ASME
SETTING THE STANDARD

Journal of Engineering for Gas Turbines and Power

Published Bimonthly by ASME

VOLUME 131 • NUMBER 2 • MARCH 2009

Transactions of the ASME, Journal of Engineering for Gas Turbines and Power (ISSN 0742-4795) is published bimonthly (Jan., Mar., May, July, Sep, Nov.) 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).

RESEARCH 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.

(Contents continued)

Journal of Engineering for Gas Turbines and Power MARCH 2009

Volume 131, Number 2

(Contents continued on facing page)

Journal of Engineering for Gas Turbines and Power **MARCH 2009**

Volume 131, Number 2

TECHNICAL BRIEFS

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

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

Foreign Object Damage in an Oxide/Oxide Composite at Ambient Temperature

Sung R. Choi[1]
e-mail: sung.choi1@navy.mil

Donald J. Alexander

Robert W. Kowalik

Naval Air Systems Command,
Patuxent River, MD 20670

Foreign object damage behavior of an oxide/oxide (N720/AS) ceramic matrix composite was determined at ambient temperature using impact velocities ranging from 100 m/s to 400 m/s by 1.59 mm diameter steel-ball projectiles. Two different support configurations of target specimens were used: fully supported and partially supported. The degree of post-impact strength degradation increased with increasing impact velocity and was greater in a partially supported configuration than in a fully supported one. For the fully supported configuration, frontal contact stress played a major role in generating composite damage, while for the partially supported case, both frontal contact and backside flexure stresses were the combined sources of damage generation. The oxide/oxide composite was able to survive high energy (~1.3 J) impacts without complete structural failure. The degree of relative post-impact strength degradation of the oxide/oxide composite was similar to that of an advanced SiC/SiC composite observed from a previous study, regardless of the type of specimen support. Like the SiC/SiC composite, impact-damage tolerance was greater in the oxide/oxide than in monolithic silicon nitride ceramics for impact velocities >300 m/s. [DOI: 10.1115/1.2969091]

Keywords: oxide/oxide composite, ceramic matrix composite, foreign object damage (FOD), impact test, ballistic impact

1 Introduction

Ceramics, because of their brittle nature, are susceptible to localized surface damage and/or cracking when subjected to impact by foreign objects. It is also true that ceramic components may fail structurally even by soft particles when the kinetic energy of impacting objects exceeds certain limits. The latter case has been often encountered in aeroengines in which combustion products, metallic particles, or small foreign objects ingested cause severe damage to airfoil components, resulting in serious structural problems. Therefore, foreign object damage (FOD) associated with particle impact needs to be considered when ceramic materials are designed for structural applications. In view of this importance, a considerable amount of work on impact damage of brittle materials by sharp particles as well as by "blunt" particles or by plates has been done both experimentally and analytically, including the assessments of FOD for turbine engine applications [1–14].

In the previous studies [15–18], FOD behavior of two representative gas-turbine grade monolithic silicon nitrides, AS800 and SN282, was determined at ambient temperature using both flexure bars and disks. Fully supported ceramic target specimens were impacted at their centers by *steel-ball projectiles* with a diameter of 1.59 mm in a velocity range from 220 m/s to 440 m/s. AS800 silicon nitride exhibited a greater FOD resistance than SN282. The key material parameter affecting FOD most was found to be fracture toughness of a material. No single crack system was involved in impact with increasing impact velocity, resulting in several different types of flaw systems associated individually or simultaneously. Assessments of FOD in AS800 silicon nitride were also made using four different projectiles of 1.59-mm-diameter hardened steel, annealed steel, silicon nitride, and brass balls [19,20]. For a given target material, the hardness of projectile materials was identified as the most important property to control impact damage. The work was extended to include FOD behavior of a gas-turbine grade, melt-infiltrated (MI) Sylramic™ SiC/SiC ceramic matrix composite (CMC) at 25°C and 1316°C employing two different types of target-specimen support [21]. Unlike monolithic silicon nitrides, the MI SiC/SiC composite exhibited monotonic degradation of post-impact strength with increasing impact velocity without causing complete structural-type of failure. The degree of FOD was significantly greater in partial support than in full support. The weakening effect of the material in response to FOD resistance was also evident at 1316°C.

The current work, as an extension of the previous studies, investigates FOD behavior of a gas-turbine grade N720™ oxide fiber-reinforced aluminosilicate matrix CMC (N720/AS). This two-dimensional (2D) woven oxide/oxide composite has been considered in exhaust nozzle applications. Oxide/oxide target specimens in a flexure bar configuration were impacted at velocities from 100 m/s to 400 m/s by 1.59-mm-diameter steel-ball projectiles at ambient temperature. As done for the MI SiC/SiC composite in the previous study [21], two different types of specimen support, fully supported and partially supported, were used. Post-impact strength of each target specimen impacted was determined in four-point flexure as a function of impact velocity to evaluate the severity of impact damage. Fractography was performed before and after post-impact strength testing to determine the impact-damage morphologies and related crack systems. In view of rarity in this type of FOD study for CMCs, to the authors' best knowledge, this work would be considered as a first *complete* investigation of FOD behavior of an oxide/oxide material system.

2 Experimental Procedures

2.1 Material and Test Specimens. The material used in this work was 0 deg/90 deg, 2D woven, N720™ fiber-reinforced aluminosilicate matrix CMCs. The oxide/oxide composite was fabricated by GEAE (vintage '99), supplied in a form of about $190 \times 190 \times 3$ mm^3 panels. The oxide/oxide panels were com-

[1]Corresponding author.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 9, 2008; final manuscript received April 10, 2008; published online December 24, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008 Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

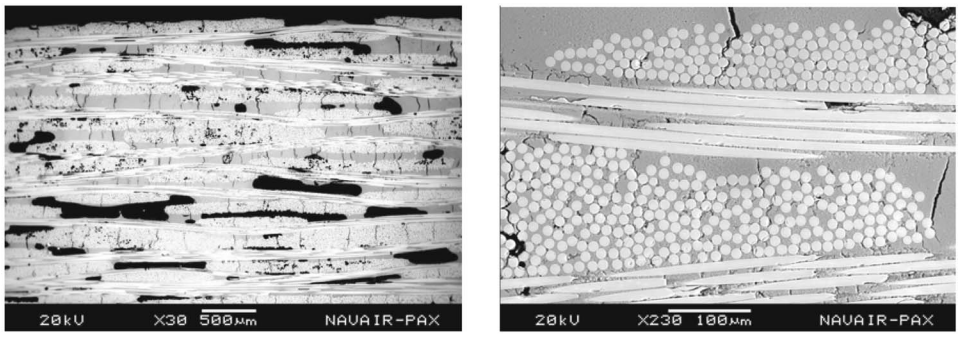


Fig. 1 SEM micrographs of oxide/oxide (N720/AS) ceramic matrix composite used in this work

posed of 12 plies with a fiber volume of about 45%. N720™ oxide fibers, produced in tow form by 3M Corp. (Minneapolis, MN), were woven into 2D eight harness-satin cloth. The cloth was cut into a proper size, slurry-infiltrated with the matrix, and 12 plystacked followed by vacuum bagging/consolidation and sintering at high temperature. No interface fiber coating was made. Typical scanning electron microcopy (SEM) micrographs of the composite are shown in Fig. 1. The presence of significant porosity and microcracks in the matrix is typical for the material, which effectively increases the damage tolerance of this class of porous oxide/oxide composites [22,23]. Porosity was found to be around 20–25%. The bulk density was 2.55 g/cm^3, and the elastic modulus was 67 GPa, determined by the impulse excitation of vibration technique [24]. The panels were machined into flexure bars measuring 10 mm in width, 50 mm in length, and 3 mm in as-furnished thickness for FOD testing.

2.2 Foreign Object Damage Testing. FOD testing was carried out using the experimental apparatus shown in Fig. 2. Detailed descriptions of the apparatus can be found elsewhere [15–18]. Hardened (HRC $\geq$ 60) chrome steel-balls with a diameter of 1.59 mm were inserted into a 300-mm-long gun barrel with an inner diameter of 1.59 mm. A helium-gas cylinder and relief valves were used to pressurize the reservoir to a specific level, depending on prescribed impact velocity. Upon reaching a specific level of pressure, a solenoid valve was instantaneously opened accelerating a steel-ball projectile through the gun barrel to impact a target specimen. The target specimen was fully or partially supported through a SiC specimen holder, as shown in Fig. 3. Each target specimen was aligned such that the projectile impacted at the center of the specimen with a normal incidence angle.

Impact velocity of each projectile was determined using two pairs of laser transmitter and receiver, incorporated with two holes in the gun barrel, as described before [15–18]. The range of impact velocity employed in this work was from 100 m/s to 400 m/s. Two test specimens were used at each velocity for a given specimen support. Impact morphologies of target specimens and some recovered projectiles were examined after impact using optical and scanning electron microscopy.

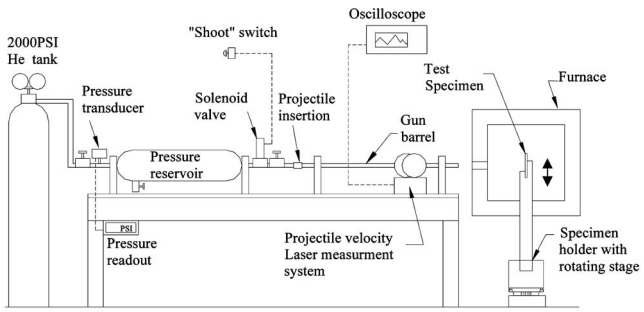


Fig. 2 Impact testing apparatus [15]

2.3 Post-Impact Strength Testing. Strength testing for impacted target specimens was carried out at ambient-temperature in air to determine the severity of impact damage using a four-point flexure fixture with 20 mm inner and 40 mm outer spans, see Fig. 3(*b*). Each impacted specimen was loaded in the flexure fixture such that its impact site was subjected to tension within the inner span. A linear variable differential transformer (LVDT) was used to determine the center deflection of specimens during strength testing. An electromechanical test frame (Model 8562, Instron, Canton, MA) was used in displacement control with an actuator speed of 0.5 mm/min. Fractographic analysis was performed after post-impact strength testing to determine failure origin, flaw configuration, and mode of fracture, if possible. "As-received" flexural strength of the composite was also determined with a total of six test specimens using the same test fixture, test frame, and test conditions that were utilized in the post-impact strength testing.

3 Results and Discussion

3.1 Post-Impact Strength. The results of strength testing for impacted target specimens are shown in Fig. 4, where post-impact flexure strength was plotted as a function of impact velocity for both types of specimen support. Included in the figure is the as-received ("As-R") flexure strength (=141 $\pm$ 4 MPa) determined. All the specimens impacted did fracture from impact sites. Post-impact strength decreased with increasing impact velocity, regard-

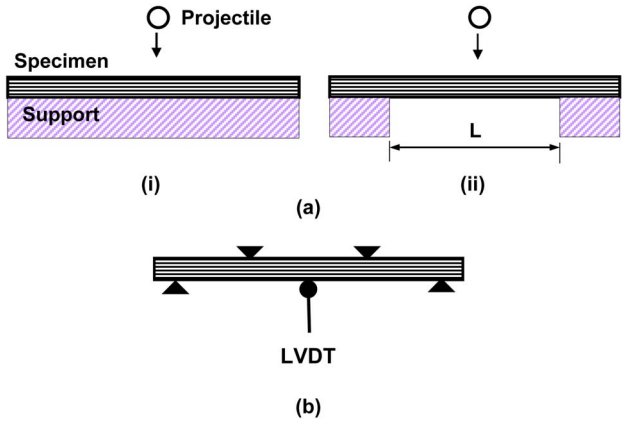


Fig. 3 (*a*) Two types of target-specimen support used in FOD testing: (i) fully supported and (ii) partially supported (*L* =20 mm); (*b*) four-point flexure configuration used in post-impact strength testing with a LVDT placed

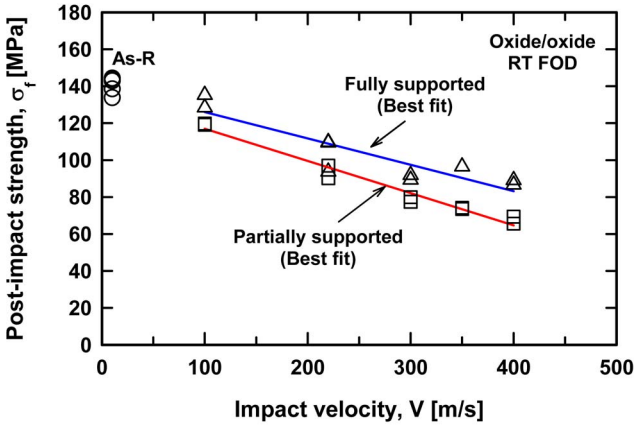


Fig. 4 Post-impact flexural strength as a function of impact velocity, determined for oxide/oxide composite impacted by 1.59-mm-diameter steel-ball projectiles with two different types of specimen support. As-R represents as-received strength with no impact.

less of the type of specimen support, attributed to increased impact damage. For a given impact velocity, the strength degradation with respect to the as-received strength was greater for the partially supported specimens than for the fully supported ones. The impact damage for the fully supported specimens was observed to be associated with frontal contact stresses, while the damage for the partially supported ones resulted from both frontal contact and backside flexure stresses, thereby resulting in more significant strength degradation in the partially supported. Strength degradation at 400 m/s was about 38% and 52%, respectively, for the fully and the partially supported. Impact damage morphology will be described in more details in Sec. 3.2.

A similar trend in strength degradation was also observed from the previous study using a MI Sylramic™ fiber-reinforced SiC matrix composite [21], as shown and compared in Fig. 5. Since as-received strength (=578 MPa in flexure) of the MI SiC/SiC composite was much greater than that (=141 MPa) of the oxide/oxide composite used here, it is necessary to normalize post-impact strength with respect to their as-received strengths so that a trend in relative strength degradation can be directly compared. Figure 5(*b*) compares normalized strength as a function of impact velocity for the two composites. Surprisingly, for a given type of specimen support, the degree of strength degradation was almost identical in either the SiC/SiC or the oxide/oxide composite. For both the composites, the degree of strength degradation was greater in the partial support than in the full support.

3.2 Impact Morphology. It has been observed that steel-ball projectiles were flattened or severely deformed or fragmented, depending on impact velocity, upon impacting monolithic silicon nitrides or the MI SiC/SiC composite, although the overall projectile damage was less in the MI SiC/SiC composite than in monolithic silicon nitride ceramics [15–21]. By contrast, the projectiles that impacted the oxide/oxide composite in this work were not flattened or noticeably deformed even at the highest impact velocity of 400 m/s, regardless of the type of specimen support, which is shown in Fig. 6. This is due to the material's "softness" and open structure, compared with the hard-and-dense MI SiC/SiC or monolithic AS800 and SN282 silicon nitride counterparts. Sometimes, at higher impact velocities $\geq$300 m/s, projectiles were embedded into the oxide/oxide composite, similar to the case occurring in which ceramic "sharp" particles impacted soft metal targets.

The frontal impact damages generated in target specimens were in the form of indents, or craters, and/or spallation with their size depending on impact velocity. Figure 7 shows the frontal impact-

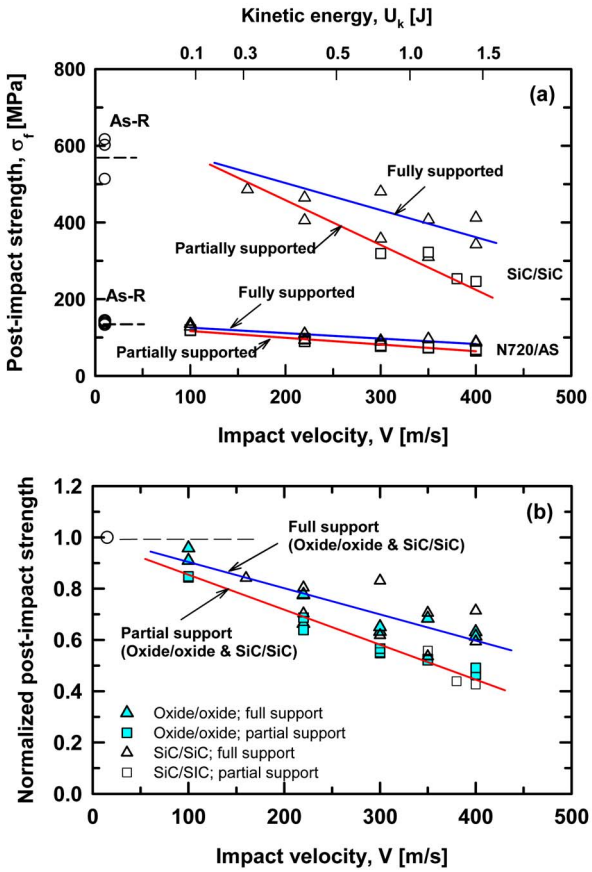


Fig. 5 (*a*) Comparison of post-impact flexural strength as a function of impact velocity between the oxide/oxide (this study) and the MI SiC/SiC composites [21], impacted by 1.59-mm-diameter steel-ball projectiles with two different types of specimen support. (*b*) Comparison in normalized post-impact strength between the two composites.

damage size as function of impact velocity. For comparison, the data on the SiC/SiC composite [21] were also included. The size of damage for the oxide/oxide increased monotonically with increasing impact velocity and was independent of the type of specimen support. By contrast, the MI SiC/SiC composite exhibited increased sensitivity of damage size to impact velocity. This is again indicative of the oxide/oxide's softness/openness with which impact size was fully developed early on at low impact velocities $\leq$200 m/s. However, it should be noted that the frontal impact size alone does not contribute to post-impact strength since the impact damage was associated with several different features such as frontal damage, delamination, and/or backside cracking, depending on the type of specimen support, as will be elaborated below. Although not detailed here, it was observed that the mate-

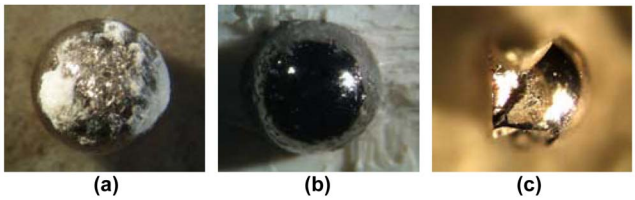


Fig. 6 1.59-mm-diameter steel-ball projectiles after impact on (*a*) oxide/oxide at 400 m/s (in full support), (*b*) oxide/oxide at 400 m/s (in partial support), and (*c*) AS800 silicon nitride at 350 m/s (in full support) [15]

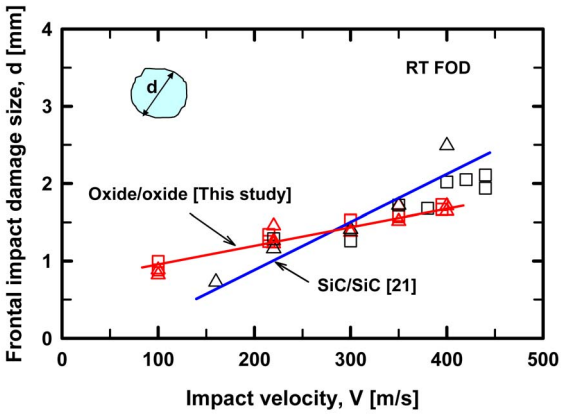


Fig. 7 Frontal impact-damage size as a function of impact velocity for the oxide/oxide composite in fully and partially supported configurations. The data on the MI SiC/SiC composite [21] were included for comparison.

rial beneath the impact site was subjected to densification/compaction, due to the material's porous nature [25]. The frontal impact-damage size (d in millimeters) in the oxide/oxide composite was approximated with impact velocity (V in m/s) as follows:

$$d \approx 2 \times 10^{-3}\ \mathrm{V} + 0.7 \tag{1}$$

At lower impact velocity (<200 m/s), the impact damage was typically associated with generation of frontal impact site with a little interlaminar delamination. With increasing impact velocity, frontal impact size, cone cracking, and delamination increased in the full support, whereas both additional backside cracking and delamination were augmented in the partial support. At higher impact velocities ≥ 350 m/s, backside damage became significant for the partially supported specimens with a presence of severe backside cracking and delamination, due to the increased backside tensile stress. Typical examples showing such impact damage at a high impact velocity of 350 m/s for both fully and partially supported specimens are presented in Fig. 8. Progressive impact damage with increasing impact velocity for both fully supported and partially supported specimens are illustrated in Fig. 9. Also, cross-sectional views of impact sites are shown in Fig. 10. It is noted from Fig. 10 that regardless of material, cone cracking

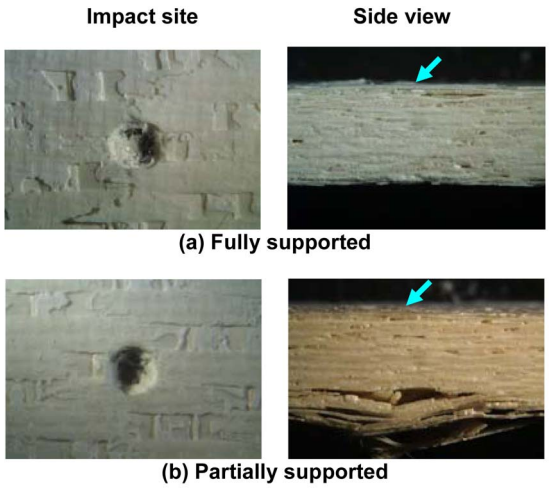


Fig. 8 Typical examples of damage showing impact sites and sides of specimens at 350 m/s: (*a*) fully supported specimen; (*b*) partially supported specimen. The arrows indicate the impact sites.

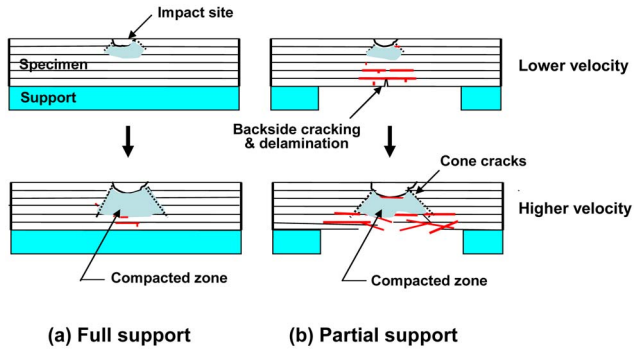


Fig. 9 Progression of impact damage with velocity in (*a*) fully supported and (*b*) partially supported specimens. The vertical arrow indicates a case of increasing impact velocity.

was one of the dominant impact-damage mechanisms, which can be seen in the figure from the SiC/SiC composite and the AS800 silicon nitride [18]. In particular, note that a complete cone has been separated from the AS800 target specimen, an evidence of the occurrence of a well-developed cone crack upon impact. More detailed analysis on impact damage and morphology of the oxide/oxide composite is needed using pertinent methodologies. The optical and scanning electron microscopies may not be sufficient for ceramic matrix composites in some cases, in which nondestructive techniques such as computed tomography (CT) and/or pulsed thermography techniques should be sought to better characterize impact damage of the materials [26].

3.3 Load-Displacement Curves. Figure 11 shows the load-deflection curves determined from post-impact strength testing for fully and partially supported specimens. As impact velocity increased, both stiffness and strength decreased. For a given impact velocity, the decrease in stiffness (i.e., the slope of the curves) was much greater in the partially supported than in the fully supported, due to the added damage of backside cracking and delamination as aforementioned. It is also noted that despite some strength degradation at 100 m/s due to frontal damage, overall integrity of the material is still retained, resulting in almost unchanging stiffness when compared with the as-received virgin material. The shape of load-displacement curves, together with strength and microstrucural observations, provides consistent support as to how the target material responded to projectile impact in the velocity range explored.

3.4 Comparison with Other Ceramics. Figure 12 shows the normalized post-impact strength ($\bar{\sigma}_f$) as a function of impact ki-

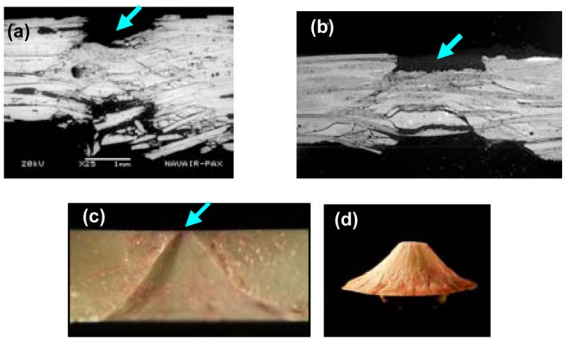


Fig. 10 Cross-sectional views of impact sites showing occurrence of cone cracking in (*a*) oxide/oxide (400 m/s), (*b*) SiC/SiC (400 m/s), (*c*) AS800 silicon nitride at 400 m/s [18], and (*d*) a cone separated from (*c*). The arrows indicate the impact sites.

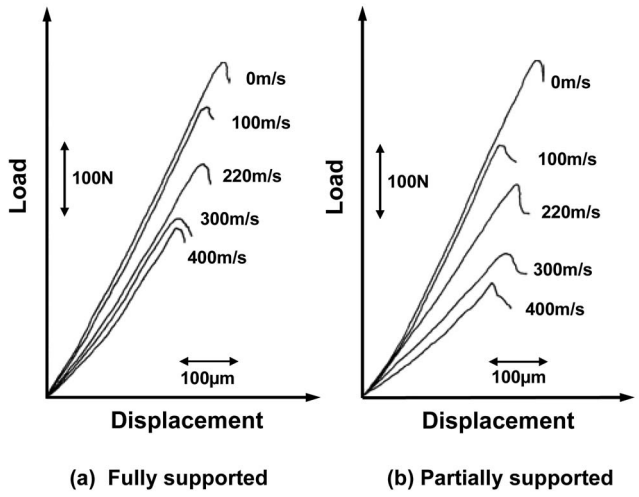


Fig. 11 Load-deflection curves determined in flexure for both fully and partially supported specimens impacted from 0 m/s to 400 m/s. The curves with zero velocity indicate the as-received specimen with no impact.

netic energy (U_k) for the current oxide/oxide, MI SiC/SiC composite [21], and AS800 and SN282 silicon nitrides [16], all impacted in a fully supported configuration. A monotonic degradation in strength with increasing impact energy was typified for the oxide/oxide and the SiC/SiC composites, yielding a linear relationship between log ($\bar{\sigma}_f$) and log (U_k). By contrast, monolithic AS800 and SN282 silicon nitrides exhibited greater strength degradation, approaching their respective "critical" impact energies of about 1.3 J (with V=400 m/s) and 0.8 J (at V =310 m/s), in which flexure test specimens fractured completely upon impact. However, at lower kinetic energy, $U_k \leq 0.4$ J (corresponding to V=220 m/s), the two silicon nitrides showed a negligible strength degradation, while the two composites yielded about 20–30% degradation. The best-fit relation for the two CMCs was approximated as follows:

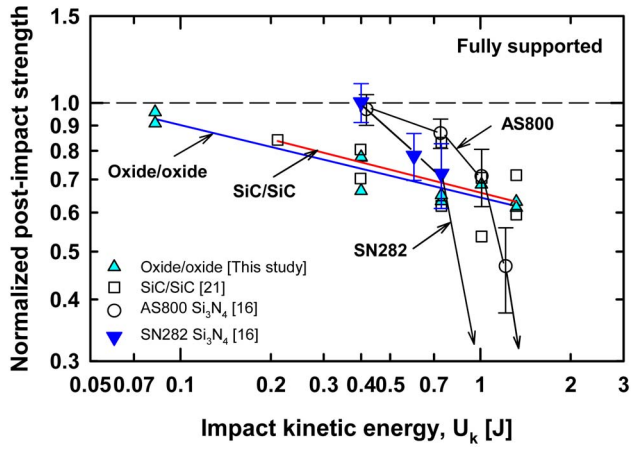


Fig. 12 Comparison of normalized post-impact strength as a function of impact kinetic energy among the oxide/oxide, the MI SiC/SiC [21], and AS800 and SN282 silicon nitrides [16], impacted by 1.59-mm-diameter steel projectiles in a fully supported configuration

$$\sigma_f = 0.65\ U_k^{-0.15}[\sigma_{f/\mathrm{as}}] \tag{2}$$

where σ_f is the post-impact strength, $\sigma_{f/\mathrm{as}}$ is the as-received strength with no impact, and the impact kinetic energy is expressed as

$$U_k = 1/2\ (mV^2) \tag{3}$$

where m is the mass of the projectile. The Hertzian contact/fracture mechanics analysis for dense brittle homogeneous elastic solids gives the following [1]:

$$\sigma_f = \Phi'' U_K^{-0.20} \tag{4}$$

The oxide/oxide and the MI SiC/SiC composites showed reasonable agreement in slope with the Hertzian contact theory, although no physical insight was provided in this work. The previous studies [15–18] indicated that AS800 and SN282 silicon nitrides exhibited a significant deviation from the contact theory, as also can be seen from Fig. 12.

3.5 Further Considerations. Designing aeroengine components to withstand FOD events is a complex task. Consideration of many factors is required, both in the generation of FOD data as well as in actual component design efforts [27]. A sample of these various factors includes the following:

- effect of projectile material/geometry
- effect of test-specimen material/geometry/architecture
- effect of test-specimen support and component attachment
- effect of temperature/environment
- effect of FOD damage on continued exposure to elevated temperature and/or environment
- appropriate protective coatings
- geometrical design of components to enhance FOD resistance
- modeling of FOD, reliability, and life prediction

4 Conclusions

The overall impact damage of the oxide/oxide composite was found to be greater in the partially supported specimens than for the fully supported ones.

For the fully supported specimens, frontal contact stresses together with cone cracking played a major role in generating composite damage, whereas for partially supported ones, both frontal contact and backside flexure stresses were combined sources of damage generation. At higher impact, velocity in full support was also characterized due to the porous nature of the composite.

The degree of relative strength degradation upon impact for the oxide/oxide composite was similar to that of the MI SiC/SiC CMC observed previously, either in the full support or in the partial support.

Unlike monolithic AS800 and SN282 silicon nitrides, the oxide/oxide composite did not exhibit a critical impact velocity in which a target-specimen fractures completely into two pieces. The same was true for the MI SiC/SiC observed from a previous study. This indicates that enhanced FOD resistance can be achieved with these CMCs. However, the damage generated in the oxide/oxide composite at $\geq$350 m/s should be taken into account with care when aeroengine components are in a partially supported configuration, from a structural and functional point of view.

Acknowledgment

This work was supported by the Aircraft Propulsion Materials Project, the Office of Naval Research.

References

[1] Wiederhorn, S. M., and Lawn, B. R., 1977, "Strength Degradation of Glass Resulting From Impact With Spheres," J. Am. Ceram. Soc., **60**(9-10), pp. 451–458.
[2] Wiederhorn, S. M., and Lawn, B. R., 1979, "Strength Degradation of Glass Impact With Sharp Particles: I. Annealed Surfaces," J. Am. Ceram. Soc., **62**(1-

2), pp. 66–70.
[3] Ritter, J. E., Choi, S. R., Jakus, K., Whalen, P. J., and Rateick, R. G., 1991, "Effect of Microstructure on the Erosion and Impact Damage of Sintered Silicon Nitride," J. Mater. Sci., **26**, pp. 5543–5546.
[4] Akimune, Y., Katano, Y., and Matoba, K., 1989, "Spherical-Impact Damage and Strength Degradation in Silicon Nitrides for Automobile Turbocharger Rotors," J. Am. Ceram. Soc., **72**(8), pp. 1422–1428.
[5] Knight, C. G., Swain, M. V., and Chaudhri, M. M., 1977, "Impact of Small Steel Spheres on Glass Surfaces," J. Mater. Sci., **12**, pp. 1573–1586.
[6] Rajendran, A. M., and Kroupa, J. L., 1989, "Impact Design Model for Ceramic Materials," J. Appl. Phys., **66**(8), pp. 3560–3565.
[7] Taylor, L. N., Chen, E. P., and Kuszmaul, J. S., 1986, "Microcrack-Induced Damage Accumulation in Brittle Rock Under Dynamic Loading," Comput. Methods Appl. Mech. Eng., **55**, pp. 301–320.
[8] Mouginot, R., and Maugis, D., 1985, "Fracture Indentation Beneath Flat and Spherical Punches," J. Mater. Sci., **20**, pp. 4354–4376.
[9] Evans, A. G., and Wilshaw, T. R., 1977, "Dynamic Solid Particle Damage in Brittle Materials: An Appraisal," J. Mater. Sci., **12**, pp. 97–116.
[10] Liaw, B. M., Kobayashi, A. S., and Emery, A. G., 1984, "Theoretical Model of Impact Damage in Structural Ceramics," J. Am. Ceram. Soc., **67**, pp. 544–548.
[11] van Roode, M., Jimenez, O., McClain, O., Price, J., Parthasarathy, V., Poormon, K. L., Ferber, M. K., and Lin, H-T., 2002, "Ceramic Gas Turbine Materials Impact Evaluation," ASME Paper No. GT2002-30505.
[12] Richerson, D. W., and Johansen, K. M., 1982, "Ceramic Gas Turbine Engine Demonstration Program," Final Report, DARPA/Navy Contract No. N00024-76-C-5352, Garrett, Report No. 21-4410.
[13] Boyd, G. L., and Kreiner, D. M., 1987, "AGT101/ATTAP Ceramic Technology Development," *Proceedings of the 25th Automotive Technology Development Contractors' Coordination Meeting*, Dearborn, MI, October 26–29, 1987, SAE, Warrendale, PA, p. 101.
[14] van Roode, M., Brentnall, W. D., Smith, K. O., Edwards, B., McClain, J., and Price, J. R., 1997, "Ceramic Stationary Gas Turbine Development—Fourth Annual Summary," ASME Paper No. 97-GT-317.
[15] Choi, S. R., Pereira, J. M., Janosik, L. A., and Bhatt, R. T., 2002, "Foreign Object Damage of Two Gas-Turbine Grade Silicon Nitrides at Ambient Temperature," Ceram. Eng. Sci. Proc., **23**(3), pp. 193–202.
[16] Choi, S. R., Pereira, J. M., Janosik, L. A., and Bhatt, R. T., 2004, "Foreign Object Damage in Flexure Bars of Two Gas-Turbine Grade Silicon Nitrides," Mater. Sci. Eng., A, **379**, pp. 411–419.
[17] Choi, S. R., Pereira, J. M., Janosik, L. A., and Bhatt, R. T., 2003, "Foreign Object Damage of Two Gas-Turbine Grade Silicon Nitrides in a Thin Disk Configuration," ASME Paper No. GT2003-38544.
[18] Choi, S. R., Pereira, J. M., Janosik, L. A., and Bhatt, R. T., 2004, "Foreign Object Damage in Disks of Gas-Turbine-Grade Silicon Nitrides by Steel Ball Projectiles at Ambient Temperature," J. Mater. Sci., **39**, pp. 6173–6182.
[19] Choi, S. R., Racz, Z., Bhatt, R. T., Brewer, D. N., and Gyekenyesi, J. P., 2005, "Effect of Projectile Materials on Foreign Object Damage of a Gas-Turbine Grade Silicon Nitride," ASME Paper No. GT2005-68866.
[20] Choi, S. R., Racz, Z., Bhatt, R. T., and Brewer, D. N., 2006, "Foreign Object Damage in a Gas-Turbine Grade Silicon Nitride by Spherical Projectiles of Various Materials," NASA Glenn Research Center, NASA, Report No. TM-2006-214330.
[21] Choi, S. R, Bhatt, R. T., Perrira, J. M., and Gyekenyesi, J. P., 2004, "Foreign Object Damage Behavior of a SiC/SiC Composite at Ambient and Elevated Temperatures," ASME Paper No. GT2004-53910.
[22] Mattoni, M. A., Yang, J. Y., Levi, C. G., Zok, F. W., and Zawada, L. P., 2005, "Effects of Combustor Rig Exposure on a Porous-Matrix Oxide Composite," Int. J. Appl. Ceram. Technol., **2**(2), pp. 133–140.
[23] Simon, R. A., 2005, "Progress in Processing and Performance of Porous-Matrix Oxide/Oxide Composites," Int. J. Appl. Ceram. Technol., **2**(2), pp. 141–149.
[24] ASTM C 1259, 2007, "Test Method for Dynamic Young's Modulus, Shear Modulus, and Poisson's Ratio for Advanced Ceramics by Impulse Excitation of Vibration," *Annual Book of ASTM Standards*, Vol. 15.01, ASTM, West Conshohocken, PA.
[25] Choi, S. R., and Alexander, D. J., 2008, "Characterization of Foreign Object Damage in an Oxide/Oxide Composite at Ambient Temperature," Ceram. Eng. Sci. Proc., to be published.
[26] Cosgriff, L. M., Bhatt, R., Choi, S. R., and Fox, D. S., 2005, "Thermographic Characterization of Impact Damage in SiC/SiC Composite Materials," Proc. SPIE, **5767**, pp. 363–372.
[27] Peralta, A. D., and Yoshida, H., 2003, *Ceramic Gas Turbine Component Development and Characterization*, Vol. 2, M. van Roode, M. K. Ferber, and D. W. Richerson, eds., ASME, New York, pp. 665–692.

Experimental Study on Laser-Induced Ignition of Swirl-Stabilized Kerosene Flames

Klaus G. Moesl[1]
e-mail: moesl@td.mw.tum.de

Klaus G. Vollmer

Thomas Sattelmayer

Lehrstuhl für Thermodynamik,
Technische Unversität München,
D-85748 Garching, Germany

Johannes Eckstein

Herbert Kopecek

Alternative Energy and Environmental Technologies,
Global Research Center,
General Electric,
D-85748 Garching, Germany

Conventional ignition systems of aeroengines are an integral part of the combustion chamber's structure. Due to this hardware-related constraint, the ignition spark has to be generated in the quench zone of the combustion chamber, which is far from the optimum regarding thermo- and aerodynamics. An improved ignitability of the fuel-air mixture can be found in the central zone of the combustor, where higher local equivalence ratios prevail and where mixing is favorable for a smooth ignition. It would be a major advancement in aeroengine design to position the ignition kernel in these zones. A laser system is able to ignite the fuel-air mixture at almost any location inside of the combustion chamber. Commercial laser systems are under development, which can replace conventional spark plugs in internal combustion engines and gas turbines. This study was conducted to evaluate the applicability of laser ignition in liquid-fueled aeroengines. Ignition tests were performed with premixed natural gas and kerosene to evaluate the different approaches of laser and spark plug ignition. The experiments were carried out on a generic test rig with a well-investigated swirler, allowing sufficient operational flexibility for parametric testing. The possibility of the free choice of the laser's focal point is the main advantage of laser-induced ignition. Placing the ignition kernel at the spray cone's shear layer or at favorable locations in the recirculation zone could significantly increase the ignitability of the system. Consequently, the laser ignition of atomized kerosene was successfully tested down to a global equivalence ratio of 0.23. Furthermore, the laser outperformed the spark plug at ignition locations below axial distances of 50 mm from the spray nozzle. [DOI: 10.1115/1.2981181]

1 Introduction

Since spark plugs (SPs) are an integral part of the combustor liner, the ignition kernel is usually located in the suboptimal quench zone of the combustor. Lean mixtures along the liner increase the demand on ignition energy, leading to an increased erosion of the spark plug electrodes, and thus to a reduced reliability and lifetime of the igniter. The electrodes of the spark plug must moreover be located close to the combustor wall to avoid disturbance of the precisely designed aerodynamic flow. This, in turn, requires inclusion of the limitations of possible ignition locations into the process of combustor design, which is a further restriction for the development engineer. Since spark plug ignition shows a reduced ignitability of lean mixtures below an equivalence ratio of $\Phi=0.6$ on the one hand and lean premixed prevaporized (LPP) combustion will become increasingly important in aeroengines on the other, nonconventional ignition sources, such as lasers, can contribute to a future performance optimization.

Laser ignition is a possible candidate to solve the mentioned problems because it allows uncoupling of the limiting link between the location of the ignition source and the ignition kernel. Lasers are able to ignite the mixture at the best thermodynamic and aerodynamic conditions from almost any installation location. Therefore laser ignition is more independent from variations of the local equivalence ratio than other ignition concepts. It is known that lasers are able to ignite leaner mixtures compared with spark plug ignition because there are no electrodes surrounding the initial flame kernel, which, in the case of the spark plug, cool down the kernel and prevent it from evolving further into the combustion chamber [1,2]. In the case of fuel spray ignition in direct injection (DI) gasoline engines, the laser-induced plasma kernel could successfully be positioned at the border of the fuel spray, utilizing an optimum equivalence ratio for a successful ignition [3,4]. The optical window needed to transmit the laser beam into the combustion chamber has to be kept clean, which seems to be the major challenge of this technology. However, several approaches are feasible to manage this issue, including a heating of the transmission window as well as a laser head position where no critical wetting effect will occur. Indeed, the utilized laser unveiled positive lens-cleaning characteristics during the experiment campaign, which were capable of counteracting soot deposition and lens wetting. In particular for optical systems, such as a laser setup, the self-cleaning ability is of major importance to maintain the system's efficiency over time.

Three different concepts of laser-induced ignition are generally feasible in order to ignite a combustible fuel-air mixture [5,6].

- Linear absorption of the laser radiation by the gaseous mixture, leading to a heating of the gas and, as a consequence, to an ignition of the surrounding gas.
- Absorption of the laser radiation on the surface of a solid body that is positioned inside the combustion chamber. The heating of the body leads to a local heating and ignition of the mixture.
- Nonlinear breakdown of the gaseous media by high field laser radiation, which yields a spark inside the mixture. Hence, this very hot plasma ball initiates ignition of the local mixture.

The aim of this study was to exclusively investigate the concept of high field laser radiation. To reach the necessary peak intensities, a *Q*-switched Nd:YAG (yttrium aluminum garnet) laser was

[1]Corresponding author.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 28, 2008; final manuscript received April 28, 2008; published online December 22, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

Table 1 Comparison: spark plug versus laser ignition

	Spark plug	Laser
State of the art	High development level	Laser ignition– a "young technology"
System durability	Extensive experience Lifetime: 0.3×10^9 pulses	Successful feasibility studies Expected lifetime: 10^9 pulses
Range of output energy	100–1000 mJ	10–100 mJ
Installation location of system	Limitations	No limitations
Location of ignition kernel	Directly at liner (quench zone)	Free spatial location

applied, which is reliable and robust. Gas breakdown and spark generation were possible for almost any location inside the combustion chamber with a minimum intensity in the range of $1-100$ GW/cm^2 for a reliable plasma formation [5,7,8]. In order to ensure the ignitability of very lean mixtures, the laser energy had to be increased well above the breakdown threshold since lean mixtures demand higher ignition energies. The breakdown threshold itself is a function of the gas pressure and temperature, the laser wavelength, the temporal pulse shape, and the spatial beam profile. The optimum spark position depends on the local flow conditions and should be distant from quenching metal walls (i.e., in the center of the combustion chamber).

Table 1 recalls the characteristics of conventional spark plugs and laser-induced ignition, with the latter being a well-known concept but at the same time a rather "young technology" in respect to technical applicability.

2 Experimental Setup

This study was conducted to evaluate the potential of laser ignition in aeroengines. The experiments were performed with a generic combustion test rig at atmospheric pressure, allowing a high operational flexibility for parametric screening. Optical access into the combustion chamber was possible for the ignition laser and for the purpose of high-speed video recording. The laser system and the spark plug were selected as being state of the art representatives in their specific technology. Hence, the pulse energy of the laser (100 mJ, 1064 nm) was significantly lower than the spark energy of the spark plug (1800 mJ). Furthermore, a homemade electrode igniter (EI) (5–140 mJ) was used as a complementary substitute for the laser ignition system. This electrode igniter could generate an ignition kernel under almost identical conditions as the laser (e.g., in the center of the combustion chamber) by utilizing the spark plug's "electrode concept." Thus an enhancement of the parametric study was possible.

