data indicated that, as one might expect, increasing tensile stresses reduce the Young's moduli.

In order to study the influence of the microstructure on Young's modulus in detail, the three different TBCs were measured as free-standing layers in the three-point-bending arrangement using the dynamical mechanical analysis (DMA). The results are shown in Fig. 7.

Generally, modulus values increase with increasing tempera-

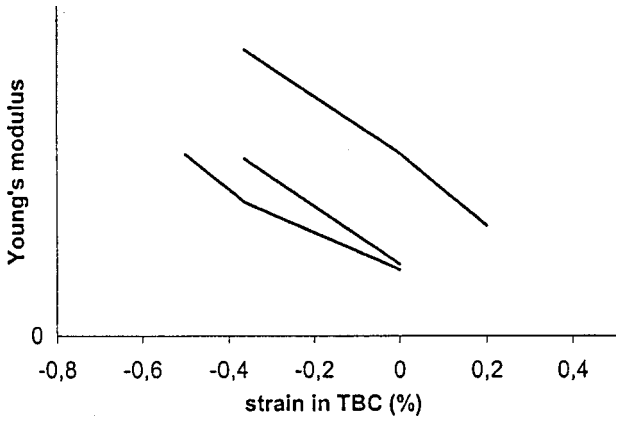



**Fig. 6 Young's modulus of various EB-PVD TBCs having different thickness as a function of mechanical load**

ture. Furthermore, a more pronounced increase of the modulus above 900°C is noticeable for all microstructures. This corresponds to the start of high-temperature related phenomena like sintering and creep. Among the three microstructures compared in this study, the intermediate structure exhibits the lowest modulus and the longest cyclic lifetime, too. If the columnar structure is closer packed and less inter-columnar spacing exists higher stress levels as well as higher Young's moduli will occur. This was observed, indeed, for the fine columnar structure which was assumed to fail early because of higher stresses at the same strain (Fig. 5). On the other hand, the coarse structure exhibited a comparable low modulus but suffered from early failure. In this case, the regular and parallel columns may provide vertical crack paths that lower the resistance against crack propagation. As shown by Johnson [13] low residual stress levels and low in-plane moduli result in a higher TBC strain tolerance as similarly observed in the present study. Contrarily, the inferior behavior of the coarse structure with a low elastic modulus showed that the Young's modulus is only one factor among others that controls the cyclic life of a TBC system. Substrate material, bond coat oxidation, and sintering of the TBC are examples for other important life-determining factors. For the coupon samples, the TGO formation during cyclic testing was nearly identical for all three microstructures (Leyens et al. [2]). The choice of the substrate alloy has a large influence on lifetime with nearly a doubled number of cycles to failure for IN 100 compared to CMSX-4 (Kaden et al. [18]), indicating that chemical factors affecting the adherence of the TGO to the bond coat play a critical role.

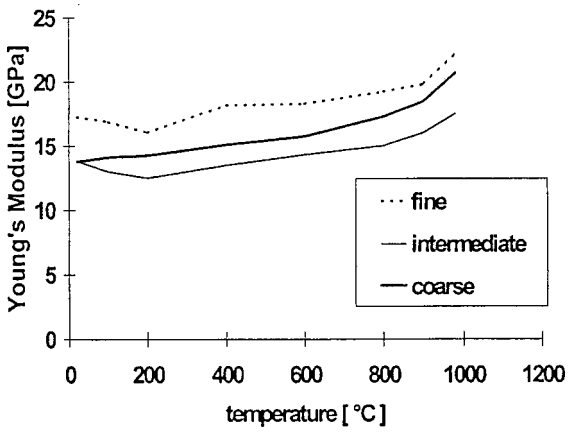



**Fig. 7 Young's modulus of free-standing EB-PVD TBCs with different microstructures**

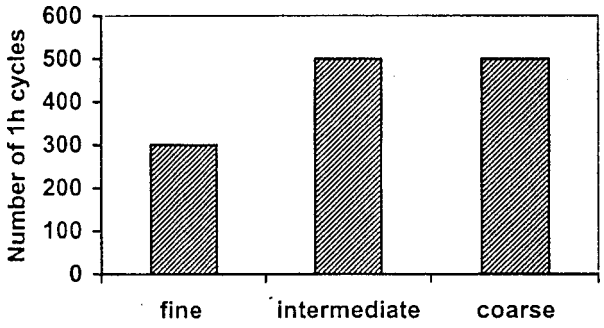



**Fig. 8 Lifetime of different PYSZ microstructures tested under sodium sulfate hot corrosion conditions at 950°C**

A recent study by Walter and Eigenmann [19] on the same three microstructures showed also significant differences of the TBC behavior under mechanical load in a four-point-bending arrangement. Two micro transducers were mounted on the metallic side of flat samples. Various information obtained from acoustic emission (AE) were analyzed, in particular signals above a set threshold (a cascaded hit), counts, amplitude, and energy of the counts. Uncoated bar metal specimens were used to separate the AE data associated with the coating from that of the metal or of the loading machine.

By summing the hits for the duration of each loading, the different microstructures have been compared for the TBC in tension and compression as well. Under tension, events are recorded earlier in the loading history for the fine structure compared to the coarse microstructure. Whereas the coarse microstructure behaved better under tension, it performed significantly worse than the other two microstructures under compression. It was argued by Walter and Eigenmann [19] that the smaller inter-columnar volume in the coarse microstructure absorbs compressive deformation, and that therefore the coarse microstructure is unable to accommodate as much compressive deformation as the intermediate and the fine microstructures. The intermediate structure showed an ordinary behavior under both tension and compression and may therefore be favored under the complex loading during thermal cycling where, for example, during rapid cooling of the TBC from the outside tension may occur in the ceramic while it is under compression at room temperature. Thus, the behavior in the acoustic emission monitored four-point-bending experiment provided some additional information regarding the ranking in cyclic testing shown in Fig. 5.

**Cyclic Hot Corrosion Testing.** Recent results on the hot corrosion behavior of TBC systems at 950°C indicate that the PYSZ microstructure may play a critical role (Leyens et al. [20]). The fine columnar microstructure failed much earlier than the intermediate and coarse microstructures (Fig. 8). However, unlike for oxidation alone, the presence of molten deposits on the surface of the PYSZ coatings affects the durability of the TBCs with regard to its resistance to deposit infiltration, rather than regarding the mechanical properties of the ceramic top layer itself. It was observed that TBC failure by spallation of the PYSZ occurred earlier if the inter-columnar volume was higher, as in the case of the fine columnar microstructure, since a larger amount of deposits penetrates the PYSZ and subsequently attacks the alumina scale at the bond-coat PYSZ interface.

Clearly, failure occurred, once the deposits had reached the metal-oxide interface and had caused hot corrosion attack of the bond coat (Fig. 9). Formation of voluminous oxide with high porosity facilitates crack propagation within the oxide scale, finally leading to delamination of the ceramic top layer.

At present, it is believed that the effect of different PYSZ microstructures is rather indirect by simply providing more or less

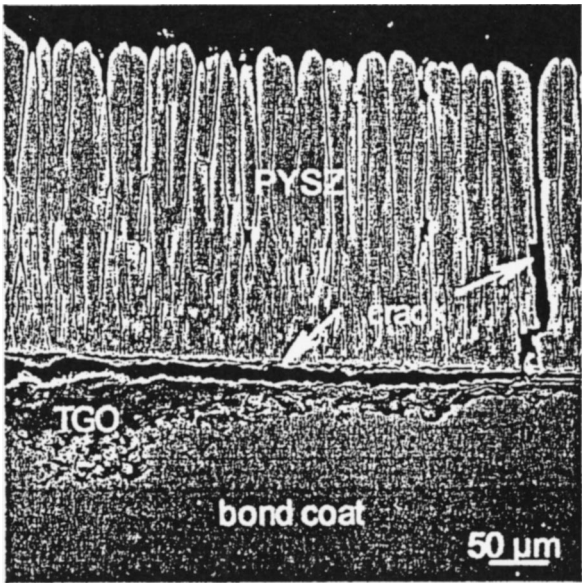



**Fig. 9 Typical failure of a TBC system tested under hot corrosion conditions**

pathways for deposit infiltration. However, promotion of sintering-like effects observed in the presence of nominally corrosive deposits may also affect the PYSZ properties directly by changing the mechanical properties of the ceramic layer similar to mechanisms observed for oxidation exposure alone.

## Conclusions

Partially yttria-stabilized zirconia TBCs produced by EB-PVD have been examined. Formation of a thermally grown oxide already during processing and the development of a zone with equiaxed grains close to the interface to the bond-coat TBC followed by layers of bent columns are the consequence of the deposition conditions prevailing during the manufacture of the coatings. Three different microstructures of the TBC, namely a fine, intermediate, and coarse-columnar structure, were achieved by appropriate combinations of substrate temperature and rotational speed during the EB-PVD process. The investigations lead to the following conclusions:

- The cyclic life of the intermediate structure demonstrated superior behavior in cyclic testing at 1100°C on two different substrate alloys. The ranking of the fine and coarse-columnar microstructures was inconsistent for different sample geometries and substrate alloys.
- Cyclic lifetime and Young's modulus correspond to each other to a great extent with longer lifetimes for the low modulus TBCs.
- The Young's modulus of EB-PVD TBCs is strain-dependent with significantly higher values under compression than under tension. This leads to differences in the acoustic emission behavior of the three microstructures between compression and tension.
- In hot corrosion test the inter-columnar spacing may be a dominating life-determining factor.

## Acknowledgments

The authors gratefully acknowledge careful manufacturing of coatings by J. Brien, C, Kröder, H. Mangers, and H. Schurmann and the TEM work of M. Schmücker. They owe special thanks to F. Szücs (TU Berlin) and M. E. Walter (Ohio State University) for the valuable collaboration. Part of this work was performed by CL on sabbatical leave to Oak Ridge National Laboratory, Metals and Ceramics Division. The financial support of the HGF strategy found in the project ''Gasturbine'' is highly acknowledged.

## References

[1] Schulz, U., Kröder, C., Brien, J., Schurmann, H., and Fritscher, K., 1999, U.S. Patent 6063435.
[2] Leyens, C., Schulz, U., Pint, B. A., and Wright, I. G., 1999, ''Influence of EB-PVD TBC Microstructure on Thermal Barrier Coating System Performance Under Cyclic Oxidation and Hot Corrosion Conditions,'' Surf. Coat. Technol., **120–121**, pp. 68–76.
[3] Fritscher, K., Schulz, U., and Schmücker, M., 1999, ''EB-PVD TBC lifetime Response to Various Bond Coat Pretreatments,'' *Cyclic Oxidation of High Temperature Materials*, M. Schütze and W. J. Quadakkers, eds., The Institute of Materials, London, EFC public. No. 27, pp. 383–391.
[4] Leyens, C; Wright, I. G., Pint, B. A., and Tortorelli, P. F., 1999, ''Significance of Experimental Procedures on the Hot Corrosion Behavior of Nickel-Base Alloys Under Cyclic Conditions,'' *Cyclic Oxidation of High Temperature Materials*, M. Schütze and W. J., Quadakkers, eds., The Institute of Materials, London, EFC Pub. No. 27, pp. 169–186.
[5] Szücs, F., 1998, ''Thermomechanische Analyse und Modellierung plasmagespritzter und EB-PVD aufgedampfter Wärme-dämmschicht-Systeme für Gasturbinen,'' Fortschr.-Ber. VDI Verlag Düsseldorf, Serial **5**, No. 518.
[6] Clarke, D. R., Sergo, V., and He, M.-Y., 1999, ''Precursor to TBC Failure by Constrained Phase Transformation in the Thermally Grown Oxide,'' *Elevated Temperature Coatings: Science and Technology III*, J. M. Hampikian and N. B. Dahotre, eds., TMS, Warrendale, PA, pp. 67–78.
[7] Unal, O., Mitchell, T. E., and Heuer, A. H., 1994, ''Microstructures of $Y_2O_3$-stabilized $ZrO_2$ Electron Beam-Physical Vapor Deposition Coatings on Ni-Base Superlloys,'' J. Am. Ceram. Soc., **77/4**, 984–992.
[8] Schulz, U., and Schmücker, M., 2000, ''Microstructure of $ZrO_2$ Thermal Barrier Coatings Applied by EB-PVD,'' Mater. Sci. Eng., A, **276**, pp. 1–8.
[9] Fritscher, K., and Bunk, W., 1990, ''Density-Graded TBC's Processed by EB-PVD,'' *1st International Symposium on Functionally Gradient Material, Proc.* M. Yamanouchi et al., eds., Society of Non. Traditional Technology, Tokyo, Japan pp. 91–96.
[10] Schulz, U., Fritscher, K., Leyens, C., Peters, M., and Kaysser, W. A., 1997, ''Thermocyclic Behavior of Differently Stabilized and Structured EB-PVD Thermal Barrier Coatings,'' Material-wissenschaft und Werkstofftechnik, **28**, pp. 370–376.
[11] Schulz, U., Fritscher, K., Rätzer-Scheibe, H.-J., Peters, M., and Kaysser, W. A., 1997, ''Thermocyclic Behavior of Microstructurally Modified EB-PVD Thermal Barrier Coatings,'' Mater. Sci. Forum, **251–254**, pp. 957–964.
[12] Movchan, B. A., and Demchishin, A. V., 1969, ''Study of the Structure and Properties of Thick Vacuum Condensates of Nickel, Titanium, Tungsten, Aluminum Oxide and Zirconium Dioxide,'' Fiz. Met. Metalloved., **28**, pp. 83–90.
[13] Johnson, C. A., Ruud, J. A., Bruce, R., and Wortman, D., 1998, ''Relationship Between Residual Stress, Microstructure, and Mechanical Properties of Electron Beam-Physical Vapor Deposited Thermal Barrier Coatings.'' Surf. Coat. Technol., **108–109**, pp. 80–85.
[14] Marci, G., Bartsch, M., and Mull, K., 1999, ''Ermittlung des Elastizitätsmoduls von keramischen Wärmedämmschichten für thermisch und mechanisch hoch belastete Gasturbinenkomponenten,'' *Proceedings DVM-conference Werkstoffprüfung 1999*, Dec. 2–3, Bad Nauhelm, pp. 271–280.
[15] Kirchhoff, G., Langmeier, P., and Hennig, J., 1998, unpublished results.
[16] Morell, P., and Rickerby, D. S., 1999, ''Advantages/Disadvantages of Various TBC Systems as Perceived by the Engine Manufacturer,'' AGARD report 823 ''Thermal barrier coatings'' 20-1/20-9.
[17] Schulz, U., Oettel, H., and Bunk, W., 1996, ''Texture of EB-PVD Thermal Barrier Coatings Under Variable Deposition Conditions,'' Z. Metallkd., **87/6**, pp. 488–492.
[18] Kaden, U., Leyens, C., Peters, M., Kaysser, W. A., 1999, ''Thermal Stability of an EB-PVD Thermal Barrier Coating System on a Single Crystal Nickel-Base Superalloy,'' *Elevated Temperature Coatings: Science and Technology III*, J. M. Hampikian and N. B. Dahotre, eds., TMS, pp. 27–38.
[19] Walter, M. E., and Eigenmann, B., 2000, ''Mechanical Response of Three As-Received EB-PVD Zirconia Thermal Barrier Coating Microstructures,'' Mater. Sci. Eng., A, **282**, 49–58.
[20] Leyens, C., Wright, I. G., and Pint, B. A., 2000, ''Hot Corrosion of an EB-PVD Thermal Barrier Coating System at 950°C,'' Oxidation of Metals, **54**(5/6), pp. 401–424.
[21] Meier, S. M., Nissley, D. M., Sheffler, K. D., 1991, ''Thermal Barrier Coating Life Prediction Model Development,'' Phase II Final Report, NASA CR 189111.

# Development of Hexaaluminate Catalysts for Combustion of Gasified Biomass in Gas Turbines

**E. M. Johansson**

**K. M. J. Danielsson**

**A. G. Ersson**

**S. G. Järås**

KTH-Royal Institute of Technology,
Department of Chemical Engineering
and Technology,
Chemical Technology, S-100 44 Stockholm,
Sweden

*There is an increasing interest in catalytic combustors fuelled by low-heating value (LHV) gases, with a LHV of 5–7 MJ/Nm$^3$. This is because catalytic combustion could be advantageous compared to flame combustion with respect to stable combustion of LHV-gases and low conversions of fuel-N (mainly $NH_3$) to $NO_X$. In the present project, funded by the EU Joule Program, catalytic combustion of gasified wood for gas turbine applications is studied. A synthetic gas mixture of $H_2$, CO, $CO_2$, $H_2O$, $CH_4$, $N_2$, and $NH_3$, that resembles the output from a fluidized bed gasifier using biomass as raw material, is used. The gas mixture is mixed with air at atmospheric pressure and combusted over washcoated cordierite monoliths in a bench-scale laboratory quartz-reactor. The objectives of the work described here are twofold. To begin with, improvement of the thermal stability of hexaaluminate washcoats by substitutions of rare earth or transition metal compounds is being studied. Secondly, catalytic combustion of gasified biomass over these washcoats has been studied in a bench-scale unit. In this on-going project, obtained result show that it is possible to improve the surface area of hexaaluminate compounds up to 17 m$^2$/g after careful synthesis and calcination up to 1400°C for four hours. The selectivity of $NH_3$-conversion to $N_2$ is at present at 60 percent, but varies strongly with temperature. Fuel components such as $H_2$, CO, $C_2H_4$, and $NH_3$ ignite at temperatures close to compressor outlet temperatures. This means that a pilot-flame may not be needed for ignition of the fuel. A comparison between a Pd-impregnated lanthanum hexaaluminate and a Mn-substituted lanthanum hexaaluminate showed that the ignition temperature and the $NO_X$-formation varied strongly over the two different catalysts.*
[DOI: 10.1115/1.1335478]

## Introduction

High-temperature catalytic combustion is a challenging and very promising alternative to ordinary flame combustion in gas turbine combustors [1]. The main advantages with a catalytic combustor are ultra-low emissions of nitrogen oxides ($NO_X$), carbon monoxide (CO), and unburned hydrocarbons (UHC). This has been shown in pilot tests with hybrid combustors fuelled by natural gas in the U.S. [2] and Japan [3,4]. The emission levels are ultra-low, since the catalyst stabilizes combustion of a lean fuel-to-air mixture. Consequently, thermal-$NO_X$ produced from $N_2$ and $O_2$ in the combustion air at peak temperatures is heavily reduced. Catalytic combustors are now being tested in commercial applications in California and are believed to be very cost-competitive. This is in comparison to the cost for a commercial gas turbine equipped with dry low-$NO_X$ technique coupled with SCR to reach the same ultra-low emission standards.

In parallel to the worldwide activities in the field of catalytic combustion of natural gas, there is an increasing interest in catalytic combustors fuelled by low-heating value (LHV) gases, such as gasified biomass. This is because catalytic combustion could be advantageous compared to flame combustion with respect to stable combustion of LHV gases [5] and low conversions of fuel-N to fuel-$NO_X$ [6–8]. The gas is produced from thermochemical conversion of biomass [9], i.e., renewable materials such as wood, fast-growing energy crops or agricultural residues, in a gasifier. In the gasifier, usually a pressurized or atmospheric fluidized bed reactor, biomass is partly combusted and gasified in air at high temperatures, which yields a gas mixture, a liquid residue and a solid residue. After gas cleaning, the gas mixture could be used as a fuel for heat and electricity production or as raw material for production of chemicals.

Substituted hexaaluminates are very promising materials for use as high-temperature combustion catalysts [10–13]. This is true especially for hydrogen and carbon-monoxide-rich LHV-fuels, since fuel components such as $H_2$, CO, $C_2H_4$, and $NH_3$ ignite at low temperatures over these catalysts, compared to methane-rich natural gas.

The objective with the present work was to study catalytic combustion of a synthetic gasified biomass over lanthanum hexaaluminates, substituted with transition metals or impregnated with palladium. The hexaaluminates were washcoated, i.e., applied as a thin layer, on cordierite monoliths. The monolith substrates were mounted in a bench-scale reactor, and tested with emphasis on ignition of gasified biomass, especially combustion of $CO$-$H_2$-$CH_4$-mixtures, and the formation of $NO_X$ from fuel-$NH_3$. In addition to the combustion tests, a series of calcinations was made to study the thermal stability of lanthanum hexaaluminates substituted by barium, manganese, and iron.

This work is part of the European ULECAT-project, a French-Italian-Swedish program with the objective to develop and assess the feasibility of an ultra-low emission gas turbine in the range of 1 to 5 MWe. The gas turbine should be able to run with both biomass derived and liquid fuels. The overall progress in the first three-year phase of this project that started in May 1996 is reported elsewhere [14].

## Experimental

**Preparation and Testing of the Catalysts.** Hexaaluminate powders, were prepared by adding aluminum, rare earth, and transition metal nitrates dissolved in water to an aqueous carbonate solution (pH≈9.0). This carbonate preparation is described in

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Stockholm, Sweden, June 2–5, 1998; ASME Paper 98-GT-338. Manuscript received by IGTI Mar. 18, 1998; final revision received by the ASME Headquarters Mar. 23, 1999. Associate Editor: R. Kielb.

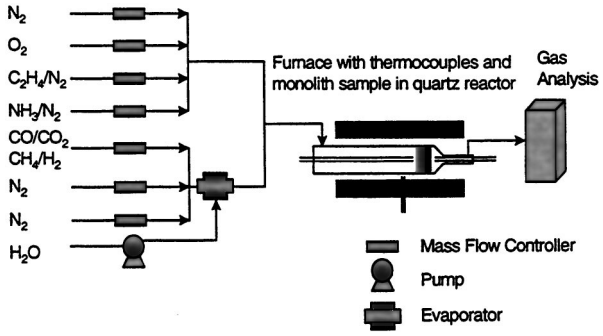



**Fig. 1 Test unit for gas mixing, furnace with reactor and gas analysis**

more detail elsewhere [15]. The precipitate was washed, dried, calcined in air to 1000°C, and ball-milled in a water slurry over night. The cordierite monoliths (400 cpsi) were dipped into the ball-milled slurry to apply a thin layer of the hexaaluminate powder uniformly onto the monoliths according to common washcoating procedures. Samples of pure lanthanum hexaaluminate ($LaAl_{11}O_{18}$), calcined to 1200°C, were impregnated with palladium, by addition of dissolved Pd-nitrate to the ball-milled slurry of hexaaluminate before the washcoating step. The washcoated monoliths were calcined to 1000°C in air and thereafter mounted in a bench-scale quartz reactor. In this reactor, outlined in Fig. 1, it is possible to mix, combust, and analyze a synthetic gas mixture that resembles real gasified biomass mixed with air. The gas hourly space velocity (GHSV) over the monolith segment was 50 000 $h^{-1}$ ($m^3$ gas/$m^3$ monolith volume).

**Characterization of the Catalysts.** For characterization of the thermal stability of the hexaaluminate powders, a series of calcinations to 1000, 1200, and 1400°C in dry air was made in a tubular furnace. The x-ray diffraction patterns (XRD) were recorded with a Siemens Diffraktometer 5000 scanning $2\theta$ from 15°–80° using monochromatized Cu-$K_\alpha$ radiation. Specific surface area was measured according to the Brunauer-Emmet-Teller (BET) method with $N_2$ adsorption at liquid $N_2$ temperature on a Micromeritics ASAP 2000/2010.

**Specifications of the Low-Heating Value Fuel Gas.** Composition of the gasified biomass depends on the method of gasification and composition of the raw biomass [9]. Therefore, the composition of the typical gasified biomass given in Table 1 is an attempt to summarize the composition of a gas with an LHV of 5–7 MJ/$Nm^3$. In addition to the compounds covered in Table I ($N_2$, $H_2$, CO, $CH_4$, $H_2O$, $C_2$-$C_{12}$ hydrocarbons and $NH_3$,) traces of HCN, sulfur, and alkali compounds could be present in the gas.

In the laboratory testing, not all fuel impurities and trace components could be included in the synthetic gas. However, the effects of some of these components are shortly addressed here. The amount of trace components and fuel impurities depend on the gasification process and could effect the catalyst thermal stability and activity. In this project, separate experiments have been done to simulate $C_2$-$C_{12}$ components with $C_2H_4$. These experiments have shown that ethene has a lower light-off temperature than methane, but higher than CO and $H_2$. Since higher hydrocarbons will be present in relatively low concentrations, they are not considered to cause any harmful effects with respect to the catalyst. HCN have been excluded since it is assumed to behave similarity to $NH_3$ and it is difficult to handle in the laboratory scale. Alkali components can cause a change of the solid-state of the catalyst by adsorption and reaction at the catalyst surface. Thereby, the thermal stability and catalytic activity will be improved or weakened. Preliminary results show that these effects are very small. It is also likely that the alkali concentrations at the inlet of the combustor after state-of-the-art gas cleaning will be lower than the 1 ppm range, and needs to be to meet gas turbine specifications. Separate experiments will be performed with $H_2S$, since sulfur compounds are usually considered reversible catalyst poisons at low temperatures. However, the high temperatures in gas turbine catalytic combustor are likely to weaken these effects over most catalysts. Other trace components were excluded in the synthetic gas, but the effects from these trace elements will be treated separately in pilot-scale testing during the project and covered in later publications.

**Table 1 Composition of $N_2$-diluted synthetic gasified biomass, mixed with air, based on a typical biomass gas composition that resembles the gas composition after air gasification of wood**

| Gas Compound | Typical Biomass | $N_2$-diluted Synthetic Biomass | |
|---|---|---|---|
| | Raw | Raw | with air |
| | [%] | [%] | [%] |
| $N_2$ | 44.5 | 81.8 | 84.1 |
| $O_2$ | 0.0 | 0.0 | 9.6 |
| $H_2$ | 10.2 | 3.4 | 1.2 |
| CO | 14.7 | 4.5 | 1.6 |
| $CO_2$ | 13.8 | 4.1 | 1.4 |
| $CH_4$ | 4.6 | 1.4 | 0.5 |
| $C_2$-$C_{12}$ | 1.0 | 0.0 | 0.0 |
| $H_2O$ | 11.2 | 4.7 | 1.6 |
| $NH_3$ | 0.165 | 0.10 | 0.034 |

As shown in Table 1 the synthetic gasified biomass is mixed with air and diluted with excess $N_2$ to decrease the heat evolution from the exothermic combustion processes during the tests. Hereby a more uniform temperature profile over the catalyst could be used in the laboratory experiments. An equivalence ratio of 1/4 was used throughout the whole study.

## High-Temperature Catalyst Development

A series of La, Ba, Mn, and Fe-catalysts were prepared to study the effect of substitution on the thermal stability of the hexaaluminates. $LaMnAl_{11}O_{19}$ (LMA) was chosen as the initial composition, since this catalyst has been previously reported to be very promising for catalytic combustion, considering both catalytic activity and thermal stability [13,16].

The effects of substituting half of the La with Ba and half of the Mn with Fe, compared to pure and Pd-impregnated $LaAl_{11}O_{18}$ are given in Table 2, together with the target compositions. Initially, at 1000°C, all samples have similar surface areas (68–95 $m^2$/g).

After calcination at 1200°C, the LMA-sample is most promising, with a surface area of 75 $m^2$/g. The Pd-based catalyst was only characterized after calcination at 1200°C, wet ball-milling and the washcoating step. Here, the surface areas were comparable to nonimpregnated washcoats (34 $m^2$/g). Further calcinations of the substituted hexaaluminates up to 1400°C causes an appreciable decrease in the surface area down to 8 $m^2$/g for all substitutions except the $Ba_{0.5}La_{0.5}Fe_{0.5}Mn_{0.5}Al_{11}O_{19}$-sample and the pure $LaAl_{11}O_{18}$-sample. The latter two both retained a surface area around 16–17 $m^2$/g. These results suggest that highly substituted hexaaluminates would show higher sintering-resistance than less substituted hexaaluminates. Also, the pure $LaAl_{11}O_{18}$ has a high thermal stability, but has very low catalytic activity. A series of activity tests over the Ba, Mn, and Fe, substituted lanthanum hexaaluminates in the bench-scale unit used for the combustion tests described below, showed that their activities for catalytic combustion of gasified biomass were quite similar.

**Table 2 BET specific surface area after consecutive calcination in flowing air at 1000 °C (4h), 1200 °C (4h), and 1400 °C (4h), respectively**

| Composition | 1000 [°C] | 1200 [°C] | 1400 [°C] |
|---|---|---|---|
| $LaMnAl_{11}O_{19}$ | 96 | 75 | 8 |
| $LaFeAl_{11}O_{19}$ | 73 | 26 | 8 |
| $Ba_{0.5}La_{0.5}MnAl_{11}O_{19}$ | 83 | 26 | 8 |
| $Ba_{0.5}La_{0.5}Fe_{0.5}Mn_{0.5}Al_{11}O_{19}$ | 82 | 54 | 16 |
| $LaAl_{11}O_{18}$ | 68 | 37 | 17 |
| $LaAl_{11}O_{18}$* | - | 34 | - |
| 0.8 wt % $PdO/LaAl_{11}O_{18}$ *# | - | 34 | - |

* after drying of ball-milled slurry
\# after Pd-impregnation and calcination to 1000 °C

The formation of the pure hexaaluminate crystal phase was confirmed by XRD analysis. All samples showed pure hexaaluminate phases and no other single metal oxide or mixed metal oxide phases, such as perovskites or spinels, could be detected.

## Bench-Scale Laboratory Testing

A study on catalytic combustion of synthetic gasified biomass mixed with air was made in a small bench-scale test unit, depicted in Fig. 1. The results reported here are over Mn-substituted lanthanum hexaaluminate ($LaMnAl_{11}O_{19}$) washcoated cordierite monoliths and lanthanum hexaaluminate-supported palladium oxide ($PdO/LaAl_{11}O_{18}$) washcoated cordierite monoliths, referred to as **LMA** and **PLA**, respectively.

Figure 2(*a*) shows conversion of CO, $H_2$, and $CH_4$ over LMA washcoats. The sequence of ignition is CO, $H_2$, and $CH_4$, reflecting the higher catalytic activity of CO and $H_2$ combustion compared to $CH_4$ combustion. The results shown here in a monolith reactor correspond to previous results shown in single channel annular microreactor [13]. A temperature above 500, 550, and 650°C is needed to convert more than 50 percent of the inlet concentration of CO, $H_2$, and $CH_4$, respectively. At temperatures above 800°C all CO, $H_2$, and $CH_4$ are completely converted to the combustion products $CO_2$ and $H_2O$.

The conversion of $NH_3$ and the yield of NO and $NO_2$ is shown in Fig. 2(*b*). Conversion and yield is defined as follows:

$$\text{conversion} = 1\text{-(outlet concentration/inlet concentration)}$$
$$\text{yield} = \text{outlet concentration/inlet concentration}$$

Here, the yield of NO and $NO_2$ is based on the inlet concentration of $NH_3$ and the conversion of the fuel component is based on the inlet concentration of each fuel component, respectively. The results show that $NH_3$ is converted at relatively low temperatures, actually lower than CO, to the main products $N_2$, NO, $NO_2$, and $N_2O$. In the current unit, $N_2$, NO, and $NO_2$ can be detected, but not $N_2O$. However, formation of $N_2O$ is assumed relatively low and this specie is not stable at higher combustion temperatures [17].

Figure 2(*b*) shows that NO and $NO_2$ are the major compounds formed at inlet temperatures between 500 and 800°C. The two NO peaks at 500 and 700°C occur at ignition of CO-$H_2$ and $CH_4$, respectively. This suggests that the catalyst promote formation of NO at the light-off temperatures of CO-$H_2$ and $CH_4$. At higher temperatures, above 700°C, the formation of NO and $NO_2$ decreases down to a minimum yield of ca. 40 percent. This means that circa 60 percent of the fuel-bound nitrogen ($NH_3$) is converted to $N_2$.

Results from the catalytic combustion of the synthetic gas mixture over the Pd-based catalyst PLA are shown in Fig. 3(*a*). The ignition of the fuel components and the formation of nitrogen oxides is different from the results shown for the LMA-catalyst.

Ignition of CO and $H_2$ is instantaneous at 230°C and is followed by a smooth ignition of $CH_4$, until the PdO is reduced from an oxidized highly active state to a less oxidized and less active state, at temperatures around 550–700°C. This transformation is well-known from literature and discussed in more detail for methane combustion elsewhere [18]. The results shown here confirm

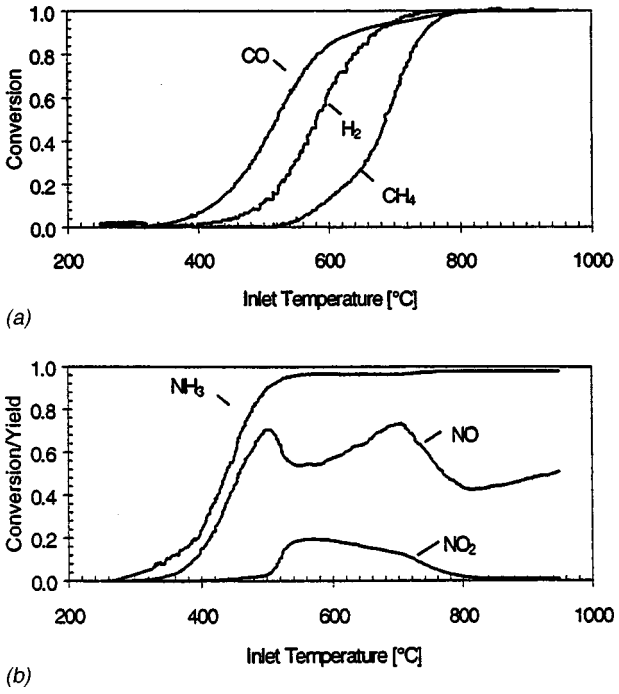



**Fig. 2 (*a*) Conversion of $H_2$, CO, and $CH_4$ over $LaMnAl_{11}O_{19}$; (*b*) conversion of $NH_3$, yield of NO and $NO_2$ over $LaMnAl_{11}O_{19}$**

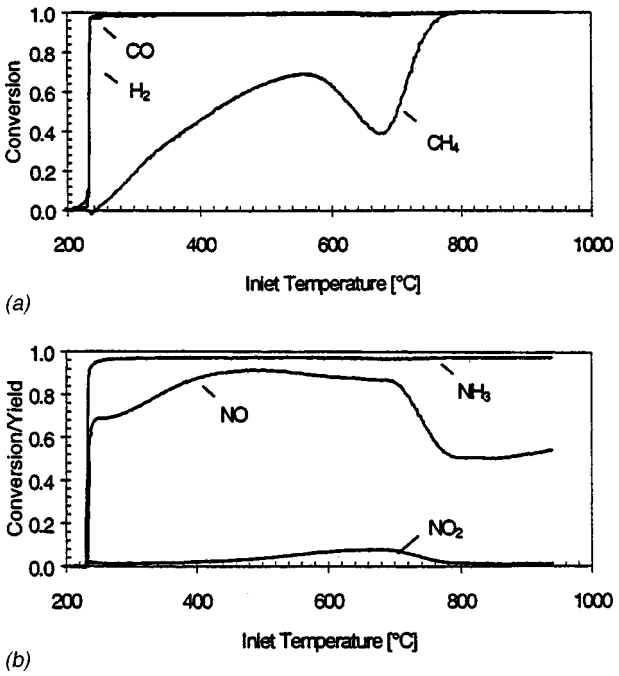



**Fig. 3 (*a*) Conversion of $H_2$, CO, and $CH_4$ over $PdO/LaAl_{11}O_{18}$; (*b*) conversion of $NH_3$, yield of NO and $NO_2$ over $PdO/LaAl_{11}O_{18}$**

earlier result for methane, i.e., a strong decrease in methane combustion activity as the highly active palladium oxide is converted to a less active palladium species with less oxygen.

However, no decrease in activity for $H_2$ and CO combustion is seen, suggesting that both palladium species are active enough for 100 percent conversion of $H_2$ and CO at temperatures above 600°C. $NH_3$ is ignited instantaneously together with CO and $H_2$ over the Pd-catalyst. The formation of $N_2$ is typically low and reaches less than 10 percent yield, but increases appreciably above 700°C up to a maximum ca. 50 percent yield, cf. Figure 3(*b*). The strong increase occurs at the same temperature that PdO is transformed to a less oxidized state. These results indicates that this less active Pd-species would have a stronger selectivity towards formation of molecular nitrogen than the more active PdO-species. The formation of $NO_2$ is less than 10 percent and decreases to insignificant levels compared to NO above 800°C.

## Pilot-Scale Testing

The catalysts developed in this project are also going to be tested in two pilot-scale test units. The tests in the first unit are accomplished in cooperation with TPS AB in an atmospheric combustion chamber, 78 mm in diameter, connected to a small-scale gasifier. The second unit is situated at The Division of Heat and Power Technology at KTH (Royal Institute of Technology) with a pressurized (4 bar) combustion chamber, 150 mm in diameter. Here, a synthetic gasified biomass fuel is produced by catalytic reforming of a mixture of air, LPG and $CO_2$ and addition of $NH_3$. The first results from these pilot-scale units will be reported during the spring 1998. In both these pilot-scale tests higher space velocities will be applied. Approximately, an industrial gas turbine catalytic combustor will run at space velocities around 200,000 $h^{-1}$, i.e., three to four times higher than in the laboratory scale test reported here, corrected for inlet pressure. This increase in space velocity is likely to increase the ignition temperature of the catalyst. Higher pressures will be applied in the second unit, but the effect of a higher pressure is more difficult to foresee, especially considering the synergistic effects from the LHV-fuel mixture. However, no large pressure effects are expected for the catalytic reactions. In the TPS reactor, a real gasifier gas will be used, hence effects from fuel impurities will be considered in these tests.

## Concluding Remarks

It has been shown possible to prepare single-phase hexaaluminate powders that retained surface areas up to 17 $m^2/g$ after calcination to 1400°C. The best results were achieved for nonsubstituted lanthanum hexaaluminate and for a Ba, Mn, and Fe substituted lanthanum hexaaluminates.

Catalytic combustion of a synthetic gasified biomass in small bench-scale tests with monolithic samples showed that it was possible to ignite CO and $H_2$ at 230°C over hexaaluminate-supported Pd catalyst (PLA) and at 400 to 600°C over a Mn-substituted lanthanum hexaaluminate catalyst (LMA). The LMA-catalyst showed a higher yield of $N_2$ from fuel-$NH_3$ compared to PLA, 60 percent $N_2$ compared to 50 percent $N_2$, at temperatures around 800°C. The overall objective in this project is to increase the conversion of fuel-N to $N_2$ to at least 90 percent, by utilizing a catalytic combustion chamber. The results on fuel-$NO_X$ formation shown here for a Pd-based catalyst are somewhat lower than could be expected for a noble metal catalyst compared to transition metal oxide-based catalysts [1], since noble metal-based catalysts typically have a high formation of fuel-$NO_X$. However, the tests described here have been made at elevated temperatures, up 950°C, with transformation of the active Pd-species and with a fairly complex gas mixture, indicating that synergistic reactions at high temperatures might influence the $NO_X$-formation. However, the transition metal substituted hexaaluminates still showed higher conversions of fuel-$NH_3$ towards $N_2$, as expected. It was also shown that the catalyst composition influenced the conversion of the CO and $H_2$-rich low heating value fuel gas largely. Hence, by varying of the catalyst composition it should be possible to influence the conversion of different fuel-components at different stages and thereby control the temperature evolution in the combustion chamber. This could be used to tailor the design of segmented catalytic combustion chamber to obtain lower emissions of $NO_X$, and stabilize the combustion of the dilute LHV-gas.

The reaction route from $NH_3$ to molecular $N_2$ or $NO_X$ is not fully understood, but it is clear that it involves several steps. One possible route from $NH_3$ to $N_2$ can be the formation of $NO_X$ that is consecutively reduced to $N_2$ by remaining fuel components. These reactions are usually very temperature sensitive, both in heterogeneous [8] and homogeneous combustion [9], which implies that it will be important to control the temperature in the different segments in the catalytic combustor.

## Acknowledgment

Funding from the European Union, contract JOR3-CT96-0071, and NUTEK, the Swedish National Board for Industrial and Technical Development, is gratefully acknowledged. The authors wish to recognize the dedication from the other partners in the ULECAT project, i.e., IFP Institute Francçais du Pérole, Turbomeca SA, Politecnico di Milano, Volvo Aero Corporation, TPS AB, Vattenfall AB, Sydkraft AB, and Perstorp AB.

## References

[1] Zwinkels, M. F. M., Järås, S. G., Menon, P. G., and Griffin, T. A., 1993, ''Catalytic Materials for High-Temperature Combustion,'' Catal. Rev. Sci. Eng., **35**, pp. 319–358.
[2] Dalla Betta, R. A., Scalatter, J. C., Nickolas, S. G., Razdan, M. K., and Smith, D. A., 1995, ''Application of Catalytic Combustion Technology to Industrial Gas Turbines for Ultra-Low $NO_X$ Emissions,'' ASME Paper 95-GT-65.
[3] Sadamori, H., Nishida, T., Yamashita, T., Furuya, A., and Matsuhisa, T., 1995, ''The Development of a High-Temperature Catalytic Combustion System,'' Proceedings, Yokohama Int. Gas Turbine Cong., pp. 247–250.
[4] Ozawa, Y., Fujii, T., Sato, M., Kanazawa, T., and Inou, H., 1996, ''Development of a Catalytically Assisted Combustor for a Gas Turbine,'' *Book of Abstracts*, 3rd Int. Workshop on Catalytic Combustion, Amsterdam.
[5] Tucci, E. R., 1982, ''Use Catalytic Combustion for LHV Gases,'' Hydrocarbon Process., May, pp. 159–166.
[6] Sung, C. J. M., Kennedy, L. A., and Ruckenstein, E., 1984, ''The Effect of Nitrogen Content on the Oxidation of Fuel Bound Nitrogen in a Transition Metal Oxide Catalytic Combustor,'' Combust. Sci. Technol., **41**, pp. 315–325.
[7] Johansson, E. M., and Järås, S. G., 1996, ''Circumventing Fuel-$NO_X$ in Catalytic Combustion of Gasified Biomass,'' *Book of Abstracts*, 3rd Int. Workshop on Catalytic Combustion, Amsterdam.
[8] Johansson, E. M., and Järås, S. G., 1996, ''Circumventing Fuel-$NO_X$ in Catalytic Combustion of Gasified Biomass,'' submitted to Catal. Today.
[9] Bridgwater, A. V., 1995, ''The Technical and Economic Feasibility of Biomass Gasification for Power Generation,'' Fuel, **74**, pp. 631–653.
[10] Machida, M., Eguchi, K., and Arai, H., 1989, ''Catalytic Properties of $BaMAl_{11}O_{19\text{-}alfa}$ (M=Cr, Mn, Fe, Co, and Ni) for High-Temperature Catalytic Combustion,'' J. Catal., **120**, pp. 377–386.
[11] Lowe, D. M., Gusman, M. I., and McCarty, J. G., 1994, ''Synthesis and Characterization of Sintering Resistant Aerogel Complex Oxide Powders,'' in *Preparation of Catalysts VI*, G. Poncelet, et al., eds., Elsevier Science, Amsterdam, pp. 445–452.
[12] Groppi, G., Belloli, A., Tronconi, E., and Forzatti, P., 1996, ''Catalytic Combustion of CO-$H_2$ on Manganese-Substituted Hexaaluminates,'' Catal. Today, **29**, pp. 403–407.
[13] Johansson, E. M., and Järås, S. G., 1997, ''Catalytic Combustion of Low-Heating Value Fuel-Gases: Ignition of CO, $H_2$, and $CH_4$,'' ACS Symp., Div. Petrol Chem., San Francisco, **42**, pp. 146–150.
[14] Le Gal, J.-H., Martin, G., and Durand, D., 1998, ''Development of a Dual Fuel Catalytic Combustor for a 2.3 MWe Gas Turbine,'' submitted to ASME Turbo Expo Land, Sea and Air 98.
[15] Ersson, A. G., Johansson, E. M., and Järås, S. G., 1998, ''Techniques for Preparation of Manganese-Substituted Lanthanum Hexaaluminates,'' in *Studies in Surface Science and Catalysis, Preparation of Catalysts VII*, G. Poncelet, eds., Elsevier Science, Amsterdam.
[16] Eguchi, K., Takahara, H., Inoue, H., Sekizawa, K., and Arai, H., 1997, ''Catalytic Properties of Transition Metal Substituted Hexaaluminate for High Temperature Combustion,'' ACS Symp., Div. Petrol Chem., San Francisco, 1997 **42**, pp. 180–182.
[17] Miller, J. A., and Bowman, C. T., 1989, ''Mechanism and Modeling of Nitrogen Chemistry in Combustion,'' Prog. Energy Combust. Sci., **15**, pp. 287–338.
[18] McCarty, J. G., 1995, ''Kinetics of PdO Combustion Catalysis,'' Catal. Today, **26**, pp. 283–293.

# Measurement of Transfer Matrices and Source Terms of Premixed Flames

**C. O. Paschereit**
e-mail: oliver.paschereit@power.alstom.com

**B. Schuermans**
e-mail: bschuermans@power.alstom.com

**W. Polifke**
e-mail: wolfgang.polifke@td.mw.tum.com

**O. Mattson**
e-mail: oscar.fungerar@hotmail.com

ALSTOM (Switzerland) Ltd.,
5405 Baden, Switzerland

*An experimental method to determine the thermoacoustic properties of a gas turbine combustor using a lean-premixed low emission swirl stabilized burner is presented. To model thermoacoustic oscillations, a combustion system can be described as a network of acoustic elements, representing for example fuel and air supply, burner and flame, combustor, cooling channels, suitable terminations, etc. For most of these elements, simple analytical models provide an adequate description of their thermoacoustic properties. However, the complex response of burner and flame (involving a three-dimensional flow field, recirculation zones, flow instabilities, and heat release) to acoustic perturbations has—at least in a first step—to be determined by experiment. In our approach, we describe the burner as an active acoustical two-port, where the state variables pressure and velocity at the inlet and the outlet of the two port are coupled via a four element transfer matrix. This approach is similar to the "black box" theory in communication engineering. To determine all four transfer matrix coefficients, two test states, which are independent in the state vectors, have to be created. This is achieved by using acoustic excitation by loudspeakers upstream and downstream of the burner, respectively. In addition, the burner might act as an acoustic source, emitting acoustic waves due to an unsteady combustion process. The source characteristics were determined by using a third test state, which again must be independent from the two other state vectors. In application to a full size gas turbine burner, the method's accuracy was tested in a first step without combustion and the results were compared to an analytical model for the burner's acoustic properties. Then the method was used to determine the burner transfer matrix with combustion. An experimental swirl stabilized premixed gas turbine burner was used for this purpose. The treatment of burners as acoustic two-ports with feedback including a source term and the experimental determination of the burner transfer matrix is novel.*
[DOI: 10.1115/1.1383255]

## Introduction

Thermoacoustic combustion instabilities are a cause for concern in many combustion applications such as household burners, gas turbines, and rocket engines. Combustion instabilities cannot only increase emissions of noise or pollutants such as unburnt hydrocarbons or oxides of nitrogen, but also lead to very high-pressure pulsations, resulting in structural damage of the combustor. If they occur, combustion instabilities may interfere with the operation of the combustion system or limit at least its operational range. Pressure amplitudes are, i.e., in the order of 1 percent of the mean operating pressure but can also be as large as 10 percent of the mean operating pressure as observed in solid propellant rocket motors ([1]). As the oscillation amplitudes scale linearly with the mean pressure and operating pressure ranges up to 35 MPa, the resulting pressure oscillation can reach considerable amplitudes, on the order of atmospheric pressure. For stationary gas turbines the drive for lower emissions of oxides of nitrogen has lead during the last decade to the wide spread of lean premix burners and convectively cooled combustion chambers. These technological changes have resulted in a reduced stability of flame anchoring and lower acoustic damping. Consequently, modern gas turbines are more susceptible to combustion driven oscillations and the importance of thermoacoustic phenomena in gas turbine combustors has increased sharply.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Indianapolis, IN, June 7–10, 1999; ASME Paper 99-GT-133. Manuscript received by IGTI, Oct. 1998; final revision received by the ASME Headquarters, Mar. 1999. Associate Editor: D. Wisler.

Thermoacoustic instabilities involve a feedback cycle comprising fluctuations in acoustic pressure, velocity, heat release rate, and the acoustic characteristics of the combustor and air supply, which act as resonators. Lord Rayleigh [2] deduced a fundamental stability criterion for thermoacoustic oscillations: nonsteady heat release can enhance acoustic oscillations if the heat release rate peaks at the moment of greatest compression. Although this criterion provides merely a necessary but not a sufficient condition for instability, as it does not take into account the stabilizing influence of dissipation of acoustic energy, it expresses essential physical insight into the problem of thermoacoustic instability. The Rayleigh criterion also illustrates that the stability analysis of a combustion system requires detailed knowledge of the phase relationship between pressure oscillations and the various processes controlling the rate of heat release.

Thermoacoustic systems can be represented conveniently as a network of acoustic elements which correspond to various components of the system, e.g., fuel and air supply, burner and flame, combustor, cooling channels, suitable terminations, etc. ([3,4]). This approach has often been used in the acoustic analysis of ducts and mufflers ([5]). The elements in such a network can be described as acoustic multiports. For many of these elements, simple analytical models, based on linearization of the governing equations, provide an adequate description of their thermoacoustic properties. However, the complex response of burner and flame to acoustic perturbations has—at least in a first step—to be determined by experiment. In our approach the "black box" theory used in communications engineering has been extended to treat problems of combustion generated sound. This method has been successfully applied to describe fluid machines as acoustical multi ports (e.g., ([6–10]). We describe the burner and the flame as an

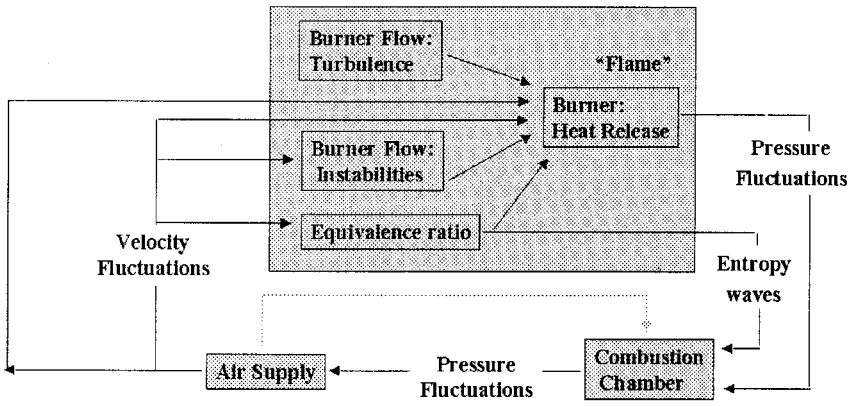



**Fig. 1 Noise generation in combustors**

active acoustical two-port, where the state variables pressure and velocity at the inlet and the outlet are coupled via a four element transfer matrix ([11]).

The interaction between large scale coherent structures which are related to flow instabilities ([12]), the heat release and the combustor acoustics was shown to be a driving mechanism leading to unstable combustion ([13–18]). Theoretical flame models have been derived to describe the relationship of unsteady heat release, acoustics and flow perturbations, e.g., in afterburners ([19]) flat flames in open tubes ([20]) and in simple experimental premix burners ([21]). However, the highly three-dimensional flow field of a swirl stabilized burner and the interaction between heat release and the flow field still requires experimental determination.

The suggested approach of experimental determination of the burner transfer function in this paper can be used to shed light into flame/acoustic interaction and will allow the development of an empirical flame model for swirl stabilized premixed combustion.

## Noise Generation in Combustors

An overview illustrating some basic mechanisms of thermoacoustic instability is given in Fig. 1. The definition of a simple combustion system consisting of a hood, the swirl stabilized burner, and the combustion chamber is given in Fig. 2. Burner flow turbulence and flow instabilities (which are manifested in the shedding of vortices influence the momentary rate of fuel consumption at the flame front. As a consequence, fluctuations in heat release occur that result in fluctuations in pressure and velocity. The downstream traveling waves will be reflected by the acoustic boundary conditions of the combustor exit and lead to pressure and velocity oscillations at the burner exit. This feedback cycle will again cause fluctuations in heat release. Furthermore, the fluctuations in pressure and velocity may result in inhomogeneities in the equivalence ratio, which are convected by the mean flow towards the flame. As these inhomogeneities are consumed by the flame, they will influence the momentary heat release rate, and therefore also the rate of volume production, which couples via the impedance at the burner exit with the system acoustics. In addition to this direct interaction, there is a second mechanism that couples equivalence ratio fluctuations with the system acoustics: the adiabatic flame temperature of richer (leaner) "pockets" of mixture is higher (lower) than average. Thus, equivalence ratio fluctuations will lead to fluctuations in the hot gas temperature downstream of the flame, i.e., so-called entropy waves. The entropy waves are convected downstream by the mean flow and couple at the choked combustor exit with the system acoustics. Entropy waves are considered in our simulation of the combustor but will not be discussed in this paper.

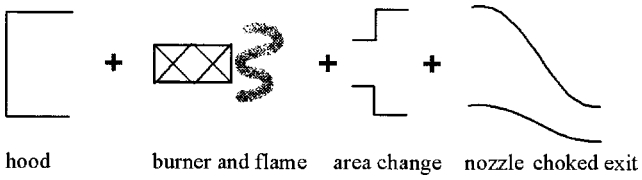



**Fig. 2 Example of modeling the combustor as a network of acoustic elements**

## Modeling of Thermoacoustic Systems

**The Network Representation of Thermoacoustic Systems.** A thermoacoustic system can be represented as a network of acoustic elements, which correspond to various components of the system, e.g., ducts, area changes, nozzles and diffusers, and the burner with flame. The network representation of a generic gas turbine combustor is shown in Fig. 2.

Under the assumption of harmonic time disturbances $\exp(i\omega t)$, the unknowns of the system are the amplitudes of acoustic variables at the end points of these elements. The equations of linear acoustics are then used to generate coupling relations—i.e., fluctuations of velocity $v$ and pressure $p$—across every element. As a simple example, consider a duct of length $L$. The propagation of acoustic waves traveling in up and downstream direction yields the relation

$$\begin{pmatrix} p_d \\ v_d \end{pmatrix} = \begin{pmatrix} \cos(kL) & iZ_0 \sin(kL) \\ \dfrac{1}{Z_0} i \sin(kL) & \cos(kL) \end{pmatrix} \begin{pmatrix} p_u \\ v_u \end{pmatrix}, \qquad (1)$$

between pressure $p$ and velocity $v$ at the up and downstream end (subscripted $u$ and $d$, respectively). Here, $Z_0 = \rho c$ is the characteristic impedance and $k$ is the wave number, equal to the ratio of angular frequency $\omega$ and sound speed $c$ in the case of plane waves without mean flow and dissipative effects. The matrix in Eq. (1) is called the transfer matrix $T$ of the element.

Each of the elements in Fig. 2 relates acoustic pressure and velocity (or incoming and outgoing Riemann in variants, which represent the up and downstream traveling waves) on both sides in a simple algebraic way as shown in Fig. 3. Physically it does not matter if an element relates the Riemann-invariants or the acoustic pressure and velocity, since they can be related to each other using Eqs. (2) and (3)

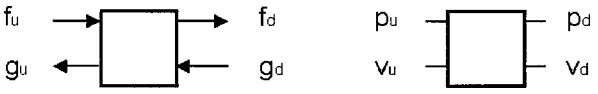



**Fig. 3 Acoustic elements relating Riemann-invariants or pressure and velocity fluctuatins of both sides of the element**

$$\frac{p(\omega)}{\rho c} = f + g, \tag{2}$$

$$v(\omega) = f - g. \tag{3}$$

It is merely a matter of convention as to which of these representations is used.

The representation of the transfer matrix of a simple tube in terms of Riemann invariants is

$$\begin{pmatrix} f_d \\ g_d \end{pmatrix} = \begin{pmatrix} e^{-ikL} & 0 \\ 0 & e^{ikL} \end{pmatrix} \begin{pmatrix} f_u \\ g_u \end{pmatrix}, \tag{4}$$

Equivalently, the Riemann invariants can be related by a scattering matrix

$$\begin{pmatrix} f_d \\ g_u \end{pmatrix} = \begin{pmatrix} e^{-ikL} & 0 \\ 0 & e^{-ikL} \end{pmatrix} \begin{pmatrix} f_u \\ g_d \end{pmatrix}, \tag{5}$$

which relates the incoming signal $(f_u, g_d)$ to the outgoing signal $(f_d, g_u)$.

The thermoacoustic description of a complete combustion system of equations $\mathbf{S} \cdot \mathbf{u} = \mathbf{d}$ is obtained by combining the coefficients of all the transfer matrices of all elements into one System matrix **S** and all unknowns in a vector **u**. The vector **d** represents the driving forces in the system.

**The Flame Transfer Function.** Noise generation problems in enclosures like duct systems are of special interest because of the interaction (feedback) of the wave field in the enclosure with the source. This becomes especially true in thermoacoustic problems, where a self-excitation mechanism resulting from the interaction of the sound field in the enclosure and the burner as a sound source may lead to high pulsation amplitudes. To model noise generation problems, especially those involving a feedback cycle as in thermoacoustics, it is essential to characterize the noise sources. Their strength and interaction with the acoustic wave field have to be known.

Many noise sources can be calculated from first principles. However, the highly three-dimensional flow field of a swirl stabilized burner and the interaction between heat release and the flow field still requires experimental determination. One possibility which is especially useful in the low-frequency regime is to describe a burner as an active acoustic two-port. In this approach the burner is treated as a ''black box.'' Questions about the internal character of the system can then be put aside. However, the data obtained for the burner give valuable information for the formulation of premix flame models.

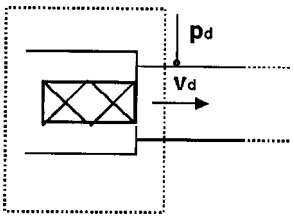



**Fig. 4 A burner modeled as an acoustic one-port**

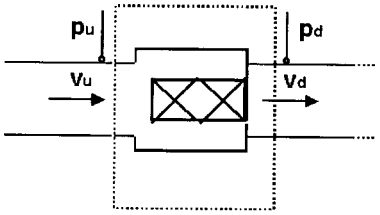



**Fig. 5 A burner modeled as an acoustic two-port**

In the acoustic one-port model the burner and the upstream conditions are treated as a black box (Fig. 4). For a measured source impedance $Z_s$ and source strength $v_d$ the upstream conditions are not allowed to change because they will influence these quantities. For practical applications this implies that the upstream acoustic boundary conditions being dictated by the plenum in which the burner is mounted as well as everything upstream of the plenum has to remain unchanged (as long as they are acoustically coupled). Only then, can the measured results be used to model the combustor.

The shortcoming of the one-port treatment of the burner can be overcome by describing *only* the burner as a black box allowing ''two-ports,'' designated as input and output (Fig. 5). Even more ports are allowed, if, e.g., fluctuations in the fuel supply rate have a reasonable influence on the acoustic behavior of the burner. In the frequency domain the two-port can be described by the equation

$$\begin{pmatrix} p_d \\ v_d \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} p_u \\ v_u \end{pmatrix}. \tag{6}$$

Here the subscripts $u$ and $d$ indicate upstream and downstream pressures $p$ and velocities $v$, respectively.

In the transfer matrix above the noise generation of the flame has not been considered. In order to include a source term the representation in terms of Riemann invarients is more convenient. The principle is shown in Fig. 6. In addition to merely coupling the Riemann invariants across the burner (and flame), the burner itself may act as a noise source generating waves $f_s$ and $g_s$. These waves which are emitted by the burner are uncorrelated with the soundfield in the combustor, i.e., they do not depend on the incoming waves $f_u$ and $g_d$ and are thus not included in the transfer matrix $T$. This yields the following relation;

$$\begin{pmatrix} f_d \\ g_u \end{pmatrix} = \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} f_u \\ g_d \end{pmatrix} + \begin{pmatrix} f_s \\ g_s \end{pmatrix}. \tag{7}$$

The matrix in the above representation is the scattering matrix $S$ as described by Eq. (5) in the previous section. Equation (7) is the complete description of an active two-port with a source term. For each combination of incoming signals $(f_u, g_d)$ it provides the outgoing signal $(f_d, g_u)$. Note that for $f_u = g_d = 0$ the outgoing signal is only the source $(f_s, g_s)$. The problem here is how to determine the six unknowns (four elements of the scattering matrix and two source terms) in Eq. 7 by simply measuring acoustic pressures? A detailed answer on how to separate the different terms will be given in the next section; however, since the system consists of only two equations, three independent test states are required to solve for the unknowns. Different methods to create these different test states are described in Paschereit and Polifke ([22]). Here, the test states were created by forcing upstream of the burner, then downstream and successively at both sides of the burner at the same time.

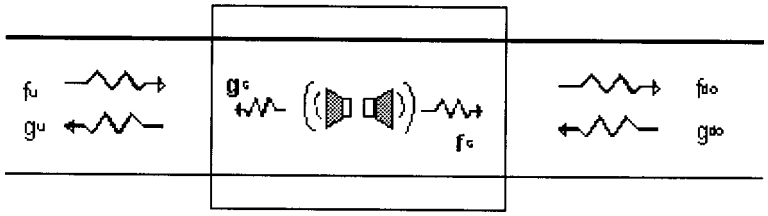



**Fig. 6 Representation of the flame as a sound source**

## Experimental Methods

**Experimental Set-Up.** The cmbustion facility is shown in Fig. 7. The atmospheric test rig consists of a plenum chamber upstream of the swirl-inducing burner and a combustion chamber downstream of the burner. The plenum chamber contains perforated plates to reduce the turbulence level of the flow. The circular combustion chamber consists of an air-cooled double-wall quartz glass to provide full visual access to the flame. The exhaust system is an air cooled tube with the same cross section as the combustion chamber to avoid acoustic reflections at area discontinuities. The acoustic boundary conditions of the exhaust system could be adjusted from almost anechoic (reflection coefficient $|r|<0.15$) to open-end reflection. An experimental swirl stabilized burner was used in the experiments which were performed using two operational modes—premixed combustion and partially premixed combustion (quasi-diffusion flame). In the premixed mode, natural gas was injected upstream of the swirling air to premix the fuel with the air. The flame was stabilized in a recirculation region near the burner outlet. During nonpremixed combustion a pilot flame was utilized and the fuel was injected into the recirculation region resulting in a quasi diffusion flame. Controlled excitation of the burner flow was accomplished by a circumferential array of four loudspeakers equally spaced in polar angle. One set of loudspeakers was placed at a distance of $x/D=4.2$ upstream of the dump plane and the second set at $x/D=9.6$ downstream of the dump plane. The nominal power of the 330-mm speakers was 120 W. Only axisymmetric excitation with the upstream loudspeakers was used and the loudspeakers were operated at zero-phase difference.

Pressure fluctuations were measured using water-cooled 1/4 in. condenser microphones.

*Measurement of Acoustic Quantities in Ducts With Mean Flow.* To determine the thermoacoustic characteristics of the burner/flame the up and downstream propagating waves (Riemann invariants) have to be determined. At least two microphones are required for this type of measurement. Using two microphones at two different axial positions $x_1$ and $x_2$ in a tube the Riemann invariants $f$ and $g$ at cross section $x_1$ are given by

$$\begin{pmatrix} f(x_1) \\ g(x_1) \end{pmatrix} = \frac{1}{\Phi^- - \Phi^+} \begin{pmatrix} \Phi^- & -1 \\ -\Phi^+ & 1 \end{pmatrix} \begin{pmatrix} p(x_1) \\ p(x_2) \end{pmatrix}. \tag{8}$$

where $\Phi^{\pm}=\exp(-ik_{\pm}(x_2-x_1))$ and $k_{\pm}=\omega/c/M\pm 1$. Once the Riemann invarients are known at location $x_1$ they can be stepped into any location $x_i$, using the relations

$$f(x_i)=f(x_1)e^{-ik-(x_1-x_1)},$$
$$g(x_i)=g(x_1)e^{-ik-(x_1-x_1)}. \tag{9}$$

In our approach an extension of the two-microphone method, the multi-microphone method is used to improve the accuracy of the measured data. Using two pressure signals it is possible to measure exactly the two quantities, i.e., the incident and the reflected wave components. By making more pressure measurements the number of equations is larger than the number of unknowns, thus the problem is overdetermined. The calculated incident $f(x_i)$ and reflected $g(x_i)$ wave components are fitted to the measured quantities $p(x_i)$ by using the nonlinear Levenberg-Marquardt method to minimize the $\chi^2$ quantity in the frequency domain

$$\chi^2=\sum_{i=0}^{N-1}[(f(x_i)+g(x_i))-p(x_i)]^2, \tag{10}$$

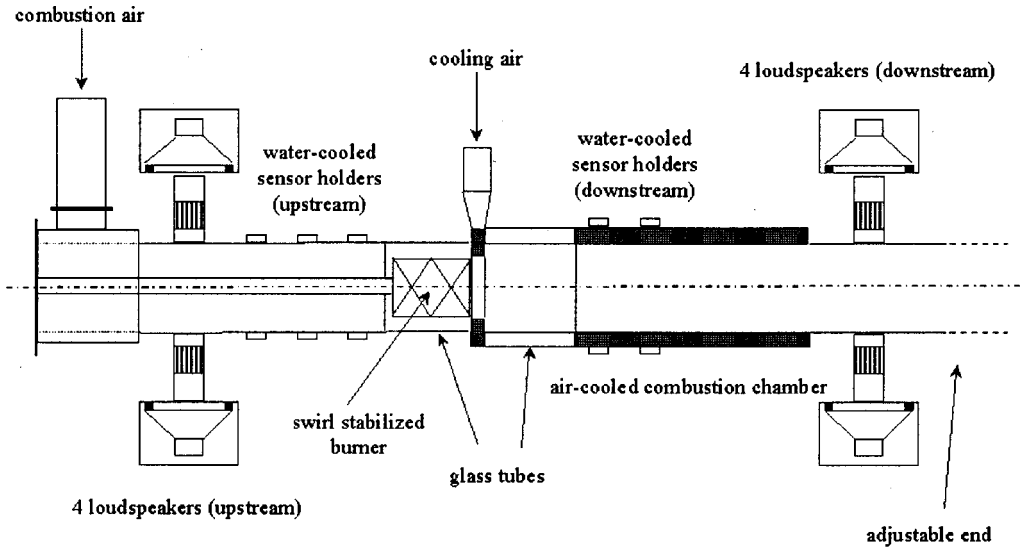



**Fig. 7 Experimental arrangement of the combustor**

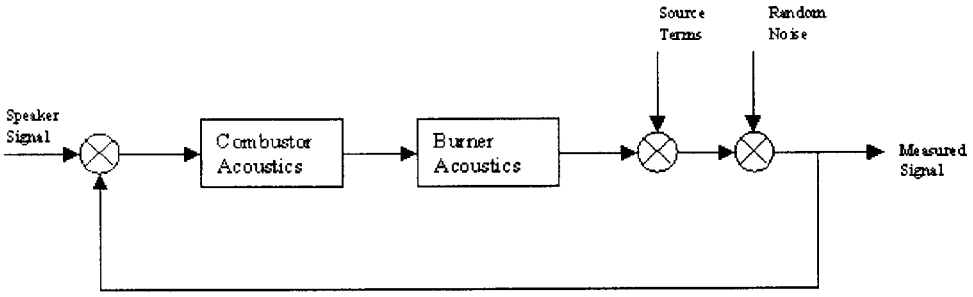



**Fig. 8 Different contributions to the measured signal**

where $N$ corresponds to the number of microphones used for the multi-microphone method.

*Determination of the Flame Transfer Function.* Assume that the response of the flame depends only on the acoustic field incident upon the burner the flame transfer function can be determined by using, e.g., the two source method. Here, acoustic forcing upstream and downstream of the flame provides two different acoustic test states to furnish a system of four linear independent equations with a unique solution for the coefficients of $T$:

$$\begin{pmatrix} p_u^{(1)} \\ u_u^{(1)} \\ p_u^{(2)} \\ u_u^{(2)} \end{pmatrix} = \begin{pmatrix} T_{11} & T_{12} & 0 & 0 \\ T_{21} & T_{22} & 0 & 0 \\ 0 & 0 & T_{11} & T_{12} \\ 0 & 0 & T_{21} & T_{22} \end{pmatrix} \begin{pmatrix} p_d^{(1)} \\ u_d^{(1)} \\ p_d^{(2)} \\ u_d^{(2)} \end{pmatrix}. \quad (11)$$

where the superscripts refer to the two test states 1 and 2.

The method described above is only applicable if the acoustic field incident upon the burner was only the result of the loudspeakers. However, the turbulent flow through the burner and flow instabilities traveling into the flame may generate sound waves, which are incoherent with the forcing signal, so that the burner itself acts as a sound source. Thus the method to measure the flame transfer function works only when the signal brought in by the loudspeakers exceeds the noise generated by flow and combustion. A method has been developed which deals with this problem. Here, the pressure signal picked up by the microphone is considered to be the sum of three different contributions (Fig. 8):

- Correlated signal: This contribution is correlated with the excitation signal. These are the signals needed to determine the transfer matrix. This signal contains the passive contribution of the burner/flame to an acoustic excitation. Compared to a simple acoustic element such as the duct described in the previous section, the burner/flame might also actively respond to the excitation. A feedback mechanism between, e.g., acoustic waves and flow instabilities (or equivalence ratio fluctuations) exists which might trigger a periodic combustion as the vortices are convected into the flame. Consequently the transfer matrix also contains an active part which is coupled to the sound field.
- Coherent uncorrelated signal: This is sound generated by the flame with a fixed frequency and phase. It is uncorrelated with the forcing signal but is coherent between multiple axially distributed microphones and not random. This contribution is referred to as the source term, and consists of sound generated by the flame or the flow field in a periodic—not random—way. This contribution is, e.g., generated by naturally occurring flow instabilities forming vortices which in turn influence the heat release of the flame. Equivalence ratio fluctuations leading to (periodic) unsteady heat release as discussed in Fig. 1 may be an additional mechanism leading to sound generation by the flame
- Random noise: This contribution is uncorrelated with the forcing signal and has no periodic component. It can be attributed to the turbulent flow and subsequent combustion noise. The length scale of this contribution is much smaller than the axial spacing of the microphones. As a consequence there is no coherence between multiple axial distributed microphones

In order to measure the transfer matrix of the flame and a possible source term a decomposition of the different contributing signals has to be performed. To determine the transfer function the only part of the signal needed is the one which is correlated with the forcing signal. All other contributions are part of the source term.

The uncorrelated coherent signal and the random noise, which is generated by turbulence and combustion, can be eliminated from the measured signal by a cross correlation between the measured signals and the forcing signal. This measurement gives the scattering matrix which contains no source term. The source term can then be determined by an additional measurement without any excitation and a subsequent substraction of measured signals and calculated signals using the ''source free'' scattering matrix which was described above.

Another method is to use three different acoustic test states to solve Eq. (7) for the six unknowns. This is the method which was applied here. Note that this method works as well if more than three test states are available. The equations can then be solved by applying the method of least squares. Once the scattering matrix is found, the source term can be determined by

$$\begin{pmatrix} f_s \\ g_s \end{pmatrix} = \begin{pmatrix} f_d \\ g_u \end{pmatrix} - \begin{pmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{pmatrix} \begin{pmatrix} f_u \\ g_d \end{pmatrix}. \quad (12)$$

From the source-free scattering matrix the transfer matrix can be determined by a simple matrix transformation.

## Results and Discussion

**The Transfer Matrix Without Combustion.** In a first step the transfer matrix was measured without combustion and without flow using a pure tone excitation. A model based on the nonsteady Bernoulli equation for compact elements has been developed ([22]) and led to the following theoretical coefficients of the transfer matrix

$$T = \begin{pmatrix} 1 & \rho_u c_u \left[ M_u \left( 1 - \zeta - \left( \frac{A_u}{A_d} \right)^2 \right) - i \frac{\omega}{c} L_{\text{red}} \right] \\ 0 & \frac{A_u}{A_d} \end{pmatrix}, \quad (13)$$

where $p$ and $u$ are set as variables. The term $A_u/A_d$ is the ratio between the areas upstream and downstream of the burner, $\zeta$ the pressure loss coefficient of the burner, $L_{\text{red}}$ a reduced length accounting for inertia effects of the accelerated air column within the burner, $\rho_u$ the density upstream of the burner, and $c_u$ the speed of sound upstream of the burner. Good agreement between this theoretical model was found for low frequencies and confirmed the validity of the measurement technique. As acoustic radiation

and scattering effects might become important for higher frequencies and have not been considered in the model a deviation between measured and analytical results for the higher frequencies could be observed.

**The Transfer Matrix With Combustion.** To determine the transfer matrix with combustion a total of seven microphones was used. Three microphones were located at different axial locations upstream of the burner and four microphones were located downstream of the burner. To obtain a low background noise level the combustion chamber was equipped with an anechoic end. The multiple source (forcing) method was used to create three different test states, i.e., forcing was applied either upstream or downstream of the burner or at both sides of the burner simultaneously.

The coherence between the forcing signal and the measured microphone signals is an indicator for the data quality. The coherence between the forcing signal and the closest microphone to the burner is displayed in Fig. 9 for the three different test states. Except for the forcing upstream of the burner the quality of the

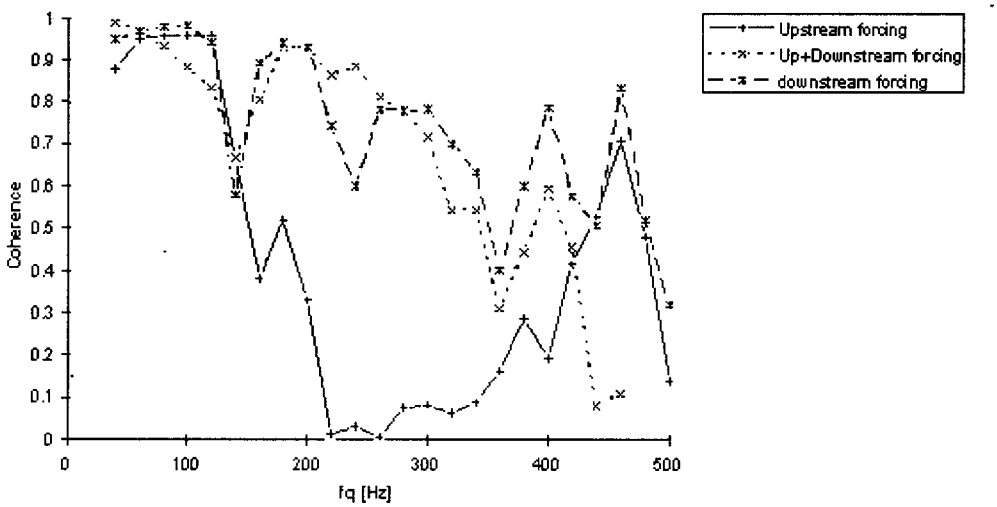



**Fig. 9 Coherence function $\Gamma_{\overline{xy}}$ between excitation signal and a microphone downstream of the burner for three different test states**

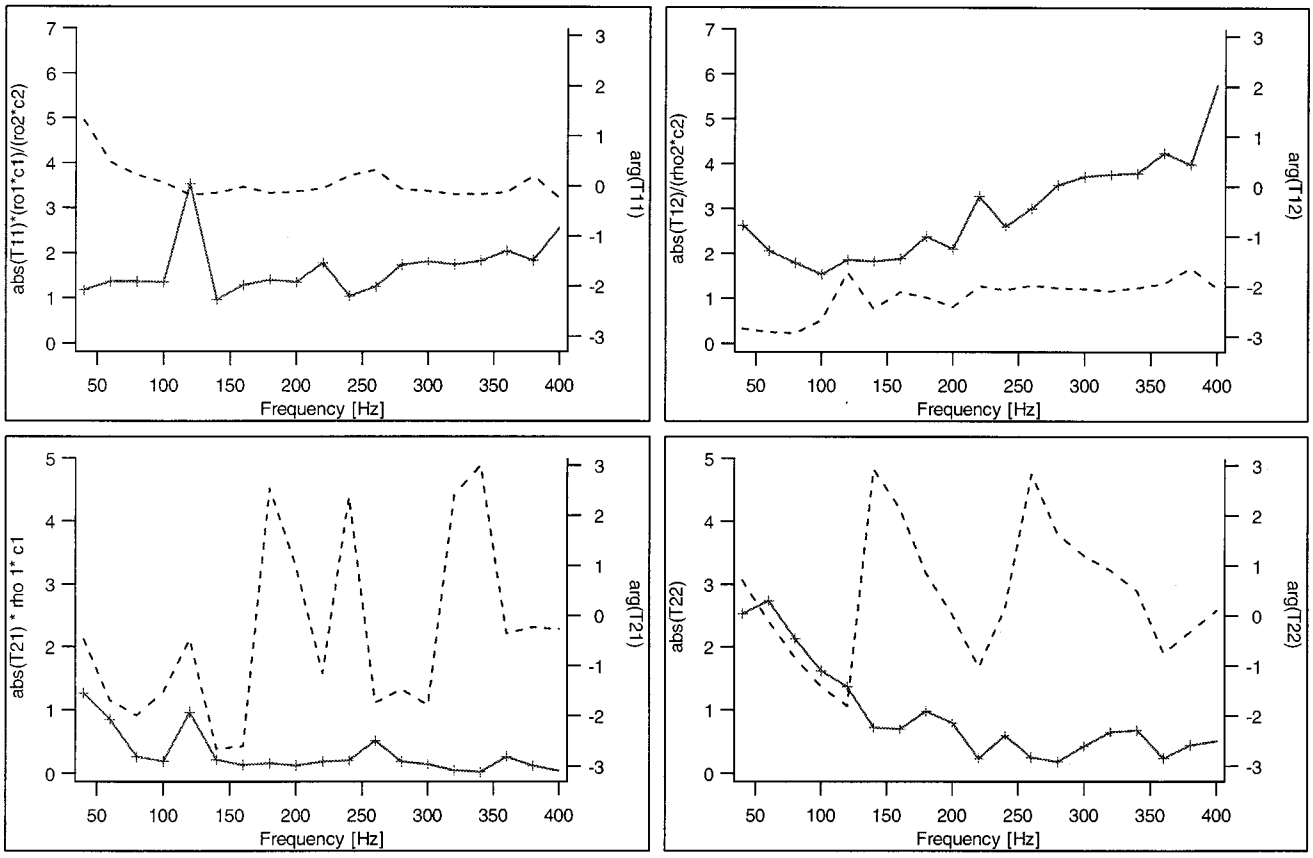



**Fig. 10 The transfer matrix of the burner/flame including the source term. Solid line: absolute value; dashed line: phase.**

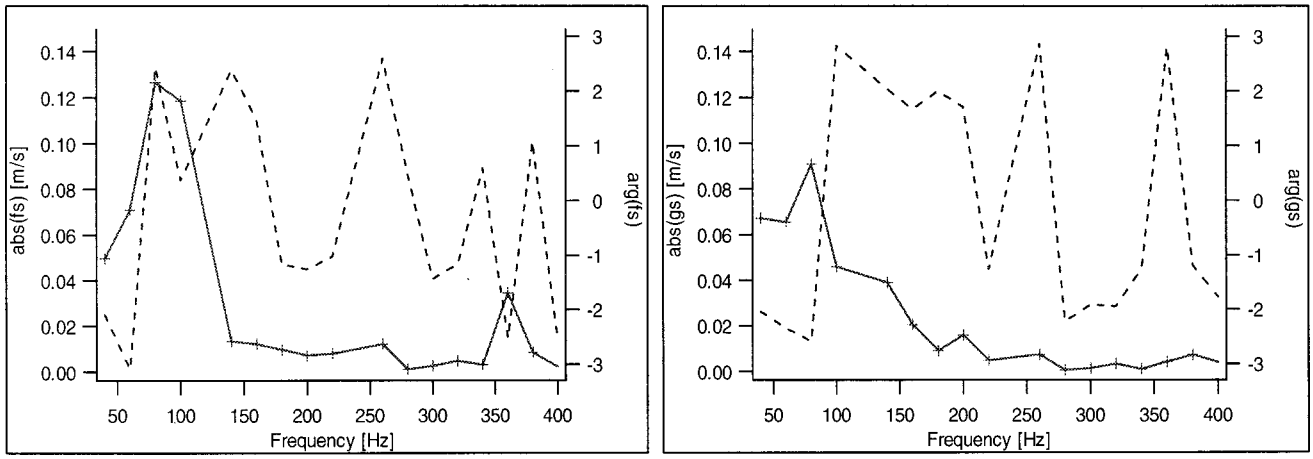



**Fig. 11 The source term of the flame. Solid line: absolute value; dashed line: phase.**

measured signal is good as indicated by a coherence value larger than 0.5. When forcing upstream the acoustic waves are reflected by the impedance of the burner resulting in a low level of the measured signal in the combustion chamber. In addition, low values only at specific frequencies, as in Fig. 9 are indicative of microphones placed in pressure nodes. This will be reflected in the coherence signal between the forcing signal and the specific microphone. However, for the determination of the transfer matrix a low coherence in one of the used microphones is not a problem as multiple microphones are used to determine the wave fields in the combustor. Using this technique the measurement was less sensitive to errors and a low coherence in one of the microphone signals, as displayed in Fig. 9, did not cause the method to be inaccurate.

In a first step the transfer matrix was calculated without separating the source term using Eq. (1). The coefficients are displayed in Fig. 10 in a frequency range of $40<f<400$ Hz. The coefficient $T_{11}$ is close to unity for frequencies $f<250$ Hz. The spike at $f=120$ Hz is believed to be related to the source term contribution, which has not been separated from the transfer matrix displayed in Fig. 10. For $f>250$ Hz an increase to 2 can be observed. $T_{12}$ shows almost linear behavior for frequencies $f>250$ Hz as predicted by the model without flame Eq. (13). However, in the low-frequency range a deviation between the model and the measurement can be observed. $T_{12}$ does not decrease with decreasing frequency as expected from the model but remains constant and shows even a slight increase. This behavior might be attributed to a source term. A similar behavior can be observed for $T_{22}$.

It was thus of further interest to determine the source term of the burner and the flame. Again, to measure this quantity a third test state is required which was provided by acoustic excitation upstream and downstream of the burner. The source term, consisting of $f_s$ and $g_s$ calculated from Eq. (12), is plotted as a function of the frequency in Fig. 11. The source term contains a significant contribution in the low-frequency region $40<f<160$ Hz with a maximum at $f=80$ Hz. For frequencies above 160 Hz the source term is very small, e.g., the burner does not act as a source at these high frequencies. The generation of noise in the low-frequency region is related to the burner flow turbulence leading to fluctuations in heat release and thus to acoustic waves. For comparison the wave components $f_u$ and $g_d$ incident on the burner are displayed in Fig. 12. For distinct frequencies the source term is of the same order of magnitude as the incident sound field on the burner indicating a significant source character of the burner/flame.

Having measured the "source free" scattering matrix the "source free" transfer matrix could then be determined by a simple matrix transition and is displayed in Fig. 13. The transfer matrix without source term differs in the low-frequency region

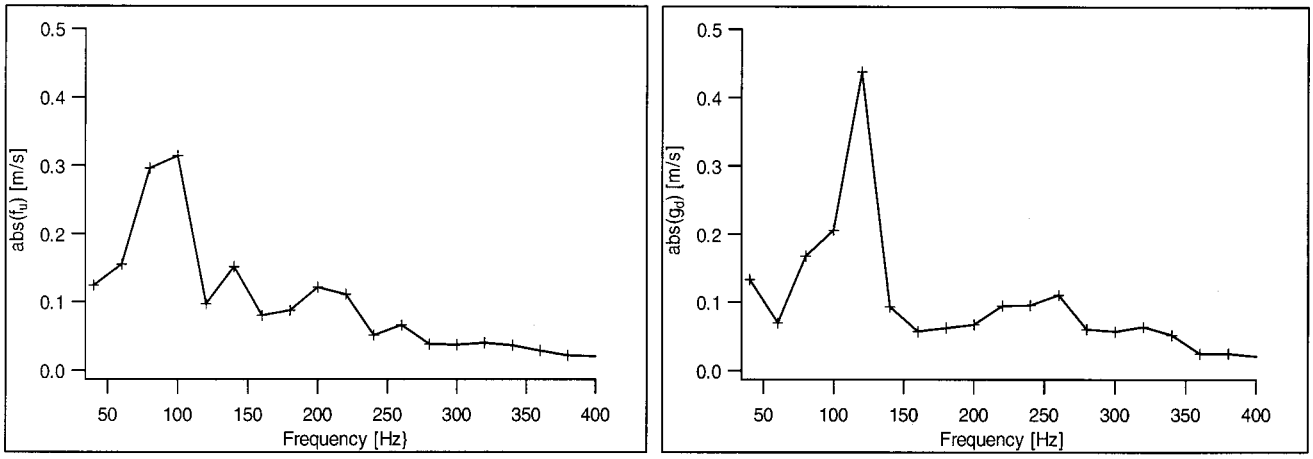



**Fig. 12 Incident wave components $f_u$ (upstream of the burner) and $g_d$ (downstream of the burner). Forcing was from the upstream and downstream of the burner simultaneously.**

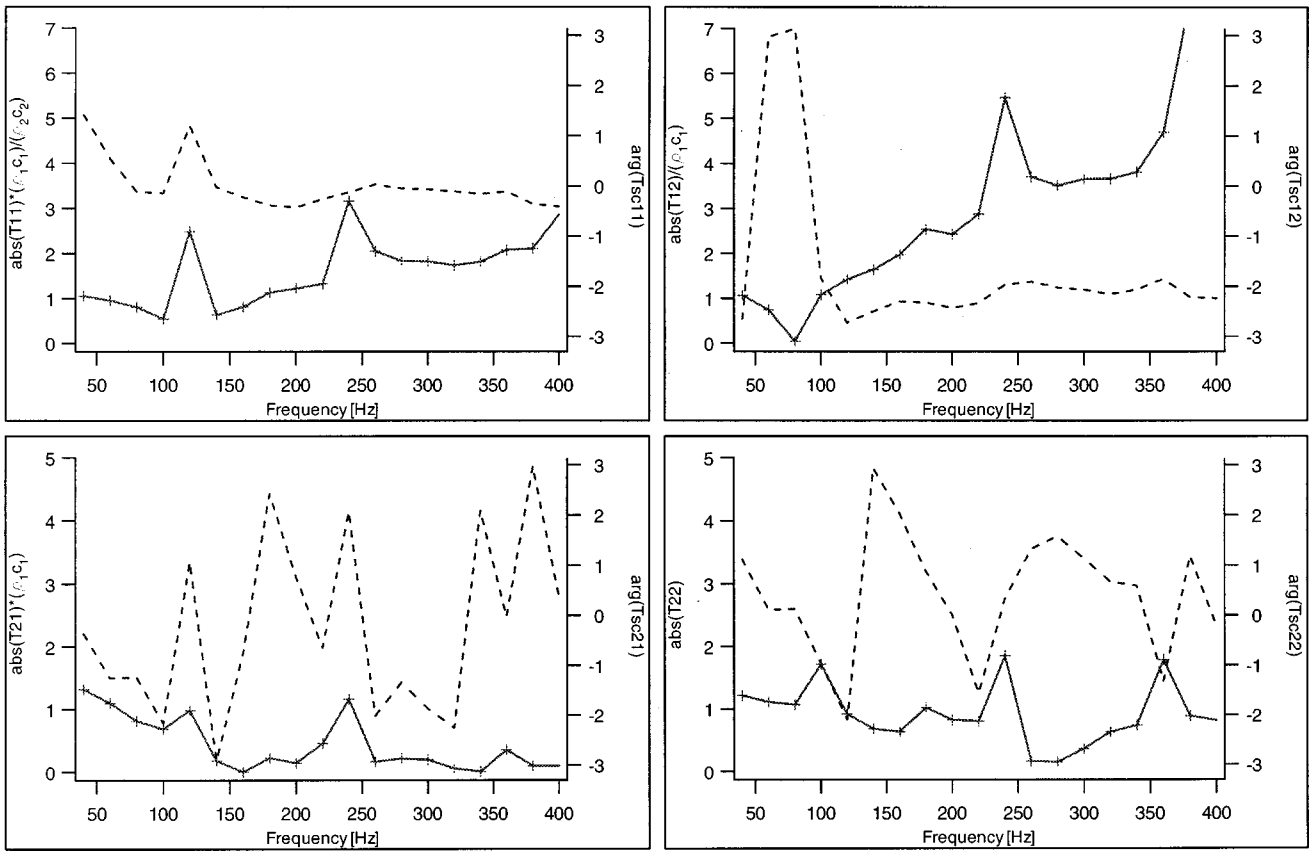



**Fig. 13 The transfer matrix of the flame without source term. Solid line: absolute value; dashed line: phase.**

from the transfer matrix which includes the source term. In the frequency region $40<f<160$ Hz the flame generates noise which is reflected in the source term.

The data for the source free transfer matrix shows a smaller peak at $f=120$ Hz but looks less smooth when compared to the transfer matrix which includes the source term. Both transfer matrices have been determined from three linear independent test states. As only two test states were required to calculate the transfer matrix which includes the source the system was overdetermined and thus its sensitivity to measurement errors reduced. The calculation of the source free transfer matrix, on the other hand, required all three test states and was thus more sensitive to measurement errors. The quality of the data can be improved by measuring more than the required three test states.

## Conclusions

The response of burner/flame to acoustic perturbations plays a crucial role when modeling thermoacoustic systems. In our approach we describe the burner and the flame as an active acoustic two-port ("black box") in which the state variables pressure and velocity are coupled via a four-element transfer matrix. The four coefficients of the transfer matrix required two linear independent test states which were achieved by acoustic forcing upstream and downstream of the burner, respectively. The measurement technique has been verified without combustion by comparing it to a simple model with good agreement.

The measurement technique was then applied to an experimental swirl stabilized premixed burner with combustion. Here, the turbulent flow through the burner and flow instabilities traveling into the flame may generate sound waves, which are incoherent with the forcing signal, so that the burner itself acts as a sound source. To understand the interaction between sound and heat release the source character of the burner has to be separated from the plain transfer function. In the case of the burner transfer function formulation with source term six unknowns have to be determined: the four coefficients of the transfer matrix $T$ and two coefficients $g_s$ and $f_s$ describing the source character of the flame. The Riemann invarients $g_s$ and $f_s$ represent here waves which are emitted by the burner acting as a source. To solve for all six unknowns a third linear independent test state is required which was achieved by acoustic excitation upstream and downstream of the burner/flame simultaneously.

The measured source term indicated that the burner/flame acts as a sound source mainly in the low-frequency region $40<f<160$ Hz. For frequencies above 160 Hz the source term is very small, e.g., the burner does not act as a sound source at these higher frequencies.

The measured flame transfer function and source term data will be used for the formulation of premix flame models and the stability analysis of thermoacoustic systems.

## Acknowledgment

We gratefully acknowledge the assistance of Wolfgang Weisenstein. The authors would also like to acknowledge the fruitful discussions with Dr. J. J. Keller, ABB Power Generation Ltd., and Dr. K. Doebbeling, ABB Corporate Research Ltd. on the subject of thermoacoustics.

## Nomenclature

$A$ = cross-sectional area
$L_{\mathrm{red}}$ = reduced length
$M$ = Mach number
$S$ = scattering matrix
$T$ = transfer matrix
$c$ = speed of sound

$f$ = Riemann invariant traveling in downstream direction
$g$ = Riemann invariant traveling in upstream direction
$k$ = wave number
$p$ = pressure fluctuations
$v$ = velocity fluctuations
$x$ = axial position
$\omega$ = angular frequency
$u$ = upstream
$d$ = downstream

## References


[1] Rann, R. L., Beckstead, M. W., Finlinson, J. C., and Brooks, K. P., 1993, "A Review of Rijke Tubes, Rijke Burners and Related Devices," Prog. Energy Combust. Sci., **19**, pp. 313–364.
[2] Rayleigh, J. W. S., 1878, Nature (London), **18**.
[3] Keller, J. J., 1995, "Thermoacoustic Oscillations in Combustion Chambers of Gas Turbines," AIAA J., **33**, No. 12
[4] Polifke, W., Paschereit, C. O., and Sattelmayer, T., 1997, "A Universally Applicable Stability Criterion for Complex Thermoacoustic Systems," VDI-Ber., pp. 455–460.
[5] Munjal, M. L., 1986, *Acoustics of Ducts and Mufflers*, John Wiley and Sons, New York.
[6] Cremer, L., 1971, "The Second Annual Fairy Lecture: The Treatment of Fans as Black Boxes," J. Sound Vib., **16**, pp. 1–15.
[7] Bodén, H., and Abom, M., 1995, "Modelling of Fluid Machines as Sources of Sound in Duet and Pipe Systems," Acta Acust. (Beijing), pp. 549–560.
[8] Lavrentjev, J., and Åbom, M., 1996, "Characterization of Fluid Machines as Acoustic Multi-port Sources," J. Sound Vib., **197**, pp. 1–16.
[9] Abom, M., 1992, "A note on the Experimental Determination of Acoustical Two Port Matrices," J. Sound Vib., **155**, pp. 185–188.
[10] Abom, M., Bodén, H., and Lavrentjev, J., 1992, "Source Characterization of Fans Using Acoustic 2-Port Models," *Proceedings of Fan Noise Symposium*, CETIM, France.
[11] Paschereit, C. O., and Polifke, W., 1997, "Characterization of Len Premixed Gas-Turbine Burners as Acoustic Multi-ports," *Bulletin of the American Physical Society/Division of Fluid Dynamics, Annual Meeting*, San Francisco, CA.
[12] Paschereit, C. O., Wygnanski, I., and Fielder, H. E., 1995, "Experimental Investigation of Subharmonic Resonance in an Axisymmetric Jet," J. Fluid Mech., **283**, pp. 365–407.
[13] Schadow, K. C., and Gutmark, E., 1992, "Combustion Instability Related to Vortex Shedding in Dump Combustors and Their Passive Control," Prog. Energy Combust. Sci., **8**, pp. 17–132.
[14] McManus, K. R., Poinsot, T., and Candel, S. M., 1993, "A Review of Active Control of Combustion Instabilities," Prog. Energy Combust. Sci., **19**, pp. 1–29.
[15] Paschereit, C. O., Gutmark, E., and Weisenstein, W., 1998, "Control of Thermoacoustic Instabilities and Emissions in an Industrial Type Gas-Turbine Combustor," *27th International Symposium on Combustion*, The Combustion Institute, Pittsburgh, PA, Aug. 2–7.
[16] Paschereit, C. O., Gutmark, E., and Weisenstein, W., 1998, "Flow-Acoustic Interactions as a Driving Mechanism for Thermoacoustic Instabilities," *4th AIAA/CEAS Aeroacoustics Conference*, Toulouse, France, June 2–4.
[17] Paschereit, C. O., Gutmark, E., and Weisenstein, W., 1996, "Role of Coherent Structures in Acoustic Combustion Control," *29th AIAA Fluid Dynamics Conference*, Albuquerque, NM, June 15–18.
[18] Paschereit, C. O., Gutmark, E., and Weisenstein, W., 1998, "Structure and Control of Thermoacoustic Instabilities in a Gas-Turbine Combustor," 36th AIAA Aerospace Science Meeting and Exhibit, Reno, NV, Jan. 12–15.
[19] Bloxsidge, G. J., Dowling, A. P., and Langhorne, P. J., 1989, "Reheat Buzz: An Acoustically Coupled Combustion Instability," J. Fluid Mech., **193**, pp. 445–473.
[20] McIntosh, A., 1990, "On Flame Resonances in Tubes," Combust. Sci. Technol., **69**, pp. 147–152.
[21] Janus, M. C., and Richards, G. A., 1996, "Results of a Model for Premixed Combustion Oscillations," *Proceedings of the 1996 AFRC International Symposium*, Baltimore, MD, Sept. 30–Oct. 2.
[22] Paschereit, C. O., and Polifke, W., 1998, "Investigation of the Thermoacoustic Characteristics of a Lean Premixed Gas Turbine Burner," ASME Turbo Expo '98, Stockholm, Sweden, June 2–5.

# Multicomponent and High-Pressure Effects on Droplet Vaporization

**S. K. Aggarwal**
Department of Mechanical Engineering,
University of Illinois at Chicago,
Chicago, IL 60607

**H. C. Mongia**
GE Aircraft Engines,
Cincinnati, OH 45215

*This paper deals with the multicomponent nature of gas turbine fuels under high-pressure conditions. The study is motivated by the consideration that the droplet submodels that are currently employed in spray codes for predicting gas turbine combustor flows do not adequately incorporate the multicomponent fuel and high-pressure effects. The quasi-steady multicomponent droplet model has been employed to investigate conditions under which the vaporization behavior of a multicomponent fuel droplet can be represented by a surrogate pure fuel droplet. The physical system considered is that of a multicomponent fuel droplet undergoing quasi-steady vaporization in an environment characterized by its temperature, pressure, and composition. Using different vaporization models, such as infinite-diffusion and diffusion-limit models, the predicted vaporization history and other relevant properties of a bicomponent droplet are compared with those of a surrogate single-component fuel droplet over a range of parameters relevant to gas turbine combustors. Results indicate that for moderate and high-power operation, a suitably selected single-component (50 percent boiling point) fuel can be used to represent the vaporization behavior of a bicomponent fuel, provided one employs the diffusion-limit or effective-diffusivity model. Simulation of the bicomponent fuel by a surrogate fuel becomes increasingly better at higher pressures. In fact, the droplet vaporization behavior at higher pressures is observed to be more sensitive to droplet heating models rather than to liquid fuel composition. This can be attributed to increase in the droplet heatup time and reduction in the volatility differential between the constituent fuels at higher pressures. For ignition, lean blowout and idle operations, characterized by low pressure and temperature ambient, the multicomponent fuel evaporation cannot be simulated by a single-component fuel. The validity of a quasi-steady high-pressure droplet vaporization model has also been examined. The model includes the nonideal gas behavior, liquid-phase solubility of gases, and variable thermo-transport properties including their dependence on pressure. Predictions of the high-pressure droplet model show good agreement with the available experimental data over a wide range of pressures, implying that quasi-steady vaporization model can be used at pressures up to the fuel critical pressure.*
[DOI: 10.1115/1.1423640]

## Introduction

The design of advanced gas turbine combustors is increasingly relying upon CFD-based methodologies. To this end, significant advances have been made in the computational capabilities for predicting the detailed structure of reacting two-phase flows in gas turbine combustors. However, several critical issues with regards to the physical-numerical modeling of these flows still remain unresolved. One such issue pertains to the realistic representation of the multicomponent nature of gas turbine fuels. The spray codes that are currently employed in the gas turbine industry are based on a single-component droplet vaporization submodel, although it has been recognized that gas turbine fuels are multicomponent with a wide distillation curve. This raises several questions regarding the applicability of these codes. First, the gasification behavior of multicomponent fuel sprays may be qualitatively different from that of corresponding pure fuel sprays. This would imply that a single-component spray model would not be adequate to predict the realistic fuel vapor distribution in a gas turbine combustor. This would clearly have an impact on the predicted ignition, flame stability, combustion characteristics, and pollutant levels. Second, the detailed and/or reduced chemistry models for hydrocarbon fuels ([1,2]) have generally been developed and validated for pure fuels, and their extension to multicomponent gas turbine fuels may not be straightforward. Third, the methodologies to model turbulent-chemistry interactions for single-component fuels would require modifications for jet fuels, especially for predicting soot.

Another issue pertains to the gasification behavior of an isolated fuel droplet. A multicomponent droplet ([3,4]) is known to exhibit a significantly different gasification behavior compared with that of a pure fuel droplet. These differences have been attributed ([3–9]) to transient liquid mass transport in the droplet interior, volatility differential between the constituent fuels, phase equilibrium at the droplet surface, and thermo-transport properties that are functions of mixture composition, temperature, and pressure. In order to address these complex issues for jet fuels in a systematic manner, we have taken a first step, i.e., study the vaporization behavior of a multicomponent fuel droplet and examine if it can be simulated by a surrogate pure fuel droplet.

In this paper, we examine conditions under which the gasification behavior of a multicomponent fuel droplet may be represented by a surrogate single-component fuel droplet. The range of conditions considered corresponds to the operating range of gas turbine combustors. The state-of-the-art vaporization models ([4–9]) are employed for both the multicomponent and single-component fuel droplets. In particular, the two commonly employed liquid-heating models, namely the infinite-diffusion ([7,9]) model and the diffusion-limit model ([7,9]), are used for both the single and multicomponent-fuel droplets. The objective of using

Contributed by the Combustion and Fuels Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the C&F Division, October 2000; final revision, March 2001. Editor: S. R. Gollahalli.

these models is to assess the effect of transient liquid transport, especially that of temperature and concentration distributions within the droplet interior, on the prediction of droplet vaporization rates under conditions relevant to gas turbine combustors. The thermophysical and transport properties of both the gas phase and liquid phase are calculated in a comprehensive manner. It should be noted that numerous previous studies have examined the gasification behavior of multicomponent fuel droplets. However, none of these studies have focused on the issue of simulating the vaporization behavior of a multicomponent fuel droplet by a surrogate single-component droplet under conditions relevant to gas turbine combustors.

Another objective of the present study is to extend the low-pressure quasi-steady droplet vaporization model to high-pressure conditions including the thermodynamic supercritical state of the liquid fuel. This is motivated by several considerations. First, the modern turbo-propulsion gas turbine combustors operate at pressures approaching or exceeding 40 atmospheres. Second, in high-performance military engines, the liquid fuel is being looked at as the primary coolant for on-board heat sources, and may attain a critical state before it is ''atomized.'' Third, the high-pressure gasification phenomena exhibit characteristics that are distinctly different from those in a low-pressure environment. For example, the gas-phase nonidealities and the liquid-phase solubility, which are negligible at low pressures, must be taken into considerations as the ambient pressure approaches the critical state of the liquid fuel. Finally, the liquid-vapor equilibrium at the droplet surface and the latent heat of evaporation are markedly different at low and high-pressure conditions. Consequently, the conventional low-pressure droplet models may not be valid at pressures near and exceeding the fuel critical pressure.

The transcritical droplet gasification phenomena are also relevant to diesel and liquid rocket engines. Consequently, a number of experimental, analytical, and computational studies have been reported in this area. Experimental studies have employed the porous-sphere ([10]) (for droplet combustion), suspended-droplet ([11]), and freely falling droplet ([12]) configurations to examine the supercritical droplet gasification behavior for a variety of liquid fuels. Computational studies ([13–17]) have often considered an isolated, stationary fuel droplet which is suddenly placed in a high-pressure environment. Detailed simulations based on the solution of transient, spherically symmetric gas-phase equations have been reported. Major differences between the various theoretical approaches have been in the treatment of liquid-phase transport processes and the representation of high-pressure effects. These effects in general have included (i) a nonideal equation of state with appropriate mixing rules, (ii) liquid-phase solubility of ambient gases, (iii) high-pressure treatment of liquid-phase equilibrium based on the fugacity of each phase, and (iv) effect of pressure on thermo-transport properties. An extensive review of the published experimental and computational investigations is provided by Gilver and Abraham [18].

It is important to emphasize that our objective is to evaluate a high-pressure droplet vaporization model within the framework of the quasi-steady approximation, such that the model can be economically incorporated in high-pressure spray algorithms appropriate for 1–10 million node calculations for gas turbine simulations. While more advanced vaporization models based on a transient gas-phase analysis may be more desirable under high-pressure conditions, it is currently not feasible to employ them in comprehensive spray computations.

## The Physical Model

An isolated multicomponent (containing $N$ components) fuel droplet evaporating in a high-temperature high-pressure environment is analyzed. The droplet size, ambient temperature, and pressure are considered in a range that corresponds to a wide range of power requirements for a gas turbine combustor. The gas-phase processes are assumed to be quasi-steady, which implies that the characteristic gas-phase time is much shorter compared to the liquid-phase transient time as well as the time associated with the surface regression rate. This requires ([6]) that the ratio of gas density to liquid density be at least an order of magnitude smaller than unity. Other assumptions include spherical symmetry, phase-equilibrium at the droplet surface, and negligible secondary diffusion and radiation. Then, the energy and fuel-vapor species conservation equations can be written as

$$\frac{d}{dr}\left( r^2\rho v c_p(T-T_s) - r^2\rho D c_p \mathrm{Le}\frac{d(T-T_s)}{dr}\right)=0 \tag{1}$$

$$\frac{d}{dr}\left( r^2\rho v Y_i - r^2\rho D\frac{dY_i}{dr}\right)=0 \tag{2}$$

where $r$ is the radial coordinate, $v$ is the gas or Stefan flow velocity, Le is the gas-phase Lewis number, and $Y_i$ is the mass fraction of $i$th fuel species with $i=1,2,\ldots N$. In addition, $T$ is the ambient temperature, $T_s$ the droplet surface temperature, $\rho$ the gas density, $c_p$ the gas specific heat at constant pressure, and $D$ the diffusion coefficient, With appropriate boundary conditions, the solution of Eq. (1) at $r=r_\infty$ yields

$$\frac{\dot{m}}{4\pi\rho D\mathrm{Le}}[1/r_s-1/r_\infty]=\ln\left[1+\frac{c_p(T_\infty-T_s)}{H}\right] \tag{3}$$

while the solution of Eq. (2) at $r=r_s$ yields

$$\frac{\dot{m}}{4\pi\rho D}[1/r_s-1/r_\infty]=\ln\left[\frac{\varepsilon_i-Y_{i\infty}}{\varepsilon_i-Y_{is}}\right]. \tag{4}$$

$H$ in Eq. (3) represents the energy supplied to the droplet (per unit mass of fuel vaporized) for heating and vaporization, and $\dot{m}$ represents the vaporization rate. Further, $r_\infty$ is the radial location representing the ambient conditions, which are assumed to be specified at infinity for an isolated droplet, and $\varepsilon_i$ is the fractional vaporization rate of species $i$, given by $\varepsilon_i=\dot{m}_i/\dot{m}$. Summing Eq. (2) over all the fuel species and integrating the resulting equation yields

$$\frac{\dot{m}}{4\pi A\rho D}[1/r_s-1/r_\infty]=\ln(1+B) \tag{5}$$

where $B$ is the transfer number given by

$$B=\frac{Y_{fs}-Y_{f\infty}}{1-Y_{fs}}. \tag{6}$$

Equations (3), (4), and (5) can be combined to obtain expressions for $\varepsilon_i$ and $H$. Details are provided in Ref. [10]. The droplet size history is computed using

$$\frac{dr_s^2}{dt}=-\frac{\dot{m}}{2\pi r_s\rho_l}. \tag{7}$$

In summary, the low pressure model for both the single and multicomponent fuel droplets considers gas-phase processes to be quasi-steady phase, equilibrium at the droplet surface, and transient liquid-phase processes represented by either the infinite-diffusion model or the diffusion-limit model ([7,9]). As discussed in the next section, the phase equilibrium is represented by the Clausius-Clapeyron relation for each fuel component, supplemented by the Raoult's law for the multicomponent case.

## High-Pressure Models

Two high-pressure models are employed. The first model is based on the ideal equation of state and does not consider the solubility of gases into liquid. Consequently, the phase equilibrium at the droplet surface is represented by using the Clausius-Clapeyron relation for each fuel component, supplemented by the Raoult's law for the multicomponent case. The Clausius-Clapeyron relation can be written as

$$X_{oi} = \exp\left\{\frac{L_i M_{fi}}{R_o}\left(\frac{1}{T_{bi}} - \frac{1}{T_s}\right)\right\} \tag{8}$$

where the subscript $i$ refers to the fuel component $i$, $X_{oi}$ is the equilibrium fuel vapor mole fraction, $T_{bi}$ the boiling temperature at a given pressure $p$, $L_i$ the heat of vaporization (considered function of $T_s$ and pressure), $M_{fi}$ the molecular weight, and $R_o$ the universal gas constant. The vapor mole fraction of each fuel component at the surface is calculated using the Raoult's law

$$X_{is} = X_{ils} X_{oi} \tag{9}$$

where $X_{ils}$ is the liquid mole fraction of fuel species $i$ at the surface. The vapor mole fraction of each fuel species at the droplet surface ($Y_{is}$) can be calculated from $X_{ils}$ and $M_i$, while $T_{bi}$ and $L_i$ are calculated as

$$T_{bi} = A_1/(A_2 - \ln(P)) \tag{10}$$

$$L_i = \left(\frac{T_{ci} - T_s}{T_{ci} - T_{bni}}\right)^{0.38} L_{ni} \tag{11}$$

where $A_1 = L_{ni} \cdot M_{fi}/R_o$, $A_2 = A_1/T_{bni}$, $L_{ni}$ is the heat of vaporization at normal boiling temperature ($T_{bni}$), and $P$ is the pressure in atmosphere. The liquid-phase transient processes appear through the variables $X_{ils}$ and $T_s$, and are represented by the infinite-diffusion and diffusion-limit models. These two models are described in Ref. [19]. The important difference between the low-pressure model and the first high-pressure model is that the liquid boiling temperature and heat of vaporization are considered to be pressure-dependent in the latter model.

The second high-pressure model considers the real gas behavior, the solubility of gases into the liquid, and the effect of pressure on thermo-transport properties. The real gas behavior is represented by using the Peng-Robinson equation of state ([20,21]) in the following form:

$$Z = \frac{PV}{RT} = \frac{V}{V-b} - \frac{c/RT + d - 2\sqrt{cd/RT}}{(V+b) + (b/V)(V-b)} \tag{12}$$

where $Z$ is the compressibility factor, $R$ the universal gas constant (82.052 atm.cm$^3$/mole.K), $T$ the temperature ($K$), $P$ the pressure (atm), and $V$ the molar volume (cm$^3$/mole). For a pure fluid, the constants $b$, $c$, and $d$ are given as

$$\begin{aligned} b &= 0.0778 R T_c / P_c \\ c &= a(T_c)(1+k)^2 \\ d &= a(T_c) k^2 / R T_c \\ a(T_c) &= 0.45724 R^2 T_c^2 / P_c . \end{aligned} \tag{13}$$

Here the subscript $c$ represents critical state. Equation (12) is solved as a cubic in $Z$ to calculate the ambient compressibility factor and density. The mixture properties ($Z$ and $\rho$) at the droplet surface are obtained by employing the appropriate mixing rules ([22]). Details are provided in Ref. [23]. The vapor-liquid equilibrium at the droplet surface is expressed by the equality of chemical potential of each species in the liquid and vapor phases, and can be written as

$$x_i^V \phi_i^V = x_i^L \phi_i^L \tag{14}$$

where the fugacity coefficients $\phi_i$ for both the vapor and liquid phases can be written as

$$RT \cdot \ln \phi_i = \int_V^\infty [(\partial P/\partial n_i)_{T,V,n_i} - (RT/V)]dV - RT \ln Z. \tag{15}$$

Using the Peng-Robinson equation of state and the appropriate mixing rules, the fugacity coefficients can be expressed as functions of $T$, $P$, and $Z$. Details are provided in Ref. [23].

The latent heat of evaporation is given by

$$h_i = h_i^o - RT^2(\partial \ln \phi_i/\partial T)_{P,x_i} \tag{16}$$

where the enthalpy $h_i$ is in J/mol, and the universal gas constant $R$ is in J/mol/K. The vapor and liquid phases refer to the same state $o$. Then, the latent heat of evaporation is defined by the difference of the vapor and liquid-phase enthalpies

$$L_i = h_i^V - h_i^L \tag{17}$$

## Thermophysical and Transport Properties

A detailed algorithm is developed for calculating the variable thermophysical and transport properties of both phases. For the liquid phase, the variable properties are considered for both the single-component and multicomponent fuels. For a majority of cases, the data is compiled from various sources ([20,24–28]) and is employed in the form of polynomials. All the thermo-transport properties are considered to be pressure, temperature, and species-dependent. The procedure along with the relevant equations are provided in Refs. [19] and [23].

The calculation of thermo-transport properties was confirmed by independently using the Chemkin subroutines ([16,28]) as well as by comparing with experimental data ([25–27]). As discussed in Refs. [19] and [23], the property algorithm was shown to reproduce the experimental data quite well.

The quasi-steady droplet vaporization model requires that the thermo-transport properties of the gas film in the droplet vicinity be calculated continuously as the droplet evaporates. All the gas-film mixture properties are computed at the weighted-averaged temperature and species mass fractions, obtained from the temperature and composition at the droplet surface and those at infinity, as

$$\Phi_{\text{avg}} = \alpha \Phi_{gs} + (1-\alpha)\Phi_g \tag{18}$$

where $\Phi$ is a generic quantity representing either temperature or mass fraction, and the $\alpha$ is selected to be as 0.7. The subscripts $gs$ and $g$ represent the gas-phase property at the droplet surface and outside the gas film (infinity), respectively.

## The Solution Procedure

The theoretical model described above is applicable to both single-component and multicomponent fuel droplets. In the present study, the results are obtained for a bicomponent ($N=2$) and an equivalent single-component ($N=1$) fuel droplet. A general procedure involves calculating $T_{bi}$, $L_i$, and $Y_{is}$ by using $T_s$ at the old time-step. Then, the average gas temperature and species mass fractions are obtained from Eq. (18), and the thermo-transport properties of the gaseous mixture are calculated by using the equations described in Refs. [19] and [23]. Similarly, the liquid fuel properties, such as specific heat, thermal conductivity, and density, are computed. Note that liquid temperature used for calculating these properties is the droplet surface temperature for the infinite-diffusion model, while an average of the liquid temperatures at the droplet surface and center are used for the diffusion-limit model. The new $T_s$ is then calculated by using the infinite-diffusion or the diffusion-limit model. Finally, the droplet radius is calculated by using Eq. (7).

## Results and Discussion

First, we focus on the comparison of the vaporization characteristics of a bicomponent fuel droplet and an surrogate pure fuel droplet for conditions that correspond to the operating range of a typical gas turbine combustor. For the base case, a bicomponent fuel containing equal amounts of n-$C_{10}H_{22}$ and n-$C_{14}H_{30}$ by mass is considered, and its vaporization behavior is compared with that of pure n-$C_{12}H_{26}$ droplet. The boiling temperature of n-$C_{12}H_{26}$ is approximately equal to that of Jet-A fuel. At one atmospheric pressure, the boiling temperatures of $C_{10}H_{22}$, $C_{12}H_{26}$ and $C_{14}H_{30}$ are 447.3 K, 489.5 K, and 526.7 K. These temperatures correspond, respectively, to the initial boiling, midpoint boiling (50

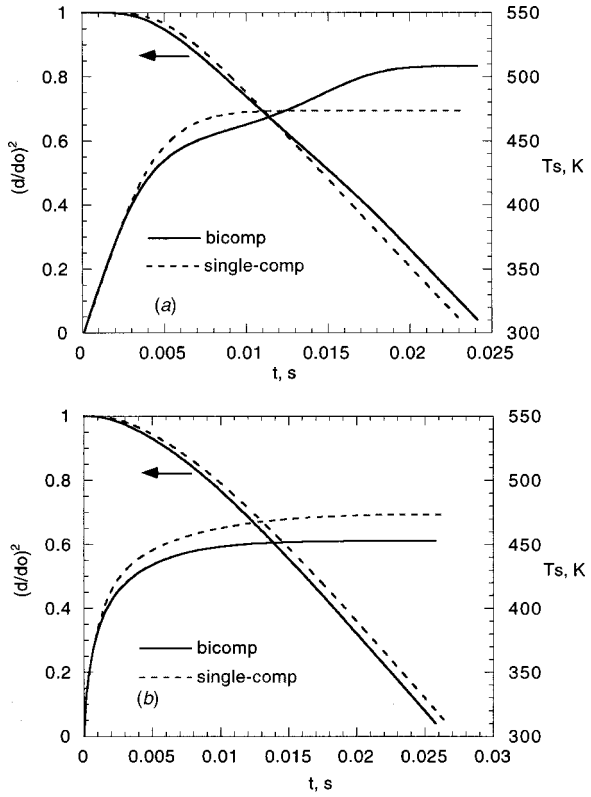



**Fig. 1 Temporal histories of droplet surface area (nondimensional) and temperature for a bicomponent ($n\text{-}C_{10}H_{22}$ and $n\text{-}C_{14}H_{30}$) and an equivalent single-component ($n\text{-}C_{12}H_{26}$) fuel droplet. Predictions of infinite-diffusion and diffusion-limit models are shown in Figs. 1(*a*) and 1(*b*), respectively. Ambient temperature=1500 K, pressure=1 atm, and initial droplet diameter and temperature are 100 $\mu$m and 300 K, respectively.**

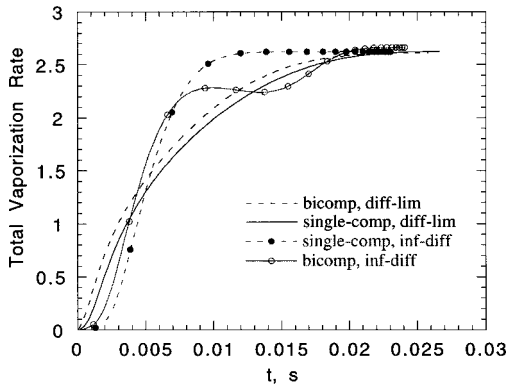



**Fig. 2 Comparison of total vaporization rates (nondimensional) for a bicomponent and an equivalent single-component fuel droplet predicted using the infinite-diffusion and diffusion-limit models for the conditions of Fig. 1**

percent) and final boiling for Jet-A fuel. The calculations presented for the base case are for a droplet with initial diameter of 100 $\mu$m, which is placed in an ambient with temperature and pressure of 1500 K and one atmosphere respectively. The 1500 K represents a typical incipient lean blowout temperature for ground-idle operation. Consequently, these calculations are representative for estimating the effect of heatup and vaporization models on lean blowout fuel-air ratios.

Figure 1 presents the vaporization and surface temperature histories of a bicomponent ($n\text{-}C_{10}H_{22}$ and $n\text{-}C_{14}H_{30}$) and an equivalent single-component ($n\text{-}C_{12}H_{26}$) droplets. The predictions of infinite-diffusion and diffusion-limit modes are shown in Figs. 1(*a*) and 1(*b*), respectively. The overall differences in the vaporization histories of bicomponent and single-component droplets are small, indicating that a suitably selected single-component fuel can be used to represent the vaporization behavior of a bicomponent fuel. There are, however, differences in the temporal variation of surface temperature ($T_s$) for the bicomponent and single-component fuels. For the bicomponent case, the infinite-diffusion model predicts a batch-distillation type behavior, see Fig. 1(*a*). Initially, the bicomponent droplet has lower surface temperature compared to the single-component droplet, as $T_s$ is limited by the boiling temperature of volatile component, but higher vaporization rate since the surface vapor concentration of volatile component is higher (even though the surface temperature is lower) than that for the single-component fuel. However, later in the lifetime, the bicomponent droplet has higher $T_s$, which is now governed by the boiling temperature of less volatile component, but a lower vaporization rate due to the presence of less volatile fuel at the surface. For the diffusion-limit model, the bicomponent droplet has lower surface temperature but higher vaporization rate compared to the single-component droplet throughout its lifetime. This is due to the presence of more volatile component at the bicomponent droplet surface.

The comparison of the total vaporization rates of bicomponent and single-component fuels predicted using the infinite-diffusion and diffusion-limit models is shown in Fig. 2. The total vaporization rate is a better indicator of how well a single-component fuel can represent the vaporization behavior of a bicomponent fuel spray. Results indicate that by using the diffusion-limit model, the vaporization rate of a bicomponent fuel droplet can be well represented by an equivalent single-component droplet. In contrast, the use of infinite-diffusion model leads to more discrepancies between the vaporization rates of bicomponent and single-component fuel droplets. This is attributable to the batch-distillation behavior predicted by the infinite-diffusion model for the bicomponent droplet. The batch-distillation behavior is characterized by the vaporization of the more volatile component, followed by leveling of the vaporization rate since the less volatile component is being heated during this time, and subsequent resumption of the higher vaporization rate due to the vaporization of less volatile component.

In order to examine whether this behavior is observed for a wide range of conditions in a gas turbine combustor, results were obtained for different ambient temperature (800 K) and initial droplet diameter (25 $\mu$m). These conditions are relevant to premixing, prevaporizing systems at high-power conditions, especially for dry low-emission combustors. For these cases, using the diffusion-limit model, the total vaporization rate of a bicomponent fuel droplet was observed to be well represented by that of an equivalent single-component droplet.

## Multicomponent Effects During Initial Starting and Ignition

As known from experience and confirmed by simple simulations here, spray-droplet heatup and evaporation rates during start (ignition and flame propagation) are impacted significantly by the fuel properties. Model results for heatup and evaporation are presented for a gas temperature of 373 K. Figure 3 shows the temporal variation of the droplet diameter squared and surface temperature for the bicomponent and single-component cases. Contrary to the high-gas-temperature cases discussed above, there are now significant differences in the vaporization histories of bicomponent and the corresponding single-component fuel droplets.

Figure 4 presents the comparison of the total vaporization rates for the bicomponent and single-component fuel droplets predicted using the two models. It is evident that for these conditions, the

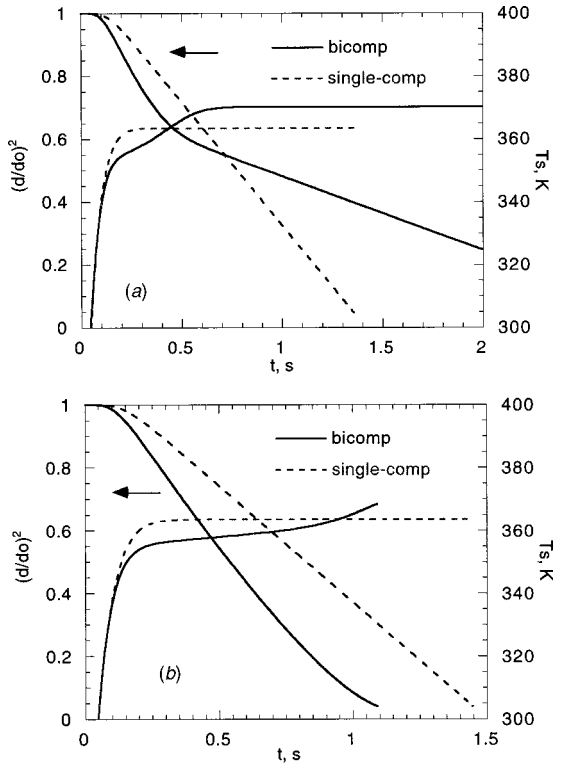



**Fig. 3 Temporal histories of droplet surface area (nondimensional) and temperature for a bicomponent ($n\text{-}C_{10}H_{22}$ and $n\text{-}C_{14}H_{30}$) and an equivalent single-component ($n\text{-}C_{12}H_{26}$) fuel droplet. Predictions of infinite-diffusion and diffusion-limit models are shown in Figs. 3(*a*) and 3(*b*), respectively. Ambient temperature=373 K, pressure=1 atm, and initial droplet temperature=233 K.**

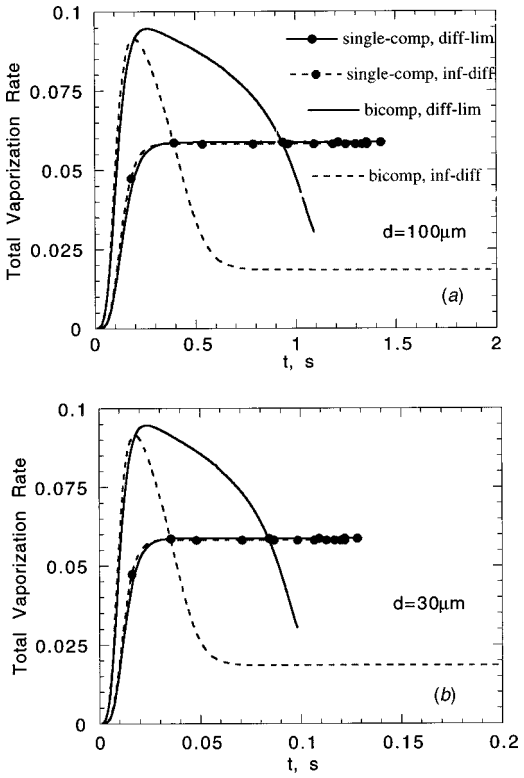



**Fig. 4 Total vaporization rates (nondimensional) for a bicomponent and an equivalent single-component fuel droplet predicted using the infinite-diffusion and diffusion-limit models. Initial droplet diameter is 100 $\mu$m for Fig. 4(*a*) and 30 $\mu$m for Fig. 4(*b*). Other conditions are the same as those in Fig. 3.**

vaporization behavior of a bicomponent droplet cannot be simulated by an equivalent pure fuel droplet, as both the instantaneous vaporization rate and the droplet life time differ considerably for the bicomponent and single-component cases. This can be attributed to the fact that the ambient temperature for this case is much lower than the boiling temperature of the less volatile component. In addition, it is indicated that while the differences in the predictions of the two models for the bicomponent case are quite significant, they are negligible for the corresponding single-component case. A similar behavior is observed when the initial droplet diameter is reduced to 30 microns, except that the life time of the 30-micron droplet is reduced by a factor of about 10 compared to that of the 100-micron droplet. An important implication of this result is that the ignition characteristics of a multicomponent fuel spray cannot be simulated by using a single-component fuel spray model, since the ignition behavior for the former is governed by the presence of the volatile component ([29]).

## Effect of Pressure

Figure 5 presents a comparison of the vaporization characteristics of bicomponent and equivalent single-component fuel droplets at different pressures. The corresponding plots showing the temporal variation of surface temperature are given in Fig. 6. The results in Figs. 5 and 6 are obtained using the first high-pressure vaporization model. Important observations from these figures are as follows.

1 For both the bicomponent and single-component fuel droplets, the droplet lifetime is not altered significantly as the ambient pressure is increased. However, the droplet heatup time becomes a more significant part of droplet lifetime, increasing from about 20 to 40 percent of the lifetime as $p$ is increased from 1 to 15 atm. The increase in droplet heatup time is caused by the increase in the fuel boiling temperature with pressure. The boiling temperature of n-dodecane increases from 489.5 K to 654.9 K as pressure is increased from $p=1$ atm to 15 atm. Correspondingly, for the bicomponent case, the boiling temperature of $n\text{-}C_{10}H_{22}$ (volatile component) increases from 447.3 K to 601.6 K, while that of $n\text{-}C_{14}H_{30}$ increased from 526.7 K to 701.3 K. Obviously, this would increase the droplet heatup time for a fixed ambient temperature for both the bicomponent and single-component cases. Note that an increase in $p$ decreases the latent heat of evaporation, which would increase the vaporization rate. However, this effect is largely compensated by an increase in the droplet heatup time. Note that for the convective case, another effect of pressure appears through its influence on the droplet Reynolds number which increases as the pressure is increased. This would enhance the vaporization rate.

2 The difference between the predictions of infinite-diffusion and diffusion-limit models is small at low pressures, but becomes increasingly more noticeable at higher pressures. As pressure increases, the infinite-diffusion model predicts increasingly higher vaporization rate compared to that by the diffusion-limit model. For example, at $p=15$ atm, the droplet lifetime predicted by this model is about 40 percent shorter. This can be attributed to two effects. First, it overpredicts the droplet surface temperature (during most of its lifetime) compared to that by the diffusion-limit model. Second, the heat of evaporation is reduced at higher pressures, so that even small differences in surface temperature can cause significant changes in the transfer number $B$.

3 For the bicomponent case, the batch distillation process, predicted by the infinite-diffusion model at $p=1$ atm, becomes less noticeable at higher pressures. This can be attributed to a decrease in relative volatility differential at higher pressures. Note that

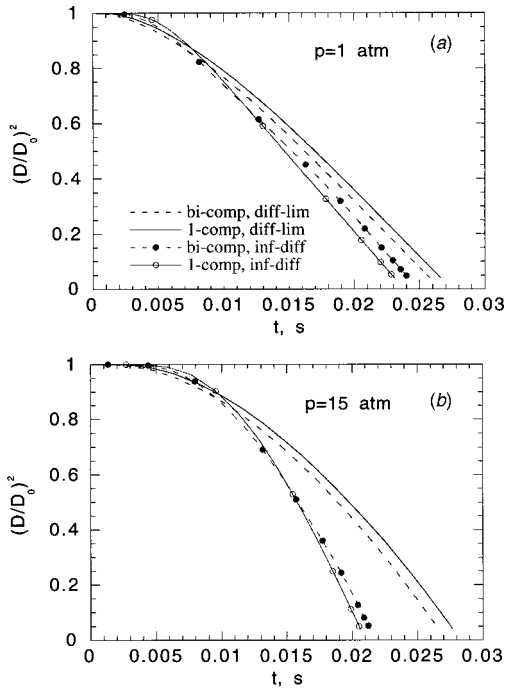



**Fig. 5 Temporal variation of droplet surface area (nondimensional) for a bicomponent ($n\text{-}C_{10}H_{22}$ and $n\text{-}C_{14}H_{30}$) and an equivalent single-component ($n\text{-}C_{12}H_{26}$) fuel droplet for pressure of 1 atm (Fig. 5(*a*)) and 15 atm (Fig. 5(*b*)). Predictions of the infinite-diffusion and diffusion-limit models are shown. The gas temperature is 1500 K, and initial droplet diameter is 100 $\mu$m.**

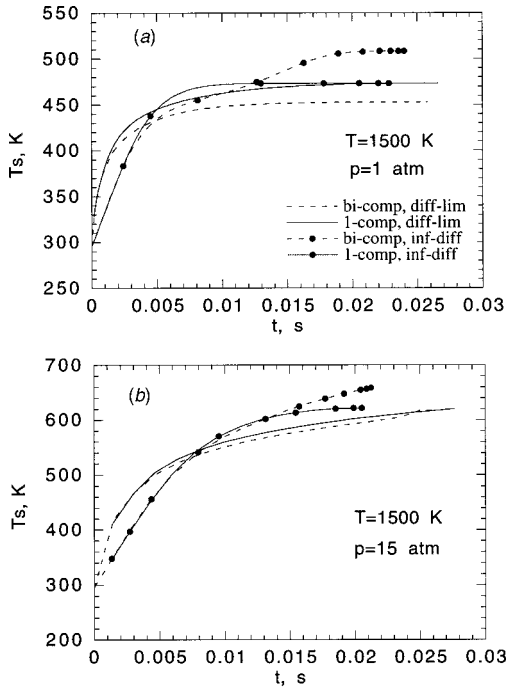



**Fig. 6 Temporal variation of droplet surface temperature for a bicomponent and an equivalent single-component fuel droplet for the conditions of Fig. 5**

batch distillation process at $p=1$ atm is easily discernible in the temperature plot for the infinite-diffusion model in Fig. 6(*a*).

4 The comparison of the temporal variations of droplet diameter squared and surface temperature for the bicomponent and single-component cases indicates that for high ambient temperatures, a single-component droplet model can simulate the vaporization behavior of a bicomponent fuel droplet reasonably well over a wide range of pressures. In fact, the predicted vaporization behavior is significantly more sensitive to the droplet heating model used rather than to the liquid fuel composition. In addition, the representation of a bicomponent fuel droplet by an equivalent single-component seems to become increasingly better as the pressure is increased. This is apparently due to the increased droplet heatup time and the reduced volatility differential at higher pressures.

## Validation of the High-Pressure Vaporization Model

The second quasi-steady high-pressure model was validated by comparing its predictions with the available experimental data. Figure 7 presents the predicted liquid-vapor equilibrium for nitrogen/n-heptane mixtures in terms of the equilibrium nitrogen mole fraction as a function of pressure and temperature. The predicted vapor and liquid mole fractions of $N_2$ show excellent agreement with the experimental data of Knapp et al. [24].

The vaporization history predictions are validated using the experimental data of Stengele et al. [12]. In the experiential study, freely falling n-heptane droplets in a stagnant, high-pressure, high-temperature nitrogen environment were considered. The droplet diameter and velocity were measured along the droplet trajectory using a high-speed video and stroboscope lamp. Since the measured data were provided along the droplet trajectory, the following equations are employed in the computational model to calculate the temporal variation of droplet position and velocity:

$$\frac{dx_d}{dt}=u_d \tag{19}$$

$$\frac{du_d}{dt}=\frac{3\rho C_d}{4\rho_l d_d}|u-u_d|(u-u_d)+(1-\rho/\rho_l)g \tag{20}$$

and the drag coefficient is calculated using the same correlation as used in the cited study ([12])

$$C_d=[0.36+5.48\,\mathrm{Re}_d^{-0.573}+24/\mathrm{Re}_d](1+B)^{-0.2} \tag{21}$$

$$\mathrm{Re}_d=\rho|u-u_d|d_d/\mu. \tag{22}$$

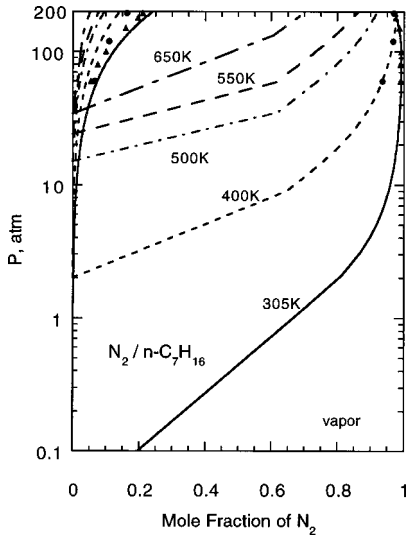



**Fig. 7 Pressure-mole fraction diagram for nitrogen/n-heptane mixtures, calculated using the Peng-Robinson equation of state. Experimental data from Ref. [24] at 305 K and 400 K are also shown.**

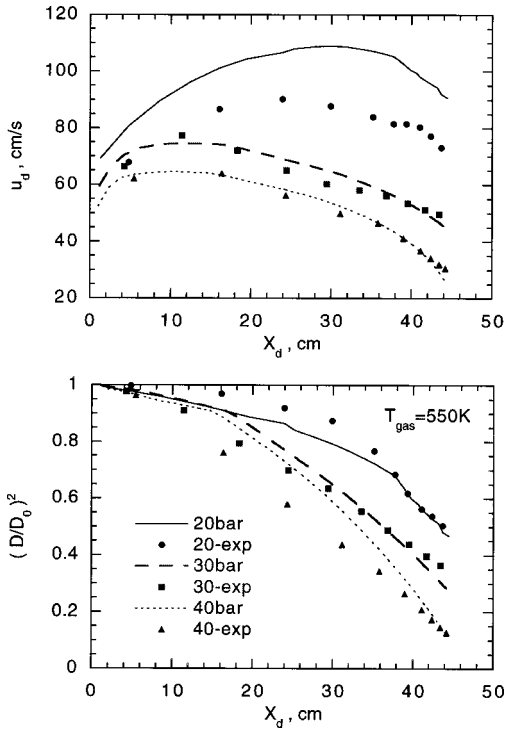



**Fig. 8 Comparison of the predicted and measured (n-heptane) droplet velocity and nondimensional surface area along the trajectory for three different ambient pressures. Ambient temperature is 550 K and initial droplet diameter is 780 $\mu$m.**

The gas velocity ($u$) in the above equations was taken to be zero in the experimental study. The gas density ($\rho$) and viscosity ($\mu$) are assumed to be ambient density and gas film viscosity, respectively.

Figure 8 presents the comparison of the predicted and measured droplet velocity and surface area for three different ambient pressures. In general, predictions of the second high-pressure model are in reasonably good agreement with experimental data over a wide range of pressures. As $p$ is increased, the computed and measured droplet velocity decrease, which can be attributed to the increased gas density at higher pressures. As indicated in Eq. (20), the increased gas density diminishes the gravitational force, but enhances the viscous drag force. Note the droplet Reynolds number increases as $p$ is increased, but $C_d$ is nearly independent of $\mathrm{Re}_d$ in the high $\mathrm{Re}_d$ limit. For $p=20$ atm, differences in the predicted and measured surface area are attributable to the underprediction of the droplet velocity, with the implication that an accurate calculation of droplet velocity and trajectory is critical for an accurate prediction of droplet vaporization history. At $p=40$ atm, the numerical model underpredicts the vaporization rate compared to the experimental data. This is indicative of the limit of the quasi-steady vaporization model at high pressures.

## Conclusions

In this paper, we have examined the multicomponent and high-pressure effects on the vaporization behavior of gas turbine fuels. The state-of-the-art vaporization models along with a detailed algorithm for the computation of liquid and gas-phase thermo-transport have been employed to examine whether the vaporization behavior of a multicomponent fuel droplet can be simulated by a surrogate pure fuel droplet. The vaporization history and other relevant properties of a bicomponent fuel droplet have been compared with those of an equivalent single-component droplet under a wide range of conditions that exist in a typical gas turbine combustor. In particular, three typical operating conditions have been considered: (1) lean blowout for ground-idle operation; (2) premixing, prevaporizing system at high-power conditions; and (3) startup (ignition and flame propagation) conditions. In addition, two high-pressure vaporization models within the quasi-steady framework have been investigated. The first model considers the effect of pressure on thermodynamic and transport properties, while the second model also considers the nonideal gas effects and dissolution of gases into the liquid. Important observations are as follows:

1 Under high-power conditions, the vaporization behavior of a gas turbine fuel is well represented by an equivalent single-component (50 percent boiling point) fuel. However, the comparison of the total vaporization rates of bicomponent and single-component fuels indicates significant differences in the predictions of the infinite-diffusion and diffusion-limit models. Using the diffusion-limit model, the vaporization rate of a bicomponent fuel droplet can be well represented by an equivalent single-component droplet for a wide range of conditions. In contrast, the use of the infinite-diffusion model indicates discrepancies between the vaporization rates of bicomponent and single-component fuel droplet. Since the total vaporization rate is a better indicator of how well a single-component fuel can represent the vaporization behavior of a multicomponent spray, it is recommended that a diffusion-limit model or a effective-diffusivity model be employed in the spray models. This conclusion also applies to the prediction of vaporization rates of gas turbine fuels for premixing, prevaporizing systems at high-power conditions.

2 Liquid and gas-phase properties change considerably during the droplet lifetime. Consequently, an accurate calculation of the thermo-transport properties should be an essential part of spray computations.

3 The droplet lifetime is relatively insensitive to pressure. However, the droplet heatup time becomes a more significant part of droplet lifetime at higher pressures. Consequently, differences between the predictions of the infinite-diffusion and diffusion-limit models become increasingly more noticeable at elevated pressures. The infinite-diffusion model predicts significantly higher vaporization rate compared to the diffusion-limit model for both single-component and bicomponent fuel droplets. For example, at $p=15$ atm, this model predicts 40 percent shorter droplet lifetime compared to the diffusion-limit model.

4 The representation of a bicomponent fuel droplet by an equivalent single-component becomes increasingly better at higher pressures. In fact, the predicted vaporization behavior is significantly more sensitive to the droplet heating model rather than to the liquid fuel composition. This can be attributed to a significant increase in droplet heatup time, and a reduction in the relative volatility differential between the constituent fuels at high pressure.

5 For ignition, LBO, and idle operation, the multicomponent fuel effects become relatively important, i.e., using a single-component droplet to represent multicomponent effects lead to unacceptable results, especially when the infinite-diffusion model is employed.

6 The predicted vaporization histories of n-heptane droplets using a quasi-steady high-pressure model show good agreement with the measured data over a wide range of pressures. The high-pressure model incorporates the nonideal gas behavior, dissolution of gases into the liquid, and dependence of thermo-transport properties on pressure. At $p=40$ atm, which is above the critical pressure of the fuel, the model underpredicts the vaporization rate compared to the experimental data, which is perhaps indicative of the high-pressure limit of our quasi-steady vaporization model.

## Acknowledgment

This work has been funded by the GE Aircraft Engines.

## References


[1] Westbrook, C. K., Pitz, W., and Warnatz, J., 1988, ''A Detailed Chemical Kinetic Reaction Mechanism for the Oxidation of Iso-Octane and n-Heptane over an Extended Temperature Range,'' Proc. Combust. Inst., **22**, pp. 893–902.

[2] Callahan, C. V., Held, T. J., Dryer, F. L., Minetti, R., Ribaucour, M., Sochet, L. R., Faravelli, T., Gaffuri, P., and Ranzi, E., 1996, ''Experimental Data and Kinetic Modeling of Primary Reference Fuel Mixtures,'' Proc. Combust. Inst., **26**, pp. 739–746.

[3] Newbold, F. R., and Amundson, N. R., 1973, ''A Model for Evaporation of a Multicomponent Droplet,'' AIChE J., **19**, pp. 22–30.

[4] Landis R. B., and Mills, A. F., 1974, ''Effects of Internal Resistance on the Vaporization of Binary Droplets,'' Fifth International Heat Transfer Conference, Paper B7-9, Tokyo, Japan.

[5] Law, C. K., 1982, ''Recent Advances in Droplet Vaporization and Combustion,'' Prog. Energy Combust. Sci., **8**, pp. 169–195.

[6] Tong, A. Y., and Sirignano, W. A., 1986, ''Multicomponent Droplet Vaporization in a High Temperature Gas,'' Combust. Flame, **66**, pp. 221–235.

[7] Abramzon, B., and Sirignano, W. A., 1989, ''Droplet Vaporization Models for Spray Combustion Calculations,'' **32**, No. 9, pp. 1605–1618.

[8] Chen, G., Aggarwal, S. K., Jackson, T. A., and Switzer, G. L., 1997, ''Experimental Study of Pure and Multicomponent Fuel Evaporation in a Heated Air Flow,'' Atomization Sprays, **7**, pp. 317–337.

[9] Aggarwal, S. K., 1987, ''Modeling of Multicomponent Fuel Spray Vaporization,'' Int. J. Heat Mass Transf., **30**, No. 9, pp. 1949–1961.

[10] Canada, G. S., and Faeth, G. M., 1974, ''Fuel Droplet Burning Rates at High Pressures,'' Proc. Combust. Inst., **14**, pp. 1345–1354.

[11] Nomura, H., Ujiie, Y., Rath, H. J., Sato, J., and Kono, M., 1996, ''Experimental Study on High-Pressure Droplet Evaporation Using Microgravity Conditions,'' Proc. Combust. Inst., **26**, pp. 1267–1273.

[12] Stengele, J., Willmann, M., and Wittig, S., 1997, ''Experimental and Theoretical Study of Droplet Vaporization in High Pressure Environment,'' ASME Paper 97-GT-151.

[13] Hsieh, K. C., Shuen, J. S., and Yang, V., 1991, ''Droplet Vaporization in High Pressure Environments: Near Critical Conditions,'' Combust. Sci. Technol., **76**, pp. 111–132.

[14] Curtis, E. W., and Farrel, P. W., 1992, ''A Numerical Study of High-Pressure Droplet Vaporization,'' Combust. Flame, **90**, pp. 85–102.

[15] Jia, J. D., and Gogos, G., 1993, ''High-Pressure Droplet Vaporization: Effects of Liquid-Phase Gas Solubility,'' Int. J. Heat Mass Transf., **36**, pp. 2403–2415.

[16] Stengele, J., Bauer, H.-J., and Wittig, S., 1996, ''Numerical Study of Bicomponent Droplet Vaporization in a High Pressure Environment,'' ASME Paper 96-GT-442.

[17] Zhu, G., and Aggarwal, S. K., 2000, ''Transient Supercritical Droplet Evaporation With Emphasis on the Effects of Equation of State,'' Int. J. Heat Mass Transf., **43**, No. 7, pp. 1157–1171.

[18] Gilver, S. D., and Abraham, J., 1996, ''Supercritical Droplet Vaporization and Combustion Studies,'' Prog. Energy Combust. Sci., **22**, pp. 1–28.

[19] Aggarwal, S. K., Shu, Z., Mongia, H., and Hura, H., 1998, ''Multicomponent Fuel Effects on the Vaporization of a Surrogate Single-Component Fuel Droplet,'' Paper 98-0157, 36th Aerospace Sciences Meeting, Reno, NV, Jan. 12–15.

[20] Reid, R. C., Prausnitz, J. M., and Polin, B. E., 1987, *The Properties of Gases and Liquids*, 4th Ed., McGraw-Hill, New York.

[21] Ruszalo, R., and Hallett, W. L. H., 1992, ''A Model for the Autoignition of Single Liquid Droplets at High Pressure,'' Combust. Sci. Technol., **86**, pp. 183–197.

[22] Benmekki, E. H., 1988, ''Fluid Phase Equilibria with Theoretical and Semi-Empirical Equation of State Model,'' Ph.D. thesis, University of Illinois at Chicago, Chicago, IL.

[23] Aggarwal, S. K., Shu, Z., Mongia, H., and Hura, H. S., 1998, ''Multicomponent and Single-Component Fuel Droplet Vaporization Under High Pressure Conditions,'' Paper 98-3933, 34th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, Cleveland, OH, July 13–15.

[24] Knapp, H., Doring, R., Oellrich, L., Plocker, U., and Prausnitz, J. M., 1982, ''Vapor-Liquid Equilibria for Mixtures of Low Boiling Substances,'' J. Chem. Eng. Data **VI** DECHEMA Frankfurt.

[25] Vargaftik, N. B., and Touloukian, Y. S., 1983, *Handbook of Physical Properties of Liquids and Gases: Pure Substances and Mixture*, 2nd Ed., Hemisphere, Washington, DC.

[26] Lide, R. David, 1994, *CRC Handbook of Chemistry and Physics 1913–1995*, 74th Ed., Chemical Rubber Publishing Company, Boca Raton, FL.

[27] Ho, C. Y., Liley, P. C., Makita, T., and Tanaka, Y, 1988, CINDAS, *Data Series on Material Properties Volume V-1: Properties of Inorganic and Organic Fluids*, Hemisphere, Washington, DC.

[28] Kee, R. J., Miller, J. A., and Warnatz, J., 1983, ''A Fortran Program Package for the Evaluation of Gas-Phase Viscosities, Conductivities, and Diffusion Coefficients,'' Sandia National Laboratories Report SAND83-8209.

[29] Aggarwal, S. K., 1989, ''Ignition Behavior of a Dilute Multicomponent Fuel Spray,'' Combust. Flame, **76**, pp. 5–15.

**K. Mathioudakis**
Associate Professor

**A. Stamatis**
Senior Research Associate

Laboratory of Thermal Turbomachines,
Mechanical Engineering Department,
National Technical University of Athens,
Iroon Polytechniou 9,
Athens 15773, Greece

**E. Bonataki**
Public Power Corporation,
Aristotelous 30-32,
Athens 10433, Greece

# Allocating the Causes of Performance Deterioration in Combined Cycle Gas Turbine Plants

*A method for defining which parts of a combined cycle gas turbine (CCGT) power plant are responsible for performance deviations is presented. When the overall performances deviate from their baseline values, application of the method allows the determination of the component(s) of the plant, responsible for this deviation. It is shown that simple differentiation approaches may lead to erroneous conclusions, because they do not reveal the nature of deviations for individual components. Contributions of individual components are then assessed by separating deviations due to permanent changes and deviations due to change of operating conditions. A generalized formulation is presented together with the way of implementing it. Test cases are given, to make clearer the ideas put forward in the proposed method.* [DOI: 10.1115/1.1426407]

## 1 Introduction

Combined cycle gas turbine (CCGT) power plants are being the type of plant that is mostly built today for power generation. The technology of the components has matured and the overall plant thermal efficiencies are much higher than the conventional steam turbine plant. The simple analysis demonstrating the net efficiency gain offered by this type of plant can be found in many textbooks.

The thermodynamics of different types of combined cycle plants have been presented in detail by different authors, as, for example, in the textbook of Horlock [1] or the paper of Crane [2]. Principles, design considerations, and further matters related to practical experience on the behavior of the components and off design operation are covered by Kehlhofer [3]. Recent advances and prospects have been reported by Horlock [4] and Frutchi [5].

Most of the analyses and works published today deal with design studies, focusing on the optimization for a new plant. Off-design behavior has been studied to a lesser degree (see, for example, Fantozzi and Desideri [6]). A few studies have appeared which focus more on problems related to the monitoring of the performance (e.g., Haub et al. [7] and Rickli [8]). An aspect which has received considerable attention in recent years is the assessment of the effects of component malfunctions on overall plant performance. Thermoeconomic analysis has been used as a tool for estimating the economic impact of component malfunctions (e.g., Stoppato and Lazzaretto [9], Arena and Borchielini [10], Valero et al. [11], Torres et al. [12] Lerch et al. [13], and Lazzaretto et al. [14]).

The present paper comes to cover an aspect, which is of practical interest to the users of CCGT plants: the determination of the causes of an observed deviation of performance from nominal values. A typical situation in which such a problem arises is the commissioning of a newly built CCGT plant. If the overall target performances are not achieved, it is of interest to identify how much of observed deviations is attributed to each of the plant components. This problem is of more interest for plants built from retrofitting existing components, for example, gas and steam turbines. Another situation is the case in which the performance deteriorates during service. It is again useful to estimate how much of the overall performance deviation is attributed to each plant component. Such information may be useful when deciding which components should be serviced or replaced. Although similar information can be provided by the thermoeconomic type of analyses, as mentioned above, the information provided by the method proposed here has the advantage of being quantitatively more precise, as highlighted in the discussion section later in the paper.

A method of evaluating the contributions of the plant components to overall performance deviations is presented. The theoretical analysis on which the method is based is given first, followed by a description of how the analysis can be implemented.

## 2 Definition of the Problem

The basic mathematical formulation will be first presented for a simple CCGT plant, without supplementary firing, consisting of the following three components: a gas turbine, a heat recovery steam generator (HRSG) and a steam turbine (see Fig. 1). The combination of an HRSG and steam turbine will be termed ''steam plant,'' and will be treated as one component in the analysis of the simple CCGT plant that follows. Once the features of

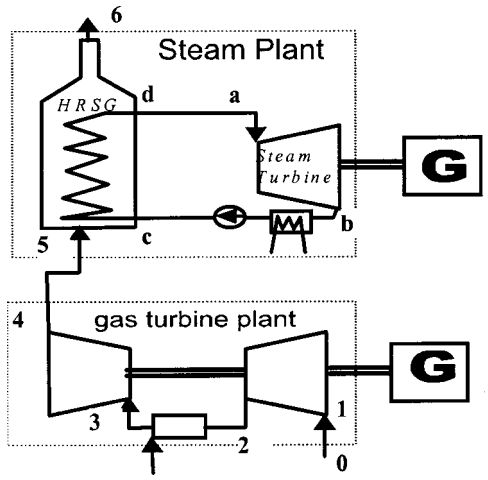



**Fig. 1 Layout of a simple combined-cycle gas turbine (CCGT) plant**

Contributed by the Turbomachinery Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Turbomachinery Division, October 2000; final revision, August 2001. Associate Editor: G. Sheard.

the method have been demonstrated, extensions to cover plants with multiple GTs and HRSGs will be discussed. The analysis is formulated in such a way that it can be implemented with data from measurements performed either permanently on the plant or during typical performance tests.

The problem which has to be solved is as follows: The overall performances of a CCGT plant, namely power output and heat rate or thermal efficiency, are evaluated from test data and they are found to be different from nominal values. (The case of most interest is when reduced performances occur.) The question is how much of the deviation is caused by alterations in each one of the components. Is an observed deviation caused by the gas turbine, the steam plant, or both, and in what proportions?

As will be shown later, even this may look a straightforward task, the responsibilities may be different from what a simple glance at individual component performances may indicate.

## 3 Relating Overall and Component Performance Deviations

Overall plant efficiency of a plant is related to component efficiencies at any operating point through the following relation, derived in the Appendix:

$$\eta_{CP} = \eta_{GT} + (1 - \eta_{GT} - \varepsilon_{GT})\,\eta_S . \tag{1}$$

Since overall efficiency is a function of gas and steam plant efficiencies, its deviations from nominal values can be caused by corresponding deviations of either gas or steam plant efficiencies. The deviations of overall efficiency can be linked to individual deviations, by differentiating Eq. (1):

$$d\eta_{CP} = \frac{\partial \eta_{CP}}{\partial \eta_{GT}} d\eta_{GT} + \frac{\partial \eta_{CP}}{\partial \eta_S} d\eta_S$$

$$= (1 - \eta_s) d\eta_{GT} + (1 - \eta_{GT}) d\eta_S . \tag{2}$$

(The term containing $\varepsilon_{GT}$ has been omitted, since it is of very small magnitude and its product with the small deviations gives a term of higher order). If deviations are small (which is the case in practical situations), differentials can be changed for differences, and this relation is transformed for fractional deviations as follows:

$$\Delta\eta_{CP} = \frac{(1 - \eta_s)\,\eta_{GT}}{\eta_{CP}} \Delta\eta_{GT} + \frac{(1 - \eta_{GT})\,\eta_S}{\eta_{CP}} \Delta\eta_S . \tag{3}$$

The first term represents the deviation of the overall efficiency, resulting from the fact that the gas turbine operates with an efficiency different by $\Delta\eta_{GT}$ from the value it would have if the plant was operating according to design. This deviation, however, in not necessarily caused by the gas turbine itself. It can be the result of a shift in its operating point to off-design, forced by a change in the condition of the steam plants that follows it. Similar arguments apply to the second term, which represents the steam plant.

It is thus noted that the differentiation gives the amounts of deviation related to the changes of efficiencies of the components, but does not show what the actual cause is. We will now see how the deviations caused by each component can be isolated.

**Estimation of Component Contributions.** The deviation in gas turbine efficiency $\Delta\eta_{GT}$ can be split into two terms: one representing a change in the condition of the gas turbine (e.g., deteriorated, fouled etc) and one representing a shift from its nominal operating point, at the nominal operating point of the entire plant. The first term will be denoted $\Delta\eta_{GT,D}$, indicating that it represents a permanent shift, present even at operation of the gas turbine at its design point. The second term will be denoted $\Delta\eta_{GT,OD}$ indicating a change due to the fact that the gas turbine is forced to operate at off-design conditions. We will thus have

$$\Delta\eta_{GT} = \Delta\eta_{GT,D} + \Delta\eta_{GT,OD} . \tag{4}$$

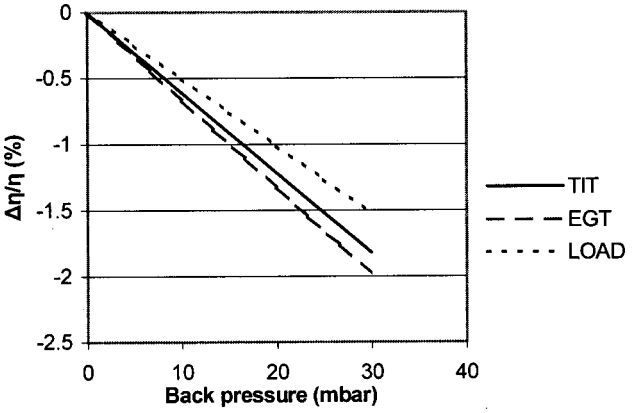



**Fig. 2 Example of influence of back pressure to the thermal efficiency of a gas turbine**

To make this distinction clear we give an example. Fouling of the compressor of a gas turbine will result in a drop of its thermal efficiency, which is present at any operating condition, and is thus termed $\Delta\eta_{GT,D}$. Fouling of the heat recovery steam generator, on the other hand, will force the gas turbine to operate at lower efficiency, due to increased back pressure. So, while the gas turbine has not suffered any deterioration and is at intact condition it operates at lower efficiency, simply because it is forced to operate at an off design point. In this second case, even though the drop in the overall plant efficiency will be caused by the drop in both gas turbine and steam plant efficiencies, the actual cause of the drop lies exclusively with the steam plant.

Backpressure is the parameter through which changes in the condition of an HRSG influences the performance of the gas turbine preceding it. Backpressure is determined by the loss factor of the HRSG and the dynamic head at its inlet. Possible variation in both should be taken into account in order to correctly estimate the components of efficiency deviation. On the other hand, the way backpressure influences efficiency is also related to the control logic of the gas turbine. For example, efficiency deviations for the same backpressure may be different if the engine operates at constant load or constant exhaust gas temperature. An example of how the thermal efficiency of a gas turbine changes with backpressure is shown in Fig. 2. Change is shown to be different for different modes of operation, namely if turbine inlet temperature, exhaust gas temperature, or turbine load is kept constant. The difference is nevertheless small.

The curves of this figure are obtained by running a performance model, in the different models of operation. The dependence shown here is typical for a gas turbine, although the slopes may change for specific turbines.

Similarly, steam plant efficiency can be expressed as the sum of two terms, one related to the physical condition of the steam plant, and one to movement of operating point:

$$\Delta\eta_S = \Delta\eta_{S,D} + \Delta\eta_{S,OD} . \tag{5}$$

Heat recovery steam generator effectiveness, and therefore steam plant efficiency, depends on mass flow rate and temperature of the flue gases entering the HRSG. An example of such a dependence is shown in Fig. 3. The curves have been obtained by running a computer model of the HRSG and varying the flue gas conditions. It is through these two parameters that gas turbine influences the performance of the subsequent HRSG.

It is noticed that sensitivity of effectiveness to inlet temperature is much larger than to mass flow rate. This feature presents an advantage in practical applications: mass flow measurement or estimation may be linked to relatively large uncertainty. The figure shows that such uncertainty will have a much smaller influence on the uncertainty of assessing effectiveness deviation.

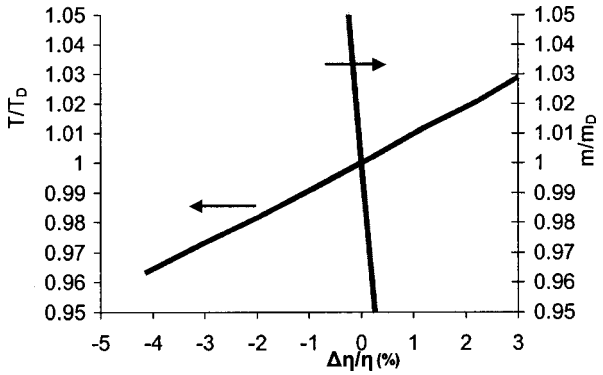



**Fig. 3 Example dependence of heat recovery effectiveness on flue gas mass flow rate and pressure**

Substituting now Eqs. (4) and (5) into Eq. (3) gives

$$\Delta\eta_{CP}=((1-\eta_s)\Delta\eta_{GT,D}+(1-\eta_S)\Delta\eta_{GT,OD})\,\eta_{GT}/\eta_{CP}$$
$$+((1-\eta_{GT})\Delta\eta_{S,OD}+(1-\eta_{GT})\Delta\eta_{S,D})\,\eta_S/\eta_{CP}.$$

We can now group the terms of this expression to reflect changes caused by individual components, according to the preceding discussion

$$(\Delta\eta_{CP})_{GT}=(1-\eta_s)\,\eta_{GT}/\eta_{CP}\Delta\eta_{GT,D}$$
$$+(1-n_{GT})\,\eta_S/\eta_{CP}\Delta\eta_{S,OD}$$

$$(\Delta\eta_{CP})_S=(1-\eta_s)\,\eta_{GT}/\eta_{CP}\Delta\eta_{GT,OD}+(1-\eta_{GT})\,\eta_S/\eta_{CP}\Delta\eta_{S,D}$$

which can be written in a matrix format

$$\begin{bmatrix}(\Delta\eta_{CP})_{GT}\\(\Delta\eta_{CP})_S\end{bmatrix}=\begin{bmatrix}\Delta\eta_{GT,D} & \Delta\eta_{S,OD}\\ \Delta\eta_{GT,OD} & \Delta\eta_{S,D}\end{bmatrix}\begin{bmatrix}(1-\eta_S)\,\eta_{GT}/\eta_{CP}\\(1-\eta_{GT})\,\eta_S/\eta_{CP}\end{bmatrix}. \tag{6}$$

**Generalization for More Components.** A generalization of this procedure is possible for a CCGT plant with any number of components. The influence of each component can be derived by taking partial derivatives of the relation linking overall efficiency to individual component efficiencies (for example, relation (*A*9) of the Appendix):

$$\Delta\eta_{CP}=\sum_i\frac{\eta_i}{\eta_{CP}}\frac{\partial\eta_{CP}}{\partial\eta_i}\Delta\eta_i. \tag{7}$$

The deviation in the efficiency of each component can then be considered the sum of individual deviations as follows:

$$\Delta\eta_i=\sum_j\Delta\eta_{ji}. \tag{8}$$

$\Delta\eta_{ji}$ expresses the deviations in efficiency of the $i$th component caused by the fact that it is forced to operate at off-design conditions by the $j$th component. $\Delta\eta_{ii}$ expresses deviation in the performance of a component in condition different from what it was designed (for example, it is deteriorated). It is present at any operating conditions and is thus called permanent deviation.

Using these relations, a matrix relation can then be written for the evaluation of component contributions.

$$\begin{bmatrix}(\Delta\eta_{CP})_1\\(\Delta\eta_{CP})_2\\\cdot\\\cdot\\\cdot\\(\Delta\eta_{CPO})_N\end{bmatrix}=\begin{bmatrix}\Delta\eta_{11} & \Delta\eta_{12} & \dots & \Delta\eta_{1N}\\ \Delta\eta_{21} & \Delta\eta_{22} & \dots & \Delta\eta_{01}\\ \cdot & & \cdot & \cdot\\ \Delta\eta_{N1} & \Delta\eta_{N2} & \dots & \Delta\eta_{NN}\end{bmatrix}\begin{bmatrix}\frac{\eta_1}{\eta_{CP}}\frac{\partial\eta_{CP}}{\partial\eta_1}\\ \frac{\eta_2}{\eta_{CP}}\frac{\partial\eta_{CP}}{\partial\eta_2}\\ \cdot\\ \cdot\\ \cdot\\ \frac{\eta_N}{\eta_{CP}}\frac{\partial\eta_{CP}}{\partial\eta_N}\end{bmatrix} \tag{9}$$

$(\Delta\eta_{CP})_i$ denotes the part of $\Delta\eta_{CP}$ caused by the $i$th component and the following relation holds:

$$\sum(\Delta\eta_{CP})_i=\Delta\eta_{CP}.$$

In the case of CCGT plants consisting of a number of gas turbines, HRSG combinations, operating in parallel and feeding a steam turbine, this matrix will be of a strong diagonal structure, with many off-diagonal elements equal to zero.

## 4 Evaluation of Deviations From Test Data

It will now be discussed how component contributions to overall plant performance deviations can be identified, when data from a performance test are available. The number of measurements needed will depend on the depth of analysis sought. We will take first the example of the simple combined cycle, to illustrate this point.

Overall and component efficiencies can be evaluated by measurement of three quantities: gas turbine output $P_{GT}$, steam turbine output $P_{ST}$ and heat input to the cycle $Q_f$ (which is evaluated from a measurement of fuel flow and the LHV of the fuel). Equation (*A*1) gives $\eta_{CP}$ and equation (*A*2) gives $\eta_{GT}$. Equation (1) gives $\eta_s$, provided that the value of $\varepsilon_{GT}$ is known. Therefore, measurement of these three quantities and knowledge of the nominal values makes it possible to evaluate $\Delta\eta_{CP}$, $\Delta\eta_{GT}$, $\Delta\eta_S$.

If the individual component contributions are to be allocated, the components of equations (4) and (5) have to be evaluated and additional measurements are necessary. Quantities characterizing the performance of the gas turbine and steam plant have to be measured: For the gas turbine, the turbine exit pressure has to be measured, in order to check if the HRSG imposes a back pressure different from the design one. Knowledge of this value allows the estimation of $\Delta\eta_{GT,OD,}$. For the steam plant, conditions at HRSG inlet have to be determined, namely mass flow and temperature. Temperature is measured directly, while mass flow can be evaluated from other measurements. This later evaluation can be done in a number of different ways. Alternative ways may be (a) heat balance of gas turbine, (b) heat balance of the HRSG for which steam flow and properties are measured as well as flue gas inlet and outlet temperatures, (c) from calibrated parts of the flow path, such as the compressor inlet duct.

Knowledge of these quantities and the efficiency sensitivity charts of figures 2 and 3 for the plant under study, allows determination of the off-design terms $\Delta\eta_{GT,OD}$ and $\Delta\eta_{S,OD}$. Having determined these terms and knowing already the overall deviations $\Delta\eta_{GT}$, $\Delta\eta_S$ allows determination of permanent deviation terms from equations (4) and (5). Equation (6) can then be applied to evaluate individual contributions.

In the more general case of a plant with a larger number of components, the steps to be followed are:

(a) Overall efficiency and component efficiencies are evaluated from test data. Subtracting from the nominal values gives $\Delta\eta_{CP}$ for the plant and $\Delta\eta_i$ for each component.

(b) Taking each component, the deviations of its efficiency due to conditions imposed by each one of the other components, $\Delta\eta_{ij}$ (i≠j), are evaluated. Having determined those, permanent efficiency deviation $\Delta\eta_{ii}$ is calculated, as the only unknown of equation (8) for the component.
(c) The values of the influence coefficients $\partial\eta_{CP}/\partial\eta_i$ are calculated by substituting efficiency values into the expressions for the partial derivatives.
(d) Equation (9) is applied for the calculation of component contributions.

Finally it should be commented that prior to application of any of the above procedures it is essential that measurement data are validated and it is ensured that measurement sensor errors are not present.

## 5 Application to Test Cases, Discussion

The use of the method will now be demonstrated through application to test cases. The cases have been set up by using plant simulation software, with parameters of realistic magnitude for the plant components. The method has successfully been applied to test data from operating plants, however the relevant data are of proprietary nature and cannot be included in the paper. The success of the predictions was demonstrated by inspections, which verified that the components pointed out by the method had actually suffered deterioration.

**Test Case 1: A Simple CCGT Plant.** A simple CCGT plant, of the layout of figure 1, has nominal performances as follows: $P_{GT}=80$ MW, $P_{ST}=42$ MW, heat rate 7975 kJ/kWh (which means that $\eta_{CP}=0.451$).

Performance tests have been conducted to the plant and the following performances, reduced to reference ambient conditions, were measured: $P_{GT}=77.7$ MW, $P_{STT}=42.1$ MW, heat rate 8106 kJ/kWh (which means that $\eta_{CP}=0.444$). The question is how much are the deviations $(\Delta\eta_{CP})_{GT}$, $(\Delta\eta_{CP})_S$.

First efficiency deviations of the gas turbine, steam plants and the entire plant are evaluated, by employing equations (*A*1), (*A*2), and (*A*9) and taking $\varepsilon_{GT}=0.02$. The resulting values are as follows:

| | Design | Current | Deltas (%) |
|---|---|---|---|
| $\eta_{CP}$ | 0.451 | 0.444 | $\Delta\eta_{CP}=-1.55\%$ |
| $\eta_{GT}$ | 0.296 | 0.288 | $\Delta\eta_{GT}=-2.7\%$ |
| $\eta_s$ | 0.227 | 0.225 | $\Delta\eta_S=-0.9\%$ |

Equation (6) will now be applied for estimation of component induced deviations. Additional measurements have to be used in order to quantify the efficiency deviation components.

*Gas Turbine.* Back pressure has been measured and found the same as for nominal conditions, therefore $\Delta\eta_{GT,OD}=0$, $\Delta\eta_{GT,D}=-2.7\%$,.

*Steam Plant.* The temperature of the flue gas existing the gas turbine has been measured 6°C above nominal, while flue gas mass flow rate is .3% lower than nominal. These deviations give an off-design operation of the steam plant, resulting into an increase of its efficiency by $\Delta\eta_{S,OD}=1.2\%$ (the increase is caused by increased HRSG effectiveness due to larger temperature at its entry). Applying now equation (5) and using the steam plant overall efficiency deviation, gives $\Delta\eta_{S,D}=-2.1\%$.

Using these values for the component deviations in equation (6) gives the contributions for the gas and steam turbine plants:

$$(\Delta\eta_{CP})_{GT}=-0.9\% \quad (\Delta\eta_{CP})_S=-0.6\%$$

It should be noted that if the influence of each component was accounted by simply taking the corresponding partial derivative, then the reduction due to the gas turbine should have been estimated to be −1.37% and due to the steam plant −0.18%. Such results would lead to the erroneous conclusion that the gas turbine is mostly responsible for the plant deficiency, while the actual situation is that steam plant deterioration contributes significantly to the overall plant deficiency. A schematic of the different deviations is presented in Fig. 4.

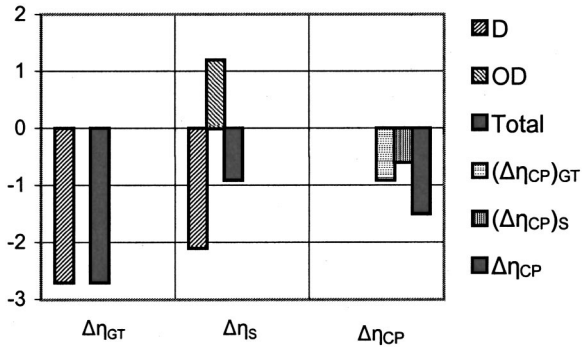



**Fig. 4 Deviations of component and overall plant efficiencies for a simple combined-cycle gas turbine (CCGT) plant**

Looking at the deviations of the steam plant, it is observed that its permanent deviation could pass unnoticed, as it is partially offset by the improvement caused by it driven to operate at a better efficiency (due to the increase is gas turbine outlet temperature). This observation shows that, the present approach is also useful as an improved diagnostic tool to identify malfunctions of the plant components.

**Test Case 2: A CCGT Plant With Two Gas Turbines.** The nominal performances of a CCGT plant with two gas turbines, with the layout shown in Fig. 5, are as follows: $P_{GT1}=P_{GT2}=80$ MW, $P_{ST}=84$ MW, heat rate 7975 kJ/kWh. Performance tests have been conducted and the following performances were measured: $P_{GT1}=77.7$ MW, $P_{GT2}=74.9$ MW, $P_{ST}=85.83$ MW, heat rate 8156 kJ/kWh. Measurements of gas turbine exhaust gas temperature, exhaust pressure, flue gas flow rate, HRSG gas inlet and outlet gas and steam temperatures as well as steam flow rate have been performed. The contribution of all five plant components to overall deficiency are to be evaluated.

Since the individual component contributions are now considered, the matrix relation (9) will be used. The component numbering for this plant is as shown in Fig. 5. The matrix relation for this particular plant takes the following form:

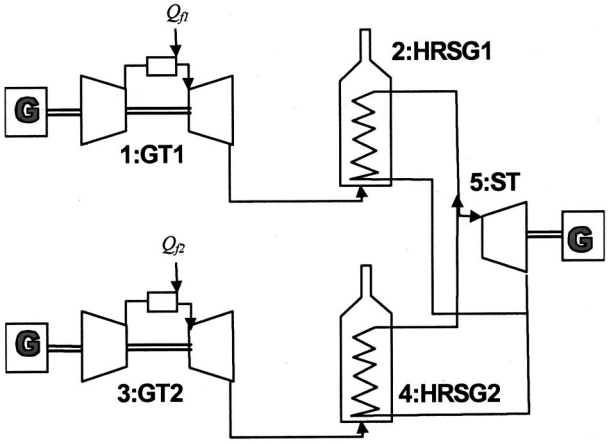



**Fig. 5 Layout of the combined power plant with two gas turbines**

$$\begin{bmatrix} (\Delta\eta_{CP})_{GT1} \\ (\Delta\eta_{CP})_{HR1} \\ (\Delta\eta_{CP})_{GT2} \\ (\Delta\eta_{CP})_{HR2} \\ (\Delta\eta_{CP})_{ST} \end{bmatrix} = \begin{bmatrix} \Delta\eta_{11} & \Delta\eta_{12} & 0 & 0 & \Delta\eta_{15} \\ \Delta\eta_{21} & \Delta\eta_{22} & 0 & 0 & \Delta\eta_{25} \\ 0 & 0 & \Delta\eta_{33}\Delta\eta_{34} & \Delta\eta_{35} & \\ 0 & 0 & \Delta\eta_{43}\Delta\eta_{44} & \Delta\eta_{45} & \\ 0 & 0 & 0 & 0 & \Delta\eta_{55} \end{bmatrix}$$

$$\times \begin{bmatrix} a_1(1-\eta_{HR1}\eta_{ST})\dfrac{\eta_{GT1}}{\eta_{CP}} \\ a_1(1-\eta_{GT1})\eta_{ST}\dfrac{\eta_{HR1}}{\eta_{CP}} \\ a_2(1-\eta_{HR2}\eta_{ST})\dfrac{\eta_{GT2}}{\eta_{CP}} \\ a_2(1-\eta_{GT2})\eta_{ST}\dfrac{\eta_{HR2}}{\eta_{CP}} \\ \sum_{i=1}^{2} a_i(1-\eta_{GTi})\eta_{HRi}\dfrac{\eta_{ST}}{\eta_{CP}} \end{bmatrix}$$

where $a_1=Q_{GT1}/Q_{GT1}+Q_{GT2}$, $a_2=Q_{GT2}/Q_{GT1}+Q_{GT2}$. In this formulation it has been assumed that the HRSG does not influence the steam turbine cycle performance, which is a good approximation when the feed steam properties are controlled.

The observed values of efficiencies of the plant and its components are as follows:

| | Design | Current | Deltas |
|---|---|---|---|
| $\eta_{CP}$ | 0.451 | 0.441 | $\Delta\eta_{CP}=-2.22\%$ |
| $\eta_{GT1}$ | 0.296 | 0.288 | $\Delta\eta_{GT1}=-2.7\%$ |
| $\eta_{HR1}$ | 0.611 | 0.606 | $\Delta\eta_{HP1}=-0.75\%$ |
| $\eta_{GT2}$ | 0.296 | 0.277 | $\Delta\eta_{GT2}=-6.4\%$ |
| $\eta_{HR2}$ | 0.611 | 0.618 | $\Delta\eta_{HR2}=1.14\%$ |
| $\eta_{STT}$ | 0.361 | 0.361 | $\Delta\eta_{STT}=0.0$ |

Efficiency deviations for each one of the components are evaluated in the same way as in Test Case 1. The previous matrix relation is then applied and the following contributions to overall efficiency deviation are calculated:

$$(\Delta\eta_{CP})_{GT1}=-0.46, \quad (\Delta\eta_{CP})_{HR1}=-0.32,$$

$$(\Delta\eta_{CP})_{GT2}=-0.90, \quad (\Delta\eta_{CP})_{HR2}=-0.54, \quad (\Delta\eta_{CP})_{ST}=0.0$$

It should be noted that in this case too the picture is quite different from what would have been deduced by simply taking partial derivatives. It can be easily verified that the HRSG2 would appear to have a positive contribution to efficiency, while its actual contribution is clearly negative. In this case too although HRSG2 was deteriorated, it only appeared to perform better because the gas turbine preceding it, being deteriorated also, fed it with hotter flue gases. This kind of situation would have remained unnoticed, a fact which could have negative consequences: if the performance of only the gas turbine was restored (by overhaul, for example), it would then be discovered that this does not lead to recovery of overall plant efficiency.

**Discussion.** As mentioned already, in order to identify component contributions, apart from the overall performance measurements, additional ones are needed, in order to characterize the performance of each component. It is useful at this point to comment on the fact that deviations $\Delta\eta_{GT,OD}$ are not expected to be very large, because the sensitivity of the gas turbine to back pressure (Fig. 2) is relatively small, for the amount of changes expected to occur in practice during operation. On the other hand, a quantity which can be significant is $\Delta\eta_{HR,OD}$, due to its increased sensitivity to EGT variations.

Application of the method requires also the existence of performance data for each component, and in particular data on how it responds to changes of certain quantities. Such data are usually provided by the manufacturer, in the form of correction curves (see, for example, Cloyd and Harris [15]). For example the curve of Fig. 2, describing sensitivity of gas turbine efficiency to backpressure, is typically provided with the set of test data correction curves of a newly commissioned gas turbine. Plant simulation software, capable of adapting to particular components and reliably representing their operation, can also be used to support application of the method.

The method presented here can be used for allocation of changes observed during operation, but also for evaluating the initial performance of a plant, namely to check whether the design targets have been achieved and if not, what is the reason. Situations in which existing old components are retrofitted for building a new plant may be cases where such evaluations can be useful. When the performance of components are evaluated, it may be necessary to take into account constraints existing on the design of components. For example, efficiency gains from operation of steam plants fed with flue gases hotter than designed, may have to be restricted due to limits in the capacity of the steam generators or the steam turbine.

The present method uses a formalism similar to techniques known as linearized Gas Path Analysis (GPA) methods, for detection of gas turbine engine faults. Typical such methods are described by Volponi and Urban [16] while a review of existing techniques has been given by Simon [17]. Although these methods too employ the concept of small deviations, they deal with ways to solve a linear system interrelating measurement deviations and component parameter deviations. For the problem studied here, they would provide $\Delta\eta_{GT}, \Delta\eta_S$ from measurements data. This approach is not essential in the present case, because the relations between measurements and parameters are simple enough to be applied directly. The component parameter deviations determined by GPA, on the other hand, are assumed to result from faults or deterioration. Changes due to mismatching are considered to be negligible, an assumption fulfilled with very good accuracy for the components of a gas turbine. On the other hand, the originality of the present method is that it goes one step further, as it permits isolation of deviations due to component alteration and estimation of the actual amount each component contributes to an overall deviation.

Some of the notions used in the present work have also been used by the works performing thermoeconomic diagnosis (Arena and Borchielini [10], Valero et al. [11], Torres et al. [12], Lerch et al. [13], Stoppato and Lazzaretto [9], and Lazzaretto et al. [14]). The idea that individual component malfunction influences overall performance not only through the change of component parameters but also though the ''dysfunction'' it causes on other components performance, has been employed. These methods focus, however, mainly onto the ''direct'' problem, namely the assessment of economic impact of individual malfunctions on the performance of a thermal system. Although criteria for allocating malfunctions to components are provided, they are not quantitative. For example Stoppato and Lazzaretto [9] state that their analysis offers indicative suggestions about the causes of malfunctions, while Lazzaretto et al. [14] mention that their analysis provides the component ''probably guilty'' for a malfunction. On the other hand Valero et al. [11] conclude that given the data of performances of a plant, thermoeconomic analysis in not sufficient to find precisely the causes of deviation. The originality of the present method is that it provides exactly this possibility. This is achieved by using data in addition to the overall performances, and in particular quantities interrelating the performance of individual components (e.g. gas turbine exhaust gas temperature, back pressure etc). On the other hand, the specific component performance data needed are identified (e.g. the data of Figs. 2, 3). The present method provides thus a quantitative estimation of causes

of plant malfunctions, going a step deeper than the above mentioned analyses. This quantification is useful, as it can serve as a basis for precise estimation of the outcome of corrective actions, while it can serve as a basis for cost allocation in cases of estimating the responsibilities of individual components on an overall plant observed inefficiency.

## 6 Conclusions

A method for diagnosing the sources of deficiencies in combined cycle gas turbine plant has been presented. The principles of the method have been introduced by examining a simple plant comprising a single gas turbine, while a generalized representation was derived for plants with more than one gas turbines.

Implementation of the method on the basis of test data was discussed and measurement needs for its application were identified. Component performance sensitivity data were also identified as necessary for implementing the method, and the way of obtaining them was discussed.

The contribution of plant components to overall plant performance deviations have been analyzed, and it was shown that the actual effect of a change in a component is not only directly through its efficiency, but also through the way it forces neighboring components to operate at conditions out of design. It was demonstrated that the relative magnitude of the causes of deficiency can be significantly different from what observation of deviations simply evaluated by differentiation indicates.

It was shown that evaluating component contribution to overall plant deficiency by the proposed method provides information useful, if corrective action is to be taken for restoring plant performance.

## Appendix

### Derivation of Relationships Between Efficiencies

*Performance Parameters Definitions.* The efficiencies used by the present analysis for a plant as well as its components are defined in the following:
Overall Plant Efficiency:

$$\eta_{CP}=(W_{GT}+W_{ST})/Q_f \tag{A1}$$

Gas Turbine:

$$\eta_{GT}=W_{GT}/Q_f \tag{A2}$$

The heat available to the HRSG is

$$H_4-H_0=Q_f-Q_{GT\,\mathrm{loss}}-W_{GT}=Q_f(1-\varepsilon_{GT})-W_{GT}.$$

A coefficient for heat loss between gas turbine exit and HRSG inlet is also defined:

$$\varepsilon_{\mathrm{CON}}=(H_4-H_5)/(H_4-H_0)=1-(H_5-H_0)/(H_4-H_0). \tag{A3}$$

The efficiency of the HRSG is expressed through effectiveness, defined as follows:

$$\eta_{HR}=(H_5-H_6)/(H_5-H_0). \tag{A4}$$

This quantity expresses the amount of heat effectively used as a fraction of the available heat, which is the heat that could be recovered if the gases were cooled to ambient temperature.

Heat losses in the HRSG (e.g., radiation to the environment, leaks) are expressed as a fraction of the heat absorbed from the flue gases:

$$\varepsilon_{HR}=1-(H_d-H_c)/(H_5-H_6). \tag{A5}$$

Steam Turbine Cycle:

$$\eta_{ST}=W_{ST}/(H_d-H_c) \tag{A6}$$

The combination of HRSG and steam turbine (“steam plant”) can be characterized by means of the steam plant efficiency $\eta_S$:

$$\eta_S=W_{ST}/(H_5-H_0). \tag{A7}$$

It is easily verified that

$$\eta_S=\eta_{HR}\eta_{ST}(1-\varepsilon_{HR}). \tag{A8}$$

*Derivation of Efficiency Interrelations.* In order to derive the relation of the overall plant efficiency to the efficiencies of its components we express the first law in the form of heat balances of the components, and apply the definitions of the relevant quantities. The heat given to the steam, evaluated in this way, is

$$Q_{ST,IN}=\eta_{HR}(1-\eta_{GT}-\varepsilon_{GT})(1-\varepsilon_{\mathrm{CON}})(1-\varepsilon_{HR})Q_f.$$

By using the definitions of efficiency of the steam turbine cycle and the overall efficiency we obtain

$$\eta_{CP}=\eta_{GT}+(1-\eta_{GT}-\varepsilon_{GT})(1-\varepsilon_{\mathrm{CON}})(1-\varepsilon_{HR})\,\eta_{HR}\eta_{ST}.$$

The connecting ducts losses are very small and can be considered to be included in the HRSG heat losses. Substituting then from Eq. (*A*8) for steam plant efficiency gives Eq. (1) of the main text.

*Generalization for More Than One Gas Turbines.* The foregoing analysis can be extended for plants comprising more than one gas turbine, each one connected to a heat recovery steam generator and all of the steam generators feeding one steam turbine. Using the same approach, the following relation can then be derived for overall efficiency:

$$\eta_{CP}=\sum_i\,(\eta_{GTi}+(1-\eta_{GTi}-\varepsilon_{GTi})\,\eta_{HRi}\eta_{ST})\,\frac{Q_{f_i}}{Q_f}. \tag{A9}$$

$i$ indicates the branch consisting of a gas turbine and the HRSG that follows it. $Q_{fi}$ is the heat input to the $i$th gas turbine and $Q_f$ the total heat input to the plant $Q_f=\sum_i Q_{fi}$.

## Nomenclature

| | | |
|---|---|---|
| $m$ | = | mass flow rate |
| $h$, $H$ | = | specific enthalpy, enthalpy |
| HRSG | = | heat recovery steam generator |
| LHV | = | fuel lower heating value |
| $P$ | = | power output |
| $Qf$ | = | heat input to the plant |
| $Q_{\mathrm{loss}}$ | = | parasitic losses, such as casing radiation and convection, mechanical losses to the lube oil |
| $W$ | = | work output of a turbine |
| $\varepsilon$ | = | heat loss coefficient |
| $\eta$ | = | efficiency of a component |
| $\Delta\eta$ | = | efficiency percentage deviation, $\Delta\eta=100\times(\eta'-\eta)/\eta$ |

### Subscripts

| | | |
|---|---|---|
| 0,1,2,3,4,5 | = | station along the gas flow path, Fig. 1 |
| $a$, $b$, $c$, $d$ | = | station along the steam flow path, Fig. 1 |
| $CP$ | = | quantity referred to the overall combined plant |
| $D$ | = | design |
| $f$ | = | fuel |
| $g$ | = | gas |
| $GT$ | = | quantity referred to a gas turbine |
| $HR$ | = | quantity referred to an HRSG |
| $i$, $j$ | = | index of plant component |
| $OD$ | = | off-design |
| $S$ | = | quantity referred to steam plant |
| $ST$ | = | quantity referred to steam turbine |

## References


[1] Horlock, J. H., 1992, *Combined Power Plants*, Pergamon Press, Oxford, UK.
[2] Crane, R. I., 1993, *Combined Cycles for Power Plants: Thermodynamics of Combined Cycle Plants*, Von Karman Institute, Lecture Series 1993-08.
[3] Kehlhofer, R., 1999, *Combined Cycle Gas and Steam Turbine Power Plants*, 2nd Ed., PennWell, Tulsa, OK, Aug.

[4] Horlock, J. H., 1995, ''Combined Power Plants—Past, Present, and Future,'' ASME J. Eng. Gas Turbines Power, **117**, pp. 608–616.
[5] Frutschi, H. U., 1999, ''Highest Efficiencies for Electrical Power Generation With Combined Cycle Plants,'' ABB Rev., **3**, pp. 12–18.
[6] Fantozi, F., and Desideri, U., 1998, ''Simulation of Power Plant Transients With Artificial Neural Networks: Application to an Existing Combined Cycle,'' Proc. Inst. Mech. Eng. A, J. Power, **212**, pp. 299–313.
[7] Haub, G. L., Hauhe, W. E., Jr., Myers, C. O., 1989, ''Real Time Performance Analysis and Condition Monitoring of a Gas Turbine Based 300 MwMW Base Load Cogeneration Plant,'' ASME paper 89-GT-28.
[8] Rickli, J. P., 1993, *Combined Cycles for Power Plants: Controls, Monitoring and Performance Measurements*, Von Karman Institute, Lecture Series 1993-08.
[9] Stoppato, A., and Lazzaretto, A., 1996, ''The Exergetic Analysis for Energy System Diagnosis,'' *ASME Engineering Systems Design and Analysis Conference*, Vol. 1, July 1–4, Montpelier, France, ASME, New York, pp. 191–198.
[10] Arena, A. P., and Borchielini, R., 1999, ''Application of Different Productive Structures for Thermoeconomic Diagnosis of an Acombined Cycle power plant.'' Int. J. Thermal Sci., **38**, pp. 601–602.
[11] Valero, A., Torres, C., and Lerch, F., 1999, *Structural Theory and Thermoeconomic Diagnosis. Part III: Intrinsic and Induced Malfunctions*, ECOS'99, June 8–10, Tokyo, pp. 35–41.
[12] Torres, C., Valero, A., Serra, L., and Royo, J., 1999, *Structural Theory and Thermoeconomic Diagnosis. Part I: On Malfunctions and Dysfunction Analysis*, ECOS'99, June 8–10, Tokyo, pp. 368–373.
[13] Lerch, F., Roys, J., and Serra, L., 1999, *Structural Theory and Thermoeconomic Diagnosis. Part II: Application to an Actual Power Plant*, ECOS'99, June 8–10, Tokyo, pp. 374–379.
[14] Lazzaretto, A., Macor, A., Mirandola, A., and Stoppato, A., 1999, *Proceedings of the ASME Advanced Energy Systems Division*, AES-Vol 39, ASME, New York. pp. 611–619.
[15] Cloyd, S. T., Harris, A. J., 1995, Gas Turbine Performance—New Application and Test Correction Curves, ASME Paper 95-GT-167.
[16] Volponi, A. J., and Urban, L. A., 1992, ''Mathematical Methods for Relative Engine Performance Diagnostics,'' SAE Trans., **101**, Paper No. 922048.
[17] Simon, D. L., 2000, ''An Overview of the NASA Aviation Safety Program Propulsion Health Monitoring Element,'' Paper No. AIAA-2000-3624.

# A Recuperated Gas Turbine Incorporating External Heat Sources in the Combined Gas-Steam Cycle

**R. Chodkiewicz**
e-mail: rchod@imph.p.lodz.pl

**J. Krysinski**

**J. Porochnicki**

Technical University of Lodz,
Institute of Turbomachinery,
Wolczanska 219/223,
93-005 Lodz, Poland

*The recuperation by means of external waste heat sources, as opposed to the recuperation of the turbine exhaust gases (to preheat the compressed air), allows one to utilize the hot exhaust gases of the gas turbine in the bottoming steam cycle to produce steam in order to generate additional power. Such a combined gas/steam energy system, closely integrated with the industrial process, can produce electric power (and useful heat) with high efficiency and very low atmospheric air pollution. In the present paper two examples of applications of this new technology have been analyzed from the economic and ecological viewpoint.* [DOI: 10.1115/1.1448325]

## Introduction

In many industrial processes significant amounts of waste heat are liberated, often with tail or waste gases (combustible or incombustible). The waste heat carrier is in most cases the water steam, generated in waste heat boilers and then utilized in steam turbines to produce the mechanical or electric power for a compressor or a generator drive. The tail gas enthalpy is utilized for power generation in gas expanders, which are used as supplementing power units in steam power systems. Industrial power plants with mostly low steam conditions and high temperature differences between hot process or tail gases and steam produce great exergy losses, which are the reason of low overall efficiency of the steam power cycle. An introduction of a recuperated gas turbine fired with natural gas into all industrial processes with waste heat or waste gas production can enhance the overall system efficiency of industrial power plants, and contribute to minimize the atmospheric air pollution. The recuperation by means of external heat sources (to preheat the compressed air before the combustor) gives an important advantage that allows one to utilize the hot exhaust gases of the gas turbine to produce steam in order to generate additional power. That leads to a combined cycle power plant of the specific kind of gas turbine integration with the production process. The integration can spread from the preheating compressed air for the combustor, through the compressor air discharge for process purposes, to hot tail gas injection into the gas turbine. The compressor cooling can also be necessary for better adjustment of the compressed air temperature to the waste heat source temperature.

The external heat recuperation can also be applied in such energy systems, where a cheap or low BTU waste fuel is available. In such systems the external recuperator converts into a waste fuel burning air preheater for the gas turbine combustor fired with high BTU fuel.

For the economic and ecological analysis of the new power systems with a recuperated gas turbine, two industrial processes are chosen. The first one is the catalyst regeneration process of the catalytic cracking of crude oil, i.e. a process with very high $CO_2$ emission and nearly without $NO_x$ discharge. The nitric acid production, as the second one under discussion is characterized by $CO_2$ free exhaust gas, and by significant $NO_x$ emission.

In the paper a theoretical analysis of these application examples of the new technology will be considered on the basis of the combined cycle calculation results.

## Recuperated Gas Turbine Process and Efficiency Definitions

A simplified gas turbine process without heat recuperation (a simple cycle) and with heat recuperation, both with an internal source (exhaust gases from a turbine), as well as with an external source, is shown on the entropy diagram, Fig. 1.

As can be seen in Fig. 1, an introduction of heat recuperation in the gas turbine process brings about (on the assumption of the constant cycle pressure ratio $\pi = p_2/p_1$, and the constant temperature ratio $\Theta = T_3/T_0$):

- a decrease in fuel consumption of the gas turbine owing to an increased air temperature upstream the combustion chamber in comparison with a simple cycle;
- a decrease in the gas turbine power resulting from additional flow resistance in air and gas paths leading to a decrease in the turbine pressure ratio $\pi_T = (1-\varepsilon)\pi$;
- an increase in the turbine exhaust gas temperature ($T_{4RZ} > T_4, T_{4RW} > T_4$).

It can also be seen (Fig. 1) that, as opposed to the recuperation of the turbine exhaust gases, the recuperation with external heat sources (to preheat the compressed air before the combustor) allows one to utilize fully the exhaust gases in the bottoming steam cycle with a steam turbine. In such a combined gas/steam process with two different heat sources as fuel energy and process waste heat, the effectiveness of electric power generation must be adequately defined, namely

- gas turbine fuel efficiency, understood as the effective thermal efficiency on the generator clamps

$$\eta_F \equiv \eta_{the} = P_{el\ GT}/\dot{Q}_F; \tag{1}$$

- combined gas/steam cycle fuel efficiency

$$\eta_{F\ \mathrm{COMB}} = (P_{e1\ GT} + P_{e1\ ST})/\dot{Q}_F; \tag{2}$$

- overall thermal efficiency of the energy system

$$\eta_0 \equiv \eta_{the0} = (P_{e1\ GT} + P_{e1\ ST} + P_{e1\ EX})/(\dot{Q}_F + \dot{Q}_P), \tag{3}$$

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-593. Manuscript received by IGTI, November 1999; final revision received by ASME Headquarters, February 2000. Associate Editor: D. Wisler.

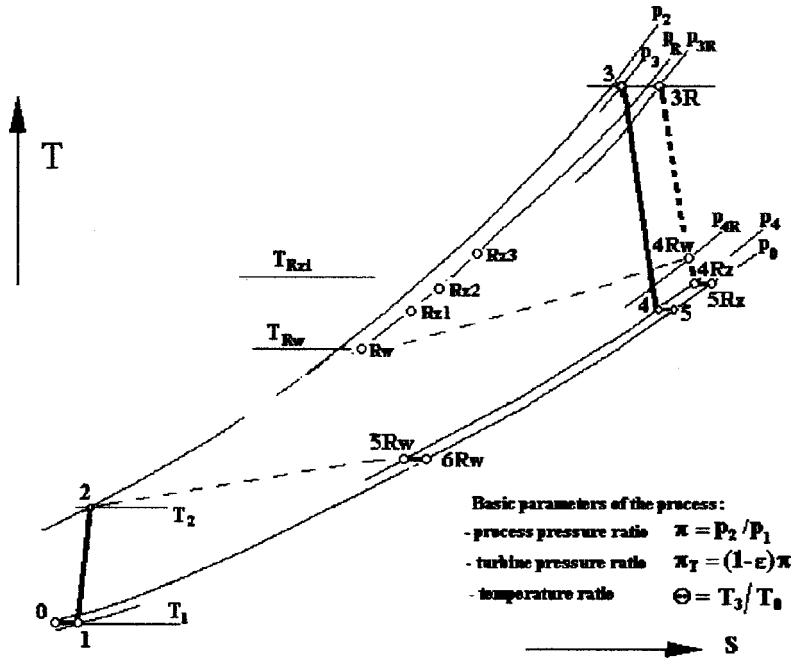



**Fig. 1 Entropy diagram of gas turbine processes. Simple cycle: 0-1-2-3-4-5; cycle with heat recuperation with turbine exhaust gases: 0-1-2-Rw-3R-4Rw-5Rw-6Rw; cycle with heat recuperation with an external source 0-1-2-Rzi-3R-4Rz-5Rz, $i=1, \ldots -n$.**

when in the energy system, besides gas and steam turbines, also gas expansion turbines (expanders) are used, and $\dot{Q}_p$ is the total waste heat output from the industrial process.

## Recuperated Gas Turbine in Fluid Catalytic Cracking Process (FCC)

**Conventional Power Systems in FCC Process.** The fluid catalytic cracking process is widely used in petroleum refining for the conversion of oil feedstocks into lighter hydrocarbons. The key element in that process is the catalyst, which breaks down the heavy oil molecules by stripping away some of the carbon atoms. The accumulated deposit of carbon atoms (coke) is burned off the catalyst in a regenerator producing high temperature gas. The regeneration process requires a substantial amount of air to support the combustion of the coke and fluidize the bed in the regenerator.

In the conventional FCC process with power recovery (Fig. 2), the hot gas enthalpy is utilized for power generation in expansion turbines (expanders). The steam for the condensing steam turbine is produced in the waste heat boiler, from the remainder part of the recoverable thermal energy.

The motor-generator included in the machine train produces the electricity and supplies the compressor power requirements.

The results of calculations of the conventional energy system in the FCC process (with power recovery, Fig. 2) are presented in Table 1.

The conventional energy system in a FCC plant is characterized by low overall efficiency (about 25 percent depending on the process pressure level), and very high carbon dioxide emission.

At such surplus power level, the economic advantages provided by the power recovery systems are difficult to ignore for the petroleum refiner.

**Recuperated Gas Turbine Set Integrated With FCC Process** The overall thermal efficiency of the FCC energy system can be significantly enhanced if the gas turbine with external process heat recuperation is applied (Fig. 3). A reduction of the relative $CO_2$ emission is inherently connected with efficiency growth.

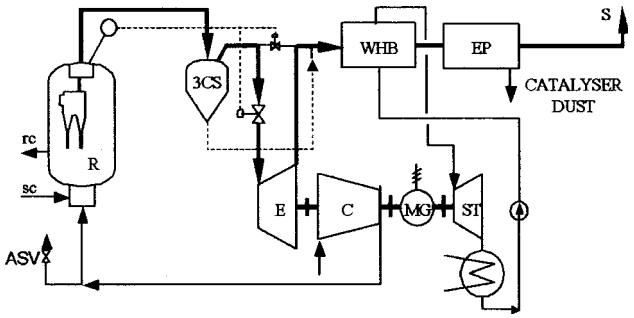



**Fig. 2 Combined gas-steam cycle utilizing the catalyst regeneration energy ([5]). E–expansion turbine, ST–steam turbine, MG-motor-generator, WHB–waste heat steam boiler, sc–spent catalyst, rc–regenerated catalyst, R–Regenerator, 3CS–third cyclone stage, EP–electrostatic precipitator, S–stack.**

**Table 1 The results of conventional FCC process calculations**

| Variant | 1 | 2 | 3 |
|---|---|---|---|
| Process heat flux $Q_p$ (MW) | 30.44 | 37.46 | 49.17 |
| Expander power $P_{ex}$ (MW) | 11.57 | 11.81 | 11.92 |
| Steam turbine power $P_{ST}$ (MW) | 5.280 | 7.576 | 11.54 |
| Compressor power $P_C$ (MW) | 10.06 | 9.966 | 9.604 |
| Electric power $P_{el}$ (MW) | 6.584 | 9.138 | 13.43 |
| Relative $CO_2$ emission (Mg/MWh) | | 11.5 | |
| Process efficiency (percent) | 21.9 | 24.4 | 28.16 |
| Expander inlet temperature (°C) | 730.0 | 730.0 | 730.0 |
| Expander pressure ratio | 4.8 | 3.6 | 2.6 |
| Air compressor pressure ratio | 5.65 | 4.235 | 3.059 |
| Expander mass flow (kg/s) | 39.0 | 48.0 | 63.0 |
| Process air mass flow (kg/s) | 44.1 | 54.2 | 71.2 |
| Expander polytropic efficiency (percent) | 81.1 | 80.0 | 80.0 |
| Compressor polytropic efficiency (percent) | 86.0 | 85.0 | 85.0 |
| Pressure drop factor | 0.15 | 0.15 | 0.15 |
| Steam mass flow (kg/s) | 5.4 | 7.75 | 11.8 |
| Live steam pressure (bar) | 40.0 | 40.0 | 40.0 |
| Live steam temperature (°C) | 430.0 | 430.0 | 430.0 |
| Process gas inlet temperature (upstream the waste heat boiler) (°C) | 504.3 | 531.0 | 587.3 |

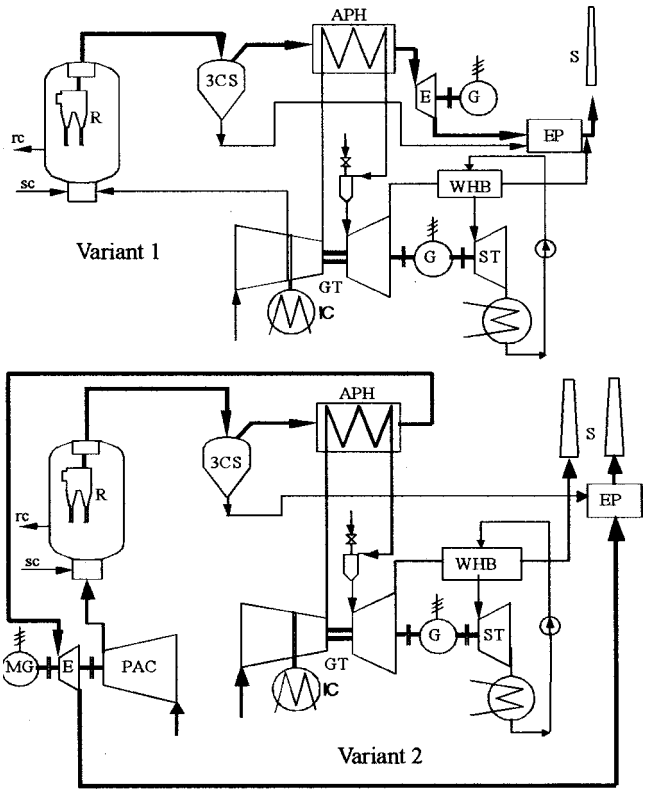



**Fig. 3 Combined gas-steam cycle using catalyst regeneration energy: Variant 1–strictly integrated GT with FCC process, Variant 2–with conventional machine train (without ST) parallel to the recuperated GT. PAC–process air compressor, E–expander, MG–motor-generator, GT–gas turbine, APH–air preheater (recuperator), ST–steam turbine, G–generator, WHB–waste heat steam boiler, IC–intersection cooler, sc–spent catalyst, rc–regenerated catalyst, R–regenerator, 3RD–third cyclone stage, EP–electrostatic precipitator, S–stack.**

The recuperated gas turbine strictly incorporated into the FCC process should provide

- process air extraction from the gas turbine compressor,
- utilization of the process heat to preheat the compressed air before the combustor,
- utilization of the exhaust gases heat for the steam generation to produce the additional electric power,
- introduction of the intersection cooler for better process heat utilization.

Computations of both the FCC energy systems (Fig. 3) have been made on the assumptions presented in Table 2.

For the pollution calculations, the following assumptions have been made:

**Table 2 Gas turbine data used in the calculations (installation and process machine parameters—the same as in Table 1)**

| | Unit | Value |
|---|---|---|
| Turbine polytropic efficiency (GT set) $\eta_{pt}$ | (%) | 88 |
| Compressor polytropic efficiency (GT set) $\eta_{pC}$ | (%) | 90 |
| Gas turbine inlet pressure drop | | 0.030 |
| Compressor cooler pressure drop | | 0.035 |
| Combustion chamber pressure drop | | 0.065 |
| Recuperator pressure drop | | 0.040 |

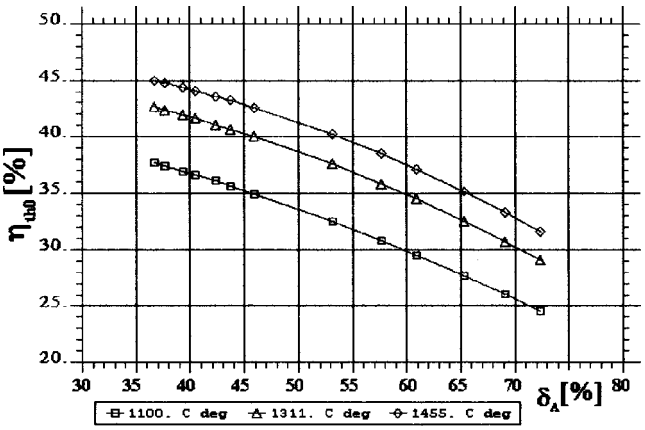



**Fig. 4 Overall thermal efficiency of the advanced FCC energy system versus relative process air extraction ($T_3$=1100°C, 1311°C, 1455°C; $\pi$=20)**

- the sulfur and its compounds are removed during earlier phases of the technological process and do not deposit on the catalyst grains,
- in the fluidized bed the burning temperatures are low enough so that the $NO_x$ production can be omitted in the calculations of that process phase,
- the $NO_x$ emission in the gas turbine does not depend on the burning temperature and is equal to $50 \cdot 10^{-6}$ kg/MJ, which corresponds to 25 ppm, and
- LHV-es for the gas and carbon are equal to 35000kJ/Nm$^3$ and 4120 kJ/kg (necessary for the $CO_2$ emission calculations).

Relative extraction mass flows differ in both the cases under examination. For the gas turbine set with process air extraction this value is defined as a rate of the process air mass flow and the gas turbine compressor inlet mass flow:

$$\delta_A = \dot{m}_A / \dot{m}_V . \tag{4}$$

For the gas turbine set without process air extraction the process air mass flow is divided by a sum of both the compressor mass flows (process air compressor and gas turbine compressor):

$$\delta_A = \dot{m}_A / (\dot{m}_A + \dot{m}_V). \tag{5}$$

The overall thermal efficiencies and fuel efficiencies for gas turbine set with process air extraction and waste heat utilization combined with steam cycle (variant 1) are presented in Figs. 4 and

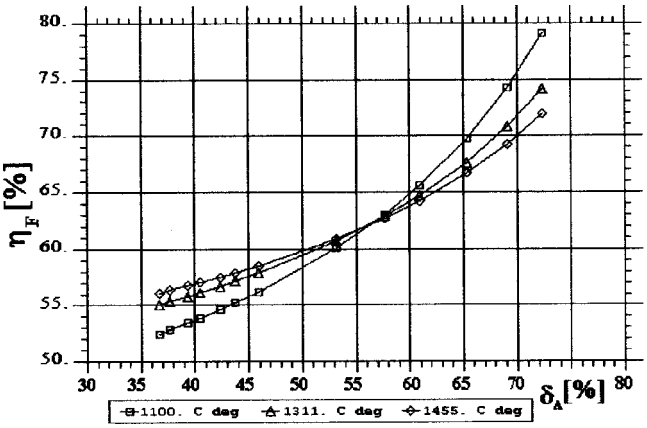



**Fig. 5 Combined gas/steam cycle fuel efficiency of the advanced FCC energy system versus relative process air extraction ($T_3$=1100°C, 1311°C, 1455°C; $\pi$=20)**

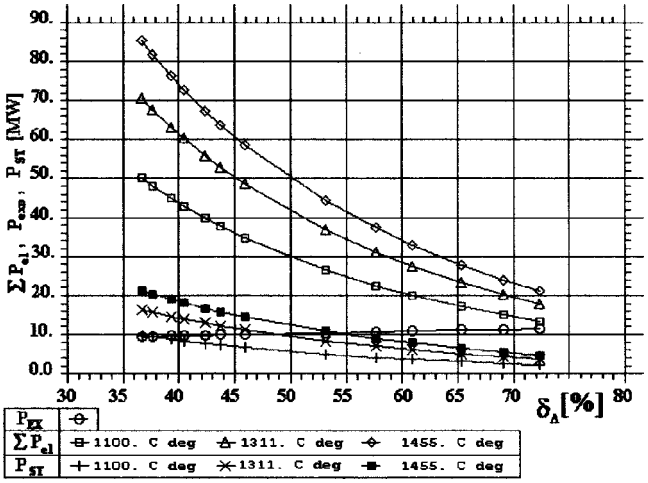



**Fig. 6 Electric power of the combined gas/steam FCC energy system, expander power and steam turbine power versus relative process air extraction ($T_3$=1100°C, 1311°C, 1455°C; $\pi$=20)**

5 as a function of relative process air extraction from the gas turbine compressor. In Fig. 6 the relation between the electric power and relative process air extraction is shown. In this figure the overall electric power of the energy system (the sum of GT, ST, and EX output) and the steam turbine and expander powers are presented as well.

The overall thermal efficiency $\eta_0$ growths for lower relative process air extractions $\delta_A$. In that region, the highly effective combined steam-gas turbine cycle dominates in the whole power generation process (Fig. 4).

The opposite situation exists for the fuel efficiency $\eta_F$, Fig. 5. For higher relative extractions, the optimized gas turbine has low power and needs less fuel, because a lot of heat required for the GT is taken from the FCC process. It means that the expander power is greater than the gas turbine power (Fig. 6).

When the gas turbine plays a significant role (small $\delta_A$ values), greater turbine inlet temperatures lead to higher fuel efficiencies. In case of overall efficiency the higher combustion temperatures are always connected with greater efficiencies.

Lower relative extractions mean a greater air mass flow at the gas turbine compressor inlet. The expander power increases insignificantly with the relative extraction growth. The steam turbine power can be optimized in some cases because the expander outlet gases have a temperature in 300–400°C range.

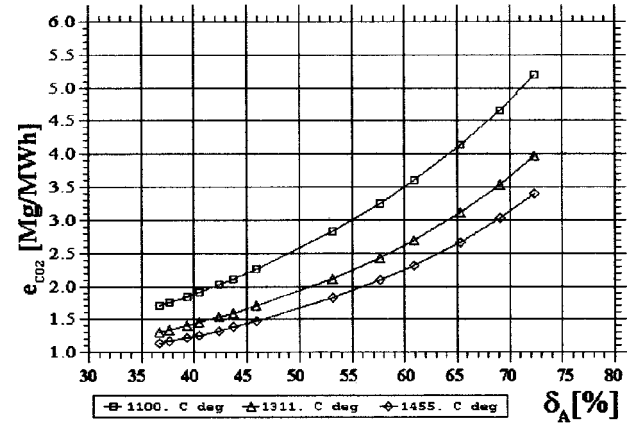



**Fig. 7 Oxygen dioxide emissions for the combine gas/steam cycle of the advanced FCC energy system**

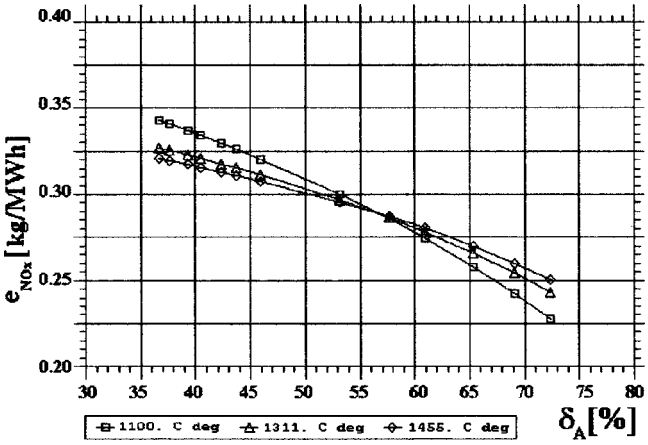



**Fig. 8 $NO_x$ emissions for the combined gas/steam cycle of the advanced FCC energy system**

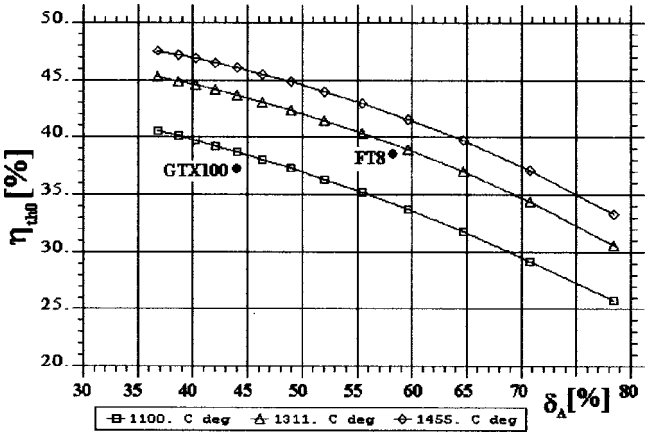



**Fig. 9 Overall thermal efficiency of the advanced FCC energy system versus relative process air extraction**

The fuel efficiency in the range of 60 percent, not available up till now in the conventional combined power plants, can be easily reached in the case of a recuperated gas/steam system incorporated into the FCC process.

The relative emissions are presented in Figs. 7 and 8. In the

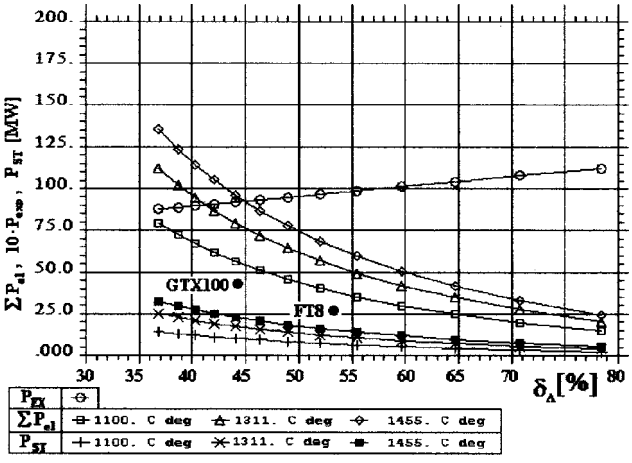



**Fig. 10 Electric power of the combined gas/steam FCC energy system, expander power, and steam turbine power versus relative process air extraction**

conventional FCC process, $CO_2$ relative emissions achieve about 11.5 Mg/MWh, which is a value few times higher than values presented in Fig. 7. For comparison, $CO_2$ emissions in the lignite-based technology used in power plants achieve about 1.2 Mg/MWh (the lowest values presented in Fig. 7).

In hard coal fired power plants, $NO_x$ emissions are about 3 kg/MWh, which means about ten times greater values than those presented in Fig. 9 for the new power system.

In variant 2, the old process air compressor remains in the new system (coupled with a expander and a motor-generator). There is no process air extraction from the gas turbine and the gas turbine under consideration differs only by the intersection cooling from conventional heavy-duty gas turbines. Some of them (e.g., RB-211) can be used in the FCC process without any significant adaptations ([1]), except replacing an internal recuperator at the turbine outlet by a FCC process heat recuperator. In Fig. 9 and 10, the parameters obtained in selected gas turbine sets with the pressure ratio $\pi=20$ (FT8 and GTX100) are marked. The higher overall efficiencies than in variant 1 are connected with the above mentioned differences of $\delta_A$ definitions in both the variants. The fuel efficiencies are lower than in variant 1, due to the fact that the electric generator and electric motor efficiencies are taken into account.

## Recuperated Gas Turbine in a Nitric Acid Plant

**Today's Power Systems in Nitric Acid Plants.** The raw materials required for the production of nitric acid ($HNO_3$) are atmospheric air and ammonia ($NH_3$). These raw materials are combined at elevated pressure and temperature and pass over a platinum-rhodium catalyst in the converter (Fig. 11), where ammonia reacts with oxygen from the air to form the nitric oxide (NO).

The reaction is highly exothermic and the generated high temperature process gas enters the heat exchanger train where a major portion of the reaction energy is recovered as heat, and the nitric oxide oxidizes further to nitrogen dioxide ($NO_2$). $NO_2$, oxygen, and water are combined in the absorption column to form nitric acid. These reactions depend highly on temperature and pressure, namely low temperatures and high pressures favor the formation of nitric dioxide together with the nitric acid. Parts of the exothermic reaction heat and the total heat flux are shown in Table 3.

The tail gas (mostly nitrogen with about two percent oxygen) from the absorption column is reheated with a part of the reaction heat in order to provide power for the air compressor by driving the hot gas expansion turbine (expander). Fig. 12 shows the temperature diagram for the heat exchanger train in the $HNO_3$ plant of 1000 Mg $HNO_3$ daily capacity.

The results of calculations of the conventional energy systems in the nitric acid production process are shown in Table 4. It is clearly visible that the overall thermal efficiencies of the electric power generation do not exceed 12 percent.

In many $HNO_3$ plants the overall thermal efficiency is much lower than the computed one, while the steam cycle dominates in the heat recovery system, and the expander inlet temperatures do not exceed 500°C. Many $HNO_3$ plants are so designed that the whole process, including the machine train, is self sufficient energywise. When the surplus energy is at disposal, then it is mainly in the form of the exported steam.

**Recuperated Gas Turbine Set Integrated With Nitric Acid Plant.** In earlier works (e.g., [2]) connected with possibilities of the gas turbine set integration with the nitric acid plant, only the tail gas injection into the combustion chamber was analyzed. The significant amounts of the process heat were used for the steam generation, like in the conventional nitric acid plant (Fig. 12).

The compressed air preheating by the process gas heat is not discussed in that paper. In the gas turbine integrated with the $HNO_3$ plant, the hot tail gas (reheated in the heat exchanger train) is injected upstream the combustion chamber.

The steam was generated in the waste heat boiler supplied by both the gas turbine exhaust gases and in the topping high pressure waste heat boiler in the heat exchanger train of the $HNO_3$ plant.

A suitable combination of the waste heat boiler downstream the gas turbine with the high pressure steam generating unit in the heat exchanger train allows one to obtain a high level of steam production.

In Fig. 13, the possibilities of using the high-pressure waste heat boiler in the heat exchanger train to enhance the steam pa-

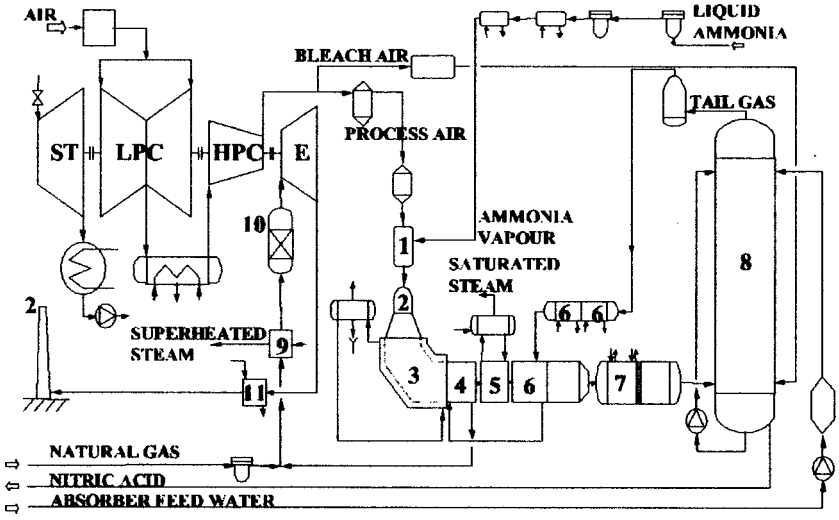



**Fig. 11 Schematic arrangement of a nitric acid plant. 1–ammonia air mixer, 2–converter catalyst, 3–converter elbow, 4–expander gas heater, 5–high pressure waste heat boiler, 6–tail gas heater, 7–waste heat exchanger, 8–absorber, 9–steam superheater, 10–combustor ($NO_x$ catalytic reduction), 11–economizer, 12–stack, ST–steam turbine, LPC–low pressure compressor, HPC–high pressure compressor, E–expander.**

**Table 3 Energy of exothermic reactions in the nitric acid production**

| Conversion Process | Liberated Reaction Heat GJ/MG $HNO_3$ | Liberated Reaction Heat % | Thermal Conditions |
|---|---|---|---|
| $4NH_3+5O_2=4NO+6H_2O$ | 3.671 | 51.4 | high |
| $4NH_3+3O_2=2N_2+6H_2O$ | 0.214 | 3.0 | temperatures |
| $2NO+O_2=2NO_2$ | 0.918 | 12.8 | |
| $4NO_2+2H_2O+O_2=4HNO_3$ | 1.022 | 14.3 | low |
| Nitric acid dilution | 0.378 | 5.3 | temperatures |
| Steam condensation | 0.946 | 13.2 | |
| Total liberated heat | 7.149 | 100.0 | |

$HNO_3$ World Production [3]: 60 mln Mg per year
Liberated Reaction Heat: 429 mln GJ per year
Waste Heat Output: $429 \cdot 10^8/(8000 \cdot 3600) = 14.9$ GW

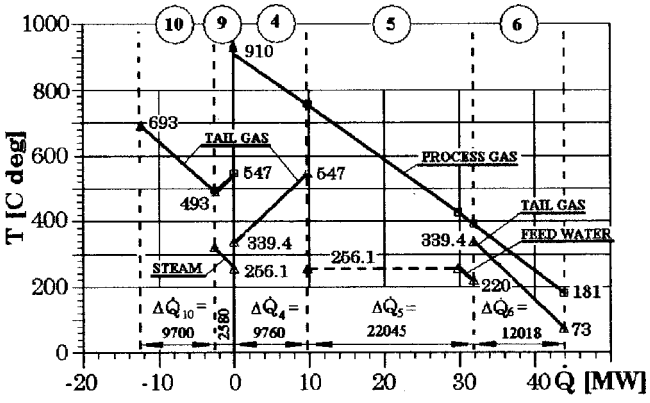



**Fig. 12 Temperature diagram for the heat exchanger train in the $HNO_3$ plant of 1000 Mg $HNO_3$ daily capacity (the exchanger and apparatus numbers correspond to numbers shown in Fig. 11)**

rameters in the combined cycle power generation are presented. The high-pressure waste heat boiler is located between the tail gas heater and the expander gas heater and reaches the thermal output of approximately 24 MW ($HNO_3$ plant of 1000 Mg/day capacity).

The recuperated gas turbine set introduced to the $HNO_3$ plant promises higher efficiencies of the gas cycle. The exhaust gases, which leave the gas turbine, have temperature high enough, which allows one to get the steam parameters like in the classical combined steam/gas cycle. In this solution, the high-pressure waste heat boiler is replaced by the compressed air preheater for the gas turbine combustor. The new heat exchanger train consists only of the compressed air preheater and tail gas low and high

**Table 4 The calculations results of conventional nitric acid production process**

| Variant | 1 | 2 |
|---|---|---|
| Process heat flux $\dot{Q}_p$ (MW) | 82.743 | 82.743 |
| Process pressure ratio | 11.0 | 15.0 |
| Expander inlet temperature | 693.0 | 693.0 |
| Process air mass flow (kg/s) | 55.0 | 55.0 |
| Tail gas mass flow (kg/s) | 43.175 | 43.175 |
| Process air compressor power (MW) | 16.404 | 19.052 |
| Expander power $P_{ex}$ (MW) | 18.902 | 20.611 |
| Steam turbine power $P_{ST}$ (MW) | 7.047 | 7.047 |
| Electric power $P_{el}$ (MW) | 9.259 | 8.348 |
| Process efficiency (percent) | 11.2 | 10.1 |
| Pressure drop factor (installation) | 0.175 | 0.175 |

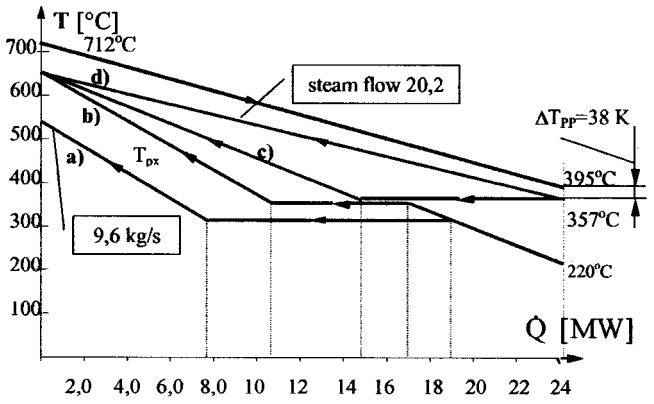



**Fig. 13 Temperature diagram of the high pressure waste heat boiler for different arrangements of boiler heating surfaces (economizer, evaporator, superheater)**

temperature preheaters. The suitable energy system is schematically presented in Fig. 14, and corresponding temperature diagram in Fig. 15.

A new energy system with the recuperated GT has been computed (Table 5) on the basis of the scheme presented in Fig. 14

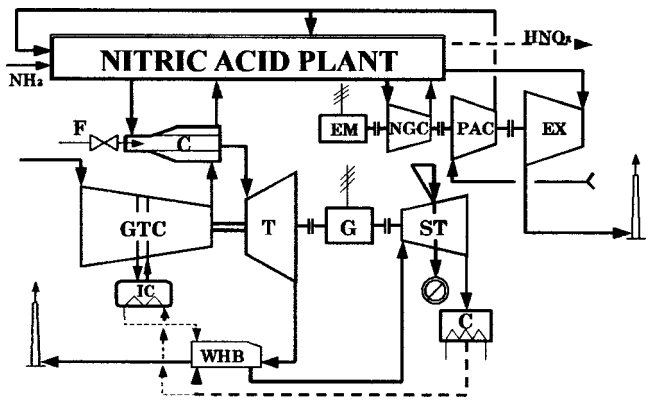



**Fig. 14 Recuperated gas turbine with compressed air preheating combined with steam cycle for an advanced nitric acid plant (with electric power coproduction). Compressor-expander train like in conventional energy system.**

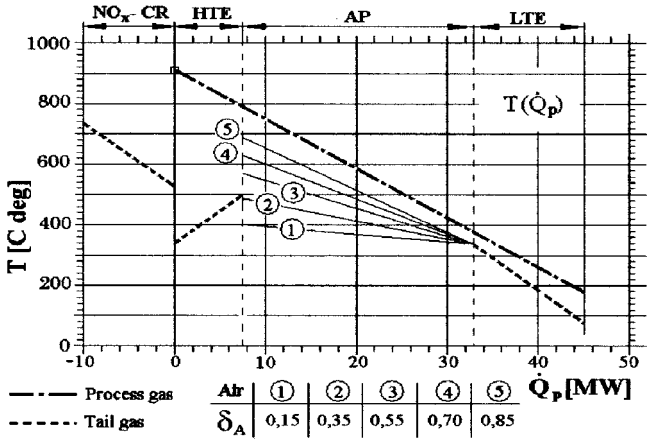



**Fig. 15 Temperature diagram for the heat exchanger train in the $HNO_3$ plant of 1000 Mg $HNO_3$ daily capacity. $NO_x$-CR–catalytic combustor for $NO_x$ red., HTE–high temperature exchanger, AP–air preheater for the GT combustor, LTE–low temperature tail gas exchanger.**

**Table 5 Results of calculations for the energy system of the combined cycle with a recuperated gas turbine ($HNO_3$ plant capacity—1000 Mg/day)**

| Gas Turbine Set (Virtually Adapted) | FT8 | GTX100 |
|---|---|---|
| Total process heat flux $Q_p$ (MW) | 82.743 | 82.743 |
| Process air compressor power (MW) | 9.299 | 9.299 |
| Nitric gases compressor power (MW) | 6.108 | 6.108 |
| Expander power $P_{ex}$ (MW) | 16.281 | 16.281 |
| Process air ratio $\delta_A$ | 0.392 | 0.312 |
| Process air mass flow (kg/s) | 55.0 | 55.0 |
| Gas turbine mass flow (kg/s) | 85.35 | 121.2 |
| Expander inlet temperature (°C) | 500.0 | 500.0 |
| Gas turbine power (MW) | 25.47 | 43.00 |
| Steam turbine power (MW) | 12.1 | 19.8 |
| Electric power (MW) | 38.418 | 63.648 |
| GT fuel heat flux | 44.400 | 93.766 |
| Heat flux transferred in air preheater | 22.45 | 22.45 |
| Overall thermal efficiency (percent) | 30.2 | 36.1 |
| Relative gas consumption ($Nm^3$/MWh) | 118.87 | 151.5 |
| Combined gas/steam cycle fuel efficiency (percent) | 86.5 | 67.9 |

and air and tail gas exchanger heat output (Fig. 15). The calculations have been made on the following assumptions:

- the tail gas heating in the catalytic combustor for $NO_x$ reduction is not taken into consideration;
- compressed air preheating is selected for two GT sets virtually adapted to the external heat recuperation;
- the selected gas turbines power is taken without any alterations, however the fuel heat output was reduced by the heat exchanged in the air preheater (recuperator).

Extremely low overall thermal efficiency of the conventional nitric acid energy system growths about three or four times due introducing the combined gas/steam cycle. In the combined cycle power plant with the recuperated gas turbine and integrated with the $HNO_3$ plant (1000 Mg daily capacity), the fuel efficiency $\eta_{F\text{comb}}$ can achieve 70 percent and even 85 percent depending on the effective electric power (38 MW, respectively, to 63.7 MW).

In the whole world about 60 mln Mg/year of $HNO_3$ is produced ([3]). It allows one to get the possible power about 17 GW with the combined gas/steam cycle fuel efficiency of 70 percent. In comparison with future combined gas/steam power plants with 60 percent fuel efficiency, it enables one to reduce the $CO_2$ world emission by about 240,000 Mg per year.

## Conclusions

An introduction of a recuperated gas turbine incorporating external heat sources in the combined gas-steam cycle into the industrial process allows for

- electric power production with an efficiency which will not be attainable before the first decade of the 21 century in combined gas-steam power plants;
- significant reduction of the atmospheric air pollution (thermal, $CO_2$, and $NO_x$ emissions).

It should not be difficult to make a realistic estimation of the investment costs for this new energy system. The investment costs depend on the integration scale. By means of the full integration of the combined gas/steam power plant with the FCC process (Fig. 3, variant 1) the investment costs can be reduced by excluding the process air compressor. In that case the WHB downstream the hot gas expander will be replaced by the air preheater for the gas turbine. It balances those costs.

There are many more possibilities of $HNO_3$ plant integration with the combined gas/steam power plant. In the simplest case, a high pressure WHB located in the heat exchanger train is replaced by an air preheater of the gas turbine, without any changes in the whole machine train. In a wider case, the whole conventional machine train of the $HNO_3$ plant is replaced by the industrial gas turbine (with process air extraction, combustor air preheating and tail gas injection into the gas turbine). The investment costs for such integration can be the lowest for newly erected $HNO_3$ plants.

The external heat recuperation requires some design adaptations of present heavy duty gas turbines referring to compressed air extraction and preheated air injection. The heavy duty gas turbine pressure ratio suits well in most industrial processes (e.g., $\pi = 15$ is the value which is used most often in two-pressure $HNO_3$ installations).

These adaptations can be carried out perhaps only within the frames of future international programs or by an extension of the existing U.S. Department of Energy Program ([4]).

## Nomenclature

$c_p$ = specific heat at constant pressure J/(kg·K)
$e$ = relative emission, kg/MWh
$\dot{m}$ = mass flow rate, kg/s
$p$ = absolute pressure, Pa
$P$ = power, W
$\dot{Q}$ = heat rate $dQ/dt$, W
$q$ = specific heat, J/kg
$\beta$ = fuel-to-air mass ratio
$\varepsilon$ = relative pressure drop
$\delta_A$ = relative air extraction
$\eta$ = efficiency
$\eta_p$ = polytropic efficiency
$\pi$ = cycle/compressor pressure ratio

### Subscripts

0 = basic conditions, overall
1,2,3... = state points
$C, c$ = compressor, cold
$e$ = effective, stagnation conditions
$el$ = electric
$EX$ = expander
$F$ = fuel
$GT$ = gas turbine
$P$ = process
$R$ = recuperator
$ST$ = steam turbine
$T$ = turbine
$th$ = thermal

## References


[1] Wieler, C. L., 1998, *WR-21 Intercooled Recuperated Gas Turbine*, http://www.gas-turbines.com.RANDD/ICR-WRDS.htm.
[2] Chodkiewicz, R., Porochnicki, J., and Potapczyk, A., 1998, *Electric Power and Nitric Acid Coproduction: A New Concept in Reducing the Energy Costs*, Vol. III, PowerGen Europe'98, Milan, Italy pp. 611–625.
[3] *Ullman Encyclopedia of Industrial Chemistry*, 1991, Vol. A17.
[4] Lane, A. W., and Hoffman, P. A., 1998, "The U.S. Dept. of Energy Advanced Turbine Syst. Program," ISROMAC-7, Hawaii.
[5] Miller, H. F., 1989, *Blade Erosion—FCCU Power Recovery Expanders*, D-R Turbo Products Division, Olean, NY.

# Aluminizing Behaviors of Vacuum Plasma Sprayed MCrAlY Coatings

**Y. Itoh**

**M. Saitoh**

**Y. Ishiwata**

Metal and Ceramics Technology
R&D Department,
Toshiba Corporation,
2-4 Suehiro-cho, Tsurumi-ku,
Yokohama 230-0045, Japan

*The objective of this study is aluminide overlay coatings of MCrAlY sprayed by a vacuum plasma spraying (VPS) process for the protection against high-temperature corrosion and oxidation of gas turbine components. Diffusion coating processes have been applied for many years to improve similarly the environmental resistance by enriching the surface of nickel-based superalloys with chromium, aluminum, or silicon element. Recently, aluminizing of MCrAlY coatings is used for improving further the high-temperature oxidation resistance. However, the aluminizing properties of plasma-sprayed MCrAlY coatings, which have an important effect on the coating performance, have not been clarified. In this study, five kinds of plasma-sprayed MCrAlY (CoCrAlY, CoNiCrAlY, CoNiCrAlY+Ta, NiCrAlY, and NiCoCrAlY) coating were selected for pack-aluminizing tests. The as sprayed and the heat-treated (1393 K, 2 h, argon cooled and 1116 K, 24 h, argon cooled) MCrAlY specimens were Al–Cr–$Al_2O_3$–$NH_4Cl$ pack-aluminized at 1173, 1223, and 1273 K for 5, 10, and 20 h, respectively. The experimental results showed that the aluminizing process formed the aluminum rich layers of NiAl or CoAl phase. It also indicated that the thickness of the aluminum rich layer showed a parabolic time-dependence in all MCrAlY coatings. The order of reaction diffusion rate was NiCoCrAlY=NiCrAlY>CoNiCrAlY >CoNiCrAlY+Ta>CoCrAlY. There was a tendency that the reaction diffusion rate by aluminizing increased with increasing nickel content in the MCrAlY coatings and the reaction diffusion rate of as sprayed MCrAlY coatings is faster than that of the heat-treated MCrAlY coatings.* [DOI: 10.1115/1.1423914]

## Introduction

High-temperature protective coatings used in gas turbine components must withstand extremely severe environmental conditions ([1,2]). There are two basic coating systems, which are currently being used in gas turbines to provide improved corrosion and oxidation resistance, and thereby extend the component life. These are called diffusion coatings and overlay coatings ([3]). The diffusion coatings provide a surface enrichment of aluminum, chromium, or silicon, and aluminizing behavior of nickel-based superalloys has been fully reported ([4]). As operating temperatures of gas turbines are increased, it becomes impossible to achieve the service lives using the diffusion coatings. On the other hand, the overlay coatings are specifically designed corrosion and oxidation resistant MCrAlY alloys, where M is iron, nickel, and/or cobalt. In recent years, a vacuum plasma spraying (VPS) process has been used to apply the overlay coatings of MCrAlY alloys ([2]). This coating process has been found to be very effective for gas turbine components, such as blades and duct segments, which are exposed to aggressive gases at high temperature. As operating temperatures are increased, the excellent performance of diffusion coatings applied on top of MCrAlY overlay coatings has been realized. The over-aluminized MCrAlY coatings have been used for improving further the high-temperature oxidation resistance.

On the other hand, thermal barrier coatings, which consisted of an air-plasma sprayed yttria partially-stabilized zirconia layer applied directly over an air-plasma sprayed MCrAlY bond coat layer, provide to improve the thermal resistivity for air-cooled gas turbine combustors. Nowadays, thermal barrier coatings, which consisted of a vacuum-plasma sprayed MCrAlY bond coat layer, are developed for gas turbine blades and vanes ([5,6]). In addition to extending the life of thermal barrier coatings, the aluminized process to enrich the outer surface of MCrAlY bond coats in aluminum content has been developed, and thereby high-temperature oxidation resistance was highly improved owing to the formation of the alumina layer at the interface between the yttria partially stabilized zirconia layer and the bond coat layer ([7,8]). As mentioned above, the over-aluminized MCrAlY coatings were one of the important technologies for improving the coating life. However, the aluminizing behaviors of MCrAlY coatings have not been fully clarified.

In the present paper, five typical kinds of vacuum plasma sprayed MCrAlY (CoCrAlY, CoNiCrAlY, CoNiCrAlY+Ta, NiCrAlY, NiCoCrAlY) coatings were selected for aluminizing experiments, and the vacuum-plasma sprayed and the heat-treated (1393 K, 2 h, argon cooled and 1116 K, 24 h, argon cooled) MCrAlY coatings were pack-aluminized at 1173, 1223, and 1273 K for 5, 10, and 20 h, respectively. The aluminizing behaviors of these MCrAlY coatings were investigated. Further, the effect of nickel content on the aluminizing behaviors of MCrAlY coatings was discussed.

## Experimental Procedure

This investigation used IN738LC as a coating substrate, which is widely used for gas turbine blades. The chemical compositions of IN738LC is given in Table 1. Also, the standard heat-treatment for IN738LC is shown in Table 1. Five commercial spraying powders CoCrAlY, CoNiCrAlY, CoNiCrAlY+Ta, NiCrAlY, and NiCoCrAlY were selected. The CoCrAlY coating is used for corrosion protection and applied at medium temperature (1000–1200 K). The NiCrAlY coating is used for oxidation protection and applied at high temperature (>1200 K) and as the bond coat layer of thermal barrier coatings. The CoNiCrAlY, CoNiCrAlY+Ta, and NiCoCrAlY are advanced coating powders that have the superior properties against the more severe environmental conditions. These spraying powders are produced by argon atomization. The chemical compositions and mesh size of these powders are given in Table 2.

The VPS coatings were carried out with the A-2000V VPS system (Plasma Technik), under the following conditions: preheating temperature 843 K during transferred arc treatment, voltage 64 V, current 685 A, spraying distance 270 mm, argon gas atmo-

Contributed by the Materials, Manufacturing, and Metallurgy Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the MM&M Division, Dec. 2000; final revision received by the ASME Headquarters Feb. 2001. Editor: H. D. Nelson.

**Table 1 Chemical composition and mechanical properties of IN738LC**

| Material | Chemical composition (mass %) | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | Ni | Cr | W | Mo | Co | Al | Ti | Nb | Ta |
| IN738LC | Bal. | 15.8 | 2.46 | 1.78 | 8.11 | 3.45 | 3.60 | 0.92 | 1.92 |

| Tensile test at room temperature | | | |
|---|---|---|---|
| 0.2% proof stress | Tensile strength | Elongation | Hardness (HRC) |
| 936 MPa | 1149 MPa | 7% | 41.5 |

Heat treatment; 1393K × 2h Ar cooled
1116K × 24h Ar cooled

**Table 2 Chemical composition of spraying powders used**

| Materials | Chemical composition (mass %) | | | | | |
|---|---|---|---|---|---|---|
| | Co | Ni | Cr | Al | Y | Ta |
| CoCrAlY −400 mesh | Bal. | <0.10 | 28.9 | 6.15 | 0.28 | — |
| CoNiCrAlY −400 mesh | Bal. | 32.08 | 21.22 | 8.33 | 0.39 | — |
| CoNiCrAlY+Ta −400 mesh | Bal. | 9.7 | 24.6 | 7.0 | 0.48 | 5.5 |
| NiCrAlY −400 mesh | — | Bal. | 31.45 | 11.55 | 0.825 | — |
| NiCoCrAlY −325 mesh | 23.23 | Bal. | 16.94 | 12.47 | 0.49 | — |

sphere 6 kPa. The single face of all IN738LC substrates, which were $15\times15\times5$ mm$^t$, was sprayed MCrAlY coatings with 0.3 mm in thickness after the pretreatment of blasting and transferred arcing. As sprayed and heat-treated MCrAlY specimens (1393 K, 2 h, argon cooled and 1116 K, 24 h, argon cooled) were subjected to an aluminizing treatment at 1173, 1223, and 1273 K for 5, 10, and 20 h in a pack-cemented retort purged with pure argon, respectively. The heating rate was 380 K/h, and the cooling rate was 160 K/h until 900 K and furnace cooled after that. The pack mixture composition is an Al+Cr alloy powder (20 mass percent) as the aluminum source, $NH_4Cl$ (0.5 mass percent) as the activator, and alumina powder (79.5 mass percent) as the inert filler.

After the aluminizing treatment, the cross section of test specimens was observed and the thickness of aluminizing layer was measured using an optical microscope. Intermetallic compounds precipitated in the MCrAlY coatings and the aluminizing layers were observed by a scanning electron micrography (SEM) and an electron probe microanalyzer (EPMA).

A typical microstructure of VPS CoCrAlY coating is shown in Fig. 1(*a*) by the comparison between the as sprayed and the heat-treated coating. A typical microstructure of VPS NiCoCrAlY coating is shown in Fig. 1(*b*), and the EPMA line analysis results of the aluminum element are shown. It was found that the microstructure of all MCrAlY coatings was composed of precipitated aluminum compound phase in the Ni and/or Co matrix ([9]). In case of the VPS CoCrAlY, the darker matrix was identified by EPMA as $\alpha$–CoCr solid solution phase whereas the light precipitated phase consisted of $\beta$–CoAl. Similarly, the VPS NiCoCrAlY is composed of two phases, which are precipitated $\beta$–(Ni,Co)Al intermetallic compounds and $\alpha$–NiCoCr solid solution matrix ([9,10]). It was recognized that the microstructure in the as-sprayed MCrAlY was fine for the rapid cooling of sprayed particles during spraying process. However, the microstructure of heat-treated MCrAlY became coarse due to gathering up the precipitated intermetallic compounds.

## Microstructure of Aluminized Coating Layer

A cross-section microstructure of the aluminized layer formed at the surface of the IN738LC substrate after aluminizing treatment at 1223 K for 10 h is shown in Fig. 2(*a*). Two reaction diffusion layers can be clearly observed using EPMA analysis, consisting of an aluminum-rich outer layer and an inner diffusion layer. The aluminum-rich outer layer is nearly 35 $\mu$m in thickness. It was confirmed by an X-ray diffraction pattern that the outer layer mainly consisted of a NiAl precipitated phase. The inner diffusion layer is nearly 5 $\mu$m in thickness, consisting of an aluminum and chromium-rich layer. There is a tendency that the NiAl precipitates decreases from the outer layer to the IN738LC substrate. Namely, it is because the NiAl precipitates in the aluminized layer are formed by inner diffusion of the aluminum element, and chromium, titanium, and tungsten, etc. enriched by exceeding their dissolution limit in a nickel-based superalloy.

Next, a microstructure of the aluminized layer formed at the surface of as-sprayed VPS CoCrAlY coating after aluminizing treatment at 1223 K for 10 h is shown in Fig. 2(*b*). It was found that the VPS CoCrAlY was composed of two phases, which were precipitated CoAl intermetallic compounds (gray zone) and a CoCr solid solution matrix (white zone) ([10]). And large size pores can be observed in the as-sprayed VPS CoCrAlY. It is clear that two reaction diffusion layers can be observed in the same way as shown in IN738LC. An outer reaction diffusion layer of aluminum rich is nearly 20 $\mu$m in thickness. It was confirmed by an X-ray diffraction pattern that the outer layer mainly consisted of a

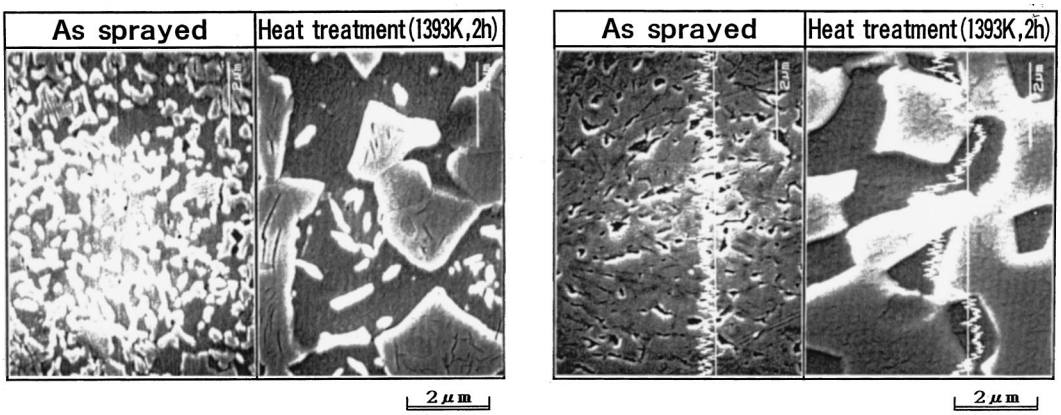
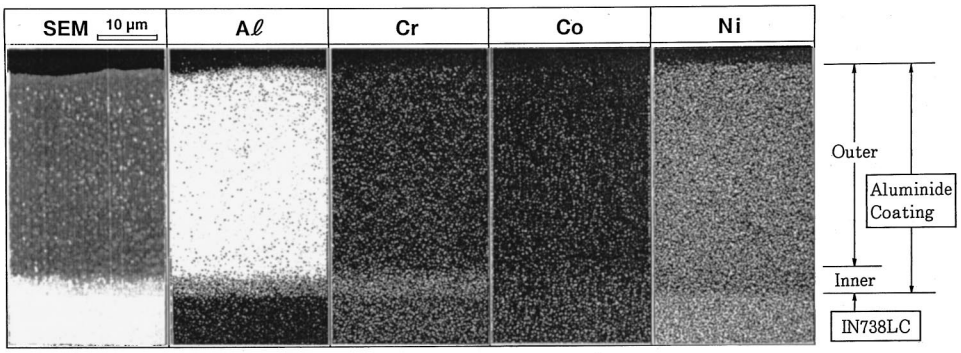
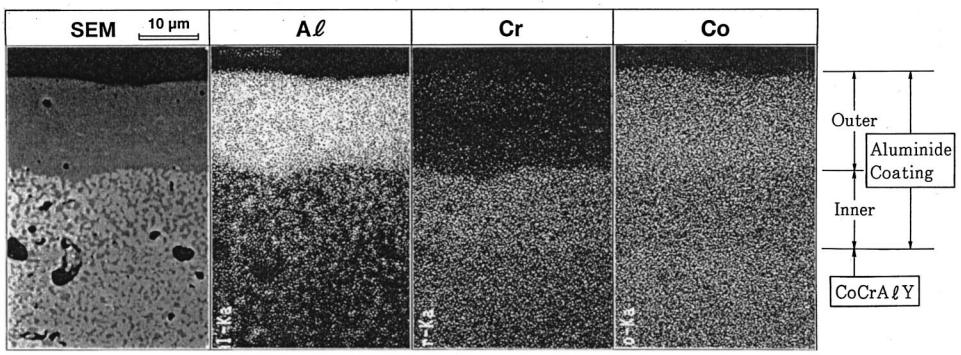


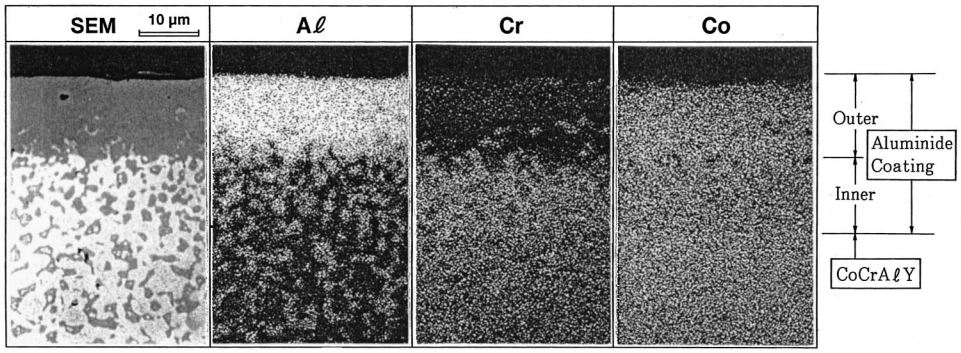

**Fig. 1 SEM observation of vacuum-plasma sprayed MCrAlY coatings before aluminizing; (*a*) CoCrAlY coating, (*b*) NiCoCrAlY coating**




(*a*)




(*b*)




(*c*)

**Fig. 2 SEM observation and EPMA analysis of aluminide coatings at 1223 K for 10 h; (*a*) aluminized layer of IN738LC, (*b*) over-aluminized layer of as-sprayed CoCrAlY coating, (*c*) over-aluminized layer of heat-treated CoCrAlY coating**

CoAl precipitated phase. The inner diffusion layer is nearly 15 $\mu$m in thickness, consisting of an aluminum and chromium rich layer. In the inner diffusion layer, the CoAl precipitates like an island and are decreased toward the inside and become small islands gradually. It is because the CoAl precipitates in an aluminized layer are formed by the inner diffusion of the aluminum element, and exceeding their dissolution limit in cobalt-based superalloy enriches chromium.

On the other hand, a microstructure of the aluminized layer formed at the surface of heat-treated VPS CoCrAlY coating after aluminizing treatment at 1223 K for 10 h is shown in Fig. 2(*c*). In this case, two reaction diffusion layers can be clearly observed consisting of an aluminum-rich outer layer and an inner diffusion layer. The pore size is smaller than that of an as-sprayed VPS CoCrAlY. The aluminum-rich layer mainly consisted of the CoAl phase and can be observed on the cross-section view of the aluminized VPS CoCrAlY, and is nearly 15 $\mu$m in thickness. Also, the inner diffusion layer is nearly 15 $\mu$m in thickness, consisting of an aluminum and chromium-rich layer. There is no qualitative difference of the aluminizing behaviors between the as-sprayed and the heat-treated VPS CoCrAlY, basically. However, the thickness of the aluminized layer is slightly thin in comparison with that of the as-sprayed VPS CoCrAlY. Within this experiment, it seems that the pores do not have an important effect upon the aluminizing behaviors. These aluminizing behaviors are almost the same as the other VPS MCrAlY coatings, except for the thickness of aluminized layer. However, it was confirmed using the X-ray diffraction method that (Ni, Co) Al intermetallic compounds were formed at the surface of VPS CoNiCrAlY, VPS CoNiCrAlY+Ta and VPS NiCoCrAlY coatings and NiAl intermetallic compounds were formed at the surface of VPS NiCrAlY coating by the aluminizing treatment.

## Aluminizing Behavior of MCrAlY Coatings

The application of Fick's first law to polyphase diffusion in binary systems leads to a simple physical interpretation of various

possible types of behaviors in the growth of diffusion layers ([11,12]). The rate equation for reaction diffusion in the *j*th layer in an *n*-phase system has been given as Eq. (1) ([13]).

$$\omega_j = k_j \cdot t^{1/2} \tag{1}$$

where $k_j$ is called the rate constant for reaction diffusion layer growth of the *j*th phase and $\omega_j$ is the thickness of the *j*th layer. This equation indicates that each diffusion thickness formed by the reaction diffusion can be generally observed to follow the parabolic time, *t* dependence.

The growth behaviors of aluminized layers formed at the surface of IN738LC are shown in Fig. 3. The thickness of the aluminized layer is fixed as that of the aluminum-rich layer and the inner diffusion layer in the experiments. There is a tendency for the aluminizing thickness to increase with increasing the aluminizing temperature and the aluminizing time. It is clear that the linear relationships between the aluminizing thickness, $\omega$ and aluminizing time, $t^{1/2}$ can be observed for each aluminizing temperature. Namely, within the $\pm 15$ $\mu$m margin of error in the measurement of aluminized thickness, the parabolic growth rate of the aluminizing layers can be observed within these experimental results.

In the same manner, the growth behaviors of aluminized layers formed at the surface of VPS MCrAlYs are shown in Fig. 4(*a*) as as-sprayed VPS MCrAlYs and Fig. 4(*b*) as heat-treated VPS MCrAlYs. It is clear from Fig. 4(*a*) that the linear relationships between the aluminizing thickness, $\omega$ and aluminizing time, $t^{1/2}$ can be observed for all kinds of VPS MCrAlY coatings. This shows that the aluminizing behaviors of each VPS MCrAlY coatings follow a simple parabolic time law. These aluminizing behaviors of the as-sprayed VPS MCrAlYs are similar to those of the heat-treated VPS MCrAlYs as shown in Fig. 4(*b*). Also, it is confirmed that the particular growth behaviors of the aluminum-rich outer layer and the inner diffusion layer follow a simple parabolic time law individually.

From the experimental results, it is confirmed that the aluminizing thickness, $\omega$ is given as Eq. (2)

$$\omega = k \cdot t^{1/2}, \tag{2}$$

where *k* is called rate constant for the growth of aluminizing layer.

The temperature dependence of rate constants that are obtained from the slope of straight lines in Fig. 4(*a*) is shown in Fig. 5(*a*). It is clear that the Arrhenius-type temperature dependence in all as-sprayed VPS MCrAlY coatings applies. The effect of the as-sprayed VPS MCrAlYs on the rate constants can be clearly obtained. The order of the aluminizing rate for VPS MCrAlYs was NiCoCrAlY = NiCrAlY > CoNiCrAlY > CoNiCrAlY+Ta >CoCrAlY. The same order of aluminizing rate can be obtained in the case of the heat-treated VPS MCrAlYs. Namely, Fig. 5(*b*) shows the temperature dependence of rate constants that are obtained from the slope of straight lines in Fig. 4(*b*). The Arrhenius-type temperature dependence can be also confirmed in all heat-treated VPS MCrAlYs. The experimental results of IN738LC are also shown in Fig. 4(*b*). The rate constants of IN738LC are in agreement with that of the heat-treated VPS CoNiCrAlY coating.

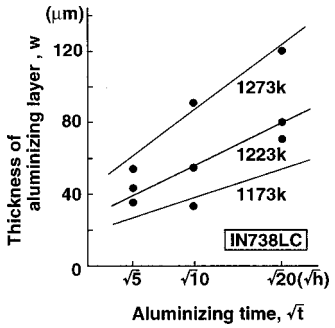



**Fig. 3 Aluminizing behavior for IN738LC**

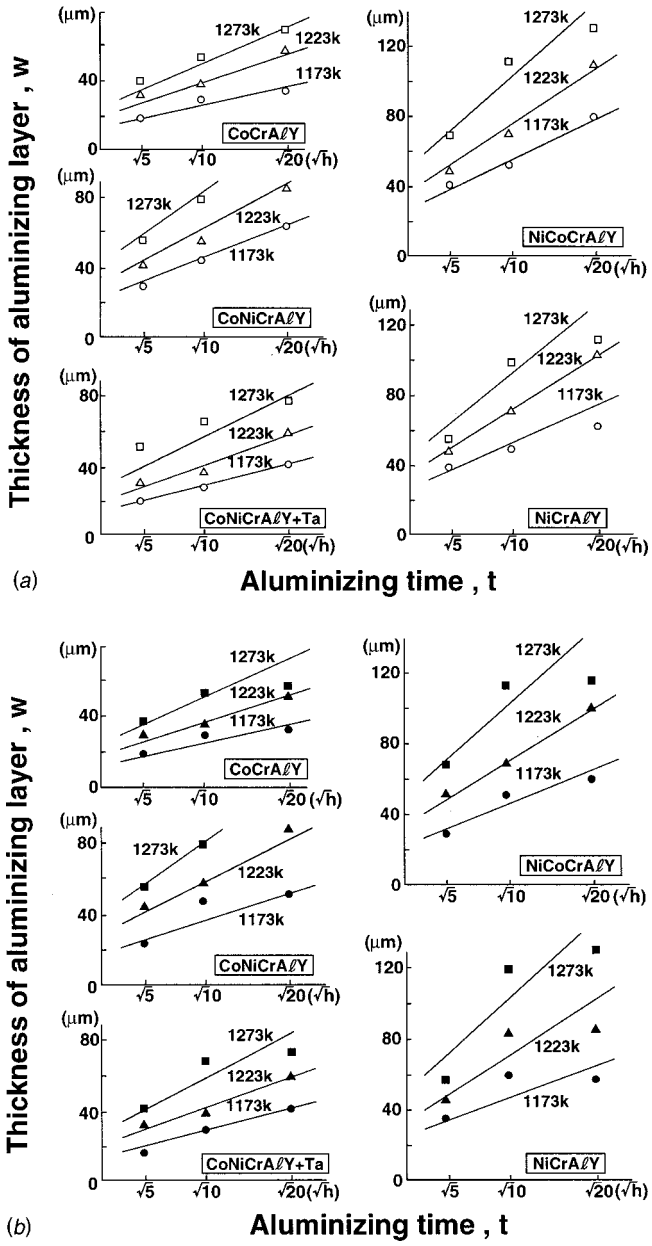






**Fig. 4 Over-aluminizing behaviors for various kinds of vacuum-plasma sprayed MCrAlY coatings; (*a*) as-sprayed MCrAlY coatings (*b*) heat-treated MCrAlY coatings**

From the experimental results obtained above, the rate constant, *k* for the aluminizing layer growth is given as equation (3).

$$k = k_0 \cdot \exp[c(1/T - 1/1223)] \tag{3}$$

where $k_0$ and *c* are material constants.

The material constants, $k_0$ and *c* obtained by the experimental data in Figs. 5(*a*) and (*b*), are shown in Figs. 6(*a*) and (*b*). The horizontal axis shows the nickel content in VPS MCrAlYs. There is a tendency that the material constants, $k_0$ increases with increasing the nickel content in the VPS MCrAlYs. Using cobalt content in the VPS MCrAlYs, the material constants, $k_0$, decrease with increasing the cobalt content in the VPS MCrAlYs. Namely, it is clear that the inner diffusion of aluminum from the surface layer to the VPS MCrAlYs is mainly affected by the nickel and

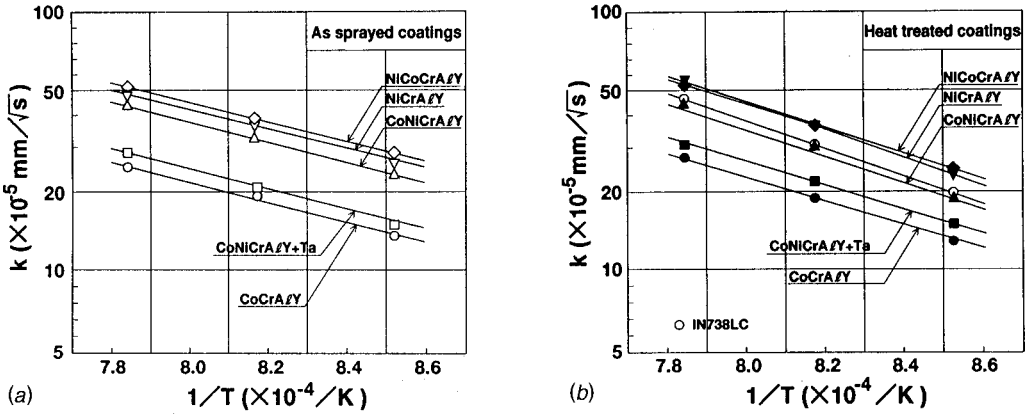



**Fig. 5 Temperature dependency of diffusion rate constant on a parabolic time law; (*a*) as-sprayed MCrAlY coatings (*b*) heat-treated MCrAlY coatings**

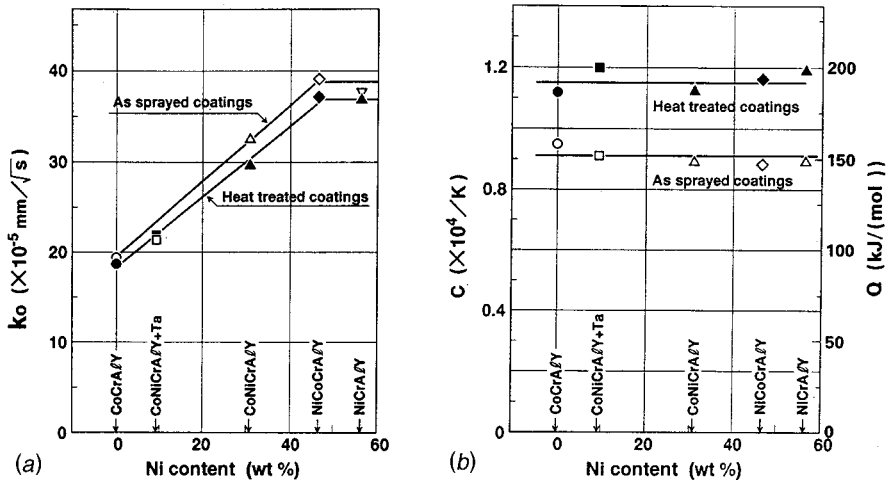



**Fig. 6 Effect of nickel content in as-sprayed MCrAlY coating on aluminizing behaviors in comparison with heat-treated MCrAlY coatings; (*a*) material constant, $k_0$ (mm/s$^{1/2}$), (*b*) material constant, *c* (/K)**

cobalt content. However, it seems that the material constants, *c*, that correspond to apparent activation energy are not remarkably affected by the nickel and cobalt content.

In Figs. 6(*a*) and (*b*) the effect of the nickel content in the as-sprayed VPS MCrAlYs on the material constants, $k_0$ and *c*, in comparison with the heat-treated VPS MCrAlYs are shown. The material constants, $k_0$, of the as-sprayed VPS MCrAlYs show nearly higher values in comparison with the heat-treated VPS MCrAlYs. On the contrary, the material constants, *c*, of the as-sprayed VPS MCrAlYs show lower values in comparison with the heat-treated VPS MCrAlYs.

By the way, it is well known that diffusion rate constants can be generally shown by the Arrhenius-type equation described below ([13]).

$$k^2 = A \cdot \exp(-Q/RT) \tag{4}$$

where *A* is a material constant, *Q* is activation energy, and *R* is a gas constant.

From the comparison between Eqs. (3) and (4), the material constant, *c*, corresponds to the material constant, $Q$ ($=2$Rc), and the material constant, *A* is the function of $k_0$ and *c* as $A = k_0^2 \cdot \exp(c/611.5)$.

## Discussion

From Figs. 6(*a*) and (*b*), it was confirmed that there was a difference of aluminizing behaviors between the as-sprayed VPS MCrAlYs and the heat-treated VPS MCrAlYs.

First, we discuss the reason of the quantitative difference of aluminizing behaviors as follows. It was thought from Figs. 2(*b*) and (*c*) that the pores had no effect on the aluminizing behavior. Namely, the aluminizing behavior through the pores by chloride gas of aluminum produced in retort cannot be observed in any as-sprayed VPS MCrAlYs. We show in Fig. 7 the variation of porosity among the VPS MCrAlY coatings investigated. These are the measurement results of freestanding VPS MCrAlYs in case of the as-sprayed and heat-treated (1393 K, 2 h and 1523 K, 2 h in argon atmosphere) coatings. Though there is little difference among the VPS MCrAlYs, the pores including in the as-sprayed VPS MCrAlYs are reduced by the heat treatment at high temperature. There is a tendency that the relative density increases above 95 percent with increasing the heating temperature. It is thought that these densification behaviors of VPS MCrAlYs should arise during the aluminizing process. The sintering of VPS MCrAlYs can be confirmed from the roundish pores observed in Figs. 2(*b*) and (*c*). The laminar microstructure of as sprayed coatings cannot be observed in the aluminized VPS MCrAlYs. The open pores in

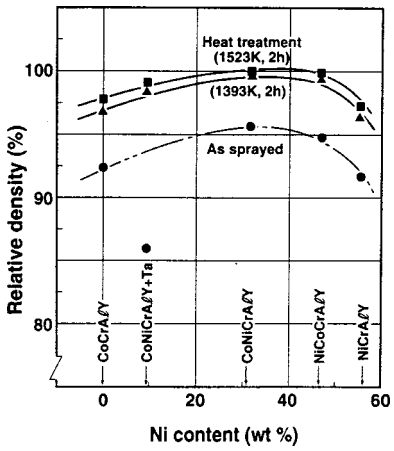



**Fig. 7 Relative density of MCrAlY coatings measured by Archimedes method**

the as sprayed VPS MCrAlYs disappear during the aluminizing process, and the local delay effect of aluminizing by the existence of pores cannot be remarkably observed.

Secondly, we investigated the influence of composition of MCrAlY coatings on aluminizing behaviors in Figs. 6(*a*) and (*b*), especially by paying attention to the nickel content. The material constants $k_0$ and $c$ prescribed the aluminizing behaviors could be arranged in order as shown in Fig. 6(*a*) using the nickel content. There is a tendency that the material constants, $k_0$ show higher value with increasing the nickel content in the VPS MCrAlYs. This shows that the reaction diffusion between nickel and aluminum increases and a lot of aluminum compounds arise. However, the material constants, $c$, that correspond to the apparent activation energy have not been influenced by the nickel content in the VPS MCrAlYs because the reaction diffusion mechanism between nickel and aluminum elements is almost the same in the VPS MCrAlYs. It was found by the comparison between Figs. 1(*a*) and (*b*) that the microstructure of the as-sprayed VPS MCrAlYs is fine in the rapid cooling during the spraying process. However, the microstructure of the heat-treated VPS MCrAlYs is coarse for gathering up the intermetallic compounds. When the microstructure becomes fine like the as-sprayed VPS MCrAlYs, it is considered that the effect of intergranular diffusion cannot be ignored. Generally, it is well known that the intercrystalline diffusion rate is faster than the transcrystalline and this tendency becomes remarkable with decreasing the test temperature. From this point of view, it might be thought that the difference of aluminizing behaviors between the as-sprayed and the heat-treated VPS MCrAlYs is influenced by the number of grain boundaries and the size and quantity of precipitated intermetallic compounds.

## Summary

Five VPS MCrAlY coatings, such as CoCrAlY, CoNiCrAlY, CoNiCrAlY+Ta, NiCrAlY, and NiCoCrAlY, were selected for a pack-aluminizing treatment. The aluminizing behaviors of these VPS MCrAlY coatings were made clear. Especially, the difference of aluminizing behaviors between the as-sprayed and the heat-treated VPS MCrAlYs is discussed.

In the case of the nickel-based superalloy, IN738LC, two reaction diffusion layers can be clearly observed, consisting of an aluminum-rich outer layer and an inner diffusion layer. In the case of the five kinds of VPS MCrAlYs, it is also confirmed that two reaction diffusion layers can be observed, consisting of an aluminum-rich outer layer and an inner diffusion layer. Using the X-ray diffraction method, CoAl intermetallic compounds were observed at the surface of VPS CoCrAlY as the aluminum-rich outer layer. Similarly, (Ni, Co) Al intermetallic compounds were observed at the surface of VPS CoNiCrAlY, VPS CoNiCrAlY+Ta and VPS NiCoCrAlY and NiAl intermetallic compounds were observed at the surface of VPS NiCrAlY by the aluminizing treatment. From observation of the microstructure, there is no qualitative difference of the aluminizing behaviors between the as sprayed and the heat-treated VPS MCrAlYs except for the thickness of the aluminized layer. It is confirmed that the linear relationships between the aluminizing thickness, $\omega$ and aluminizing time, $t^{1/2}$ can be observed for all kinds of VPS MCrAlYs. This shows that the aluminizing behaviors of each VPS MCrAlYs follow a simple parabolic time law. Also, the Arrhenius-type temperature dependence of the aluminizing rate can be confirmed in all VPS MCrAlYs. In case of the as-sprayed VPS MCrAlYs, the order of aluminizing rate for the VPS MCrAlYs was NiCoCrAlY $=$NiCrAlY$>$CoNiCrAlY$>$CoNiCrAlY+Ta$>$CoCrAlY. The same order of aluminizing rate can be obtained in the case of the heat-treated VPS MCrAlYs. However, the thickness of the aluminized layer of heat-treated VPS MCrAlYs is slightly thin in comparison with that of the as-sprayed VPS MCrAlYs. There is a tendency that the reaction diffusion rate by aluminizing increases with increasing nickel content in the MCrAlY coating and the reaction diffusion rate of the as-sprayed MCrAlYs is faster than that of the heat-treated MCrAlYs. It might be thought that the difference in aluminizing behaviors between the as-sprayed and the heat-treated VPS MCrAlYs is influenced by the number of grain boundaries and the size and quantity of precipitated intermetallic compounds.

## References


[1] Bettridge, D. F., and Ubank, R. G., 1986, "Quality Control of High-Temperature Protective Coatings," Mater. Sci. Technol., **2**, pp. 232–242.

[2] Mevrel, R., 1989, "State of the Art on High-Temperature Corrosion-Resistant Coatings," Mater. Sci. Eng., A, **A120**, pp. 13–24.

[3] Nicholls, J. R., and Stephenson, D. J., 1991, "High Temperature Coatings for Gas Turbine," Surf. Eng., **22**, pp. 156–163.

[4] Das, D. K., Singh, V., and Joshi, S. V., 1998, "Evolution of Aluminide Coating Microstructure on Nickel-Base Cast Superalloy CM-247 in a Single-Step High-Activity Aluminizing Process," Metall. Mater. Trans. A, **29A**, pp. 2173–2188.

[5] Duvall, D. S., and Ruckle, D. L., 1982, "Ceramic Thermal Barrier Coatings for Turbine Engine Components," ASME Paper 82-GT-322.

[6] Miller, A. M., 1987, "Current States of Thermal Barrier Coatings—An Overview," Surf. Coat. Technol., **30**, pp. 1–11.

[7] Wortman, D. J., Nagaraj, B. A., and Duderstadt, E. C., 1989, "Thermal Barrier Coatings for Gas Turbine Use," Mater. Sci. Eng., A, **A121**, pp. 433–440.

[8] Wu, B. C., Chao, C. H., and Chang, E., 1990, "Effect of Bond Coat Pre-Aluminizing Treatment on the Properties of ZrO2–8wt.%Y2O3/Co-29Cr-6Al-1Y Thermal Barrier Coatings," Mater. Sci. Eng., A, **A124**, pp. 215–221.

[9] Mazars, P., Manesse, D., and Lopvet, C., 1987, "Degradation of MCrAlY Coatings by Interdiffusion," ASME Paper 87-GT-58.

[10] Itoh, Y., 1999, "Reaction Diffusion Behaviors for Interface Between Ni-Based Super Alloys and Vacuum Plasma Sprayed MCrAlY Coatings," ASME J. Eng. Gas Turbines Power, **121**, pp. 476–483.

[11] Heumann, T., 1952, "Zur Berechnung von Diffusions Koeffizienten bei Ein- und Mehrphasiger Diffusion in Festen Legierungen," Zeit. Physik. Chem., **201**, pp. 168–189.

[12] Wagner, C., 1956, "Reaktionstypen bei der Oxydation von Legierungen," Z. Elektrochem., **63**, No. 7, pp. 772–782.

[13] Kidson, G. V., 1961, "Some Aspects of the Growth of Diffusion Layers in Binary Systems," J. Nucl. Mater., **3**, No. 1, pp. 21–29.

# CFD Predictions and Field Measurements of $NO_x$ Emissions From LM1600 Gas Turbine During Part Load Operation

**G. R. Price**

**K. K. Botros**

NOVA Research & Technology Centre,
Calgary, Canada

**G. M. Goldin**
Fluent Inc.,
Lebanon, NH 03766

*CFD simulations of the combustion process and formation of emissions in an industrial GE LM1600 gas turbine have been performed over a range of unit loads. Two combustion models were considered here to characterize the combustion process, the chemical equilibrium model and the nonequilibrium laminar flamelet model. The flamelet model predictions of nitric oxide concentration demonstrated much closer agreement with field measurements of continuous emission monitoring systems, due to accurate modeling of the oxygen radial concentration. The predictions made with this model are within 15 percent at maximum load and considerably better at lower loads. Field measurements also showed that nitric oxide constitutes about 95 percent of the total $NO_x$ measured. Unburnt hydrocarbons and carbon monoxide emissions are significantly overpredicted, however, arguably as a result of neglecting their oxidation in the high temperature, fuel-lean environment of the turbine and exhaust stack downstream of the combustor.*
[DOI: 10.1115/1.1413769]

## Introduction

Air pollution control programs in the United States increasingly require industry to continuously monitor and report emissions of pollutants, specifically oxides of nitrogen ($NO_x$) and sulfur ($SO_x$). These components are generally acknowledged as the major contributors to acid rain, while $NO_x$ also contributes along with volatile organic compounds to the photochemical generation of ground level ozone or smog.

Current U.S. monitoring requirements are generally biased towards continuous emissions monitoring systems (CEMS) based on calibrated measurement devices. For example, the regional clean air incentives program (RECLAIM) ([1]) in southern California imposes stringent CEMS and reporting requirements for $NO_x$ and $SO_x$ emissions, particularly for large $NO_x$ sources. On a federal level, the U.S. Environmental Protection Agency stipulates CEMS for installations specified under the acid rain program ([2]).

The situation is quite different in Canada, where predictive models or one-time site measurements are employed to characterize these emissions as part of the initial licensing application to the provincially based environmental regulatory bodies. Continuous monitoring of emissions is generally not a licensing requirement. Some proactive initiatives have been introduced in recent years including the Voluntary Challenge and Registry Program ([3]), which encourages private and public sector organizations across Canada to voluntarily limit and disclose their net greenhouse gas emissions. However, further regulatory changes with some form of ongoing emission monitoring could be tabled in the near future in order to meet Canada's Kyoto commitment to reduce greenhouse gases to 3 percent below 1990 levels by the year 2010.

TransCanada Transmission ([4]), which operates numerous industrial gas turbines as part of its natural gas transmission pipeline, has sponsored a program of predictive emissions monitoring systems (PEMS) in anticipation of future regulatory changes. PEMS are often based on semi-empirical expressions or one-dimensional models in an attempt to correlate emission rates to operating parameters for a given gas turbine ([5–7]). Suffice it to say, PEMS offer a very cost-effective means for estimating continuous emission levels with minimal operational and maintenance requirements, i.e., no calibration gases or expensive analyzers. A good account of the benefits of PEMS over CEMS is given by Hung ([5]), where a quantitative comparison between the two methods is provided.

As a part of this larger PEMS program, this study focuses on the ability of more sophisticated tools such as computational fluid dynamics (CFD) to accurately predict $NO_x$ emissions. Clearly, due to the complexity and large computational requirements, CFD is not feasible for direct application as a PEMS. However, it can be used to provide some valuable data that can be implemented in PEM models.

In this paper, CFD predictions of the nitric oxide (NO) emissions from a stationary General Electric LM1600 gas turbine engine are presented. We begin with a brief review of related CFD studies of gas turbine (GT) combustion chambers. This is followed by details of the CFD modeling effort with particular emphasis on the combustion process and NO models along with the combustor geometry, grid generation, and prescribed boundary conditions. Two popular combustion models were considered in the simulations, both based on the mixture fraction probability density function (PDF) approach. The first model assumes chemical equilibrium, whereas the second laminar flamelet based model accounts for nonequilibrium effects. The results of the CFD simulations are compared with field measurements of the unit emissions taken over a range of part load operating conditions on two separate occasions.

## Related CFD Studies

Many researchers have used CFD simulations to study the fundamental combustion process and heat transfer characteristics in gas turbine combustors ([8–10]). In fact, aeroengine manufacturers now commonly apply CFD during the design of new combustion chambers. Mongia ([10]) provides an excellent review of

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-350. Manuscript received by IGTI Nov. 1999; final revision received by ASME Headquarters Feb. 2000. Associate Editor: D. Wisler.

the application of CFD to gas turbine combustors along with some recommendations for future research from an industry perspective.

Some noteworthy CFD related $NO_x$ studies include a series of papers with detailed experimental and numerical results for a propane-fired laboratory-scale can combustor ([11–14]). Isothermal flow simulations by Koutmos and McGuirk ([13]) demonstrated that the standard $k$-$\varepsilon$ turbulence model was acceptable for predicting the velocity field despite some deficiencies in the jet-on-jet impingement region. More recently, Mueller and Knill ([14]) compared the effect of two different combustion models; namely the eddy dissipation model ([15]) and the laminar flamelet model ([16]), on predicted temperature field, species concentration and NO emissions. Their $NO_x$ model was based on the extended Zeldovich mechanism to describe thermal NO formation along with an empirical expression by De Soete ([17]) to account for the prompt NO formation. They concluded that simple combustion models like the eddy dissipation model tend to overpredict the temperature field and NO concentrations, denoted by [NO], due to the lack of intermediate species. Conversely, the flamelet model, which better accounts for nonequilibrium effects, was much more accurate.

Further advances in CFD modeling of pollutant formation in gas turbine combustors are reported by Barths et al. ([18]). These authors introduced a series of unsteady flamelets based on a complex reaction mechanism involving 106 reactions and 14 species to describe the thermal, prompt and nitrous oxide contributions to $NO_x$ formation and account for $NO_x$ reburn. For n-hexane fuel, the agreement between predicted and measured NO values was excellent, within 12 percent. An analysis of the reaction paths showed that the thermal NO contributes 67 percent to the total production with the remaining 33 percent formed via the prompt mechanism. NO was globally consumed by a reburn mechanism (22 percent) and the nitrous oxide path (8 percent).

## Description of CFD Model

Numerical simulations of the time-averaged turbulent flow field in the combustion chamber of the LM1600 were conducted using FLUENT 5.1 ([19]), a commercial CFD package. Calculations of the Reynolds number for the fuel, primary and secondary air injection indicate the flow conditions are fully turbulent. FLUENT solves the incompressible time-averaged governing equations for mass and momentum using a conservative control volume formulation. Closure of the Reynolds stress terms is achieved using the standard $k$-$\varepsilon$ turbulence model. In spite of its inherent assumptions, i.e., isotropy, the $k$-$\varepsilon$ model has been widely utilized in combustor simulations and generally yields acceptable results ([10,13,14,18]).

**Mixture Fraction Approach.** The mixture fraction approach was developed to model turbulent nonpremixed flames. The mixture fraction, commonly denoted by $f$, is defined as the atomic mass fraction that originated from the fuel stream. Under the assumptions of uniform pressure (which is valid for low speed flames), and equal turbulent species diffusivities (which is valid at high Reynolds number), the species equations can be reduced to a single equation for $f$. Since atomic elements are conserved during chemical reactions, the mixture fraction equation does not have a source term.

In turbulent combustion, all flow variables at a point fluctuate in time due to turbulent eddies. Turbulence is modeled in this simulation by Reynolds averaging, which may be considered as a time average of the fluctuating quantities at a point. The effects of this time varying turbulent field on the mixture fraction are accounted for by using an assumed shape for the PDF. The PDF, denoted by $P$, describes the proportion of time spent at each turbulent state. Hence, mean quantities, such as the mean temperature (means are denoted by a "$\sim$" overbar) are determined by a PDF weighted integration over all states,

$$\tilde{T} = \int_0^1 T(f)P(f)df. \tag{1}$$

Here, a beta PDF is used, which is specified by its first two moments, namely the mean of $f$ and its variance. The CFD code solves equations for these first two moments, from which the mean temperatures and species can be computed according to (1)). In order to avoid the computational expense of evaluating these integrals, they are calculated a priori, tabulated and then interpolated from the table during the actual CFD simulations.

**Equilibrium Model.** When chemical reactions are fast relative to turbulent mixing, the chemistry may be approximated as being in chemical equilibrium. For adiabatic flames, the entire temperature and species fields are completely determined by the mixture fraction; nonadiabatic flames also depend on the enthalpy. As the residence time in the combustor decreases, however, the turbulent mixing time-scales become competitive with the chemical reaction time-scales and the chemistry deviates from thermodynamic equilibrium. The flamelet model relaxes this assumption to cases of moderate nonequilibrium.

**Flamelet Model.** The laminar flamelet model approximates a turbulent flame brush as an ensemble of laminar flames. Each flamelet is assumed to have the same structure as the stretched laminar flame that occurs between opposing fuel and air nozzles. For this configuration, the species and temperature in the flame can be completely described by two parameters, namely the mixture fraction, $f$ and the scalar dissipation rate, $\chi$. The scalar dissipation rate is a measure of the strain on the flame, and hence the extent that it departs from chemical equilibrium ([19]),

$$\chi = 2D\left(\frac{df}{dx}\right)^2. \tag{2}$$

Individual laminar flamelets can be calculated from equations governing individual species transport and an overall energy balance,

$$\rho\frac{\partial Y_j}{\partial t} = \frac{1}{2}\rho\chi\frac{\partial^2 Y_j}{\partial f^2} + \dot{w}_i - \frac{1}{2}\left[\rho\chi\frac{1}{\mathrm{Le}_j^2}\frac{\partial \mathrm{Le}_j}{\partial f} + \frac{1}{2}(1-\mathrm{Le}_j) \times\left(\frac{\partial \rho\chi}{\partial f} + \rho\chi\frac{c_p}{\lambda}\frac{\partial}{\partial f}\frac{\lambda}{c_p}\right)\right]\frac{\partial Y_j}{\partial f} \tag{3}$$

$$\rho\frac{\partial T}{\partial t} = \frac{1}{2}\rho\chi\frac{\partial^2 T}{\partial f^2} + \frac{1}{2}\rho\chi\left[\frac{\partial c_p}{\partial f} + \sum_j \frac{1}{\mathrm{Le}_j}c_{p_j}\frac{\partial Y_j}{\partial f}\right] - \frac{1}{c_p}\sum_j \dot{w}_j h_j - \frac{1}{c_p}\left[4\sigma p\sum_j X_j a_j (T^4 - T_b^4)\right]. \tag{4}$$

Note that the last term in the energy Eq. (4) represents heat loss from the flamelet due to radiation in the optically thin limit. The flamelet Eqs. (3) and (4) are solved over the range of scalar dissipation rates, integrated over the beta PDF according to Eq. (1) and tabulated to relate mean scalar quantities to the mixture fraction, its variance, and the scalar dissipation rate.

**$NO_x$ Model.** Since $NO_x$ occurs in trace quantities, it does not influence the flow field, and is accordingly post-processed. In this study, only NO is considered. The governing equation for the transport of NO is given by

$$\frac{\partial}{\partial t}(\rho Y_{\mathrm{NO}}) + \frac{\partial}{\partial x_i}(\rho u_i Y_{\mathrm{NO}}) = \frac{\partial}{\partial x_i}\left(\rho D_{\mathrm{eff}}\frac{\partial Y_{\mathrm{NO}}}{\partial x_i}\right) + \widetilde{w_{\mathrm{NO}}} \tag{5}$$

where $u_i$ is the fluid velocity and $D_{\mathrm{eff}}$ is the effective (turbulent+laminar) diffusivity. The last term in Eq. (5), the mean NO source term, is computed via Eq. (1) based on the instantaneous NO reaction rate and an assumed beta PDF. The instantaneous NO reaction rate has contributions from thermal and prompt sources. The thermal pathway is modeled by the extended

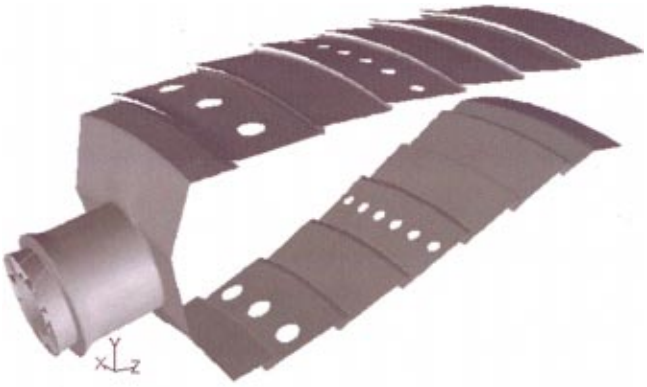


**Fig. 1 Computational domain—20 deg annular secton of LM1600 combustion chamber**

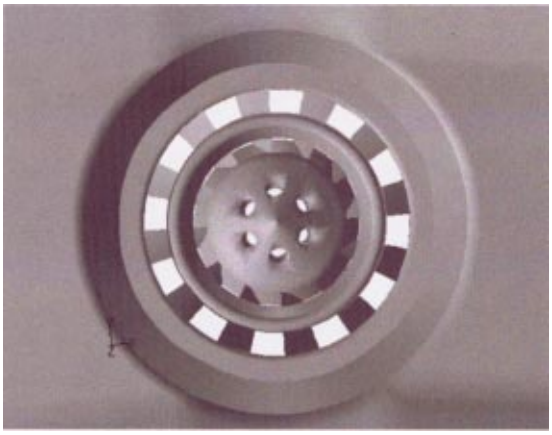


**Fig. 2 Close up of the fuel injection port, primary air swirl vanes, and venturi nozzle details**

Zeldovich mechanism with the steady-state assumption for the nitrogen radical. The prompt pathway is modeled via the empirical De Soete mechanism ([17]). Owing to the relatively high temperatures in the combustor, the thermal pathway is the predominant mechanism for NO formation.

**Geometry, Grid Generation, and Boundary Conditions.** The LM1600 gas turbine installed at the TransCanada compressor station is rated for 13.6 MW at standard site conditions. The annular combustion module consists of 18 fuel injection nozzles equally spaced about the periphery. The inherent geometric symmetry of the combustor was exploited by considering a 20-deg sector with a single fuel injection nozzle for the computational domain and employing periodic boundary conditions. The fuel port consists of a tube capped by a conical head with six small holes. Primary air enters the combustion chamber through a concentric annulus surrounding the fuel port. The primary air is swirled by twelve, 30 deg swirl vanes in the annulus before passing through a venturi nozzle and subsequently mixing with the natural gas jets. Dilution air is introduced further downstream through a series of holes and slots located along the top and bottom wall of the combustion liner. The combusted gases exit through an annular outlet, which leads to the multistaged high-pressure turbine. An overview of the 20-deg section of combustion chamber geometry is shown in Fig. 1. A close up front view of the fuel injection nozzle assembly, i.e., fuel injection port, primary air swirl vanes and venturi nozzle, is shown in Fig. 2.

A block-structured grid consisting of about 286,000 hexahedral control volumes with body-fitted coordinates was generated for the three-dimensional computational geometry. The boundary surfaces were uniquely identified during this process for later specification of boundary conditions. Second-order upwind differencing was employed for the convective terms in the discretized conservation equations in order to minimize numerical diffusion and achieve reasonably accurate solutions for this mesh size. Although a comprehensive grid refinement study was not performed, tests with coarser and slightly finer meshes showed only minor changes in predicted results. Unfortunately, complete grid independence is generally beyond available computing resources for most three-dimensional combustion problems.

The boundary conditions for the CFD simulations include periodic boundaries to connect the azimuthally rotated side planes, a constant pressure outlet, thin walls for the swirl vanes, and finally adiabatic, hydraulically smooth walls for the external surfaces such as the combustion liner and venturi nozzle. The mass flow rate of fuel, which varies with unit load, was specified from direct measurements with a conventional orifice meter.

Specification of the airflow boundary conditions requires some additional explanation, however, because the flow rate of air is not directly measured by the on-line data acquisition system that monitors the engine performance. Single point measurements of the exhaust gas velocity with a Pitot tube during the field tests did not provide reasonable estimates of the airflow. Instead, the air flow rate was indirectly determined by dividing the reported power output by the enthalpy change across the power turbine; the enthalpy change being derived from the measured pressure and temperature at the unit inlet and outlet. The actual breakdown of the total air flow rate into prescribed mass flow rates at the primary air inlet and the various secondary air slots and holes, shown in Fig. 1, was based on proprietary information supplied by General Electric.

## Field Measurements

Field measurements were taken in order to determine the variation of emission levels from an LM1600 at various loads, i.e., engine power output. The tests were performed at the same TransCanada compressor station on two occasions, separated by almost one year. The ambient conditions at the site on these dates were quite similar: the air temperature for the first test ranged from 6–9°C compared to 11–13°C for the second test, while the barometric pressures remained within 1 percent.

The levels of $NO_x$, NO, CO, $CO_2$, $O_2$ and moisture in the exhaust stack were measured by a CEMS using instruments and on-site calibration techniques approved by U.S. EPA. For example, the $NO_x$ measurements were made with an approved chemiluminescent continuous analyzer. Pertinent engine performance parameters such as the unit power output, fuel flow rate, turbine stage pressures and temperatures, and ambient conditions at the site were recorded by the data acquisition system at the station. Gas samples were also taken in order to determine the composition of the fuel gas. The results of the gas analysis are summarized below in Table I.

At each operating point, the unit speed was adjusted to reach the desired power output. Then the unit was allowed to stabilize before the CEMS measurements were taken over a 20-minute pe-

**Table 1 Natural gas composition**

| Component | Field Test #1 Mole % | Field Test #2 Mole % |
|---|---|---|
| $CH_4$ | 96.92 | 96.75 |
| $C_2H_6$ | 0.86 | 0.80 |
| $C_3H_8$+ | 0.30 | 0.38 |
| $N_2$ | 1.03 | 1.12 |
| $CO_2$ | 0.89 | 0.95 |

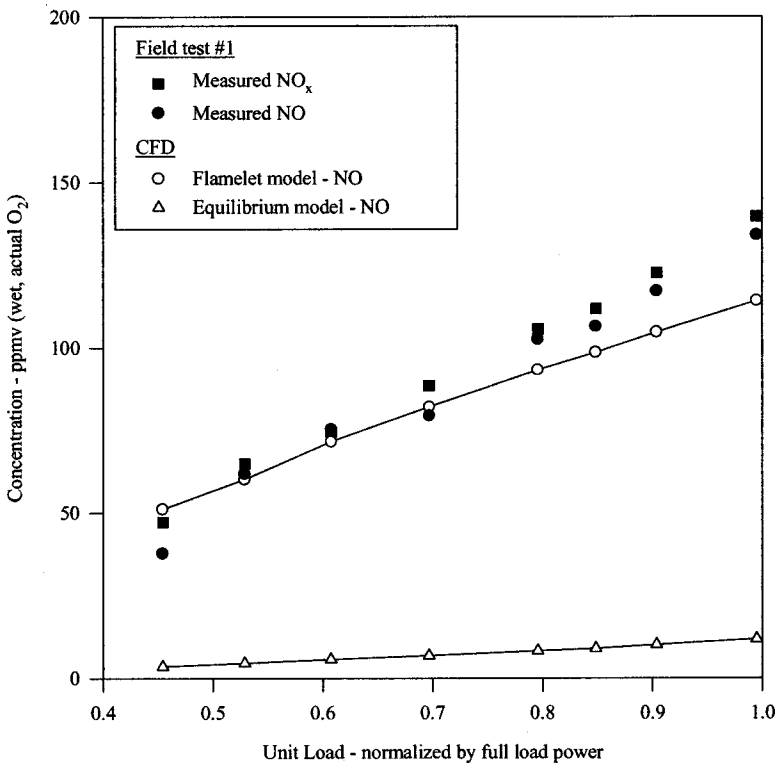



**Fig. 3 Comparison of predicted (NO) with first set of CEMS based field measurements**

riod. The unit load was varied from roughly 40 percent to maximum load to give eight operating points on each test date, for a total of 16 points.

## Results and Discussion

CFD simulations were performed at each of the 16 operating points corresponding to the field measurements. At each operating point, multiple laminar flamelet files were generated to cover the range of strain rates in the combustor. Property look-up tables were also generated for both combustion models at each operating point. This was necessary because the absolute pressure and incoming compressed air temperature vary directly with the air compression ratio, which declines with decreasing unit load. Due to the extremely high methane content of the fuel gas, ~97 per-

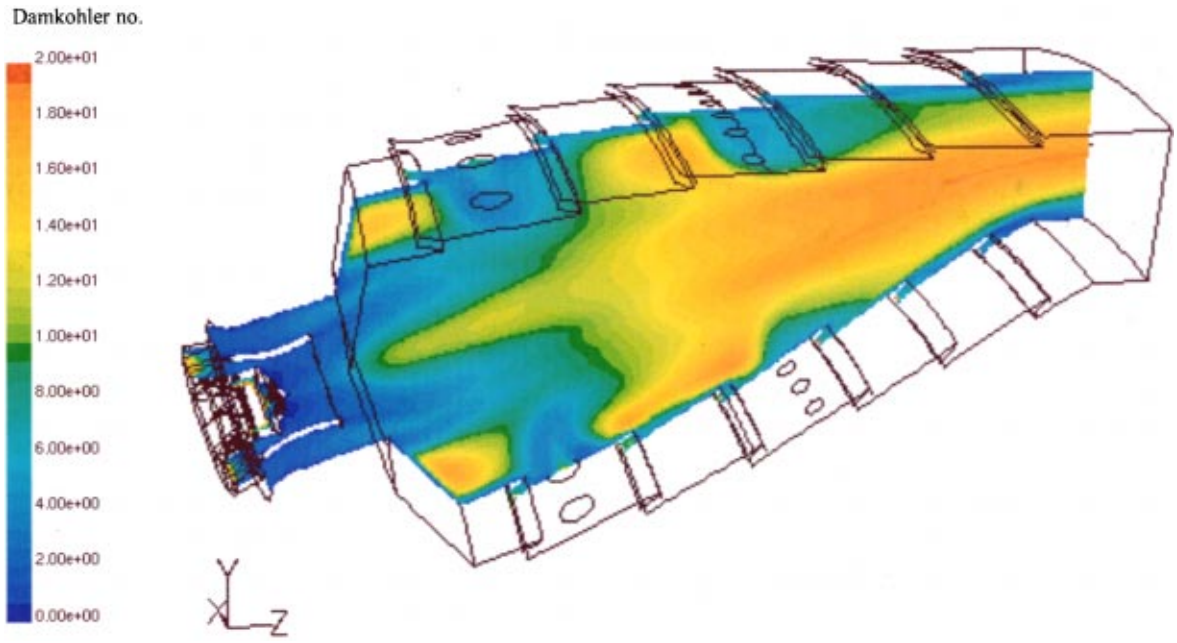



**Fig. 4 Contours of the local Damkohler number on the central plane**

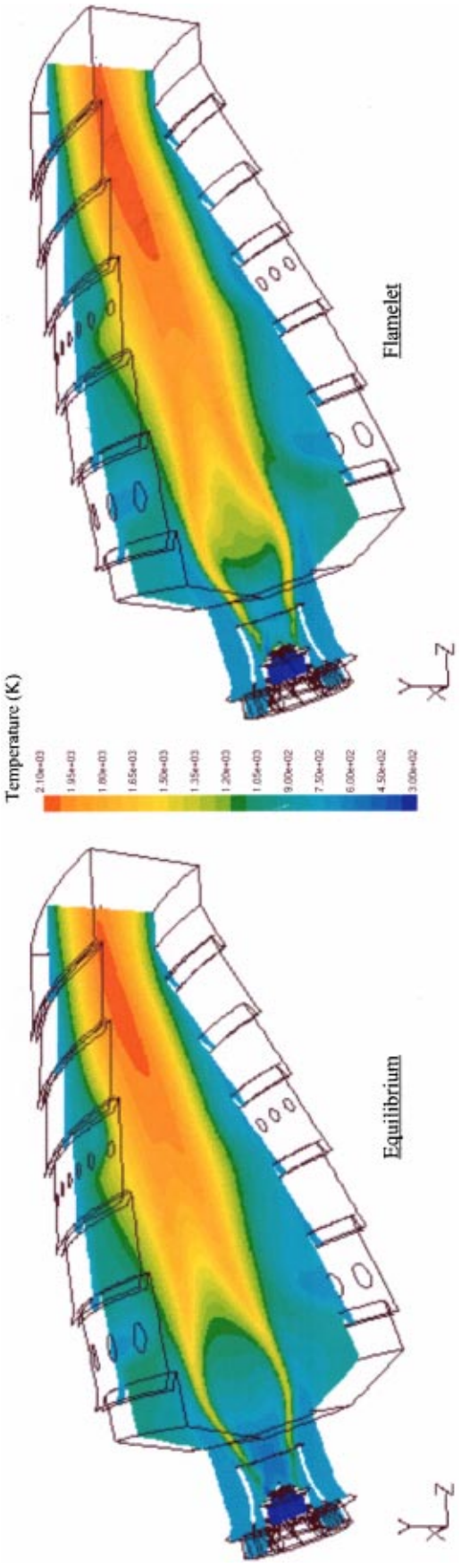



Fig. 5 Contours of static temperature on the central plane

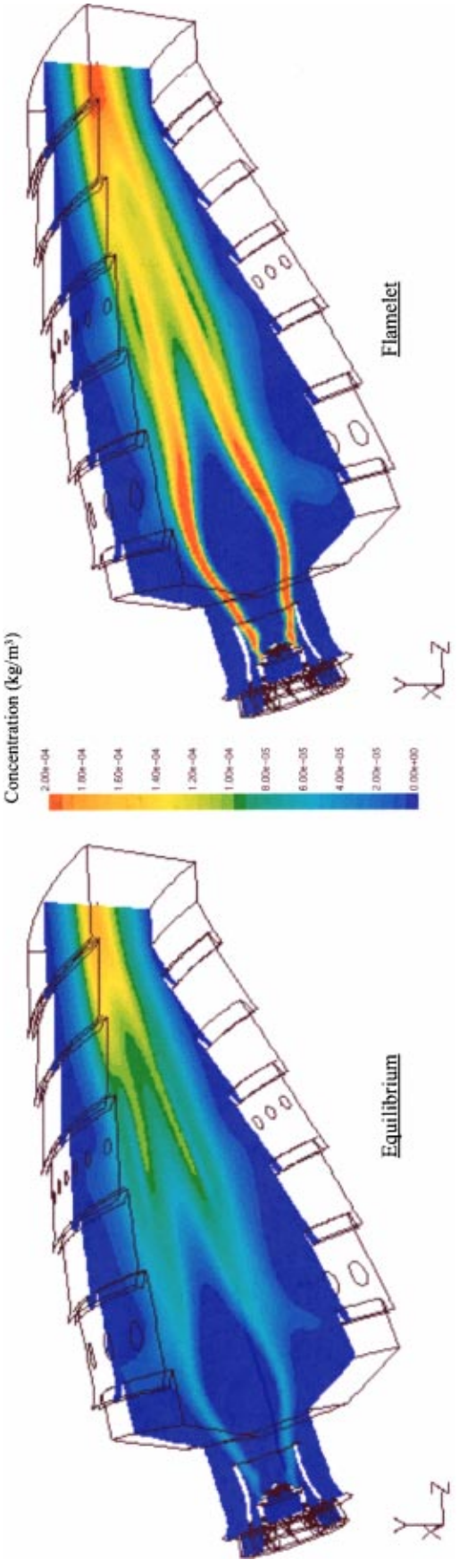



Fig. 6 Contours of oxygen radical concentration on the central plane

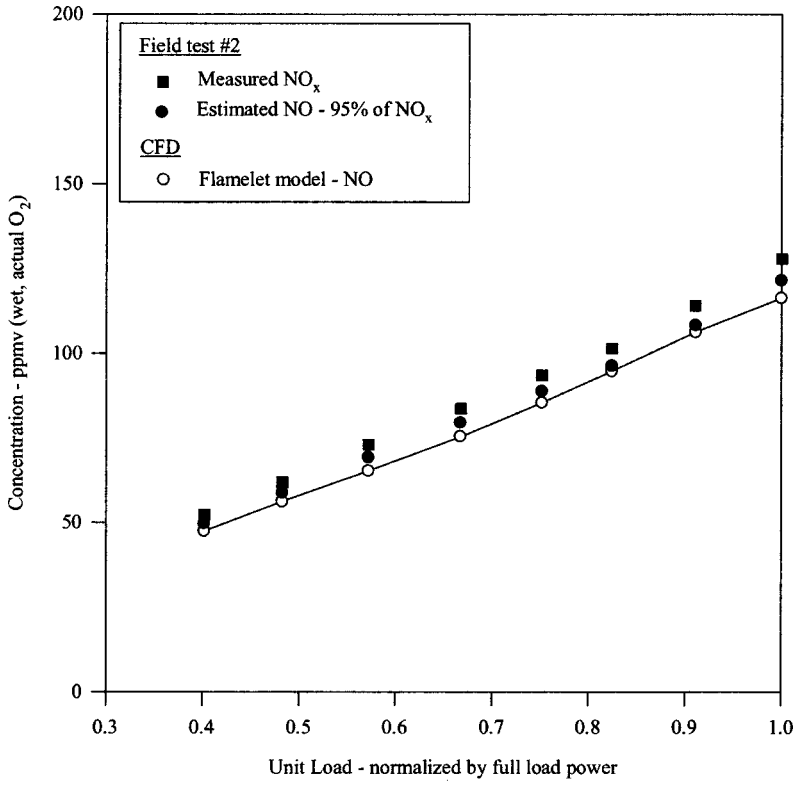



**Fig. 7 Comparison of predicted (NO) with second set of CEMS based field measurements**

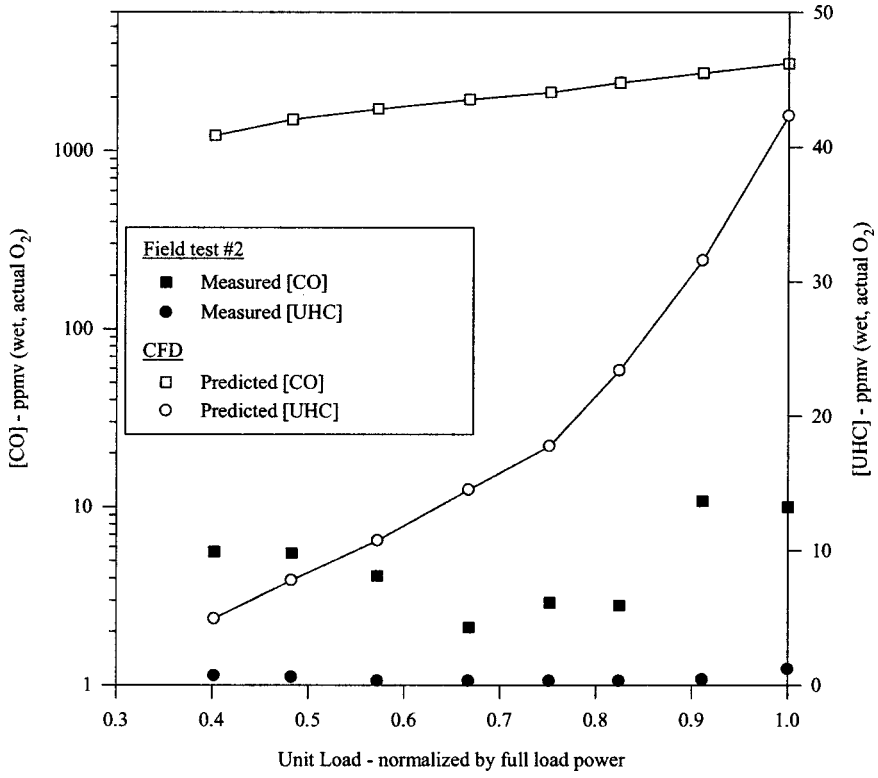



**Fig. 8 Comparison of predicted (CO) and (UHC) with second set of CEMS based field measurements**

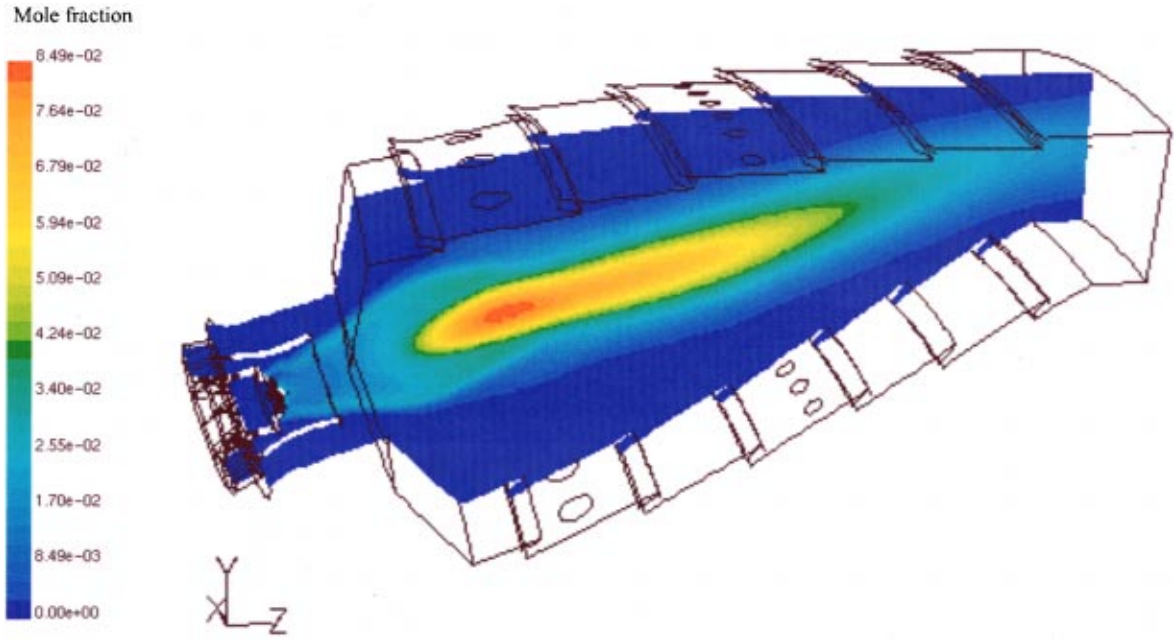



**Fig. 9 Contours of CO mole fraction on the central plane**

cent in Table 1, the reduced GRI-MECH 1.22 combustion mechanism ([20]) was implemented in the present analysis. This simplified methane mechanism considers 104 reactions and 22 species.