3 Definition of the Measurement Campaign

Ignitability can generally be measured by the minimum ignition energy, ignition/flammability limits, and the ignition delay time. The ignition performance depends on the ignition system (including pulse/spark energy and duration, and the location of the ignition kernel), the flow conditions (including air pressure, temperature, and velocity), and the fuel preparation and mixing (including fuel type and temperature, equivalence ratio, and the spray characteristics in the case of liquid fuels). The following list summarizes the inputs for the selected parameter space of this study:

- ambient temperature: $T_{amb}=290$ K
- ambient pressure: $p_{amb}=0.098-0.103$ MPa
- combustor preheating temperature: $T_{pre}=300-473$ K
- combustor pressure: constant at ambient conditions, open exhaust (0.1 MPa $\pm$ 10%)
- air mass flow (AMF): $\dot{m}_{air}=8.2-37.7$ g/s
- swirl number: $S=0.55, 0.63$
- flame type: spray flame (Jet-A1), technically premixed (natural gas (NG))
- droplet spectrum: $D_{32}=28.3-77.9$ μm
- spray angle (axial hollow cone nozzle): 60 deg, 80 deg
- equivalence ratio: $\Phi=0.23-1.08$
- thermal power: $\dot{Q}_{th}=9.5-70.0$ kW
- ignition technology: spark plug (Champion), laser (LOT-Oriel), and electrode igniter (homemade)
- nominal ignition energy per pulse: 1800 mJ (spark plug), 15–100 mJ (laser and electrode igniter)
- ignition location: vertical midplane, radial offset locations

In order to evaluate the technical applicability of laser ignition in aeroengines, three major results are presented within this paper:

- ignitability limits at different flow/mixture conditions
- combustor ignition maps (mainly for spray combustion)
- similarities and differences between technically premixed and spray combustion

4 Test Facility

Figure 1 gives an overview of the test rig and the laboratory's infrastructure. The main functional units are the air and fuel supplies, the air preheater, the plenum, the swirler (including a hollow lance for the kerosene supply), and the combustion chamber, which will be discussed in more details in the following subsection (Fig. 2). The in-house air supply (1.2 MPa) provides the mass flow for the combustion air as well as for the convective cooling of the combustion chamber. The electrical air preheater has a power of 32 kW and provides a maximum air temperature of

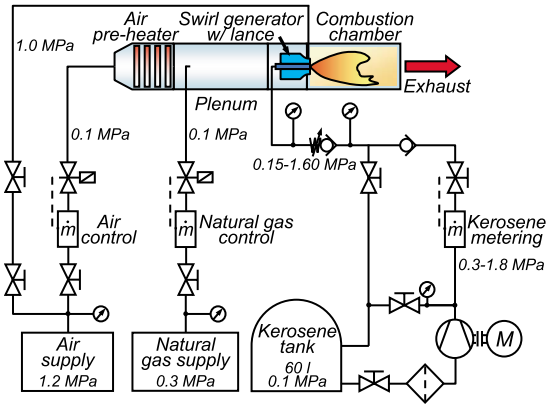


Fig. 1 Test rig for the measurement campaign

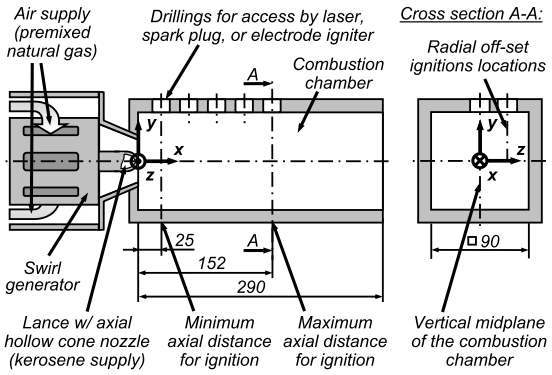


Fig. 2 Sectional view of the combustion chamber

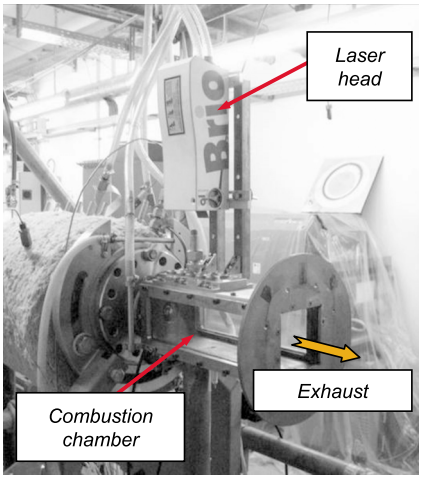


Fig. 3 Installation of the laser head above the combustion chamber

$T_{\text{pre}|\max}=800$ K. As shown in Fig. 1, the fuel inlet locations vary for natural gas and kerosene. Natural gas is injected farther upstream from the combustion chamber to ensure fuel premixing, whereas kerosene is injected directly from the swirler lance into the combustion chamber (Figs. 1 and 2). The natural gas (98% methane, 0.3 MPa) is supplied by the laboratory infrastructure. Jet-A1 is used for the spray combustion experiments. It is stored in a 60 l tank and a fuel pump is used to raise its pressure up to 1.8 MPa. In order to atomize the kerosene, an axial hollow cone nozzle is mounted in the swirler lance. Four different nozzle types were utilized during the experiments to study the influences of the spray's Sauter mean diameter (SMD) and the fuel mass flow (FMF). Ignition was carried out by a compact laser, a homemade electrode igniter, or a gas turbine spark plug unit. The main advantage of the employed laser was the laser head's compact construction so that a straightforward installation on top of the combustion chamber was possible.

The combustion chamber has a square cross section and measures $90\times90\times290$ mm^3 (Fig. 2), which is comparable to the dimensions of a single sector in an annular aeroengine combustor [1]. The lateral panes of the combustion chamber are equipped with exchangeable silica glass for optical access (e.g., for high-speed video recording). In order to allow access for the different ignition systems, the combustion chamber's top plate has a grid of 2×5 drillings. However, only one drilling is used at a time; the spare ones are sealed by blanking plugs.

5 Laser and Optics

The laser ignition experiments were performed with a commercial laser. Figure 3 shows the head of the Q-switched Nd:YAG laser mounted above the combustion chamber. It is used with a graded reflectivity mirror at 1064 nm wavelength, resulting in a laser beam energy of $E_{\text{laser}|\max}=100$ mJ [9].

A set of six planoconvex lenses (focal length $f=25.0-80.0$ mm) was available to focus the laser beam inside of the combustion chamber and to induce gas breakdown [10]. Every lens was assembled in a separate optics holder, which could be mounted in the combustion chamber's top plate. In order to vary the focal point location, two approaches were pursued: the variation of the focal length and the axial adjustment of the optics/lens holder. As the latter option allows a focal point variation of 13 mm in total, an almost complete spatial coverage of the combustor volume could be achieved by combining both approaches.

An electrode igniter was developed to supplement the laser ignition tests (Fig. 4). It is capable of reaching any spatial location within the upper half of the combustion chamber and its power spectrum is adaptable to the utilized ignition laser, as well as to the currently available miniaturized laser systems. Its electrodes are made of tungsten and provide a separation distance of 2 mm in which the ignition spark is generated during the system's discharge process. Furthermore, the igniter design aims for the least possible disturbance of the combustor's flow field. The ignition energy can be adjusted in the range of $E_{\text{EI}}=15-140$ mJ by variably charging an ignition coil.

6 Comparison of Laser and Spark Plug

Contrary to conventional spark ignition, a laser releases its energy within a few nanoseconds, leading to a shockwave generation and the characteristic ellipsoidal shape of the flame kernel. The laser-induced ignition generates a local ignition kernel of very small dimensions. On the other hand, the ignition kernel dimensions of the gas turbine spark plug are significantly bigger (Fig. 5). It is capable of yielding an energy of 1800 mJ, which was obtained by measurements on the secondary electric circuit. Evaluating the performance of both ignition systems, the individual energy transfer efficiencies have to be considered. According to Maly and Borgnakke [11] and Rager [2], the glow discharge of a spark plug ignition is the dominant phase in regard to the total energy balance because it is in the time range of a few micro- to milliseconds. Breakdown and arc discharge only last from a few nanoseconds up to one microsecond. During the glow discharge an efficiency of 30–34% can be assumed [2,4,11]. Thus, the effective pulse energy of the spark plug is around $E_{\text{SP}|\text{eff}}=600$ mJ. For laser ignition, energy losses of approximately 8% are expected due to the utilization of one focal lens, which results in $E_{\text{laser}|\text{eff}}=92$ mJ. Further energy losses are evident but not considered hereinafter. Even though the ignition energies of the two

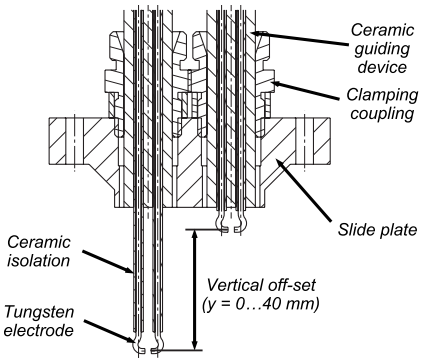


Fig. 4 Sectional view of the electrode igniter. The drawing displays the vertical displacement capability of the system by showing two pairs of electrodes although only one pair was used for the ignition tests at a time.

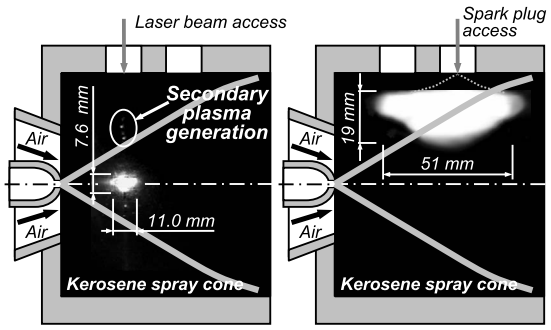


Fig. 5 Ignition kernel of the laser (left plot) and the spark plug ignition (right plot)

systems differ by one magnitude ($E_{\text{laser}} \ll E_{\text{SP}}$), both of them represent the state of the art. This was considered as a main aspect for a reasonable evaluation of the two ignition systems. Despite its lack in absolute power, the laser offers one additional degree of freedom: the free spatial allocation of the ignition kernel inside of the combustor. With respect to the ignition kernels' shape, size, and appearance, both ignition systems showed the typical characteristics reported before [5,8,12–14].

7 Ignition Tests

7.1 Ignition Tests With Premixed Natural Gas. The initial ignition tests for this study were performed with premixed natural gas. The collected results were used as a reference for the succeeding ignition tests with atomized kerosene. During these initial experiments the fuel mass flow was increased up to $\dot{m}_{\text{NG|max}} = 1.4$ g/s, which corresponds to $\dot{Q}_{\text{th}} = 70$ kW. The main objective was to determine the combustor's lean ignition limit for natural gas at ambient as well as elevated temperatures ($T_1 = 300$ K, $T_2 = 473$ K). The air mass flow was kept constant (three levels: $\dot{m}_{\text{air}} = 8$ g/s, 15 g/s, and 25 g/s) and the fuel mass flow was increased in small increments, until a successful ignition could be achieved. Figure 6 illustrates representative results for the ambient temperature condition (T_1) and the air mass flow of $\dot{m}_{\text{air}} = 25$ g/s=const.

The ignition tests at the lean ignition limit unveiled a better performance for the conventional aeroengine spark plug ($E_{\text{SP}} = 1800$ mJ) than the laser-induced ignition ($E_{\text{laser}} = 100$ mJ) and the electrode igniter ($E_{\text{EI}} = 100$ mJ). Due to the difference in pulse energy in one order of magnitude, the expected performance shift between the two concepts could be confirmed by the experiments. Nevertheless, the lean ignition limits show a similar trend for all three tested systems (Fig. 6). For a successful ignition, the heat

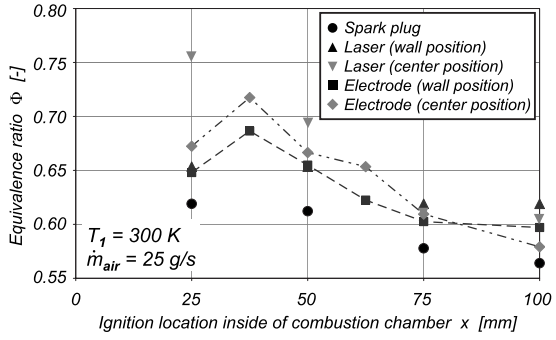


Fig. 6 Lean ignition limits of the different ignition systems for technically premixed natural gas. The plot shows ignition locations close to the upper wall as well as in the center of the combustion chamber.

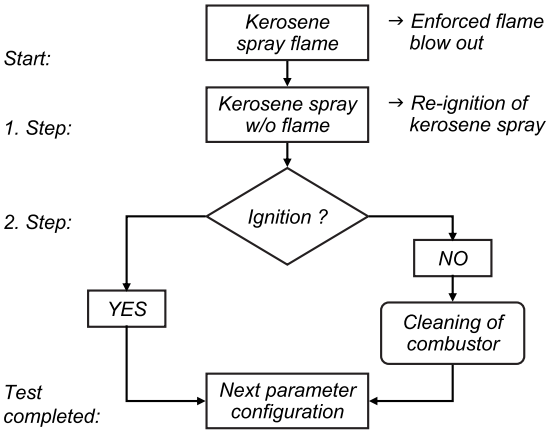


Fig. 7 Procedure for the kerosene spray ignition experiments

release rate of the initial chemical reactions must approximately balance the rate of heat loss [15]. High flow velocities increase the thermal convection, and thus increase the rate of the heat loss. Since the spark plug is characterized by a high spark energy and exhibits a prominent ignition kernel, it outperforms the laser and the electrode igniter at swirl-stabilized flames with technical/ideal premixing. It is able to provide more energy to the combustor so that the chemical start reactions can ensure a sufficiently high temperature increase and establish a stable flame. Ignition locations close to the combustor walls offer an improved flow field condition due to boundary layer effects [1,16]. The resulting ignition characteristics are reflected in low equivalence ratios without excessive spatial scatter. Farther downstream in the combustion chamber ($x \geq 75$ mm), the overall flow field is generally characterized by a relatively uniform turbulence level, implying an improved ignitability of the premixed natural gas [17].

Regarding laser ignition at focal point locations near the combustor walls, AMF variations have no major impact on the lean ignition limit. However, a strong influence is evident in the center of the combustion chamber. Additionally, this impact is discernable only in the near field of the combustion chamber inlet (up to $x < 75$ mm). Downstream of $x = 75$ mm the lean ignition limit is at the lowest level for all three AMF levels ($\dot{m}_{\text{air}}$=const.). Moreover, comparing the two vertical ignition locations ("wall" and "center") of Fig. 6 for the laser and the electrode igniter, a reversed trend can be identified for ignition positions downstream of $x = 75–80$ mm. This is due to a stronger decrease in the lean ignition limit for the center than for the wall locations. In order to better understand the ignition characteristics within the combustion chamber, it will be helpful to refer to Fig. 8 later, which shows the axial velocity of the isothermal flow field.

7.2 Ignition Tests With Atomized Kerosene. The ignition tests for kerosene were performed according to the sequence of Fig. 7. The starting point is a technically premixed natural gas flame. Kerosene is added to this flame at a progressive rate with a simultaneous cutback of the natural gas flow until a mere kerosene spray flame prevails. In this way, a kerosene flame can be generated without soot deposition onto the nozzle and without excessive kerosene contamination of the combustion chamber. After the enforced flame blowout (by a supplementary air jet) and a few seconds of waiting time, the reignition of the pure kerosene spray is initiated. If the ignition is successful with the actual parameter set in this "first step" (Fig. 7), the experiments can be continued right away with the next parameter configuration. If the reignition fails, the combustion chamber must be cleaned from the kerosene residues before continuing on with the next test run. In this case, a premixed natural gas flame is used for cleaning purposes.

Axial hollow cone nozzles for spray angles of 60 deg and 80 deg [18] were chosen for changing the spray cone's shear layer

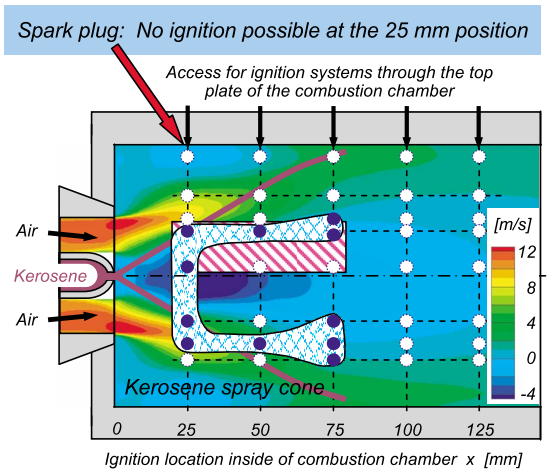


Fig. 8 Example of an ignition map for kerosene spray combustion ($\dot{m}_{air}$=14.0 g/s, Φ_{global}=0.40, D_{32}=43.8 μm, S=0.55, and T_{pre}=300 K). The tested ignition locations are marked by circles. Filled (blue) circles indicate successful ignitions by laser, resulting in the cross-hatched area. The (pink) shaded area displays successful ignitions by using the electrode igniter as a laser substitute. The background shows the numerical result for the axial velocity component of the combustor flow field.

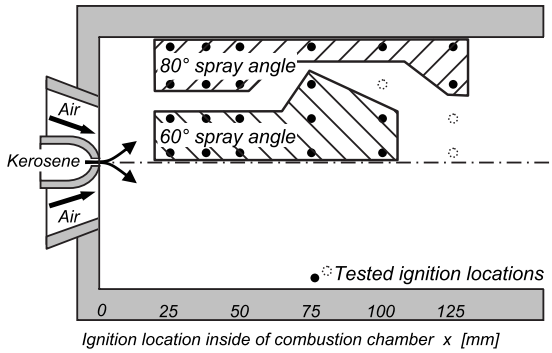


Fig. 9 Variation of the kerosene spray angle ($\dot{m}_{air}$=14.0 g/s, Φ_{global}=0.47, and T_{pre}=300 K). These results were obtained by using the electrode igniter (E_{EI}=100 mJ).

and the SMD of the spray. The provided SMD depends mainly on the pressure drop across the lance/nozzle setup (Figs. 1 and 8). Due to the different hole diameters, nozzle variations result in different volumetric flow rates at constant pressure. An increasing injection pressure increases the kerosene flow rate and generates an improved atomization by simultaneously lowering the SMD. The droplet spectra were analyzed by a particle sizer from Malvern directly at the nozzle orifice. The SMD was adjusted within the range of D_{32}=28.3–77.9 μm, depending on the specific experiment conditions.

The experiments on kerosene ignition were also carried out with all three ignition systems: the gas turbine spark plug from Champion, the Nd:YAG laser from LOT-Oriel, and the homemade electrode igniter. Depending on the ignition system, a pulse/spark energy between 15 mJ and 1800 mJ was provided for the combustion process. The majority of the tests were conducted at ambient conditions in order to determine the ignitability of the spray without the positive influence of an elevated temperature level. In comparison to typical engine startup conditions, very low global equivalence ratios ($\Phi_{global|min}$=0.23) were investigated during the test runs as the soot generation was high for moderate equivalence ratios in the investigated combustor setup. Therefore, the lens-cleaning characteristics of the laser played an important role to counteract soot deposition or lens wetting. Both typically reduced the ignition system's efficiency.

The test results are illustrated in the form of ignition maps as shown in Fig. 8. The axial flow direction is from left to right. The blank circles represent failed ignitions. All successful ignitions are indicated by filled (blue) circles. Hence, the cross-hatched area corresponds to the area where successful ignition can be expected. In order to further discuss the influence of the burner's flow field, the numerical result for the axial velocity at isothermal flow conditions is plotted as a background. Considering the inner and outer recirculation zones, the hollow spray cone overlaps with the highest flow velocities. This results in an improved mixing within the shear layers of the spray. Since laser ignition is generally possible in all areas with favorable equivalence ratios and good mixing, focusing right onto these optimum spark locations yields excellent results. The ability of the free selection of the ignition kernel's location is a major advantage in this regard. Consequently, successful ignitions could be performed down to global equivalence ratios of $\Phi_{global|min}$=0.23 and from installation positions, where no ignition was possible by the spark plug (i.e., x=25 mm).

Recalling Fig. 8, a successful laser ignition was achievable up to 75 mm downstream of the combustor. The ignition map shows a characteristic U-shape (Fig. 8, cross-hatched area) due to the hollow spray cone. An ignition in the combustor center is not possible farther downstream than x=50 mm, which also arises from the nozzle type and an insufficient recirculation of the kerosene/air mixture. The cross-hatched ignition area is slightly asymmetric and eccentric. This is an optical effect caused by the utilization of lenses of different focal lengths f.

Due to the symmetric flow field and the quasisymmetric test results (Fig. 8), the electrode igniter was designed to allow spark generation at every location within the upper half of the combustion chamber. Furthermore, the system's power spectrum was adjusted to the laser performance. Comparing the contours of both ignition maps (laser: cross-hatched and blue; electrode igniter: shaded and pink), the outer edges expose an exact concurrence. However, the inner edges show a deviation between the two ignition approaches; and, an improved ignitability can be stated for the electrode igniter along the combustor's center axis. This is due to a wetting of the electrodes and fuel enrichment around their spikes, implying an improved ignitability.

Since the electrode ignition resembles the nonresonant breakdown ignition by laser in its essential physical processes, and the conformance of the experimental results from laser and electrode igniter is high, the following parametric study is based on the test results of these two ignition systems.

8 Parametric Study on Kerosene Spray Flames

A variation of the kerosene spray angle results in a spatial shift of the ignitable area, as illustrated in Fig. 9. By employing a 60 deg hollow cone nozzle, ignition is successful in the vertical range of y=0–27 mm. On the other hand, a 80 deg nozzle facilitates ignition close to the combustor walls due to an increased expansion of the spray cone. The aforementioned correlation between the spray angle, the flow field, and the resulting ignitable area is evident here. The global equivalence ratio is kept constant at the value of Φ_{global}=0.47, which is close to the lean ignition limit for premixed combustion (Fig. 6).

Changing the air mass flow $\dot{m}_{air}$ has a considerable effect on the flow field of the combustor. Figure 10 displays the results for four AMF values ($\dot{m}_{air}$=8.2 g/s, 14.0 g/s, 27.6 g/s, and 37.7 g/s). Since the global equivalence ratio Φ_{global} and the Sauter mean diameter D_{32} were kept constant for this test, it represents a wide AMF variation. Two 60 deg and two 80 deg nozzles had to be utilized for the technical realization. This has to be considered

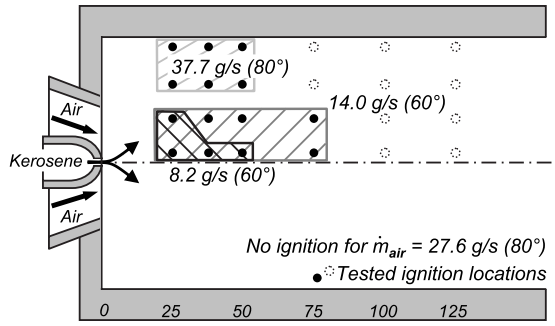


Fig. 10 Variation of the air mass flow $\dot{m}_{air}$ (Φ_{global}=0.40, D_{32} =43.8 μm, and T_{pre}=300 K). These results were obtained by using the electrode igniter (E_{EI}=100 mJ).

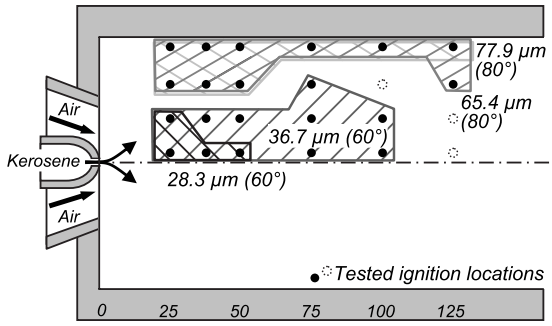


Fig. 12 Variation of the Sauter mean diameter D_{32} ($\dot{m}_{air}$ =14.0 g/s, Φ_{global}=0.47, and T_{pre}=300 K). These results were obtained by using the electrode igniter (E_{EI}=100 mJ).

regarding the spatial allocation of the ignitable areas. Once more, the electrode igniter was used to obtain the results. The nominal spark energy was E_{EI}=100 mJ.

Figure 11 presents the results of a variation of the equivalence ratio Φ. It emphasizes the expected effect that an increase in the equivalence ratio yields an increase in the ignitable area. For the minimum equivalence ratio of $\Phi_{global|min}$=0.23 an ignition is only possible in the near field of the nozzle. Referring to lean mixtures in general, the fuel/air mixing and thus the ignition kernel's axial location are of major importance for the performance of the ignition system. The initially fuel-rich areas close to the nozzle are subject to an intense mixing and prevaporization so that local peaks of the equivalence ratio Φ are leveled toward a spatially uniform Φ. If this, in turn, is below the lean ignition limit, no ignition will be observed farther downstream in the combustion chamber.

The influence of the Sauter mean diameter D_{32} on swirl-stabilized kerosene spray flames is summarized in Fig. 12. Small droplets with an SMD of D_{32}=28.3 μm can only be ignited in the near field of the nozzle ($x \leq 50$ mm). However, increasing droplet diameters imply an increase in the ignitable area, as illustrated by the corresponding ignition maps. Since the vaporization time increases with the square of the initial droplet diameters [16], larger droplets promote ignition locations in the rear of the combustion chamber. Even though the initial flame kernel is located farther downstream, the succeeding flame front propagates upstream until it reaches the stagnation point of the internal recirculation zone. Müsing et al. [19] showed that the size of the plasma kernel during the gas breakdown is affected by the size of nearby droplets. Larger droplets increase the plasma kernel, and thus facilitate the ignition process.

The ignition energy was varied in order to comply with the performance of the currently available miniaturized ignition lasers and to show the potential of future laser developments [9]. Figure 13 reflects the results for ignition energies of E_{EI}=15 mJ and 100 mJ. The ignition map indicates a successful ignition process even at E_{EI}=15 mJ and despite the lean fuel conditions of Φ_{global} =0.40.

Figure 14 shows the effect of combustion air preheating on laser ignition. The laser energy was set at E_{laser}=60 mJ for these tests. A temperature increase resulted in an increased ignitability. Indeed, an initiation of successful ignitions could be observed at elevated temperatures (T_2=473 K) at locations where no ignition was possible without preheating. The ignitability and flammability limits are generally related to the heat losses from the combustion

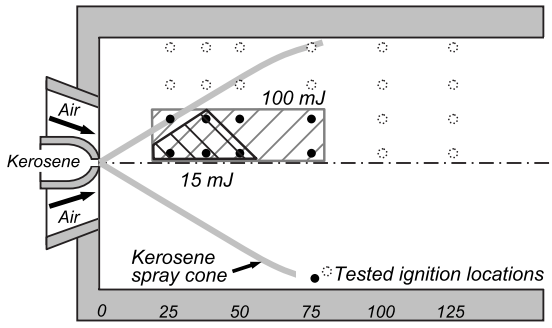


Fig. 13 Variation of the ignition energy ($\dot{m}_{air}$=14.0 g/s, Φ_{global}=0.40, D_{32}=43.8 μm, and T_{pre}=300 K). These results were obtained by using the electrode igniter.

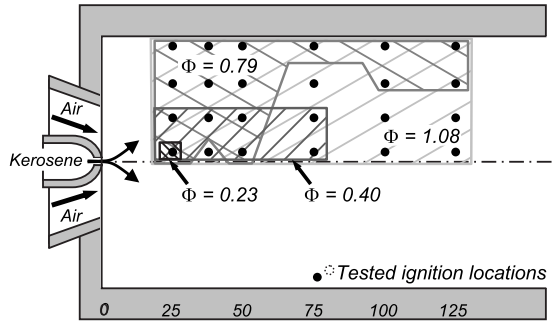


Fig. 11 Variation of the global equivalence ratio Φ ($\dot{m}_{air}$ =14.0 g/s, D_{32}=43.8 μm, and T_{pre}=300 K). These results were obtained by using the electrode igniter (E_{EI}=100 mJ).

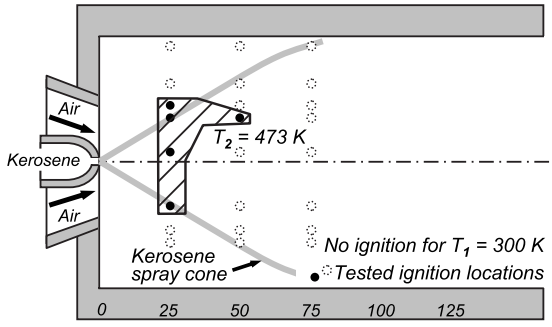


Fig. 14 Variation of the combustion air temperature T_{pre} ($\dot{m}_{air}$ =14.0 g/s, Φ_{global}=0.40, and D_{32}=43.8 μm). These results were obtained by using laser-induced ignition (E_{laser}=60 mJ).

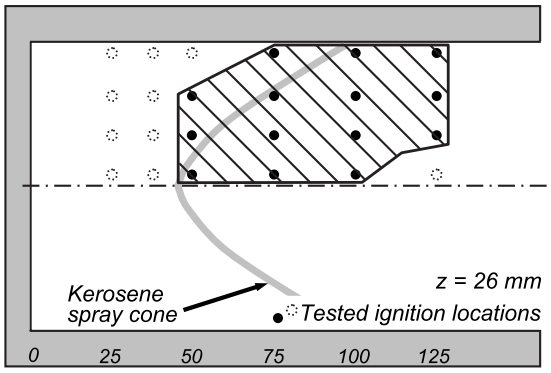


Fig. 15 Variation of the vertical ignition plane (i.e., *z*-axis offset), ($\dot{m}_{air}$=14.0 g/s, Φ_{global}=0.50, D_{32}=34.5 μm, and T_{pre}=300 K). These results were obtained by using the electrode igniter.

system. Thus, the negative temperature gradient toward the combustor liner has a major influence on the ignitability. According to Williams' criteria, the ignition source needs to supply sufficient energy in order to heat the critical gas volume for the combustion from its initial state to flame temperatures [1,15]. Ergo, the ignitable area can be expanded when the preheating temperatures are increased. Farther downstream of the combustion chamber (x >50 mm) the ignitability deteriorates as the vaporized kerosene mixes well with the swirled preheated air and the local equivalence ratios converge with Φ_{global}, which is below the lean ignition limit.

Recalling the design of the combustor setup, the flow field can be considered to be axis-symmetric (Fig. 8). As presented earlier, the hollow cone nozzle is installed on the axis. A parallel shift of the ignition plane in direction of the z-axis confirms the ignition characteristics: Fig. 15 displays the ignition map for an offset of z=26 mm (Fig. 2). The results are in excellent agreement with the results for the vertical midplane of the chamber. Ignition is facilitated in the shear layers of the kerosene spray cone due to locally rich areas and an intensified mixing. This mechanism is predominant and supports ignition, even at global equivalence ratios below the lean ignition limit.

9 Conclusions

In order to establish as a new technology on the market for aeroengine combustion, laser-induced ignition has to compete with the well-developed spark plug ignition. This study was conducted for a better understanding of the performance and the potentials of both ignition systems. Ignition tests of premixed natural gas and kerosene sprays were carried out with a conventional gas turbine spark plug, a compact Q-switched Nd:YAG laser, and a homemade electrode igniter. Spark plug and laser were selected as being state of the art representatives in their specific technology. The laser-induced spark was obtained by focusing the laser beam down to several micrometers up to several hundreds of micrometers, thereby exceeding the specific breakdown threshold intensity of the gaseous medium within the focal region. The electrode igniter was used to complement the measurement campaign by uniting the advantages of spark plug and laser. The major difference between those two ignition concepts is the pulse energy with significantly lower values for the laser (E_{laser}=100 mJ) in comparison to the spark plug (E_{SP}=1800 mJ).

The location of the ignition kernel in the combustor is one of the most relevant parameters for the overall performance of an ignition system. This was confirmed by the ignition tests with premixed natural gas. Spatial differences in the determined lean ignition limit were mainly due to the combustor aerodynamics. The lowest equivalence ratios were identified for ignition locations in the downstream center of the combustion chamber for the laser as well as the electrode igniter, (T_1=300 K: Φ=0.57; T_2 =473 K: Φ=0.52). Higher equivalence ratios were required close to the wall of the combustion chamber and at areas with an increased turbulence level. The lean ignition limits show similar characteristics for laser and spark plug ignitions, but slightly lower absolute numbers for the spark plug. This is due to its remarkably higher ignition energy ($E_{SP|eff} \gg E_{laser|eff}$).

The ignition of atomized kerosene can be performed successfully in the near field of the spray cone's shear layer if flammable equivalence ratios prevail at this location. Consequently, the main advantage of laser ignition is the possibility of a spatially free positioning of the focal point and thus of the ignition kernel. For that reason, locations could be unveiled where the laser outperformed the spark plug ($x<50$ mm). Moreover, successful laser ignition was observed down to a global equivalence ratio of Φ_{global}=0.23. This aspect is significant for aeroengines during startup because of the high potential for emission reduction and during reignition in high altitudes. Considering all restricting effects and the severe experiment conditions, the tests demonstrated the high potential of laser-induced ignition. Characterized by a remarkably lower output energy, laser ignition is already competitive with conventional ignition systems, even though it is still under development.

Acknowledgment

This research was performed in the framework of a study on laser ignition initiated by General Electric. The authors would like to thank General Electric for its technical and financial support and the permission to publish the scientific results.

Nomenclature

D_{32} = Sauter mean diameter, μm
E = energy, mJ
f = focal length, mm
$\dot{m}$ = mass flow, g/s
p = pressure, MPa
P = power, W
$\dot{Q}$ = heat flow, W
S = swirl number
T = temperature, K
Φ = equivalence ratio

Subscripts

amb = ambient
max = maximum
min = minimum
pre = preheating
th = thermal

References

[1] Lefebvre, A. H., 1983, *Gas Turbine Combustion*, Hemisphere, Washington, DC.
[2] Rager, J., 2006, "Funkenerosion an Zündkerzenelektroden," Naturwissenschaftlich-Technische Fakultät III, Universität des Saarlandes.
[3] Robert Bosch GmbH, 2002, Zündung im Ottomotor.
[4] Pischinger, F., 2001, "DFG Sonderforschungsbereich 224 'Motorische Verbrennung': Chapter 3.2—Zündung," Lehrstuhl für Verbrennungskraftmaschinen, RWTH Aachen.
[5] Kopecek, H., Charareh, S., Lackner, M., Forsich, C., Winter, F., Klausner, J., Herdin, G., Weinrotter, M., and Wintner, E., 2005, "Laser Ignition of Methane-Air Mixtures at High Pressures and Diagnostics," ASME J. Eng. Gas Turbines Power, **127**(1), pp. 213–219.
[6] Ronney, P. D., 1994, "Laser Versus Conventional Ignition of Flames," Opt. Eng., **33**(2), pp. 510–521.
[7] Kröner, M., 2003, "Einfluss lokaler Löschvorgänge auf den Flammenrückschlag durch verbrennungsinduziertes Wirbelaufplatzen," Lehrstuhl für Thermodynamik, Technische Universität München.
[8] Lackner, M., Winter, F., Graf, J., Geringer, B., Weinrotter, M., Kopecek, H., Wintner, E., Klausner, J., and Herdin, G., 2004, "Laser Ignition in Internal Combustion Engines: A Contribution to a Sustainable Environment," *Proceed-*

ings of the 14th IFRF Members Conference, Noordwijkerhout, The Netherlands, May 11–14.
[9] LOT-Oriel, 2006, Kompakter Nd:YAG-Laser Brio.
[10] Linos, 2006, Plano-Concave Lenses.
[11] Maly, R., and Borgnakke, C., 1984, *Fuel Economy in Road Vehicles Powered by Spark Ignition Engines*, Plenum, New York.
[12] Carleton, F. B., Klein, N., Krallis, K., and Weinberg, F. J., 1990, "Laser Ignition of Liquid Propellants," *Proceedings of the 23rd Symposium International on Combustion*, Orleans, France, July 22–27, pp. 1323–1329.
[13] Chen, Y.-L., and Lewis, J. W. L., 2001, "Visualization of Laser-Induced Breakdown and Ignition," Opt. Express, **9**(7), pp. 360–371.
[14] Lackner, M., Charareh, S., Winter, F., Iskra, K. F., Ruedisser, D., Neger, T., Kopecek, H., and Wintner, E., 2004, "Investigation of the Early Stages in Laser-Induced Ignition by Schlieren Photography and Laser-Induced Fluorescence Spectroscopy," Opt. Express, **12**(19), pp. 4546–4557.
[15] Turns, S. R., 2000, *An Introduction to Combustion: Concepts and Applications*, McGraw-Hill, New York.
[16] Lefebvre, A. H., 1989, *Atomization and Sprays*, Hemisphere, New York.
[17] Wäsle, J., Winkler, A., and Sattelmayer, T., 2005, "TD1: Experimental Investigation of the TD1 Swirl Burner," Lehrstuhl für Thermodynamik, Technische Universität München.
[18] Lechler, 2006, Axial Hollow Cone Nozzles.
[19] Müsing, A., Riedel, U., Warnatz, J., Herden, W., and Ridderbusch, H., 2007, "Laser-Induced Breakdown in Air and Behind Droplets: A Detailed Monte-Carlo Simulation," Proc. Combust. Inst., **31**, pp. 3007–3014.

Assessment of Different Actuator Concepts for Acoustic Boundary Control of a Premixed Combustor

Mirko R. Bothien
Jonas P. Moeck
Christian Oliver Paschereit

Institut für Strömungsmechanik
und Technische Akustik,
Technische Universität Berlin,
10623 Berlin, Germany
e-mail: mirko.bothien@tu-berlin.de

In the design process, new burners are generally tested in combustion test rigs. With these experiments, computational fluid dynamics, and finite element calculations, the burners' performance in the full-scale engine is sought to be predicted. Especially, information about the thermoacoustic behavior and the emissions is very important. As the thermoacoustics strongly depend on the acoustic boundary conditions of the system, it is obvious that test rig conditions should match or be close to those of the full-scale engine. This is, however, generally not the case. Hence, if the combustion process in the test rig is stable at certain operating conditions, it may show unfavorable dynamics at the same conditions in the engine. In previous works, the authors introduced an active control scheme, which is able to mimic almost arbitrary acoustic boundary conditions. Thus, the test rig properties can be tuned to correspond to those of the full-scale engine. The acoustic boundary conditions were manipulated using woofers. In the present study, proportional valves are investigated regarding their capabilities of being used in the control scheme. It is found that the test rig impedance can be tuned equally well. In contrast to the woofers, however, the valves could be used in industrial applications, as they are more robust and exhibit more control authority. Additionally, the control scheme is further developed and used to tune the test rig at discrete frequencies. This exhibits certain advantages compared with the case of control over a broad frequency band.
[DOI: 10.1115/1.2969088]

Keywords: burner development, impedance tuning, acoustic boundary condition, thermoacoustic

1 Introduction

One of the main issues in gas turbine development is the stability of the combustion process. In order to follow stringent NO_x emission restrictions, lean-premixed combustion was introduced by the gas turbine industry. Combustion systems operating in this mode are, however, susceptible to self-excited oscillations arising due to the interaction of the unsteady heat release and the acoustic field in the combustion chamber. If the two mechanisms constructively interfere, high amplitude pressure and heat release fluctuations occur, which have a detrimental effect on the combustion process. These so-called thermoacoustic instabilities cause structural wear, increase noise and pollutant emissions, and can even lead to engine failure [1–3].

The most important component influencing the tendency to thermoacoustic instabilities is the burner, making its development a crucial task. However, not in each step of the early design phase the quality of new burner generations can be assessed in the full-scale engine. This is usually accomplished by means of extensive experimental investigations in combustion test rigs [4,5]. These experiments in conjunction with computational fluid dynamics and finite element calculations are then used to predict the burner's performance in the full-scale engine. The quality of this prediction, however, can only be verified at the very end of the design process after the burner is tested in the full-scale engine. At this point, changes in component design are very expensive and will cause severe delays in the schedule [4]. This underlines the need for a procedure allowing a reliable prediction of burner behavior in the engine. Alternatives to component design modifications then are only active and passive control strategies. Numerous research groups have been working on these topics. However, only a few publications exist reporting the implementation of these concepts in the real engine. Engine applications of active control strategies were shown by Seume et al. [6] and Richards et al. [7]. Passive means are more common and were, for example, presented by Bellucci et al. [8] and Berenbrink and Hoffmann [9]. More information and references for engine applications were summarized in a comprehensive review on passive control strategies in gas turbines conducted by Richards et al. [10]. Different industrial approaches can also be found in Ref. [1].

Generally, the acoustic characteristics of the test rig are very different from those of the full-scale engine, thus complicating the prediction. This is due to higher static pressures and temperatures and, in particular, due to a different geometry. Thermoacoustic instabilities usually appear close to the acoustic resonance frequencies of the combustion chamber, whenever the flame response at these frequencies is sufficiently large and the phase relationships involved promote constructive interference [1,11]. For this reason, a burner that shows a desirable behavior with respect to pressure pulsations and emissions in the test rig might exhibit essentially different characteristics in the engine. This is due to the fact that the thermoacoustics of the engine are properly represented by the test rig only if both the dynamic flame response (significantly influenced by the burner) and the acoustic boundary conditions match.

Accounting for the different acoustic boundary conditions by means of design modifications of the test rig is expensive and time consuming. Therefore, Mongia et al. [4] presented a laboratory scale combustor, whose resonance frequencies could conveniently be changed. The test rig basically consisted of an upstream tube, the burner, and a downstream tube. The tubes' lengths and thus the test rig's resonances could be continuously varied using perforated

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 31, 2008; final manuscript received April 21, 2008; published online December 24, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

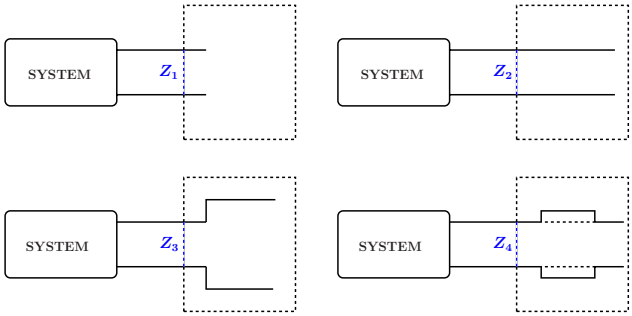


Fig. 1 System with different acoustic boundary conditions induced by a change in geometry or by implementation of a liner

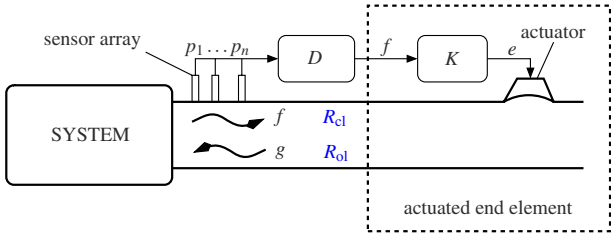


Fig. 2 Schematical setup of the control concept for manipulation of the acoustic boundary condition of a duct

pistons. The combustion group at Georgia Tech uses a similar setup to tune the upstream end of their combustion test facility [12]. Although in both cases a change of test rig resonances is quite straightforward, the degree of reflectivity remained constant and problems are to be expected when applying the system to industrial test rigs with higher power, i.e., higher mass flows and elevated pressure.

Another way to adjust the boundary conditions to those of the engine is followed here by actively tuning, i.e., artificially changing them. This can be done by using an active control scheme, which detects the acoustic field at the combustion chamber outlet and feeds back a suitable signal to drive an acoustic actuator. The actuator modifies the acoustic field in the system to match the desired one. For example, the acoustic boundary condition can be changed to that of a choked flow without actually needing the flow to be choked. An additional advantage of this method is that the same test rig can be used to simulate the acoustic conditions of different full-scale engines. Also, by incorporating the real boundary conditions to the test rig, it is possible to have a realistic estimate about the whole combustion process subject to the engine acoustics, e.g., NO_x emissions. Bothien et al. [13] introduced and applied this active control scheme to study the thermoacoustic behavior of a premixed combustor, exposed to different acoustic boundary conditions.

The underlying idea is that the impact of a geometry change on the plane wave acoustic field *physically* can be represented by a change of the impedance at a certain position. In the following, the setup depicted in Fig. 1 (top left) is considered. The system is connected to a duct, in which only plane waves propagate, terminated with an acoustic boundary condition Z_1. Here, Z denotes the acoustic impedance, a complex valued function of frequency. It is defined as

$$Z = \frac{\hat{p}}{\hat{v}} \tag{1}$$

where $\hat{p}$ and $\hat{v}$ denote Fourier transforms $\hat{(\cdot)}$ of acoustic pressure and particle velocity normal to the boundary, respectively. Here, as in the following, the acoustic pressure has been scaled by the characteristic impedance ρc, ρ being the fluid density and c the speed of sound.

The system considered in Fig. 1 has certain characteristics depending on the acoustic field associated with Z_1. For other acoustic boundary conditions, e.g., induced by a change in length (top right), an area change (bottom left), or by some damping device mounted in the duct (bottom right), the system may exhibit essentially different characteristics. This can be due to higher or lower resonance frequencies or increased damping, for example. The aim of the control scheme is to mimic certain prescribed impedances. In this way, the impact of different acoustic boundary conditions on the system characteristics can be studied, without actually implementing the devices that alter the boundary impedance.