The predicted concentrations of NO from CFD were obtained by dividing the integrated volumetric flow rate of NO by the total integrated volumetric flow rate over the combustor outlet plane for the two different combustion models. The predicted concentrations were subsequently corrected to account for additional cooling air that bypasses the combustion chamber. These predictions are compared with the first set of CEMS measurements in Fig. 3. Note that NO was also measured during this first test along with the total $NO_x$ to allow direct comparison with the CFD results. With the exception of two points in Fig. 3, NO accounts for about 95 percent of the total $NO_x$ produced. The flamelet model results are clearly superior to the chemical equilibrium model, which under predicts the NO concentrations by an order of magnitude. The flamelet model only under predicts the NO concentrations by 15 percent at full load. Although this discrepancy diminishes at part loads, the rate of increase for the measured NO with increasing load is somewhat steeper than the predicted trend. The overall trend of increasing NO with increasing load is due to the higher temperatures in the combustor, which leads to additional NO via the thermal mechanism. Note these higher combustor temperatures result from higher compressed air temperatures and lower air-to-fuel ratios as the unit load is increased.

The inadequacy of the equilibrium based CFD model in this situation can be elucidated by considering the deviation from chemical equilibrium as measured by the Damkohler number, Da, which is defined as the ratio of characteristic turbulent time scale, $k/\varepsilon$, to chemical time scale, $1/b_s$. Here $b_s$ is the laminar strain rate of a methane flame at extinction, $b_s = 11{,}700\ \mathrm{s}^{-1}$. Contours of the local Damkohler number are plotted on the central plane, i.e., vertical plane dissecting the computational domain, in Fig. 4. Ideally, this dimensionless parameter should be much greater than unity, or in other words the turbulent time scale should be much longer than the chemical time scale, for chemical equilibrium to be valid. This is clearly not the case in Fig. 4, where $5 < \mathrm{Da} < 20$ over a majority of the combustion chamber. Hence, there is insufficient time for chemical equilibrium to prevail.

Further insight into this problem is gained by comparing the temperature and oxygen radical concentrations for these two combustion models. In Fig. 5, contours of temperature on the central plane show only slight deviations between the two cases, although the high temperature region is larger for the flamelet model. The primary difference lies in the predicted concentration of oxygen radical [O] as shown in Fig. 6. Here, the flamelet model yields significantly higher [O] than predicted by equilibrium chemistry, particularly in the thin flame region near the injection nozzle. Higher [O] increase the thermal NO formation rate, which is directly proportional to the concentration of oxygen radial as shown by Mueller and Knill ([14]).

CFD simulations based on the second set of field measurements were performed only with the flamelet model. The integrated [NO] at the combustor outlet are compared with the average CEMS measurements in Fig. 7. Unfortunately, nitric oxide measurements were unavailable for this set of tests, so the reported [NO] in Fig. 7 was estimated at 95 percent of the measured $NO_x$ based on the results of our previous set of field tests. Again, flamelet based CFD results slightly underpredict the measured trend but it is within 10 percent of the measured $NO_x$ and within 5 percent of the estimated NO. This level of agreement is surprisingly good and the discrepancy can easily be attributed to additional NO formation via residence time during the time the gas leaves the combustion chamber and its subsequent expansion in the high-pressure turbine stage.

We also compared the CFD predictions of the unburned hydrocarbons (UHC) and carbon monoxide at the combustor outlet with the CEMS field measurements for the second set of tests in Fig. 8. Clearly the agreement between the predicted and measured values of UHC and CO is poor, particular for CO where the predicted levels are almost three orders of magnitude higher than the measurements. The inability of CFD to accurately predict these components was recognized by Nicol et al. ([21]).

Direct measurements of the [CO] and [UHC] at the outlet of a laboratory combustor chamber were reasonably well predicted by a flamelet-based model in Ref. [14]. Hence, we offer that these order of magnitude discrepancies can be explained by the discrepancy in the locations from which the CFD and measured values are derived. From Fig. 9, which shows the CO mole fraction, it appears that CO is produced in the fuel rich region of the flame. Further downstream in the fuel lean zone where the temperatures are higher (refer to Fig. 5) it is oxidized into $CO_2$. This oxidiza-

tion process is relatively slow and will continue beyond the combustor, through the downstream turbine and even into the exhaust stack, where the pressure is atmospheric and the temperature ranges from 420–492°C depending on the unit load. Chemical equilibrium calculations show that negligible amounts of CO would be present under these conditions and the residence time to reach the measurement location on the exhaust stack is on the order of several seconds versus milliseconds in the combustor. Of course, better mixing could also reduce the UHC and CO emissions in the actual unit.

One last comment, the mass flow rate of air is an important parameter in the NO predictions but the indirect calculation method used to derive it is prone to inaccuracies due to uncertainties in the temperature measurements across the turbine. This is particularly true for an engine with a single turbine because the Turbine Inlet Temperature (TIT) is notoriously difficult to measure accurately due to hot/cold streaks in the flow from the combustion chamber. The LM1600 consists of two separate turbines: a high-pressure turbine to drive the air compression and a low-pressure turbine for power generation. Fortunately, the temperature measurements across the low-pressure turbine are significantly more accurate than the TIT and these are used in the air flow calculation. An additional CFD simulation at full load conditions revealed that a 10 percent increase in the total air flow rate reduced the predicted [NO] by 13 percent. This decrease is largely due to effect of increased air dilution along with slightly enhanced mixing in the combustion chamber. The sensitivity analysis also provides a gross estimate of the uncertainty in [NO] associated with the inaccuracy in the calculated flow rate of air through the combustion chamber.

## Concluding Remarks

The following conclusions can be drawn from the present study:

1 The predicted [NO] with the flamelet model was in much closer agreement with CEMS field measurements than the chemical equilibrium model. This is likely because it more accurately predicts the [O] within the combustor, which is an important factor in thermal NO mechanism.

2 Even with all the extra details and efforts incorporated in the flamelet model, the best CFD predictions were shown to be within 15 percent at maximum load.

3 The field measurements showed that NO constitutes about 95 percent of the total $NO_x$ measured in this type of engine.

4 UHC and CO emissions were greatly overpredicted, arguably as a result of neglecting their post-combustor oxidation in the high-temperature fuel-lean environment of the turbine and exhaust stack.

## Acknowledgments

This work was conducted as part of a research program sponsored by TransCanada Pipeline Ltd. (TCPL) and permission to publish it is gratefully acknowledged. We would also like to acknowledge Entect Environmental Services led by Mrs. Jackie Hudson and Conor Pacific Environmental Technologies Inc., particularly Mr. Sean Miner and Mr. Francis Lucas, for the CEMS measurements. Thanks are also due to Dr. Golam Kibrya and Mr. Gene Bayrock, both of TCPL, for their assistance in coordinating and monitoring the field tests.

## Nomenclature

$a$ = absorption coefficient
$b_s$ = laminar flamelet strain rate at extinction
$c_p$ = specific heat
$D$ = diffusion coefficient
Da = Damkohler number, $\mathrm{Da}=b_s k/\varepsilon$
$f$ = mixture fraction
$k$ = turbulent kinetic energy
Le = Lewis number, $\mathrm{Le}=\rho c_p D/\lambda$
$p$ = static pressure
$P$ = probability density function
$T$ = static temperature
$u_i$ = velocity components, $i=1,2,3$
$\dot{w}$ = reaction rate
$X$ = mole fraction
$Y$ = mass fraction

### Greek Symbols

$\chi$ = scalar dissipation rate
$\varepsilon$ = turbulence dissipation rate
$\lambda$ = thermal conductivity
$\rho$ = fluid density
$\sigma$ = Stefan-Boltzmann constant

### Subscripts

$b$ = background conditions
$j$ = current species

## References

[1] South Coast Air Quality Management District, details of RECLAIM program at web site, http://www.aqmd.gov/
[2] Environment Protection Agency—Acid Rain Program, details at web site, http://www.epa.gov/acidrain/
[3] Voluntary Challenge and Registry Program, details at official web site, http://www.vcr-mvr.ca/
[4] TransCanada Transmission Ltd, details at official web site, http://www.transcanada.com/
[5] Hung, W. S., 1991, "A Predictive $NO_x$ Monitoring System for Gas Turbines," ASME Paper No. 91-GT-306.
[6] Hung, W. S., 1995, "Predictive Emission Monitoring System (PEMS): The Established $NO_x$ Monitoring System for Industrial Gas Turbines," Air & Waste Management Association Paper No. 95-MP16A.01.
[7] Sullivan, D. A., 1997, "A Simple Gas Turbine Combustor $NO_x$ Correlation Including the Effect of Vitiated Air," ASME J. Eng. Gas Turbines Power, **119**, pp. 145–152.
[8] Tolpadi, A. K., Hu, I. Z., Correa, S. M., and Burrus, D. L., 1997, "Coupled Lagrangian Monte Carlo PDF-CFD Computation of Gas Turbine Combustor Flowfields With Finite-Rate Chemistry," ASME J. Eng. Gas Turbines Power, **119**, pp. 519–526.
[9] Danis, A. M., Burrus, D. L., and Mongia, H. C., 1997, "Anchored CCD for Gas Turbine Combustor Design and Data Correlation," ASME J. Eng. Gas Turbines Power, **119**, pp. 535–545.
[10] Mongia, H. C., 1994, "Combustor Modeling in Design Process: Applications and Future Directions," AIAA Paper No. 94-0466.
[11] Jones, W. P., and Toral, H., 1983, "Temperature and Composition Measurements in a Research Gas Turbine Combustion Chamber," Combust. Sci. Technol., **31**, pp. 249–275.
[12] Heitor, M. V., and Whitelaw, J. H., 1986, "Velocity, Temperature and Species Characteristics of the Flow in a Gas Turbine Combustor," Combust. Flame, **64**, pp. 1–32.
[13] Koutmos, P., and McGuirk, J. J., 1991, "Isothermal Modeling of Gas Turbine Combustors: Computational Study," J. Propul. Power, **7**, No. 6, pp. 1064–1071.
[14] Mueller, C. M., and Knill, K. J., 1996, "Numerical Simulation of Emissions in a Gas Turbine Combustor," *Proceedings of Combustion Canada '96*, National Resources Canada, Environment Canada, and National Research Council Canada, Ottawa, Canada.
[15] Magnussen, B. F., and Hjertager, B. H., 1976, "On Mathematical Models of Turbulent Combustion With Special Emphasis on Soot Formation and Combustion," *16th Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, pp. 719–729.
[16] Peters, N., 1984, "Laminar Diffusion Flamelet Models in Non-Premixed Turbulent Combustion," Prog. Energy Combust. Sci., **10**, pp. 319–339.
[17] De Soete, G. G., 1975, "Overall Reaction Rates of NO and $N_2$ Formation From Fuel Nitrogen" *15th Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, pp. 1093–1102.
[18] Barths, H., et al., 1998, "Simulation of Pollutant Formation in a Gas-Turbine Combustor Using Unsteady Flamelets," *27th Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, pp. 1841–1847.
[19] FLUENT™ Release 5.1, 1999, Fluent Inc., Lebanon, N.H.
[20] Kazakov, A., and Frenklach, M., "A Simplified 19 Species Methane Oxidation Mechanism," please refer to website at http://www.me.berkeley.edu/drm/ for further details.
[21] Nicol, D. G., et al., 1999, "Development of a Five-Step Global Methane Oxidation—NO Formation Mechanism for Lean-Premixed Gas Turbine Combustion," ASME J. Eng. Gas Turbines Power, **121**, pp. 272–280.

# A Demonstration of Artificial Neural-Networks-Based Data Mining for Gas-Turbine-Driven Compressor Stations

**K. K. Botros**
Nova Research and Technology Corporation,
Calgary, Alberta, Canada

**G. Kibrya**

**A. Glover**

TransCanada Pipelines, Ltd.,
Calgary, Alberta, Canada

*This paper presents a successful demonstration of application of neural networks to perform various data mining functions on an RB211 gas-turbine-driven compressor station. Radial basis function networks were optimized and were capable of performing the following functions: (a) backup of critical parameters, (b) detection of sensor faults, (c) prediction of complete engine operating health with few variables, and (d) estimation of parameters that cannot be measured. A Kohonen SOM technique has also been applied to recognize the correctness and validity of any data once the network is trained on a good set of data. This was achieved by examining the activation levels of the winning unit on the output layer of the network. Additionally, it would also be possible to determine the suspicious, faulty or corrupted parameter(s) in the cases which are not recognized by the network by simply examining the activation levels of the input neurons.*
[DOI: 10.1115/1.1414130]

## 1 Introduction

Within the area of stationary gas turbines, machine health monitoring, fault prognosis, and diagnosis rely on effective processing of data which are often unpredictable and imprecise. To tackle such tasks, current approaches tend to be based on some ''model'' of the engine in question, combined with a rule-based expert system. The biggest challenge with this approach is the knowledge base required to produce the rules. It requires extensive analysis of process and monitoring data, as well as documentation of the experience of a human expert. Both are extremely time-consuming. Additionally, real failures often don't follow exact rules and are not repeatable which impair the effectiveness of a rule-based system.

Artificial neural network (ANN, or simply NN) is an alternative means to present knowledge about the health of the engine. This tool has emerged during the past decade, from an obscure field that had been discredited by perceived inadequacies, into one of the fastest growing technologies for information processing ([1,2]). Unlike a ruled-based system, it is based on binary logic which can autonomously store knowledge by learning from historical data and has the characteristics of associative memory ([1]). Information about faults can be discerned by training the network on a set of data such as the state variables for the normal condition and those for identified faulty conditions. Neural computing systems are more adept at many classification and identification tasks than both traditional statistical and expert systems ([2]).

In gas turbine (GT) operation, few papers have applied NNs to machinery health monitoring and diagnosis. DePold and Gass [3] used ANNs in the development of GT condition monitoring system in three areas: (i) filtering of raw data to improve its quality, and (ii) trend change recognition based on parameter threshold, and (iii) combinations of above which lead to root cause identification. Another system for condition monitoring based on Kohonen self-organizing networks is described by MacIntyre et al. [4]. The system is also trained on vibration data recorded from machine operating at normal conditions. Kim et al. [5] demonstrated the capability of ANNs in identifying the location and severity of machinery faults, including multiple component faults.

An adaptive control system for a GT engine, which diagnoses conditions of axial compressor faults, was proposed and analyzed by Lombardo [6]. Zhang and Ganesan [7] used on-line vibration measurements as the diagnostic signal from which four variables, called indices, were derived for trend analysis using self-organizing mapping (SOM) and constrained topological mapping (CTM) based NN. Another specific operational application of ANN is in the detection of compressor stall, which has been demonstrated by Lo and Shi [8]. Similar work has been reported in [9] where the input signals were taken from the magnetic bearing controller which have proven to be sensitive to shaft disturbances.

It was also shown that a neural network could be successfully used in place of an actual model to estimate key unmeasured parameters needed for the operation of a GT. For example, Chbat [10] studied the combustion reference temperature, a parameter that is currently estimated via a nonlinear model inside the controller and is used in a number of critical mode-setting functions within the controller such as calculating the fuel-split between various manifolds. The neural network matches the accuracy of the current estimate; and it is more robust to errors in its internal parameters.

The potential applications of neural networks in the operation of pipeline facilities are numerous. A recent review of these application was conducted by Botros and Glover [11]. The present paper demonstrates the application of NN's in five specific areas of GT operation. These are

- backup of critical parameters,
- detection of sensor faults,
- prediction of complete engine operating parameters with few variables,
- estimation of parameters that cannot be measured, and
- Kohonen self-organizing map (SOM) for data mining.

Three sets of data were obtained from an RB211-driven compressor station on the TransCanada Pipelines (TCPL) system in Alberta, Canada. Each set represents one full year of operation,

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-351. Manuscript received by IGTI, November 1999; final revision received by ASME Headquarters, February 2000. Associate Editor: D. Wisler.

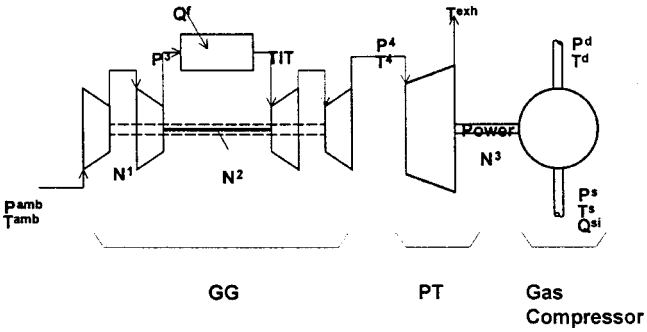



**Fig. 1 Schematic of a three-spool gas turbine driving a gas compressor**

namely 1996, 1997, and 1998. The first year was used for training the NN architecture while 1997 and 1998 sets were used for testing. Each set contains pertinent engine parameters designated on a sketch of a three-spool RB211 gas-turbine-based compressor station shown in Fig. 1.

## 2 Basic Concept of ANN

Most neural networks are based on feed-forward architectures with backpropagation optimization during training of the network. These architectures are commonly used because they are very versatile, and perform well on a wide variety of problems. In these networks, the processing elements are organized in layers. Neurons in one layer receive signals from neurons in the layer directly below and send signals to the neurons in layer directly above. Connections between neurons in the same layer are not allowed. The connection function is defined as a simple multiplication and an activation function $\Phi$, which is nonlinear and differentiable. The activation of a neuron in one of the layers is given by

$$z_{li}=\Phi\left(\sum_{j=1}^{M(l-1)} W_{lij}\cdot Z_{l-1j}\right)$$

where

$l$ = layer number
$i, j$ = neuron indices
$M(l)$ = total number of neurons in layer $l$
$z_{li}$ = output of neuron $i$ in layer $l$
$w_{lij}$ = weight between neuron $j$ in layer $l-1$ and neuron $i$ in layer $l$
$\Phi(\cdot)$ = function to be applied at each neuron.

A variety of optimization techniques can be used to find the weights ($w$). Let $J$ be the cost function of the optimization problem defined as

$$J=\frac{1}{2}\sum_{i=0}^{k}\sum_{j=0}^{n}(y_j(i)-y_j(i))^2.$$

For off-line computation the most common algorithms are back propagation, conjugate gradient descent, and quick propagation. Detailed account of these algorithms are given in [1]. Several papers survey the various types of networks, describing their architecture, learning algorithms and applications, e.g., Jain et al. [2] and Hunt et al. [12].

It should be worth mentioning that the independent variables used in traditional multi-regression analysis become the inputs and the dependent variable becomes the output. There are several key differences between ANN and traditional multi-regression analysis, however. In ANN the equation of a line or surface to fit the given data need not to be given a priori, i.e., the independent variables need not to be specified to interact in any particular fashion. Secondly, ANN can have more than one dependent variable (output) in one network interconnected with the input. Thirdly, it is always possible to get a tighter fit with ANN than with regression depending on the number of hidden neurons and error factor. Finally, ANN is able to function quite nicely with a huge number of training patterns, even those with ''noisy'' or slightly incorrect data in them.

## 3 Most Optimum NN Architecture

The training data for the present study were taken from the condition monitoring system of the station mentioned before for 1996. The data contains a total of 1844 cases representing a data point every four to five hours. Only 15 out of 132 measured parameters were selected for the present study: seven of which were input to the NN and eight output. The seven input parameters were considered to be the most thermodynamically representative parameters to describe the operating condition of the GT/centrifugal compressor set, while the output parameters were picked at random from the remaining parameters (Table 1).

Five NN architectures were considered (see Fig. 2) and the respective training algorithms were applied as shown in Table 2. The number of neurons in each layer is given under the structure in sequence indicating the number of neurons in the input layer, then in the first hidden layer, . . . , then output layer. The neurons in all output layers are linear with linear activation function. The hidden neurons in the multilayer perceptrons (MLP) NN's are linear type neurons with sigmoidal activation function. The hidden

**Table 1 Input/output parameters used in NN**

| Input | Output |
|---|---|
| Pamb | Qse |
| Tamb | P3 |
| N1 | P4 |
| N3 | T4 |
| Ts | Texh |
| Ps | Power |
| Pd | N2 |
| | Qf |

**Table 2 Overall normalized training RMS error for the various NN architectures of Fig. 2**

| NN Architecture | Structure | Training Technique | Overall Normalized RMS Error |
|---|---|---|---|
| MLP (Fig. 2(*a*)) | 7-14-8 | back propagation | 0.4695 |
| | | conj. grad. descent | 0.4213 |
| | | quick propagation | 0.7207 |
| MLP (Fig. 2(*b*)) | 7-28-8 | back propagation | 0.4309 |
| | | conj. grad. descent | 0.4958 |
| | | quick propagation | 0.4348 |
| MLP (Fig. 2(*c*)) | 7-14-14-8 | back propagation | 0.4396 |
| | | conj. grad. descent | 0.4205 |
| | | quick propagation | 0.9275 |
| RBF (Fig. 2(*d*)) | 7-128-8 | K-mean/K-nearest | 0.2410 |
| GRNN (Fig. 2(*e*)) | 7-1844-9-8 | sample/K-mean | 0.4068 |

**Table 3 RBF training accuracy**

| | Data Mean | Data S.D. | Error Mean | Error S.D. |
|---|---|---|---|---|
| Qse (inches H2O) | 578.9 | 119.2 | 3.41 | 36.31 |
| P3 (kPa-g) | 1464 | 173.2 | −1.338 | 16.22 |
| P4 (kPa-g) | 196.3 | 37.94 | −0.2336 | 4.294 |
| T4 (C) | 676.4 | 74.51 | −0.1619 | 8.469 |
| Texh (C) | 409.1 | 0.1614 | 0.02034 | 0.09148 |
| S.Power (kW) | 1.61E+04 | 4028 | −40.41 | 375.9 |
| N2 (rpm) | 8802 | 293.1 | −2.664 | 34.52 |
| Qf (m3/hr) | 5101 | 1206 | −58.38 | 525.3 |

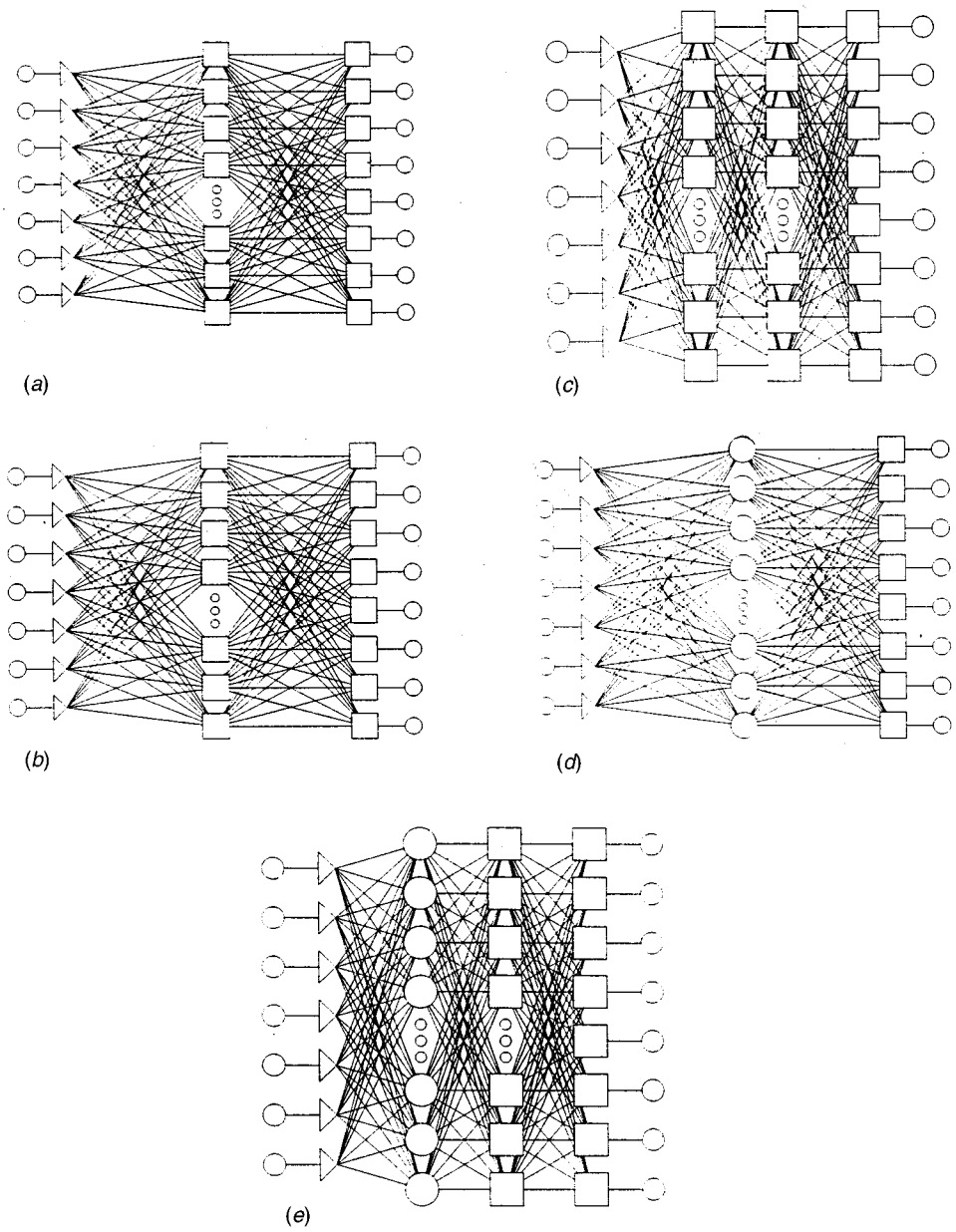



**Fig. 2 (*a*) MLP (7-14-8) architecture, (*b*) MLP (7-28-8) architecture, (*c*) MLP (7-14-14-8) architecture, (*d*) RBF (7-128-8) architecture, (*e*) GRNN (7-1844-9-8) architecture**

neurons in the radial basis function (RBF) NN's are of hypershere type with negative exponential activation function. The generalized regression NN (GRNN) is like RBF except it has a second hidden layer of linear type, and output layer's neurons of a division type.

Table 2 also shows the resulting normalized RMS error of the last iteration (often called "epoch") submitted to the respective NN after successful convergent training. It indicates that the RBF architecture of 7-128-8 structure is the most optimized NN to use in the present application, and hence was selected for the present investigation. The number of neurons in the hidden layer was determined following an iterative optimization process using various number of neurons. The resulting statistics for this NN in terms of the error in each of the output parameter (mean and standard deviation) is given in Table 3. It should be mentioned that the GRNN's though was extremely fast to train, were prone to overlearning and hence higher error when used for testing.

Using this RBF network architecture, the same data were then submitted to the network and the eight output parameters were compared to the respective measured parameters. The results of this comparison are shown in indicative Figures 3 (*a* through *h*) for an arbitrary time interval between cases #300 through #400 (representing approx. 500 hours). During this period, the station was idling for five hours (standby) which is represented by case #327. It is clear that the NN was able to predict the associated output parameters for this particular case very well. The fluctuations and trends of all 1844 cases in this data set were very well predicted by the NN in a manner similar to those demonstrated in these figures. Furthermore, in examining Fig. 3(*h*), it is shown that the NN was able to predict the fuel gas flow parameter at times (cases) when it appeared that the fuel flow sensor failed to provide a signal (see cases 411 through 433, and case 562, for example). This will be elaborated later.

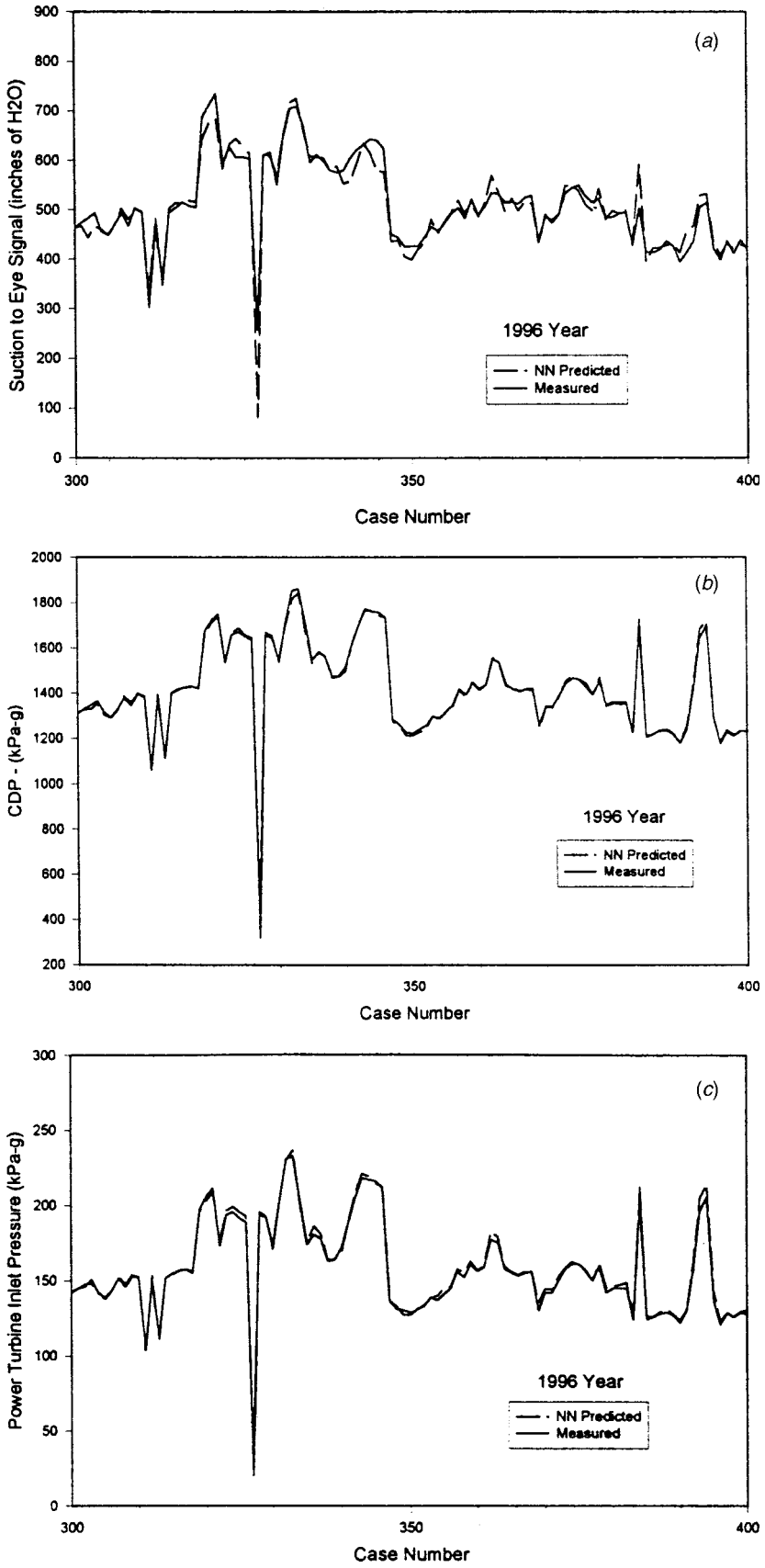



**Fig. 3 Results of testing the RBF network architecture of Fig. 2(*d*) using 1996 data set for training and then testing**

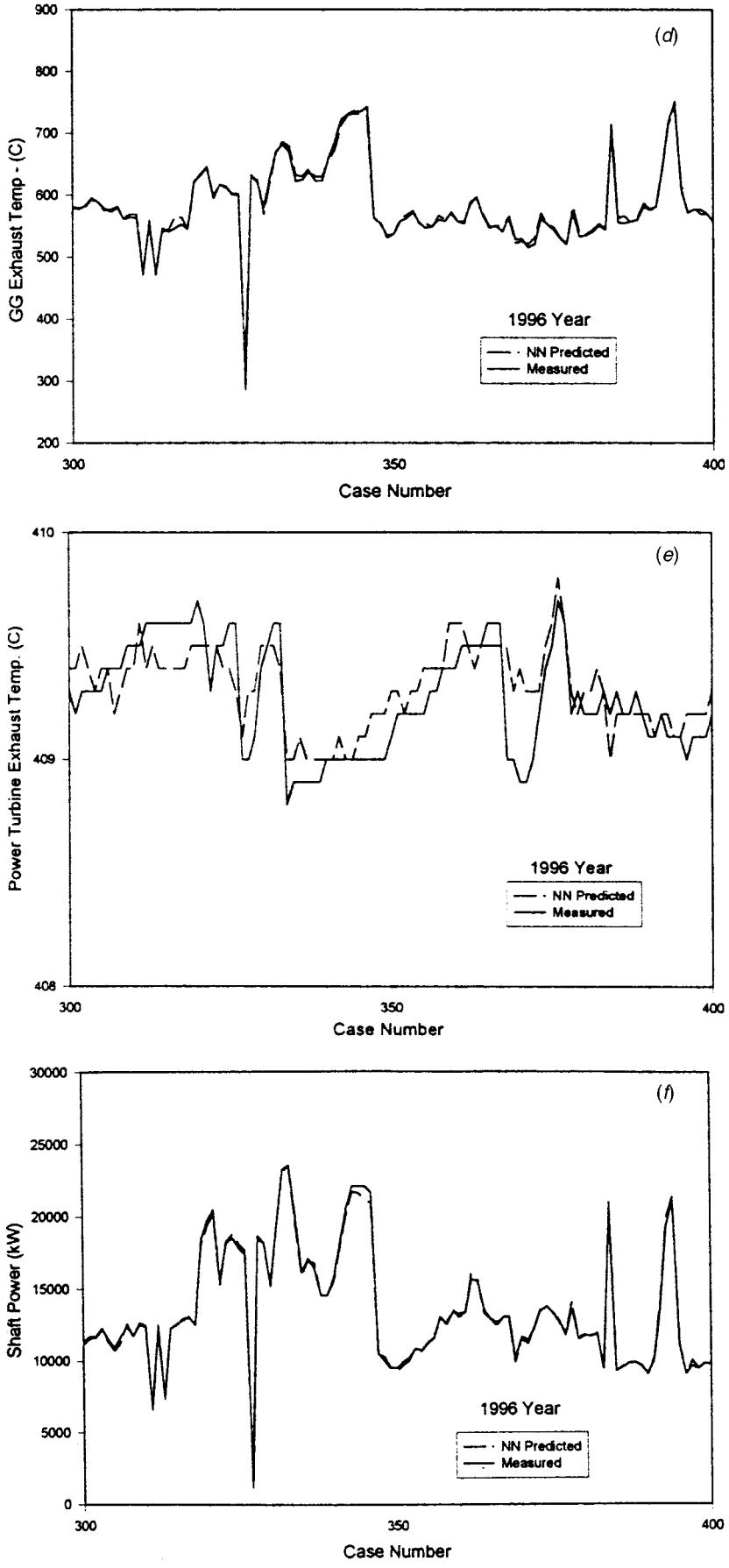
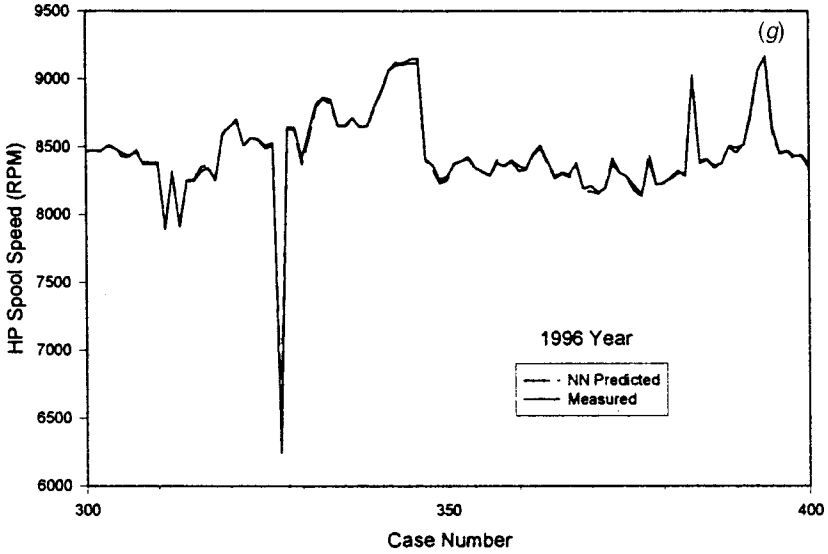



**Fig. 3 (continued)**



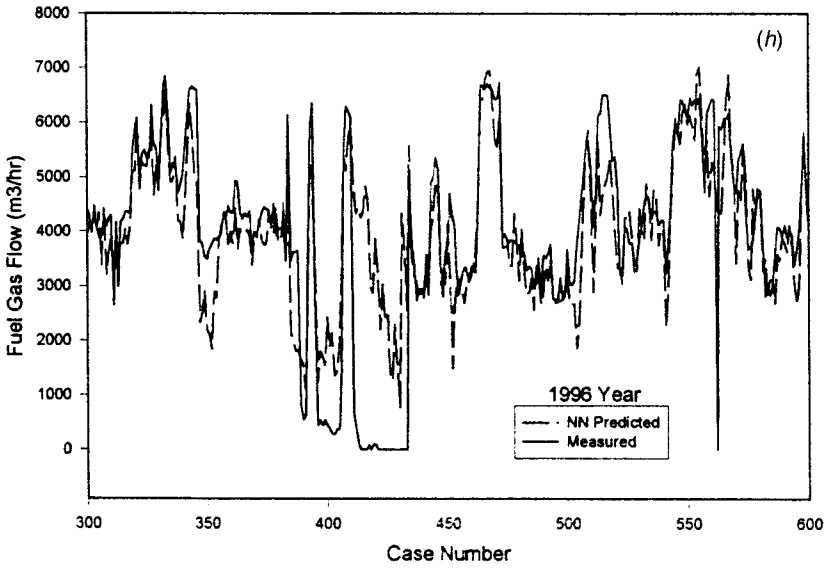




**Fig. 3 (continued)**

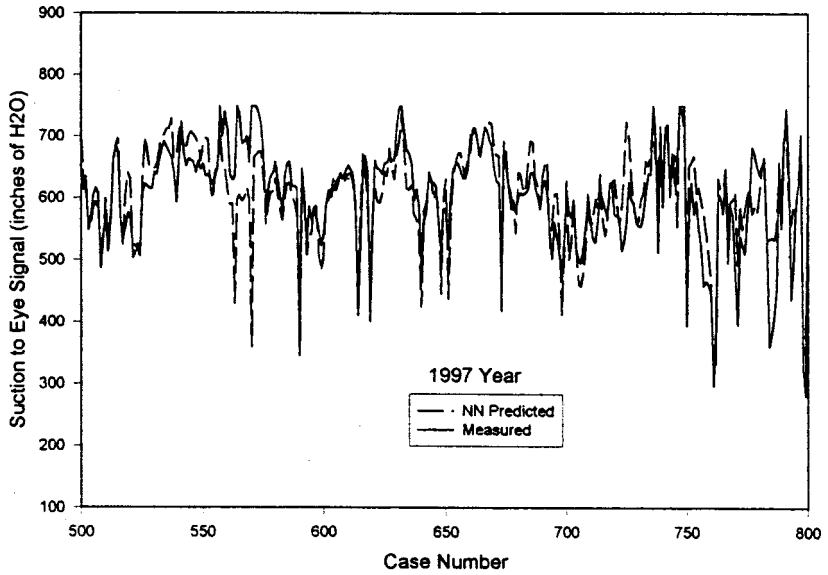



**Fig. 4 Prediction of Qse for a period in 1997 versus measurements**

## 4 Backup of Critical Parameters

The above NN, trained with 1996 data set, can now be used to predict the value of any of the output eight parameters in later years using only the fundamental seven parameters described above. For example, since Qse is very critical in the surge control procedure for the process compressor, it would be advantageous to have a backup procedure to determine this parameter in case of instrument failure or malfunction. The optimized NN can be used to provide this backup value. Figure 4 shows the prediction of Qse for a period in 1997 using the trained 1996 RBF network. Comparison with measured Qse showed good agreement between the two values for the entire year. This demonstrates one useful application of having such a trained NN. Obviously, the main assumption here is that the GT/compressor set is assumed to be at the same condition as the previous year for which the NN has been trained.

## 5 Detection of Sensor Faults

Another use of NN prediction capabilities is that it can provide information in case of a sensor failure or malfunction. In fact this happened over the first half of 1997 when the CDP sensor (P3) failed to operate and hence the corresponding data reported zero during this time (see Fig. 5). When the trained NN was employed with the seven input parameters from 1997 set, this missing parameter was reasonably predicted. Obviously, there were no measured data to compare with up to case #720 where the CDP sensor failed. However, once the sensor was corrected, the remaining cases demonstrate good agreement between predicted and measured values as shown in Fig. 5.

Another example is from 1998 data set where the fuel gas flow sensor failed completely for the entire year as shown in Fig. 6. In this case the 1996 trained NN can be used to provide an estimate of the fuel gas flow to a degree of confidence in the order of one

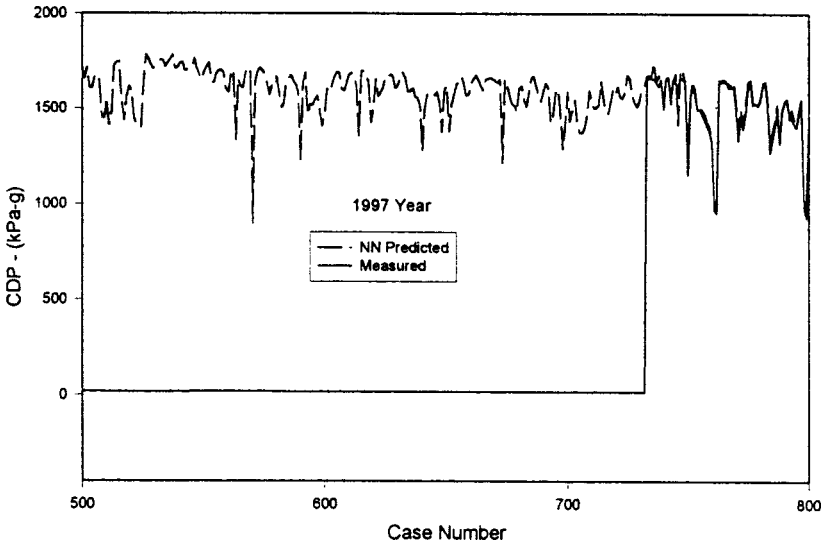



**Fig. 5 Detection of CDP sensor failure up to Case 720**

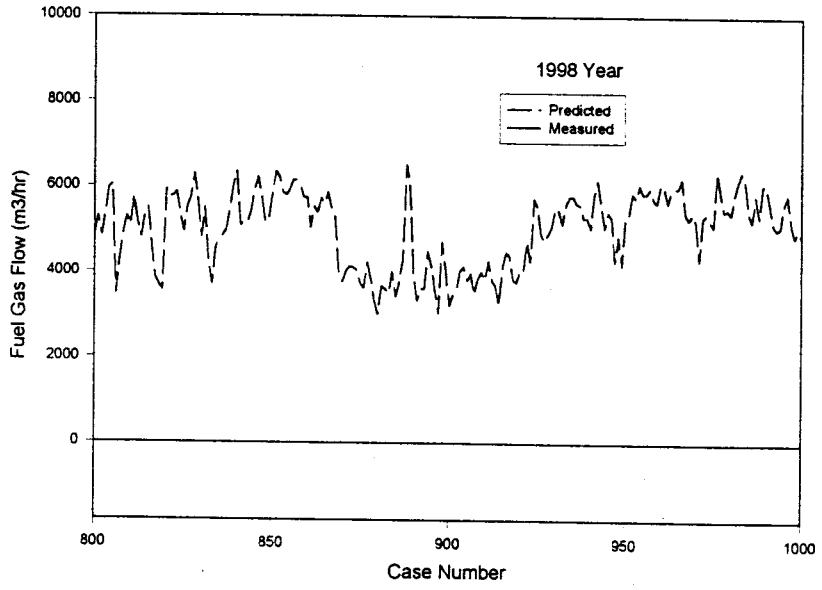



**Fig. 6 Estimate of the fuel gas flow for 1998 when the flow sensor failed in the entire year**

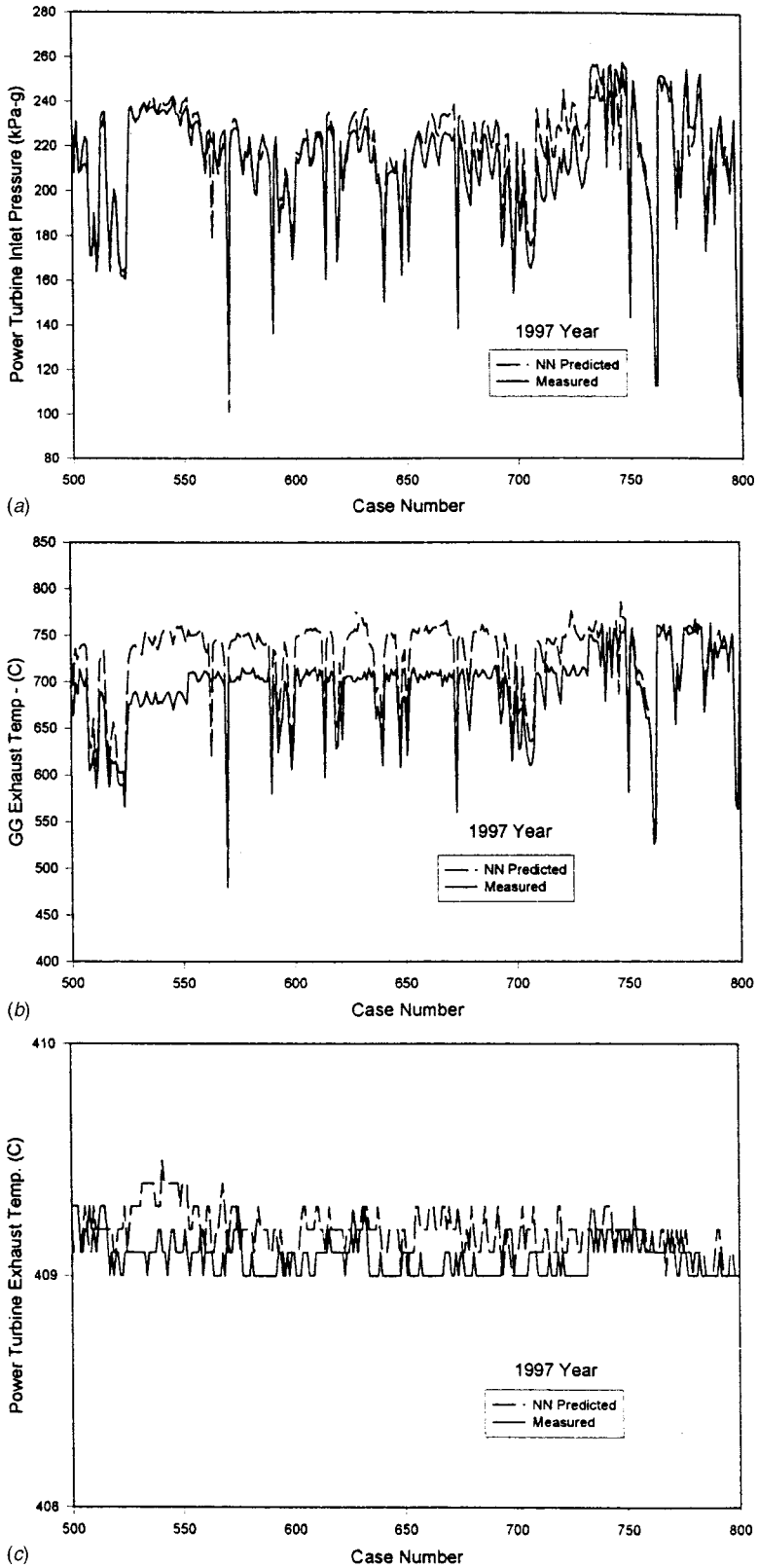
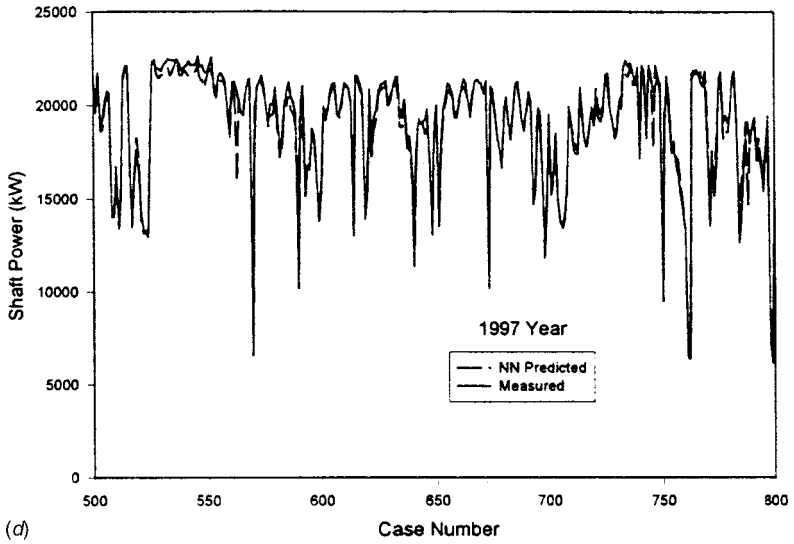
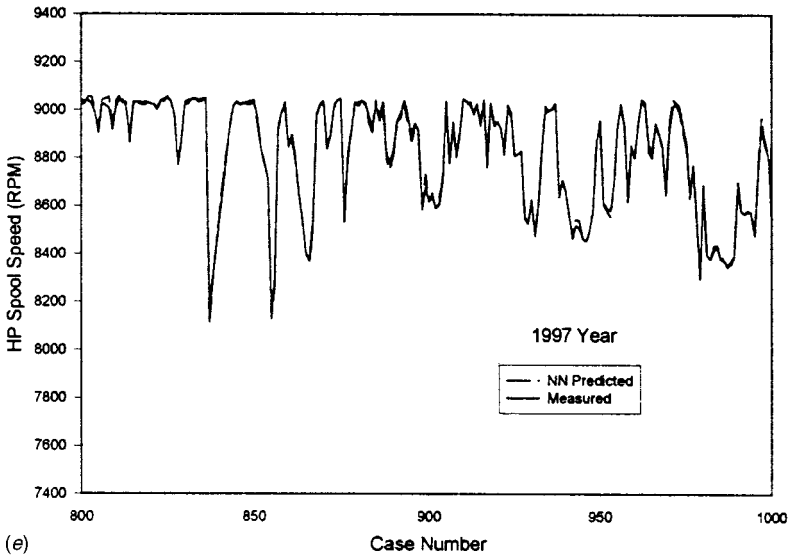


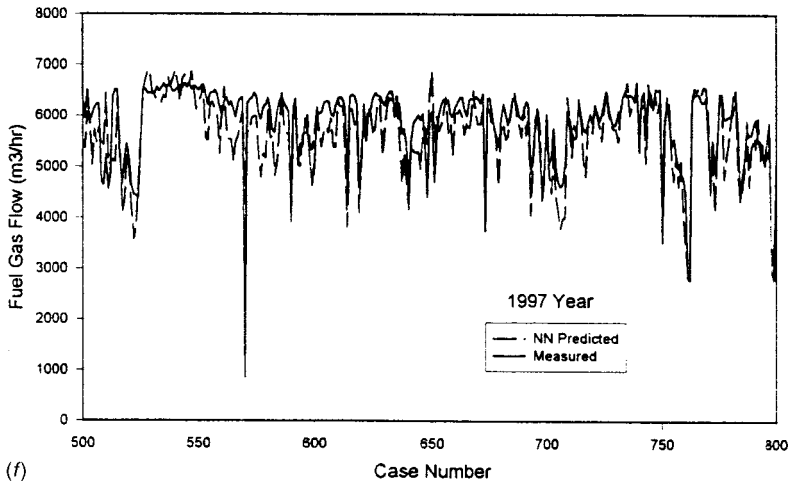

**Fig. 7** (*a*) **Comparison between predicted versus measured power turbine inlet pressure for 1997 year based on 1996 trained RBF NN;** (*b*) **comparison between predicted versus measured GG exhaust temperature for 1997 year based on 1996 trained RBF NN;** (*c*) **comparison between predicted versus measured power turbine exhaust temperature for 1997 year based on 1996 trained RBF NN;** (*d*) **comparison between predicted versus measured shaft power for 1997 year based on 1996 trained RBF NN;** (*e*) **comparison between predicted versus measured HP spool speed for 1997 year based on 1996 trained RBF NN;** (*f*) **comparison between predicted versus measured HP fuel gas flow for 1997 year based on 1996 trained RBF NN**




(*d*)




(*e*)




(*f*)

**Fig. 7 (continued)**

percent mean error and ten percent standard deviation as suggested by Table 3.

## 6 Complete Engine Health Monitoring With Few Variables

Typically, condition monitoring systems rely on measurements of over 130 parameters on the GT/Process Compressor set. A useful application of Neural Networks is to determine these parameters with good degree of confidence with only few measured variables. For example, the same seven input parameters previously mentioned can be used to predict all other 100+ parameters if trained with measured data of these parameters. In order to demonstrate this, the above RBF network has been trained on 7-input/8-output parameters using 1996 data set. This trained NN is then used to predict the same eight-output parameters for 1997 and 1998 years.

Figure 7 (*a* through *f*) show these predictions and comparisons with the corresponding actual measured parameters in 1997. The confidence levels of these prediction lie within the training accuracy specified earlier in Table 3. Obviously, the NN can be used to predict other parameters from the 100+ if trained with those parameters in the first place.

## 7 Estimation of Parameters That Cannot be Measured

A really useful application of NN can be to estimate parameters such as TIT, which cannot be measured due to instrumentation limitation. In this case, what is called a first principle model (FPM) has to be employed in parallel with the NN during the training stage. The FPM will provide the output data (TIT) to the NN as targeted values. Once trained, the NN can be used to predict this data without the use of the FPM.

The concept is illustrated by the procedure shown in Fig. 8(*a*). The same seven-input parameters are used again for the NN which is taken as an RBF network with one hidden layer of 128 neurons and one output parameter (TIT)—see Fig. 9. The FPM utilizes three measured parameters, namely P3, P4, and T4 to calculate TIT. The combustor pressure loss is assumed to be three percent of its inlet pressure, while a typical value for the GG turbine stage efficiency of 88 percent was taken. The 1996 data set was used in training the RBF network of Fig. 9 and the results of the training regression is given in Table 4.

Additionally, the predicted values of TIT using the trained NN is compared to those estimated from the FPM for the same year in Fig. 10. Now, this trained NN can be used to predict TIT for other cases in other years, provided that the GT has maintained the

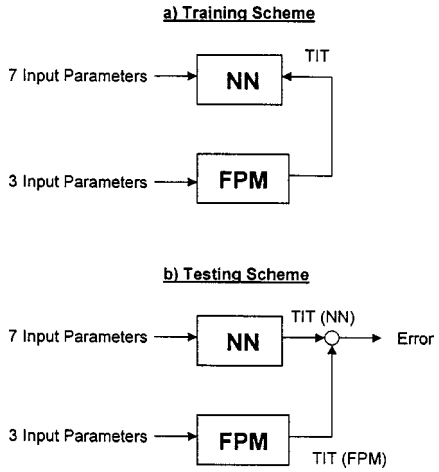



**Fig. 8 Hybrid NN–FPM scheme for estimation of parameters that cannot be measured**

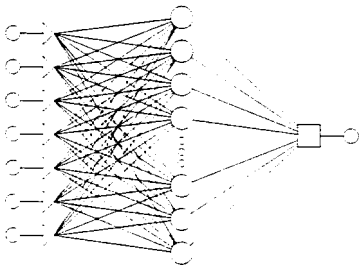


**Fig. 9 RBF architecture used in TIT training/prediction**

**Table 4 RBF Training accuracy for TIT prediction**

| | °C |
|---|---|
| Data mean | 1074 |
| Data S.D. | 94.56 |
| Error mean | −0.5602 |
| Error S.D. | 13.66 |

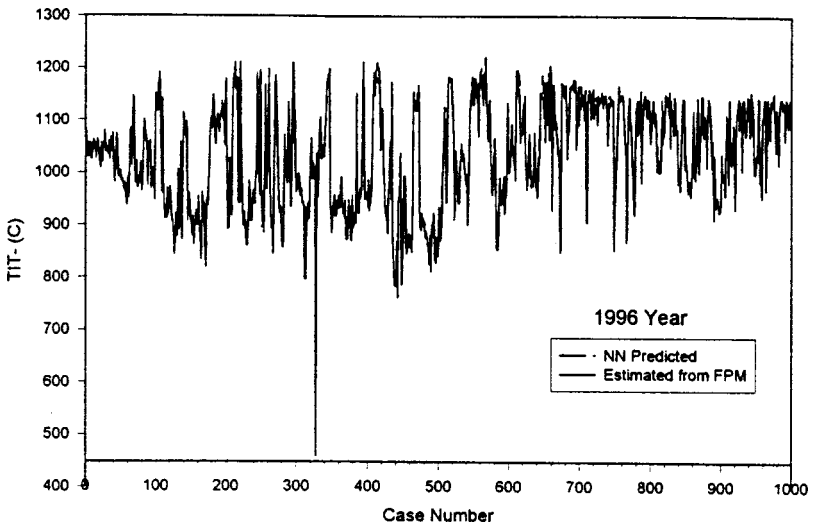



**Fig. 10 NN predicted versus estimated TIT from FPM for the same year**

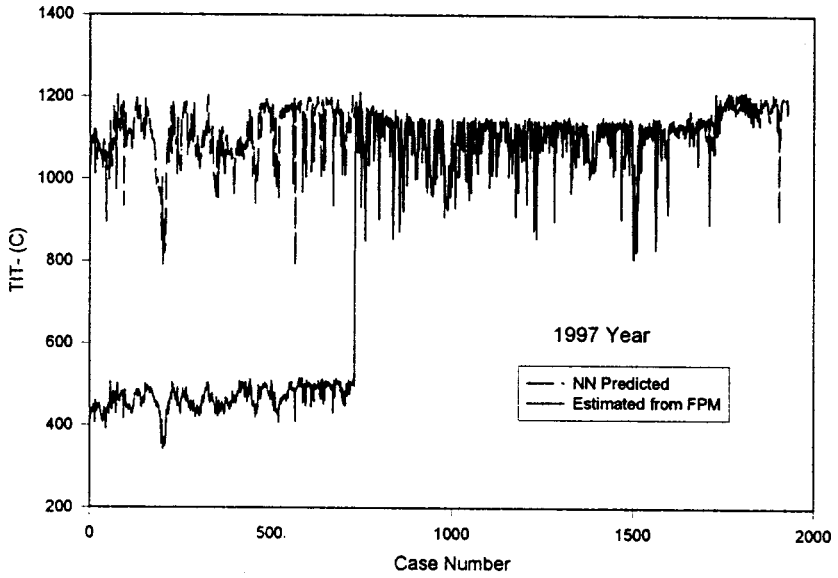



**Fig. 11 NN predicted versus estimated TIT from FPM for a subsequent year**

same health condition. Figure 11 shows a demonstration of this application to predict TIT for 1997 year. The predicted values were also compared to the FPM prediction independent of the NN using measured values of P3, P4, and T4 for each case. The comparison is also shown in Fig. 11. Recall the fact that the CDP (or P3) sensor was faulted during the first 720 cases in 1997, hence the FPM prediction failed to give reliable data to compare with (refer to Fig. 5). However, for the rest of the year, comparison between NN prediction of TIT and those independently estimated from an FPM show excellent agreement as shown in Fig. 11.

## 8 Kohonen SOM for Data Mining

Quite recently, several new computational techniques that make use of NN's have been proposed for nonparametric regression problems, often called ''unsupervised.'' Among them is the self-organizing topological mapping (SOM) by Kohonen [13]. All of the above-mentioned NN applications are called ''supervised'' NN where the training data contains cases featuring input variables together with the associated outputs. In ''unsupervised'' NN's, the training contains only input variables and the SOM algorithm topologically maps these high-dimensional sample space onto a lower dimensional (usually two-dimensional) that is represented by self-organized neuron as a discrete lattice of units. The procedure consists of two steps. The first step is to find the best matching unit (winning unit) to the current input case vector based on the Euclidean distance in the training sample vector. The second step is to modify the weight vectors of both the winning unit and its neighboring units, so as to reduce the distance between these units and the input vector.

The whole data set consisting of the 15 variables of Table 1 (previously seven-input and eight-output) were taken as the input vector for a Kohonen network of $3\times15$ output layer (Fig. 12). The choice of this two-dimensional layer was arbitrary, although the main rational was to allow for three categories (winning units) to capture possible variations in 15 input variables. The 45 output neurons are of RBF type with a square root activation function.

The above Kohonen network was trained on 1996 data set, and after convergence, each case in the set corresponded to a winning unit in the output layer defined as the unit with least activation level. The following is called ''the winning frequency'' of the $3\times15$ lattice of the output layer. (See Table 5.)

Figure 13 plots the winning unit activation level (A) for each case in this 1996 data set. This set is considered to be a good set with no apparent faults in any of the sensors as was the case with some cases in 1997 and 1998 data sets. It appears that the majority of the cases have activation levels between 0.3 and 0.7. The lower the A, the higher the destination of the winning unit. There are three instances, however, where the Kohonen map seemed to have failed to identify some cases from the data set. This is signified by the relatively higher levels of A. These are

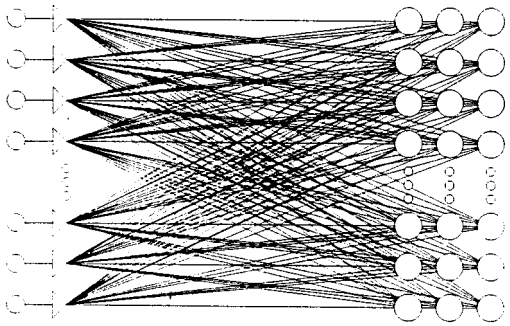


**Fig. 12 Kohenen SOM of 15 input units and $3\times15$ output units**

**Table 5 Winning frequency of 1996 data set**

| | Unit [J] | | |
|---|---|---|---|
| Unit [I] | 1 | 2 | 3 |
| 1 | 55 | 34 | 104 |
| 2 | 12 | 1 | 2 |
| 3 | 9 | 3 | 21 |
| 4 | 18 | 15 | 95 |
| 5 | 17 | 7 | 540 |
| 6 | 14 | 1 | 22 |
| 7 | 10 | 3 | 4 |
| 8 | 0 | 3 | 5 |
| 9 | 6 | 1 | 8 |
| 10 | 15 | 3 | 6 |
| 11 | 49 | 2 | 1 |
| 12 | 5 | 0 | 2 |
| 13 | 0 | 22 | 163 |
| 14 | 91 | 307 | 29 |
| 15 | 52 | 34 | 53 |

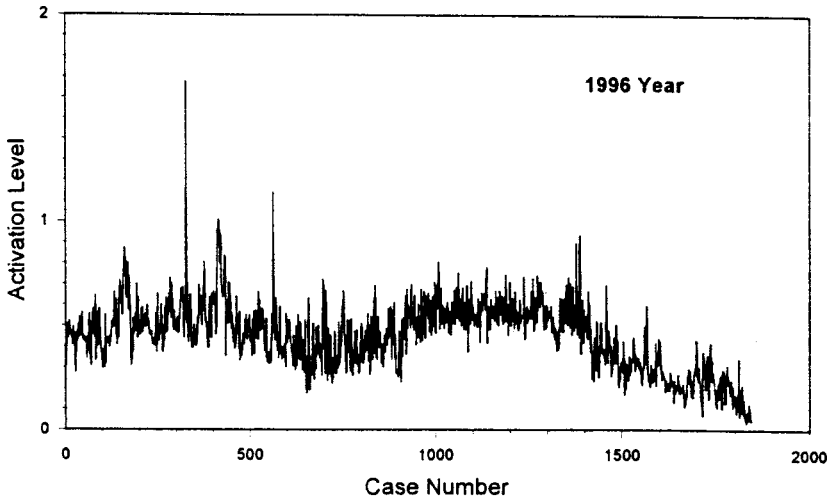



**Fig. 13 Winning unit activation level for each case of 1996 data set**

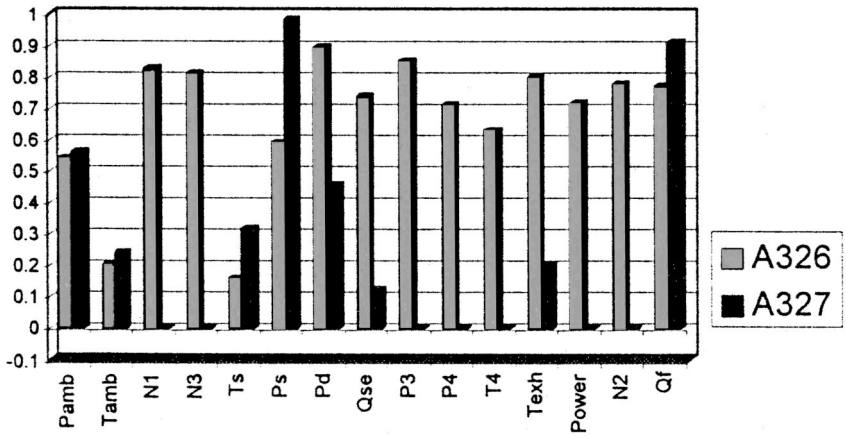



**Fig. 14 Activation levels of the input units for cases 326 and 327 of 1996 data set**

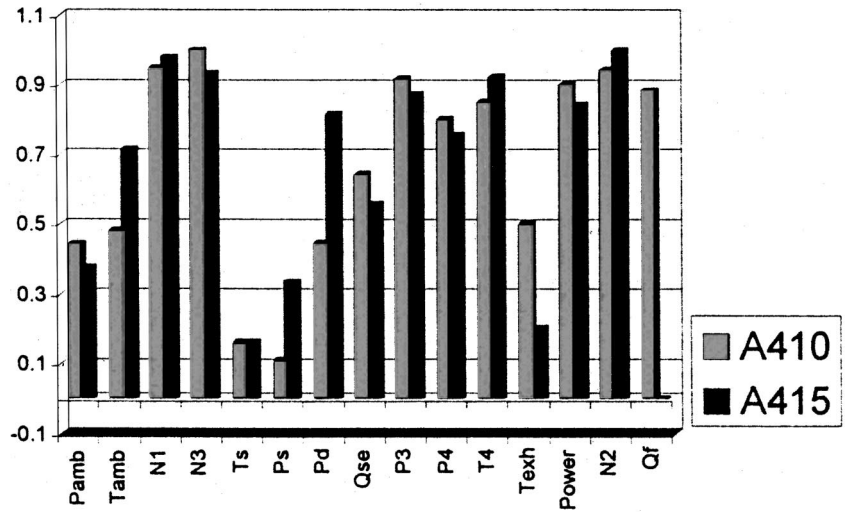



**Fig. 15 Activation levels of the input units for cases 410 and 415 of 1996 data set**

| | Activation Level |
|---|---|
| Case #327 | 1.7 |
| Cases #411–433 | 0.9–1.0 |
| Case #562 | 1.2 |

As was mentioned earlier, case #327 has been identified as one during a shutdown period. If one does not know this fact, a quick and effective technique to establish which parameter(s) effected this failure to recognize the case can be determined from the activation levels of the input layer. Recall that there are 15 input parameters. One can then compare A's for the unrecognized case # 327 to that of a recognized one immediately before (say # 326). Figure 14 shows a bar chart comparison between the 15 A's for

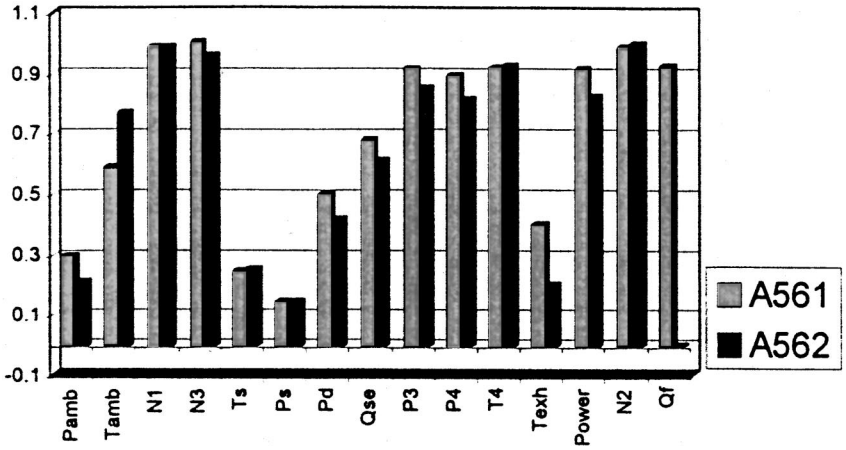



**Fig. 16 Activation levels of the input units for cases 561 and 562 of 1996 data set**

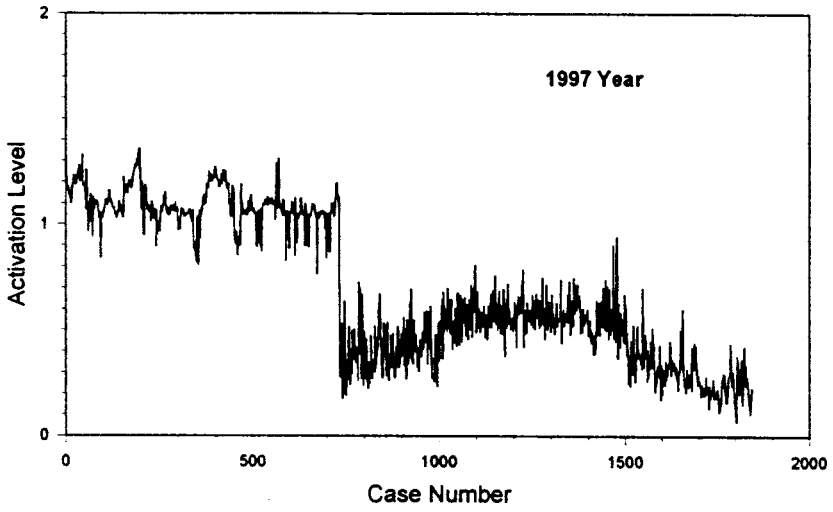



**Fig. 17 Winning unit activation level for each case of 1997 data set**

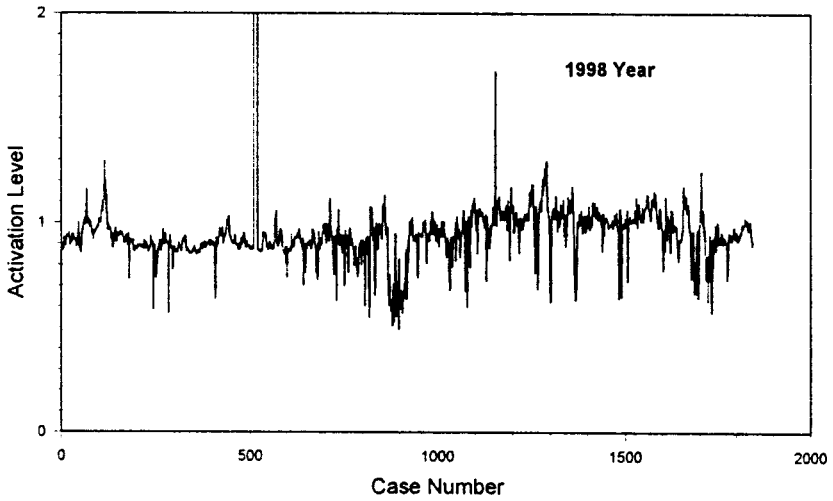



**Fig. 18 Winning unit activation level for each case of 1998 data set**

these two cases. It is apparent that the main cause for the failure of NN to recognize case # 327 is the extremely low activation level of the following parameters: N1, N3, P3, P4, T4, Power, and N2. Examination of the values of these parameters in the condition monitoring system for this case indeed confirmed the above finding.

A similar finding was arrived at when comparing the activation levels of a recognized case # 10 versus an unrecognized case # 15, as well as case # 661 versus case # 562, shown in Fig. 15 and 16, respectively. It is shown that in these cases failure to recognize the mentioned cases can clearly be attributed to the faulty fuel gas flow signal.

When the above trained Kohonen network was applied to 1997 data set, different activation levels of the winning units are determined and plotted in Fig. 17. The first 720 cases showed higher activation levels than the rest of the cases. As mentioned earlier, during these cases, the CDP (P3) signal failed to operate. This demonstrates the validity and usefulness of Kohonen SOM technique in recognizing the fact that there is something wrong (or different) in the data set during this period without even looking at the set itself. Figure 18 gives the corresponding activation levels for the 1998 data set showing higher levels of activation across the entire set. This was recognized earlier as a result of failure of the fuel gas flow sensor all across the set.

## 9 Conclusions

It was demonstrated that neural network is a very useful tool when applied to a gas turbine driven compressor station. RBF networks were found to be the most effective networks which provide the maximum degree of confidence. They are also relatively faster to train than MLP networks. The generalized regression NN's are fast to train, but are prone to overlearning and hence can result in higher errors. The optimized RBF network was capable of performing the following functions: (a) backup of critical parameters, (b) detection of sensor faults, (c) complete engine health monitoring with few variables, and (d) estimation of parameters that cannot be measured.

It was demonstrated that Kohonen SOM technique can easily be applied to recognize the correctness and validity of any data set once the network is trained on a good set. This is achieved by examining the activation level of the winning unit on the output layer of the network. Additionally, it is possible to detect any doubtful, faulty or corrupted parameter(s) in the cases which are not recognized by the network by simply examining the activation levels of the input neurons.


## Acknowledgments

The work presented here is part of a research program sponsored by TransCanada Pipelines Ltd., and permission to publish it is gratefully acknowledged. Valuable discussion and review of the paper by Mr. Glenn Price of NRTC is greatly appreciated. The authors also wish to acknowledge the use of StatSoft Neural Networks in the present investigation.


## Nomenclature

GG = gas generator
GT = gas turbine
N1 = low-pressure gas generator spool speed
N2 = high-pressure gas generator spool speed
N3 = power turbine speed
P3 = GG compressor discharge pressure (CDP)
P4 = intermediate pressure between GG and PT
Pamb = ambient pressure
Pd = process gas discharge pressure
Power = PT shaft power
Ps = process gas suction pressure
PT = power turbine
Qf = fuel gas flow rate
Qse = suction to eye flow signal
T4 = intermediate temperature between GG and PT
Tamb = ambient temperature
Texh = PT exhaust gas temperature
TIT = GG turbine inlet temperature (exit from combustor)
Ts = process gas suction temperature
$\eta_{is}$ = GG stage efficiency

## References


[1] Bishop, C. M., 1995, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford, UK.
[2] Jain, A. K., et al., 1996, "Artificial Neural Networks—A Tutorial," Computer, **29**, No. 3, pp. 31–44.
[3] DePold, H. R., and Gass, F. D., 1988, "The Application of Expert Systems and Neural Networks to Gas Turbine Prognostics and Diagnostics," ASME Turbo-Expo, Stockholm, Sweden, June 2–5.
[4] MacIntyre, J., Tait, J., Kendal, S., Smith, P., Harris, T., and Brason, A., 1994, "Neural Networks Applications in Condition Monitoring," *Applications of Artificial Intelligence in Engineering, Proceedings of the 9th International Conference*, Pennsylvania, July 19–21, pp. 37–48.
[5] Kim, D. S., Shin, S. S., and Carison, D. K., 1991, "Machinery Diagnostics for Rotating Machinery Using Backpropagation Neural Network," *Proceedings of the 3rd International Machinery Monitoring and Diagnostics Conference*, Las Vegas, NV, Dec. 9–12, pp. 309–320.
[6] Lombardo, G., 1996, "Adaptive Control of a Gas Turbine Engine for Axial Compressor Faults," *Proceedings of the 1996 International Gas Turbine and Aeroengine Congress & Exhibition*, Burmingham, UK, June 10–13, ASME, New York.
[7] Zhang, S., and Ganesan, R., 1997, "Multivariable Trend Analysis Using Neural Networks for Intelligent Diagnostics of Rotating Machinery," ASME J. Eng. Gas Turbines Power, **119**, pp. 378–384.
[8] Lo, C. F., and Shi, G. Z., 1992, "Wind-Tunnel Compressor Stall Monitoring Using Neural Networks," J. Aircr., **29**, No. 4, pp. 736–738.
[9] Gibson, S., 1994, "Detection of Surge Utilizing Magnetic Bearings," Revolve 1994, Sponsored by ASME, National Petroleum Show.
[10] Chbat, N. W., Rajamani, R., and Ashley, T. A., 1996, "Estimating Gas Turbine Internal Cycle Parameters Using a Neural Network," *Proceedings of the 1996 International Gas Turbine and Aeroengine Congress & Exhibition*, Burmingham, UK, Jun 10–13, ASME, New York.
[11] Botros, K. K., and Glover, A., 1999, "Neural Networks & Fuzzy Logic—Overview of Pipeline Applications (Part I and II)," 18th International Conference on Offshore Mechanics and Arctic Engineering, St. John's, Newfoundland, July 11–16.
[12] Hunt, K. J., Sbarbaro, D., Zbikowski, R., and Gawthrop, P. J., 1992, "Neural Networks for Control Systems—A Survey," Automatica, **28**, No. 6, pp. 1083–1112.
[13] Kohonen, T., 1989, "The Self-Organizing Map," Proc. IEEE, **78**, No. 9, pp. 1464–1480.

# Twin Web Disk: A Step Beyond Convention

**R. R. Cairo**[1]
Pratt and Whitney

**K. A. Sargent**
Air Force Research Laboratory

*This paper will discuss a study of an innovative design for an advanced turbine rotor that could have a great impact on future engines. The design challenge is to provide a minimum weight turbine rotor system that can withstand beyond state-of-the-art levels of* $AN^2$ *(turbine annulus area multiplied by speed squared). An* $AN^2$ *limit has been reached for high-pressure turbine (HPT) disks configured in conventional (single web) geometry with state-of-the-art nickel alloys. The problem has reached the point where increased* $AN^2$ *has been declared a "break-through" technology. The twin-web disk has the potential to provide this break through. This paper will present the history of this turbine rotor design, analytical results, material/component processing, and concept validation results. All work was performed under an Air Force sponsored program entitled "Composite Ring Reinforced Turbine" (CRRT).* [DOI: 10.1115/1.1445440]

## Introduction

The technology developed and described in this paper was funded by the Air Force under a program entitled ''Composite Ring Reinforced Turbine (CRRT) ([1]).

The CRRT program was structured to expand upon a concept that was conceived during the last phase of the prior Air Force sponsored Advanced Turbine Rotor Design (ATRD) program ([2]). The concept, an internal ceramic matrix composite (CMC) ring reinforced nickel alloy disk capable of satisfying advanced engine demonstrator requirements, showed significant weight and stress reductions relative to a conventional nickel alloy HPT disk during initial trade studies. The CRRT program started with this concept and progressively added more details to the design study and began to include manufacturing considerations as they impacted the disk geometry. The program also remained flexible to the increasing demands of the demonstrator engine. However, the turbine rotor inlet temperature and associated HPT body temperature increased to the point where contemporary CMCs could no longer offer a structural advantage due to the thermal incompatibility between the nickel and CMC. This concept was subsequently replaced by a nonring reinforced twin-web disk (TWD) that eventually became the focus of the program. The TWD continued to offer significant advantages over conventional HPT disks in terms of structural load path efficiency and $AN^2$.

## Design Studies

**Design Criteria/Philosophy.** Design constraints included a 1.94-in. (4.93 cm) minimum bore radius and a 5.2-in. (13.2 cm) maximum bore width. The boundary conditions reflected simultaneous application of maximum T4 (turbine rotor inlet) temperature and $AN^2$ ($600\times10^{+8}$ $in^2$ $rpm^2$ ($3871\times10^{+8}$ $cm^2$ $rpm^2$)). The temperatures at other disk locations were interpolated using a simple conduction analysis. All structural analyses were performed using ANSYS 5.0.

Allowable stresses were based on surface flaw crack growth. Lives for metallic components were computed for a range of stresses using a crack growth analysis program. Based on an advanced engine duty cycle, the disks were required to meet 1000 total accumulated cycles (TACs) of crack growth life in the inspectable areas to support a 500 TAC inspection interval. Internal areas (twin-web cavity), which could not be inspected after disk assembly, were required to have 2000 TAC cycles crack growth life.

**Initial Design Studies with CMC Ring.** The initial configurations of disk designs used geometric parameters that were output from a designed experiment. This study suggested that some of the critical variables were disk bore width, composite ring width, disk cavity width, disk bore height, and thickness at ring inner diameter, disk bore height at appendages for flange and bearing attachments, and appendage structural flexibility. The unique loading situation of transferring radial load into the composite ring results in an unusual state of stress in the rotor system. The studies all revolved around transferring load from the nickel disk to the CMC reinforcing ring(s) by radial compression, loading the ring inner diameter (ID). Load transfer by this mechanism avoids the weak interlaminar tension and shear load paths of the CMC. The rings were designed with fiber architectures that specifically avoid axial splitting and local ID crushing. The rings were developed and demonstrated under the ATRD program. As the design evolved, added realism was built into the finite element model. This was to ensure that the analytical results most appropriately reflected the rotor response. The stresses for the initial and final CRRT configurations are presented in Table 1 for comparison.

**Final Design Studies Without Ring.** As operating temperatures for the engine increased due to cycle changes, the viability of the CRRT HPT came under question. The thermal property mismatch between the CMC and nickel alloy was pushed beyond the threshold for structural efficiency. The additional temperature caused the higher expansion rate nickel bore to grow too much relative to the CMC ring resulting in excessive axial bending. As a result, a design comparison was completed between the CRRT

**Table 1 Stresses from design trade studies for CRRT**

| Stress Location | Initial Design | Final Design |
|---|---|---|
| Upper-Rim Stress in Web, Radial Stress | 202 ksi (1392 MPa) | 139 ksi (958 MPa) |
| Bore Forward Edge Hoop Stress | 220 ksi (1517 MPa) | 158 ksi (1089 MPa) |
| Radial Bearing Stress, Disk Load Ring Interface | -165 ksi (-1138 MPa) | -121 ksi (-834 MPa) |
| CMC Ring Stresses | -33 ksi (-228 MPa) Radial 72 ksi (496 MPa) Hoop | -31 ksi (-214 MPa) Radial 79 ksi (545 MPa) Hoop |

[1]Currently at GE Power Systems, General Electric Company, 300 Garlington Road, GTTC-M/S 174D, Greenville, SC 29602-0648. e-mail: ron.cairo@ps.ge.com.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Stockholm, Sweden, June 2–5, 1998; ASME Paper 98-GT-505. Manuscript received by the IGTI Division October 1997; final revision received by the ASME Headquarters March 1998. Associate Editor: R. Kielb.

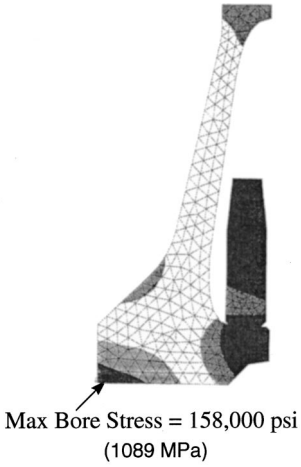



**Fig. 1 Hoop stress, CRRT**

and another design concept that removed the CMC ring, the TWD. The design requirements discussed in the design criteria section were used for this study.

Both disk configurations were iterated upon until strength and durability requirements were met. Comparing hoop stress plots, as seen in Figs. 1 and 2, the bore stress distribution for the TWD was more uniform than the CRRT disk since the CMC ring provided localized restraint due to a combination of reduced inertial and thermal response relative to the nickel disk. This restraint induced a small degree of axial bending, in the nickel bore, about the ring. Also, there was a higher peak radial stress at the bore/web juncture for the CRRT due to the limited space available to develop a smooth geometric transition via a generous ''blend'' radius. Both configurations had desirable axial stress distributions; however, the TWD had preferred distributions. The bore interface stresses for the TWD were entirely in compression at 30 ksi (207 MPa), compared to a gradient of 45 ksi (310 MPa) compression to 8 ksi (55 MPa) tension for the CRRT. Another result of the design change was that the TWD was 7 lbs (3.2 kg) lighter. Figure 3 shows an overlay of the TWD and the CRRT disk half geometries. At the completion of this study it was decided to drop the CRRT and continue with the TWD design.

The design effort now focused on meeting two additional objectives: (1) analyzing the current TWD design to determine stress levels to set bond strength requirements and (2) determining interface requirements for the spin demo disk/spin arbor system, the


**Fig. 2 Hoop stress, TWD**

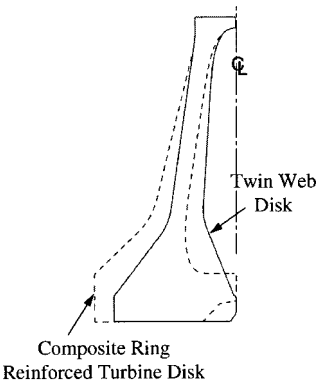



**Fig. 3 Overlay of CRRT and TWD geometries**

details of which will not be discussed in this paper. However, for the first objective, the finite element analysis (FEA) revealed several critical stress locations. One location is the bond joint itself, which is axially stressed by virtue of the twin webs and the corresponding secondary moment, brought about by the eccentricity in load path induced by the blade load line of action and the centroid of each web. Axial tensile stresses of 65 ksi (448 MPa) are predicted. Another critical stress location is the web, which is currently designed to a maximum allowable stress of 135 ksi (931 MPa) to minimize weight. This stress can be reduced, as required, if producibility becomes a concern. The third critical location is the bore since it is required to carry all disk loading beyond the ''self-sustaining'' radius (the radius beyond which a rotating mass becomes parasitic rather than load carrying). This area is currently designed to conventional design allowable stresses for weight optimization. Again, stresses can be reduced, if required, by the addition of bore volume at the expense of weight savings.

## Processing

Several different processing methods were considered for producing this HPT disk. As the TWD emerged as the HPT concept of choice, new bond process requirements developed. There was a need for a bond process that minimized material deformation, to minimize ''fish mouthing'' at bond boundaries and maintain an interior bond zone that had mirror image symmetry about the bond plane for spin dynamics/balance control. The remainder of the program focused on forge joining and the development of a derivative process known as activated forge joining (AFJ). Activated forge joining utilizes features of both transient liquid phase (TLP) and forge joining to produce high quality metallurgical bonds while imparting low distortion. Similar to TLP, AFJ utilizes a foil interlayer to suppress the melting point of the alloys being joined. Similar to forge joining, it utilizes an interface pressure, though greatly reduced relative to pure forge joining. The key feature of AFJ, when compared to TLP, is its ability to transition to production more effectively due to shorter heat cycle times, retention of parent mechanical properties across the bond, and the ability to join hardware in final or near final machined condition due to minimal material upset.

Bond development activities were replanned to include AFJ using a combination of coupons, subscale rings and full-scale rings. A design of experiments approach (DOX) was included in this plan utilizing the AFJ parameters, to determine the optimum process parameters such as load, heating rate, bond cycle length, and post-bond heat treatment. Mechanical specimens were extracted from each bonded unit to assess tensile, creep-rupture, and impact response. The results of the creep rupture tests exceeded the specification requirement for that of the parent material, which is 75 hrs. Tensile test results also compared favorably with the parent material baseline properties.

## Full-Scale Bonding Trials

Up to this point, the viability of AFJ had been demonstrated with coupons, rings, scaled disks, and the designed experiment. Realizing that some of the bonding problems may be unique to the subscale rings that lack the heat path of a disk web and the heat sink of a disk bore; the project expanded the study to include a combination of sub and full-scale simulated disk halves. Applying the processing parameters defined in the designed experiment to full-scale geometry and controlling part positioning in addition to mechanical properties, proved most challenging. Bond trials were run on six disks that were designated FSD-02 through FSD-08 (FSD-04 was not bonded) with bond integrity varying from somewhat porous to excellent.

The results for the first full-scale simulated disk bonding trial (FSD-02) revealed that the interface pressure was too low. Both micrographs and an ultrasonic evaluation showed bonding occurred at the outer and inner edges of the rim, with porosity through the center regions. Further investigation showed error in the equipment setup, which caused the applied load to be approximately one third the desired level. This was corrected for all subsequent trials. A second bonding trial was run successfully (FSD-03: also a simulated full-scale disk) with the proper full bonding load. The next trial, FSD-05, was the first fully processed, monolithic disk-half bonding trial performed. The bonded disk was subjected to ultrasonic and dimensional inspection and then put through a stress relief and age heat treatment. There was very little dimensional change in the bonded disk before and after stress relief. A photomicrograph showing grain growth across the interface is included as Fig. 4.

FSD-06 was the first full-scale, full-geometry, pedigree alloy disk bonded. Ultrasonic scans (C-scans) indicated that the rim outer region was bonded more effectively than the center region, showing a decrease in density going from the midbond plane region to the ID. Specimens were extracted from various locations per the test plan shown in Table 2. Test results for monotonic tension and stress rupture for bonded specimens were extremely impressive in terms of property level and consistency. Figure 5 summarizes monotonic tension test results for in situ bond strength and in situ parent alloy strength. As can be seen, tension ultimate and yield strength for both bond and parent alloy are essentially the same. As expected, elongation and reduction in area for the joint are approximately 50 percent of parent response. These levels are, however, more than adequate for the HPT application. Stress rupture results, shown in Fig. 6, are equally impressive. The test conditions were selected for comparison to established data for the parent alloy. The minimum requirement is ten hours at this condition. As can be seen, the in situ bonded stress rupture life is clustered between 26.6 and 32.3 hours.

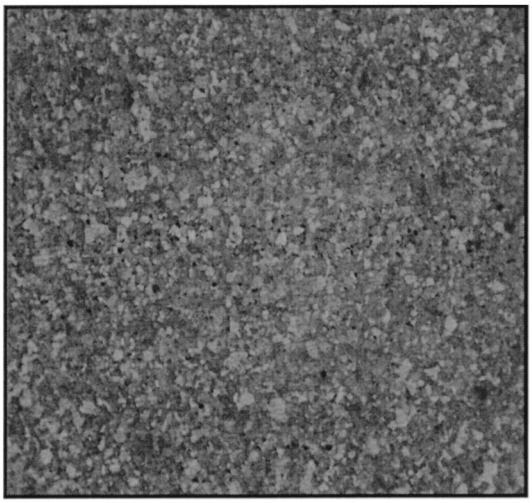


**Fig. 4 Photomicrograph showing grain growth across interface on FSD-05**

**Table 2 FSD-06 test specimen summary**

| Specimen/Test Type | Quantity | Test Conditions |
|---|---|---|
| Tension | 14 | 1300° F (704 °C) |
| Stress Rupture | 4 | 1350° F/106 ksi (732 °C/731 MPa) |
| Low Cycle Fatigue | 20 | 1000°F (538 °C)/ 20 cpm/R=0.05 |
| Crack Growth Rate | 2 | 1200° F (649 °C)/R=0.1 |

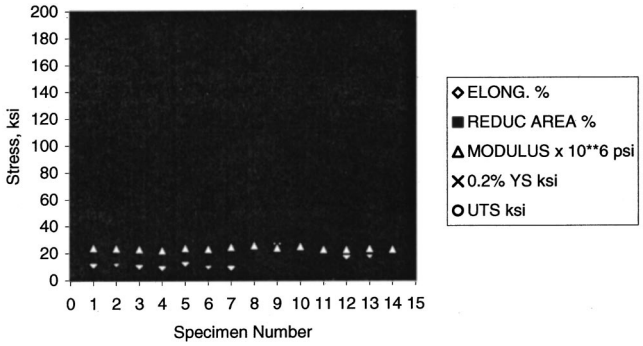



**Fig. 5 Monotonic tension results for FSD-06 specimens. Specimens 1–7: rim-bonded, 8–10: rim parent, 11–14: bore.**

LCF testing was completed on specimens cut from various sections of the disk to include bonded and nonbonded regions in both the rim and the bore. The LCF test data showed that the bond lives are lower than the parent lives, but acceptable for the design. The LCF test data were generated for a maximum stress of 140 ksi (965 MPa) at $R=0.05$ and 1000°F (538°C) for comparison to the available database of parent material. In actuality, the engine HPT design exposes the bond to an average tensile stress of 10 ksi (69 MPa) and a local peak tension stress of 37 ksi (255 MPa). The spin demonstrator has an average axial stress on the bond plane of about 1 ksi (6.8 MPa) in compression and a peak axial tensile stress of 29 ksi (200 MPa). LCF data is shown in Fig. 7 for the rim section only, since this is the main point of interest. In general, it is readily apparent that a significant debit in LCF capability is associated with the bonded specimens. It was inferred that improvements in this capability were likely to come with lower porosity levels. Therefore, reduction of bond porosity was viewed as a key element in future development work and would be pursued by optimizing the chemistry of the bond activation layer.

Crack growth testing was performed on two specimens from the rim section, with the bond aligned with the specimen notch. The first specimen was threshold tested twice at 20 Hz, followed by a 0.167 Hz Region II crack growth test, which is the region of

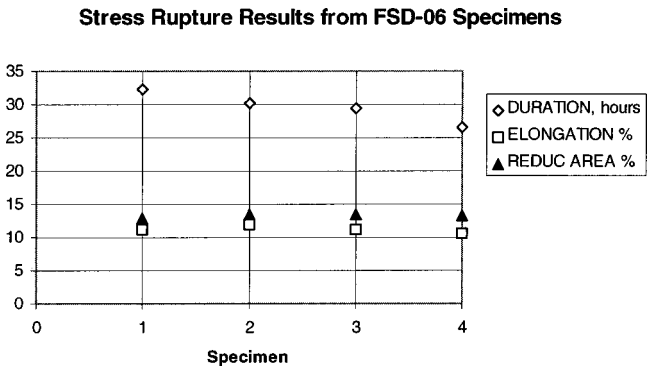



**Fig. 6 Stress rupture results for FSD-06 specimens**

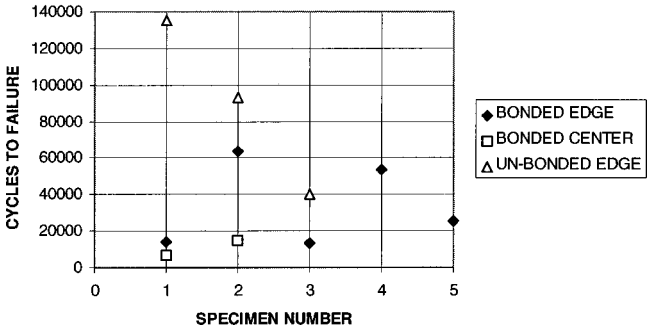



**Fig. 7 Low-cycle fatigue data for FSD-06 rim**

stable, slow, controlled crack growth. The second specimen was tested in Region II only to allow sufficient crack length to obtain data at higher crack growth rates. Test results were compared to both a baseline threshold model and baseline data. The test data compared to the model favorably, with slightly slower crack growth rates in the threshold and a slightly faster growth rate in Region II. Compared to baseline data, the tests show slightly faster growth rates at the higher stress intensities. Examination of the broken specimen halves revealed the crack grew within or at the edge of the bondline, and did not propagate into the parent material. Photomicrographs revealed desirable transgranular rather than intergranular failure morphology.

The bonding trial for FSD-07, which was originally targeted as the spin demonstrator disk, appeared to go smoothly. Due to a malfunction of the pneumatic cylinder, an unusual spread in the rim temperatures during the trial was noted and a section of the rim appeared to be overheated. Inspection of the rim bond-line seemed to substantiate a low or no-load condition during the trial. Loss of load was consistent with the unusual temperature profile seen along the rim of the part since poor surface contact along the rim causes nonuniform heat transfer, which could result in localized hot spots. This part was completely unsuitable for use as the spin demonstrator or for bond property characterization.

The final rotor, FSD-08, is shown in Fig. 8. It was bonded successfully and was designated as the spin demonstrator. Witness material testing revealed properties similar to FSD-06 (reference Figs. 5–7).

It must be emphasized that the characterization data developed during this program were focused to support the design of an experimental spin disk. Much more, and costly, characterization is required to support man-rated applications. This is part of a typical evolutionary approach to evaluate an innovative concept. Once validated, the concept would receive more detailed design and characterization to enable it to transition to engine applications.

## Concept Validation

The validation program utilized a full-scale twin-web disk with an oversize, interrupted rim, to simulate, the effect of blade loading. The spin test program consisted of three phases: (I) a room temperature strain survey to maximum speed; (II) an elevated temperature cyclic spin test that was interrupted for NDE of the bond plane; and (III) an overspeed demonstration taken to burst conditions.

**Room Temperature Strain Survey.** For Phase I, the spin demonstrator was instrumented with 16 strain gages to capture the strain distribution across each bore and at the 0 deg and 180 deg polar positions. The strain survey was run twice to 19,600 RPM taking gage readings accelerating to and decelerating from that RPM. The maximum attained stress was 138 ksi in the bore region closest to the web cavity. The load paths in the disk halves were very symmetric and the polar variation (0 and 180 deg) in stress was very low. The gages recorded strains that were typically within three percent of the average gage reading. Fig. 9 shows a typical plot of strain versus speed squared using the highest reading bore gage as an example.

**Interrupted, Elevated Temperature Cyclic Spin Test.** In Phase II, the test called for a total of 1000 cycles (triangular waveform pattern, zero dwell time) with an $R=0.05$ at a temperature of 1000°F with periodic inspections. The $R=0.05$ corresponds to a maximum rotor speed of 19,600 RPM and a minimum speed of 4400 RPM. Inspection intervals at 250, 500, and 750 cycles were selected. A portable ultrasonic probe and tank (providing a drip feed of water to couple the probe with the disk) were assembled to perform the interval inspections. No discernable changes in bond plane integrity were noted during any of the inspection points including one at the completion of 1000 cycles.

**Over-Speed Demonstration Followed by Burst.** The Phase III test to demonstrate overspeed and burst margin required a dwell time at both maximum and overspeed RPM of five minutes, before decelerating for a residual growth measurement. Since burst is frequently driven by high rate-of-change acceleration events, it was more representative, for the actual burst test, to accelerate at the limit of the drive turbine until burst (essentially zero dwell time).

The maximum design speed demonstration was accomplished without incident. The disk was accelerated to 21,500 RPM ($AN^2=600\times10^8$ in$^2$ rpm$^2$) at a rate of 110 RPM/sec, the limit of

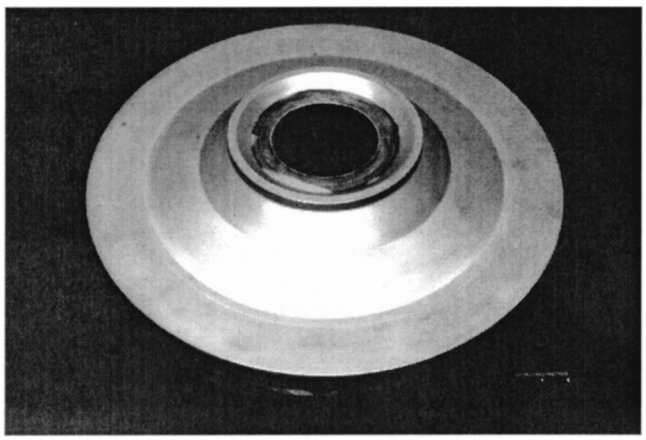


**Fig. 8 FSD-08 post bond assembly**

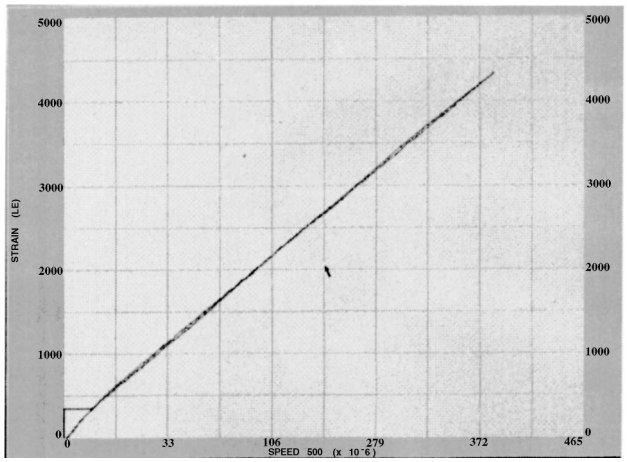



**Fig. 9 Typical strain survey result using highest reading bore gage (adjacent to web cavity) at 0-deg position**

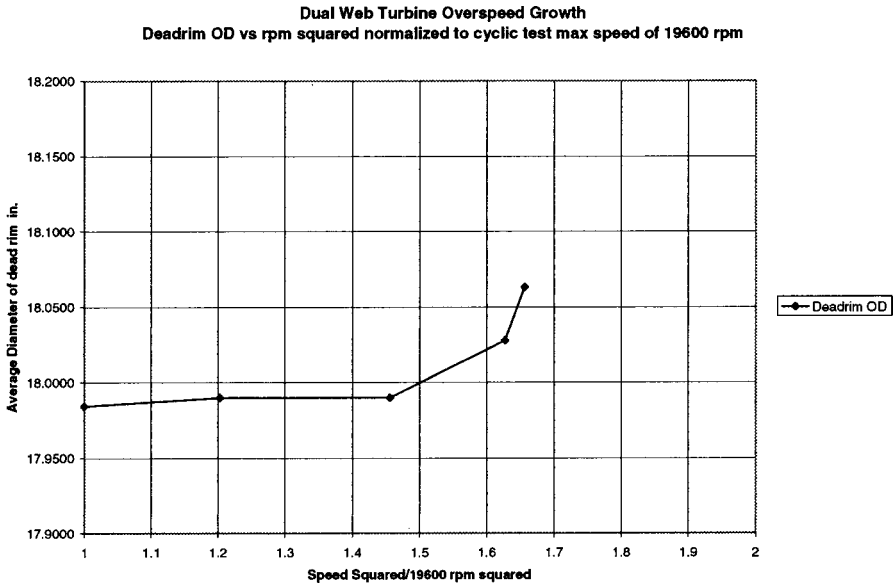



**Fig. 10 Residual growth of dead rim versus normalized RPM squared**

the drive turbine in the spin pit. After a five minute dwell at 21,500 RPM, the disk was brought back to zero RPM. The pit was opened for dimensional inspection of the disk and dishpan outer diameters to document plastic growth. There was no measurable change in geometry.

The next demonstration point, ten percent overspeed (23,650 RPM-$AN^2 = 725 \times 10^8$ in$^2$ rpm$^2$) followed the same procedure. Once again, there was no appreciable growth detected and spin dynamics remained very stable. At this point, it appeared the disk was capable of considerably higher RPM. Hence, a revision to the test plan, to include several additional RPM dwells prior to burst, was adopted for plastic growth documentation.

The next selected point was at 25,000 RPM (16.3 percent overspeed- $AN^2 = 811 \times 10^8$ in$^2$ rpm$^2$), also went without incident. The subsequent residual growth assessment did show the start of measurable plastic deformation. The next dwell point was 26,296 RPM (22.3 percent overspeed). The acceleration was halted at 25,225 RPM (17.3 percent overspeed-$AN^2 = 825 \times 10^8$ in$^2$ rpm$^2$) due to rotor imbalance. Upon shutdown, the spindle connecting the disk to the drive turbine was found to be bent and appreciable plastic deformation (rim growth) had occurred. Figure 10 shows dead rim (OD) growth versus speed squared (normalized to 19,600 RPM-the maximum cyclic speed). The trend indicates a robust disk concept that is capable of functional failure (excessive growth) rather than catastrophic structural failure (burst). Post-spin evaluation revealed two things worthy of note: (1) radial cracks at mid web position; and (2) bi-directional cracks in the web beneath the live rim. As the disk speed exceeded the point of inducing elastic hoop stress in the bore, a zone of plastic stress began to develop. This zone, and its associated front, grew radially with subsequent, increased, speed excursions to the point where the midweb location was carrying a much higher hoop stress than at maximum design speed. This high hoop stress initiated the radially growing cracks. Concurrent with this, was the increased external radial load imparted by the segmented rim (simulated blades) as speed increased. This induced increasingly higher radial stress in the live rim and upper regions of the web, resulting in hoop oriented cracks. As rotor speed continued to increase, the plastic front from the bore advanced radially further outward until it coincided with the zone of higher radial stress progressing radially inward from the rim. This confluence of high biaxial stresses was manifested as the network of biaxial cracks observed in the post spin evaluation.

## Summary

The twin-web disk represents a radical departure from conventional HPT design practice and was conceived to meet the aggressive $AN^2$ goals of the future. The disk was developed by an integrated technology team that produced a novel metallurgical joining process. The joining development process utilized a ''building block'' approach starting with small 2-in. (5.08 cm) square blocks (in a statistically designed experiment to optimize joining parameters), progressing to scaled rings (disk rim simulation); then scaled disks; and finally, full-scale/full-geometry disks. Mechanical properties were evaluated in all critical locations to ensure application to demonstrator engine hardware and correlation made to ultra sonic evaluations of bond integrity. A full-scale spin demonstrator was fabricated and subjected to structural validation spin testing which was a complete success. It demonstrated not only load path symmetry, bore stress uniformity, cyclic stress capability and over-speed capability, but also the ability to exceed an $AN^2 = 600 \times 10^8$ in$^2$ rpm$^2$ without catastrophic burst—truly a step beyond convention. Future activities will include studies to introduce shot peening to the internal web walls and the rim inner diameter to enhance life, and further validation in a demonstrator engine.

## References

[1] Cairo, R. R., 1999, ''Composite Ring Reinforced Turbine,'' Program Final Report, Contract F33615-92-C-2201, Report No. AFRL-PR-WP-TR-1999-2050.

[2] Cairo, R. R., 1992, ''Advanced Turbine Rotor Design,'' Program Final Report, Contract F33615-85-C-2594, Report No. AFWAL-TR-88-2129.

# An Improved Transfer Matrix Method for Steady-State Analysis of Nonlinear Rotor-Bearing Systems

**J. W. Zu**

**Z. Ji**

Department of Mechanical and Industrial Engineering, University of Toronto, Five King's College Road, Toronto, ON M5S 3G8, Canada

*An improved transfer matrix method is developed to analyze nonlinear rotor-bearing systems. The rotating shaft is described by the Timoshenko beam theory which considers the effect of the rotary inertia and shear deformation. A typical roller bearing model is assumed which has cubic nonlinear spring and linear damping characteristics. Transfer matrices for the Timoshenko shaft element, disk element, and nonlinear bearing element are derived and the global transfer matrix is formed. The steady-state response of synchronous, subharmonic, and superharmonic whirls is determined using the harmonic balance method. Two numerical examples are presented to demonstrate the effectiveness of this approach.* [DOI: 10.1115/1.1447235]

## Introduction

The steady-state response analysis of rotor-bearing systems is one area of interest in rotor dynamics to many researchers. Typical rolling element bearings possess nonlinear spring characteristics due to Coulomb friction and the radial clearance between rollers and rings. Yamamoto et al. [1] showed that the nonlinear force in a bearing can be approximated using a nonlinear function related to the third power of the deflection in single-row deep-groove ball bearings. Ishida et al. [2] obtained a fourth-order power nonlinear bearing model in double-row angular contact ball bearings.

The harmonic balance method, introduced by Nayfeh and Mook [3], has been successfully applied to analyze the forced vibration and stability of nonlinear systems. Yamamoto et al. [4,5] and Ishida et al. [2] obtained superharmonic and subharmonic vibrations of a two-degree-of-freedom rotor mounted on nonlinear bearing using this method. They showed good agreement between analytical and experimental results.

The transfer matrix method is a popular technique used to analyze rotor-bearing systems. However, the method has mainly been applied to linear systems, such as the early work done by Prohl [6] and Lund [7], where an Euler-Bernoulli beam element was used to model the shaft segment. Gu [8] combined the transfer matrix method with the direct integration method. Lee et al. [9] proposed a modified transfer matrix to incorporate the Timoshenko beam element. A Rayleigh-Euler beam with the effects of rotary inertia and gyroscopic moments was introduced by Kang et al. [10] for asymmetric linear rotor-bearing systems. Not until recently was the transfer matrix method extended to nonlinear systems by Lee et al. [11], who performed steady-state vibration analyses of a flexible rotor supported by bearings with power nonlinearity. The global equation for the transfer matrix was not derived, however, and a reformulation was necessary for different systems. In addition, a large number of unknowns, i.e., 24 unknowns state vectors must be solved from 24 sets of equations using iterative procedures.

In this paper, an improved transfer matrix is proposed for rotor-bearing systems with nonlinear bearings. The shaft is modeled based on the Timoshenko beam theory. A typical rolling element bearing model is assumed, which has cubic nonlinear spring and linear damping characteristics given by

$$F_x = K_1 u_x + K_3 u_x^3 + C\ddot{u}_x$$
$$F_y = K_1 u_y + K_3 u_y^3 + C\dot{u}_y \tag{1}$$

where $K_1$ and $K_3$ are linear and nonlinear stiffness coefficients and $C$ is the linear damping coefficient of the bearings. A global equation for the transfer matrix is derived and steady-state response of the system is approximated using the harmonic balance method with initial values predicted using a numerical scheme. Numerical simulations are carried out to demonstrate the effectiveness of the method and to study the influence of the nonlinear bearings on the dynamic response characteristics of the system.

## Transfer Matrix for a Timoshenko Shaft Segment

Consider the $i$th shaft segment as shown in Fig. 1. The left end coordinate is $z_{i-1}$ and the right end coordinate is $z_i$, thus the total length of the shaft segment is $l = z_i - z_{i-1}$. The transverse displacements of the shaft are denoted by $u_x$ and $u_y$ along the $x$ and $y$-directions and the bending angles in two perpendicular planes of motion $Oxz$ and $Oyz$ are represented by $\psi_x$ and $\psi_y$, respectively. Introducing the nondimensional space variable $\zeta = (z - z_{i-1})/l$ and the complex notations, $u = u_x + iu_y$, $\psi = \psi_x + i\psi_y$; the governing equations of the rotating shaft are (Zu and Han [12])

$$\frac{\partial^2 u}{\partial t^2} + \frac{\kappa G}{\rho l^2}\left[l\frac{\partial \psi}{\partial \varsigma} - \frac{\partial^2 u}{\partial \varsigma^2}\right] = 0 \tag{2}$$

$$\frac{\partial^2 \psi}{\partial t^2} - i\frac{\Omega J_z}{\rho I l}\frac{\partial \varphi}{\partial t} - \frac{E}{\rho l^2}\frac{\partial^2 \psi}{\partial \varsigma^2} + \frac{\kappa AG}{\rho I l}\left(l\psi - \frac{\partial u}{\partial \varsigma}\right) = 0 \tag{3}$$

where $\rho$ is the mass density, $A$ the cross-sectional area, $I$ the transverse moment of inertia, $J_z$ polar mass moment of inertia and $\Omega$ is the shaft angular velocity. $E$, $G$, and $\kappa$ are Young's modulus, shear modulus and shape factor, respectively.

Due to the existence of nonlinear bearing forces, the steady-state response may consist of not only synchronous vibration at the rotating speed, but also subharmonic and superharmonic vibrations. There are many subharmonic and superharmonic components in the system. For cubic nonlinear bearing forces, the dominant subharmonic vibration is a vibration at the frequency of

Contributed by the Structure and Dynamics Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the S&D Division February 1998; final revision received by the ASME Headquarters October 2001. Technical Editor: M. Mignolet.

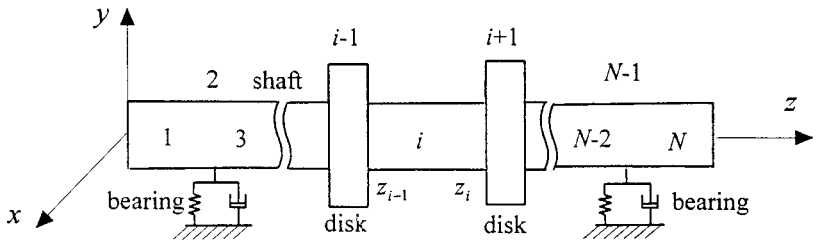



**Fig. 1 A shaft-bearing system with divided elements**

$\Omega/3$ and the dominant superharmonic vibration is a vibration at the frequency of $3\Omega$. Therefore, a solution to Eqs. (2) and (3) for steady-state response may be assumed as

$$u(\varsigma,t)=\sum_p U_p(\varsigma)e^{ip\Omega t}$$
$$\psi(\varsigma,t)=\sum_p \Psi_p(\varsigma)e^{ip\Omega t} \qquad p=-\frac{1}{3},\frac{1}{3},-1,1,-3,3. \tag{4}$$

Substituting Eq. (4) into Eqs. (2) and (3) and decoupling the subsequent set of ordinary differential equations,the following equations are obtained:

$$l_2U_p^{(4)}+l_1U_p''+l_0U_p=0$$
$$l_2\Psi_p^{(4)}+l_1\Psi_p'+l_0\Psi_p=0. \tag{5}$$

The coefficients in Eq. (5) are given by

$$l_0=\frac{\rho I}{\kappa AG}\omega^4-\frac{\Omega J_z}{\kappa AG}\omega^3-\omega^2$$
$$l_1=\frac{I}{l^2A}\left(1+\frac{E}{\kappa G}\right)\omega^2-\frac{\Omega J_z}{l^2\rho A}\omega \tag{6}$$
$$l_2=\frac{EI}{\rho Al^4}$$

in which $\omega=p\Omega$.

There are two cases for the solutions to Eq. (5), $\sqrt{l_1^2-4l_2l_0}>l_1$ and $\sqrt{l_1^2-4l_2l_0}<l_1$. In general, $\sqrt{l_1^2-4l_2l_0}>l_1$, where the solutions are expressed as

$$U_p(\varsigma)=A_{p1}\cosh(s_1\varsigma)+A_{p2}\sinh(s_1\varsigma)+A_{p3}\cos(s,\varsigma)+A_{p4}\sin(s_2\varsigma)$$
$$\Psi_p(\varsigma)=A'_{p1}\sinh(s_1\varsigma)+A'_{p2}\cosh(s_1\varsigma)+A'_{p3}\sin(s_2\varsigma)+A'_{p4}\cos(s_2\varsigma) \tag{7}$$

in which

$$s_1=\sqrt{\frac{-l_1+\sqrt{l_1^2-4l_2l_0}}{2l_2}}, \quad s_2=\sqrt{\frac{l_1+\sqrt{l_1^2-4l_2l_0}}{2l_2}}.$$

Coefficients $A_{p1}-A_{p4}$ and $A'_{p1}-A'_{p4}$ in Eq. (7) are arbitrary constants and only four of them are independent. Their relationship can be derived from Eqs. (2) and (4) as

$$A'_{p1}=c_{p1}A_{p1}, \quad A'_{p3}=c_{p2}A_{p3}$$
$$A'_{p2}=c_{p1}A_{p2}, \quad A'_{p4}=-c_{p2}A_{p4} \tag{8}$$

where

$$c_{p1}=\frac{1}{s_1}\left(\frac{\rho l}{\kappa G}p^2\Omega^2+\frac{1}{l}s_1^2\right), \quad c_{p2}=\frac{1}{s_1}\left(\frac{\rho l}{\kappa G}p^2\Omega^2-\frac{1}{l}s_2^2\right).$$

Having obtained displacements of the shaft segment in Eq. (4) and (7), the stress resultants, which are the bending moments $M_x$, $M_y$ and the shear forces $Q_x$, $Q_y$, can be written in terms of the displacements in complex form as

$$M(\varsigma,t)=M_x(\varsigma,t)+iM_y(\varsigma,t)=EI\psi'(\varsigma,t)/l$$
$$Q(\varsigma,t)=Q_s(\varsigma,t)+iQ_y(\varsigma,t)=\kappa AG\left(\frac{1}{l}u'(\varsigma,t)-\psi(\varsigma,t)\right). \tag{9}$$

Using Eq. (7), Eq. (9) can be expressed by

$$M(\varsigma,t)=\sum_p M_p(\varsigma)e^{ip\Omega t}$$
$$Q(\varsigma,t)=\sum_p Q_p(\varsigma)e^{ip\Omega t} \tag{10}$$

where

$$M_p(\varsigma)=\frac{EI}{l}(A'_{p1}s_1\cosh(s_1\varsigma)+A'_{p2}\sinh(s_1\varsigma)+A'_{p3}s_2\cos(s_2\varsigma)-A'_{p4}s_2\sin(s_2\varsigma))$$
$$Q_p(\varsigma)=\kappa AG\left(\left(\frac{s_1}{l}A_{p1}-A'_{p1}\right)\sinh(s_1\varsigma)+\left(\frac{s_1}{l}A_{p2}-A'_{p2}\right)\cosh(s_1\varsigma)-\left(\frac{s_1}{l}A_{p3}-A'_{p3}\right)\sin(s_2\varsigma)+\left(\frac{s_2}{l}A_{p4}-A'_{p4}\right)\cos(s_2\varsigma)\right). \tag{11}$$

Introducing the following node parameter vector

$$\{S(z,t)\}=\{u(z,t)\ \ \psi(z,t)\ \ M(z,t)\ \ Q(z,t)\}^T,$$

Eqs. (4) and (10) can be compactly written by

$$\{S(z,t)\}=\sum_p \{S_p(z)\}e^{ip\Omega t} \quad p=-\frac{1}{3},\frac{1}{3},-1,1,-3,3 \tag{12}$$

where

$$\{S_p(z)\}=\{U_p(z)\ \ \psi_p(z)\ \ M_p(z)\ \ Q_p(z)\}^T. \tag{13}$$

Substitution of Eqs. (7), (8), and (11) into the vector $\{S_p(z)\}$ in Eq. (13) leads to

$$\{S_p(z)\}=[K(\varsigma)]\{A\} \tag{14}$$

in which

$$\{A\}=\{A_{p1}\ \ A_{p2}\ \ A_{p3}\ \ A_{p4}\}^T$$

is a coefficient vector. At $z=z_{i-1}$, $\varsigma=0$ and at $z=z_i$, $\varsigma=1$. $\{A\}$ can be expressed by $\{S_p(z_{i-1})\}$ by letting $\zeta=0$ in Eq. (14):

$$\{A\}=[K(0)]^{-1}\{S_p(z_{i-1})\}. \tag{15}$$

Substituting Eq. (15) into Eq. (14), the vector $\{S_p(z)\}$ can be written as

$$\{S_p(z)\}=[T_p(z)]\{S_p(z_{i-1})\} \tag{16}$$

where $[T_p(z)]$ is the transfer matrix for the shaft segment. It is given by

$$[T_p(z)]==[K(\zeta)][K(0)]^{-1}. \tag{17}$$

Expanding Eq. (17) yields

$$[T_p(z)]=\begin{bmatrix} \frac{c_1 s_1\cos(s_2\varsigma)-c_2 s_2\cosh(s_1\varsigma)}{A_1} & \frac{l[A_4\sinh(s_1\varsigma)-A_3\sin(s_2\varsigma)]}{A_2} \\ \frac{c_1c_2[s_1\sin(s_2\varsigma)-s_2\sinh(s_1\varsigma)]}{A_1} & \frac{l[c_1A_4\cosh(s_1\varsigma)-c_2A_3\cos(s_2\varsigma)]}{A_2} \\ \frac{-EIc_1c_2s_1s_2[\cosh(s_1\varsigma)-\cos(s_2\varsigma)]}{A_1 l} & \frac{EIl[c_1s_1A_4\sinh(s_1\varsigma)-c_2s_2A_3\sin(s_2\varsigma)]}{A_2 l} \\ \frac{\kappa AG[c_2s_2A_3\sinh(s_1\varsigma)-c_1s_1A_4\sin(s_2\varsigma)]}{A_1} & \frac{\kappa AGA_3A_4l[-\cosh(s_1\varsigma)+\cos(s_2\varsigma)]}{A_2} \end{bmatrix}$$

$$\begin{bmatrix} \frac{l[\cosh(s_1\varsigma)-\cos(s_2\varsigma)]}{EIA_1} & \frac{l[c_2\sinh(s_1\varsigma)+c_1\sin(s_2\varsigma)]}{\kappa AGA_2} \\ \frac{l[c_1\sinh(s_1\zeta)-c_2\sin(s_2\varsigma)]}{EIA_1} & \frac{c_1c_2l[\cosh(s_1\varsigma)-\cos(s_2\varsigma)]}{\kappa AGA_2} \\ \frac{c_1s_1\cosh(s_2\varsigma)-c_2s_2\cos(s_1\varsigma)}{A_1} & \frac{EIc_1c_2l[s_1\sinh(s_1\varsigma)+s_2\sin(s_2\varsigma)]}{\kappa AGA_2 l} \\ \frac{\kappa AGl[-A_3\sinh(s_1\varsigma)+A_4\sin(s_2\varsigma)]}{EIA_1} & \frac{-l[c_2A_3\cosh(s_1\varsigma)+c_1A_4\cos(s_2\varsigma)]}{A_2} \end{bmatrix} \tag{18}$$

where

$$c_1=\frac{1}{s_1}\left(\frac{\rho l}{\kappa G}p^2\Omega^2+\frac{1}{l}s_1^2\right),\quad c_2=\frac{1}{s_2}\left(\frac{pl}{\kappa G}p^2\Omega^2+\frac{1}{l}s_2^2\right),$$

$$A_1=c_1s_1-c_2s_2,\quad A_2=c_1s_2-c_2s_1,$$

$$A_3=c_1-\frac{s_1}{l},\quad A_4=c_2-\frac{s_2}{l}$$

and $\zeta=(z-z_{i-1})/l$. The transfer matrix formula for the shaft segment can be written as

$$\{S_p(z_i)\}=[T_p(z_i)]\{S_p(z_{i-1})\}. \tag{19}$$

## Transfer Matrix for a Disk

A disk is assumed to be a rigid body with gyroscopic effects and is located at the $i$th station. The thickness of the disk is considered small compared with the length of the shaft and is thus ignored. Unbalance mass of the disk is a source of excitation, which induces synchronous whirling and nonsynchronous whirling for nonlinear shaft bearing systems. From the relationships between forces and deflections of the disk, the equilibrium and compatibility conditions may be written as

$$\begin{aligned} u(z_i,t)&=u(z_{i-1},t)\\ \Psi(z_i,t)&=\Psi(z_{i-1},t)\\ M(z_i,t)&=M(z_{i-1},t)+I_D\ddot{\Psi}(z_{i-1},t)-iJ_p\Omega\dot{\Psi}(z_{i-1},t)\\ Q(z_i,t)&=Q(z_{i-1},t)+M_D\ddot{u}(z_{i-1},t)-M_De\Omega^2\exp(i\Omega t) \end{aligned} \tag{20}$$

where $M_D, I_D$ are the total mass of the disk and its moment of inertia about $x$, $y$-axes, $J_p$ is the polar moment of inertia, and $e$ is the disk mass eccentricity. In Eq. (20), the gyroscopic moment of the disk is included in the moment $M$.

Substituting Eq. (12) into Eq. (20) and equating the coefficients of the like terms associated with $e^{ip\Omega t}$ leads to

$$\begin{aligned} U_p(z_i)&=U_p(z_{i-1})\\ \psi_p(z_i)&=\psi_p(z_{i-1})\\ M_p(z_i)&=M_p(z_{i-1})-(I_Dp^2\Omega^2-J_pp\Omega^2)\psi_p(z_{i-1})\\ Q_p(z_i)&=Q_p(z_{i-1})-M_D(p^2\Omega^2U_p(z_{i-1})+\delta_{1p}e\Omega^2). \end{aligned} \tag{21}$$

Thus, the transfer matrix formula for a disk element may be written as

$$\{S_p(z_i)\}=[T_p(z_i)]\{S_p(z_{i-1})\}+\{P_{pi}\} \tag{22}$$

where the transfer matrix $[T_p(z_i)]$ is given by

$$[T_p(z_i)]=\begin{bmatrix} 1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & -I_Dp^2\Omega^2+J_pp\Omega^2 & 1 & 0\\ -M_Dp^2\Omega^2 & 0 & 0 & 1 \end{bmatrix} \tag{23}$$

and the load vector $\{P_{pi}\}$ is

$$\{P_{pi}\}=\delta_{1p}\{0\ \ 0\ \ 0\ \ -eM_D\Omega^2\}^T. \tag{24}$$

## Transfer Matrix for Nonlinear Bearings

Consider the $i$th element being a nonlinear bearing element, in which the bearing force is expressed in Eq. (1). The equilibrium condition for the bearing force element is

$$Q(z_i,t)=Q(z_{i-1},t)-K_1u(z_{i-1},t)-C\dot{u}(z_{i-1},t)-\tilde{Q} \tag{25}$$

where $\tilde{Q}$ is the nonlinear term in the bearing force expressed as

$$\tilde{Q}=\tilde{Q}_x+i\tilde{Q}_y=K_3[u_x^3(z_{i-1},t)+iu_y^3(z_{i-1},t)]. \tag{26}$$

To determine $\tilde{Q}$, it is necessary to split $u$, defined in Eq. (4), back into $u_x$ and $u_y$ so that $u_x$ and $u_y$ can be expressed in terms of the space and time variables. Separating the forward and the backward whirl components, the displacement $u(z_{i-1},t)$ from Eq. (4) may be rewritten as

$$u(z_{i-1},t)=\sum_{p=1/3,1,3}(U_p(z_{i-1})e^{ip\Omega t}+U_{-p}(z_{i-1})e^{-ip\Omega t}). \tag{27}$$

Rearranging Eq. (27) leads to

$$u(z_{i-1},t)=\sum_{p=1/3,1,3}((U_{cpx}(z_{i-1})+iU_{cpy}(z_{i-1}))\cos(p\Omega t) +(U_{spx}(z_{i-1})+iU_{spy}(z_{i-1}))\sin(p\Omega t)) \tag{28}$$

where

$$U_{cpx}=\mathrm{Re}(U_p+U_{-p}),\quad U_{cpy}=\mathrm{Im}(U_p+U_{-p})$$
$$U_{spx}=\mathrm{Im}(U_{-p}-U_p),\quad U_{spy}=\mathrm{Re}(U_p-U_{-p}).$$

Thus, from $u=u_x+iu_y$, each displacement component, $u_x$ and $u_y$, are obtained as

$$u_x(z_{i-1},t)=\sum_{p=1/3,1,3}(U_{cpx}(z_{i-1})\cos(p\Omega t)+U_{spx}(z_{i-1})\sin(p\Omega t)) \tag{29}$$
$$u_y(z_{i-1},t)=\sum_{p=1/3,1,3}(U_{cpy}(z_{i-1})\cos(p\Omega t)+U_{spy}(z_{i-1})\sin(p\Omega t)).$$

Introducing the harmonic balance method, which assumes that the harmonic components of the bearing force are the same as those in the displacements, the nonlinear forces can be calculated as

$$\tilde{Q}_x=K_3u_x^3(z_{i-1},t)\approx\sum_{p=1/3,1,3}(Q_{cpx}\cos(p\Omega t)+Q_{spx}\sin(p\Omega t)) \tag{30}$$
$$\tilde{\Omega}_y=K_3u_y^3(z_{i-1},t)\approx\sum_{p=1/3,1,3}(Q_{cpy}\cos(p\Omega t)+Q_{spy}\sin(p\Omega t))$$

where $Q_{cpx}$, $Q_{spx}$, $Q_{cpy}$, and $Q_{spy}$ are given by Zu and Ji [13]. Using the following relationships

$$\cos(p\Omega t)=\frac{1}{2}(e^{ip\Omega t}+e^{-ip\Omega t}),$$
$$\sin(p\Omega t)=-\frac{i}{2}(e^{ip\Omega t}-e^{-ip\Omega t}), \tag{31}$$

$\tilde{Q}$ may be computed as

$$\tilde{Q}=\tilde{Q}_x+i\tilde{Q}_y=\sum_{p=1/3,1,3}(\tilde{Q}_pe^{ip\Omega t}+\tilde{Q}_{-p}e^{-ip\Omega t}) \tag{32}$$

where

$$\tilde{Q}_p=\frac{1}{2}[(Q_{cpx}+Q_{spy})+i(Q_{cpy}-Q_{spx})]e^{ip\omega t}$$
$$\tilde{Q}_{-p}=\frac{1}{2}[(Q_{cpx}-Q_{spy})+i(Q_{cpy}+Q_{spx})]e^{-ip\omega t} \qquad p=1/3,1,3. \tag{33}$$

Substituting Eqs. (12) and (32) into Eq. (25) and equating the coefficients of like terms associated with $e^{ip\Omega t}$, the following expression can be obtained:

$$Q_p(z_i)=Q_p(z_{i-1})+(K_1+iCp\Omega)U_p(z_{i-1})+\tilde{Q}_p \tag{34}$$
$$p=-\frac{1}{3},\frac{1}{3},-1,1,-3,3.$$

From Eq. (34), along with the continuity conditions of the displacements and the moment for the bearing element, the transfer matrix formula for the nonlinear bearing force element may be written as

$$\{S_p(z_i)\}=[T_p(z_i)]\{S_p(z_{i-1})\}+\{P_{pi}\} \tag{35}$$

where the transfer matrix is

$$[T_p(z_i)]=\begin{bmatrix}1&0&0&0\\0&1&0&0\\0&0&1&0\\K_1+iCp\Omega&0&0&1\end{bmatrix}\quad p=-\frac{1}{3},\frac{1}{3},-1,1,-3,3 \tag{36}$$

and the force vector is

$$\{P_{pi}\}=\{0\ \ 0\ \ 0\ \ \tilde{Q}_p\}^T. \tag{37}$$

Note that the linear term of the bearing force is reflected in the transfer matrix $[T_p(z_i)]$ and the nonlinear term of the bearing force is reflected by the force vector $\{P_{pi}\}$ in Eq. (37).

## Global Transfer Matrix of the System and Numerical Algorithm

The nonlinear rotor-bearing system considered includes linear shaft segments and rigid disks supported by nonlinear rolling element bearings. The bearing forces are defined by a power nonlinearity in the displacements. The system is divided into $N$ elements, as shown in Fig. 1, in which each bearing, disk and shaft segment is considered as a single element. On the $i$th element, the initial and terminal nodal numbers are $i-1$ and $i$. The transfer matrix of the global system with nonlinear bearings may be written as

$$\{S(z_i,t)\}=\sum_p\{S_p(z_i)\}e^{ip\Omega t}=\sum_p([T_p(z_i)]\{S_p(z_{i-1})\} +\{P_{pi}\})e^{ip\Omega t}\quad i=1,2,\cdots,N \tag{38}$$

where $[T_p(z_i)]$ and $\{P_{pi}\}$ for shaft element, disk element and bearing element are given by Eqs. (18), (23), (24), (36), and (37), respectively.

For linear rotor systems, a constant global matrix may be obtained to relate the initial nodal vector $\{S_p(z_0)\}$ and the terminal nodal vector $\{S_p(z_N)\}$. Thus, from two known boundary conditions at the end point, the two remaining unknown initial parameters can be solved via two linear algebraic equations obtained from the global matrix. For nonlinear rotor-bearing systems, however, the global matrix is a function matrix of the two unknown initial parameters. It is not only difficult to form such a matrix but also meaningless to do so. Hence a numerical scheme is designed to solve for the unknown initial parameters directly without explicit expressions obtained from the global matrix. This scheme is based on a quasi-Newton method for solving simultaneous, nonlinear equations.

The method begins by taking initial guesses for the two unknown parameters from the initial nodal vector $\{S_p(z_0)\}$, denoted by $S_{p,i}(z_0)$ $(p=-1/3,1/3,-1,1,-3,3;i=1,2)$. Thus, $\{S_p(z_i)\}$ up to $\{S_p(z_N)\}$ can be calculated from Eq. (38). Two implicit nonlinear algebraic equations for each $p$ value can be formed as

$$f_{1p}(S_{-1/3,1}(z_0),S_{1/3,1}(z_0),\cdots,S_{3,1}(z_0),S_{-1/3,2}(z_0),\cdots,S_{3,2}(z_0))=S_{p,1}(z_N)$$
$$f_{2p}(S_{-1/3,1}(z_0),S_{1/3,1}(z_0),\cdots,S_{3,1}(z_0),S_{-1/3,2}(z_0),\cdots,S_{3,2}(z_0))=S_{p,2}(z_N) \quad p=-\frac{1}{3},\cdots,3 \tag{39}$$

where $S_{p,i}(z_N)$ $(i=1,2)$ are the two known parameters from the terminal nodal vector $\{S_p(z_N)\}$. Note that there are six $p$ values for each unknown parameter $S_{p,i}(z_0)$, hence Eq. (39) represents six sets of nonlinear algebraic equations for 12 complex unknowns.

First a small tolerance number $\epsilon$ is assumed. Based on the initial guesses for $S_{p,i}(z_0)$ $(p=-1/3,1/3,-1,1,-3,3;\ i=1,2)$, the value for $f_{ip}(i=1,2)$ is calculated. Presumably $|f_{ip}-S_{p,i}(z_N)|$ is greater than $\epsilon$ and therefore $S_{p,i}(z_0)$ is adjusted by a fixed increment. $f_{ip}$ is then re-calculated with the new value for $S_{p,i}(z_0)$. This step is repeated until all values of $|f_{ip}-S_{p,i}(z_N)|$ are less than $\epsilon$. A flowchart of the basic solution scheme is outlined in Fig. 2.

The quasi-Newton method was chosen as the basic root solver over other methods because the quasi-Newton method does not require the derivative of the function. This is crucial since the derivatives of Eq. (39) are unattainable analytically due to the numerical nature of the solution.

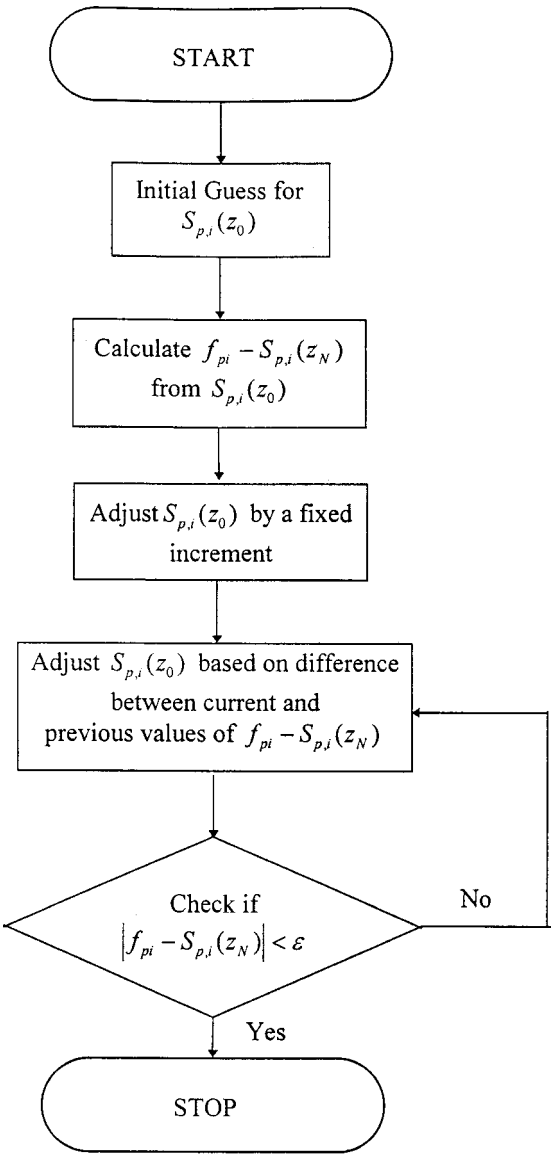



**Fig. 2 A flowchart of solution scheme**

A satisfactory choice of the initial estimate value $S_{p,i}(z_0)$ at the starting point of the frequency range of interest is obtained from the steady-state response of the linear system for synchronous whirl, by putting in zero for the nonlinear bearing stiffness coefficient $K_3$. Substituting the arbitrary initial guesses and $K_3=0$ into Eq. (39), the initial estimate value $S_{p,i}(z_0)$ can be obtained in one iteration.

It can be seen that the proposed transfer matrix method allows a direct transfer of matrices from one end of the system to the other end as expressed in Eq. (38). This procedure is similar to the conventional transfer matrix method for linear systems. In addition, there are only two unknowns for each $p$ value involved in the calculation, and in comparison, 24 unknown vectors need to be solved using iterative procedures in the transfer matrix method proposed by Lee et al. [11] to solve the same nonlinear rotor-bearing system. The initial two unknown boundary conditions are obtained from a set of nonlinear algebraic equations by applying the known boundary conditions. Thus, the proposed method can be applied to any nonlinear rotor-bearing systems without reformulating the equations.

It should be mentioned that the transfer matrix method has certain limitations. The main disadvantage of the method is that the method is not so accurate and may run into numerical problems if the length of individual stations is too long. This is true for both linear and nonlinear systems using the transfer matrix method. However, the method is still a popular method due to its numerical efficiency. The number of stations and the number of variables compared with the finite element method, for example, are much fewer for the same numerical accuracy, and thus the proposed method takes much less computational time.

## Numerical Examples

Two nonlinear rotor-bearing systems are simulated to verify the effectiveness of the method and to demonstrate the nonlinear effect of the bearings on the system's steady-state response. The first example, as shown in Fig. 3, is a shaft supported by two bearings and there is a rotor at the middle. The system is divided into four shaft elements, two nonlinear bearing elements, and one disk element. The physical parameters of the system are given in the following:

$$E=2.07\text{e}7\ \text{N cm}^{-2},\quad K_1=1.5\text{e}4\ \text{N/cm}^{-1},$$
$$G=8.1\text{e}6\ \text{N cm}^{-2},\quad C=600.0\ \text{N s cm}^{-1}$$
$$\kappa=0.68,\quad L_1=1.2\ \text{cm}$$
$$\rho=7.75\text{e-}3\ \text{kg cm}^{-3},\quad L_2=15\ \text{cm}$$
$$M_d=60\ \text{kg},\quad \text{Dia.}=1.2\ \text{cm}$$
$$J_z=2000\ \text{kg cm}^2,\quad e=66.67\ \mu\text{m}$$
$$I_D=1000\ \text{kg cm}^2$$

The maximum amplitude of the shaft response is plotted against the rotating speed in Fig. 4. Due to the sharpness at the resonance, only response curves near the synchronous resonance are shown. Two different nonlinear bearing values are chosen, $K_3=100\ \text{GN/cm}^3$ and $K_3=-100\ \text{GN/cm}^3$. The response curve for linear bearings when $K_3=0$ is also plotted for comparison. In order to verify the solutions, the steady-state response of a simplified system is derived (see Zu and Ji [13]) and presented in Fig.

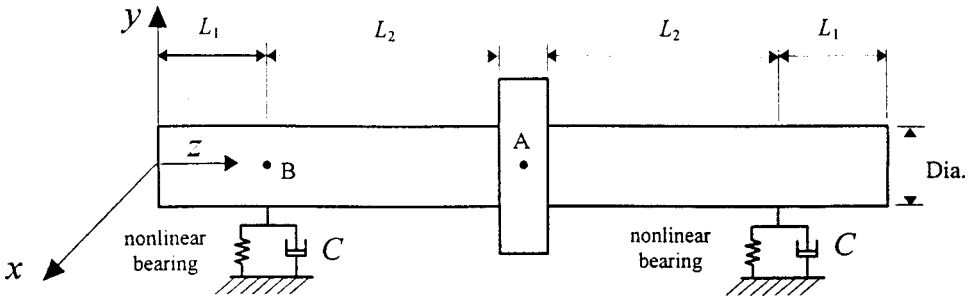



**Fig. 3 A shaft-bearing system with an intermediate rotor**

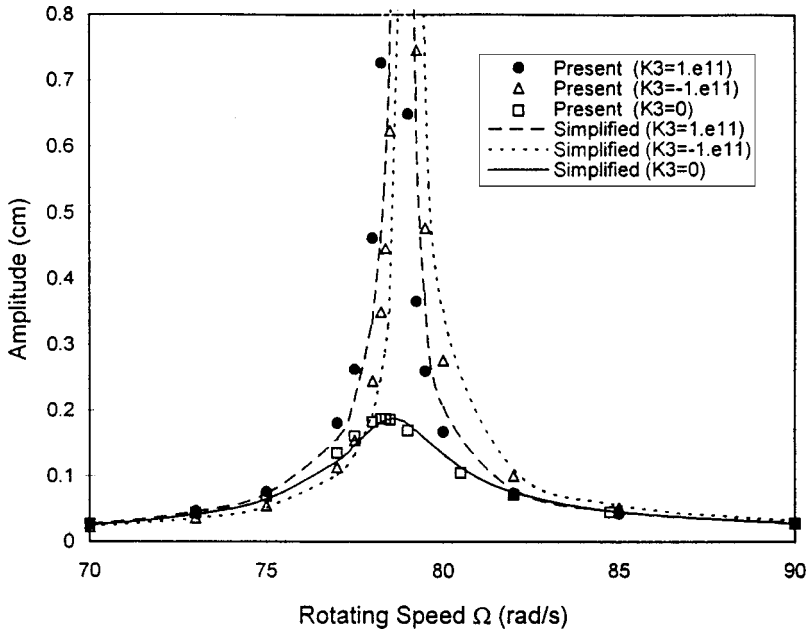



**Fig. 4 Frequency response for various $K_3$**

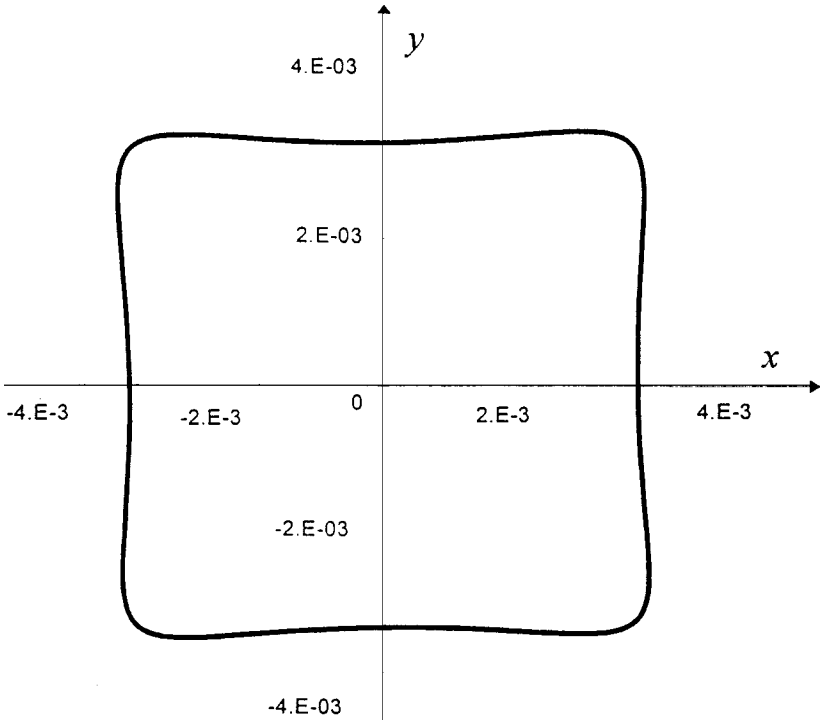



**Fig. 5 Whirling orbit at point *B* for $K_3 = 1.0e1$ 1N/cm$^3$ and $\Omega = 78.8$ rad/s**

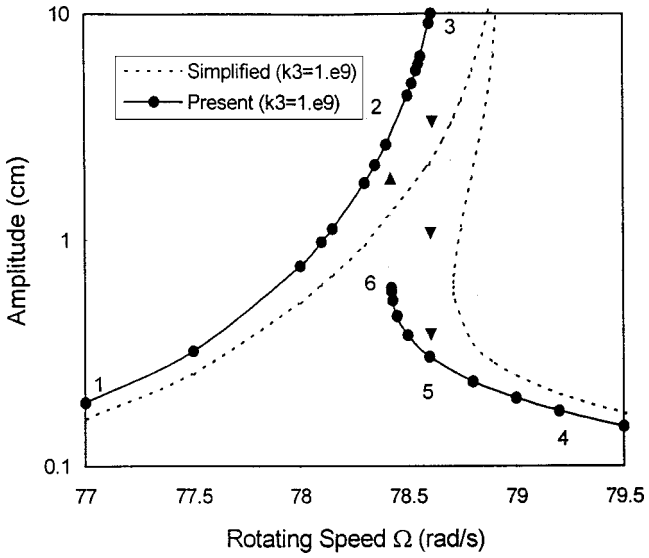



**Fig. 6 Frequency response at $K_3 = 1.0e9 N/cm^3$**

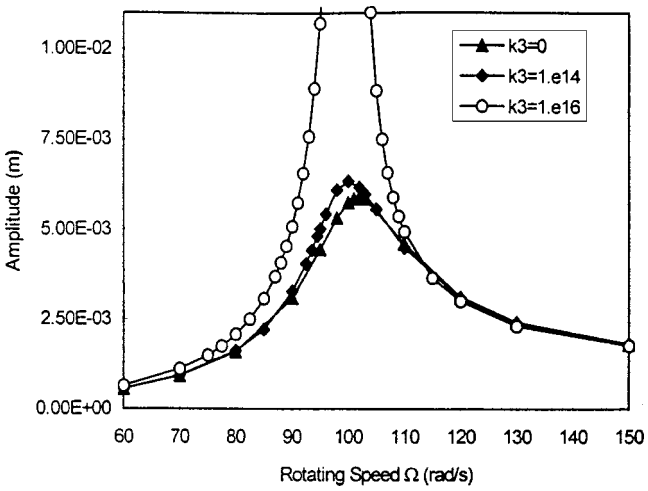



**Fig. 8 Frequency response for various $K_3$**

4, in which only the stiffness of the shaft is considered while the mass of the shaft is ignored. It is shown that the responses obtained from the present solution using the transfer matrix method and from the simplified solution agree quite well. It is observed that the nonlinearity of the bearings increases the vibration level substantially at the resonance compared to the linear system. This occurs because the support from the bearings becomes stiffer with the incorporation of nonlinear bearings, resulting in larger vibrations at resonance. Furthermore, the peak at the resonance is much sharper for nonlinear bearing systems. The peak value for the nonlinear system is also shifted towards the right compared to the linear system.

It should be mentioned that the harmonic balance method is accurate only for small vibrations and thus it does not describe well at the resonance proximity where displacements are large. However, the response curves reflect the trend and the characteristics of the system near the resonance.

The whirling orbit is shown in Fig. 5 at the location of the bearing, point *B*, for $K_3 = 100\ GN/cm^3$ and the resonant frequency $\Omega = 78.8$ rad/s. A noncircular orbit is observed in Fig. 5 due to the influence from the nonlinear bearings that alter the orbit of the shaft. Instead of a circular orbit as is normally the case for linear systems, both circular and noncircular orbits are possible for nonlinear systems.

In Fig. 6, the frequency response, similar to that of Fig. 4, is plotted for $K_3 = 1\ GN/cm^3$. However, a multivalue response curve is seen at this particular $K_3$ value. In another words, there exists a jump phenomenon for this nonlinear system. The response follows curves 1-2-3 as the rotor increases in rotating speed. Conversely, when the speed decreases, the response follows curve 4-5-6 is generated.