In 1984, Guicking and Karcher [14] introduced the general concept of changing the impedance through active control of the acoustic boundary condition. However, they only considered fixed frequency harmonic signals and tuned the control parameters empirically. Furthermore, they only considered a system without mean flow and combustion. Application of this concept was also presented in Ref. [15], where a similar method was used to reduce the resonator length of a thermoacoustic cooler.

Bothien et al. [16] used the concept of active boundary control to validate linear stability analysis with experimental data. By tuning the downstream reflectivity of a combustion test rig, the predictive capabilities of a network model regarding frequency of instability, linear growth rate, and transition from stability to instability could be assessed. Furthermore, it was shown that impedance tuning can be used to suppress thermoacoustic instabilities. By decreasing the downstream reflection coefficient or by adequately changing its phase the thermoacoustic feedback cycle can be interrupted.

2 Impedance Tuning Concept

In this section, the basic features of the impedance tuning concept will be described. A more detailed derivation of the underlying control concept can be found in Refs. [13,17].

2.1 General Control Concept. To develop the control scheme, the setup in Fig. 2 is considered. The duct, whose end-impedance is to be manipulated, is equipped with a sensor array, consisting of microphones at different axial positions and an actuator. To achieve the control objective, the sensor signals are fed to a control scheme, which generates the command signal and acts on the acoustic field via the actuator. Instead of the acoustic pressures, it is more convenient to consider the downstream propagating wave f as the actual control input. For this reason, the control scheme depicted in Fig. 2 has been split into two parts: (i) a wave decomposer D that extracts the downstream propagating wave from the pressure signals p_i; and (ii) the actual control law K. Accordingly, finding a wave decomposition scheme D that works accurately over a range of frequencies is necessary.

The problem now is to find a suitable K that, given the downstream propagating wave f, is able to drive the actuator in such a way that the closed-loop reflection coefficient R_{cl} is close to the desired one. The control law K can be built if the uncontrolled (or open-loop) reflection coefficient R_{ol} and the actuator transfer function G are known, as will be shown below. In contrast to Guicking and Karcher's approach [14], there is no need for manual parameter tuning and the controller acts over a range of frequencies. Also, the scheme presented here will not only work for time-harmonic but also for transient signals. This is particularly important when considering combustion noise, where no distinct frequencies prevail.

2.2 Controller Identification. In this study, only plane wave acoustic fields are considered, where the acoustic pressure $\hat{p}$ and the particle velocity $\hat{v}$ only depend on the axial coordinate. Practically, this means that only frequencies below the cut-on frequency for the first no-planar mode can be accounted for. For the

control approach considered here, a more convenient notation makes use of the Riemann invariants $\hat{f}$ and $\hat{g}$, which can be thought of as the down- and upstream traveling waves. They are related to the primitive acoustic variables by

$$\hat{p} = \hat{f} + \hat{g} \tag{2a}$$

$$\hat{v} = \hat{f} - \hat{g} \tag{2b}$$

Due to linearity, Eqs. (2*a*) and (2*b*) hold in frequency as well as in the time domain.

In Eq. (1), the acoustic boundary condition was expressed as the impedance Z, the ratio of acoustic pressure and particle velocity. Equivalently, the reflection coefficient R is defined as the ratio of the reflected to the incident wave, $R=\hat{g}/\hat{f}$. Using Eqs. (2*a*) and (2*b*), it is found that the reflection coefficient is related to the impedance by

$$R = \frac{Z-1}{Z+1} \tag{3}$$

To derive a suitable control law, a model for the actuated end element (Fig. 2) is written as

$$\hat{g} = R_{\mathrm{ol}}\hat{f} + G\hat{e} \tag{4}$$

where $\hat{e}$ is the control command. Here, it is assumed that the reflected $\hat{g}$-wave can be obtained as a linear superposition of the geometrical reflection of the incident wave $\hat{f}$ and the wave generation due to the actuator. The validity of this assumption was previously shown in Ref. [18].

Given that $\hat{f}$ can be extracted from the microphone signals and is fed to the controller ($\hat{e}=K\hat{f}$), the closed-loop reflection coefficient is given by

$$R_{\mathrm{cl}} = \frac{\hat{g}}{\hat{f}} = R_{\mathrm{ol}} + GK \tag{5}$$

If the uncontrolled reflection coefficient R_{ol} and the actuator transfer function G are known, then, given a desired reflection coefficient R_{cl}, the control law K can be calculated from

$$K = \frac{R_{\mathrm{cl}} - R_{\mathrm{ol}}}{G} \tag{6}$$

The uncontrolled reflection coefficient and the actuator transfer function can be identified using the multimicrophone method (MMM) [19,20] and two independent excitation states. Excitation with the upstream woofer and application of the MMM yield the uncontrolled reflection coefficient R_{ol} as $\hat{e}=0$ in Eq. (4). Subsequent excitation with the downstream speaker and further application of the MMM allow computation of the speaker transfer function G from Eq. (4) as follows:

$$G = \frac{\hat{g} - R_{\mathrm{ol}}\hat{f}}{\hat{e}} \tag{7}$$

taking into account the reflection coefficient R_{ol} calculated in the previous step.

The procedure explained above provides the controller transfer function K as discrete frequency response data. However, to apply it in a control scheme, a model accurately capturing the response is necessary. In this study, the models were obtained by using frequency domain system identification algorithms [21]. When doing so, it has to be made sure that the resulting controller model is stable and proper.

2.3 Controller for Discrete Frequency Impedance Tuning. Identifying a stable controller transfer function K for a broad frequency range can sometimes be cumbersome if not impossible. This is due to the reason that the expression for K might be too complex or even unphysical. Consider, for example, an actuator transfer function G with a time-delay larger than that of the numerator in Eq. (6). This would result in a phase increase of K, which physically would correspond to the fact that the controller would have to generate a signal prior to getting its input. Another problem is that K is determined from measured frequency response data and therefore is only prescribed in a certain frequency range. A model, however, is valid for all frequencies between zero and infinity. In addition to identifying an accurate model, it has to be made sure that the K-model does not overshoot at frequencies approaching zero or infinity. Otherwise, the control output would have to be saturated and control is not possible.

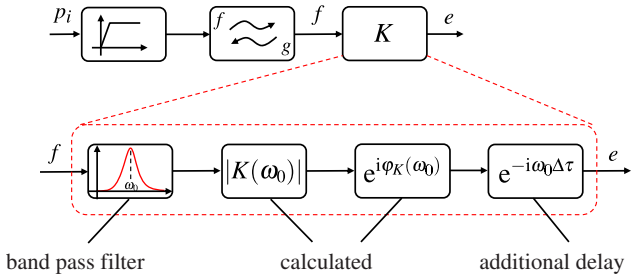


Fig. 3 Control schematic for impedance tuning at discrete frequencies

These problems can be avoided if the impedance is only tuned at discrete frequencies. The idea is to compose the controller of the transfer function K, calculated at a single frequency ω_0, say, from Eq. (6) and a bandpass filter centered at ω_0. This modified controller setup is shown in Fig. 3.

2.4 Method to Generate an Instability at ω_0. In the following, it is outlined how the concept can be used to generate an instability at a discrete frequency ω_0. It is common practice to represent the linear behavior of thermoacoustic systems by so-called network models [18,22–24]. The most simple form of such a network model is obtained if the system properties are described by up- and downstream reflection coefficients, which are then connected to each other. This is shown in Fig. 4, where the upstream reflection coefficient comprises the flame response and the downstream reflection coefficient is the one to be tuned.

This system may get unstable if g_{in} and g_{out} are in phase and the reflection coefficients' magnitudes are high enough, i.e.,

$$\varphi_{\mathrm{cl}}(\omega_0) + \varphi_{\mathrm{us}}(\omega_0) = 2n\pi, \quad n = 0,1,2,\ldots \tag{8a}$$

$$|R_{\mathrm{cl}}(\omega_0)||R_{\mathrm{us}}(\omega_0)| \geq 1 \tag{8b}$$

This procedure is not totally correct, as the frequency of a growing oscillation is complex, $\omega=\omega_{\mathrm{osc}}+i\omega_{\mathrm{GR}}$, ω_{osc} being the frequency of oscillation and ω_{GR} being the growth rate. In contrast to this, K is calculated for ω_0, a real frequency. The higher the growth rate, i.e., the larger the negative imaginary part, the more the experimentally observed frequency will differ from ω_0. Moreover, if the system gets into a limit-cycle, either nonlinearities or the impact of high amplitude oscillations on the operating parameters may slightly change the frequency of instability as was observed in Ref. [25]. In most cases, however, this difference was

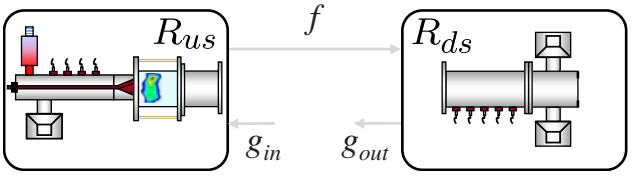


Fig. 4 Network representation of the test rig

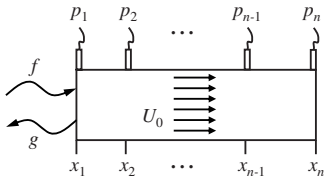


Fig. 5 Setup for online wave identification with multiple microphones

negligible. If this would not be the case, deviations might be adjusted by an additional delay $\exp(-i\omega_0\Delta\tau)$ implemented in the controller (Fig. 3). Using this delay, the phase of K and therefore the phase of R_{cl} might be changed, causing the system to get unstable at the desired frequency ω_0.

If the system is tuned at discrete frequencies, no controller identification is necessary, as the controller only consists of a bandpass filter, a gain, and a delay. Moreover, tuning the system to higher resonance frequencies with the conventional approach is limited, as this physically corresponds to a reduction of test rig length. Theoretically, the shortest length realizable is the one at which the upstream traveling wave generated by the actuator and the downstream traveling f-wave, serving as the control input, meet. Tuning at discrete frequencies avoids this problem, as Eq. (8*a*) can be solved for all frequencies. Another important advantage is that the delay of the actuator response to its input signal is not important, as the phase has only to be shifted by a constant value. Regarding a controller for a frequency band, this is not the case, as its phase normally cannot be described by a single time-delay.

Besides these important advantages, one has to account for some disadvantages, too. As the test rig has to be tuned for each frequency of interest, measurement campaigns will be extended. Furthermore, tuning R_{cl} to fully reflecting at discrete frequencies might lead to limit-cycle oscillations at two different modes, say, ω_1 and ω_2. If R_{cl} is adjusted to fully reflecting in a frequency band covering ω_1 and ω_2, the mode with the larger growth rate (ω_1, say) might completely dominate the limit-cycle due to nonlinear effects [26]. Thus, at ω_2, no limit-cycle oscillations will be observed. The burner's behavior to this respect therefore cannot be observed when tuning at discrete frequencies.

2.5 Scheme for Online Wave Decomposition. As the f-wave is the control input, a suitable algorithm is required that separates the up- and downstream traveling waves from the pressure sensor signals. In the frequency domain, this can be accomplished by using the MMM. However, in contrast to the conventional procedure, here, the wave decomposition cannot be performed as a postprocessing step. The f-wave has to be identified *online* and its calculation from the sound pressure sensors has to be accomplished in the time domain. Therefore, a scheme analogous to the MMM in the frequency domain was developed [17], allowing to map multiple pressure measurements to the Riemann invariants. This can be done by first noting that the Riemann invariants in a uniform duct are only functions of retarded arguments,

$$f = f(t - x/(c + U_0)), \quad g = g(t + x/(c - U_0)) \tag{9}$$

With respect to the setup shown in Fig. 5, n equations for the pressure signals are obtained as

$$p_1(t) = f(t) + g(t) \tag{10a}$$

$$p_2(t) = f(t - \tau^+_{1,2}) + g(t + \tau^-_{1,2}) \tag{10b}$$

$$\vdots$$

$$p_n(t) = f(t - \tau^+_{1,n}) + g(t + \tau^-_{1,n}) \tag{10c}$$

where $\tau^{\pm}_{n,m} = (x_m - x_n)/(c \pm U_0)$.

The aim is now to find a causal relation for $f(t)$ that accounts for all n pressure recordings. Therefore, the pressure at location x_k is written at time $t - \tau^-_{1,k}$ as follows:

$$p_1(t) = f(t) + g(t) \tag{11a}$$

$$p_2(t - \tau^-_{1,2}) = f(t - \tau^+_{1,2} - \tau^-_{1,2}) + g(t) \tag{11b}$$

$$\vdots$$

$$p_n(t - \tau^-_{1,n}) = f(t - \tau^+_{1,n} - \tau^-_{1,n}) + g(t) \tag{11c}$$

Equations (11*a*)–(11*c*) are now considered as an overdetermined linear system for $f(t)$ and $g(t)$, whose least-squares solution takes the form

$$\begin{bmatrix} f(t) \\ g(t) \end{bmatrix} = \begin{bmatrix} 1 & -1/(n-1) & \dots & -1/(n-1) \\ 0 & -1/(n-1) & \dots & -1/(n-1) \end{bmatrix} \cdots \cdot \begin{bmatrix} p_1(t) \\ p_2(t - \tau^-_{1,2}) - f(t - \tau^+_{1,2} - \tau^-_{1,2}) \\ \vdots \\ p_n(t - \tau^-_{1,n}) - f(t - \tau^+_{1,n} - \tau^-_{1,n}) \end{bmatrix} \tag{12}$$

The feedback scheme, mapping the n pressures online to the downstream traveling wave, is shown in Fig. 6.

3 Experimental Setup

3.1 Combustion Test Rig. Experiments were conducted in an atmospheric combustion test rig, which is schematically shown in Fig. 7. The combustor features a swirl-stabilized burner and is

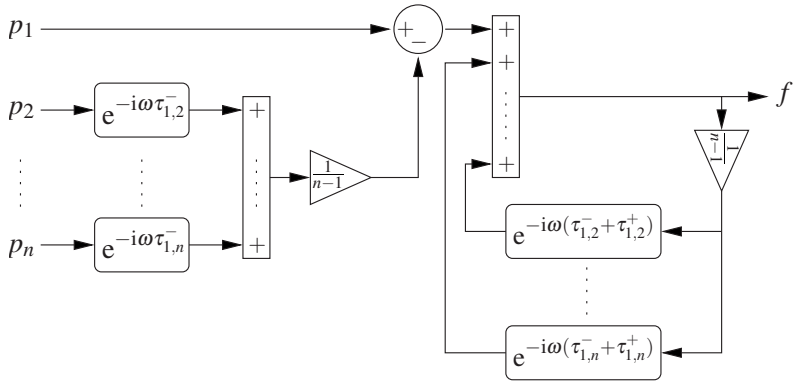


Fig. 6 Block diagram mapping *n* measured pressures to the downstream traveling wave *f*

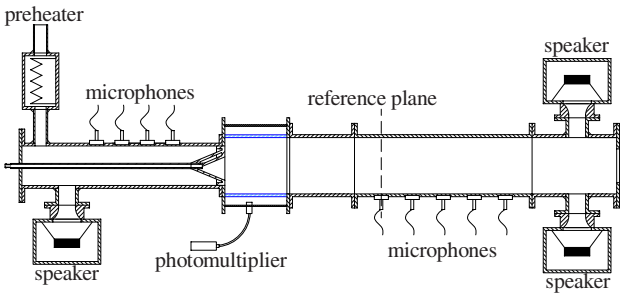


Fig. 7 Schematic setup of the atmospheric test rig

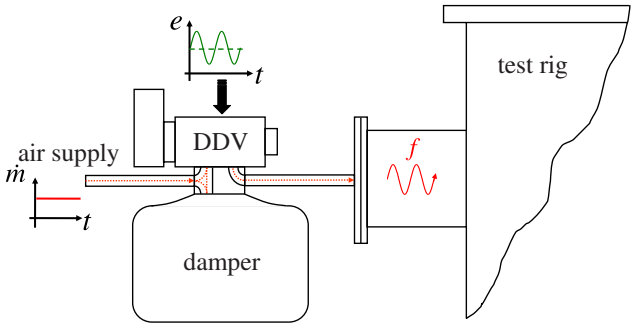


Fig. 8 Setup of the DDV mounted at the downstream end of the test rig

operated in lean-premixed mode. The diameter of the exhaust duct is 0.2 m and the burner has an area expansion ratio of approximately 4. To simulate conditions that are closer to those in full-scale engines an electrical preheater can be used to provide higher air temperatures at the burner inlet. Several water-cooled condenser microphones up- and downstream of the flame allow for full identification of the plane wave pressure field. The OH-chemiluminescence in the flame is monitored with a photomultiplier equipped with an optical bandpass filter that is connected to the combustion chamber via a fiber optic cable. In the zone where the flame stabilizes, a quartz glass tube allows for full optical access. In the setup shown, woofers are mounted at both ends of the rig providing for acoustic excitation. These woofers can be replaced with other actuators. The measurement tube, where the downstream microphones are mounted, and the section with the woofers have water-cooled walls. Additionally, the speaker casings are purged with air to prevent hot gases from entering. The combustor outlet can be equipped with different terminations, such as, e.g., an orifice to reduce the reflection of low-frequency acoustic waves [27,28].

The reference plane drawn in Fig. 7 marks the location where the downstream traveling wave, serving as the control input, was identified. The up- and downstream reflection coefficients shown in Sec. 4 were also determined at this position. Note that this plane is located upstream of the actuator position. Therefore, the influence of the actuator setup will be included in all reflection coefficients of the downstream end shown.

The controller was implemented in Simulink and ran on a DS1103 PPC control board (dSPACE), which generated the command signal for the actuators at a sampling frequency of 2^{13} Hz.

3.2 Actuators. In this work, results obtained with two different types of actuators will be presented. When the concept was first introduced [13], woofers mounted at the downstream end (see Fig. 7) were used to tune the test rig. It was shown that they were capable of modifying the acoustic boundary condition in a broad frequency range. However, woofers suffer from certain shortcomings if to be applied to industrial high power test rigs. Besides lacking sufficient power and thus control authority, they will not withstand the harsh environmental conditions, in particular, at elevated pressures.

To avoid these shortcomings, the woofers were replaced with proportional valves. These so-called direct-drive-valves (DDVs) (manufactured by Moog Inc., Germany) were used to modulate an air mass flow to excite the test rig. Seume et al. [6] used similar valves for active instability control in heavy duty gas turbines by modulating the pilot fuel. The valves were mounted at the downstream end of the test rig at the same location where the woofers were installed. A schematic of the setup is shown in Fig. 8. The DDV is connected to the pressurized air supply. A damper, i.e., a volume, is mounted to prevent the imposed modulation from traveling upstream. The valves presented here are a special version of the ones used in Ref. [6], manufactured to modulate higher mass flows. With this special design, it is possible to modulate 0.25 kg/s of air up to 300 Hz. Moreover, they fulfill another major actuator requirement, the linearity between actuator input and generated perturbation. This linearity implies both the increase of acoustic excitation with increasing control command and increasing air mass flow.

4 Experimental Results

4.1 Woofers. The test rig was operated in lean-premixed mode burning natural gas with an equivalence ratio of 0.65 at a thermal power of 110 kW. Two different configurations of the downstream termination were used. One was an open end, thus, generating a pressure node at the downstream exit. In the other case, the test rig was equipped with an orifice. This configuration produced a relatively small magnitude of the reflection coefficient as acoustic energy was dissipated into vortices, shedding from the orifice's edge [27,28]. Figure 9 depicts the reflection coefficients for these two configurations. For the open end (◇) the reflection coefficient is similar to the one described by the formula of Levine and Schwinger [29] for long wavelengths. With the orifice mounted (□), R was decreased to values around 0.3. The controller was applied to the configuration with orifice and its transfer function (Eq. (6)) was identified for frequencies between 50 Hz and 300 Hz.

In the top frame of Fig. 10, the acoustic pressure spectra for the two configurations without control are shown. For the open end (◇), distinct peaks at 82 Hz and its multiples can be observed, representing the $\lambda/4$-mode of the rig. The effect of the orifice can clearly be seen (□). By reducing the downstream reflection coefficient it stabilized the combustion process. Mounting the orifice

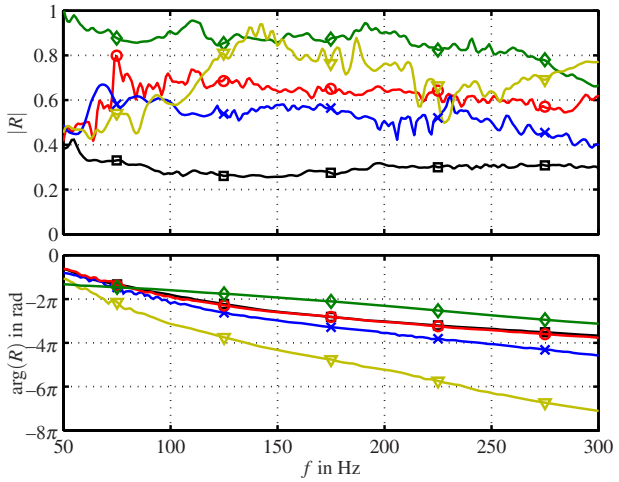


Fig. 9 Downstream reflection coefficient. Without control (□—with orifice and ◇—without orifice); $|R_{cl}|=1$—with different additional lengths ($\Delta l=0$ m, ○; $\Delta l=0.5$ m, ×; and $\Delta l=2$ m, ▽).

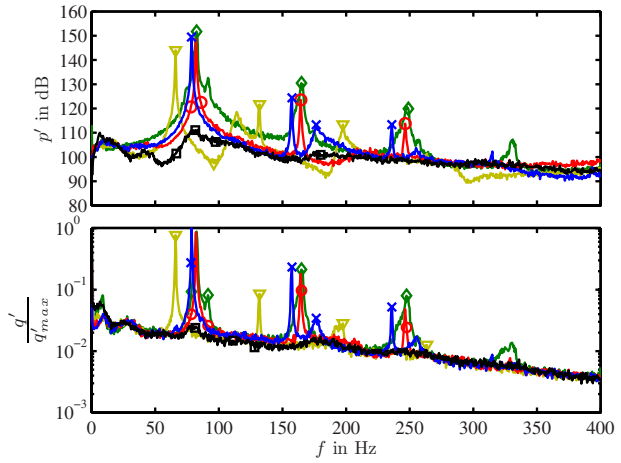


Fig. 10 Top: spectra of acoustic pressure; bottom: spectra of OH-chemiluminescence. Without control (□—with orifice and ◇—without orifice); $|R_{cl}|=1$—with different additional lengths ($\Delta l=0$ m, ○; $\Delta l=0.5$ m, ×; and $\Delta l=2$ m ▽). Downstream reflection coefficient is tuned for frequency range.

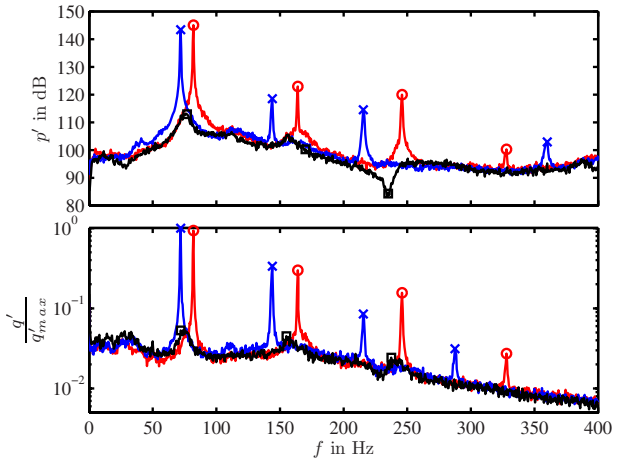


Fig. 11 Top: spectra of acoustic pressure; bottom: spectra of OH-chemiluminescence. Without control (□); $|R_{cl}|=1$—with different additional lengths (○ and ×). Downstream reflection coefficient is tuned for frequency range.

resulted in a peak amplitude reduction from 153 dB to 111 dB, i.e., a reduction to less than 1% of the pressure amplitude for the open end.

Using the impedance tuning concept, the reflection coefficient was changed from highly absorbing to fully reflecting as follows:

$$R_{cl} = 1 \cdot e^{i(\varphi_{ol} - \omega 2\Delta l/c)} \tag{13}$$

Different virtual lengths could be simulated by adding an additional length Δl.

At first, the controller was used to generate an instability at 82 Hz corresponding to the frequency of the $\lambda/4$-mode of the open end. As can be seen in the phase plots of Fig. 9 (bottom), the reflection coefficients' phases of both uncontrolled cases were nearly the same at the instability frequency. Therefore, the controller transfer function was calculated to produce $|R_{cl}|=1$ without changing the phase of the uncontrolled R_{ol} ($\Delta l=0$ m). The ○ curve shows the result for this case. In Fig. 10, distinct peaks at 82 Hz and its multiples can be observed, proving that the anechoic end was virtually tuned to one with a higher reflection coefficient. Note that the test rig was not additionally excited by the upstream loudspeaker for all spectra shown. A peak amplitude of 150 dB was generated, 3 dB less than in the uncontrolled case without orifice.

The bottom plot of Fig. 10 shows the impact of the generated acoustic field on the heat release fluctuations. Here, the scaled OH-chemiluminescence signals are shown. Large heat release fluctuations occurred at the instability frequency of 82 Hz (open end, ◇), whereas no peaks are visible for the anechoic termination without control (□). In case of control (○), the effect of the generated acoustic field on the heat release was nearly the same as for the open end without control.

The phases of the tuned reflection coefficients (Fig. 9, bottom) were the same for the case of control and the uncontrolled case, as prescribed. The magnitude of R_{cl}, however, lay between 0.6 and 0.7, thus, differing from the desired value of $|R_{cl}|=1$. This was due to the fact that the output calculated by the controller had to be restricted to 70% of the actual value, as otherwise the woofers' amplifiers went into saturation. Hence, the saturation was not system-inherent compared with the uncontrolled case without orifice, in which the nonlinear saturation could be attributed to the flame response. Here, the actuator power was the limiting factor.

In a next step, the test rig length was virtually extended by manipulation of the reflection coefficient's phase. Consequently, the instability peak was shifted to lower frequencies. The × and ▽ curves in Fig. 10 show the spectra for additional lengths of $\Delta l=0.5$ m and $\Delta l=2$ m, shifting the pressure peaks to 78 Hz and 66 Hz, respectively. The same holds for the heat release fluctuations. Virtually extending the test rig length resulted in a larger slope of the reflection coefficient's phase as the propagation time of, and thus the time-delay between, the incident and the reflected wave became larger (Fig. 9, bottom). Additional tests with the woofers are documented in Refs. [13,16], where also anechoic ends were simulated and the concept was used to suppress thermoacoustic instabilities.

4.2 Proportional Valve. The combustion test rig was operated at an equivalence ratio of 0.75 and a thermal power of 100 kW. Without control, no pressure oscillations were present as can be seen in Fig. 11 (□).

The reflection coefficient was first tuned in a frequency range. According to the results shown previously, the magnitude of R_{cl} was again set to unity. The additional length (see Eq. (13)) was adjusted to produce an instability at the same frequencies as it was the case for the woofers. The pressure and OH-chemiluminescence spectra for the controlled cases, depicted in Fig. 11 (○ and ×), show peaks at 82 Hz and 72 Hz. The imposed pressure amplitudes were approximately of the same magnitude as it was observed in Fig. 10. Figure 12 shows the corresponding

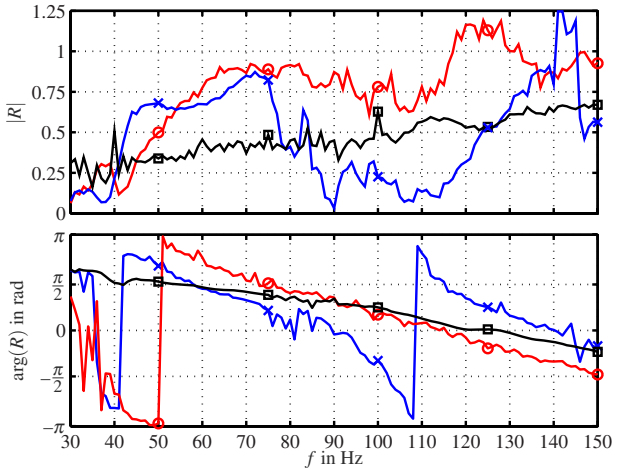


Fig. 12 Downstream reflection coefficient. Without control (□); $|R_{cl}|=1$—with different additional lengths (○ and ×).

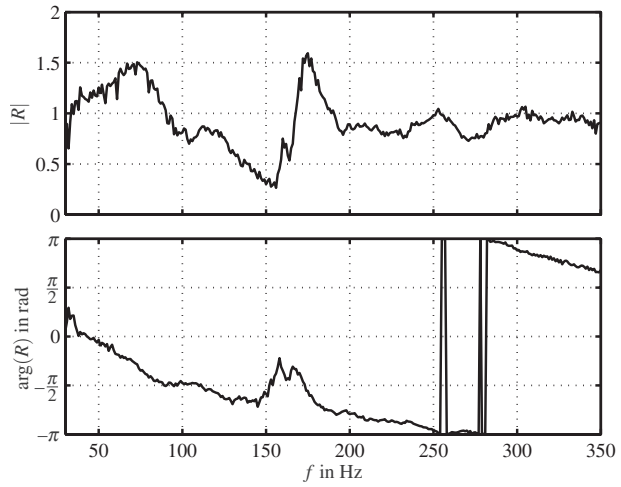


Fig. 13 Upstream reflection coefficient R_{us} comprising the flame response

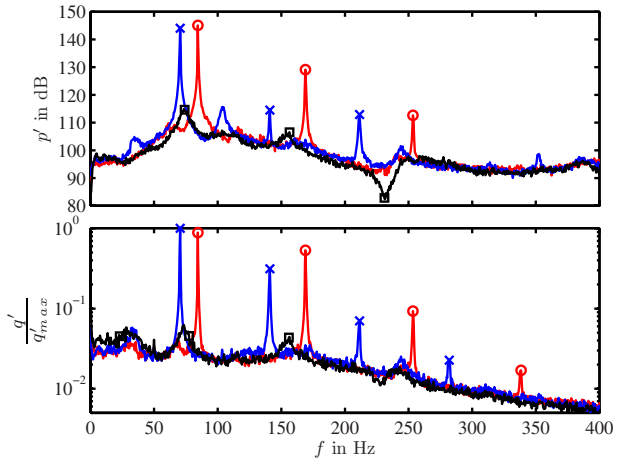


Fig. 14 Top: spectra of acoustic pressure; bottom: spectra of OH-chemiluminescence. Without control (□); $|R_{cl}|=1-85$ Hz (○) and 70 Hz (×). Downstream reflection coefficient is tuned at discrete frequencies.

downstream reflection coefficients R_{cl}.

Comparison of R_{ol} of the base line cases for both actuator setups reveals some deviations. With the valves mounted, the reflection coefficient was higher. These differences were to be expected and can be attributed to the different actuator setups and the slight differences in operating parameters. The mean air mass flow through the valves was set to 60 g/s, equal to the main combustion air. In contrast to that, the speakers were only purged with 10 g/s. This difference of cold air insertion influences the downstream reflectivity as the characteristic impedance of the medium, ρc, is changed abruptly. Even if the air insertion was the same, the different actuator setups would cause different reflection coefficients.

Regarding the controlled downstream reflection coefficients, a major difference between the actuator concepts becomes obvious. Using the valves, it was very cumbersome to identify a controller for a broad frequency range. The frequency band was limited to 35–75 Hz (×) and 60–150 Hz (○), respectively. With the woofers, it was possible to adjust a controller bandwidth of 250 Hz (50–300 Hz). As described earlier, this is due to the actuator transfer function G of the valves. Hot-wire measurements conducted at the valve exit showed that the modulation amplitude was maximal at 290 Hz continuously decreasing with decreasing frequency. To improve the response at low frequencies, a tube was mounted between the valve exit and the test rig. This increased the propagation time of the generated acoustic wave traveling from the valve exit to the reference plane and had a significant effect, as the speed of sound was much lower in the side branch than in the main duct. The low-pass characteristic of the valve caused an additional time-delay. However, especially for low frequencies, this influence was much smaller than that of the increased duct length. Due to the different setup, the phase decrease in the valve transfer function was higher than that of the woofers. As stated previously, the larger the time-delay of G, the harder is the identification of the controller transfer function, calculated from Eq. (6). Still, if the magnitudes of R_{cl} at the frequencies of oscillation are considered, better results were obtained with the valves. In both cases, the magnitudes lay above 0.9.

The controller identification can be avoided if the system is tuned at discrete frequencies. Besides this, another major advantage lies in tuning at high frequencies. Generally, thermoacoustic instabilities might arise if the phase relation between heat release and pressure fluctuations provides constructive interference and the flame response is sufficiently high. The flame response can be thought of as a contribution to the upstream reflection coefficient R_{us} if it is determined at the reference plane in Fig. 7. In Fig. 13, R_{us} is shown, measured for the uncontrolled stable base line case by sweep excitation with the valves. The influence of the flame as an acoustically active element caused R_{us} to be larger than unity in two frequency bands (30–90 Hz and 170–190 Hz). These were the frequencies for which the system could get unstable if the downstream reflection coefficient was adequately manipulated.

The downstream reflection coefficient was first tuned at 70 Hz and 85 Hz to

$$R_{cl}(\omega_0) = 1 \cdot e^{-i\varphi_{us}(\omega_0)} \tag{14}$$

With this, $K(\omega_0)$ can be calculated from Eq. (6) to set up the band-pass controller. At both frequencies, R_{us} was larger than unity and with the phase of R_{cl} being set to $-\varphi_{us}$, the phase relationship for constructive interference between heat release and pressure fluctuation was given (see Eq. (8)).

The spectra shown in Fig. 14 prove that for the low-frequency region similar results were obtained compared with multifrequency tuning. The resulting peak amplitudes were found to be equally high (compared with Fig. 11). The magnitudes of the tuned reflection coefficient lay again around 0.9 and the phases were correctly adjusted to $-\varphi_{us}$ (not shown), proven by the fact that the system oscillated at the correct frequencies. In the case of 70 Hz, the additional delay (see Fig. 3) was set to 0.24 ms, as otherwise the instability frequency was 71 Hz.

The underlying idea of tuning the downstream reflection coefficient in a frequency band is to virtually change the length of the test rig (see Eq. (13)). As was stated before, this implies the limitation of a reduction of the test rig length. This, therefore, limits the ability to examine the system behavior at higher oscillation frequencies. If the test rig is tuned at a discrete frequency, these restrictions are overcome. The results for tuning in the second frequency band (170–190 Hz), in which R_{us} is also larger than unity, are depicted in Fig. 15.

The combustion system was forced to oscillate at 175 Hz and 185 Hz, respectively. Neither were the actuators in saturation, nor was it necessary to introduce the additional delay. This was due to the fact that the valves have a higher response at these frequencies. Additionally, it is supposed that the growth rates and therefore the imaginary parts of the instability frequencies were smaller than at frequencies around 70 Hz.

The valves were also used to tune the downstream reflection coefficient to 0 (not shown). Particularly at frequencies between 150 Hz and 200 Hz it was possible to decrease R_{ds} to values below 0.05.

4.3 Saturation of the Actuator. If the actuator power is not high enough to meet the control objective, the calculated control

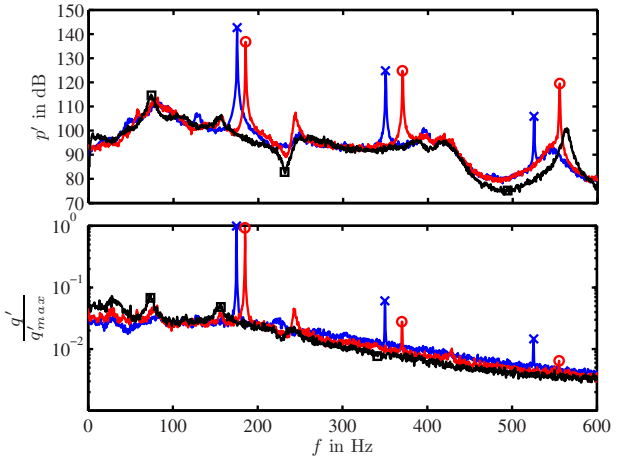


Fig. 15 Top: spectra of acoustic pressure; bottom: spectra of OH-chemiluminescence. Without control (□); $|R_{cl}|=1-185$ Hz (○) and 175 Hz (×). Downstream reflection coefficient is tuned at discrete frequencies.

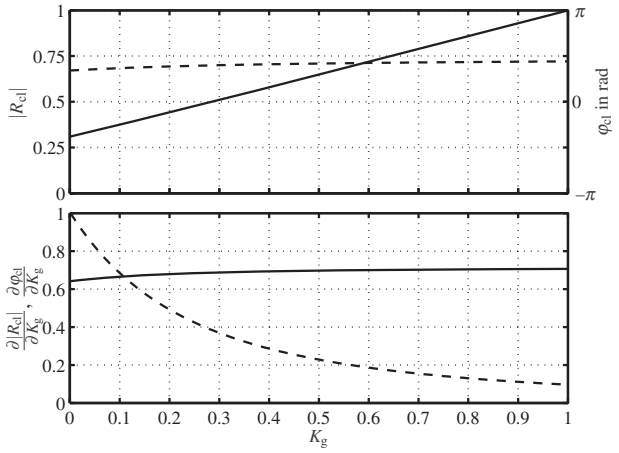


Fig. 16 Influence of limitation of control signal on closed-loop reflection coefficient ($|R_{cl}|$ solid; φ_{cl} dashed). The system is tuned to generate an instability at 85 Hz.

output has to be decreased. Otherwise, the actuator could be damaged or the signal is saturated and therefore not processed correctly. An additional gain K_g was inserted into the controller to avoid this. In the results shown previously, this gain limited the control signal to ~70–90% of its actual value. Nonetheless, the system was forced to oscillate at the prescribed frequency (although the amplitude of oscillation would be higher in case of $K_g=1$). This raises the question, why the adjustment of the phase of the downstream reflection coefficient, which determines the oscillation frequency, is less sensitive to K_g than it is the case for the magnitude. Note that the accuracy of the phase adjustment also has a governing influence on the amplitude of oscillation as for the thermoacoustic feedback cycle the phase relationship between R_{us} and R_{cl} is crucial.

If K_g is introduced to the controller, Eqs. (5) and (6) can be rearranged to yield

$$R_{cl} = R_{ol}(1-K_g) + R_{set}K_g \tag{15}$$

where R_{set} is the prescribed reflection coefficient to be adjusted and R_{cl} is the one that is actually adjusted. R_{set} is used to calculate K in Eq. (6). If $K_g=1$, i.e., if the control signal would not have to be limited, they are identical. However, if the actuator does not exhibit sufficient control authority, the actual reflection coefficient will differ from the prescribed one. The influence of K_g on $|R_{cl}|$ and φ_{cl} can be seen, if Eq. (15) is split into

$$|R_{cl}| = [(\Re(R_{ol})(1-K_g) + \Re(R_{set})K_g)^2 \ldots + (\Im(R_{ol})(1-K_g) + \Im(R_{set})K_g)^2]^{1/2} \tag{16a}$$

$$\varphi_{cl} = \arctan\frac{\Im(R_{ol})(1-K_g) + \Im(R_{set})K_g}{\Re(R_{ol})(1-K_g) + \Re(R_{set})K_g} \tag{16b}$$

Calculating the derivatives of Eqs. (16*a*) and (16*b*) with respect to K_g produces

$$\frac{\partial|R_{cl}|}{\partial K_g} = \frac{\Re(R_{cl})(-\Re(R_{ol})+\Re(R_{set}))}{|R_{cl}|} \ldots + \frac{\Im(R_{cl})(-\Im(R_{ol})+\Im(R_{set}))}{|R_{cl}|} \tag{17a}$$

$$\frac{\partial\varphi_{cl}}{\partial K_g} = -\frac{\Im(R_{ol})\Re(R_{set}) - \Im(R_{set})\Re(R_{ol})}{|R_{cl}|^2} \tag{17b}$$

Figure 16 shows how $|R_{cl}|$ and φ_{cl} (top) and their derivatives (bottom) develop with increasing control gain K_g. The depicted curves represent the case for which the downstream reflection coefficient was tuned to result in an instability at 85 Hz.

For $K_g=0$, $|R_{cl}|$ and φ_{cl} correspond to those of the uncontrolled case. If the actuator would have sufficient control authority, it would be possible to impose the full control gain $K_g=1$. The adjusted downstream reflection coefficient $|R_{cl}|$ would then match the prescribed one $|R_{set}|$ (see Eq. (15)). In the example shown, $|R_{cl}|$ (solid) increased almost linearly with K_g, whereas φ_{cl} (dashed) had a larger slope for small K_g, which decreased with K_g. Consequently, if the control gain is high enough, the deviation of the actual phase to the prescribed one is relatively small compared with the deviation in magnitude. To produce an instability at 85 Hz, K_g had to be limited to 0.8. In this example, the difference between the uncontrolled phase and the one to be adjusted was quite small. Figure 17 depicts results for generating an instability at 175 Hz. Although the initial difference was much larger in this case, similar observations were made.

Increasing K_g at first resulted in a decrease of $|R_{cl}|$. The phase was only slightly changed. Between $K_g=0.3$ and 0.5, φ_{cl} was rapidly changed and almost tuned to its end value at $K_g=0.6$. For this control gain, however, the magnitude of R_{cl} was still relatively low. This means that the phase relationship and therefore the frequency of oscillation do not change when K_g is further increased

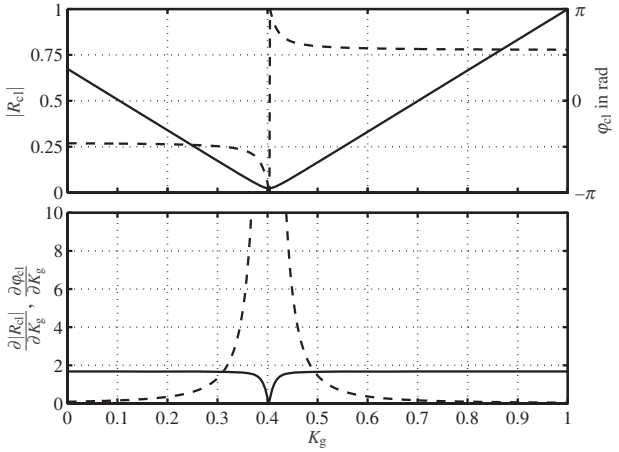


Fig. 17 Influence of the limitation of control signal on closed-loop reflection coefficient ($|R_{cl}|$ solid; φ_{cl} dashed). The system is tuned to generate an instability at 175 Hz.

from 0.6 to 1. In contrast to that, the amplitude of oscillation might be strongly influenced. Note that the actuator in this case had sufficient power and K_g could be adjusted to 1.

In both examples, the sensitivity of φ_{cl} decreased with increasing K_g. If K_g was high enough, the influence on φ_{cl}, and therefore on the frequency of oscillation, could be neglected. This explains why the system oscillated at the prescribed frequency, although the control gain was limited. For all experiments described in this work, for which the control gain had to be limited, a similar behavior was found. However, it should be noted that the dependences of $|R_{cl}|$ and φ_{cl} may differ for other cases.

5 Conclusions and Outlook

An active control scheme was used to manipulate the acoustic boundary condition, i.e., the downstream reflection coefficient, of a combustion test rig. This was achieved by decomposing the plane acoustic field in the system *online* by means of a time domain formulation of the multimicrophone method. The pressure measurements were fed back to an algorithm, which generated a control signal to drive an acoustic actuator. Through an adequate adjustment of the control law, the system behavior subject to quasi-arbitrary acoustic boundary conditions could be studied. This could be a crucial step in the development process of new burner generations, as new burners could be tested under realistic engine conditions regarding their thermoacoustic behavior. Since the occurrence of thermoacoustic instabilities in the full-scale engine has to be ruled out, the impedance tuning concept could save a lot of time and result in financial benefits in the design process.