The second example is a rotor-bearing system with multiple bearings and disks, as shown in Fig. 7. It includes six beam segments, three nonlinear bearing elements, and two disk elements. The physical parameters of the system are chosen as

$$E = 207\ GN/m^2, \quad K_t = 15\ KN/m^1,$$

$$G = 8.1\ Mn/m^2, \quad C = 6.0\ MNs/m^1$$

$$\kappa = 0.68, \quad L_1 = 0.2\ m$$

$$\rho = 7750\ kg\ m^{-3}, \quad L_2 = 0.5\ m$$

$$M_d = 1.0e4\ kg, \quad Dia. = 0.1\ m$$

$$J_z = 200\ kg\ m^2, \quad e = 0.001\ m$$

$$I_D = 100\ kg\ m^2$$

The response near synchronous resonance is shown in Fig. 8, in which the maximum amplitude on the shaft versus the rotating speed of the shaft is plotted. Response curves corresponding to $K_3 = 0$, $K_3 = 1.0e14\ N/m^3$ and $K_3 = 1.0e16\ N/m^3$ are presented. Similar to the first example, it is seen that as the value for $K_3$ increases, i.e., the strength of the bearing nonlinearity increases, the response may increase dramatically. A whirl orbit at the loca-

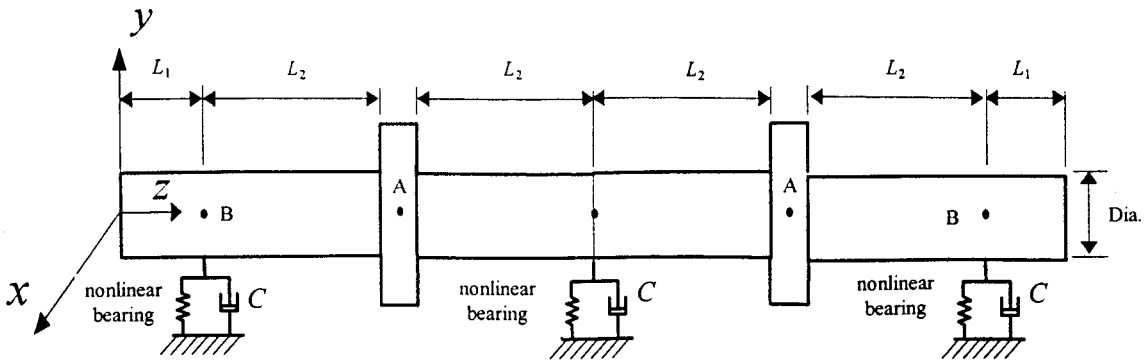



**Fig. 7 A shaft-bearing system with multiple bearings and disks**

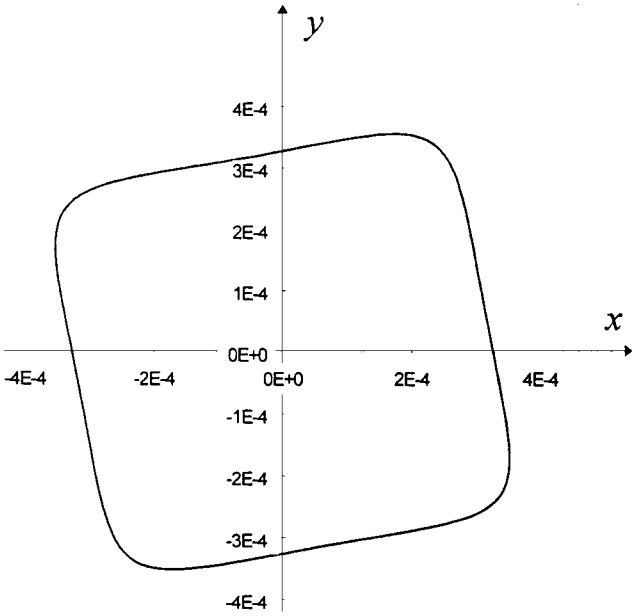



**Fig. 9 Whirling orbits at point *B* for $K_3 = 1.0e16\ N/m^3$ and $\Omega = 96$ rad/s**

tion of the nonlinear bearing, point *B*, is shown in Fig. 9 for $K_3 = 10^7$ GN/m$^3$ and $\Omega = 96$ rad/s. Once again, a noncircular orbit is observed.

## Conclusions

An improved transfer matrix method is presented in this paper for calculating the steady-state response of nonlinear rotor-bearing systems. The systems consist of nonlinear bearings, rigid disks, and Timoshenko shaft segments. Transfer matrices of each component are derived and the global transfer matrix is formed. The harmonic balance method is used to obtain the steady-state response of the nonlinear systems possessing synchronous, subharmonic, and superharmonic vibrations. The main advantage of the proposed method is that it allows a direct transfer of states from one end of the system to the other end, which is similar to the conventional transfer matrix method for linear systems. Thus, the proposed improved transfer matrix method can be applied to any nonlinear rotor-bearing system without reformulating the equations. In addition, there are only two unknown parameters involved in the calculation for each harmonic number. Numerical results show that the shape of natural whirl orbits of the shaft is affected by the presence of the nonlinear bearings and much larger displacements are observed near the synchronous resonant speeds.

## References

[1] Yamamoto, T., Ishida, Y., Ikedam, T., and Yamada, M., 1981, ''Subharmonic and Summed-and-Differential Harmonic Oscillations of an Unsymmetrical Rotor,'' Bull. JSME, **24**, pp. 192–199.
[2] Ishida, Y., Ikeda, T., and Yamamoto, T., 1990, ''Nonlinear Forced Oscillations Caused by Quadratic Nonlinearity in a Rotating Shaft,'' ASME J. Vibr. Acoust., **112**, pp. 288–297.
[3] Nayfeh, A. H., and Mook, D. T., 1979, *Nonlinear Oscillations*, John Wiley and Sons, New York, pp. 59–61.
[4] Yamamoto, T., Ishida, Y., and Aizawa, K., 1979, ''On the Subharmonic Oscillations at Unsymmetrical Shafts,'' Bull. JSME, **22**, pp. 164–173.
[5] Yamamoto, T., Ishida, Y., Ikeda, T., and Yamamoto, M., 1982, ''Nonlinear Forced Oscillations of a Rotation Shaft Carrying on Unsymmetrical Rotor at the Major Critical Speed,'' Bull. JSME, **25**, pp. 1969.
[6] Prohl, M. A., 1945, ''A General Method for Calculating Critical Speeds of Flexible Rotors,'' ASME J. Appl. Mech., **67**, p. 142.
[7] Lund, J. W., 1974, ''Stability and Damped Critical Speeds of a Flexible Rotor in Fluid-Film Bearings,'' J. Eng. Ind., **96**, p. 509.
[8] Gu, J., 1986, ''An Improved Transfer Matrix-Direct Integration Method for Rotor Dynamics,'' ASME J. Vibr. Acoust., **108**, pp. 182–188.
[9] Lee, A. C., Kang, Y., and Liu, S. L., 1991, ''A Modified Transfer Matrix for the Linear Rotor-Bearing System,'' ASME J. Appl. Mech., **58**, pp. 776–783.
[10] Kang, Y., Lee, A. C., and Shih, Y. P., 1994, ''A Modified Transfer Matrix Method for Asymmetric Rotor-Bearing Systems,'' ASME J. Vibr. Acoust., **116**, pp. 309–317.
[11] Lee, A. C., Kang, Y., and Liu, S. L., 1993, ''Steady-State Analysis of a Rotor Mounted on Nonlinear Bearings by the Transfer Matrix Method,'' Int. J. Mech. Sci., **35**, pp. 479–490.
[12] Zu, J. W. and Han, R. P. S., 1994, ''Dynamic Response of a Spinning Timoshenko Beam With General Boundary Conditions and Subjected to a Moving Load,'' ASME J. Appl. Mech., **61**, pp. 152–160.
[13] Zu, J. W. and Ji, Z. Y., 1998, ''Steady-State Response of Rotor Systems With Timoshenko shaft and Nonlinear Bearings,'' ASME J. Eng. Gas Turbines Power, **120**, pp. 751–758.

# Dynamic Response Predictions for a Mistuned Industrial Turbomachinery Rotor Using Reduced-Order Modeling

**R. Bladh**
Graduate Student

**C. Pierre**[1]
Professor,
Fellow ASME

**M. P. Castanier**
Assistant Research Scientist

Department of Mechanical Engineering,
The University of Michigan,
Ann Arbor, MI 48109-2125

**M. J. Kruse**
Senior Engineer
Honeywell Engines & Systems,
Product Safety & Integrity, M/S 2101-121,
Phoenix, AZ 85034-3440

*This paper explores the effects of random blade mistuning on the dynamics of an advanced industrial compressor rotor, using a component-mode-based reduced-order model formulation for tuned and mistuned bladed disks. The technique uses modal data obtained from finite element models to create computationally inexpensive models of mistuned bladed disks in a systematic manner. Both free and forced responses of the rotor are considered, and the obtained results are compared with "benchmark" finite element solutions. A brief statistical study is presented, in which Weibull distributions are shown to yield reliable estimates of forced response statistics. Moreover, a simple method is presented for computing natural frequencies of noninteger harmonics, using conventional cyclic symmetry finite element analysis. This procedure enables quantification of frequency veering data relevant to the assessment of mistuning sensitivity (e.g., veering curvatures), and it may provide a tool for quantifying structural interblade coupling in finite element rotor models of arbitrary complexity and size. The mistuned forced response amplitudes and stresses are found to vary considerably with mistuning strength and the degree of structural coupling between the blades. In general, this work demonstrates how reduced order modeling and Weibull estimates of the forced response statistics combine to facilitate thorough investigations of the mistuning sensitivity of industrial turbomachinery rotors.* [DOI: 10.1115/1.1447236]

## 1 Introduction

In a dynamic analysis of a turbomachinery rotor, one traditionally assumes that the blades are identical. The assumption of cyclic symmetry enables analysts to reduce the computational time considerably by modeling a single sector (e.g., Fig. 1(*b*)) rather than modeling the entire blade assembly (e.g., Fig. 1(*a*)). However, previous studies indicate that cyclic symmetry analysis are seldom adequate for predicting actual blade response—see for instance the work by Ewins [1,2] and Irretier [3]. In practice, there are small differences in the structural properties of individual blades, due to manufacturing and material tolerances or in-service degradation, which are referred to as blade mistuning. These variations destroy the cyclic symmetry of the system and may change qualitatively its dynamic behavior. In particular, mistuning may lead to a confinement of the vibration energy to a few blades or even a single blade. As a result, one or more blades may experience deflections and stresses that are much larger than those predicted by a tuned analysis. Thus, mistuning effects must be included in the analysis in order to obtain accurate predictions of vibration response. Furthermore, blade mistuning is a random quantity, and therefore statistical analyses such as Monte Carlo simulations are critical for safely predicting the response amplitudes in the design process. However, a Monte Carlo simulation for a full finite element blade assembly, such as that shown in Fig. 1(*a*), is enormously costly, and not even feasible for most industrial turbomachinery rotors. Therefore, low-order models and efficient response prediction methods are needed to facilitate the dynamic analysis of mistuned bladed disks.

The primary focus of this paper is to demonstrate that mistuned response statistics can be accurately and efficiently predicted via reduced-order modeling. In order to be useful for industrial applications, the reduced-order model formulation should be based on parent finite element models of arbitrary complexity and size. The technique employed in this paper was originally developed by Castanier et al. [4] and Kruse and Pierre [5], and it was extended to shrouded rotors by Bladh et al. [6]. The procedure involves a component mode analysis of the rotor using a single blade and the cyclic disk as components. This method represents the first effort to generate reduced-order models (ROMs) of turbomachinery rotors directly from their finite element models in a systematic fashion. Since the introduction of this technique, several finite-element-based reduced-order modeling techniques for mistuned bladed disks have followed using modal expansions in terms of component modes ([7–10]), nominal (tuned) system modes with projection of mistuning data ([10,11]), or a combination of both approaches in the form of a secondary modal analysis reduction technique (SMART) ([8,9]). Some of the more recent techniques exhibit significant improvements over the chosen approach in accuracy and/or computational efficiency. However, the selected method has gained popularity through its computer program implementation, REDUCE, which has been used by several members of the U.S. gas turbine community. Furthermore, a wealth of mistuned forced response simulation data—including blade stresses—has been obtained with the REDUCE code for the industrial rotor considered in this study. It is therefore relevant to employ this method to demonstrate the suitability of reduced order models to capture mistuned bladed disk dynamics.

In this paper, the reduced-order modeling technique is employed to study the free and forced dynamic responses of an industrial turbomachinery rotor with blade mistuning. Comparisons are made with ''benchmark'' solutions obtained for the finite element model of the full assembly in Fig. 1(*a*) in order to validate the ROM results. For selected cases of mistuning, the ROM results are projected back into full-field finite element coordinates and post-processed in order to calculate the maximum blade

[1]To whom correspondence should be addressed. E-mail: pierre@umich.edu

Contributed by the Gas Turbine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the GT Division, October 2000; final revision, August 2001. Associate Editor: M. Mignolet.

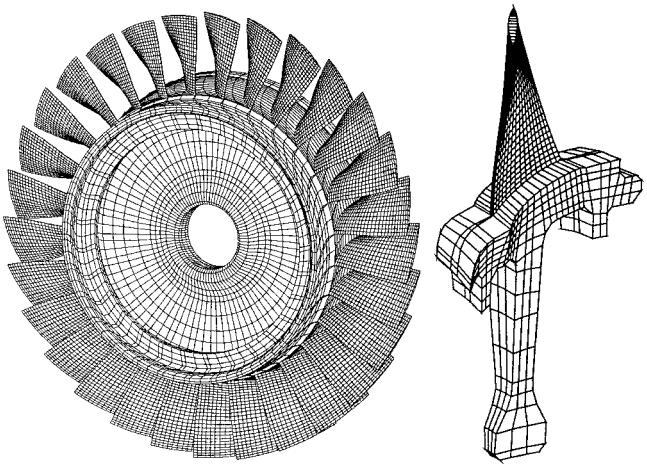



**Fig. 1 Finite element meshes for the industrial 29-blade compressor rotor: (*a*) the full model; (*b*) the fundamental sector**

stresses. Again, the ROM results are compared to those from a full finite element analysis. This is the first ROM-based investigation of the increase in blade stresses due to mistuning, a calculation that is critical to assessing the safety of a rotor design.

In addition, the reduction of the cost of Monte Carlo simulations is considered. Based on the theory of the statistics of extreme values ([12]), it is proposed that the simulation process can be accelerated by assuming that the distribution of the maximum blade amplitudes tend toward a Weibull distribution. A technique from the field of reliability analysis is used to estimate the parameters of the Weibull distribution from relatively few Monte Carlo realizations of mistuned rotors. Results from several 50-rotor Monte Carlo simulations are compared with those from a 1000-rotor simulation in order to demonstrate the effectiveness of this approach.

Once the validity of the ROM and Weibull predictions has been established, these tools are used to examine the effects of interblade coupling on mistuned forced response amplitudes. In unshrouded rotors, the two predominant forms of interblade coupling are aerodynamic coupling and structural coupling through the disk. However, aerodynamic coupling will not be considered in this study. Instead, the focus will be on the effects of structural interblade coupling, which governs the transmission of vibration energy between blades through the disk. Ottarsson and Pierre [13] showed that interblade coupling plays a key role for mistuning sensitivity, such that there is an intermediate range of coupling stiffness at which the system is more susceptible to the effects of mistuning compared to both lower and higher values. So far, this has been shown only for relatively simple lumped parameter systems for which the structural interblade coupling is an easily accessed and controlled quantity. However, for more elaborate finite element models, identifying an effective measure of interblade coupling is much more complicated.

In this work, it is suggested that the curvature of the eigenfrequency veerings when plotted versus the number of nodal diameters can be used to assess the interblade coupling strength ([5,14]). To this end, a simple method is presented for calculating eigenfrequencies for interblade phase angles that are between those corresponding to integer numbers of nodal diameters. By treating the interblade phase angle as a continuous variable rather than a discrete variable, the frequency veerings may be fully captured, and the curvature of the veerings may be quantified. This type of approach was used by Kruse and Pierre [5], but here it is shown how the eigenfrequency calculations may be performed using commercially available finite element software. This extended eigenvalue analysis may provide a tool for quantifying the interblade coupling, which would aid in determining the rotor's sensitivity to mistuning in terms of forced response.

This paper is organized as follows. The reduced-order modeling technique is outlined in Section 2. Section 3 describes the studied rotor and discusses the finite element and reduced-order models used in subsequent analyses. Selected finite element models versus ROM validation results from free and forced response of tuned and mistuned rotors are compared in Section 4. For the forced response, both displacements and stresses are examined. Section 5 discusses the use of Weibull distributions as an approximation for the forced response statistics of randomly mistuned rotors. The method for computing continuous interblade phase angle eigenfrequencies is described in Section 6. This section also includes a brief exploration of the effects of structural interblade coupling on forced response amplitude increases due to mistuning. Finally, conclusions are given in Section 7.

## 2 Reduced-Order Modeling Technique

**2.1 Reduced-Order Model Basis.** It is assumed that the disk (d) and blade (b) degrees-of-freedom are ordered in such a manner as to give the following block-diagonal form for the assembled stiffness matrix of the entire structure:

$$\mathbf{K}=\begin{bmatrix}\mathbf{K_d} & \mathbf{0}\\ \mathbf{0} & \mathbf{K_b}\end{bmatrix}=\begin{bmatrix}\mathbf{I}\otimes\tilde{\mathbf{K}}_\mathbf{d} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\otimes\tilde{\mathbf{K}}_\mathbf{b}\end{bmatrix}, \tag{1}$$

where $\mathbf{I}$ is an identity matrix, and the symbol $\otimes$ denotes the Kronecker product, which is defined in Appendix A. The structure of the mass matrix is identical to that of the stiffness matrix. The ''tilde'' notation is used to indicate that a quantity refers to a single blade or disk sector. Note that the decoupled disk and blade matrix setup in Eq. (1) implies that disk-blade interface degrees-of-freedom appear in both disk and blade matrices, and that degrees-of-freedom on both cyclic boundaries appear in the disk sector matrices. The necessary constraints—i.e., equality between the redundant sets of disk-blade interface degrees-of-freedom and harmonic phase shifts between the cyclic disk boundaries—are applied by including all applicable degrees-of-freedom in the retained mode shapes that form the reduced-order basis. Furthermore, note that the location of the disk-blade interface can be chosen completely arbitrarily. In practice, though, this choice may affect the accuracy of the approximate solutions.

A key idea for the reduced-order modeling technique of Castanier et al. [4] is to describe the motion of the bladed disk assembly using two particular sets of component modes. Figure 2 depicts the two fundamental component mode types for a greatly simplified finite element model of a bladed disk sector.

The first mode set consists of the modes of a single blade that is clamped at the chosen disk-blade interface location (Fig. 2(*a*)). For unshrouded blades, the modal matrix $\mathbf{U^b}$ for all $N$ identical blades is block-diagonal and is assembled as $\mathbf{I}\otimes\tilde{\mathbf{u}}^\mathbf{b}$, where $\tilde{\mathbf{u}}^\mathbf{b}$ is the cantilever mode shapes of a single blade.

The second mode set is comprised of the modes of the disk with massless blades attached (Fig. 2(*b*)). More precisely, these are the cyclic modes of the bladed disk when the density of every blade

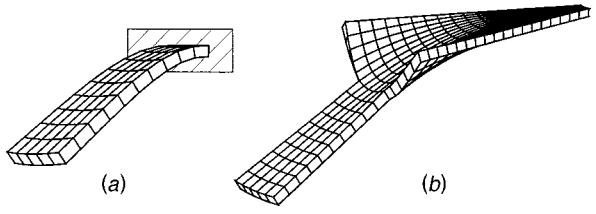



**Fig. 2 Employed component modes: (*a*) normal modes of a cantilevered blade; (*b*) cyclic modes for a fundamental disk-blade sector, where the blade is massless**

finite element is set to zero. Although these are not classical component modes, they serve as the disk component modes. Note that these mode shapes also include disk-induced blade motion, which is comprised of rigid-body motion plus elastic deformation due to the disk-blade boundary motion.

The set of these cyclic disk modes can be written as

$$\mathbf{V}=[\mathbf{V_0}\ \ \mathbf{V_1}\ \ \cdots\ \ \mathbf{V_P}], \tag{2}$$

where the integer $P$ is the highest possible harmonic given by int$[N/2]$, and $N$ is the total number of blades (sectors). In this formulation, the assembly modal vectors associated with the $m$th cosine ($c$) and sine ($s$) mode pair of the $h$th harmonic are expanded as

$$\mathbf{V_{h,m}}=[\mathbf{f_h^c}\otimes\tilde{\mathbf{v}}^{\mathbf{c}}_{\mathbf{h,m}}-\mathbf{f_h^s}\otimes\tilde{\mathbf{v}}^{\mathbf{s}}_{\mathbf{h,m}}\ \ \mathbf{f_h^s}\otimes\tilde{\mathbf{v}}^{\mathbf{c}}_{\mathbf{h,m}}+\mathbf{f_h^c}\otimes\tilde{\mathbf{v}}^{\mathbf{s}}_{\mathbf{h,m}}], \tag{3}$$

where $\mathbf{f_h^{c/s}}$ is the appropriate column of the real Fourier matrix $\mathbf{F}$ defined in Appendix B, and $\tilde{\mathbf{v}}^{\mathbf{c/s}}_{\mathbf{h,m}}$ are the disk mode shapes associated with the $m$th repeated eigenvalue of the $h$th harmonic for a fundamental sector. Note that there is only one mode per eigenvalue for the 0th and, for $N$ even, the $P$th harmonic. Consequently, there is only one vector per mode in these two special cases. The disk mode shapes (recall Fig. 2($b$)) are partitioned into the disk-induced blade motion, $\mathbf{U^d}$, and the disk motion, $\mathbf{V^d}$.

By superposing the two sets of component modes and using the node ordering of Eq. (1), the resulting nodal displacements of the entire assembly can be expanded as

$$\begin{Bmatrix}\mathbf{x_d}\\ \mathbf{x_b}\end{Bmatrix}=\begin{bmatrix}\mathbf{V^d}\\ \mathbf{U^d}\end{bmatrix}\mathbf{a}+\begin{bmatrix}\mathbf{0}\\ \mathbf{U^b}\end{bmatrix}\mathbf{b}, \tag{4}$$

where $\mathbf{a}$ and $\mathbf{b}$ are modal coordinates for the disk modes and the cantilever blade modes, respectively.

The external excitation force vector $\mathbf{Q}$ defines the forcing on all the blade degrees-of-freedom of the assembly. The restriction to blade degrees-of-freedom is not an absolute requirement, but leads to a more compact formulation, and it should also be sufficient from a practical perspective. An engine-order excitation is assumed, which is harmonic in time and differs only in phase from blade to blade. The phase at the $n$th blade, $\phi_n$, is given by

$$\phi_n=\frac{2\pi C(n-1)}{N},\quad n=1,\ldots,N, \tag{5}$$

where $C$ is the engine order of the excitation. The external force vector can then be expressed as

$$\mathbf{Q}=\begin{Bmatrix}\tilde{\mathbf{f}}e^{j\phi_1}\\ \tilde{\mathbf{f}}e^{j\phi_2}\\ \vdots\\ \tilde{\mathbf{f}}e^{j\phi_N}\end{Bmatrix}=\sqrt{N}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{f}}, \tag{6}$$

where $\mathbf{e_{C+1}}$ is the $(\mathbf{C}+1)$th column of the complex Fourier matrix $\mathbf{E}$ defined in Appendix B, and $\tilde{f}$ is the force vector on a single blade. The external virtual work done by the time-harmonic engine order excitation force, $\mathbf{Q}$, may be formulated in component modal quantities as

$$\delta W^{\text{ext}}=\delta\mathbf{x_b^T}\mathbf{Q}=\delta\mathbf{a^T}\mathbf{U^{dT}}\mathbf{Q}+\delta\mathbf{b^T}\mathbf{U^{bT}}\mathbf{Q}. \tag{7}$$

Applying Hamilton's principle yields the governing equations of motion for the reduced-order model, in the absence of aerodynamic coupling. They are conveniently written in matrix form as

$$\mathcal{M}\ddot{\mathbf{z}}+\mathcal{C}\dot{\mathbf{z}}+(1+\gamma j)\mathcal{K}\mathbf{z}=\mathcal{Q}, \tag{8}$$

where

$$\mathcal{M}=\begin{bmatrix}\mathbf{I}+\underset{\mathbf{h}=0,\ldots,\mathbf{P}}{\tilde{\mathbf{B}}\text{diag}}\ [\tilde{\mathbf{u}}_{\mathbf{h}}^{\mathbf{dT}}\tilde{\mathbf{M}}_{\mathbf{b}}\tilde{\mathbf{u}}_{\mathbf{h}}^{\mathbf{d}}] & \mathbf{U^{dT}}(\mathbf{I}\otimes\tilde{\mathbf{M}}_{\mathbf{b}}\tilde{\mathbf{u}}^{\mathbf{b}})\\ (\mathbf{I}\otimes\tilde{\mathbf{u}}^{\mathbf{bT}}\tilde{\mathbf{M}}_{\mathbf{b}})\mathbf{U^d} & \mathbf{I}\end{bmatrix}$$

$$\mathcal{C}=\begin{bmatrix}\mathbf{0} & \mathbf{0}\\ \mathbf{0} & \left(\mathbf{I}\otimes\underset{k=1,\ldots,m_b}{\mathbf{diag}}\ [2\zeta^k]\right)\sqrt{\hat{\mathbf{K}}_{\mathbf{b}}}\end{bmatrix}$$

$$\mathcal{K}=\begin{bmatrix}\hat{\mathbf{K}}_{\mathbf{d}} & \mathbf{U^{dT}}(\mathbf{I}\otimes\tilde{\mathbf{K}}_{\mathbf{b}}\tilde{\mathbf{u}}^{\mathbf{b}})\\ (\mathbf{I}\otimes\tilde{\mathbf{u}}^{\mathbf{bT}}\tilde{\mathbf{K}}_{\mathbf{b}})\mathbf{U^d} & \mathbf{B}\,\underset{n=1,\ldots,N}{\mathbf{diag}}\left[\underset{k=1,\ldots,m_b}{\text{diag}}\ [1+\delta_n^k]\right]\hat{\mathbf{K}}_{\mathbf{b}}\end{bmatrix}$$

$$\mathcal{Q}=\{\mathcal{Q}_{\mathbf{d}}^{\mathbf{T}}\vdots\mathcal{Q}_{\mathbf{b}}^{\mathbf{T}}\}^{\mathbf{T}}=\{\mathbf{0}\cdots\mathbf{0}\mathcal{Q}_{\mathbf{d}}^{\mathbf{C}^{\mathbf{T}}}\mathbf{0}\cdots\mathbf{0}\vdots\mathcal{Q}_{\mathbf{b}}^{\mathbf{T}}\}^{\mathbf{T}}$$

$$\mathcal{Q}_{\mathbf{d}}^{\mathbf{C}}=\begin{Bmatrix}\sqrt{N}\{\mathbf{f}_{\mathbf{C,c}}^{\mathbf{T}}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C,1}}^{\mathbf{d,c}^{\mathbf{T}}}\tilde{\mathbf{f}}-\mathbf{f}_{\mathbf{C,s}}^{\mathbf{T}}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C,1}}^{\mathbf{d,s}^{\mathbf{T}}}\tilde{\mathbf{f}}\}\\ \sqrt{N}\{\mathbf{f}_{\mathbf{C,s}}^{\mathbf{T}}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C,1}}^{\mathbf{d,c}^{\mathbf{T}}}\tilde{\mathbf{f}}+\mathbf{f}_{\mathbf{C,c}}^{\mathbf{T}}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C,1}}^{\mathbf{d,s}^{\mathbf{T}}}\tilde{\mathbf{f}}\}\\ \vdots\\ \sqrt{N}\{\mathbf{f}_{\mathbf{C,c}}^{\mathbf{T}}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C,m_d}}^{\mathbf{d,c}^{\mathbf{T}}}\tilde{\mathbf{f}}-\mathbf{f}_{\mathbf{C,s}}^{\mathbf{T}}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C,m_d}}^{\mathbf{d,s}^{\mathbf{T}}}\tilde{\mathbf{f}}\}\\ \sqrt{N}\{\mathbf{f}_{\mathbf{C,s}}^{\mathbf{T}}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}_{\mathbf{,m_d}}^{\mathbf{d,c}^{\mathbf{T}}}\tilde{\mathbf{f}}+\mathbf{f}_{\mathbf{C,c}}^{\mathbf{T}}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}_{\mathbf{C,m_d}}^{\mathbf{d,s}^{\mathbf{T}}}\tilde{\mathbf{f}}\}\end{Bmatrix}$$

$$\mathcal{Q}_{\mathbf{b}}=\sqrt{N}\mathbf{e_{C+1}}\otimes\tilde{\mathbf{u}}^{\mathbf{b}^{\mathbf{T}}}\tilde{\mathbf{f}};\quad \mathbf{z}=\begin{Bmatrix}\mathbf{a}\\ \mathbf{b}\end{Bmatrix}.$$

$\tilde{\mathbf{B}}$**diag** $[\cdot]$ denotes a pseudo-block-diagonal matrix, with the argument being the $h$th "block," and the range of $h$ is shown. Note that the disk portion of the modal force vector is zero everywhere, except for the $C$th harmonic mode shapes, due to orthogonality between harmonics of differing orders. Moreover, structural damping with coefficient $\gamma$, as well as viscous modal damping of the individual cantilever blade modes, $\zeta^k$, have been added to the reduced-order model to facilitate more realistic modeling of the structure's dynamic response.

The reduced-order model represented by Eq. (8) also includes blade stiffness mistuning. The modal stiffness of each cantilever blade mode for each blade is isolated in the diagonal matrix $\hat{\mathbf{K}}_{\mathbf{b}}$. Therefore, the formulation lends itself to a straightforward input of blade mistuning via the random mistuning parameters $\delta_n^k$:

$$(\omega_n^k)^2=(1+\delta_n^k)(\omega^k)^2 \tag{9}$$

where $\omega_n^k$ represents the mistuned natural frequency of the $k$th mode of blade $n$, and $\omega^k$ is the corresponding nominal, or tuned, natural frequency. Note that this implementation of mistuning incorporates two approximations:

- The mistuned characteristics of a blade are restricted to its stiffness (lower-right quadrant of $\mathcal{K}$). While stiffness mistuning is sufficient for the purposes of this study, it may be more accurate to model mistuning in other structural parameters as well; for instance, by using the mixed least squares–maximum likelihood method of Mignolet and Lin [15].
- The effects of stiffness mistuning on the other three quadrants of $\mathcal{K}$ are neglected. To a large extent, this is justified by assuming that the disk-induced blade motion is mostly rigid-body motion.

As a concluding remark, the reduced-order model for unshrouded bladed disks requires the following input from finite element analyses:

$\hat{\mathbf{K}}_{\mathbf{b}}, \tilde{\mathbf{u}}^{\mathbf{b}}$: The modal stiffnesses and corresponding mode shapes of a single blade, clamped at the chosen disk-blade interface.

$\hat{\mathbf{K}}_{\mathbf{d}}, \tilde{\mathbf{u}}^{\mathbf{d}}$: The modal stiffnesses and corresponding cyclic mode shapes from a cyclic symmetry eigenanalysis for a bladed disk sector, where the density of the blade's elements are set to zero.

$\tilde{\mathbf{M}}_{\mathbf{b}}, \tilde{\mathbf{K}}_{\mathbf{b}}$: The finite element mass and stiffness matrices for an unconstrained blade.

**2.2 Cantilever Blade Eigenvalue Adjustments.** It should be noted that the disk motion is described only by the disk portion from the second mode set (Fig. 2($b$)), and no separate set of constraint modes is employed to yield a complete description of the disk-blade interface motion. This causes the representation of the disk-blade interface to be too stiff, which degrades the performance (modal convergence) of the method. However, it has been

found that an artificial softening of the cantilevered blade modes yields significant accuracy improvements for both free and forced response ([4,6]).

The artificial softening of the cantilevered blade is achieved by adjusting the eigenvalues (modal stiffnesses) of the cantilevered blade modes in an iterative fashion. The cantilevered blade eigenvalues are scaled linearly based on the ratio of the tuned ROM and exact (in the finite element sense) system eigenvalues for blade-dominated modes. For the $i$th eigenvalue adjustment iteration, the scaled cantilevered blade eigenvalues are computed as

$$(\omega_b^k)^2|_i = \frac{(\omega_{fe}^{P,k})^2}{(\omega_r^{P,k})^2|_{i-1}} (\omega_b^k)^2|_{i-1}, \quad k=1,\ldots,m_b. \tag{10}$$

$\omega_{fe}^{P,k}$ is the exact eigenvalue of the tuned assembly mode characterized by the $k$th cantilevered blade mode family at the highest possible harmonic ($P$), $\omega_r^{P,k}$ is the corresponding eigenvalue from the tuned ROM, and $m_b$ is the number of retained cantilever blade modes. Although heuristic, this technique has been found to be very efficient. In one case study, it was seen that residual errors among the blade-dominated mode eigenvalues were reduced to less than 0.001 percent after as few as two iterations ([6]).

Finally, note that the problems associated with incomplete disk-blade interface descriptions are alleviated in Yang and Griffin's second technique ([11]), as well as in the ''true'' component mode synthesis method and the secondary modal analysis reduction technique (SMART) developed by Bladh et al. [8,9]. Yang and Griffin's second technique employs the true modes of the tuned assembly and thus captures exactly—with respect to the finite element model—the dynamic characteristics of the tuned system, such as the frequency veerings. With the same objective, the methods of Bladh et al. use a complete set of constraint modes to yield an exact disk-blade interface description, such that the truncation of component normal modes is the only source of (small) errors in the tuned free response.

## 3 Rotor Description and Computational Models

The industrial rotor illustrated in Fig. 1 is the second stage of a four-drum compressor rotor used in an advanced gas turbine application. There are 29 blades in the rotor. The design is referred to as a blisk, since the blades and disk are machined from a single, continuous piece of material. The material properties are listed in Table 1.

The tuned finite element model is represented by the single sector model in Fig. 1($b$), using MSC/NASTRAN™ cyclic symmetry routines. The sector model is clamped at the ribs located at the outer edges of the disk, which is a rough approximation of boundary conditions due to neighboring stages. The sector finite element model is constructed with standard linear brick elements (eight-noded solids). The disk portion of the model contains 528 elements, and the blade portion has 375 elements. There are 4374 degrees-of-freedom per sector in the finite element model. In contrast, the ROM used for technique validation consists of five disk modes and ten cantilever blade modes, for a total of 15 degrees-of-freedom per sector.

The mistuned finite element model consists of the entire blade assembly, as shown in Fig. 1($a$). The same element type and mesh pattern is used in the single-sector cyclic symmetry model and the full mistuned model. There are thus 126,846 degrees-of-freedom in the mistuned finite element model, as compared to 435 degrees-of-freedom in the mistuned ROM. Mistuning is introduced into the assembly by allowing each blade to have a different Young's modulus:

$$E_n = E_o(1+\delta_n), \quad n=1,\ldots,N, \tag{11}$$

where $E_o$ is Young's modulus for a tuned blade, and $\delta_n$ is a dimensionless mistuning parameter associated with the $n$th blade. A specific mistuning distribution, or pattern, for the industrial rotor that was used to obtain the finite element model versus ROM validation results is listed in Table 2. These mistuning parameters are based on experimental natural frequency measurements on a prototype rotor.

**Table 1 Industrial blisk material properties**

| Material Property | Notation | Value | Unit |
|---|---|---|---|
| Young's modulus | $E_o$ | 203.4 | GPa |
| Mass density | $\rho$ | 7909 | kg/m$^3$ |
| Poission's ratio | $\nu$ | 0.305 | - |
| Structural damping | $\gamma$ | 0.006 | - |

**Table 2 Mistuning pattern for the case study rotor, based on natural frequency measurements on a prototype rotor**

| Blade | Mistuning $\delta_n$ | Blade | Mistuning $\delta_n$ |
|---|---|---|---|
| 1 | 0.05704 | 16 | 0.04934 |
| 2 | 0.01207 | 17 | 0.04479 |
| 3 | 0.04670 | 18 | 0.03030 |
| 4 | −0.01502 | 19 | 0.00242 |
| 5 | 0.05969 | 20 | 0.01734 |
| 6 | −0.03324 | 21 | 0.02919 |
| 7 | −0.00078 | 22 | −0.00328 |
| 8 | −0.01688 | 23 | 0.00086 |
| 9 | 0.00242 | 24 | −0.03654 |
| 10 | −0.02747 | 25 | −0.03631 |
| 11 | −0.03631 | 26 | −0.01665 |
| 12 | −0.03570 | 27 | 0.00783 |
| 13 | −0.03631 | 28 | −0.01169 |
| 14 | −0.03631 | 29 | −0.01332 |
| 15 | 0.00242 | | |

## 4 Reduced-Order Model Validation

**4.1 Free Response.** It is convenient to describe the mode shapes of a tuned rotor in terms of nodal diameters (nodal lines across the diameter of the disk) and nodal circles (nodal lines in the circumferential direction). The number of nodal diameters for a mode shape corresponds to a certain phase shift between adjacent blades, called the interblade phase angle, which is given by

$$\phi_h = \frac{2\pi h}{N}, \quad h=0,\ldots,P, \tag{12}$$

where $h$ is the number of nodal diameters, and $N$ is the number of blades in the assembly. These modes are accordingly referred to as constant interblade phase angle modes.

The characteristics of the free vibration modes are conveniently summarized by a plot of natural frequencies versus the number of nodal diameters, as shown in Fig. 3. This plot reveals two interesting features of the tuned rotor's modal structure. First, as the number of nodal diameters increases, the disk stiffens rapidly, and the slanted lines in Fig. 3 thus correspond to disk-dominated modes. In the absence of blade-to-blade shrouds, the blade-dominated modes do not stiffen significantly as the number of nodal diameters increases. Hence, lines that are approximately horizontal represent families of blade-dominated modes. For instance, the family of modes around 2150 Hz features motion that is dominated by the first flexural mode (1F) of a cantilevered blade, while motion in the second family of modes around 7400 Hz is dominated by the first torsional mode (1T). Neither of these families of modes exhibit nodal circles. Motion in the third family of modes (9100 Hz) is dominated by the second flexural blade mode (2F), which features one nodal circle, and so on. The physical descriptions of the blade-dominated mode families are listed on the right-hand side of Fig. 3, where S (for ''Stripe'') denotes plate-type modes characterized by flexural motion along the chord of the blade.

A second notable feature in Fig. 3 is the numerous eigenfrequency veerings—regions in which blade and disk-dominated

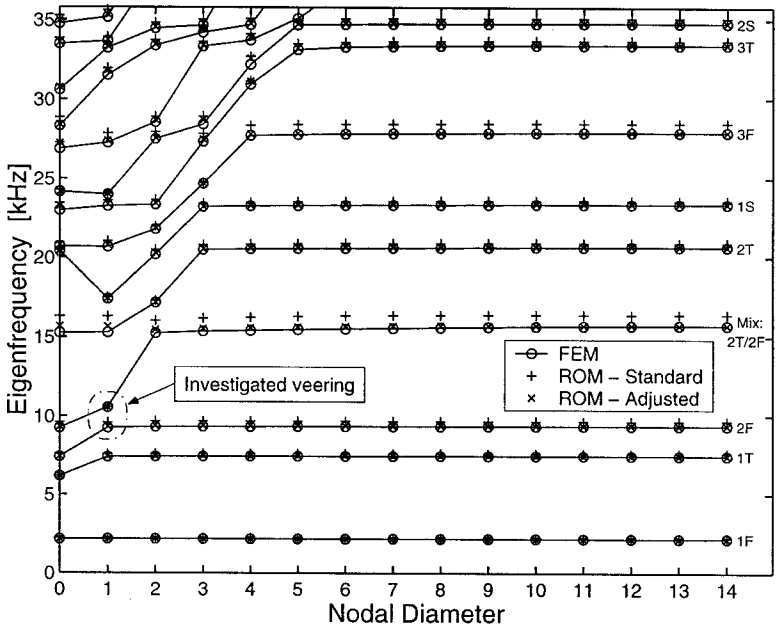



**Fig. 3 Natural frequencies versus number of nodal diameters for the tuned rotor by finite element and reduced-order modeling**

mode families appear to veer away from each other. Physically, eigenfrequency veerings indicate the degree of coupling between families of disk and blade modes. The strength of a veering may be measured by the distance between the natural frequencies and the local curvature in the veering region ([14]). The impact of the eigenfrequency veerings on the forced response of mistuned rotors will be highlighted in later sections.

The average error for the standard (unadjusted) ROM among the natural frequencies in the frequency range shown in Fig. 3 is 2.2 percent, where the maximum error of 6.9 percent is found at the fifth zero-nodal diameter mode. The corresponding errors after cantilever blade eigenvalue adjustments are 0.3 percent and 2.6 percent, respectively. Note that all ROM results from here on refer to the adjusted ROM. The demonstrated accuracy of the adjusted ROM in terms of tuned natural frequencies also extends to mode shapes. This is evidenced in Figs. 4 and 5, which depict example four and one nodal diameter mode shapes, respectively, obtained from finite element and reduced order models. The four nodal diameter mode is characterized by 1F motion, while the one nodal diameter mode is dominated by 2F motion. The mode shapes are represented by a scalar denoted "relative blade Euclidean displacement norm." The norm, $\overline{u}_n^r$, for each blade $n$, is defined as

$$\overline{u}_n^r = \frac{\sqrt{\sum_{j=1}^{N_b} u_{j,n}^2}}{\sqrt{\sum_{n=1}^{N}\left[\sum_{j=1}^{N_b} u_{j,n}^2\right]}}, \quad n=1,\ldots,N, \tag{13}$$

where $u_{j,n}$ is the displacement of the $j$th degree-of-freedom of the $n$th blade, $N_b$ is the number of degrees-of-freedom in one blade,

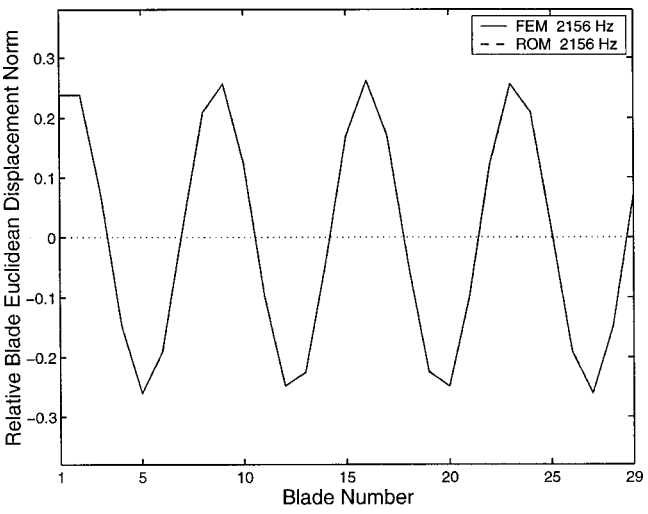



**Fig. 4 Comparison of tuned finite element and ROM four nodal diameter mode shapes (dominated by 1F motion)**

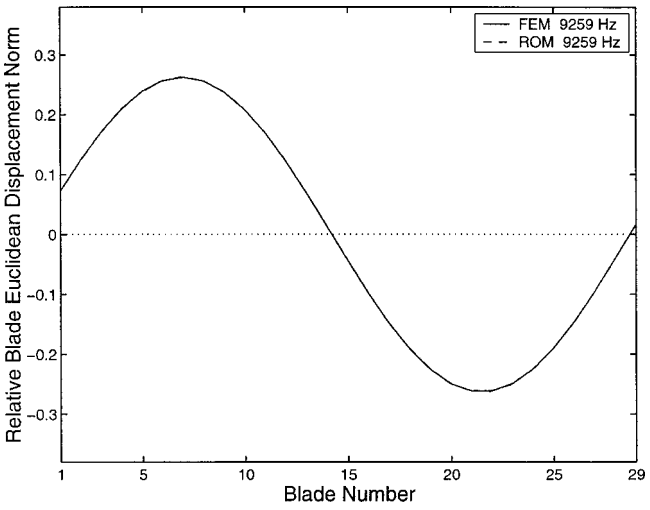



**Fig. 5 Comparison of tuned finite element and ROM one nodal diameter mode shapes (dominated by 2F motion). This mode is located in the investigated veering.**

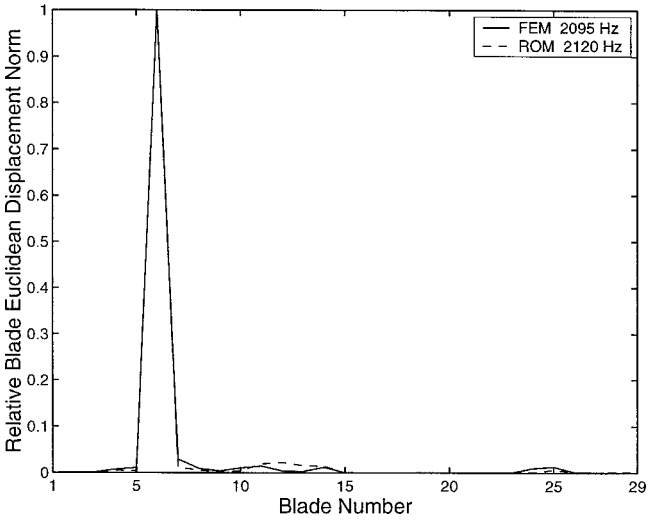



**Fig. 6 Comparison of mistuned finite element and ROM mode shapes in the frequency region encompassing the 1F blade-dominated modes. The mode shape is spatially localized about blade number six.**

and $N$ is the number of blades in the assembly. For these tuned mode shapes, the signs of the blade deflections were identified to enable a more familiar sinusoidal representation, since the motion extends uniformly throughout the blade assembly. Note that the one nodal diameter mode is located in the center of the investigated eigenfrequency veering (see Fig. 3).

The nodal diameter description implies that the mode shapes of the rotor are spatially periodic, which is true for tuned rotors. However, small blade mistuning may alter mode shapes and cause confinement of vibration energy to only a few blades—the so-called localization phenomenon. It has been shown that the first-order mode shape perturbation due to mistuning is inversely proportional to the difference in the tuned system's natural frequencies, which indicates that mode localization is most acute in frequency regions of high modal density ([14]). From Fig. 3, it can be deduced that all families of blade-dominated modes in the depicted range exhibit high modal density and are therefore susceptible to mode localization upon the introduction of mistuning.

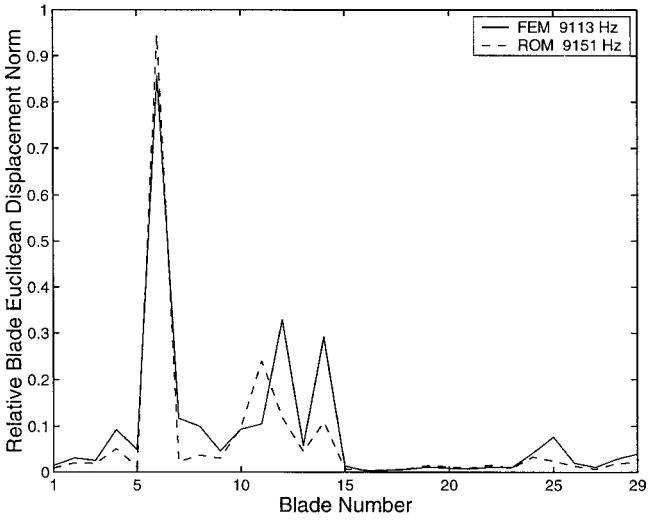



**Fig. 7 Comparison of mistuned finite element and ROM mode shapes in the frequency region encompassing the investigated veering. Motion is dominated by the 2F blade mode and is localized about blade number six.**

As expected, the selected mode shapes from the mistuned rotor model depicted in Figs. 6 and 7 demonstrate significant mode localization in both finite element and ROM representations. The ROM mode shapes correlate well with those of the much larger finite element model, even though the degrading effects of mode truncation and incomplete disk-blade interface description are clearly visible. In particular, the ROM captures adequately the location of peak deflection, as well as the rate of spatial decay in blade participation.

**4.2 Forced Response.** The external excitation force vector used for the forced response analysis consists of point loads located at the tip and midchord of each blade. These loads are applied in a direction normal to the surface of the blade. Of course, this is a simplified version of an actual loading case, in which the blades would be subject to distributed pressure loads. More realistic pressure loads can be applied to the ROM by specifying the equivalent nodal loading on all blade degrees-of-freedom. Structural damping, $\gamma$, which had been experimentally determined to be 0.006, is used in both the finite element and ROM forced response calculations. A scalar representation of blade deflection amplitude similar to Eq. (13) is utilized to represent the forced response of the assembly:

$$\overline{u}_n = \sqrt{\sum_{j=1}^{N_b} u_{j,n}^2}, \quad n = 1, \ldots, N. \tag{14}$$

In the tuned system response, all blades have identical peak displacement norms. However, this is not true for a mistuned rotor, where the various blades may have vastly different response amplitudes. In the mistuned case, the maximum blade displacement norm throughout the assembly is selected at each driving frequency, defining the maximum response amplitude.

Figure 8 depicts the frequency response of the tuned assembly using both finite element and reduced-order models in the frequency region encompassing the 1F family of blade-dominated modes. An engine-order four excitation is applied ($C=4$). Figure 8 shows that the resonant response amplitude of the tuned system using the ROM formulation is less than two percent higher than the corresponding finite element result.

Figure 9 depicts a comparison of mistuned finite element and ROM maximum blade responses. The highly complex response pattern is very well captured by the ROM. Only small discrepancies in resonant frequencies and response amplitudes are observed. The response sampling for the mistuned finite element model is relatively coarse, due to the tremendous computational

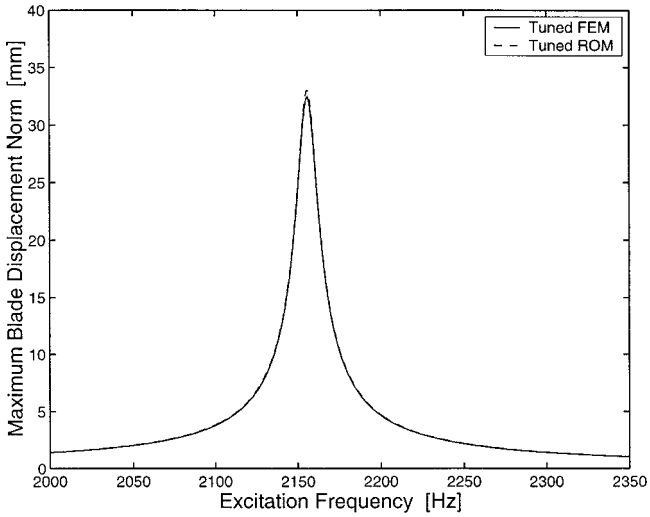



**Fig. 8 Comparison of tuned finite element and ROM forced responses, for blade tip excitation with $C=4$**

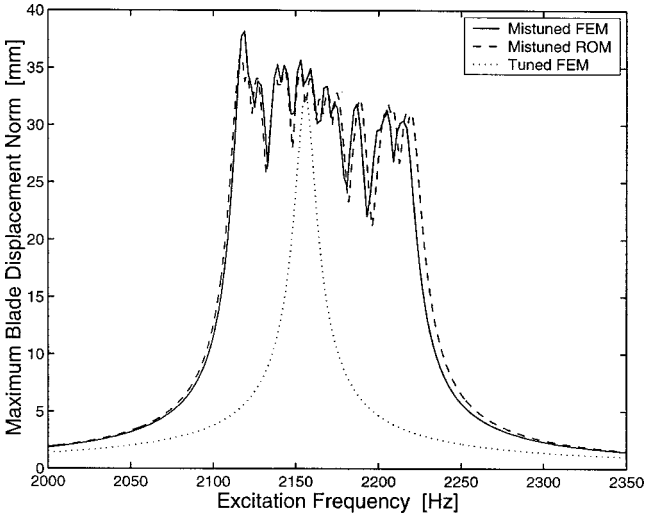



**Fig. 9 Comparison of mistuned finite element and ROM maximum blade forced responses, for blade tip excitation with $C=4$**

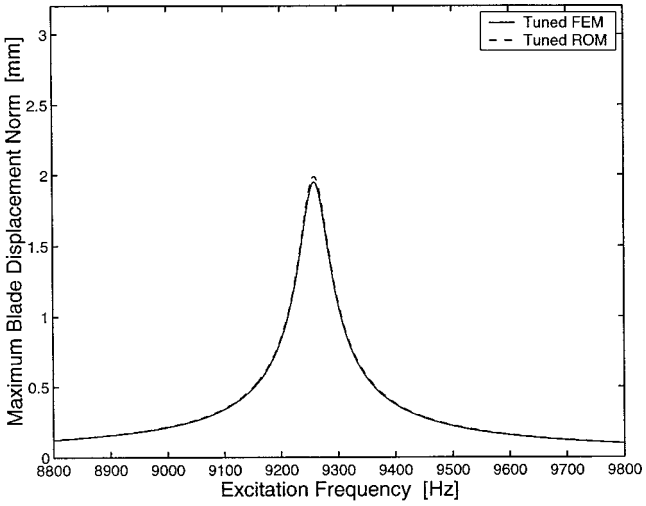



**Fig. 10 Comparison of tuned finite element and ROM forced responses, for blade tip excitation with $C=1$**

expense associated with solving a full blade assembly. Each sampled response (one data point) for the mistuned finite element model requires approximately 27 cpu minutes on a 360 MHz UNIX workstation. In contrast, each ROM response sample requires less than 2.7 cpu seconds on the same workstation.

Regarding Fig. 9, note that the maximum mistuned response amplitude is only 17 percent higher than the tuned resonant amplitude. Recall, from the work by Wei and Pierre [16] and Ottarsson and Pierre [13], that significant amplitude increases occur when there is moderately weak interblade coupling. From Fig. 3 it is seen that this particular frequency region exhibits no eigenfrequency veerings. This implies minimal disk-blade modal interaction and very weak interblade coupling. Hence, the mode shapes are dominated by cantilever blade motion, where the individual blades are isolated from each other by the disk and cannot effectively communicate energy to one another. The mistuned system thus responds very much like a collection of uncoupled blades, yielding a relatively modest response amplitude increase due to mistuning.

An excitation close to an eigenfrequency veering is considered next, since it is crucial that interblade coupling effects be well captured by the ROM. Figure 10 illustrates the tuned finite element and ROM responses to an engine order one excitation. The resonance in Fig. 10 corresponds to the 2F mode of a cantilever blade. It is located in the heart of the investigated eigenfrequency veering, between the third family of blade-dominated modes and the first family of disk-dominated modes, as can be seen in Fig. 3. As with the 1F mode family, the correlation between tuned finite element and ROM responses in Fig. 10 is excellent.

Figure 11 depicts a comparison of mistuned finite element and ROM maximum blade response amplitudes for engine order one excitation. It is again seen that the ROM yields a good representation of the mistuned response pattern. However, the ROM response amplitudes do not compare as well as in the previous case. The accuracy of the mistuned response of blade-dominated mode families is critically dependent on the accurate capture of any existing frequency veerings. Hence, even the very small residual errors that remain after eigenvalue adjustments among the modes comprising the investigated veering give rise to clearly visible discrepancies in the mistuned response, as shown in Fig. 11. However, the maximum mistuned response amplitude, which is the most critical quantity, is still captured by the ROM to within six percent of the full finite element solution, with only a small fraction of the computational effort. Note that the maximum mistuned response exceeds the tuned response by 57 percent. As mentioned previously, this is due to the increased disk-blade modal interaction found in veering regions, enabling vibration energy to be transferred from lower to higher-responding blades.

Once the full displacement field has been solved with the efficient ROM, displacements can be imported back into the finite element model for post-processing of stresses and strains. For the engine-order one excitation, the maximum principal stresses in the blade at tuned resonance are 507 and 527 MPa (+4 percent) as obtained with the finite element model and the ROM, respectively. The corresponding maximum principal stresses in the mistuned case are 777 and 751 MPa (−3 percent). Hence, the maximum principal stresses obtained with the ROM are in this case within ±5 percent of the principal stresses determined with the finite element model. Furthermore, note that the maximum principal stress for the mistuned finite element model is as much 53 percent higher than that for the tuned model. This 53 percent increase in principal stress corresponds to the 57 percent increase in resonant response amplitude observed in Fig. 11.

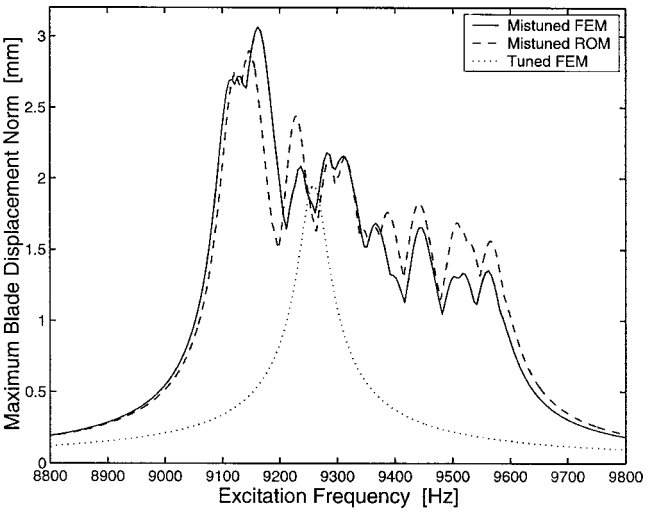



**Fig. 11 Comparison of mistuned finite element and ROM maximum blade forced responses, for blade tip excitation with $C=1$**

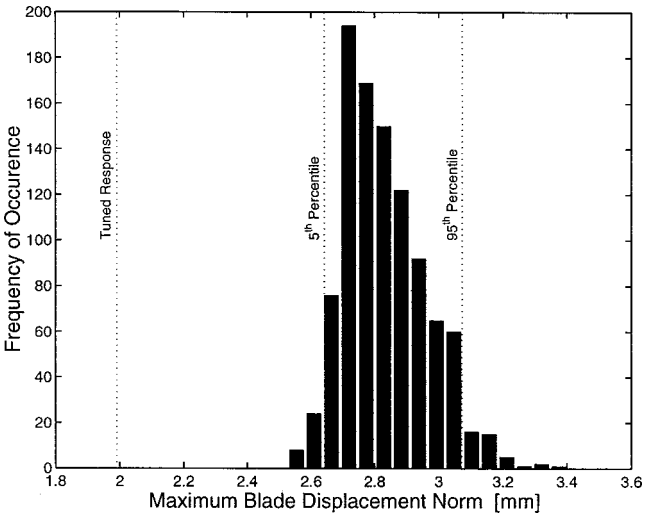



**Fig. 12 Histogram of the maximum blade response amplitudes for engine order one excitation. Obtained by Monte Carlo simulation of 1000 different mistuned systems with uniform distributions of zero mean and three percent standard deviation.**

## 5 Forced Response Statistics

**5.1 Monte Carlo Simulations.** As the above free and forced response results illustrate, reduced-order models of bladed disks can correlate well with much larger finite element models. In addition, reduced-order modeling enables engineers to estimate the statistical characteristics of blade forced response amplitudes for randomly mistuned rotors. The harmful effects of random mistuning must be compensated for by increasing the overall fatigue strength of the blades in order to meet some statistically determined stress level. Alternatively, if feasible, the designer could choose an intentional mistuning pattern in an attempt to minimize the increase in mistuned vibratory stresses. Recent work by Castanier and Pierre [17,18] indicates that this is a promising approach. Note that in both instances, reduced-order modeling can aid the designer in capturing mistuning effects.

Figure 12 shows results from an ROM-based Monte Carlo simulation for engine-order one excitation over the frequency range 8750–10,750 Hz, which encompasses the 2F family of blade modes. The simulation consists of frequency sweeps for 1000 different mistuning patterns, obtained from a uniform distribution with zero mean and three percent standard deviation. In Fig. 12, the 95th percentile of the maximum blade response amplitude—the amplitude such that 95 percent of all mistuned rotors are expected to exhibit lower maximum blade amplitudes—is seen to correspond to a 54 percent increase over the tuned system's resonant response amplitude. The corresponding maximum principal stress for one 95th percentile mistuning pattern is 830 MPa. Recall that the tuned maximum principal stress is 527 MPa. The stress level for the 95th percentile of the maximum blade response amplitude is thus as much as 58 percent higher than the tuned stress level. Hence, if the current rotor design is based on tuned stress magnitudes, fatigue properties of the blades should be increased significantly to compensate for the effect of random mistuning.

**5.2 Accelerated Monte Carlo Simulations.** A simulation using 1000 mistuning patterns is computationally expensive, even with the ROM. In addition, if compliance of 95 percent of the rotors is not acceptable, and, for instance, 99.9th percentile compliance is required, 1000 realizations will not suffice. Determining the 99.9th percentile forced response amplitudes with Monte Carlo simulations requires analyses of an estimated 50,000 mistuned rotors, which is a formidable task.

Therefore, it is proposed that the number of Monte Carlo realizations may be reduced by estimating the distribution of the forced response amplitudes from a relatively small set of simulated rotors. From the theory of the statistics of extremes ([12]), the distribution of the maximum blade amplitudes for a population of mistuned rotors will tend to one of three extreme value distributions. Furthermore, since the vibration response is bounded, the distribution of the largest-responding blade amplitudes will asymptotically approach the third extreme value distribution, which is the Weibull distribution. The cumulative distribution function (cdf) for the three-parameter Weibull distribution for largest values is ([12,19])

$$F(x) = e^{-((\lambda - x)/\delta)^{\beta}}, \tag{15}$$

and the associated probability density function (pdf) is

$$f(x) = \frac{\beta}{\delta}\left(\frac{\lambda - x}{\delta}\right)^{\beta - 1} e^{-((\lambda - x)/\delta)^{\beta}}, \tag{16}$$

where $\delta$, $\beta$, and $\lambda$ are the scale, shape, and location parameters, respectively, and $x$ is the random variable under investigation—in this case, the peak amplitude of the largest-responding blade on a mistuned rotor.

In order to estimate the three parameters of the Weibull distribution, one may begin by applying the operator $\ln(-\ln(\cdot))$ to each side of Eq. (15) and rearranging to retrieve the form

$$\ln(\lambda - x) = \frac{1}{\beta}\ln(-\ln(F)) + \ln(\delta). \tag{17}$$

The results from a small set of Monte Carlo simulations may be used for the $x$ values, and the cdf values (percentiles) are found by sorting this data set.

Next, an estimated value is assigned to the location parameter, $\lambda$, which represents the upper limit of $x$. For the case of mistuned forced response amplitudes, an approximation derived by Whitehead [20,21] is employed

$$\lambda \approx \frac{1}{2}(1 + \sqrt{N}). \tag{18}$$

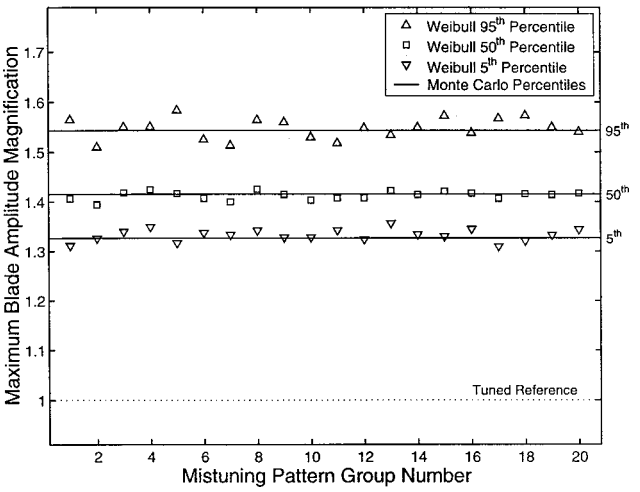



**Fig. 13 Comparison of Weibull and Monte Carlo determined responses for 5th, 50th, and 95th percentiles of maximum blade response amplitude magnification. The approximate percentiles from the Weibull distributions conform well with the Monte Carlo percentiles.**

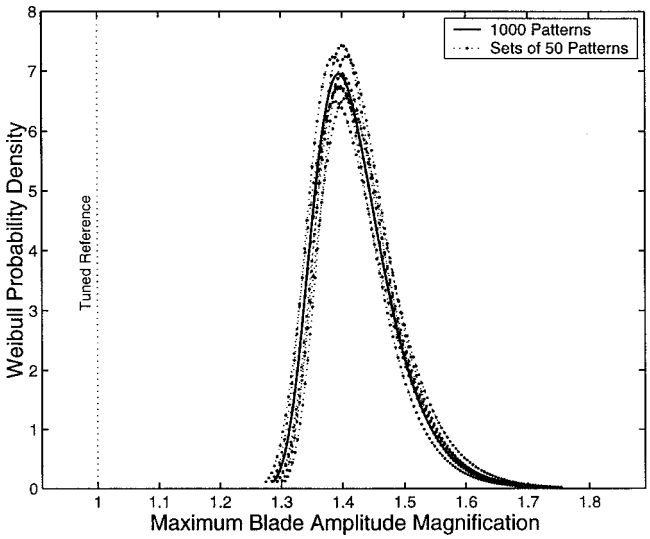



**Fig. 14 Comparison of Weibull estimates of the probability density function for several sets of 50 mistuning patterns each and the full set of 1000 mistuning patterns. The Weibull approximations based on the smaller sets conform well with the probability density function obtained with the full set.**

In Eq. (18), λ is normalized by the peak blade amplitude for the tuned system. The ratio of the peak mistuned amplitude to the peak tuned amplitude is referred to as the amplitude magnification factor.

Using the above approximation for λ, there are two unknowns in Eq. (17), $\delta$ and $\beta$. Now it can be seen that the form of this equation is that of a line with unknown slope and intercept. Therefore, $\delta$ and $\beta$ may be estimated from a least-squares linear regression ([19]).

To determine the correlation between Weibull estimates of response statistics and ''true'' response statistics, the results from the 1000-rotor Monte Carlo simulation were compared to numerous Weibull approximations that utilize only 50 mistuned rotors. This comparison is presented in Fig. 13 for the 5th, 50th, and 95th percentiles of the maximum blade response amplitude. Note how closely the approximated Weibull distributions match the ''true'' distribution. In addition, the quality of the Weibull estimates of the probability density function based on several subsets containing 50 mistuning patterns each is displayed in Fig. 14. Clearly, the probability density is well captured using a significantly smaller number of patterns than the full set of 1000 patterns.

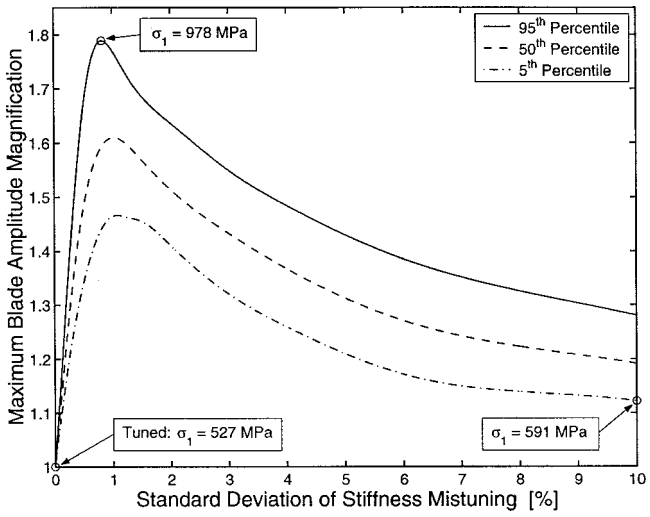



**Fig. 15 Variation in blade amplitude magnification factor with standard deviation of mistuning. Mistuned maximum principal stresses are as much as 86 percent higher than the tuned maximum principal stress.**

Using Weibull distributions to calculate approximate maximum forced response statistics in the described manner, the effect of mistuning strength was investigated for the 5th, 50th, and 95th percentile responses in the frequency veering considered here (see Fig. 3). As shown in Fig. 15, there is a local maximum in the maximum blade response amplitudes for mistuning distributions with approximately one percent standard deviation. The maximum principal stress for one mistuning pattern corresponding to the 95th percentile response amplitude at 0.8 percent mistuning standard deviation is $\sigma_1 = 978$ MPa. This principal stress represents an 86 percent increase over the tuned principal stress reported earlier.

## 6 Effects of Structural Interblade Coupling

As mentioned earlier, two important factors that determine response increases for a given level of mistuning are mode localization and interblade coupling. Using a single-degree-of-freedom per-sector bladed disk model, Wei and Pierre [16] and Ottarsson and Pierre [13] determined that moderately weak interblade coupling is required for significant increases in forced response amplitudes. If there is no interblade coupling, then each blade acts as an individual mistuned oscillator, and the mistuned response does not deviate significantly from the tuned response. As coupling increases, an avenue is created for the blades to communicate vibration energy, which raises the possibility of confining energy to a few blades. The mistuned response may then deviate significantly from the tuned response, until further increases in coupling prohibit the confinement of energy, yielding tuned-like response for large coupling values.

In a structural model of an unshrouded bladed disk, the only means of communication from one blade to the next is through the disk. Therefore, it is reasonable to assume that the interblade coupling is indicated by the amount of interaction between disk and blade-dominated modes. Such modal interaction appears as veering regions in a natural frequency versus nodal diameter plot, where the lines that connect natural frequencies veer away from each other. This relationship between the veerings and the interblade coupling was investigated for a simple model by Kruse and Pierre [5]. From curve veering theory ([14]), it is known that the strength of the disk-blade interaction, and thus the interblade coupling, is a function of the veering curvature. If the interaction between disk-dominated and blade-dominated motion is negligible, then the disk-dominated and the blade-dominated frequency curves will appear to ''pass through'' each other, or cross. Thus, the veering is extremely sharp, with high curvature. In contrast, a lower-curvature veering indicates a higher level of modal interaction, and thus, stronger interblade coupling. This does not imply that the maximum sensitivity to mistuning is always found in the veering region, since the interblade coupling here might fall above the critical range. In this case, a critical amount of interblade coupling may exist at a harmonic that is not adjacent to the veering, since the effective interblade coupling decays with increasing ''distance'' from the veering region.

For the tuned system, cyclic symmetry arguments lead to a defined number of natural modes at discrete interblade phase angles, representing integer harmonics. That is, the modes shapes associated with natural frequencies form standing waves that undergo, over the assembly, integer multiples of $2\pi$ phase shift. However, from these discrete points representing the true modes of the system, it is difficult to obtain reliable estimates for the veering curvatures and mode distances, since the center point of the veerings may be missed.

To enhance the veering representation, one may compute modes based on arbitrary, intermediate interblade phase angles (noninteger harmonics), which allows better approximations of curvatures

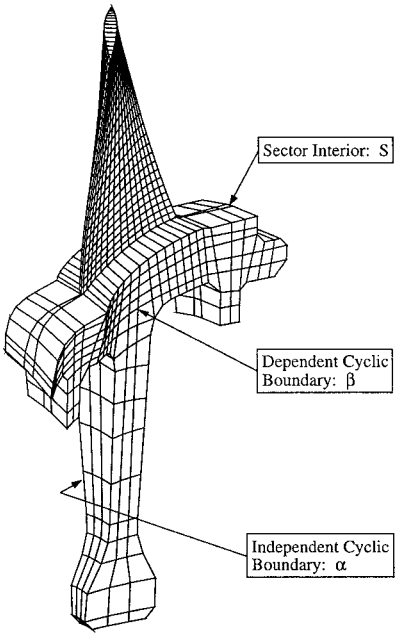



**Fig. 16 Definition of index notation**

and mode distances in the veering regions. Note that this leads to displacement discontinuity at the interface between the $N$th and the first cyclic sector, and therefore, these modes cannot occur in reality. Instead, they correspond to standing waves of a structure comprised of a multiple of $N$ sectors. Hence, the integer harmonics of the $N$-bladed system may be considered to be a subset of the rich dynamic characteristics pertaining to any system with the given disk-blade modal interaction properties. Indeed, when aliased down to the actual $N$-bladed system, the noninteger harmonic eigenfrequencies can also be viewed as being associated with circumferentially propagating waves ([22]). Note that in this interpretation, the aforementioned displacement discontinuity over the $N$th-to-first sector interface is manifested by the time-lag of one wave-revolution.

Following the indices defined in Fig. 16, the eigenvalue problem for a cyclic structure may be formulated in real-valued matrix form as

$$\left[\begin{bmatrix} \tilde{\mathbf{K}}_0^{\mathbf{h}} & \tilde{\mathbf{K}}_1^{\mathbf{h}} \\ \tilde{\mathbf{K}}_1^{\mathbf{h}^{\mathrm{T}}} & \tilde{\mathbf{K}}_0^{\mathbf{h}} \end{bmatrix} - \omega^2 \begin{bmatrix} \tilde{\mathbf{M}}_0^{\mathbf{h}} & \tilde{\mathbf{M}}_1^{\mathbf{h}} \\ \tilde{\mathbf{M}}_1^{\mathbf{h}^{\mathrm{T}}} & \tilde{\mathbf{M}}_0^{\mathbf{h}} \end{bmatrix}\right] \begin{Bmatrix} \tilde{\mathbf{u}}_{\mathbf{h}}^{\mathbf{c}} \\ \tilde{\mathbf{u}}_{\mathbf{h}}^{\mathbf{s}} \end{Bmatrix} = \mathbf{0}, \tag{19}$$

where

$$\tilde{\mathbf{K}}_0^{\mathbf{h}} = \begin{bmatrix} \mathbf{K}_{\alpha\alpha} + (\mathbf{K}_{\alpha\beta} + \mathbf{K}_{\alpha\beta}^{\mathbf{T}})\cos\phi_h + \mathbf{K}_{\beta\beta} & \mathbf{K}_{\alpha\mathrm{S}} + \mathbf{K}_{\mathrm{S}\beta}^{\mathbf{T}}\cos\phi_h \\ \mathbf{K}_{\alpha\mathrm{S}}^{\mathbf{T}} + \mathbf{K}_{\mathrm{S}\beta}\cos\phi_h & \mathbf{K}_{\mathrm{SS}} \end{bmatrix}$$

$$\tilde{\mathbf{K}}_1^{\mathbf{h}} = \begin{bmatrix} (\mathbf{K}_{\alpha\beta} - \mathbf{K}_{\alpha\beta}^{\mathbf{T}})\sin\phi_H & -\mathbf{K}_{\mathrm{S}\beta}^{\mathbf{T}}\sin\phi_h \\ \mathbf{K}_{\mathrm{S}\beta}\sin\phi_h & \mathbf{0} \end{bmatrix},$$

and the mass matrix is partitioned in exactly the same fashion. The trigonometric argument, $\phi_h$, is the interblade phase angle defined in Eq. (12), from which one observes that the interblade phase angle is a function of the ratio $h/N$. Thus, a continuous interblade phase angle mode description may be obtained either by letting $h$ be a ''continuous'' variable (i.e., using small, noninteger steps), or, equivalently, by specifying integer values of $h$ and $N$ that yield the same ratio. The latter option offers a possibility to use commercial finite element software featuring cyclic symmetry to compute these intermediate interblade phase angle modes, as long as the finite element code does not check for geometric consistency. The continuous solid lines in Fig. 17 were obtained using MSC/NASTRAN™ for an assembly of 580 blades (compared to the actual number, 29) to get 20 data points per integer harmonic. Moreover, Fig. 18 provides a detail view of the highly complex region of modal interaction at lower harmonics (i.e., the lower left portion of Fig. 17). From Fig. 18 it is particularly clear that much of the complex interaction between blade and disk-dominated modes will be missed by viewing integer harmonics alone.

Using a continuous interblade phase angle representation with sufficiently small step size, one may invoke, for instance, a finite

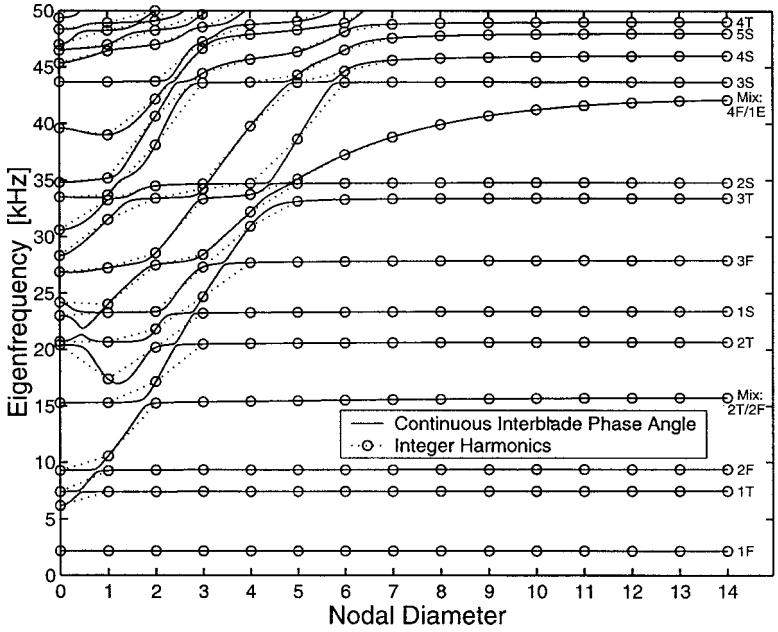



**Fig. 17 Natural frequencies versus number of nodal diameters as a continuous variable**

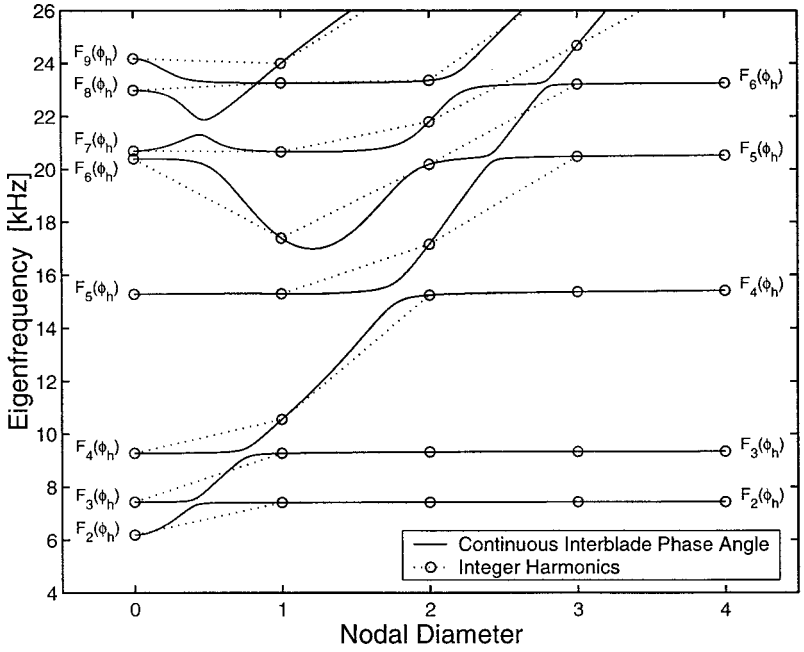



**Fig. 18 Close-up view of intense frequency veering region in Fig. 17**

difference scheme to compute the curvature of the eigenfrequency "function," $F_k(\phi_h)$, in the neighborhood of the veerings as

$$\kappa(\phi_h) = \frac{\dfrac{d^2F_k(\phi_h)}{d\phi_h^2}}{\left[1+\left(\dfrac{dF_k(\phi_h)}{d\phi_h}\right)^2\right]^{3/2}}, \tag{20}$$

for each mode set $k$ (ordered by ascending frequencies as shown in Fig. 18). As an example, consider the veering region at the 2nd harmonic for the 4th and 5th mode sets seen in Fig. 18. Using a fourth-order finite difference scheme, approximations of the first and second derivatives with respect to the interblade phase angle were obtained to compute the veering curvatures depicted in Fig. 19. This plot quantifies the relatively sharp veering between mode sets 4 and 5 occurring close to the 2nd harmonic, where the two mode sets veer away from each other with curvatures of comparable magnitudes but of opposite signs. Also visible are the obtained curvatures of the veering between mode sets 3 and 4 (near harmonic 1), and of the veering between mode sets 5 and 6 (between harmonics 2 and 3).

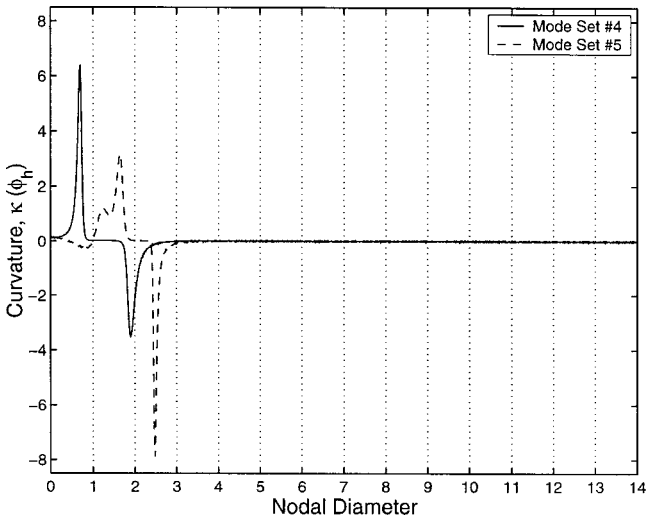



**Fig. 19 Finite difference approximation of curvature for mode sets 4 and 5**

Comprehensive statistical analyses similar to the one presented in Fig. 15 were carried out for several veerings among the low-order mode families. The analyses use Weibull fits based on 50 random mistuning patterns, and the selected results are shown in Fig. 20. Figure 20 shows the locations and maximum curvatures ($\kappa$) of the analyzed veerings. Note that the maximum veering curvatures are averaged between the lower and upper mode sets. In addition, Fig. 20 shows the engine order of the excitation (C), the maximum 99.9th percentile response amplitude magnification factor relative to the tuned response (A), and the standard deviation of the mistuning distribution yielding maximum amplification ($\sigma$) at selected locations.

Several interesting observations can be made in Fig. 20. As shown, the response increases due to mistuning can be very large—close to 200 percent above tuned levels in some cases. Based on the large magnification factors found for excitations near the veerings, it may be deduced that the veering curvatures are such that the corresponding interblade coupling values are near critical levels relatively close to the analyzed veerings. The depicted results also indicate that the amplitude magnification decays significantly with the distance from the veering. This important observation is consistent with the decay of interblade coupling with increasing distance from the veering. Moreover, comparing the 2F and 2T veerings, it is observed that these two veerings have comparable curvatures. However, both maximum amplifications and the mistuning strength for which maximum occurs are vastly different. In fact, based on the displayed results, the blade-dominated mode families with torsion (T) content appear to be significantly more sensitive to mistuning than purely flexural mode families (F and S). This may be explained by significant differences in disk-blade interface motion and how far this motion extends in the disk. Another interesting observation is that the maximum amplifications are found at a much higher level of mistuning strength for both families exhibiting torsion content. Hence, these results indicate that the blade mode type may also be a factor influencing the sensitivity to mistuning.

One focus of future work is to be able to relate mistuning sen-

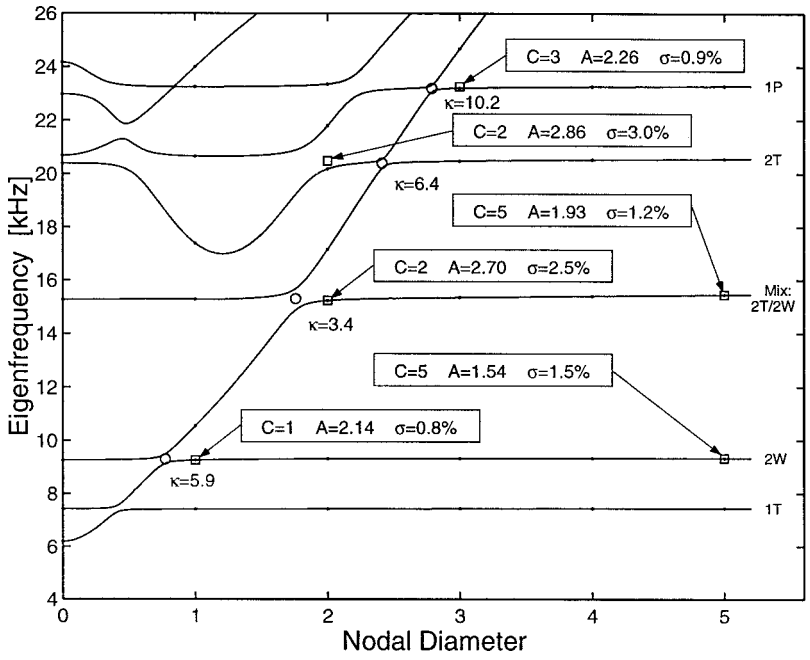



**Fig. 20 Forced response statistical data for selected frequency veerings**

sitivity to veering curvature, distance between eigenfrequencies, and distance of eigenfrequencies from the veering. If such a relationship can be thoroughly understood, this extended eigenvalue analysis may provide an inexpensive design tool for early assessments of mistuning sensitivity.

## 7 Conclusions

This paper demonstrated the feasibility of using a reduced-order modeling technique for comprehensive dynamic analyses of realistic mistuned bladed disks. A reduced-order model (ROM) was employed to investigate mistuned free and forced response characteristics for an industrial turbomachinery rotor. In all investigations, very good correlation was observed between finite element and ROM responses.

A brief statistical study demonstrated the potential of approximating the mistuned response amplitude spectra utilizing Weibull distributions. The computationally inexpensive ROM was used in combination with estimated Weibull distributions to determine statistically the variation in response amplitudes due to random blade mistuning. This combined approach was shown to accelerate significantly the Monte Carlo simulations of the mistuned forced response statistics, while retaining sufficient accuracy.

This investigation verified the existence of the localization phenomenon in industrial turbomachinery rotors, which can cause significant amplitude increases in the forced response of the rotor. In particular, it was shown that mistuned responses can exceed tuned response levels by nearly 200 percent, if appropriate levels of mistuning and interblade coupling are present. In addition, by projecting the ROM results back into full-field finite element displacement coordinates, estimates of the peak stresses were obtained. This ROM-based analysis procedure makes it possible to predict the increase in maximum blade stresses due to mistuning, which is crucial for assessing the safety of a rotor design.

The relationship between blade mistuning for a particular rotor and the associated mistuned forced response is not simple. Prior research efforts have shown that there exists a critical, intermediate level of interblade coupling that leads to maximum mistuned forced response amplitudes. In this study, the analysis of the curvature of eigenfrequency veerings was considered as a means of determining the interblade coupling strength for a bladed disk. A simple method was presented for calculating eigenfrequencies for interblade phase angles corresponding to noninteger harmonics. This calculation may be performed using commercial finite element software. Since this analysis allows one to quantify the veering curvature, it might also allow one to quantify the interblade coupling strength. Using this technique, a brief forced response statistical study indicated that there are four principal factors governing mistuned forced response: (a) modal density, (b) interblade coupling, (c) mistuning strength, and (d) characteristic blade motion. Relating these factors to mistuned forced response amplifications in a quantitative manner will be the subject of future work.

## Acknowledgments

This work is supported by the GUIde Consortium on blade durability at Carnegie Mellon University.

The authors gratefully acknowledge graduate student Mahmoud I. Hussein for obtaining and compiling forced response statistical data for the industrial rotor model used in this study.

## Appendix A

**The Kronecker Product.** The Kronecker product of two matrices is defined as

$$\mathbf{A}\otimes\mathbf{B}=\begin{bmatrix} a_{11}\mathbf{B} & a_{12}\mathbf{B} & \dots & a_{1N}\mathbf{B} \\ a_{21}\mathbf{B} & a_{22}\mathbf{B} & \dots & a_{2N}\mathbf{B} \\ \vdots & \vdots & \ddots & \vdots \\ a_{N1}\mathbf{B} & a_{N2}\mathbf{B} & \dots & a_{NN}\mathbf{B} \end{bmatrix}. \tag{A1}$$

Selected useful properties of the Kronecker product:

$$(\mathbf{A}\otimes\mathbf{B})(\mathbf{C}\otimes\mathbf{D})=(\mathbf{AC})\otimes(\mathbf{BD}) \tag{A2}$$

$$(\mathbf{A}\otimes\mathbf{B})^{-1}=\mathbf{A}^{-1}\otimes\mathbf{B}^{-1} \tag{A3}$$

$$(\mathbf{A}\otimes\mathbf{B})^{\mathbf{T}}=\mathbf{A}^{\mathbf{T}}\otimes\mathbf{B}^{\mathbf{T}}. \tag{A4}$$

## Appendix B

**Circulant Matrices.** The mass and stiffness matrices of any linear cyclic structure may be cast in block-circulant form. Hence, the reduced-order model formulation outlined in this paper makes frequent use of the properties of circulant matrices and their

eigenvectors, as applied to cyclic symmetry problems. The properties of circulant matrices are thoroughly examined in Davis [23]. Moreover, a detailed description of modes of vibration for cyclic structures is contained in the important work by Thomas [24], although certain related mathematical aspects, such as the fundamentals of circulant matrices, appear to have been unrealized at the time.

The general form of a square circulant matrix is

$$\mathbf{C}=\mathbf{circ}(c_1,c_2,\ldots,c_N)=\begin{bmatrix} c_1 & c_2 & \ldots & c_N \\ c_N & c_1 & \ldots & c_{N-1} \\ \vdots & \vdots & \ddots & \vdots \\ c_2 & c_3 & \ldots & c_1 \end{bmatrix}. \quad (B1)$$

All circulant matrices of order $N$ possess $N$ independent eigenvectors. In particular, they share the same set of eigenvectors that make up the complex Fourier matrix, **E**:

$$\mathbf{E}=[e_{ki}];\ \ e_{ki}=\frac{1}{\sqrt{N}}e^{j\alpha(i-1)(k-1)}\ \ k,i=1,\ldots,N \qquad (B2)$$

where $j=\sqrt{-1}$, and $\alpha=2\pi/N$. In addition, there exists an ''almost-equivalent'' real-valued form of Eq. (*B*2):

$$\mathbf{F}=[\mathbf{f_0}\ \ \mathbf{f_1^c}\ \ \mathbf{f_1^s}\ \ \cdots\ \ \mathbf{f_n^c}\ \ \mathbf{f_n^s}\cdots\mathbf{f_{N/2}}]=\begin{bmatrix} \frac{1}{\sqrt{N}} & \sqrt{\frac{2}{N}} & 0 & \cdots & \frac{1}{\sqrt{N}} \\ \frac{1}{\sqrt{N}} & \sqrt{\frac{2}{N}}\cos\alpha & \sqrt{\frac{2}{N}}\sin\alpha & \cdots & -\frac{1}{\sqrt{N}} \\ \frac{1}{\sqrt{N}} & \sqrt{\frac{2}{N}}\cos 2\alpha & \sqrt{\frac{2}{N}}\sin 2\alpha & \cdots & \frac{1}{\sqrt{N}} \\ \vdots & \vdots & \vdots & & \vdots \\ \frac{1}{\sqrt{N}} & \sqrt{\frac{2}{N}}\cos(N-1)\alpha & \sqrt{\frac{2}{N}}\sin(N-1)\alpha & \cdots & \frac{(-1)^{N-1}}{\sqrt{N}} \end{bmatrix} \qquad (B3)$$

where the last column only exists if $N$ is even.

Note that both **E** and **F** are *orthonormal*, or *unitary*, such that $\mathbf{E}^*\mathbf{E}=\mathbf{F^T}\mathbf{F}=\mathbf{I}$, where **I** is an identity matrix of size $N$, and $*$ denotes the Hermitian adjoint (complex conjugate transpose). In addition, this implies that $\mathbf{E^{-1}}=\mathbf{E^*}$ and $\mathbf{F^{-1}}=\mathbf{F^T}$, such that the typical transformation products $\mathbf{E^*CE}$ and $\mathbf{F^TCF}$ are *similarity transformations* (Strang [25]).

The reason behind calling **F** ''almost-equivalent'' to **E** is that the columns of **F** are not true eigenvectors of **C**, and hence, the similarity transformation $\mathbf{F^TCF}$ will not yield a diagonalized matrix. However, it will result in a matrix where all nonzero elements will be grouped into $2\times2$ blocks (''double'' harmonics) on the diagonal, except for the (1,1) and, for $N$ even, the ($N$, $N$) elements (''single'' harmonics). This matrix type is referred to as *pseudo-block-diagonal*.

These properties are readily extended to the case of *block-circulant* matrices by expanding **E** and **F** as $\mathbf{E}\otimes\mathbf{I}$ and $\mathbf{F}\otimes\mathbf{I}$, respectively. The scalar $c_i$ then represents a matrix block $\mathbf{C_i}$, where $\mathbf{C_i}$ and **I** are of the same size. The symbol $\otimes$ denotes the Kronecker product, which is defined in Appendix A.

## References


[1] Ewins, D. J., 1969, ''The Effects of Detuning Upon the Forced Vibrations of Bladed Disks,'' J. Sound Vib., **9**, pp. 65–79.

[2] Ewins, D. J., 1973, ''Vibration Characteristics of Bladed Disc Assemblies,'' J. Mech. Eng. Sci., **15**, pp. 165–186.

[3] Irretier, H., 1983, ''Spectral Analysis of Mistuned Bladed Disk Assemblies by Component Mode Synthesis,'' *Vibrations of Bladed Disk Assemblies*, ASME, New York, pp. 115–125.

[4] Castanier, M. P., Ottarsson, G., and Pierre, C., 1997, ''A Reduced-Order Modeling Technique for Mistuned Bladed Disks,'' ASME J. Vibr. Acoust., **119**, pp. 439–447.

[5] Kruse, M. J., and Pierre, C., 1996, ''Forced Response of Mistuned Bladed Disks Using Reduced-Order Modeling,'' *Proc. 37th AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference*, Vol. 4, AIAA, New York, pp. 1938–1950.

[6] Bladh, R., Castanier, M. P., and Pierre, C., 1999, ''Reduced Order Modeling and Vibration Analysis of Mistuned Bladed Disk Assemblies with Shrouds,'' ASME J. Eng. Gas Turbines Power, **121**, pp. 515–522.

[7] Yang, M.-T., and Griffin, J. H., 1997, ''A Reduced Order Approach for the Vibration of Mistuned Bladed Disk Assemblies,'' ASME J. Eng. Gas Turbines Power, **119**, pp. 161–167.

[8] Bladh, R., Castanier, M. P., and Pierre, C., 2001, ''Component-Mode-Based Reduced Order Modeling Techniques for Mistuned Bladed Disks, Part I: Theoretical Models,'' ASME J. Eng. Gas Turbines Power, **123**, pp. 89–99.

[9] Bladh, R., Castanier, M. P., and Pierre, C., 2001, ''Component-Mode-Based Reduced Order Modeling Techniques for Mistuned Bladed Disks, Part II: Application,'' ASME J. Eng. Gas Turbines Power, **123**, pp. 100–108.

[10] Moyroud, F., Jacquet-Richardet, G., and Fransson, T., 2000, ''A Comparison of Two Finite Element Reduction Techniques for Mistuned Bladed-Disks,'' *Proc. 45th ASME Gas Turbine and Aeroengine Technical Congress, Exposition and Users Symposium*, ASME, New York.

[11] Yang, M.-T., and Griffin, J. H., 1999, ''A Reduced Order Model of Mistuning Using A Subset of Nominal System Modes,'' *Proc. 44th ASME Gas Turbine and Aeroengine Technical Congress, Exposition and Users Symposium*, ASME, New York.

[12] Gumbel, E. J., 1958, *Statistics of Extremes*, Columbia University Press, New York.

[13] Ottarsson, G. S., and Pierre, C., 1995, ''On the Effects of Interblade Coupling on the Statistics of Maximum Forced Response Amplitudes in Mistuned Bladed Disks,'' *Proc. 36th AIAA/ASME/ASCE/AHS Structures, Structural Dynamics, and Materials Conference*, Vol. 5, AIAA, New York, pp. 3070–3078, also, Journal of Sound and Vibration (in print).

[14] Pierre, C., 1988, ''Mode Localization and Eigenvalue Loci Veering Phenomena in Disordered Structures,'' J. Sound Vib., **126**, pp. 485–502.

[15] Mignolet, M. P., and Lin, C.-C., 1997, ''Identification of Structural Parameters in Mistuned Bladed Disks,'' ASME J. Vibr. Acoust., **119**, pp. 428–438.

[16] Wei, S. T., and Pierre, C., 1988, ''Localization Phenomena in Mistuned Assemblies With Cyclic Symmetry, Part I: Free Vibrations; Part II: Forced Vibrations,'' ASME J. Vibr. Acoust., Stress, Reliab. Des., **110**, pp. 429–449.

[17] Castanier, M. P., and Pierre, C., 1997, ''Consideration on the Benefits of International Blade Mistuning for the Forced Response of Turbomachinery Rotors,'' *Proc. ASME International Mechanical Engineering Congress and Exposition*, Vol. 55, ASME, New York, pp. 419–425.

[18] Castanier, M. P., and Pierre, C., 1998, ''Investigation of the Combined Effects of Intentional and Random Mistuning on the Forced Response of Bladed Disks,'' *Proc. 34th AIAA/ASME/SAE/ASEE Joint Propulsion Conference and Exhibit*, AIAA, New York.

[19] Castillo, E., 1988, *Extreme Value Theory in Engineering*, Academic Press, San Diego, CA.

[20] Whitehead, D. S., 1966, ''Effect of Mistuning on the Vibration of Turbomachine Blades Induced by Wakes,'' J. Mech. Eng. Sci., **8**, pp. 15–21.
[21] Whitehead, D. S., 1998, ''The Maximum Factor by Which Forced Vibration of Blades Can Increase Due to Mistuning,'' ASME J. Eng. Gas Turbines Power, **120**, pp. 115–119.
[22] Mead, D. J., 1975, ''Wave Propagation and Natural Modes in Periodic Systems, I: Mono-Coupled Systems,'' J. Sound Vib., **40**, pp. 1–18.
[23] Davis, P. J., 1979, *Circulant Matrices*, John Wiley and Sons, New York.
[24] Thomas, D. L., 1979, ''Dynamics of Rotationally Periodic Structures,'' Int. J. Numer. Methods Eng., **14**, pp. 81–102.
[25] Strang, G., 1988, *Linear Algebra and Its Applications*, 3rd Ed., Saunders, Philadelphia, PA.

# Contact Stresses in Dovetail Attachments: Physical Modeling

**G. B. Sinclair**
Department of Mechanical Engineering,
Louisiana State University,
Baton Rouge, LA 70803-6413

**N. G. Cormier**
General Electric Aircraft Engines,
1 Neumann Way,
Cincinnati, OH 45215

*Simple physical models for the stresses in dovetail attachments are developed. These models address: to slip or not to slip, nominal stresses during loading up, peak contact and shear stresses during loading up, hoop stresses during loading up, peak contact and shear stresses during unloading, and hoop stresses during unloading. Comparisons are made with a previous paper on companion finite element modeling. Generally there is good agreement between the simple physical models and the finite element analysis. Together the two identify a pinching mechanism as leading to large fluctuations in hoop stresses at the edges of contact. These fluctuating hoop stresses can be expected to be a major contributor to the fatigue of dovetail attachments.* [DOI: 10.1115/1.1415740]

## 1 Introduction

**1.1 Motivation and Background.** Single tooth attachments or ''dovetails'' are used to secure fan and compressor blades to disks in gas turbines. A section through a typical dovetail is shown in Fig. 1(*a*). Herein the base of the blade is pulled as a result of the centripetal acceleration of its remainder, while it is restrained by contact with the disk on two flats (e.g., $C-C'$ in Fig. 1(*a*)). At the edges of these contact regions, fretting can occur when loads vary. This fretting can lead to fatigue crack initiation and ultimately to failures (e.g., in the disk at $C'$, the blade at $C$). In order to understand these failures better, an appreciation of the stresses occurring in these critical regions, together with the physical mechanisms causing them, is required. The primary intent of this work is to begin to identify these mechanisms.

This research builds on an earlier stress analysis of a dovetail attachment ([1], hereafter simply referred to as DA1). The attachment treated in DA1 is as in Fig. 1(*a*): Further specifications are given in DA1. Even though the stress analysis in DA1 is limited to two-dimensional elasticity, it is not without challenge. In the first place, challenge stems from resolving the stress gradients present. An indication of these gradients is given in Fig. 2. This figure displays the contact stress $\sigma_c$ (Fig. 1(*c*)), normalized by a nominal stress for the angular motion $\sigma_0$, throughout the contact region (from DA1, for maximum load without friction). In the second, challenge stems from the nonlinearities that attend both expanding contact and contact in the presence of friction.

These challenges, as well as the general complexity of the configuration of Fig. 1, lead to the use of finite element analysis in DA1. By using a submodeling approach, [2], grids that are highly refined in the contact region can be run with reasonable levels of computational effort. The average element size for these grids in the contact region is of the order of one percent of the contact radius $R$ (Fig. 1(*b*)). This is more than an order of magnitude more refined than earlier finite element analysis of dovetail attachments (see DA1 for a review). These refined grids enable convergent and otherwise verified stresses to be obtained. The closeup in Fig. 2 illustrates the convergence of peak $\sigma_c$ for three grids with element sizes that are successively halved.

**1.2 Objective and Scope.** Here we seek to take advantage of the verified stress fields determined in DA1 to improve our understanding of their sources and companion failure modes. We do this by generating a series of fairly simple models for responses at dovetail attachments during loading up and unloading. After each model is developed, comparisons are made between its predictions and the finite element results of DA1.

Specifically the models we develop address the following aspects of dovetail attachment response: whether the blade slips during loading up; nominal stresses during loading up; peak contact and shear stresses during loading up ($\sigma_c$, $\tau_c$ of Fig. 1(*c*)); hoop stresses during loading up ($\sigma_h$ of Fig. 1(*c*)); peak contact and shear stresses during unloading; and hoop stresses during un-

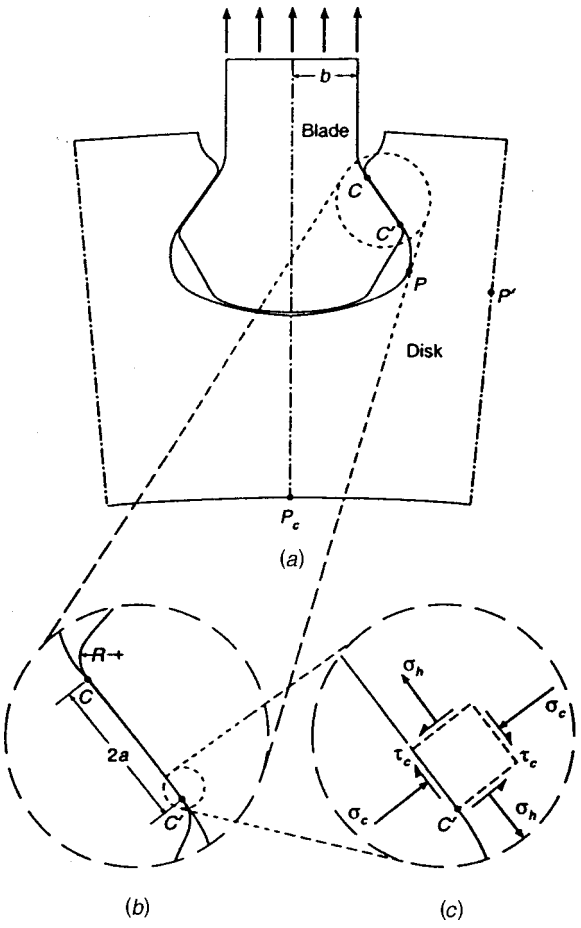



**Fig. 1 Dovetail attachment configuration: (*a*) section of overall attachment, (*b*) closeup of contact region, (*c*) closeup of disk near lower contact point with stresses acting**

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-356. Manuscript received by IGTI Nov. 1999; final revision received by ASME Headquarters Feb. 2000. Associate Editor: D. Wisler.

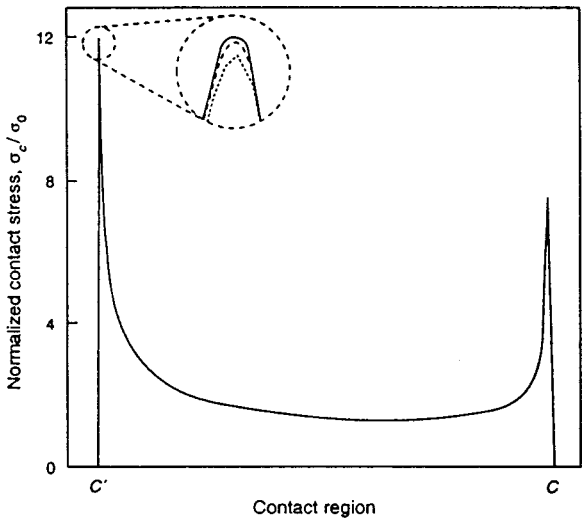



**Fig. 2 Contact stress distribution ($\mu$=0)**

loading. These models are described in Section 2. The paper closes with some concluding remarks in Section 3.

## 2 Models

**2.1 To Slip or Not to Slip.** The question here is whether or not the blade slips relative to the disk during loading up. To begin to answer this question, we consider slip, in the presence of lateral deformation, for a cork pushed into a tapered hole (Fig. 3(*a*)). The applied force on the cork is $F$ and its taper angle is $\alpha$. The hole surface is rough and the material comprising it taken to be relatively rigid to keep analysis simple.

In response to $F$, the resultant normal and tangential forces on each side of the cork are $N$ and $T$, respectively (Fig. 3(*b*)). The fact that the cork is deformable is reflected by the single spring with stiffness $k$ in Fig. 3(*b*).

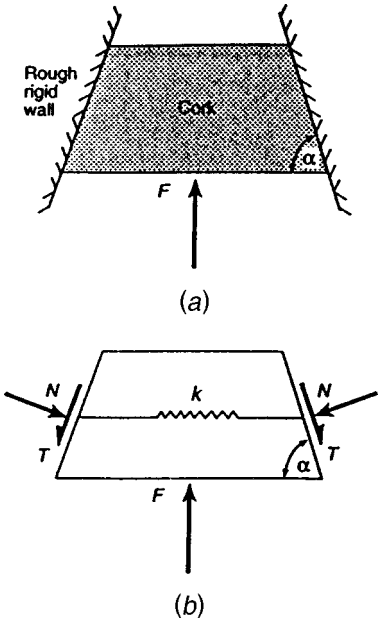



**Fig. 3 Slipping in the presence of lateral deformation: (*a*) cork pressed into two inclined walls, (*b*) free-body diagram for the cork**

For the cork not to slip, no ''gap'' can open up between the cork and the walls.[1] For the configuration of Fig. 3, this means that there can be no compressive force in the spring. Thus

$$N \sin \alpha \leqslant T \cos \alpha. \tag{1}$$

Under Amonton's (or Coulomb's) law of friction,

$$T \leqslant \mu N, \tag{2}$$

where $\mu$ is the coefficient of friction. Together (1) and (2) have

$$\tan \alpha \leqslant \mu, \tag{3}$$

our *condition for no slip*.

The condition in (3) continues to hold for somewhat more complex configurations than that of Fig. 3. If horizontal deformation of the cork is entertained at two vertically spaced stations, (3) holds at each station. If horizontal deformation of the walls is admitted, (1) becomes an equality for no gap to open up, but (3) still holds.

The condition in (3) is, however, a one-dimensional simplification of a two-dimensional configuration. In particular, it ignores the deformation caused by $F$ which would tend to produce interference between the cork and the walls. Hence in reality there would be some lateral compressive force in the cork to remove the interference, in contrast to the development of (3). It follows we can only use (3) to give an indication of when slipping occurs in our dovetail attachment.

To check (3) for the dovetail attachment, we run our finite element analysis with $\mu/\tan\alpha$=0.8, 1.3, where $\alpha$ is now the angle of inclination of the contact flat. We find slipping for the lower value of $\mu$ and sticking throughout most of the contact region for the upper value of $\mu$, both in accord with (3).[2] For our actual dovetail, $\alpha$ is approximately 45 deg so that (3) predicts slip for the highest value of $\mu$ considered ($\mu$=0.4). Our finite element analysis produces slip throughout the contact region during loading up, and thus is in accord with this prediction.

**2.2 Nominal Stresses During Loading Up.** A free-body diagram for half of the blade section of Fig. 1 is given in Fig. 4. The loading of this half blade comes from the centripetal acceleration of its self mass, $F_\omega$, and that of the portion of the blade not shown, $F_b$. These forces are balanced by normal and shear forces, $N$ and $T$, acting on a contact region of extent $2a$ inclined at an angle $\alpha$. By virtue of symmetry, there are no counterbalancing forces on the centerline and only a horizontal force there, $H$. In general, there can be a moment reaction $M$ on the contact region. This is balanced by an equal and opposite moment on the centerline, which also has a further moment $M_F$ to offset the moment produced by $F_\omega$ and $F_b$ about the head of $N$.

Given the slip already noted to occur for our dovetail attachment, equality holds in (2). Then resolving vertically leads to the following expressions for the *nominal contact* and *shear stress*:

$$\bar{\sigma}_c = \frac{F_\omega + F_b}{2a(\cos\alpha + \mu \sin\alpha)}, \quad \bar{\tau}_c = \mu \bar{\sigma}_c. \tag{4}$$

In (4), $\bar{\sigma}_c$ is taken to be positive when compressive (as in Fig. 1(*c*)), and bars atop stresses denote nominal values.

There is also a nominal bending stress in the contact region. The moment $M$ producing this stress is *statically indeterminate*, and consequently this nominal stress cannot be as easily determined. An indication of its sense, though, can be obtained by considering the stiffnesses of the blade section at the upper and

[1] ''Gap'' because no complete gap could occur as a result of compressive contact forces since they would go to zero once contact is completely lost. In reality, adhesive stresses act here and continue to do so with reduced values if the cork starts to contract. Then the accompanying reduced values of $N$, $T$ let $F$ push the cork further into the walls and slip occurs.-

[2] Two-dimensional conforming contact problems cannot stick throughout the entire contact region and must have some slip near the edges of contact if contact inequalities are to be policed (see, e.g., DA1, Section 2.2, for a discussion).

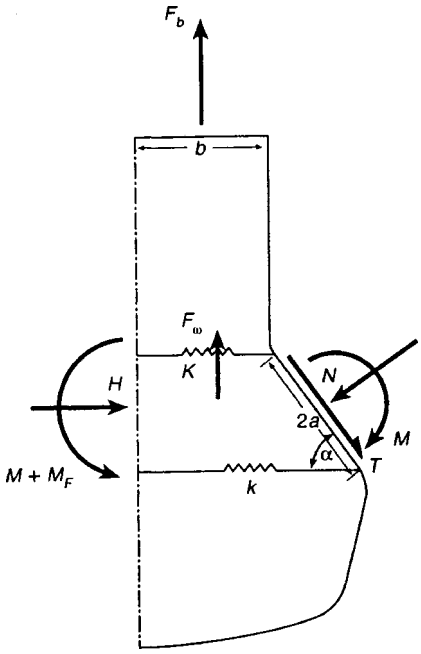



**Fig. 4 Free-body diagram of half of the blade section**

lower edges of contacts, $K$ and $k$, respectively, in Fig. 4. As a result of the section width being smaller for $K$ than $k$, $K>k$. The converse result holds for corresponding stations on the disk (i.e., the disk is stiffer at the lower edge of contact than the upper, see Fig. 1($a$)). Thus when the blade and disk are pressed together, there tends to be a reduction in compressive deflection of the blade at $C$ with a corresponding increase in compressive deflection of the disk at $C$ in order to balance contact forces there, and vice versa at $C'$. Hence we can expect a moment acting in the direction shown in Fig. 3.

To check (4), Fig. 5 shows a plot of finite element results overlaid on its predictions for the case of maximum friction ($\mu$=0.4). Herein, $\omega$ is the angular velocity of the attachment, $\omega_{\max}$ its maximum value, and $\sigma_0$ is the normalizing nominal stress given by

$$\sigma_0=\frac{F_\omega+F_b}{b}, \tag{5}$$

where $b$ is the thickness of the half blade at its end (Fig. 4). Results are almost indistinguishable on the scale of Fig. 5. The same level of agreement is found for $\bar{\sigma}_c$ for $\mu$=0, and $\bar{\sigma}_c$, $\bar{\tau}_c$ for $\mu$=0.2.

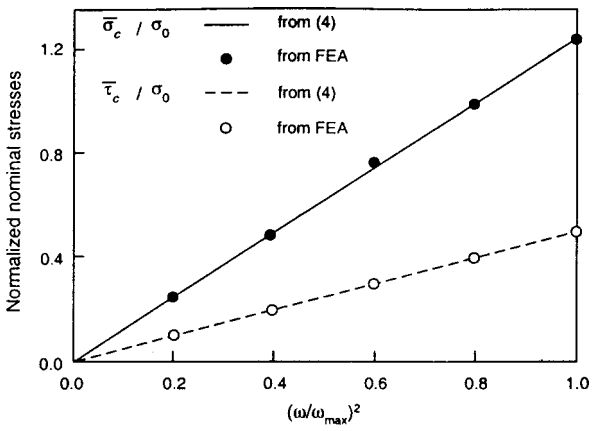



**Fig. 5 Comparison of nominal contact stresses ($\mu$=0.4)**

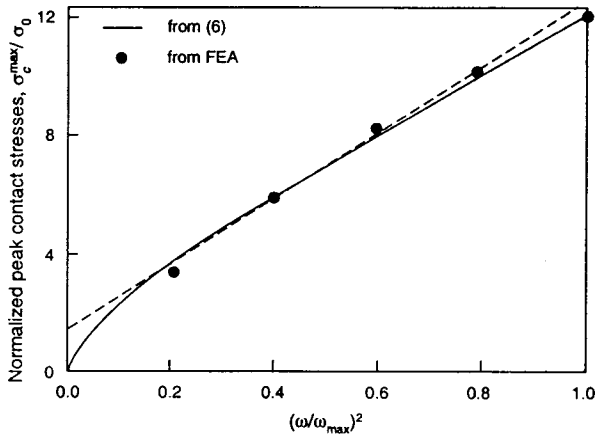



**Fig. 6 Comparison of peak contact stresses ($\mu$=0)**

For the nominal bending stress, the finite element results do agree with the sense shown in Fig. 4 (see Fig. 2). At maximum load, the finite element analysis values of the nominal bending stress range from 36 percent of $\bar{\sigma}_c$ without bending when $\mu$=0, to 58 percent when $\mu$=0.4. These are not the sort of values that one could safely ignore. It follows that any complete nominal stress analysis of a dovetail attachment requires the use of finite elements or a like capability.

**2.3 Peak Contact and Shear Stresses During Loading Up.** For a given angular velocity, the peak contact stress $\sigma_c^{\max}$ occurs within the contact region just inside of $C'$ because of the additional compressive stress there due to bending (cf., Figs. 2, 4). Values of these $\sigma_c^{\max}$ are shown in Fig. 6 as a function of $\omega^2$ for $\mu$=0. For $\omega>0$, a straight line fits the finite element results reasonably well (the dashed line in Fig. 6). However, in contrast to the nominal stresses (Fig. 5), this straight line does not pass through the origin. Clearly $\sigma_c^{\max}$ must be zero for no load, so the straight line fit in Fig. 6 is not appropriate for $\omega$=0.

To explain this phenomenon, we consider the approximate superposition indicated in Fig. 7. Near the edges of contact, as at $C$ of Fig. 1($a$), the two contacting solids have a geometry like that on the top of Fig. 7. The peak contact stress occurring just inside the edge of contact at $C$ on the top of Fig. 7 can be approximated by contributions from a frictionless roller and a lubricated flat punch shown on the bottom of Fig. 7 (ignore the broken lines in this part of the figure for the present). ''Approximated'' because the inden-

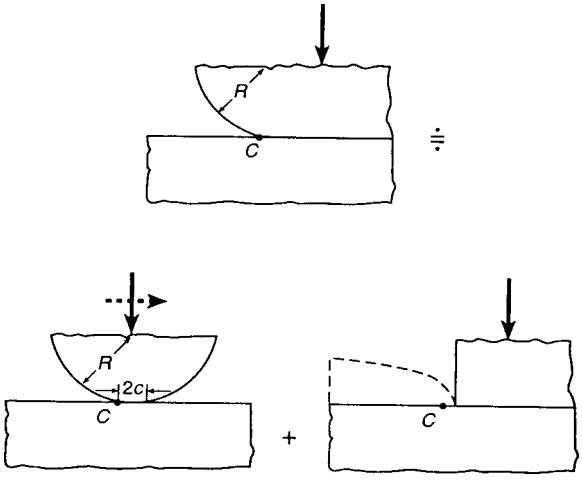



**Fig. 7 Approximate decomposition near the edge of contact**

tor profiles on the bottom in Fig. 7 have a surplus when compared with the indentor on the top. We introduce a further approximation by taking the roller and punch to be rigid. Then we can approximate peak contact stresses in our dovetail attachment using known solutions for the rigid frictionless roller developed by Hartz [3], and for the rigid, flat, lubricated punch derived in Sadowsky [4].

For the cylindrical roller on an elastic half-space, the contact stresses throughout the contact region are proportional to the square root of the applied vertical load because of the expanding contact involved. For the flat punch, the contact stresses are directly proportional to the applied load, with the exception of right at the edge of contact where they are singular. This is because there is no expanding contact. The point *C*, though, is outside the contact region so that the contribution to the peak contact stress from the flat punch in isolation is zero. However, because of the stress singularity present for the flat punch, the displacements outside the flat punch rise with an infinite gradient (indicated with the broken surface profile in Fig. 7). Thus, when the roller acts in concert with the flat punch, it first must flatten this displacement before it gets to indent a flat half-space. The amplitude of this displacement is directly proportional to the applied load, so that the contact stresses needed to flatten it can also be expected to be directly proportional. In combination, therefore, we can expect the peak contact stress from the roller and the punch to have contributions which are proportional to the square root of the applied load as well as contributions which are directly proportional. Hence we expect the *peak contact stresses* in our dovetail attachment to depend upon the angular velocity as in

$$\sigma_c^{\max} = c_1 \omega + c_2 \omega^2. \tag{6}$$

In (6), $c_1$, $c_2$ are constants to be determined by fitting finite element results of DA1.

Applying the fit of (6) leads to the solid line in Fig. 6. This fit must, of course, pass through the origin. In addition, it matches peak contact stresses throughout the angular frequency range to within five percent. Moreover, both the fitted constants are positive and of comparable magnitude, supporting the need for both terms in (6).

Results for $\mu \neq 0$ are all similar to that of Fig. 6. That is, a straight line fits the finite element results well for $\omega > 0$, but does not pass through the origin. The height of the intercept at $\omega = 0$ is less for $\mu \neq 0$ than $\mu = 0$, but nonetheless clearly not zero. Once the fit of (6) is applied, the intercept is removed, and results throughout the $\omega$ range again fitted to within five percent. Shear stress results can be similarly fitted for $\mu \neq 0$.

One consequence of the fits in Figs. 5 and 6 is that there are no stress concentration factors which are independent of $\omega$ over its entire range. For the upper level of $\omega$, though, stress concentration factors which are nearly independent of $\omega$ can be extracted.

At maximum load, we can obtain an approximate analytical solution for the frictionless case for our dovetail attachment from Steuermann [5]. This paper contains a solution for a rigid lubricated punch with a polynomial profile. If $y$ is the depth of penetration of this profile, then

$$\frac{y}{\delta} = 1 - \left(\frac{x}{a}\right)^n \quad (|x| \leqslant a), \tag{7}$$

where $x$ is a horizontal coordinate with origin centered in the contact region of extent $2a$, and $\delta$ is the maximum penetration. We can fit (7) to the profile in our dovetail attachment, at least on one side as in Fig. 7, then compare results.

In making the fit we can select the power in the polynomial profile $n$, as well as its coefficients. We actually choose $n$ so that the contact stress in Steuermann [5] reproduces the stress concentration found for our dovetail attachment once bending is removed. Thus we take $n = 50$ for a common $\sigma_c^{\max}/\overline{\sigma_c} = 4.9$ without bending at maximum load. Then we choose $\delta$ so that the profile in (7) and our dovetail attachment share a common radius of curvature at the edge of contact. No doubt other choices could be made: The preceding is but one fairly reasonable one.

In comparison, the profiles are not the same but they are similar, with a large, central, essentially flat region. Over the interior 90 percent, they are within one percent of each other. The stress distributions, too, are quite similar. Of course, both are zero at the edge of contact and share the same peak value as a result of the choice of $n$. In addition, however, they are comparable at the center (within nine percent), and the location of the peak contact stress is the same (one percent of $a$ inside the edge of contact).

With bending contributions restored in our dovetail attachment, the peak contact stress occurs at $C'$ rather than $C$, as noted earlier. As with the nominal stresses, the bending contribution is significant, ranging from 29 percent of $\sigma_c^{\max}$ without bending when $\mu = 0$, to 41 percent when $\mu = 0.4$.

**2.4 Hoop Stresses During Loading Up.** At the outset in considering the hoop stress $\sigma_h$ (Fig. 1($c$)), we distinguish contributions as being with and without friction. That is, we take

$$\sigma_h = \sigma_{h0} + \sigma_{h\mu}, \tag{8}$$

where $\sigma_{h0}$ is the hoop stress without friction, $\sigma_{h\mu}$ the hoop stress due to friction. We make this distinction because of the different physics producing the two contributions.

To explain further, we consider a roller in contact with a half-space over a strip of extent $2c$ which is small compared to the roller's radius (on the bottom left in Fig. 7). The roller and half-space now are taken to have the same elastic moduli. The roller is both pressed into the half-space and slid sideways (indicated by the dashed arrow in Fig. 7). From Hertz [3], *without friction* on the surface of the half-space,

$$\sigma_{h0} = -\sigma_c = \begin{cases} -\sigma_c^{\max}\sqrt{[1-(x/c)^2]}, & |x| \leqslant c \\ 0, & |x| > c \end{cases}, \tag{9}$$

where $x$ is again a horizontal coordinate with origin $O$ centered in the contact region (Fig. 8($a$)). In (9), we continue to take $\sigma_c$ as positive when compressive, but take $\sigma_h$ as positive in tension (Fig. 1($c$)). Hence without friction, hoop stresses in the surface of the half-space are nowhere tensile. *With friction and sliding*,

$$\tau_c = \mu \sigma_c. \tag{10}$$

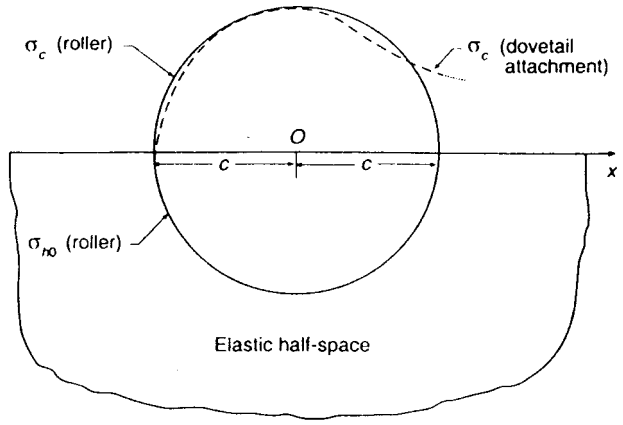
Then, from Poritsky [6], the hoop stress due to $\tau_c$ alone is given by

$$\sigma_h^{\tau} = -2\tau_c^{\max} \begin{cases} x/c, & |x| \leqslant c \\ \{(x/c) - \mathrm{sgn}(x)\sqrt{[(x/c)^2 - 1]}\}, & |x| \geqslant c \end{cases}, \tag{11}$$

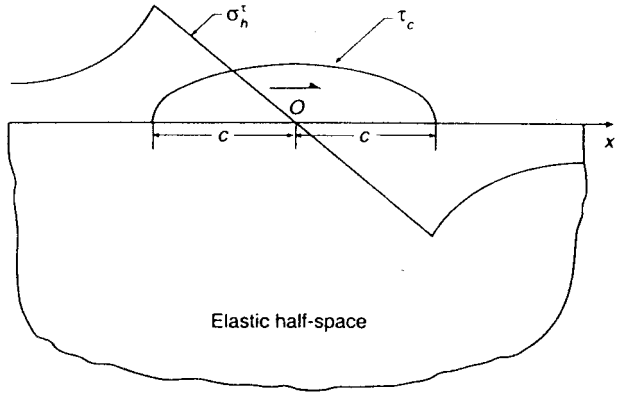
where $\tau_c^{\max}$ is the peak value of $\tau_c$ and sgn is the signum function. Hence with friction, the hoop stresses ahead of the shear ($x > 0$) are compressive, those behind it ($x < 0$) are tensile (Fig. 8($b$)). It follows that, as far as local, contact, hoop stresses are concerned, tensile contributions only occur when friction is present.

How does this example relate to our dovetail attachment? Figure 8($a$) includes the contact stress distribution at $C'$ from DA1 (shown as a dashed line). Near the edge of contact, this matches the Hertzian distribution of (9) well. In addition in this vicinity, (10) holds (see Fig. 6 in DA1). Accordingly we can expect the results for the roller to indicate the local hoop stresses in our dovetail attachment.

From henceforth we concentrate on the production of *tensile hoop stresses* because we expect these stresses to contribute most to fatigue failures of dovetail attachments. We justify this expectation subsequently in Section 2.6. For the local stresses attending contact, tensile hoop stresses stem from the effects of friction and $\tau_c$. In addition, there can be tensile hoop stresses induced by stress resultants acting on the individual blade and disk geom-




(*a*)




(*b*)

**Fig. 8 Surface stresses for sliding roller: (*a*) contact and hoop stresses without friction, (*b*) shear and hoop stresses due to friction**

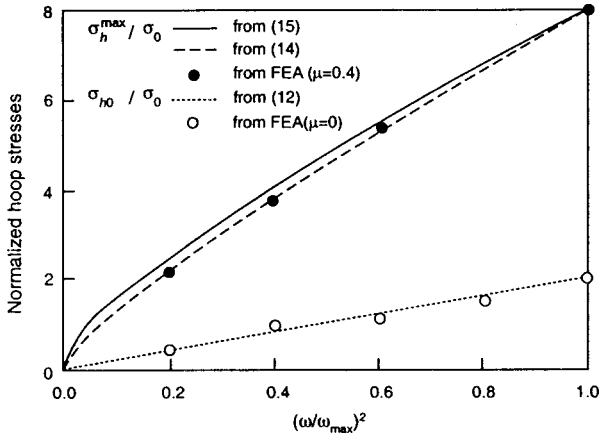



**Fig. 9 Comparison of hoop stresses**

etries of our dovetail attachment. We look for locations where these two sources in combination maximize the magnitudes of tensile hoop stresses.

For the blade, $\sigma_{h0}$ has no tensile contribution from contact alone (see (9)). However, $F_b$ exerts a tension in the upper blade section of Fig. 4 which can be expected to induce a tensile $\sigma_{h0}$ of the order of $\sigma_0$ at $C$. This force may, as well, produce some tensile $\sigma_{h0}$ at $C'$, but, even if tensile, the magnitude of this stress can be expected to be far lower because the normal resultant $N$ has now counterbalanced $F_b$. With friction, the tangential force $T$ produces tensile $\sigma_{h\mu}$ at $C$, compressive at $C'$. Thus for the blade the maximum tensile hoop stress occurs at the edge of contact at $C$.

For the disk, again $\sigma_{h0}$ has no tensile contribution from contact alone. However, the counterparts of $N$ and $M$ of Fig. 4 for the disk tend to open the disk notch region near $P$ in Fig. 1(*a*), and thus induce a tensile $\sigma_{h0}$ of the order of $\overline{\sigma_c}$ at $C'$. The resultant of $N$ and $M$ may also induce some tensile $\sigma_{h0}$ at $C$ in the disk, but this can be expected to be less in magnitude for the disk geometry by virtue of being further from the notch root at $P$. With friction, the tangential force on the disk has the same magnitude as the $T$ shown in Fig. 4, but acts in the opposite direction. Hence, for the disk, it produces compressive stress at $C$, tensile at $C'$. Thus for the disk the maximum tensile hoop stress occurs at the edge of contact at $C'$. Moreover, because $\sigma_c$ is greater at $C'$ than $C$ due to the bending contributions, this location is the one with the maximum tensile $\sigma_{h\mu}$. We therefore concentrate on $C'$ in the disk in what follows, though very similar physics can be expected to occur at $C$ in the blade.

For just outside contact at $C'$ in the disk, $\sigma_{h0}$ is due to $N$ and $M$ alone ($T$ being zero when $\mu=0$). Thus $\sigma_{h0}$ is statically indeterminate. We can, though, use our frictionless finite element analysis to estimate $\sigma_{h0}$.

In Section 2.2, we have that $N$ is proportional to $F_\omega$ and $F_b$, hence $\omega^2$. This follows simply from statics. We do not have the same result for $M$ because it is statically indeterminate. The finite element analysis of DA1, though, shows that $M$ too is proportional to $\omega^2$. Consequently we expect

$$\sigma_{h0}=c_3\omega^2, \tag{12}$$

where $c_3$ is a constant to be fitted to our finite element analysis for $\mu=0$.

For just outside contact at $C'$ in the disk, from (10), (11), we can expect

$$\sigma_{h\mu}^{\max}=c_4\mu\sigma_c^{\max}, \tag{13}$$

where $c_4$ is a further constant for fitting. In (10), (11), $c_4=2$ when $x=-c$. Because our dovetail attachment is not an identical configuration to the roller and half-space that lead to (10), (11), we admit the possibility of a different value, although we expect something comparable to 2 given the similarity demonstrated in Fig. 8(*a*). From (6), (13) implies $\sigma_{h\mu}^{\max}$ depends upon both $\omega$ and $\omega^2$. Then from (12) and (8), we expect the *peak, tensile, hoop stress* to vary with $\omega$ as in

$$\sigma_h^{\max}=c_5\omega+c_6\omega^2, \tag{14}$$

where $c_5$ and $c_6$ are two more constants to be fitted.

As a further model for $\sigma_h^{\max}$, we can also combine finite element results for $\sigma_{h0}$ with $\sigma_{h\mu}^{\max}$ of (13). In doing this we are assuming that the hoop stress induced by stress resultants does not change much when friction is introduced. This is quite possibly so since the vertical resultant on the contact flat remains the same as without friction for a given $\omega$. Accordingly we have the following approximate solution for the *peak, tensile, hoop stress*:

$$\sigma_h^{\max}=\sigma_{h0}(\mathrm{FEA})+c_4\mu\sigma_c^{\max}. \tag{15}$$

We fit (12), (14), and (15) to finite element results of DA1.

Figure 9 shows a comparison of $\sigma_{h0}$ of (12) with finite element hoop stresses for varying $\omega$. Agreement is good. In addition, Fig. 9 shows a comparison of $\sigma_h^{\max}$ of (14) and (15) with finite element results when $\mu=0.4$. Agreement is fair. The value of $c_4$ obtained by the fit is 1.6, and therefore quite comparable to the 2 of ana-

**Table 1 Normalized stresses during unloading ($\mu$=0.4)**

| $(\omega/\omega_{\max})^2$ | $\sigma_c^{\max}/\sigma_0$ | $\tau_c^{\max}/\sigma_0$ |
|---|---|---|
| 1.0 | 9.0 | 3.5 |
| 0.8 | 10.2 | 1.1 |
| 0.6 | 11.3 | 0.5 |
| 0.4 | 11.2 | 0.0 |
| 0.2 | 6.7 | 0.0 |
| 0.0 | 0.0 | 0.0 |

lytical solution for the roller. Also apparent in Fig. 9 is the large contribution to tensile hoop stresses caused by friction when $\mu$=0.4.

**2.5 Peak Contact and Shear Stresses During Unloading.** *Without friction, N* and $\sigma_c^{\max}$ during unloading take on the same values as during loading up for a corresponding $\omega$ (to within 0.5 percent, DA1, Section 4.2). *With friction, N* and $\sigma_c^{\max}$ during unloading *increase*. This result is contrary to any one-dimensional physical reasoning, although it can be explained with two-dimensional reasoning.

Consider what happens at the section through the disk at $PP'$ in Fig. 1(*a*). With reduced $\omega$, $F_b$, and $F_\omega$ of Fig. 4 are reduced. Therefore so is the radial pull on the disk at $PP'$. The material above $PP'$ in the disk then retracts radially inward. Because the periodic symmetry line through $P'$ is not parallel to the central symmetry line through $P_c$ (Fig. 1(*a*)), this retraction is accompanied by a tendency for the disk contact region to move laterally towards the centerline. Without friction, the blade can slip inwards and thereby accommodate lateral motion of the disk contact region. With friction, the blade can stick and get pinched by lateral motion of the disk contact region. It is this *pinching mechanism* that increases $N$ and $\sigma_c^{\max}$ during unloading.

To explain further how sticking during unloading when friction is present is consistent with an increase in $N$, we consider what happens to the tangential resultant under these circumstances. Thus we reconsider Fig. 4 with $F_b \rightarrow F_b - \Delta F_b$, $F_\omega \rightarrow F_\omega - \Delta F_\omega$, and $N \rightarrow N + \Delta N$. Here $\Delta F_b$ and $\Delta F_\omega$ are positive because we are unloading, and $\Delta N$ is positive under the assumption of an increase in $N$. Now resolving vertically gives

$$F_b - \Delta F_b + F_\omega - \Delta F_\omega - (N + \Delta N)\cos\alpha - (T + \Delta T)\sin\alpha = 0, \tag{16}$$

where $\Delta T$ is the change in $T$ accompanying unloading. The $F_b$, $F_\omega$, $N$, and $T$ terms balance from resolving vertically when $\omega = \omega_{\max}$. Thus we have

$$\Delta T = -(\Delta F_b + \Delta F_\omega)\csc\alpha - \Delta N \cot\alpha. \tag{17}$$

Hence $\Delta T$ is negative and $T$ is reduced on unloading. Because $T = \mu N$ at maximum load, it follows that

$$T + \Delta T < \mu(N + \Delta N) \tag{18}$$

on unloading, since the left-hand side is smaller than $T$ while the right is larger than $\mu N$. The inequality in (18) is the condition for sticking in Amonton's law. Accordingly an increase in $N$ on unloading is consistent with sticking and pinching.

What (17) also shows is that, with pinching, there are two sources of reduction in the tangential resultant. The first is the expected reduction occurring because there is less load to be balanced—the $F$ terms in (17). The second is the reduction occurring because the normal reaction has increased and consequently is balancing a greater share of the load—the $N$ term in (17).

Table 1 gives peak contact and shear stresses during unloading for the case of maximum friction (from DA1 for $\mu$=0.4). The values in Table 1 demonstrate the significant increases in $\sigma_c^{\max}$ that accompany the increases in $N$ due to pinching. These increases occur up until the blade starts to slip inwards as a whole when about 60 percent unloaded ($(\omega/\omega_{\max})^2$=0.4).

Table 1 also demonstrates the significant reductions in $\tau_c^{\max}$ that are due to pinching during unloading. Only the peak shear stress acting in the same direction as that at maximum load is given in Table 1.

As noted earlier, conforming contact problems cannot completely stick throughout the contact region. Hence, even with pinching and sticking at 20 percent unloaded, there is some local slipping. Near $C'$, this slipping occurs in the extension of the contact region that accompanies the increased normal contact stress due to unloading. It is possible to obtain an indication of the direction of this slipping and thus the shear stress in this region by drawing on solutions in the literature. This can be done by approximating the local geometry near $C'$ (Fig. 1(*a*)) with a half-space for the disk together with a quarter-space for the blade. Then we can use solutions in Michell [7] for a uniform strip of pressure on a half-space and in Levy [8] for a uniform pressure on a quarter-space. These solutions indicate that the surface near $C'$ wants to move outwards on the disk (i.e., towards $C$, Fig. 1(*a*)) in response to the increase in contact, while that on the blade wants to move inwards (i.e., away from $C$). This is the opposite direction to relative motion between the disk and the blade during loading up. Consequently during unloading we have slip and *shear reversal* near the edges of contact. The finite element results of DA1 confirm these phenomena.

Eventually during unloading, the blade starts to slip throughout the entire contact region with the shear being completely reversed. This complete reversal is reflected by the zeros in Table 1 for the shear in the original direction. It occurs here for $\mu$=0.4 when the dovetail attachment is about 60 percent unloaded. The magnitudes of these reversed shears comply with (10) (to within 0.5 percent, DA1).

**2.6 Hoop Stresses During Unloading.** *Without friction,* $\sigma_h$, including $\sigma_h^{\max}$, during unloading takes on the same values as during loading up for a corresponding $\omega$ (to within 0.5 percent, DA1, Section 4.2). *With friction,* $\sigma_h^{\max}$ during unloading can be expected to *decrease*, and possibly by a *significantly greater* amount than the *percentage unload*. There are two reasons for this expectation. First, peak shear stresses due to friction drop appreciably during unloading and consequently contribute less to the tensile $\sigma_h^{\max}$. Second, normal contact stresses increase during unloading and consequently contribute compressive hoop stresses to be subtracted off $\sigma_h^{\max}$.

With respect to the first reason, we note that frictional shear stresses are the major source of tensile hoop stress when $\mu$=0.4 (75 percent at maximum load, Fig. 9). Moreover, in response to a 20 percent unload, $\tau_c^{\max}$ decreases by 68 percent of its value at maximum load, and for a 40 percent unload by 86 percent (Table 1). Hence the greater part of $\sigma_h^{\max}$ due to frictional shear stress can be expected to be reduced considerably on unloading.

With respect to the second reason, we note that the expanding contact attending the increase in $N$ with unloading moves points of peak and near peak hoop stress just outside of contact at maximum load to within the contact region. This subtracts from $\sigma_h^{\max}$ a compressive hoop stress as in (9) (Fig. 8(*a*)). It can also subtract a compressive hoop stress due to shear reversal near the edge of the expanded contact region.

To confirm these expectations we examine finite element results for the hoop stress in the disk near $C'$ (for $\mu$=0.4 from DA1). Figure 10 shows the hoop stress distributions near $C'$ at maximum load and when 20 percent unloaded. The fluctuation in $\sigma_h$ at the edge of contact is dramatic. In fact, if $\Delta\sigma_h$ is the magnitude of this fluctuation, then for Fig. 10,

$$\Delta\sigma_h = 67\Delta\sigma_0, \quad \Delta\sigma_0 = \sigma_0/5. \tag{19}$$

Also indicated in Fig. 10 is the increase in contact region extent which leads to compressive stresses at the original edge of contact on unloading. It is this reversal of sign for the peak tensile hoop

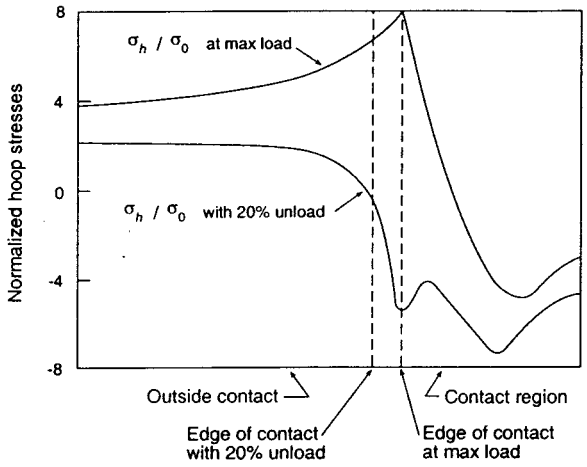



**Fig. 10 Hoop stress variation near the edge of contact for 20 percent unloading ($\mu$=0.4)**

stress which was anticipated in Section 2.4 when we chose to focus on tensile $\sigma_h$ during loading up: No such reversal occurs for compressive $\sigma_h$ at maximum load.

In addition to the two reasons already discussed for the amplification of $\Delta\sigma_0$ by $\Delta\sigma_h$ in (19), there is the fact that $\sigma_h^{\max}$ is appreciably greater than $\sigma_0$ so that there is a lot more hoop stress to fluctuate. This is because $\sigma_h^{\max}$ is concentrated and thus greater than $\overline{\sigma_c}$, and $\overline{\sigma_c}$ is somewhat greater than $\sigma_0$ (Fig. 5).

## 3 Conclusions

Using simple physical reasoning together with solutions in the literature and some elementary analysis, the physical response of dovetail attachments during loading and unloading is inferred. These physical predictions are confirmed by finite element results from DA1 (typically to within five percent).

The simple models confirmed by finite element analysis build upon one another. They identify contact shears due to friction as a major source of peak tensile hoop stresses at the edges of contact during loading up (at $C$ in the blade, $C'$ in the disk; Fig. 1($a$)). Thereafter they identify a pinching mechanism that can occur with sufficient friction during unloading. This mechanism explains the counter-intuitive response of an increase in normal contact stress during unloading. It also explains the significant fluctuations that can occur in the hoop stresses at the edges of contact. These fluctuations can be an order of magnitude greater than oscillations in the loading due to rotation. Such oscillations can accompany mission cycles or be due to variations in stiffwise bending. Consequently the fluctuations in hoop stresses are seen as being likely to be a major contributor to either the low-cycle or the high-cycle fatigue of dovetail attachments.

## Acknowledgments

We are pleased to acknowledge the benefit of technical discussions with J. H. Griffin of Carnegie Mellon and with engineers at General Electric Incorporated. We are also grateful for the financial support of this work by General Electric Aircraft Engines.

## References

[1] Sinclair, G. B., Cormier, N. G., Griffin, J. H., and Meda, G., 2002, "Contact Stresses in Dovetail Attachments: Finite Element Modeling," J. Eng. Gas Turbines Power, **124**, 182–191.
[2] Cormier, N. G., Smallwood, B. S., Sinclair, G. B., and Meda, G., 1999, "Aggressive Submodeling of Stress Concentrations," Int. J. Numer. Methods Eng., **46**, pp. 889–909.
[3] Hertz, H., 1882, "On the Contact of Elastic Solids," J. Reine Agnew. Mat., **92**, pp. 156–171 (in German: for an account in English, see Johnson, 1985, Ch. 4.2). Johnson, K. L., 1985, *Contact Mechanics*, Cambridge University Press, Cambridge, U.K.
[4] Sadowsky, M. A., 1928, "Two-Dimensional Theory of Elasticity Theory," Z. Angew. Mat. Mech., **8**, pp. 107–121 (in German: for an account in English, see Johnson, 1985, Ch. 2.8).
[5] Steuermann, E., 1939, "To Hertz's Theory of Local Deformations in Compressed Elastic Bodies," C. R. (Dokl.) Acad. Sci. URSS, **25**, pp. 359–361.
[6] Poritsky, H., 1950, "Stresses and Deflections of Cylindrical Bodies in Contact With Applications to Contact of Gears and of Locomotive Wheels," ASME J. Appl. Mech., **17**, pp. 191–201.
[7] Michell, J. H., 1902, "The Inversion of Plane Stress," Proc. London Math. Soc., **34**, pp. 134–142.
[8] Levy, M., 1898, "On the Elastic Equilibrium of a Masonary Dam of Triangular Cross Section," C. R. de l'Académie des Sciences, **127**, pp. 10–15.

# Application of Model Order Reduction to Compressor Aeroelastic Models

**K. Willcox**

**J. Peraire**

**J. D. Paduano**

Gas Turbine Laboratory,
Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
Cambridge, MA 02139

*A model order reduction technique that yields low-order models of blade row unsteady aerodynamics is introduced. The technique is applied to linearized unsteady Euler CFD solutions in such a way that the resulting blade row models can be linked to their surroundings through their boundary conditions. The technique is applied to a transonic compressor aeroelastic analysis, in which the high-fidelity CFD forced-response results are better captured than with models that use single-frequency influence coefficients. A low-speed compressor stage is also modeled to demonstrate the multistage capability of the method. These examples demonstrate how model order reduction can be used to systematically improve the versatility, fidelity, and range of applicability of the low-order aerodynamic models typically used for incorporation of CFD results into aeroelastic analyses.* [DOI: 10.1115/1.1416152]

## Introduction

One of the foremost challenges in the study of free and forced blade vibrations in turbomachinery is encapsulation of aerodynamic effects in a form suitable for coupled aerostructural analysis. It is generally agreed that in modern aeroengines the unsteady aerodynamic effects governing the severity of vibrations are complex; at least some of the flow details (shock motion, blade loading, viscosity, boundary conditions, etc.) must be modeled to obtain realistic analyses. Because very little data exist to isolate the most important of these details, the current state of the art utilizes computational fluid dynamic (CFD) analysis to capture as much of the physics as possible. The CFD results are encapsulated as unsteady aerodynamic forces resulting from (usually prescribed sinusoidal) blade vibrations, which can then be coupled to the structural dynamics. Forced vibration analyses are often done in a similar two-step fashion, in which the aerodynamic effects are computed first and then incorporated into a coupled analysis ([1]). Nonreflecting boundary conditions are commonly used in these analyses for simplicity, although the boundary conditions can have a significant effect on the results ([2]).

Pre-computation of aerodynamic effects, using canonical motions and forcing functions, invokes various assumptions based on the problem at hand. For stability analysis of a rotor, the fundamental assumption is that the blade vibration frequencies change by only a small amount due to aerodynamic coupling. Thus analysis of the aerodynamic loads using the most important structural dynamic eigenfunctions is justified. Forced vibration analyses often compute gust loads assuming sinusoidal gusts and fixed blades, assuming that the incoming gusts are not modified by unsteady blade motion ([3]). Mistuning analyses commonly use simplified aerodynamic models (as in Crawley and Hall [4]), although incorporation of CFD-based influence coefficients (based on sinusoidal, single-frequency runs) has been reported ([5]).

Experience and order-of-magnitude analyses (mass ratio arguments, etc.) suggest that the engineering approximations discussed above are often valid, allowing trends and insight to be more easily obtained. However, the framework for encapsulation of aerodynamic effects is limited, and in some cases inconsistent with the goals of the analysis. Nonsinusoidal forcing, forced response with mistuning, and other analyses that involve combinations of effects stretch the assumptions of standard methods. In addition, the effects of upstream and downstream boundary conditions (nearby blade rows, inlet duct acoustics) are increasingly coming under scrutiny.

It is in these cases that a reduced-order model with generalized boundary conditions can play an important role. The CFD code is viewed as an input-output model: blade motions, incoming flow perturbations, and mean flow properties are viewed as inputs; outputs include outgoing flow perturbations and the blade forces and moments. A reduced-order model is a low-order model that duplicates the behavior of the CFD analysis over a limited range of conditions. The range of validity of the reduced-order model is determined by the set of CFD experiments performed during the model order reduction procedure. By making the model building procedure systematic, a single tool can yield reduced-order models for a wide range of applications.

In the sense defined above, models based on spatially and temporally sinusoidal motion at a single frequency (we will call these ''assumed-frequency'' methods) are simple reduced-order models. They are limited to the boundary conditions chosen in the CFD analysis and are only valid at the precise mean flow conditions analyzed. They are limited to a small range of frequencies about the assumed frequency, and this range is unknown. Owing to their limited range of validity, assumed-frequency models are the most compact: they are coupled to the structural model as constant coefficients that are independent of changes in the mean flow, boundary conditions, frequency, etc. More detailed aerodynamic models, motivated by external aeroelastic (wing flutter) methods, have been suggested ([6]), but do not appear to be in broad use.

In this paper we present a systematic method for model order reduction from CFD, which can be applied to a broad range of problems. The method allows the analyst to specify the range of interblade phase angles, reduced frequencies, forcing functions, and boundary conditions that are appropriate to encapsulate into the problem at hand. This range is spanned by a series of ''canonical'' runs of the CFD code, which are then used to derive a low-order model. The model is in ordinary differential equation form, requiring approximately 100 to 200 states per blade row (in our examples), suitable for incorporation in a fully coupled aeroelastic analysis. Examples of application of this model are given for illustrative purposes.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-377. Manuscript received by IGTI November 1999; final revision received by ASME Headquarters February 2000. Associate Editor: D. Wisler.

## Computational Model

The starting point for the procedure is a computational model. In this paper, the nonlinear two-dimensional Euler equations, expressed on an unstructured triangular grid, are used:

$$\frac{d\mathbf{U}}{dt}+\mathbf{R}(\mathbf{U},\mathbf{U}_b,\mathbf{x})=0, \tag{1}$$

where $\mathbf{R}$ represents the nonlinear flux contributions at each node, which are functions of the problem geometry $\mathbf{x}$, the unknown flow quantities on the grid points $\mathbf{U}$ and the boundary values $\mathbf{U}_b$. For a two-dimensional bladed disk with $r$ blades with rigid cross sections, $\mathbf{x}$ can be written in terms of plunge displacement $h_i$ and pitch displacement $\alpha_i$:

$$\mathbf{x}=\mathbf{x}(\mathbf{q}), \quad \mathbf{q}=[\mathbf{q}_1^T,\mathbf{q}_2^t,\ldots,\mathbf{q}_r^T]^T, \tag{2}$$

where $\mathbf{q}_i=[h_i\ \alpha_i]^T$. The vector $\mathbf{U}_b$ allows external flow disturbances and impedance properties to be modeled. These could be, for example, time-varying pressure or velocity distortions due to a neighboring blade row, inhomogeneity in the incoming flow, or acoustic boundary conditions. Given blade motion $\mathbf{q}$ and boundary disturbance $\mathbf{d}$, we write $\mathbf{U}_b$ as

$$\mathbf{U}_b=\mathbf{U}_p(\mathbf{q},\dot{\mathbf{q}},\mathbf{d},\mathbf{x}), \tag{3}$$

where $\mathbf{U}_p$ is a vector containing the appropriate prescribed quantities.

Linearization of the Euler equations, assuming all motions and disturbances are small, results in the following equation:

$$\frac{d\mathbf{U}'}{dt}+\frac{\partial\mathbf{R}}{\partial\mathbf{U}}\mathbf{U}'+\frac{\partial\mathbf{R}}{\partial\mathbf{U}_b}\mathbf{U}_b'+\frac{\partial\mathbf{R}}{\partial\mathbf{x}}\mathbf{x}'=0, \tag{4}$$

where prime denotes perturbation quantities, and all partial derivatives are computed at the steady state. Due to the linearity assumption, it is not necessary to actually deform the grid; the final term on the left-hand side of Eq. (4) accounts for the first-order effects of grid motion. After linearizing the boundary terms

$$\mathbf{U}_b'=\frac{\partial\mathbf{U}_p}{\partial\mathbf{q}}\mathbf{q}+\frac{\partial\mathbf{U}_p}{\partial\dot{\mathbf{q}}}\dot{\mathbf{q}}+\frac{\partial\mathbf{U}_p}{\partial\mathbf{d}}\mathbf{d}+\frac{\partial\mathbf{U}_p}{\partial\mathbf{x}}\mathbf{x}', \tag{5}$$

and writing the grid displacement as a linearized function of blade displacement $\mathbf{x}'=T\mathbf{q}$, where the transformation matrix $T$ is computed from the mesh geometry, we obtain a high-order state space system for the perturbation quantities

$$\frac{d\mathbf{U}'}{dt}+\frac{\partial\mathbf{R}}{\partial\mathbf{U}}\mathbf{U}'=\left[-\frac{\partial\mathbf{R}}{\partial\mathbf{x}}T-\frac{\partial\mathbf{R}}{\partial\mathbf{U}_b}\left(\frac{\partial\mathbf{U}_p}{\partial\mathbf{q}}+\frac{\partial\mathbf{U}_p}{\partial\mathbf{x}}T\right)\right]\mathbf{q}-\frac{\partial\mathbf{R}}{\partial\mathbf{U}_b}\frac{\partial\mathbf{U}_p}{\partial\dot{\mathbf{q}}}\dot{\mathbf{q}}-\frac{\partial\mathbf{R}}{\partial\mathbf{U}_b}\frac{\partial\mathbf{U}_p}{\partial\mathbf{d}}\mathbf{d}, \tag{6}$$

which can be written equivalently as

$$\frac{d\mathbf{U}'}{dt}=A\mathbf{U}'+B\mathbf{u}+E\mathbf{d}. \tag{7}$$

Here $\mathbf{u}=[\mathbf{q}\ \dot{\mathbf{q}}]^T$ is the input vector containing the displacement and velocity of each blade, and the matrices $B$ and $E$ contain the appropriate forcing terms from Eq. (6).

To determine the unsteady response of the cascade, the inputs $\mathbf{u}(t)$ and $\mathbf{d}(t)$ are specified and the large system (7) is time-marched to determine the resulting flow. Often we are interested not in the flow itself, but in relevant output quantities. These are typically the forces and moments acting on the blades, but could be any feature of the response. We define an output vector $\mathbf{y}$ as

$$\mathbf{y}=C\mathbf{U}', \tag{8}$$

which for the analyses presented here contains the aerodynamic force and moment acting on each blade and the outgoing flow perturbations. $C$ is a matrix containing the linearized force calculation.