The feasibility of actively tuning the acoustic boundary condition to a prescribed value strongly depends on the actuator's control authority. In this work, two different actuator concepts were investigated and compared with each other. In a first step, woofers were used, exhibiting the best characteristics concerning the response over a broad frequency range and the time-delay between the control signal and the generated acoustic wave. With the woofers, excellent results were obtained for tuning the downstream reflection coefficient in a broad frequency range ($\sim$250 Hz) to fully reflecting or anechoic. However, application of the woofers to industrial test rigs is limited, as they do not exhibit enough acoustic power and will not withstand the harsh environmental conditions. These shortcomings could be overcome by the second type of actuators, proportional valves, used to modulate an air mass flow to generate an acoustic wave. The results obtained with these valves proved the concept to work equally well. The capability to work over a broad frequency range at once, however, was limited due to the actuator time-delay. The maximal frequency range possible was 90 Hz. Therefore, the control concept was modified to work at discrete frequencies. Through this, several difficulties, e.g., tuning at high frequencies and the controller identification, could be avoided.

Application of the impedance tuning concept to Siemens PG test rigs at ambient and elevated pressure is a subject of ongoing work.

Acknowledgment

The work presented in this paper was conducted within the framework of AG TURBO COOREFF-T 2.2.2. The authors gratefully acknowledge financial support from the Federal Ministry of Economics and Technology (BMWi) and SIEMENS AG Power Generation.

Nomenclature

c = speed of sound
D = wave decomposer
e = control command
f = frequency
f = downstream traveling wave
G = actuator transfer function
g = upstream traveling wave
i = imaginary unit
K = controller transfer function
l = length
n = integer
p = acoustic pressure (scaled by ρc)
q = fluctuation of the OH-radical light emission intensity
R = reflection coefficient
t = time
U_0 = mean flow velocity
v = axial acoustic particle velocity
x = axial coordinate
Z = (specific) acoustic impedance

Symbols

λ = wavelength
ρ = density
τ = time-delay
φ = phase angle
ω = angular frequency

Subscripts and Superscripts

cl = closed loop
ds = downstream
g = gain
ol = open loop
us = upstream
$(\cdot)^-$ = against the mean flow direction
$(\cdot)^+$ = in the mean flow direction
$\hat{(\cdot)}$ = Fourier transform

References

[1] Lieuwen, T. C., and Yang, V., eds., 2005, *Combustion Instabilities in Gas Turbine Engines* (Progress in Astronautics and Aeronautics Vol. 210), AIAA, Reston, VA.
[2] Rea, S., James, S., Goy, C., and Colechin, M. J. F., 2003, "On-Line Combustion Monitoring on Dry Low NO_x Industrial Gas Turbines," Meas. Sci. Technol., **14**(7), pp. 1123–1130.
[3] Sewell, J., Sobieski, P., and Beers, C., 2004, "Application of Continuous Combustion Dynamics Monitoring on Large Industrial Gas Turbines," ASME Paper No. GT2004-54310.
[4] Mongia, H., Held, T., Hsiao, G., and Pandalai, R., 2003, "Challenges and Progress in Controlling Dynamics in Gas Turbine Combustors," J. Propul. Power, **19**(5), pp. 822–829.
[5] Streb, H., Prade, B., Hahner, T., and Hoffmann, S., 2001, "Advanced Burner Development for the VX4.3A Gas Turbines," ASME Paper No. 2001-GT-0077.
[6] Seume, J. R., Vortmeyer, N., Krause, W., Hermann, J., Hantschk, C.-C., Zangl, P., Gleis, S., Vortmeyer, D., and Orthmann, A., 1998, "Application of Active Combustion Instability Control to a Heavy Duty Gas Turbine," ASME J. Eng. Gas Turbines Power, **120**, pp. 721–726.
[7] Richards, G. A., Thornton, J. D., Robey, E. H., and Arellano, L., 2007, "Open-Loop Active Control of Combustion Dynamics on a Gas Turbine Engine," ASME J. Eng. Gas Turbines Power, **129**(1), pp. 38–48.
[8] Bellucci, V., Flohr, P., Paschereit, C. O., and Magni, F., 2004, "On the Use of Helmholtz Resonators for Damping Acoustic Pulsations in Industrial Gas Turbines," ASME J. Eng. Gas Turbines Power, **126**(2), pp. 271–275.
[9] Berenbrink, P., and Hoffmann, S., 2000, "Suppression of Dynamic Combustion Instabilities by Passive and Active Means," ASME Paper No. 2000-GT-0079.
[10] Richards, G. A., Straub, D. L., and Robey, E. H., 2003, "Passive Control of Combustion Dynamics in Stationary Gas Turbines," J. Propul. Power, **19**(5), pp. 795–810.
[11] Poinsot, T. J., Trouvé, A. C., Veynante, D. P., Candel, S. M., and Esposito, E. J., 1987, "Vortex-Driven Acoustically Coupled Combustion Instabilities," J. Fluid Mech., **177**, pp. 265–292.
[12] Torres, H., Lieuwen, T C., Johnson, C., Daniel, B. R., and Zinn, B. T., 1999, "Experimental Investigation of Combustion Instabilities in a Gas Turbine Combustor Simulator," AIAA Paper No. 99-0712.
[13] Bothien, M. R., Moeck, J. P., and Paschereit, C. O., 2007, "Impedance Tuning of a Premixed Combustor Using Active Control," ASME Paper No. GT2007-27796.
[14] Guicking, D., and Karcher, K., 1984, "Active Impedance Control for One-Dimensional Sound," ASME J. Vib., Acoust., Stress, Reliab. Des., **106**, pp. 393–396.
[15] Li, Y., Chiu, G.-C., and Mongeau, L., 2004, "Dual-Driver Standing Wave

Tube: Acoustic Impedance Matching With Robust Repetitive Control," IEEE Trans. Control Syst. Technol., **12**(6), pp. 869–880.
[16] Bothien, M. R., Moeck, J. P., and Paschereit, C. O., 2007, "Experimental Validation of Linear Stability Analysis in Premixed Combustors Supported by Active Control," *Proceedings of the 14th International Congress on Sound and Vibration*, Cairns, July 9–12.
[17] Moeck, J. P., Bothien, M. R., and Paschereit, C. O., 2007, "An Active Control Scheme for Tuning Acoustic Impedances," AIAA Paper No. 2007-3540.
[18] Bothien, M. R., Moeck, J. P., Lacarelle, A., and Paschereit, C. O., 2007, "Time Domain Modelling and Stability Analysis of Complex Thermoacoustic Systems," Proc. Inst. Mech. Eng., Part A, **221**(5), pp. 657–668.
[19] Schuermans, B., Bellucci, V., Guethe, F., Meili, F., Flohr, P., and Paschereit, C. O., 2004, "A Detailed Analysis of Thermoacoustic Interaction Mechanisms in a Turbulent Premixed Flame," ASME Paper No. 2004-GT-53831.
[20] Paschereit, C. O., Schuermans, B., Polifke, W., and Mattson, O., 2002, "Measurement of Transfer Matrices and Source Terms of Premixed Flames," ASME J. Eng. Gas Turbines Power, **124**(2), pp. 239–247.
[21] Gustavsen, B., and Semlyen, A., 1999, "Rational Approximation of Frequency Domain Responses by Vector Fitting," IEEE Trans. Power Deliv., **14**(3), pp. 1052–1061.
[22] Schuermans, B., Bellucci, V., and Paschereit, C. O., 2003, "Thermoacoustic Modeling and Control of Multi Burner Combustion Systems," ASME Paper No. 2003-GT-38688.
[23] Krebs, W., Flohr, P., Prade, B., and Hoffmann, S., 2002, "Thermoacoustic Stability Chart for High-Intensity Gas Turbine Combustion Systems," Combust. Sci. Technol., **174**(7), pp. 99–128.
[24] Kopitz, J., Huber, A., Sattelmayer, T., and Polifke, W., 2005, "Thermoacoustic Stability Analysis of an Annular Combustion Chamber With Acoustic Low Order Modeling and Validation Against Experiment," ASME Paper No. GT2005-68797.
[25] Moeck, J. P., Bothien, M. R., Paschereit, C. O., Gelbert, G., and King, R., 2007, "Two-Parameter Extremum Seeking for Control of Thermoacoustic Instabilities and Characterization of Linear Growth," AIAA Paper No. 2007-1416.
[26] Zhao, D., and Morgans, A. S., 2007, "Tuned Passive Control of Combustion Instabilities Using Multiple Helmholtz Resonators," AIAA Paper No. 2007-3423.
[27] Bechert, D. W., 1980, "Sound Absorption Caused by Vorticity Shedding, Demonstrated With a Jet Flow," J. Sound Vib., **70**(3), pp. 389–405.
[28] Paschereit, C. O., Gutmark, E., and Weisenstein, W., 2000, "Excitation of Thermoacoustic Instabilities by Interaction of Acoustics and Unstable Swirling Flow," AIAA J., **38**(6), pp. 1025–1034.
[29] Levine, H., and Schwinger, J., 1948, "On the Radiation of Sound From an Unflanged Circular Pipe," Phys. Rev., **73**(4), pp. 383–406.

The Reheat Concept: The Proven Pathway to Ultralow Emissions and High Efficiency and Flexibility

Felix Güthe[1]
e-mail: felix.guethe@power.alstom.com

Jaan Hellat

Peter Flohr

Alstom,
Brown-Boveri-Strasse 7,
CH-5400 Baden, Switzerland

Reheat combustion has been proven now in over 80 units to be a robust and highly flexible gas turbine concept for power generation. This paper covers three key topics to explain the intrinsic advantage of reheat combustion to achieve ultralow emission levels. First, the fundamental kinetic and thermodynamic emission advantage of reheat combustion is discussed, analyzing in detail the emission levels of the first and second combustor stages, optimal firing temperatures for minimal emission levels, as well as benchmarking against single-stage combustion concepts. Second, the generic operational and fuel flexibility of the reheat system is emphasized, which is based on the presence of two fundamentally different flame stabilization mechanisms, namely, flame propagation in the first combustor stage and autoignition in the second combustor stage. This is shown using simple reasoning on generic kinetic models. Finally, the present fleet status is reported by highlighting the latest combustor hardware upgrade and its emission performance.
[DOI: 10.1115/1.2836613]

1 Introduction

In the mid 1990s, Alstom introduced two similar sequential combustion gas turbines: the GT24 for the 60 Hz market and the GT26 for the 50 Hz market. Since its first launching in 1995 [1], the advanced class GT24/GT26 engines have demonstrated that this technology platform does offer superior operating flexibility, low emissions, and high part-load efficiency with world-class levels of reliability and availability.

The main technology differentiator of Alstom's GT24/GT26 (Fig. 1) gas turbines is the sequential combustion principle, which was already introduced in 1948 into the market as a way of increasing efficiency at low turbine inlet temperature levels. The GT24/GT26 combustion system is based on a well-proven Alstom combustion concept using the environmental (EV) burner in an annular combustor followed by the SEV (sequential environmental) burner in the second combustion stage (see Fig. 1). This dry low NO_x EV burner has a long operating history and is used in the whole Alstom gas turbine range. Sequential combustion, "the reheat principle for gas turbines," had already been applied to earlier (Brown Boveri) engines but using two side-mounted silo combustors. Integrating the concept of dry low NO_x EV burner and sequential combustion into a one shaft engine resulted in the GT24/GT26—a machine with a high power density and a small footprint. With a pressure ratio of $>30:1$, the compressor delivers nearly double the pressure ratio of a conventional compressor.

The compressed air is heated in a first combustion chamber (EV combustor) by adding about 50% of the total fuel (at base load). After this, the combustion gas expands through the high-pressure (HP) turbine, which lowers the pressure by approximately a factor of 2. The remaining fuel is added together with some additional cooling air in the second combustion chamber (SEV combustor), where the combustion gas is heated a second time to the maximum turbine inlet temperature and finally expanded in the four-stage low-pressure (LP) turbine.

Relative to a conventional nonreheat cycle, the same specific power output is achieved at lower turbine inlet temperature. This is illustrated in Fig. 2.

2 Basic Reheat Features for Low Emissions

The low emission levels, which can be achieved with a reheat system, are the combined effect of three key mechanisms: First, a reheat combustor makes a more efficient use of the oxygen by burning twice in the lean premix mode. Second, there exists a chemical advantage of reheat combustion, which can be exploited. Third, the flame stabilization by autoignition leads to increased flexibility, which allows operating in low emission mode at a wide load range by avoiding high peak flame temperatures T_{flame}, which lead to exponential increase in NO_x, for both combustors. We explain these effects in detail below.

2.1 Impact of Lower O_2 Level. The reheat concept also makes very efficient use of the O_2 from the air by operating the second stage at approximately 15% O_2 at the inlet. Overall this leads to very high power densities. The O_2 content in the exhaust is in the order of ~10%, yielding a beneficial normalization factor (for 15% O_2) of $(21-15\%)/(21-10\%)$ ~0.55 for the absolute NO_x, as shown in Fig. 3 on the right hand axis. The correction factor decreases with higher temperature (~ consumed fuel ~O_2) and is lower for the reheat engines due to its increased O_2 consumption.

This leads to varying absolute levels for a fixed "15%" guarantee value depending on reheat or nonreheat combustion and T_{flame} and allows higher absolute emission values after two combustion chambers. This correction has been introduced to make a fair comparison of power plants based on their emissions relative to consumed air or fuel. Emissions quantified in the widely used

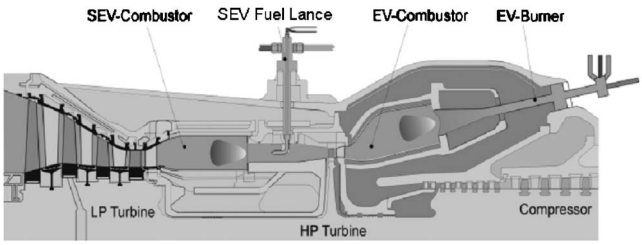


Fig. 1 GT24/GT26 sequential combustion system

[1]Corresponding author.

Manuscript received May 30, 2007; final manuscript received October 8, 2007; published online December 24, 2008. Review conducted by Dilip R. Ballal.

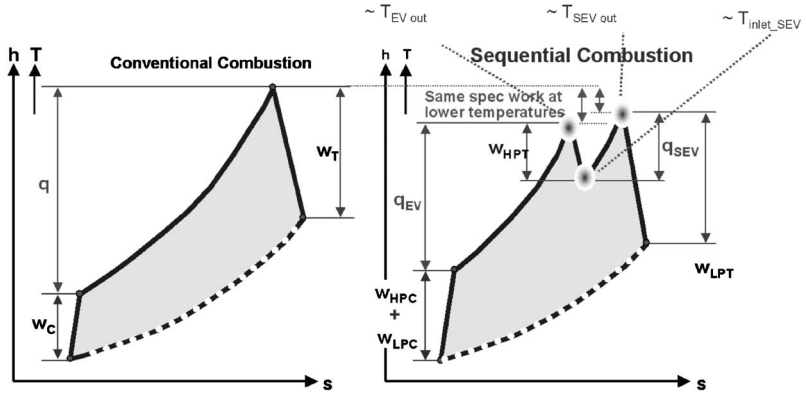


Fig. 2 Principle of the reheat cycle compared to a standard GT cycle

emission index defined as (atmospherically produced) $g(NO_2)/kg$ fuel are reduced accordingly.

2.2 NO_x Characteristics of Reheat Combustion. A second reason for the low NO_x emission is the combustion in the SEV at reduced O_2 levels and increased H_2O contents from the first combustor ($\sim 5\%$). This allows the SEV combustor to operate at higher firing temperatures and produce less NO_x than an EV combustor would produce at the same temperature.

Results of CHEMKIN [2] simulations using the GRI-MECH [3] for the different flame types are shown in Fig. 3, where the laminar flame speed module PREMIX has been used for the EV [4] and the plug flow reactor (PFR) module for SEV. A detailed network model based on PREMIX and PFR has been developed for an accurate emission prediction for realistic engine conditions, but its details are out of the scope of this paper.

A simple CHEMKIN calculation shows that NO_x produced in the second combustor is essentially unaffected by NO_x levels arising from the EV combustor as long as they are of moderate magnitude. The SEV calculation seems simply to be offset by the amount of EV NO_x given at the start. Therefore, a chemical model of the engine treats both combustors as independent NO_x producers. Although the GRI is assumed to describe the SEV combustion sufficiently, further efforts are taken to investigate possible effects of NO_x and water as well as reduced O_2 content in ongoing projects in connection with shock tube studies.

Shown in Fig. 3 are the NO_x contributions versus T_{flame} of a single combustor (EV) at approximately 30 bars and of a SEV-type combustor at approximately 15 bars, as well as a combination of the two with fixed T_{flame} for the first combustor (EV) for the idealized situation of perfect mixing without heat losses. NO_x emissions are given on an absolute scale normalized to the value for a single combustor (EV) at T_{ref}. This allows the SEV contribution (at reduced O_2) to be shown on the same plot. The curves are analytical fits to CHEMKIN II results of the GRI3.0. The reheat engine results in lower emissions at the same T_{flame}. The calibration factor for 15% O_2 refers to the right hand scale. The NO_x emissions of the full reheat engine are taken as the sum of the absolute emissions of the two combustors assuming no interaction of NO_x from the first combustor on the second.

The reduced NO_x production of the reheat combustor is among other possible factors due to the decreased concentration of O radicals, which is influenced by the reduced O_2 and the increased

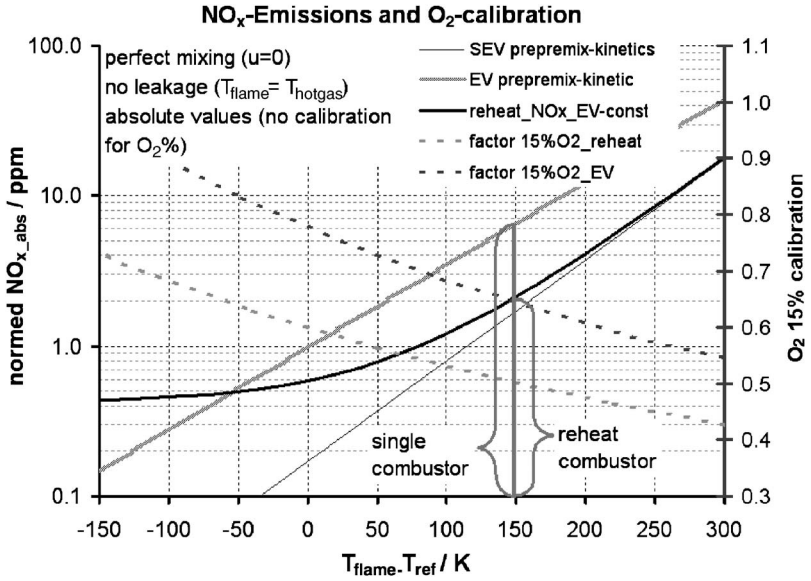


Fig. 3 NO_x emissions on log scale (normalized, noncalibrated) for a single combustor (EV), the SEV combustor, and a reheat combustion system. The factors calibrating to 15% O_2 depend on T_{flame} and are shown in green and refer to the right hand scale.

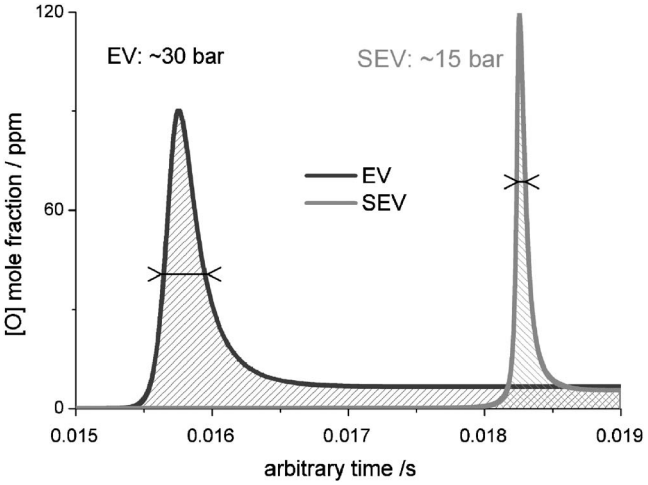


Fig. 4 O-radical concentration profiles for EV and SEV flames (taken from PREMIX and PFR calculations) for similar T_{flame}. The arbitrary time scale extends over 4 ms. The shaded area in the integral $\int t[O]_{dt}$ determines the most important NO_x formation routes and is approximately three times higher in the EV than in the SEV.

H_2O content. The O radicals are an essential part of the main route to produce NO_x by oxidizing the otherwise stable N_2 molecule according to the kinetic rate law:

$$\frac{d[\mathrm{NO}]}{dt} = k(T,p)[\mathrm{N_2}][\mathrm{O}] \tag{1}$$

$$[\mathrm{NO}] = k(T,p)[\mathrm{N_2}]t[\mathrm{O}] \tag{2}$$

Since k and the N_2 concentration do not vary much for given temperatures and pressures, the NO_x formation depends mainly on the time t and the concentration O radicals. The time profiles of this species are plotted in Fig. 4, revealing the integral to be approximately three times smaller, which is the combination of short residence times and low O concentration due to H_2O and O_2 contents of the vitiated air as explained above. For such laminar cases, the prompt NO_x formation due to the O radicals is estimated to be approximately three to six times smaller at given T_{flame} considering O-concentration profiles. To invoke the often invoked reburn chemistry to explain the low NO_x emissions seems not to be necessary and is believed to be unlikely since the heat release is occurring entirely in fuel lean conditions.

Another factor reducing SEV-NO_x emissions is the significantly reduced residence time in the SEV compared to the EV. Some additional benefit might arise from the fact that the fast SEV gas velocities of part of the rich combustion products in the SEV result in faster quenching of this hot pockets and subsequently less time to produce NO_x. It can be summarized that the NO_x production in the SEV starts at a significantly lower value.

2.3 Reheat Combustion With Autoignition. The reheat concept allows a very flexible operation of the engine, enabling us to find an optimized operating range for the given conditions and fuel compositions. The reason for that is described in the following section.

While the EV inlet temperature T_{inlet_EV} is determined by the high-pressure compressor, in the SEV the inlet temperature T_{inlet_SEV} is governed by the outlet temperature of the HP turbine. The SEV flame temperature, T_{flame_SEV}, determines the inlet temperature of the LP turbine. Accordingly, the HP-turbine inlet temperature (and, along with it, T_{inlet_SEV}) is controlled by the EV flame temperature, T_{flame_EV}.

The reactor inlet temperature T_{inlet} in the chemical model is derived from mixing cold fuel and hot combustion air before the reaction, yielding an effective T_{mix} as T_{inlet}. The shorter residence time of the products in the combustor is accounted for in the model.

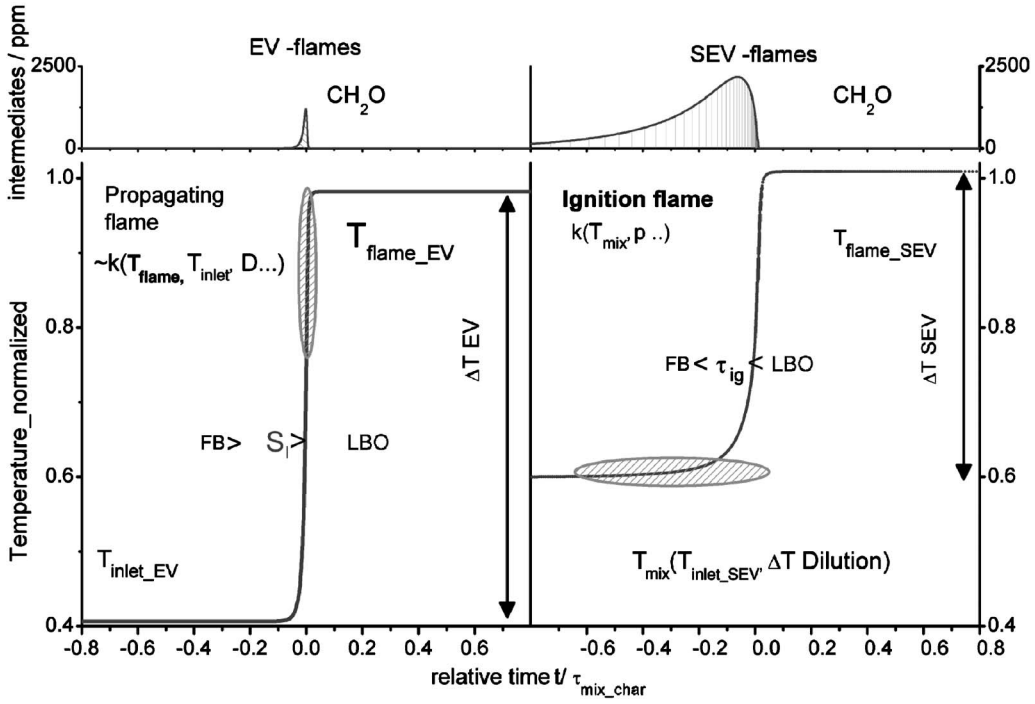


Fig. 5 Schematic diagram explaining the physics and chemistry of the EV flame (left) and the SEV flame (right) for natural gas. The intermediates indicating preflame zones (CH_2O) are displayed in the upper graphs. The shaded areas indicate the most relevant regions of reactivity. The different time axes were given in reverence to the typical mixing times.

The higher $T_{\text{inlet\_SEV}}$ results in a much higher reactivity of the fresh gases, a different mechanism being important for the description of the flame. This higher reactivity is partly compensated by higher burner velocity leading to shorter characteristic mixing times $\tau_{\text{mix\_SEV}}$ ($\sim$ ignition time $\tau_{\text{ignition}} \sim 1-2$ ms) available for mixing for the SEV combustor, leaving some residual unmixedness in the flame region compared to the EV flame, where a characteristic mixing time $\tau_{\text{mix\_EV}}$ ($\sim 5-10$ ms) is given by the traveling time from fuel injection to flame front.

It should be pointed out that Fig. 3 indicates ideal conditions. For a real system, effects of local unmixedness and heat losses need additional consideration.

This high reactivity at high combustor inlet temperatures in the SEV has the consequence that flame stabilization occurs always under premixed conditions by autoignition, and no piloting is needed to stabilize the SEV flame. The T_{flame} of the two combustors are ideally optimized primarily with respect to performance and turbine lifetime, but also at a given load for equal amounts of NO_x produced, deriving an absolute minimum in engine emissions.

2.4 Optimal Firing of Combustor for Minimal Emissions and Maximum Flexibility. The reheat concept enables us to operate the two combustors at different temperatures with only little impact on the overall power output. This increases the flexibility of the combustion system, allowing us to burn more or less reactive fuels in the same engine by adjusting the relative load between the two combustors and to mitigate ranges where operation would become difficult. This can be utilized for different fuels (C2+ or even H_2 containing fuels), exhaust gas recirculation, or changes in ambient conditions.

In general, the engines are adjusted so that the first combustor runs at low firing temperatures, keeping the emissions as low as possible for conventional premix combustion, while the second combustor is operated at higher temperatures, producing similar emissions (on absolute scale) due to the reasons mentioned above. Note that while the EV runs more or less at constant conditions (see Fig. 10), keeping low NO_x even at very low loads, the loading is achieved by varying the fuel input in the SEV starting at as low as $\sim$25% GT load (as well as increasing the mass flow of air). The high $T_{\text{inlet\_SEV}}$ allows operation without piloting for the entire operation range, entirely relying on autoignition for the SEV flame.

The conceptual difference between EV and SEV flames is presented in Fig. 5. The flame has to be stabilized between limits for flashback (FB) and lean blowoff (LBO) determined by the laminar flame speed S_L for the EV flame, which depends on conditions around the adiabatic T_{flame} and less on T_{inlet}. The SEV flame is dominated by the chemistry occurring in the preflame zone at T_{inlet}. This includes the buildup of a radical pool up to a critical threshold. The flame position is here determined by the ignition time. To compare the two different combustion regimes, the time has been scaled to zero for the region of maximum heat release and normalized to the characteristic mixing times as defined before.

For the inlet conditions of the EV, autoignition times are much too long to play a role even for the most reactive fuels. Flame propagation is the leading mechanism for a stationary reaction zone stabilization, as can be calculated with PREMIX for the laminar case (laminar flame speed S_L). The description of the turbulent flame propagation is derived from that concept. As for every conventional premix burner, the EV flame stabilization happens within the boundaries of flame blowoff at a low flame temperature limit and flame FB at a high flame temperature limit. These flame propagation limits are sketched on the left side of Fig. 5. The reactivity is controlled by transport across the flame front and determined by the kinetic rates at high temperatures close to the adiabatic one. Intermediates are abundant only close to the flame front. The reactivity (Fig. 6) depends on the kinetics of the fuel

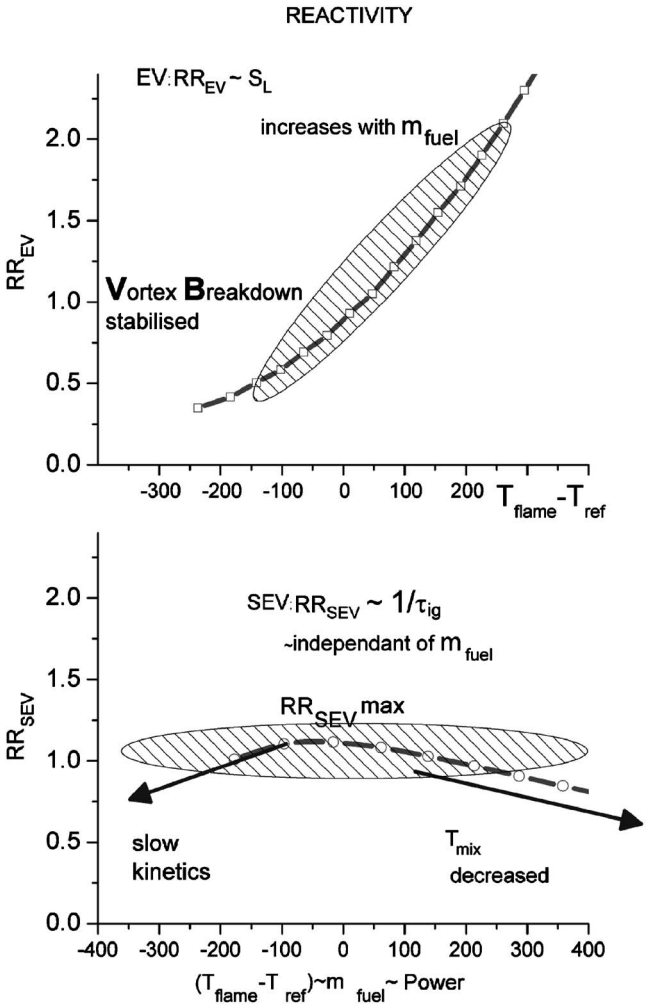


Fig. 6 EV and SEV reactivities as defined by Eqs. (3) and (4) for varied fuel contents and constant T_{inlet} (compare with Fig. 8). The shaded area defines an operational range.

and its transport properties at high temperatures close to the final $T_{\text{flame\_EV}}$ and only slightly on pressure and $T_{\text{inlet\_EV}}$. For the EV, the relative reactivity RR_{EV} is defined with reference to one operating point as the ratio of laminar flame speeds in

$$\text{RR}_{\text{EV}} = S_L / S_{L-\text{ref}} \tag{3}$$

This flame stabilization is fundamentally different for the SEV flame, which is controlled by autoignition (calculated with PFR) arising from reactive intermediates forming and accumulating up to a critical threshold before the flame front. The chemistry at lower temperatures ($\sim T_{\text{inlet\_SEV}}$) is responsible for the reactivity. This means that in the engine the SEV flame is controlled and governed by the $T_{\text{flame\_EV}}$, while $T_{\text{flame\_SEV}}$ has only little influence on the flame position. The relative reactivity RR_{SEV} is defined in Eq. (4) from the ignition times as

$$\text{RR}_{\text{SEV}} = \tau_{\text{ig\_ref}} / \tau_{\text{ig}} \tag{4}$$

These reactivity definitions enable comparing combustor specific definitions on one scale (Figs. 6–8).

Plotting RR over T_{flame} ($\sim 1/\phi \sim \lambda$ for lean conditions) in Fig. 6 leads to the contrary effects: For this laminar approximations, the EV flame becomes more reactive with T_{flame}, but the SEV flame becomes less reactive because the cooling effect of adding more cold fuel to the heated air from the first combustor controls the reactivity while the reactivity (toward ignition) is not increased by a higher concentration of fuel. In the very lean region, an increase

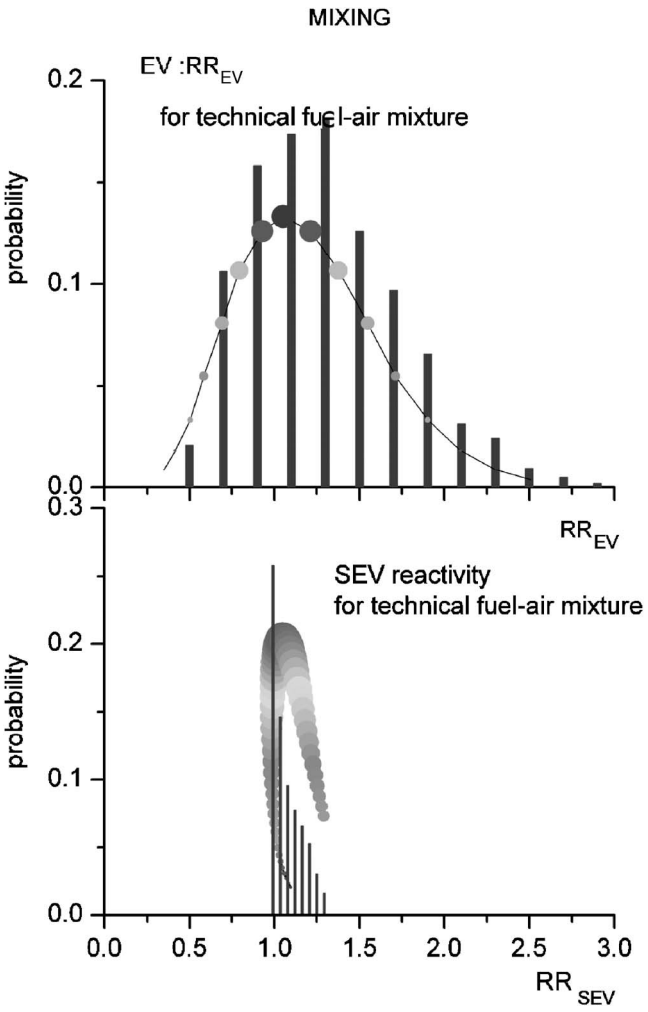


Fig. 7 Probability of reactivity derived from a distribution of fuel for EV and SEV combustors at fixed inlet conditions. The histograms show the probability for a given reactivity, while the bubbles are calculated reactivities for a given fuel-mass ratio.

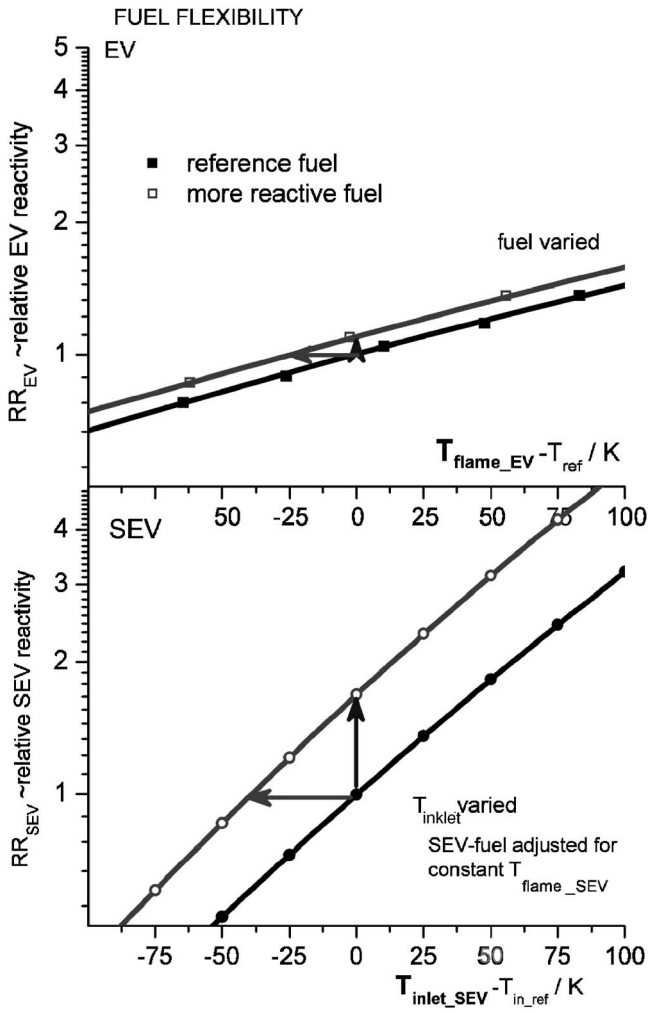


Fig. 8 EV reactivity for varied $T_{\text{flame\_EV}}$ (upper graph) for two different fuels on logarithmic scale. In the lower graph, the reactivity of the SEV for varied inlet conditions ($T_{\text{inlet\_SEV}}$) is presented, referenced to a standard inlet temperature $T_{\text{in\_ref}}$. A reactivity increase due to the fuel is compensated by reduced $T_{\text{inlet\_SEV}}$. Note that the RR_{SEV} is plotted versus T_{flameSEV} in Fig. 6.

of fuel molecules can increase RR_{SEV} slightly. Taken together, these effects counterbalance to give a reactivity maximum and, overall, very small dependence on T_{flame}. The small variation for the SEV reactivity with fuel is the cause for the high flexibility of the reheat concept. The reactivity increases with fuel only for very lean conditions. The chemical acceleration of rich mixtures levels off at intermediate levels, and the reduction of T_{mix} (due to the mixing with cold fuel) outruns this effect, slowing down the reactivity toward rich mixtures.

The EV reactivity RR_{EV} varies much more with fuel; therefore, the EV flame position has to be controlled aerodynamically by attaching it to a flame anchoring point, in this case a stagnation point in the flow field (vortex breakdown) of the swirled EV burner. Effects of autoignition in the EV combustor or flame propagation in the SEV are assumed to be minor and are neglected here.

The operating range of a burner is limited by a minimum reactivity ($\sim S_L$ for the EV), causing a LBO of the flame and a maximum acceptable reactivity preventing the flame to flash back. The EV is governed by the aerodynamic stabilization, allowing this combustor to hold the flame at the same positions over a wide range of reactivity.

The reheat combustion system has the particular property that the flame stabilization in both premix combustors, EV and SEV, is controlled by a single physical quantity—the flame temperature of the first EV combustor. This way, the high flexibility of the engine can be maintained by being able to operate the EV over a range of T_{flame}, therefore enabling us to find an optimum HP-turbine outlet temperature for SEV operation. This optimization can be used for the lowest emissions as well as to handle a variety of fuels and inlet conditions.

The SEV flame position is mainly controlled by its chemical reactivity, which is mainly influenced by $T_{\text{inlet\_SEV}}$ and the fuel composition. The different control mechanisms of flame positions are utilized in the ALSTOM gas turbine design by attaching the EV flame aerodynamically and by controlling the SEV flame through $T_{\text{inlet\_SEV}}$. The fact that the SEV flame is relatively insensitive to T_{flame} leads to a high operational range: SEV load can be increased without increasing FB risk (emissions and pulsations still have to be considered, of course).

This leads to an interesting effect for not perfectly mixed flames. The reactivity probabilities resulting from a given fuel-air unmixedness are plotted in Fig. 7. The same variance in the upper and lower graph lead to a wider distribution of flame speeds for the EV and a very narrow distribution for the SEV (at given HP-turbine outlet temperature). Here the different mechanisms become clear once more by spreading out over a wide range of

reactivities for the EV flame but only a very narrow range for the SEV-flame.

3 Generic Operational and Fuel Flexibility

3.1 Fuel Flexibility for Highly Reactive Fuels. While the reactivity for the EV is controlled by $T_{\text{flame\_EV}}$, the SEV reactivity is mostly controlled by the inlet conditions and $T_{\text{inlet\_SEV}}$. This has some interesting conclusions for the engine operation.

1. The SEV can be varied in $T_{\text{flame\_SEV}}$ without changing its flame position much.
2. An increased reactivity due to a more reactive fuel can be easily compensated by reducing the $T_{\text{flame\_EV}}$ and, therefore, $T_{\text{inlet\_SEV}}$.

This is also illustrated in Fig. 8, which demonstrates the fuel flexibility achievable by varying $T_{\text{flame\_EV}}$. The increase in reactivity (vertical arrow) due to a more reactive fuel can be compensated by reducing $T_{\text{flame\_EV}}$ (horizontal arrow). The reduction of the EV fuel leads to a $T_{\text{inlet\_SEV}}$ reduction, which simultaneously reduces the SEV reactivity. In the SEV, a reactivity increase due to the fuel is compensated by reduced $T_{\text{inlet\_SEV}}$.

The described insensitivity of flame reactivity to $T_{\text{flame\_SEV}}$ allows the inlet temperature of the LP turbine to be constant (in terms of protection for the combustor). Therefore, only little loss in overall power and efficiency due to derating the EV since the power output is largely recovered by adding more fuel to the SEV. This is in contrast to the nonreheat engines, where a derating due to increased fuel reactivity leads to decreased turbine inlet temperature resulting in severe loss of overall power and efficiency.

This key feature can be used to operate the engine [5] such that changes in fuel composition do not impact emission performance and reliability significantly. For example, if for given T_{flame} higher hydrocarbons ("C2+") or even H_2 or CO are added to the fuel, reactivity will increase (i.e., flame speed will increase and ignition time will decrease) and, as a result, the EV combustor will likely produce higher NO_x. Likewise, the SEV combustor may encounter FB risk due to shorter ignition times if no adjustment is foreseen.

For highly reactive fuels, the fuel input in the EV combustor is reduced, which leads to lower SEV burner inlet temperature, to stay below the burner protection limit. This is feasible because the EV burner lean extinction limit is also reduced for increased reactivity in the fuel.

However, a simple adjustment by decreasing the fuel input in the EV combustor and simultaneously increasing the fuel input in the SEV combustor (to keep the total fuel flow constant) will stabilize combustion performance: EV emissions remain low, and SEV flame stabilization is back at its design point. Due to the inherently low emissions of the SEV combustor, the emission penalty will be negligible if the engine is operated in such a way as schematically shown in Fig. 9.

3.2 Dual Fuel Capability. The primary fuel of the GT24/GT26 fleet is gaseous fuel, while fuel oil is often needed as a backup. The sequential combustion system can also be operated with fuel oil, engines in this case being delivered with dual fuel EV and SEV lances instead of the normal gas-only lances. As is normal practice, demineralized water is used for NO_x control, this being combined with the fuel oil in a mixing block before being fed to the lance as an emulsion. The GT24/GT26 engines also have the ability to switch between the two fuels while remaining synchronized to the grid and so providing the operator with arbitration capabilities and full backup security.

3.3 Part-Load Flexibility. The sequential combustion system offers the intrinsic advantage that premix combustion can be applied across the entire load range because the EV combustor is operated at a constant flame temperature through the entire load range, while the (premix) SEV combustor is used to vary loads from approximately 12–100%. This is schematically shown in Fig. 10.

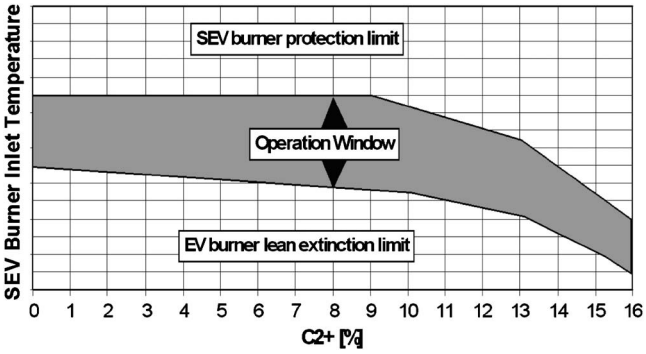


Fig. 9 GT26/GT24 operation concept as a function of fuel reactivity, here given by the content of higher hydrocarbon content (C2+) in the fuel

The so-called "low load operation concept" utilizes the possibility of shutting down the sequential SEV combustor at low part loads and is therefore a unique feature of the GT24/GT26 technology. This concept allows the plant to be operated in a combined cycle mode at a very low combined cycle load (about 15–25%) with the EV combustor operating in the lean premix mode, giving low emission levels as well as a homogenous HP-turbine inlet temperature distribution while keeping the water-steam cycle up and running. The result is a concept that has a number of advantages to the operator because it avoids start-stop cycles and the related cyclic thermal stress to the top and bottom cycle equipment. It enables the operation at low emission levels close to base load levels and reduces cumulative emissions compared to parking a plant at a higher partial load. Also, it assures a homogenous turbine inlet temperature distribution, which is more difficult to achieve for nonreheat engines due to burner piloting at part load.

4 Engine Experience

4.1 Combustor Hardware Status. The continuous drive [6] toward higher performance and lower emissions motivated an upgrade of the EV combustor. The combustor upgrade development, based on the so-called "staged EV burner" is described in detail in a recent paper [7]. Its main feature is the replacement of the central pilot injector by a premix injector. This leads to a burner system with two fuel stages, which can be operated both in the premix mode to vary the fuel profile. In the premix mode, the

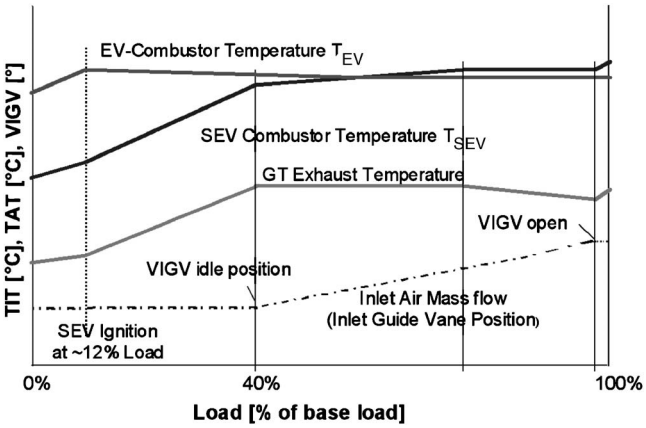


Fig. 10 Schematic of the operating concept of the GT24/GT26 with load variation done through the SEV combustor fuel flow

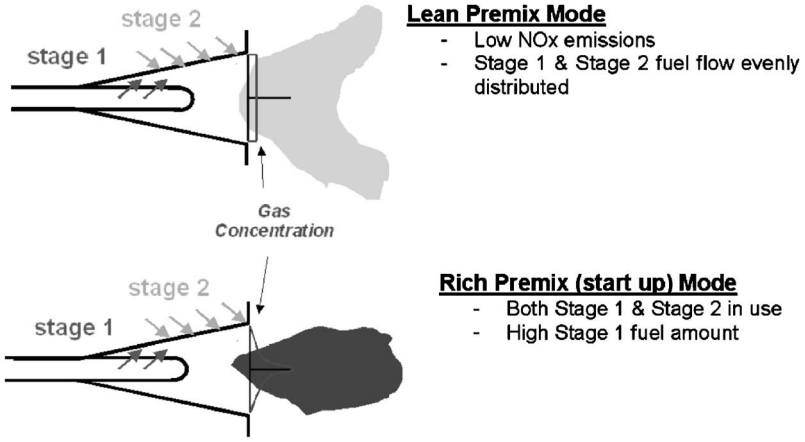


Fig. 11 Staged EV-burner principle. Details of its development are described in Ref. [7].

profile is optimized for minimal emission performance, while in the startup mode, the staging is optimized for maximal stability (see Fig. 11). As a result, this burner has much reduced emission levels in part-load operation, as well as in the base load where the fuel staging between the two stages can always be chosen to be optimal to the base load adjustment point of the engine.

4.2 Emission Characteristics. The emission performance of the staged EV combustor against the standard EV combustor is shown in Fig. 12. This graph shows two effects: First, emission levels are reduced and thus confirm the expected improvement of the staged EV burner. Second, the emission gradient is changed over a wide load range (in this plot between the measured points between 65% and 100% loads). This change in gradient is interesting because it is a direct consequence of the optimization of the EV combustor.

The emission levels of the EV combustor are now so low that the total emissions of the engine start to become dominated by the SEV combustor. As was shown in Fig. 10, the fuel input in the SEV is increased with load and, as a result, now visible in the engine emission performance even if absolute values are very low. Indeed, a more detailed analysis of the combustion system, which includes effects of heat losses and burner imperfections, shows that the relative emission contribution of the SEV combustor increases.

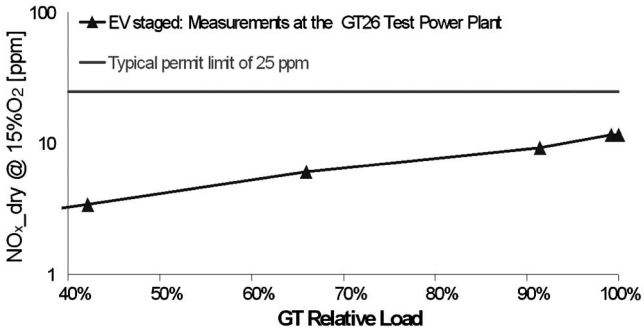


Fig. 12 Emission levels achieved with the staged EV burner on logarithmic scale

5 Summary

In this paper, the key advantages of the reheat system as implemented in the GT24/GT26 engines have been highlighted, namely, its inherently low emission levels as well as its operational flexibility. The contribution of the two combustors to overall emission levels has been discussed in detail and highlights how the latest EV combustor upgrade has improved the overall engine emission performance. Also, it has been explained how the proven operational flexibility of these engines are inherently linked to the complementary combustion kinetics in the two combustion systems.

The achievement of ultralow emission levels while keeping maximal operational flexibility (both in load variation and fuel composition changes) and reliability will remain the key driver for ALSTOM's combustor development. The sequential combustion system is suited very well for this challenge, both today and for the future.

Acknowledgment

The support of Reaction Design in using the CHEMKIN code is gratefully acknowledged.

References

[1] Joos, F., Brunner, P., Schulte-Wenning, B., Syed, K., and Eroglu, A., 1996, "Development of the Sequential Combustion System for the ABB GT24/26 Gas Turbine Family," ASME Paper No. 1996-GT-315.
[2] Kee, R. J., Rupley, F. M., Miller, J. A., Coltrin, M. E., Grcar, J. F., Meeks, E., Moffat, H. K., Lutz, A. E., Dixon-Lewis, G., Smooke, M. D., Warnatz, J., Evans, G. H., Larson, R. S., Mitchell, R. E., Petzold, L. R., Reynolds, W. C., Caracotsios, M., Stewart, W. E., and Glarborg, P., 1999, CHEMKIN Collection, Reaction Design, Inc., San Diego, CA.
[3] GRI-MECH (http://www.me.berkeley.edu/grimech).
[4] Biagioli, F., and Güthe, F., 2007, "Effect of Pressure and Fuel-Air Unmixedness on NO_x Emissions From Industrial Gas Turbine Burners," Combust. Flame, **151**, pp. 274–288.
[5] Riccius, O., Smith, R., Güthe, F., and Flohr, P., 2005, "The GT24/GT26 Combustion Technology and High Hydrocarbon (C2+) Fuels," ASME Paper No. GT2005-68799.
[6] Philipson, S., Ladwig, M., Lindvall, K., and Schmidli, J., 2006, "A Further Uprate for Alstom's Sequential Combustion GT26 Gas Turbine," Power Gen Europe 2006, Cologne, Germany.
[7] Zajadatz, M., Lachner, R., Bernero, S., Motz, C., and Flohr, P., 2007, "Development and Design of ALSTOM's Stage Fuel Gas Injection EV Burner for NO_x Reduction," ASME Paper No. GT2007-27775.

Flow Field and Combustion Characterization of Premixed Gas Turbine Flames by Planar Laser Techniques

Ulrich Stopper
e-mail: ulrich.stopper@dlr.de

Manfred Aigner

Wolfgang Meier

Rajesh Sadanandan

Michael Stöhr

German Aerospace Center,
Institute of Combustion Technology,
Pfaffenwaldring 38-40,
D-70569 Stuttgart, Germany

Ik Soo Kim
Siemens Industrial Turbomachinery Ltd.,
P.O. Box 1,
Waterside South,
Lincoln LN5 7FD, UK

Lean premixed natural gas/air flames produced by an industrial gas turbine burner were analyzed using laser diagnostic methods. For this purpose, the burner was equipped with an optical combustion chamber and operated with preheated air at various thermal powers P, equivalence ratios Φ, and pressures up to p = 6 bars. For the visualization of the flame emissions OH chemiluminescence imaging was applied. Absolute flow velocities were measured using particle image velocimetry (PIV), and the reaction zones as well as regions of burnt gas were characterized by planar laser-induced fluorescence (PLIF) of OH. Using these techniques, the combustion behavior was characterized in detail. The mean flow field could be divided into different regimes: the inflow, a central and an outer recirculation zone, and the outgoing exhaust flow. Single-shot PIV images demonstrated that the instantaneous flow field was composed of small and medium sized vortices, mainly located along the shear layers. The chemiluminescence images reflected the regions of heat release. From the PLIF images it was seen that the primary reactions are located in the shear layers between the inflow and the recirculation zones and that the appearance of the reaction zones changed with flame parameters.*
[DOI: 10.1115/1.2969093]

Keywords: lean premixed combustion, swirl flow, gas turbine combustor, high pressure, NO_x reduction, OH chemiluminescence, laser diagnostics, OH PLIF, PIV, reaction zone, heat release, turbulent flow field, vortex core, validation data*

1 Introduction

The concept of lean premixed combustion is nowadays most often applied in stationary gas turbines (GTs) in order to achieve very low NO_x levels. However, the full potential of this concept for NO_x reduction can often not be exploited because effects of unmixedness, coherent flow structures, or combustion instabilities limit its applicability. In order to gain a better understanding of the complex interactions between the turbulent flow field and chemical reactions in GT-flames, detailed measurements under GT-relevant conditions are necessary. Here, the best available measurement techniques are nonintrusive laser-based methods. The results of such measurements play also an important role for the verification of numerical simulations. Especially in the field of turbulent combustion under high-pressure conditions there is a need of numerical tools, which are capable of correctly determining the flame behavior by taking into account phenomena such as the local extinction and re-ignition. To be able to validate existing models and to develop new numerical approaches, diverse comprehensive sets of measurement data are required. The ideal validation database would contain information about all basic flame properties such as the distributions of species concentrations, temperature, and flow velocity. Furthermore, the simulation of instationary combustion needs highly time-resolved measurements or discrete single-shot data, allowing statements about statistical and coherent deviations from the average flame behavior.

Laser-based measurements in high-pressure GT model combustors are expensive and difficult to perform [1,2]. Due to the limited optical access and the high thermal loads, robust laser techniques are preferably applied. For temperature measurements, coherent anti-Stokes Raman scattering (CARS) [3] and a two-line OH laser-induced fluorescence (LIF) [4] have been used. For the visualization of hot and reacting gases, the planar LIF of OH was frequently applied [5–9]; the LIF was also used for the detection of NO [10,11]. Flow field measurements were mostly performed by laser Doppler velocimetry (LDV) [12–14]. The addition of seeding particles to the flow and their contamination of the windows present an additional challenge in high-pressure GT model combustors. This is even more serious for the application of particle image velocimetry (PIV) and is a main reason for the relatively few publications describing PIV measurements at high-pressure GT conditions [15–17]. However, 2D images of the instantaneous flow field enable new insights into the dynamics of turbulent swirl flames, which are important for the understanding of combustion instabilities.

In the experiments described in this paper, an industrial GT burner was equipped with an optical combustion chamber and was installed in the high-pressure test rig of the German Aerospace Center in Stuttgart. The following diagnostic methods were applied in this project: Chemiluminescence imaging provided information about the distribution of electronically excited OH radicals (OH^*), which is a marker for the reaction zones. OH radicals in the electronic ground state were mapped using PLIF, and the flow field was measured by PIV. In addition to the optical techniques, CO, CO_2, NO, NO_2, and SO_2 were determined by probe techniques in the exhaust, and pressure variations in the combustion chamber were registered by dynamic pressure sensors. By these measurements, the behavior of lean premixed flames was characterized in detail for different pressures, thermal powers, and equivalence ratios. This paper focuses on the implementation and the results of the optical measurements. Parameter studies are pre-

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 10, 2008; final manuscript received April 14, 2008; published online December 30, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

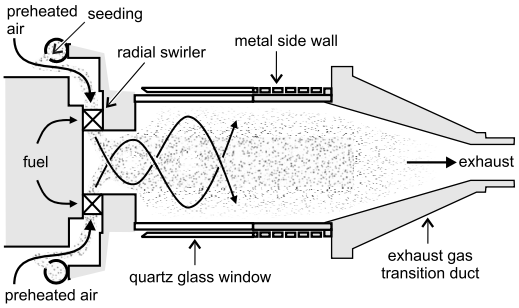


Fig. 1 Cut-away drawing of the combustor. The preheated air is seeded with TiO_2 particles for PIV measurements. A swirl burner mixes the air with natural gas. Each window consists of two quartz plates with a cooling air flowing in between.

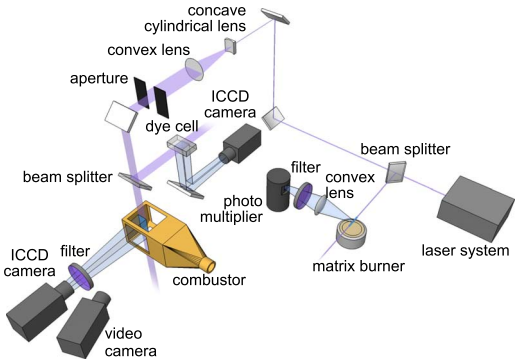


Fig. 2 Experimental setup for the imaging of the OH PLIF and the OH^* chemiluminescence. One ICCD camera images the fluorescence, while the other one records the laser intensity profile.

sented in extracts, whereas probe measurements are not published. One main goal of this complete study was a better understanding of this combustor and the flame characteristics, e.g., the region of flame stabilization, the turbulent structures, the mixing characteristics, and pollutant emissions.

The work presented here is part of an ongoing research project, which will be continued by laser Raman measurements of the joint probability density functions (PDFs) of the temperature and major species concentrations. The final goal is the establishment of an experimental database, which allows also the identification and quantification of effects of unmixedness and finite-rate chemistry. Such "complete" data sets are currently available only for flames at atmospheric pressure, e.g., for turbulent nonpremixed flames [18] or (partially) premixed swirling flames [19,20].

2 Combustor and Test Rig

One aim of this project is to perform the experiments under conditions as close to the industrial case as possible. To meet this objective, an unscaled industrial swirl burner [21] was equipped with a specially designed optical combustion chamber (Fig. 1). It is capable of operation at a thermal power of 1 MW and provides extensive optical access for laser diagnostics. The inner dimensions of the square combustor are based on the size of a commercial gas turbine combustor. Inside, natural gas with preheated air is burnt in a lean premixed combustion. The air preheating temperature was 400°C for all operating points. To reduce the thermal load of the quartz glass, double-glazed windows with cooling air flowing between the glass plates have been installed. The rest of the combustion chamber sidewalls are cooled by impingement cooling. A water-cooled transition duct guides the exhaust gas into the rig's exhaust channel. For the flow field measurements, the gas in the combustion chamber had to be seeded with particles with a diameter of 1 μm. Therefore, a perforated ring tube was installed in the plenum upstream of the radial swirler. During the particle imaging it was used to add a constant flow of seeded air (at a mass flow of up to 4 g/s) to the preheated air. All measurements were performed at pressures p between 1.5 bars and 6 bars and burner pressure drops Δp between 1% and 3%, where the percentage refers to the absolute pressure upstream of the burner.

This combustor was installed at the high-pressure test rig facility HBK-S of the Institute for Combustion Technology in Stuttgart. The pressure vessel of the rig was equipped with glass windows for optical access and several probes for the measurement of temperature, absolute pressure, differential pressure, and acoustics. The concentrations of several species in the exhaust gas were monitored and recorded during every operation.

3 Experimental Implementation

All activities at the test rig facility are organized in measurement phases of several weeks. In the first phase of the project, the imaging of chemiluminescence and PLIF was accomplished, whereas the flow field studies took place during a second phase.

3.1 OH^* Chemiluminescence and OH PLIF Imaging Setup. Figure 2 shows a schematic of the diagnostic system, which was used in the first measurement phase. The laser system consisted of a dye laser pumped by a Nd:YAG laser. The output of the dye laser was frequency doubled and tuned to a wavelength of 283.55 nm, which is resonant with the $Q_1(8)$ line of the A-X (1–0) band of the OH radical. In order to facilitate the adjustment of the laser wavelength and to control its stability, a small portion of the beam was coupled out by a beam splitter and irradiated into a flat laminar methane-air flame stabilized on a matrix burner. The OH LIF signal from this flame was detected by a photomultiplier tube equipped with a spectral bandpass filter that is transparent for wavelengths between 300 nm and 325 nm. The majority of the laser light passes the beam splitter and reaches two lenses, which are set up with a distance of about 64 mm from each other. The first lens ($f=-50$ mm) is shaped cylindrically, whereas the second one ($f=750$ mm) is a spherical lens. This optics produces a laser light sheet, which is guided vertically through the center of the combustion chamber. The focus of the sheet is located in the symmetry axis of the chamber. The LIF of OH radicals is imaged with an intensified charge-coupled device (ICCD) camera equipped with a bandpass filter identical to the one in front of the photomultiplier tube. To correct the PLIF images for the intensity distribution within the laser light sheet, a beam splitter redirects a small portion of the laser sheet into a dye cell. The fluorescence from this cell was recorded by a second ICCD and used for the correction. As the width of the laser sheet is about two or three times smaller than the size of the field of view, it was shifted in the axial direction in order to perform measurements at two different sheet positions. This is done by simultaneously moving all optical elements and detectors except the first beam splitter and the matrix burner with the photomultiplier. The processing of the raw images includes background correction, camera sensitivity correction, and laser intensity profile correction. After these processes the two OH concentration maps from the two corresponding measurements with different sheet positions were merged into one mapping of the whole region of interest.

The imaging of the OH^* chemiluminescence was performed with the same detection system as used for PLIF. Details of the measuring technique are listed in Table 1.

3.2 PIV Setup. The PIV method calculates the vectors of the local streaming velocity by analyzing the displacement of very small seeding particles during a short time span. For measuring this displacement, the particles are illuminated with a double-pulsed laser sheet. A schematic of the PIV setup is shown in Fig. 3. The flashlamp-pumped Nd:YAG laser used in this setup emits

Table 1 Technical details of the PLIF and chemiluminescence diagnostics

OH* chemiluminescence measurement	
Exposure time	40 μm
Frame rate	3 Hz
Effective image size	512×446 pixels
Bandpass filter	300–325 nm
OH PLIF measurement	
Exposure time	400 ns
Laser pulse duration	8 ns
Laser pulse energy	4.6–5.5 mJ
Laser wavelength	283.55 nm
Effective image size	359×446 pixels
Bandpass filter	300–325 nm
ICCD camera specifications (both cameras)	
Dynamic range	16 bits
CCD chip size	512×512 pixels
Objective (flame)	UV Halle, 100 mm, $f/2$
Objective (dye cell)	UV Nikkor, 105 mm, $f/4.5$

short double-pulses at $\lambda=532$ nm that are triggered synchronously with the exposure gates of the PIV camera (Table 2). A combination of two cylindrical lenses ($f_1=-12.7$ mm, $f_2=200$ mm) converts the shape of these pulses into a laser sheet wide enough to cover the whole field of view. 418 mm behind the cylindrical lenses a spherical lens ($f=1000$ mm) focuses the sheet and reduces its planar divergence. Similar to the PLIF setup, a large bending mirror guides the laser light vertically through the central plane of the combustor. The objective of the PIV camera is covered with a narrow band filter (532 nm), reducing the influence of disturbing light sources. The accuracy of the measured velocities depends on the seeding particle density, the degree of window

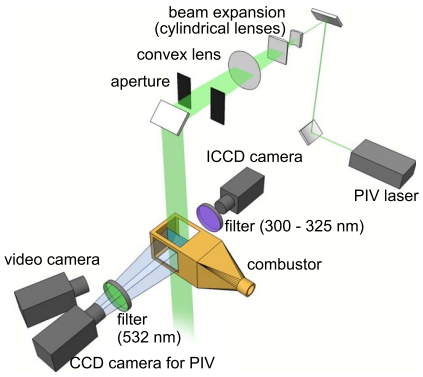


Fig. 3 Experimental setup for the PIV. An additional ICCD camera observes the OH* chemiluminescence. The video camera is used for control purposes only.

Table 2 Technical details of the PIV setup

PIV measurement	
Laser pulse duration	5 ns
Laser pulse energy	120 mJ
Laser sheet width	1 mm
Double-pulse delay	10–15 μs
PIV repetition rate	5 Hz
Seeding material	TiO_2(1 μm)
PIV camera objective	Nikkor, 50 mm, $f/5.6$

contamination, and the turbulence of the flow. On the opposite side of the test rig an ICCD camera with an adequate filter for OH* chemiluminescence is mounted, providing images that can be compared with the results of the previous measurement phase, in order to check whether the flame behavior is identical.

At good PIV measurement conditions, an interrogation window size of 12×12 pixels could be achieved. The accuracy of the single-shot velocity estimates can reach ±0.7 m/s (assuming a subpixel accuracy of particle displacement of 0.1 pixel and neglecting the influence of unrecognized erroneous vectors).

3.3 Experimental Challenges. As temperatures in the combustion chamber increase above 1600°C, the windows of the chamber have to sustain very high heat loads. Although the cooling system reduces the temperature of the glass, the surface of the interior window is still heated up to 1000°C. At this temperature the glass begins to soften and its surface becomes rough and increasingly opaque. Impurities on the window surface glow red and orange. This thermal radiation does not affect the measured chemiluminescence, which is in the UV regime and can be isolated using a spectral filter. More interfering is the effect the high temperature has on the opacity of the glass. The temperature distribution over the glass window is asymmetric as the swirled flow rotates through a noncylindrical combustion chamber. This in turn leads to an asymmetry in the opacity of the windows as the aforementioned roughness increases. The majority of the asymmetry in the chemiluminescence data can be ascribed to this asymmetric window transparency. Additionally, single particles on the window surface are reproduced as small spots in the chemiluminescence images.

During the PIV measurements the window staining was significantly increased by seeding particles adhering to the quartz glass. As a consequence, the visibility of the particles flowing through the central plane of the combustor decreased, while impurities accumulated on the window, producing more and more stationary spots in the image. Although stationary background signals in the series of single shots could be mostly eliminated by applying special filter routines, the described process caused a rapid degradation of the image quality. Below a certain signal-to-noise ratio the cross correlations calculated by the analyzing software were too falsified to enable the detection of the correct flow vector, resulting in an erroneous vector field. Depending on the window opacity this caused the single-shot analysis either to produce single blank regions or to completely fail. The former case might still be usable for calculating the temporal average of the flow field. However, if an averaged flow field is composed out of many low-quality single shots, it can contain artificially lowered absolute values of the velocity due to erroneous vectors that statistically do not favor any direction. This problem often can be bypassed by applying a method called "ensemble correlation" or "sum of correlation" (SOC). This method first sums up the cross correlations of all single shots and calculates the vector field afterward [22]. The result is the flow field with the maximum likelihood, which for the very most cases is practically identical with the temporal average of the flow field. The section of this paper about the analysis of the temporally averaged data does not differentiate between the temporal average and the results produced by the use of the SOC.

4 Experimental Results

The following four remarks should be kept in mind while interpreting the experimental results.

•Chemiluminescence imaging is a line-of-sight technique. This means that the detected signal intensity is not emitted from a certain excitation plane (as with PLIF or PIV). The emission is rather integrated along the whole line-of-sight. However, for the central parts of the flame we can assume a cylinder symmetry of the averaged OH* distribution, which allows us to apply the

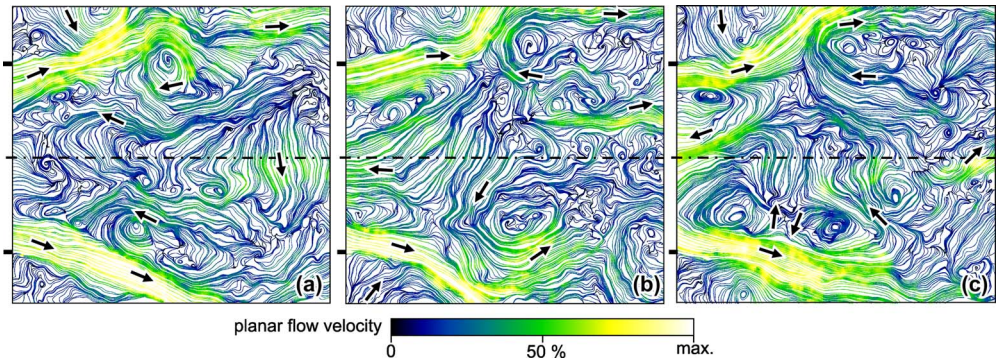


Fig. 4 Streamline plots of three successive instantaneous PIV measurements at $p=3$ bars, $\Delta p=1\%$, and $\Phi=0.59$. The size of the burner nozzle is indicated by two marks on the left side of each image. The PIV repetition rate was 5 Hz. Arrows indicate the local flow direction.

Abel inversion. This algorithm deconvolutes the image data in order to calculate the distribution of the signal intensity on a radial section plane.

•The color scale of all figures showing chemiluminescence or PLIF in this paper is scaled from zero intensity (black) to the maximum intensity of the particular image (white).

•The averaged data in this article that are described as temporally averaged are, in fact, ensemble averaged over a set of single-shot measurements that were performed in succession under constant operating conditions. Since each averaged image is composed out of 200 single shots, which were recorded at a frame rate of about 3–5 Hz, we can assume the ensemble-average to be identical with the temporal average. For clarity reasons the latter expression will be used here.

•All images are background corrected by subtracting an image with background intensity. The background images for chemiluminescence were obtained by using a shutter in front of the camera. The background measurements for the PLIF diagnostic were done by shifting the laser wavelength to a nonresonant value. Each background image is an average of 200 shots. Additionally, a detector sensitivity correction was performed by dividing all raw images by a flatfield image of the particular camera.

4.1 Instantaneous Flow Field. The results of three typical single-shot PIV measurements at $p=3$ bars, $\Delta p=1\%$, and $\Phi=0.59$ are visualized in Fig. 4. The flow field is depicted in the form of a streamline plot, making the absolute planar velocity of the flow (color coded) and its orientation visible. The first feature that catches one's eye is the inflow region. Its shape changes significantly from one laser pulse to the next. This instationary behavior matches well with the occurrence of vortices distributed statistically in the shear layers. The size of these vortices approximately ranges from 1 cm to 5 cm, and they appear more often in the inner shear layer. The generation of these structures is caused by the different velocities of the injected fuel-air mixture and the hot gas between the inflow regions. The latter moves much slower and its mean flow is directed toward the burner mouth, which in turn is caused by the swirl of the flow and the corresponding conical orientation of the inflow. The backflow rate of reaction products into the burner nozzle is remarkably varying, for example, between Figs. 4(*a*) and 4(*c*). Fewer statements can be done about the outer recirculation zone because only about two-thirds of it are visible due to vignetting by elements of the combustor. The vortices that are involved with the outer recirculation are rarely seen in the PIV single shots. Therefore, it may be presumed that they are generally located very close to the walls of the combustor. Further downstream, in the exhaust gas region on the right-hand side of the analyzed field of view, the planar flow velocity sometimes reaches relatively high values. This flow structure mostly appears with an axially stretched shape. Its radial location as well as its intensity strongly varies. At about half of the single-shot measurements it does not exist at all, whereas other images of the same series come up with a global maximum of velocity in this region. First interpretations of the effect assume this acceleration to be correlated with a fast rotating axial tornadolike vortex core that dominates the transportation of the exhaust gas to the exhaust nozzle of the combustor.

4.2 Instantaneous OH* and OH Distributions. The presence of a high OH* concentration is an indicator of an intensive heat release and therefore usually can be found close to the flame front. The aim of imaging the chemiluminescence of electronically excited radicals is to get an impression of the location and the shape of the zone of the primary reactions. Figure 5 shows a typical instantaneous image of the OH* chemiluminescence at $p=4$ bars, $\Delta p=1\%$, and $\Phi=0.55$. The burner exit is marked by a white ellipse. The inhomogeneous intensity distribution gives an impression of a turbulent flame zone, leading to the assumption of a combustion that is affected by events of local ignition and extinction. Measurements using laser Raman scattering will show whether variations in the mixture fraction might also be a source for the observed inhomogeneities. A good portion of the heat release seems to happen in a region located between 5 cm and 8 cm downstream of the burner exit.

These findings are supplemented by the mappings of the laser-induced OH fluorescence. Detectable concentrations of OH radi-

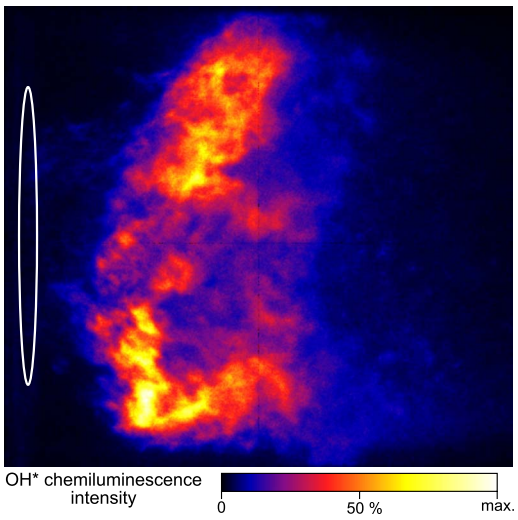


Fig. 5 Instantaneous distribution of the OH* chemiluminescence (line-of-sight) at $p=4$ bars. The white ellipse represents the burner nozzle.

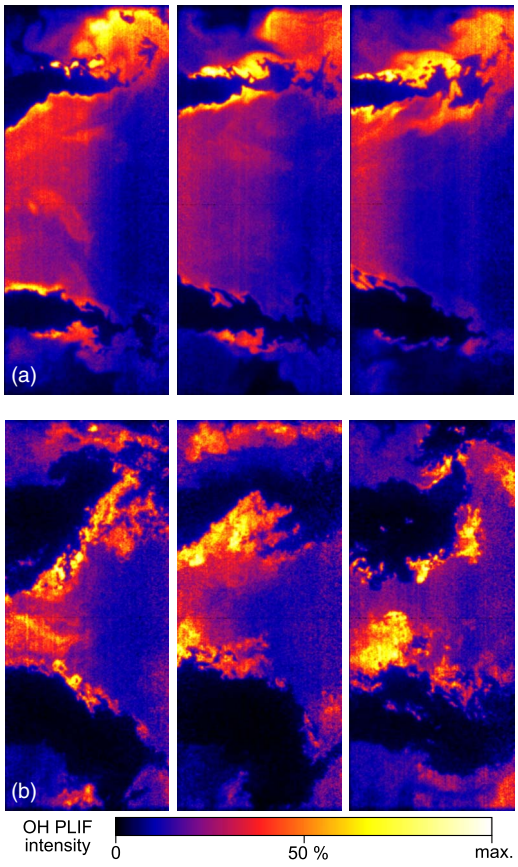


Fig. 6 Single-shot OH PLIF mappings at (*a*) $p=1.5$ bars, $\Delta p=1.1\%$, and $\Phi=0.53$ and (*b*) $p=4$ bars, $\Delta p=3\%$, and $\Phi=0.50$

cals are present not only in the reaction zones but also in the hot exhaust gas when the local temperature is approximately above 1400 K. Since OH is formed in superequilibrium concentrations in the reaction zone, its highest concentrations are found in the flame front. The relaxation toward the equilibrium concentration proceeds via slow three-body reactions and takes a few milliseconds at $p=1$ bar. At a higher pressure it is significantly faster. An extract of a set of single-shot PLIF images is shown in Fig. 6(*a*). The interval between the successive images was about 0.33 s. Only the images from the first laser sheet position are displayed. The pressure drop over the burner was $\Delta p=1.1\%$ at $p=1.5$ bars and $\Phi=0.53$. The black areas near the nozzle exit reflect the region of the inflowing gas of relatively low temperatures. In the inner and outer recirculation zones, medium OH concentrations (blue and red areas, b&w print: dark gray) prevail representing hot gas, probably close to chemical equilibrium. The highest OH concentrations (yellow areas, b&w print: light gray) occur in the shear layers between the inflow and the recirculation zones. These are indicative of the superequilibrium OH and reflect the presence of reaction zones. The "snapshots" clearly show a fragmented flame front.

An example for the influence of the operating parameters on the flame properties can be seen by comparing this set of single shots with the set shown in Fig. 6(*b*). The latter is an extract of a series of PLIF images recorded at $p=4$ bars, $\Delta p=3\%$, and $\Phi=0.50$. The turbulence-related characteristics are much more pronounced as with the operating point from Fig. 6(*a*). The regions that are directly involved in the shear flow are clearly extended, and the superequilibrium OH appears as numerous small spots, each the size of a few pixels. Moreover, excessive OH concentrations occur more frequently in the inner shear layer than in the outer. This indicates a pressure-dependent displacement of the mean distribution of the reaction zones. The distribution of the equilibrium OH in the inner recirculation zone close to the burner changed too. This region has become smaller in diameter. Similar pressure-dependent phenomena were discovered in the temporally averaged OH^* chemiluminescence data, which are treated in Sec. 4.3. A clearer impression of the instationary shape of the reaction zone can be obtained by calculating the gradient of the PLIF intensity. This image processing technique is very sensitive to the rapid growth of the OH concentration between the unburnt mixture and the superequilibrium zone [23]. Such a visualization of a flame at $p=4$, $\Delta p=1\%$, and $\Phi=0.53$ is given in Fig. 7. Here, again, one can see the strongly wrinkled nature of the fragmented flame fronts.

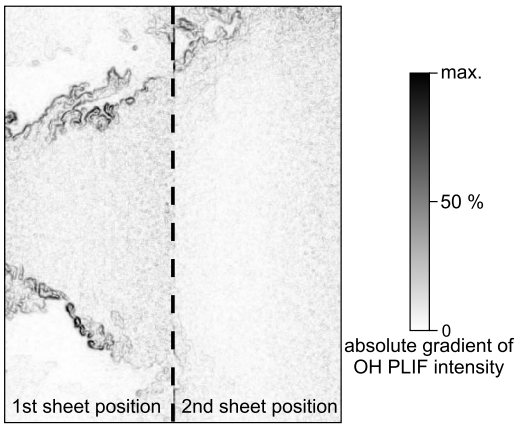


Fig. 7 The absolute gradient of the OH PLIF intensity gives an impression of the instantaneous shape of the flame front. Collage of two measurements with shifted laser sheet positions at different points in time ($p=4$ bars).

4.3 Mean Flow Field. The averaged flow field in Fig. 8 illustrates how the hot exhaust gas in the inner recirculation zone (IRZ) is transported backward until it remixes with the inflow outside or inside of the burner nozzle. A clear distinction between IRZ, outer recirculation zone (ORZ), and inflow region can be obtained by regarding the planar flow direction only (Fig. 9). The sharp contrast in this visualization allows a precise definition of these zones. Remarkably, the complete appearance of the flow field is neither dependent on the pressure inside the combustion chamber nor on the equivalence ratio. The same applies to the variation of the pressure drop over the burner, apart from the

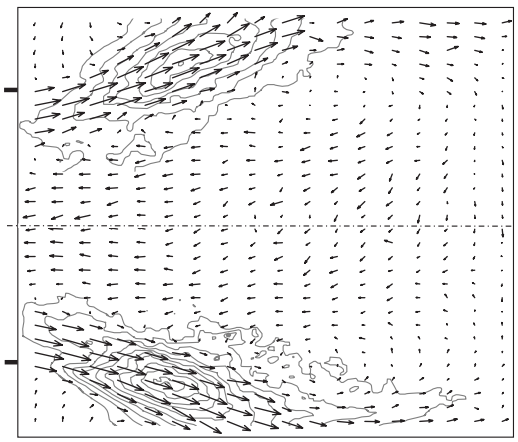


Fig. 8 Mean flow field (sum of correlation) at $p=4$ bars, $\Delta p=1\%$, and $\Phi=0.56$. Background: contour plot of the corresponding mean OH^* density.

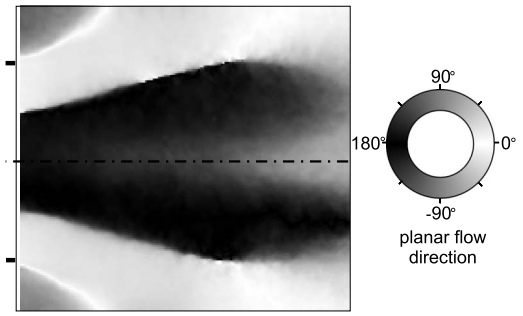


Fig. 9 The inflow region can be clearly distinguished from the recirculation zones by mapping the averaged flow direction

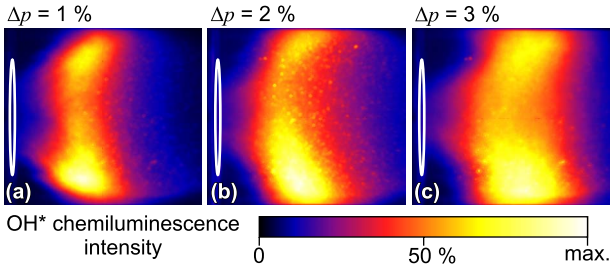


Fig. 12 Line-of-sight mappings of the mean OH* chemiluminescence distribution for different pressure drops over the burner at p=4 bars and $\Phi \approx 0.50$. The burner nozzle is marked with a white ellipse.

mass-flow-dependent scaling of the absolute flow speed. In Fig. 10 this scaling is illustrated by plotting the radial profile of the axial flow velocity component averaged over a range of 7–16 mm downstream of the burner.

Explicit findings concerning the turbulence of the flow are obtained by calculating the standard deviation from the mean flow velocity. An appropriate mapping together with the averaged flow vectors is shown in Fig. 11. As expected, strong variations of the velocity can be found within the shear layers. The highest absolute rms values do not coincide exactly with the boundary layer between the IRZ and the inflow standing out in Fig. 9. The maxima are rather located further outside. Another highly turbulent zone is detected further downstream along the central axis of the combustor, although there is no notable shear flow in the temporal average. This can be explained with the existence of a fluctuating axial vortex core, which already was considered at the interpretation of the PIV single-shot results.

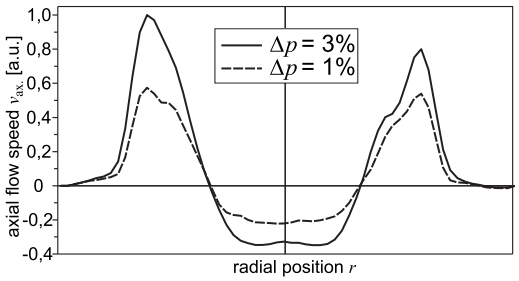


Fig. 10 Radial distribution of the axial component of the flow velocity averaged over an axial range of 7–16 mm downstream of the burner exit

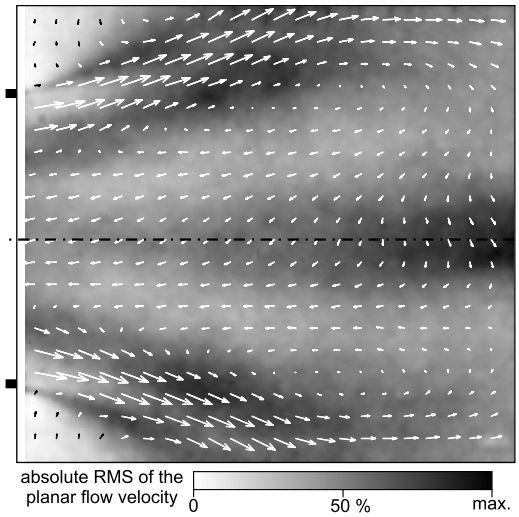


Fig. 11 Mapping of the standard deviation of the flow velocity at p=3 bars, Δp=1%, and Φ=0.59

4.4 Mean OH* Distribution. In Fig. 12 the temporally averaged chemiluminescence images of three operating points with different mass flows at p=4 bars and $\Phi \approx 0.50$ are shown. The asymmetry between the upper and the lower half of the spatial intensity distribution is caused by the inhomogeneous opacity of the optical access. The images reveal the mean flame to stabilize some centimeters downstream of the burner mouth in a ring-shaped region. The position and size of this region are dependent on the mass flow. All of the examined test conditions showed an expansion of the mean flame zone in both the axial and the radial directions when the burner pressure drops, and thus the thermal load is increased. At Δp=3% the region of maximum heat release is located very close to the sidewalls (Fig. 12(c)), which may result in a strong thermal load for the combustion chamber. The increase in the flow velocity of the unburnt fuel-air mixture causes a clearly visible downstream shift of the region with maximum heat release and a narrowing of the mean OH* distribution in the IRZ. The latter effect is particularly visible at high pressures.

To be able to quantitatively analyze the spatial information of the chemiluminescence images, the Abel inversion has to be applied first. In order to bypass difficulties with the vignetted flame periphery, it is possible to perform the deconvolution only in an axial range where the radial extension of the flame is small enough for the application of this image processing method. Here, the pressure influence on the radial distribution of the OH* concentration at a distance of 22 mm from the burner was studied. The plots in Fig. 13 were attained by deconvolving the corresponding chemiluminescence images and applying a 9 $\times$9 pixel^2 smoothing filter. Additionally, an integration over an axial range of $\pm$3 mm was done for further noise reduction. Due to the spatial symmetry, this results in two curves of which the arithmetical mean is plotted in the figure. In this case, the window opacity is negligible since the radial extent of the flame is relatively small. The curves of this parameter study (measured at Δp=2%, Φ=0.53) clearly reflect a pressure-induced displacement of the OH* concentration from the outer to the inner shear layer. The raw data of the 6 bar measurement had a low signal-to-noise ratio and a considerable inconsistency between the two curves from the opposite sides of the symmetry axis. Therefore, the slight deviation of the 6 bar case from the general tendency must not be overrated. However, a distinct bimodal distribution is only existent at pressures below 4 bars and close to the burner. Further downstream, both peaks merge into a distribution with a single maximum, which coincides with the maximum of the radial distribution of the planar flow velocity. Figure 8 shows a mapping of the mean OH* density in the form of a contour plot together with an overlaid vector field of the mean planar flow velocity. It can be seen from the superimposition that the allocation of the reaction zones is directed by the flow field.

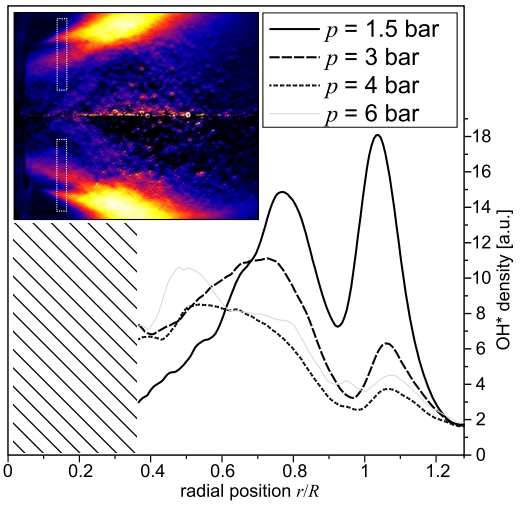


Fig. 13 The radial OH* distribution is influenced by the pressure. The shaded range cannot be analyzed due to deconvolution artifacts. *R* is the burner nozzle radius. Inset: Abel-inverted OH* chemiluminescence. The marked zones were analyzed for the graphs.

5 Summary

An industrial gas turbine burner for lean premixed natural gas flames was equipped with an optical combustion chamber and operated in a high-pressure test rig. The general flame behavior, the structures, and the flow field have been measured using OH* chemiluminescence imaging, planar LIF of OH, and particle image velocimetry. Systematic variations of the mass flow, equivalence ratio, and pressure have been performed. The main goals of this project were a better understanding of the physical and chemical processes dominating the flame behavior and the establishment of an experimental database for the validation of numerical flame simulations.

From the averaged PIV measurement results it was known that the inflow of fresh gases was conically shaped and that the mean flow field had a small ORZ and a large IRZ transporting the hot reaction products back toward the inflow region. Instantaneous PIV measurements exposed the flow in the shear layers to be dominated by small vortices. The PLIF images also showed the unburnt reactants mixing with the combustion products in these regions. By analyzing the distributions of electronically excited OH and superequilibrium concentrations of OH, it could be seen that the ignition happens within the shear layers. Further downstream the primary reactions also took place throughout the complete ring-shaped region with a positive axial flow component. The axial and radial extensions of the averaged flame zone are strongly dependent on the total mass flow. The instantaneous shape of the flame front is disturbed and fragmented by the turbulent flow, especially for conditions with high pressures and high flow velocities. Deconvoluted chemiluminescence images showed a pressure-dependent displacement of the heat release zone from the outer to the inner shear layer. The IRZ of the averaged flow field reaches into the burner nozzle occupying about a quarter of its cross-sectional area. Furthermore, the mean structure of the flow field, e.g., its partition into inflow and recirculation, is independent of the operating conditions. The standard deviation of the flow velocity exposed a relatively strong fluctuating flow within an axially expanded region downstream of the actual flame. This can be interpreted as an effect of a fast rotating exhaust gas swirl.