The system of ordinary differential Equations (7), (8) can be coupled into an aeroelastic model, since the input $\mathbf{u}$ and output $\mathbf{y}$ can be directly linked to a structural model. However, Eq. (7) has tens of thousands of states per blade passage (four unknowns for every point on the computational domain). Suitable models for aeroelastic analyses must have a much lower number of states, and can be derived using systematic model order reduction procedures. Such techniques are introduced in the next section.

## Model Order Reduction Procedure

Reducing the order of a system such as (7), (8) is a problem which has been studied in great detail in various contexts ([7]). Many procedures are not suitable for the problem at hand, however, because they require computation of the eigenvalues and eigenvectors of $A$; this is not practical for the problem sizes of interest in turbomachinery aeroelasticity. In addition, such approaches do not extend to nonlinear systems. Other methods, such as the proper orthogonal decomposition (POD), rely on robust computational procedures. These ''projection'' methods follow the following basic steps:

- Compute a small number of solutions using the full model. The conditions for these full solutions or ''snapshots'' are chosen to span the operating range over which fidelity to the CFD results is desired.
- Create a set of basis vectors which capture the behavior in the solutions. ''Behavior'' can be quantified by spatial patterns in the flow, snapshots of the unsteady response (frequency domain), or the input-output transfer functions. The number of basis vectors required depends on the problem and the desired degree of fit.
- Project the original equations onto the space represented by the basis vectors. The result is a set of ordinary differential equations (ODEs) for the time-varying coefficients of the basis vectors. The low-order system is completely defined by these coefficients or ''states''; approximate reconstruction of the entire flow field can be accomplished by multiplying the states by the basis vectors. Thus any output quantity (blade pressure distribution, downstream disturbances, forces and moments) can be approximated.

This procedure has been developed and reported on in previous work. In Willcox et al. [8], a unique application of the POD method for linearized systems is described, which uses frequency domain instead of time domain solutions to create a low-order model. In Willcox et al. [9], a similar method based on Arnoldi vectors, and their connection with the POD, is discussed. Appendix A outlines how the above steps are accomplished using these two methods. The ODEs for the reduced-order model are written

$$\dot{\mathbf{v}}=\mathcal{A}\mathbf{v}+\mathcal{B}\mathbf{u}+\mathcal{E}\mathbf{d} \tag{9}$$

$$\mathbf{y}=\mathcal{C}\mathbf{v} \tag{10}$$

where the states in $\mathbf{v}$ are akin to generalized coordinates, being the amplitudes of the individual basis vectors.

Using these projection methods, a model of the form (9), (10) can be obtained for an entire blade row. In Willcox [10], it is shown that each interblade phase angle can typically be represented over a broad frequency range using 6 to 10 states. Thus for a full blade row analysis, model orders of 100–500 states are required, depending on the number of blades and the frequency range of interest. This represents a three order of magnitude reduction in model order from the full model, allowing many coupled aerostructural and aeroacoustic analyses that would otherwise be impossible. Here, results are presented for two-dimensional inviscid flow, although the model order reduction methodology is applicable to any linearized CFD model.

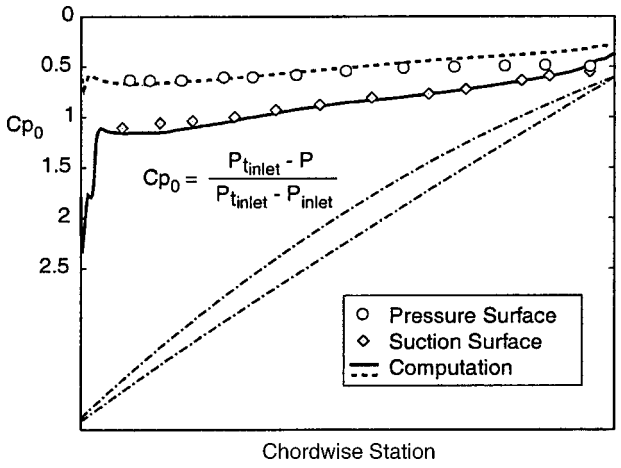



**Fig. 1 DFVLR transonic rotor from Youngren [11] and steady pressure coefficient profile computation compared to data from Schreiber and Starken [12]**

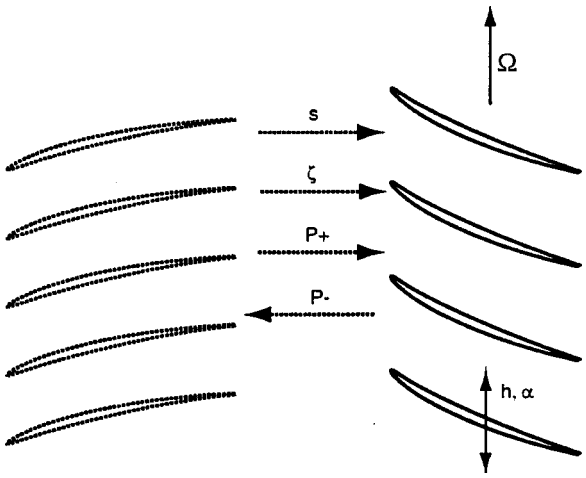



**Fig. 2 Low-speed airfoil cross sections from Silkowski [13], shown for one quarter of the computational domain used for stage analysis. Arrows represent the characteristics waves: entropy wave (*s*), vorticity wave ($\zeta$), and downstream and upstream running pressure waves (*P*+ and *P*−).**

## Examples

In this section the model order reduction procedure is applied, and compared to the results obtained by applying assumed-frequency methods to the same problems. The intent is to show the qualitative differences between the methods for relevant examples, rather than to draw quantitative conclusions. All plotted quantities represent nondimensional parameters. In the ensuing results, lengths have been nondimensionalized by the rotor chord length, and velocities by the ambient speeds of sound. Thus nondimensional frequency $\omega$ is related to reduced frequency $k$ by $\omega = kM$, where $M$ is the inlet Mach number.

### Test Cases

*Transonic Rotor.* A transonic fan blade design is used in our first two examples. Figure 1 shows the airfoil chosen for our test cases ([11]), together with a comparison of the steady pressure distribution data ([12]). The test conditions are inlet Mach number 0.82, relative flow angle 58.5 deg, blade angle 49 deg, and ambient sea level pressure and temperature. The structural system consists of 20 blades each with a mass ratio of 100, decoupled bending and twist motion, and a first bending (plunge motion) reduced frequency referenced to the chord of 0.12.

*Low-Speed Stage.* For our third example, a low-speed compressor stage is analyzed. The airfoil coordinates represent the third stage of a four-stage low-speed research compressor described in Silkowski [13]. The test conditions are inlet Mach numbers of 0.165 and 0.123, and inlet flow angles of 59.3 deg and 47.3 deg in the rotor and stator, respectively. The gap between the stator and rotor is 47.5 percent of the rotor chord. Both the rotor and stator are analyzed, assuming that the rotor has 16 blades and the stator has 20 (the actual compressor stage has 54 rotor blades and 74 stator vanes, but a smaller problem was chosen to simplify the analysis). One quarter of the periodic domain (containing five stator blades and four rotor blades) is shown in Fig. 2, although for generality, the analysis is carried out on the full stage.

**Example 1: Frequency Dependence of Aerodynamic Response Coefficients.** To illustrate how the low-order model captures the aerodynamic response to blade motion, a series of CFD solutions near the blade bending frequency $\omega = 0.1$ ($k = 0.12$) for the transonic rotor test case were obtained. These snapshots were taken for each interblade phase angle at zero frequency and at 90, 95, 100, 105, 110, 150, and 200 percent of the blade structural frequency. This resulted in 16 snapshots of the frequency-domain response of the system per interblade phase angle (8 frequencies $\times 2$ snapshots per frequency—real and imaginary parts). The boundary conditions for these CFD trials were nonreflecting, but the procedure yields models suitable for coupling to other blade rows.

Unsteady aerodynamic lift forces on the blade were computed for each interblade phase angle, using both the CFD results and reduced-order models. Figure 3 shows the results for interblade phase angles of 90, 180, and 270 deg. Both the 90 deg and the 180 deg cases show a linear variation in forced response amplitude, with almost no phase variation. It was determined that by retaining the first four POD modes computed from the available 16 snapshots, both the phase and the magnitude of the response could be accurately captured over the entire sample range (an exact match would be obtained if all 16 snapshots were used, but the resulting model order would be unnecessarily high). The dynamics for the 270-deg mode are more complicated, with a local resonance present near a frequency of 0.14. In this case, by choosing the snapshots at zero frequency and at $\omega = 0.15$, the phase and magnitude can be (fairly accurately) captured using a second-order model.

At interblade phase angles of 90 deg and 180 deg, the blade unsteady forces are not strongly dependent on reduced frequency. In the range of reduced frequencies from 0.09 to 0.11, an

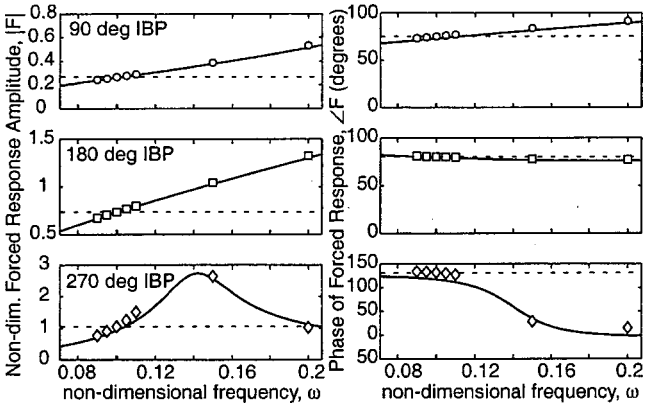



**Fig. 3 Unsteady aerodynamic plunge force on the DFVLR transonic rotor, computed at several reduced frequencies (symbols) and plotted against reduced-order model results (solid lines). Dashed lines show assumed frequency predictions, which are constant with frequency.**

assumed-frequency method would introduce an error of about ten percent in unsteady forcing. However, at an interblade phase angle of 270 deg, there is strong frequency dependence. Using the unsteady force calculated at $\omega=0.1$ would introduce an error of 36 percent at 0.09 and 30 percent at 0.11. The reduced-order model, on the other hand, tracks this strong variation of blade force magnitude. At 0.09 the reduced-order model error is nine percent, at 0.1 it is 12 percent, and at 0.11 it is 18 percent. Thus the model trades fidelity at the assumed frequency for better reproduction of the CFD results over a broad range of frequencies, at the expense of additional computations. One could further improve the accuracy by incorporating additional POD modes in the reduced-order model. This would require no additional snapshot computations, but would make the resulting model order higher.

As the range of frequencies of interest broadens further, the disparity between the two models becomes more severe. Figure 3 shows that the low-order model faithfully reproduces the CFD force results over the entire range of interblade phase angles and frequencies it is designed to model.

**Example 2: Forced Response Comparisons.** The low-order model from the previous analysis can be coupled to a structural model, to determine the effect of aerodynamics on system stability and forced response. Such calculations can be compared to those computed using an assumed frequency for both the forcing function and the blade motion.

To create a coupled system, the structural dynamics are first represented by a system of ODEs forced by aerodynamic blade forces **y** (from the aerodynamic model), whose states/outputs are the blade displacements and motions **u**:

$$\dot{\mathbf{u}}=S\mathbf{u}+T\mathbf{y}. \tag{11}$$

This structural model is then coupled to the aerodynamic reduced-order state-space system (9),(10). The coupled system is as follows:

$$\begin{bmatrix}\dot{\mathbf{v}}\\ \dot{\mathbf{u}}\end{bmatrix}=\begin{bmatrix}\mathcal{A} & \mathcal{B}\\ TC & S\end{bmatrix}\cdot\begin{bmatrix}\mathbf{v}\\ \mathbf{u}\end{bmatrix}+\begin{bmatrix}\mathcal{E}\\ 0\end{bmatrix}\cdot\mathbf{d}. \tag{12}$$

Note that the coupling of the low-order model to the upstream and downstream flowfields is not explicitly shown in this equation. If the perturbation flow inputs at the boundaries are assumed to be zero or are prescribed (for forced response analysis), and the perturbation flow outputs are not allowed to excite any external dynamical systems (such as duct acoustics), the system is uncoupled to the upstream and downstream flow fields, i.e., the boundaries are nonreflecting.

The eigenvalues of this system can be compared to those computed using aerodynamic influence coefficients computed at the blade natural frequency of $\omega=0.1$. Figures 4 and 5 show that both methods yield similar eigenvalues over a broad range of interblade phase angles, but in the critical region near a nodal diameter of 15, the methods differ by ten percent in frequency and 50 percent in damping. For a mode with $l$ nodal diameters, the corresponding interblade phase angle is $\sigma=2\pi l/r$, where $r$ is the number of blades (in this example $r=20$). The mode $l=15$ therefore corresponds to an interblade phase angle of 270 deg. To show that the reduced-order model yields higher fidelity results, the influence coefficients are recomputed at $\omega=0.09$, and the resulting eigenvalues are also shown in Fig. 5. Note that the eigenvalue at $l=15$, whose frequency is about 0.09, matches that predicted by the low-order model (this is also shown in Fig. 4), and the eigenvalue at $l=16$ is more poorly predicted than in either of the previous calculations.

The effect of the change in eigenvalues on the forced response behavior is shown in Fig. 6. To obtain these results using the assumed-frequency method, influence coefficients must be computed for the specific forcing function being modeled. A sinusoidal axial velocity defect with 15 nodal diameters ($\sigma=270$ deg), rotating at a nondimensional frequency $\omega=0.1$,

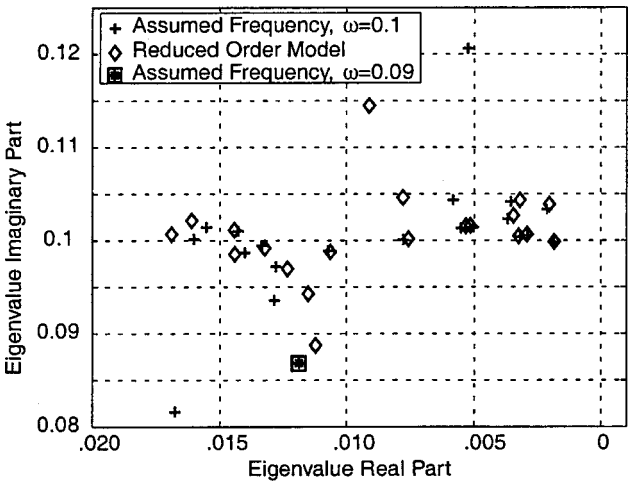



**Fig. 4 Comparison of eigenvalues obtained using reduced-order model and assumed-frequency method**

was applied in the CFD code, assuming no blade motion. The CFD results were incorporated into the assumed-frequency model, resulting in a model that assumes both the temporal *and* spatial frequency of the excitation. Both models were then excited at the same nodal diameter over a range of temporal frequencies. The resulting blade displacement predictions illustrate the inaccuracies that result from assuming the temporal forcing frequency; similar inaccuracies result from assuming the nodal diameter and spatial shape (sinusoidal) of the forcing.

Figure 6 shows that the assumed-frequency method introduces error in two ways. First, because it underpredicts the frequency of the coupled system eigenvalue, the peak forced response is at a lower frequency than the actual peak. The response of the coupled system at the forcing frequency is thus underpredicted. Second, because the damping ratio is overpredicted, the overall response is less severe. The combined effect is a 26 percent underprediction of the forced response at $\omega=0.1$ when compared to the reduced-order model. The peak response of the system is underpredicted by about 20 percent. Our conjecture is that the errors introduced by the assumed-frequency method would be even higher for lower system damping ratio, because the frequency peaks would be narrower and more sensitive to small changes in damping ratio.

To show that the reduced-order model results are correct, the aerodynamic influence coefficients were recomputed at an as-

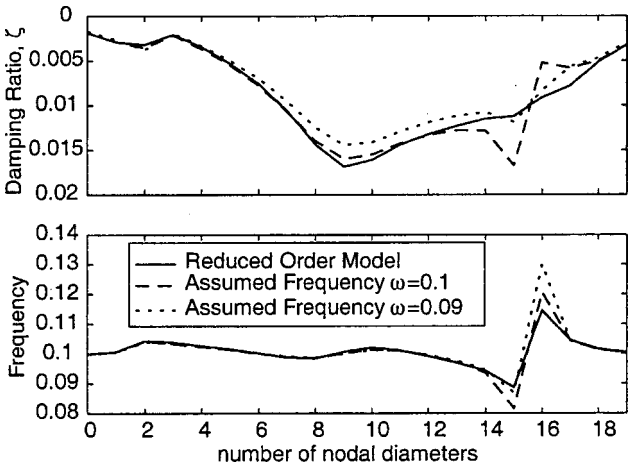



**Fig. 5 Comparison of damping and frequency of eigenvalues obtained using reduced-order model and assumed-frequency method**

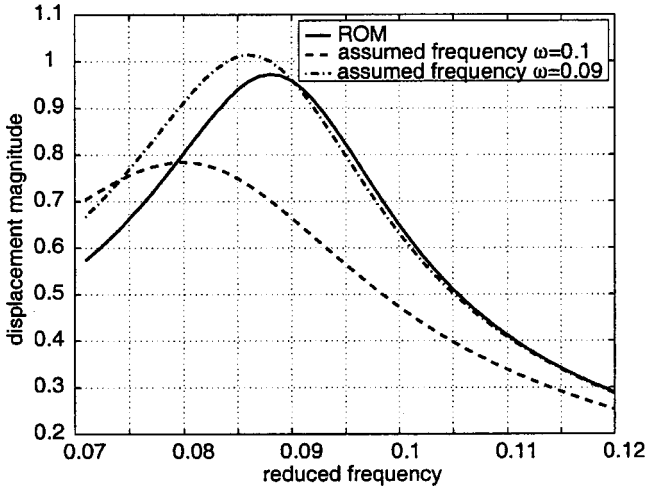



**Fig. 6 Displacement response of the transonic rotor to excitation over a broad range of reduced frequencies**

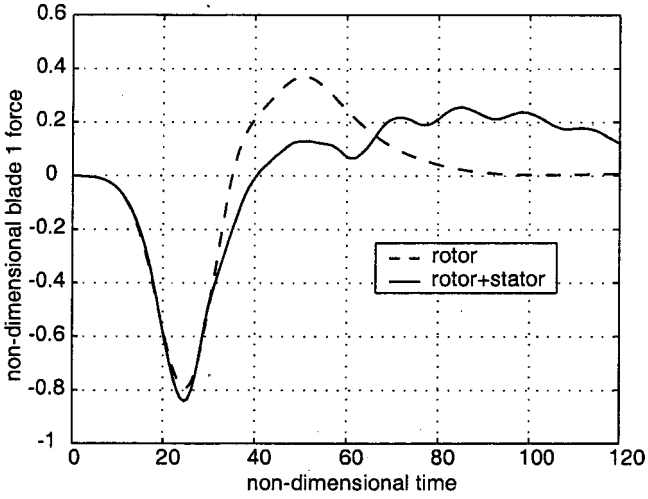



**Fig. 7 Aerodynamic force on blade 1 of the low-speed rotor due to a prescribed pulse input on its plunge displacement. Blades are considered rigid for this run.**

sumed frequency of 0.09, which is much closer to the actual eigenvalue frequency for the fifteenth nodal diameter mode. The computed forced response is very close to that computed by the reduced-order model, and the eigenvalue for the 15th nodal diameter mode is much closer to that predicted by the reduced-order model as shown by the asterisk on Fig. 4. Note that applying the assumed-frequency method iteratively would probably yield the correct answer, but would require at least two additional iterations if the reduced-order model information were not available. We avoid the iterative procedure by encapsulating several CFD runs into one low-order model. Similarly, determining the effect of various nonsinusoidal forcing functions is cumbersome using assumed-frequency methods, but straightforward using reduced-order models.

**Example 3-Multistage Analysis.** Our model order reduction procedure incorporates two important features—a model whose complexity is consistent with a user-specified range of validity, and a structure that allows blade rows to be coupled to their surroundings. To illustrate how these two features are exercised in a multistage analysis, this example analyzes a low-speed compressor stage, comparing the stability and response properties of a rotor before and after the inclusion of an upstream stator row.

Our example is the low-speed compressor stage described at the beginning of this section, which has a nondimensional period of $T=316$ for one rotor revolution. Because the Mach number of this example is low, the system's actual reduced frequency is quite high; for this example we choose a lower reduced frequency ($k=0.5$) to increase the interaction between the structural dynamics and aerodynamics.

To create a multistage model, the rotating and stationary blade rows must be coupled. To accomplish this, reduced-order models for each blade row are first developed in the relative reference frame using the Arnoldi-based approach (see Appendix). The models are then coupled in a time-marching simulation by moving the grids at each time step. The generalized boundary conditions in the reduced-order models for each blade row allow them to communicate as shown in Fig. 2. The resulting system is time varying because of the relative motion of the blade rows.

To look at the blade force during transients, a time marching simulation is performed with no coupling of the aerodynamic force to the structure. The first blade of the compressor is prescribed to execute a pulse-like plunging motion, which excites all reduced frequencies $k<11$ (see Willcox [10] for details).

The resulting aerodynamic plunge force on the blade is plotted in Fig. 7 for the rotor-only and the rotor-stator case. In the rotor-only case, the pulse input results in an initial plunging force on the blade. Subsequent to this, there is a reaction force due to the impedance of the overall fluid system, which persists for approximately 0.3 rotor revolutions. When the stator is coupled to the system, the picture changes in two ways. First, a periodic force is superimposed on the basic response, whose frequency corresponds to the blade passing frequency between the stages. Second, the response of the system is sustained over a much longer period of time, indicating that the phase properties of the fluid dynamic system have changed.

The blade forces are next examined for the coupled aerostructural system. Figure 8 shows the coupled system's response when the first blade is given an initial deflection, and the system is allowed to ''ring down'' over multiple rotor revolutions. The initial condition excites all interblade phase angles simultaneously, so the response will evolve toward the least stable interblade phase angle as time proceeds. The decaying oscillations that exist for the rotor only become sustained oscillations when the stator is added, with superimposed high frequency forces due to the stator passing.

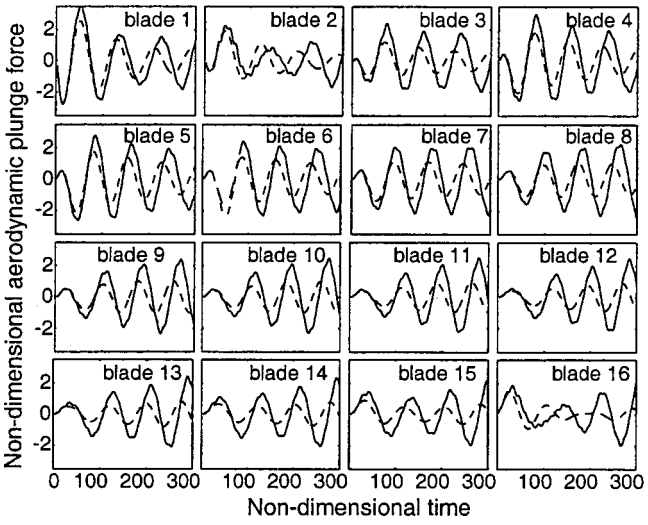



**Fig. 8 Force on every blade of low-speed rotor due to an initial displacement of the coupled system. The structure and aerodynamics are coupled in this run. Solid—with stator, dashed—no stator.**

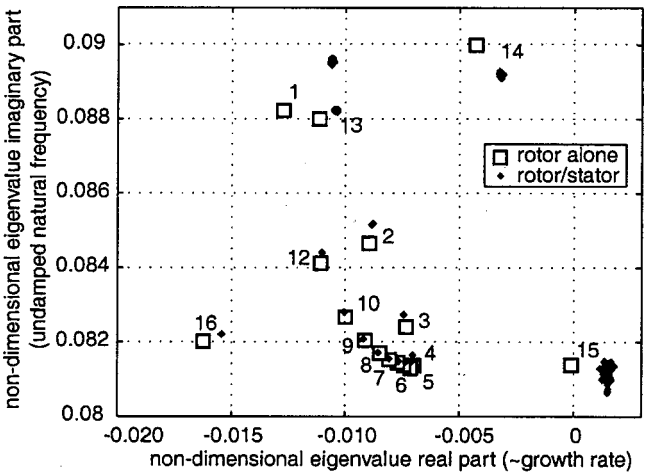



**Fig. 9 Eigenvalues of the low-speed rotor. Eigenvalues are numbered by their nodal diameter ($\sigma = 2\pi 1/16$).**

Although the system we have created is time varying, so that linear time invariant analyses are not strictly valid, a crude stability analysis can be performed by computing the frozen-time eigenvalues over the range of relative positions of the blade rows. Figure 9 shows the eigenvalues computed in this way, comparing the rotor-only eigenvalues and the eigenvalues for the stage with a stator-rotor gap of 47.5 percent of the rotor chord. Note that the rotor-stator eigenvalues appear in closely packed clusters, indicating that the system is only weakly time-varying. All of the eigenvalue clusters lie entirely in either the left or right half-plane, so stability of the system is assumed to depend on the locations of these clusters. Based on this assumption, Fig. 9 indicates that the eigenvalues associated with a nodal diameter of $l=15$ (interblade phase angle of 337.5 deg) are destabilized by bringing the stator into the proximity of the rotor. This explains the difference between the responses in Fig. 8.

## Discussion and Conclusions

Because the models created using projection methods are much lower order than their CFD counterparts, modeling the behavior of entire compressors is feasible. Such modeling must be done with a clear goal in mind, so that the appropriate range of validity is built in to the reduced-order models. But because the procedure is systematic, and because the low-order models allow interconnection of upstream and downstream impedance conditions, one can envision a design system that allows many types of interactions to be studied using the same baseline codes.

In the examples presented in this paper, we showed potential areas where such a design system would allow higher fidelity, more complete aeroelastic analysis to be carried out. The fact that the solutions obtained using reduced-order models capture the CFD results better than those obtained with simpler models is not surprising: more detailed information is computed and incorporated, yielding improved results in a direct way.

While derivation of models using the POD technique will in general be as expensive as an assumed-frequency approach, the Arnoldi method can be used efficiently to derive high-fidelity reduced-order models. For example, with $r$ blades, we require just $r$ matrix factorizations to obtain the Arnoldi reduced-order model. The assumed-frequency method will require $r$ factorizations per frequency iteration, while the POD requires $r$ factorizations per snapshot frequency. To obtain an accurate model, the number of snapshot frequencies will be comparable to (or may exceed) the number of frequency iterations one might use to obtain the desired fidelity using an assumed-frequency approach. Clearly, the Arnoldi method is more desirable; however, it is a new approach and has not yet been implemented for all types of problems. Hence we have obtained some of the results here using the more expensive POD. For a more complete comparison of POD and Arnoldi methods, see Willcox et al. [9].

Using either the POD or Arnoldi approach, the resulting models are valid over a (chosen) range of frequencies for general forcing inputs. Methods that assume both the shape and frequency of motion and forcing have unknown, potentially small ranges of validity. In contrast, reduced-order models can be used over the appropriate frequency range with the confidence of the original high-fidelity CFD code. Whether the improvements obtained are worth the computational effort remains to be determined, and will depend on the specific problem considered. Therefore the question of whether reduced-order models are necessary and cost effective is not addressed here; we have instead demonstrated the following:

- Model order reduction can be applied in a generalized framework that is independent of the CFD code used. This capability is achieved through the use of canonical snapshots of the flow, rather than eigenvalue analyses or methods involving frequency domain fitting. Using this framework, application of reduced-order modeling to a turbomachinery Euler code was demonstrated for the first time.
- A variety of problems can be studied using reduced-order modeling, provided that the range of desired model validity is carefully defined. In particular, analyses involving multiple simultaneous effects, broadband temporal/spatial excitation, or multiple stages are feasible.
- Because of their matrix ordinary differential equation form (also known as ''state space'' form), reduced-order models can be easily introduced into existing aeroelastic analyses. This allows one to assess the impact of assumed-frequency models on fidelity, and easily incorporate additional computational results into the analysis when warranted.
- Initial results using realistic blade rows indicate that the fidelity of existing aeroelastic methods may in fact be insufficient, especially if the blade row is assumed to be isolated. Further research is required in this area.

## Acknowledgments

The authors would like to thank Asif Khalak for his comments and conceptual contributions. This research was supported by AFOSR Grant #F49620-95-1-0409, Marc Jacobs, technical monitor.

## Appendix

Some details of the steps outlined in the discussion of model order reduction procedures are given in this Appendix. For further details, see Willcox [10].

**Generation of Snapshots.** Typically, projection methods for model order reductions begin with *snapshots* of the flow field obtained from a time-marching simulation. This is impractical for an entire bladed disk, and unnecessary for linearized flows. A more efficient approach is to decompose the forcing and aerodynamic response into spatial and temporal Fourier modes, as is typical in aeroelastic analysis. Blade motion is decomposed into $r$ discrete spatial Fourier modes with interblade phase angles $\sigma_j = 2\pi j/r$, where $r$ is the number of blades in the compressor, and an infinity of temporal frequencies $\omega_k$. External disturbance forcing is similarly viewed as a summation over an infinite number of spatial and temporal modes with frequencies $\sigma_j$ (now for any $j$ since the disturbance is a continuous function) and $\omega_k$, respectively. Blade motion $\mathbf{u}$ and disturbances $\mathbf{d}$ are thus written for any blade passage $n$ as follows:

$$\mathbf{u}_n^{jk}(t) = \tilde{\mathbf{u}}_{jk} e^{i\omega_k t} e^{i(n-1)\sigma_j} \tag{13}$$

$$\mathbf{d}_n^{jk}(b,t) = \tilde{\mathbf{d}}_{jk} e^{i\omega_k t} e^{i(n-1)}\sigma_j \tag{14}$$

where $b$ denotes the boundary along which $\mathbf{d}$ is defined for a given blade passage. The corresponding perturbation flow solution in each passage will also be harmonic of the form

$$\mathbf{U}_n^{jk}(\mathbf{x},t)=\tilde{\mathbf{U}}_{jk}(\mathbf{x})e^{i\omega_k t}e^{i(n-1)\sigma_j}, \tag{15}$$

where the vector $\mathbf{U}_n$ represents the unknown perturbation flow variables associated with blade $n$. Thus if the response of the first blade $(n=1)$ is known, the response of all subsequent blades can be determined by using (15). Therefore we solve the linearized Euler Eq. (7) in the first blade passage only, casting them in the frequency domain as

$$[i\omega_k-A_j]\tilde{\mathbf{U}}_{jk}=B\tilde{\mathbf{u}}_{jk}+E\tilde{\mathbf{d}}_{jk}, \tag{16}$$

where $A_j$ represents the the original matrix $A$ in Eq. (7) modified to to impose complex periodicity of the solution; i.e., the solution must be identical if phase shifted by the interblade phase angle $\sigma_j$. Solutions of the frequency domain CFD Eqs. (16) provide images of the unsteady flow field at each temporal frequency $\omega_k$, for each spatial frequency $\sigma_j$. The real and imaginary parts of these images form the snapshots for the projection analysis.

Although (16) implies that the snapshots are taken at imposed sinusoidal motions, as in assumed-frequency methods, this is merely a means of exciting the system dynamics to obtain a reduced-order subspace (next section). In the final projection step, all of the dynamics are projected onto this subspace, and a reduced-order model is obtained that is approximately valid for any general forcing function, provided its frequency content lies within the values of $\omega_k$ sampled.

**Generation of an Optimal Set of Basis Vectors.** To obtain the reduced-order model, the unknown flow solution vector is projected onto a small number of orthonormal basis functions $\{\Psi_i\}$, so that the flow dynamics can be represented as

$$\mathbf{U}'(t)=\sum_{i=1}^{m} v_i(t)\Psi_i. \tag{17}$$

If the basis set is chosen to be efficient, it can accurately represent the flow dynamics even though $m$ is small. The following subsections outline two alternative methods to create the basis set.

*Proper Orthogonal Decomposition.* The POD computes a basis set by seeking functions which optimally represent the flow solution $\mathbf{U}'(t)$. More specifically, they maximize the following cost ([14]):

$$\max_{\Phi}\frac{\langle|(\mathbf{U}',\Phi)|^2\rangle}{(\Phi,\Phi)}=\frac{\langle|(\mathbf{U}',\Psi)|^2\rangle}{(\Psi,\Psi)}, \tag{18}$$

where $(\mathbf{U}',\Psi)$ denotes the inner product of the basis vector with the field $\mathbf{U}'(t)$ and $\langle\ \rangle$ represents a time-averaging operation. It can be shown that a necessary condition for (18) to hold is that $\Psi$ is an eigenfunction of the kernel $K$ defined by

$$K(\mathbf{x},\mathbf{x}')=\frac{1}{n}\sum_{i=1}^{n}\mathbf{U}_i(\mathbf{x})\mathbf{U}_i(\mathbf{x}'), \tag{19}$$

where $\mathbf{U}_i(\mathbf{x})$ is the instantaneous perturbation flow field or snapshot at a frequency $\omega_i$, and the number of snapshots $n$ is sufficiently large ([15]). The eigenvectors of $K$ are of the form

$$\Psi=\sum_{i=1}^{n}B_i\mathbf{U}_i, \tag{20}$$

where $B_i$ have been shown to be the eigenvectors of the eigenvalue decomposition of the correlation matrix $C$:

$$CB=\Lambda B \tag{21}$$

where the correlation matrix $C$ is given by

$$C_{ij}=\frac{1}{n}(\mathbf{U}_i,\mathbf{U}_j). \tag{22}$$

The function that achieves the maximum in (18) corresponds to $\lambda_1$, the largest eigenvalue of $C$ ([15]), and is constructed using the eigenvector associated with $\lambda_1$ via the equation $\Psi_1=\Sigma_{j=1}^{n}\beta_j^1\mathbf{U}_j$, where $B_i=[\beta_1^i\beta_2^i\ldots\beta_n^i]$. Subsequent functions correspond to the eigenvalues $\lambda_i$ ordered according to size: $\lambda_1\geqslant\lambda_2\geqslant\ldots\geqslant\lambda_n$. The subspace spanned by the resulting vectors $\Psi_i$ minimizes the ''averaged energy'' or the 2-norm of the error between the exact and projected data.

*Arnoldi-Based Approach.* An alternative approach for computing an efficient basis is to approximately match the transfer functions of the low-order system to those of the high-order system. The transfer function of the high-order system (7), (8) for a single input $u$ and single output $y$ is

$$H(s)=\mathbf{c}(sI-A)^{-1}\mathbf{b}, \tag{23}$$

which can also be represented as a rational function

$$H(s)=\frac{N(s)}{D(s)}, \tag{24}$$

where the numerator $N(s)$ and denominator $D(s)$ are both polynomials in $s$. A $q$th order Padé approximation to the transfer function is defined as

$$H_q(s)=\frac{b_{q-1}s^{q-1}+\ldots+b_1s+b_0}{a_qs^q+a_{q-1}s^{q-1}+\ldots+a_1s+1}, \tag{25}$$

where the $2q$ coefficients $a_j$, $b_j$, are selected to match the first $2q$ terms in a McLaurin expansion of the transfer function (23), written

$$H(s)=-\sum_{k=1}^{\infty}m_ks^k \tag{26}$$

where

$$m_k=\mathbf{c}^T\mathbf{A}^{-(k+1)}\mathbf{b} \tag{27}$$

is called the $k$th moment of $H(s)$. The $q$th order Padé approximation (25) can be constructed via the Lanczos process and will match the first $2q$ moments of $H(s)$ ([16]). However, this process carries with it no guarantees concerning the stability properties of the approximation.

Another option is to use the Arnoldi method to generate a set of vectors which spans the $q$th order Krylov subspace defined by

$$\mathcal{K}_q(A,\mathbf{b})=\text{span}\{A^{-1}\mathbf{b},A^{-2}\mathbf{b},\ldots,A^{-q}\mathbf{b}\}. \tag{28}$$

This set of $q$ ''Arnoldi vectors'' matches $q$ moments of the system transfer function, only half the number matched by the Padé approximation. However, since the Arnoldi approach has the advantage of generating a congruent transformation, in many cases it generates models with guaranteed stability.

It is also possible to reduce systems with multiple inputs using the block Arnoldi method. For example, if we consider a system with two inputs $u_1$ and $u_2$,

$$\dot{\mathbf{U}}'=A\mathbf{U}'+\mathbf{b}_1u_1+\mathbf{b}_2u_2, \tag{29}$$

then the block Arnoldi method is used to generate vectors which span the Krylov subspace

$$\mathcal{K}_q(A,\mathbf{b}_1,\mathbf{b}_2)=\text{span}\{A^{-1}\mathbf{b}_1,A^{-1}\mathbf{b}_2,A^{-2}\mathbf{b}_1,A^{-2}\mathbf{b}_2,\ldots,A^{-q}\mathbf{b}_1,A^{-q}\mathbf{b}_2\}. \tag{30}$$

We also note that it is not necessarily the first $q$ moments which must be matched. If we were to consider a Taylor series expansion of the transfer function about some nonzero value of $s$, a model could be obtained which would give a better approximation of the

system dynamics for higher frequencies. These multiple frequency point Arnoldi methods are described in Grimme [17].

**Projection of Equations Onto Optimal Basis.** Once the basis vectors $\Psi_i$ have been computed, by whichever method chosen, the assumed expansion for the perturbation solution (17) is substituted into the governing Eqs. (7). Making use of orthogonality, a system of ordinary differential equations for the modal coefficients is obtained as follows:

$$\frac{dv_i}{dt} = \Psi_i^T A \sum_{j=1}^{m} v_j \Psi_j + \Psi_i^T B\mathbf{u} + \Psi_i^T E\mathbf{d}. \tag{31}$$

Since solutions at different interblade phase angles are orthogonal to one another, the system of Eq. (31) can be constructed for each value of $\sigma_j$ and combined to form a block diagonal state-space system for the entire bladed disk. For each pair of interblade phase angles $\pm\sigma_j$, a set of equations is obtained which has the following form:

$$\dot{\mathbf{v}}_{\sigma_j} = \mathcal{A}_{\sigma_j}\mathbf{v}_{\sigma_j} + \mathcal{B}_{\sigma_j}\mathbf{u}_{\sigma_j} + \mathcal{E}_{\sigma_j}\mathbf{d}_{\sigma_j}. \tag{32}$$

In (32) the aerodynamic state and input vectors are given in interblade phase angle coordinates, whereas we require the blade motion to be expressed in blade coordinates for coupling with a structural model. Therefore we perform a transformation from traveling wave to standing wave coordinates as described in Dugundji and Bundas [5], which can be written $\mathbf{u}_\sigma = \mathcal{T}\mathbf{u}$. As a result, the block diagonal state-space matrix $\mathcal{B}_\sigma$ becomes block circulant. The final set of aerodynamic equations can then be written

$$\frac{d\mathbf{v}}{dt} = \mathcal{A}\mathbf{v} + \mathcal{B}_\sigma\mathcal{T}\mathbf{u} + \mathcal{E}\mathbf{d} \tag{33}$$

and

$$\mathbf{y} = \mathcal{C}\mathbf{v}, \tag{34}$$

where $\mathbf{v}$ contains the aerodynamic states for all interblade phase angles and the matrix $\mathcal{A}$ is block diagonal with the submatrices $\mathcal{A}_{\sigma_j}$. The vector $\mathbf{y}$ is any desired output, but for the problems considered here contains the forces and moments acting on the blades, and outgoing perturbation waves at the passage inlet and exit.

Note that static corrections may be used to aid in capturing high-frequency dynamics as described in Romanowski and Dowell [18]. This was done for the POD model presented here. In this case the reduced-order model has a slightly different form. The static correction functions are particular solutions of the flow system and are precomputed by solving for steady flows with unit boundary conditions on blade position, velocity, and external disturbance.

## References


[1] Chiang, H. D., and Kielb, R. E., 1993 "An Analysis System for Blade Forced Response," ASME J. Turbomach., **115**, pp. 762–770.

[2] Hall, K., and Silkowski, P., 1995 "The Influence of Neighboring Blade Rows on the Unsteady Aerodynamic Response of Cascades." ASME Paper 95-GT-35.

[3] Manwaring, S. R., and Wisler, D. C., 1993, "Unsteady Aerodynamics and Gust Response in Compressors and Turbines," ASME J. Turbomach., **115**, pp. 724–740.

[4] Crawley, E., and Hall, K., 1985 "Optimization and Mechanisms of Mistuning in Cascades," ASME J. Eng. Gas Turbines Power, **107**, pp. 418–426.

[5] Dugundji, J., and Bundas, D., 1984 "Flutter and Forced Response of Mistuned Rotors Using Standing Wave Analysis," AIAA J., **22**, No. 11, pp. 1652–1661.

[6] Crawley, E., 1988, "Aeroelastic Formulation for Tuned and Mistuned Rotors," *AGARD Manual on Aeroelasticity in Axial-Flow Turbomachines*, Vol. 2. AGARD-AG-298.

[7] Dowell, E., Hall, K., Thomas, J., Florea, R., Epureanu, B., and Heeg, J., 1999, "Reduced Order Models in Unsteady Aerodynamics," AIAA Paper 99–1261.

[8] Willcox, K., Paduano, J., Peraire, J., and Hall, K., 1999, "Low Order Aerodynamic Models for Aeroelastic Control of Turbomachines." AIAA Paper 99-1467.

[9] Willcox, K., Peraire, J., and White, J., 1999, "An Arnoldi Approach for Generation of Reduced-Order Models for Turbomachinery," Comput. Fluids, **31**(3), pp. 369–389.

[10] Willcox, K. E., 2000 "Reduced-Order Aerodynamic Models for Aeroelastic Control of Turbomachines," Ph.D. thesis, Department of Aeronautics and Astronautics, MIT.

[11] Youngren, H., 1991 "Analysis and Design of Transonic Cascades with Splitter Vanes," Master's thesis, Department of Aeronautics and Astronautics, M.I.T.

[12] Schreiber, H., and Starken, H., 1984 "Experimental Investigation of a Transonic Compressor Rotor Blade Section," ASME J. Eng. Gas Turbines Power, **106**, pp. 288–294.

[13] Silkowski, P., 1995 "Measurements of Rotor Stalling in a Matched and Mismatched Compressor." GTL Report 221, M.I.T.

[14] Sirovich, L., 1987 "Turbulence and the Dynamics of Coherent Structures, Part 1: Coherent Structures," Q. Appl. Math., **45**, No. 3, pp 561–571.

[15] Berkooz, G., Holmes, P., and Lumley, J., 1993, "The Proper Orthogonal Decomposition in the Analysis of Turbulent Flows," Annu. Rev. Fluid Mech., **25**, pp. 539–575.

[16] Feldmann, P., and Freund, R., 1995, "Efficient Linear Circuit Analysis by Padé Approximation via the Lanczos Process," IEEE Trans. Comput.-Aided Des., **14**, pp. 639–649.

[17] Grimme, E., 1997, "Krylov Projection Methods for Model Reduction," Ph.D. thesis, Coordinated-Science Laboratory, University of Illinois at Urbana-Champaign.

[18] Romanowski, M., and Dowell, E., 1994 "Using Eigenmodes to Form an Efficient Euler Based Unsteady Aerodynamics Analysis," *Proceedings of the Special Symposium on Aeroelasticity and Fluid/Structure Interaction Problems*, Vol. AD-Vol. 44, ASME, New York, pp. 147–160.

# Rotor/Seal Experimental and Analytical Study on Full Annular Rub

**J. J. Yu**[1]
e-mail: john.yu@bently.com

**P. Goldman**

**D. E. Bently**

**A. Muzynska**

Bently Rotor Dynamics Research Corporation,
1631 Bently Parkway South,
Minden, NV 89423

*Rotor/seal full annular rub, including synchronous (forward) and reverse (backward) precessions, has been investigated both experimentally and analytically. Of particular interest is the finding of reverse precessional full annular rub (dry whip) that occurs repeatedly in small clearance cases without any outside disturbance. The experimental results include rub triggering mechanism, mass unbalance, and rotative speed effects. A simplified mathematical model is used to interpret experimental results. Nonlinear solutions for both synchronous and reverse precessions are obtained along with instability zones. Mass unbalance effect on shifting from synchronous response to reverse rub and destabilizing factors such as dry friction, rotor damping, and seal stiffness, are discussed.*
[DOI: 10.1115/1.1416691]

## 1 Introduction

Rotor to stationary part rubbing is a serious malfunction in rotating machinery that may lead to a machine catastrophic failure. The physical contact of the rotating shaft with a stationary element, which in the normal operation should not occur, involves several major physical phenomena, such as friction, impacting, and modification of the system stiffness. As a result of rub, the machine normal operation is affected.

The rub is a secondary phenomenon resulting from a primary cause, such as rotor high lateral vibrations and high lateral displacement of the rotor centerline due to misalignment or working fluid-related constant radial forces. The trend to increase efficiency of fluid-handling machines leads to minimizing rotor/seal clearances, which in turn creates a higher hazard of rubbing.

There are mainly two types of rub. In the ''light partial rub'' ([1,2]), the contact forces are minimal. The rotor occasionally touches the stationary element and maintains contact during a fraction of its precessional period. It is usually not too dangerous to the integrity of the machine. The second, and more serious phenomenon in its destructive effects to the machine, is a full annular rub ([3]), which can occur in seals or in auxiliary bearings ([4]). During full annular rub, the rotor maintains contact with the seal continuously or almost continuously. Self-excited reverse precessional full annular rub, known as ''dry whip,'' is more severe than synchronous full annular rub.

Recently, Yu et al. [5] and Goldman and Bently [6] have conducted experimental and analytical studies on the full annular rub, which has been a subject of many papers. Muszynska [7] did literature survey on this subject. Black [8] examined the synchronous interaction of a rotor and stator due to contact across a clearance annulus. Ehrich [9] indicated that ratio of reverse whirl frequency to natural frequency only depended on stator-to-rotor natural frequency and damping ratios. Muszynska [3] discussed full annular rub with a 2-degree-of-freedom model and presented measured results of reverse precessional rub triggered by external impacts. Lingener [10] and Crandall [11] demonstrated dry-friction whirl and whip for a case when ratio of rotor radius to radial clearance $r/C_r$ reached 2.66. Childs [12] summarized their work and indicated that dry-friction whirl and whip were only likely to occur at a large clearance and that triggering of contact normally required an outside disturbance.

However, new experiments show in this paper that this type of self-excited reverse precessional rub with an almost constant frequency can be ensured to occur in a rotor kit attached with a seal. No outside disturbance or impact is needed to trigger the reverse precession. Moreover, it has been found that reverse precessional rub can even occur for small clearance cases where ratio of rotor radius to radial clearance $r/C_r$ reaches 40.

This paper presents the results of experimental and analytical studies on rotor/seal full annular rub. Both synchronous and reverse precessions are included, with emphasis on the latter and the transition from the former to the latter without any outside disturbance.

## 2 Test Rig

As shown in Fig. 1, a shaft with the diameter and length of 0.01m and 0.56m, respectively, was supported by two brass bushing bearings and driven by a 75 W motor. One or two 0.8-kg disks were attached to the shaft in the midspan. A seal was located at the outboard side of the shaft. The seals used in the experiments were either tightly or flexibly fitted with an ''O-ring'' in the seal support. The rotor was originally centered in the seal and well balanced. A known mass unbalance was then added to one of the disks at the radius of 0.03 m. Without contact with the seal, the rotor first balance resonance speed of synchronous ($1\times$) motion was around 1500 rpm with two disks attached, and around 2000 rpm with only one disk. Slight system anisotropy existed for both uncoupled (without a seal contact) and coupled (with a seal contact) rotors. Teflon™ and bronze seals with diametral clearances varying from 250 to 1000 $\mu$m were used. The data acquisition and processing system consisted of four $XY$ displacement proximity probes (probes 1 and 2 were located near the seal, and probes 3 and 4 were close to the rotor midspan), one speed probe and one Keyphasor® probe for speed and phase measurement.

The coefficient of static dry friction between steel and Teflon surfaces, as quoted from a handbook, is usually less than 0.1. During the experiment, however, a high contact force and friction-related heat severely deformed the Teflon seal and changed the normal contact direction on the rotor/seal surface, thus yielding a high equivalent friction coefficient. The reverse precessional full annular rub, accompanied by a higher contact normal force, corresponded to a higher value of the equivalent coefficient of sliding friction than the forward precessional rub.

[1]To whom correspondence should be addressed.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-389. Manuscript received by IGTI Nov. 1999; final revision received by ASME Headquarters Feb. 2000. Associate Editor: D. Wisler.

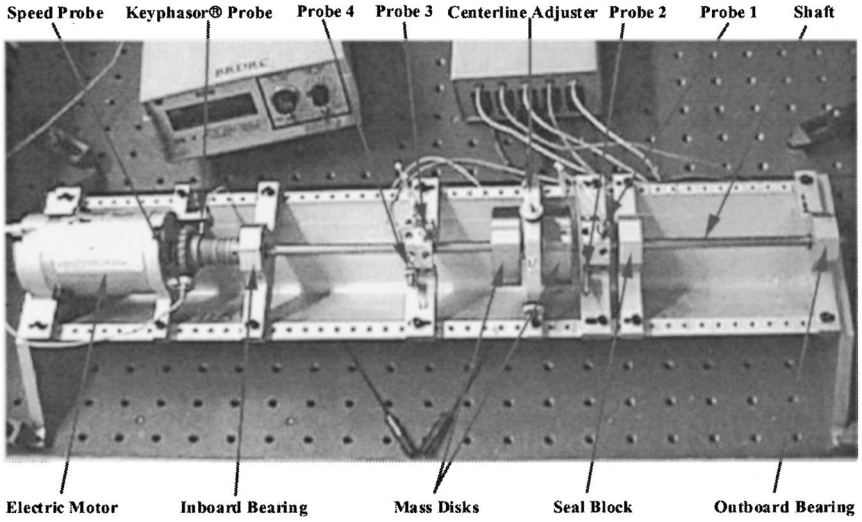



**Fig. 1 Rotor/seal full annular rub test rig**

## 3 Experimental Study

**3.1 Synchronous (Forward Precessional) Rub.** When the friction between the shaft and the seal was not high (such as cases where the contact surface is contaminated with oil), the rotor 1× forced response to mass unbalance was the major vibration component during runup and rundown. This regime is also called the "synchronous rub." An interaction with the seal was seen in the first-balance-resonance speed range. When the rotor/seal contact occurred, the seal started acting as an additional bearing, thus yielding a coupled rotor/seal system with increased stiffness and natural frequencies. The rotor slid along the seal surface without rolling at a particular speed. Since the friction force was small, the rub resulted in only minor surface damage.

Figure 2 shows the direct and 1× Bode plots of rotor horizontal (*X*-direction) responses during runup and rundown, with and without rubbing. Due to differences in the rotor angular acceleration, the runup and rundown curves differ slightly without contacting the seal. In the range of the original resonance, responses changed when rubbing was involved. The resonance range of speeds was much wider during the runup than rundown. Peak response am-

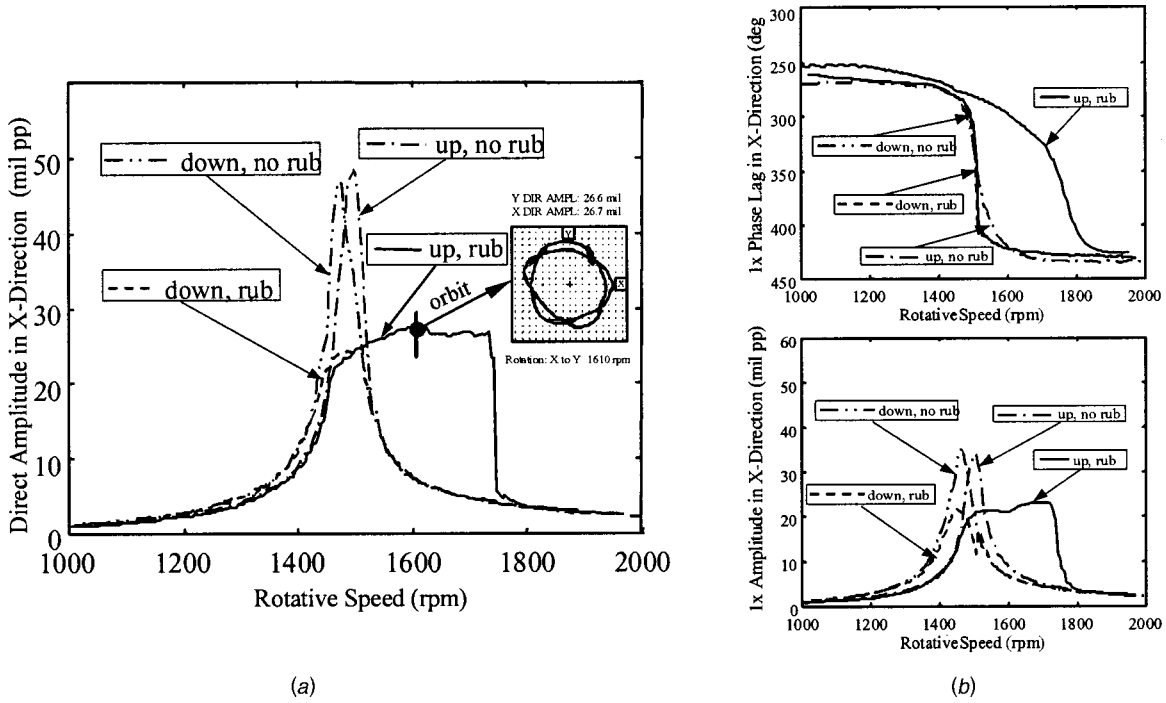



**Fig. 2 Comparison of rotor lateral responses of the two-disk rotor during runup and rundown tests with/without rubbing, respectively. Teflon seal with diametral clearance of 500 $\mu$m; two disks with mass unbalance of 1.1 grams at 0 degree. (*a*) Direct responses, (*b*) 1× Bode plots near the seal location. An insert in (*a*) displays the rotor orbit indicating multicontact intermittent rub.**

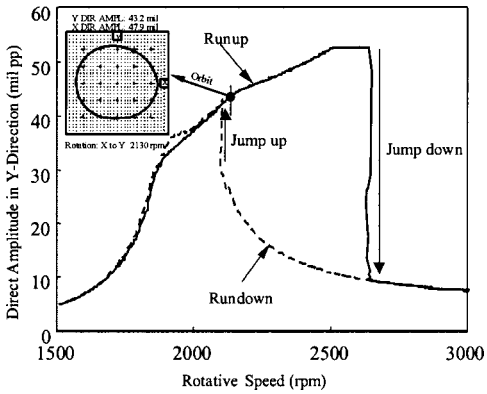



**Fig. 3 Synchronous rub showing that amplitude jumps down during runup and jumps up during rundown. Teflon seal with diametral clearance of 1000 $\mu$m; one disk with mass unbalance of 1.6 grams.**

plitudes were reduced due to the seal acting as a bearing. In this case, the rotor was not in contact with the seal continuously. The orbit in Fig. 2 shows that the rotor bounced inside the seal, but it maintained the forward precession. As the rotative speed increased or decreased, the rotor conditions changed from noncontact to contact and then to noncontact again, with jumping amplitude.

During rundown changes in 1× phase values were minor, but there was a clear amplitude "jump-up" discontinuity phenomenon which was often overlooked. The fact that a sudden decrease in 1× amplitude (see Fig. 2), accompanied by onset of other frequency components when the rotor was just contacting the seal during rundown, indicates an impact due to amplitude discontinuity. Figure 3 makes it clear that a "jump-up" did exist during rundown. In this case, lubrication oil was dropped onto the contact surface to avoid severe reverse precessional rubbing. Unlike the case in Fig. 2 where the contact surface was not fully lubricated, the rotor orbit pattern in Fig. 3 was very smooth. The rotor "bounce" pattern depended significantly on surface friction.

During rubbing there was a "third bearing" effect, as described by Yu et al. [5]. The rotor mode shape was changed due to this physical restriction.

**3.2 Reverse Precessional Rub.** Without an outside disturbance, reverse precessional full annular rub occurred repeatedly for the ratio of rotor radius-to-radial clearance $r/C_r$ ranging from 10 to 40. Around the resonance speed, synchronous forward precessional rub developed into reverse precessional full annular rub when rotor/seal surface friction was high. Once generated, the reverse precessional rub could be sustained over the whole speed range as long as slippage was maintained between the contact surface. With a hammer impact, the reverse rub could be disengaged at nonresonance speed, and engaged above very low speed (for instance, ten percent of the resonance speed).

*3.2.1 Rub Triggering Without Outside Disturbance.* Figures 4(*a*) and 4(*b*) show the transition process from the synchronous forward precessional rub to reverse precessional full annular rub. When the rotative speed was approaching the first balance resonance, the rotor started bouncing inside the seal, showing elliptical orbits slowly rotating their major axes due to lateral stiffness anisotropy effect. Multicontact intermittent rub came first with forward precessional orbits. Afterwards, reverse-dominant precession occurred from around 1550 rpm to 1650 rpm. Then suddenly, around 1700 rpm, the full annular rub appeared at a frequency of 2300 cpm. The much higher vibration component in reverse frequency (−2300 cpm) than that in forward frequency (2300 cpm), as shown in Fig. 4(*b*), also indicates an almost circular reverse orbit motion. With the increase in rotative speed, the rubbing frequency remained almost unchanged, though the amplitude was slightly increased due to thermal expansion of the seal.

The reverse precessional full annular rub could also be triggered during rundown after the rotor contacted the seal near the resonance speed (see Fig. 5). Then the rub was maintained until very low speed, replacing the synchronous 1× motion.

*3.2.2 Mass Unbalance Effect.* For the same experimental setup as it is discussed in Section 3.2.1, when the amount of mass unbalance was reduced to 0.4 grams, the reverse precessional full annular rub was not generated although the rotor bounced inside the seal around the resonance speed.

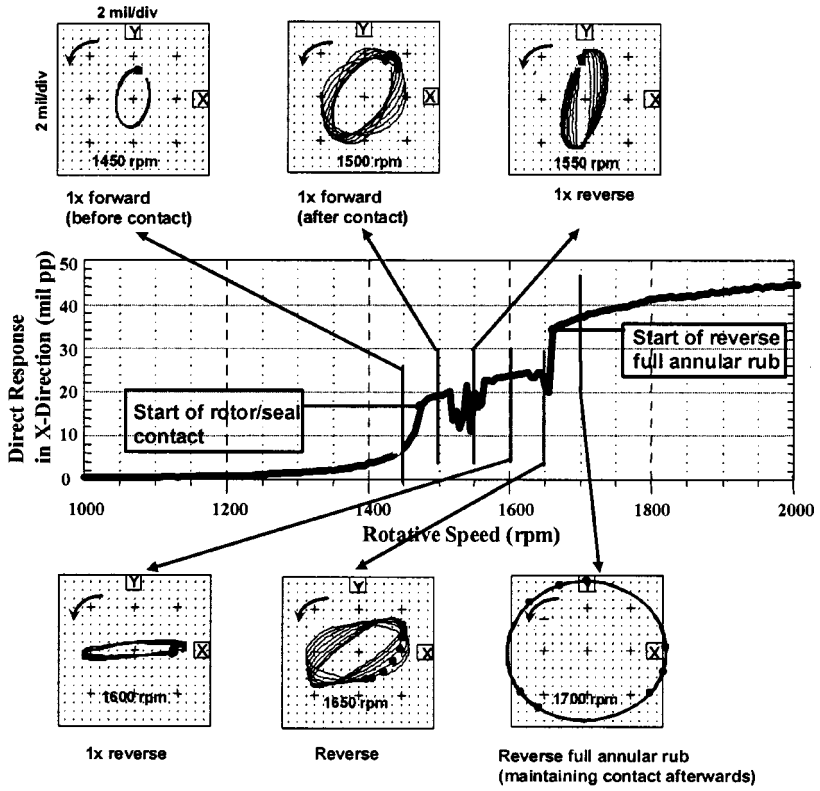
In addition to triggering of reverse precessional full annular rub, amount of mass unbalance can also affect the starting point of the rub in terms of rotative speed, as shown in Fig. 6 where the seal clearance was 250 $\mu$m. The first resonance speed of the original uncoupled rotor system was around 2000 rpm. In this case, for lower mass unbalance, the reverse rub occurred after the resonance speed, while for higher mass unbalance it occurred before the resonance speed, with the same precessional frequency of 3120 cpm. Although the rotor with higher mass unbalance started to contact the seal at around 1600 rpm, the reverse rub was not engaged quickly until around 1900 rpm where an increasing contact force due to synchronous 1× rub motion was able to trigger it.

*3.2.3 Rotative Speed Effect.* Once generated, the reverse precessional full annular rub could not be disengaged by changes in rotative speed unless at very low speed where slippage could not be maintained. Figures 7(*a*) and 7(*b*) show a waterfall plot and direct response, respectively, for an entire running period, including startup and shutdown. The reverse full annular rub started around 1600 rpm during runup. A frequency of 2280 cpm (38 Hz) and amplitude of around 1270 $\mu$m (50 mils) were maintained with the increase of speed until 2500 rpm. They were also maintained during almost entire rundown. As the speed decreased to around 500 rpm, the amplitude started to decrease slightly for lack of energy. The vibration frequency was accordingly reduced due to the reduced seal stiffness resulted from the decreased amplitude and contact force. Figure 7(*c*) shows the reverse rub at rotative speed of 170 rpm. The reverse full annular rub did not vanish until at speed of 120 rpm (compared with resonance speed of around 1500 rpm). Above 120 rpm, a hammer impact could easily generate the reverse rub, and also release it.

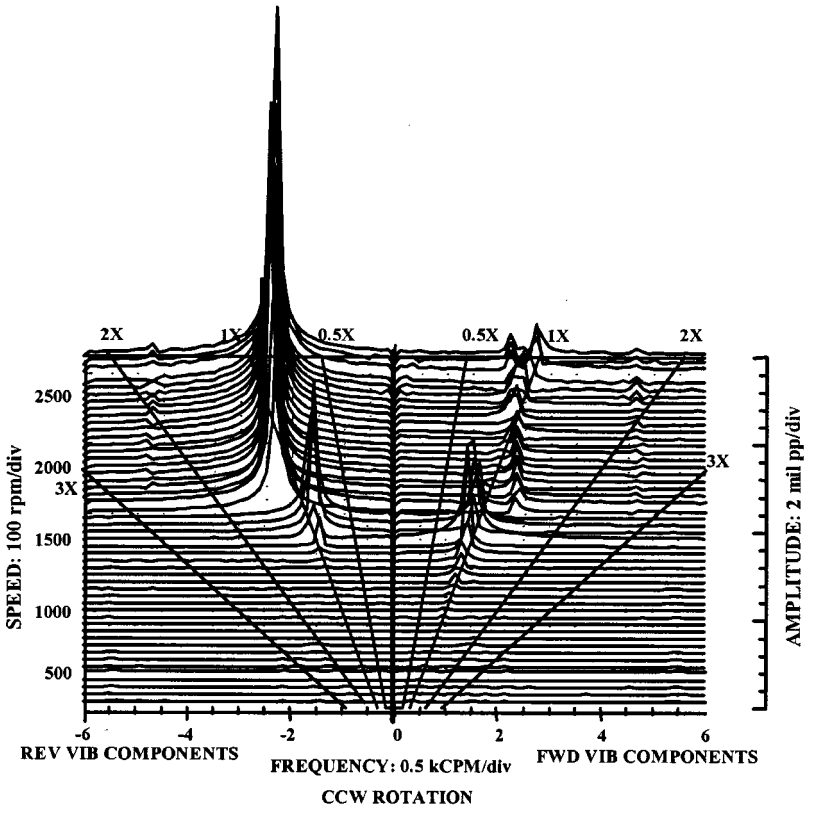
The ratio of reverse rub frequency to speed $\omega/\Omega$ was always less than the ratio of radius to radial clearance $r/C_r$ (=13.3 in this case). Assuming that the seal vibrates harmonically at the same frequency $\omega$, the relative slip velocity ([5]) can be given by

$$\Delta v = r\Omega - C_r\omega = C_r\Omega\left(\frac{r}{C_r} - \frac{\omega}{\Omega}\right). \tag{1}$$

Figure 8 shows the slippage that occurred in Fig. 7. It was found that the relief of reverse rub occurred near zero slippage velocity where the friction force was approaching zero and could not balance the damping force. Obviously, the slip velocity was proportional to the rotative speed, thus yielding more wear and heat between the rotor and the seal at higher speed. When the slippage could not be maintained as $r/C_r$ approached the ratio $\omega/\Omega$ at low speed, as shown in Fig. 8(*b*), the rotor disengaged from the seal and the reverse rub then ceased.




(*a*) **Rotor lateral response and orbits**




(*b*) **Rotor full spectrum response based on both X and Y probe data**

**Fig. 4 Process of generating reverse precessional rub (self-excited vibration, also called "dry whip") without any outside disturbance. Two-disk rotor/Teflon seal diametral clearance of 500 $\mu$m with mass unbalance of 0.5 grams.**

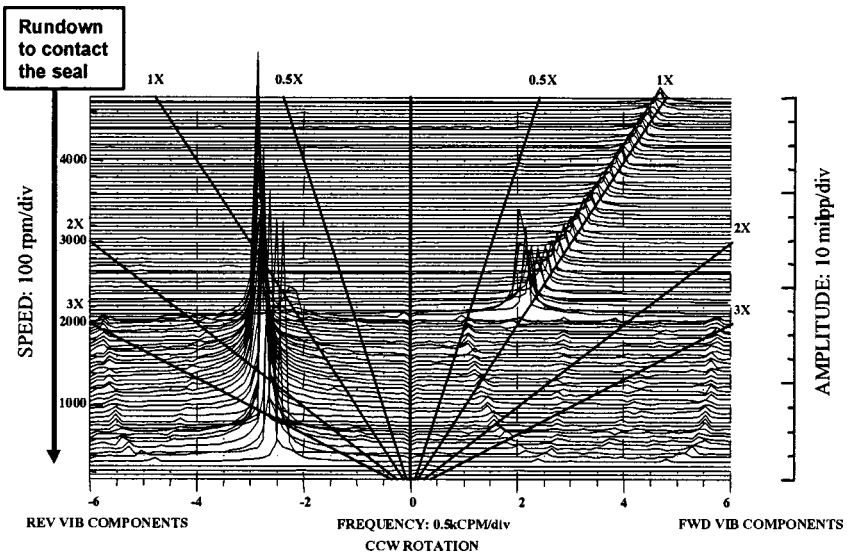



**Fig. 5 Reverse precesional rub triggered during rundown. One-disk rotor/Teflon seal diametral clearance of 1000 $\mu$m with mass unbalance of 1.1 grams.**

## 4 Analytical Study

**4.1 Mathematical Model.** A simplified mathematical model of the rotor/seal system is shown in Fig. 9, which reflects a contact through the rub-related normal force and dry friction when the rotor response amplitude $|\rho|$ is higher than the radial clearance $C_r$. The equation of motion can be given by

$$M\ddot{\rho}+D\dot{\rho}+K\rho=Mh\Omega^2e^{j\varphi}-\frac{K_s\rho}{2}(1+jf)\left[\left(1-\frac{C_r}{|\rho|}\right)+\left|1-\frac{C_r}{|\rho|}\right|\right] \tag{2}$$

where $\rho=x+jy$, $j=\sqrt{-1}$, $M$, $D$, $K$ are modal mass, damping, and stiffness of the rotor system without rub, $h$ is a center of mass eccentricity, $\varphi=\Omega t$, $K_s$ is a stiffness added to the system due to the contact with the stator, and $f$ is a friction coefficient between the rotor and the seal. Equation (2) can be rewritten in nondimensional form as

$$w''+2\varsigma w'+w=an^2e^{j\varphi}-\frac{pw}{2}(1+jf)\left[\left(1-\frac{1}{|w|}\right)+\left|1-\frac{1}{|w|}\right|\right] \tag{3}$$

where

$$w=r/C_r,\nu=\sqrt{\frac{K}{M}},\tau=\nu t,w'=\frac{dw}{d\tau},w''=\frac{d^2w}{d\tau^2},$$

$$\varsigma=\frac{D}{2\sqrt{KM}},a=h/C_r,n=\Omega/\nu,p=K_s/K.$$

**4.2 Synchronous Response and Stability.** Equation (3) allows for the following synchronous response in the noncontact case:

$$w=\frac{an^2}{1-n^2+2j\varsigma n}e^{jn\tau},\quad \text{if } |w|=\frac{an^2}{\sqrt{(1-n^2)^2+4\varsigma^2n^2}}\leqslant 1. \tag{4}$$

Obviously, if $a<2\varsigma\sqrt{1-\varsigma^2}$, the contact will never occur. If $a>1$, the contact will be reached at the rotative speed below the natural frequency, and will never cease.

In the case of contact between the rotor and the seal, assume synchronous response in the form

$$w=Ae^{j(n\tau+\alpha)}. \tag{5}$$

Substitution of Eq. (5) into Eq. (3) yields the following solution:

$$A=\frac{1}{S}\left[p\frac{\cos(\delta+\beta)}{\cos\delta}\pm\sqrt{a^2n^4-p^2\frac{\sin^2(\delta+\beta)}{\cos^2\delta}}\right] \tag{6}$$

$$\tan\alpha=\frac{AS\sin\beta+p\tan\delta}{AS\cos\beta-p}$$

where

$$S=\sqrt{(1-n^2)^2+4\varsigma^2n^2+p^2/\cos^2\delta+2p[(1-n^2)+2\varsigma n\tan\delta]}$$

$$\tan\beta=-\frac{2\varsigma n+p\tan\delta}{1-n^2+p},\quad \tan\delta=f.$$

The second solution with the sign "$-$" is unstable (as it can be shown later based on the characteristic equation).

The stability of Eq. (6) needs to be examined by introducing a small perturbation $w_1(\tau)$ as follows:

$$w=Ae^{j(n\tau+\alpha)}+w_1. \tag{7}$$

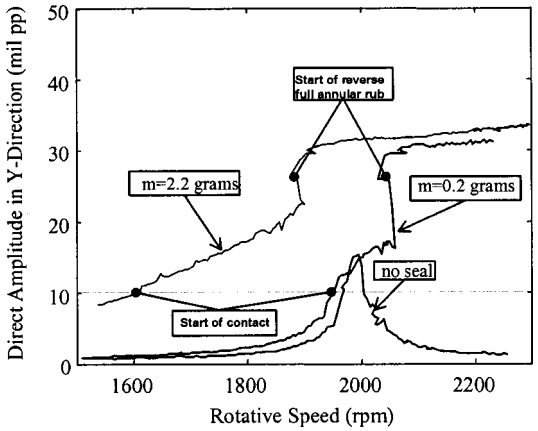
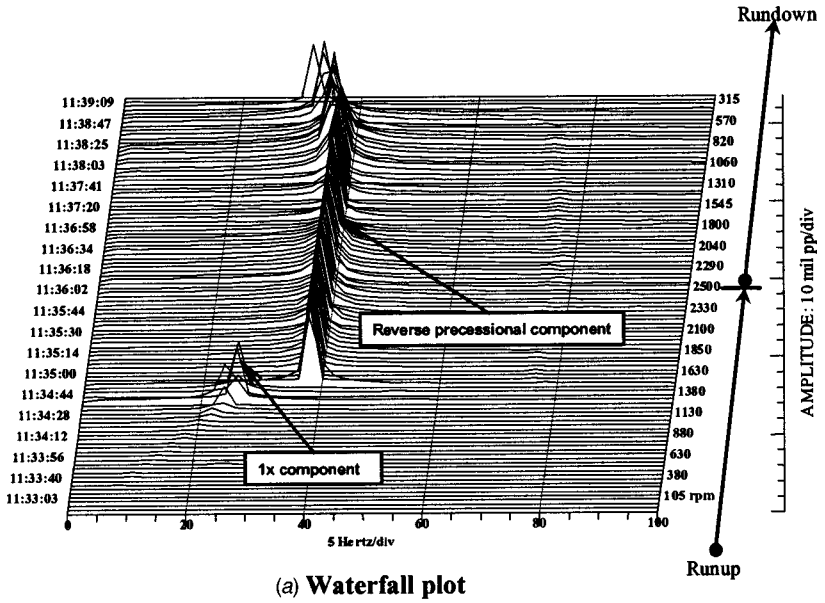
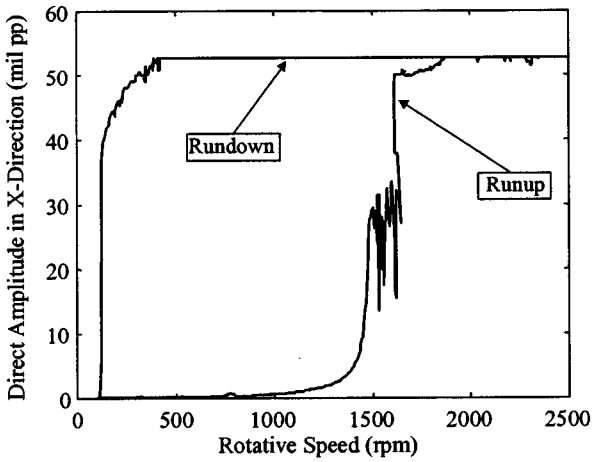


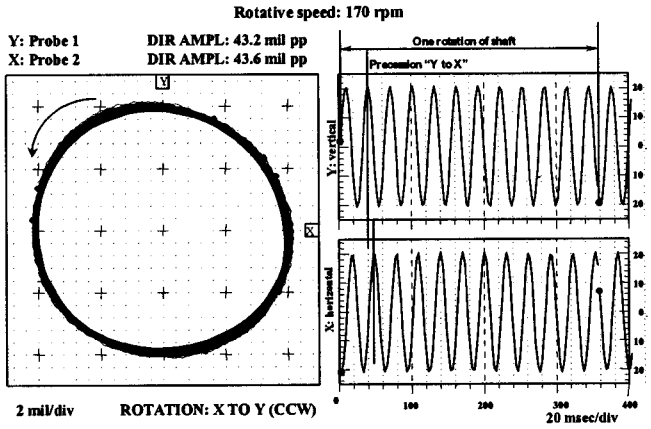

**Fig. 6 Effect of mass unbalance on the starting point where reverse full annular rub occurred without any outside disturbance. One disk rotor, Teflon seal diametral clearance of 250 $\mu$m.**




(*a*) **Waterfall plot**




(*b*) **Direct response**




(*c*) **Reverse rub at rotative speed of 170 rpm**

**Fig. 7 Reverse precessional rub versus time and speed during the whole running process including runup and rundown. One disk rotor, Teflon seal diametral clearance of 750 $\mu$m.**

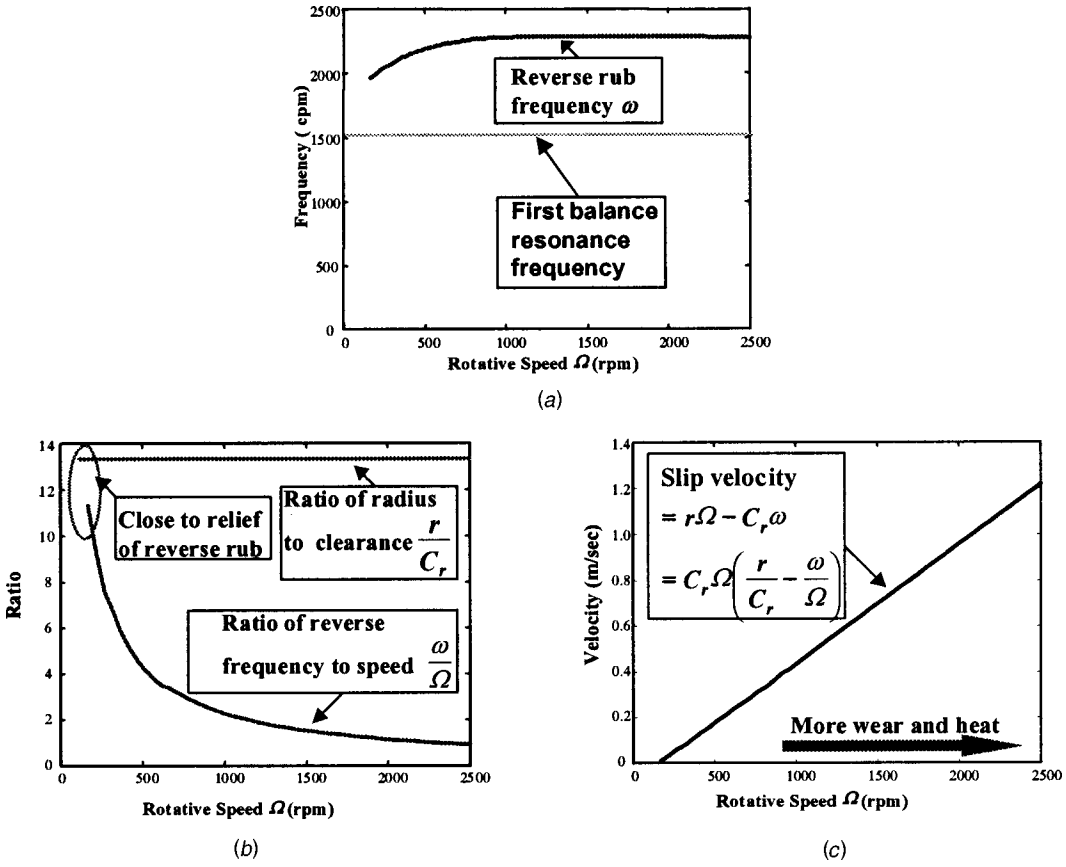



**Fig. 8 Parameters of slippage of rotor against seal with changes in rotative speed; (*a*) reverse rub frequency, (*b*) ratio of reverse frequency to speed, and (*c*) slip velocity**

Assuming that the rotor/seal contact is maintained all the time, the linearized equation for a small perturbation can be given by

$$w_1'' + 2\varsigma w_1' + \left[1 + \frac{p}{\cos\delta}\left(1 - \frac{1}{2A}\right)e^{j\delta}\right]w_1 + \frac{p}{2A\cos\delta}\overline{w}_1 e^{j(2\varphi+2\alpha+\delta)} = 0 \tag{8}$$

where $\overline{w}_1$ is complex conjugate of $w_1$. To avoid solving differential Eq. (8) with periodic coefficients, the perturbation $w_1$ can be expressed in the rotating frame as

$$w_1 = (\xi + j\eta)e^{j(\varphi+\alpha)}. \tag{9}$$

The eigenvalue solution can be written as

$$\xi = C_1 e^{\chi t}, \quad \eta = C_2 e^{\chi t} \tag{10}$$

where $\chi$ is an eigenvalue, $C_1$, $C_2$ are components of an eigenvector. A substitution of Eqs. (9) and (10) into Eq. (8) yields the following characteristic equation:

$$(\chi+\varsigma)^4 + A_2(\chi+\varsigma)^2 + A_1(\chi+\varsigma) + A_0 = 0 \tag{11}$$

where

$$A_2 = 2(1+n^2-\varsigma^2+p\Phi); \quad A_1 = 4npf\Phi;$$

$$A_0 = (1-n^2-\varsigma^2)^2 + 2p\Phi(1-n^2-\varsigma^2) - p^2(1+f^2)(1-2\Phi);$$

$$\Phi = 1 - \frac{1}{2A}.$$

For the case of contact-noncontact situation, a characteristic equation can be generalized from Eq. (11) by assuming

$$p = \begin{cases} 0, & \text{if } |w| \le 1; \\ p, & \text{if } |w| > 1. \end{cases}$$

Corresponding to the synchronous response given by Eq. (6), root locus plots based on Eq. (11) can be created. Figures 10–11 present these plots calculated for the case of damping factor $\varsigma = 0.025$, unbalance ratio $a = h/C_r = 0.2$, stiffness parameter $p = K_s/K = 0.5$. They differ only by the value of the dry friction coefficient $f$. Figure 10 presents a root locus plot in case of low friction ($f = 0.1$). It shows a situation when the synchronous forward precession, after establishing full annular rub, remains stable until the rotor jumps out of the contact against the seal. The unrealized branch of amplitude-speed curve (the second solution in Eq. (6)) is unstable due to the real root turning positive. Figure 11 presents a case of $f = 0.2$, when synchronous forward precession is stable at the beginning of rub, then becomes unstable and then stable again with increased rotative speed. Note that after the con-

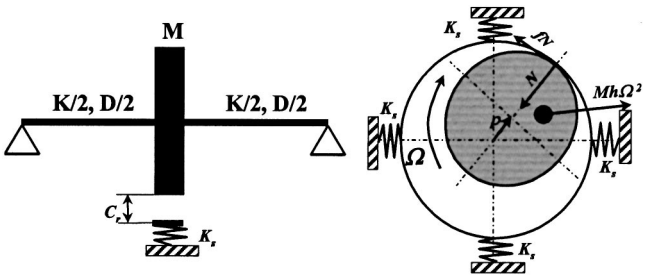



**Fig. 9 Diagram of the rotor/seal system with full annular rub**

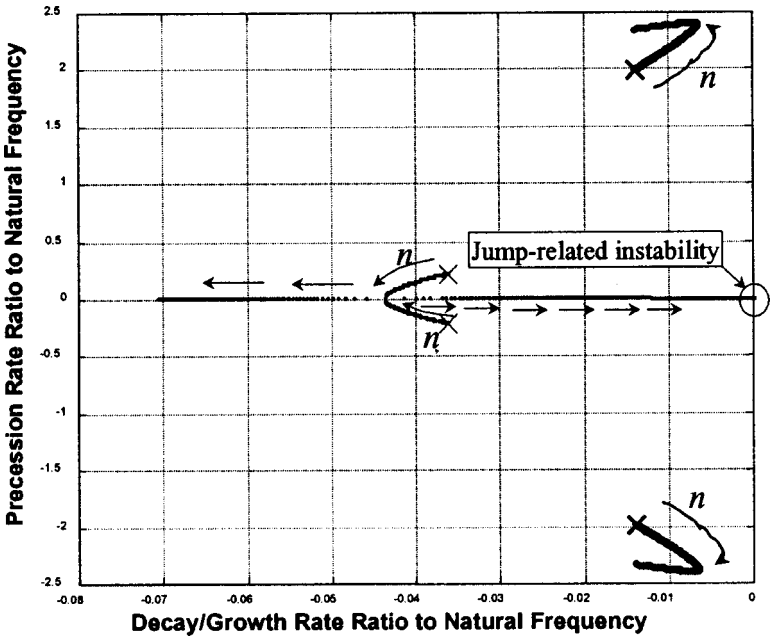



**Fig. 10 Root locus of synchronous response under full annular rub conditions. $a=h/C_r=0.2$, $\varsigma=0.025$, $K_s/K=0.5$. Note that for low value of friction coefficient, $f=0.1$, the synchronous response is always stable.**

tact between rotor and stator takes place, the rotor synchronous response is stable, and then it becomes unstable. Figure 12 depicts the case of high friction ($f=0.3$), when the synchronous response is unstable from the start of rubbing. Figure 13 shows how the stability of synchronous response changes with unbalance parameter $a$. The boundary of instability region does not change with $a$, but crossings of the synchronous response into instability zone are affected.

Four analytically found roots are reflected on the root locus plots in Figs. 10–12, with modifications due to friction. In general two roots are complex conjugated at all rotative speeds. These two roots can become unstable due to friction. This instability leads to the reverse precessional rub observed in the experimental study. The other two roots start as complex conjugated and then turn real. Those two roots are not strongly affected by the change in friction. That instability simply removes the possibility of the second solution in Eq. (6).

In order to investigate the stability boundaries of synchronous response, the eigenvalue $\chi$ is assumed purely imaginary ($\chi=js$). In this case Eq. (11) is equivalent to the following two equations:

$$\begin{aligned}(1+n^2-s^2+p\Phi)^2-4\varsigma^2s^2+4n\varsigma pf\Phi-4n^2(1-\varsigma^2+p\Phi)\\ =p^2[\Phi^2+(1+f^2)(1-2\Phi)]\\ s\left(1+n^2-s^2+p\Phi+\frac{f}{\varsigma}np\Phi\right)=0.\end{aligned} \tag{12}$$

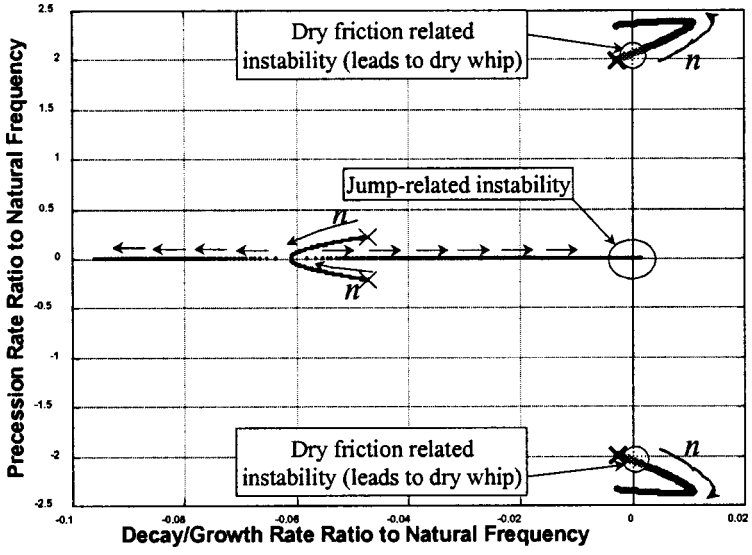



**Fig. 11 Root locus of synchronous response stability under full annular rub conditions. $a=h/C_r=0.2$, $\varsigma=0.025$, $K_s/K=0.5$, $f=0.15$. Note that with rotative speed increase the synchronous response experiences transition from being stable to unstable then to stable again.**

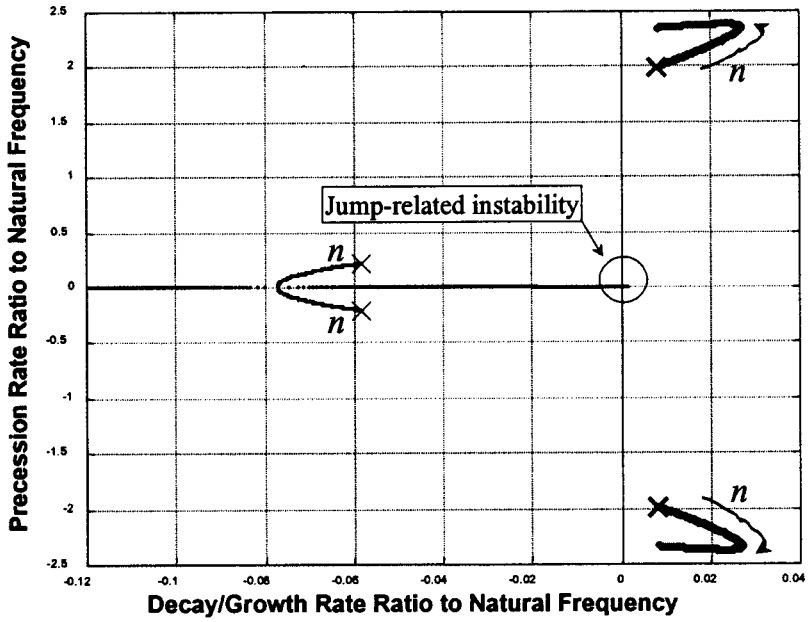



**Fig. 12 Root locus of synchronous response under full annular rub conditions. $a=h/C_r=0.2$, $\varsigma=0.025$, $K_s/K=0.5$. Note that for high value of friction coefficient, $f=0.3$, the synchronous regime is always unstable.**

Obviously, $s=0$ and $s>0$ satisfy Eq. (12), and yield the boundaries of jump and dry friction related instability regions, respectively.

For $s=0$, its saddle-type instability boundary is given in terms of rotative speed to natural frequency ratio $n_{stj}$ by

$$n_{stj}^4-2n_{stj}^2(1+p\Phi-2\varsigma^2)+4n_{stj}fp\varsigma\Phi+1 +2p\Phi-p^2(1+f^2)(1-2\Phi)=0 \tag{13}$$

where $\Phi=1-1/(2A)$. The resulting curve of $A(n_{stj})$ is shown in Fig. 13 where its encompassed area is marked with jump-related instability region. This zero real root determines the boundary between stable and unstable branches of synchronous full annular rub response. During runup when the response reaches to this curve, the rotor will jump down to end rubbing. During rundown it will jump up, and its rub response will follow the first branch (the first solution in Eq. (6)) once contacting the seal. Even if the rotor is being held at a point along this curve, its rub response will surely follow the first branch with decrease in rotative speed.

For $s>0$, its focus type instability boundary is given by

$$n_{stf}^2=\varsigma^2\frac{4\varsigma^2(1+p\Phi)+p^2[\Phi^2+(1+f^2)(1-2\Phi)]}{p^2f^2\Phi^2-4\varsigma^2(1+p\Phi)} \qquad s^2=1+p\Phi+n_{stf}^2+n_{stf}p\Phi\frac{f}{\varsigma} \tag{14}$$

where $\Phi=1-1/(2A)$. Higher friction coefficient $f$ corresponds to a larger dry friction related instability region and makes lower amplitude response go unstable. $f=0.3$ in Fig. 12, for instance, makes the whole contact region unstable, thus making the boundary $A(n_{stf})$ less than 1 according to Eq. (14). When $f$ is increased to 0.15, the instability boundary $A(n_{stf})$ becomes more than 1, as

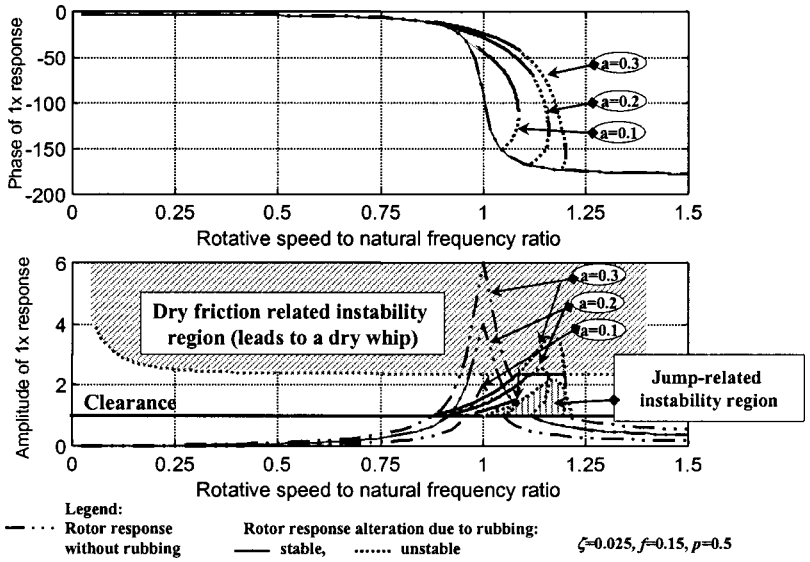



**Fig. 13 Synchronous response for different values of unbalance parameter *a* with and without rub and its stability**

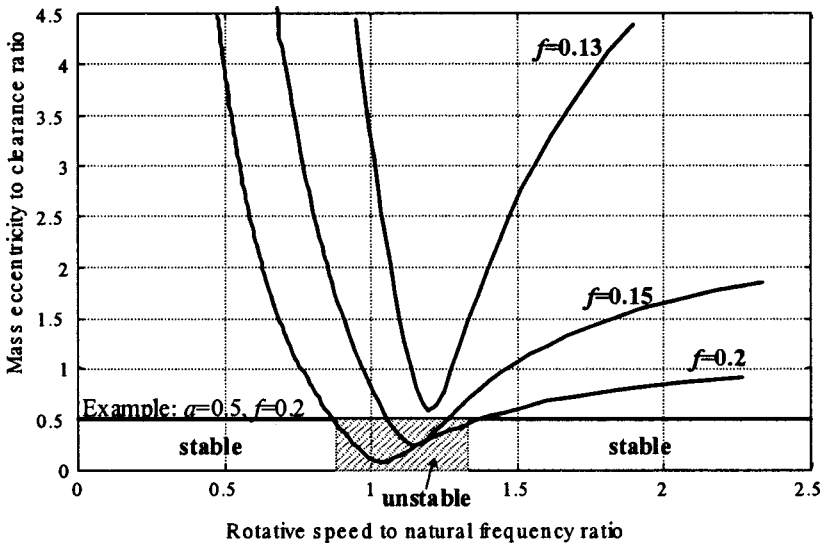



**Fig. 14 Synchronous response instability boundaries for different values of $f$ ($\varsigma$=0.025, $p$=0.5). Instability leads to self-excited reverse precession. Note destabilizing effect with increase of $f$.**

shown in Fig. 13. Considering the case where the boundary is located within $1 \leq A < \infty$, the positive right-side term of the first expression in Eq. (14) requires

$$f > 2\frac{\varsigma}{p}\sqrt{1+p}. \tag{15}$$

This inequality is a necessary condition for dry-friction-related instability to occur, as it can be easily confirmed in all the cases shown in Figs. 10–13.

The dry friction related instability boundary can also be expressed in terms of unbalance parameter $a$. The two crossing points, as shown in Fig. 13, can be determined from Eqs. (6) and (14) at a given value $a$. Figure 14 shows that increase of $a$ makes instability region larger for a given friction $f$. It is important to note that damping $\varsigma$ parameter plays a stabilizing role, while friction coefficient $f$ and stiffness ratio $p$ have destabilizing effect. For different values of $\varsigma$ and $p$, similar figures as in Fig. 14 are expected. Qualitatively, the increase of parameters $f$ and $p$, as well as mass eccentricity ratio $a$, leads to the increase of the friction force in the direction opposite to the rotation direction. That contributes to the destabilization of the synchronous response that has a forward direction. From Eq. (15), which describes a necessary condition of the reverse self-excited precession known as a "dry whip," minimum required friction coefficient $f$ increases with the rotor damping and decreases with stiffening of the seal. This result shows the stabilizing nature of damping and destabilizing effect of the seal stiffness.

**4.3 Reverse Precession and Stability.** Since mass unbalance has no effect on maintaining the reverse precessional rub, Eq. (3) can be simplified into

$$w'' + 2\varsigma w' + w + pw(1+jf)\left(1 - \frac{1}{|w|}\right) = 0. \tag{16}$$

Assume a particular solution

$$w = Be^{j(\omega\tau+\beta)} \tag{17}$$

where $B$, $\omega$, can be determined from Eq. (16), while $\beta$ is arbitrary. Two solutions satisfy Eq. (16). One, however, shows a forward precession and has an amplitude less than 1. The other meaningful reverse precession solution is as follows:

$$\omega = -\left(\frac{\varsigma}{f} + \sqrt{1+\frac{\varsigma^2}{f^2}}\right), \quad B = \frac{p}{\left|p - 2\frac{\varsigma^2}{f^2}\left(1+\sqrt{1+\frac{f^2}{\varsigma^2}}\right)\right|}. \tag{18}$$

In order to investigate the stability of this self-excited solution, a linearized equation for a small perturbation must be considered. Following the procedure described in the previous analysis for synchronous rub, a linearized perturbation equation in the rotating frame with precession frequency $\omega$ can be derived by replacing $n$ with $\omega$. The characteristic equation is then given by

$$(\chi+\varsigma)^4 + A_2(\chi+\varsigma)^2 + A_1(\chi+\varsigma) + A_0 = 0 \tag{19}$$

where

$$A_2 = 2(1+\omega^2-\varsigma^2+p\Phi); \quad A_1 = 4\omega p f \Phi;$$

$$A_0 = (1-\omega^2-\varsigma^2)^2 + 2p\Phi(1-n^2-\varsigma^2) - p^2(1+f^2)(1-2\Phi);$$

$$\Phi = 1 - \frac{1}{2B}.$$

Substitution of Eq. (18) into Eq. (19) yields

$$\chi\{\chi^3 + 4\varsigma\chi^2 + (4\varsigma^2+3\omega^2+p+1)\chi + 2\varsigma(p+1+3\omega^2) + 2f\omega(p-1+\omega^2)\} = 0. \tag{20}$$

The zero root in Eq. (20) is caused by system autonomy and does not affect stability. Substitution of $\chi = js$ in Eq. (20) yields the following stability boundary:

$$\begin{gathered} s(s^2 - 4\varsigma^2 - 3\omega^2 - p - 1) = 0 \\ \frac{\varsigma}{f}(2s^2 - 3\omega^2 - p - 1) = \omega(p - 1 + \omega^2). \end{gathered} \tag{21}$$

Substitution of Eq. (18) into Eq. (21) yields

$$\frac{\varsigma}{f} = \frac{p}{2\sqrt{1+p}}. \tag{22}$$

Equation (22) determines the stability boundary of the self-excited reverse precession region. Obviously, it is stable if

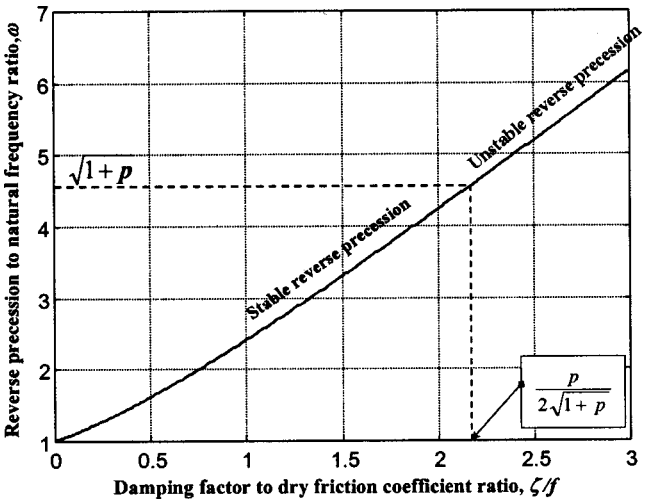



**Fig. 15 Reverse precessional full annular rub frequency**

$$\frac{\varsigma}{f} < \frac{p}{2\sqrt{1+p}}. \tag{23}$$

It coincides with the inequality (15), which determines the necessary conditions for the full annular rub synchronous response becoming unstable. Substitution of Eq. (18) into Eq. (23) shows

$$|\omega| < \sqrt{1+p}, \quad \text{or} \quad \nu|\omega| < \sqrt{\frac{K+K_s}{M}} \tag{24}$$

Equations (18) and (24) show that self-excited reverse precessional frequency is always higher than the natural frequency of the rotor without rub and lower than the natural frequency of the rotor with added seal stiffness (see Fig. 15).

## 5 Conclusions

Based on experimental and analytical studies presented in this paper, the following conclusions are stated concerning rotor/seal full annular rub:

1 Reverse precession (dry whip) can occur without any outside disturbance for small rotor/seal clearances. It is triggered near resonance speed during runup or rundown after experiencing unstable synchronous rub. Its occurrence is enhanced with higher mass unbalance for a given rotor/seal system.

2 Once generated, the reverse precession can be sustained with almost the same frequency and amplitude over the whole speed range until at very low speed when rotor/seal slippage cannot be maintained. It can also be generated or released with a hammer impact, independent of mass unbalance. Its frequency is between the natural frequencies of rotor without and with the seal.

3 Synchronous response involves jump phenomena. During runup the rotor response jumps down when it disengages from the seal, while during rundown the response jumps up when it starts to contact the seal. Stability analysis yields a jump-related instability zone that is related to an always-unstable branch, thus determining a unique synchronous solution.

4 Instability boundary leading to the self-excited reverse precession is given by $f=2\varsigma\sqrt{1+p}/p$ where $f$ is friction coefficient, $\varsigma$ is rotor damping ratio, and $p$ is seal-to-rotor stiffness ratio. Case $f<2\varsigma\sqrt{1+p}/p$ corresponds to stable synchronous response for any amount of mass unbalance. Case $f>2\varsigma\sqrt{1+p}/p$ indicates stable reverse precession. In this case, synchronous response can become unstable and lead to reverse recession for above certain amount of mass unbalance.

5 $f<2\varsigma\sqrt{1+p}/p$ can be used as a criterion for choosing a seal for a rotor system, thus excluding the possibility of severe reverse precession.

## References


[1] Bently, D. E., 1974, "Forced Subrotative Speed Dynamic Action of Rotating Machinery," ASME Paper No. 74-PET-16.

[2] Bently, D. E., Grant, J. W., and Goldman, P., 1992, "A 'Butterfly' Rub Response," Bently Rotor Dynamics Research Corporation, BRDRC Report 3.

[3] Muszynska, A., 1984, "Rotor/Seal Full Annular Rub," Senior Mechanical Engineering Seminar, Bently Nevada Corp., Carson City, NV.

[4] Lawen, Jr., J. L., and Flowers, G. T., 1999, "Interaction Dynamics Between a Flexible Rotor and an Auxiliary Clearance Bearing," ASME J. Vibr. Acoust., **121**, pp. 183–189.

[5] Yu, J. J., Muszynska, A., and Bently, D. E., 1998, "Dynamic Behavior of Rotor With Full Annular Rub," Bently Rotor Dynamics Research Corporation, BRDRC Report 7.

[6] Goldman, P., and Bently, D. E., 1998, "Analytical Study of Full Annular Rub," Bently Rotor Dynamics Research Corporation, BRDRC Report 6.

[7] Muszynska, A., 1989, "Rotor-to-Stationary Element Rub-Related Vibration Phenomena in Rotating Machinery, Literature Survey," Shock Vib. Dig., **21**, pp. 3–11.

[8] Black, H. F., 1968, "Interaction of a Whirling Rotor with a Vibrating Stator Across a Clearance Annulus," J. Mech. Eng. Sci., **10**, pp. 1–12.

[9] Ehrich, F. F., 1969, "The Dynamics Stability of Rotor/Stator Radial Rubs in Rotating Machinery," J. Eng. Ind., pp. 1025–1028.

[10] Lingener, A., 1990, "Experimental Investigation of Reverse Whirl of a Flexible Rotor," *IFToMM Third International Conference on Rotordynamics*, Lyon, France, pp. 13–18.

[11] Crandall, S., 1990, "From Whirl to Whip in Rotordynamics," *IFToMM Third International Conference on Rotordynamics, Lyon*, France, pp. 19–26.

[12] Childs, D., 1993, *Turbomachinery Rotordynamics: Phenomena, Modeling, and Analysis*, John Wiley and Sons, New York.

# Reconciliation of Rotordynamic Models With Experimental Data

**E. H. Maslen**
e-mail: ehm7s@virginia.edu

**J. A. Vázquez**
e-mail: jose@virginia.edu

Department of Mechanical and Aerospace Engineering,
University of Virginia,
Charlottesville, VA 22903

**C. K. Sortore**
Synchrony, Inc.,
7777 Bent Mountain Road,
Roanoke, VA 24018
e-mail: Chris.Sorte@Synchrony.com

*A computationally efficient strategy is presented for adjusting analytic rotordynamic models to make them consistent with experimental data. The approach permits use of conventional rotordynamic models derived using finite element methods in conjunction with conventional plant identification models derived from impact or sine sweep testing in a transfer function or influence coefficient format. The underlying assumption is that the predominant uncertainties in engineered models occur at discrete points as effects like shrink fits, seal coefficients or foundation interactions. Further, it is assumed that these unmodeled or poorly modeled effects are essentially linear (at least within the testing and expected operating domains). Matching is accomplished by deriving a dynamic model for these uncertain effects such that the resulting composite model has a transfer function which matches that obtained experimentally. The derived augmentations are computationally compatible with the original rotor model and valid for stability or forced response predictions. Further, computation of this augmentation is accomplished using well developed and widely disseminated tools for modern control. Background theory and a complete recipe for the solution are supported by a number of examples.*
[DOI: 10.1115/1.1416692]

## Introduction

Rotating machine analyses are based on engineering models. These models are assembled from mathematical models of various components and physical effects such as the rotor, bearings, seals, bearing support structure, rotor couplings, aerodynamic cross-coupling, fluid-rotor interaction, etc. The accuracy of the model of the machine, therefore, depends on accuracy of the models of the individual components and of the models of the physical effects acting on the rotor. Unfortunately, not all the components and physical effects acting on a rotor can be modeled accurately or identified precisely. The accuracy of a model is determined by comparing the response of the model and the response of the real machine to the same excitation.

To improve the accuracy of the machine models, work has been carried out on improving modeling techniques of components like the rotor, bearing, and seals. In the case of rotors, experimental methods have been used to improve the quality of the models by testing the rotor, extracting the modal parameters and then modifying the model of the rotor until the modal parameters of the model match the experimental data. In the case of components where the modeling is difficult or very expensive, techniques have been developed to extract models from experimental data or to use the experimental data directly in the analysis.

The process of extracting a model from experimental data is called system identification. Modal identification is a special case of this technique. All the current literature regarding system identification presents a means of constructing a model whose outputs match the measured response for a known set of inputs. The ''identified'' model is typically presented as a transfer function. This approach generally makes no reference to any *a priori* model of the system.

This paper presents a method that bridges the gap between the well-developed modeling techniques in rotating machinery and the powerful identification techniques available. The end result is that the engineer can use the modeling techniques for those components and effects that are well known and then use experimental data to modify (reconcile) the model at those locations where error is entering the model because of unidentified effects or because the modeling techniques cannot provide enough accuracy ([1]).

The literature in the fields of component modeling in rotating machinery and system identification is very extensive. Lund [2] presented a method to calculate unbalance response and stability analysis of a rotor on flexible supports. This method is now known as the transfer matrix method. Nelson and McVaugh [3] developed a finite element formulation for rotor-bearing systems. Vance et al. [4] present the idea of using experimental modal data in the refinement of models of the rotors of rotating machinery. The texts by Vance [5], Ehrich [6], and Childs [7] present some of the most common methods to write the equations of motion of rotor-bearing systems. The text by Ljung [8] includes a review of system identification techniques. The texts by Ewins [9] and Maia et al. [10] include a review of the modal identification techniques most commonly used.

Direct system identification in rotating machinery has been applied to the bearing support structure and casing. The dynamic behavior of the support structure is measured in the form of frequency response functions (FRF) measured at the bearing locations. This information is later used to include the flexible supports effects into rotordynamic calculations. Barrett et al. [11] and Nicholas et al. [12] used FRF data to produce equivalent single mass support models. The bearing coefficients were modified to include the effects of the supports. Redmond [13,14] used frequency response data of the support structure to produce single mass support models including the effects of cross-talk between supports. The author showed that substantial errors could arise if the experimental data was taken with the rotor installed in the machine and proposed a method to subtract the rotor influence from the experimental data. Rouch et al. [15] and Stephenson et al. [16] used FRF to create equivalent finite elements models of the support structure. Vázquez and Barrett [17] showed the effects of flexible supports and casings of rotating machinery can be included in the rotordynamic analyses with polynomial transfer functions calculated from experimental data. Vázquez [18] and Vázquez et al. [19,20] used this technique to include the support structure of an experimental flexible rotor supported by fluid film bearings on flexible bearing supports. The polynomial transfer functions were extracted from FRF of the support structure mea-

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-395. Manuscript received by IGTI, Nov. 1999; final revision received by ASME Headquarters, Feb. 2000. Associate Editor: D. Wisler.

sured at the bearing locations using an identification technique introduced by Sanathanan and Koerner [21] and later applied to magnetic bearings by Gähler and Herzog [22].

The work on the bearing support structures has shown that the bearing support structure can be represented as transfer functions acting at the bearings of rotating machinery. This is an important conclusion. In the field of modern control, these transfer functions are represented as local feedback mechanisms acting on the plant (rotor-bearing model). The importance of this representation is that there are formal mathematical tools that permit the synthesis of transfer functions of local feedback mechanisms from experimental data. Therefore, we can use the synthesis tools used in modern controls to identify effects acting in rotor models without having to measure these effects independently.

Parallel work in the area of modern controls includes Wolodkin et al. [23] and Fisher et al. [24]. Wolodkin et al. showed that local feedback can be used successfully for the identification of the parameters (elements) of a constant matrix. Fisher et al. [24] used local feedback for the refinement of the model obtained during the identification of high-order lightly damped systems. The authors obtained a transfer function from frequency response data and then used local feedback to refine the model with time-domain data as it became available. We use a similar idea on this work but we start with a mathematical model developed using finite elements. Also, the corrections to the mathematical model are evaluated at discrete points. In this manner, it is possible to keep the mathematical model in physical coordinates and only modify the parts of the model that are expected to contribute error to the model.

This work presents a method that will allow the automation of model reconciliation by making use of existing mathematical models. The system is proposed as a nominal model with uncertainties at known locations. The idea is to identify the transfer functions of the mechanisms acting at the uncertainty locations such that the response of the mathematical model matches the experimental data.

## Model Reconciliation

The reconciliation method starts with a nominal engineering model of a rotating machine and experimental data measured at some known locations. The purpose of the method is to modify the engineering model such that the output of the model matches the experimental data. The assumption is that the error in the engineering model occurs at discrete points and/or components. These components could be aerodynamic effects, seal effects, or interference fits between disks and journal—these are features or conditions which may not have an accurate accounting within the model and are usually difficult to identify precisely. These effects are represented by a local feedback mechanism as shown in Fig. 1. The "controller" $\mathbf{K}(s)$ is the correction included in the nominal engineering model. It can be interpreted as complex frequency-dependent stiffness or transfer functions acting at discrete points in the model (these points are selected by the engineer).

## State-Space Model

The equations of motion of the engineering model will be written using a state-space representation. This representation of the equations of motion is assumed by the modern control tools that will be used in the reconciliation process. The transformation of the equations of motion of a rotor-bearing system from the standard **M**, **C**, **K** representation into the state-space representation is straightforward and it is presented by Maslen and Bielk [25].

The state-space equations for the nominal rotor model with the addition of some uncharacterized disturbance forces can be expressed as

$$\dot{z} = \mathbf{A}z + \mathbf{B}_c u_c + \mathbf{B}_k f_k \tag{1}$$

$$y_m = \mathbf{C}_m z \tag{2}$$

$$y_k = \mathbf{C}_k z. \tag{3}$$

The vector $f_k$ represents the suspected uncharacterized disturbance forces (i.e., uncertainties) and $B_k$ is a model of how they enter the model ($B_k$ is a matrix with the location of the uncertainties). The vector $y_k$ is the response at the disturbance locations. In the present method, the response at these locations and the disturbances are related through (presently unknown) feedback as

$$f_k = \mathbf{K}(s) y_k. \tag{4}$$

The model reconciliation strategy is to compute a dynamic system, $\mathbf{K}(s)$, which minimizes the error between the actual (identified) response and the corresponding response computed by the nominal model at the same locations. Figure 2 illustrates the construction of the identification problem.

The state-space model for this system is

$$\begin{Bmatrix} \dot{\hat{z}} \\ \dot{z} \end{Bmatrix} = \begin{bmatrix} \hat{\mathbf{A}} & 0 \\ 0 & \mathbf{A} \end{bmatrix} \begin{Bmatrix} \hat{z} \\ z \end{Bmatrix} + \begin{bmatrix} \hat{\mathbf{B}}_c \\ \mathbf{B}_c \end{bmatrix} u_c + \begin{bmatrix} 0 \\ \mathbf{B}_k \end{bmatrix} f_k \tag{5}$$

$$e_{ym} = [\hat{\mathbf{C}}_m - \mathbf{C}_m] \begin{Bmatrix} \hat{z} \\ z \end{Bmatrix} + \varepsilon \mathbf{D}_{12} f_k \tag{6}$$

$$y_k = [0 \quad \mathbf{C}_k] \begin{Bmatrix} \hat{z} \\ z \end{Bmatrix} + \varepsilon \mathbf{D}_{21} u_c \tag{7}$$

$$f_k = \mathbf{K}(s) y_k. \tag{8}$$

The constant matrices $\hat{\mathbf{A}}$, $\hat{\mathbf{B}}_c$, and $\hat{\mathbf{C}}_m$ represent the state-space realization of the identified rotor. This state-space realization is calculated from the experimental FRF (Gähler and Herzog [22]). $\mathbf{D}_{12}$ and $\mathbf{D}_{21}$ are full-rank matrices and $\varepsilon$ is a very small number. These values do not affect the results but are needed to avoid numerical problems.

Any suitable method for designing the feedback controller may be used: $\mathcal{H}_\infty$, $\mathcal{H}_2$, etc. ([26]). The identification process is a matter of identifying parameters (sources of model error), constructing the error supermodel (Fig. 2) and then turning the "crank" of a chosen controller design algorithm. If the controller is designed properly, the reconciled model will be matched to the measured

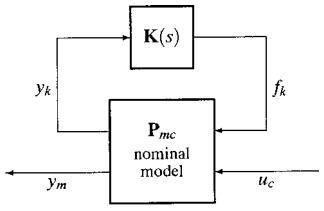



**Fig. 1 Reconciled model representation**

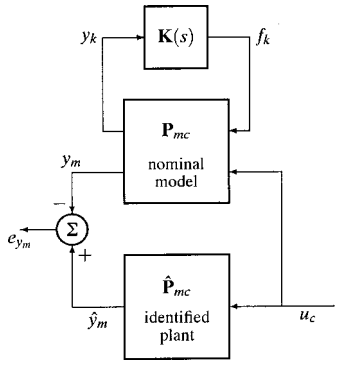



**Fig. 2 Model identification block diagram**

data in the sense that, when the original test signals are applied to this modified model, its output will match the original data.

If the uncertainty locations are chosen appropriately, then it is expected that the internal states of the system will also be matched to the measured plant, not just the output measurements. This method does not ignore (as other currently available identification methods do) but takes advantage of existing well-developed models of the rotor, bearings, seals, etc. If the particular choice of uncertainty locations was chosen poorly (e.g., assuming the wrong location of a disturbance effect), the designed controller, $\mathbf{K}(s)$, may still force a match between the response of the modeled and measured plants. In this case, however, it is less likely that the subsequent use of this reconciled model to construct transfer functions to and from other points (points other than that which were used to perform the measurement) will be accurate.

## Steps in the Reconciliation Process

1 Measure FRFs of the actual rotor-bearing system.
2 Identify a dynamic system from FRF (state-space model for the transfer functions).
3 Construct a nominal engineering model (rotor, bearing, seals, etc.).
4 Identify the points of likely modeling uncertainty (interference fits, seal locations, aerodynamic cross-coupling, etc.).
5 Assemble the error supermodel (consists of engineering model coupled to the experimentally identified model, Fig. 2).
6 Turn $H_\infty$ crank⇒produces $\mathbf{K}(s)$.

$\mathbf{K}(s)$ is a representation of the physical effects operating at the points of uncertainty. It will be an $n\times n$ matrix where $n$ is the number of uncertainties. The elements of this matrix are transfer functions of order equal to the sums of the orders of the identified plant and the mathematical model.

## Examples

Consider a long, flexible three-bearing rotor supported by magnetic bearings. The geometry of this rotor is illustrated in Fig. 3.

The bearing layout of this rotor is intended to be representative of a large boiler feed pump. The rotor is approximately 2.46 m long and weights 1036 N. The expected continuous running speed of the pump is 5200 rpm. The stiffness and damping coefficients for all three bearings are $8.75\times10^6$ N/m and 17,500 N s/m, respectively.

The initial measurement of FRF and subsequent model identification is not actually carried out here because we are using a numerical model. Instead, for the purposes of this example, the ''identified model'' is the same as the ''nominal model'' but with the actual values of the unknown effects included.

The model of the rotor is assembled from undamped mode shapes. Two rigid-body mode shapes and four flexible mode shapes are used. For the purposes of this work the model is considered planar and no gyroscopic effects are included. This limitation is imposed to simplify the examples. The method, however, can be used directly for coupling between the vertical and horizontal direction.

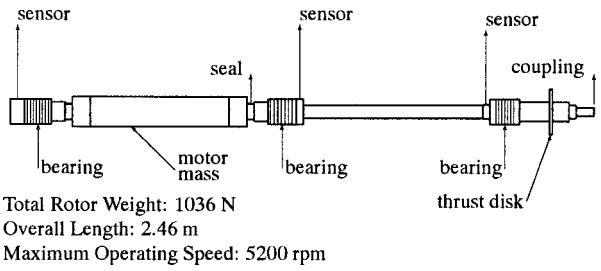



**Fig. 3 Three-bearing rotor**

**Example 1.** It is posited that the true rotor contains some dynamics at the thrust bearing end of the shaft, say due to some interaction with a coupling at the end of the shaft. In this example a stiffness of 900,000 N/m is found in the ''identified'' model at this location. In the nominal model, only the *location* of the unknown spring is assumed to be known. This is the only assumption made in the model reconciliation process. The value of $\varepsilon$ in Eqs. (6) and (7) is selected to be small enough such that the error in the transfer function of the dynamic system does not change with the value of $\varepsilon$ and big enough to allow good convergence of the $H_\infty$ crank. The actual value of $\varepsilon$ depends on the dynamics of the rotor being analyzed. For the rotor in these examples, values of $\varepsilon$ between $10^{-9}$ and $10^{-10}$ are appropriate.

Figures 4–6 show the results of this process. The maximum singular value of each system is used for comparison. The maximum singular value is a composite measure of the component transfer functions in a matrix and provides useful and direct comparison of the overall model of the system ([27]). Figure 4 shows

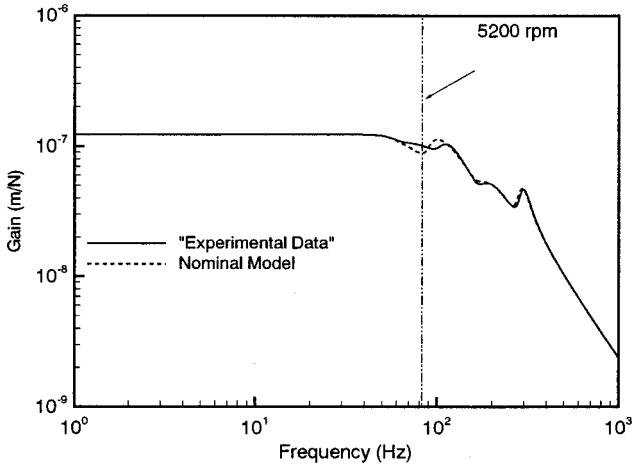



**Fig. 4 Maximum singular value comparison between nominal model and the identified rotor transfer functions from measurable inputs (bearings) to measurable outputs (sensors), Example 1**

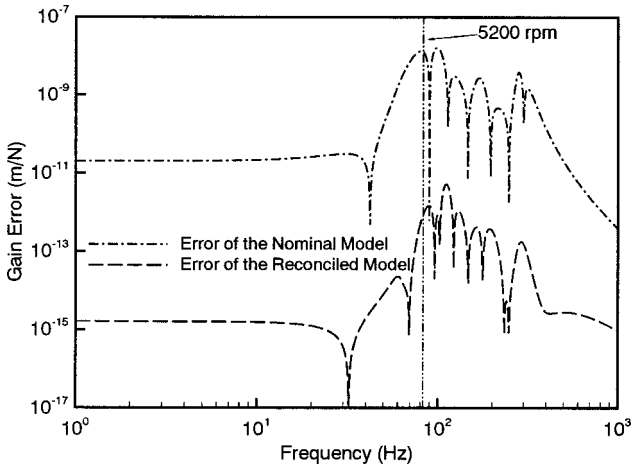



**Fig. 5 Comparison of the error of the nominal model and the reconciled model, Example 1**

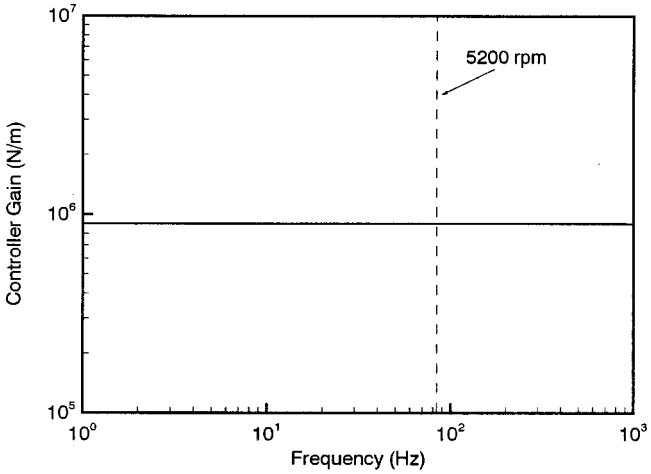



**Fig. 6 Matching controller transfer function gain, Example 1**

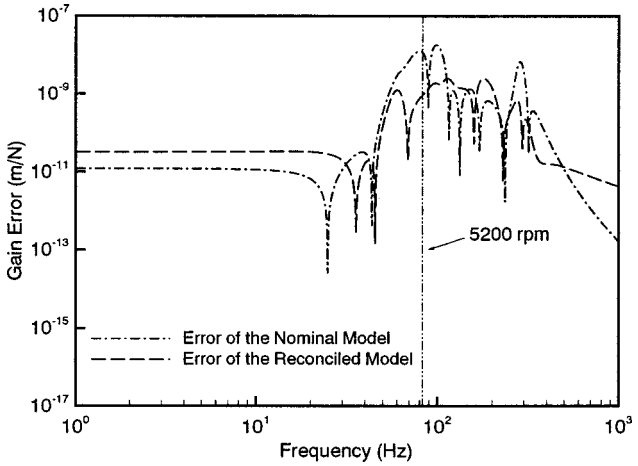



**Fig. 8 Comparison of the error of the nominal model and the reconciled model, Example 2**

a noticeable difference between the identified transfer function and nominal model transfer function at frequencies in the vicinity of the operating speed.

Figure 5 shows the error between the identified rotor and the nominal and reconciled rotor models. At the running speed, the difference between the identified rotor model and the nominal model is 12.35 percent. At the same speed, the difference between the identified rotor and the reconciled model is 0.00083 percent. The reconciliation method decreases the error by five orders of magnitude. As shown in Fig. 6, the $H_\infty$ controller synthesis exactly identifies the unknown spring with a stiffness of 900,000 N/m.

**Example 2.** As a second example, the identified rotor now includes the effects of the coupling described in Example 1 and the effects of a seal located at midspan. The seal has a stiffness coefficient of 350,000 N/m and damping coefficient of 4,300 N s/m. For the reconciliation process, the location of the coupling effect is identified but the seal is ignored. Figure 7 shows the comparison between maximum singular value of the identified rotor system and the nominal model. There is a noticeable difference in the vicinity of the operating speed of the rotor. This is expected because the model does not include the effects of the seal and the stiffness at the coupling end. Figure 8 shows the error

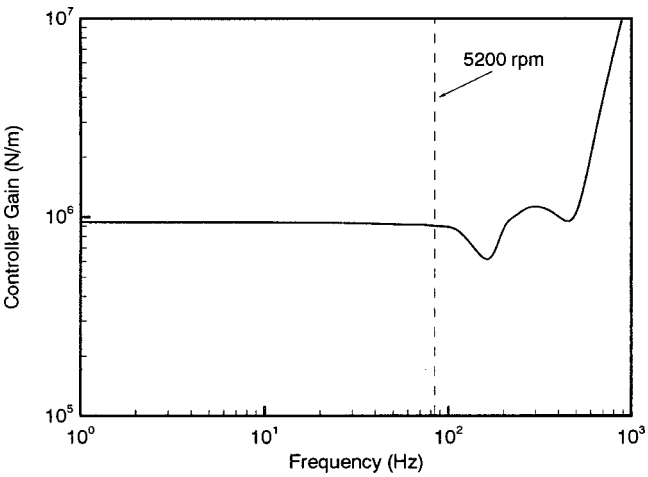



**Fig. 9 Matching controller transfer function gain, Example 2**

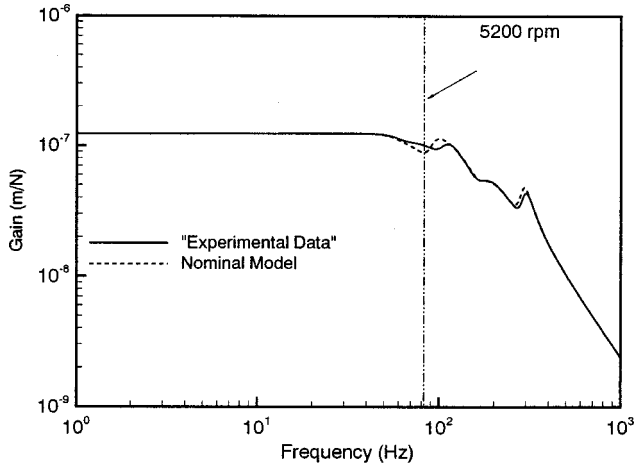



**Fig. 7 Maximum singular value comparison between nominal model and the identified rotor transfer functions from measurable inputs (bearings) to measurable outputs (sensors), Example 2**

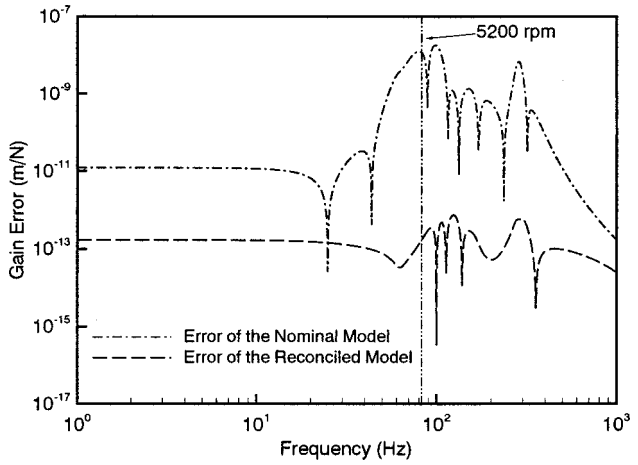



**Fig. 10 Comparison of the error of the nominal model and the reconciled model, Example 3**

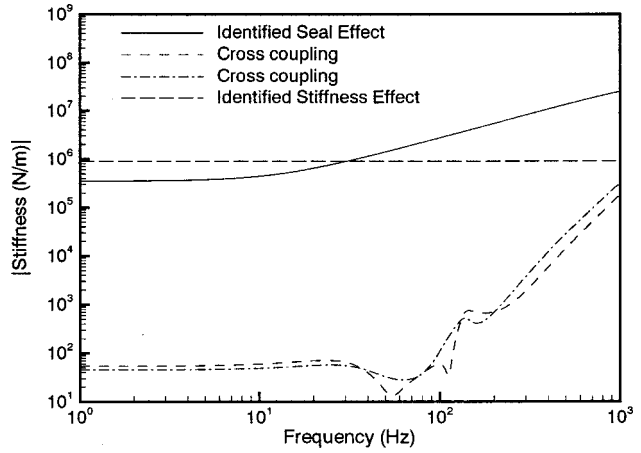



**Fig. 11 Matching controller transfer function gain, Example 3**

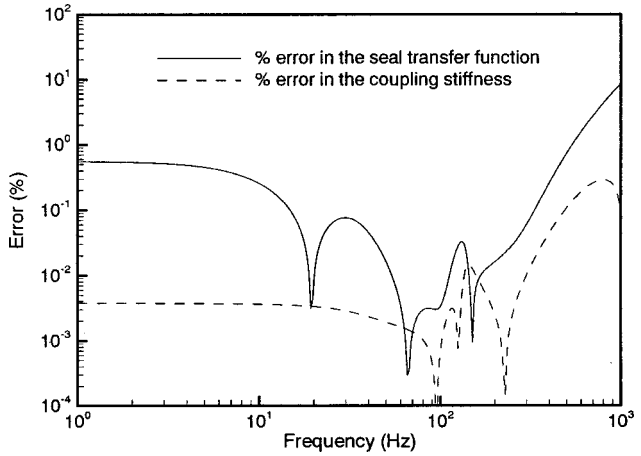



**Fig. 12 Percentage error of matching controller transfer function gain, Example 2**

of the nominal model and the error of the reconciled model. The reconciliation process (which is unaware of the seal) does not usefully improve the overall error of the model. Figure 9 shows the transfer function of the matched dynamic system. In this case, the synthesis method is trying to match the effect of the coupling and the seal with a transfer function at the coupling location. This example shows both a limitation and strength of this method. The limitation is that the sources of uncertainty must be known in order for the method to improve the model. The strength is that, if the sources of uncertainty are poorly chosen, then the resulting poor match suggests that other sources should be included.

**Example 3.** As a last example, the true model includes the effects of the coupling and the seal as introduced in Example 2. For the reconciliation process, however, the locations of the uncertainties (i.e., the location of the coupling and seal) are both known. Figure 10 shows the error in the maximum singular value of the nominal model and the reconciled model. The error in this case is also reduced for the reconciled model, showing a large improvement. The error is now four orders of magnitude smaller than the error of the nominal model. Figure 11 shows the transfer function of the matching dynamic system. The transfer functions of the coupling effects and the seal are clearly identified.

The synthesis method also produces cross-coupling between the two uncertainty locations, although their magnitudes are much smaller than the direct, physically significant transfer functions. Generation of undesired cross-couplings is a drawback of using standard $H_\infty$ or $H_2$ cranks to compute $\mathbf{K}(s)$. The cross-coupling is an inherent result of the synthesis method and cannot be avoided at present. However, forcing the synthesis of uncoupled transfer functions is an active area of research in the controls community. Once such methods become available, they can be used directly in this reconciliation technique.

**Table 1 Comparison between the eigenvalues of the nominal model, identified rotor-bearing system, and the reconciled model**

| Nominal Model | | Identified System | | Reconciled Model | |
|---|---|---|---|---|---|
| Real Exp (1/s) | Frequency (rad/s) | Real Exp (1/s) | Frequency (rad/s) | Real Exp (1/s) | Frequency (rad/s) |
| | | | | −1255.6 | 142.73 |
| | | | | −122.24 | 184.13 |
| | | | | −1870.8 | 351.84 |
| −105.40 | 407.23 | −131.12 | 424.51 | −131.12 | 424.51 |
| −284.03 | 479.69 | −302.85 | 497.53 | −302.85 | 497.53 |
| −80.116 | 549.56 | −474.39 | 575.84 | −474.39 | 575.84 |
| −473.07 | 590.60 | −117.07 | 670.79 | −117.08 | 670.79 |
| −311.28 | 1073.1 | −301.82 | 1050.3 | −301.82 | 1050.3 |
| −161.96 | 1808.7 | −173.79 | 1842.9 | −173.80 | 1842.9 |
| | | | | −55804. | 2827.6 |
| | | | | −1384.8 | 23698. |
| | | | | −220.77 | 102950 |
| | | | | −326.41 | |
| | | | | −348.87 | |
| | | | | −421.84 | |
| | | | | −1140.0 | |
| | | | | −1822.1 | |
| | | | | −2929.3 | |
| | | | | −40650. | |
| | | | | −60953. | |

Figure 12 shows the percentage error between the transfer functions of the coupling stiffness and seal effects and the transfer functions calculated by the $H_\infty$ synthesis. The error of the coupling stiffness is below 0.004 percent up to 130 Hz. The maximum error in the range of interest is 0.28 percent at 811 Hz. The error of the seal transfer function is below 0.55 percent up to 470 Hz.

Table 1 compares the eigenvalues of the nominal model, the identified rotor-bearing system and the reconciled model. The reconciliation process improves the eigenvalues of the system. The reconciled model includes all the eigenvalues of the true rotor-bearing system plus some eigenvalues that are artifacts of the reconciliation process. The additional eigenvalues are all well damped and do not significantly affect the stability analyses.

## Conclusions

A method is presented to reconcile engineering models of rotating machinery with experimental data. The underlying assumption is that the difference between the engineering models and the experimental data is caused by uncharacterized disturbances (uncertainties) acting at discrete points in the model. The forces caused by the uncertainties are represented as local feedback mechanisms. The method makes use of the existing engineering model of the system and likely location of uncertainties where error enters the model. Transfer functions acting at the location of uncertainties are calculated so that the response of the reconciled engineering model matches the experimental data.

This method has several advantages over other parameter identification schemes. One is that it effectively shrinks the number of identification variables (uncertainties) to a reasonable size and allows the engineer direct control over where the model adjustments are made. Another advantage is that the uncertainties have a direct analogue to real, physical effects—interference fits, aerodynamic behavior, etc. The identification of the location of the uncertainties is a critical step. Given the current state of the art in rotordynamics modeling, it is expected that the experienced rotordynamicist

should be capable of sensibly identifying the likely sources of uncharacterized disturbances acting upon the system.

An interesting benefit to this method is that it can actually produce a scaled measure of the disturbance effect. Numerical examples are presented where coupling stiffness and seal effects are identified from frequency response data of the rotor-bearing system. The only assumption used in the process is the location of the uncertainties in the model. The reconciliation process reduces the error between the engineering model and the experimental data by four orders of magnitude. The eigenvalues (stability analysis) of the reconciled model match the eigenvalues of the actual rotor-bearing system. The reconciliation process adds some eigenvalues as the product of the added transfer function. However, these added eigenvalues are all well damped.

If the locations of all uncertainties acting in the system are not included in the reconciliation process, the method will try to match the response of the nominal model to the experimental data using only the identified locations of uncertainties. In the example presented, the reconciliation process does not usefully improve the error between the reconciled model and the experimental data. This indicates that other sources of uncertainty need to be identified.

The reconciliation process produces a measure of the disturbance effects acting on the engineering model plus some undesired cross-coupling effects. The cross-coupling is an inherent result of the synthesis method and cannot be avoided at present. However, forcing the synthesis of uncoupled transfer functions is an active area of research in the controls community. Once such methods become available, they can be used directly in this reconciliation technique.

## References


[1] Sortore, C. K., 1999, ''Observer Based Critical Response Estimation in Rotating Machinery,'' Ph.D. dissertation, University of Virginia, Jan.

[2] Lund, J. W., 1974, ''Stability and Damped Critical Speeds of a Flexible Rotor in Fluid-Film Bearings,'' J. Eng. Ind., **96**, pp. 509–517.

[3] Nelson, H. D., and McVaugh, J. M., 1976, ''The Dynamics of Rotor-Bearing Systems Using Finite Elements,'' J. Eng. Ind., **98**, pp. 593–600.

[4] Vance, J. M., Murphy, B. T., and Tripp, H. A., 1987, ''Critical Speeds of Turbomachinery: Computer Predictions vs. Experimental Measurements—Part I: The Rotor Mass–Elastic Model,'' ASME J. Vib., Acoust., Stress, Reliab. Des., **109**, pp. 1–7.

[5] Vance, J. M., 1988, *Rotordynamics of Turbomachinery*, John Wiley and Sons, New York.

[6] Ehrich, F. F., 1992, *Handbook of Rotordynamics*, McGraw-Hill, New York.

[7] Childs, D., 1993, *Turbomachinery Rotordynamics*, John Wiley and Sons, New York.

[8] Ljung, L., 1999, *System Identification, Theory for the user*, 2nd Ed., Prentice-Hall, Englewood Cliffs, NJ.

[9] Ewins, D. J., 1984, *Modal Testing: Theory and Practice*, Research Studies Press Ltd., Tauton, Somerset, UK.

[10] Maia, N. M. M., Silva, J. M. M., He, J., Lieven, N. A. J., Lin, R. M., Skingle, G. W., To, W. M., and Urgueria, A. P. V., 1997, *Theoretical and Experimental Modal Analysis*, Research Studies Press Ltd., Tauton, Somerset, UK.

[11] Barrett, L. E., Nicholas, J. C., and Dhar, D., 1986, ''The Dynamic Analysis of Rotor-Bearing Systems Using Experimental Bearing Support Compliance Data,'' *Proceedings of the Fourth International Modal Analysis Conference*, Union College, Schenectady, NY, pp. 1531–1535.

[12] Nicholas, J. C., Whalen, J. K., and Franklin, S. D., 1986, ''Improving Critical Speed Calculations Using Flexible Bearing Support FRF Compliance Data,'' *Proceedings of the 15th Turbomachinery Symposium*, Texas A&M University, College Station, TX.

[13] Redmond, I., 1995, ''Practical Rotordynamics Modeling Using Combined Measured and Theoretical Data,'' *Proceedings of the 13th International Modal Analysis Conference*, Nashville, TN, Society for Experimental Mechanics, Bethel, CT.

[14] Redmond, I., 1996, ''Rotordynamic Modelling Utilizing Dynamic Support Data Obtained From Field Impact Tests,'' *Proceedings of Sixth International Conference on Vibrations in Rotating Machinery*, Paper C500/055/96.

[15] Rouch, K. E., McMains, T. H., and Stephenson, R. W., 1989, ''Modeling of Rotor-Foundation Systems Using Frequency-Response Functions in a Finite Element Approach,'' *Proceedings of the 1989 ASME Design Technical Conference 12th Biennial Conference on Mechanical Vibration and Noise*, Montreal, Canada, ASME, New York, pp. 157–166.

[16] Stephenson, R. W., and Rouch, K. E., 1992, ''Generating Matrices of the Foundation Structure of a Rotor System From Test Data,'' J. Sound Vib., **154**, No. 3, pp. 467–484.

[17] Vázquez, J. A., and Barrett, L. E., 1999, ''Transfer Function Representation of Flexible Supports and Casings of Rotating Machinery,'' *Proceedings of the 17th International Modal Analysis Conference*, Orlando, FL, Feb., Society for Experimental Mechanics, Bethel, CT.

[18] Vázquez, J. A., 1999, ''Using Transfer Functions to Model Flexible Supports and Casings of Rotating Machinery,'' Ph.D. dissertation, University of Virginia, Charlottesville, VA, Jan.

[19] Vázquez, J. A., Barrett, L. E., and Flack, R. D., 1999, ''A Flexible Rotor on Flexible Bearing Supports. Part I: Stability,'' *Proceedings of the 1999 Vibration Conference*, Paper DETC99/VIB-8285. Las Vegas, NV, Sept. 12–15, ASME, New York.

[20] Vázquez, J. A., Barrett, L. E., and Flack, R. D., 1999, ''A Flexible Rotor on Flexible Bearing Supports. Part II: Unbalance Response,'' *Proceedings of the 1999 Vibration Conference*, Paper DETC99/VIB-8286. Las Vegas, NV, Sept. 12–15, ASME, New York.

[21] Sanathanan, C. K., and Koerner, J., 1963, ''Transfer Function Synthesis as a Ratio of Two Complex Polynomials,'' IEEE Trans. Autom. Control, **8**, pp. 56–58.

[22] Gähler, C., and Herzog, R., 1994, ''Identification of Magnetic Bearing Systems,'' *Proceedings of the Fourth International Symposium on Magnetic Bearings*, ETH Zurich, Aug., University Press, Swiss Federal Technical University, Zurich, Switzerland, pp. 293–298.

[23] Wolodkin, G., Rangan, S., and Poolla, K., 1997, ''LFT Approach to Parameter Estimation,'' *Proceedings of the American Control Conference*, Vol. 3, June 4–6, IEEE, Piscataway, NJ, pp. 2088–2092.

[24] Fisher, D., Jue, D., Packard, A., and Poolla, K., 1999, ''On the Identification of High-Order Lightly-Damped Multivariable Systems,'' *Proceedings of the 1999 American Control Conference*, San Diego, CA, June 2–4, IEEE, Piscataway, NJ, pp. 848–853.

[25] Maslen, E. H., and Bielk, J. R., 1992, ''A Stability Model for Flexible Rotors With Magnetic Bearings,'' ASME J. Dyn. Syst., Meas., Control, **114**, pp. 172–175.

[26] Zhou, K., Doyle, J. C., and Glover, K., 1996, *Robust and Optimal Control*, Prentice Hall, Englewood Cliffs, NJ.

[27] Maslen, E. H., 1991, ''Magnetic Bearing Synthesis for Rotating Machinery,'' Ph.D. dissertation, University of Virginia, Jan.

# Synchronous Response Estimation in Rotating Machinery

**E. H. Maslen**
Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, VA 22903
e-mail: ehm7s@virginia.edu

**C. K. Sortore**
Synchrony, Inc., 777 Bent Mountain Road, Roanoke, VA 24018
e-mail: Chris.Sortore@synchrony.com

**J. A. Vázquez**
e-mail: jose@virginia.edu

**C. R. Knospe**
e-mail: crk4y@virginia.edu

Department of Mechanical and Aerospace Engineering, University of Virginia, Charlottesville, VA 22903

*Synchronous response estimation attempts to determine the forced response (displacement) of a rotor at critical points which cannot be measured directly. This type of prediction, if accurate and reliable, has broad potential use in the rotating machinery industry. Many machines have close clearance points on their shafts, such as seals, which can easily be damaged by excess vibration. Accurate estimates of the actual level of vibration at these points could usefully assist machine operators in troubleshooting and in protecting the equipment from expensive damage. This type of response information can be used both to generate less conservative alarm limits and, if magnetic bearings are available, to directly guide the bearing controllers in restricting the rotor motion at these critical points. It is assumed that the disturbance forces acting upon the rotor are predominantly synchronous. The response estimate is constructed using the measurable response in conjunction with an estimator gain matrix derived from a model of the transmissibilities of the rotor system. A fundamental performance bound is established based on the single-speed set of measurements by bounding the response to the unmeasurable component of the disturbance force. Acknowledging that some model uncertainty will always exist, a robust performance analysis is developed using structured singular value ($\mu$) analysis techniques. Assuming some reasonable levels of uncertainty for the model parameters (natural frequencies, modal dampings, mode shapes, bearing stiffnesses, and dampings) the results of the estimator construction and analysis establish feasibility of the proposed estimation. Two reference rotor models that are representative of industrially sized machines are used to demonstrate and evaluate the estimation. The unmeasurable response estimation errors consistently lie below 25 $\mu$m for the examples examined.*
[DOI: 10.1115/1.1417482]

## Introduction

In monitoring the synchronous response of rotating machinery, it is common to try to infer the motion at a point of interest from measurements taken elsewhere in the machine. This type of response estimation capability, if it is accurate and reliable, has broad potential use in the rotating machinery industry. Many machines have close clearance points on their shafts, such as at a seal location, which can be easily damaged by excess vibration. An accurate estimate of the actual level of vibration at these points could usefully assist machine operators in troubleshooting and in protecting the equipment from expensive damage. Some specific uses of this type of synchronous response information might be to generate less conservative vibration alarm limits or to directly guide magnetic bearing controllers in restricting the rotor motion at these critical points.

The theoretical basis for this work is straightforward. Fundamentally, it must be shown that measurements of sensor signals can be used to accurately determine the magnitude and phase of both a variety of disturbance forces acting upon the rotor and, ultimately, motion at points other than those that are directly measured. In general, rotordynamic disturbance forces can be categorized into three main groups: those that are constant, those that are synchronous and those that are random or unpredictable. This work focuses on response due to synchronous forces. Dominant synchronous disturbance forces may be produced by mass unbalance at distributed locations along the rotor, seal interaction, coupling misalignment, shaft bow, disk skew, or fluid excitation. There are many machines which are subject to nonsynchronous disturbance forces, however, synchronous forces are by far the dominant forcing mechanism in most rotating machinery.

It is well established that just a few sensors can be used to identify a relatively large number of well characterized synchronous disturbance forces, such as mass unbalance, if measurements are taken at many different speeds ([1]). Uncertainty in the rotor model can substantially degrade this capability as demonstrated in ([2]). In the present work, we accept the common constraint that the machine speed must be held constant for sufficiently long periods of time and multispeed measurements cannot be used to enhance the estimation process.

Crucial to the utility of any estimation process is a reliable bound on the error in the estimate. It is axiomatic that *some* estimate of the motions of interest can always be formed, but for this estimate to be useful, a good bound on how accurate this estimate is must be established. If the bound is overly conservative (too large), then the estimate will always appear to be useless. If the bound underestimates the actual error, then it can't be relied on. Thus, the really critical issue in formulating the estimation process is assessment of its accuracy. The present paper develops a baseline bound on estimation error under the condition that the model has no errors. This analysis is then extended to consider the influence of model error on this bound and a clear method is presented for obtaining a sensible bound given estimates of the size of the model errors and forces acting on the rotor.

## Basic Estimator Formulation

There are two main components which influence the synchronous response estimation process. First, the nature of all the dominant unknown disturbance force inputs must be identified. This entails identifying the type of force, the location at which it acts and a realistic bound on its magnitude. Second, the transmissibilities between the input forces and the measurable and unmeasurable output displacements must be determined. With these components, an estimator can be constructed which attempts to

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-397. Manuscript received by IGTI Nov. 1999; final revision received by ASME Headquarters Feb. 2000. Associate Editor: D. Wisler.

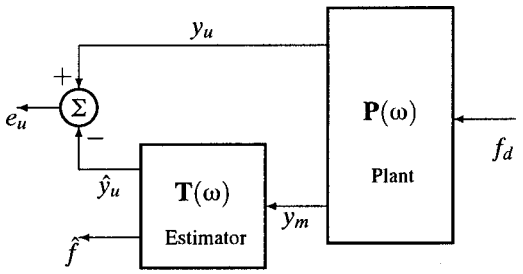



**Fig. 1 Estimation process block diagram**

minimize the error between the actual unmeasured response, $y_u$, and its estimate, $\hat{y}_u$. The estimate is extracted solely from the measurable responses, $y_m$, using the plant model. This process is summarized in Fig. 1.

The frequency-dependent transmissibility functions or transfer functions for a rotating machine between the measurable and unmeasurable response and any disturbance inputs (forces), can be modeled ([2]) as

$$\begin{Bmatrix} y_m(\omega) \\ y_u(\omega) \end{Bmatrix} = \begin{bmatrix} \mathbf{P}_m(\omega) \\ \mathbf{P}_u(\omega) \end{bmatrix} f(\omega). \tag{1}$$

In this model, the individual elements of the disturbance vector, *f*, can represent localized forces (like mass unbalance) or they can represent the scaling of distributed forces like shaft bow. The elements of the measurable response, $y_m$, will typically be measures of rotor motion at discrete points as determined by a physical measurement device such as a proximity probe or an accelerometer. The elements of the unmeasurable response, $y_u$, may either be similar measures of rotor motion at discrete points (perhaps at a seal clearance, for example) or scales of distributed measures such as the material strain at a point where fatigue is a concern. It is assumed that the elements of $y_u$ are not directly measurable and must be inferred from $y_m$.

In order to accurately predict the unmeasurable response, $y_u(\omega)$, the above transmissibility coefficient submatrices, $\mathbf{P}_m(\omega)$ and $\mathbf{P}_u(\omega)$, must be identified through a combination of modeling and careful system identification. Once these transmissibilities are identified, the unmeasurable displacements can apparently (assuming that $\mathbf{P}_m^{-1}$ exists) be computed as

$$y_u(\omega) = \mathbf{P}_u(\omega)[\mathbf{P}_m(\omega)]^{-1} y_m(\omega). \tag{2}$$

The unknown disturbance force can also be identified, as an intermediate step to (2), and is computed as

$$f(\omega) = [\mathbf{P}_m(\omega)]^{-1} y_m(\omega). \tag{3}$$

The result, as can readily be seen from (2) and (3), is highly dependent on the quality of the estimated transmissibility coefficient submatrices. In addition, $\mathbf{P}_m(\omega)$ must be invertible which, as will be shown, is not necessarily possible.

## First Bound on Estimation Error

The most fundamental limit to the performance of a fixed speed synchronous estimator arises because, in general, only a subset of *f* can be estimated. This is because the dimension of *f* is generally greater than that of $y_m$: There are more forcing functions than sensors. More precisely, only those forces lying in the domain of $\mathbf{P}_m(\omega)$ produce measurable effects at $y_m$ so only they can be measured. Those forces which cannot be estimated through $y_m$ may still produce a part of the response at $y_u$, and this part of the response fundamentally cannot be estimated.

Assume that an orthonormal basis for the domain of $\mathbf{P}_m$ can be computed (using a singular value decomposition):

$$\mathbf{H}_\parallel = [H_1, H_2, \ldots, H_p] : \mathbf{P}_m H_i \neq 0, \quad \mathbf{H}_\parallel^* \mathbf{H}_\parallel = \mathbf{I}$$

and that its orthogonal complement can be computed:

$$\mathbf{P}_m \mathbf{H}_\perp = 0, \quad \mathbf{H}_\perp^* \mathbf{H}_\perp = \mathbf{I}, \quad \text{rank}[\mathbf{H}_\parallel | \mathbf{H}_\perp] = \dim(f).$$

With this, the disturbance force can be decomposed as

$$f = [\mathbf{H}_\parallel | \mathbf{H}_\perp] \begin{Bmatrix} f_\parallel \\ f_\perp \end{Bmatrix} \tag{4}$$

and the response (1) becomes

$$\begin{Bmatrix} y_m \\ y_u \end{Bmatrix} = \begin{bmatrix} \mathbf{P}_m \mathbf{H}_\parallel & 0 \\ \mathbf{P}_u \mathbf{H}_\parallel & \mathbf{P}_u \mathbf{H}_\perp \end{bmatrix} \begin{Bmatrix} f_\parallel \\ f_\perp \end{Bmatrix}. \tag{5}$$

Since the product $\mathbf{P}_m \mathbf{H}_\parallel$ has full-column rank, $\mathbf{H}_\parallel f_\parallel$ can be computed through the pseudo-inverse,

$$\mathbf{P}_m^\dagger \triangleq \mathbf{H}_\parallel (\mathbf{H}_\parallel^* \mathbf{P}_m^* \mathbf{P}_m \mathbf{H}_\parallel)^{-1} \mathbf{H}_\parallel^* \mathbf{P}_m^* \; : \; \mathbf{H}_\parallel f_\parallel = \mathbf{P}_m^\dagger y_m. \tag{6}$$

Finally, the unmeasurable response is *estimated* by assuming that $f_\perp = 0$ so that, from (5),

$$\hat{y}_u = \mathbf{P}_u \mathbf{H}_\parallel f_\parallel = \mathbf{P}_u \mathbf{P}_m^\dagger y_m. \tag{7}$$

Using this decomposition of *f*, the estimator can now be *properly* constructed as

$$\mathbf{T}(\omega) = \begin{bmatrix} \mathbf{P}_u(\omega) \\ \mathbf{I} \end{bmatrix} \mathbf{P}_m^\dagger(\omega). \tag{8}$$

The residual error in the estimated position is then

$$e_u = \hat{y}_u - y_u = \mathbf{P}_u \mathbf{H}_\perp f_\perp.$$

The 2-norm of this error can be bounded to measure the accuracy of the estimator:

$$\max_{f_\perp} |e_u|_2 = \max_{f_\perp} |\mathbf{P}_u \mathbf{H}_\perp f_\perp|_2 \leq \bar{\sigma}(\mathbf{P}_u \mathbf{H}_\perp \mathbf{H}_\perp^*) \max |f|_2. \tag{9}$$

This means that, given a bound on the forces which act on the rotor and a dynamic model of the rotor, a meaningful bound on the estimation accuracy can be determined. This bound assumes that the model has no error and it is an *upper bound on error under this assumption*: If $f_\perp = 0$, then the error would be zero.

## Scaling

Missing in the formulation at this point is attention to the relative scales of the components of *f* and of the relative importance of errors in $y_u$. These scales are significant because error bounds are composite measures of multiple signals: their 2-norm. Further, as a dimensioned measure, the bound is hard to interpret. A better way to carry out the analysis is to scale the vectors *f* and *e* to encapsulate physical bounds on individual elements of *f* and engineering requirements on *e*.

Typically, limits to *f* can be described as

$$|f_i| \leq \alpha_i \; : \; i = 1 \ldots n \tag{10}$$

and the target limits to error in position estimation are

$$|e_{u,i}| \leq \beta_i \; : \; i = 1 \ldots m. \tag{11}$$

If the $f_i$ represent discrete forces, then each corresponding $\alpha_i$ is the largest value that force is expected to exhibit. For instance, rotor balancing specifications typically limit the number of allowed gram-millimeters of unbalance at each balance plane. This limit is easily converted to a limit to the unbalance force which can act at that plane at a given speed. In a similar manner, the $\beta_i$ would typically reflect how precisely the measurements need to be known in order to be useful. If the maximum rotor motion at a seal needs to be on the order of 25 $\mu$m, then the corresponding $\beta_i$ might be 10 $\mu$m.

To conveniently implement this scaling concept, we introduce a pair of scaling matrices which are simply diagonal with each diagonal element being the reciprocal of the corresponding scale: $\mathbf{Q}_f = \text{diag}(1/\alpha_1 \ldots 1/\alpha_n)$, $\mathbf{Q}_e = \text{diag}(1/\beta_1 \ldots 1/\beta_m)$. Given these

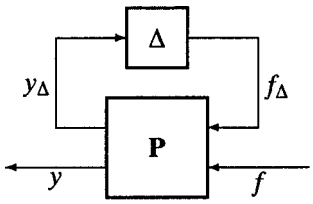



**Fig. 2 Feedback connection of uncertainties**

conventions, it is expected that $|\mathbf{Q}_f f|_2 \leq 1$ and desired that $|\mathbf{Q}_e e_u|_2 \leq 1$. These scales can then be inserted into the performance bound computation (9) to obtain

$$\max_{f_\perp} |\mathbf{Q}_e e_u|_2 \leq \bar{\sigma}(\mathbf{Q}_e P_u H_\perp H_\perp^* \mathbf{Q}_f^{-1}) \max |\mathbf{Q}_f f|_2$$

$$\leq \bar{\sigma}(\mathbf{Q}_e P_u H_\perp H_\perp^* \mathbf{Q}_f^{-1}).$$

The nominal performance of the estimator will be acceptable as long as

$$\bar{\sigma}(\mathbf{Q}_e P_u H_\perp H_\perp^* \mathbf{Q}_f^{-1}) \leq 1.0. \tag{12}$$

To find the upper bound on the error, the values of $\beta_1, \ldots, \beta_m$ are decreased until (12) is equal to 1. At this point, $\beta_1, \ldots, \beta_m$ are the upper bounds on the error.

## Model Uncertainty Representation

In order to accurately evaluate the effects of the errors between the model and the true plant on the critical response estimation process a suitable description of the expected or likely sources of error is required. The choice of a particular model uncertainty representation will reflect the a priori knowledge of the physical mechanism causing the error. This choice will also depend upon the ability to produce a sensible mathematical representation of the mechanism which will permit the accurate evaluation of the error.

A powerful tool used in the analysis of control system robustness is the structured singular value ($\mu$) matrix function ([3]). This analysis is predicated on the notion that the effect of most uncertainties can be represented by the feedback connection diagrammed in Fig. 2. The matrix $\Delta$ contains all of the uncertainties: Its elements are assumed to be unknown but to vary within specific bounds. Such an arrangement represents a family of systems falling within the uncertainty bounds defined by $\Delta$. The structure and bounds of the uncertainty are selected so that the transfer function of the true plant, $\mathbf{P}_t$, is believed to lie within this family of possible systems.

Modeling uncertainties is a process of determining where model uncertainties lie and unravelling those uncertainties to produce the structure indicated in Fig. 2. A detailed explication of this process is provided in [2] but a very simple example will serve to illustrate the main points. Suppose that a rotor model has the form

$$\mathbf{M}\ddot{x} + \mathbf{D}\dot{x} + \mathbf{K}x = \mathbf{B}f$$

$$y_m = \mathbf{C}_m x$$

$$y_u = \mathbf{C}_u x.$$

The transfer functions from $f$ to $y_m$ and $y_u$ are

$$\mathbf{P}_m(\omega) = \mathbf{C}_m[\mathbf{K} - \mathbf{M}\omega^2 + j\mathbf{D}\omega]^{-1}\mathbf{B}$$

$$\mathbf{P}_u(\omega) = \mathbf{C}_u[\mathbf{K} - \mathbf{M}\omega^2 + j\mathbf{D}\omega]^{-1}\mathbf{B}.$$

Further, a seal produces stiffness and damping at a certain location so that $\mathbf{D} = \mathbf{D}_0 + p_s d_s p_s^*$, $\mathbf{K} = \mathbf{K}_0 + p_s k_s p_s^*$ in which the vector $p_s$ selects the mass station on which the seal acts while $d_s$ and $k_s$ are the nominal seal rotordynamic coefficients. Now, suppose that the two coefficients are known only within some uncertainty: $d_s = d_{s,0} + \Delta_d$, $k_s = k_{s,0} + \Delta_k$ where $\Delta_d$ and $\Delta_k$ are unknown

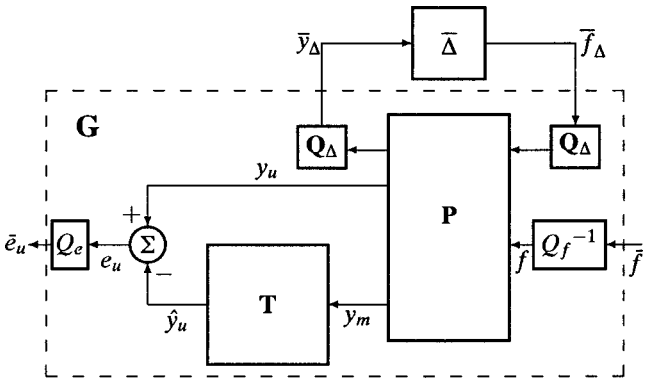



**Fig. 3 Estimation process with plant incorporating structured uncertainty**

but reasonable bounds on their magnitudes are known. Denote $\mathbf{D}_n = \mathbf{D}_0 + p_s d_{s,0} p_s^*$, $\mathbf{K}_n = \mathbf{K}_0 + p_s k_{s,0} p_s^*$ so that the rotor model including seal uncertainty becomes

$$\mathbf{M}\ddot{x} + \mathbf{D}_n\dot{x} + \mathbf{K}_n x = \mathbf{B}f - p_s\Delta_d p_s^*\dot{x} - p_s\Delta_k p_s^* x$$

$$y_m = \mathbf{C}_m x$$

$$y_u = \mathbf{C}_u x.$$

Define the uncertainty output, $y_\Delta$, and the uncertainty input, $f_\Delta$:

$$y_\Delta \triangleq \begin{Bmatrix} p_s^* x \\ p_s^* \dot{x} \end{Bmatrix} \quad f_\Delta \triangleq \begin{bmatrix} \Delta_k & \\ & \Delta_d \end{bmatrix} y_\Delta$$

and let $K(\omega) = [\mathbf{K}_n - \omega^2\mathbf{M} + j\omega\mathbf{D}_n]$ to generate the system

$$\begin{Bmatrix} y_\Delta \\ y_m \\ y_u \end{Bmatrix} = \begin{bmatrix} \begin{Bmatrix} 1 \\ j\omega \end{Bmatrix} p_s^* \\ \mathbf{C}_m \\ \mathbf{C}_u \end{bmatrix} K(\omega)^{-1}[\mathbf{B}|-p_s|-p_s] \begin{Bmatrix} f \\ f_\Delta \end{Bmatrix}$$

$$f_\Delta = \begin{bmatrix} \Delta_k & \\ & \Delta_d \end{bmatrix} y_\Delta.$$

In this manner, the effect of the uncertainty is separated from the dynamics of the nominal plant as a feedback process readily represented by Fig. 2.

As illustrated in Sortore [2], this approach can be applied to virtually any uncertainty of interest in a rotor system including localized effects like seals, bearings, or disk fits as well as more distributed effects like rotor free-free modal frequency, damping, and even mode shapes.

## Structured Robust Performance Through $\mu$ Analysis

The model of the plant including the uncertainties as shown in Fig. 3 is in the canonical form used in $\mu$ analysis. This canonical form requires that the signals $\bar{f}$ and $\bar{f}_\Delta$ have expected 2-norm bounds of 1.0 and that the signals $\bar{e}_u$ and $\bar{y}_\Delta$ have target 2-norm bounds of 1.0 also. Consequently, the previously discussed scaling matrices $\mathbf{Q}_f$ and $\mathbf{Q}_e$ are introduced to scale $f$ and $e_u$, respectively. In addition, the matrix $\mathbf{Q}_\Delta$ is introduced to rescale the uncertainty block $\Delta$ so that each diagonal element has an expected magnitude bound of 1.0. Typically, if the unscaled $\Delta$ block has diagonal elements $|\Delta_{i,i}| \leq \gamma_i$, then $\mathbf{Q}_\Delta$ has diagonal elements $\sqrt{\gamma_i}$.

A generalized and complete presentation of $\mu$ is given by Young and Doyle [4]. In simplified terms, the $\mu$ analysis tool determines whether or not the 2-norm of $\bar{e}_u$ will be less than 1.0 for all unit norm bounded inputs $\bar{f}$ and all permissible uncertainties $\bar{\Delta}$. This criterion is effectively the same as (12) except that it holds true not only for the nominal model, but for all possible perturbed models:

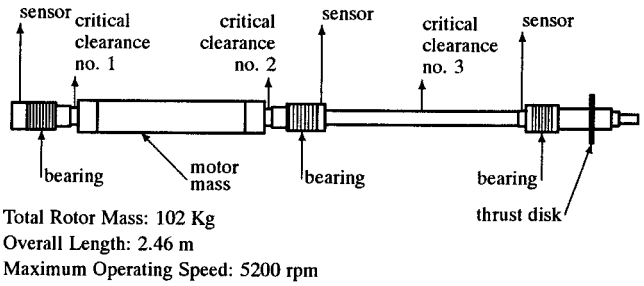



**Fig. 4 Three-bearing rotor**

$$\mu(\mathbf{G},\overline{\Delta}):\quad \frac{|\overline{e_u}|_2}{|f|_2} \overset{?}{\leq} 1 \quad \forall \overline{\Delta}:|\overline{\Delta}_{i,i}| \leq 1.0. \tag{13}$$

Finding a close upper bound on the error of the response estimation of a system with uncertainties can be accomplished in an efficient manner using the MU command in the MATLAB $\mu$ toolbox ([5]). The scaling matrix $\mathbf{Q}_e$ which satisfies this condition is easily found by a bisection routine or similar algorithm. The $\mu$ test only tells whether or not (13) is satisfied: it does not produce an actual norm. Thus, a bound can be determined by repeatedly adjusting the elements of $\mathbf{Q}_e$ until they are as large as possible while still satisfying (13). At this point, the actual error magnitude bounds are the reciprocals of the diagonal elements of $\mathbf{Q}_e$.

## Example 1

For the first example, consider the long flexible three-bearing rotor depicted in Fig. 4. ''Unmeasurable'' points along the rotor have been selected and identified which represent the location of seals or other potentially close clearance points. The bearing layout of this rotor is intended to be representative of a large boiler feed pump. The free-free natural frequencies (at zero speed) of this rotor are summarized in Table 1.

The bearing stiffness and damping coefficients for all three bearings are given in Table 2. For this rotor configuration, response predictions are based on the measured response at the three sensor locations along the rotor using (7). The magnitude of the predicted response depends on the magnitude of the measured response at the sensor locations. We will concentrate on the uncertainty of the prediction since they can be calculated independently of the actual measurements.

In order to calculate a bound on the estimation error, it is necessary to identify the forces that may act on the rotor and set a bound on them. For this example, two synchronous forces are considered, unbalance and shaft bow. While the actual unbalance distribution is not known we can assess a maximum limit. The API Standard 617 for centrifugal compressors section 2.9.2.4 defines the minimum allowable unbalance weight used for shop testing as

**Table 1 Rotor free-free natural frequencies**

| Pump | Compressor |
|---|---|
| 0 | 0 |
| 0 | 0 |
| 1817.3 cpm (30.2 Hz) | 6300.7 cpm (105.0 Hz) |
| 5885.3 cpm (98.1 Hz) | 13485.5 cpm (224.8 Hz) |
| 10506.3 cpm (175.2 Hz) | 26556.1 cpm (442.6 Hz) |
| 17641.9 cpm (294.1 Hz) | 38938.1 cpm (649.0 Hz) |

**Table 2 Bearing and seal coefficients**

| model | coefficient | central value | percent error |
|---|---|---|---|
| three-bearing | $K_b$ | $8.7510^7$ N/m | 10% |
| rotor | $C_b$ | 17,512. N-s/m | 10% |
| compressor | $K_b$ | $3.5010^7$ N/m | 10% |
| rotor | $C_b$ | 35,025. N-s/m | 10% |
| seal | $K_s$ | $1.7510^5$ N/m | 20% |
| | $k_s$ | 0 - $1.7510^7$ N/m | 20% |
| | $C_s$ | 1751. N-s/m | 20% |
| | $c_s$ | 0 | 20% |

$$U_{\max} = 6350W/N \tag{14}$$

where $W$ is the weight of the rotor in kilograms and $N$ is the speed of the rotor in revolutions per minute. This value will normally be the maximum acceptable unbalance in a machine operating in the field. For the sake of discussion and to make a conservative estimate, we will use this value (124.5 gr-mm) at all the balance planes of the rotor. That is, this unbalance is assumed to be acting at the left bearing, at the left end of the motor, at the right end of the motor, at the midspan bearing, at the right bearing location and at the thrust bearing location. The actual unbalance distribution is unknown but only the maximum bound is necessary for the calculation. Bow was added to the rotor in the form of the first bending mode of the rotor with a maximum amplitude of 127 $\mu$m.

Let's first consider the case of a perfect model (i.e., there are no uncertainties in the model). The error on the response prediction is due to the unknown (and unidentifiable) forces acting on the system. The upper bound of the error of the response prediction is calculated using (12). Figure 5 shows the results of this calculation for the critical clearance number 3. This is the maximum error expected in the response prediction at any frequency. The actual error is likely to be much smaller.

Figure 5 also include the maximum bound on the error in the case where the model is not perfect. In this case, uncertainties are assigned to the various parts of the model. The rotor modal parameters were assigned an error according to Table 3. The uncertainties of the bearing stiffness and damping coefficients are summarized in Table 2. The upper bound in the estimation error in this case is calculated using $\mu$ analysis (13).

It is observed, as expected, that the error bound increases throughout the speed range from that predicted from the nominal optimal model with only unidentified forces. At lower frequencies,

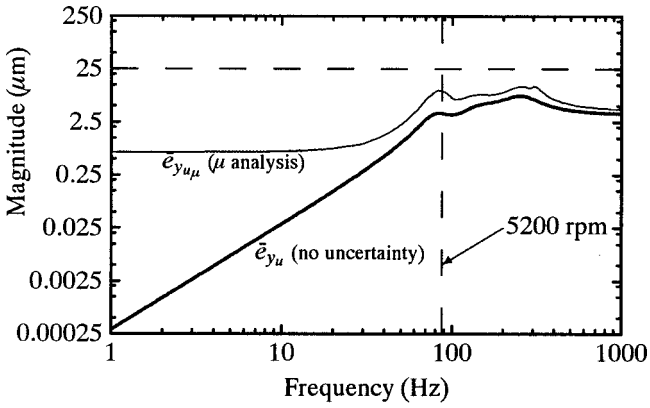



**Fig. 5 Maximized robust performance error and nominal maximized estimation error. Three-bearing rotor, critical point #3.**

**Table 3 Modal parameter uncertainties**

| mode | frequency error | damping ratio error | mode shape error |
|---|---|---|---|
| 1 | 2% | 3% | .04 |
| 2 | 2% | 4% | .06 |
| 3 | 3% | 5% | .08 |
| 4 | 3% | 7% | .10 |
| 5 | 4% | 9% | .12 |
| 6 | 4% | 11% | .14 |

the error increases by a substantial percentage, but is still acceptable. At higher frequencies, including the operating speed, the robust error bound adds 12 $\mu$m or less to the error predicted with no uncertainty, remaining within acceptable bounds (below the 25 $\mu$m threshold.)

## Example 2: Effect of Strong Localized Model Error

The second example explores a large turbomachine rotor such as that would be used in a single-stage centrifugal pipeline compressor. The geometry of this rotor is illustrated in Fig. 6. "Unmeasurable" points along the rotor have been selected and identified which represent the location of seals or other potentially close clearance points. The free-free natural frequencies (at zero speed) of this rotor are summarized in Table 1.

Of particular interest is the case when the synchronous response must be estimated at a point which has a strong localized model uncertainty and/or a large localized excitation. It was postulated that this particular scenario could exhibit unusually poor estimator performance due to the combined uncertainties of both the response estimation and the disturbance force acting at the same point. An example of this which has practical interest is a seal component such as a balance piston seal in a compressor, which can contribute substantial destabilizing forces to the rotor. Annular liquid seals in pumps typically contribute substantial and very uncertain forces to the rotor. Marscher [6], Childs [7], and Zhao [8] all provide insight into this issue. If unaccounted, this can introduce a substantial difference between the actual and modeled dynamics of the rotor. This scenario was examined for the compressor rotor in which a seal was introduced into the model at the balance piston location.

An analysis was carried out on the compressor rotor model with a cross-coupled seal acting on the critical point of estimation. The results of this analysis are presented in Fig. 7. The value of the cross-coupled coefficient, $k_s$, was increased linearly with frequency from 0 (at 0 rpm) to $1.7510^7$ N/m at 7000 rpm. The peaks in this plot are due to the uncertainty in the lightly damped higher modes (frequencies) of the system.

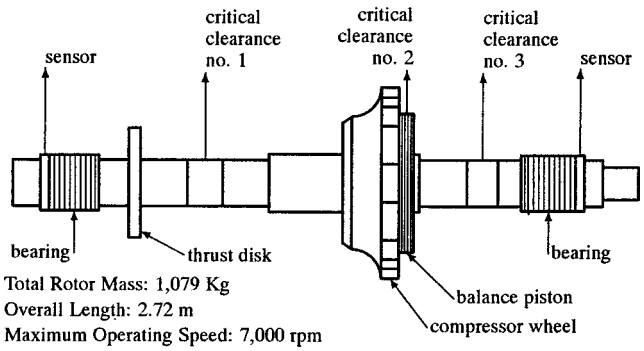



**Fig. 6 Compressor rotor**

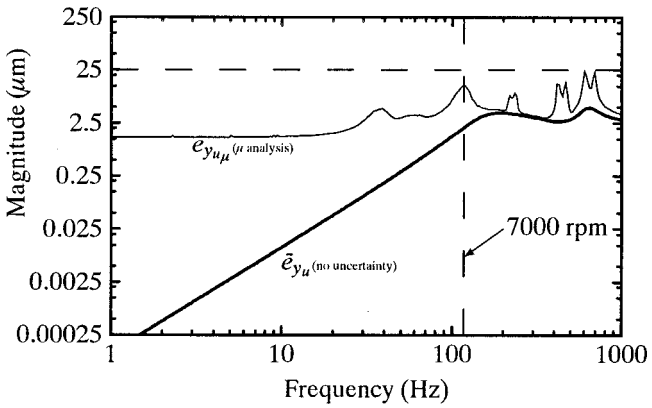



**Fig. 7 Maximized robust performance error and nominal maximized estimation error. Cross-coupled seal acting at point of estimation. Compressor rotor, critical point #2.**

Of particular interest in this example is the clearly visible split peaks in the error bound associated with the higher flexible modes. These peaks occur because the family of models inspected includes models with flexible modes on either side of the nominal value. If the actual mode lines up with the modeled mode, then the error at that frequency will be very low (leading to a dip in the error bound at the nominal mode frequency). If the actual mode is on either side, then the actual response at the actual mode will substantially exceed the predicted response, leading to the two side peaks in the error bound. In addition, it will be noted that, despite the combination of large potential imbalance and substantial model uncertainty, the estimator error bound is still acceptable.

## Discussion

A method has been presented which permits construction of a synchronous response estimator, *T*, from the elements of a conventional rotordynamic model for a rotating machine. This estimator is used to predict the vibration of the machine at points which are inaccessible for direct measurement, basing this estimate on measurements which can be made. More importantly, a well-developed assessment of the accuracy or reliability of this estimator has been developed which assesses the impact not only of that part of the exogenous loading which cannot be determined from the measurements but also includes the effect of uncertainty in the engineering model. This assessment is based on the powerful $\mu$ analysis tool developed over the last decade by the controls community for (in part) evaluation of dynamic system performance in the face of model uncertainty.

Several examples illustrate the important features of the estimator and its sensitivity to uncertainty. First, a very simple analysis based on maximum singular value provides a bound on performance without model uncertainty. This bound arises because not

all of the exogenous loads can be estimated. Next, the same problem was examined with numerous sources of model uncertainty including modal uncertainties as well as bearing parameter uncertainties. It was seen that the resulting error bounds are everywhere higher than for the simpler analysis: as would be expected. In particular, the error bound at low frequencies increased significantly due to bearing stiffness uncertainty. Finally, the influence of a strong model uncertainty acting directly at the point of estimation was examined. In this case, the localized uncertainty was due to aerodynamic coefficients of a balance piston seal. The estimator error bound was still acceptable under these circumstances.

All of the examples attempted to use reasonable uncertainty and force bounds based on engineering judgment, experience, and published standards, where applicable. The examples all showed the estimator uncertainties to be less than 25 $\mu$m, which would probably be acceptable for use in driving alarms or for sensing in a response minimizing control scheme.

## Nomenclature

$\mathbf{B}$ = force input matrix
$\mathbf{C}_m$ = measurable output matrix
$\mathbf{C}_u$ = unmeasurable output matrix
$C_b$ = bearing damping
$C_s$ = seal direct damping
$c_s$ = seal cross-coupled damping
$\mathbf{D}$ = damping matrix
$d_s$ = seal damping
$e_u$ = vector of error in unmeasurable response estimate
$f$ = vector of forces, including unbalance or bow effects
$f_\perp$ = unmeasurable force components
$f_\parallel$ = measurable force components
$\hat{f}$ = estimate of force vector
$f_\Delta$ = vector of ''uncertainty inputs''
$\mathbf{G}$ = composite scaled system transfer function for $\mu$ analysis
$\mathbf{H}_\perp$ = null space of $\mathbf{P}_m$
$\mathbf{H}_\parallel$ = domain of $\mathbf{P}_m$
$H_i$ = a column of the matrix $\mathbf{H}$
$\mathbf{K}$ = stiffness matrix
$K_b$ = bearing stiffness
$K_s$ = seal direct stiffness
$k_s$ = seal stiffness, seal cross-coupled stiffness
$\mathbf{I}$ = identity matrix
$\mathbf{M}$ = mass matrix
$\mathbf{P}$ = nominal machine transfer function matrix
$\mathbf{P}_m$ = machine transfer function matrix from forces to $y_m$
$\mathbf{P}_u$ = machine transfer function matrix from forces to $y_u$
$p_s$ = vector selecting seal location
$\mathbf{Q}_e$ = scaling matrix for estimate error
$\mathbf{Q}_f$ = scaling matrix for forces
$\mathbf{Q}_\Delta$ = scaling matrix for uncertainty bounds
$\mathbf{T}$ = estimator transfer function matrix
$x$ = vector of generalized rotor displacements
$y_m$ = vector of measurable responses
$y_u$ = vector of unmeasurable responses
$\hat{y}_u$ = vector of estimates of unmeasurable responses
$y_\Delta$ = vector of ''uncertainty outputs''
$\alpha$ = bound on magnitude of force
$\beta$ = bound on acceptable estimate error
$\gamma$ = uncertainty scale
$\Delta$ = uncertainty matrix for $\mathbf{P}$
$\Delta_k$ = uncertainty in seal stiffness
$\Delta_d$ = uncertainty in seal damping
$\omega$ = excitation frequency
$\overline{\sigma}(\cdot)$ = maximum singular value of $(\cdot)$
$(\cdot)^*$ = conjugate transpose of $(\cdot)$
$(\cdot)^\dagger$ = pseudo-inverse of $(\cdot)$
$\overline{a}$ = unit normalized version of $a$

## References


[1] Vance, J. M., 1988, *Rotordynamics of Turbomachinery*, John Wiley and Sons, New York.
[2] Sortore, C. K., 1999, ''Observer Based Critical Response Estimation In Rotating Machinery,'' Ph.D. dissertation, University of Virginia, Charlottesville, VA, Jan.
[3] Doyle, J., Packard, A., and Zhou, K., 1991, ''Review of LFTs LMIs and $\mu$,'' *Proceedings of the 30th IEEE Conference on Decision and Control*, IEEE, Piscataway, NJ, pp. 1227–1232.
[4] Young, P. M., and Doyle, J. C., 1990, ''Computation of $\mu$ with Real and Complex Uncertainties,'' *Proceedings of the 29th IEEE Conference on Decision and Control*, IEEE, Piscataway, NJ, pp. 1230–1235.
[5] Balas, G. J., Doyle, J. C., Glover, K., Packard, A., and Smith, R., 1995, *$\mu$-Analysis and Synthesis Toolbox*, The Mathworks, Inc., Natick, MA.
[6] Marscher, W. D., 1991, ''Analysis and Test of Multistage Pump ''Wet'' Critical Speeds,'' Tribol. Trans., **34**, No. 3, pp. 445–457.
[7] Childs, D., 1993, *Turbomachinery Rotordynamics*, John Wiley and Sons, New York.
[8] Zhao, Y., 1998, ''Dynamic Analysis of Annular Liquid Seals,'' M. S. Thesis, University of Virginia, Charlottesville, VA.

# Measurements of Rotordynamic Coefficients of Hybrid Bearings With (a) a Plugged Orifice, and (b) a Worn Land Surface

**F. Laurant**
Graduate Research Assistant

**D. W. Childs**
Leland T. Jordan Professor of Mechanical Engineering,

Turbomachinery Laboratory,
Texas A&M University,
College Station, TX 77843-3254

*Test results are presented for the rotordynamic coefficients of a hybrid bearing that is representative of bearings for liquid-rocket-engine turbopump applications. The bearing is tested in the following two degraded conditions: (a) one of five orifices plugged, and (b) a locally enlarged clearance to simulate a worn condition. Test data are presented at 24,600 rpm, with supply pressures of 4.0, 5.5, and 7.0 MPa, and eccentricity ratios from 0.1 to 0.5 in 0.1 increments. Overall, the results suggest that neither a single plugged orifice nor significant wear on the bearing land will "disable" a well-designed hybrid bearing. These results do not speak to multiple plugged orifices and are not an endorsement for operations without filters to prevent plugging orifices.*
[DOI: 10.1115/1.1403461]

## Introduction

Hybrid bearings are being actively considered to replace rolling-element bearings for future high-speed cryogenic turbopumps. They will be supplied with the cryogenic working fluid of the turbopump at high pressures. Because of the high rotational speed, hydrodynamic effects will be pronounced, despite the low-viscosity working fluids. The ''hybrid'' designation refers to the combined actions of hydrostatic and hydrodynamic effects. Their attractive features include: a greatly elevated DN limit, the ability to use split-stator bearings, and large direct stiffness and damping force coefficients. A considerable amount of test data have been obtained and published for these bearings operating in a nominal mode. This paper concerns two degraded operating conditions, namely, a plugged orifice and enlarged clearances around an orifice recess, simulating wear.

Although some turbopump manufacturers plan to use filters to prevent debris contamination from reaching a bearing, debris coming through the feed-line and plugging one or more orifices is a possible failure mode. To examine this possibility, tests were conducted on a five-orifice hybrid bearing, with one plugged orifice.

For a reusable turbopump, every start and shutdown will be accompanied by wear between the rotor and stator surfaces. Scharrer et al. [1] published a theoretical study on wear using a Reynolds-equation-based model. The available evidence suggests that lift off is achieved quickly and wear is minimal; however, the tests conducted here are designed to simulate fairly heavy wear, centered around an orifice recess.

The theory of hybrid bearings and the different reaction forces acting on the rotor are widely discussed in the literature. San Andrés [2,3] provides summaries of relevant papers on the subject. However, available analyses only apply to ''pristine'' bearing geometries. The authors are not aware of applicable analyses for altered geometries due to a plugged orifices or wear.

The rotordynamic performance of a hybrid bearing is determined by the stiffness, damping and inertia coefficients of the motion/reaction-force model established for a small displacement about a static equilibrium position:

$$-\begin{Bmatrix}\Delta F_x\\ \Delta F_y\end{Bmatrix}=\begin{bmatrix}K_{xx}(\epsilon_0) & K_{xy}(\epsilon_0)\\ K_{yx}(\epsilon_0) & K_{yy}(\epsilon_0)\end{bmatrix}\begin{Bmatrix}\Delta x\\ \Delta y\end{Bmatrix}+\begin{bmatrix}C_{xx}(\epsilon_0) & C_{xy}(\epsilon_0)\\ C_{yx}(\epsilon_0) & C_{yy}(\epsilon_0)\end{bmatrix}\times\begin{Bmatrix}\Delta\dot{x}\\ \Delta\dot{y}\end{Bmatrix}+\begin{bmatrix}M_{xx}(\epsilon_0) & M_{xy}(\epsilon_0)\\ M_{yx}(\epsilon_0) & M_{yy}(\epsilon_0)\end{bmatrix}\begin{Bmatrix}\Delta\ddot{x}\\ \Delta\ddot{y}\end{Bmatrix}. \quad (1)$$

Diagonal terms are called direct coefficients, while off-diagonal terms are called cross-coupled coefficients. Cross-coupled coefficients result from fluid rotation within the bearing.

## Test Facility

The test rig and facilities have been fully described by Kurtin et al. [4] and Childs and Hale [5]. Childs and Hale [5] performed uncertainty analysis on the experimental data. Figure 1 provides an end view of the test bearing with an applied static side load. The load direction establishes the $+X$-axis. The test rotor is supported by hydrostatic bearings, and the test bearing supports itself on the test rotor.

## Bearing Designs

Separate modifications have been made on a reference bearing to obtain plugged-orifice and worn conditions. The reference bearing is 76.4 mm in diameter, 76.2 mm in length, with a 0.1016-mm radial clearance. The five square recesses are 27 mm wide and 2.032 mm deep. The nominal radial clearance is 0.1016 mm.

The basic geometry of the bearing stator is shown in Fig. 2. The end view of the bearing shows five orifice/recess locations. Starting at 0 degrees and proceeding in a counterclockwise direction around the bearing, recess 1, recess 2, etc., are encountered. The load direction ($X$-axis) is at 225 degrees, opposite recess 1. Hence, orifice/recess 1 will have the smallest film thickness

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-398. Manuscript received by IGTI, Nov. 1999; final revision received by ASMA Headquarters, Feb. 2000. Associate Editor: D. Wisler.

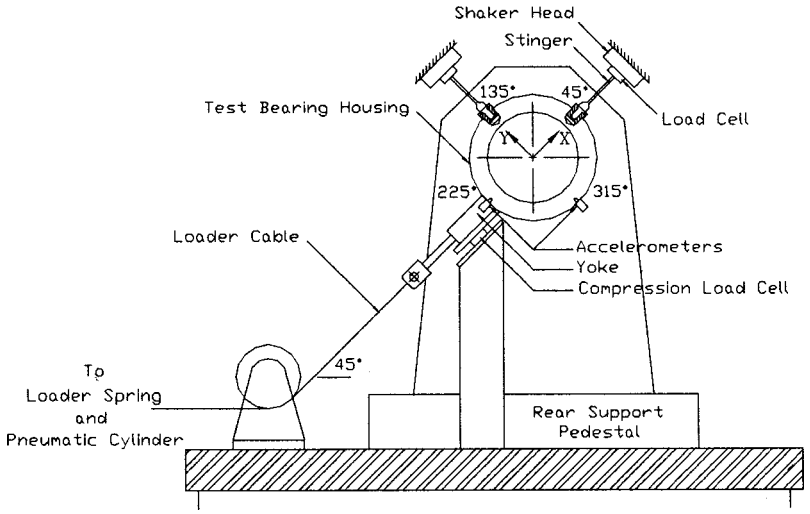



**Fig. 1 Test-section end view**

during a loaded (eccentric) condition. The bearing is expected to be most sensitive to degradation of this orifice/recess; hence, orifice 1 was plugged, and stator wear was centered about this recess.

Figure 3 illustrates the simulated worn bearing design. The clearance is increased locally to produce a maximum increase of 65 percent of the concentric radial clearance at the center of recess 1. The circular arc of the simulated wear surface has the radius of the shaft. The increased clearance varies in the circumferential direction but is uniform across the bearing-land surface. All orifices are open when testing this bearing.

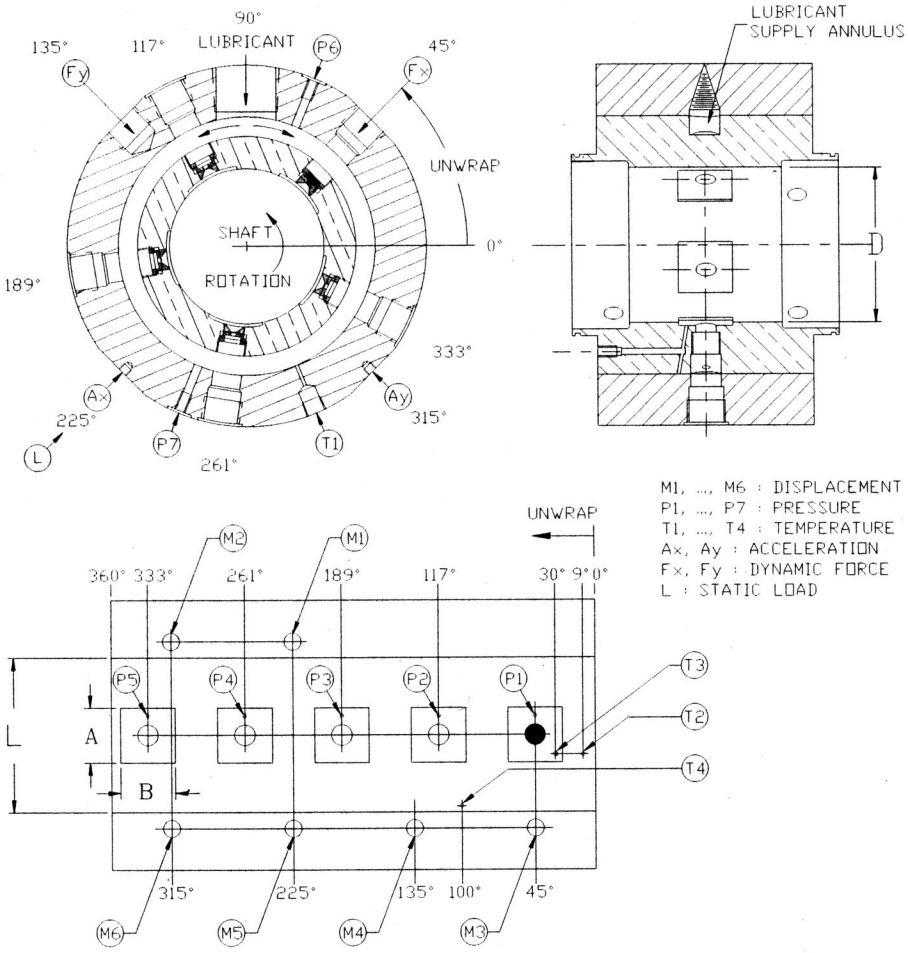



**Fig. 2 Plugged-orifice test bearing geometry**

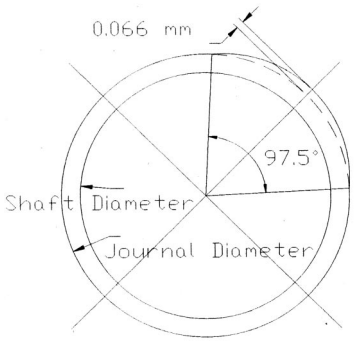



**Fig. 3 Wear on the bearing land geometry**

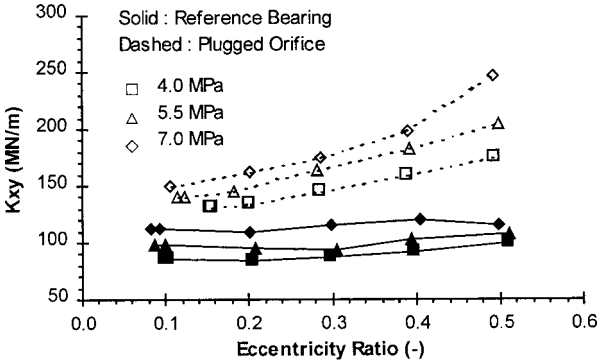



**Fig. 5 Cross-coupled stiffness at high-speed versus eccentricity ratio and supply pressure for the plugged-orifice test**

## Test Procedures

The following parameter ranges define the test matrix: (a) three running speeds (10,200, 17,400, and 24,600 rpm), (b) three supply pressures (4.0, 5.5, and 7.0 MPa), and (c) six eccentricity ratios (0.0 to 0.5 in 0.1 increments).

This test matrix produces considerably more data than can be presented here. Hence data are only presented at the highest speed, which is most representative of the intended turbopump application. Results are presented versus eccentricity ratio and supply pressure. Each figure shows experimental data for both the reference bearing and the damaged bearing. For the worn bearing, the eccentricity-ratio definition is referenced to the undamaged geometry.

The results presented focus on the main parameters and do not include comments on cross-coupled damping and mass coefficients due to their negligible changes with bearing degradations and high uncertainties on their measurements.

## Experimental Results With the Plugged Orifice

During tests of the damaged bearings, the authors had some difficulties holding a static eccentric position, presumably because of the asymmetric orifices distribution. These difficulties were most prominent near the centered position and did not arise with the reference bearing. Observed changes in bearing coefficients are strongest in the direct stiffness and damping ($K_{xx}$ and $C_{xx}$) coefficients and less evident in the cross-coupled stiffnesses coefficients ($K_{xy}$ and $K_{yx}$). $K_{yy}$ and $C_{yy}$ are largely unchanged when the orifice is plugged. Given this outcome, results are only presented for $K_{xx}$, $K_{yy}$, and $C_{xx}$.

**Pressure and Supply Flow Rate.** Plugging orifice 1 while maintaining a constant supply pressure reduces all recess pressures in comparison to the reference bearing. For recess 1, the drop is 60 percent (concentric positions) to −30 percent (high eccentricity). For the remaining recesses the drop is 30 percent (all cases).

The flow rate drops by 17 percent when one orifice is plugged. This difference holds over the entire test matrix and is near the expected drop due to a 20 percent reduction in flow area.

Absent an analysis for bearings with a plugged orifice or wear, no comparisons are presented between theory and measurements.

**Direct Stiffness.** Figure 4 shows that $K_{xx}$ is much more sensitive to changes in the eccentricity ratio with a plugged orifice. Although not shown here, $K_{yy}$ increased slightly, by an overall average of 9 percent, and does not show the same sensitivity to eccentricity ratio. Plugging orifice 1 reduces $K_{xx}$ at low eccentricities; in the centered position, the average difference is −23 percent. Increasing the eccentricity ratio causes a rapid increase in $K_{xx}$, which reaches 450 MN/m (+130 percent) at the highest pressure. This difference at high eccentricity ratios increases with increasing running speed.

**Cross-Coupled Stiffness.** Figure 5 shows the sensitivity of the cross-coupled stiffness $K_{xy}$ to changes in the eccentricity ratio. Trends are the same for $K_{yx}$, but less pronounced. Overall, within the test matrix $K_{xy}$ is higher by 62 percent and $K_{yx}$ by 38 percent. The percentage difference does not vary significantly with running speed or supply pressure. Plugging the orifice clearly increases the sensitivity of $K_{xy}$ to changes in the eccentricity ratio.

**Direct Damping.** From Fig. 6, plugging orifice 1 increases $C_{xx}$ generally, and causes $C_{xx}$ to increase more rapidly with increasing eccentricity ratio. Although not illustrated, $C_{yy}$ is not changed appreciably. At high eccentricity ratios, the increase in $C_{xx}$ is more pronounced at low and medium speeds (up to 400

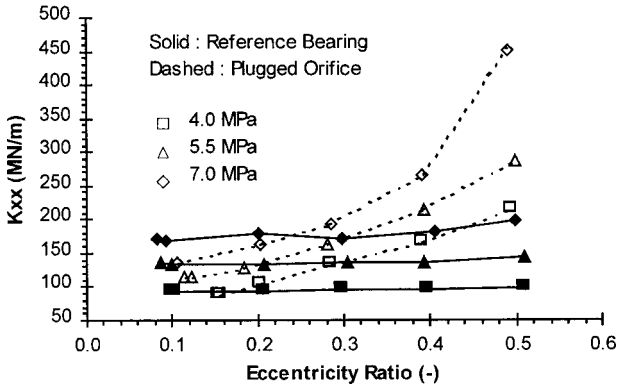



**Fig. 4 Direct stiffness at high-speed versus eccentricity ratio and supply pressure for the plugged-orifice test**

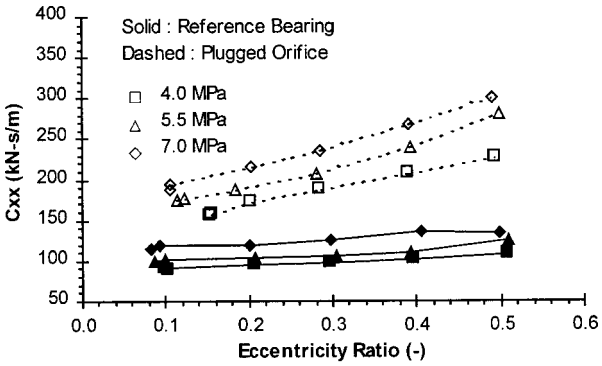



**Fig. 6 Direct damping at high-speed versus eccentricity ratio and supply pressure for the plugged-orifice test**

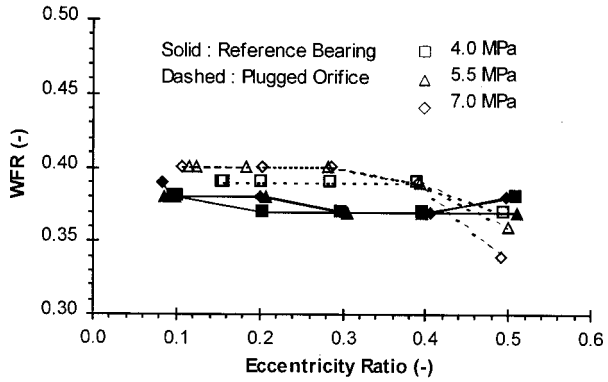



**Fig. 7 Whirl frequency ratio at high-speed versus eccentricity ratio and supply pressure for the plugged-orifice test**

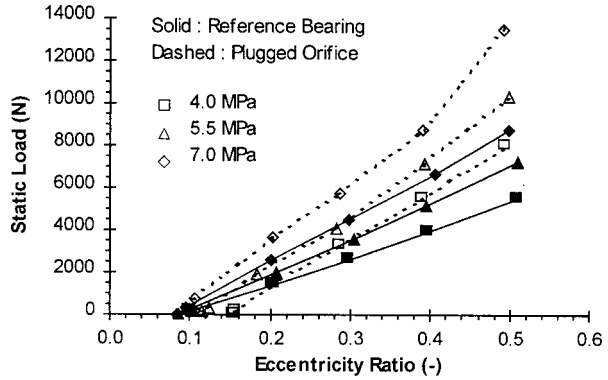



**Fig. 8 Static load at high-speed versus eccentricity ratio and supply pressure for the plugged-orifice test**

kN-s/m) than at high speed. The difference goes from +50 percent at concentric positions to +150 percent at high eccentricity ratio.

**Whirl Frequency Ratio.** As presented in Fig. 7, plugging orifice 1 causes only a slight increase in WFR (as defined by Lund [6]). The increases of cross-coupled stiffness and direct damping largely compensate, so that stability as predicted by the WFR is largely unchanged.

**Load Capacity.** Figure 8 presents static load versus eccentricity data for the reference and plugged-orifice bearings. Note that both bearings have the same eccentricity at zero load. This result is somewhat misleading because the reference bearing would center itself, and the plugged-orifice bearing would not. Keeping this in mind, Fig. 8 suggests that the reference bearing has better "centering" capacity at low eccentricities, and the plugged-orifice bearing has better load capacity at high eccentricities.

## Experimental Results for the Worn Bearing

**Pressure and Supply Flow Rate.** Recess pressure (P1) where the damaged sector is centered drops by 15 percent (less resistance due to the larger clearance). Opposite the wear area, pressures drop 8 percent at recess 4 and 30 percent at recess 3.

The worn-bearing flow is slightly higher (from +5 percent to 10 percent) than the reference bearing. This result holds for all the supply pressure/running speed cases.

**Direct Stiffness.** Figure 9 shows that the simulated wear does not change $K_{xx}$ much at 24,600 rpm. The eccentricity ratio in this and subsequent figures relates to the original bearing geometry, not the damaged geometry. At lower speeds and supply pressures, $K_{xx}$ and $K_{yy}$ are both reduced. The maximum reduction is 20 percent at low speed and high pressure. The direct stiffness is not sensitive to changes in the eccentricity ratio.

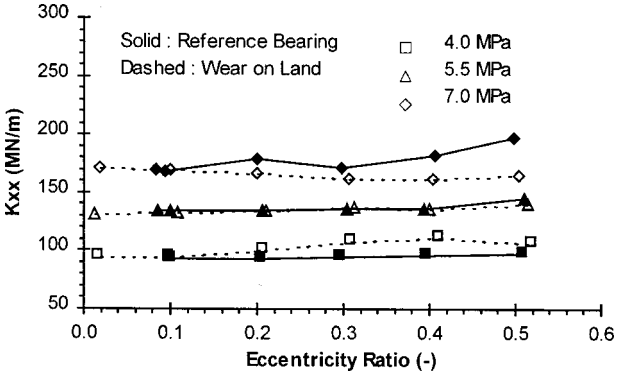



**Fig. 9 Direct stiffness at high-speed versus eccentricity ratio and supply pressure for the worn land surface test**

**Cross-Coupled Stiffness.** Figure 10 shows that $K_{xy}$ is also largely unmodified by the presence of wear at 24,600 rpm. Both $K_{xy}$ and $K_{yx}$ are reduced by 5–10 percent.

**Direct Damping.** Figure 11 shows $C_{xx}$ to be modestly sensitive to changes in the eccentricity ratio. The direct damping coefficients for the worn bearing are slightly higher near the centered position (+5 percent) and slightly lower at an eccentricity ratio of 0.5 (−10 percent) than the reference bearing.

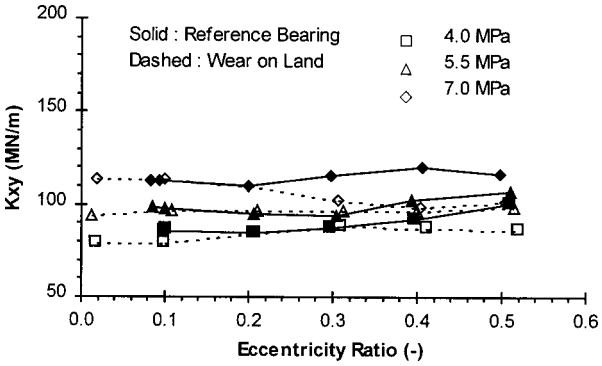



**Fig. 10 Cross-coupled stiffness at high-speed versus eccentricity ratio and supply pressure for the worn land surface test**

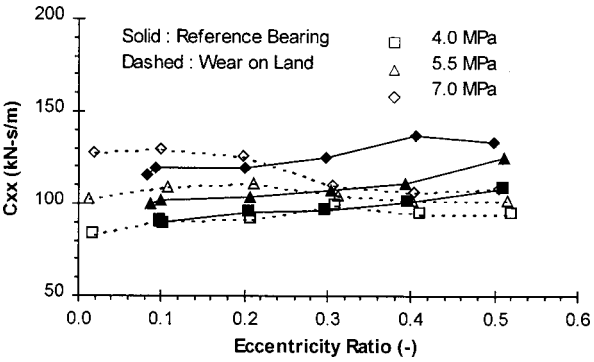



**Fig. 11 Direct damping at high-speed versus eccentricity ratio and supply pressure for the worn land surface test**

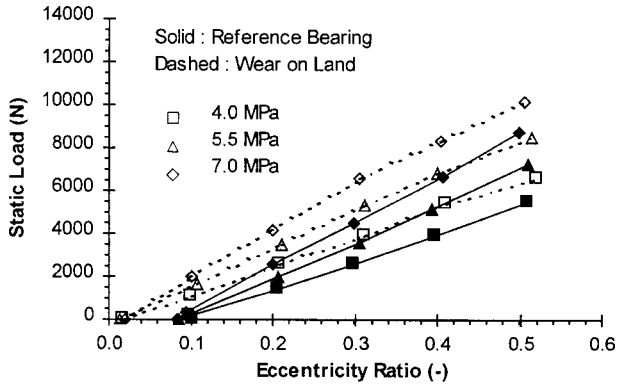



**Fig. 12 Static load at high-speed versus eccentricity ratio and supply pressure for the worn land surface test**

**Whirl Frequency Ratio.** The WFR is unchanged by the simulated wear. The constancy of the cross-coupled stiffness and direct-damping coefficients explains this result.

**Load Capacity.** The slopes of the static load versus eccentricity ratio presented in Fig. 12 are similar for both bearings. Curves plotted from the test results are offset at their starting points. When the base (zero load) points are superposed, the load capacity is generally not influenced by the simulated wear on the land. These results are consistent with the direct-stiffness measurements.

## Uncertainty Analysis

Experimental results depend on the following two types of errors: uncertainties from the calibration of each transducer and the analog-to-digital conversion process in the data-acquisition system and deviations from curve fitting impedances to determine the rotordynamic coefficients. Childs and Hale [5] performed an uncertainty analysis in agreement with the ANSI/ASME standard on Measurement and Uncertainty (1986). Table 1 gives the average standard deviations of the present tests with the percentages from initial values. Errors are acceptable for stiffness and direct damping coefficients but larger for the cross-coupled damping and inertia terms.

The difficulties in centering the damaged bearing are not directly reflected in the uncertainties. Since none of the rotordynamic coefficients are sensitive to the static eccentricity ratio, the small offsets experienced with the degraded bearings would have a minor impact on the uncertainties.

## Discussion

**Plugged-Orifice Bearing.** Experimental results for one plugged orifice for a five-recess bearing give a good feeling of the limits of hydrostatic bearings in case the orifice of the minimum clearance recess is plugged. The dynamic parameters are not much affected near the concentric position. Increasing the eccentricity in the direction of the load and the plugged orifice causes a strong increase in the direct stiffness and damping coefficients with an increase in eccentricity ratio, presumably because of an enhancement in the hydrodynamic behavior of the bearing.

**Table 1 Average standard deviations and percentages from nominal values for the rotordynamic coefficients**

| | Curve-Fit Deviations | Percentages |
|---|---|---|
| $K$ (MN/m) | ±2.4 MN/m | ±2.0 percent |
| $k$ (MN/m) | ±1.4 MN/m | ±1.7 percent |
| $C$ (kN-s/m) | ±2.5 kN-s/m | ±1.2 percent |
| $c$ (kN-s/m) | ±0.9 kN-s/m | ±7.3 percent |
| $M$ (kg) | ±2.2 kg | ±15.4 percent |
| $m$ (kg) | ±0.8 kg | ±11.0 percent |

**Worn Bearing.** The static and dynamic coefficients were only slightly altered by simulated wear. The observed differences should be within the running margin of the machines. The flow rate increases due to a larger leakage flow area. Dynamics coefficients decrease by around 5 percent to 15 percent (the decrement increases with increasing eccentricity). The trends versus eccentricity ratio, running speed, and supply pressure are similar with and without wear.

Scharrer et al. [1] have published the only theoretical results known by the authors. Their solution is based on the Reynolds equation with an analytical expression of the film clearance, and they analyzed an 8-recess hybrid-bearing running at 23,600 rpm with cryogenic fluid (LH2) for a 0 to 100 percent of clearance-wear range. Their theoretical case and the present experimental case are not comparable due to completely different fluid characteristics. Specifically, their model does not include the convective and temporal acceleration terms of the Navier-Stokes equations, and these terms have been shown to be significant for the range of Reynolds numbers produced in this test program (Childs and Hale [5]). San Andres' models are entirely appropriate but do not include either plugged-orifice or worn-bearing capabilities. The experimental results show negligible reductions in the direct parameters, versus Scharrer's predictions of a 25 percent reduction.

## Summary and Conclusions

From the tests performed, the plugged-orifice bearing showed large increases in direct and cross-coupled stiffnesses and direct damping at higher values of eccentricity ratios. No similar increases were observed with the worn bearing. Load capacity was not significantly influenced by large simulated wear on the land.

Overall, these test data suggest that neither a single plugged orifice nor significant wear on the bearing land will "disable" a well-designed hybrid bearing. These results do not speak to multiple plugged orifices and are not an endorsement for operations without filters to prevent plugging orifices. Also, the wise selection of materials to minimize wear continues to be advisable.

## Acknowledgments

The work reported in this paper was carried out while the first author was supported at the Turbomachinery Laboratory by SEP, a division of SNECMA.

## Nomenclature

$C_{xx}, C_{yy}$ = direct damping coefficients $(FT/L)$
$C_{xy}, C_{yx}$ = cross-coupled damping coefficients $(FT/L)$
$C, c$ = damping coefficients at $\epsilon_0 = 0$ $(FT/L)$
$C_r$ = radial clearance $(L)$
$e_0$ = journal eccentricity $(L)$
$K_{xx}, K_{yy}$ = direct stiffness coefficients $(F/L)$
$K_{xy}, K_{yx}$ = cross-coupled stiffness coefficients $(F/L)$
$K, k$ = stiffness coefficients at $\epsilon_0 = 0$ $(F/L)$
$M_{xx}, M_{yy}$ = direct inertia coefficients $(M)$
$M_{xy}, M_{yx}$ = cross-coupled inertia coefficients $(M)$
$M, m$ = inertia coefficients at $\epsilon_0 = 0$ $(M)$
$P_r$ = pressure ratio $P_{ra}/P_s$
$P_{ra}$ = average recess pressure $(F/L^2)$
$P_s$ = supply pressure $(F/L^2)$
$\dot{Q}$ = supply flow rate $(L^3/T)$
WFR = whirl frequency ratio
$\epsilon_0$ = eccentricity ratio $e_0/C_r$

$\Delta F_x, \Delta F_y$ = reaction forces ($F$)
$\Delta x, \Delta y$ = relative displacement between stator and rotor ($L$)
$\Delta \dot{x}, \Delta \dot{y}$ = relative velocity between stator and rotor ($L/T$)
$\Delta \ddot{x}, \Delta \ddot{y}$ = relative acceleration between stator and rotor ($L/T^2$)

## References


[1] Scharrer, J. K., Hecht, R. I., and Hibbs, Jr., R. I., 1991, "The Effects of Wear on the Rotordynamic Coefficients of a Hydrostatic Journal Bearing," ASME J. Tribol., **113**, pp. 210–213.

[2] San Andrés, L. A., 1990, "Turbulent Hybrid Bearings With Fluid Inertia Effects," ASME J. Tribol., **112**, pp. 699–707.

[3] San Andrés, L. A., 1992, "Analysis of Turbulent Hydrostatic Bearings With a Barotropic Fluid," ASME J. Tribol., **114**, pp. 755–765.

[4] Kurtin, K. A., Childs, D. W., San Andrés, L. A., and Hale, R. K., 1991, "Experimental Versus Theoretical Characteristics of a High-Speed Hybrid (Combination Hydrostatic and Hydrodynamic) Bearing," ASME J. Tribol., **115**, pp. 160–169.

[5] Childs, D. W., and Hale, R. K., 1994, "A Test Apparatus and Facility to Identify the Rotordynamic Coefficients of High-Speed Hydrostatic Bearings," ASME J. Tribol., **116**, pp. 337–344.

[6] Lund, J. W., 1966, "The Stability of an Elastic Rotor in Journal Bearings With Flexible Damped Supports," ASME J. Appl. Mech., **27**, pp. 911–920.

**J. A. Vázquez**[1]
e-mail: Jose@virginia.edu

**L. E. Barrett**
e-mail: leb@virginia.edu

**R. D. Flack**
e-mail: rdf@virginia.edu

Department of Mechanical, Aerospace and Nuclear Engineering and Applied Science, University of Virginia, Charlottesville, VA 22903

# Flexible Bearing Supports, Using Experimental Data

*A laboratory rotor, representing a scaled-down model of a three-stage compressor supported by fluid film bearings on anisotropic flexible supports was analyzed. The support characteristics were measured at the bearing locations by exciting the bearing housings with electromechanical shakers and measuring the acceleration. Direct, cross-coupled, and cross-talk accelerance between supports were measured. Unbalance response and stability analyses of the rotor were performed using polynomial transfer functions extracted from the measured accelarance data. The predicted critical speeds and instability threshold agree with measured data. Predictions using other support models are included to show the effectiveness of this method.* [DOI: 10.1115/1.1426085]

## Introduction

The bearing support structure may have a large effect on the behavior of rotating machinery. The dynamic behavior of the support structure combines with the bearings stiffness and damping coefficients and modifies the impedance observed by the rotor at the bearing locations. API standards 617 for compressors ([1]), Section 2.9.2.4.f, recommends including the effects of the support structure in unbalance response analyses when the ratio of the support to bearing stiffness ratio is less than or equal to 3.5. This section specifies the use of calculated frequency-dependent support stiffness and damping values or the values derived from modal testing. Unfortunately, the standard does not indicate how these stiffness and damping values should be calculated or what type of modal testing should be used. Also, the API standard does not require stability analyses for the rotor system.

The calculation of frequency-dependent support stiffness and damping values is, at best, very difficult because of the complexity of the casings and bearing supports. In most cases, the analyst's only recourse is to determine the dynamic characteristics of the support structure experimentally. The most common (and least expensive) experimental testing is the measurement of frequency response functions (FRFs) of the support structure at the bearing locations. For these types of measurements, the supports are excited (by impact or with electromechanical shakers) at the bearing locations and the response is measured at the bearings. These experiments are performed in the vertical and horizontal directions and between supports. The measured FRFs are then used to create equivalent physical models of the supports, either as mass supports ([2–6]) or with many degrees-of-freedom ([7,8]). Instead of FRF, Feng and Hahn [9], Lees and Friswell [10], Lees et al. [11], and Edwards et al. [12] use machine unbalance response to calculate the foundation modal parameters. These parameters are then used to create equivalent foundation models.

An alternative method to include the dynamics of the support structure is to calculate polynomial transfer functions from the measured FRF's. These transfer functions can be used for unbalance response and stability analysis ([13,14]).

A research work was initiated to answer the question of how much the dynamics of the support structure affects the rotor bearing system and how accurately one can predict this effect. To that end, a flexible rotor supported by fluid film bearings on flexible supports was tested for unbalance response and stability. The research included one set of tilting pad bearings and two sets to three-lobe bearings and 15 support configurations. The support structure was designed such that the support stiffness in the horizontal direction could be parametrically changed while the support stiffness in the vertical direction remained nearly constant. Vázquez [15] reported the complete results of this research. Vázquez et al. [16,17] reported the results for one set of three-lobe bearings for all the support configurations.

The work presented here concentrates on one set of three-lobe bearings and one support configuration. More detailed information is provided for the support configuration presented. The numerical predictions have been extended to study the effects of cross coupling in the support structure between the vertical and horizontal direction and the effect of cross talk between pedestals. Also, a single mass support representation of the support structure was created using the static stiffness of the supports at the bearing locations and the mass of the bearings and bearing housings.

## Experimental Apparatus

The test apparatus consists of a flexible rotor with three disks equally spaced between two identical fluid film bearings. The bearings are supported by anisotropic flexible supports. The entire apparatus is mounted on a heated, oil-filled reservoir and clamped to a 900-kg concrete block. This concrete foundation is isolated from the laboratory floor by rubber pads.

The rotor represents a scaled-down model of a three-stage compressor. The shaft has a maximum diameter of 25.4 mm and a minimum diameter of 19.0 mm. The total length of the shaft is 654.0 mm with a bearing span of 520.7 mm. Three rigid disks, 152.4 mm in diameter, representing the compressor stages, are rigidly attached to the rotor between the bearing locations. The total weight of the rotor is 12.64 kg. The first and second critical speeds of the system are located at 2550 rpm and 4130 rpm, respectively. A one-horsepower motor drives the rotor through a flat belt and flexible coupling. Any rotor speed up to 11000 rpm can be obtained. Figure 1 shows the experimental setup.

Two identical three lobe bearings were used in this work. Table 1 shows the bearing characteristics and Fig. 2 shows the eight calculated linear bearing stiffness and damping coefficients ([18]). The bearing supports consist of rigid housings mounted on flexible elements. The flexible element design was based on a beam-type construction with stiffening plates connected at the ends. It was constructed of aluminum. The oil drained from the bearing housing passed through drainpipes in the flexible element to the oil reservoir located under the assembly. Figure 3 shows a detail of the bearing housing with the flexible element. Figure 4 shows the design of the flexible support element with the stiffening plates. Changing the stiffening plates parametrically varies the horizontal stiffness of the support element while the vertical stiff-

[1]Currently at DuPont Engineering Technologies, Wilmington, DE.

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-404. Manuscript received by IGTI, Nov. 1999; final revision received by ASME Headquarters, Feb. 2000. Associate Editor: D. Wisler.

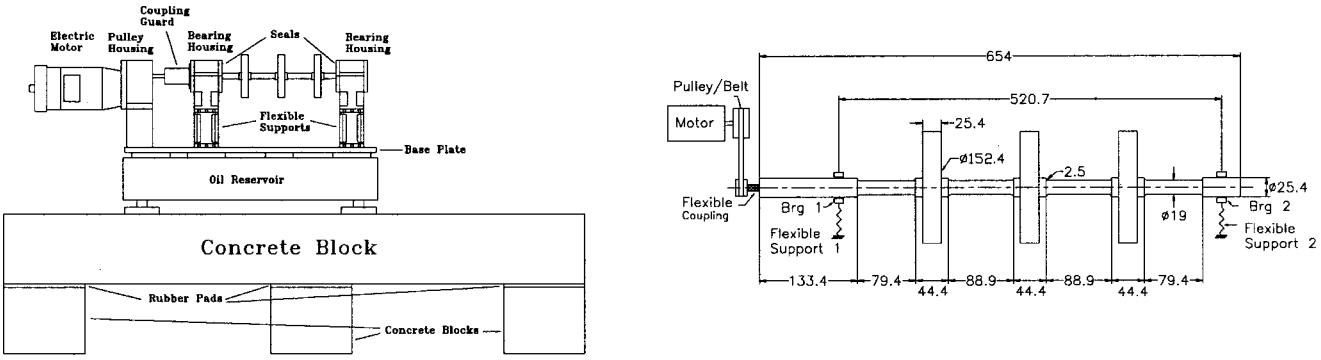



**Fig. 1 Experimental setup**

ness remains constant. The stiffening plates used in this work were made of aluminum with a width of 31.75 mm and a thickness of 0.762 mm.

## Flexible Supports Characteristics

The dynamic characteristics of the support were determined experimentally. Electromechanical shakers were attached to the bearing housing through piezoelectric force transducers and a thin

**Table 1 Bearing characteristics**

| | |
|---|---|
| Journal Diameter (mm) | 25.400 (±0.0025) |
| Lobe Radial Clearance (mm) | 0.0833 (±0.0017) |
| Bearing Radial Clearance (mm) | 0.0330 (±0.0017) |
| Lobe Length (mm) | 12.700 (±0.0025) |
| Bearing preload factor | 0.604 (±0.023) |
| Pivot offset factor | 0.505 (±0.012) |
| Lobe arc length (deg.) | 93 (±1) |
| Lobe loading condition (deg.) | 0.0 (±2.5) |
| Static load (N) | 62.2 (±0.1) |
| Inlet Pressure (kPa) | 20.7 (±0.1) |
| Inlet Temperature (° C) | 48.3 (±0.5) |
| Oil viscosity (Pa.s)@ 40 ° C | 0.0262 |
| Oil viscosity (Pa.s)@ 55 ° C | 0.0159 |
| Oil density (kg/m$^3$) | 908 |

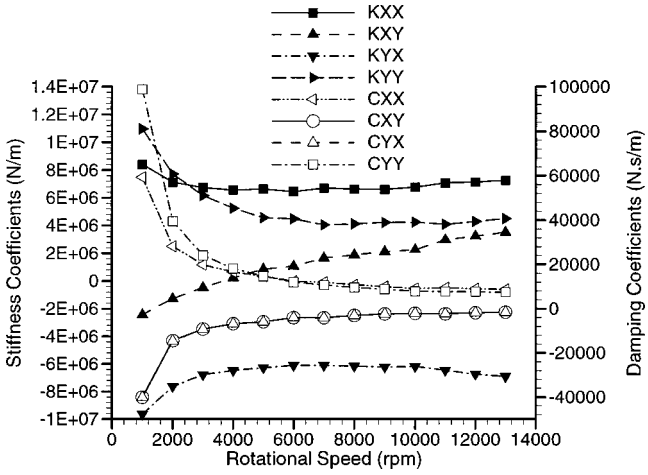



**Fig. 2 Bearing stiffness and damping coefficients**

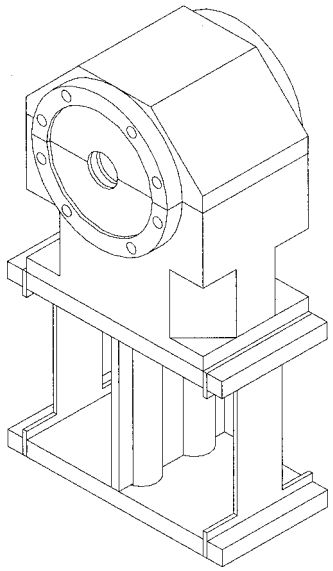


**Fig. 3 Details of the bearing housing and flexible support**

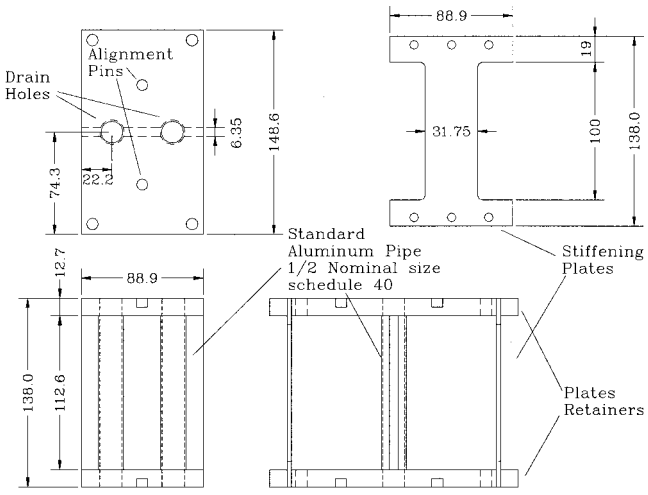



**Fig. 4 Flexible support element design (dimensions are in mm)**

rod. The use of the thin rod minimized the effects of misalignment between the shaker and the bearing housing. A sine sweep excitation was used and the acceleration of the bearing housings was measured in the vertical and horizontal directions. The excitation was applied to each support, one direction at a time. A total of 16 accelarances were measured, corresponding to the direct, cross-coupling, and cross-talk between the bearing housings. An accelerance matrix can be assembled for this support configuration as

$$\begin{Bmatrix} \ddot{x}_1 \\ \ddot{y}_1 \\ \ddot{x}_2 \\ \ddot{y}_2 \end{Bmatrix} = \underbrace{\begin{bmatrix} a_{x1x1} & a_{x1y1} & a_{x1x2} & a_{x1y2} \\ a_{y1x1} & a_{y1y1} & a_{y1x2} & a_{y1y2} \\ a_{x2x1} & a_{x2y1} & a_{x2x2} & a_{x2y2} \\ a_{y2x1} & a_{y2y1} & a_{y2x2} & a_{y2y2} \end{bmatrix}}_{[A(\omega)]} \begin{Bmatrix} F_{x1} \\ F_{y1} \\ F_{x2} \\ F_{y2} \end{Bmatrix}. \quad (1)$$

The subscripts 1 and 2 refer to support 1 and 2, respectively. In order to use the experimental data several steps are necessary. First, the accelerance data is integrated twice to get the dynamic compliance of the supports. This is not strictly necessary but it is convenient because most rotordynamic programs use displacements and rotations as the unknown variables. After the integration, Eq. (1) becomes

$$\begin{Bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{Bmatrix} = \underbrace{\begin{bmatrix} dc_{x1x1} & dc_{x1y1} & dc_{x1x2} & dc_{x1y2} \\ dc_{y1x1} & dc_{y1y1} & dc_{y1x2} & dc_{y1y2} \\ dc_{x2x1} & dc_{x2y1} & dc_{x2x2} & dc_{x2y2} \\ dc_{y2x1} & dc_{y2y1} & dc_{y2x2} & dc_{y2y2} \end{bmatrix}}_{[DC(\omega)]} \begin{Bmatrix} F_{x1} \\ F_{y1} \\ F_{x2} \\ F_{y2} \end{Bmatrix} \quad (2)$$

where $[DC(\omega)]$ is the dynamic compliance matrix of the support structure. This matrix is defined for each of the frequencies used for the excitation of the supports. The dynamic compliance matrix can be used directly for unbalance response but fails to capture the support dynamics for stability analysis ([5]). An extra step is necessary. Polynomial transfer functions are calculated from each of the terms of the dynamic compliance matrix. These transfer functions can be calculated using one of the several system identification techniques available ([19–23]). Equation (2) becomes

$$\begin{Bmatrix} x_1 \\ y_1 \\ x_2 \\ y_2 \end{Bmatrix} = \underbrace{\begin{bmatrix} g_{x1x1} & g_{x1y1} & g_{x1x2} & g_{x1y2} \\ g_{y1x1} & g_{y1y1} & g_{y1x2} & g_{y1y2} \\ g_{x2x1} & g_{x2y1} & g_{x2x2} & g_{x2y2} \\ g_{y2x1} & g_{y2y1} & g_{y2x2} & g_{y2y2} \end{bmatrix}}_{[G(s)]} \begin{Bmatrix} F_{x1} \\ F_{y1} \\ F_{x2} \\ F_{y2} \end{Bmatrix}. \quad (3)$$

Equation (3) defines the transfer function matrix (TFM) of the support structure. The TFM is defined as a function of the complex variable $s$ and it is valid in the whole complex plane. The elements of the TFM are of the form

$$g_{ij} = \frac{a_n s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0}{s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_0}. \quad (4)$$

Richardson and Formenti [20,21] present a method to select the order of the polynomials for the fitting process. A good starting point for the order of the denominator is to count the number of peaks (modes) in the range of interest and multiply the result by two (second order differential equations).

Table 2 shows the coefficients for a typical polynomial transfer function. This table shows the direct transfer function of support 1 in the horizontal direction ($g_{x1x1}$). The denominator is of order 24 and the numerator of order 22.

For unbalance response $s = i\omega_s$. For stability analysis $s = p \pm i\omega_d$. Flexible support equations require the stiffness and damping coefficients (they can also be combined to form complex stiffness coefficients). These coefficients are obtained by inverting the transfer function matrix for each complex frequency $s_i$ and extracted as

**Table 2 Polynomial coefficients for $G_{x1x1}$**

| Power | Numerator | Denominator |
|---|---|---|
| $s^0$ | -7.5883E+52 | -8.1376E+58 |
| $s^1$ | -2.9126E+50 | -3.4452E+56 |
| $s^2$ | -3.2889E+49 | -4.2405E+55 |
| $s^3$ | -1.0143E+47 | -1.4275E+53 |
| $s^4$ | -5.2538E+45 | -8.1402E+51 |
| $s^5$ | -1.2383E+43 | -2.1028E+49 |
| $s^6$ | -3.8553E+41 | -7.1844E+47 |
| $s^7$ | -6.5225E+38 | -1.3568E+45 |
| $s^8$ | -1.3427E+37 | -3.0525E+43 |
| $s^9$ | -1.5143E+34 | -4.0103E+40 |
| $s^{10}$ | -2.1393E+32 | -6.2515E+38 |
| $s^{11}$ | -1.4771E+29 | -5.4496E+35 |
| $s^{12}$ | -1.4199E+27 | -6.0282E+33 |
| $s^{13}$ | -4.7764E+23 | -3.3079E+30 |
| $s^{14}$ | -2.1526E+21 | -2.4479E+28 |
| $s^{15}$ | 1.0983E+18 | -5.3914E+24 |
| $s^{16}$ | 1.7481E+16 | 7.7228E+21 |
| $s^{17}$ | 1.0298E+13 | 3.2525E+19 |
| $s^{18}$ | 8.5890E+10 | 4.1830E+17 |
| $s^{19}$ | 2.1541E+07 | 1.6895E+14 |
| $s^{20}$ | 1.3894E+05 | 1.4336E+12 |
| $s^{21}$ | 1.5569E+01 | 2.8297E+08 |
| $s^{22}$ | 7.6516E-02 | 1.9865E+06 |
| $s^{23}$ | 1.8540E-06 | 1.6216E+02 |
| $s^{24}$ | 0.0000E+00 | 1.0000E+00 |

$$[K_{dyn}(s)] = [G(s)]^{-1}$$
$$[C_{\sup}] = \frac{\mathrm{Im}([K_{dyn}(s)])}{\mathrm{Im}(s)} \quad (5)$$
$$[K_{\sup}] = \mathrm{Re}([K_{dyn}(s)]) - \frac{\mathrm{Re}(s)}{\mathrm{Im}(s)} \mathrm{Im}([K_{dyn}(s)]).$$

These support stiffness and damping coefficients are used with the standard methods to include the support effects in rotordynamic analysis. However, the variation of the coefficients with the complex frequency has to be included in the analysis tool.

Figures 5 through 7 show the magnitude of the measured dynamic compliance of the support structure (the dynamic compliance is the double integration of the accelerance and therefore considered as measured data). Because the support structure is a passive system, the dynamic compliance matrix (Eq. (2)) must be symmetric. Therefore, it is sufficient to show the upper right half of the matrix. Figure 5 shows the magnitude of the direct and cross talk dynamic compliance in the horizontal direction. Both $|DC_{x1x2}|$ and $|DC_{x2x1}|$ are shown to verify the symmetry condition. This figure shows that the direct response in both supports is similar. Support 2 is more flexible at the 4200 cpm resonance. It is important to note that the magnitude of the crosstalk dynamic compliance is in the same order of magnitude as the magnitude of the direct dynamic compliance. This indicates that the crosstalk between supports has a large influence in the response of the rotor.

Figure 6 shows the magnitude of the measured direct and crosstalk support dynamic compliance in the vertical direction. This figure also confirms the symmetry of the dynamic compliance matrix. The dynamic compliance in the vertical direction is one order of magnitude smaller than in the horizontal direction. This indicates that the stiffness of the support is one order of

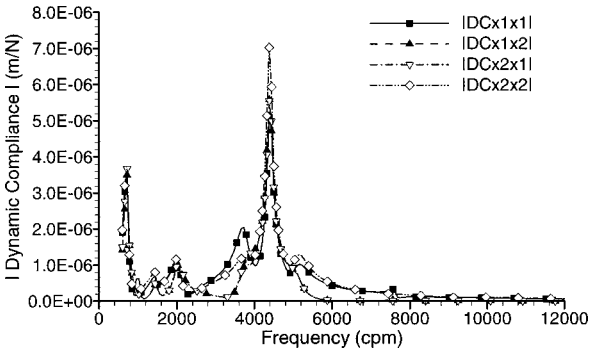



**Fig. 5 Magnitude of the measured direct and crosstalk support dynamic compliance in the horizontal direction**

magnitude larger in the vertical direction than in the horizontal direction. This shows the anisotropy of the support structure.

Figure 7 shows the magnitude of the cross coupling terms of the dynamic compliance matrix. The magnitude of the cross-coupling term is in the same order of magnitude as the direct response in the vertical direction and one order of magnitude smaller than the direct response in the horizontal direction. This indicates (and will be shown later) that the cross-coupling does not have an effect as large as the crosstalk between supports.

The static stiffness of the supports was obtained by extrapolating the measured dynamic compliance data to zero frequency. This is easily accomplished once polynomial transfer functions are fitted to the dynamic compliance data. To get the support static stiffness just calculate

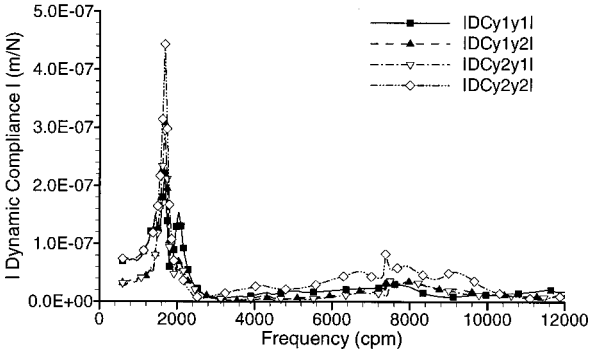



**Fig. 6 Magnitude of the measured direct and crosstalk support dynamic compliance in the vertical direction**

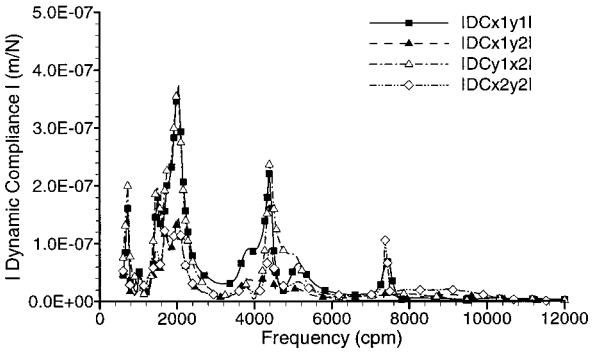



**Fig. 7 Magnitude of the measured cross-coupling dynamic compliance**

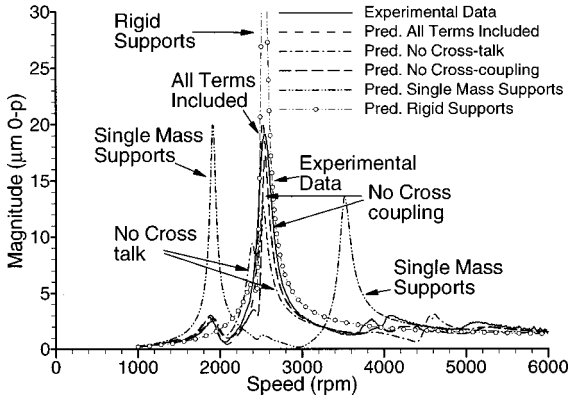



**Fig. 8 Measured and predicted unbalance response near the middle disk in the horizontal direction. Unbalance distribution 1**

$$[K_{dyn}(0)]=[G(0)]^{-1}$$
$$\text{if } \mathrm{Im}([K_{dyn}(0)])=0 \tag{6}$$
$$[K_{\sup}]=\mathrm{Re}([K_{dyn}(0)]).$$

Equation (6) is a special case of Eq. (5). For this equation to be valid, the support dynamic stiffness matrix must be real at zero frequency.

The static stiffness of support 1 is $1.07\ 10^6$ N/m in the horizontal direction and $1.66\ 10^7$ N/m in the vertical direction. The measured static stiffness of support 2 is $1.08\ 10^6$ N/m in the horizontal direction and $1.39\ 10^7$ N/m in the vertical direction. The mass of the bearing and bearing housing (12.7 kg) was used as the support mass for single mass supports. Damping equal to two percent of the critical damping was assumed. The support stiffness to bearing stiffness ratio for support 1 is within 0.13 and 0.17 in the horizontal direction and within 1.5 and 3.7 for the vertical direction. For support 2 the support stiffness to bearing stiffness ratio is within 0.13 and 0.17 in the horizontal direction and within 1.3 and 3.1 in the vertical direction. These values were used for the coefficients of the single mass supports because they are the most easily obtainable values and are the most likely used for an engineer in the field.

## Unbalance Response

The unbalance response of the rotor was measured during runup from 1000 rpm to 6000 rpm with a speed resolution of 10 rpm. The average acceleration of the rotor was 3.5 rad/s$^2$. Known unbalance distributions were applied to the rotor and the response was measured twice. The second time, the unbalance weights were located at 180 deg from the original location. The unbalance response of the rotor to the applied unbalance distribution is obtained by subtracting the response of the second run from the response of the first run and dividing the result by 2. This operation eliminates the effects of mechanical and electrical runout, shaft bow, and residual unbalance.

Figure 8 shows the unbalance response near the middle disk for an unbalance weight of 8.92 gr-mm located at 225 deg (lead) on the middle disk. The unbalance distribution is designed to excite the first critical speed. This figure shows three main responses. The response at 1900 rpm corresponds to a structural resonance. The response at 2550 rpm corresponds to the first critical speed and the response around 4000 rpm corresponds to a coupling with the second critical speed of the rotor. This figure also shows predicted responses using different models for the support structure. The predicted response using single mass supports does not agree with the experimental data, showing two responses, one at the structural resonance of 1900 rpm and the other at 3500 rpm.

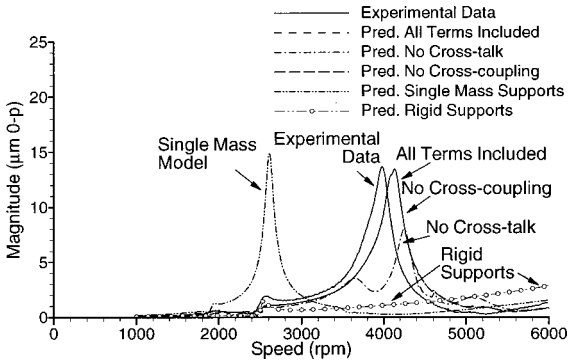



**Fig. 9 Measured and predicted unbalance response near the right disk in the horizontal direction. Unbalance distribution 2.**

The analysis using rigid support model predicts the first critical speed at 2530 rpm (0.7 percent below the measured critical speed) but the predicted amplitude is 134 percent larger than the measured amplitude at the first critical speed.

The predictions using the transfer function representation of the support structure show better agreement with the experimental data. Using all the terms of the transfer function matrix predicts the first critical speed at 2520 rpm (1.2 percent below the measured critical speed). The predicted magnitude at the critical speed is six percent larger than the measured response. Neglecting the crosstalk between supports predicts a double critical speed, 2390 rpm and 2520 rpm. The predicted magnitude at the critical speed is 33.1 percent smaller than the magnitude of the measured response.

The analysis ignoring the cross coupling between the horizontal and vertical directions predicts the first critical speed at 2550 rpm. The predicted magnitude at the critical speed is ten percent smaller than the magnitude of the measured response.

Figure 9 shows the unbalance response near the right disk for an unbalance distribution of 22.86 gr-mm at 30 deg (lead) on the left disk and 22.86 gr-mm at 210 deg (lead) on the right disk. This unbalance distribution is designed to excite the second critical speed. The measured second critical speed of the rotor is 3980 rpm.

The predicted unbalance response using single mass supports does not agree with the measured data. It shows a peak at 2600 rpm and does not predict the second critical speed within the range of operation (the second critical speed is predicted at 9860 rpm).

The analysis using the rigid support models predicts the small coupling with the first mode of vibration. The second critical speed is predicted at 10,350 rpm (outside the range of operation and 160 percent above the measured second critical speed) with amplitude of 56.1 $\mu$m (310 percent above the magnitude of the measured response at the second critical speed).

The analysis using all the terms of the transfer function matrix predicts the second critical speed at 4130 rpm (3.8 percent above the measured second critical). The predicted magnitude at the second critical speed is 1.5 percent smaller than the magnitude of the measured response. The predictions neglecting the cross-coupling effects between the horizontal and vertical directions are identical to the prediction including all the terms in the transfer function matrix.

The analysis assuming independent supports (cross talk between supports is not considered) predicts the second critical speed at 4250 rpm (6.8 percent above the measured second critical speed). The predicted magnitude at the second critical speed is 41.4 percent smaller than the magnitude of the measured response at the second critical speed.

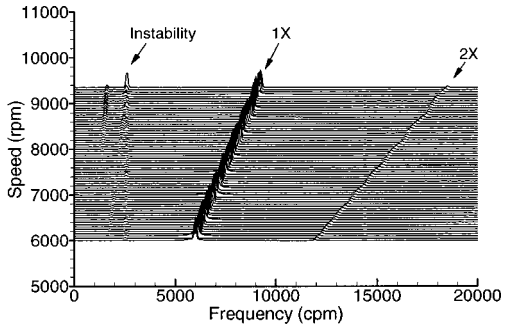



**Fig. 10 Spectral map of the vibration displacement at the middle disk**

## Stability

To determine the stability threshold, the rotor was accelerated until the rotor became unstable. Figure 10 shows the spectral map of the vibration displacement at the center disk. The spectral map shows the frequency contents of the vibration signal for different rotor speeds. This spectral map permits the identification of vibration frequency components at different rotor speeds. Figure 10 was created during rotor runup from 6000 rpm until the stability threshold of 9350 rpm with a frequency resolution of 60 cpm (1 Hz) and a speed resolution of 50 rpm.

The stability threshold was defined as the speed were sub-synchronous vibrations were self-sustained and grew with time. For a lightly damped system, this definition of the stability threshold is important because sub-synchronous vibrations may be present but the overall vibration does not grow in time. The spectral map shows the instability (at 2640 cpm) and the 1X and 2X vibrations components (synchronous and two times the running speed). The vibration component at the running speed includes the effects of mechanical and electrical runout and residual unbalance.

Figure 11 shows the stability map calculated for the rotor system using different support models. The logarithmic decrement is plotted against the rotor speed because the cross coupling in the bearing is the only instability mechanism present in the system. The measured stability threshold is included in the figure, including the experimental uncertainty.

The analysis using rigid support models predicts the instability threshold at 5210 rpm, 44.28 percent below the measured instability threshold. Using single mass supports predicted the stability threshold at 13000 rpm, 39 percent above the measured stability threshold. The stability analysis including all the terms in the transfer function matrix predicts the instability threshold at 9320 rpm, 0.32 percent below the measured threshold and within the experimental uncertainty. Analysis neglecting the effects of cross coupling between the vertical and horizontal directions predicts the stability threshold at 9230 rpm, 1.2 percent below the mea-

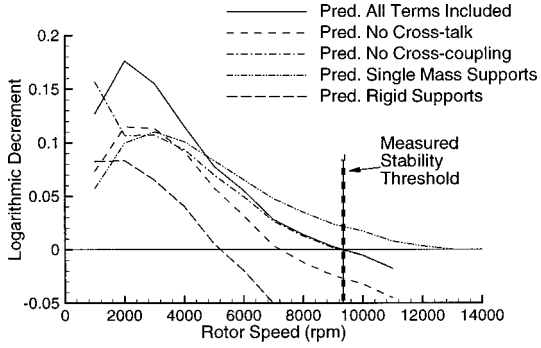



**Fig. 11 Stability maps comparing different support models**

sured stability threshold. The analysis assuming independent supports (does not include the cross talk between supports) predicts the stability threshold at 7224 rpm, 22.7 percent below the measured stability threshold.

## Summary and Conclusions

This work presents the unbalance response and stability of a flexible rotor supported by identical fluid film bearings on anisotropic flexible supports. The dynamic behavior of the support structure was measured and introduced into the rotordynamic analyses in the form of polynomial transfer functions.

Specific conclusions of this work are as follows:

1. The support structure may have a very important influence in the behavior of rotating machinery. Analysis assuming rigid supports predicts the first critical speed within 0.7 percent of the measured critical speed but with the amplitude 134 percent larger than the measured amplitude. The second critical speed is predicted outside the range of operation at 10350 rpm, 160 percent above the measured critical speed and with amplitude 310 percent larger than the measured amplitude. The stability analysis predicted the instability threshold at 5210 rpm, 44.28 percent below the measured instability threshold.
2. The dynamic effects of the support structure can be introduced into unbalance response and stability analyses with polynomial transfer functions.
3. Rotordynamic analyses using polynomial transfer functions agree with the experimental data. The predicted first and second critical speeds agree with the measured critical speed within 1.2 percent and 3.8 percent, respectively. The predicted magnitude of the response at the first and second critical speeds agree with the magnitude of the measured response within 6 percent and 1.5 percent, respectively. The predicted stability threshold agrees with the measured threshold within 0.32 percent (within the experimental uncertainty of instability threshold).
4. Cross-coupling between the horizontal and vertical directions does not have a large effect in the analysis. The predicted first and second critical speeds are within 1.2 percent and 3.8 percent of the measured critical speeds while the predicted magnitudes at these critical speeds are within 10 percent and 1.5 percent of the magnitude of the measured response at the critical speeds. The predicted instability threshold was within 1.2 percent of the measured threshold.
5. The crosstalk between supports has an important effect in the rotor-bearing systems used in this work. The analyses assuming independent supports predicted a split first critical speed with a magnitude 33.1 percent smaller than the magnitude of the measured response. The predicted second critical speed was 6.8 percent above the measured critical speed with a predicted magnitude 41.4 percent smaller than the magnitude of the measured response at the second critical speed. The predicted instability threshold was 22.7 percent below the measured instability threshold. The effect of the cross talk between support depends on the configuration of the support structure and may not be important in all cases.
6. Single mass support models created using the static stiffness of the support and the mass of the bearing and bearing housing do not capture the dynamic behavior of the support structure for unbalance response and stability analyses.

## Acknowledgments

This research was sponsored by the Rotating Machinery and Controls Laboratories (ROMAC) at the University of Virginia.

## References

[1] American Petroleum Institute, 1995, ''Centrifugal Compressors for Petroleum, Chemical and Gas Service Industries,'' API Standard 617, Sixth Ed.

[2] Barrett, L. E., Nicholas, J. C., and Dhar, D., 1986, ''The Dynamic Analysis of Rotor-Bearing Systems Using Experimental Bearing Support Compliance Data,'' *Proceedings of the Fourth International Modal Analysis Conference*, Union College, Schenectady, NY, Society for Experimental Mechanics, Bethel, CT, pp. 1531–1535.

[3] Nicholas, J. C., and Barrett, L. E., 1986, ''The Effect of Bearing Support Flexibility on Critical Speed Prediction,'' ASLE Trans., **29**, No. 3, pp. 329–338.

[4] Nicholas, J. C., Whalen, J. K., and Franklin, S. D., 1986, ''Improving Critical Speed Calculations Using Flexible Bearing Support FRF Compliance Data,'' *Proceedings of the 15th Turbomachinery Symposium*, Texas A&M University, College Station, TX.

[5] Redmond, I., 1995, ''Practical Rotordynamics Modeling Using Combined Measured and Theoretical Data,'' *Proceedings of the 13th International Modal Analysis Conference*, Nashville, TN, Society for Experimental Mechanics, Bethel, CT.

[6] Redmond, I., 1996, ''Rotordynamic Modelling Utilizing Dynamic Support Data Obtained From Field Impact Tests,'' *Proceedings of Sixth International Conference on Vibrations in Rotating Machinery*, Oxford, Sept., Paper C500/055/96.

[7] Rouch, K. E., McMains, T. H., and Stephenson, R. W., 1989, ''Modeling of Rotor-Foundation Systems Using Frequency-Response Functions in a Finite Element Approach,'' *1989 ASME Design Technical Conference 12th Biennial Conference on Mechanical Vibration and Noise*, Montreal, Canada ASME, New York, pp. 157–166.

[8] Stephenson, R. W., and Rouch, K. E., 1992, ''Generating Matrices of the Foundation Structure of a Rotor System From Test Data,'' J. Sound Vib., **154**, No. 3, pp. 467–484.

[9] Feng, M. S., and Hahn, E. J., 1998, ''On the Identification of a Flexibly Supported Rigid Foundation With Unknown Location of the Principal Axes of Inertia,'' *Proceedings of ISROMAC-7, the 7th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery*, Honolulu, HI, Feb., Pacific Center of Thermal Fluid Engineering, Bird Rock Publishing, pp. 705–714.

[10] Lees, A. W., and Friswell, M. I., 1997, ''The Evaluation of Rotor Imbalance in Flexibly Mounted Machines,'' J. Sound Vib., **208**, No. 5, pp. 671–683.

[11] Lees, A. W., Friswell, M. I., Smart, M. G., and Prells, U., 1998, ''The Identification of Foundation Vibration Parameters From Running Machine Data,'' *Proceedings of ISROMAC-7, the 7th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery*, Honolulu, HI, Feb., Society for Experimental Mechanics, Bethel, CT, pp. 715–724.

[12] Edwards, S., Lees, A. W., and Friswell, M. I., 1999, ''The Identification of Rotor Unbalance From Measured Foundation Response Data,'' *Proceedings of the 17th International Modal Analysis Conference*, Feb. 8–11, Kissimmee, FL, Pacific Center of Thermal Fluid Engineering, Bird Rock Publishing, pp. 1610–1615.

[13] Vázquez, J. A., and Barrett, L. E., 1998, ''Representing Flexible Supports by Polynomial Transfer Functions,'' ASME Paper 98-GT-27.

[14] Vázquez, J. A., and Barrett, L. E., 1999, ''Transfer Function Representation of Flexible Supports and Casings of Rotating Machinery,'' *Proceedings of the 17th International Modal Analysis Conference*, Feb., 8–11, Kissimmee, FL, Society of Experimental Mechanics, Bethel, CT.

[15] Vázquez, J. A., 1999, ''Using Transfer Functions to Model Flexible Supports and Casings of Rotating Machinery,'' Ph.D. dissertation, University of Virginia, Charlottesville, VA, Jan.

[16] Vázquez, J. A., Barrett, L. E., and Flack, R. D., 1999, ''A Flexible Rotor on Flexible Bearing Supports. Part I: Stability,'' *Proceedings of the 1999 Vibration Conference*, Sept. 12–16, Las Vegas, NV, Paper DETC99/VIB-8285.

[17] Vázquez, J. A., Barrett, L. E., and Flack, R. D., 1999, ''A Flexible Rotor on Flexible Bearing Supports. Part II: Unbalance Response,'' *Proceedings of the 1999 Vibration Conference*, Sep. 12–16, Las Vegas, NV, Paper DETC99/VIB-8286.

[18] Branagan, L. A., 1988, ''Thermal Analysis of Fixed and Tilting Pad Journal Bearings Including Cross-Film Viscosity Variations and Deformations,'' Ph.D. dissertation, University of Virginia, Charlottesville, VA.

[19] Sanathanan, C. K., and Koerner, J., 1963, ''Transfer Function Synthesis as a Ratio of Two Complex Polynomials,'' IEEE Trans. Autom. Control, Jan., Society for Experimental Mechanics, Bethel, Ct, pp. 56–58.

[20] Richardson, M. H., and Formenti, D. L., 1982, ''Parameter Estimation From Frequency Response Measurements Using Fraction Polynomials,'' *Proceedings of the 1st International Modal Analysis Conference*, Orlando, FL, Society for Experimental Mechanics, Bethel, CT, pp. 167–181.

[21] Richardson, M. H., and Formenti, D. L., 1985, ''Global Curve Fitting of Frequency Response Measurement using Rational Fraction Polynomial Method,'' *Proceedings of the 3rd International Modal Analysis Conference*, Orlando, FL, Nov., Society for Experimental Mechanics, Bethel, CT, pp. 390–397.

[22] Friswell, M. I., and Penny, J. E. T., 1993, ''The Choice of Orthogonal Polynomials in the Rational Fraction Polynomial Method,'' Int. J. Anal. Exp. Modal Anal., **8**, No. 3, pp. 257–262.

[23] Maia N. M. M., and Silva, J. M. M., 1997, *Theoretical and Experimental Modal Analysis*, Research Studies Press LTD.

# Performance of a Foil-Magnetic Hybrid Bearing

**E. E. Swanson**

**H. Heshmat**
Fellow ASME

**J. Walton, II**

Mohawk Innovative Technology, Inc.,
1037 Watervliet-Shaker Road,
Albany, NY 12205-2033

*To meet the advanced bearing needs of modern turbomachinery, a hybrid foil-magnetic hybrid bearing system was designed, fabricated, and tested in a test rig designed to simulate the rotor dynamics of a small gas turbine engine (31 kN to 53 kN thrust class). This oil-free bearing system combines the excellent low and zero-speed capabilities of the magnetic bearing with the high-load capacity and high-speed performance of the compliant foil bearing. An experimental program is described which documents the capabilities of the bearing system for sharing load during operation at up to 30,000 rpm and the foil bearing component's ability to function as a backup in case of magnetic bearing failure. At an operating speed of 22,000 rpm, loads exceeding 5300 N were carried by the system. This load sharing could be manipulated by an especially designed electronic control algorithm. In all tests, rotor excursions were small and stable. During deliberately staged magnetic bearing malfunctions, the foil bearing proved capable of supporting the rotor during continued operation at full load and speed, as well as allowing a safe rotor coastdown. The hybrid system tripled the load capacity of the magnetic bearing alone and can offer a significant reduction in total bearing weight compared to a comparable magnetic bearing.* [DOI: 10.1115/1.1417485]

## Introduction

Advancement in turbine engine performance has been, and continues to be, a critical element in maintaining U.S. technological superiority. In the near future advanced turbomachinery is expected to require bearings capable of continuous operation at between three to four million DN (where D is the diameter in mm, N is the shaft speed in rpm) and temperatures of up to 500 °C. Even current state of the art liquid lubricated, ceramic rolling element bearings are severely challenged by these conditions. Consequently, what is needed is not a mere improvement of present day practices or gradual progress in bearing design, but a new approach to bearing technology. One approach is to develop a new category of oil-free bearings.

One such oil free system is represented by active magnetic bearings (AMB). Magnetic bearings offer a number of benefits, including:

- operation at zero and low speeds,
- active control of rotor system dynamics,
- potential for high-temperature capability, and
- low losses.

Offsetting these are a number of concerns:

- finite load capacity (no overload beyond maximum design load),
- generally poor response to transient conditions and shocks (control system bandwidth is frequently limited),
- problems with high-frequency structural dynamics (noncollocation issues, actuator location constraints and finite bandwidth), and
- large number of potential failure modes due to number of components in control loop.

Much of the current research aimed at increasing AMB reliability, is directed towards fault-tolerant redundant systems. These systems frequently are significantly more complex. However, even with redundant AMB systems, many end-users will require some form of backup bearing. Furthermore, even if a completely fail-safe magnetic bearing system were developed, there must still be a bearing-like mechanical element to support the rotor when the AMB system is not active.

Historically, rolling element bearings (REB) with a fixed clearance between the inner race and the rotor surface have been used as backup bearings in most AMB applications. The success of this approach has been mixed. In small and/or low-speed machines, generally acceptable performance is obtained. In larger and higher speed machines, there have been many problems. Much of the difficulty with AMB backup bearings is associated with the initial transient when the journal drops onto the heretofore nonrotating, unloaded REB ([1]). The sudden REB acceleration from zero to full speed is likely to lead to ball skidding and cage damage, if not outright failure, as well as galling and destructive wear of the journal and inner race due to the large inertia of the components during the initial spinup of the bearing. The impact loads during a transient can also be orders of magnitude higher than the steady load. In addition, rotor rebounds in the REB clearance space can excite rotor system natural frequencies, and result in a violent backward whirl ([2]). Even if the transient problems can be overcome, REB based backup bearings have limited life under these conditions due to high speeds and generally poor lubrication conditions. Thus, as a backup bearing expected to substitute for a failed magnetic bearing, a REB is often an inadequate solution, especially in larger machinery. Even alternatives, such as the zero clearance auxiliary bearing ([3,4]), or cone profiles on the shaft and inner race, which solve many of the dynamic problems associated with the fixed clearance, have operating life limitations. Likewise, solid bushing-type backup bearings, which are sometimes used in place of a REB, are generally not applicable to larger machinery due to the high loads and surface speeds involved.

Another oil-free bearing which has been applied successfully in advanced high-speed machinery, is the air lubricated, compliant surface foil bearing (see, for example, [5–9]). A recent summary of progress in the state of the art for foil bearings is presented in Heshmat and Heshmat [10]. These bearings have been applied to a wide range of high-speed machinery, with operating environments ranging from cryogenic to high temperature and speeds from 30,000 RPM to well in excess of 100,000 RPM.

As shown in Fig. 1, the essential feature of such a bearing is its

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-411. Manuscript received by IGTI November 1999; final revision received by ASME Headquarters February 2000. Associate Editor: D. Wisler.

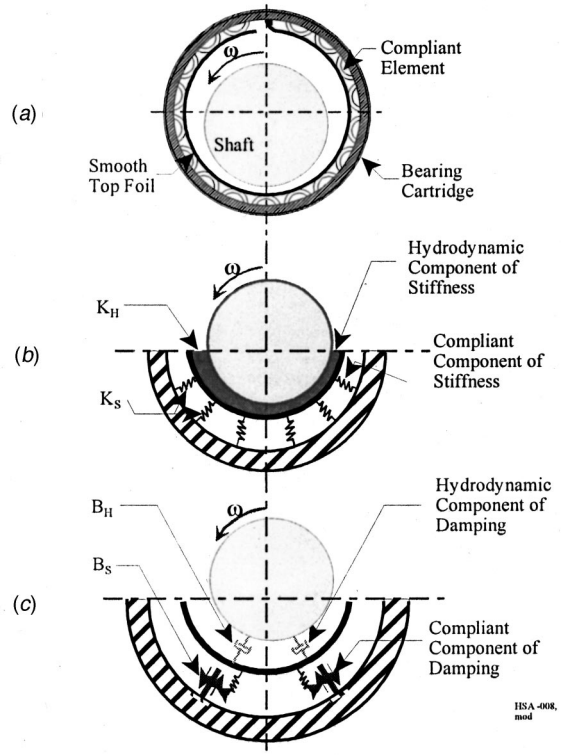



**Fig. 1 Components of foil bearing**

twofold mechanism of imparting stiffness and damping to the system. One is via the geometry and materials of the complaint support structure. The other is due to the hydrodynamic film between journal and top foil. The compliant support can consist of one or more corrugated bump foils, which offer great flexibility in designing dynamic characteristics for a specific system. Additional design flexibility is afforded through construction as a single top foil, or a multipad bearing with multiple top foil segments.

The complaint construction allows the foils to be forced radially outwards by the hydrodynamic pressure as speed increases, thus forming a converging wedge. This self-generated converging wedge becomes more pronounced with increasing speed and load thereby increasing load capacity. The complaint foil surface readily accommodates itself to rotor centrifugal and thermal growth, as well as thermal and mechanical deformations of the bearing housing. These bearings have also demonstrated good performance under shock loading [11].

From a historical perspective there has been a considerable advance in CSFB technology over the last decade. The capabilities of advanced designs now meet the requirements for advanced applications, with unit loadings in excess of 689 kPa ([7]), and adequate damping to allow successful operation of a foil bearing supported rotor at shaft speeds above the first system bending critical speed ([12,13]).

While CSFB's have many advantages, they are a hydrodynamic bearing, hence the journal contacts the bearing surface at rest, and at low speeds during startup and shutdown. An appropriate wear-resistant coating is usually required to minimize wear. Thus, in general, whenever there is a high load on startup, or persistent start-stop cycling, or prolonged low-speed operation, the foil bearing may present operational problems. Additionally, foil bearing do not offer the option of active shaft control as is available with a magnetic bearing.

## Foil-Magnetic Hybrid Bearing

Since the foil bearing has a number of advantages at high speeds and the magnetic bearing at zero and low speeds, it is natural to combine the two technologies. A hybrid bearing combining the two bearings takes advantage of the strengths of each while compensating for each other's weaknesses. Such a hybrid bearing exhibits the characteristics of high load capacity of the CSFB along with the high static stiffness and control versatility of the AMB ([6]). Such a hybrid system represents a significant technological advancement in terms of range of operation and reliability. This work documents the initial tests of a commercially useful hybrid bearing system which combines a 121 mm diameter active magnetic bearing with a 100-mm diameter compliant surface foil bearing. This hybrid bearing has demonstrated load support capability of 5300 N at 22,000 rpm, and is believed to have an ultimate load capacity approaching 7000 N. This hybrid bearing can be configured as a nested design, by either inserting the foil bearing components in the radial gap between the AMB poles and the rotor or, in the case of a homopolar AMB, using the axial space between the sets of poles. The nested configuration is only slightly heavier than the magnetic bearing alone, and of comparable size. Alternately, the foil-magnetic bearing could be configured in a side-by-side arrangement, allowing a reduction in the length and weight of the AMB, but maintaining the same operating load capacity, with a small increase in total shaft length required for the bearing. A number of tradeoffs can be envisioned.

## Experimental Setup

The test rig constructed for the development of the foil bearing is shown in Fig. 2. The test rig rotor was designed to simulate the rotor dynamics of the main rotor of a small gas turbine engine. It is described in detail in Swanson, Walton, and Heshmat [14]. A brief description is as follows: The drive end of the test rig is supported by a 20-mm bore, grease lubricated, deep-groove ball bearing installed in a compliant mount. The test rig is driven from the ball bearing end with a 37 kW induction motor through a high-speed flexible coupling. Speed is controlled with a variable frequency drive which allows bi-directional operation from zero to 36,000 rpm with provisions for dynamic braking. The shaft is coupled to the motor through a disk pack-type coupling with low angular stiffness to minimize its influence on the test rig. The entire rig is mounted on a heavy tee-slot base which allows considerable flexibility for mounting test components. The rotor is 985 mm long, with a mass of approximately 63 kg. A 100-mm diameter sleeve is mounted at one end of the shaft as a running surface for the bearing. Due to heavy shrink fits and the relatively thin sleeve wall, the running surface of the sleeve departed significantly from a true cylinder. Figure 3 shows the measured profile of the installed sleeve with the location of the foil bearing indicated. Two eddy-current displacement sensors are located adjacent to the foil bearing to monitor shaft vertical and horizontal displacements. Two additional eddy-current displacement sensors are located inboard of the magnetic bearing for AMB control and shaft position monitoring. Based on calibrations performed, the displacement readings are accurate within $\pm 5$ $\mu$m over the range of shaft excursion.

In order to apply a load on the test bearing, a hydraulic loading system was attached outboard the foil bearing, as shown in Fig. 2. The loader system consists of a duplex pair, angular contact bearing attached to the shaft, an oil-jet lubrication system, oil seals, load cells, and hydraulic cylinders to apply the load. The system is operated with a hand pump, and is capable of applying unidirectional loads of more than 6.7 kN to the shaft both vertically and horizontally. Each axis of the loader was provided with a load cell to allow the total applied load to be measured. Each bearing was also mounted on load cells to allow the individual bearing loads to be measured separately. Problems with dynamic loads and load cell zero drift, however, made both of these load measurements inaccurate with the rig operating. Thus, an experimentally deter-

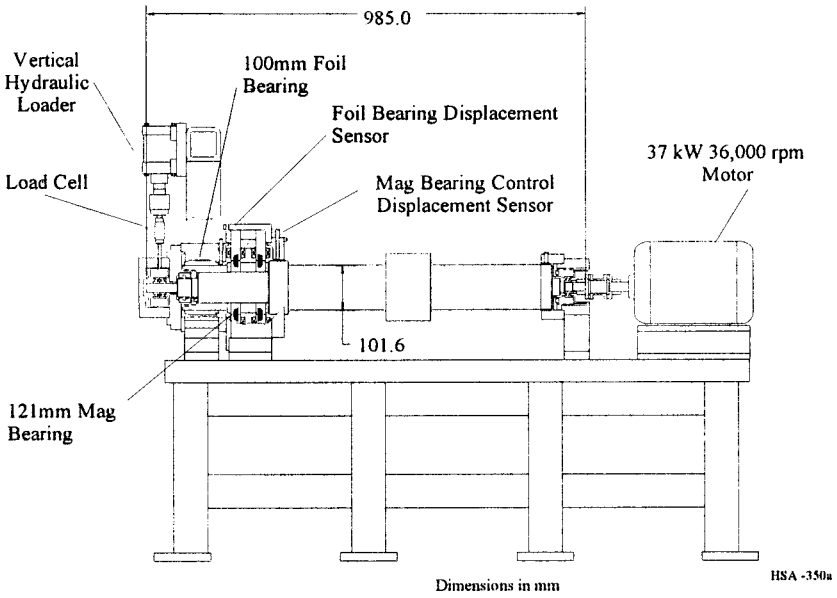



**Fig. 2 Test rig outline**

mined relationship for magnetic bearing load based on control current and shaft position was used to determine the magnetic bearing load component. This information, along with the measured total load, could then be used to compute the foil bearing load. The load measurements are accurate to at least ±10 percent, which was deemed sufficient for the purposes of this study.

A 24-channel 16-bit high-speed digital data acquisition system which samples all channels simultaneously was used in parallel with a 14-channel analog data recorder to acquire and store test results. A user interface allowed for real time as well as post-test data analysis. The analog data recorder was used a backup to the digital data acquisition system as well as to monitor rig performance between tests.

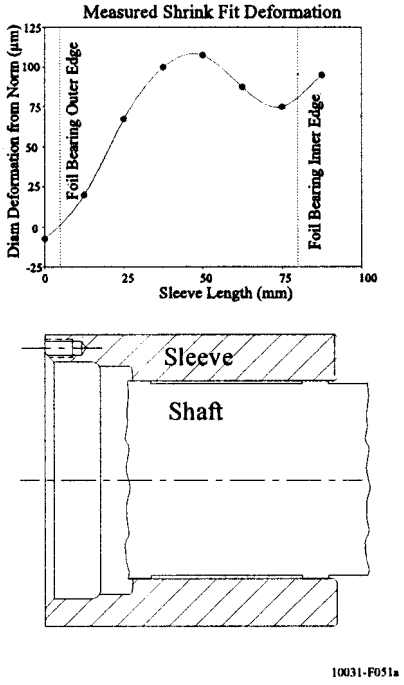



**Fig. 3 Measured journal sleeve deformation**

## Test Bearing

Two possible modes of combining the magnetic and foil bearing were considered: nested and side by side. To simplify this initial study, a side-by-side arrangement was selected as shown in Fig. 2, although this is not optimal from the standpoint of size and weight. Due to the presence of high eddy currents at high speeds, an eight-pole homopolar magnetic design with was chosen. The parameters for the magnetic bearing are as shown in Table 1. The foil bearing parameters are as shown in Table 2. The foil bearing was also tested separately, as is discussed in ([15]). In the side-by-side arrangement of the two bearings used for these tests, the foil bearing was located at the outboard end of the test rig as shown in Fig. 2. Note should also be taken of the different clearances in the two bearings; in a concentric position the CSFB's

**Table 1 AMB parameters**

| | |
|---|---|
| Load capacity | 980 N in normal operation<br>1700 N at saturation |
| Shaft diameter | 121 mm |
| Axial length | 2×21.8 mm |
| Control current | 8 A nominal<br>16 A for maximum load |
| Bias current | 8 A nominal<br>16 A for maximum load |
| Bias winding | 200 turns |
| Bias design flux | 6 k Gauss nominal<br>10 k Gauss for maximum load |
| Laminates | 177 $\mu$m silicon-iron |
| Pole area | 1600 mm$^2$ |
| Centered air Gap | 635 $\mu$m |

**Table 2 Foil bearing parameters**

| | |
|---|---|
| Load capacity | 5300 N @ 22,000 rpm (predicted) |
| Shaft diameter | 100 mm |
| Axial length | 75 mm |
| Clearance ratio | 0.002 |
| Foil extent | 355 deg. |

radial clearance is 101 $\mu$m when the shaft is not rotating, whereas the concentric radial clearance in the AMB is 635 $\mu$m. Thus, even when heavily loaded, the foil bearing will limit shaft motion to a fraction of the magnetic bearing clearance.

## Control Algorithm

An important aspect in a CSFB-AMB hybrid operation is the development of a control system. Using a conventional proportional-integral-derivative (PID) control algorithm, the controller would attempt to force the journal to a centered position at all times, regardless of load. However, in a loaded foil bearing the journal runs at some eccentricity and attitude angle when supporting a load. One approach to coordinating the characteristics of the two bearings is a supervisory control system which is described in Heshmat, Chen, and Walton [6]. Briefly, this algorithm functions as follows:

**Step 1.** At a given speed the AMB takes all the steady-state load so that total load and its direction can be ascertained.

**Step 2.** Using the stored data the eccentricity and attitude angle of the foil bearing are determined for a chosen share of the load.

**Step 3.** The AMB reference location is moved to the foil bearing eccentric position. Once rotor center has been relocated the currents are measured and the load carried by the AMB calculated. Also, magnetic stiffnesses as well as the dynamic stiffness as a function of excitation frequency can be obtained from the PID gains.

**Step 4.** Calculate the stiffness and damping of the foil bearing for the steady-state load. These dynamic coefficients are added to the those of the magnetic bearing.

**Step 5.** Check the adequacy of the combined dynamic properties in terms of rotor stability. If not adequate, adjust the PID controller to improve them.

## Operational Characteristics

An extensive series of tests was conducted to verify the combined performance of the two bearings. These tests focused on several issues: (1) system operation from 0 to 30,000 RPM with shaft load, (2) system operation at varying load magnitudes, (3) system operation for varying load angles, (4) system operation during transition from AMB to CSFB, and (5) system operation during simulated failures of the magnetic bearing.

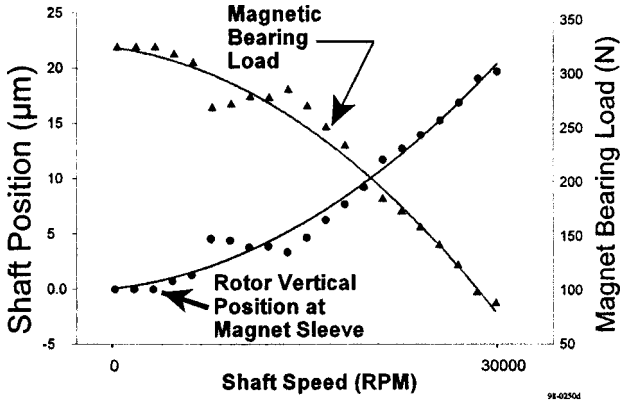



**Fig. 4 Run-up to 30,000 rpm, 355 N load**

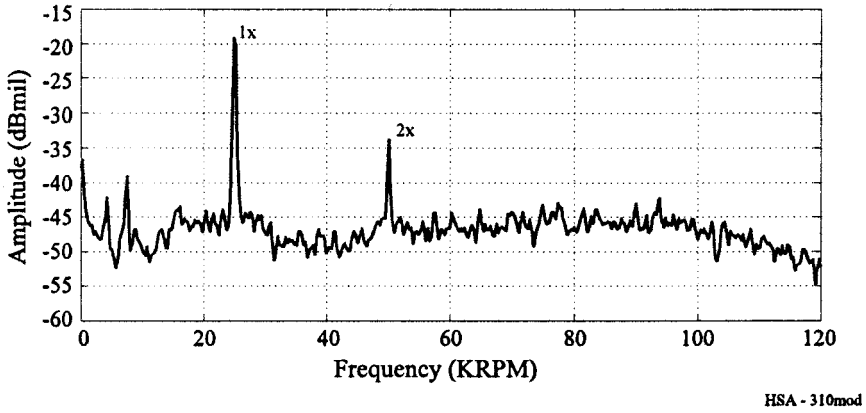



**Fig. 5 Averaged vertical spectrum operating at 25,000 rpm, 355 N load**

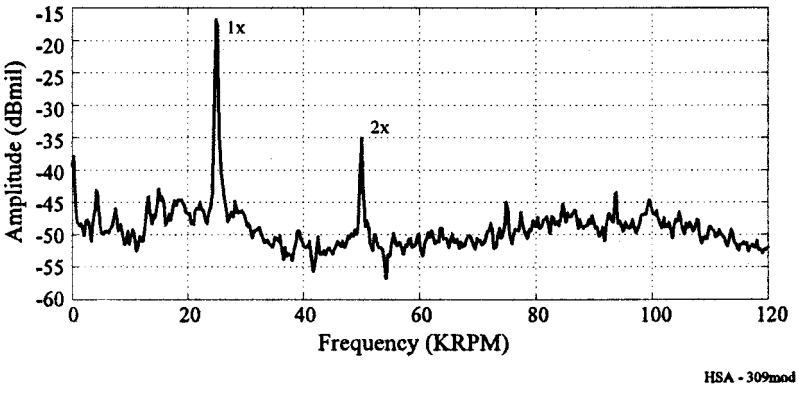



**Fig. 6 Averaged horizontal spectrum operating at 25,000 rpm, 355 N load**

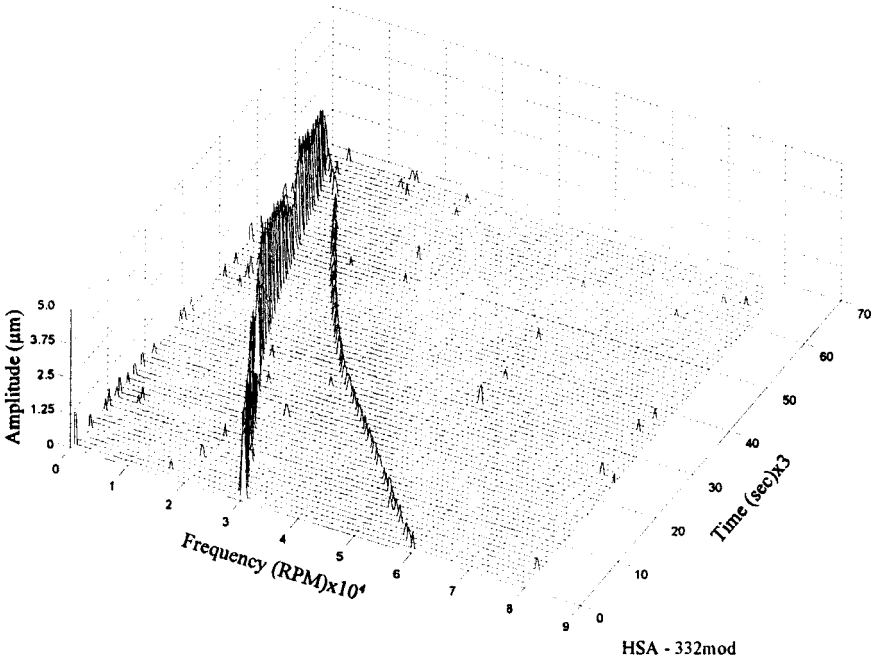



**Fig. 7 Vertical waterfall for coastdown from 30,000 rpm, 355 N load**

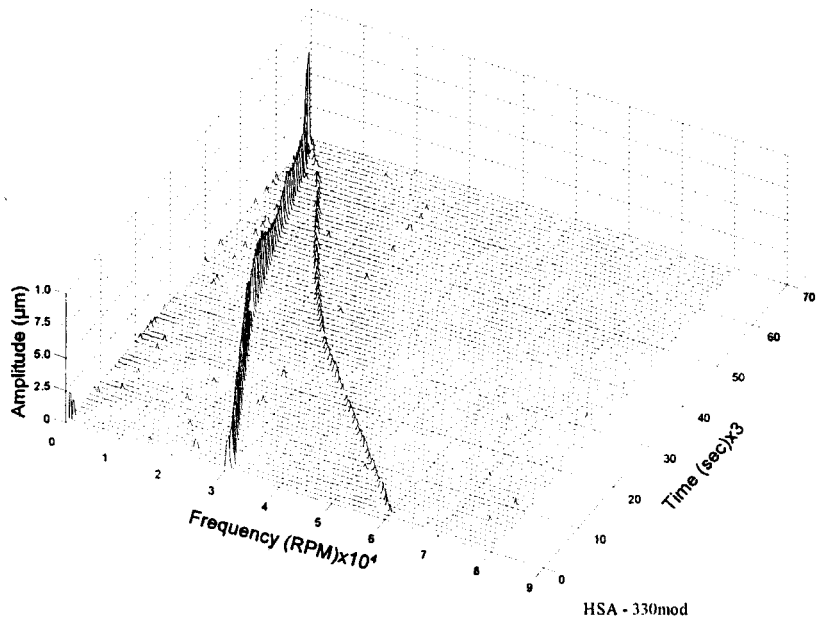



**Fig. 8 Horizontal waterfall for coastdown from 30,000 rpm, 355 N load**

**Shaft Load Operation.** For this series of tests, operation with the shaft dead load of 355 N was examined. Load sharing was specified to be approximately all magnetic bearing at low speed, gradually increasing to 80 percent foil bearing at 30,000 RPM. Fig. 4 presents the measured magnetic bearing load and shaft position versus speed for a test starting from zero and accelerating to 30,000 rpm. As speed increased the portion of load carried by the AMB decreased, the remainder being carried by the CSFB. As shown in Fig. 4, the hybrid bearing operated over the entire speed range, smoothly transferring the load from the magnetic bearing to the foil bearing as desired.

Figures 5 and 6 present typical steady state displacement spectra when the AMB was supporting most of the load at 25,000 RPM. Figures 7 and 8 present the peak-held spectra for a coast down on the hybrid bearing under shaft load of 355 N. The data shown are for the magnetic bearing probes. These figures show that the hybrid bearing is stable over a range of operating speeds, with the primary synchronous response. The similarity between the vertical and horizontal motions suggests that the bearing is nearly isotropic under this loading condition.

**Load Sharing/Higher Loads.** For this series of tests, a range of discrete load sharing/speed points were targeted. Figures 9–11 present some of the points examined. The total loads ranging from 750 N to 5300 N. Each of these test points specified was achieved through a supervisory control system similar to that described above, which balanced the load between the two bearings. The wide range of load combinations shown demonstrates the ability

of the hybrid bearing to control the amount of load shared, as well as stable operation over a wide range of load sharing splits.

**Operation During AMB Failure.** As discussed previously, one of the advantages of a foil-magnetic hybrid bearing is the redundancy in the bearing system. In most applications, this redundancy can eliminate the need for an auxiliary/backup bearing for the magnetic bearing. Two distinct regimes of operation must be considered. The first is failure of the magnetic bearing at operating speed, with a potential requirement for the foil bearing to continue to operate as the primary load support. The second is a failure at lower speed, wherein the foil bearing would need to provide adequate shaft support for a safe machine shutdown. Note that continued operation will, of necessity, also require the ability for eventual regimes, a number of tests were performed. To simplify testing, these tests were conducted under shaft load alone. To increase the severity of the test, the system was operated with the magnetic bearing component supporting most of the shaft load prior to the simulated failure. A total of 13 different failure modes at two different speeds were considered. A typical worst case rotor response is shown in Fig. 12. As shown in this figure, consideration was also given to the possibility that the magnetic bearing fault could be removed, thus requiring re-activation of the magnetic bearing. Again, to increase the severity of these tests, the operating point was shifted from total foil bearing load support to primarily magnetic bearing support. Additionally, the magnetic bearing was not operated in a soft-start mode, resulting in a substantial overshoot at re-activation. Note the substantial shaft overshoot. Indeed, if this response is compared with the failure transient, it can be seen that the larger dynamic loads appear to occur at bearing start-up in these tests.

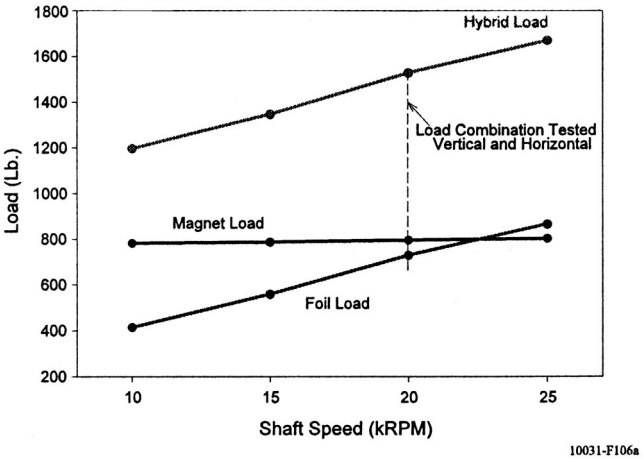



**Fig. 9 Hybrid loads 1**

As discussed in Swanson and Heshmat [15], which considers operation with the foil bearing alone, this bearing had no problem supporting loads of up to 4200 N, and was operated for periods of up to 1 hour to demonstrate thermal stability. Based on these results, the foil bearing in a foil-magnetic hybrid bearing can provide a continuous full load, full-speed operation capability. Additionally, the testing reported in this work demonstrated a coast-down to a full stop under shaft load without bearing damage. Thus, except possibly in the case of extremely high bearing loads, the foil bearing also provides a safe shutdown capability.

## Summary/Conclusions

Experimental tests of a foil-magnetic hybrid bearing has demonstrated the viability of the concept. This bearing offers great

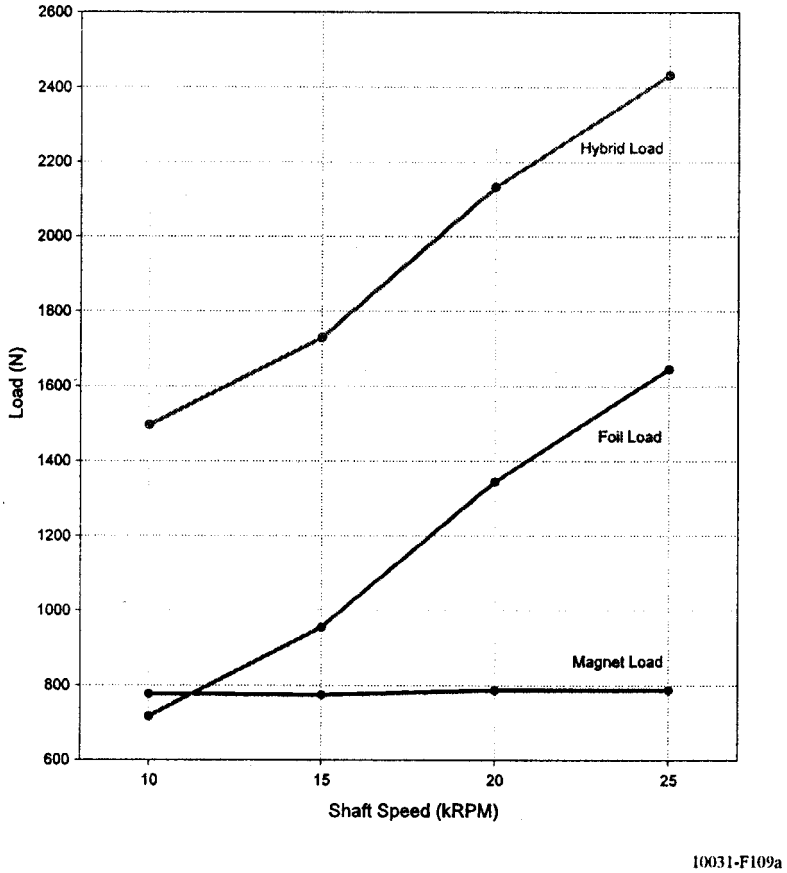



**Fig. 10 Hybrid loads 2**

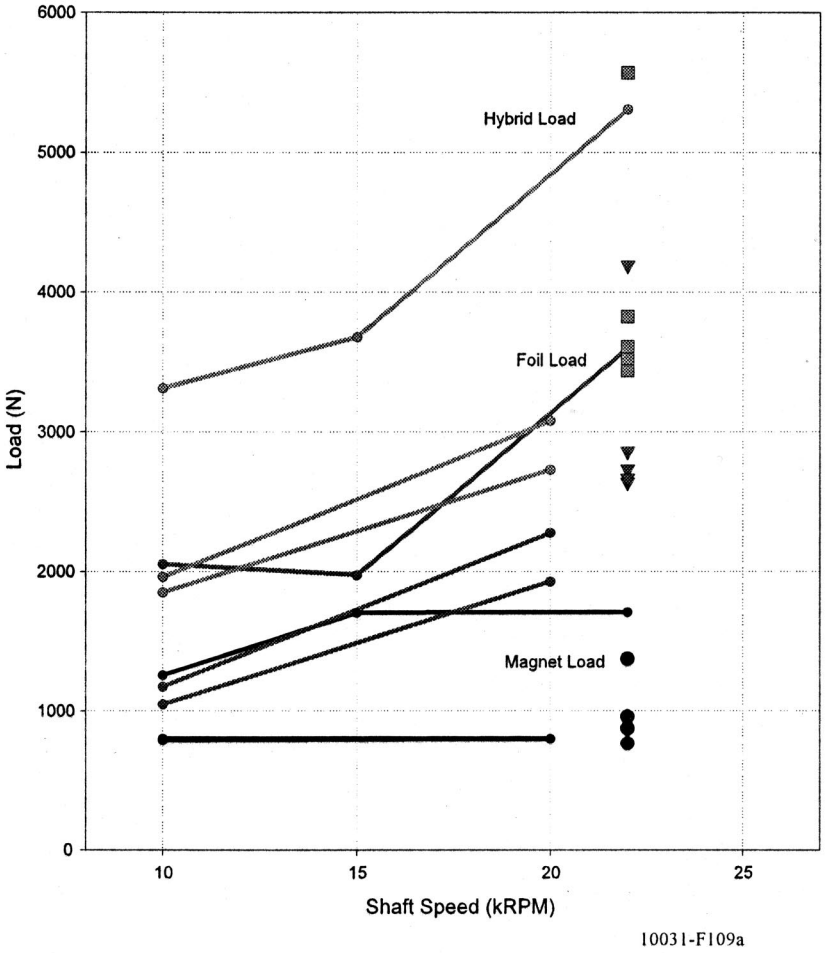



**Fig. 11 Hybrid loads 4**

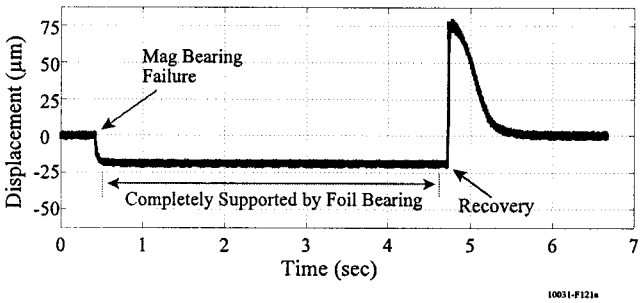



**Fig. 12 Typical worst case simulated failure response 5300 N at 22,000 rpm**

promise for aerospace applications, especially in light of the potential to triple the bearing load capacity relative to a similar size and weight magnetic bearing ([6]). This work has demonstrated several key features:

1 The hybrid arrangement functioned without problems to a maximum speed of 30,000 rpm, with a maximum tested load of 5300 N at 22,000 rpm.

2 The CSFB functioned well as a backup device in absorbing the transient impact of an AMB shutdown. It proved capable of either supporting the shaft during coastdown, or during continued operation.

3 Under the control of a proper monitoring system successful load sharing could be achieved. Within the constraint of limited foil bearing load capacity at low speed, considerable flexibility in load sharing was observed.

## Acknowledgments

The authors would like to thank Mohawk Innovative Technology, Inc, especially Mr. Mike Tomaszewski, Mr. Mike Albertini, and Prof. Oscar Pinkus; as well as Dr. Chris Della Corte of NASA Glenn Research Center and Dr. Mark Valco of the U.S. Army Research Laboratory for their support of this work.

## References

[1] Swanson, E. E., Raju, K. V. S., and Kirk, R. G., 1996, "Test Results and Numerical Simulation of an AMB Rotor Drop," *Proc. 6th Int. Conf. on Vibrations in Rotating Machinery*, I Mech Eng., UK, pp. 119–131.
[2] Ehrich, F. F., 1996, "The Dynamic Stability of Rotor/Stator Rubs in Rotating Machinery," J. Eng. Ind., **90**, pp. 1025–1028.
[3] Chen, H. M., Walton, J. F., and Heshmat, H., 1997, "Zero Clearance Auxiliary Bearings for Magnetic Bearing Systems," ASME Paper No. 97-GT-112.
[4] Chen, H. M., Walton, J. F., and Heshmat, H., 1997, "Test of a Zero Clearance Auxiliary Bearing," *Proceedings of the Mag '97 Industrial Conference and Exhibition on Magnetic Bearings*, pp. 111–119.
[5] Suriano, F. J., Dayton, R. D., and Woessner, F. G., 1983, "Test Experience With Turbine-End Foil Bearing Equipped Gas Turbine Engines," ASME Paper No. 83-GT-73.
[6] Heshmat, H., Chen, H. M., and Walton, J. F., 1998, "On the Performance of Hybrid Foil-Magnetic Bearings," ASME Paper No. 98-GT-376.
[7] Heshmat, H., 1994, "Advancement in the Performance of Aerodynamic Foil Journal Bearings—High Speed and Load Capacity." ASME J. Tribol., **116**, pp. 287–295.
[8] Walton, II, J. F., and Heshmat, H., "Application of Foil Bearings to Turbomachinery Including Vertical Operation," ASME Paper No. 99-GT-391.

[9] Heshmat, H., and Ku, C.-Pr., 1994, ''Structural Damping of Self-Acting Compliant Foil Journal Bearings,'' ASME J. Tribol., **116**, pp. 76–82.
[10] Heshmat, C. A., and Heshmat, H., 1995, ''An Analysis of Gas-Lubricated, Multileaf Foil Bearings with Backing Springs,'' ASME J. Tribol., **117**, pp. 437–443.
[11] Heshmat, H., and Hermel, P., 1992, ''Compliant Foil Bearing Technology and Their Application to High-Speed Turbomachinery,'' 19th Leeds-Lyon Symp., Leeds, UK.
[12] Walton, J. F., and Heshmat, H., 1994, ''Compliant Foil Bearings for use in Cryogenic Turbopumps,'' NASA CP3232, Vol. 1, pp. 372–381.
[13] Heshmat, H., 1999, ''Operation of Foil Bearings Beyond the Bending Critical Mode,'' ASME Paper No. 99-TRIB-48.
[14] Swanson, E. E., Walton, J. F., and Heshmat, H., 1999, ''A 35,000 RPM Test Rig for Magnetic, Hybrid, and Back-Up Bearings,'' ASME Paper No. 99-GT-180.
[15] Swanson, E. E., and Heshmat, H., 2000, ''Capabilities of Large Foil Bearings,'' ASME Paper 2000-GT-387.

# Thermoelastic Stress Analysis: An NDE Tool for Residual Stress Assessment of Metallic Alloys

**A. L. Gyekenyesi**
Resident Research Scientist,
Ohio Aerospace Institute,
NASA Glenn Research Center,
21000 Brookpark Road,
MS 6-1,
Cleveland, OH 44135

*The primary objective of this report involves studying and developing various experimental techniques for accurate measurement of the mean stress effect in thermoelastic stress analysis (TSA, also recognized as SPATE: stress pattern analysis by thermal emission). The analysis of cyclic mean stresses at the coupon level directly relates to the measurement of residual stresses in structures. In a previous study by the authors, it was shown that cyclic mean stresses significantly influenced the TSA results for titanium and nickel-based alloys, although, difficulties were encountered concerning the quantification of the mean stress effect because of large test-to-test variations. This study continues the effort of accurate direct measurements of the mean stress effect by implementing various experimental modifications. In addition, a more in-depth analysis is included which involves analyzing the second harmonic of the temperature response. By obtaining the amplitudes of the first and second harmonics, the stress amplitude and the mean stress at a given point on a structure subjected to a cyclic load can be simultaneously obtained. The rather complex analysis of the temperature response involves obtaining the first and second harmonic amplitudes for 16384 infrared detectors (128×128 focal plane array). Upon establishing a protocol for mean stress measurements in the laboratory using the TSA technique, the next step is to utilize the method to assess residual stress states in complex structures during manufacturing and life.* [DOI: 10.1115/1.1417982]

## Introduction

Thermoelastic stress analysis (TSA) is a full-field noncontacting technique for surface stress mapping of structures. TSA is based on the fact that materials experience a temperature change when compressed or expanded (i.e., experience a change in volume). If the load causing the volumetric change is removed and the material returns to its original temperature and shape, the process is deemed reversible. This reversibility is achieved when a material is loaded elastically at a high enough rate so as to eliminate significant conduction of heat. The thermoelastic temperature change, in steel for instance, as a result of an applied cyclic load of 1.0 MPa (145 psi) is on the order of 0.001°C ([1]).

Lord Kelvin [2] first quantified an analytical relationship between the change in temperature and the change in stress. The formulation is as follows:

$$\frac{\delta T}{T_o} = \frac{-\alpha \delta I_1}{\rho C_p} \tag{1}$$

$$= -K \delta I_1 \tag{2}$$

where $\delta T$ is the cyclic change in temperature; $\alpha$ is the coefficient of linear thermal expansion; $T_o$ is the absolute temperature of the specimen; $\delta I_1$ is the change in the sum of the principal stresses; $\rho$ is the material density; $C_p$ is the specific heat at constant pressure; and $K$ is the thermoelastic constant. Note that the formulations of Eqs. (1) and (2) indicate that $\delta T$ is independent of the mean stress. Therefore, $\delta T$ is assumed to remain constant for a given stress range, $\delta I_1$, and absolute temperature, $T_o$, regardless of the applied mean stress.

In recent years, experimental data as well as theoretical formulations have shown a mean stress dependence concerning the thermoelastic constant, $K$. Such stress dependence was illustrated experimentally by Machin et al. [3] and Dunn et al. [4]. The theoretical explanation for this nonlinear response was defined by Wong et al. [5]. Additional experimental evidence of the mean stress effect was provided in Wong et al. [6]. In Wong et al. [5], the reformulated theory shows that the mean stress dependence of the thermoelastic parameter (earlier referred to as a constant) is fully accounted for by the temperature dependence of the elastic moduli of the material.

In this paper, further experimental verification of the mean stress effect is provided. This includes both a direct measure of the TSA signal as a function of the mean stress as well as an in-depth analysis of the first and second harmonics of the TSA signal. The theory, as described in the following section, shows that the thermoelastic response of a structure subjected to a pure sinusoidal mechanical load with a frequency, $\omega$, produces a TSA signal with frequency components at the primary frequency, $\omega$, as well as the second harmonic, $2\omega$. The first harmonic of the thermal response is a function of the cyclic stress amplitude and the mean stress while the second harmonic is a function of the square of the stress amplitude. By obtaining the TSA amplitudes of the first and second harmonics, the stress amplitude and the mean stress at a given point on a structure subjected to a cyclic load can be simultaneously obtained. Furthermore, comparisons are made between the theoretical predictions and the experimental data. Once confidence is achieved concerning the measured TSA response, then steps can be taken for developing empirical relationships between the TSA signal and the cyclic stress state of various materials.

It should be noted that the past thermoelastic data (see above references) were produced utilizing TSA systems with infrared (IR) cameras based on a single detector. This single detector camera incorporates a network of scanning mirrors to provide an image of an area. In the above references, the scanning devices were disengaged and only the information from a single point on the specimens was analyzed. For this study, the TSA system uses an IR camera with a 128×128 focal plane array (FPA) of detectors. As a result, the reported IR detector response is based on the average of an array of detectors representing a larger two-dimensional area on the specimen surface. To obtain the first and

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-424. Manuscript received by IGTI Nov. 1999; final revision received by ASME Headquarters Feb. 2000. Associate Editor: D. Wisler.

second harmonics of a TSA signal composed of an array of IR detectors, hardware modifications of the current TSA system were required and are detailed in the Experimental Methods section.

The alloy of interest in this study is a titanium-based alloy, TIMETAL 21S. This material is utilized by the aviation industry in propulsion components due to its relative lightweight, high strength, and stiffness, as well as its property retention at elevated temperatures.

During manufacturing, certain components, which are subjected to a cyclic fatigue environment, are fabricated so as to contain compressive residual stresses on the surface. These compressive stresses inhibit the nucleation of cracks. As a result of overloads and elevated temperature excursions, the induced residual stresses dissipate while the component is still in service, in turn, lowering its resistance to crack initiation. Once confidence is achieved concerning reliable TSA measurements of the mean stress effect, research can focus on the application of the method to residual stress assessment. Such measurements will assist in the characterization of materials in the laboratory as well as in situ monitoring of the current residual stress state in actual structural components during fabrication and service.

## Theory

The revised thermoelastic equation (derived in Wong et al. [5]) relating the rate of temperature change and the rate of change in the stress state in a homogeneous Hookean material under adiabatic conditions can be written as

$$\rho_o C_\varepsilon \frac{\dot{T}}{T} = -\left[\alpha + \left(\frac{\nu}{E^2}\frac{\partial E}{\partial T} - \frac{1}{E}\frac{\partial \nu}{\partial T}\right) I_1\right]\dot{I}_1 + \left[\frac{(1+\nu)}{E^2}\frac{\partial E}{\partial T} - \frac{1}{E}\frac{\partial \nu}{\partial T}\right]\sum_{i=1}^{3} \sigma_{ii}\dot{\sigma}_{ii} \tag{3}$$

where $T$ is the thermodynamic temperature (Kelvin); $\sigma_{ii}$ are the principal stresses ($N/m^2$); $I_1$ is the sum of the principal stresses, ($N/m^2$); $\rho_o$ is the density ($kg/m^3$); $C_\varepsilon$ is the specific heat under constant strain (N m/°C kg); $\alpha$ is the coefficient of thermal expansion ($°C^{-1}$); $E$ is the Young's modulus ($N/m^2$); and $\nu$ is the Poisson's ratio. The dotted symbols represent derivatives with respect to time. Notice that for the reformulated thermoelastic equation, the elastic moduli are not assumed constant, but are taken as functions of temperature. Hence, the difference when comparing the new formulation to the classical thermoelastic theory of Eqs. (1) and (2). It is also seen that the temperature response is now dependent on both the stress rate and the stress state.

Following the procedure set forth in Wong et al.[6], the uniaxial case is taken as an example. The uniaxial case is defined as having

$$\sigma_{11} = I_1, \quad \sigma_{22} = \sigma_{33} = 0 \tag{4}$$

and

$$\dot{\sigma}_{11} = \dot{I}_1, \quad \dot{\sigma}_{22} = \dot{\sigma}_{33} = 0. \tag{5}$$

Utilizing Eqs. (4) and (5), Eq. (3) simplifies to

$$\rho_o C_\varepsilon \frac{\dot{T}}{T} = -\left(\alpha - \frac{1}{E^2}\frac{\partial E}{\partial T}\sigma_{11}\right)\dot{\sigma}_{11}. \tag{6}$$

In keeping with the approach taken by Wong et al.[6], a sinusoidal stress input of the following form is defined:

$$\sigma = \sigma_m + \sigma_{amp}\sin \omega t \tag{7}$$

where $\sigma_m$ and $\sigma_{amp}$ are the mean and amplitude of the applied stress, respectively, and $\omega$ is the loading frequency. Note that for ease of presentation the subscripts on the stress terms have been omitted with the understanding that $\sigma = \sigma_{11}$. Next, Eq. (7) is substituted into Eq. (6) to obtain the following solution for the uniaxial case:

$$\rho_o C_t \frac{\delta T}{T_o} = -\left(\alpha - \frac{1}{E^2}\frac{\partial E}{\partial T}\sigma_m\right)\sigma_{amp}\sin \omega t - \frac{1}{4E^2}\frac{\partial E}{\partial T}(\sigma_{amp})^2 \cos 2\omega t. \tag{8}$$

The above equation was achieved by assuring that a valid adiabatic solution was maintained for $\delta T$. This was accomplished by equating to zero any terms that presented a steady-state component of $\delta T$. Such components cannot be maintained due to energy transfers to the surroundings (see Ref. [6]). Note that the thermal response of a structure subjected to a pure sinusoidal mechanical excitation produces a thermal signature function that has components at both the primary loading frequency, $\omega$, as well as the second harmonic, $2\omega$. The first harmonic of the thermal response is a function of the cyclic stress amplitude and the mean stress while the second harmonic is a function of the square of the stress amplitude. It is now apparent that by obtaining the first and second harmonics, it is possible to obtain both the cyclic stress amplitude and the mean stress.

Lastly, comparing Eqs. (1) and (2) with Eqs. (6) and (8), the effective $K$ for a homogeneous Hookean material subjected to a uniaxial stress is derived as

$$K = \left(\alpha - \frac{1}{E^2}\frac{\partial E}{\partial T}\sigma_m\right)(\rho C_\varepsilon)^{-1}. \tag{9}$$

Equation (9) shows that the thermoelastic parameter is a linear function of the mean stress, $\sigma_m$. A normalized measure of the thermoelastic parameter's mean stress dependence can be expressed as

$$\frac{1}{K_o}\frac{\partial K}{\partial \sigma_m} = \frac{-1}{\alpha E^2}\frac{\partial E}{\partial T} \tag{10}$$

where

$$K_o = \frac{\alpha}{\rho C_\varepsilon}. \tag{11}$$

Equation (10) simply gives the slope of the thermoelastic constant (in a dimensionless, normalized form) in respect to the mean stress, $\sigma_m$. Note that Eq. (11) is the zero mean stress solution of Eq. (9).

## Experimental Methods

The titanium alloy utilized for this study was TIMETAL 21S (heat-treat stabilized at 621°C (1150°F) for eight hours). It was supplied in the form of cylindrical, reduced gage section coupons machined from 0.79 cm (0.31 in.) thick, flat plates. The coupons had a constant diameter gage length of 1.91 cm (0.75 in.) and a nominal gage diameter of 0.635 cm (0.25 in.). Three specimens were used during the testing.

The liquid nitrogen cooled infrared camera of the TSA system employs a 128×128 focal plane array of InSb detectors (3–5 $\mu$m sensitivity). The system operates by recording a periodic temperature change, as viewed by the IR camera, of a specimen subjected to a cyclic mechanical load (see Fig. 1). A reference signal from the load cell is used by the software to allow it to monitor only the true thermoelastic change in temperature and to disregard any noise and environmental effects that do not correlate with the reference signal's primary frequency. This is accomplished by using a lock-in detection method (see Ref. [7] for an explanation). The lock-in technique filters the thermal response allowing the analysis of only the TSA frequency component that corresponds to the primary frequency of the reference signal. To further improve the TSA signal to noise ratio, the IR signal is monitored, accumulated, and averaged over a period of hundreds or thousands of load cycles. When an image is captured, the computer display depicts a dimensionless, digitized value of the average camera signal range corresponding to the cyclic load range for each of the 16,384

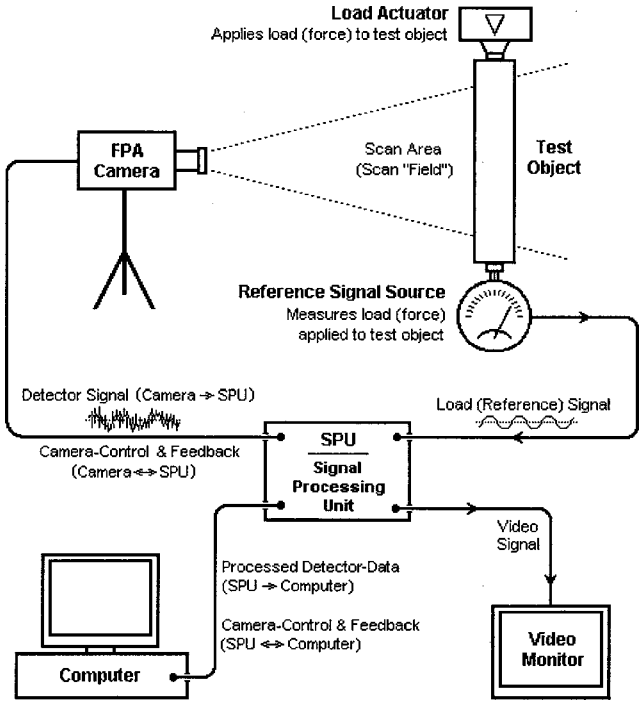



**Fig. 1 Schematic diagram of TSA system (DeltaTherm 1000 User's Manual)**

pixels. For typical structural stress measurements, the dimensionless digital IR values are correlated to a known cyclic stress amplitude (e.g., as measured by a strain gage) and then utilized to map the surface stress amplitudes of the entire structure. Further details concerning the TSA system and the specific settings are given in the Appendix.

Two techniques were utilized to compare the experimental results with the theory. First, direct experimental measurements were conducted for analyzing the mean stress effect. This was accomplished by comparing the camera IR value at various discrete mean stresses while maintaining a fixed cyclic stress amplitude.

The test platform employed for inducing the mechanical excitation was a digitally controlled ($\pm$50 kN dynamic/$\pm$100 kN static) servo-hydraulic test system. The experiments were conducted in load control with a 10-Hertz sinusoidal waveform. The TIMETAL 21S specimens had mean stresses varying from $-276$ MPa ($-40$ ksi) to 276 MPa (40 ksi) while maintaining a 34 MPa (5 ksi) cyclic stress amplitude. A test repetition involved stepping through each of the discrete mean stresses and recording an IR signal. The mean stress stepping order was randomized for each test. The camera distance and angle were held constant throughout the experiments. The camera's line of sight was perpendicular to the specimen face, while the distance from the camera lens to the specimen face was maintained at 12.38 cm (4.88 in.). Lastly, it should be noted that all loads were well within the material's elastic regime, hence, no plastic deformation occurred.

Next, a comparison was conducted between theory and experiment that concerned the analysis of the second harmonic of the thermoelastic temperature response. Since the TSA system filters all components of the TSA signal except the primary frequency of the reference signal, hardware was developed which allowed the electronics to focus in on the second harmonic of the IR signal. The reference signal from the load cell was passed through a linear frequency doubler (developed for this study and placed inline with the reference signal) prior to its use by the TSA system, thereby, allowing the TSA system to focus in on the second harmonic. In the earlier studies referred to in the Introduction, a fast Fourier transform (FFT) spectrum analyzer was employed. This was possible because the IR signal was composed of only a single detector. Here, 16,384 detectors were monitored by the lock-in electronics. To maintain the practicality of the IR array TSA system, the lock-in electronics were utilized during the capture of the second harmonic. Finally, a FFT spectrum analyzer was used throughout the study to monitor the purity of the sinusoidal loading function.

Test preparation included painting the specimens with an ultraflat black paint to improve surface emissivity. In addition, K-type thermocouples were spot welded to the back of each specimen (slightly below the gage section) so as to allow temperature monitoring throughout the tests. Since the IR flux is dependent on the source (i.e., specimen) temperature, an empirical equation was used to correct the camera signals to values corresponding to a specimen temperature of 23°C (73.4°F). This empirical equation was developed during this study. Specimens were cycled at various amplitudes (103 and 207 MPa; 15 and 30 ksi) while maintaining a zero mean stress. During any particular fixed amplitude test, the absolute specimen temperature was varied between 21 and 27°C (70 and 80°F) while the IR response was monitored. Each set of data was normalized by its respective value at 23°C.