Altogether, these findings contribute to the understanding of the interaction between the turbulent flow and the chemical reactions in gas turbine combustors.

6 Outlook

In the third measurement phase of this project, laser Raman spectroscopy will be applied to determine the temperature and the concentrations of N_2, O_2, CH_4, H_2, CO_2, CO, and H_2O simultaneously by single-shot measurements with a one-dimensional spatial resolution. For this purpose, a specially designed diagnostic system has already been set up and tested at the German Aerospace Center [24]. Finally, the results of all three measurement phases will be incorporated into a comprehensive database, which will be available for the verification of numerical simulations.

Acknowledgment

The presented work was conducted at the German Aerospace Center (DLR) in cooperation with SIEMENS Industrial Turbomachinery. We gratefully acknowledge the financial support from the State of Baden-Württemberg within the KW21 research program. We would like to thank Richard Noden, Khawar Syed, Lorin Wehr, and Jan Withag for their fruitful collaboration in this project.

References

[1] Löfström, C., Engström, J., Richter, M., Kaminski, C., Johansson, P., Nyholm, K., Nygren, J., and Aldén, M., 2000, "Feasibility Studies and Application of Laser/Optical Diagnostics for Characterisation of a Practical Low-Emission Gas Turbine Combustor," ASME Paper No. 2000-GT-0124.

[2] Gord, J., Brown, M., and Meyer, T., 2002, "Optical Diagnostics for Characterizing Advanced Combustors and Pulsed-Detonation Engines," *Proceedings of the 22nd AIAA Aerodynamic Measurement Technology and Ground Testing Conference*, St. Louis, MO, June 24–26, Paper No. AIAA2002-3030.

[3] Zarzalis, N., Ripplinger, T., Hohmann, S., Hettel, M., Merkle, K., Leuckel, W., Klose, G., Koch, R., Wittig, S., Carl, M., Berendt, T., Hassa, C., Meier, U., Lückerath, R., and Stricker, W., 2002, "Low-NO_2 Combustor Development Pursued With the Scope of the Engine 3E German National Research Program in a Cooperative Effort Among Engine Manufacturer MTU, University of Karlsruhe and DLR German Aerospace Center," Aerosp. Sci. Technol., **6**, pp. 531–544.

[4] Meier, U., Wolff-Gaßmann, D., and Sticker, W., 2000, "LIF Imaging and 2D Temperature Mapping in a Model Combustor at Elevated Pressure," Aerosp. Sci. Technol., **4**, pp. 403–414.

[5] Hicks, Y., Locke, R., Anderson, R., Zaller, M., and Schock, H., 1997, "Imaging Fluorescent Combustion Species in Gas Turbine Flame Tubes: On Complexities in Real Systems," *Proceedings of the 33rd Joint Propulsion Conference*, Seattle, WA, July 6–9, Paper No. AIAA-97-2837.

[6] Frank, J., Miller, M., and Allen, M., 1999, "Imaging of Laser-Induced Fluorescence in a High-Pressure Combustor," *Proceedings of the 37th AIAA Aerospace Science Meeting and Exhibition*, Reno, NV, Jan. 11–14, Paper No. AIAA 99-0773.

[7] Griebel, P., Siewert, P., and Jansohn, P., 2007, "Flame Characteristics of Turbulent Lean Premixed Methane/Air Flames at High Pressure: Turbulent Flame Speed and Flame Brush Thickness," Proc. Combust. Inst., **31**, pp. 3083–3090.

[8] Lee, S.-Y., Seo, S., Broda, J., Pal, S., and Santoro, R., 2000, "An Experimental Estimation of Mean Reaction Rate and Flame Structure During Combustion Instability in a Lean Premixed Gas Turbine Combustor," Proc. Combust. Inst., **28**, pp. 775–782.

[9] Janus, B., Dreizler, A., and Janicka, J., 2007, "Experiments on Swirl-Stabilized Non-Premixed Natural Gas Flames in a Model Gasturbine Combustor," Proc. Combust. Inst., **31**, pp. 3091–3098.

[10] Cooper, C., and Laurendau, N., 2000, "Comparison of Laser-Induced and Planar Laser-Induced Fluorescence Measurements of Nitiric Oxide in a High-Pressure, Swirl-Stabilized, Spray Flame," Appl. Phys. B: Lasers Opt., **79**, pp. 903–910.

[11] Fink, R., Hupfer, A., and Rist, D., 2002. "Non Intrusive Measurements of a LPP Combustor Under Elevated Pressure Conditions," ASME Paper No. GT-2002-30078.

[12] Carl, M., Behrendt, T., Fleing, C., Frodermann, M., Heinze, J., Hassa, C., Meier, U., Wolff-Gaßmann, D., Hohmann, S., and Zarzalis, N., 2001, "Experimental and Numerical Investigation of a Planar Combustor Sector at Realistic Operating Conditions," ASME J. Eng. Gas Turbines Power, **123**, pp. 810–816.

[13] Soika, A., Dinkelacker, F., and Leipertz, A., 2003, "Pressure Influence on the Flame Front Curvature of Turbulent Premixed Flames: Comparison Between Experiment and Theory," Combust. Flame, **132**, pp. 451–462.

[14] Janus, B., Dreizler, A., and Janicka, J., 2005, "Experimental Study on Stabilization of Lifted Swirl Flames in a Model GT Combustor," Flow, Turbul. Combust., **75**, pp. 293–315.

[15] Geigle, K.-P., Meier, W., Aigner, M., Willert, C., Jarius, M., Schmitt, P., and Schürmans, B., 2006, "Phase Resolved Laser Diagnostic Measurements of a Downscaled, Fuel Staged Gas Turbine Combustor at Elevated Pressure and Comparison With LES Predictions," ASME Paper No. GT2006-90875.

[16] Hassa, C., Willert, C., Fischer, M., Stockhausen, G., Röhle, I., Meier, W.,

Wehr, L., and Kutne, P., 2006, "Nonintrusive Flowfield, Temperature and Species Measurements on a Generic Aeroengine Combustor at Elevated Pressures," ASME Paper No. GT2006-90213.
[17] Willert, C., Hassa, C., Stockhausen, G., Jarius, M., Voges, M., and Klinner, J., 2006, "Combines PIV and DGV Applied to a Pressurized Gas Turbine Combustor Facility," Meas. Sci. Technol., **17**, pp. 1670–1679.
[18] Barlow, R., 2007, Library of the TNF workshop, http://public.ca.sandia.gov/TNF
[19] Weigand, P., Meier, W., Duan, X., Stricker, W., and Aigner, M., 2006, "Investigations of Swirl Flames in a Gas Turbine Model Combustor: I. Flow Field, Structures, Temperature, and Species Distribution," Combust. Flame, **144**, pp. 205–224.
[20] Meier, W., Weigand, P., Duan, X., and Giezendaner-Thoben, R., 2007, "Detailed Characterization of the Dynamics of Thermoacoustic Pulsations in a Lean Premixed Swirl Flame," Combust. Flame, **150**, pp. 2–26.
[21] Ng, W., Syed, K., and Zhang, Y., 2004, "The Study of Flame Dynamics and Structures in an Industrial-Scale Gas Turbine Combustor Using Digital Data Processing and Computer Vision Techniques," Meas. Sci. Technol., **15**, pp. 2303–2309.
[22] Meinhart, C., Wereley, S., and Santiago, J., 2000, "A PIV Algorithm for Estimating Time-Averaged Velocity Fields," ASME J. Fluids Eng., **122**, pp. 285–289.
[23] Sadanandan, R., Stöhr, M., and Meier, W., 2008, "Simultaneous OH-PLIF and PIV Measurements in a Gas Turbine Model Combustor," Appl. Phys. B: Lasers Opt., **90**, pp. 609–618.
[24] Wehr, L., Meier, W., Kutne, P., and Hassa, C., 2007, "Single-Pulse 1D Laser Raman Scattering Applied in a Gas Turbine Model Combustor at Elevated Pressure," Proc. Combust. Inst., **31**, pp. 3099–3106.

A Simple Mathematical Approach to Data Reconciliation in a Single-Shaft Combined Cycle System

S. Can Gülen
e-mail: can.gulen@ps.ge.com

Raub W. Smith
e-mail: raub.smith@ps.ge.com

GE Energy,
1 River Road,
Schenectady, NY 12345

The ultimate proof of the soundness and viability of a novel technology is a full-scale demonstration test in which actual components are run successfully over the entire operating envelope. Consequently, the collection of accurate and meaningful test data is of utmost importance to the success of the test. An analysis of such data will validate the original design concepts and will lead to paths of further improvement for the next generations thereof. Statistical fundamentals to determine the accuracy and precision of measured data are amply documented and readily available in well-established standards. The yardstick that should be used for the "meaningfulness" of the measured test data is the satisfaction of the fundamental laws of conservation. While it is known that the "true" values of the sensor data when inserted into the governing equations for the tested component will result in perfect balances, "actual" measured values will always result in "imbalances." Therefore, reconciliation of the individual measurements with the governing conservation equations is a must prior to the actual analysis of the data. Reconciliation in this context is an estimation of the true values of the sensor data from the actual sensor data by using statistical concepts. This paper describes the development of a data reconciliation concept that is universally applicable to any process or power plant system where sensor data are used. The usefulness and power of the technique are demonstrated by its application to a single-shaft combined cycle with both gas turbine and steam turbine driving a common generator. In the absence of a reliable and accurate measuring system to individually determine gas and steam turbine shaft outputs, data reconciliation is vital to an accurate analysis of the data. [DOI: 10.1115/1.2978994]

Introduction

Data reconciliation ensures that a set of measurements is thermodynamically consistent. Such a set of measurements results in a closed heat balance; i.e., when measured variables are inserted into the steady-state, steady-flow mass and energy conservation equations for a control volume (CV), the net total of mass and energy streams each is exactly equal to zero. If this is not the case, the heat/mass balance is not closed and the imbalance is called a heat/mass balance error. Furthermore, if the error is greater in magnitude than the value obtained from measurement errors via the principle of error propagation, the problem is especially severe. In such cases, the measurement system should be investigated for defects leading to the implied violation of the fundamental laws of nature.

In general, the heat balance error indicates the presence of one or more of the following conditions:

1. One or more of the measured quantities have errors; i.e., they deviate significantly from their "true" values.
2. Test conditions do not represent a truly steady-state, steady-flow operating condition, and unaccounted transient effects are in play.
3. The components are treated as adiabatic and isolated, whereas significant heat and/or mass leaks are present.

Measurement errors due to calibration, data acquisition, and reduction errors are unavoidable even in the most rigorously designed tests with high-precision instrumentation. While it is not realistic to expect that a real complex system such as a combined cycle power plant will attain a "perfect" steady state per textbook definition, performance tests conducted under the well-established stability and duration guidelines should not suffer from transient effects. Similarly, cycle isolation guidelines in the performance test codes will ensure that leakage flows that enter or exit the test cycle or bypass components are minimized if not entirely eliminated. Obviously, tests or online (continuous) monitoring systems that rely on less accurate station instrumentation with calibration drifts and missing or broken sensors will suffer even more from stability and cycle isolation defects. Regardless of the type of the test, before using the measured quantities, effort has to be made to close the heat and balances, i.e., reconcile the test data.

Data reconciliation has been widely accepted and commonly used in the chemical process industry. There is such a large body of literature available dealing with the subject that even a limited survey is beyond the scope of the present work. A detailed description of the underlying concepts, application examples, and a relevant bibliography can be found in Romagnoli and Sanchez [1]. Statistical and mathematical principles and methodology that form the foundation of data reconciliation are also applied to aircraft gas turbine (GT) engine health monitoring and diagnostics. Linear and nonlinear versions of the gas path analysis (GPA) have been introduced more than three decades ago [2,3] and later improved upon significantly [4]. The TEMPER algorithm [5] that uses a weighted-least-squares (WLS) principle to determine the allocation of sensor deviations to hardware issues and measurement errors is such an example.

The application of data reconciliation in the power generation industry is not as widespread as in the chemical process industry.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 15, 2006; final manuscript received July 15, 2008; published online December 18, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2006: Land, Sea and Air (GT2006), Barcelona, Spain, May 8–11, 2006.

A survey of published work shows that as in myriad process industry variants and linear or nonlinear incarnations of GPA, the problem statement is almost always the same: the solution to a WLS problem where the weights are determined by measurement uncertainties. The solution methodology is either model based (utilizing a commercially available heat balance software) [6] or balance equation based [7]. In most cases, the WLS technique is applied to the simultaneous solution of the combined problem of sensor validation and model parameter matching.

The fundamental principle underlying GPA and all similar WLS algorithms can be traced back to perhaps the most famous sensor validation principle: the *Kalman* filter [8,9]. A Kalman filter combines all available data and data acquisition system information, past and present, and generates an estimate of the measured variable such that the resulting error is statistically a minimum . The method that is developed in this paper can also be reduced to the basic Kalman filter expression. A simple formula is developed starting from a basic engineering intuition and by carrying it to its logical conclusion using straightforward differential calculus. It is then demonstrated that the formula can be applied to any power plant test system without resorting to any "brute force" method to solve a large complicated system of equations consisting of covariance matrices, inverse matrices, and determinants. It is shown that no complex numerical solution algorithms or detailed plant model runs are needed for large-scale number crunching. In that sense, a clear separation is made between the reconciliation of the measured test data and the "tuning" of the particular hardware model design parameters so that the model prediction matches the measured performance.

Theory of Data Reconciliation

Each major component of the system that is being tested constitutes a control volume, e.g., GT, steam turbine (ST), and heat recovery steam generator (HRSG). For each CV, the three fundamental conservation equations for mass, energy (or heat), and momentum (angular and linear) can be written in the following generic form:

$$\varepsilon = f(x_i), \quad i = 1,2, \ldots, n \tag{1}$$

Since power plant performance test measurements rarely, if ever, provide data that are amenable to a momentum balance check, the present discussion is limited to energy and mass balances. However, the principles are equally applicable if data pertinent to the evaluation of momentum conservation are available. In Eq. (1) ε is the nonzero heat or mass balance error that results when the measured values, x_i, of the material stream data (pressure, temperature, and mass flow rate) and/or energy stream data (shaft work and heat transfer rates) that cross the boundaries of the CV are substituted into the equation.

While ε in Eq. (1) should in theory yield exactly zero, this will not be the case in practice, except by pure coincidence, due to a combination of reasons listed in the preceding section. Ideally, what one can hope for is that Eq. (1) will yield an error whose absolute value, $|\varepsilon|$, is no greater than its uncertainty, U_ε, obtained from the error propagation principle,

$$U_\varepsilon = \left\{ \sum_{i=1}^{n} \left(\frac{\partial f(x_i)}{\partial x_i} \cdot U_i \right)^2 \right\}^{1/2} = \left\{ \sum_{i=1}^{n} (S_i \cdot U_i)^2 \right\}^{1/2} \tag{2}$$

where U_ε is the uncertainty of the calculated heat/mass balance error, U_i is the measurement uncertainty of x_i, and n is the number of independent measurements. The sensitivity of ε to a change in x_i, S_i, is obtained from Eq. (1) via respective partial derivatives.

The derivations and analysis from this point on are based on the definitions of "measurement" and "uncertainty" outlined in great detail in ASME PTC 19.1 [10]. As such, x_i is the average value of a set of individual measurements (observations and readings) from one or more sensors measuring the same quantity over a certain period of time. Uncertainty, U_i, is a figure of merit that represents the total error associated with the measurement of x_i and is equal to the half-width of a band within which the true value of the measured parameter, x_i^*, is expected to lie with a certain confidence level, i.e., $x_i \pm U_i$. Calculations of x_i and U_i are outside the scope of the current analysis. They are obtained via the methodology outlined in ASME PTC 19.1 such that U_i represents the uncertainty calculated using the root-sum-squares (RSS) model with a 95% coverage of x_i.

If the absolute value of the calculated error is greater than U_ε, a correction to the measured variables is introduced so that $|\varepsilon| < U_\varepsilon$; i.e., the measured data are *reconciled*. In fact, in most practical cases one has to reconcile the measured data even when $|\varepsilon| < U_\varepsilon$ to begin with. This is so because typically inaccurate measurements such as large gas or liquid flow rates have large errors that amplify U_ε. In such cases, the large mass and/or energy imbalance, even though it is within the measurement error limits, may render the unreconciled data unusable in applications such as control logic software, system optimization, and model calibration.

Consider the Taylor expansion of ε around $\varepsilon = 0$,

$$\varepsilon = \sum_{i=1}^{n} \left\{ \frac{\partial f(x_i)}{\partial x_i} \cdot \Delta x_i + \frac{\partial^2 f(x_i)}{\partial x_i^2} \cdot (\Delta x_i)^2 + \cdots \right\} \tag{3}$$

where Δx denotes the difference between the measured value of the parameter x and its true value, x^*, that would exactly yield

$$\varepsilon = f(x_i^*) = 0$$

The ultimate goal in reconciliation is to estimate the true value of x and substitute it for the measured value. To a very good approximation, second and higher order terms in Eq. (3) can be neglected, thus yielding

$$\varepsilon = \sum_{i=1}^{n} \left\{ \frac{\partial f(x_i)}{\partial x_i} \cdot \Delta x_i \right\} = \sum_{i=1}^{n} \{ S_i \cdot \Delta x_i \} \tag{4}$$

Each term on the right-hand side (RHS) of Eq. (4) represents the contribution of an error in the respective measured variable to the total error. The approximation presupposes one or both of the two following conditions:

1. The error term, Δx_i, is small (i.e., the sensor quality is high).
2. The functional relationship between ε and x_i, $f(x_i)$, does not have a strong curvature in the region of interest (i.e., second and higher order derivatives are 0).

Since the true values of the measured parameters are not known, all that can be said a priori is that each individual contribution term represents a fraction, w_i, of the total error such that

$$w_i \cdot \varepsilon = S_i \cdot \Delta x_i = S_i \cdot (x_i - x_i^*) \tag{5}$$

$$\sum_{i=1}^{n} w_i = 1 \tag{6}$$

Equation (5) can be recast so that for the true value of a particular measurement, x_i, it becomes

$$x_i^* = x_i - \frac{w_i \cdot \varepsilon}{S_i} = x_i - \frac{w_i \cdot f(x_i)}{S_i} \tag{7}$$

Data reconciliation is essentially a problem of estimating the fractions, w_i, in Eq. (7) that will yield the estimated true value, x_i^*, such that when all x_i^* are substituted into Eq. (1) the resulting mass or heat balance error will be zero. Since Eq. (7) contains "measured" values of stream data x_i with an uncertainty of U_i, the estimate itself will have an uncertainty. Based on the principle of error propagation, the uncertainty of x_i^* can be calculated as

$$(U_i^*)^2 = \left(\frac{\partial x_i^*}{\partial x_i}\cdot U_i\right)^2 + \sum_{j\neq i}\left(\frac{\partial x_i^*}{\partial x_j}\cdot U_j\right)^2 \tag{8}$$

The sensitivity coefficients for x_i^* that appear in Eq. (8) are obtained via partial derivatives of Eq. (7) as

$$\frac{\partial x_i^*}{\partial x_i} = 1 - w_i\cdot\left[1 + f(x_i)\cdot\frac{\partial S_i^{-1}}{\partial x_i}\right] \tag{9a}$$

$$\frac{\partial x_i^*}{\partial x_j} = -w_i\cdot\left[\frac{\partial f(x_i)}{\partial x_j}\cdot\frac{1}{S_i} + f(x_i)\cdot\frac{\partial S_i^{-1}}{\partial x_j}\right] \tag{9b}$$

The "best" estimate of x_i^* will have an uncertainty, U_i^*, as given by Eq. (8), that will be a minimum. This estimated best value is referred to as the reconciled value of the measured parameter. Thus, the task of data reconciliation is reduced to the determination of the set of fractions, w_i^*, for which U_i^* is a *minimum*. Taking the first derivative of U_i^* with respect to w_i in Eq. (8) and setting it to 0 result in

$$0 = -U_i^2\cdot[1 - w_i\cdot(1+\phi)]\cdot(1+\phi) + w_i\cdot\sum_{j\neq i}U_j^2\cdot\left[\frac{\partial f(x_i)}{\partial x_j}\cdot\frac{1}{S_i} + \phi\cdot\frac{\partial x_i}{\partial x_j}\right]^2$$

$$\phi = f(x_i)\cdot\frac{\partial S_i^{-1}}{\partial x_i} \tag{10}$$

The linearity approximation that is used to arrive at Eq. (4) suggests that

$$\frac{\partial S_i^{-1}}{\partial x_i} = -\frac{1}{S_i^2}\cdot\frac{\partial S_i}{\partial x_i} = -\frac{1}{S_i^2}\cdot\frac{\partial^2 f(x_i)}{\partial x_i^2} \approx 0$$

Thus, Eq. (10) becomes

$$0 = -U_i^2\cdot(1-w_i) + w_i\cdot\sum_{j\neq i}U_j^2\cdot\left(\frac{S_j}{S_i}\right)^2 \tag{11}$$

Solving Eq. (11) for w_i, the result is

$$w_i = \frac{(S_i\cdot U_i)^2}{\sum_{i=1}^{n}(S_i\cdot U_i)^2} = \frac{(S_i\cdot U_i)^2}{U_\varepsilon^2} \tag{12}$$

Note that the denominator of Eq. (12) is equal to the RSS uncertainty of the heat/mass balance error given by Eq. (1). Equation (12) shows that the optimal fraction that represents the contribution of each individual measurement to the total heat balance error is equal to the contribution of the uncertainty of that measurement to the overall heat balance error's RSS uncertainty. Substituting Eq. (12) into Eq. (7), the following fundamental data reconciliation equation is obtained:

$$x_i^* = x_i - \frac{(S_i\cdot U_i)^2}{U_\varepsilon^2}\cdot\frac{f(x_i)}{S_i} \tag{13}$$

The inherent analogy of Eq. (13) to the Kalman filter principle can be easily recognized by considering a linear system with $n=2$ and solving for x_1^* and x_2^* [9]. The associated uncertainty for x_i^* defined by Eq. (13) is obtained by substituting Eq. (12) into Eq. (8), which after some algebraic manipulation yields

$$(U_i^*)^2 = (1-w_i)\cdot U_i^2 \tag{14}$$

Equation (13) gives the reconciled measurement of stream data, x_i^*, that is the best estimate of the true value of the actual measurement, x_i, in the sense that

1. x^* has an uncertainty, U^*, that is less than the original measurement uncertainty, U, of x.
2. U^* is the minimum possible uncertainty.
3. x^* satisfies the energy and mass balances.

Application to Control Volumes. For any generic control volume with material (inlet, outlet, and leakage flows) and energy (generator output and losses) streams that is representative of most subsystems encountered in power plant performance tests, the conservation equations can be written as follows:

$$\sum_i(\pm)\cdot\dot{m}_i = \varepsilon_{\text{mass}} \tag{15}$$

$$\sum_i(\pm)\cdot\dot{m}_i\cdot h_i + \sum_k(\pm)\cdot\dot{Q}_k - \sum_l(\mp)\cdot\dot{W}_l = \varepsilon_{\text{heat}} \tag{16}$$

The set is completed by an equation of state that enables us to calculate the enthalpy h_i of the material stream i as a function of the measured pressure and temperature, p_i and T_i, respectively. ASME 1967 steam property tables and JANAF properties [11] can be used for water/steam and air/gas properties, respectively.

The validity of the approximation leading to Eq. (11) and the applicability of Eqs. (12)–(14) to the general nonlinear system described by Eq. (16) can be easily confirmed. Consider the following relationships for typical control volume (CC) steam turbine throttle conditions:

$$\frac{\partial^2\varepsilon}{\partial T^2} = \dot{m}\frac{\partial^2 h}{\partial T^2} = \dot{m}\frac{\partial c_p}{\partial T} = \dot{m}\cdot O(10^{-4})\frac{\text{kW}}{^\circ\text{K}^2}$$

$$\frac{\partial^2\varepsilon}{\partial p^2} = \dot{m}\frac{\partial^2 h}{\partial p^2} = \dot{m}\cdot O(10^{-6})\frac{\text{kW}}{\text{bara}^2}$$

Even for typical steam turbine throttle flows of 45–90 kg/s (100–200 pps), the largest of these terms is of the order of 0.01. Furthermore, for even a very large error in temperature, e.g., 6 °C (10 °F), the error by the omission of the larger order terms in Eq. (3) would be approximately 1 kW. Thus the validity of Eq. (4), which is the starting premise of the methodology leading to Eqs. (12)–(14), is also ascertained.

Once a test system or subsystem CV is identified and a complete set of measurements for all energy and flow streams into and out of the said CV is available, the solution of the data reconciliation problem is straightforward:

1. Conservation equations similar to Eqs. (15) and (16) are written for each of them.
2. The measurement uncertainty of each measurement is determined per ASME PTC 19.1.
3. Mass flow rate measurements are reconciled by applying Eqs. (12) and (13) to each of them.
4. Reconciled mass flow rate measurements, $\dot{m}_i^*$, are substituted into Eq. (16).
5. Sensitivities S_i for each pressure and temperature measurement, p_i and T_i, are determined from Eq. (16) by evaluating the partial derivatives, $\partial\varepsilon_{\text{heat}}/\partial p_i$ and $\partial\varepsilon_{\text{heat}}/\partial T_i$.
6. Heat balance error allocation weights for each pressure and temperature measurement, p_i and T_i, are determined from Eq. (12).
7. Reconciled pressure and temperature measurements, p_i^* and T_i^*, are determined by applying Eq. (13) to each of them.
8. No further correction is applied to reconciled mass flow rate measurements, $\dot{m}_i^*$.

The calculation of sensitivity coefficients, S_i, in Eq. (12) requires the evaluation of partial derivatives of ε with respect to individual measured parameters using Eq. (16). While this is straightforward, in principle, it can become tedious to set up and run due to the complex functional form of the state equations that will typically require a numerical evaluation of derivatives. An example of the calculation is provided in the Appendix.

Table 1 Typical CC performance test measurement errors for steam and water FPT

	Precision calibrated	Station
Pressure (full gauge scale)	±0.1%	±0.5%
Differential pressure	±0.1%	±0.5%
Temperature	±1°F	±7°F
Water flows (ASME throat tap nozzles)	±0.8% to ±1.25%	
Steam flows (8–20 in. venturis)	±2.00% to ±2.50%	

Single-Shaft CC System Test Data Analysis

The data reconciliation methodology developed in the preceding section is suitable for application to a single CV (e.g., GT or ST) or a collection of interconnected CVs such as a CC power plant. The method has been actually deployed during the characterization test of the GE's H-System™ in Baglan Bay, Wales in 2003. The salient features of the methodology are demonstrated below using a generic single-shaft CC system as an example.

During the typical base-load performance test, data are acquired during a several-hour period to assess plant thermal performance. Perfect plant stability is typically impossible to achieve in any given test run. A good example is the variation in ambient air temperature during a test duration of a couple of hours. A cycle isolation check, e.g., per guidelines outlined in ASME performance test code PTC 46 [12], prior to the test identifies significant system leakage flows that are quantified and, if possible, eliminated. It should be pointed out that the data reconciliation process described herein can confirm and account for all major leakage flows that are detected during the cycle isolation during the analysis of the data.

Measured data include all of the usual fluid state points for air, water, steam, and fuel, as well as energy streams that cross the system boundaries (e.g., shaft work and losses.) Ideally, all steam cycle water and steam flows, pressures, and temperatures (FPT) are measured using precision-calibrated test instrumentation. In many instances, precision measurements are supplemented by station instrumentation although as few parameters as possible should be determined solely from them. Representative uncertainties of individual FPT sensors are shown in Table 1.

In cases where there are multiple measurements of the same parameter, they are statistically combined via the RSS method. The responsible test team determines an overall measurement uncertainty for each measured parameter following the guidelines in relevant test codes, e.g., ASME PTC 6R-1985 [13]. In addition to stream FPT, measurements of several other key items are also required to establish complete energy balances for subsystem control volumes (see Table 2).

Single-shaft configuration presents a unique challenge of distributing the generator output measurement between the gas and steam turbines. Direct measurement techniques that can accomplish this goal are proposed, e.g., optical torque sensors [14], but they have not been proven in the field. The determination of the turbine shaft outputs from their respective heat balances is the standard method. ST heat balance presents a difficulty because it is not possible to accurately determine the enthalpy of the wet steam at the low-pressure (LP) section exhaust. The ST exhaust energy can be calculated according to the procedure described by Cotton [15]. Calculations utilize the design values of the LP efficiency and the design exhaust loss curve for the last stage bucket (LSB). In subsequent analyses, LP efficiency and exhaust loss curves should be updated if or when detailed accurate analysis results become available. The data can then be used to obtain an indirect measurement of the LP efficiency to be used in the ST heat balance calculations. The uncertainty of the LP efficiency from a detailed rigorous analysis utilizing such data can be accurate to ±1%. If not, ±2% is conceivable and practically dominates the ST heat balance by itself.

Since actual data obtained during actual tests are proprietary and cannot be disclosed, "pseudodata" generated from a commercial heat balance modeling software will be used for demonstration [17]. The numbers generated by the model are based on the default performance of a GE Frame 9 heavy-duty industrial gas turbine (9351FA) from the software's GT library. The steam system is also based on the heat balance software's default bottoming cycle design assumptions. The system is qualitatively representative of a modern single-shaft CC plant both in terms of the general configuration and the order of magnitude of measured quantities.

The model is run to generate "expected" performance data. Thereafter some random error is introduced to the models to generate "pseudotest" data shown in Fig. 1. The second set represents what is encountered in an actual performance test run. The analysis is based on the realistic measurement uncertainties per Tables 1 and 2 and represents the salient features of the data reconciliation process accurately.

The step-by-step CC power plant performance test data reconciliation process is as follows:

1. Once the test is completed, raw instrument data are analyzed and reduced into final measured data with individual measurement uncertainties.
2. The plant performance model is run at the particular test site ambient and loading conditions to generate expected parameter data.
3. The measured and expected data sets are compared to identify questionable and/or missing measurements whose values deviate significantly from their expected counterparts.
4. Once the measured parameter set passes the initial inspection, the data reconciliation process is started.

The validation of each steady-state test point data against the pretest model run at the respective test site ambient and loading conditions is essential. An item-by-item comparison highlights any large disparity between measured and expected data that may

Table 2 Typical CC performance test measurement errors for key parameters

	Method	Error
Generator output	Precision watt-hour meter meeting ASME PTC 6 code requirement	±0.25%
Fuel flow	Flange tap orifices with an upstream flow straightener	±0.50%
GT inlet air flow	Multiple flow probes (calibrated to factory rig tests) placed in the compressor inlet and measurements of compressor inlet (bell mouth) total temperature, pressure, and static pressure in a compressible flow formula [16]	±1.5–2.5%
GT exhaust temperature	GT exhaust thermocouple rakes	±7–10°F
HRSG stack temperature	Stack thermocouple rakes	±10°F
Generator loss	Coolant flow and temperature increase across the hydrogen coolers and the armature winding cooler (more than 90% of the total losses)	±5%
Mechanical loss	Measurements of the bearing lubricating oil flows and inlet and drain temperatures	±10%
GT enclosure heat loss	Measurements of the enclosure inlet and outlet temperatures and the constant volumetric flow rate of the ventilation fan	±10%

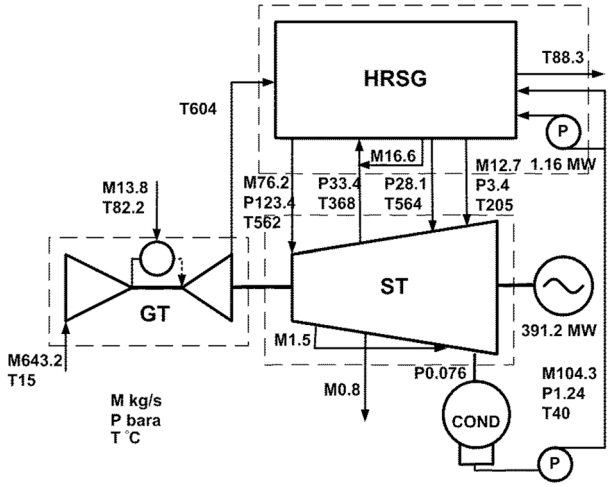


Fig. 1 Sample heat and mass balance data for a single-shaft CC system

indicate faulty instrumentation and/or human error in data reduction and transmittal. Especially in large-scale testing programs, readings that are available at the start of testing will be lost to instrument failure as testing progresses. Such "outliers" and missing items are replaced by their expected values and subjected to data reconciliation in order to prevent the possibility of "bad data driving out the good data." The validated data are then subjected to the data reconciliation process described herein, which enforces overall and subsystem heat and mass balances for each control volume (GT, HRSG, ST, mixing tees, etc.)

Individual heat and mass balance tables for each subsystem are interconnected to represent the full plant level system heat and mass balances. The reading of the data files into preset tables, the transfer of data from those tables into subsystem heat and mass balance tables, and the reconciliation of measured data via Eqs. (12)–(14) can easily be automated in an Excel spreadsheet.

The modular structure of the process enables individual examinations of critical subsystems. This flexibility is extremely useful in detecting unexpected valve leaks, heat losses, and myriad small mishaps that are unavoidable in such a large and complex system. The data reconciliation process comprises the following semiautomated steps:

1. Calculation of the GT shaft output from the GT heat balance
2. Closure of mass flow rate imbalance via major water-steam circuit mass balances
3. Calculation of the ST shaft output from the ST heat balance
4. Evaluation of the single-shaft energy balance and reconciliation of GT and ST shaft outputs
5. Closure of the single-shaft energy imbalance

The heat balance around the GT control volume in Fig. 1 indicates that the GT shaft output is 259,771 kW (see the Appendix for details.). Similarly, using the heat balance around the ST control volume in Fig. 1, the ST shaft output is 139,929 kW. Combining the calculated shaft outputs, one can easily obtain the single-shaft CC heat balance, as shown in Table 3.

Table 3 Shaft energy balance (kW)

	As tested	U (%)	Expected	Delta (%)
Generator output	391,200	0.26%	391,588	–0.10%
Generator losses	4000	5%	4111	–2.70%
Mechanical losses	550	5%	529	3.95%
ST gross output	139,929	1.77%	139,807	0.09%
ST mechanical losses	515	10%	464	11.06%
GT shaft output	259,771	3.04%	256,885	1.12%
Error	–3435		0	

The single-shaft CC energy balance in Table 3 indicates that the measured generator output is approximately 3.5 MW lower than that indicated by GT and ST shaft outputs from their respective heat balances. Using the reconciliation method per Eqs. (12)–(14), the heat balance error is closed, as shown in Table 4.

Table 4 highlights the following salient facts:

1. GT shaft output error ($\pm 3.04\%$) accounts for almost 90% of the shaft error. In fact, in a carefully conducted performance test, it is quite acceptable to use the shaft output from the ST heat balance (with measured LP efficiency) with no correction.
2. The most likely GT output is 256.7 MW ± 2.5 MW ($\pm 0.99\%$), indicating a GT heat balance error of approximately 3.1 MW.

The remaining task is to identify and reconcile individual flow, pressure, and/or temperature measurements that are consistent with the shaft energy balance. This requires a close look at the GT heat balance in conjunction with the HRSG heat balance. Note that the HRSG heat balance is based on a heat utilization factor (HUF) of 99%; i.e., 1% of the heat recovered from the GT exhaust gas is lost due to radiation, convection, and conduction. Such losses can be calculated using guidelines such as those outlined in ASME PTC 4.4 [18]. However, the effort and cost involved in such an endeavor is not justified by the resultant contribution to the overall test uncertainty, and the manufacturer-supplied guarantee number is used in the actual test as a "reconciliation knob" with $\pm 1\%$ uncertainty. Similarly, the heat loss in the transition duct is represented by a temperature loss with a $\pm 1°$F uncertainty. Both parameters are reconciled just like all the other measured parameters. The heat balance for the HRSG control volume in Fig. 1, which uses steam and mass flow rates from a separate mass balance reconciliation in Step 2, indicates that the heat recovered from the GT exhaust is approximately 1.4 MW (1328 Btu/s) higher than the heat consumption indicated by the steam generation. The imbalance is quite significant at approximately 0.35% of the GT exhaust energy.

While many individual measurements alone or in combination

Table 4 Shaft energy balance closure (kW)

	S	U	$(SU)^2$	w	Reconciled	U^*
Generator output	+1	1017	1.03×10^{06}	1.49%	391,251	1010
Generator losses	+1	200	4.00×10^{04}	0.06%	4002	200
Mechanical losses	+1	28	7.56×10^{02}	0.00%	550	27
ST wheel output	–1	2475	6.13×10^{06}	8.82%	139,626	2363
ST mech'l losses	+1	52	2.65×10^{03}	0.00%	515	51
GT shaft output	–1	7891	6.23×10^{07}	89.63%	256,692	2541
Error					0	

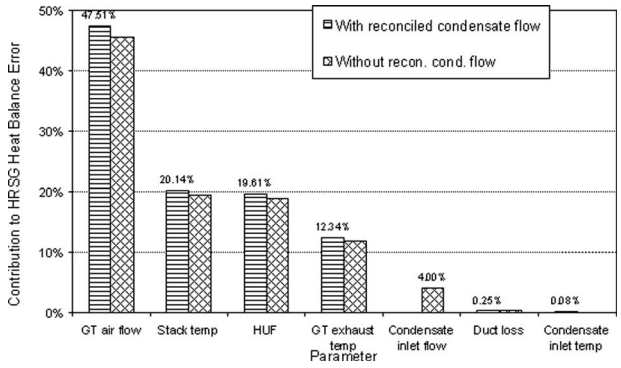


Fig. 2 Contribution of measured parameters to the HRSG heat balance error

are possible culprits for this imbalance, error analysis points to a few likely candidates, as can be seen in Fig. 2, which shows the w values obtained from Eq. (12) in bar-chart form.

Note that many steam/water pressures and temperatures with $w_i \approx 0$ are omitted for clarity. Therefore, they are not to be included in the HRSG heat balance reconciliation and used at their reconciled values. Incorporating mass conservation into the HRSG heat balance does not lead to a significant change in the Pareto chart. The flow measurement that would have the largest impact on the HRSG heat balance is the condensate flow that is also the most accurate flow measurement, and it would account for about 4% of the HRSG heat balance error, as shown in Fig. 2. Similarly, the Pareto chart in Fig. 3 shows that there are only a few parameters that are significant in their impact on the GT heat balance error. In particular, air and fuel flows account for three quarters of the GT heat balance error. The reader is referred to the Appendix for the details of the calculations that generated the data displayed in Fig. 3.

Two parameters that show up prominently in HRSG and GT heat balances are the GT airflow and the exhaust temperature. The GT airflow is the largest contributor to the heat balance error in both cases.

The Pareto charts in Figs. 2 and 3 illustrate the most challenging aspect of the data reconciliation process that also makes it more an art than an exact science. If an inspection of the instrumentation and data acquisition system does not show any defects that may render one or more of the pressure and/or temperature readings useless, it is next to impossible to discern a "bad" reading. One way to detect such faulty measurements is to compare the measured parameters with their expected values from a reliable system model, which is the very first step in the process. Typically this is helpful in detecting gross errors of a rather large magnitude. More subtle errors in ultimately faulty readings can be detected by comparing the correction term and the measurement

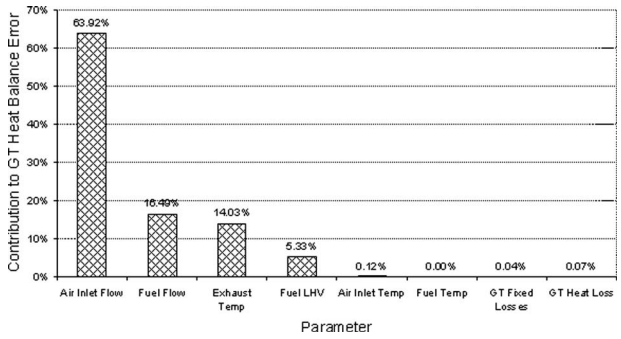


Fig. 3 Contribution of measured parameters to the GT heat balance error

uncertainty. If the reconciliation of the heat and/or mass balance indicates that for a particular measured parameter

$$|x_i^* - x_i| > U_i$$

a closer look into that subsystem is warranted. In general, a correction that is larger in magnitude than the measurement uncertainty should not be taken at its face value. This may indicate a measurement problem (i.e., bad sensor, calibration drift, etc.), either for the parameter in question or for other parameters, as well as a calculation error in the data reduction process. A modular and semi-automated spreadsheet-based structure of the data reconciliation process is particularly useful in that it enables the engineer to focus on the individual systems on a test-by-test basis and to take corrective action during the process.

Two approaches to the final data reconciliation can be identified:

1. *HRSG-to-GT*: Close the HRSG heat balance error to determine the reconciled values of GT air flow and exhaust temperature. Using these reconciled parameters, close the GT heat balance error.
2. *GT-to-HRSG*: Close the GT heat balance error to determine the reconciled values of GT air flow and exhaust temperature. Using these reconciled parameters, close the HRSG heat balance error.

Using the first method, HRSG-to-GT, airflow is determined with an uncertainty of +1.16%. When the airflow and exhaust temperature values from HRSG reconciliation are substituted into the GT heat balance, about 75% of the heat balance error becomes attributable to the fuel flow.

Using the second method, GT-to-HRSG, airflow is determined with an uncertainty of +0.94%. When the airflow and exhaust temperature values from GT reconciliation are substituted into the HRSG heat balance, the HRSG heat balance error becomes attributable in about equal measure to the HUF and stack temperature.

The ST heat balance closure requires the reconciliation of a 300 kW difference between the shaft outputs from the ST heat balance and the single-shaft energy balance. Substituting unreconciled values for mass flow rates into the ST heat balance, it is seen that, by far, the largest contributor to the ST heat balance error is the LP efficiency, followed by the condensate flow. These two measurements account for nearly 85–90% of the ST heat balance error (or, equivalently, of the ST shaft output uncertainty.) Substituting reconciled values for mass flow rates from Step 2 into the ST heat balance, however, the LP efficiency is essentially the sole contributor to the ST shaft output uncertainty.

A summary of the data reconciliation process is shown in Table 5. Based on the sensitivity analysis, the GT shaft output is identified as the largest contributor to the observed shaft energy balance error. The value calculated from the GT heat balance is nearly 3 MW larger than its expected value and has an uncertainty of $\pm 3.04\%$ equivalent to about 8 MW. Data reconciliation led to the best estimate of the GT shaft output that is within a few hundred kilowatts of its expected value and has a much lower uncertainty of less than $\pm 1\%$ (about 2.5 MW). Furthermore, the possibility of underestimation of true plant performance based on raw fuel flow and generator outputs is also highlighted. The reconciled heat consumption that represents the best estimate of the true value is about 2.34 MW (8 MMBtu/h) less than the raw measurement with ± 1.47 MW (5 MMBtu/h) better uncertainty. This is equivalent to about 0.4% of the gross heat rate. The measured parameters and performance numbers (output and heat consumption) constitute a thermodynamically consistent set satisfying mass and energy conservation for all subsystems and the overall system. As such, the reconciliation gives the best-possible estimate of the plant's performance (i.e., the lowest possible uncertainty) and provides reliable data for individual component characterization and design model tuning.

Table 5 Summary of reconciled measurements

		As measured		GT→HRSG		HRSG→GT	
	Units	X	U	X^*	U^*	X^*	U^*
GT airflow	kg	643.2	10.3	646.3	6.1	644.0	7.5
	(pps)	(1418.0)	(22.8)	(1424.8)	(13.4)	(1419.8)	(16.5)
GT fuel flow	kg	13.80	0.07	13.79	0.06	13.76	0.03
	(pps)	(30.42)	(0.15)	(30.40)	(0.14)	(30.34)	(0.07)
GT shaft output	kW	259,771	7891	256,692	2541	256,692	2541
GT exhaust temperature	°C	603.9	3.9	604.4	3.6	604.1	3.7
	(°F)	(1119.0)	(7.0)	(1119.9)	(6.5)	(1119.3)	(6.6)
Fuel LHV	kJ/kg	50,000	134.9	49,988	132.6	49,958	118.6
	(Btu/lb)	(21,496)	(58)	(21,491)	(57)	(21,478)	(51)
HRSG HUF	-	0.990	0.010	0.989	0.007	0.991	0.009
Stack temperature	°C	88.3	5.6	88.8	4.1	88.1	4.9
	(°F)	(191.0)	(10.0)	(191.9)	(7.3)	(190.5)	(8.9)
Duct loss	°C	1.11	0.56	1.12	0.56	1.10	0.56
	(°F)	(2.00)	(1.00)	(2.01)	(1.00)	(1.99)	(1.00)
CC generator loss	kW	4000	200	4002	200	4002	245
CC mechanical loss	kW	550	28	550	27	550	36
ST gross output	kW	139,929	2475	139,626	2363	139,626	2363
ST mechanical loss	kW	515	52.0	515	51.5	515	51.5
CC generator output	kW	391,200	1017	391,251	1010	391,251	1010
Heat consumption	MW	689.92	3.87	689.24	3.60	687.46	2.34
	(MMBtu/h)	(2354.1)	(13.2)	(2351.8)	(12.3)	(2345.7)	(8.0)
Gross heat rate	kJ/kWh	6348.9	39.4	6342.0	36.9	6325.6	27.0
	(Btu/kWh)	(6017.6)	(37.3)	(6011.0)	(35.0)	(5995.5)	(25.6)

Conclusions

A simple but very powerful mathematical method for reconciling measured data is derived from the fundamental principles of error propagation. The method uses fundamental conservation equations and measurement (sensor) uncertainties to develop corrections to raw data. For systems governed by linear conservation equations, the method is exact and leads to the least possible measurement uncertainty for each parameter while ensuring thermodynamic consistency. The method is equally robust even for more complex systems governed by nonlinear conservation equations. It is shown that ignoring second order effects leads to negligible error, and the method can be applied to all power plant systems and subsystems via mass and energy conservation equations.

Using simulated (not the actual) data, the application of the method to a single-shaft combined cycle power generation system has been demonstrated. In the absence of widely accepted and field-proven techniques, data reconciliation is shown to determine the GT output with less than $\pm 1.0\%$ uncertainty. It is also shown that, provided that the precision data for water and temperature flows are available, the HRSG can be utilized to determine the GT inlet airflow with an uncertainty of $\pm 1.15\%$.

The method is amenable to full automation and should be particularly useful for online continuous performance monitoring systems. These systems employ station data from the plant's distributed control system (DCS) or data historian that uses data from station instrumentation that are typically less accurate than precision-calibrated sensors used in performance tests. However, the best application for the technique is a single-shaft combined cycle with both gas and steam turbines driving a common generator. The absence of a reliable and accurate measuring system to individually determine gas and steam turbine shaft outputs makes the method described herein vital for an accurate analysis of data in such systems. In fact, the method has been successfully deployed during the characterization testing of GE's H-System™ in Baglan Bay, Wales.

Disclaimer

The plant data cited in this paper are generated using a commercially available heat balance simulation tool for the sole purpose of demonstrating the data reconciliation techniques that can be used in any power plant system performance test. As such, they constitute a "pseudo-data set" that does not represent an actual plant performance of any vintage, past, present, or future.

Nomenclature

ε = heat or mass balance error
σ = standard deviation
x_i = generic ith measured parameter
U_i = measurement uncertainty of x_i
w_i = fraction of error ε attributable to x_i
n = total number of measured parameters
S_i = sensitivity coefficient; partial derivative of ε with respect to x_i
$\dot{W}$ = shaft work transfer across the control volume (CV) boundaries; (–) into the CV, (+) out of CV, Btu/s or kW
$\dot{Q}$ = heat transfer across the CV boundaries; (+) into the CV, (–) out of CV, Btu/s or kW
$\dot{m}$ = mass flow rate across the CV boundaries; (+) into the CV, (–) out of CV, kg/s or lb/s
p = pressure, bara or *psi* (absolute)
T = temperature, °C or °F
h = enthalpy, kJ/kg or Btu/lb
LHV = fuel lower heating value, kJ/kg or Btu/lb

Subscripts

i = measured parameter index, $i=1,2,\ldots,n$
1 = GT compressor inlet
4 = GT exhaust
F = GT combustor fuel inlet
mech = mechanical
elec = electrical

Superscripts

* = reconciled parameter

Table 6 GT heat balance sensitivity parameters and measurement uncertainties

	X	S	U	$(SU)^2$	w
Air inlet flow, kg/s (pps)	643.2 (1418)	−268.11	22.80	37,366,265	63.92%
Fuel flow, kg/s (pps)	13.80 (30.42)	20,696.0	0.15	9,637,300	16.49%
Exhaust temperature, °C (°F)	603.9 (1119)	−409.16	7.00	8,203,069	14.03%
Fuel LHV, kJ/kg (Btu/lb)	50,000 (21,496)	30.42	58.04	3,117,176	5.33%
Air inlet temperature, °C (°F)	15.0 (59.0)	264.00	1.00	69,699	0.12%
Fuel temperature, °C (°F)	82.2 (180.0)	17.28	3.00	2688	0.00%
GT bearing losses, kW	1475	−1.00	147.50	21,756	0.04%
GT heat loss, kW	2050	−1.00	205.00	42,025	0.07%

Appendix

A numerical example that illustrates the application of the data reconciliation principle is given below. Note that the numbers in formulas are in English units in order to keep them simple and easy to check with reference to the data in the text. The gas turbine heat balances, similar to Eq. (16), can be written as

$$\dot{m}_1 \cdot h_1 + \dot{m}_f \cdot (\mathrm{LHV} + h_f) - (\dot{m}_1 + \dot{m}_f) \cdot h_4 - \dot{Q}_{\mathrm{heat}} - \dot{Q}_{\mathrm{mech}} - \dot{W}_{\mathrm{GT}} = 0 \tag{A1}$$

Since an independent measurement for $\dot{W}_{\mathrm{GT}}$ is not available, the heat balance closure (i.e., $\varepsilon_{\mathrm{heat}}=0$) will be assumed by calculating it from Eq. (A1). Substituting the measured values into Eq. (A1), the GT shaft output is calculated as 259,771 kW$\pm$3.04%. This is the measured value of the GT shaft output. In a similar fashion, the ST shaft output is obtained from the ST heat balance as 139,929 kW$\pm$1.77%. GT and ST shaft uncertainties are calculated from the RSS formula that is applied to the heat balance equations. The single-shaft energy balance can then be written as follows:

$$\{\dot{W} + \dot{Q}_{\mathrm{mech}} + \dot{Q}_{\mathrm{elec}}\}_{\mathrm{gen}} - \{\dot{W} - \dot{Q}_{\mathrm{mech}}\}_{\mathrm{ST}} - \dot{W}_{\mathrm{GT}} = \varepsilon_{\mathrm{shaft}} \tag{A2}$$

Substituting the measured values from Table 3 into Eq. (A2), it can be seen that the error is −3435 kW. In other words, the measured generator output is nearly 3.5 MW lower than the sum of the GT and ST outputs calculated from their respective heat balances. Since Eq. (A2) is a linear equation, sensitivity coefficients for the measured parameters therein are ± 1. Therefore, the contribution of the measurement uncertainty of each parameter in Eq. (A2) to the shaft energy balance error, $\varepsilon_{\mathrm{shaft}}$, can be calculated from Eq. (12); e.g., for the GT,

$$w = \frac{7891^2}{1017^2 + 200^2 + 28^2 + 2475^2 + 52^2 + 7891^2} = 0.8963$$

Thus, approximately 90% of the observed shaft heat balance error can be attributed to the uncertainty in the GT shaft output. Applying Eq. (13), the reconciled value of the GT shaft output can be found as

$$\dot{W}^*_{\mathrm{GT}} = 259{,}771 - \frac{-3435 \cdot 0.8963}{-1} = 256{,}692$$

The uncertainty of the reconciled GT shaft output measurement is obtained from Eq. (14) as

$$U^*_{\mathrm{GT}} = \sqrt{(1 - 0.8963)} \cdot 7891 = 2541$$

Returning to the GT heat balance equation, Eq. (A1), it can be rewritten by substituting $\dot{W}^*_{\mathrm{GT}}$ for $\dot{W}_{\mathrm{GT}}$,

$$\dot{m}_1 \cdot h_1 + \dot{m}_f \cdot (\mathrm{LHV} + h_f) - (\dot{m}_1 + \dot{m}_f) \cdot h_4 - \dot{Q}_{\mathrm{heat}} - \dot{Q}_{\mathrm{mech}} - \dot{W}^*_{\mathrm{GT}} = \varepsilon_{\mathrm{heat}} \tag{A3}$$

Equation (A3) does show a heat balance error that is exactly equal to the difference between $\dot{W}_{\mathrm{GT}}$ and $\dot{W}^*_{\mathrm{GT}}$, i.e., 3079 kW (or 2918 Btu/s). The task now is to calculate the reconciled values for individual measurements in Eq. (A3) that will render the heat balance error zero. This will be accomplished by applying Eqs. (12) and (13) to each independent measurement in Eq. (A3). Let us consider the GT inlet airflow, $\dot{m}_1$, that is measured as 1418 ± 22.8 pps ($\pm 1.61\%$). The sensitivity coefficient can be calculated as

$$S_{\dot{m}_1} = \frac{\partial \varepsilon_{\mathrm{heat}}}{\partial \dot{m}_1} = h_1 - h_4 - (\dot{m}_1 + \dot{m}_f) \cdot \frac{\partial h_4}{\partial \dot{m}_1} = -4.34 - 275.97 + 12.21 = -268.11 \tag{A4}$$

The enthalpies are calculated from the JANAF tables at T_1 and T_4 with the exhaust gas composition that is calculated from the combustion stoichiometry with specified fuel and air compositions. Note that the third term on the RHS of Eq. (A4) is the impact of airflow on exhaust gas enthalpy h_4 via exhaust gas composition. For the exhaust temperature, the sensitivity coefficient is

$$S_{T_4} = \frac{\partial \varepsilon_{\mathrm{heat}}}{\partial T_4} = -(\dot{m}_1 + \dot{m}_f) \cdot \frac{\partial h_4}{\partial T_4} = -(1418 + 30.4) \cdot 0.2825 = -409.16$$

The partial derivative of the exhaust gas enthalpy with respect to the exhaust temperature is obtained numerically. The sensitivities for all measured parameters are obtained similarly. The results are shown in Table 6. The contribution of the measurement uncertainty of the GT inlet airflow in Eq. (A3) to the GT shaft energy balance error, $\varepsilon_{\mathrm{heat}}$, is calculated from Eq. (12),

$$w_{\dot{m}_1} = \frac{(-268.11 \cdot 22.8)^2}{5.85 \cdot 10^7} = 0.6392$$

The denominator is the sum of the $(S_i U_i)^2$ terms in the last column of Table 6. Applying Eq. (13), the reconciled value of the GT inlet airflow is found as

$$\dot{m}^*_1 = 1418 - \frac{2918 \cdot 0.6392}{-268.11} = 1425.0$$

The uncertainty of the reconciled GT inlet airflow measurement is obtained from Eq. (14) as

$$U^*_{\dot{m}} = \sqrt{(1 - 0.6392)} \cdot 22.8 = 13.70$$

In summary, the original GT inlet airflow measurement of 1418 pps ($\pm 1.61\%$) has been reconciled with other FPT data for a consistent GT and single-shaft CC energy balance to result in a new value of 1425.0 pps. The new value has a lower uncertainty than the original one, i.e., $13.7/1425.0 = 0.96\%$, because it incorporates more information provided by the conservation equations.

As an alternative, w_i can be evaluated using a Monte Carlo (MC) simulation, e.g., Crystal Ball by Oracle (Denver, CO). Using this Excel-based tool, the heat balance equation, Eq. (A3), is set up in Excel with individual measurements identified as independent variables with mean (set to the measurement itself) and standard (set to the measurement uncertainty) deviations based on a normal probability distribution. After running the simulation, w_i can be read from the sensitivity chart that shows the "contribution to variance" of each independent variable (i.e., measured parameter). The analogy between a measurement x with an uncertainty U and a normal probability distribution with a mean μ and a standard deviation σ is another confirmation of the linearity assumption.

References

[1] Romagnoli, J. A., and Sanchez, M. C., 2000, *Data Processing and Reconciliation for Chemical Process Operations*, Process System Engineering Vol. II, 1st ed., Academic, New York.

[2] Urban, L. A., 1972, "Gas Path Analysis Applied to Turbine Engine Condition Monitoring," AIAA/SAE Eighth Joint Propulsion Specialist Conference, New Orleans, LA, Nov. 29–Dec. 11, Paper No. 72-1082.

[3] Volponi, A. J., 1982, "Gas Path Analysis: An Approach to Engine Diagnostics," 35th Symposium of the Mechanical Failures Prevention Group, Gaithersburg, MD, Apr. 20–22.

[4] Escher, P. C., 2002, "Gas Turbine Data Validation Using Gas Path Analysis," ASME Paper No. GT-2002-30024.

[5] Doel, D. L., 1994, "TEMPER: A Gas Path Analysis Tool for Commercial Jet Engines," ASME J. Eng. Gas Turbines Power, **116**, pp. 82–89.

[6] Hartner, P., Petek, J., Pechtl, P., and Hamilton, P., 2005, "Model-Based Data Reconciliation to Improve Accuracy and Reliability of Performance Evaluation of Thermal Power Plants," ASME Paper No. GT-2005-68937.

[7] Cheng, P.-C., and Andersen, H., 2005, "The Implementation of the Data Validation Process in a Gas Turbine Performance Monitoring System," ASME Paper No. GT-2005-68429.

[8] Kalman, R. E., 1960, "A New Approach to Linear Filtering and Prediction Problems," ASME J. Basic Eng., **82**, pp. 35–45.

[9] Maybeck, P. S., 1979, *Stochastic Models, Estimation and Control*, Vol. 1, Academic, New York, pp. 1–16.

[10] ASME PTC 19.1, 1998, *Measurement Uncertainty*, The American Society of Mechanical Engineers, New York.

[11] Stull, D. R., and Prophet, H., 1971, *JANAF Thermodynamic Tables*, 2nd ed., National Bureau of Standards No. NSRDS-NBS 37.

[12] ASME PTC 46, 1996, *Performance Test Code on Overall Plant Performance*, The American Society of Mechanical Engineers, New York.

[13] ASME PTC 6, 1985, *Guidance for Evaluation of Uncertainty in Performance Tests of Steam Turbines*, The American Society of Mechanical Engineers, New York.

[14] Umezawa, S., 2005, "Diagnosis of Thermal Efficiency of Combined Cycle Power Plants Using Optical Torque Sensors," JSME Int. J., Ser. B, **48**(1), pp. 129–135.

[15] Cotton, K. C., 1998, *Evaluating and Improving Steam Turbine Performance*, 2nd ed., Cotton Fact Inc., Rexford, NY.

[16] Shapiro, H., 1958, *The Dynamics and Thermodynamics of Compressible Fluid Flow*, Ronald, New York, Vol. 1, Pts. I and II, p. 104.

[17] GT PRO Version 14.0.0, ©1987–2004, Thermoflow, Inc., 29 Hudson Road, Sudbury, MA 01776.

[18] ASME PTC 4.4, 1981, *Gas Turbine Heat Recovery Steam Generators*, The American Society of Mechanical Engineers, New York.

A Conceptual Design Study for a New High Temperature Fast Response Cooled Total Pressure Probe

Jean-François Brouckaert
e-mail: brouckaert@vki.ac.be

Mehmet Mersinligil
e-mail: mersinli@vki.ac.be

Marco Pau
e-mail: pau@vki.ac.be

Turbomachinery and Propulsion Department,
von Kármán Institute for Fluid Dynamics,
72, Chaussée de Waterloo,
B-1640 Rhode-Saint-Genèse, Belgium

The present paper proposes a concept for a water-cooled high temperature unsteady total pressure probe intended for measurements in the hot sections of industrial gas turbines or aero-engines. This concept is based on the use of a conventional miniature piezoresistive pressure sensor, which is located at the probe tip to achieve a bandwidth of at least 40 kHz. Due to extremely harsh conditions and the intention to immerse the probe continuously into the hot gas stream, the probe and sensor must be heavily cooled. The short term objective of this design is to gain the capability of performing measurements at the temperature conditions typically found at high pressure turbine exit (1100–1400 K) and in the long term at combustor exit (2000 K or higher). [DOI: 10.1115/1.2969092]

1 Introduction

Accurate hot gas path measurements (including combustion diagnostics) are recognized as a major need for the assessment of engine component health and performance. This fact has been highlighted in a recent study, "The Lab Gap Matrix," carried out by the European Virtual Institute for Gas Turbine Instrumentation (EVI-GTI) [1]. Besides the interest to validate numerical predictions of engine performance, efficiency, and component life, the availability of high temperature steady and unsteady pressure and temperature data would also enable much progress in the field of condition monitoring and active control technologies, as outlined by Simon et al. [2] in their review of sensor needs for control and health management of future intelligent aircraft engines.

Due to the limitations of current experimental techniques and of the severe flow conditions, the measurement of pressure and temperature in the hot sections of modern gas turbines still remains a true challenge for test engineers.

Regarding unsteady pressure measurements above temperatures of 1300 K, few, if any, accurate and reliable technique exists to measure the time-resolved gas path total pressure. The long term goal of this probe development is the measurement of combustion instabilities inside the combustion chamber at temperatures up to 2000 K and above, and in the short term, proof of the concept through measurements at the temperature conditions typically found at high pressure turbine exit (1100–1400 K) with a high bandwidth. A high temperature total pressure probe would yield the unsteady pressure field between turbine blade rows, including turbulence level, shock position, and wake passing events. The single sensor probe could also be used in a virtual three-hole mode, yielding the periodic flow angle fluctuations [3].

To resolve the highly three-dimensional and unsteady flow field within turbine stages, it is required to record continuous time-series of pressure. At those elevated temperatures, it is therefore mandatory to design a highly efficient cooling system for a probe continuously immersed into the gas path.

The objective is to investigate different probe arrangements and cooling solutions aiming at gas temperatures of 1400 K. Both gas and liquid coolants may be considered but emphasis was given on water cooling because of its higher heat capacity. A thermal barrier coating (TBC) is envisaged to reduce the probe metal temperature. Besides the cooling of the probe stem, which requires creating a uniform cooling stream, avoiding stagnating fluid regions inside the probe channels, the most challenging part of the cooling design is the reduction of the gas temperature in the cavity ahead of the sensor diaphragm.

2 State of the Art

Measurements of unsteady pressure in gas-turbine combustors or in hot turbine sections have traditionally been accomplished using two different types of devices: flush mounted wall pressure sensors or immersion type instruments.

2.1 High Temperature Pressure Sensors. Although flush mounted wall static pressure sensors do not yield the gas path total pressure as intended in the present development, some high temperature sensors have recently been developed in either cooled or uncooled configurations. Commercially available uncooled piezoelectric transducers allow the measurement of pressure fluctuations up to 1020 K with a bandwidth of 15 kHz. Some other uncooled sensor developments based on fiber optic Fabry–Perot interferometry allow the combined measurement of pressure and temperature up to 1000 K. Although those technological achievements greatly helped in increasing the limit of temperature at which measurements can be done, the size of those transducers still remains prohibitive for use in aerodynamic probes [4].

Semiconductor pressure sensors are probably most widely used because of their small size and their ability to measure both steady and unsteady components of pressure with a high bandwidth (150 kHz–1 MHz) over a wide pressure range (0.35–70 bars). Up to present, the use of miniature pure silicon-based piezoresistive pressure sensors in fast response pressure probes was limited to hardly 420 K due to leakage current appearing at the *p-n* junction. The development of silicon-on-insulator (SOI) sensors pushed the limit to beyond 780 K [5]. Silicon carbide (SiC)-based sensors allow extended use up to 880 K [6] and have been demonstrated in a cantilever beam configuration at that temperature in a running gas turbine [7]. Nevertheless, stability of the transducers remains an issue at those elevated temperatures, and in order to reach higher temperatures, as encountered in the hot parts of gas turbines, sensor cooling becomes mandatory.

In recent years, some air- or water-cooled sensors became commercially available with quoted temperature capabilities up to

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 10, 2008; final manuscript received April 10, 2008; published online December 29, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

1370 K in the water-cooled configuration and up to 1070 K in the air-cooled configuration. Ferguson and Ivey [8,9] have conducted numerous studies regarding cooling of these transducers, have presented an air-cooled jacket design to protect an unsteady pressure transducer, and demonstrated survival up to 1170 K [10]. However, once again, the overall dimensions of those cooled sensors remain inappropriate for implementation in a probe assembly.

2.2 High Temperature Pressure Probes. Historical development of probe techniques in high temperature gases is to be found in early plasma research. This development began with Langmuir's [11] classical studies in ionized gases of very low density and therefore low heat flux to the probes. The first example of a relatively sophisticated diagnostic probe utilizing forced cooling methods was the double-sonic-orifice probe developed in the late 1940s and described in detail by Grey [12]. Subsequent developments have led to the evolution of sophisticated cooled probe techniques mainly for calorimetric, electrostatic, heat flux, or gas sampling probes [13].

Regarding probe assemblies continuously immersed into the flow path, Ashby [14] and Lagen et al. [15] were among the first researchers to patent a pneumatic pressure probe employing water cooling for thermal protection. In 1988, Ashby has patented a water-cooled pressure probe, which is used to measure the average total pressure inside a hot gas stream. The design employs a pressure chamber, which contains the sensor that is connected to an external cooling system via tubes. However, the system patented is not suitable for fast response measurements since any change of pressure therein requires a settling-time to appear in the closed pressure chamber. Lagen et al. developed a water-cooled probe for measuring average static pressure operable up to 1650 K. Regarding average pressure measurements, it is worth mentioning the recent work of Warnack [16], who presented a pneumatic four-hole probe designed to be operated up to 1070 K, including a thermocouple, where the probe shaft is water cooled and the 6 mm diameter probe tip is not cooled.

Currently some attempt is being made to develop a high temperature intermittent aspirated probe to measure simultaneously pressure and temperature. A laboratory capability has been demonstrated by Yang et al. [17] in a furnace up to 1800 K but the maximum service temperature of the probe designed for engine environments is currently limited to 1300 K.

Regarding higher response designs, an original design of a transient pressure probe for the measurement of detonation wave characteristics inside rocket thrust chambers experiencing high-frequency combustion instabilities was presented by Grey in 1970 [13]. A 5.6 mm diameter piezoelectric crystal transducer was mounted in the probe tip and the transducer diaphragm was splash-cooled by a ring of tiny re-entrant jets directed at different angles so as to completely cover the transducer face.

Nieberding and Englund [18] presented a Pitot-type combined steady state and dynamic pressure probe with a far recessed water-cooled transducer in 1972. In 1977, Moore [19] successfully demonstrated the capability to acquire high response (100 Hz–10 kHz) fluctuating pressure data in a high temperature flowing gas stream such as an aircraft engine exhaust. Miniature pressure transducers are mounted in a 12.7 mm diameter water-cooled probe with their sensing surfaces parallel to the flow and flush with the probe outer wall, while the probe is immersed in a high temperature gas stream (1250 K, 2 atm). A similar approach is also presented by Larguier [20].

A new approach, pioneered by the University of Oxford, is the fast immersion technique where an uncooled probe is fitted onto a rapid injection mechanism and an external air cooling system. Pressure and temperature probes have been built, incorporating a commercially available piezoresistive sensor, a pneumatic pressure port, a thermocouple, and a dual thin film sensor on an 8 mm outer diameter probe head, yielding the steady and unsteady total pressure and temperature. Measurements have been performed in an industrial gas-turbine combustor by Passaro et al. [21] However, the drawback of this method is the short injection time, of the order of 100 ms, which is mandatory for the survival of the probe. The transient, which both fast response sensor and pneumatic pressure port undergo during the injection, makes it difficult to obtain statistically representative pressure data at a given position of the traverse. The response of the pneumatic line must be sufficiently high to be able to use this value as a reference for the pressure given by the fast response sensor, which is submitted to a very strong thermal transient during the injection and retrieval.

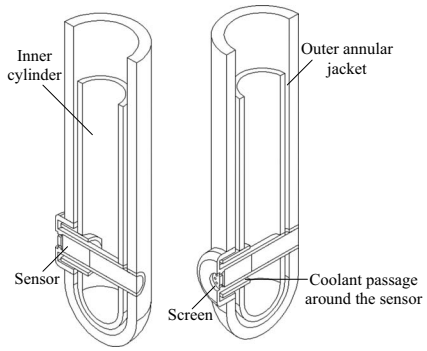


Fig. 1 Fast response cooled total pressure probe concept

3 Design Requirements

The primary objective of this design study is to assess the feasibility of a probe concept based on the following requirements:

- the measurement of both steady and unsteady components of the gas path total pressure with high accuracy ($\pm 0.1\%$ full-scale) and high signal-to-noise ratio
- the measurement of both periodic (by phase-locked averaging over several blade passages or rotor revolutions) and random fluctuations
- the highest possible frequency bandwidth, ideally up to 100 kHz
- the smallest possible probe head size for optimum spatial resolution and minimal flow disturbance
- a high robustness and reliability of design considering the extremely harsh conditions of the hot turbine environment

4 Probe Concept

The intention to obtain both the periodic and random fluctuations, respectively, to obtain statistically representative data for phase-locked averaging and for the proper measurement of turbulence levels has favored the choice of a cooled probe design for continuous immersion into the hot gas stream.

The objective of measuring both the steady and unsteady components of total pressure has favored the use of a piezoresistive sensor as opposed to piezoelectric sensors. The small dimensions (1.65 mm outer diameter and 2.0 mm length), high signal-to-noise ratio, good temperature behavior, and availability of a wide variety of pressure ranges are also important factors for the choice of this type of sensors. The design is based on a commercially available high temperature ultraminiature Kulite sensor (XCE-062) with an operating temperature range from 218 K to 546 K, temperature compensated between 298 K and 508 K. This upper operating temperature limit, 546 K, is actually set by the connecting Teflon insulated leads; the sensor itself being based on the leadless SOI technology is able to withstand temperatures above 773 K. The sensor is bare, without any room-temperature vulcanizing (RTV) coating to maintain the highest frequency response, but is protected by a perforated screen, which is cooled as part of the probe body rather than as part of the sensor itself (Fig. 1).

One of the objectives of the cooling system design is to keep the sensor operating temperature below 373 K, to restrict tempera-

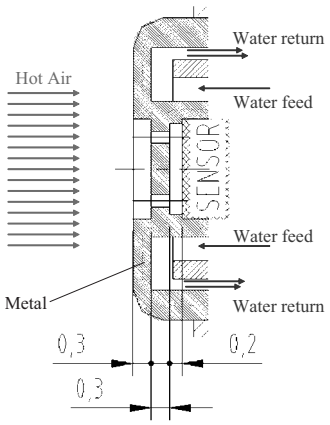


Fig. 2 Sensor head and screen cooling configuration

ture effects. The sensor will be operated without the passive temperature compensation commercially available but rather with an active monitoring of the bridge temperature by an electronic circuit incorporating a sense resistor in series with the Wheatstone bridge [22]. This procedure requires a calibration of the sensor both in pressure and temperature prior to measurements, but the output signal proportional to the diaphragm temperature can be used as a safety indication of the cooling effectiveness during operation.

The probe concept is illustrated in Figs. 1 and 2. The probe is essentially based on a concentric tube assembly to bring and return the coolant fluid from the sensor region. The sensor itself is located transversally to the probe axis inside a tube, which also incorporates the protective screen ahead of the sensor diaphragm. The sensor and screen cooling are performed by creating a narrow return channel around the sensor by another transversal concentric tube assembly.

5 Environmental Constraints

Combustion chamber or high pressure turbine sections of gas turbines are probably among the most hostile environments to perform measurements.

Regarding flow conditions at those locations, the highest pressure and temperature of the latest aero-engine thermodynamic cycles exceeds 50 bars and 2000 K, even up to 2200 K in peak operation. The huge heat flux seen by a probe immersed in such conditions is not only caused by the high temperature and high density but is also enhanced by the high level of turbulence due to the blade passing events, the high velocity of the gas, radiation heat transfer, and obviously by the probe geometry and immersion depth itself.

It is therefore crucial to have a correct estimate of the external heat transfer to the probe in order to determine the power to be extracted from the probe by the cooling system to assess the feasibility of this design.

The external heat transfer to the probe has been estimated from various correlations between Nusselt and Reynolds numbers for a cylinder in cross-flow. A detailed review of the many correlations that have been developed for the circular cylinder is provided by Morgan [23]. However, many correlations do not include the effect of turbulence intensity, which was shown to have a substantial influence on the heat transfer (Ref. [24]) and may reach values as high as 15% at the combustion chamber exit [21]. The correlations considered in the present study are those that are probably the most widely used within the gas-turbine heat transfer community, namely, Smith and Kuethe [25] in Eq. (1), Kestin and Wood [26] in Eq. (2), Lowery and Vachon [27] in Eq. (3), Zukauskas and Ziugzda [28] in Eq. (4), and Zukauskas [29] in Eq. (5).

$$\overline{\mathrm{Nu}_d} = \mathrm{Re}_d^{0.5}(1 + 0.0277(1 - e^{-2.91\times10^5\mathrm{Re}_d})(\mathrm{TuRe}_d^{0.5})) \quad (1)$$

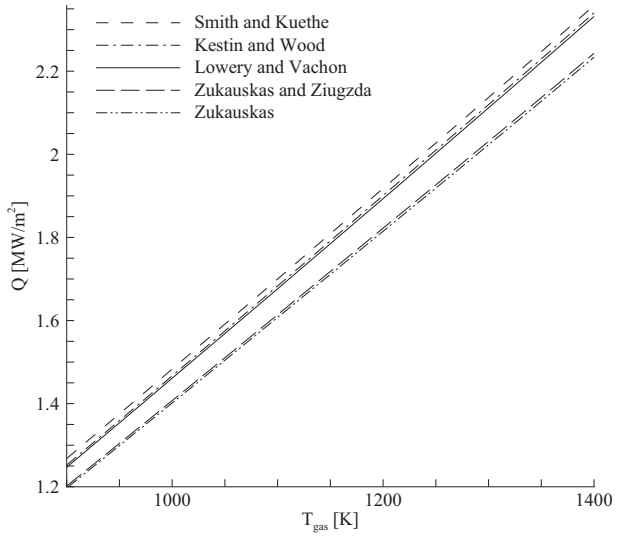


Fig. 3 Comparison of heat flux evolution as a function of gas temperature using different Nusselt number correlations, P_o = 40 bars, Ma = 0.5, and Tu = 0.15

$$\overline{\mathrm{Nu}_d} = \mathrm{Re}_d^{0.5}\left(0.945 + 3.48\left(\frac{\mathrm{TuRe}_d^{0.5}}{100}\right) - 3.99\left(\frac{\mathrm{TuRe}_d^{0.5}}{100}\right)^2\right) \quad (2)$$

$$\overline{\mathrm{Nu}_d} = \mathrm{Re}_d^{0.5}\left(1.010 + 2.64\left(\frac{\mathrm{TuRe}_d^{0.5}}{100}\right) - 3.070\left(\frac{\mathrm{TuRe}_d^{0.5}}{100}\right)^2\right) \quad (3)$$

$$\overline{\mathrm{Nu}_d} = 0.068\left(\frac{\mathrm{Re}_d\pi}{2}\right)^{0.73}(\mathrm{Tu}^{0.17})(\mathrm{Pr}_g^{0.4})\left(\frac{T_g}{T_{\mathrm{wall}}}\right)^{0.12} \quad (4)$$

$$\overline{\mathrm{Nu}_d} = C\mathrm{Re}_d^m\mathrm{Pr}_g^n\left(\frac{\mathrm{Pr}_\infty}{\mathrm{Pr}_s}\right)^{1/4} \quad (5)$$

The effect on the calculated heat flux by using the different correlations is shown in Fig. 3. The worst case scenario is given by the relation of Smith and Kuethe. However, this model, as well as the model of Zukauskas, is not validated for a wide range of Reynolds and Mach numbers, which limit their applicability. For instance, the relation by Smith and Kuethe is valid only for incompressible flows, where the relation by Kestin and Wood is valid for low Reynolds numbers. Therefore, the only model that is applicable is the one by Lowery and Vachon, and has therefore been selected for estimating the average Nusselt number around the probe. The perfect agreement between the three models is worth mentioning although they are not shown to be valid in the applied conditions. As it can be seen from Fig. 3, the heat flux experienced by the probe reaches values in excess of 2 $\mathrm{MW/m^2}$ above 1300 K.

Besides those heat transfer considerations, some other secondary constraints linked to gas-turbine applications demand an accounting of the accessibility (bore diameter, probe length under engine cover, and available space for cooling pipes to the probe) as well as temperatures just outside of the core engine casing, which may still be at the order of 600–700 K.

6 Aerodynamic Design

The quality of fast response aerodynamic probe measurements depends not only on static considerations, typical of pneumatic probes, but also on a series of dynamic characteristics. As noted by Gossweiler [30], all of those characteristics are strongly related to the probe geometry and often to the probe size as well, essentially when unsteady effects are concerned.

In terms of static characteristics, and in the case of total pressure measurements only, one may consider factors such as angular

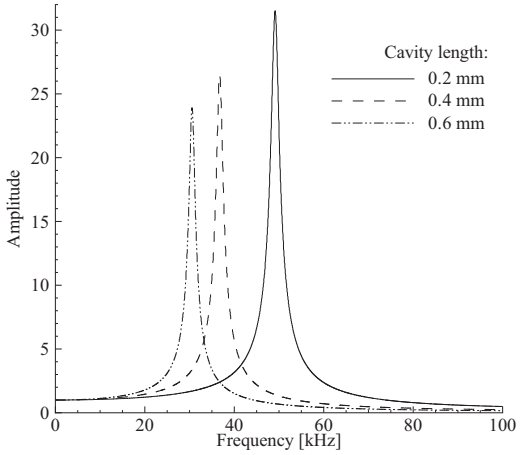


Fig. 4 Comparison of theoretical transfer functions for different cavity lengths for the frequency response of screen-cavity configuration

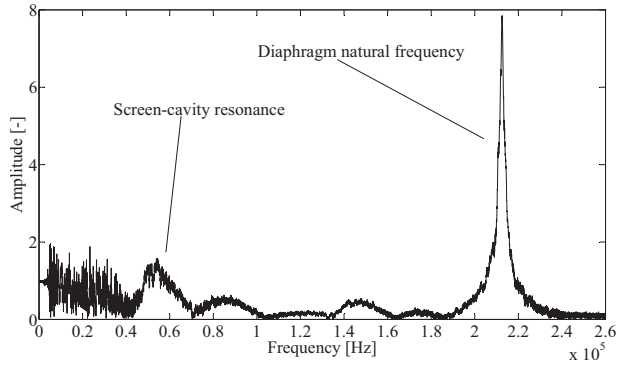


Fig. 5 Experimental transfer function from shock tube tests for the frequency response of screen-cavity configuration

insensitivity range, Mach number, Reynolds number and turbulence effects, as well as velocity gradient and blockage effects. Unfortunately, most of those effects are unavoidable and are usually taken into account by an appropriate calibration. They can be reduced by minimizing the probe size and choosing an appropriate probe geometry.

In terms of dynamic characteristics, other than the frequency response itself, one may consider the reduction of unsteady errors due to unsteady lift, dynamic stall, added mass effects, or vortex shedding interactions. Cylindrical probe geometries have been shown to be the least sensitive to those unsteady effects. Furthermore, total pressure measurements are much less affected by the corresponding errors than side pressure measurements in the case of unsteady angle measurements by multisensor probes [31].

The present design has been carried out with the priority of keeping the probe diameter to a minimum, i.e., 8 mm outer diameter, keeping in mind the complexity of the internal cooling channel arrangements.

Regarding frequency response, the flush mounting of the sensor into a Pitot probe or into a cylinder is of course the best solution, but the recorded value is in general far from the true value because of nonuniform pressure distribution on the diaphragm.

A slight retreat of the diaphragm from the probe mouth cures this problem and improves angular insensitivity. There is, however, no firm indication as to an optimum value of the retreat. Epstein [32] mentioned l/d ratios of 1, Larguier [33] proposed $l/d=0.5$, and Sieverding et al. [34] used $l/d=0.2$–0.3. Of course, any retreat will affect the frequency response. A retreat of 0.3 mm of the screen relative to the probe head (Fig. 2) has been chosen in the present design ($l/d=0.18$) in order to minimize the loss of bandwidth while still maintaining a good angular response.

The frequency response of the corresponding geometry has been estimated by the line-cavity theory of Bergh and Tijdeman [35], assuming a line with an equivalent diameter equal to the sum of the area of the eight holes of 0.3 mm diameter in the screen. The transfer function obtained is presented in Fig. 4 and calculated for two other cavity lengths taking into account the small recess of the sensor with respect to the sensor housing itself. The experimental result from shock tube tests for the configuration shown in Fig. 2 is presented in Fig. 5. Good agreement is found for both theoretical and experimental values with a screen-cavity resonance around 50 kHz, which would yield a flat frequency response at frequencies as high as 20 kHz. It has to be noted that the modeling of this complex configuration requires some strong simplifications, and additionally, the temperature of the gas in the recess cavity can affect the frequency response in such a way to increase the resonance frequency [10].

7 Internal Cooling Design

7.1 Parametric Study: One-Dimensional Model. Both the external and internal heat transfers to the probe are influenced by a number of important parameters, which may also vary significantly depending on the application. Not only the gas path flow conditions (pressure, temperature, Mach number, Reynolds number, turbulence level, etc.) but also outside casing temperature, probe length and diameter, immersion depth, probe material and wall thickness, as well as coolant type (air or water), coolant flow rate, and pressure constitute a wide matrix of possibilities to explore. In view of the complexity of representative 3D Navier–Stokes calculations, a monodimensional prediction tool has been developed. The motivation for the implementation of this one-dimensional model was to create a fast calculation tool for obtaining a good estimate of the global thermal parameters. The main objective was to carry out parametric studies to optimize the design of a probe tailored for a given application or to assess the behavior of a given probe geometry under different flow conditions.

This prediction tool is based on solving the heat balance between the heat flux entering the probe on the hot gas side (for which the appropriate correlations have been discussed above), and transmitted to the coolant by conduction through the probe walls and evacuated by forced convection. To determine the heat transfer coefficients on the wall-water boundary the following Dittus–Boelter [36] equation can be used:

$$\overline{\mathrm{Nu}_d} = 0.023\mathrm{Re}_d^{4/5}\mathrm{Pr}^n = \frac{\overline{h_1}d}{k_w} \tag{6}$$

where $n=0.4$ for heating and $n=0.3$ for cooling. Equation (6) is valid for turbulent flows in circular tubes and it has been experimentally confirmed for the range of conditions: $\mathrm{Re}_d > 10{,}000$, $0.7 \leq \mathrm{Pr} \leq 160$, and $l/d > 10$. As a first approximation, Eq. (6) may be used to obtain a reasonable estimate of $\overline{\mathrm{Nu}_d}$ evaluating all fluid properties at the arithmetic average of the coolant mean inlet and outlet temperatures

$$\overline{T_w} = \frac{T_{w\,\mathrm{in}} + T_{w\,\mathrm{out}}}{2} \tag{7}$$

The underlying model assumes the probe will act as a coaxial counter-flow heat exchanger. The outer gas temperature is represented by T_g, while the probe wall surface temperatures are represented as T_{p_j}, with j denoting the location starting from the innermost wall through the outer probe surface as shown in Figs. 6 and 7, with the possibility to model a TBC layer if required. The heat exchange in the vicinity of the sensor is not modeled, which also means the sensor temperature is not modeled. Instead, the coolant temperature at the specific location, T_{ws}, is calculated, which is lower than the actual sensor temperature by definition.

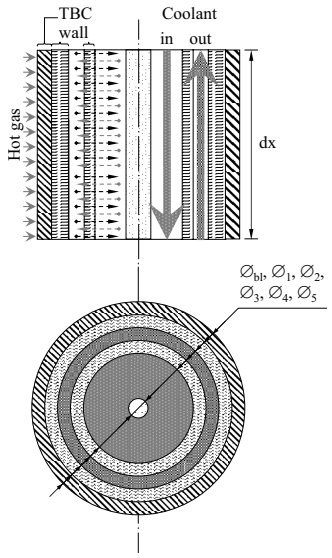


Fig. 6 Schematic representation of the 1D model theory

The probe length is defined as L and the whole heat balance written for an incremental length dx is therefore integrated in order to reach global values. While the wall temperatures are assumed to be uniform, the heat transfer is assumed to take place normal to the surfaces, and therefore normal to the flow direction, i.e., axial heat transfer is neglected. Contact resistance (e.g., between the TBC layer and the probe outer wall) is also not taken into consideration.

One can easily write the basic heat balance shown schematically in Fig. 7, which results in the set of equations given from

$$dq_5 = dq_4 = dq_{\text{out}} - dq_{\text{in}} \tag{8}$$

$$dq_5 = dq_4 = dq_3 = dq_2 + dq_{\text{out}} \tag{9}$$

$$dq_2 = dq_1 = dq_{\text{in}} \tag{10}$$

When these equations are expanded with the appropriate heat transfer mode relation and then combined, the following is obtained:

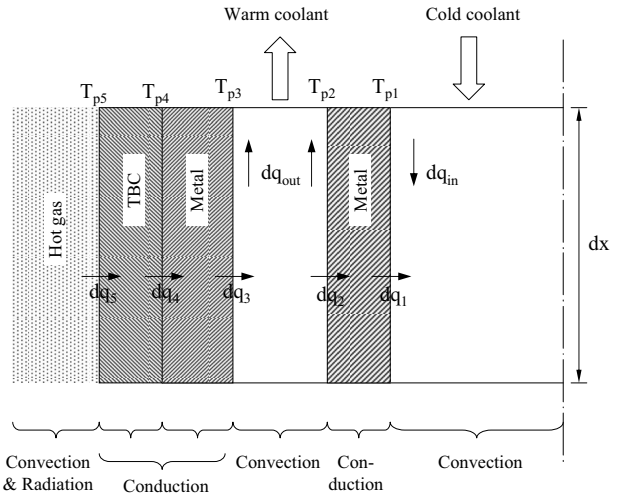


Fig. 7 Schematic representation of the heat flow through the probe

$$dq_5 = h_g \pi d_5 dx(T_g - T_{p5}) = \frac{2\pi k_{\text{TBC}} dx(T_{p5} - T_{p4})}{\ln d_5/d_4}$$

$$= \frac{2\pi k_o dx(T_{p4} - T_{p3})}{\ln(d_4/d_3)} = h_3 \pi d_3 dx(T_{p3} - T_{\text{out}}(x)) \tag{11}$$

One shall note that the term T_{out} stands for the outlet temperature of the incremental element of length dx as shown in Fig. 7. For the sake of simplicity, the common term dx in all equations is dropped and two variables are defined as follows:

$$\alpha = \frac{2\pi k_{\text{TBC}}}{\ln(d_5/d_4)}, \quad \beta = \frac{2\pi k_o}{\ln(d_4/d_3)} \tag{12}$$

The resultant equation is given in Eq. (13) and maybe rewritten as a system of equations as shown in Eq. (14). Using this given system of equations one can proceed to solve for surface temperatures starting with the gas temperature T_g. However, conductivity and convective heat transfer coefficient values are dependent on temperature and pressure. So as new temperature values are obtained, an iterative approach has to be followed in order obtain correct results.

$$h_g d_5(T_g - T_{p5}) = \alpha(T_{p5} - T_{p4}) = \beta(T_{p4} - T_{p3}) = h_3 d_3(T_{p3} - T_{\text{out}}) \tag{13}$$

$$h_g d_5 T_g - h_g d_5 T_{p5} = \alpha T_{p5} - \alpha T_{p4}$$

$$\beta T_{p4} - \beta T_{p3} = \alpha T_{p5} - \alpha T_{p4}$$

$$\beta T_{p4} - \beta T_{p3} = h_3 d_3 T_{p3} - h_3 d_3 T_{\text{out}} \tag{14}$$

When Eq. (14) is multiplied with L, it is valid for a probe of length L. T_{out} is the arithmetic mean of the water temperature near sensor (at the probe tip) and outlet water temperature. Hence the problem reduces to obtaining the outlet water temperature and water temperature near sensor.