## Results

Figure 2 displays a typical thermoelastic image as captured by the software. The dimensionless IR camera signal value used throughout this study was an average of an image box in the gage area of the specimen. For the TIMETAL 21S, the box consisted of 300 pixels representing the individual IR detectors (10 wide$\times$30 long) which were focused on the center portion of the specimen. The boxed area of interest was 0.318 cm$\times$0.953 cm (0.125 in. $\times$0.375 in.). Figure 3 displays the data used to develop the empirical temperature correction curve. The curve is composed of eight tests utilizing three specimens. After conducting a first-order linear regression, the temperature-dependent relationship was defined as

$$S_{23^\circ C}=\frac{S}{(0.337+0.0288T)} \tag{12}$$

where $S_{23^\circ C}$ and $S$ are the corrected and original IR signals, respectively, and $T$ is the specimen temperature in degrees Celsius. All IR data presented in this paper are in the temperature corrected form. The relationship of Eq. (12) is physically a function of the specimen's IR flux, and assumed to be independent of the

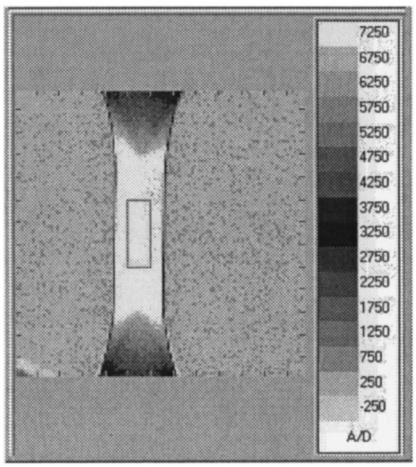



**Fig. 2 Typical IR test image for TIMETAL 21S specimen. The rectangular box within the specimen indicates the area where the average signal was obtained. The scale displays the dimensionless digital values of the IR camera signal.**

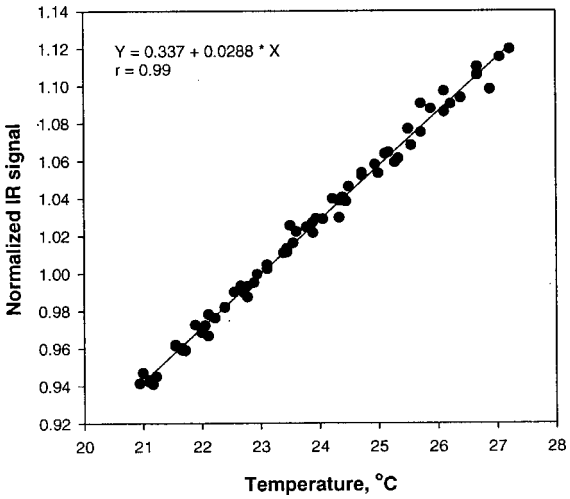



**Fig. 3 Temperature dependence of TSA IR signal. Data consists of eight tests utilizing three specimens. All tests normalized by corresponding IR signals at 23°C. Equation represents first-order linear regression fit of all data.**

subject material. Even so, a cautious approach should be taken on its general use since it is dependent on various experimental factors (e.g., thermocouple location when calibrating TSA response to the monitored temperature).

In the following paragraphs various normalized mathematical expressions are defined and utilized for both the experimental data as well as for the theoretical formulations. The normalized expressions are required so as to allow a direct comparison between the experimental data, which is in the form of nondimensional digital values representing the IR camera's electrical output, and the theoretical formulations, which provide a direct value for the temperature range resulting from the thermoelastic effect. The goal here is to verify the experimental behavior with the theory, thereby providing confidence in the experimental measurements. Once confidence is achieved, direct empirical relationships can be developed between the IR camera's digital values and a structure's stress state.

Figure 4 shows the thermoelastic response for the TIMETAL 21S specimens as a function of the mean stress. It is seen that a change in the mean stress from $-276$ MPa ($-40$ ksi) to 276 MPa (40 ksi) induced, on average, a 21 percent increase in the IR signal. Subsequently, the data of the three specimens were pooled and fitted with a linear first-order regression. The constants of the regression fit are presented in Fig. 4. The regression constants were used to develop the experimentally determined measure of the mean stress dependence of the thermoelastic parameter (see Eq. (10)). This was accomplished by dividing the slope of the regression line by the intercept of the same regression line. Table 1 compares the experimentally measured normalized mean stress effect with the theoretical value obtained using Eq. (10) and material properties from Castelli et al. [8]. The comparison indicates a 2.4 percent difference between the theoretical and experimental values.

**Table 1 Comparison of theoretical and experimental mean stress dependence of *K***

| Material* | $\alpha$, $10^{-6}$/°K | $E_{23°C}$, GPa | $\partial E/\partial T$, MPa/°K | Theory: Eq. (10), $MPa^{-1}$ | Experiment, $MPa^{-1}$ |
|---|---|---|---|---|---|
| TIMETAL 21S | 8.24 | 116 | $-37.6$ | $3.39\times10^{-4}$ | $3.47\times10^{-4}$ |

*Material properties obtained from Castelli et al. [8]

Next, the experimental results were compared with the theoretical prediction of Eq. (8). Presented in Figs. 5 and 6 are the first harmonic and second harmonic response curves, respectively. It is seen that the first harmonic is linearly related to the stress amplitude. In keeping with the approach of Wong et al. [6], an equation of the form, $S_\omega = a(\Delta\sigma)$ was used to describe the data. A curve fit of the data in Fig. 5 yielded a slope of $a = 103.0\ \mathrm{MPa}^{-1}$.

Now focusing on the second harmonic response of Fig. 6, it can be seen that the IR response is linearly proportional to the square of the cyclic amplitude and can best be described by an equation of the form, $S_{2\omega} = b(\Delta\sigma^2)$. The calculated slope was $b = 6.53 \times 10^{-3}\ \mathrm{MPa}^{-2}$. Next, the ratio of the slopes of the experimentally determined first and second harmonics was calculated as $b/a = 0.634\times10^{-4}\ \mathrm{MPa}^{-1}$. This was done so as to allow a comparison with Eq. (8). The theoretical ratio, using Eq. (8) with a zero mean stress, is as follows:

$$\frac{b}{a} = \left|\frac{1}{4\alpha E^2}\frac{\partial E}{\partial T}\right|. \tag{13}$$

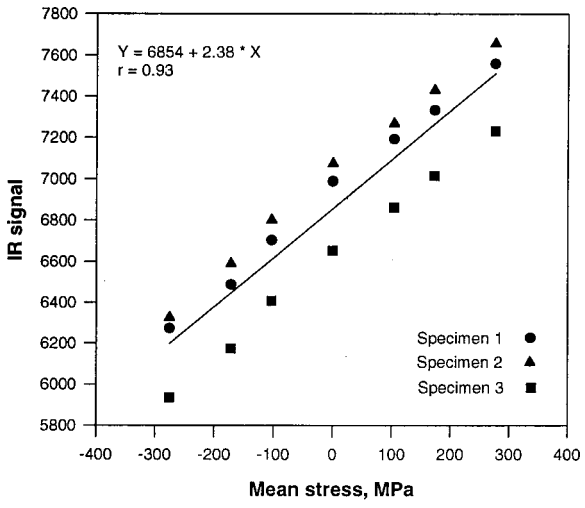



**Fig. 4 Mean stress dependence of IR signal for three specimens. Equation represents first-order linear regression fit of all data.**

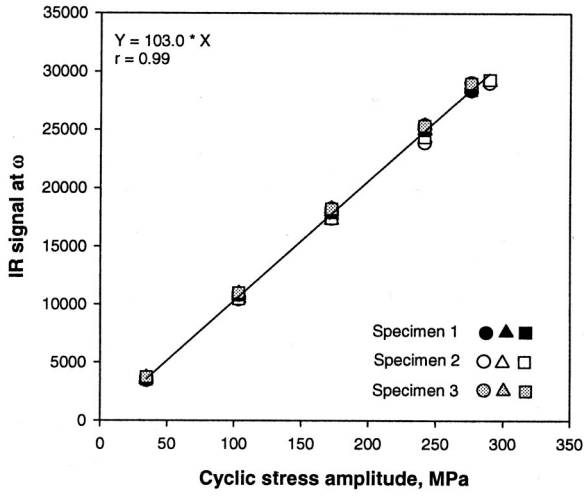



**Fig. 5 First harmonic IR signal output at various stress amplitudes ($\sigma_m = 0$). The plot contains data from three specimens; each with three repetitions. Equation represents first-order linear regression fit of all data.**

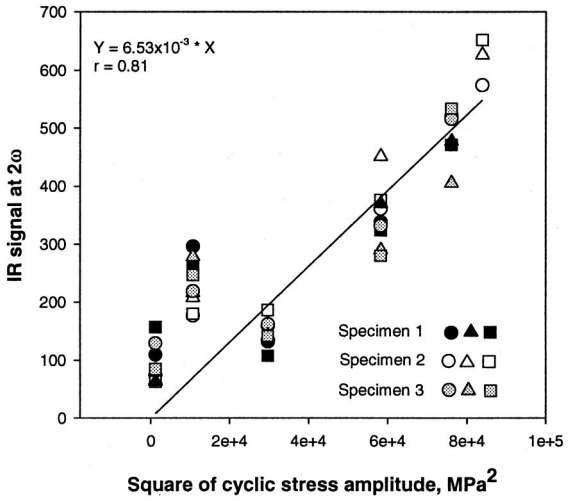



**Fig. 6 Second harmonic IR signal output at various amplitudes ($\sigma_m=0$). The plot contains data from three specimens; each with three repetitions. Equation represents first-order linear regression fit of all data.**

Utilizing the material properties in Table 1 (obtained from Castelli et al. [8] Eq. (13) gives a predicted value of $b/a=0.846\times10^{-4}\,\mathrm{MPa}^{-1}$ ([9]). The difference between the theoretical and experimental values is 25 percent. The following section discusses the experimental results.

## Discussion and Conclusion

The mean stress dependence has been shown for the current material system, TIMETAL21S, both experimentally in Fig. 4 as well as theoretically by Eq. (10). The error between the theory and experiment is reasonable and presents an improvement over the error reported in the previous study by the authors (Ref. [9]). The improvement is a result of using the temperature correction curve defined during this study as well as utilizing more accurate material properties for the theoretical solution.

By utilizing a linear frequency doubler in line with the reference signal, the second harmonic of the IR signal was successfully measured using the TSA system's lock-in electronics. As a result, the practicality of obtaining simultaneous solutions for a $128\times128$ array of IR detectors was maintained. The linear relationship observed in the Fig. 5 agrees with the theory where the slope, $a$, is directly related to the thermoelastic parameter, $K$. Changing the mean stress would result in a changing slope according to Eq. (10). Therefore, it is possible at this point to utilize plots of the first harmonic such as the one in Fig. 5 and to use the slopes of such curves to empirically solve for the mean stress. However, a value for the stress amplitude is required. This is only practical when characterizing coupons in a laboratory environment or when testing complex structures which have been fitted with strain gages.

It is seen in Fig. 6 that the second harmonic data has a larger spread than the first harmonic data of Fig. 5. This is due to the fact that the second harmonic signal is only about 2 percent of the primary signal. As a result, the second harmonic IR signal is closer to the noise band of the system especially when utilizing small cyclic amplitudes. Also, it is seen that some aberrations in the data did occur. Note that the second harmonic IR values at the 103 MPa (15 ksi) level are larger than the IR values at 34 MPa (5 ksi) (see Fig. 6). No explanation for this behavior is available, although, it could be due to the low value of the IR signal and its proximity to the noise band of the TSA system. Although the error between the theoretical value of $b/a$ and the experimental $b/a$ is reasonable, improvements can be achieved. The current error is attributed to the sensitivity of the theoretical value to the material properties, and the fact that the material properties utilized here were obtained from Ref. [8] and not directly measured. Therefore, it appears that confidence was achieved concerning the accuracy of the experimentally obtained data. It is now reasonable to develop empirical equations that directly correlate the TSA results (i.e., the nondimensional digital values of the IR camera) to the cyclic stress amplitudes and mean stress states of components.

Difficulties do arise with the introduction of multiaxial stresses. The multiaxial problem is very complex, although, some attempts have been made to address the issue of solving for the individual stress components. Past researchers have combined the bulk stress TSA data with numerical methods to obtain stress separation in limited cases. Also, the possibility does exist for combining the results of TSA, which provides information concerning the sum of the principal stresses, and photoelasticity, which provides information on the differences and directions of principal stresses. With on going research in these areas, the TSA method does show promise as a residual stress assessment tool.

## Acknowledgment

The author would like to thank Dr. George Y. Baaklini of the NASA Glenn Research Center at Lewis Field as well as the Aviation Safety Program for providing the funding for this study.

## Appendix

The thermoelastic stress analysis was conducted using the DeltaTherm 1000 system manufactured by Stress Photonics Incorporation. The following system software settings were utilized during IR data acquisition: accumulation time=4.7 seconds; gain=1; and AC channel integration time constant=120 seconds. Collection time for any given test condition was 240 seconds. This consisted of 120 seconds to collect the first harmonic and then activating the frequency doubler and obtaining a 120 seconds reading for the second harmonic. Note the timer was reset after each 120 second capture.

## References

[1] Lesniak, J. R., and Boyce, B. R., 1994, "A High-Speed Differential Thermographic Camera," *SEM Spring Conference Proceedings*, Baltimore, MD, June 6–8, pp. 491–497.
[2] Thomson, W. (Lord Kelvin), 1853, "On the Dynamical Theory of Heat," Trans. - R. Soc. Edinbrgh, **20**, pp. 261–283.
[3] Machin, A. S., Sparrow, J. G., and Stimson, M. G., 1987, "Mean Stress Dependence of the Thermoelastic Constant," Strain, **23**, pp. 27–30.
[4] Dunn, S. A., Lombardo, D., and Sparrow, J. G., 1989, "The Mean Stress Effect in Metallic Alloys and Composites," *SPIE Stress and Vibration: Recent Developments in Industrial Measurement and Analysis*, Vol. 1084, pp. 129–142.
[5] Wong, A. K., Jones, R., and Sparrow, J. G., 1987, "Thermoelastic Constant or Thermoelastic Parameter?" J. Phys. Chem. Solids, **48**, pp. 749–753.
[6] Wong, A. K., Sparrow, J. G., and Dunn, S. A., 1988, "On the Revised Theory of the Thermoelastic Effect," J. Phys. Chem. Solids, **49**, pp. 395–400.
[7] Lesniak, J. R., and Boyce, B. R., 1995, "Differential Thermography Applied to Structural Integrity Assessment," *Proceedings of the International Conference on Thermal Sensing and Imaging Diagnostic Applications*, Vol. 2473, Orlando, FL, Apr. 19–21, pp. 179–189.
[8] Castelli, M. G., Arnold, S. M., and Saleeb, A. F., ?, "Specialized Deformation Tests for the Characterization of a Viscoplastic Model: Application to a Titanium Alloy," NASA TM-106268, in print.
[9] Gyekenyesi, A. L., and Baaklini, G. Y., 1999, "Thermoelastic Stress Analysis: The Mean Stress Effect in Metallic Alloys," *Proceedings of SPIE Conference on the NDE of Aging Materials and Composites*, Newport Beach, CA, Mar. 1–4.

**T. Berruti**
**S. Filippi**
**L. Goglio**
**M. M. Gola**

Department of Mechanical Engineering,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
10129 Turin, Italy

**S. Salvano**
FIAT AVIO,
Via Nizza 312,
10127 Turin, Italy

# A Test Rig for Frictionally Damped Bladed Segments

*The paper describes the development of an instrumented test rig suitable to carry out vibration testing of the bladed segments of a turbine (belonging to a stator low pressure stage) in which the friction occurring at the contact of neighboring segments is used in service to damp the vibrations. The system is able to assign the desired preload to the segment and to measure the forces exerted in the contact and the relative displacements due to slipping. The main results that can be obtained concern the type of motion that takes place (stick/slip) and the effect of the interlocking on it, the friction coefficients, the contact tangential stiffness, and, more generally, the frequency response.*
[DOI: 10.1115/1.1419015]

## Introduction

The contact among blades or bladed segments of turbomachinery is difficult to model, mainly due to the nonlinearity of the phenomenon and to the uncertainty affecting some data. The topic is of considerable interest, since the friction developed in the contact is used to damp the vibration of the blades. The literature shows that experimental and theoretical results are often not in accordance, thus it is required to have at disposal both of them. Yet more than 15 years ago, Srinivasan and Cutts [1] tested a complete rotor blade assembly to measure rubbing at the shrouds. The motion was measured by means of special sensors based on photoemitting diodes.

More commonly, test rigs are used to simulate the blade assembly, or part of it, reproducing and measuring its behavior.

Griffin [2] used a single blade with a friction damper, vibrated on a shaker table, to measure stresses and resonant frequencies for different contact normal loads. Srinivasan and Cutts [3] built a loading device to reproduce the shroud rubbing of a blade with the neighboring ones. It allowed for measuring strain and accelerations for different contact normal loads.

Very recently, Berthillier et al. [4] used a simple cantilever beam with friction dampers on opposite sides to reproduce the behavior of a rotor blade. The point-to-point frequency response (acceleration/force) was measured. Yang and Menq [5] described a test rig to reproduce the behavior of the wedge dampers placed under the blade root. The wedges were pulled by wires to simulate the centrifugal force and the excitation was applied by an air jet. The stress at the root blade was measured. A similar system, but with mechanical excitation, has been used by Sanliturk et al. [6] and by Csaba [7]. In both cases the blade accelerations were measured.

The present paper considers the stator blades of a low pressure stage of an aeronautical engine ([8]). In this construction the blades are grouped in segments, each of them formed by six blades connected at their ends by the ''upper platform'' (external stator radius) and by the ''lower platform'' (internal stator radius), respectively, (Fig. 1). The upper platform of each segment is fixed to the engine casing, the lower platform is in contact (on surfaces normal to the turbine axis) with the lower platforms of the two neighboring segments. The mating surfaces are mounted with interference, referred to as ''interlocking,'' in axial direction. Thus a normal force is exerted and the segments are elastically predeformed. During operation the relative slipping of the surfaces generates the friction that is used to reduce the vibrations.

A test rig has been developed aiming at the following scopes:

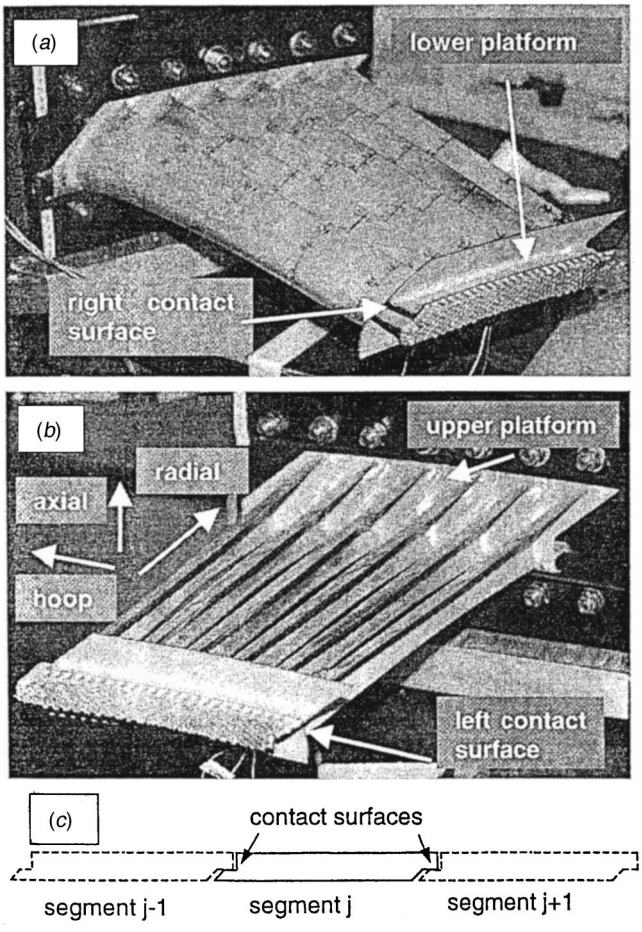



**Fig. 1 Stator segment; (*a*) right side view, (*b*) left side view, (*c*) schematics of the assembly**

Contributed by the International Gas Turbine Institute (IGTI) of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Munich, Germany, May 8–11, 2000; Paper 00-GT-538. Manuscript received by IGTI November 1999; final revision received by ASME Headquarters February 2000. Associate Editor: D. Wisler.

- to reproduce in laboratory, but on the actual parts, the rubbing phenomenon that is expected to occur in service, measuring the quantities of interest, i.e., contact force (normal and tangential components) and relative displacement;
- to obtain the friction coefficients and the tangential stiffness at the contact interface;
- to investigate on the effect of the interlocking; and
- to allow for validating the theoretical modeling of the phenomenon.

The paper is focussed on the first two topics of the previous list. The first two sections discuss the features of the test rig and the design and calibration of the transducers, the third section presents the early results.

The work is a part of a research program carried out under a contract between Politecnico di Torino and FIAT Avio.

## Test Rig

The test rig has been designed to hold the segment during vibration testing, apply the desired interlocking and carry the force and displacement transducers. A sketch is shown in Fig. 2.

Basically the test rig consists in

- a frame to which the upper platform of the segment is clamped (see Fig. 1) and
- a cradle pivoting on the frame and carrying two elastic supports, in contact with the segment, which simulate the presence of the neighboring segments.

The displacements of the contact points, both of the platform and of the supports, are not directly measurable since the points are obviously not accessible. To overcome such inherent difficulty the following strategy has been adopted. Assuming that, for this purpose, the lower platform behaves as a rigid body, its hoop displacement is measured in a reachable zone by means of a laser transducer. The displacements of the contact points of the supports are measured by means of the same strain gages which are used to measure the contact forces (as described in a following section).

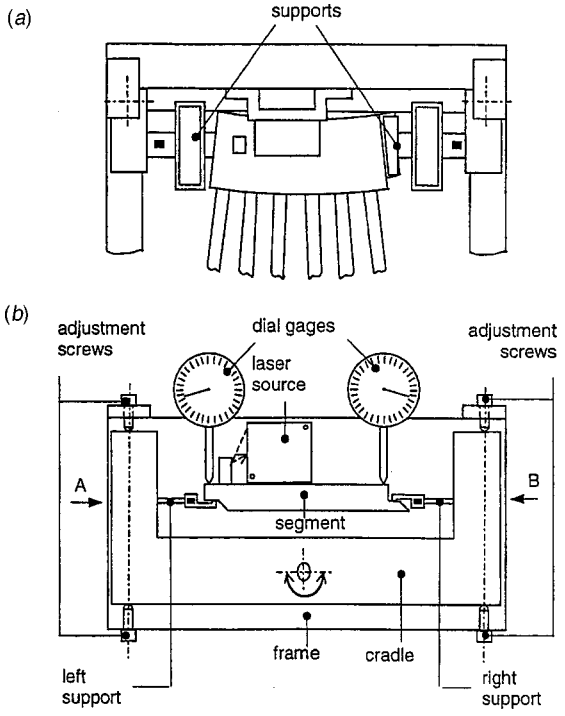



**Fig. 2 Sketch of the test rig; (*a*) top view, (*b*) detail of the preloading (A, B accelerometer locations)**

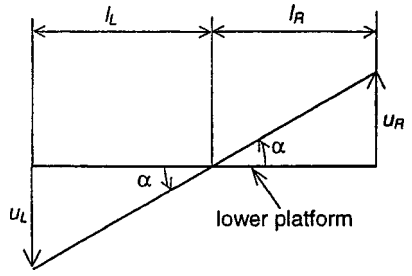



**Fig. 3 Lower platform rotation of a twist angle $\alpha$**

Both displacements are measured relatively to the test rig. Then, for each contact the hoop relative displacement is obtained as difference of the displacements measured by laser and strain gages.

**Frame.** It is a very rigid aluminum alloy block, fixed to the shaker table, on which all mechanical parts and sensors are mounted.

The clamping system of the segment consists in two slots machined in the frame in which two ribs protruding from the upper platform are inserted. Two rows of seven bolts ensure fastening.

On the frame (see Fig. 2) is locked the laser source (Micro-Epsilon, model LD 1605-4, measuring range $\pm 2$ mm, max nonlinearity $\pm 0.3$ percent, approximate size $50\times50\times20$ mm), while the reflector is glued on the lower platform of the segment.

**Cradle.** The cradle (approximately U shaped, see Fig. 2), rotates with respect to the frame about an axis parallel to the segment radial direction. Adjustment is performed by means of four screws. As sketched in Fig. 3, a rotation of the cradle of the angle $\alpha$ induces respectively at the right and left ends of the lower platform (placed at distances $l_R$ and $l_L$ from the rotation center) two opposite displacements $u_R$ and $u_L$. The angle $\alpha$ is the "angle of twist," while the value $i=u_R+u_L$ is the interlocking. Therefore, the desired value of interlocking and, as a result, the values of the normal forces exerted on the mating surfaces of segment and supports can be applied by rotating the cradle. The value of imposed interlocking is measured by two dial gages (resolution 0.01 mm).

When the desired value of interlocking has been reached, the cradle is locked by six bolts perpendicular to the adjustment screws.

## Instrumented Supports

The cradle carries two identical supports in contact with the slipping surfaces of the segment as shown in Fig. 2. A sketch is shown in Fig. 4. The end to be fixed to the cradle is cylindrical in order to allow the alignment (by small rotation of the cylinder) of the contact surfaces during mounting. The slipping surface in contact with the segment is covered with the same material, ceramic coating applied by plasma spray technology, laid on the mating surfaces on the segment. In this way the same friction behavior

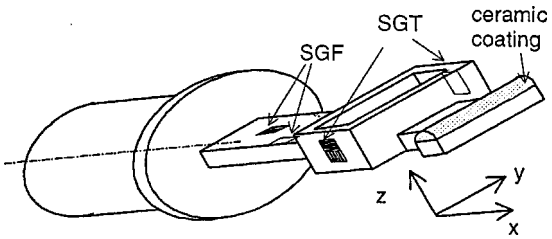



**Fig. 4 Sketch of the support instrumented with strain gages**

between mating surfaces of neighboring segments is reproduced. The slipping surface of the supports have a single curvature in order to ensure linear contact.

A local reference system $x$, $y$, $z$ for the support is defined (see Fig. 4). After the assembly of the supports on the frame, the correspondence of the support local reference system and the segment global reference system (see Fig. 1) is the following:

- support $x \leftrightarrow$ segment hoop,
- support $y \leftrightarrow$ segment radial, and
- support $z \leftrightarrow$ segment axial.

The supports have been designed with a particular architecture and dimensions in order to show stiffness (at least in the $z$-direction) of the same order of magnitude of the two neighboring segments that they substitute. Thus, the same interlocking values imposed in the test rig and in the actual engine assembly will lead to normal contact forces of the same order of magnitude.

Furthermore, the supports act as transducers (instrumented with strain gages) measuring normal and tangential forces and displacements due to the contact. Obviously from this viewpoint the support must be flexible enough (in axial and hoop direction) so that the strain gages mounted on it can measure the imposed strain. The shape and dimension is as a tradeoff between such requirements: essentially the zone closest to the cylindrical end rules the $z$ stiffness, the central ring rules the $x$ stiffness.

**Finite Element Method (FEM) Analysis of the Support.** A finite element method (FEM) analysis has been performed to assess stiffness and dynamic properties of the support. The model (MSC/Nastran) is shown in Fig. 5 and consists of 5204 solid elements HEXA8.

The following FEM calculations have been developed.

- Static analysis in order to obtain the values of stiffness in $z$ and $x$ direction: a unit force has been applied in the center of the contact area in the $z$-direction and the consequent displacement has been calculated, then the stiffness (in the $z$-direction) is obtained as force to displacement ratio. The same has been done for the $x$-direction. The results are listed in Table 1 where they are compared to the experimental values of stiffness obtained by the support calibration as explained in next paragraphs. The stiffness values obtained in the same directions by the NASTRAN segment model are also listed in Table 1.
- Dynamic analysis to obtain natural frequencies and modal shapes of the support; the results are listed in Table 1 compared to the experimental values.

**Experimental Modal Analysis of the Support.** The support is fixed to cradle in the same way as in use. The end is excited by an instrumented hammer and the response is detected by means of a one-axis accelerometer placed near the contact area. The measured resonant frequencies match well (error$<$5 percent) with the numerical ones as shown in Table 2. Considering that in operation the supports are elastically sustained at the tip, the resonant frequencies are surely uncoupled from those of interest in the measurements.

**Table 1 Stiffness of the support and of the segment in $x$ (hoop) and $z$ (axial) directions**

| Stiffness (N/mm) | | | |
|---|---|---|---|
| Support | | Segment | |
| FEM | Experimental | FEM | |
| 229 | 220 | 327 | $z$ (axial) |
| 1825 | 1892 | 822 | $x$ (hoop) |

**Table 2 Natural frequency and mode shape features of the support**

| Frequency (Hz) | | |
|---|---|---|
| FEM | Experimental | Mode Shape Feature |
| 695 | 668 | $z$ displacement (axial) |
| 1136 | 1081 | $y$ displacement (radial) |
| 1621 | 1575 | $x$ displacement (hoop) |

**Support Instrumentation.** The support has been instrumented as shown in Fig. 4 as follows:

- two strain gages (half bridge connection) measuring the flexural deformation of the support in the $z$-direction (SGF): one SGF on the upper surface and the other on the lower surface of the bending element;
- four strain gages (full bridge connection), measuring the tensile deformation of the support in the $x$-direction (SGT): two SGT (perpendicular each other) are placed on one side of the central frame, the other two on the opposite side.

The strain gages are by Measurements Group, type EA-06-062AQ-35, nominal resistance$=350\,\Omega$.

The supports have been calibrated as explained in the next subsection. Each one acts as a transducer where the normal contact force and $z$-displacement are deduced from SGF and the tangential contact force and the $x$ displacement are deduced from SGT.

## Support Calibration

**Tensile Calibration.** As shown in Fig. 6($a$) the support is positioned on a steel ball (ball bearing ball) and loaded by weights

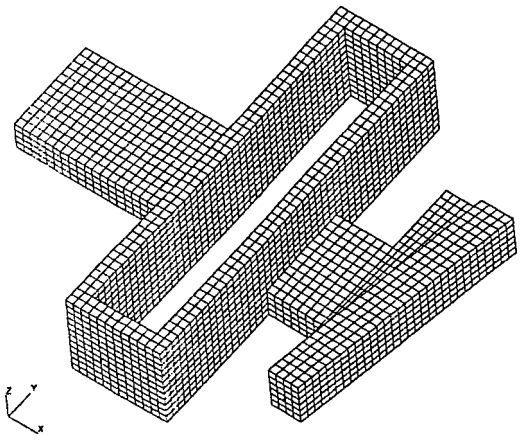


**Fig. 5 Support finite element method model**

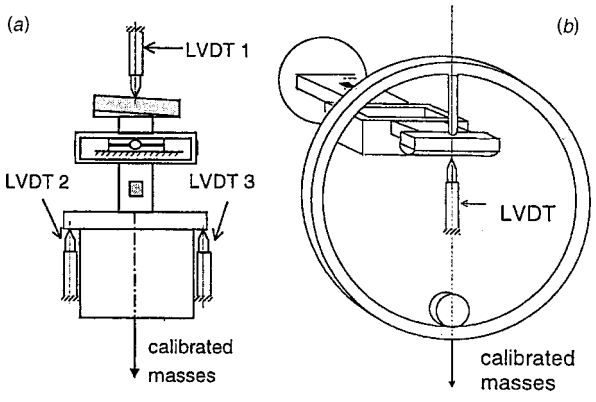



**Fig. 6 Sketch of the support calibration; (*a*) tensile, (*b*) flexural**

hanged on the cylindrical end. Three LVDT have been used in order to measure the displacement of the support. The difference between the mean value (to clear the effects of rotation) of signals from LVDT 2 and LVDT 3 and the value of the signal from LVDT 1 gives the elongation of the support during loading. Masses of 1 kg have been applied for four cycles each of loading and unloading to a maximum of 15 kg. After these settling cycles, two calibration loading cycles have been applied and strain gages signals recorded. The strain gages signals were considered not only from SGT but also from SGF to detect spurious coupling between tension and flexural measurement.

**Flexural Calibration.** The loading application setup for flexion calibration is sketched in Fig. 6(*b*). The cylindrical end of the support is fixed in the cradle which is locked to the frame. The LVDT measures the deflection of the support. Also in this case after four settling cycles (from 1 to 15 kg loading and unloading) two calibration loading cycles have been applied. The strain gages signals were considered not only from, SGF but also from SGT to detect spurious coupling between flexion and tensile measurement.

The following symbols are used:

- $Z_j$ (mV) is the electrical signal from the strain gages, the index $j$ stands for the type of strain gage (*F*-flexural, *T*-tensile);
- $P_T$ ($N$) is the tensile load, $u_T$ ($\mu$m) is the corresponding elongation (from LVDT) of the support (Fig. 6(*a*));
- $P_F$ ($N$) is the flexural load, $u_F$ (mm) is the corresponding deflection (from LVDT) of the support (Fig. 6(*b*));
- $a_{jk}$ (mV/N) is the element of the sensitivity matrix $[A]$ relating the electrical signal $Z_j$ to the loads, the index $j$ stands for the type of strain gage (*F*-flexural, *T*-tensile), the index $k$ for the type of load (*F*-flexural, *T*-tensile);
- $b_{jk}$ (mV/$\mu$m) is the element of the sensitivity matrix $[B]$ relating the electrical signal $Z_j$ to the displacements, as in $[A]$ the index $j$ stands for the type of strain gage, the index $k$ for the displacement in tensile calibration ($T$) or in flexural calibration ($F$).

From each calibration of Fig. 6 four experimental sets of data (two relating the SG signals to the load, and two relating the SG signals to the displacement) are obtained. The data are least squares fitted with straight lines. The values of their slopes form the elements of the sensitivity matrices $[A]$ and $[B]$ as follows:

$$\{Z\}=\begin{Bmatrix} Z_F \\ Z_T \end{Bmatrix}=\begin{bmatrix} a_{FF} & a_{FT} \\ a_{TF} & a_{TT} \end{bmatrix}\begin{Bmatrix} P_F \\ P_T \end{Bmatrix}=[A]\{P\}$$
$$\{Z\}=\begin{Bmatrix} Z_F \\ Z_T \end{Bmatrix}=\begin{bmatrix} b_{FF} & b_{FT} \\ b_{TF} & b_{TT} \end{bmatrix}\begin{Bmatrix} u_F \\ u_T \end{Bmatrix}=[B]\{u\}. \quad (1)$$

By inverting the two matrix Eqs. (1) the calibration factors, to be used in the dynamic tests to evaluate forces and displacements from the signals of the strain gages, are obtained as elements of the inverse matrices $[A]^{-1}$ and $[B]^{-1}$:

$$\{P\}=[A]^{-1}\{Z\} \quad \{u\}=[B]^{-1}\{Z\}. \quad (2)$$

The values for the two supports are listed in Table 3. The smallness of the off-diagonal terms confirms that the coupling between flexural and tensile signals is very limited. In the table are also reported, arranged in the same order of the related matrix coefficients, the standard deviations of the calibration factors. These have been calculated (see the Appendix) from the standard deviations of the sensitivity coefficients which are obtained from the linear fitting of the experimental data. It can be noticed that, as far as the diagonal terms are concerned, the estimates of the calibration factors are much larger than their respective standard deviation, and this shows that they are obtained with good precision. Conversely, for the off-diagonal terms, the estimates are not so much larger than their standard deviation. As it could be expected this terms, which are less relevant, are obtained with lower relative precision.

**Table 3 Calibration factors and related standard deviations**

| $[A]^{-1}$ (N/mV) | | $[B]^{-1}$ ($\mu$m/mV) | |
|---|---|---|---|
| **Estimates of the calibration factors** | | | |
| Support L (segment left side) | | | |
| $1.62\cdot10^{-1}$ | $-2.90\cdot10^{-3}$ | $7.38\cdot10^{-1}$ | $-1.30\cdot10^{-2}$ |
| $4.62\cdot10^{-3}$ | $-4.43\cdot10^{-1}$ | $2.31\cdot10^{-3}$ | $-2.23\cdot10^{-1}$ |
| Support R (segment right side) | | | |
| $1.58\cdot10^{-1}$ | $-8.54\cdot10^{-4}$ | $7.33\cdot10^{-1}$ | $-3.96\cdot10^{-3}$ |
| $4.23\cdot10^{-3}$ | $-4.43\cdot10^{-1}$ | $2.48\cdot10^{-3}$ | $-2.59\cdot10^{-1}$ |
| **Standard deviations of the calibration factors** | | | |
| Support L (segment left side) | | | |
| $3.39\cdot10^{-4}$ | $2.91\cdot10^{-4}$ | $5.32\cdot10^{-3}$ | $1.49\cdot10^{-3}$ |
| $2.43\cdot10^{-4}$ | $8.07\cdot10^{-4}$ | $1.40\cdot10^{-4}$ | $4.23\cdot10^{-3}$ |
| Support R (segment right side) | | | |
| $2.28\cdot10^{-4}$ | $1.31\cdot10^{-4}$ | $3.61\cdot10^{-3}$ | $6.05\cdot10^{-4}$ |
| $1.41\cdot10^{-4}$ | $4.75\cdot10^{-4}$ | $8.24\cdot10^{-5}$ | $1.59\cdot10^{-3}$ |

## Tests

All vibration tests have been performed with the test rig mounted on a shaker (Unholtz-Dickie N. TA250D-206, force 1700 lbs pk) and applying the excitation in the segment hoop direction. The data have been recorded by means of an InstruNET system (sampling rate 20 kHz).

**Natural Frequencies of the Segment.** Preliminarily, the lowest natural frequencies of the segment, clamped at the upper platform and not in contact with the elastic supports, have been identified. An accelerometer has been placed on the lower platform (measuring axis in axial direction) in order to acquire the frequency response of the segment. The frequency sweep imposed by the shaker ranged from 50 to 1000 Hz. The mode shape features were observed by means of a stroboscopic system. The first three mode shape features concern the deformation of the whole segment, the higher frequency modes concern mainly deformation of the blades whilst the lower platform remains almost still. Table 4 lists the measured values of the first frequencies, together with the results of the finite element model ([8]).

**Verification of the Frame Stiffness.** The stiffness of the frame, in the region carrying the laser source and the cradle with the instrumented supports was experimentally verified. Figure 7 shows the accelerations measured in the positions A e B (as indicated in Fig. 2) due to an excitation of 1 g peak applied by the shaker. The signal of the accelerometer A was used as feedback for the shaker control.

It can be noticed that below 400 Hz the accelerations values are identical for the two measurement points, thus in the frequency range of interest the region of the frame carrying the laser source and the cradle behaves rigidly in the hoop direction.

**Table 4 Natural frequencies of the segment (without supports, upper platform constrained)**

| Frequency (Hz) | | |
|---|---|---|
| **FEM** | **Experiment** | **Mode shape feature** |
| 141 | 140 | hoop translation |
| 276 | 282 | radial rotation |
| 341 | 326 | axial translation |

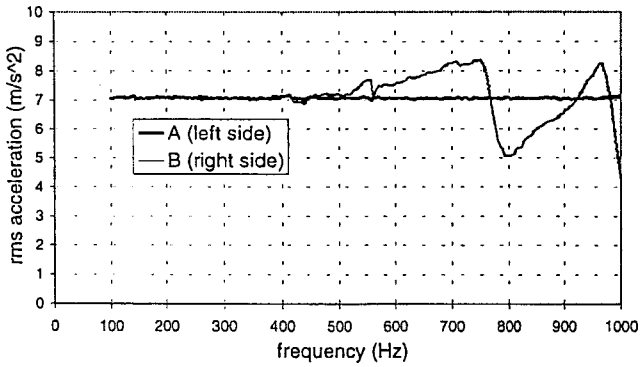



**Fig. 7 Accelerations measured in hoop direction for the positions A and B of the frame (excitation 1 g pk)**

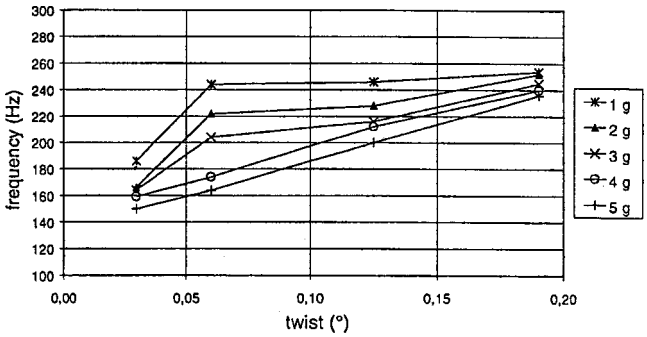



**Fig. 8 First mode resonant frequency of the segment versus twist value under different levels of excitation**

**Behavior Under Different Interlocking Values.** Vibration tests were performed with different values of imposed interlocking. As previously done without supports, at first for each case of interlocking the response to a frequency sweep (50–1000 Hz) was measured under five levels of excitation (1, 2, 3, 4, 5 g). At this stage of the research, among all resonant peaks only the behavior of the first one (displacement mainly in hoop direction) was examined. This mode, in fact, is the most involved in the phenomena under study of slipping contact in hoop direction. In Fig. 8 the value of the first resonant frequency versus the twist angle, under different excitation levels, is plotted. As it could be expected, by increasing the twist angle the segment is more constrained, since higher friction force is applied, as a consequence its first resonant frequency increases. Similarly, by reducing the level of excitation the segment moves less and it results to be more constrained so that its first resonant frequency increases. As shown in Fig. 8, the first resonant frequency increases up to a value of about 250 Hz. Since some doubts could arise about the nature of this resonance, it has been verified that the features of the mode shape associated with this increasing frequency value are those of the first one. In

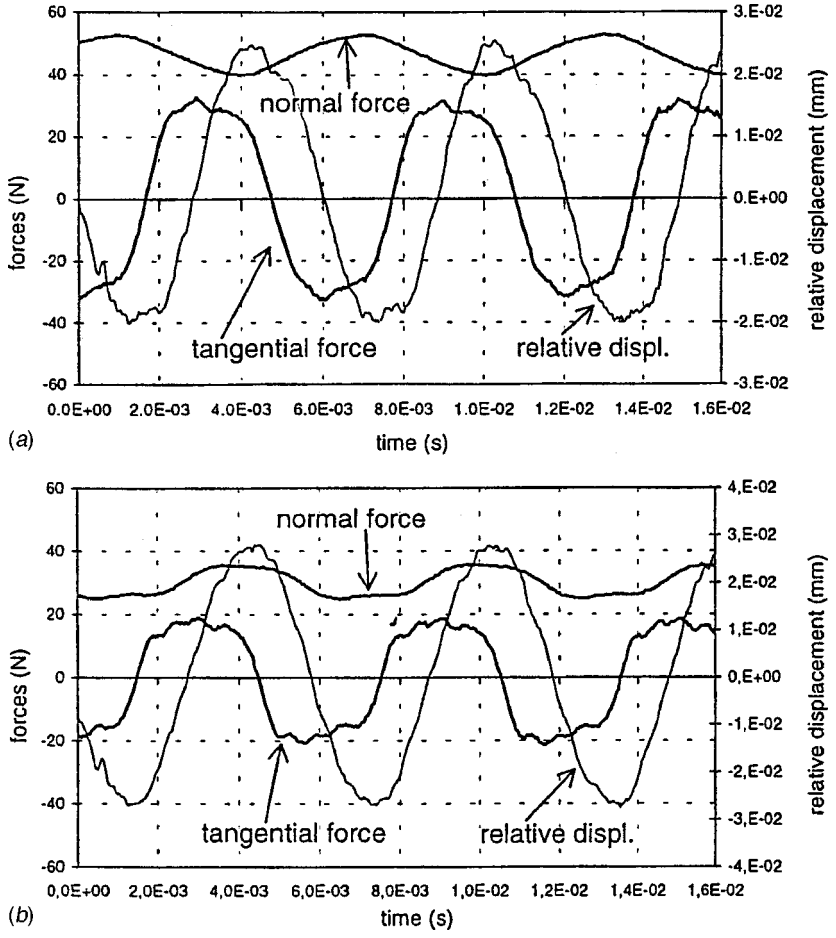



**Fig. 9 Measured time histories of relative hoop displacement normal and tangential contact forces. Twist angle 0.10 angle frequency 165 Hz, excitation 5 g; (*a*) left support/segment contact, (*b*) right support/segment contact.**

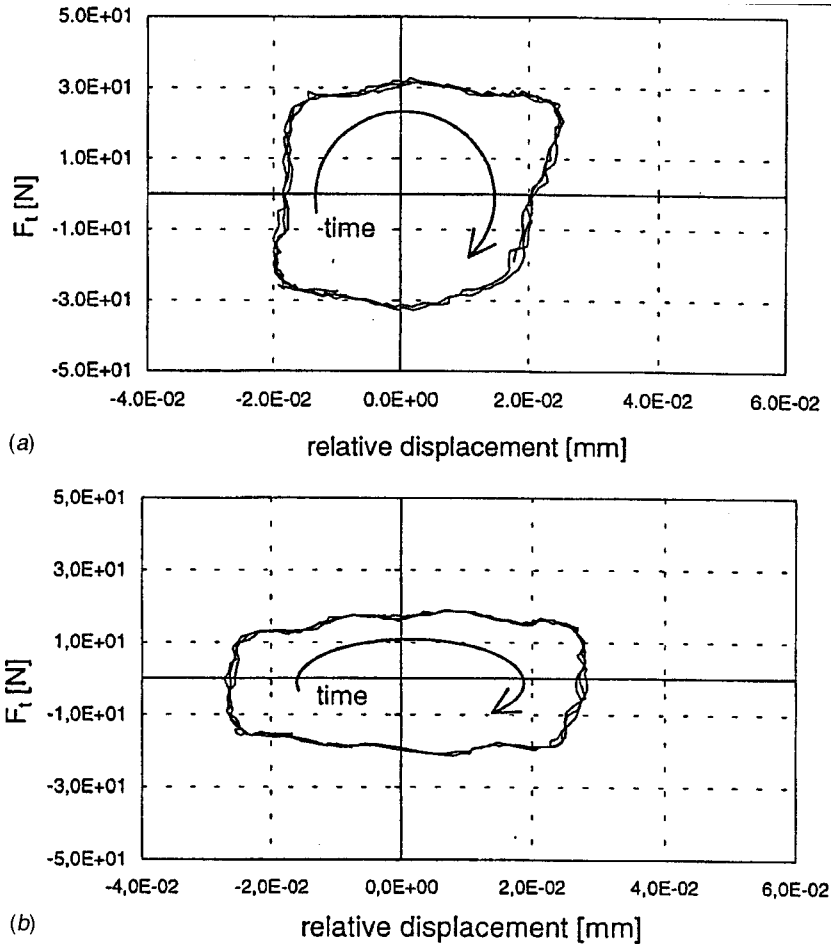



**Fig. 10 Measured tangential force versus relative hoop displacement. Twist angle 0.10 deg, frequency 165 Hz, excitation 5 g; (*a*) left support/segment contact, (*b*) right support/segment contact.**

fact the comparison of the signals from SGT and SGF shows that the displacement is mainly along the hoop direction.

As example of the experimental results that are obtained by the test rig, Fig. 9 shows (for the contact on both left and right side) the recorded time history of relative hoop displacement, normal force, and tangential force over almost three periods. The angle of twist is 0.10 deg, the excitation amplitude is 5 g and the frequency is 165 Hz, corresponding to resonant conditions as detected by the above-mentioned frequency sweep. It should be noted that the orders of magnitude are some tens of N for the forces and some hundredths of mm for the relative displacement, the latter value allows to appreciate the test rig sensitivity.

It can be noted from Fig. 9 that the parts never lose contact, since the normal force never drops to zero. Furthermore, by comparison of the curves of tangential force and relative displacement it can be noticed that, with this combination of interlocking, excitation amplitude, and frequency, slipping conditions are reached on both segment sides. During slip the tangential force does not remain exactly constant, due to the simultaneous variation of the normal force and also to a high frequency noise possibly due to the transducer hoop resonance (Table 2). Also the relative displacement is continuously changing in time.

The data related to relative displacement and tangential force are also reported in Fig. 10 for the contact on both segment sides. The plots show the well-known cyclic behavior, approximately quadrilateral shaped with two branches almost horizontal and the other two almost vertical. The observation of the former branches, corresponding to slipping conditions, suggests that adopting of the Coulomb's law, as usually done in the theoretical modeling of the contact, works anyway reasonably well. The friction coefficient evaluated in any instant as the ratio of tangential to normal force is in the range 0.6–0.7.

On the contrary the apparent tangential stiffness, evaluated as slope of the almost vertical branches of the cyclic diagrams in Fig. 10 assumes different values. For the left contact (Fig. 10(*a*)) the values are 16.5 and 6 kN/mm, respectively, for the ascending and descending branches. For the right contact (Fig. 10(*b*)) the corresponding values are 18.6 and 8.2 kN/mm. The different slope values for the ascending and descending branches of the same cycle can be justified by the following argument. In the case of mono-dimensional relative motion the reversal of the motion requires that a slip-stick-slip transition occurs. Therefore the tangential contact stiffness is involved when sticking occurs and the force varies from $+\mu N$ to $-\mu N$, or vice versa (where $\mu$ is the friction coefficient and $N$ is the normal contact force).

If the relative motion occurs on an elongated orbit, i.e., the trajectory is two-dimensional (as it was found in [8]), the motion reversal does not implies the slip-stick-slip transition. In fact, the relative motion direction changes progressively from hoop to radial and again to hoop in opposite direction. During this phase slipping never ceases, thus the tangential force is always equal to $\mu N$ and parallel to the relative velocity. In these conditions, the plot of the tangential force versus the hoop component of the relative displacement gives again a close loop, but the apparent tangential stiffness is no longer related to an elastic tangential deformation as in the mono-dimensional case. Moreover, if the

motion reversal does not occur symmetrically at the ends of the orbit (trajectory), different slopes can be detected, in the plot $F_t$ versus relative hoop displacement, respectively, for the ascending and descending branches. In a real case the situation can be even more complicated, since both two-dimensional motion and stick-slip can occur. However, further analysis is required to get more physical insight of this aspect.

## Conclusions

A test rig has been constructed, instrumented and checked to carry out in laboratory dynamic testing of bladed segments of turbine engines. The facility reproduces the type of assembly that is used in service, especially from the viewpoint of the friction that is developed and used to damp vibrations. Special supports that act as force and displacement transducers have been built and calibrated. The experimental results show that the test rig is able to reproduce the typical behavior of such a type of bladed assembly, measuring the quantities of interest.

## Appendix

**Error Propagation in Matrix Inversion.** The standard deviations of the calibration factors reported in Table 3 have been obtained in the following way. As it is well known, if $\eta$ is a function $f$ of $N$-independent random variables $\xi_1, \xi_2, \ldots, \xi_N$, than its standard deviation $\sigma_\eta$ can be obtained from the standard deviations of the random variables $\sigma_{\xi_1}, \sigma_{\xi_2}, \sigma_{\xi_N}$ as

$$\sigma_\eta = \sqrt{\left(\frac{\partial f}{\partial \xi_1}\sigma_{\xi_1}\right)^2 + \left(\frac{\partial f}{\partial \xi_2}\sigma_{\xi_2}\right)^2 + \ldots + \left(\frac{\partial f}{\partial \xi_N}\sigma_{\xi_N}\right)^2}. \tag{A1}$$

In case of matrix inversion, since each coefficient $a_{ij}^{-1}$ of the inverse matrix $[A]^{-1}$ is a function of the coefficients of $[A]$, its standard deviation is obtained from the standard deviations of the coefficients $a_{ij}$ and from the partial derivatives:

$$\frac{\partial a_{11}^{-1}}{\partial a_{11}} = \frac{-a_{22}^2}{(a_{11}a_{22}-a_{12}a_{21})^2} \qquad \frac{\partial a_{11}^{-1}}{\partial a_{12}} = \frac{a_{22}a_{21}}{(a_{11}a_{22}-a_{12}a_{21})^2}$$
$$\frac{\partial a_{11}^{-1}}{\partial a_{21}} = \frac{a_{22}a_{12}}{(a_{11}a_{22}-a_{12}a_{21})^2} \qquad \frac{\partial a_{11}^{-1}}{\partial a_{22}} = \frac{-a_{12}a_{21}}{(a_{11}a_{22}-a_{12}a_{21})^2} \tag{A2}$$

$$\frac{\partial a_{12}^{-1}}{\partial a_{11}} = \frac{a_{12}a_{22}}{(a_{11}a_{22}-a_{12}a_{21})^2} \qquad \frac{\partial a_{12}^{-1}}{\partial a_{12}} = \frac{-a_{11}a_{22}}{(a_{11}a_{22}-a_{12}a_{21})^2}$$
$$\frac{\partial a_{12}^{-1}}{\partial a_{21}} = \frac{-a_{12}^2}{(a_{11}a_{22}-a_{12}a_{21})^2} \qquad \frac{\partial a_{12}^{-1}}{\partial a_{22}} = \frac{a_{11}a_{12}}{(a_{11}a_{22}-a_{12}a_{21})^2} \tag{A3}$$

$$\frac{\partial a_{21}^{-1}}{\partial a_{11}} = \frac{a_{21}a_{22}}{(a_{11}a_{22}-a_{12}a_{21})^2} \qquad \frac{\partial a_{21}^{-1}}{\partial a_{12}} = \frac{-a_{21}^2}{(a_{11}a_{22}-a_{12}a_{21})^2}$$
$$\frac{\partial a_{21}^{-1}}{\partial a_{21}} = \frac{-a_{11}a_{22}}{(a_{11}a_{22}-a_{12}a_{21})^2} \qquad \frac{\partial a_{21}^{-1}}{\partial a_{22}} = \frac{a_{11}a_{21}}{(a_{11}a_{22}-a_{12}a_{21})^2} \tag{A4}$$

$$\frac{\partial a_{22}^{-1}}{\partial a_{11}} = \frac{-a_{12}a_{21}}{(a_{11}a_{22}-a_{12}a_{21})^2} \qquad \frac{\partial a_{22}^{-1}}{\partial a_{12}} = \frac{a_{11}a_{21}}{(a_{11}a_{22}-a_{12}a_{21})^2}$$
$$\frac{\partial a_{22}^{-1}}{\partial a_{21}} = \frac{a_{11}a_{12}}{(a_{11}a_{22}-a_{12}a_{21})^2} \qquad \frac{\partial a_{22}^{-1}}{\partial a_{22}} = \frac{-a_{11}^2}{(a_{11}a_{22}-a_{12}a_{21})^2} \tag{A5}$$

## References


[1] Srinivasan, A. V., and Cutts, D. G., 1984, "Measurement of Relative Vibratory Motion at the Shroud Interfaces of a Fan," ASME J. Vibr. Acoust., **106**, pp. 189–197.

[2] Griffin, J. H., 1980, "Friction Damping of Resonant Stresses in Gas Turbine Engine Airfoils," J. Eng. Power, **102**, pp. 329–333.

[3] Srinivasan, A. V., and Cutts, D. G., 1983, "Dry Friction Damping Mechanisms in Engine Blades," J. Eng. Power, **105**, pp. 332–341.

[4] Berthillier, M., Dupont, C., Mondal, R., and Barrau, J. J., 1998, "Bladed Force Response—Analysis with Friction Dampers," ASME J. Vibr. Acoust., **120**, pp. 468–474.

[5] Yang, B. D., and Menq, C. H., 1998, "Characterization of Contact Kinematics and Application to the Design of Wedge Dampers in Turbomachinery Blading: Part 2—Prediction of Forced Response and Experimental Verification," ASME J. Eng. Gas Turbines Power, **120**, pp. 418–423.

[6] Sanliturk, K. Y., Ewins, D. J., and Stanbridge, A. B., 1999, "Underplatform Dampers for Turbine Blades: Theoretical Modelling, Analysis and Comparison of Experimental Data," 44th ASME Gas Turbine and Aeroengine Congress and Exhibition, June 7–10, Indianapolis IN, ASME Paper No. 99-GT-335.

[7] Csaba, G., 1999, "Modelling of a Microslip Friction Damper Subjected to Translation and Rotation," 44th ASME Gas Turbine and Aeroengine Congress and Exhibition, June 7–10, Indianapolis, IN, ASME Paper No. 99-GT-149.

[8] Berruti, T., Filippi, S., Goglio, L., and Gola, M. M., "Forced Vibrations of Frictionally Damped Stator Segments," 44th ASME Gas Turbine and Aeroengine Congress and Exhibition, June 7–10, Indianapolis, IN, ASME Paper No. 99-GT-412.

**J. Hoffren**
Research Scientist,
Laboratory of Aerodynamics,
Helsinki University of Technology,
P.O. Box 4400,
FIN-02105 HUT, Finland

**T. Talonpoika**
Senior Assistant

**J. Larjola**
Professor

Department of Energy Technology,
Lappeenranta University of Technology,
P.O. Box 20,
FIN-53850 Lappeenranta, Finland

**T. Siikonen**
Professor,
Laboratory of Applied Thermodynamics,
Helsinki University of Technology,
P.O. Box 4400,
FIN-02105 HUT, Finland

# Numerical Simulation of Real-Gas Flow in a Supersonic Turbine Nozzle Ring

*In small Rankine cycle power plants, it is advantageous to use organic media as the working fluid. A low-cost single-stage turbine design together with the high molecular weight of the fluid leads to high Mach numbers in the turbine. Turbine efficiency can be improved significantly by using an iterative design procedure based on an accurate CFD simulation of the flow. For this purpose, an existing Navier-Stokes solver is tailored for real gas, because the expansion of an organic fluid cannot be described with ideal gas equations. The proposed simulation method is applied for the calculation of supersonic flow in a turbine stator. The main contribution of the paper is to demonstrate how a typical ideal-gas CFD code can be adapted for real gases in a very general, fast, and robust manner.* [DOI: 10.1115/1.1423320]

## Introduction

In recent years, the study of the boiler plant process has dealt much with Rankine processes using some other fluid than water and steam. Organic chemicals, like toluene and ammonia, have been the most important new fluids ([1]). However, such fluids exhibit significant real-gas effects in operation, which must be explicitly addresses in the related design of turbomachinery. Modern design methods involve detailed CFD calculations, which are inherently computationally intensive, even for ideal fluids. Thus, the modeling of the real-gas effects in the CFD calculations should not significantly increase the required computing effort. On the other hand the absolute accuracy of the real-gas modeling is not necessary, but the overall behavior and correct trends should be correctly predicted. The application of a CFD code developed along these ideas to the design of a small single-stage turbine stator working on toluene is the overall theme of this paper.

The viscous flow solver applied in this study has been developed at Helsinki University of Technology over the past decade. It has been utilized successfully in numerous different applications with complex geometries but with ideal gas as the flowing medium ([2,3]). For the present turbine simulations, the code was modified for arbitrary single-phase real gases. In principle, the thermodynamic properties of the fluid can be evaluated by any mathematical model called for at each computational cell at each iteration, but generality and computational efficiency are desired. To fulfill these requirements, a straightforward table interpolation scheme was adopted, as will be described.

In the following, a turbine design is briefly described. This is followed by a discussion of the evaluation of the required thermodynamic data for the present application. As the main contribution of wider interest, the flow solver and the implementation of a general real-gas model in it are described next to facilitate similar extensions to other CFD codes. The final section of the paper contains an example of a turbine stator calculation, where the extreme flow conditions demonstrate the real-gas effects.

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

## Turbine Design

The starting point of the turbine design considered here is the specified expansion of the fluid. Pressures and temperatures in the Rankine process are optimized by using a complex computation program ([4]). The optimum values depend on the heat source temperature, on the condensing temperature and on the maximum permitted temperature of the working fluid. The values shown in Table 1 represent a typical case for toluene ([5,6]). We can see that the pressure ratio over the turbine is high, in excess of one hundred.

In Fig. 1, the radial turbine design of Larjola and Nuutila is shown as a surface grid presentation. The design permits some freedom in the selection of the turbine speed. However, because of the high pressure ratio, the specific speed $N_s$ defined by

$$N_s = \frac{\omega (q_m v_2)^{1/2}}{\Delta h_s^{0.75}} \tag{1}$$

**Table 1 (*a*) Turbine Data of the Design Example**

| Working Fluid | Toluene $C_6H_5CH_3$ |
|---|---|
| Turbine power | 285.5 kW |
| Turbine efficiency | 75 percent |
| Rotational speed | 400 l/s |
| Specific speed $N_s$ | 0.46 |
| Mass flow rate | 2.03 kg/s |
| Volume flow rate at turbine exit | 2.74 $m^3/s$ |

**(*b*) Expansion Process in the Turbine**

| Node | T (K) | p (kPa) | h (kJ/kg) | v ($m^3/s$) | s (J/kgK) |
|---|---|---|---|---|---|
| 0 | 583.15 | 3400 | 978.5 | | 4.125 |
| 2s | 438.05 | 30.8 | 790.7 | | 4.125 |
| 2 | 465.45 | 30.8 | 837.6 | 1.3523 | 4.229 |

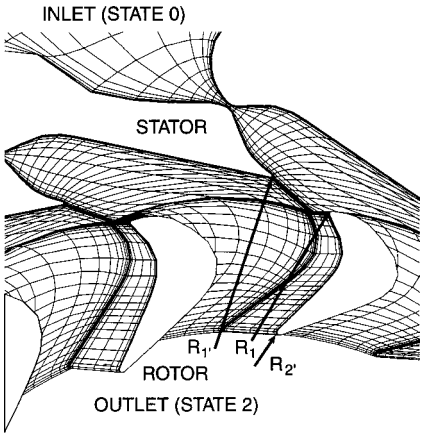



**Fig. 1 Radial turbine design considered, shown by a coarse grid**

should be relatively low, because otherwise the relative height of the blades at turbine exit would be too high. Here $\omega$ is the angular velocity, $q_m$ the mass flow, $v_2$ the outlet specific volume, and $\Delta h_s$ the isentropic enthalpy drop. In the design example of Table 1 the specific speed is $N_s=0.46$. At this $N_s$ value, a turbine efficiency $\eta_s$ of 0.86 to 0.88 could be reached ([7]). However, because of the high pressure ratio and high Mach numbers, the obtainable efficiency is normally significantly lower. As the goal in this study, $\eta_s=0.75$ to 0.80 was selected.

The selection of a suitable turbine type is based on the works of Balje [7], Glassman [8], Horlock [9], and Verneau [10]. The main idea is that the losses in the stator nozzle ring are reduced by decreasing the nozzle exit Mach number. On the other hand, the inlet velocity relative to the rotor cascade $w_1$ should be subsonic, because a supersonic inlet velocity increases losses, particularly during an off-design operation. A good compromise is to select the intermediate design pressure $p_1$ before the rotor so that $w_1$ is just subsonic. For the degree of reaction $r$, this typically gives values between 0 and 0.5.

If an axial turbine were used, there would be two problems: (1) If the number of nozzles in the stator is large, the width of the nozzle throat is very small, causing high manufacturing costs ([10]). On the other hand, if the number of nozzles is reduced in order to obtain acceptable manufacturing costs, the flow is supersonic in the nozzle ring circle, which causes losses. (2) Because $r>0$, the axial flow direction causes a significant axial force, which increases thrust bearing losses and expenses. Therefore, the turbine type selected was of a radial design. The stator nozzle ring design is two-dimensional, which reduces manufacturing expenses, and an essentially two-dimensional supersonic mixing zone after the nozzles makes the design easier, improving efficiency. Despite $r>0$, the axial force is very low, because the pressure forces are compensated for. However, the turbine design is not of a conventional radial type, because the standard radial design ([8]) would cause a very low blade height at the rotor inlet, thus increasing losses. Also, in the standard radial turbine, the rotor weight would be significant, causing problems in keeping the common generator-turbine rotor subcritical for bending oscillations.

The starting point of the design consists of the inlet pressure $p_o$, the outlet pressure $p_2$ and the enthalpy drop $\Delta h$ over the turbine (see Fig. 1 and Table 1(*b*)). The definition of $\Delta h$ is, however, an iterative process, because its calculation requires knowledge of the turbine efficiency. Also, the calculation of the intermediate pressure $p_1$ requires knowledge of the stator nozzle ring efficiency $\eta_{ST}$. In the design method, these efficiencies are first estimated and used as input for a one-dimensional design program, which gives the velocity vectors and the main geometry.

**Table 2 Main Gas and Cascade Values in the Turbine**

| Ratio of Stator Outlet and Rotor Inlet Radius | $R_{1'}/R_1$ | 1.05 |
|---|---|---|
| Ratio of blade heights | $b_1/b_{1'}$ | 1.15 |
| Degree of reaction | $r$ | 0.25 |
| Flow coefficient | $\nu$ | 0.479 |
| Rotor outer diameter | $2R_1$ | 246.1 mm |
| Height of stator blade | $b_1$ | 11.2 mm |
| Number of stator nozzles | | 20 |
| Breadth of a stator throat | | 0.89 mm |
| Mach number at stator exit | | 2.27 |
| Mach number at stator blade passage exit | | 2.16 |
| Mach number at rotor inlet | | 0.92 |
| Mach number at rotor blade passage exit | | 1.38 |

This geometry is then the starting point for the final design that utilizes CFD. After optimizing the blade passage geometry on the basis of the CFD calculations, new efficiencies are obtained, which are then used as the one-dimensional program input. The basic one-dimensional equations used are

$$r=1-\frac{h_0-h_1}{\Delta h+\frac{1}{2}c_{R_{2'}}^2} \tag{2}$$

where $\Delta h=h_1-h_2$ is the enthalpy drop in the turbine and

$$c_1=[2(h_0-h_1)]^{1/2} \tag{3}$$

$$c_1^2=c_{u_1}^2+c_{R_1}^2=\frac{\Delta h^2}{u_1^2}+\nu^2u_1^2. \tag{4}$$

Here $c_{u_1}=\Delta h/u_1$, $c_{R_1}=\nu u_1$, $\nu$ is the flow coefficient, $c$ the absolute velocity, $c_R$ the radial velocity, and $u$ the tangential velocity. From Eqs. (2) to (4) we obtain

$$u_1^2=\frac{\Delta h(1-r)-\Delta h[(1-r)^2-(k_2^2\nu^2r-(k_2^2-1)\nu^2)]^{1/2}}{k_2^2\nu^2r-(k_2^2-1)\nu^2} \tag{5}$$

where $k_2=c_{R_{2'}}/c_{R_1}$. From the velocity vectors we obtain

$$w_1=[c_{R_{1'}}^2+(c_{u_1}-u_1)^2]^{1/2} \tag{6}$$

where $c_{R_{1'}}=(b_1/b_{1'})c_{R_1}$ and $b$ is the blade height. Since $h_1=h_0-\frac{1}{2}c_1^2$ and by utilizing the defined stator isentropic efficiency

$$\eta_{ST}=\frac{h_0-h_1}{h_0-h_{1s}} \tag{7}$$

the intermediate pressure $p_1$ corresponding to $h_1$ and $h_{1s}$ can be solved. In Table 2, the design values used in this example are shown. The degree of reaction $r$ is selected to give a slightly subsonic $w_1$.

## The Principle of Modeling Real Gases

In order to study different flow media in the turbine design, a capability to create suitable continuous thermodynamic models from sparse data tables must be available. The modeling tool described here is also applied to create suitable input arrays for the Navier-Stokes solver, although the code itself is not in any way dependent on this specific tool. The principle of the formulation of the fluid properties is based on well-known thermodynamic laws. The functions based on these laws are supplemented with terms that fit the properties to the specific values of each medium, published generally in tables. The results are obtained as analytical functions of two variables, which are typically pressure and temperature ([11]). With this general method, changing the fluid and comparing different fluids is straightforward. In the following, the formulas applied for the evaluation of saturated pressure, specific enthalpy and specific volume are given.

**Saturated Pressure and Temperature.** Saturated pressure is calculated using the Antoine equation ([12])

$$p_s(T) = p_0 e^{A-B/(T-C)}. \tag{8}$$

The constants *A*, *B*, and *C* are identified using three known points of temperature and saturated pressure. An inverse function of the Antoine equation is applied without iteration for the saturated temperature

$$T_s(p) = C + \frac{B}{A - \ln(p/p_0)}. \tag{9}$$

**Specific Enthalpy.** The specific enthalpy of superheated vapor $h_g(T,p)$ is calculated from the enthalpy of saturated vapor $h_{gs}(T)$ and the enthalpy of ideal gas vapor $h_{Id}(T)$ (see Fig. 2).

$$h_g(T,p) = h_{Id}(T) - [h_{Id}(T_s(p)) - h_{gs}(T_s(p))]\left(\frac{T_s(p) - T_{tr}}{T - T_{tr}}\right)^{a_{gh}} \tag{10}$$

where $T_s(p)$ is the saturated temperature that corresponds to the pressure $p$ and $T_{tr}$ is the triple point temperature. A suitable value for the exponent for several fluids is $a_{gh} = 2.2$ ([11]). If one superheated vapor point is known, it can be used to calculate a more accurate value of the exponent for this medium.

The specific enthalpy of saturated vapor $h_{gs}(T)$ is the sum of the enthalpy of saturated liquid and the latent heat of vaporization

$$h_{gs}(T) = h_l(T) + h_{fg}(T). \tag{11}$$

The specific enthalpy of saturated liquid $h_l(T)$ as well as the specific enthalpy of vapor in the ideal gas state $h_{Id}(T)$ are calculated using the known data points $(T_{li}, h_{li})$ and $(T_{Idi}, h_{Idi})$ either by the Lagrange interpolation or by making a polynomial fit.

The latent heat of vaporation is calculated using a known data point $(T_{fg1}, h_{fg1})$ as ([13])

$$h_{fg}(T) = h_{fg1}\left(\frac{T_{cr} - T}{T_{cr} - T_{fg1}}\right)^{a_{fg}}. \tag{12}$$

A good value for the exponent for several fluids ([13]) is $a_{fg} = 0.38$. A more precise value for the exponent $a_{fg}$ can be achieved when a second point of latent heat is known.

**Specific Volume.** The specific volume of superheated vapor is calculated using the real gas equation of state as

$$v_g(T,p) = Z_g(T,p)\frac{RT}{Mp} \tag{13}$$

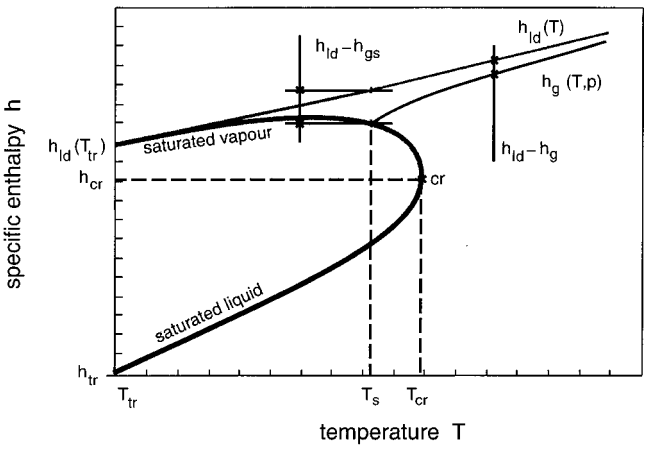



**Fig. 2 Relations for the calculation of the specific enthalpy of the superheated vapor $h_g$ using the enthalpies of the saturated vapor $h_{gs}$ and the ideal gas vapor $h_{Id}$**

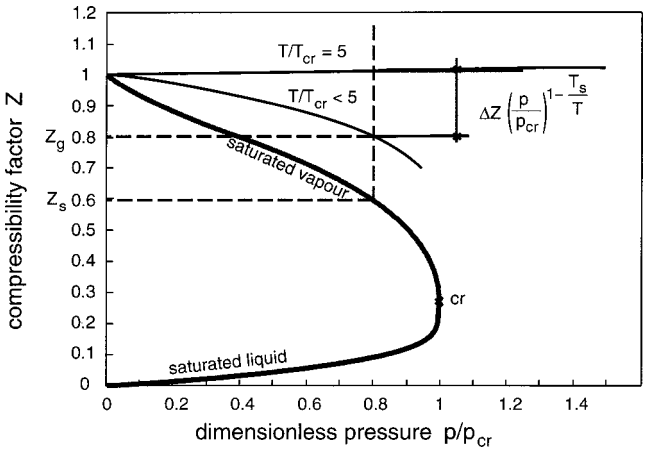



**Fig. 3 The compressibility factor $Z_g$ is obtained by modifying a linear approximation for compressibility at temperature $T/T_{cr}=5$ with a correction $\Delta Z$ depending on the compressibility factor of the saturated vapor $Z_s$**

where the compressibility factor (see Fig. 3) is

$$Z_g(T,p) = 1 + 0.015\frac{p}{p_{cr}} - \Delta Z(T,p)\left(\frac{p}{p_{cr}}\right)^{1-T_s(p)/T}. \tag{14}$$

The sum of the first and the second terms in the equation of the compressibility factor equals with a good accuracy the compressibility factor when the dimensionless temperature is $T/T_{cr}=5$, and the third term takes account of the deviation from this value. The coefficient of the third term is

$$\Delta Z(T,p) = \left[1 + 0.015\frac{p}{p_{cr}} - Z_s(T_s(p))\right]\left(\frac{5 - T/T_{cr}}{5 - T_s(p)/T_{cr}}\right)^{a_{gv}} \tag{15}$$

where the compressibility factor of the saturated vapor is

$$Z_s(T_s(p)) = \frac{Mpv_{gs}(T_s(p))}{RT_s(p)}. \tag{16}$$

If a superheated vapor point is not known, a value of $a_{gv}=10$ is used. If one point is known, the exponent $a_{gv}$ is calculated utilizing this known value.

The specific volume of saturated vapor is calculated using the specific volume of saturated liquid and the derivative of the saturated pressure with the Clausius equation ([12]) as

$$v_{gs}(T) = v_l(T) + \frac{h_{fg}(T)}{T}\frac{dT}{dp}. \tag{17}$$

The derivative of the saturated pressure can be obtained from the Antoine equation

$$1\Big/\frac{dT}{dp} = \frac{dp}{dT} = \frac{B}{(T-C)^2}p(T). \tag{18}$$

**Comparison With the Values in Tables.** Compared with values in tables for toluene ([14]), the deviation is typically less than 0.2 percent for the specific enthalpy, and 0.5 percent for the specific volume. The largest deviations are 1.5 percent close to the critical point. As the numerical values of enthalpy depend on the chosen reference state, the precision of the calculation was also tested using an inverse function, which compares the value in the table with the calculated temperature when the pressure and the value in the table of some property are known. The temperature deviation varied a lot and was found to be typically 1 to 10 K for the specific enthalpy, and 1 to 15 K for the specific volume. The precision of the functions is considered good enough for rough process modeling.

For the numerical flow simulations of the present turbine, the results of the model just described are converted into numerical data arrays defined using temperature and density, which are the common basic parameters in CFD work with compressible fluids. However, it is reminded that instead of the data modeling just described, the flow solver could apply tabular property data of a suitable basic format available from any source.

## Flow Simulation Method

**Governing Equations.** The basic equations employed in the numerical flow simulations are the Navier-Stokes equations governing a viscous compressible flow. Using Cartesian coordinates, they can be cast in a conservative form as

$$\frac{\partial U}{\partial t}+\frac{\partial F(U)}{\partial x}+\frac{\partial G(U)}{\partial y}+\frac{\partial H(U)}{\partial z}=0 \tag{19}$$

where $U=(\rho,\rho u,\rho v,\rho w,\rho E)^T$ is the vector of conservative variables consisting of density $\rho$, momentum components $\rho u,\rho v,\rho w$, and total energy per volume $\rho E$. In the flux vectors $F(U)$, $G(U)$, and $H(U)$ there are additional quantities, such as the velocity $\mathbf{V}=u\mathbf{i}+v\mathbf{j}+w\mathbf{k}$, pressure $p$, viscous stresses $\tau_{ij}$ and heat fluxes $q_i$. The viscous stresses and heat fluxes also contain viscosity coefficients $\lambda$ and $\mu$, thermal conductivity $k$ and temperature $T$. To reduce the number of variables, the assumption $\lambda=-2/3\mu$ for the viscosities is adopted, and in the present implementation, the heat conductivity is approximated via the viscosity and a constant Prandtl number. In this study, the effects of turbulence are taken into account by the Boussinesq approximation and the simple algebraic Cebeci-Smith model. More sophisticated heat transfer and turbulence models could naturally be applied if deemed necessary for each application, without affecting the main theme of the paper, the real-gas effects.

The total energy $E$ is defined as

$$E=e+V^2/2 \tag{20}$$

where the specific internal energy $e$, in turn, depends on the equation of state via $e=e(T,\rho)$. Similarly, the pressure needed for the fluxes is obtained from a function $p=p(T,\rho)$, and a corresponding expression $\mu=\mu(T,\rho)$ is applied for the viscosity.

**Finite Volume Form.** In the finite volume formulation applied, the flow equations are written in an integral form

$$\frac{d}{dt}\int_V U dV+\int_S \mathbf{F}(U)\cdot d\mathbf{S}=0 \tag{21}$$

for an arbitrary region $V$ with a boundary $S$. Performing the integrations for a grid cell $i$ yields

$$V_i\frac{dU_i}{dt}=\sum_{\text{faces}} -S\hat{F}=\mathcal{R}_i \tag{22}$$

where the sum is taken over the faces of the computational cell. A shortened form $\mathcal{R}_i$ for the residual is adopted to facilitate the following discussion. The flux on each cell face is of the form

$$\hat{F}=n_xF+n_yG+n_zH. \tag{23}$$

Here $F$, $G$, and $H$ are the Cartesian flux vectors, and $n_x$, $n_y$, $n_z$ are the components of the unit normal vector of the cell face in the $x$, $y$ and $z$-directions, respectively.

**Flux Calculation.** The calculation of fluxes is divided into two distinct parts. Details of the implementation can be found elsewhere ([15]) and only the essential features related to the adaptation of the code for real gases are described here. For the viscous and heat fluxes, the thin-layer approximation with conventional central differences is utilized. The calculation does not depend directly on the equation of state, but the actual implementation differs from a typical ideal-gas code in the evaluation of temperature gradient, viscosity, and heat conductivity.

The inviscid part, i.e., the convective and pressure terms of the fluxes, is evaluated applying the upwind principle ([16]). The calculation is formally second or third-order accurate. Roe's flux at each cell face is evaluated by computing an average flux from the states $U^r$ and $U^l$ on the different sides of the face and adding corrective flux difference terms to the result, as follows:

$$\hat{F}=\frac{1}{2}[\hat{F}(U^l)+\hat{F}(U^r)]-\frac{1}{2}(\Delta\hat{F}^+-\Delta\hat{F}^-). \tag{24}$$

The flux difference terms $\Delta\hat{F}^{\pm}$ can be written in turn as the products of the positive and negative parts of the Jacobian matrix $\partial\hat{F}/\partial U$ and the difference $U^r-U^l$. If the Jacobian matrix is further transformed to utilize the characteristic variables, the flux differences can be formulated as

$$\Delta\hat{F}^{\pm}=R\Lambda^{\pm}R^{-1}(U^r-U^l). \tag{25}$$

Here $R$ is a matrix containing suitably selected right eigenvectors of $\partial\hat{F}/\partial U$ and $R^{-1}$ is its inverse, whereas $\Lambda^+$ and $\Lambda^-$ are diagonal matrices consisting of the positive and negative eigenvalues, or the characteristic speeds $\lambda$.

The flux components $\hat{F}(U^l)$ and $\hat{F}(U^r)$ can be computed in a straightforward manner by applying the basic flux formulas, but the calculation of the flux difference requires an evaluation of additional thermodynamic quantities. Because the selection of the eigenvectors and characteristic variables is not unique, different formulations can be utilized. The original flow solver into which the real gas relations were implemented in this study utilizes the derivatives $\partial p/\partial\rho$ and $\partial p/\partial e$ as well as internal energy $e$ itself in order to evaluate $R$ and $\Lambda$.

In ideal gas calculations, the internal energy and the derivatives of pressure are available from $U^r$, $U^l$ using an analytical equation of state and the following relationships:

$$\partial p/\partial p=e(\gamma-1) \tag{26a}$$

$$\partial p/\partial e=\rho(\gamma-1). \tag{26b}$$

Simple averaging gives sufficient accuracy for the cell-face values. The local sonic speed is then available through the expression

$$c=\sqrt{\frac{\partial p}{\partial e}\frac{p}{\rho^2}+\frac{\partial p}{\partial\rho}} \tag{27}$$

if only simple turbulence models are utilized. The kinetic energy of turbulence modifies Eq. (27) ([15]) but the details are irrelevant here. The real gas computations require that the partial derivatives $\partial p/\partial\rho$ and $\partial p/\partial e$ are available for the flux calculation. The values of the derivatives for each cell are evaluated separately before the flux calculation and stored.

**Solution Method.** The solution method for the present steady-state calculations is based on an implicit time integration. The linearization of the residual $\mathcal{R}_i$ leads to an equation of the type

$$\left(1-\frac{\Delta\tau_i}{V_i}\left[\frac{\partial\mathcal{R}}{\partial U}\right]_i\right)\delta U_i=\frac{\Delta\tau}{V_i}\mathcal{R}_i \tag{28}$$

for the corrections $\delta U_i$ of the conservative variables at each local pseudo-time step $\Delta\tau_i$. For the solution of Eq. (28), an approximate factorization is performed. This results in a set of tridiagonal equations. To improve the computational efficiency, the tridiagonal equations are replaced by bidiagonal sweeps. The resulting LU-factored scheme is described in detail elsewhere ([2,15]), and here only the features relevant to the equation of state are discussed.

As in the flux calculation, the sonic speed $c$ and the derivatives $\partial p/\partial\rho$ and $\partial p/\partial e$ are needed in the evaluation of split Jacobians $\partial\mathcal{R}/\partial U$, and they are available from the precalculated stored arrays. In typical ideal gas calculations, the implicit stage gives directly the conservative corrections $\delta\rho$, $\delta\rho u$, $\delta\rho v$, $\delta\rho w$, and

$\delta\rho E$. With real gas, however, the total energy correction is replaced by $\delta e$. After the implicit stage, the flow variables are updated as

$$U_i^{n+1} = U_i^n + \delta U_i . \tag{29}$$

With ideal gas, this can be done simply by adding the obtained corrections to the conservative variables. In the real gas calculations, the density and the momentum components are updated similarly, but instead of the total energy, we are using temperature as a primary variable at this stage. The temperature at the new time level $T^{n+1}$ is calculated using the linearization

$$T^{n+1} = T^n + \frac{\delta e}{c_v^n} \tag{30}$$

where $c_v^n = (\partial e/\partial T)^n$ is the specific heat evaluated at the previous time level. This is computed and stored as the other thermodynamic properties referred to earlier.

To speed up the convergence, a multigrid algorithm based on the work of Jameson and Yoon [17] is utilized. As the multigrid cycle uses the same flux calculation and solution subroutines as the fine-grid calculations, the total energy must be available at the entry onto each grid level. To fulfill this requirement, an intermediate updated total energy is computed at the end of the implicit stage discussed above. This is done by updating the total energy from

$$\rho E^* = \rho^{n+1}(e^n + \delta e) + \frac{(\rho^{n+1}V^{n+1})^2}{2\rho^{n+1}} \tag{31}$$

before proceeding to the next coarser grid level. Furthermore, the temperature is included in the quantities transferred onto the coarse grid levels. It is to be noted that the updating of the total energy via $e(T,\rho)$ must not be performed within the multigrid system. If this is done after the implicit stage instead of the intermediate updating of Eq. (31), the calculations invariably fail. The reason is related to the break-up in the syncronization of the treatment of the internal energy that is evaluated from the conservative variables in the flux calculation.

The multigrid scheme gives corrections to the flow variables that are to be interpolated onto the finer grids. In an ideal gas solution, the corrections concern directly all the conservative variables. In a real gas solution, this is done for the density and the momentum components, but instead of total energy, the temperature is the primary variable also in the multigrid procedure. After the multigrid cycle, the final updated values for $\rho$, $\rho u$, $\rho v$, $\rho w$, and $T$ on the fine grid level are available. Finally, the total energy is updated on the densest grid level. This is done by evaluating the internal energy from the stored function $e(T,\rho)$ and assembling $\rho E$ from its components now available. Everything is then set for the beginning of a new iteration step.

## Thermodynamics in the Flow Solver

**Initialization of the Interpolation Functions.** The flow solution methods described above require the determination of pressure $p$ and specific internal energy $e$. In addition, we apply the partial derivatives $\partial p/\partial\rho$, $\partial e/\partial T = c_v$ and $\partial p/\partial e = (\partial p/\partial T)/c_v$. These quantities must be defined as functions of temperature and density. In principle, the derivatives can be evaluated numerically from the $p$ and $e$-arrays, but they should be continuous to ensure smooth operation of the code. The reasons for selecting a straightforward and general interpolation of tabulated function values in the implementation with continuous derivatives with respect to $T$ and discontinuous derivatives with respect to $\rho$ are discussed below.

The numerical data arrays for each variable, obtainable for example from the model functions described earlier, may be irregular and cover only a portion of the envelope defined by the extreme temperatures and densities. Computationally, it is more efficient to form regular equal-interval arrays from the given data

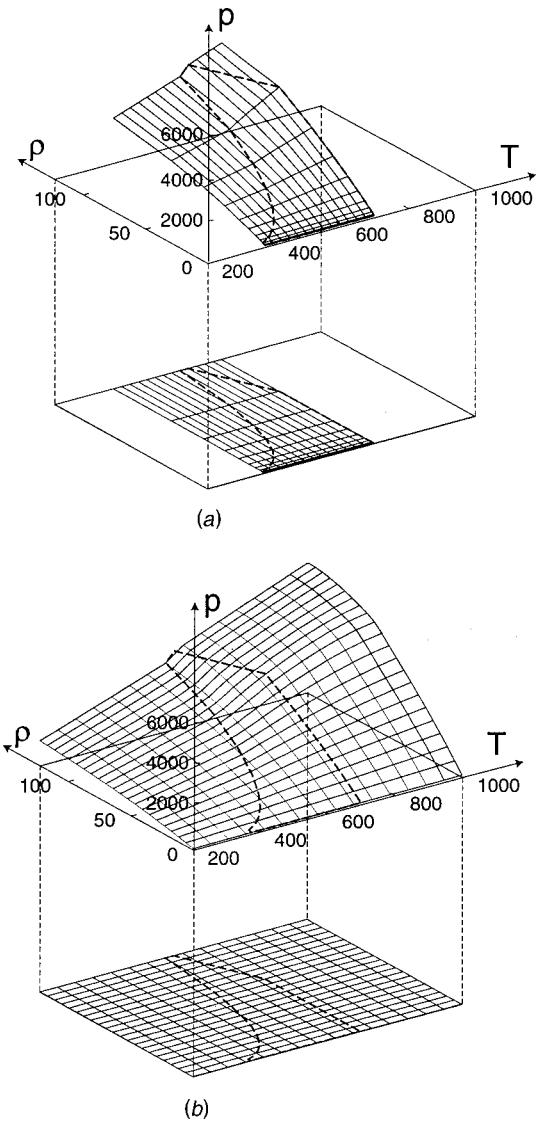



**Fig. 4 (*a*) The original definition array and (*b*) the formed regular computational array for pressure. The dashed line defines the envelope of the real physical data.**

at the initialization of each flow case and to apply these during the simulation, as illustrated in Fig. 4. Such a method with piecewise-defined simple analytical functions is utilized in this study.

Since it is difficult to form efficient two-dimensional interpolation functions with continuous derivatives, only the derivatives with respect to $T$ are evaluated numerically during the solution. For the derivative $\partial p/\partial\rho$, an extra array is formed at the initialization stage by applying pressure interpolants that are continuous with respect to density. Thus, three regular arrays for the thermodynamic variables $p$, $e$ and $\partial p/\partial\rho$ with continuous derivatives with respect to $T$ and discontinuous derivatives with respect to $\rho$ are formed, and the necessary interpolation parameters are stored. In addition to the thermodynamic properties, other quantities depending on the state can be defined similarly. In this work, the viscosity is defined as a numerical array $\mu(T,\rho)$.

The interpolation scheme adopted applies piecewise cubic Hermite polynomials with respect to temperature and linear interpolation with respect to density. The Hermite polynomials produce visually pleasing results with continuous first derivatives. The interpolated curves are less prone to oscillations caused by noisy data than cubic splines, and the scheme guarantees monotonous

approximations from monotonous data. The Hermite interpolation is based on the programs of ([18]), which are modified for the present application.

At the initialization of each flow case, potentially non-equal-interval data tables for $T$, $\rho$, $p(T,\rho)$, and $e(T,\rho)$ are read in. New regular, stored arrays for $T$ and $\rho$ are formed by dividing their ranges into suitable numbers of equal intervals. At present, the number of $T$-points is taken as the same as in the input tables, and the number of $\rho$-points is doubled to compensate for the simple interpolation in this direction. From the input $p$ and $e$-tables, equal-interval arrays are then formed by applying the Hermite interpolation in two directions at each regular $(T,\rho)$ point. The array for $\partial p/\partial\rho$ is obtained by applying the Hermite interpolation to the just-formed regular $p$-array with respect to density at each temperature. The interpolation process involves the determination of the derivatives, which are taken as the result of the operation. Finally, the Hermite interpolation is applied to the regular data arrays for $p$, $e$, and $\partial p/\partial\rho$ with respect to temperature at each density, and the parameters of the interpolation functions for each data interval are stored for later use.

**Evaluation During Iteration.** Because the piecewise interpolation functions are defined at equal intervals, very simple operations are needed during the actual calculations. At each cell with the given values for $T$ and $\rho$, the indices of the appropriate interpolation parameters are known by simple divisions without any checks. The function values and the derivatives with respect to temperature are obtained directly at the densities bordering the cell density, and the final values are obtained by fast linear interpolation.

In the actual implementation, the quantities $p$, $\mu$, $c_v$, $\partial p/\partial\rho$, $\partial p/\partial e$ and the resulting speed of sound $c$ are computed at the beginning of each iteration cycle. A separate evaluation is applied for $e$ at the end of the step, in which the calculation of the total energy is integrated. The treatment of the real gas properties in the way described is efficient. The code consumes just 6.7 percent more CPU time per cell and pseudo-time step than a corresponding ideal gas version.

Although the basic principle of the interpolation is straightforward, a word of caution is to be given. During the solution process, the density and temperature may wander far from their final expected ranges. The property definition arrays should cover such situations, too, to avoid risky extrapolations. In the present implementation, the properties are generally extrapolated linearly with respect to density and temperature if the definition range is exceeded. An exception is the derivative $\partial p/\partial\rho$, which is kept constant at excessive densities to match the treatment of pressure itself.

## Turbine Stator Simulations

**Toluene Model.** In the present simulation of the turbine stator of Fig. 1 operating with toluene vapor, the properties of the fluid in a temperature range of 375–700 K and density range of 0.46–123 kg/m$^3$ are required. In this range, interpreted as a rectangle, toluene is actually not always in a gas phase. At high densities and low temperatures it turns into a liquid, and no data related to this region were available. To be able to use the thermodynamic property evaluation method just described, the data arrays were filled manually with reasonable-looking pseudo-values to obtain a correct format. These pseudo-values have no effect on the final results, since the actual turbine always operates on the gas ghase. The input data tables contain 16 temperature points and 11 density points at nonequal intervals. Since the number of density values is doubled for the regular equal-interval arrays, 16 temperature values, and 22 density values are utilized during the solution.

The specific internal energy of toluene depends rather linearly on temperature and weakly on density. The pressure depends strongly and nonlinearly on both parameters, leading to variable slopes. Compared to air, a major difference is the low value of the specific heat ratio $\gamma$. For toluene, $\gamma$ varies at around 1.05–1.10, whereas for idealized air, $\gamma$ is a constant of 1.4. This difference should lead to large pressure and density differences in relation to temperature differences for toluene.

**Code Validation.** Before the actual turbine application, the real-gas flow solver was tested by inviscid computations around the venerable NACA 0012 airfoil ([19]). As the first stage, the properties of ideal gas were written as numerical data arrays in an external file. The real-gas code was run applying this gas data in a well-documented test case, and the convergence and results were compared with references obtained from the well-established ideal-gas version of the code. Both the convergence and results were practically identical for the two code versions, verifying the correct coding of the real-gas modifications.

As the second test stage, NACA 0012 calculations applying ideal gas and toluene were compared in three different flow cases. At a low flight Mach number, the differences between the results produced by air and toluene were small, as they should. At a transonic and supersonic Mach number, the differences were large and conformed well with expectations. Although the tests are not a rigorous validation of the overall approach because of the unavailability of any suitable reference results for toluene, they strongly support the realism of the turbine stator simulations.

**Stator Problem Setup.** The modeled geometry shown in Fig. 5 represents one channel between 17 stator vanes around the rotor. The flow comes in from the lower left and exits from the circular section at the upper right towards the rotor. The vane surfaces extend from the bulbous leading edges in the middle of the figure to the small step on the grid lower boundary and to the slope discontinuity on the upper boundary. In front of and behind the vane, the flow conditions at the grid boundaries are cyclic. The single-block grid contains 128 cells in the flow direction and 64 cells across it.

As a reference state, the well-defined throat conditions were selected. Specifically, the temperature was set to 551.95 K and the density to 51.653 kg/m$^3$. The flow is sonic, and the Reynolds number based on the throat width of 1.11765 mm is $8\times10^5$. The evaluation of the equation of state with the present method gives a

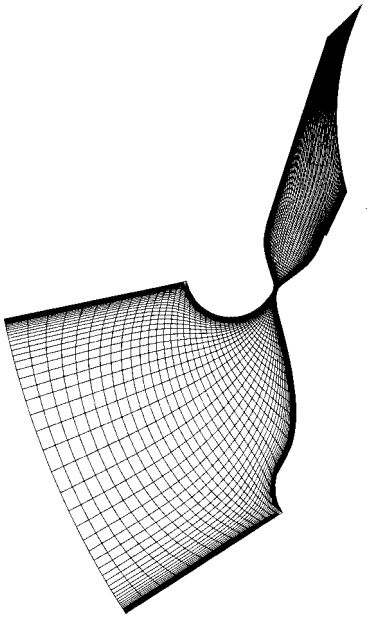


**Fig. 5 Grid for a passage between the turbine stator vanes. The flow is from lower left to upper right.**

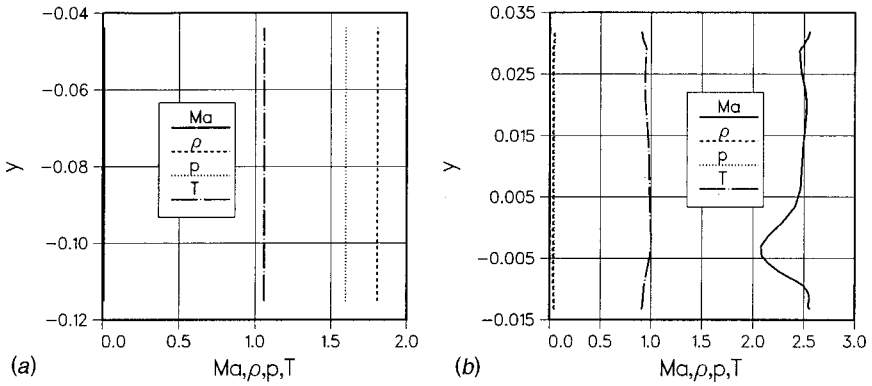



**Fig. 6 (*a*) Inflow and (*b*) outflow conditions of the turbine stator passage nondimensionalized by the throat values**

pressure of 1914.8 kPa, a specific internal energy of 922.89 kJ/kg, a specific enthalpy of 959.96 kJ/kg and a sonic speed of 171.33 m/s.

At the inflow boundary, the total enthalpy was fixed to a given one-dimensional design reservoir value of 1.02055 times the throat enthalpy. In addition, the mass flow determined by the throat was fixed, and the inflow was forced to be perpendicular to the grid boundary. The density was extrapolated from the interior, and its final value was iterated through successive corrections $\Delta\rho$ at each pseudo-time step. From the fixed momentum components and updated density, velocity corrections at each step could be calculated and an approximation to the enthalpy change $\Delta h$ obtained via a linearization $\Delta h \approx -u\Delta u - v\Delta v$. Subsequently, the temperature corrections at the inlet were calculated from the following linearization:

$$\Delta T = \frac{1}{c_v}\left[\Delta h - \frac{\Delta\rho}{\rho}\left(\frac{\partial p}{\partial\rho} - \frac{p}{\rho}\right)\right]. \tag{32}$$

With the updated temperature and density, new total energy was finally obtained.

At the outflow boundary, the flow in the steady state is supersonic and everything can be simply extrapolated from the interior. However, more delicate conditions are needed to start the calculation from slow flow. In the present calculations, the outlet pressure was gradually lowered from the extrapolated value towards a fixed low number. From the pressure changes, density corrections were obtained as $\Delta p/(\partial p/\partial\rho)$. The momentum components and temperature were extrapolated without tricks, with the exception of a lower limit for temperature. When the flow turns supersonic as the calculation proceeds, the boundary conditions set at the outflow automatically adopt their correct nature with the upwind scheme utilized.

In this type of computation, the initial conditions are important to the success of the runs. After some trial and error the density was set to 1.8 times the throat density, and the temperature was set to 1.056 times the throat temperature. The Mach number was 0.01 in a unidirectional flow roughly aligned to the channel. The conditions closely match the expected final conditions at the inflow boundary.

**Results.** With the boundary and initial conditions described above, the calculations with second-order fully upwinded differences and a flux limiter could be performed smoothly, applying three multigrid levels. The run was stopped after 10,000 steps. At this stage, the difference in the inflow and outflow mass flows was 0.19 percent, and the relative maximum density corrections were of the order of $10^{-4}$.

The final conditions at the inflow and outflow boundaries nondimensionalized by the throat values are shown in Fig. 6. The distributions at the inflow are flat, as expected. The Mach number is 0.010, and the temperature is 1.057 times the throat temperature, close to the values suggested by the one-dimensional design estimations. However, the pressure of 3062 kPa is somewhat lower than expected on the basis of the one-dimensional analysis, as is the density of 93.2 kg/m$^3$. The discrepancy is probably related to the estimates of the sonic speed at the throat, which in these simulations is some 13 percent lower than the value estimated in the simple analysis. At the outflow boundary, the distributions show some variation across the channel. The relative pressure and density are low, but still slightly higher than the values predicted by the simple analysis. The same applies to the relative temperatures. However, the Mach numbers of 2.08–2.55 conform well to the expectations.

The calculated results confirm that very large density and pressure changes exist in relation to modest temperature changes. The flowfield is illustrated by the contour plots downstream of the throat area in Fig. 7. Upstream, very little happens, which enables the plots to be truncated as shown without loss of information. The Mach contour plot is the most revealing. The increasing flow speed, the boundary layers and the wake of the vane are typical, but the apparent oblique shock emanating from the channel's lower surface is unexpected and undesired. The cause of the shock may be too steep a curvature of the wall towards the flow. Large gradients in the density and pressure plots are evident in the throat area, with little happening further downstream. The shock is clearly perceptible in the temperature plot, which otherwise displays mild variations.

On the basis of the results it can be concluded that the flow is first accelerated to a supersonic velocity quite smoothly in the nozzle, but when the jets of parallel nozzles are mixed, a deceleration takes place. Thus, the nozzle contour near the outlet should be reshaped (more space) and the computation repeated until the acceleration is smooth and continuous from the throat to the rotor inlet. It should be noted that the geometry is possible to be optimized only for one pressure ratio (design Mach number).

With a nonlinear real gas model, the code appears to be well behaved, provided that the property definition arrays cover the range required during the calculation. If no extrapolations are needed, the robustness corresponds to ideal gas calculations ([19]). Also the density residuals reach the same levels as with air. However, the situation becomes different if extrapolations are required in the equation of state. Small, temporary array limit exceedances of the order of ten percent of the density or temperature range may be easily coped with, but larger excursions are dangerous as the mismatch of the extrapolated properties grows.

## Conclusions

As a practical tool for research and development involving real gases, an extension of a Navier-Stokes solver for arbitrary equa-

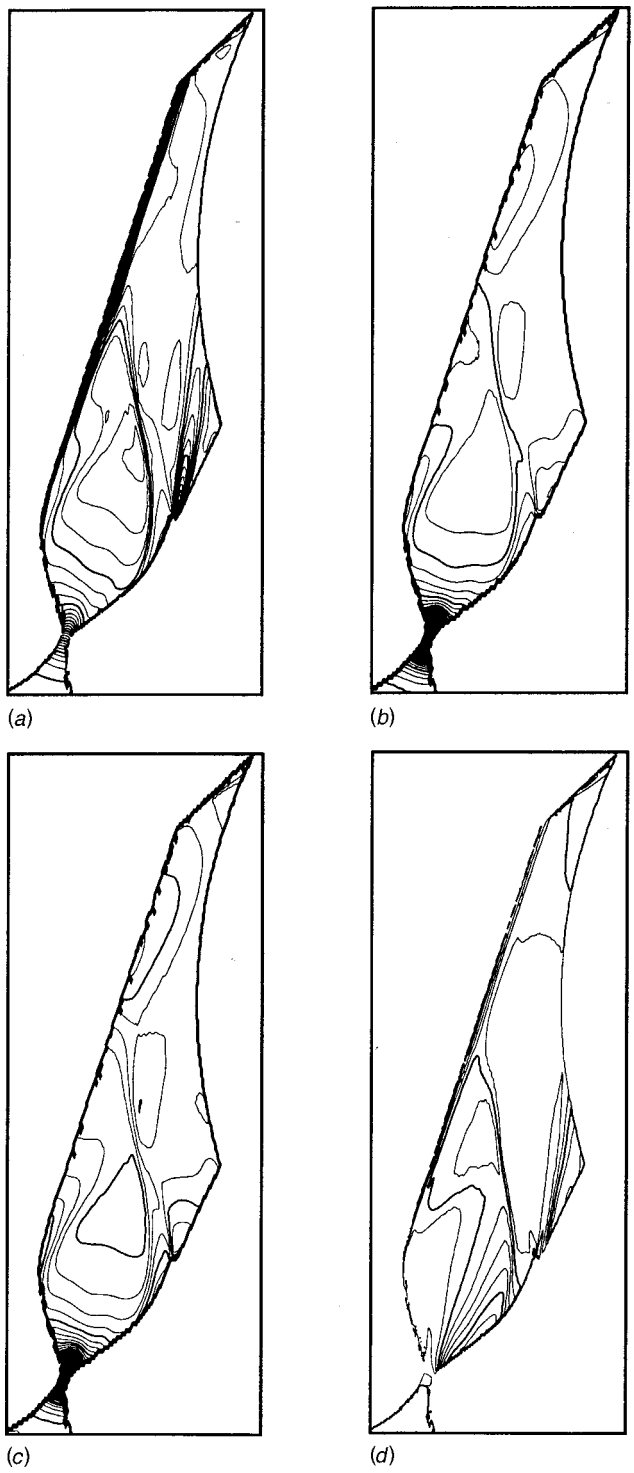



**Fig. 7 (*a*) Mach number (*b*) density, (*c*) pressure, and (*d*) temperature distributions in the turbine stator. The nondimensional contour interval is 0.1 in the Mach plot, 0.01 in the density and pressure plots, and 0.02 in the temperature plot.**

tions of state is presented. The method adopted utilizes externally defined numerical data arrays for the thermodynamic properties of the flow medium as functions of temperature and density. During the computation, two-dimensional equal-interval arrays formed at the beginning of each case are interpolated as necessary. The interpolation functions are piecewise cubic Hermite polynomials of temperature, which give continuous first derivatives in this direction. With respect to density, the interpolation is linear. Because the interpolation utilizes equal-interval arrays, direct divisions determine the indices of the function parameters to be applied at each computational cell at each solution step without any checks. The simplicity leads to a high computational efficiency of the method, and different media can be employed just by changing the data file. The method is also robust and easy to adapt to implementations employing other quantities than in the present work.

Test calculations with toluene vapor flowing in a radial turbine stator were performed. The code performed robustly, and the agreement of the results with simple one-dimensional design calculations is within expectations. Additional airfoil calculations performed with air and toluene further support the conclusion of a valid overall approach and a successful implementation. It is felt that the developed method can be used, e.g., to optimize the geometry of a supersonic stator nozzle ring. The geometry is changed according to the calculated velocity field, and the computation is repeated until a continuous, smooth acceleration takes place. Thus this combination of CFD analysis and one dimensional calculation may prove to be an efficient tool for the design of radial turbines with high pressure ratios.

## Nomenclature

SI units are applied, but any consistent unit system would do as well.

$A, B, C$ = constants in real gas model
$a$ = exponent in real gas model
$b$ = blade height
$c$ = speed of sound, absolute velocity
$c_v$ = specific heat at constant volume
$E$ = specific total energy
$e$ = specific internal energy
$F, G, H$ = flux vectors
$\hat{F}$ = flux vector in a curvilinear grid
$h$ = specific enthalpy, $e+p/\rho$
$\mathbf{i}, \mathbf{j}, \mathbf{k}$ = unit vectors in a Cartesian coordinate system
$k$ = heat conductivity
$k_2$ = constant
$M$ = molar mass
Ma = Mach number
$N_s$ = dimensionless specific speed
$\mathbf{n}$ = unit vector normal to a cell surface
$n_x, n_y, n_z$ = Cartesian normal unit vector components
$p$ = pressure
$q$ = heat flux
$q_m$ = mass flow
$R$ = gas constant, right eigenvector, radius
$\mathcal{R}$ = residual
$r$ = degree of reaction
$S$ = cell surface segment area
$s$ = specific entropy
$T$ = temperature
$t$ = time
$U$ = vector of conservative variables
$V$ = cell volume, also velocity
$v$ = specific volume
$u, v, w$ = Cartesian velocity components
$w$ = relative velocity
$x, y, z$ = spatial coordinates
$\gamma$ = specific heat ratio
$Z$ = compressibility factor
$\Delta U$ = change in conservative variables at boundaries
$\delta U$ = change in conservative variables after the implicit stage
$\delta\tau$ = pseudo-time step
$\Lambda$ = diagonal matrix containing the characteristic speeds
$\lambda, \mu$ = viscosity coefficients, $\lambda$ also eigenvalue (characteristic speed)
$\nu$ = flow coefficient

$\rho$ = density
$\omega$ = angular velocity
$\tau_{ij}$ = viscous stresses

**Subscripts**

$cr$ = critical
$g$ = gas, superheated vapor
$h$ = specific enthalpy
$Id$ = ideal gas, vapor in the real-gas state
$i, j, k$ = spatial indices
$l$ = saturated liquid
$fg$ = latent heat
$p$ = pressure
$R$ = radial
$ST$ = stator
$s$ = saturated, also isentropic
$tr$ = triple point
$u$ = tangential
$v$ = specific volume
$0$ = stator inlet
$1'$ = stator blade passage outlet
$1$ = rotor inlet
$2'$ = rotor blade passage outlet
$2$ = turbine outlet

**Superscripts**

$l, r$ = states at the sides of a cell surface (left, right)
$\pm$ = positive and negative flux contributions
$n, n+1$ = time level indices
$*$ = intermediate value between the time levels

## References


[1] Hornnes, A., and Bolland, O., 1991, "Power Cycle Working Fluids," SINTEF Report STF15 A91041, Trondheim.
[2] Siikonen, T., and Pan, H., 1992, "Application of Roe's Method for the Simulation of Viscous Flow in Turbomachinery," *Proceedings of the First European Computational Fluid Dynamics Conference*, Ch. Hirsch et al., eds., Elsevier, New York, pp. 635–641.
[3] Pitkänen, H. P., 1997, "The CFD Analysis of the Impeller and Vaneless Diffuser of an Industrial Water-Treatment Compressor," ASME International Mechanical Engineering Congress & Exposition, Dallas, TX, Nov. 16–21.
[4] Honkatukia, J., 1996, private communication.
[5] Larjola, J., and Nuutila, M., 1995, "District Heating Plant Converted to Produce Also Electric Power," Paper 228 E, 27th Unichal Congress, Stockholm, June 12–14.
[6] Larjola, J., 1995, "Electricity from Industrial Waste Heat Using High-Speed Organic Rankine Cycle (ORC)," Int. J. Production Economics, **41**, pp. 227–235.
[7] Balje, O. E., 1981, *Turbomachines: A Guide to Design, Selection and Theory*, John Wiley and Sons, New York.
[8] Glassman, A. J., ed., 1975, "Turbine Design and Application," Vol. 3, NASA SP-290.
[9] Horlock, J. H., 1985, *Axial Flow Turbines*, R. E. Krieger, Malabar, FL.
[10] Verneau, A., 1987, "Supersonic Turbines for Organic Fluid Rankine Cycles from 3 to 1300 kW, Small High Pressure Ratio Turbines," von Karman Institute for Fluid Dynamics, Lecture Series 1987-07.
[11] Talonpoika, T., 1994, "Modelling the Properties of Fluid in a Thermodynamic Cycle" (Termodynaamisen kiertoprosessin väliaineen aineominaisuuksien mallitus, in Finnish), Lappeenranta University of Technology, Research Report EN B-82, Lappeenranta.
[12] Edmister, W. C., and Lee, B. I., 1984, *Applied Hydrocarbon Thermodynamics*, Vol. 1, 2nd Ed., Gulf Publ Comp., Houston, TX.
[13] Goodwin, R. D., 1989, "Toluene Thermophysical Properties From 178 to 800 K at Pressures to 1000 Bar," J. Phys. Chem. Ref. Data, **18**, No. 4, pp. 1565–1636.
[14] ESDU, 1974, "Thermodynamic Properties of Toluene," ESDU Engineering Sciences Data Item Number 74024, Engineering Sciences Data Unit Ltd., London.
[15] Siikonen, T., 1995, "An Application of Roe's Flux-Difference Splitting for $k-\epsilon$ Turbulence Model," Int. J. Numer. Methods Fluids, **21**, pp. 1017–1039.
[16] Roe, P. L., 1981, "Approximate Riemann Solvers, Parameter Vectors, and Difference Schemes," J. Comput. Phys., **43**, pp. 357–372.
[17] Jameson, A., and Yoon, S., 1986, "Multigrid Solution of the Euler Equations Using Implicit Schemes," AIAA J., **24**, No. 11, pp. 1737–1743.
[18] Kahaner, D., Moler, C., and Nash, S., 1988, *Numerical Methods and Software*, Prentice- Hall, Englewood Cliffs, NJ.
[19] Hoffren, J., 1997, "Adaptation of FINFLO for Real Gases," Helsinki University of Technology, Laboratory of Applied Thermodynamics, Report No. 102, Espoo.

# A Kinetic Investigation of the Role of Changes in the Composition of Natural Gas in Engine Applications

**E. B. Khalil**

**G. A. Karim**
e-mail: karim@enme.ucalgary.ca

Department of Mechanical Engineering,
The University of Calgary,
Calgary T2N 1N4, Canada

*The influence of variations in the composition of natural gas on the ignition and combustion processes in engines is investigated. Particular attention is given to changes in the relatively small concentrations of high molar mass alkanes that may be present in the fuel. A detailed chemical kinetic scheme for the oxidation of the higher hydrocarbon components of up to n-heptane was used to investigate analytically the combustion reactions of different fuel mixtures under constant volume adiabatic conditions with initial states that are similar to those during the ignition delay of a typical internal combustion engine. These comprehensive simulation calculations require much computing capacity and time that would preclude their incorporation in full simulation models of engine processes. A simplification is introduced based on replacing artificially the small concentrations of any higher hydrocarbons that may be present in the natural gas by a kinetically equivalent amount of propane in the fuel mixture. This is done such that the resulting equivalent fuel has the same ignition delay as the original fuel under constant volume engine T.D.C. conditions. This "propane equivalent" concept was used in full engine simulation models while employing a relatively short scheme of 150 steps for the oxidation of fuel mixtures of propane, ethane, and methane in air.* [DOI: 10.1115/1.1445438]

## Introduction

The utilization of natural gas in diesel engines is increasing due to its potential favorable cost and lower exhaust emissions, especially of $NO_x$ and particulates, compared to conventional diesel operation. This is usually achieved while providing a comparable high efficiency, particularly at high loads. Various approaches are available for converting diesel engines into pilot ignited dual fuel engine operation (Turner and Weaver [1]).

Variations in the composition of natural gas have been shown to have a direct impact on its ignition and combustion characteristics (e.g., Vimar and Harald [2]). This tendency may render some natural gases unsuitable for applications in engines when the rated output of a specific engine is to be maintained due mainly to the incidence of knock, which produces uncontrolled combustion with exceedingly fast and unacceptably high rates of energy release. These are critically dependent on the chemical reaction activity of the charge which becomes a controlling factor in establishing the relative suitability of various gaseous fuel compositions to satisfactory engine operation. Hence, the present contribution examines the consequences of changes in the composition of a pipeline natural gas on its combustion characteristics mainly in relation to its utilization in dual fuel engines.

A typical processed pipeline natural gas is made up mainly of methane with relatively small concentrations of other hydrocarbons (C2-C7+) and incombustible components (e.g., $N_2$ and $CO_2$). Despite the fact that methane is normally resistant to autoignition, it is known that the addition of small amounts of higher alkanes such as ethane, propane and butane can lower the ignition delay times of methane-oxygen mixtures in shock tubes (e.g., Higgins and Williams [3], Lifshitz et al. [4], Crossley et al. [5], and Frenklach and Bornside [6]). This reduction in the ignition delay was initially attributed to the heat release by these additives prior to combustion (Crossley et al. [5]). Subsequent studies showed that the ignition processes of these additives are chemically coupled to those of the methane (eg. Westbrook [7], methane-ethane; Frenklach and Bornside [6], methane-propane, etc.). The present contribution examines analytically, the nature of this chemical kinetic coupling for the complex mixtures normally found in natural gases and particularly the contribution of the presence of some higher hydrocarbons to changes in the compression ignition characteristics under engine-like conditions. Some guidelines are presented for considering the reactivity of these fuel mixtures in engine performance modeling applications.

## Kinetic Modeling

There has been a significant and rapid progress made in recent years in the status of chemical kinetic modeling of the combustion of hydrocarbon fuels under engine like conditions. These enabled the derivation of comprehensive chemical kinetic models for higher hydrocarbons. Such comprehensive schemes can be incorporated into a variety of predictive models to analyze engine performance under operating and design conditions with a reasonable level of confidence. They may be used also to validate the results of other models that employ only reduced kinetic schemes for a wide range of fuel mixtures commonly encountered in natural and industrial gases.

A comprehensive kinetic model for the oxidation of both normal and isomer hydrocarbon fuels containing up to eight carbon atoms, was compiled for this investigation (Khalil [8]) from the work of others to include 1966 elementary reactions and 380 species (e.g., Axelsson et al. [9], Westbrook and Pitz [10,11] Westbrook [12], Karim et al. [13], and Samuel [14]). This comprehensively detailed kinetic model is identified in this contribution as the ''full'' scheme. Calculated results obtained using this full scheme for the autoignition of n-heptane-air mixtures under adiabatic constant volume conditions were shown to agree well with the corresponding experimental results of Ciezki and Adomeit [15]. The hierarchical nature of the comprehensive kinetic mechanism allows the simulation of the oxidation reactions of fuel mix

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, October 1998; final revision received by the ASME Headquarters, June 2001. Associate Editor: D. N. Assanis.

tures that are commonly important, containing components of a lesser number of carbon atoms such as methane, ethane, ethylene, or propane. This model was used in the present investigation to examine some key features of the homogeneous oxidation in air of different hydrocarbon gases such as methane, ethane, propane, and butane in the presence of n-heptane vapor, representing a typical higher hydrocarbon having a cetane number similar to that of a light diesel fuel, under engine-like conditions of constant volume process. The nature and extent of any chemical interaction that may take place between these markedly different hydrocarbons were also examined. On the basis of these calculations, special-purpose reduced schemes could be formulated to a varying degree of detail. The application of these suitably reduced schemes were shown to yield good agreement with the results obtained using the original full model under conditions relevant to engine applications for the major operating parameters and species that are usually of interest to common engine performance considerations.

## The Kinetic Scheme

The chemical kinetic scheme lists the likely set of significant homogeneously uncatalized reactions of the main species participating in the chemical reaction of the system, and their dependence on temperature and pressure. A scheme is considered successful if it can be applied satisfactorily over wide ranges of temperature, equivalence ratio, pressure, and other key relevant physical parameters. In order to cover the ranges of temperature and pressure normally encountered in practical combustion devices, these models are invariably made up of a very large number of elementary reactions that cater for both high and low temperature regimes. The model and the associated kinetic data are then validated indirectly through comparison of the numerically predicted values of various operating parameters with the corresponding experimental values obtained in combustion devices such as shock tubes, reactors, rapid compression machines, and motored and fired engines.

The major high temperature reactions include the kinetics of the oxidation of hydrogen and carbon monoxide, involving radical-molecule, radical-radical, and unimolecular reactions of the common species H, O, OH, $HO_2$, $H_2O_2$, $O_2$, $H_2$, $H_2O$, CO, and $CO_2$. These reactions are supplemented by those of the relatively smaller organic molecules and radicals produced during the course of the reaction, such as CHO, $CH_2O$, . . . . etc. Subsets of reactions associated with other $C_1$, $C_2$, and $C_3$ molecules or radicals are also included such as $C_2H_2$, $C_2H_4$, . . . etc. The resulting group of reactions then tends to contain repetitive themes to represent the reactions of the higher alkane and their conjugate alkyl radicals (e.g., Griffiths [16]). For higher hydrocarbons, the initiation reactions that produce the pool of radicals usually tend to involve the breaking up of the C-C bonds rather than the C-H bonds which usually have much higher bond dissociation energies. The products resulting from the decomposition of the larger hydrocarbon molecules are usually of comparable sizes. Once a pool of radicals is established, the fuel consumption occurs primarily through an H-atom abstraction process at different sites of the alkane fuel. The large alkane and alkyl radicals produced by H-atom abstraction are consumed by thermal decomposition to form smaller intermediate species such as $C_2H_2$ and $C_2H_4$ which are converted later on to $H_2$ and CO.

There are additional classes of reactions that become significant at temperatures below 1000 K. The degradation of the hydrocarbon at low temperature tends to follow a sequential breakdown such as $C_4 \mapsto C_3 \mapsto C_2 \mapsto C_1$ rather than the earlier fragmentation into smaller intermediates of comparable size that may occur at higher temperature (e.g., $C_4 \mapsto 2C_2$). The reactions involving different propagating free radicals (e.g., $RO_2$, RO, $HO_2$, OH, where R is an alkyl radical) are particularly important at lower temperatures and higher pressures (Griffiths [16]). Under these conditions the radicals (R) tend to combine with molecular oxygen to form alkylperoxy radical ($RO_2$) which in turn may produce alkyl hydroperoxides (ROOH). These hydroperoxides later on decompose to produce OH radicals and various oxygenated hydrocarbon species. These processes are described in the literature (e.g., by Westbrook [17–19] and others) and are summarized by Khalil [20] who lists the lengthy reaction kinetic scheme and data employed in this investigation.

The use of very lengthy full kinetic schemes in modeling procedures tends to be too unwieldy, requiring excessive computing capabilities and time. Also, for many engineering applications the extensive resulting information often may be of little interest. Accordingly, reduced kinetic models when developed properly can simulate adequately the main features of the combustion reactions. In the present work, the comprehensive kinetic scheme was reduced in length by a factor of around three while retaining the values of the kinetic data of the individual reaction steps. Various procedures were employed for the reduction that included sensitivity analyses that attempt to quantify the response to a prescribed extent of variation in the rate constants of each reaction step of the kinetic scheme. To perform this analysis properly and systematically for all likely sets of conditions is an unmanageable task due to the very large number of possible combinations. Hence, the progressive reduction of the number of reaction steps was made through using a sensitivity analysis based on the removal of reaction steps that involve certain reactive species that were judged to be of relatively minor importance under the conditions considered. This was checked by comparing the resulting temporal development of calculated temperature and major species concentrations with those obtained using the full scheme and ensured that only insignificantly small differences occur. Yet, the reduced scheme developed remained still relatively lengthy. Typically, it was made up of 683 reactions and 158 species for mixtures of n-heptane and lighter hydrocarbons up to propane; while 794 reactions were needed for those up to n-butane. The resulting reduction in computing time was of the order of 80 percent. Reliable calculated results were produced as long as no attempt was made to match the concentrations of all the reactive species including those of minor importance. Further reductions in the length of the scheme are obviously possible but with a serious deterioration in the accuracy of the predicted values and the extent of coverage of conditions.

Differential equations were solved at any time for the concentrations of the "$n$" reacting species present in the scheme along with the energy equation. The solution of the "$n+1$" coupled, nonlinear, first-order ordinary differential equations, while employing established subprograms (e.g., Numerical Recipe Software, Cambridge, MA), with sufficiently small steps, produced the concentrations and temperature at the end of each time increment. Then, the pressure and density of the mixture could be obtained using the gas mixture and state equations.

## Results and Discussions

Detailed calculations using the *full* kinetic scheme were made typically for n-heptane vapor in the presence of different simpler hydrocarbons with air. N-heptane was chosen since it can be considered to represent the ignition characteristics of diesel fuel (Cetane number of approx. 60).

Figure 1 shows the calculated temperature variations with time for the adiabatic constant volume combustion of n-heptane-methane mixtures in air for initial conditions relevant to those of an engine near the top dead center position. The characteristic two-stage ignition phenomenon depicted by a pressure step change before hot ignition, can be seen in the reactions of heptane and methane-heptane mixtures at high initial pressure and intermediate temperature ranges. Such two-stage ignition reactions were expectedly absent for methane. It can be seen that significant changes in the length of the ignition delay take place as the relative composition of heptane and methane in the fuel mixture is

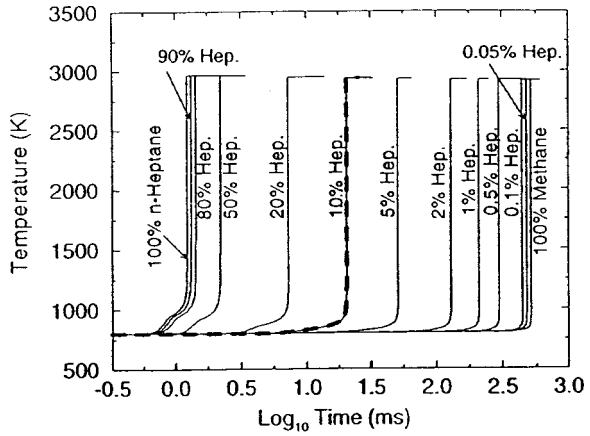



**Fig. 1 The variations of calculated temperature with reaction time for a variety of n-heptane-methane mixtures in air in an adiabatic constant volume process, $T_i$=800 K, $P_i$=2.8 MPa, and $\varphi$=1.0, dotted lines indicates results with the reduced scheme**

changed. The presence of a relatively small fraction of n-heptane in the fuel mixture reduces the ignition delay of methane very substantially.

Figure 2 shows the effects of a progressive increase in the heptane content of the fuel mixture on the ignition delay and combustion completion times for two values of initial temperature at an initial pressure of 2.8 MPa under constant volume adiabatic conditions. The delay period decreases very rapidly from that of methane with the introduction of some of heptane in the fuel mixture (Khalil and Karim [8]). When a significant amount of n-heptane is present with the methane (e.g., >10 percent by volume) then the delay and combustion times change only relatively little. Also, the value of the delay for these fuel mixtures always lies between those for pure heptane and methane. As an example, the addition of a mere 0.5 percent n-heptane to the methane (i.e., 99.5 percent methane and 0.5 percent heptane mixture by volume) resulted in 50 percent and 75 percent reduction in the ignition delay for initial temperatures of 800 K and 650 K, respectively, when compared to that for pure methane. On the other hand, for fuel mixtures with high heptane content, a substantial amount of methane (as high as 55 percent by volume) is required to be present to produce an ignition delay that is twice that observed for pure n-heptane, for an initial temperature of 800 K.

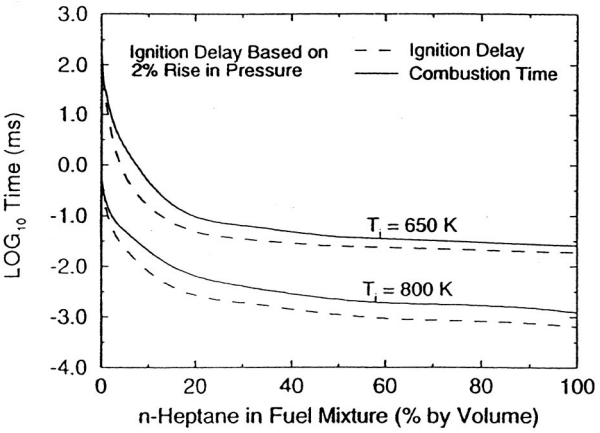



**Fig. 2 Variations of the calculated delay and total combustion times with n-heptane percentage in n-heptane-methane fuel mixture for two initial temperatures in an adiabatic constant volume process $P_i$=2.8 MPa and total $\varphi$=1.0**

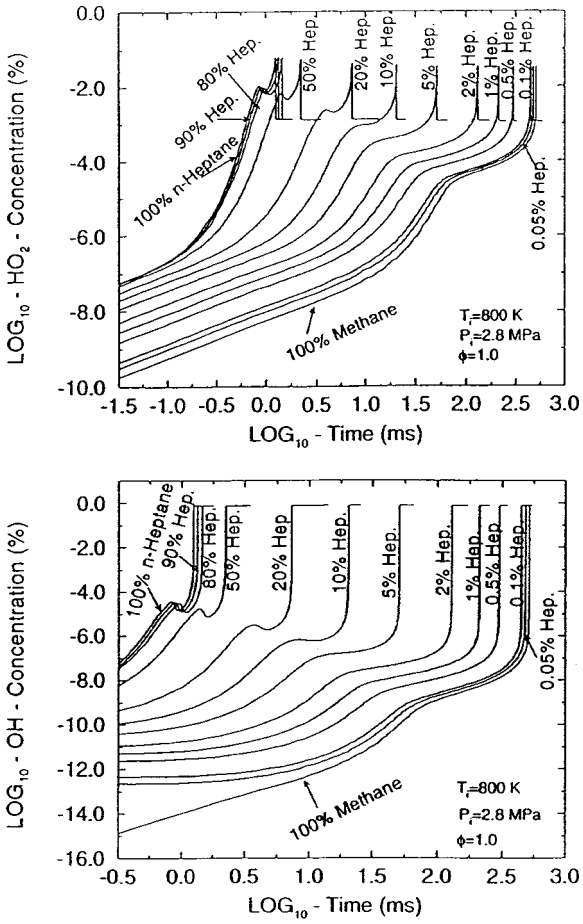



**Fig. 3 Variations of calculated OH and $HO_2$ concentrations with time for a variety of n-heptane-methane mixtures in air in an adiabatic constant volume process with $T_i$=800 K, $P_i$=2.8 MPa and total $\varphi$=1.0**

The length of the ignition delay is very sensitive to changes in the heptane content of the fuel mixture at low initial temperature levels, due to the very active thermal and kinetic roles of the n-heptane reactions then. These reactions raise the temperature of the mixture to levels when the methane reactions begin to have an effect. Also, the conversion of some heptane produces a pool of radicals in the early stages which accelerate the oxidation process. Figure 3 shows typically the effects of the progressive addition of n-heptane to methane on the concentrations of the chain branching radicals OH and $OH_2$. For example, after 1.0 ms from the commencement of the reactions, which is a representative time scale for high-speed engines applications, it can be seen that the addition of 0.5 percent heptane resulted in an increase of the concentrations of OH, $HO_2$ by orders of magnitude.

The principle product of the initiation reactions of methane is $CH_3$ via mainly the thermal decomposition of methane and reactions with oxygen molecules. These reactions are very slow except at elevated temperatures which explains the low autoignition tendency of methane. The methyl radical hence reacts slowly in the second phase (chain-branching) to produce a pool of chain propagating radicals, such as OH, H, $HO_2$, and O that accelerate the oxidation process by the destruction of the fuel molecule and reaction with intermediate products including formaldehyde. The intermediate products then are rapidly converted to the final products.

The most significant effect of the addition of a higher hydrocarbon to the methane is on the initiation phase. Frenklach and Bornside [6] showed that for relatively high initial temperatures

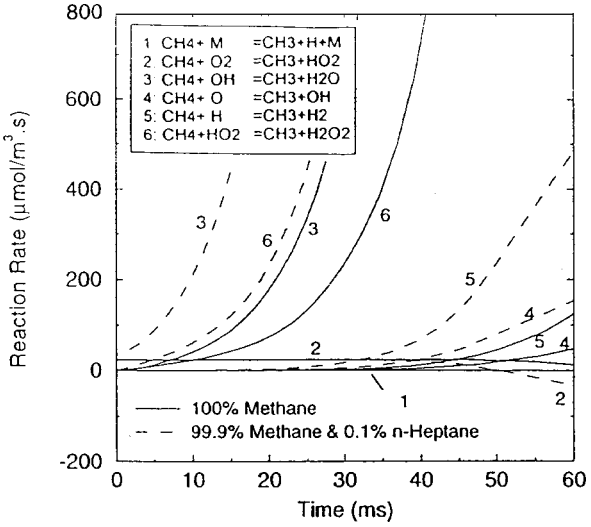



**Fig. 4 Variations of the rates of selected reactions with time for methane and a fuel mixture (99.9 percent methane and 0.1 percent n-heptane) in air in an adiabatic constant volume process with $T_i$=800 K, $P_i$=2.8 MPa and total $\varphi$=1.0**

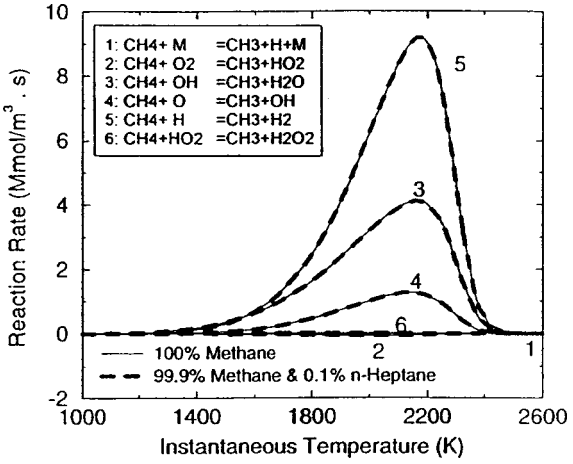



**Fig. 5 Variations of the rates of selected reactions with combustion temperature for methane and a fuel mixture (99.9 percent methane and 0.1 percent n-heptane) in air in an adiabatic constant volume process with $T_i$=800 K, $P_i$=2.8 MPa and total $\varphi$=1.0**

(e.g., over 1200 K), the abundant production of $CH_3$ and other branching radicals suppresses the much slower conventional initiation reactions of pure methane and the methane molecule destruction proceeds via H-abstraction reactions by these radicals. This could explain the dramatic sensitization of methane-air mixtures by the small addition of higher hydrocarbons at lower temperature levels. In general this strong effect is brought about by the fact that the portion of the reaction time within the low temperature region represents over 90 percent of the total ignition delay. Figure 4 shows the variation of the rates for some initiation reactions at the early stages of combustion for pure methane and a fuel mixture of 99.9 percent methane and 0.1 percent n-heptane, at a relatively low initial temperature of 800 K. It can be seen that for pure methane, a main initiation route to produce the methyl radical is via the reaction with molecular oxygen (i.e., reaction 2) rather than through thermal decomposition (reaction 1). As chain branching radicals are produced, the rates of reactions 3 through 6 show a rapid increase in comparison to reaction 2 which starts to proceed in the opposite direction (e.g., at 50 ms approx.). The maximum rate of methane depletion occurred through the reaction with OH, followed by $HO_2$, then H and O. For the sensitized mixture, the rate of reaction 3 (with OH) is initially slightly higher than that of reaction 2 (with $O_2$) and provides a main route for methane initiation reactions. The figure shows clearly the kinetic effect of the addition of a small amount of n-heptane on the initiation phase. Even though the temperatures were virtually the same, the rates of reactions 3 through 6 increased substantially due to the relative abundance of these radicals resulting from the heptane preignition reactions, which is consistent with Fig. 3. As the temperature increases over 1400 K, the rate of reaction 5 (with H) outweighed reaction 3 (with OH) as seen from Fig. 5. On the other hand the reaction of methane with $HO_2$ became unimportant due to the lower concentrations of $HO_2$ at temperatures higher than 1000 K.

The sensitizing effect described earlier could be further illustrated by comparing the effect of adding n-heptane to both dry and moist carbon monoxide as a fuel. It is well known that carbon monoxide oxidation rates are dramatically changed by the addition of minute amounts of moisture. The oxidation mechanism for dry carbon monoxide is through $CO+O=CO_2$ and $CO+O_2=CO_2+O$. However, in the presence of $H_2-O_2$ mechanism the main route is via $CO+OH=CO_2+H$ and $CO+HO_2=CO_2+OH$ which accelerate the reaction rate. This is consistent with calculated results showing that the addition of water to the extent of 0.1 percent of the carbon monoxide concentration resulted in the reduction of the ignition delay by about 20 percent at 800 K and 2.8 MPa. The presence of 0.01 percent and 0.1 percent n-heptane in dry carbon monoxide resulted in the corresponding reduction of the combustion time by a factor of around 5.5 and 15, respectively, which reflects the strong sensitizing effect of n-heptane. However, the delay time was unaffected by the addition of any n-heptane both for the dry and moist carbon monoxides combustion in air. The addition of very minute amounts of n-heptane furnished the required pool of radical that were obtained by the addition of moisture while enhanced the chain branching mechanism leading to the significant reduction in the time observed.

The effect on the combustion times of the addition of n-heptane to alkanes higher than methane (ethane, propane, and butane) and to hydrogen is shown in Fig. 6. The general trend is a decrease in the sensitizing effect of heptane with the increase in the alkane chain length in a descending order from n-butane to methane; a reflection of the higher reactivity of the longer chain alkanes. For alkane fuels heavier than methane, the initiation reactions occur mainly through the breaking of a C-C bond since the C-H bond has much higher bond dissociation energy. The longer the chain, the easier is its breaking into smaller intermediate hydrocarbons and chain propagating radicals and hence a higher reactivity and a

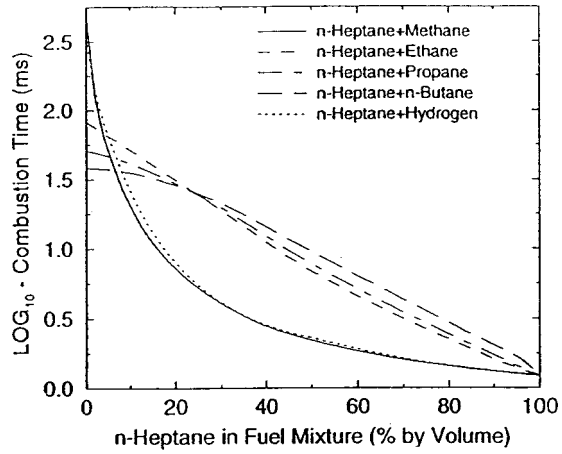



**Fig. 6 Variations of the calculated combustion times with n-heptane percentage in various binary mixtures of n-heptane in an adiabatic constant volume process with $T_i$=800 K, $P_i$=2.8 MPa and total $\varphi$=1.0**

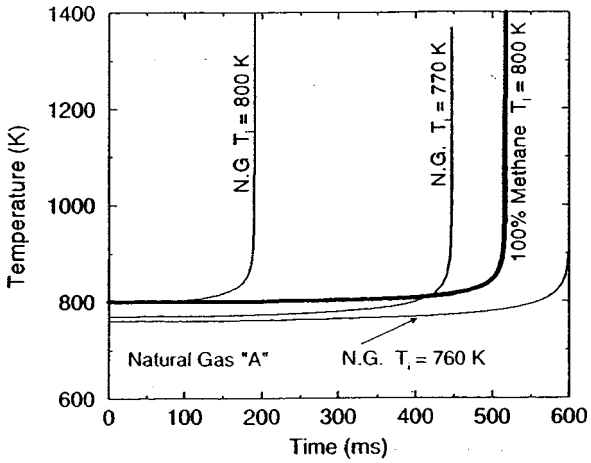



**Fig. 7 Variations with time of the calculated temperature of natural gas "A" of Table 1 for various initial temperatures. The corresponding variations for methane at an initial temperature of 800 K are shown; $P_i$=2.8 MPa**

reduction in the effects of addition of the heavier hydrocarbon. This is in contrast to methane, where a C-H bond must be broken in order to commence the initiation reactions which results in the characteristic higher resistance for autoignition of methane. This explains the dramatic reduction in the combustion times of methane shown in Fig. 6 due to the addition of the heptane that furnishes the required propagating radicals for the initiation step. Similarly, rapid reductions in the reactions of hydrogen can also be observed. The ignition delay for mixtures containing relatively high percentages of n-heptane was shown to have an opposite trend, where the delay time was shorter for lower hydrocarbons (Fig. 7). This effect is attributed to the higher percentage of n-heptane present in the fuel-air mixture for fuels of low molecular weight. For example, the percentage of n-heptane in the fuel-air mixture is 0.5 percent for a mixture of 20 percent n-heptane and 80 percent n-butane, while the corresponding value for n-heptane-methane mixture is approximately one percent.

## Some Examples Relating to Natural Gas Oxidation

To illustrate the effect of the presence of relatively small amounts of higher hydrocarbons on the reactivity of a natural gas, the calculated ignition delay for a homogeneous stoichiometric mixture in air of a typical natural gas "*A*," with a composition shown in Table 1, was compared to that of pure methane under constant volume conditions and initial temperature of 800 K and 28 bars representing engine-like conditions near T.D.C., (Fig. 7). The presence in the natural gas of approximately 6 percent by volume of a mixture of higher than methane hydrocarbons resulted in a reduction of the time for the autoignition to 65 percent of that for pure methane under the same conditions. This would correspond to a reduction in the initial temperature of the natural gas "*A*" by about 35 K to produce the same ignition delay as that of methane. If this was related to operating conditions of a typical spark ignition or compression ignition engine, then it would correspond roughly to a reduction of 13 percent in the compression ratio or a lowering of at least 15 K in the intake temperature of the charge. Due to the fact that knock in engines is kinetically controlled, the results of the adiabatic constant volume process can serve as an indication of the relative reactivity of the mixture and its tendency to autoignite.

It would be expected usually that such an extent of variations in the fuel composition would have a relatively insignificantly small effect on flame propagation processes in comparison to the corresponding tendency of the fuel-air mixture to autoignite. Thus, this relatively simple kinetic modeling approach of comparing the calculated value of the ignition delay of natural gas with those of methane obtained under the same engine like settings of temperature and pressure may be employed to provide approximate guidelines for the changes in operating conditions needed to produce a methane-like performance, in relation to the tendency of the engine to knocking with the natural gas. To our knowledge, no such convenient approach has been used before.

## An Example of a Practical Case

A number of different samples of natural gases with known compositions shown in Table 2 reflects actual reported fluctuations in the composition of a natural gas supply over a period of time. These samples were used to establish the consequences of such variation in composition on their suitability as fuels for a dual fuel engine. Gas sample No. 9 shows relatively high concentrations of the hydrocarbon components heavier than methane, while sample No. 3 has relatively high concentrations of propane and n-butane. Similarly, sample No. 2 contains lower concentra

**Table 1 Composition of Natural Gas "*A*"**

| Gas Constituents | % (mol) |
|---|---|
| $CH_4$ | 87.00 |
| $C_2H_6$ | 3.00 |
| $C_3H_8$ | 1.40 |
| n-$C_4H_{10}$ | 0.30 |
| i-$C_4H_{10}$ | 0.24 |
| n-$C_5H_{12}$ | 0.06 |
| i-$C_5H_{12}$ | 0.95 |
| CO | 0.17 |
| Air | 6.88 |
| Total | 100.00 |

**Table 2 Composition of different natural gases**

| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 |
|---|---|---|---|---|---|---|---|---|---|
| $N_2$ | 4.83 | 5.42 | 0.84 | 3.53 | 3.91 | 3.91 | 2.56 | 2.49 | 4.32 |
| $CH_4$ | 86.16 | 86.00 | 89.96 | 88.34 | 87.81 | 87.61 | 88.74 | 87.99 | 85.26 |
| $CO_2$ | 1.14 | 0.98 | 1.41 | 1.08 | 1.07 | 1.23 | 1.28 | 1.42 | 1.45 |
| $C_2H_6$ | 5.43 | 5.47 | 5.38 | 5.27 | 5.30 | 5.30 | 5.64 | 5.82 | 6.06 |
| $C_3H_8$ | 1.70 | 1.58 | 1.92 | 1.29 | 1.36 | 1.29 | 1.39 | 1.67 | 2.13 |
| i-$C_4H_{10}$ | 0.19 | 0.18 | 0.13 | 0.13 | 0.16 | 0.17 | 0.11 | 0.16 | 0.19 |
| n-$C_4H_{10}$ | 0.30 | 0.31 | 0.28 | 0.21 | 0.24 | 0.28 | 0.18 | 0.29 | 0.35 |
| i-$C_5H_{12}$ | 0.08 | 0.08 | 0.00 | 0.05 | 0.06 | 0.07 | 0.04 | 0.06 | 0.08 |
| n-$C_5H_{12}$ | 0.08 | 0.08 | 0.03 | 0.04 | 0.05 | 0.07 | 0.04 | 0.06 | 0.07 |
| C6+ | 0.10 | 0.09 | 0.01 | 0.06 | 0.06 | 0.08 | 0.04 | 0.05 | 0.09 |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

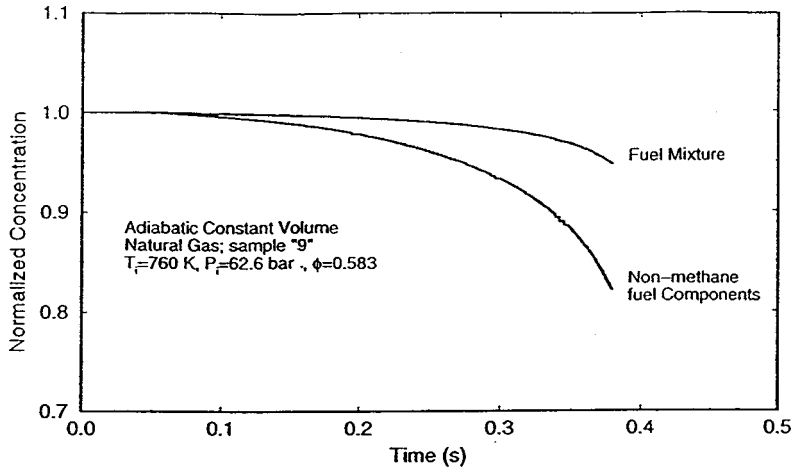



**Fig. 8 Variation with time of the normalized concentration of the fuel and for the non-methane components**

tions of propane compared to that of sample No. 1, while the concentrations of the higher hydrocarbons are almost the same in both samples.

The relative suitability of these samples for engine applications was established at conditions corresponding to the values of air-fuel ratio, temperature and pressure normally found during the ignition delay period in a dual fuel engine when operating at around its declared rated power output. Moreover, the relative tendency of these natural gases to produce knock was evaluated in terms of their pre-ignition reaction activity at constant volume and adiabatic conditions.

Figure 8 shows a plot of the variation with time of the instantaneous normalized concentrations of fuel No. 9, due to chemical reaction. The corresponding variations in the relative concentrations of the non-methane fuel components in an adiabatic constant volume reaction under engine like conditions are also shown. It can be seen that even for the very lean typical mixture considered ($\varphi = 0.583$), these components commence significant reaction activity at a very early stage leading to autoignition of the mixture due to the thermal and kinetic effects they produce.

To emphasize the high reactivity of the components of natural gas that are higher than propane, Fig. 9 shows a comparison of the temperature time development through autoignition for the actual composition of the natural gas, No. 9, and for an assumed fuel having double the concentrations of $C_4$, $C_5$, and $C_6+$. The heating value of the fuel was maintained the same, through some minor adjustment to the concentration of methane. A very substantial reduction in the autoignition time is achieved while encountering essentially no change in the resulting peak temperature. Figure 10 shows the corresponding relative preignition reactivity of the various individual components of the natural gas for the case shown in Fig. 8. It can be seen that methane is hardly reacted before autoignition while the relative reactivity of the other components increases very significantly with the increase in their molecular weight. Although propane is much more reactive than methane under these conditions, the higher hydrocarbons (C6+) are yet more reactive even at the early stages of the reaction time. Figure 11 shows the relative variations in the reactivity of the different natural gases listed in Table 2. It is evident that the fluctuations in their composition produce different levels of reactivity leading to auto-ignition. The sample designated as No. 9, produced the fastest reactions, while the sample No. 4 had the lowest. A comparison of the relative reactivity of the fuel mixtures shows that although the relative concentration of methane varied by only two percent, the reactivity of these fuel mixtures varied very significantly as a result of the variations in the concentrations of their other components. This would emphasise the important role played by these higher hydrocarbons.

Fuel samples No. 4 and No. 6 both contained the same amount of propane and virtually the same amount of ethane. Both of these

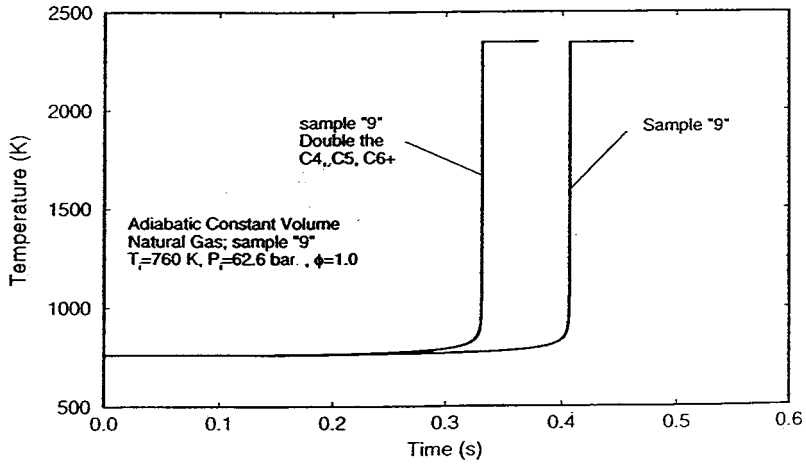



**Fig. 9 Variation in the temperature with time for a natural gas<sample "9," and for a sample having double the concentration of C4<C5 and C6+**

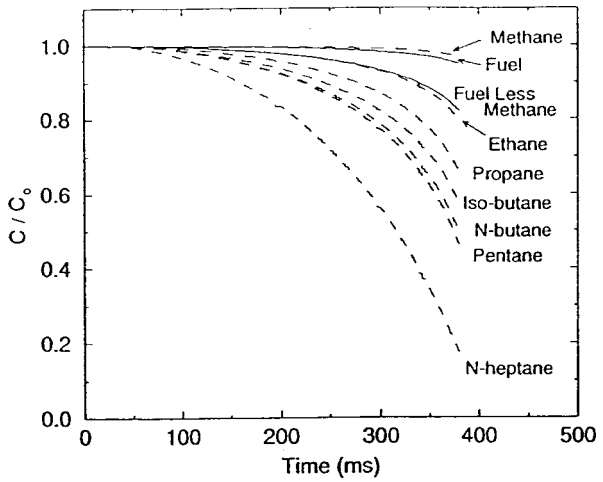



**Fig. 10 Variations with time of the relative concentrations of the various n-alkanes components of natural gas "B" in an adiabatic constant volume process. $T_i$=760 K, $P_i$=6.26 MPa and total $\varphi$=0.583**

fuels tend to shorten the ignition delay when added to methane. A higher concentration of the higher than propane hydrocarbons in sample No. 6 resulted in a significant increase in its reactivity. On the other hand, comparing samples No. 3 and No. 4 shows that the higher concentrations of propane and butane in sample No. 3 resulted in a higher reactivity despite the fact that sample No. 4 contained higher concentrations of the $C_5$ and $C_6$ hydrocarbons. Obviously, a natural gas that is essentially methane will show extremely low reactivity in comparison to those shown.

These examples help to demonstrate how detailed chemical kinetic modeling can be employed as a tool for providing guidelines towards solving practical problems in engines arising from changes in the composition of natural gas.

## A "Propane Equivalent Composition" for a Natural Gas

At present, it appears impractical to account in engine modeling comprehensively for the combustion reactions of natural gas components and reliable reduced kinetic schemes for the oxidation of higher hydrocarbons remain too lengthy to be used in extensive engine simulation calculations. For the simpler common fuels and their mixtures (e.g., methane, ethane, propane, ethylene, hydrogen, carbon monoxide, etc.) chemical kinetic reaction schemes of manageable length can be formulated that could be incorporated in engine modeling applications. Accordingly, a further simplification to the kinetic modeling of different natural gases for engine applications is proposed by us. It is suggested to replace the composition of the natural gas which contains higher hydrocarbons ($C_4$,$C_5$,$C_6$+) by an equivalent composition made up of only lower hydrocarbons which would have essentially similar reactivity and heating value as those of the original natural gas. This is done on the basis that the components of natural gas that are higher than $C_3$ are replaced by an equivalent amount of propane such that the same rate of variations with reaction time of the original fuel is obtained. Propane was chosen since it is commonly available and can be used extensively in engine testing and operation. Appropriate slight adjustment in the methane concentration may be needed to retain essentially the same heating value.

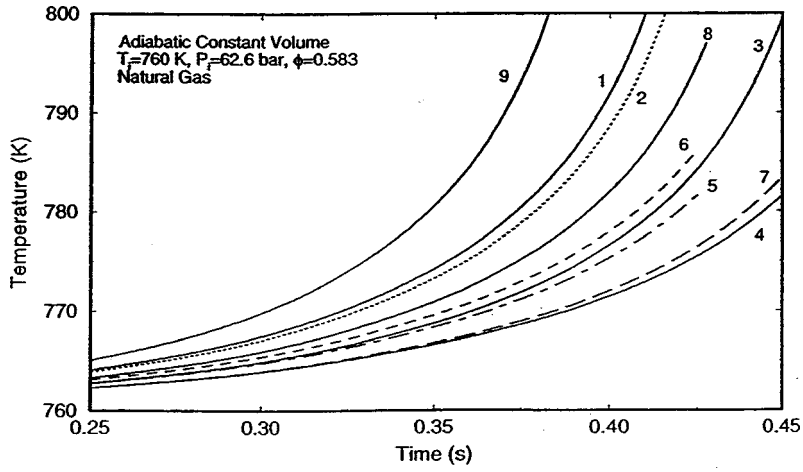



**Fig. 11 Variation in the temperature-time development for the different natural gas samples listed in Table 2**

**Table 3 Equivalent versus original compositions for four natural gases**

| | Case 9 | | Case 1 | | Case 3 | | Case 4 | |
|---|---|---|---|---|---|---|---|---|
| | Orig. | Equi | Orig. | Equi | Orig. | Equi | Orig. | Equi |
| $N_2$ | 4.43 | 4.32 | 4.83 | 4.83 | 0.84 | 0.84 | 3.53 | 3.53 |
| $CO_2$ | 1.45 | 1.45 | 1.14 | 1.14 | 1.41 | 1.41 | 1.08 | 1.08 |
| $CH_4$ | 82.56 | 84.12 | 86.16 | 5.02 | 89.96 | 88.86 | 88.34 | 87.25 |
| $C_2H_6$ | 6.06 | 6.06 | 5.43 | 5.43 | 5.38 | 5.38 | 5.27 | 5.27 |
| $C_3H_8$ | 2.13 | 4.05 | 1.70 | 3.58 | 1.92 | 3.50 | 1.29 | 2.87 |
| n-$C_4H_{10}$ | 0.19 | 0.00 | 0.19 | 0.00 | 0.13 | 0.00 | 0.13 | 0.00 |
| i-$C_4H_{10}$ | 0.35 | 0.00 | 0.30 | 0.00 | 0.28 | 0.00 | 0.21 | 0.00 |
| i-$C_5H_{12}$ | 0.08 | 0.00 | 0.08 | 0.00 | 0.03 | 0.00 | 0.05 | 0.00 |
| n-$C_5H_{12}$ | 0.07 | 0.00 | 0.08 | 0.00 | 0.03 | 0.00 | 0.04 | 0.00 |
| n-$C_7H_{16}$ | 0.09 | 0.00 | 0.10 | 0.00 | 0.01 | 0.00 | 0.06 | 0.00 |
| Total | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 | 100.0 |

Thus, the resulting equivalent fuel gas mixture would consist of only methane, ethane, propane and diluent gases such as nitrogen, carbon dioxide and water vapor.

Table 3 shows the original and calculated equivalent composition of four different natural gases (Nos 1, 3, 4, and 9 in Table 1). The simulations were carried largely for convenience and economy of computation under adiabatic constant volume conditions while using the detailed kinetic scheme. It can be seen that although the concentrations of butane and the higher fuel components tend to be very small, the equivalent amount of propane needed in this substitution is around THREE times the corresponding value of their concentrations. Moreover, the reactivity of the fuel No. 9, under these conditions will be equivalent to having as much as 4.05 percent propane present in the gas, instead of the declared original concentrations of merely 2.13 percent. Since it is well known that the increased presence of propane with methane can seriously undermine the knock resistance quality of methane, this emphasizes the need to reduce the concentration of higher hydrocarbons that are higher than propane to an absolute minimum. Otherwise, the reactivity of the natural gas will be unacceptably high for knock-free operation in engines.

Accordingly, this propane equivalent concept for replacing the higher hydrocarbon components of natural gas may be employed to render the simulation of the combustion of complex fuel mixtures such as those in natural gases more manageable in engine performance modeling.

## Conclusions

- The presence of small concentration of n-heptane with methane can produce very substantial changes in its autoignition and combustion characteristics.
- Relatively small changes in the composition of natural gas, particularly in relation to the higher hydrocarbon components (butane and heavier), can produce significant changes to the reactivity of the natural gas.
- A relatively simple approach of comparing the calculated values of the ignition delay of natural gas to that of methane under adiabatic constant volume combustion and engine-like conditions can serve as a guideline for effecting the required changes to engine operating conditions to avoid the onset of knock.
- The ''propane equivalent'' concept proposed can serve as a tool for the simulation of combustion of complex fuel mixtures such as natural gases in engine performance modeling calculations. Such an approach also demonstrates the impact of small changes in the concentration of the higher hydrocarbon components of the natural gas through the resulting substantial inflation of the concentration of propane, a more reactive fuel than methane.

## Acknowledgment

The financial support of Natural Sciences and Engineering Research Council of Canada (NSERC) and Alternative Fuel Systems, Inc. is gratefully acknowledged. The contributions of Dr. C. K. Westbrook, Prof. I. Skorokhodov, Dr. P. Samuel, and Dr. Z. Liu to this investigation are gratefully acknowledged.

## References

[1] Turner, S. H., and Weaver, C. S., 1994, ''Dual Fuel Natural Gas/Diesel Engines Technology, Performance and Emissions,'' Gas Research Institute Technical Report No. 0094.
[2] Vilmar, A., and Harald, V., 1996, ''The Influence of Natural Gas Composition on Ignition in a Direct Injection Gas Engine Using Hot Surface Assisted Compression Ignition,'' SAE Paper No. 961934.
[3] Higgin, R., and William, A., 1969, ''A Shock Tube Investigation of the Ignition of Lean Methane and n-Butane Mixtures With Oxygen,'' *12th Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA, p. 579.
[4] Lifshitz, A., Scheller, K., Burcat, A., and Skinner, G. B., 1971, ''Shock-Tube Investigation of Ignition in Methane-Oxygen-Argon Mixtures,'' Combust. Flame, **16**, p. 311.
[5] Crossley, R., Dorko, E., Scheller, K., and Burcat, A., 1972, ''The Effect of Higher Alkanes on the Ignition of Methane-Oxygen-Argon Mixtures in Shock Waves,'' Combust. Flame, **19**, p. 373.
[6] Frennklach, M., and Bornside, E., 1984, ''Shock-Initiated Ignition in Propane Mixtures,'' Combust. Flame, **56**, pp. 1–27.
[7] Westbrook, C. K., 1979, ''An Analytical Study of the Shock Tube Ignition of Mixtures of Methane and Ethane,'' Combust. Sci. Technol., **20**, pp. 5–17.
[8] Khalil, E., Samuel, P., and Karim, G. A., 1996, ''An Analytical Examination of the Chemical Kinetics of the Combustion of N-Heptane-Methane Air Mixtures,'' SAE Paper No. 961932.
[9] Axelsson, E. I., Brezinsky, K., Dryer, F. L., Pitz, W. J., and Westbrook, C. K., 1988, ''Chemical Kinetic Modelling of the Oxidation of Large Alkane Fuels: N-Octane and Iso-Octane,'' *Twenty-First Symposium (International) on Combustion*, The Combustion Institute, Pittsburgh, PA, p. 783.
[10] Westbrook, C. K., and Pitz, W. J., 1988, ''Detailed Kinetic Modeling of Autoignition Chemistry,'' Transaction of SAE, **96**, Section 7, p. 559.
[11] Westbrook, C. K., and Pitz, W. J., 1991, ''Numerical Modeling of Combustion of Complex Hydrocarbon Fuels,'' *Numerical Approaches to Combustion Modeling* (Vol. 135, Progress in Astronautics and Aeronautics), AIAA, Washington, DC, p. 57.
[12] Westbrook, C. K., 1990–1992, personal communications.
[13] Karim, G. A., Hanafi, A., and Zhou, G., 1992, ''A Kinetic Investigation of the Oxidation of Low Heating Value Fuel Mixtures of Methane and Diluents,'' *Proceedings of the 15th Annual ASME/ETCE*, Houston, TX, ASME, New York.
[14] Samuel, P., 1994, ''Computational and Experimental Investigation of Ignition and Combustion of Liquid Hydrocarbon Fuels Within Homogeneous Environments of Fuel and Air,'' Ph.D. dissertation, The University of Calgary.
[15] Ciezk, H., and Adomeit, G., 1993, ''Shock-Tube Investigation of Self-Ignition of n-Heptane-Air Mixtures Under Engine Relevant Conditions,'' Combust. Flame, **93**, p. 421.
[16] Griffiths, J. F., 1995, ''Reduced Kinetic Models and Their Application to Practical Combustion Systems,'' Prog. Energy Combust. Sci., N. A. Chigier, ed., **21**, Nov., p. 27.
[17] Westbrook, C. K., and Pitz, W. J., 1986, ''Kinetic Modelling of Autoignition of Higher Hydrocarbons: n-Heptane, N-Octane and iso-Octane,'' *Complex Chemical Reaction Systems*, Springer-Verlag, New York, pp. 45–62.
[18] Westbrook, C. K., Pitz, W. J., Thornton, M., and Malte, P. C., 1988, ''A Kinetic Modelling of n-Pentane Oxidation in a Well-Stirred Reactor,'' Combust. Flame, **72**, pp. 45–62.
[19] Westbrook, C. K., and Pitz, W. J., 1991, ''Numerical Modelling of a Combustion of Complex Hydrocarbon Fuels,'' *Numerical Approaches to Combustion Modelling* (Vol. 135, Progress in Astronautics and Aeronautics), AIAA, Washington, DC, pp. 57–76.
[20] Khalil, E., 1998, ''Modelling the Chemical Kinetics of Combustion of Higher Hydrocarbon Fuels in Air,'' Ph.D. dissertation, University of Calgary.

**R. Banerjee**

**K. M. Isaac**

University of Missouri-Rolla,
Rolla, MO 65409

**L. Oliver**

**W. Breig**

Dayco Technical Center,
Springfield, MO 65807

# Features of Automotive Gas Tank Filler Pipe Two-Phase Flow: Experiments and Computational Fluid Dynamics Simulations

*Extensive flow visualization in an automotive fuel filler pipe made visible by introducing dyes and smoke in water and air, respectively, were conducted for nominal flow rates of 4–18 liters per minute. Video and still cameras were used for imaging. Features of the flow such as laminar-to-turbulent transition, progressive development of strong swirl along filler pipe axis, air entrainment, and mixing with the liquid were observed in the experiments. The experimental observations were supported by computational fluid dynamics (CFD) simulations of the flow which also showed features such as swirl and air entrainment.* [DOI: 10.1115/1.1445439]

## Introduction

Automotive refueling process has lately drawn increased attention because of environmental concerns regarding the fuel vapor released during refueling. Increased ozone levels in urban areas are believed to be related to gasoline vapor release. Current practice of capturing refueling vapor fall into a few different types. Some installations recover vapor at the nozzle spout by vacuum assist, and return the vapor to underground storage tanks (UST). This process has the drawback that the entrained air also enters the UST along with the vapor. Beginning with the 1998 model year, U.S. pollution control laws require that vehicles be equipped with devices to capture vapors during refueling. Such vehicle-installed vapor recovery systems are designated as on-board refueling vapor recovery (ORVR). One ORVR design objective is to prevent the entrained air from entering the UST. As a first step towards ORVR designs, vapor generation and transport during refueling needs to be understood to develop recovery techniques. Some of the factors that affect vapor emission are the geometry of the nozzle-filler pipe fit, the characteristics of automatic nozzle shut off, operating environment such as temperature and humidity, and fuel dispense rates. Due to various reasons, identifying ORVR-equipped vehicles as they enter refueling stations, prior to the beginning of refueling, is another important design objective. Vacuum-assist pumps at gas stations must be turned off to prevent air flow into nozzle/hose/pump, and the fill neck should not be at negative gauge pressure. In a balance-type system, which operates in the open/closed mode, ORVR sensing is required to change the state of balance. In this type of system, the pathway between the vehicle fuel tank and the UST is either open or closed depending upon the pressure at certain locations in the pathway. The California Air Resources Board (CARB) Stage II regulations require that, when refueling a non-ORVR vehicle, vapor capture efficiencies must be at least 95 percent. CARB regulations limit the back pressure in balance-type systems to 0.5 in. $H_2O$ (125 $N/m^2$) at 60 $ft^3/hr$. (28.32 liters/minute) of $N_2$ at standard temperature and pressure (STP) [1]. Various techniques are being explored to detect ORVR-equipped vehicles. These ''smart sensor interfaces'' between the dispenser nozzle and the fill neck include sensing hydrocarbon (HC) concentration, pressure, or flow direction in the fill neck. However, more studies were recommended before implementing these technologies [1].

Refueling is a complex process that involves unsteady multiphase turbulent flow, heat transfer, and mass transfer. Because of safety concerns, fuel tank location has moved further interior in the chassis, away from wings and under-trunk locations in recent vehicle designs, resulting in longer filler pipes. Few studies to examine the fluid mechanics and mass transfer mechanisms of refueling have been conducted. There are two main objectives in examining the refueling process. The first concerns reducing design cycle time by developing design guidelines for the filler pipes of present day vehicle designs. The second objective is to predict vapor emission levels during the refueling process, and develop techniques to reduce fuel vapor release during refueling.

Some of the earlier filler pipes were manufactured by blow molding high density polyethylene, which led to noncircular, nonuniform cross sections of the interior geometry. Such irregular geometries have higher flow losses compared to a tube with a smooth circular cross section. Fortunately, more recent designs are of circular constant area cross section steel pipes with much better control of cross-sectional area, cross-sectional geometry, and surface roughness. Moreover, they are formed with bends of large radii of curvature which tend to minimize the flow losses in them.

Existing design rules are based on empirical relations to determine flow losses in the filler pipe passage. However, these empirical design rules do not allow predicting, with an acceptable degree of confidence, important aspects such as air entrainment and mass transfer during refueling. An initiative to investigate the filler pipe flow using computational fluid dynamics (CFD) has been recently reported [2]. This effort is expected to yield detailed flow field information including air entrainment and mass transfer. The study is expected to shed light on the phenomena of spit-back, i.e., fuel drops escaping into the atmosphere because of the high-momentum fuel jet impinging the filler pipe wall and forming drops that rebound; and well-back, the process of the fuel flooding the filler pipe and flowing backwards at the filler pipe mouth. The study also includes plans for experimental verification of the CFD results.

Since the filler pipe is part of the fuel system consisting of various components that include the fuel nozzle, filler pipe, gas tank, sensor tube, vapor absorption canister, and the fuel lines to the engine, the filler pipe performance would also be dependent on the characteristics of these other components. However, including all these components in the analysis would make it a complex

Contributed by the Internal Combustion Engine Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the ICE Division, July 1999; final revision received the ASME Headquarters, June 2001. Associate Editor: D. N. Assanis.

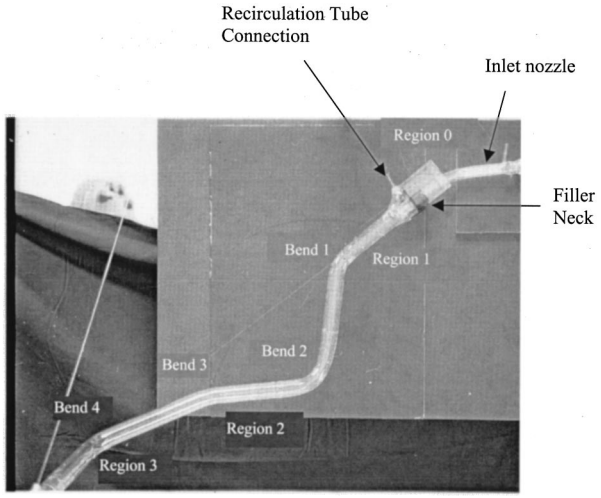



**Fig. 1 Photograph of 1:1 glass model of filler pipe and nozzle setup**

dynamical system. Lockhart et al. [3] have reported on the dispenser nozzle characteristics that can be used as input to ORVR design process. The dispenser nozzle characteristics and the filler pipe characteristics are closely linked because the nozzle shutoff mechanism is initiated when the nozzle sensor port senses the liquid level in the filler pipe when it reaches the sensor port. What ensues is a dynamic process with large variations in pressure and flow rates.

## Experimental Procedure

A photograph of the full-scale glass model of the filler pipe is given in Fig. 1. This geometry can be considered representative of a typical production filler pipe, with several bends. The second and the third bends are out of plane making the flow more complex by introducing a strong swirl component to the mean flow and high turbulence levels. The quartz 1:1 scale model has been used in all the experiments reported in the present work. The tube has 25 mm (1 in.) nominal ID and has a total length of ~1 m. It was formed from a single quartz tube bent at various angles to form five straight segments with large radii of curvature at the bends. Segments 1–5 are, respectively, 15 cm, 20 cm, 15 cm, 20 cm, and 30 cm long. The entire tube axis did not lie in one plane, but occupied a three-dimensional space. A larger diameter mouth that made a smooth transition to the smaller diameter tube was formed at the filler pipe inlet. During refueling, the fuel nozzle would be inserted into the mouth. A connection for the recirculation tube that would return gasoline vapor-air mixture from the tank to the filler pipe is provided as shown in Fig. 1. An insert with slotted openings at the side and having a stainless steel flapper valve that is part of an automotive production filler pipe was inserted into the opening. During refueling, the nozzle tip would open the flapper valve and extend into the filler pipe, about 50 mm beyond the flapper valve location.

The production filler pipe is made of steel and painted black inside. The tube inside surface equivalent roughness is estimated to be 45 $\mu$m, same as that of new commercial steel tubes. The quartz model, used in the experiments, is considered smooth with zero equivalent roughness. In the experiments, the filler pipe was oriented such that the axis of the first tube segment made an angle of 37.5 deg with the horizontal, and the first two straight segments of the tube lay in the vertical plane. Bends 1–4 had included angles of 136, 100, 163, and 121 deg, respectively.

The lighting arrangements used to obtain the flow visualization results presented in this study included dual fiber-optic lighting and a short duration (~10 ns) Nd:YAG pulsed laser beam formed into a light sheet. 35-mm still camera and S-VHS video camera were used for image acquisition. Various shutter speeds and aperture settings were used in the still photography. The S-VHS camera operated at the following settings: frame rate = 30 fps, shutter speed = 1/125 s, and aperture = 11.

City water was used as the flow medium. Test flow rates were in the range of ~4–20 liters/minute (lpm) (~1–5 gallons/minute). These flow rates correspond to a Reynolds number (based on tube diameter) range of 3236 to 16,181. The corresponding Reynolds number range, for the ~4–20 lpm flow rates, for gasoline would be 7879–39,397, which is higher because of the lower kinematic viscosity of gasoline. The tests were run at an ambient temperature of 20±1 deg C. The test conditions are summarized in Table 1.

## Results

**3.78 Liters/Minute (lpm).** At this flow rate the nozzle flow consists of a mixture of water and air entrained through the sensor port. The initially perturbed flow settles down to a smoother form towards the nozzle tip. Region 0, Frame 1, Fig. 2(*a*), shows details of the flow at the nozzle exit and the inlet to the filler pipe. In Region 1, Frame 6, Fig. 2(*a*), at 3.78 lpm, the liquid has a glassy appearance. It is interesting to note that as the liquid flows down, not all of it stays at the lowest part of the tube, but some of it sticks to the top surface as well. The corresponding flow observed using a video system shows that the flow is not completely steady, but has frame-to-frame variations in the free surface shape. At the first bend the liquid stream jumps across and impinges the opposite wall because of its momentum.

In Region 2, Frame 11, Fig. 2(*b*), near the second bend, voids are clearly visible at this flow rate. The liquid occupies mostly the lower portion of the second bend. As the liquid stream impacts the wall at the bend, it becomes brighter because of the increase in turbulence which increases the amount of scattering of the incident light.

The video camera ran at a frame rate of 30 frames/second (fps), and a shutter opening of (1/125) s. Therefore, the successive video frames are approximately at (1/30) s apart. A frame-by-frame analysis of the video shows the following features for this flow rate. Liquid stays at the bottom for the most part, the flow becomes turbulent downstream of the bend indicated by the increased brightness of the liquid stream, and the transition point shows intermittency. The flow appears to be developing a swirl component. Region 3 is the bend upstream of the filler pipe exit and lies mostly in the horizontal plane. This region, characterized by the increased momentum of the liquid stream caused by its downward gravity-assisted motion transforms into a highly swirling turbulent flow as it impacts the wall at the nearly 90-deg bend. At this low flow rate the liquid stays mostly towards the wall

**Table 1 Filler pipe experimental case summary**

| $Q$ Lpm (gpm) | $\overline{V}$ ms$^{-1}$ | Re ($H_2O$) | $\Delta p$ ($H_2O$) Nm$^{-2}$ | Re($G$) | $\Delta p$ ($G$) Nm$^{-2}$ | $\Delta p$ ($G$), steel tube Nm$^{-2}$ |
|---|---|---|---|---|---|---|
| 3.78 (1) | 0.129 | 3,236 | 11.32 | 7,879 | 10.64 | 11.77 |
| 7.56 (2) | 0.257 | 7,457 | 44.75 | 12,850 | 29.32 | 31.28 |
| 11.34 (3) | 0.385 | 11,171 | 81.22 | 19,250 | 52.64 | 65.81 |
| 18.93 (5) | 0.643 | 16,181 | 339.10 | 39,397 | 205.30 | 241.80 |

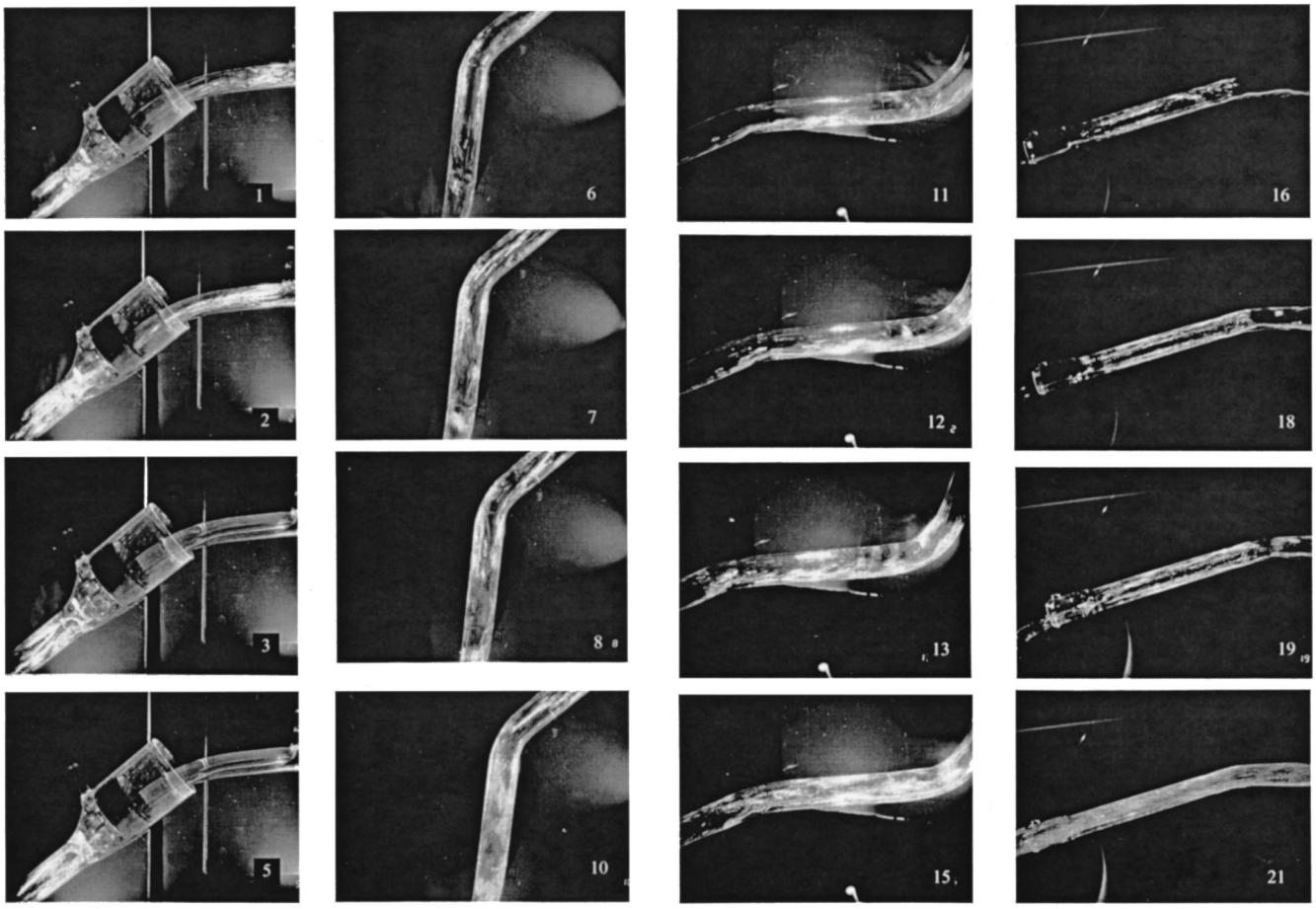



**Fig. 2** (*a*) **Flow visualization in Region 0 (left column) and Region 1 (right column). Flow rates (top to bottom): 3.78, 7.56, 11.34, and 18.93 lpm. (*b*) Flow visualization in Region 2 (left column) and Region 3 (right column). Flow rates (top to bottom): same as in (*a*).**

because of its high angular momentum, and the gas phase core is clearly visible when observed from upstream, looking down the axis of the straight exit segment of the tube.

**7.56 Liters/Minute.** The nozzle at 7.56 lpm is still not fully filled with liquid. Compared to the 3.78 lpm case, the liquid free surface in the nozzle is more irregular indicating increased turbulence. At this flow rate the liquid occupies more space, and the liquid surface in Region 0 is not as clearly demarcated as in the previous case. In Region 1, Frame 7, Fig. 2(*a*), the liquid surface moves from frame to frame in the video. In Region 2, Frame 12, Fig. 2(*b*), large-scale structures form more upstream and more frequently than the 3.78 lpm case. Downstream of the bend a large void of more or less of constant size is present. The free surface shows finger-like structures which are of smaller scale than those in the 3.78 lpm case. Most of the liquid stays at the bottom of the nearly horizontal downstream portion of the tube. Occasionally, the liquid stream breaks up into a stream of drops which are carried downstream by the gas phase. The process is highly intermittent; the detached liquid mass varies in size from small drops to large ligaments. There are also large frame-to-frame variations in the size of the detached liquid. That there is swirl developing in the flow is evident from the liquid occupying different azimuthal regions of the tube in successive frames of the video.

**11.34 Liters/Minute.** At 11.34 lpm the liquid fills the entire nozzle with little air entrained at the recirculation tube opening (Region 0, Frame 3, Fig. 2(*a*)). Transition from an air-water mixture flow to just water flow in the nozzle may be seen by comparing Frames 2 and 3 in Fig. 2(*a*). At this flow rate the flow structures are smaller in size. The void at the first bend, in Region 1, moves around from frame to frame. The liquid flowing down at higher speed now hits the wall at the bend and seems to be transitioning into turbulent flow. Liquid drops can be seen to break off from the liquid free surface.

In Region 2 the voids are much smaller and appear intermittently. The individual drops are now barely visible; instead, the gas-liquid mixture has become more homogeneous in appearance and causes scattering of bright diffused light from the flowing mixture. The edge structures are finer than those at lower flow rates. It may also be observed that there is less intermittency in the flow indicating the higher Reynolds number. Mixing of the two phases has progressed considerably indicated by the absence of large voids. Right at the bend, at the inner corner, an air pocket of nearly constant size is present.

**18.93 Liters/Minute.** At this high flow rate the flow at the first bend exhibits its turbulent nature in terms of the formation of turbulent structures and their chaotic motion as observed in the frame-to-frame variations of the flow structure. The flow is characterized by large drops and ligaments in Region 1. As the higher flow rates cause the liquid to fill the tube, there are no longer any large voids in the flow. Instead, several small voids not having well-defined boundaries develop. Because of the higher flow speed, the video images are blurred even at the top of the tube. Another interesting feature at this high flow rate is the stronger swirl component that the flow acquires right at the mouth of the

filler pipe, because of the slight misalignment between the nozzle spout axis and the filler pipe axis. Downstream of the first bend the liquid and air form a mixture because of the impact of the high momentum liquid stream with the wall which causes the liquid stream to break up into drops and ligaments, mix with the air stream, and develop a swirl component.

At the second bend the turbulent region has moved upstream and the void is present upstream of the second bend. Ligaments are seen to be breaking off from upstream of the second bend. At this location, the swirl component is well established as evident from the structures repeating themselves at the same location in about every fifth frame indicating a frequency of about 6 Hz. The swirl velocity close to the wall may be estimated from this frequency observation to be of the same order as the axial velocity. At this higher flow rates the voids are smaller and they form and convect downstream rapidly.

## General Observations

From viewing the video and still pictures of the flow over the range of ~4–18 lpm, the following general observations may be made. Except for low flow rates at the filler pipe entrance, the flow is turbulent. The flow transitions to high turbulence at the bends because of the impact and break up of the liquid stream. The flow also acquires a swirl component due to the asymmetric flow impacting the wall at the bends. At low flow rates there are large voids where the liquid free surfaces are clearly visible. At higher flow rates these voids become smaller and their boundaries become blurred. At still higher flow rates the voids are no longer visible, but a highly turbulent liquid-gas mixture fills the entire tube.

The flow losses in the filler pipe can be calculated by considering the losses in the straight segments and those in the bends. Fully developed, steady turbulent pipe flow assumptions were made for these calculations. Pipe friction coefficient and minor loss coefficients for the pipe bends were obtained from Munson et al. [4]. Table 1 gives pressure losses for the quartz model for 3.78 lpm (1 gpm), 7.56 lpm (2 gpm), 11.34 lpm (3 gpm), and 18.93 lpm (5 gpm) for water as well as gasoline. Also given are the pressure losses for gasoline in the production filler pipe. Between the inlet and the exit, there is a height difference of ~39 cm, which corresponds to 3820 $N/m^2$ for water. Therefore, the liquid accelerates as it flows down the tube because the momentum gain due to drop in height is much higher than the pipe losses. This is confirmed in the video which shows the liquid rushing down the fill tube from entrance to exit.

## Computational Fluid Dynamics (CFD) Simulation

Few successful CFD investigations of the refueling process have been conducted in the past. Sinha et al. [5] studied premature fuel shutoff by dividing the entire process into a number of unit processes. They used an implicit Roe/TVD upwind framework for the numerics. Simple straight tubes for filler pipe and vent pipes and a cylindrical geometry for fuel tank were used to model the fuel system. The unit processes included flow field near tank dome prior to vent pipe sealing, compression of trapped air/vapor, wave propagation in a uniform cylinder with piston, and the effect of contraction. The effect of coupling of the various subsystems was not attempted in their CFD simulation.

**CFD Simulation Scheme.** In the present work, CFD simulations of the filler pipe flow were conducted using the software packages GAMBIT for grid generation and FLUENT [6] for solution of the two phase flow equations, and the results have been documented. GAMBIT is a grid generation package suitable for complex three-dimensional geometries. It can be run interactively to generate structured as well as unstructured grids, and the resultant grids can be viewed and optimized for the geometry and the flow problem under consideration. Several parameters can be summoned while generating the grid to determine how well it has been optimized. Figure 3 shows a typical grid used in the present simulations. This hexagonal grid has 19,831 cells and 22,128 nodes. The grid density has been carefully adjusted to optimize grid distribution in regions of high flow gradients. Adaptive technique was used to check for independence of the solution. Based on the solution obtained from FLUENT, the grid was further refined, particularly at the gas/liquid interface such that the air entrainment at the filler neck mouth remains almost constant for two successive cases. The final number of grid cells and grid density varied for different flow rates because the gas/liquid interface area would change with flow rate. After grid adaptation, the number of grid cells ranged from approximately 36,000 for 7.56 lpm flow rate to about 96,000 for 44 lpm flow rate.

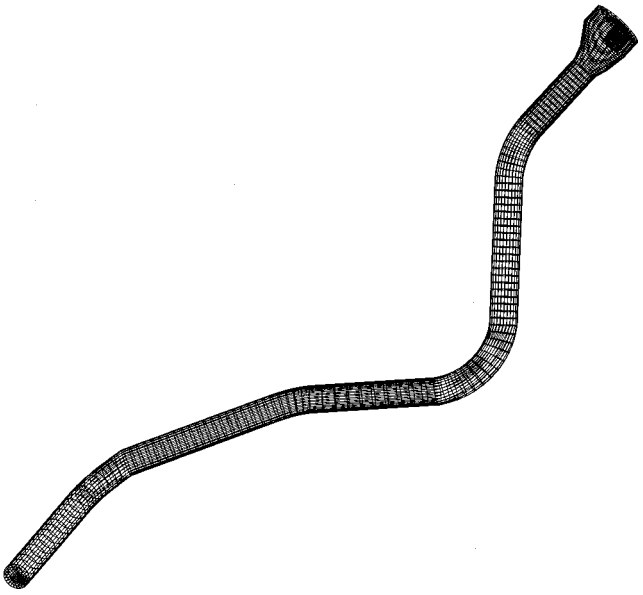


**Fig. 3 A typical grid**

Selected CFD simulation results of four flow rates, 7.56 lpm, 11.34 lpm, 34 lpm, and 44 lpm, are discussed in this work. The first two flow rates correspond to those used in the experiments. The last two fall in the range of gasoline dispense rates usually encountered at refueling stations.

The present two-phase flow problem comprising water or gasoline for the liquid phase and air-vapor mixture for the gas phase poses several challenges for the CFD practitioner. Several numerical models [7] are available for the resulting governing equations. The choice of a particular model often depends on several factors and is often guided by experience. The volume-of-fluid (VOF) model, the cavitation model and the algebraic slip model are three different options available in FLUENT [6] for multiphase flow modeling. After careful evaluation, the VOF model was chosen for the present simulations. The VOF model is designed for tracking the interface between immiscible fluids. A single set of momentum equations is shared by the fluids, and the volume fraction of each fluid in each computational cell in the domain is tracked during the solution. We can write the continuity equation for the volume fraction of the phase as

$$\frac{\partial \alpha_q}{\partial t} + u_i \frac{\partial \alpha_q}{\partial x_i} = S_{\alpha q}. \tag{1}$$

The source term, $S_{\alpha q}$, on the right-hand side is zero for the VOF model signifying that there is no mass transfer across the interface. The properties such as density and viscosity appearing in the transport equations are computed using the following general equation:

$$\phi = \alpha_2 \phi_2 + (1 - \alpha_2)\phi_1 \tag{2}$$

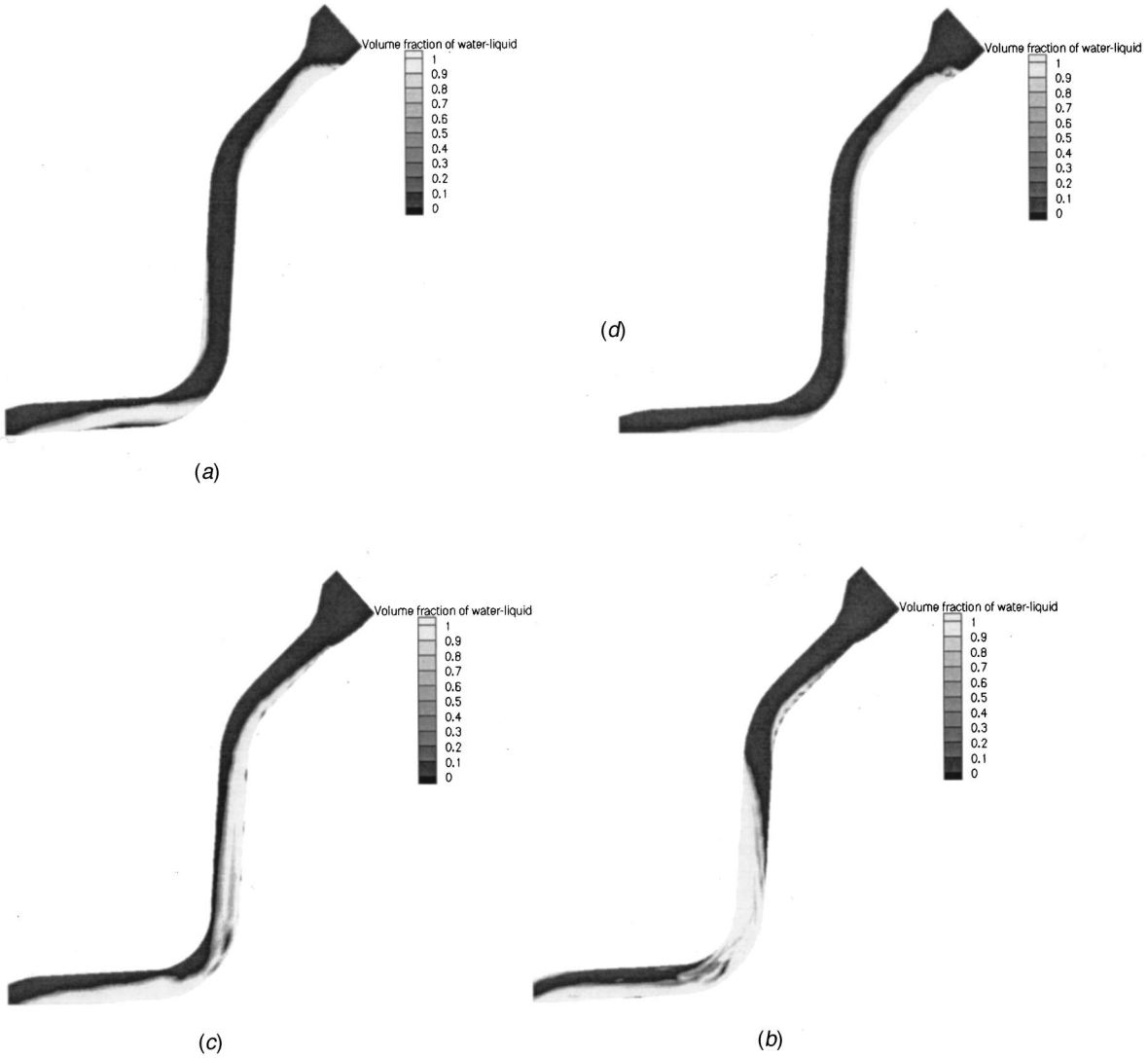



**Fig. 4 Phase distribution. Flow rates: (*a*) 7.56 lpm, (*b*) 11.34 lpm, (*c*) 34 lpm, (*d*) 44 lpm.**

where $\phi_1$ is the primary phase property. A single momentum equation is solved throughout the domain, and the phases share the resulting velocity field. The momentum equation given below is dependent on the volume fraction through the density $\rho$ and the viscosity $\mu$.

$$\frac{\partial}{\partial t}\rho u_j + \frac{\partial}{\partial x_i}\rho u_i u_j = -\frac{\partial P}{\partial x_j} + \frac{\partial}{\partial x_i}\mu\left(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial u_i}\right) + \rho g_j + F_j \tag{3}$$

One of the important decisions to be made in a CFD analysis is the choice of the turbulence model to be used. Several turbulence models are available in FLUENT [6] such as the $k$-$\varepsilon$ model and its variants like RNG–$k$-$\varepsilon$; Reynolds stress model (RSM); and large eddy simulation (LES) model. The $k$-$\varepsilon$ model by Launder and Spalding [8] is probably the most widely used turbulence model. The more recent RNG–$k$-$\varepsilon$ model of Yakhot and Orszag [9] has improved accuracy for rapidly strained flow. It is also better in predicting flow with swirl. The constants used in the RNG–$k$-$\varepsilon$ model are obtained by using the renormalization-group theory (RNG) from statistical principles applied to isotropic turbulence. A disadvantage of any turbulence model based on Boussinesq hypothesis like the standard $k$-$\varepsilon$ and RNG–$k$-$\varepsilon$ models is that the turbulent viscosity is assumed to be isotropic, which is not strictly true. An alternative approach, as embodied in the RSM model of Launder, Reece, and Rodi [10] and Gibson and Launder [11] is to solve individually all the terms of the Reynolds stress tensor. An additional scale-determining equation (normally for $\varepsilon$) is also required. Thus, seven additional equations are required to be solved for three-dimensional flows in this model. The computational time increases due to the additional equations. In the present study, $k$-$\varepsilon$, RNG–$k$-$\varepsilon$, and RSM models were tested to determine the optimum turbulence model. LES was not used because FLUENT [6] does not support this model for multiphase flow problems. It was found that RNG–$k$-$\varepsilon$ model performed better compared to the standard $k$-$\varepsilon$ model in terms of overall agreement with the flow features observed in the experiments. Air entrainment rates obtained from RNG–$k$-$\varepsilon$ and RSM models were compared for two flow rates. Air entrainment differed by about 2 percent for 7.56 lpm and about 5 percent for 34 lpm. Thus, the additional computational time required for RSM model was not justified for all the cases considered in this work. The results discussed in this paper are, therefore, based on the RNG–$k$-$\varepsilon$ model.

As swirl is introduced at the out-of-plane bends, numerical diffusion may increase because the flow may not be aligned along the grid. A second-order discretization scheme was tested for the flow rates of 7.56 lpm and 34 lpm. It was found that the difference in air entrainment at the filler neck mouth was about 2.5 percent for 7.56 lpm case and about 8 percent for 34 lpm case. The

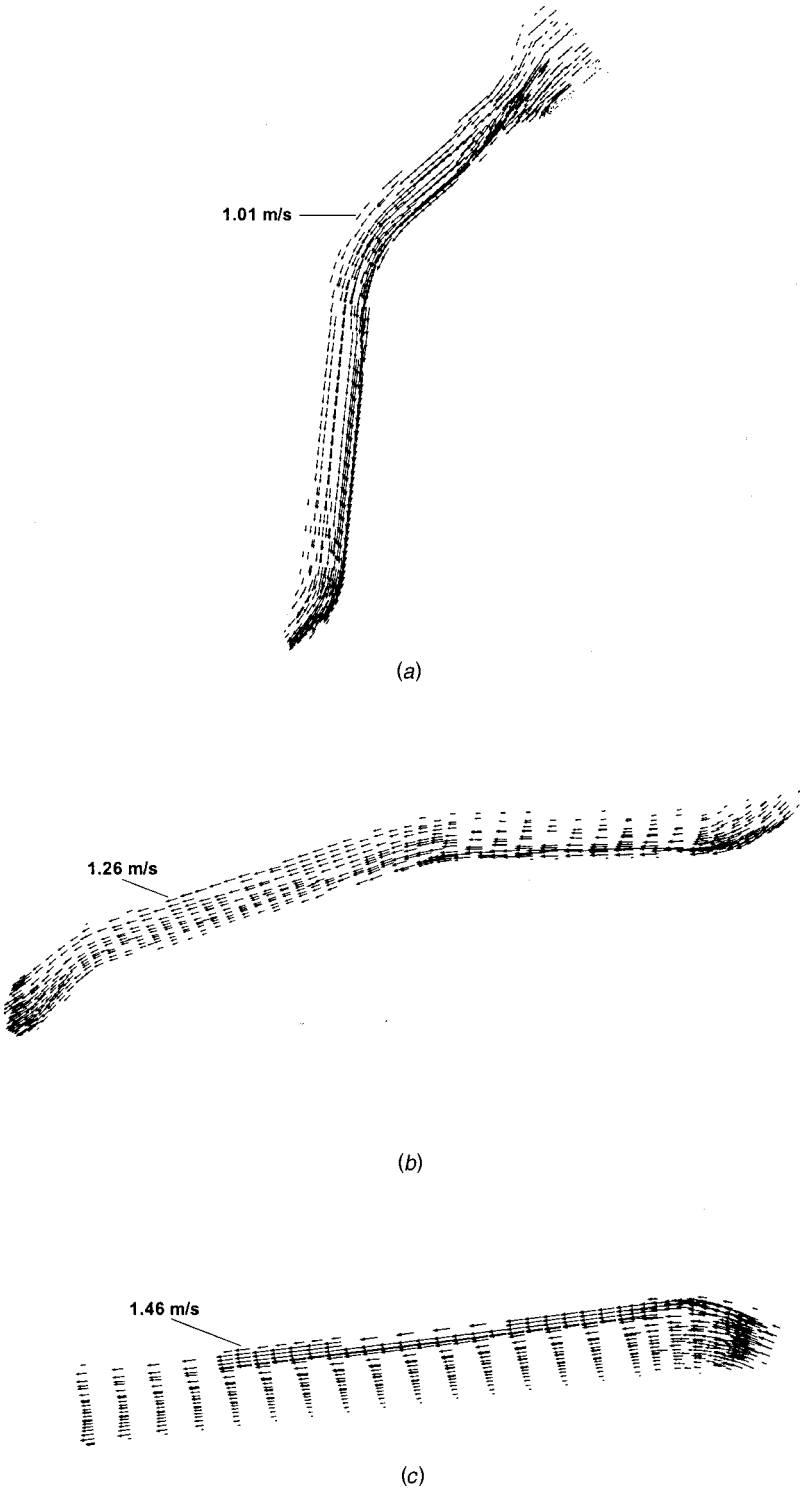



**Fig. 5 Velocity vectors. (*a*) Region 1, (*b*) Region 2, and (*c*) Region 3. Flow rate: 11.34 lpm.**

second-order scheme showed slower convergence, thereby increasing the computational time. As the difference in air entrainment rates between first-order and second-order discretization is within acceptable limits, first-order scheme was used for all the cases presented in this work.

## Simulation Results

CFD simulations for 7.56, 11.34, 34, and 44 lpm cases were conducted and the results analyzed. The simulations were started with the grid shown in Fig. 3, which was further refined by the grid adaptation technique described above. Though not clearly visible, the computational domain consisted of the cylindrical fuel nozzle inserted into the filler neck and terminated where the filler pipe begins to contract in cross-sectional area. The filler neck diameter was 54 mm and the nozzle diameter was 17 mm. The nozzle tip was 19 mm inside the filler neck entrance and the nozzle axis was parallel to the filler neck axis, but offset by 10 mm below the filler neck axis. The velocity profile at the nozzle

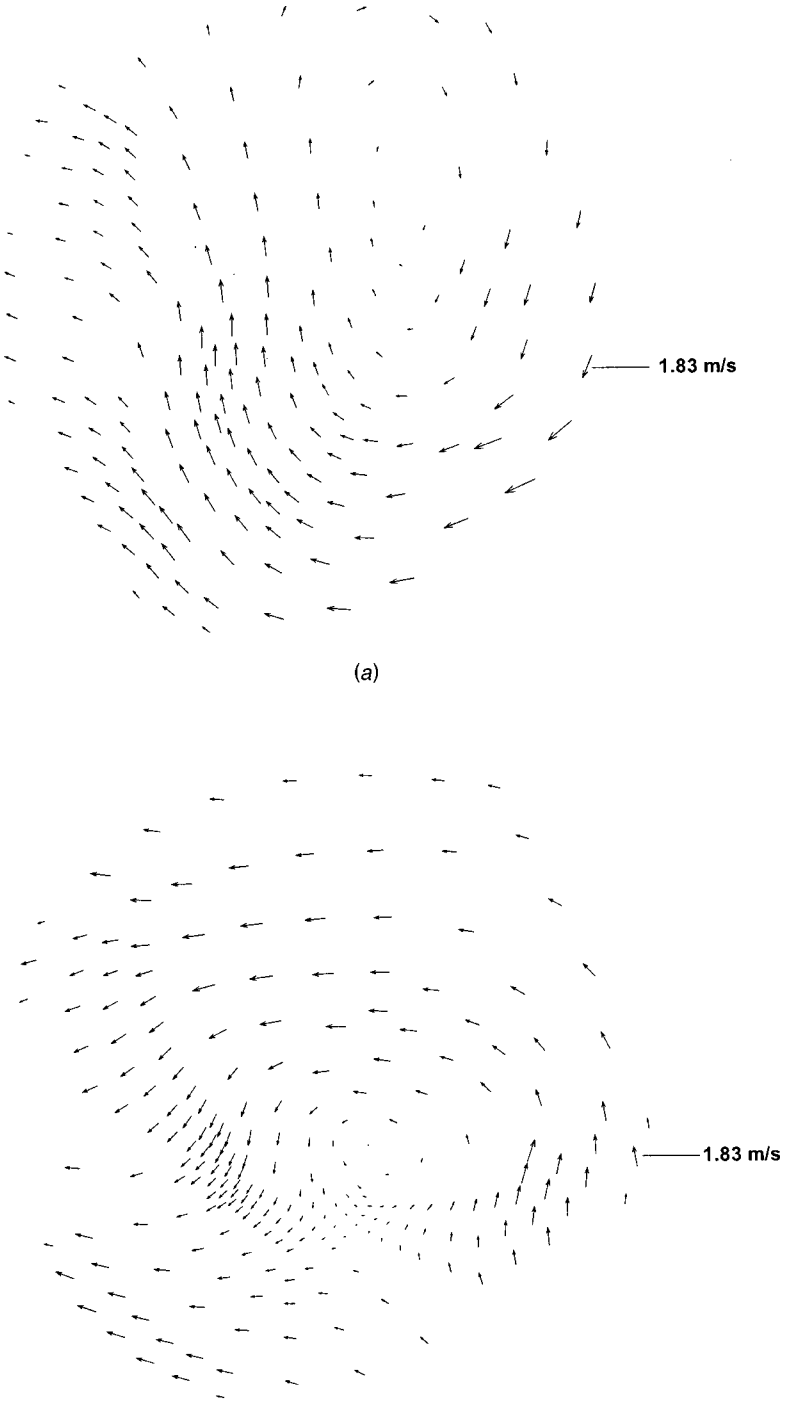



**Fig. 6 Velocity vectors at a cross section after Bend 2. Flow rates: (*a*) 7.56 lpm, (*b*) 34 lpm.**

inlet was uniform and isotropic turbulence of 10 percent was assumed at the nozzle inlet. Ambient pressure was used for inlet and outlet boundary conditions. An unsteady solver with a time step of $5\times10^{-4}$ s was used when the number of grid cells was approximately 20,000. It was further reduced to $1\times10^{-4}$ s after grid refinement.

Phase distributions were analyzed for different flow rates. The simulation results from the different flow rates had many similarities with the flow features observed in the experiments. It was found that swirl was introduced after Bend 2, and further enhanced after Bend 4. The void fraction reduced with increase in flow rate as the liquid started filling more of the cross-section. The simulations were stopped from time to time to analyze the results. As observed in the experiments, there were fluctuations in the flow features on a time-to-time basis. The turbulence level increased due to the presence of the bends. The pressure contours showed that there was a high pressure spot on the lower portion of Bend 2, and it increased in magnitude with flow rate. This high

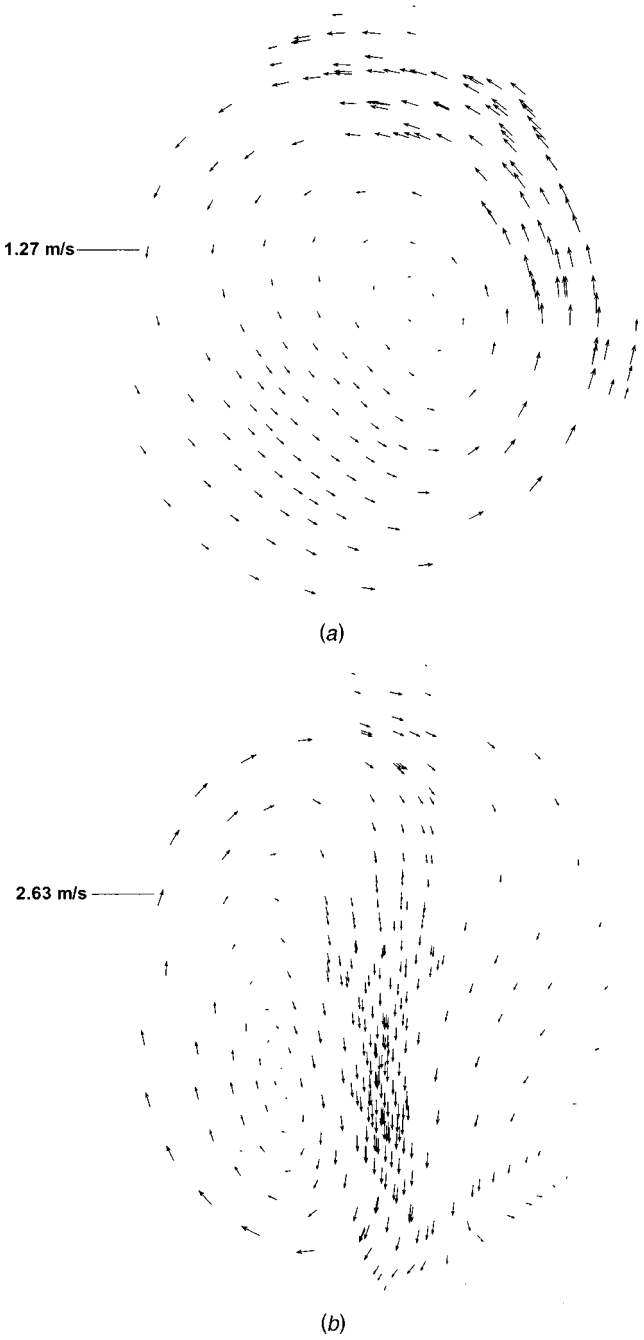



**Fig. 7 Velocity vectors at a cross section after Bend 4. Flow rates: (*a*) 7.56 lpm, (*b*) 34 lpm.**

pressure spot is due to the impact of the liquid after an almost vertical fall in the section downstream of Bend 1. A low pressure area was also noticed at the lower portion of Bend 1. The pressure in this zone decreased with increase in flow rate.

Figure 4 shows the phase distribution along the filler pipe wall in Region 1 for the four flow rates cited above. In these grayscale plots, the light regions indicate liquid phase and the dark regions indicate gas phase. As the flow rate increases, the liquid tends to fill the entire cross section of the filler pipe upstream of the second bend. In the almost vertical segment after the first bend the liquid accelerates due to gravity. However, the segment after the second bend is almost horizontal where the flow slows down. This has been confirmed by determining the cross-section average velocities just before and after the second bend for all the flow rates. This slowing down dictates the maximum flow rate that can be dispensed through the filler pipe.

**Table 2 Computational Fluid Dynamics (CFD) Simulation Results**

| $Q$ Lpm (gpm) | $V_{gp}$ m/s | $A_{gp}$ m$^2$ | $\dot{m}_{\text{air}}$ kg/min |
|---|---|---|---|
| 7.56 (2) | 1.28 | $4.77\times10^{-4}$ | 0.0449 |
| 11.34 (3) | 1.11 | $4.49\times10^{-4}$ | 0.0367 |
| 34 (8.98) | 2.18 | $3.06\times10^{-4}$ | 0.0490 |
| 44 (11.62) | 2.23 | $2.60\times10^{-4}$ | 0.0427 |

The velocity vectors along axial sections of Regions 1, 2, and 3 are shown in Fig. 5. The vector length corresponds to the velocity magnitude. As can be seen in the vector plots, air is entrained by the liquid because of the momentum exchange between the liquid and the gas phases. The mean velocity in the gas phase increases with increase in liquid flow rate.

Figures 6 and 7 show the velocity vectors for the flow rates of 7.56 lpm and 34 lpm at cross sections downstream of Bends 2 and 4, respectively. Both are out-of-plane bends and therefore strong swirl is introduced in the flow. In these two plots, the regions with smaller vectors indicate gas phase and those with larger vectors indicate liquid phase. For Bend 2, the gas phase vortex is closer to the wall for the lower flow rate of 7.56 lpm, but it moves more towards the center at 34 lpm. At the higher flow rate, a stagnation region can be clearly seen near the axis. Figure 7(*a*) shows that the swirl strength at Bend 4 is much higher compared to that at Bend 2. At this cross section, past Bend 4, the liquid stays mostly at the top, and the gas phase forms a large vortex below. At 34 lpm (Fig. 7(*b*)), the vortex pattern is more complex; the vortex is not as clearly defined as in the lower flow rate case. Note that some of these features, especially their azimuthal dependence, would be different at different axial locations.

To determine gasoline vapor emission during refueling it is necessary to first determine the variation of air entrainment rate with liquid flow rate ($Q$). Therefore, air entrainment rates were determined for four liquid flow rates and the results are summarized in Table 2. The average gas phase velocity ($v_{gp}$) and void area ($A_{gp}$), determined at a cross-section downstream of the second bend, are also given in Table 2. As can be seen, there is a steady increase in the gas phase velocity with increase in flow rate; however, the void area decreases with increase in flow rate. The product of the cross-section averaged velocity and the void area equals the volume flow rate of the air entrained. These results indicate that air entrainment cannot be correlated to a single parameter such as the Reynolds number, but may depend on several parameters.

## Conclusions

The present investigation of automotive filler pipe flow comprising both experiments and CFD simulations has been successful in understanding this complex two-phase flow. Many interesting features of the flow are uncovered, potentially leading to a better understanding of the processes of practical importance such as air entrainment and vapor emissions. The rate of air entrainment indicates an increasing trend with increasing liquid flow rate, but it is not monotonic. There is maximum flow rate that can be dispensed through a filler pipe. Beyond that maximum flow rate, well-back, liquid backing up in the filler pipe, can be expected. Out-of-plane bends introduce strong swirl in the flow which enhances air entrainment, sometimes referred to as the ''dynamic seal.'' Trends in the velocity vector plots suggest that lower fill rates may not be the best strategy for reducing emissions, but higher fill rates coupled with prevention of spit back during nozzle shut off may be more beneficial. The present work shows the key role that CFD would play in future fuel system design and analy-

sis. With the use of more powerful computers coming to market at a fast pace, additional components such as the fuel tank and the recirculation tube could be incorporated in the numerical model, obviously bringing down development cost and cycle time. Rapidly evolving physical models and numerical algorithms would yield useful parameters such as heat and mass transfer, emissions, and gas phase composition.

## Acknowledgment

This work is supported by a grant from Dayco Inc./MarkIV Automotive. We thank Joey Edson and Darrel Klein for assistance with the experiments. We also thank the anonymous reviewers for their many valuable suggestions.

## References

[1] SAE ORVR Task Force Meeting Minutes, Apr. 1998.
[2] Stoneman, S., 1997, ''On the Design of Automotive Fuel Filler Pipes,'' Automotive Engineer, pp. 32–36.
[3] Lockhart, M. C., Thompson, R. H., and Baldus, S. L., 1997, ''Methodology for Evaluating Fuel Dispenser Nozzle Characteristics,'' SAE Paper 970311, SAE International Congress & Exposition, Detroit, MI.
[4] Munson, B. R., Young, D. F., and Okiishi, T. H., 1998, *Fundamentals of Fluid Mechanics*, 3rd Ed., John Wiley and Sons, New York.
[5] Sinha, N., Thompson, R., and Harrigan, M., 1998, ''Computational Simulation of Fuel Shut-off During Refueling,'' SAE Paper 981377.
[6] FLUENT 5 User's Guide, 1998, Vol. 1–5, Fluent Inc.
[7] Hirt, C. W., and Nicholls, B. D., 1981, ''Volume of Fluid (VOF) Method for the Dynamics of Free Boundaries,'' J. Comput. Phys. **39**, pp. 201–225.
[8] Launder, B. E., and Spalding, D. B., 1972, *Lectures in Mathematical Models of Turbulence*, Academic Press, London.
[9] Yakhot, V., and Orszag, S. A., 1986, ''Renormalization Group Analysis of Turbulence: I. Basic Theory,'' J. Sci. Comput. **1**, No. 1, pp. 1–51.
[10] Launder, B. E., Reece, G. J., and Rodi, W., 1975, ''Progress in the Development of a Reynolds-Stress Turbulence Closure,'' J. Fluid Mech. **68**, No. 3, pp. 537–566.
[11] Gibson, M. M., and Launder, B. E., 1978, ''Ground Effects on Pressure Fluctuations in the Atmosphere Boundary Layer,'' J. Fluid Mech. **86**, pp. 491–511.

**T. Shamim**
Assistant Professor
e-mail: shamim@umich.edu

**H. Shen**

**S. Sengupta**

Department of Mechanical Engineering,
The University of Michigan,
Dearborn, MI 48128-1491

**S. Son**

**A. A. Adamczyk**

Chemical Engineering Department,
Ford Scientific Research Laboratory,
Ford Motor Company,
Dearborn, MI 48121

# A Comprehensive Model to Predict Three-Way Catalytic Converter Performance

*This paper describes the development of a comprehensive mathematical and numerical model for simulating the performance of automotive three-way catalytic converters, which are employed to reduce engine exhaust emissions. The model simulates the emission system behavior by using an exhaust system heat conservation and catalyst chemical kinetic submodel. The resulting governing equations are solved numerically. Good agreements were found between the numerical predictions and experimental measurements under both steady-state and transient conditions. The developed model will be used to facilitate the converter design improvement efforts, which are necessary in order to meet the increasingly stricter emission requirements.* [DOI: 10.1115/1.1424295]

## Introduction

Since 1968 when legislation first introduced emission regulations in the U.S., there has been a steady trend of lowering limitations on vehicle exhaust emissions. For example, the Inter Industry Emission Control Program (IIEC) in 1968 required carbon monoxide (CO) limitation to be 7.19 g/mile, hydrocarbon (HC) to be 0.82 g/mile, and no limitation to oxides of nitrogen ($NO_x$) ([1]). The present effective standard that was modified in 1994 requires the CO limit to be 3.4 g/mile, HC to 0.25 g/mile, and $NO_x$ to 0.4 g/mile for passenger cars ([2]). This trend of stricter regulation will continue with proposed standards for California as low as or lower than 0.01 g/mile NMOG (non-methane organic gas) and 0.02 g/mile $NO_x$. Worldwide, similar trends are occurring. Hence, engine exhaust systems need to undergo continuous modification to meet increasingly stricter regulations. In the past, much of the design and engineering process to optimize various components of engine and emission systems has involved the prototype testing. The complexity of modern systems and the resulting flow dynamics and thermal and chemical mechanisms have increased the difficulty in assessing and optimizing system operation. Due to overall complexity and increased costs associated with these factors, modeling continues to be pursued as a method of obtaining valuable information supporting the design and development process associated with the exhaust emission system optimization.

Mathematical models of catalytic converter have been employed for nearly 30 years. There are different approaches to simulate the phenomena of exhaust systems. In some studies, models of transient heat transfer have been used to assess the temperature distribution along the exhaust systems as dimensions, materials and insulating properties are modified to obtain a workable system ([3–5]). These models are used to define the state of the inlet gases into the catalyst as their positions change along the exhaust system. They are not, however, concerned with the complicated heat transfer and detail chemical reaction mechanism inside the catalytic converters. In other studies, models of flow distribution are applied to predict the nonuniformity of flow at the catalyst entrance and to predict pressure losses in the emission system ([6–8]). These parameters have been found to influence the converter performance. However, these studies emphasize only on the flow patterns in the diffusion section of the catalytic converters, and do not include the full flow and kinetic processes in the total length of the converter. Detailed models of catalyst processes have been used to assess catalyst functions ([9–17]). These models can, in principle, be used to support the understanding of the catalyst process. Many either report steady-state kinetics ([13]) or report accumulated conversion efficiencies during different driving test cycles ([14]). Some studies focused on investigating the light-off behavior of converters ([18–19]).

Most of the previous studies did not compare their predictions with experiments under realistic transient driving conditions. The model testing under realistic conditions is very important since the catalyst behavior under steady-state conditions is very different from that under transient conditions. Pattas et al. [14] used a simple five-step reaction scheme and a one-dimensional model to predict the catalytic converter performance during a regulated European driving cycle, ECE-15. A unique feature of their work was that it presented comparisons of numerical predictions and experimental measurements of both time-resolved and accumulated emissions under real engine cycle operations. Their accumulated emission results showed good agreement with the experimental measurements. However, they did not report comparison of time-resolved emissions for HC and NO. Recently, Koltsakis et al. [16,17] presented a model featuring an extended reaction scheme and an improved oxygen storage submodel for highly transient catalyst operation. The predictions and measurements were compared for accumulated emissions under the new European driving cycle (NEDC). However, rather than using a set of fully transient equations, they computed the transient behavior of the catalytic converter as a series of quasi-steady states.

A major difficulty in modeling of catalytic converters is the lack of adequate kinetics data for different type of catalysts. Furthermore, the effect of thermal degradation, which greatly influences the catalyst operation, is largely unknown and is not incorporated in the kinetic data. Most studies mentioned above relied on the classical work of Voltz et al. [20] in order to simulate chemical reactions and to calculate the reaction rates. This work involved the measurement of pellet-type Pt catalyst performances and derived kinetic reaction rate expressions for the oxidation of CO and $C_3H_6$ under lean conditions. The expressions are of the Langmuir-Hinshelwood type and take into account the inhibition due to NO.

Contributed by the Heat Transfer Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the HT Division, Dec. 1999; final revision received by the ASME Headquarters May 2001. Editor: H. D. Nelson.

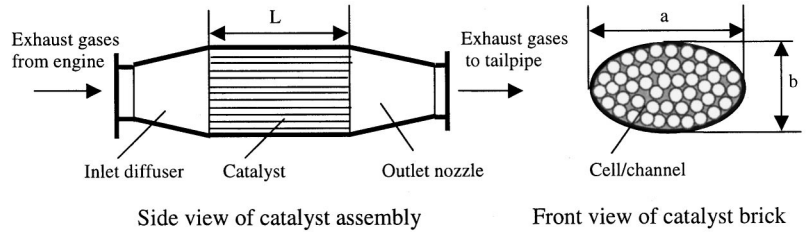



**Fig. 1 Schematic of a catalytic converter**

These rate expressions with some modifications have been used for non-Pt catalysts due to lack of data for other type of catalysts. Montreuil et al. [13] presented a methodology for updating steady-state kinetic data for Pt/Rh and Pd/Rh catalysts. Their work included a detailed reaction scheme consisting of 13 reactions and derived corresponding reaction rate expressions involving 97 independent parameters. They listed kinetic constants for four different types of catalysts. These kinetic rate expressions are to date the most detailed and are used in the present study.

The present study employs a transient heat transfer model, coupled with a transient catalyst chemical kinetic mechanism that includes oxygen storage chemistry to simulate the catalyst performance during the U.S. and the European driving test cycles. This model can be used to better understand the physical and chemical processes of gas flowing through the catalytic converter. It can also be used to support the detailed optimization of geometry and strategic control of the entire vehicle emission system.

## Model Description

**Governing Equations.** Figure 1 shows schematic of a catalytic converter. The governing equations were developed by considering the conservation of mass, energy and chemical species. The following phenomena are explicitly included in the model: (i) convective heat and mass transfer from the exhaust gas to the catalytic surface; (ii) heterogeneous chemical reactions taking place on the catalytic surface; (iii) the catalyst's capacity of storing extra oxygen under lean conditions and releasing it under rich conditions; (iv) heat losses to the surroundings; and (v) heat conduction along the catalytic converter.

The nonuniform flow distribution at the monolith's face is neglected to simplify the mathematical model and for computational expediency. This simplification results in a one-dimensional model. The governing conservation equations, which are similar to those of Montreuil et al. [13] and Kuo et al. [21], are listed below:

The gas phase energy equation:

$$\rho_g c_{P_g}\left(\varepsilon\frac{\partial T_g}{\partial t}+v_g\frac{\partial T_g}{\partial z}\right)=-h_g G_a(T_g-T_s) \tag{1}$$

The gas phase species equations:

$$\left(\varepsilon\frac{\partial C_g^j}{\partial t}+v_g\frac{\partial C_g^j}{\partial z}\right)=-k_m^j G_a(C_g^j-C_s^j) \tag{2}$$

where the superscript $j$ varies from 1 to 7 representing, respectively, the following gas species: CO, NO, $NH_3$, $O_2$, $C_3H_6$, $H_2$, and $C_3H_8$. The effect of homogeneous reactions in gas phase is very small and is, therefore, neglected.

The surface energy equation:

$$(1-\varepsilon)\rho_s c_{Ps}\frac{\partial T_s}{\partial t}=(1-\varepsilon)\lambda_s\frac{\partial^2 T_s}{\partial z^2}+h_g G_a(T_g-T_s)-h_\infty S_{\text{ext}}(T_s-T_\infty)+G_a\sum_{k=1}^{n_{\text{reaction}}}R^k(T_s,C_s^1,\ldots\ldots,C_s^{n_{\text{species}}})\cdot\Delta H^k \tag{3}$$

The surface species equations:

$$(1-\varepsilon)\frac{\partial C_s^j}{\partial t}=k_m^j G_a(C_g^j-C_s^j)-G_a R^j(T_s,C_s^1,\ldots\ldots,C_c^{N_{\text{species}}}) \tag{4}$$

where superscript $j$ varies from 1 to 7, representing the surface species in the same order as the gas species. The heat and mass transfer coefficients ($h_g$ and $k_m^j$) in the above equations are calculated from the conventional correlations of Nusselt and Sherwood numbers ([13]). The conservation equation for the surface oxygen storage mechanism can be represented by Eq. (4) excluding the convective mass transport term.

## Reaction Scheme

The heterogeneous surface reactions, which occur in the substrate of the catalyst, have been documented in several references. Kuo et al. [21] set up a model of CO and HC oxidation within a catalyst as part of the IIEC program. Pattas et al. [14] showed that a simplified reaction scheme may be used to simulate the catalyst behavior with a reasonable accuracy. Their mechanism consisted of a five-step reaction scheme comprising CO, $H_2$, and HC oxidation, and NO reduction by CO. Another kinetic mechanism by Otto and LeGray [11] was modified by Otto [22] to include a more elaborate $NO_x$ chemistry mechanism and the formation of ammonia and its effects on NO oxidation under fuel-rich conditions. This mechanism was later further modified by Montreuil et al. [13] and Li et al. [23]. This general group of chemical reactions, which represents the so far most inclusive reaction scheme ([24]), is used in the present study in order to simulate the kinetic processes in the catalytic converters. This mechanism, listed below, consists of 13 independent forward pathways for oxidation of CO, $H_2$, $C_3H_6$, $C_3H_8$, and $NH_3$ with $O_2$ and NO as oxidizing agents, and their corresponding rich and lean kinetic rate expressions. The effect of sulfur on the catalyst is not included in this mechanism.

$$\text{CO}+\frac{1}{2}\text{O}_2\rightarrow\text{CO}_2 \tag{$R$-1}$$

$$\text{CH}_\alpha+M_1\cdot\text{O}_2\rightarrow\text{CO}_2+\frac{\alpha}{2}\cdot\text{H}_2\text{O}\ \ (\text{HC}_F) \tag{$R$-2}$$

$$\text{CH}_\alpha+M_2\cdot\text{O}_2\rightarrow\text{CO}+\frac{\alpha}{2}\cdot\text{H}_2\text{O}\ \ (\text{HC}_F) \tag{$R$-3}$$

$$H_2+\frac{1}{2}O_2\rightarrow H_2O \qquad (R\text{-}4)$$

$$NH_3+\frac{3}{4}O_2\rightarrow 1.5H_2O+\frac{1}{2}N_2 \qquad (R\text{-}5)$$

$$CO+NO\rightarrow CO_2+\frac{1}{2}N_2 \qquad (R\text{-}6)$$

$$2.5CO+NO+1.5H_2O\rightarrow NH_3+2.5CO_2+\frac{1}{2}N_2 \qquad (R\text{-}7)$$

$$H_2+NO\rightarrow H_2O+\frac{1}{2}N_2 \qquad (R\text{-}8)$$

$$2.5H_2+NO\rightarrow NH_3+H_2O \qquad (R\text{-}9)$$

$$CH_\alpha+M_1\cdot O_2\rightarrow CO_2+\frac{\alpha}{2}\cdot H_2O \quad (HC_S) \qquad (R\text{-}10)$$

$$\frac{1}{2M_2}CH_\alpha+NO\rightarrow \frac{1}{2M_2}CO+\frac{\alpha}{4M_2}H_2O+\frac{1}{2}N_2 \quad (HC_F) \qquad (R\text{-}11)$$

$$CH_\alpha+M_3\cdot NO\rightarrow CO_2+\frac{\alpha}{2}\cdot H_2O+\frac{M_3}{2}\cdot N_2 \quad (HC_F) \qquad (R\text{-}12)$$

$$\frac{2.5}{2M_2}CH_\alpha+NO+\frac{(3-\alpha)}{4M_2}H_2O\rightarrow NH_3+\frac{2.5}{2M_2}CO \quad (HC_F) \qquad (R\text{-}13)$$

where $\alpha$=hydrogen-to-carbon ratio; $HC_F$=fast burning HC; $HC_S$=slow burning HC; and $M_1=[1+\alpha/4]$; $M_2=[1/2+\alpha/4]$; $M_3=[2+\alpha/2]$. In the present study, propylene and propane represent the fast and slow burning hydrocarbons, respectively. The corresponding reaction rate expressions are listed in the Appendix. For different catalyst formulations, the coefficients in the reaction rate expressions are different. They also vary with aging of catalytic converters. The present study uses the kinetic rate constants for an early palladium-based catalyst. The corresponding coefficients were taken from Montreuil et al. [13] wherein they were appropriately adjusted using experimental flow reactor measurements.

## Oxygen Storage Mechanism

A typical closed-loop control fuel supply system causes the $A/F$ to fluctuate rapidly about the stoichoimetrically balanced composition with a frequency of nearly 1 Hz. The conversion efficiency of a three-way catalytic converter can be improved by storing the extra oxygen under fuel lean conditions and releasing it under rich conditions ([25]). The released oxygen may participate in the reactions with the reducing agents, thereby increasing the conversion of CO and HC in a rich exhaust-gas environment ([22–25]). Such an oxygen storage capacity (OSC) is developed in the modern catalyst by coating its substrate with a wash-coat material containing ceria. The OSC is recognized as an important mechanism affecting catalyst behavior during vehicle acceleration and deceleration.

In addition to the pathways, which specify the kinetics over the noble metal sites, an additional kinetic mechanism is required to represent the OSC. In the present study, the OSC is simulated by designating two kinds of sites that can be oxidized and reduced through a nine-step site reaction mechanism ([22]). The metal reduction site on the surface is defined as $\langle s\rangle$ and the oxidized site is defined as $\langle os\rangle$. This oxygen storage mechanism of Otto [22], listed below, is used in the present study.

$$\langle S\rangle+\frac{1}{2}O_2\rightarrow\langle OS\rangle \qquad \text{site oxidation (OSC-1)}$$

$$\langle OS\rangle+CO\rightarrow\langle S\rangle+CO_2 \qquad \text{site reduction by CO (OSC-2)}$$

$$H_2O+CO\rightarrow H_2+CO_2 \qquad \text{water-gas shift (OSC-3)}$$

$$\langle OS\rangle+H_2\rightarrow\langle S\rangle+H_2O \qquad \text{site reduction by } H_2 \text{ (OSC-4)}$$

$$CH_\alpha+H_2O\rightarrow CO+\left(1+\frac{\alpha}{2}\right)\cdot H_2 \qquad \text{steam reforming (OSC-5)}$$

$$\frac{3}{2}CH_\alpha+\langle OS\rangle\rightarrow\langle S\rangle+\frac{1}{2}C+CO+\left(\frac{3}{4}\alpha\right)H_2$$

reduction by HC (OSC-6)

$$C+O_2\rightarrow CO_2 \qquad \text{coke burnoff (OSC-7)}$$

$$\langle S\rangle+NO\rightarrow\langle OS\rangle+\frac{1}{2}N_2 \qquad \text{NO storage (OSC-8)}$$

$$\langle OS\rangle+\frac{2}{5}NH_3\rightarrow\langle S\rangle+\frac{2}{5}NO+\frac{3}{5}H_2O$$

$NH_3$ site reduction (OSC-9)

where $\alpha$ is hydrogen-to-carbon ratio of the hydrocarbon.

Each reaction has two rate expressions, one being the fast site rate expression and the other the slow site rate expression with each being associated with the detailed materials in the catalyst washcoat. The total sites are conserved for both fast site and slow site. Therefore,

$$S_{\text{total},f}=\langle S\rangle_f+\langle OS\rangle_f \qquad (5)$$

$$S_{\text{total},s}=\langle S\rangle_s+\langle OS\rangle_s\,. \qquad (6)$$

The rates of the transient reactions are of the form

$$R_{\text{transient}}=\frac{(\text{OXSW})\cdot\text{CTR}\cdot e^{-E/RT_s}\cdot\Pi_{i=1}^{N_{\text{specie}}}(X_s^i)^{ex(i)}}{1+\Sigma_{n=1}^{N_{\text{specie}}}K_n\cdot X_s^n} \qquad (7)$$

where OXSW is an oxidation switch and equals 1 if redox ratio $<1$, and 0 if redox ratio $>1$. The redox ratio is defined as

$$\text{Redox Ratio}=\frac{[CO]+[H_2]+6\cdot\left(1+\frac{\alpha}{4}\right)[HC]}{[NO]+2\cdot[O_2]}. \qquad (8)$$

The corresponding coefficients for calculating the transient reaction rates are based on experimental data ([23]).

## Solution Procedure

The governing equations were discretized by using a nonuniform grid and employing the control volume approach and the central implicit difference scheme in the spatial direction. Since the gradients of all properties are greater near the catalyst inlet, smaller grid spacing was used near the inlet and larger spacing near the exit. A standard tridiagonal matrix algorithm with a successive line under relaxation method was used in the present study to solve the finite difference equations.

The inlet boundary conditions for the gas phase energy and species equations are of the first-kind Dirchlet type. These are determined either by experiment or through prior computation using models that calculate the exhaust properties from an engine operating under transient conditions. These data include gas temperature, mole concentrations of CO, NO, $NH_3$, $O_2$, $C_3H_6$, $H_2$, and $C_3H_8$, and total flow rate. In the present study, the available measured data are gas temperature, mole fractions of CO, unburned HC, NO, air-fuel ratio of the fresh mixture, and volume flow rate of the gases. During reactor experiments to determine kinetic rates, all hydrocarbon species are lumped into three general categories: fast burning, slow burning, and inert; and are char-

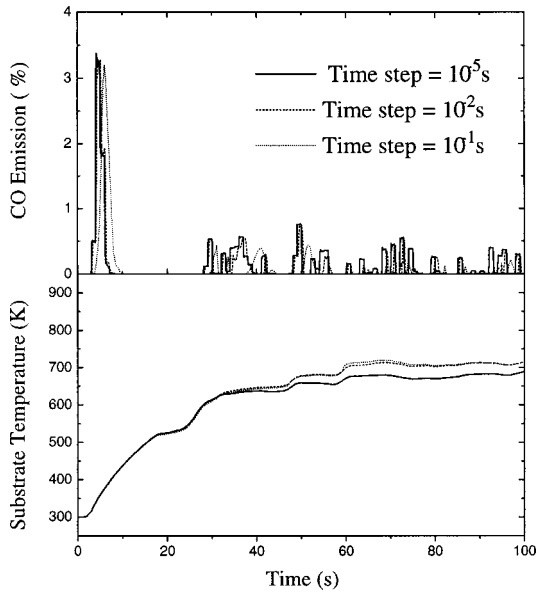



**Fig. 2 Effect of time step on transient CO emission and substrate temperature**

acterized by propylene, propane, and methane, respectively. The ratios of these three species' concentrations are determined experimentally and are specified in the model from total unburned hydrocarbons and correlation of these ratios versus air-fuel ratio of the fresh charge into the engine. The amount of ammonia gas in the engine exhaust gases is assumed to be zero. Mole fraction of hydrogen is assumed to be one third of CO mole fraction for a gasoline-fueled vehicle ([24]). This ratio must be adjusted using experimental values for fuels of different hydrogen-to-carbon and oxygen-to-carbon ratios. Mole fraction of oxygen is calculated based on the values of fractions of CO, NO, HC, $H_2$, and air-fuel ratio of the engine fresh charge. Surface species equations do not need boundary conditions, and initial conditions are enough to solve the equations. The surface energy equation has second-kind Neumann boundary conditions.

The spatial node size for this study ranged from 0.1693 mm to 19.32 mm. The grid sensitivity was tested by checking if the results remain the same when a finer grid (half of the original grid spacing) was used. A time step of 0.01 second was employed for the transient calculation. The sensitivity of temporal step size was also tested extensively. Figure 2 shows the effect of time step on catalyst substrate temperature and CO emission during the first 100 seconds of the U.S. federal test procedure (FTP). The figure shows that as the time step is decreased from $10^{-2}$ second to $10^{-5}$ second, the results remain nearly the same. This figure clearly shows the suitability of the selected time step.

## Results and Discussion

The numerical model was assessed using experimental measurements under both steady and transient conditions. The model's predictions under steady-state conditions were compared with the experimental measurements of Montreuil et al. [13]. For this case, all transient terms in the governing equations were set to be zero. The geometry of the catalytic converter and the feed gas conditions used in this case are listed in Tables 1 and 2, respectively.

Figure 3 shows the comparison of the converter pollutant conversion efficiencies as determined by the model and experimental measurements. The conversion efficiencies are plotted as a function of redox ratio. The figure depicts that the numerical model results are in good agreement with the experimental measurements at the typical engine air-fuel ratio regions. NO conversion efficiency is particularly well predicted by the model over a wide

**Table 1 Dimensions of catalytic converters**

| Dimension | Steady state | FTP | Comparison between FTP and ECE15+EUDC |
|---|---|---|---|
| Length, cm | 3.81 | 8 | 14.99 |
| Cross-sectional area, $cm^2$ | 5.067 | 86.0254 | 50.26 |
| Cell density, cells/$cm^2$ | 62 | 62 | 72.85 |
| Wall thickness, mm | 0.188 | 0.1905 | 0.1905 |

redox ratio range. However, when the redox ratio increases, i.e., in rich mixture zone, the discrepancies in the predictions of CO and HC increase. One reason for such a disagreement is the inaccuracy in chemical kinetic data, which are obtained for conditions close to redox ratio of unity. Nonetheless, the model validation in a high redox ratio region is not a matter of much concern since, for a practical engine, the operating conditions are always in the vicinity of stoichiometric air-fuel ratio (redox ratio is unity), where the model predictions are acceptable. Hence, it can be concluded that the present model well predicts the performance of catalytic converters under steady-state conditions.

The model's performance under transient conditions was assessed by using experimental measurements across the front catalyst brick from a 4.6L 2V Lincoln Towncar as it was driven during the FTP cycle. The geometry of the converter used is listed in the second column of Table 1. It was a palladium-based catalyst, with a length of 8 cm, cross-sectional area of 86.0254 $cm^2$, cell density of 62 cells/$cm^2$, and wall thickness of 0.1905 mm. The measured

**Table 2 Feed gas data of a steady-state condition for the model validation (*space velocity is the ratio of gas volumetric velocity to catalyst converter volume)**

| Parameter | Value |
|---|---|
| Gas temperature | 371 °C |
| CO fraction | 1 % |
| $CH_4$ fraction | 0 ppm C1 |
| $C_3H_6$ fraction | 1000 ppm C3 |
| $C_3H_8$ fraction | 500 ppm C3 |
| NO fraction | 1000 ppm |
| Space velocity* | 50000 $hr^{-1}$ |

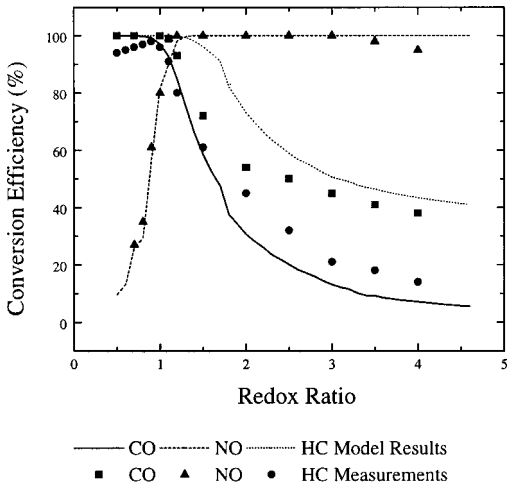



**Fig. 3 Comparison of model prediction and experimental measurements at steady-state conditions**

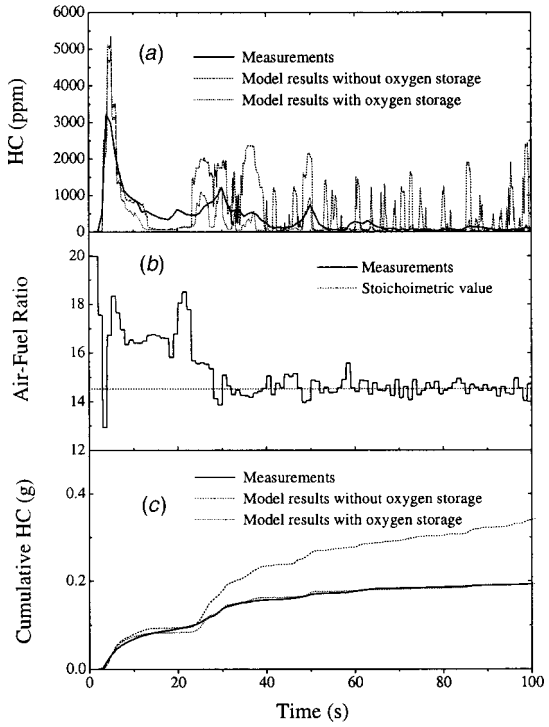



**Fig. 4 Comparison of hydrocarbon (HC) emissions between model predictions and experimental measurements during the first 100 seconds of the federal test procedure (FTP) (*a*) instantaneous hydrocarbon (HC) emissions; (*b*) measured air-fuel ratio versus time; (*c*) cumulative hydrocarbon (HC) emissions**

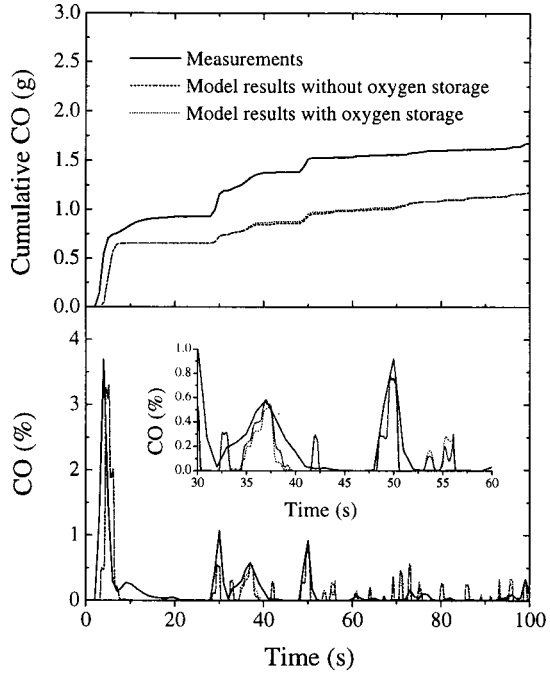



**Fig. 5 Comparison of CO emission between model predictions and experimental measurements during the first 100 seconds of the federal test procedure (FTP)**

data included the mass flow rates of CO, HC, $NO_x$ and the total gases at both inlet and outlet of the catalytic converter. The air-fuel ratio of engine's fresh charge was calculated based on the measured air flow and fuel supply rates. The ambient temperature during the experiment was 297 K. The measurements were made at an interval of one second; therefore, there were 1372 measured data for FTP-72.

Figure 4(*a*) shows the comparison of instantaneous HC emissions as determined by the model and experimental measurements during the first 100 seconds of the FTP. In order to understand the impact of the oxygen storage, the model results without the oxygen storage capacity are also displayed in the figure. As expected from prior experiments and calculations, the results show relatively large differences between the model predictions and the experimental measurements if the oxygen storage mechanism is not considered. The HC profile has spikes when the air-fuel ratio in the feed gas is near or below the stoichiometric values (Fig. 4(*b*) shows the air-fuel ratio in the feed gas versus time). This problem has been reported in literature. Moore and Mondt [3] found that their model, due to the absence of the oxygen storage mechanism, is more sensitive than real catalyst to the air-fuel ratio variations. The figure shows that the model results are substantially improved by including the oxygen storage mechanism. Hence, as has been shown previously experimentally and analytically, the oxygen storage mechanism plays an important role in the hydrocarbon oxidation during the transient operating conditions. Figure 4(*c*) shows the comparison of cumulative hydrocarbon emissions as determined by the model and experimental measurements. The results show good agreement as the difference between the computational results including oxygen storage mechanism and the experimental measurements are within 0.3 percent during this 100 second period.

Figures 5 and 6 show instantaneous and cumulative CO and NO emissions during the first 100 seconds of the FTP. The results show that the model underestimates the CO and NO emissions during the initial phase of the legislated cycle. Even though the model accuracy is not as good as in the case of HC prediction, the difference between the numerical predictions and the experimental measurements is within the variability caused by the dynamic measurement of $A/F$ ratio, which strongly influences modeling results ([26]). The influence of the oxygen storage mechanism on CO and NO conversions has been found to be less significant than

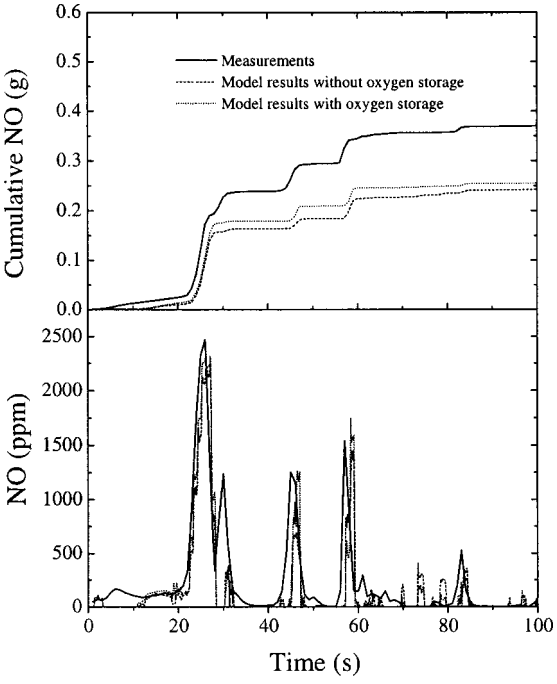



**Fig. 6 Comparison of NO emission between model predictions and experimental measurements during the first 100 seconds of the federal test procedure (FTP)**

**Table 3 Comparison of total conversion efficiencies between model predictions and experimental measurements during the federal test procedure (FTP) test**

| Efficiency (%) | CO | NO | HC |
|---|---|---|---|
| Experimental | 91.53 | 92.03 | 95.88 |
| Model prediction | 88.69 | 95.93 | 94.84 |
| Difference (%) | 3.1 | 4.24 | 1.08 |

its effect on HC using the kinetic rates for these early catalysts. Figures 5(*b*) and 6(*b*) show that the numerical model is also able to predict the instantaneous performance of the catalyst with a reasonable accuracy. Most of the emission peaks during the initial period of the FTP test cycle are accurately simulated by the numerical model.

For the total FTP test cycle, comparison of pollutant conversion efficiencies as determined by the model and experimental measurements is shown in Table 3. The table clearly shows the adequacy of the numerical model in simulating the performance of three-way catalytic converter. There is a difference of less than five percent between the conversion efficiencies predicted by the numerical model and those determined experimentally. These results were computed on a desktop computer (Pentium 233) and the whole FTP simulation was performed in about 15 minutes.

The generalization of the numerical model and the chemical kinetics were also tested by assessing the model performance for a different catalytic converter and during a different legislated cycle—the European test cycle plus extra urban driving cycle (ECE 15+EUDC). The geometry of the converter used in this analysis is listed in the third column of Table 1. It was a palladium-based catalyst, with a length of 14.99 cm, cross-sectional area of 50.26 $cm^2$, cell density of 72.85 cells/$cm^2$, and wall thickness of 0.1905 mm. Figure 7 shows the comparisons of HC emissions as determined by the model and experimental measurements during the first 100 seconds of the FTP and ECE15 +EUDC test cycles. The model results in both cases are in good agreement with experimental measurements, except that the phase of the first predicted HC peak in the ECE15+EUDC delays about four seconds. This may be due to the inaccuracy of the measurement. Good agreements were also found between the model predictions and experimental measurements for instantaneous and accumulated CO and NO emissions. These results attest the validity and robustness of the present model for analyzing the converter performance under different driving conditions.

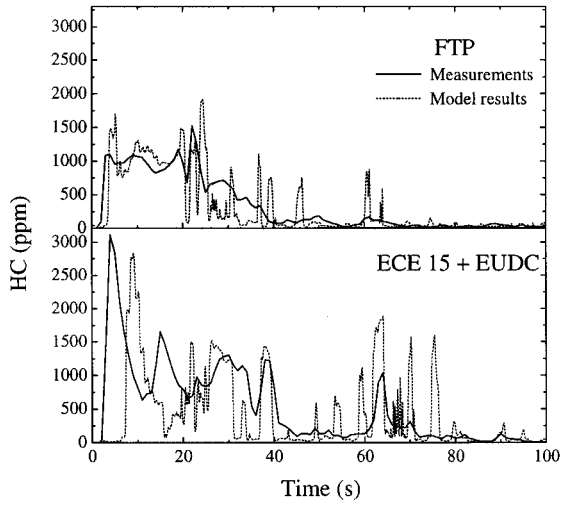



**Fig. 7 HC emissions versus time during the federal test procedure (FTP) and European test cycles**

## Conclusions

- A numerical simulation code for predicting the catalytic converter performance has been developed. This code includes the transient terms in the governing equations so that it can be used to predict the catalytic converter performance during transient operating conditions such as U.S. federal test procedure (FTP) and European regulated driving cycles.
- The computational results show good agreement with the experimental measurements at both steady-state and transient operating conditions. The model predicts the pollutant conversion efficiencies within five percent of the experimental measurements for the FTP.
- The numerical model includes the oxygen storage mechanism of the substrate. This mechanism improves the catalyst conversion efficiencies during rich operating conditions. The results show a greater influence of oxygen storage mechanism on HC than on CO and $NO_x$ due to the great sensitivity of CO and $NO_x$ to $A/F$ ratio modulation ([26]).
- The code has been found to be robust and can simulate the catalyst performance during various operating conditions. In addition, this code possesses a high computation speed. It takes about 15 minutes on a Pentium 233 PC to simulate the performance of a catalytic converter during the FTP.

## Acknowledgments

The financial support from the Chemical Engineering Department of the Ford Scientific Research Laboratory and the Center for Engineering Education and Practice (CEEP) of the University of Michigan-Dearborn is greatly appreciated.

## Nomenclature

$C_g^j$ = gas phase concentration of species $j$, mol/$m^3$
$C_s^j$ = surface concentration of species $j$, mol/$m^3$
$cp_g$ = specific heat of gas, J/kg·K
$cp_s$ = specific heat of substrate, J/kg·K
CTR = empirical constant for oxygen storage reaction rates
$EA,E$ = activation energy, Pa-$m^3$/g-mol
$ex$ = empirical constant for reaction rates
$G_a$ = geometric surface area, $m^2/m^3$
$\Delta H^k$ = heat of reaction of species $k$, J/mol
$h_g$ = heat transfer coefficient between flow and substrate, J/$m^2$·s·K
$h_\infty$ = heat transfer coefficient between substrate and atmosphere, J/$m^2$·s·K
$K$ = empirical constant for oxygen storage reaction rates
$k_m^j$ = mass transfer coefficient for species $j$, m/s
$R^k$ = reaction rate of $k$th reaction, mol/$m^2$·s
$S_{\text{ext}}$ = external surface to volume area ratio, $m^2/m^3$
$t$ = time, s
$T_\infty$ = ambient temperature, K
$T_g$ = gas temperature, K
$T_s$ = substrate temperature, K
$v_g$ = gas flow velocity, m/s
$X_s^i$ = mole fraction of species $i$ in substrate
$z$ = coordinate along catalyst axis, m
$\alpha$ = hydrogen-to-carbon ratio in the fuel
$\varepsilon$ = void volume fraction (value ranging from zero (for no void) to one)
$\lambda_s$ = thermal conductivity of substrate, J/m·s·K
$\rho_g$ = gas density, kg/$m^3$
$\rho_s$ = substrate density, kg/$m^3$

## Appendix

The chemical reaction rate expressions for the reaction scheme are as following. The values of different coefficients used in these expressions are listed elsewhere ([13]).

$$\text{RATE}\ (1)=\frac{C_1 e^{-EA(1)/RT_s} X_{\text{CO}} X_{\text{O}_2}^{ex(1)}}{[(1+C_4 X_{\text{CO}})^2(1+C_5 X_{\text{C}_3\text{H}_6})]} \tag{9}$$

$$\text{RATE}\ (2)=[C_2 e^{-EA(6)/RT_s}+(1-\beta_1)C_3 e^{-EA(3)/RT_s}] \cdot \frac{3X_{\text{C}_3\text{H}_6} X_{\text{O}_2}^{ex(2)}(1+C_{31}X_{\text{C}_3\text{H}_6}X_{\text{CO}})}{(1+C_{35}X_{\text{C}_3\text{H}_6})^2(1+C_{32}e^{-EA(7)/RT_s}X_{\text{CO}})} \tag{10}$$

$$\text{RATE}\ (3)=\beta_1 \cdot [C_3 e^{-EA(3)/RT_s}] \cdot \frac{3X_{\text{C}_3\text{H}_6} X_{\text{O}_2}^{ex(2)}(1+C_{31}X_{\text{C}_3\text{H}_6}X_{\text{CO}})}{(1+C_{35}X_{\text{C}_3\text{H}_6})^2(1+C_{32}e^{-EA(7)/RT_s}X_{\text{CO}})} \tag{11}$$

$$\text{RATE}\ (4)=\frac{C_6 e^{-EA(4)/RT_s} X_{\text{H}_2\text{O}} X_{\text{O}_2}}{[(1+C_{47}X_{\text{H}_2})^2(1+C_4 X_{\text{CO}})^2(1+C_5 X_{\text{C}_3\text{H}_6})]} \tag{12}$$

$$\text{RATE}\ (5)=\frac{C_{46} X_{\text{NH}_3}^{ex(4)}}{(1+C_{47}X_{\text{H}_2}+C_{48}X_{\text{CO}}+C_{49}X_{\text{C}_3\text{H}_6})^2} \tag{13}$$

$$\text{RATE}\ (6L)=\frac{(1-\beta_1)C_{15}X_{\text{CO}} e^{-EA(2)/RT_s} X_{\text{NO}}^{ex(10)}}{[(1+C_{16}X_{\text{O}_2})^{ex(11)}(1+C_{50}X_{\text{CO}})^2]} \tag{14}$$

$$\text{RATE}\ (6R)=\frac{\beta_1 \cdot C_{11} e^{-EA(5)/RT_s} X_{\text{CO}} X_{\text{NO}}^{ex(5)}(1+C_{12}X_{\text{O}_2})^{ex(6)}}{[F1+F2+F3-2]^2}$$
$$F1=(1+\text{CMILL}\cdot X_{\text{CO}})^{ex(7)}$$
$$F2=(1+C_{14}X_{\text{H}_2})^{ex(8)}$$
$$F3=(1+C_{34}e^{EA(10)/RT_s}X_{\text{C}_3\text{H}_6})^{ex(23)} \tag{15}$$

$$\text{RATE}\ (7)=\frac{\beta_1 \cdot C_{17}X_{\text{CO}}X_{\text{NO}}^{\text{ex}(13)}(1+\text{CMILL}\cdot X_{\text{O}_2})^{ex(14)}}{F4\cdot F5\cdot F6}$$
$$F4=[1+C_{19}|T_S-C_{20}|^{ex(12)}]$$
$$F5=[1+C_{13}X_{\text{CO}}+C_{33}X_{\text{H}_2}]$$
$$F6=[1+C_{30}X_{\text{C}_3\text{H}_6}] \tag{16}$$

$$\text{RATE}\ (8L)=\frac{(1-\beta_2)C_{22}X_{\text{H}_2}^{ex(16)}X_{\text{NO}}^{ex(15)}}{(1+C_{44}X_{\text{O}_2})^2} \tag{17}$$

$$\text{RATE}\ (8R)=\frac{C_{21}X_{\text{H}_2}X_{\text{NO}}^{ex(9)}(1+C_{23}X_{\text{O}_2})^2\cdot\beta_2}{[F7+F8+F9-2]^2}$$
$$F7=(1+C_8 X_{\text{CO}})^{ex(7)}$$
$$F8=(1+C_9 X_{\text{H}_2})^{ex(8)}$$
$$F9=(1+C_{34}e^{EA(10)/RT_s}X_{\text{C}_3\text{H}_6})^{ex(23)} \tag{18}$$

$$\text{RATE}\ (9L)=(1+\beta_2\cdot\text{CMILL}\cdot X_{\text{O}_2})^{ex(18)}\cdot\text{RATE}(9R) \tag{19}$$

$$\text{RATE}\ (9R)=\frac{C_{25}X_{\text{H}_2}X_{\text{NO}}^{ex(17)}}{F10\cdot F11\cdot F12}$$
$$F10=[1+C_{27}|T_S-C_{28}|^{ex(26)}] \tag{20}$$

$$F11=[1+C_{24}X_{\text{CO}}+C_{18}X_{\text{H}_2}]$$
$$F12=[1+C_{26}e^{EA(11)/RT_s}X_{\text{C}_3\text{H}_6}]$$

$$\text{RATE}\ (10)=\frac{3C_{10}e^{-EA(8)/RT_s}X_{\text{O}_2}X_{\text{C}_3\text{H}_8}^{ex(26)}}{(1+C_{51}X_{\text{O}_2})^{ex(27)}} \tag{21}$$

$$\text{RATE}\ (11)=\frac{\beta_1\cdot 6C_{38}X_{\text{C}_3\text{H}_6}X_{\text{NO}}^{ex(24)}(1+C_{37}X_{\text{O}_2})^{ex(3)}}{(1+C_{39}X_{\text{CO}})[F1+F2+F3-2]^2} \tag{22}$$

$$\text{RATE}\ (12)=(1-\beta_1)9C_{42}e^{-EA(9)/RT_s}X_{\text{NO}}^{ex(22)}X_{\text{C}_3\text{H}}^{ex(25)} \tag{23}$$

$$\text{RATE}\ (13)=\frac{\beta_1\cdot 2.4C_{38}X_{\text{NO}}^{ex(20)}X_{\text{C}_3\text{H}_6}(1+\text{CMILL}\cdot X_{\text{O}_2})^{ex(21)}}{[F13\cdot F14\cdot F15]}$$
$$F13=[1+C_{40}|T_S-C_{41}|^{ex(19)}]$$
$$F14=[1+C_{29}X_{\text{C}_3\text{H}_6}]^2$$
$$F15=[1+C_{43}X_{\text{CO}}+C_{45}X_{\text{H}_2}] \tag{24}$$

where $\beta_1$ and $\beta_2$ are rich-to-lean blending functions;

$$0\leqslant\beta_1\leqslant 1 \quad \text{and} \quad 0\leqslant\beta_2\leqslant 1$$
$$\beta_1=2500\,R_1+0.5$$
$$\beta_2=2500\,R_2+0.5$$
$$\text{if}\ \ \beta<0,\ \ \beta=0$$
$$\text{if}\ \ \beta>0,\ \ \beta=1.$$

Here,

$$R_1=X_{\text{CO}}+9X_{\text{C}_3\text{H}_6}+1.5X_{\text{NH}_3}-X_{\text{NO}}-2X_{\text{O}_2}-R_{\text{flip}}$$
$$R_2=X_{\text{CO}}+9X_{\text{C}_3\text{H}_6}+1.5X_{\text{NH}_3}-X_{\text{NO}}-2X_{\text{O}_2}-H_{\text{flip}}+X_{\text{H}_2}$$

$R$ indicates rich-side (Redox $>1$) rate, and $L$ indicates lean-side (Redox $<1$) rate.

## References


[1] Federal Register, 1968, Vol. 33, No. 108.
[2] U.S. Revised Clean Air Act of 1994.
[3] Moore, W. R., and Mondt, J. R., 1993, "Predicted Cold Start Emission Reductions Resulting From Exhaust Thermal Energy Conservation to Quicken Catalytic Converter Lightoff," SAE Paper No. 931087.
[4] Chen, D. K. S., 1993, "A Numerical Model for Thermal Problem in Exhaust Systems," Vehicle Thermal Management Systems Conference, Columbus, OH, Mar. 29–Apr. 1.
[5] Wendland, D. W., 1993, "Automobile Exhaust-System Steady-State Heat Transfer," SAE Paper No. 931085.
[6] Lai, M. C., Lee, T., Kim, J. Y., Li, P., Chui, G., and Pakko, J. D., 1992, "Numerical and Experimental Characterization of Automotive Catalytic Converter Internal Flows," J. Fluids Struct., **6**, pp. 451–470.
[7] Baxendale, A. J., 1993, "The Role of Computational Fluid Dynamics in Exhaust System Design and Development," SAE Paper No. 931072.
[8] Lloyd-Thomas, D. G., Ashworth, R., and Qiao, J., 1993, "Meeting Heat Flow Challenges in Automotive Catalyst Design with CFD," SAE Paper No. 931079.
[9] Young, L. C., and Finlayson, B. A., 1976, "Mathematical Models of the Monolithic Catalytic Converter, Part 1, Development of Model and Application of Orthogonal Collocation," AIChE J., **22**, pp. 331–343.
[10] Young, L. C., and Finlayson, B. A., 1976, "Mathematical Models of the Monolithic Catalytic Converter, Part 2, Application to Automobile Exhaust," AIChE J., **22**, pp. 343–353.
[11] Otto, N. C., and LeGray, W. J., 1980, "Mathematical Models for Catalytic Converter Performance," SAE Paper No. 800841.
[12] Oh, S., 1988, "Thermal Response of Monolithic Catalytic Converters During Sustained Engine Misfiring: A Computational Study," SAE Paper No. 881591.
[13] Montreuil, C. N., Williams, S. C., and Adamczyk, A. A., 1992, "Modeling Current Generation Catalytic Converters: Laboratory Experiments and Kinetic Parameter Optimization—Steady State Kinetics," SAE Paper No. 920096.
[14] Pattas, K. N., Stamatelos, A. M., Pistikopoulos, P. K., Koltsakis, G. C.,

Konstantinidis, P. A., Volpi, E., and Leveroni, E., 1994, ''Transient Modeling of 3-Way Catalytic Converters,'' SAE Transactions, Paper No. 940934.
[15] Siemund, S., Leclerc, P., Schweich, D. J., Prigent, M., and Castagna, F., 1996, ''Three-Way Monolithic Converter: Simulations Versus Experiments,'' Chem. Eng. Sci., **51**, pp. 3709–3720.
[16] Koltsakis, G. C., Konstantinidis, P. A., and Stamatelos, A. M., 1997, ''Development and Application Range of Mathematical Models for 3-Way Catalytic Converters,'' Appl. Catal., B, **12**, pp. 161–191.
[17] Koltsakis, G. C., Kandylas, I. P., and Stamatelos, A. M., 1998, ''Three-Way Catalytic Converter Modeling and Applications,'' Chem. Eng. Commun., **164**, pp. 153–189.
[18] Oh, S. H., and Cavendish, J. C., 1985, ''Mathematical Modeling of Catalytic Converter Lightoff—Part 2: Model Verification by Engine-Dynamometer Experiments,'' AIChE J., **31**, pp. 935–942.
[19] Oh, S. H., and Cavendish, J. C., 1985, ''Mathematical Modeling of Catalytic Converter Lightoff—Part 3: Prediction of Vehicle Exhaust Emissions and Parametric Analysis,'' AIChE J., **31**, pp. 943–947.
[20] Voltz, S. E., Morgan, C. R., Liederman, D., and Jacob, S. M., 1973, ''Kinetic Study of Carbon Monoxide and Propylene Oxidation on Platinum Catalysts,'' Ind. Eng. Chem. Prod. Res. Dev., **12**, pp. 294–301.
[21] Kuo, T. C., Morgan, C. R., and Lassen, H. C., 1971, ''Mathematical Modeling of CO and HC Catalytic Converter Systems,'' SAE Paper No. 710289.
[22] Otto, N. C., 1984, private communication.
[23] Li, P., Adamczyk, A. A., and Pakko, J. D., 1994, ''Thermal Management of Automotive Emission Systems: Reducing the Environmental Impact,'' The Japan-U.S. Seminar on Thermal Engineering for Global Environment Protection (A-3).
[24] Koltsakis, G. C., and Stamatelos, A. M., 1997, ''Catalytic Automotive Exhaust Aftertreatment,'' Prog. Energy Combust. Sci., **23**, pp. 1–39.
[25] Gandhi, H. S., Piken, A. G., Shelef, M., and Delosh, R. G., 1976, ''Laboratory Evaluation of Three-Way Catalysts,'' SAE Trans., **85**, pp. 201–212.
[26] Shamim, T., Medisetty, V. S., and Shen, H., 2000, ''Dynamic Response of Catalytic Converters Subjected to Changes in Air-Fuel Ratio,'' *Proceedings of the 2000 Technical Meeting of the Central States Section of the Combustion Institute*, Combution Insitute, Pittsburgh, PA, pp. 389–394.

# Triple Cycle: A Conceptual Arrangement of Multiple Cycle Toward Optimal Energy Conversion

**T. C. Hung**
Associate Professor,
Department of Mechanical Engineering,
I-Shou University,
1, Section 1,
Hsueh-Cheng Road,
Ta-Hsu Hsiang,
Kaohsiung County,
840 Taiwan, R.O.C.
e-mail: tchung@isu.edu.tw

*The purpose of this study is to find a maximum work output from various combinations of thermodynamic cycles from a viewpoint of the cycle systems. Three systems were discussed in this study: a fundamental combined cycle and two other cycles evolved from the fundamental dual combined cycle: series-type and parallel-type triple cycles. In each system, parametric studies were carried out in order to find optimal configurations of the cycle combinations based on the influences of tested parameters on the systems. The study shows that the series-type triple cycle exhibits no significant difference as compared with the combined cycle. On the other hand, the efficiency of the parallel-type triple cycle can be raised, especially in the application of recovering low-enthalpy-content waste heat. Therefore, by properly combining with a steam Rankine cycle, the organic Rankine cycle is expected to efficiently utilize residual yet available energy to an optimal extent. The present study has pointed out a conceptual design in multiple-cycle energy conversion systems.* [DOI: 10.1115/1.1423639]

## Introduction

As living standards soar, demand on energy also escalates to an unprecedented level. However, the developments of energy resources have been hampered by public concerns on environment protection and living quality that resulted in a global-wise shortage of current utility capacity. Accordingly, other than the development of reusable energy sources, recovery of waste heat and integration of energy systems have become urgent issues in an attempt to yield an optimal utilization of the limited energy resources.

Hung et al. [1] and Marciniak [2] have analyzed the organic Rankine cycle (ORC) system performances and their parametric impacts using several organic fluids such as benzene, toluene, CFC-containing refrigerants, and some newly developed substitutes for refrigerants. They have set some criteria in terms of selecting the working fluids based on the temperatures of the waste heat to be recovered. Schulitz [3] analyzed waste heat (200–500°C) recovery from the exhaust of an internal combustion engine in conjunction with an ORC as the bottoming cycle. He also found an optimal operation range for each working fluid.

There is a wide range of the heat sources, which can be applied to the ORC systems such as waste heat from the condenser of a conventional fossil or a nuclear power plant, waste heat from industrial process, solar radiation, and geothermal energy ([4,5]). Chaudoir et al. [6] have designed a medium temperature ORC system with a capacity of 37.5 kW using solar energy as the energy source and toluene as the working fluid. Manco and Nervegna [7] analyzed thermodynamic behavior of the ORC systems with the aid of computer simulations. They deduced some schemes for selecting appropriate working fluids in the application of heat recovery of an ORC system.

Lee et al. [8] studied the efficiencies of an ideal Rankine cycle using several working fluids. In their work, the fluid thermophysical properties were calculated by using the Iwai-Margerum-Lu equation of state. Their results indicated that the thermal capacity of the system directly relates to the boiling temperature under atmospheric pressure, critical pressure, and molecular weight of the fluid. Furthermore, they also found that among the fluids studied, refrigerant R-123 turned out to be the best choice for the working fluid based on both environmental concerns and system efficiencies. This is also the reason why R-123 was used as the working fluid of the ORC system in this study.

Habib [9] reviewed various studies and concluded that cogeneration plants in general exhibit a 25 percent higher thermal efficiency and a 38 percent lower irreversibility than conventional power plants. Lucia et al. [10] have performed experimental operations and tests. They found that, as the inlet temperature of the compressor drops, the system efficiency and the total economic benefit could be effectively raised because of lower power requirements by the compressor.

Woodward [11] studied the recovery of waste heat using the Brayton cycle and the Rankine cycle as alternatives to the diesel cycle. By referring to the actual conditions under operation, he derived an analytical model in his study. Briesch et al. [12], in a joint Advanced Turbine System (ATS) project sponsored by Westinghouse and the U.S. Department of Energy with an objective of raising the efficiency of combined cycle power systems from 54 percent to 60 percent before year 2000, proposed some improvements, including raising the inlet temperature of gas turbines, development of new techniques for cooling turbine blades, and some schemes for raising the efficiencies of topping cycles and bottoming cycles. However, in an attempt to raise the efficiency to 60 percent with their proposed modifications and schemes, it would require an investment that can easily exceed one billion U.S. dollars.

A detailed thermodynamic analysis of the combined power cycle without supplementary firing was given in Horlock [13]. Minimizing the exergy loss due to the temperature difference in the heat recovery steam generator (HRSG) has been a dominant role to increase the overall efficiency of a combined cycle. Thermodynamic calculations illustrated that the gains in overall efficiency are not large by the employment of double and even triple pressure steam increases. Tomlinson et al. [14] pointed out that

Contributed by the Combustion and Fuels Division of THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS for publication in the ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the C&F Division, October 2000; final revision, April 2001. Editor: S. R. Gollahalli.

while an increase of nearly four percent comes from single to dual pressure steam increase, the incremental improvement in efficiency in moving from dual pressure to triple pressure is only about one percent.

In the study of Marston and Hyre [15] for a gas turbine combined cycle power plant, they have compared a triple-pressure steam cycle with both single-stage and three-stage Kalina cycles as the bottoming sections. They found that both Kalina cycles were more efficient than the triple pressure steam cycle. Optimization of the three-stage Kalina cycle resulted in only a two percent improvement.

The ultimate purpose of this study is to find an optimal combination of the bottoming cycle(s) in order to fully utilize the waste heat of the Brayton cycle. Our approach is to use an ORC in conjunction with a combined cycle to form a triple-cycle model in order to transport heat and reuse the energy. Furthermore, two more models are analyzed based on the heating values of waste heat of the Brayton cycle, i.e., the combined cycle with a steam Rankine cycle for recovery of high-temperature waste heat and an ORC to recover low-temperature waste heat. Both are in a three-cycle-in-parallel model. Based on these models, the efficiencies of the multicycle systems were evaluated.

## Models and Concepts for Analysis

Due to different thermophysical properties, each fluid has an operational range for its highest performance. For instance, some fluids can be used as working fluids without superheating; dry fluids are suitable for higher working temperatures; and isentropic fluids have better performance at lower temperatures, in which dry fluids and isentropic fluids represent the positive slope and nearly vertical saturated-vapor curves on the T-S diagram, respectively. Therefore, system efficiency presumably can be raised by appropriately selecting working fluids in various multiple subcycle systems.

A schematic diagram with multiple bottoming cycles is shown in Fig. 1. The total waste heat of the topping cycle, $Q_L$, is exhausted to the bottoming cycle in a series of stages with a quantity of $\delta Q_L$ exhausted in each waste heat temperature range. Meanwhile, waste heat is recovered in each subcycle to generate work, as expressed by the following equations:

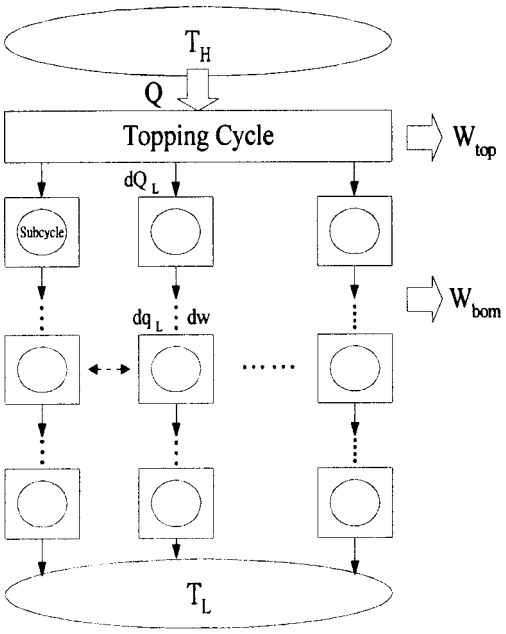



**Fig. 1 The conceptual arrangement for the energy conversion system with multiple cycles**

$$\delta Q_L = \int \delta q_L + \int \delta w \tag{1}$$

$$Q_L = Q - W_{\text{top}} = \int \delta Q_L = \int \int \delta q_L + \int \int \delta w. \tag{2}$$

In the topping cycle and each subcycle of the bottoming cycle, an appropriate operations model associated with each working fluid can be analyzed and selected to yield optimal performance. Thermophysical properties of each fluid, which are crucial to system performance, can be expressed by a mathematical model

$$\xi = f(\xi_1, \xi_2) \tag{3}$$

in which $\xi_1$ and $\xi_2$ are two arbitrary intensive thermodynamic properties.

Accordingly, the system efficiency, $\eta$, and the irreversibility, $\dot{I}$, are functions of the following factors:

$$\eta = f(\xi_{ij}) \quad \text{and} \quad \dot{I} = f(\xi_{ij}, T_k) \tag{4}$$

where $\xi_{ij}$ denotes the property function of the fluid in any subcycle of the multicycle system; $T_k$ denotes the environmental temperature.

System optimization means that the maximum efficiency and minimum irreversibility have been reached when

$$\frac{\partial \eta}{\partial \lambda} = 0 \quad \text{and} \quad \frac{\partial \dot{I}}{\partial \lambda} = 0 \tag{5}$$

where $\lambda$ is any variable parameter in the system.

In this study, efforts were focused on model development and analysis for two or three-cycle combinations. For two-cycle systems, for instance, Brayton and Rankine cycles are combined to form a so-called combined cycle based on the principle of heat exchange between cycles. The first type of the three-cycle system is a series type cycle as shown in Fig. 2. The steam Rankine cycle, which absorbs the waste heat of the Brayton cycle as the heat source, generates power and also exhausts part of its low-temperature thermal energy to an ORC to further generate power. On the other hand, a steam Rankine cycle and an ORC can be connected in parallel so that each cycle will absorb the waste heat that is most suitable for its own operation. It means that the steam Rankine cycle absorbs the waste heat of a higher temperature from the Brayton cycle while the ORC absorbs the waste heat of a lower temperature as shown in Fig. 3. These two fundamental combination types can be expanded to any multiple cycles with various configurations, depending on their applicability.

## Results and Discussion

The simple reversible subcycles are employed in the following parametric analyses and only major parameters were selected. Certainly, efficiency could be improved by modifying the processes such as intercooling to the gas turbine, dual-pressure design in the boiler of the steam Rankine cycle, reheating, etc. These discussions will be left for future study.

A triple cycle is the original concept of multiple-cycle energy conversion. Limited, but crucial parameters, are selected in the following analyses. The framework of this study was based on a combined cycle configuration and, furthermore, was extended to multicycle systems including organic Rankine cycles. The corresponding reference states and values used are listed in Table 1. In a series-typed cycle, the working fluid tested for ORC was refrigerant R-12. Other working fluids, $NH_3$, $C_6H_6$, and refrigerant R-123, were evaluated in an ORC parallel-type cycle.

**Combined Cycle.** Before analyzing the triple cycle, the combined-cycle (Brayton and steam Rankine cycles) behaviors with sample parameters are discussed a priori.

As the compression ratio, $Rp\ (P_2/P_1)$, increases under a fixed turbine inlet temperature, the compressor exit temperature will increase, which reduces fuel consumption. In the meantime, the

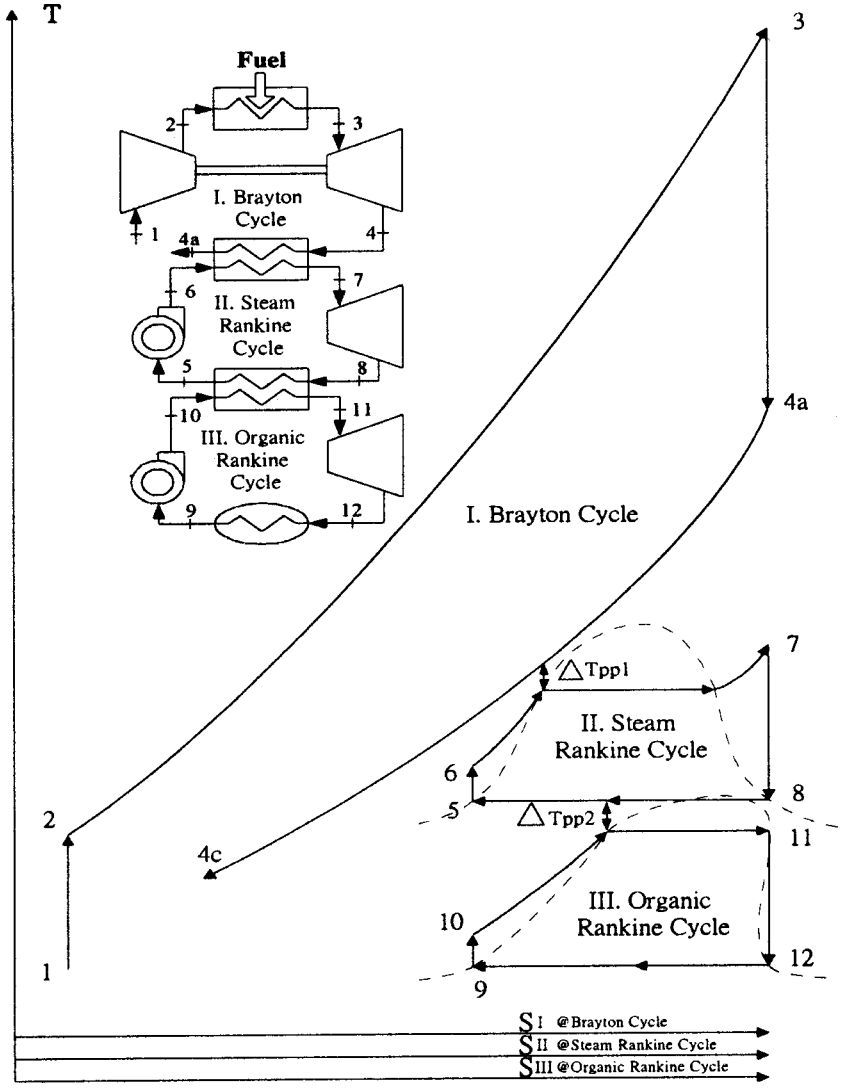



**Fig. 2 Components and T-S process diagram for an ideal series-type triple cycle**

turbine exit temperature of the Brayton cycle will drop as more work is generated. Both effects result in an increase of the efficiency of the Brayton cycle. Based on the pressure range chosen for the boiler, as *Rp* decreases, the difference between gas turbine exit temperature and steam turbine inlet temperature drops accordingly. The effects of the compression ratio on the efficiencies of the Brayton, steam Rankine, and combined cycles are shown in Fig. 4. It can be seen that, in spite of a favorable increase in efficiency in the Brayton cycle, the efficiency of the combined cycle increases only slightly as *Rp* increases. The efficiency of the steam Rankine cycle drops as *Rp* increases due to a reduction in the gas turbine exhaust temperature, which results in a lower steam turbine inlet temperature. The result shows a trend similar to that reported by Briesch et al. [12]. Therefore, it is not an effective measure to raise the efficiency of the combined cycle by raising *Rp*.

The effects of pump inlet (i.e., condenser) temperature, $T_5$, on efficiency are shown in Fig. 5, and the effects of increasing pump pressure are shown in Fig. 6. The efficiency of the Brayton cycle remains the same for both these operating changes; therefore the effects on combined cycle efficiency are only these effects on the steam cycle. Increasing the condensation temperature lowers the overall efficiency, while increasing the pump pressure causes a slight increase in overall efficiency. Of course, a lower ambient temperature would yield a greater overall efficiency.

**Series-Type Triple Cycle.** Basically a series-type triple cycle is a straight forward extension derived from the combined cycle. When compared to a dual-cycle combined system, however, the resulting steam expansion ratio is reduced, which reduces the availability. The temperature difference between steam and organic Rankine cycles enlarges the irreversibility and could result in lower power generation from the ORC. However, selection of an appropriate fluid, corresponding to the condensation temperature of the steam cycle, might compensate for this loss and create an overall greater efficiency. This is the main motivation for this series-triple-cycle study; however, to date only refrigerant R-12 has been the only working fluid analyzed in the present investigation.

As previously discussed, the efficiency of the steam Rankine cycle drops as the compression ratio, *Rp*, increases as a result of a decrease in the amount of waste heat available, as noted in Fig. 7. Also included in Fig. 7 are the corresponding Brayton cycle, ORC cycle, and series-type triple cycle efficiencies. Accordingly, in a triple cycle, an ORC serves as the third stage of power generation. The efficiency of the ORC also drops when the waste heat exhausted from the upstream steam Rankine cycle is reduced, as shown in Fig. 7.

When $\Delta T_{pp2}$ (see Fig. 3) remains unchanged, an increase of condensation temperature of the steam cycle, $T_5$, will cause a higher inlet turbine temperature to the ORC. From the T-S dia-

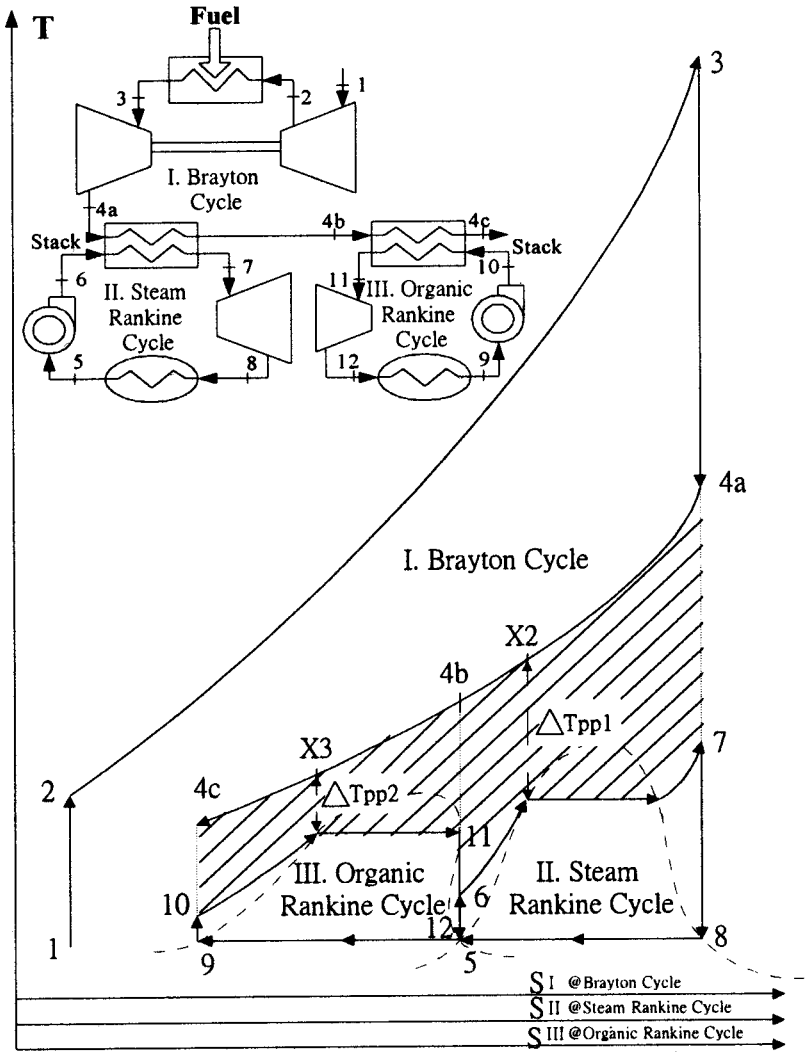



**Fig. 3 Components and T-S diagram for an ideal parallel-type triple cycle**

gram, it is found that the effective heat-to-work conversion area is decreased for the steam cycle but increased for the ORC. This increase of ORC efficiency is not significant enough to compensate for the decrease of efficiency in the steam cycle. The reason could be the use of fluid R-12. Based on the results of Hung et al. [1], R-12 is not a good candidate for this range of operational temperatures. The efficiencies are shown in Fig. 8. A further appropriate selection of working fluids for the ORC to absorb the heat from the condensing process of the steam Rankine cycle is possible in order to obtain better performance in thermal to power conversion.

**Parallel-Type Triple Cycles.** The parallel-type triple cycle proposed in this study is shown in Fig. 3, in which a Brayton cycle represents the topping cycle while a steam Rankine cycle and an ORC represent the bottoming cycle. As the temperature difference between the topping cycle and the bottoming cycle increases, available energy is utilized less and the irreversibility, i.e., the difference of work obtained from a practical (irreversible) process and an ideal (reversible) process, increases. However, in order to have an operation temperature of the bottoming cycle as close to the ambient temperature as possible, the inefficiently used energy (as indicated as the shaded area in Fig. 3) would increase and would result in a reduction of the overall efficiency. The purpose of this study is to raise the utilization of the waste heat

**Table 1 Reference condition for the states of triple cycles**

| Parameters and States | *Rp* | $T_3$ (K) | $T_5$ (K) | $P_6$ (MPa) | $T_9$ (K) | $\Delta T_{pp1}$ |
|---|---|---|---|---|---|---|
| Series type | 5 | 1300 | 353 | 10 | 293 | 10 |
| Parallel type | 15 | 1200 | 313 | 1.5 | -- | -- |

common baseline conditions: $P_1 = 0.1$ MPa; $T_1 = 300$ K; $T_{4c} = 423$ K; $\Delta T_{pp1} = 15$ K; power output of Brayton cycle = 10 MW

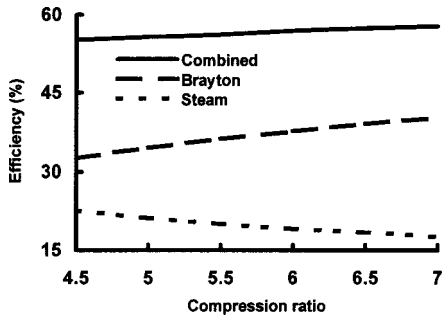



**Fig. 4 Effect of compression ratio on the efficiency in the combined cycle**

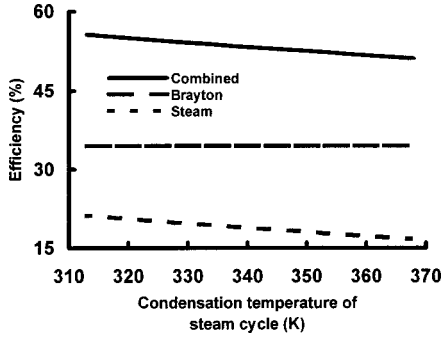



**Fig. 5 Variation of efficiency with the condensation temperature of the steam cycle in the combined cycle**

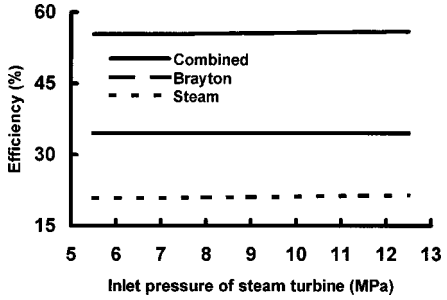



**Fig. 6 Variation of efficiency with the steam turbine inlet pressure in the combined cycle**

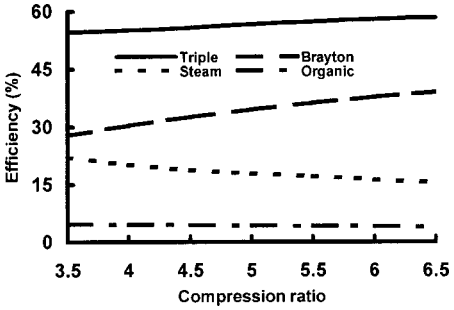



**Fig. 7 Effect of compression ratio on the efficiency in series-type triple cycle**

exhausted from the Brayton cycle by adding two or more cycles using appropriate working fluids according to their optimal operation temperatures in order to reduce the temperature difference between the topping and the bottoming cycles, and consequently, increase the overall efficiency.

As discussed previously, the series-type triple cycle using refrigerant R-12 as the working fluid in an ORC does not give a satisfactory performance. Thus, comparisons of the performance of the parallel-type triple cycles using other working fluids, i.e., $C_6H_6$, $NH_3$, and R-123, are discussed in this section.

The efficiency of an ORC using R-123 as a function of ORC turbine inlet pressure at a constant $T_{4b}$ is shown in Fig. 9. The exhausted waste heat from the gas turbine with temperature greater than $T_{4b}$ is transferred to the steam cycle or to the ORC. In order to have a proper flow rate to avoid adverse effects, e.g., a flow rate that is high enough to cause a turbine inlet temperature to drop into the saturation region or a flow rate that is too low to cause an ORC turbine inlet temperature to be higher than $T_{4b}$, the ranges of the ORC turbine inlet pressure have been properly adjusted according to the values of pinch point. For a fixed value of

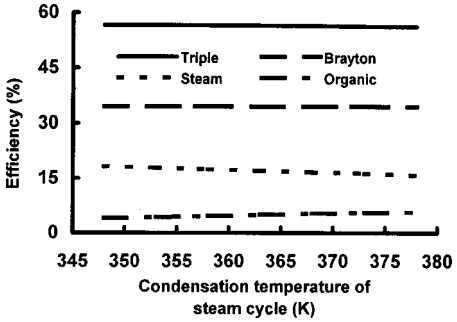



**Fig. 8 Variation of efficiency with the condensation temperature of steam cycle in series-type triple cycle**

pinch point, the efficiency of the ORC increases as the ORC turbine inlet pressure increases as a result of not only a rise in flow rate but also an increase of the total enthalpy received from the fluid. For different values of the pinch point, the effect of turbine inlet pressure on flow rate is insignificant. A lower value of pinch point would receive a higher quality of heat from the fluid and, thus a lower value of pinch point yields a higher efficiency. Therefore, under a fixed $T_{4b}$, an optimal efficiency can be achieved by properly adjusting operation ranges of turbine inlet pressure and pinch point.

The effects of $T_{4b}$ and pinch point on the efficiency of the ORC (using $C_6H_6$) as a function of turbine inlet pressure are shown in Fig. 10. It can be clearly seen that by increasing $T_{4b}$, i.e., an increase in the mean temperature of waste heat for the ORC, which efficiency increases accordingly. The effects of pinch point are also shown in Fig. 10 in which each fixed value of pinch point has its own operation range of the turbine inlet pressure. Each

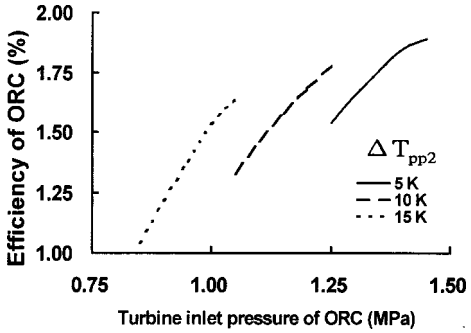



**Fig. 9 Variation of ORC efficiency corresponding to the turbine inlet pressure of ORC with $\Delta T_{pp2}$ as parameter when $T_{4b}$ is fixed in parallel-type triple cycle (R-123)**

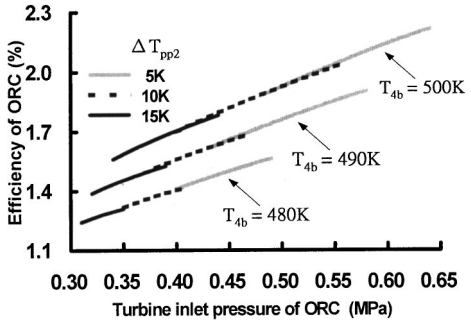



**Fig. 10 Variation of ORC efficiency corresponding to the turbine inlet pressure of the ORC with both $T_{4b}$ and $\Delta T_{pp2}$ as parameters in parallel-type triple cycle ($C_6H_6$)**

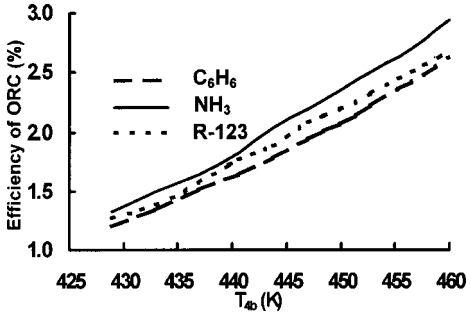



**Fig. 11 Variation of ORC efficiency corresponding to $T_{4b}$ associated with different fluids in parallel-type triple cycle**

fixed value of pinch point 2, Fig. 10, with $C_6H_6$ as working fluid, shows the same trend as that of Fig. 9 using R-123 as the working fluid.

The effects of working fluid on the efficiency of the ORC are shown in Fig. 11 in which the turbine inlet pressures and the values of pinch point have been adjusted to yield optimal efficiencies. It can be seen that $NH_3$ has the best performance among all fluids analyzed, and the efficiency increases as $T_{4b}$ increases.

As discussed previously, it can be seen that the parameter $T_{4b}$, the heat source for an ORC, plays an important role in the parallel-type triple cycle analysis. In this study, this parameter was calculated using two different calculation schemes, and was then adjusted in order to find an optimal performance for the parallel-type triple cycles.

*$T_{4b}$ as an Independent Variable.* In this calculation scheme, $T_{4b}$ is treated as an adjustable variable in order to vary the amount of waste heat transferred to the steam Rankine cycle. Based on the incoming heat source, the operation mode of the steam Rankine cycle can efficiently be regulated to yield an optimal performance. Furthermore, the effects of compression ratio and inlet temperature of the gas turbine were also studied as representative parameters in conjunction with the variation of $T_{4b}$ in order to find an optimal operation range.

Based on the conservation of energy in the heat recovery steam generator (HRSG) and under a fixed value of $T_{4b}$, the operation range of $P_6$ should be kept between 10 MPa and 19 MPa to avoid steam turbine inlet temperature higher than the gas turbine exit temperature, and a steam turbine exhaust with a high moisture content. As shown in Fig. 12(*a*), an increase in $P_6$ initially drops the efficiency of the steam Rankine cycle slightly but the efficiency increases monotonically as $P_6$ continues to increase. For example, for $\Delta T_{pp1}=35$ K, the flow rate of the steam Rankine cycle increases while the enthalpy content at the turbine inlet drops significantly when $P_6$ increases from 10 to 11 MPa. This causes a slight drop in efficiency. As $P_6$ exceeds 11 MPa, the enthalpy content at the turbine inlet changes only slightly while the increase of the flow rate becomes significant (see Fig. 12(*b*)), and results in an increase in efficiency, also shown in Fig. 12(*a*). The increase in efficiency is most evident for a small value of $\Delta T_{pp1}$. This is due to the smaller shaded area between the topping and the bottoming cycles as shown in the T-S diagram of Fig. 3.

Similar to combined-cycle system discussion previously, the efficiency of the Brayton cycle would increase if *Rp* increases under a constant $T_3$. Meanwhile, the steam Rankine cycle would receive an exhaust heat with a lower temperature and a low total energy content. Thus, it results in a lower efficiency of the steam Rankine cycle. For the ORC, its efficiency will increase slightly as a result of a lower flow rate, which yields a greater value of specific enthalpy at the turbine inlet. Both effects result in a slight increase of the ORC turbine work. From the system's point of view, under the condition of varying *Rp* only, the combined increase of the efficiencies of the Brayton cycle and the ORC overpowers the decrease of the efficiency of the steam Rankine cycle, as shown in Fig. 13. The increase of efficiency can be further amplified by using a higher value of $T_{4b}$, i.e., a higher energy absorption in the ORC than in the steam Rankine cycle. Also shown in Fig. 13, a higher value of $T_{4b}$ implies that a higher value of *Rp* range can be applied.

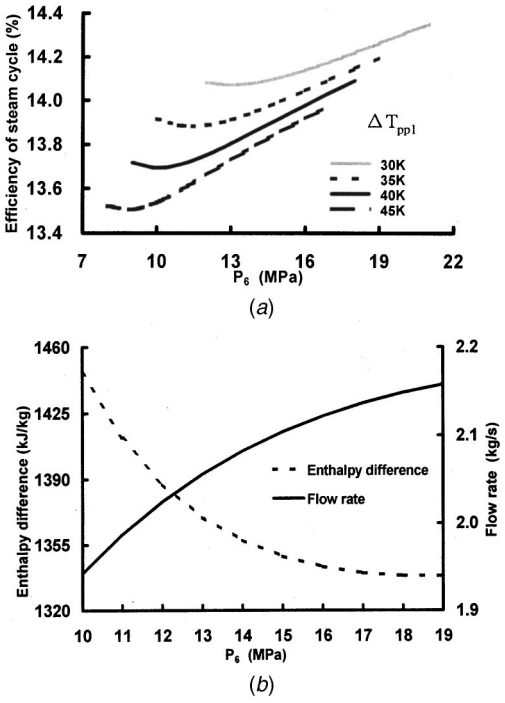



**Fig. 12 (*a*) Variation of steam-cycle efficiency corresponding to the evaporation pressure of steam cycle with $\Delta T_{pp1}$ as the parameter in the parallel-type triple cycle; (*b*) effect of the evaporation pressure of steam cycle on the enthalpy drop of steam turbine and the associated flow rate for $\Delta T_{pp1}=35$ K**

Under the condition of a fixed compression ratio, the effect of a higher gas turbine inlet temperature, $T_3$, is to raise the total thermal energy source and thus reduce the efficiency of the Brayton cycle for a constant gas turbine work output. As for the steam Rankine cycle under fixed *Rp* and $T_{4b}$, the effect of raising $T_3$ would increase its waste heat absorption from the Brayton cycle due a higher temperature gas turbine exhaust; consequently, the efficiency of the steam Rankine cycle is increased rather significantly. As for the ORC, an adverse effect will occur due to a drop of flow rate both in the Brayton cycle and the ORC. The reduction of flow rate in the Brayton cycle means a reduction of energy contained in the waste heat, which is transported and utilized in

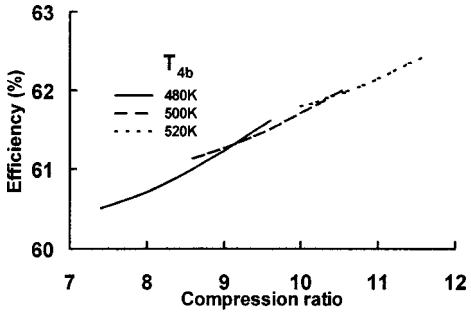



**Fig. 13 Variation of overall efficiency of the triple cycle as a function of compression ratio and $T_{4b}$ in a parallel-type triple cycle ($C_6H_6$)**

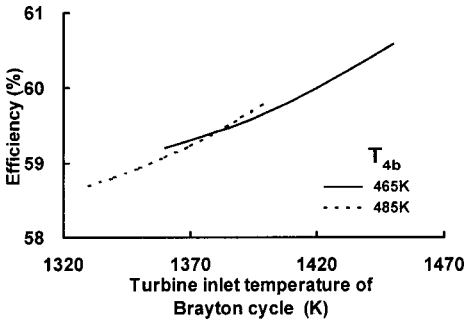



**Fig. 14 Variation of overall efficiency of triple cycle as a function of gas turbine inlet temperature and $T_{4b}$ in a parallel-type triple cycle ($C_6H_6$)**

the ORC. However, for the triple cycle as a whole, the effect of increasing $T_3$ increases the total system efficiency. Also shown in Fig. 14, the available range of $T_3$ is limited by the value of $T_{4b}$, i.e. a higher specified value of $T_{4b}$ would reduce the available range of $T_3$.

*$T_{4b}$ as a Dependent Variable Evaluated Under Specified Steam Turbine Inlet Temperature.* In this calculation scheme, $T_{4b}$ is calculated based on a specified turbine inlet temperature of the steam Rankine cycle. As the mass flow rate of the Brayton cycle is pre-determined, $T_{4b}$ is hence calculated accompanied with the determination of flow rate in the steam Rankine cycle. The values of $T_{4b}$ shown on the *x*-axis of Figs. 15–16 presented in this section were calculated based on various parametric studies. The compression ratio and the turbine inlet temperature of the Brayton cycle have been employed as representative parameters to determine $T_{4b}$ and then the performance of this triple cycle.

It has been well understood that a greater compression ratio results in a higher efficiency for the Brayton cycle. The drop of gas turbine exit temperature also means a drop of steam turbine inlet temperature and a lower specific enthalpy of the gas turbine exhaust. Due to a large heat capacity difference between air and steam, the steam flow rate in the steam Rankine cycle would drop appreciably and the efficiency of this cycle would drop also. The variations of efficiency for compression ratios between 10 and 20 are shown in Fig. 15. It can be seen that, in spite of a drop in efficiency of the steam Rankine cycle, the combined triple cycle exhibits a significant efficiency increase with an increase in both compression ratio (and accordingly $T_{4b}$).

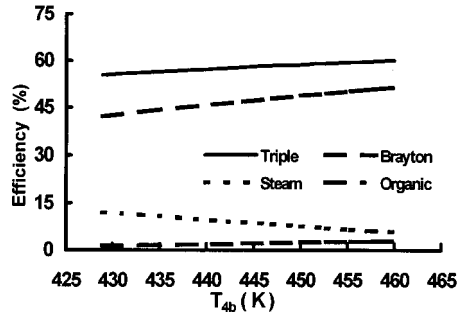



**Fig. 15 Variation of efficiency with compression-ratio dependent $T_{4b}$ in parallel-type triple cycle ($Rp$=10~20)**

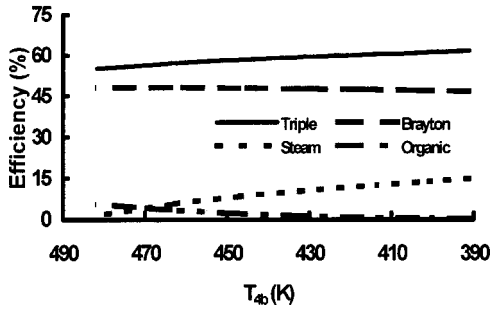



**Fig. 16 Variation of efficiency with $T_{4b}$, which is dependent on gas turbine inlet temperature (1000K ~ 1600K), in parallel-type triple cycle**

The effects of raising gas turbine inlet temperature on the efficiency of each individual cycle are virtually opposite to the effects of raising compression ratio. Under a constant compression ratio, as the gas turbine inlet temperature rises from 1000 K to 1600 K that resulted in a drop in efficiency of the Brayton cycle. However, the steam flow rate in the steam Rankine cycle increases due to a higher temperature and a greater amount of enthalpy in the waste heat exhausted from the gas turbine. Therefore, the efficiency of the steam Rankine cycle increases significantly. The higher capability in absorbing waste heat exhibited by the steam Rankine cycle also means a lower value of $T_{4b}$, i.e., it drops from 480 K to 390 K, and accordingly a lower efficiency for the ORC. However, the significant improvement exhibited by the steam Rankine cycle overcomes the adverse effects on the Brayton cycle and the ORC, and the triple cycle yields an improvement in efficiency as $T_3$ increase. This can be seen in Fig. 16.

## Conclusions

The efficiencies of combined cycles, series-type triple cycles, and parallel-type triple cycles were analyzed based on the effects of the thermophysical properties of the working fluids. Refrigerant R-12 was used in the ORC of series-type cycles while $NH_3$, $C_6H_6$, and Refrigerant R-123 were used in the ORC of parallel-type cycles. A number of parametric studies were carried out to search for optimal cycle combinations, according to key operational parameters, which will affect thermal efficiency. Additionally, two calculation schemes to separate both bottoming cycles of parallel-type triple cycle were used in performing the analysis.

The series-type triple cycle using R-12 as the working fluid was found to exhibit no significant difference compared with the combined cycle. This may be because R-12 is not a high-efficiency working fluid, or because of the temperature differences which occurred in the heat exchanger between the steam Rankine cycle and the ORC.

A further analysis of other fluids is recommended for this scheme. On the other hand, using other working fluids, i.e., R-123, $NH_3$, and $C_6H_6$, showed an increase in efficiency of the parallel-type triple cycle, especially in the application of recovering lower-enthalpy waste heat. Therefore, by properly combining a steam Rankine cycle and an ORC, an optimal configuration can be formed. The parametric study can be further applied in specific working environments and with specific hardware for future design applications. A systematic study and selection of working fluids for the ORC will be the subject of future research.

## Acknowledgment

The research reported in this paper was supported by the National Science Council, Taiwan, R.O.C. under the grants NSC 89-2212-E-214-001.

## Nomenclature

$I$ = irreversibility
$P$ = pressure
$Q$, $q$ = heat
$Rp$ = compression ratio
$S$ = entropy
$T$ = temperature
$W$, $w$ = work
$\Delta T_{pp}$ = pinch point

$\eta$ = efficiency
$\lambda$ = arbitrary parameter
$\xi$ = arbitrary intensive property

### Subscripts

bom = bottoming cycle
$H$ = high-temperature thermal reservoir
$k$ = index of thermal-energy reservoir
$L$ = low-temperature thermal reservoir
top = topping cycle
$i, j$ = index of subcycle

## References


[1] Hung, T. C., Shai, T. Y., and Wang, S. K., 1997, ''A Review of Organic Rankine Cycles (ORCs) for the Recovery of Low-Grade Waste Heat,'' Energy, **22**, pp. 661–667.
[2] Marciniak, T. J., Krazinski, J. L., Bratis, J. C., Bushby, H. M., and Buyco, E. H., 1981, ''Comparison of Rankine-Cycle Power System: Effects of Seven Working Fluids,'' Argonne National Laboratory Report, ANL/CNSV-TM-87.
[3] Schulitz, B., 1986, ''Working Fluids for ORC Plants: Add-On Circuits for Waste Heat Utilization,'' VGB Kaftwerkstechnik, **66**, pp. 419–426.
[4] Badr, O., Probert, D., and O'Callaghan, P. W., 1986, ''Selection of Operating Conditions and Optimization of Design Parameters for Multi-Vane Expanders,'' Appl. Energy, **23**, pp. 1–46.
[5] Drake, R. L., 1985, ''Organic Rankine Cycle Power Modules Recover Heat at Major Oil Refinery,'' Turbomachinery International, **26**, pp. 31–33.
[6] Chaudoir, D. W., Niggemann, R. E., and Bland, T. J., 1985, ''Solar Dynamic ORC Power System for Space Station Application,'' IEEE Service Cent (Cat No. 85CH2242-6), pp. 58–65.
[7] Manco, S., and Nervegna, N., 1985, ''Working Fluid Selection Via Computer Assisted Analysis of ORC Waste Heat Recovery Systems,'' IEEE Service Cent (Cat No. 85CH2242-6), pp. 71–83.
[8] Lee, M. J., Tien, D. L., and Shao, C. T., 1993, ''Thermophysical Capability of Ozone-Safe Working Fluids for an Organic Rankine Cycle System,'' J. Heat Recovery Syst., **13**, pp. 409–418.
[9] Habib, M. A., 1994, ''First- and Second-Law Analysis of Steam-Turbine Cogeneration Systems,'' ASME J. Eng. Gas Turbines Power, **116**, pp. 15–19.
[10] Lucia, M. D., Lanfranchi, C., and Boggio, V., 1996, ''Benefits of Compressor Inlet Air Cooling for Gas Turbine Cogeneration Plants,'' ASME J. Eng. Gas Turbines Power, **118**, pp. 598–603.
[11] Woodward, J. B., 1994, ''Evaluation of Brayton and Rankine Alternatives for Diesel Waste Heat Exploitation,'' ASME J. Eng. Gas Turbines Power, **116**, pp. 39–45.
[12] Briesch, M. S., Bannister, R. L., Diakunchak, I. S., and Huber, D. J., 1995, ''A Combined Cycle Designed to Achieve Greater Than 60 Percent Efficiency,'' ASME J. Eng. Gas Turbines Power, **117**, pp. 734–741.
[13] Horlock, J. H., 1992, *Combined Power Plants*, Pergamon Press, Oxford, UK.
[14] Tomlinson, L. O., Chase, D. L., Davidson, T. L., and Smith, R. W., 1993, ''GE Combined Cycle Product Line and Performance,'' GE Power Generation, GER-3574D.
[15] Marston, C. H. and Hyre, M., 1995, ''Gas Turbine Bottoming Cycles: Triple-Pressure Stem Versus Kalina,'' ASME J. Eng. Gas Turbines Power, **117**, pp. 10–15.