Recalling equations from Eqs. (8)–(10), one can repeat the approach demonstrated. When the heat balance at the inner side is written, the resultant equations are

$$Q_2 = \frac{\pi L(T_{w\,\text{out}} - T_{w\,\text{in}})}{\dfrac{1}{h_2 d_2} + \dfrac{1}{h_1 d_1} + \dfrac{\ln(d_2/d_1)}{2k_i}} \tag{15}$$

$$Q_5 = \frac{\pi L(T_g - T_{w\,\text{out}})}{\dfrac{1}{h_3 d_3} + \dfrac{1}{h_g d_5} + \dfrac{\ln(d_5/d_4)}{2k_{\text{TBC}}} + \dfrac{\ln(d_4/d_3)}{2k_o}} \tag{16}$$

The problem is solved in an iterative manner, starting with an initial T_{ws} guess. With the knowledge of coolant mass flow rate, the outlet water temperature is calculated using Eq. (17) and the sensor region water temperature using Eq. (18).

$$T_{w\,\text{out}} = T_{ws} + \frac{(Q_5 - Q_2)}{\dot{m}_w C_{p_{w\,\text{out}}}} \tag{17}$$

$$T_{ws} = T_{w\,\text{in}} + \frac{Q_2}{\dot{m}_w C_{p_{w\,\text{in}}}} \tag{18}$$

The radiation heat transfer is not taken into account in the 1D model. Calculation of radiation heat transfer requires the knowledge of emissivities of both the probe surface and the hot walls, and a view factor. However, using some strong assumptions, an order-of-magnitude analysis was carried out and radiation heat transfer is found to be confined within one-fifth of convection heat transfer. It is worth mentioning that this analysis is carried out assuming the probe is not subject to flames, which would greatly increase radiation amount.

Although the definition given above assumes the probe to be

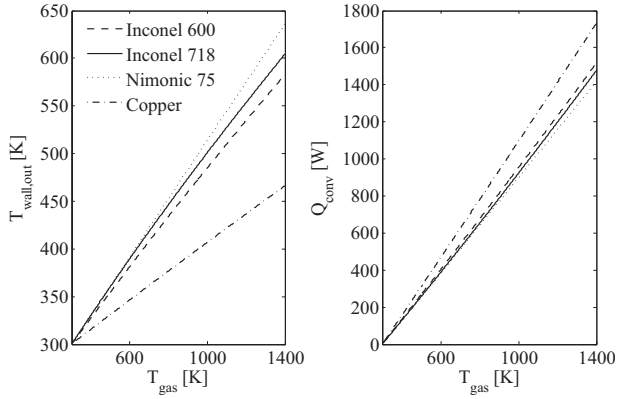


Fig. 8 Effect of probe material on probe outer wall temperature and heat to be extracted as a function of gas temperature

constructed of a single zone subject to uniform external conditions, multiple sector capability has been added, which enables the calculation of the probe stem effects and parts of the probe subject to different conditions, i.e., simulating the probe tip immersed into the gas path (e.g., at 1400 K) and the remaining probe length traversing the hot core engine casing (e.g., at 53 0K).

This simple prediction tool permitted the authors to perform various parametric studies. On the cooling side, the effect of coolant mass flow rate and coolant flow direction was investigated (i.e., cold inlet through the inner tube or through the outer annular jacket), leading to the conclusion that the coolant is best fed through the inner tube and that cooling rates of the order of 2 l/min are required.

Different typical high temperature probe materials were investigated. The results for Inconel 600, Inconel 718, and Nimonic 75 are presented in Fig. 8 and compared with copper, which has essentially no mechanical strength at high temperature but for the sole purpose of assessing the effect of a thermal conductivity, which is about 20 times higher than most high temperature alloys. This effect is best seen on the outer probe wall temperature, which is about a hundred degrees lower for copper than for Inconel 600, which has the highest thermal conductivity among the high temperature alloys investigated. The better cooling of the probe wall not only results in a higher heat power extracted but also in higher water temperatures, with the inherent risk of boiling.

On the hot gas side, Fig. 9 shows the effect of immersion depth, thermal barrier coating, heat exchange within the probe stem, and heat power to be extracted as a function of gas temperature. The reference case corresponds to an 8 mm diameter Inconel 600 probe, having a thermal barrier coating layer of 0.3 mm, immersed by 27 mm into a Ma=0.5 hot gas stream at 1400 K, and 40 bars with a turbulence intensity of 15%. The probe is water cooled with a flow rate of 2 l/min. Adding the influence of heat exchange inside the probe stem, which is 300 mm long, a slight increase in heat extracted and probe outer wall temperature is observed. The presence of the TBC greatly decreases the heat transfer through the probe from the hot gas stream. The probe surface temperature is reduced from 1070 K to 570 K through the TBC layer. Without TBC, and for 85 mm immersion depth, it may be required to extract as much as 16 kW from the probe. The probe length does not have a major influence on the surface temperature but rather on the heat extracted.

7.2 2D Navier–Stokes Computations. The objective of the numerical studies presented hereafter was not only to verify the cooling performance estimated by the analytical model but also to identify possible fluid stagnation or recirculation zones, which might cause local overheating of the probe, and finally to obtain a more detailed description of the flow field and temperature distributions around the sensor, which was not modeled as such in the analytical model.

The computations were carried out using FLUENT, using the standard $k-\varepsilon$ model, whose robustness, economy, and reasonable accuracy for a wide range of turbulent flows explain its popularity in industrial flow and heat transfer simulations. Second order upwind discretization schemes have been used to solve the momentum and energy equation.

To model the sensor cooling zone a simple 2D axisymmetric geometry has been developed, as shown in Fig. 10 with the sensor axis of symmetry perpendicular to the probe axis. The external flow field (hot gas side) was not modeled and the most critical

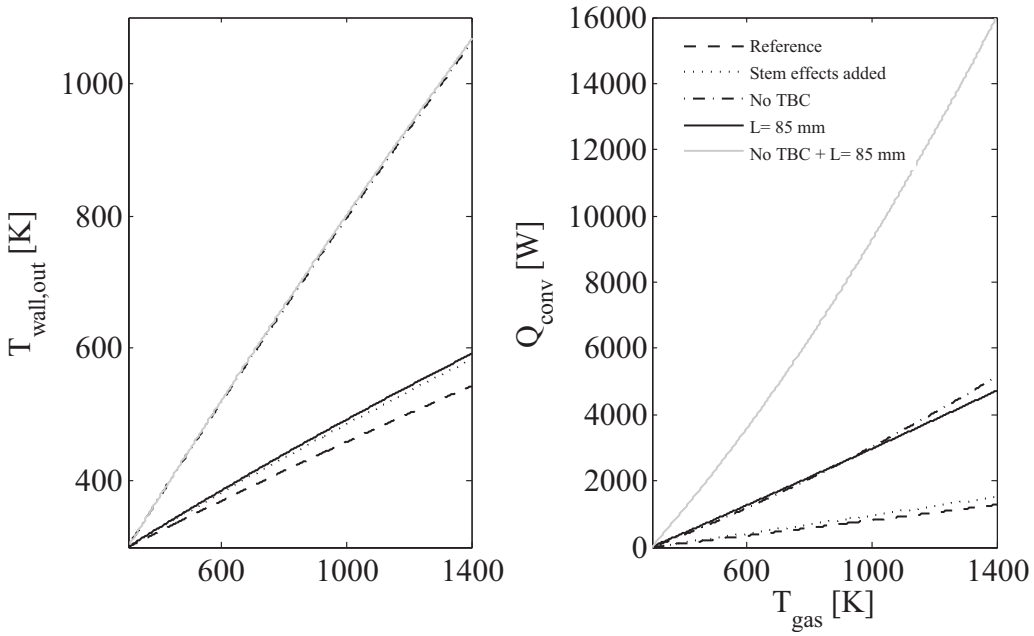


Fig. 9 Effect of immersion depth, thermal barrier coating, core engine casing temperature on the probe outer wall temperature, and heat to be extracted as a function of gas temperature (P_o =40 bars, Ma=0.5, and Tu=0.15)

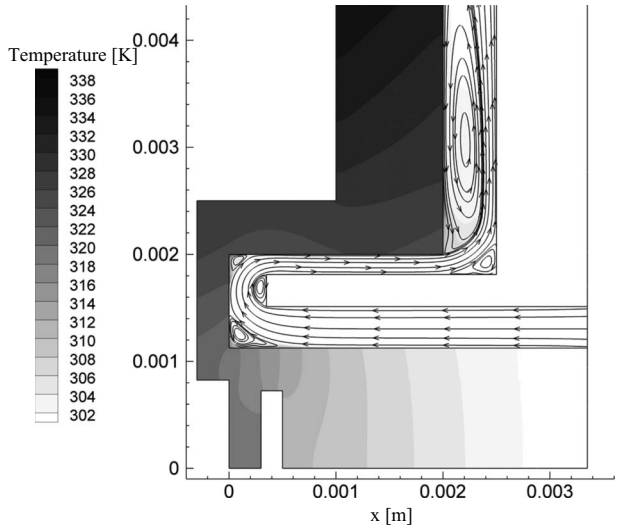


Fig. 10 2D axisymmetric model of the sensor region, temperature contour plot, and stream traces, T_g=1400 K, Ma=0.5, P_o=1 bar, h_g=2600 W/m^2 K, and water flow rate=2 l/min

boundary condition is therefore to determine the convective heat transfer coefficient. An iterative procedure was followed starting from a representative first guess of the probe outer wall temperature (from the analytical model) and imposing the corresponding heat transfer coefficient, which yields a new surface temperature after computation, used to recalculate a new heat transfer coefficient until convergence. This procedure can be done quickly; thanks to the fact that the 2D axisymmetric model requires very limited computational time. However, since the geometry is no longer representative of a cylinder in a cross-flow, Eq. (5) can be used properly only modifying the constants C, m, and n in such a way to be representative of a cylindrical plate perpendicular to a cross-flow [29]. The results in terms of temperature contour plots and stream traces show clearly that the temperature levels around

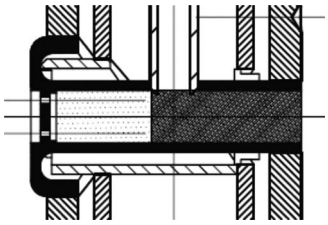


Fig. 11 Design modifications on the screen and the sensor cooling channel

the sensor region remain low; the sensor itself does not experience a temperature higher than 313 K. Different recirculation zones do not seem to affect substantially the temperature distribution. However, the final design was modified accordingly with fillet radii inside the return channel instead of sharp corners as shown in Fig. 11.

7.3 3D Navier–Stokes Computations. The full 3D probe geometry was developed modeling both fluid and solid zones and solving also the energy equation. The whole computational domain was split into 112 volumes to allow the proper meshing of the different regions of the probe tube and of the sensor region. In order to reduce the complexity of the simulation, the external flow field was not meshed and the same boundary conditions were applied as described above but with a fixed value of the heat transfer coefficient calculated by the correlation of Lowery and Vachon given in Eq. (3) according to the flow conditions.

Structured and unstructured grids have been used depending on the complexity of the geometry; small modifications of the fluid zones previously meshed have been done to allow gridding also of the solid regions. Almost 1,496,000 grid cells have been used to simulate the whole probe geometry.

Figure 12 shows the temperature contour plots on the solid region around the sensor. In both cases (water from the inner tube and from the annular jacket), the overall temperature levels remain lower than 380 K and the sensor never experiences a temperature higher than 340 K; the maximum temperature area is

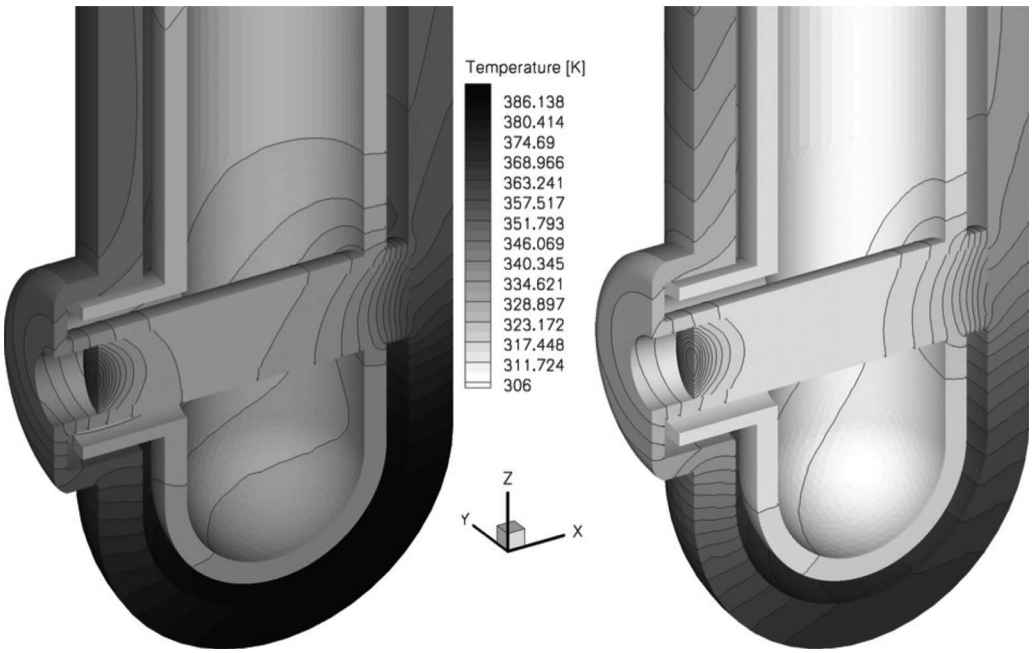


Fig. 12 Temperature distribution inside the probe walls, water inlet from the annular jacket (left), and water inlet from the inner tube (right), T_g=1400 K, Ma=0.5, P_o=1 bar, h_g=2600 W/m^2 K, and water flow rate=2 l/min

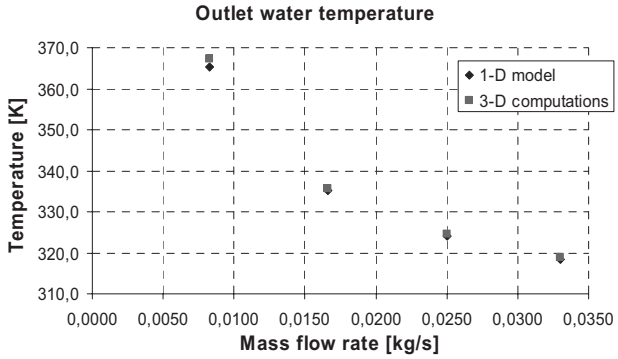


Fig. 13 Probe surface temperature as a function of mass flow rate, T_g=1400 K, Ma=0.5, and P_o=1 bar

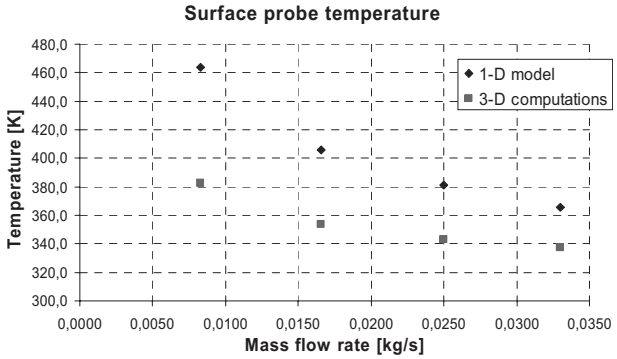


Fig. 15 Water outlet temperature as a function of mass flow rate, T_g=1400 K, Ma=0.5, and P_o=1 bar

found to be at the outer annulus at the bottom of the probe. The cooling system in which water is fed from the internal cylinder is computed to have a lower overall temperature value than the case of coolant fed from the external jacket. The heat flux transferred to the fluid along the inner tube cylinder is lower than that one exchanged with the fluid along the external jacket, directly in contact with the hot gas side. Therefore, for the case in which water feeds the probe from the inner tube, the sensor experiences a temperature lower than 310 K.

7.4 Comparison with the 1D model. The results obtained from the 1D model and from the computed full 3D geometry are compared in terms of global cooling parameters. To this end, the 3D results have been postprocessed as follows. The water temperature at the outlet of the computational domain has been computed using a mass flow weighted average. The water temperature approaching the sensor has been computed as a mass flow weighted averaged temperature in a horizontal plane located just before the sensor region. The overall heat flux and the probe surface temperature are computed as integral values on the external probe tube surface. Results have been compared for the case with hot gas freestream temperature and Mach number of 1400 K and 0.5, respectively.

Figures 13–16 compare the results obtained for different mass flow rates between 0.5 and 2.0 l/min. General agreement is found for the water temperature at the outlet shown in Fig. 15 and for the water temperature approaching the sensor shown in Fig. 14. The 1D model computes generally higher values of the external probe surface than the 3D geometry calculation. This is related to a higher heat flux calculated for the 3D geometry case than the 1D model. It is attributed to the fact that the 1D analysis models the probe as a tube without considering the cooling effects at the sensor region. In the 3D geometry, heat can pass from the probe surface to the sensor, which is like a cylinder in a water cross-flow incoming from the inner cylinder. In this way, around the sensor, the coolant can extract more heat and consequently the values computed for the 3D geometry were found to be higher than those calculated in the 1D model. Higher heat flux means higher outlet water temperature and lower external surface probe temperature as seen in Fig. 13.

Overall, the comparison indicates that the 1D model is able to predict well the global cooling parameters, with a slight overestimation, which might be used as a safety factor.

7.5 Internal Cooling Optimization. Considering the flow pattern inside the original probe design, it can be noticed that the cooling of the probe tip suffers from the recirculation inside the inner tube and the poor circulation in the space between the two hemispherical parts as shown in Fig. 17.

In order to reduce the overheating of the probe tip, and among different solutions investigated, it was found that creating a hole at the bottom of the inner hemispherical part would enhance the cooling of the tip by impingement, which is shown to reduce the temperature by approximately 40 K (Fig. 18). Regarding the flow pattern, some recirculation is still present and unavoidable (Fig. 19) but the tip cooling has been greatly improved without compromising the main cooling flow around the sensor.

To conclude on three-dimensional simulations of the probe concept, the effect of probe material is compared between copper and Inconel 600, for a probe immersed by 85 mm, without any TBC layer in an environment representative of a modern high pressure turbine, i.e., T_g=1400 K, Ma=0.5, and P_o=40 bars (Fig. 20). It is obvious that the thermal gradient across the probe wall is much higher for Inconel 600 due to the low value of thermal conductivity (14.9 W/m K at 300 K and 27.5 W/m K at 1070 K) against copper (>400 W/m K). The values of temperature obtained stay below 800 K for copper, where they exceed 1000 K in the case of

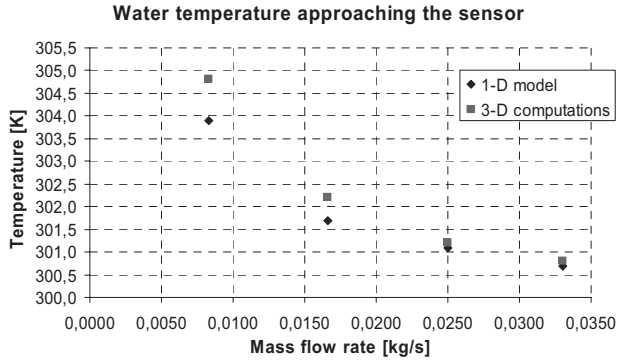


Fig. 14 Water temperature approaching the sensor as a function of mass flow rate, T_g=1400 K, Ma=0.5, and P_o=1 bar

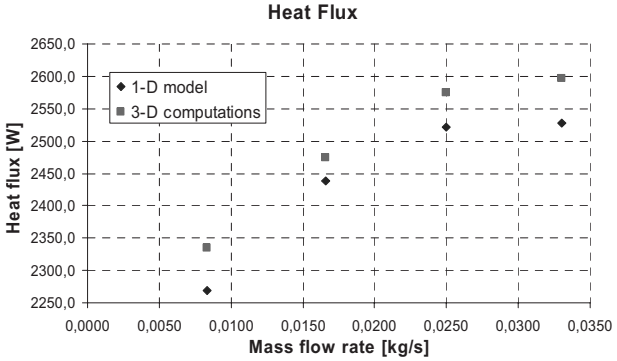


Fig. 16 Evacuated heat as a function of mass flow rate, T_g =1400 K, Ma=0.5, and P_o=1 bar

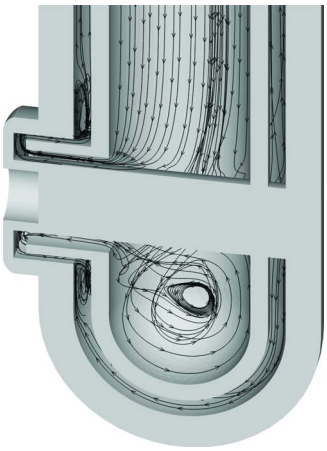


Fig. 17 Streamtraces in original probe design, T_g=1400 K, Ma=0.5, and P_o=1 bar

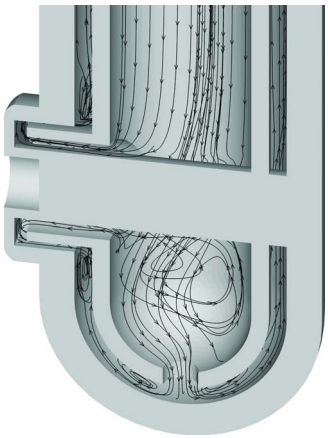


Fig. 19 Streamtraces in modified probe design, T_g=1400 K, Ma=0.5, and P_o=1 bar

Inconel 600, with a screen temperature around 700 K. As already stated earlier, the sensor diaphragm is able to withstand a maximum continuous temperature of 780 K. However, Inconel 600 is preferred because of its better mechanical properties at high temperatures. On the water side, calculations show that the water temperature locally reaches 460 K at small recirculation zones around the sharp corners near the sensor. Therefore it was decided to pressurize the coolant up to a pressure of 50 bars, at which the water boiling temperature is near 560 K. Pressurizing the probe internally also has the advantage to be able to equalize internal and external pressures depending on the environment and therefore helps minimize the structural stresses at high temperatures.

7.6 Mechanical Design Considerations. The probe is first modeled as a pressure vessel using AD-Merkblatt factsheets [37], which were employed to certify design of pressure vessels. Investigated cases include a pressurized probe (at 50 bars) in a vacuum environment and a probe under vacuum in a pressurized environment (at 50 bars).

Thermal stresses around the probe are modeled using local Nusselt number relations with respect to Reynolds number from Merker [38] (Fig. 21). Local maxima and minima corresponding to the actual Reynolds number are applied disregarding the distribution but as a discontinuity. The corresponding thermal expansion of the material is used as a strain to obtain the corresponding stress and is found to be below the yield strength of the material at the elevated temperature. The calculated thermal stress is approximately 40% of the yield strength, where at the location where thermal stress is a consideration, the von Mises stress calculated using a finite-element-method solver is found to be one order of magnitude less than the yield stress. It shall be noted that the maximum thermal stress occurs at the probe tip where the sensor is located whereas the maximum bending stresses are at the probe fixtures due to the cantilever arrangement of the probe in the traverse gear.

The structural design of the probe is assessed using a three-dimensional finite-element solver for real engine environment with respect to different immersion lengths. With the current traverse gear, which supports the probe at a location away from the engine outer casing, maximum immersion depth before yield

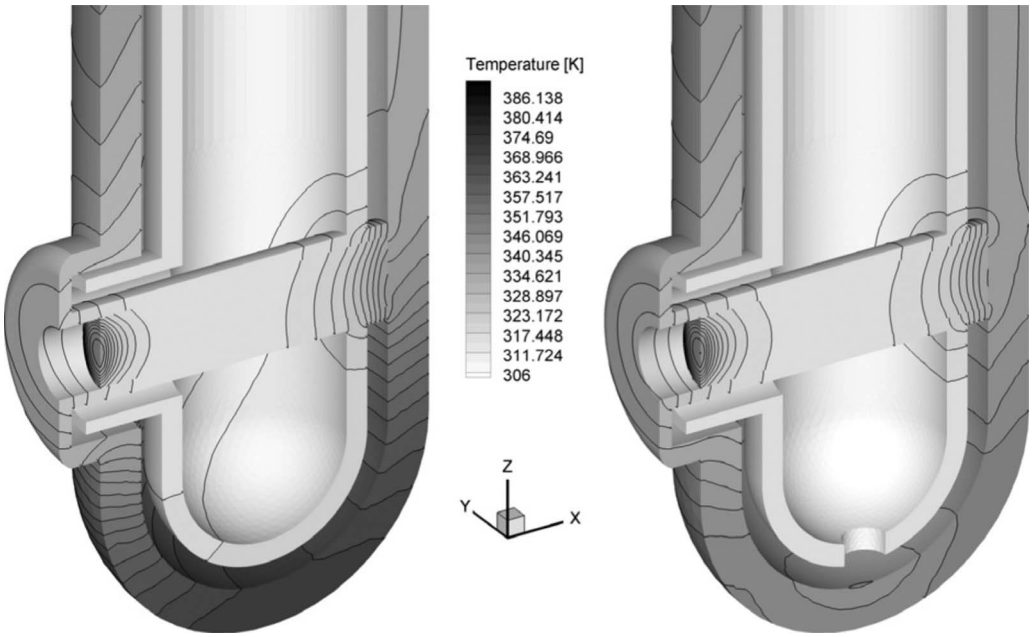


Fig. 18 Temperature distribution in original and modified probe designs, T_g=1400 K, Ma=0.5, and P_o=1 bar

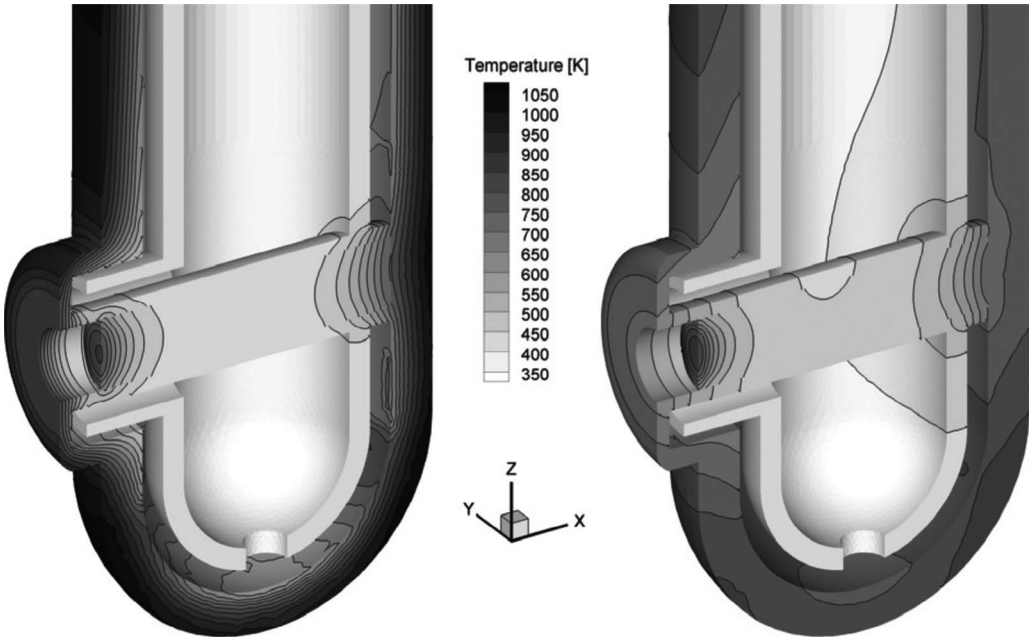


Fig. 20 Temperature distribution in modified probe design, Inconel 600 (left), copper (right), T_g =1400 K, Ma=0.5, and P_o=40 bars

is found to be around 70 mm. However, this value may be greatly enhanced using a different traverse gear, which may support the probe nearer to the casing. In Fig. 22, the maximum von Mises stress and safety factor are plotted versus immersion depth. The yield strength in all the calculations are assumed to be of an annealed bar, which is a safe approximation as the probe will undergo a hot finishing process.

8 Perspectives

Several prototypes of the presented probe design are currently being manufactured. It is planned to thoroughly validate this design before any real engine testing. A hydraulic test to ensure the structural integrity of the probe under coolant pressure of 50 bars will be carried out. First laboratory tests will then be performed in the VKI plasma-torch facility where very high temperatures ($>$2000 K) can be reached at low pressure levels in the plasma jet axis. Those first laboratory tests will also be used for assessing the thermal behavior of the sensor. It is also envisaged to apply temperature-sensitive-paint to provide experimental evidence of the actual probe surface temperature distribution for comparison with the computational fluid dynamics (CFD) data. Following this experimental validation, combustion rig and gas-turbine tests will be carried out, starting with jet-pipe measurements immediately downstream of the turbine rotor.

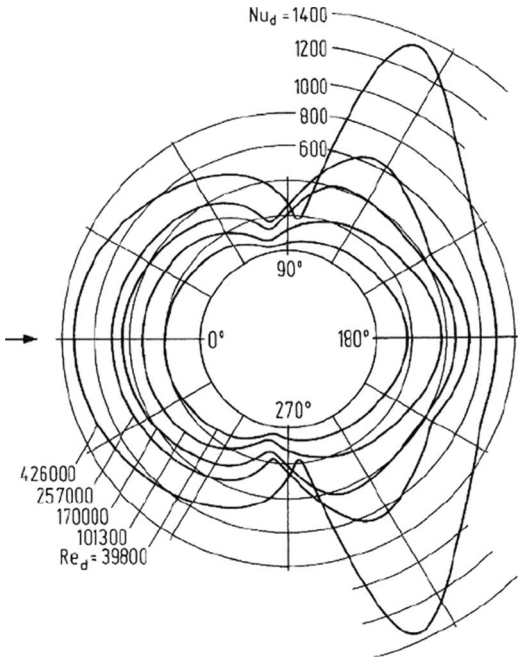


Fig. 21 Nusselt number distributions around a circular cylinder for various Reynolds numbers (Ref. [38])

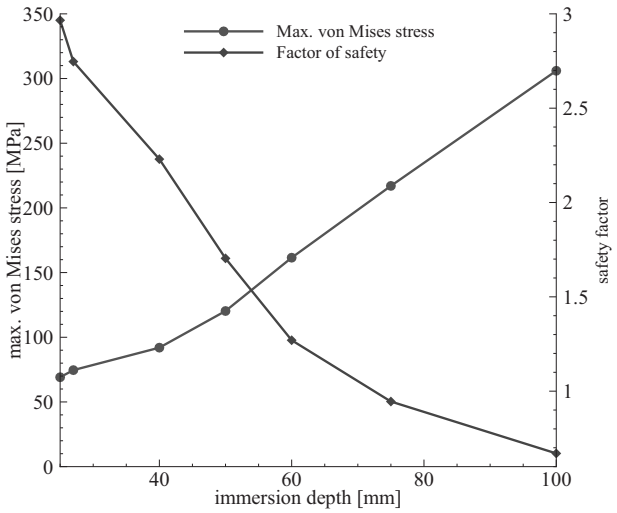


Fig. 22 von Mises stresses and safety factor as a function of immersion depth, T_g=1400 K, Ma=0.5, and P_o=1 bar

9 Conclusions

This paper first describes the predesign one-dimensional calculations, which were carried out to assess the cooling performance required in terms of heat transfer as a function of the environmental conditions.

A first water-cooled design concept is then described, mainly dictated by the objective of keeping a minimal probe size and by the cooling requirements. All efforts were concentrated to achieve the most compact design, which results in an 8 mm diameter probe tip (incorporating the sensor).

The one-dimensional model developed within this study allows one to carry out parametric studies to optimize the design of a probe tailored for a given application or to assess the behavior of a given probe geometry under different flow conditions.

2D and 3D Navier–Stokes computations were carried out to determine the optimum cooling configuration, in particular, to detect flow separation or stagnation regions causing possibly a local heating of the probe or sensor. These simulations have motivated some modifications of the probe design to improve the cooling flow pattern.

The mechanical aspects of the design are considered by calculating bending stresses and thermal stresses according to the prescribed probe geometry and material properties. The mechanical design is supported by comparison to finite-element model (FEM) calculations.

Regarding the suitability of the design, not only thermal and mechanical but also aerodynamic aspects have to be considered. In this respect, minimal probe size (8 mm diameter) for good spatial resolution and high bandwidth ($\sim$40 kHz) for good time-resolution have been the main concern in this design. In order to allow for the recording of continuous time-series, and therefore to immerse the probe continuously into the hot gas stream, the internal cooling was optimized to avoid any recirculation zones and water boiling temperature was increased to 560 K by pressurization up to 50 bars. Particular attention was given to the cooling of the perforated screen ahead of the sensor in order to restrict the heating of the sensor diaphragm, which can withstand a maximum temperature of 780 K.

Acknowledgment

This work is supported by the EU Framework Program 6 project HEATTOP (Accurate High Temperature Engine Aero-Thermal Measurements for Gas Turbine Life Optimization Performance and Condition Monitoring) under Contract No. AST5-CT-2006-030696. HEATTOP is a new collaborative research project, which aims to develop new, more durable, more accurate instrumentation for high temperature gas-turbine applications. It has a total budget of 9 million Euros, including 5 million Euros in funding from the European Commission. Siemens PG leads a consortium of 17 partners gathering major European engine manufacturers, supply chain companies, universities, and research centers. The work in this paper has been performed under WP4 "Advanced Gas Path Measurements."

Nomenclature

Roman Symbols

C = coefficient in Eq. (5)
C_p = specific heat at constant pressure, J/(kg K)
d = diameter, m
dq = incremental heat flow rate, W
dx = incremental length along the probe, m
h = convection heat transfer coefficient, W/(m^2 K)
k = conductivity, W/(m K)
l/d = length to diameter ratio
L = probe length, m
m = exponent in Eq. (5)
Ma = Mach number
n = exponent in Eq. (5)
Nu = Nusselt number
P = pressure, Pa
Pr = Prandtl number
Q = heat flow rate, W
Re = Reynolds number
T = temperature, K
Tu = turbulence intensity

Greek Symbols

α = dummy variable
β = dummy variable

Subscripts

1 = inner surface of the inner wall
2 = outer surface of the inner wall
3 = inner surface, outer wall
4 = TBC-probe interface
5 = outer surface of TBC
d = based on diameter
g = gas
i = at the inner wall
in = inlet
j = index defining location
o = at the outer wall
out = outlet
p = surface
s = sensor
TBC = at thermal barrier coating
w = water

Superscripts

$\dot{x}$ = time rate
$\bar{x}$ = averaged

References

[1] EVI-GTI, "The Lab Gap Matrix," on the www, at http://www.evi-gti.com
[2] Simon, D. L., Garg, S., Hunter, G. W., Guo, T.-H., and Semega, K. J., 2004, "Sensor Needs for Control and Health Management of Intelligent Aircraft Engines," ASME Paper No. GT2004-54324.
[3] Kupferschmied, P., Koppel, P., Gizzi, W., Roduner, C., and Gyamarthy, G., 2000, "Time-Resolved Flow Measurements With Fast-Response Aerodynamic Probes in Turbomachines," Meas. Sci. Technol., **11**(7), pp. 1036–1054.
[4] Brouckaert, J. F., 2007, "Fast Response Aerodynamic Probes for Measurements in Turbomachines," Proc. Inst. Mech. Eng., Part A, Professional Engineering Publishing, London, Vol. 221(6), pp. 811–813. Paper No. VKI RP 2007-26.
[5] Kurtz, A. D., and Ned, A. A., 2004, "Ultra High Temperature, Miniature, SOI Sensors for Extreme Environments," IMAPS International HITEC Conference, May 17–20, Santa Fe, NM.
[6] Ned, A. A., Kurtz, A. D., Beheim, G., Masheeb, F., and Stefanescu, S., 2004, "Improved SIC Leadless Pressure Sensors for High Temperature, Low and High Pressure Applications," Twenty-First Transducer Workshop, June 22–23, Lexington, MD.
[7] Okojie, R. S., Fralick, G. C., Saad, G. J., Blaha, C. A., Adamczyk, J. J., and Feiereisen, J. M., 2003, "A Single Crystal SIC Plug-and-Play High Temperature Drag Force Transducer," *TRANSDUCERS, The 12th International Conference on Solid-State Sensors, Actuators and Microsystems*, June 8–12, Boston, Vol. 1, pp. 400–403.
[8] Ferguson, D. G., and Ivey, P. C., 1995, "Unsteady Pressure Measurement in a High Temperature Environment Using Water Cooled Fast Response Pressure Transducers," ASME Paper No. 95-GT-345.
[9] Ferguson, D. G., and Ivey, P. C., 1998, "A High-Temperature Assessment of Air-Cooled Unsteady Pressure Transducers," ASME J. Turbomach., **120**(3), pp. 608–612.
[10] Ivey, P. C., and Ferguson, D. G., 2002, "An Air Cooled Jacket Designed to Protect Unsteady Pressure Transducers at Elevated Temperatures in Gas Turbine Engines," *Proceedings of ASME Turbo Expo 2002*, June 3–6, Amsterdam, The Netherlands.
[11] Langmuir, I., 1923, "The Pressure Effect and Other Phenomena in Gaseous Discharges," J. Franklin Inst., **196**, pp. 751–762.
[12] Grey, J., 1965, "Thermodynamic Methods of High Temperature Measurement," ISA Trans., **4**, pp. 102–115.
[13] Grey, J., 1970, "Probe Measurements in High-Temperature Gases and Dense Plasmas," *AGARDograph on 'Measurement Techniques in Heat Transfer'*, Vol. 130, E. R. G. Eckert and R. J. Goldstein, eds., NATO AGARD, Slough, England, pp. 269–307.

[14] Ashby, G. C., Jr., 1988, "Pressure Measuring Probe," U. S. Patent No. 4783994.
[15] Lagen, N. T., Reece, G. D., Eves, J. W., and Geissinger, S. L., 1991, "Water Cooled Static Pressure Probe," U.S. Patent No. 5076103.
[16] Warnack, D., 2002, "An Optimized Pneumatic Probe for Investigation of Gas Turbine Aerodynamics in Full Scale Gas Turbines," *Proceedings of ASME Turbo Expo 2002*, June 3–6, Amsterdam, The Netherlands.
[17] Yang, X. L., Miller, R., and Hodson, H. P., 2002, "A New Probe for the Measurement of High-Temperature Gases," XVIth Bi-Annual Symposium on Measuring Techniques in Transonic and Supersonic Flow in Cascades and Turbomachines, Sept. 23–24, Cambridge, UK.
[18] Nieberding, W. C., and Englund, D. R., Jr., 1972, "Modularized Instrument System for Turbojet Engine Test Facilities," *Symposium on Instrumentation for Airbreathing Propulsion, Progress in Astronautics and Aeronautics*, A. Fuhs and M. Kingery, eds., MIT Press, Cambridge, MA, Vol. 34, pp. 206–215, Report No. NASA-TM-X-68123.
[19] Moore, M. T., 1977, "High Temperature-High Response Fluctuating Pressure Measurements," *ASSL: Infrared and Submillimeter Astronomy*, Vol. 63, G. G. Fazio, ed., D. Reidel Publishing Co., Dordrecht, The Netherlands, pp. 469–473.
[20] Larguier, R., 1985, "Experimental Analysis Methods for Unsteady Flows in Turbomachines," *Proceedings of the Seventh International Symposium on Air Breathing Engines, Proceedings (A86-11601 02-07)*, Beijing, P.R.C., Sept. 2–6, AIAA, New York, pp. 279–289.
[21] Passaro, A., LaGraff, J. E., Oldfield, M. L. G., and Biagioni, L., 2006, "Fast Acting Probe for Measurement of Turbulent Pressure and Temperature Fluctuations in a Gas Turbine Combustor," 44th AIAA Aerospace Sciences Meeting and Exhibit, Jan. 9–12, Reno, NV, AIAA Paper No. AIAA-2006-550.
[22] Brouckaert, J. F., 2000, "Development of Single- and Multi-Hole Fast Response Pressure Probes for Turbomachinery Applications," 15th Bi-Annual Symposium on Measurement Techniques in Transonic and Supersonic Flow in Cascades and Turbomachines, Sept. 21–22, Florence, Italy.
[23] Morgan, V. T., 1975, "The Overall Convective Heat Transfer From Smooth Circular Cylinders," *Advances in Heat Transfer (A76-17076 05-34)*, Vol. 11, Academic, New York, pp. 199–264.
[24] Moss, R. W., and Oldfield, M. L. G., 1992, "Measurements of the Effect of Free-Stream Turbulence Length Scale on Heat Transfer," ASME Paper No. 92-GT-244.
[25] Smith, M. C., and Kuethe, A. M., 1966, "Effects of Turbulence on Laminar Skin Friction and Heat Transfer," Phys. Fluids, **9**(12), pp. 2337–2344.
[26] Kestin, J., and Wood, R. T., 1971, "The Influence of Turbulence on Mass Transfer From Cylinders," ASME J. Heat Transfer, **93**, pp. 321–327.
[27] Lowery, G. W., and Vachon, R. I., 1975, "The Effect of Turbulence on Heat Transfer From Heated Cylinders," Int. J. Heat Mass Transfer, **18**, pp. 1229–1242.
[28] Zukauskas, A., and Ziugzda, J., 1985. *Heat Transfer of a Cylinder in Crossflow*, Hemisphere, Washington, DC, p. 219.
[29] Zukauskas, A., 1972, "Heat Transfer From Tubes in Crossflow," Adv. Heat Transfer, **8**, pp. 93–160.
[30] Gossweiler, C., 1993, "Sonden und Messsystemen Fur Schnelle Aerodynamische Strommungsmessung mit Piezoresistiven Druckgebern," Ph.D. thesis, The Swiss Federal Institute of Technology Zurich (ETH Zurich), Zurich, Switzerland.
[31] Brouckaert, J. F., 2002, "Development of Fast Response Aerodynamic Pressure Probes for Time-Resolved Measurements in Turbomachines," Ph.D. thesis, Université Libre de Bruxelles/von Karman Institute for Fluid Dynamics, Sint Genesius-Rode, Belgium.
[32] Epstein, A. H., 1985, "High Frequency Response Measurements in Turbomachines," *VKI Lecture Series on Measurement Techniques in Turbomachines, VKI LS 1985–03*, Feb., Rhode-St-Genese, Belgium.
[33] Larguier, R., 1981, "Experimental Analysis Methods for Unsteady Flows in Turbomachines," J. Eng. Power, **103**, pp. 415–423.
[34] Sieverding, C. H., Arts, T., and Dénos, R., 1995, "Annular Cacade Tests," *Advances in Engines Technology*, R. Dunker, ed., Wiley, New York.
[35] Bergh, H., and Tijdeman, H., 1965, "Theoretical and Experimental Results for the Dynamic Response of Pressure Measuring Systems," NLR, Report No. NLR-TR-F.238.
[36] Dittus, F. W., and Boelter, L. M. K., 1930, *Publications on Engineering*, Vol. 2, University of California, Berkeley, CA.
[37] der Technischen Uberwachungs-Vereine (VdT UV), V., 1995, *Arbeitsgemeinschaft Druckbeh Alter Merkblatt B-Series (Design)*, 1995 ed., Carl Heymanns, Cologne, Germany.
[38] Merker, G. P., 1987, *Konvektive Warmeubertragung*, Springer-Verlag, Berlin.

Performance and Cost Analysis of Advanced Gas Turbine Cycles With Precombustion CO_2 Capture

Stéphanie Hoffmann
Michael Bartlett
Matthias Finkenrath

General Electric,
Global Research Center,
D-85748 Garching Bei München, Germany

Andrei Evulet
General Electric,
Global Research Center,
Niskayuna, NY 12309

Tord Peter Ursin
StatoilHydro,
Technology and New Energy,
N-4035 Stavanger, Norway

This paper presents the results of an evaluation of advanced combined cycle gas turbine plants with precombustion capture of CO_2 from natural gas. In particular, the designs are carried out with the objectives of high efficiency, low capital cost, and low emissions of carbon dioxide to the atmosphere. The novel cycles introduced in this paper are comprised of a high-pressure syngas generation island, in which an air-blown partial oxidation reformer is used to generate syngas from natural gas, and a power island, in which a CO_2-lean syngas is burnt in a large frame machine. In order to reduce the efficiency penalty of natural gas reforming, a significant effort is spent evaluating and optimizing alternatives to recover the heat released during the process. CO_2 is removed from the shifted syngas using either CO_2 absorbing solvents or a CO_2 membrane. CO_2 separation membranes, in particular, have the potential for considerable cost or energy savings compared with conventional solvent-based separation and benefit from the high-pressure level of the syngas generation island. A feasibility analysis and a cycle performance evaluation are carried out for large frame gas turbines such as the 9FB. Both short-term and long-term solutions have been investigated. An analysis of the cost of CO_2 avoided is presented, including an evaluation of the cost of modifying the combined cycle due to CO_2 separation. The paper describes a power plant reaching the performance targets of 50% net cycle efficiency and 80% CO_2 capture, as well as the cost target of 30$ per ton of CO_2 avoided (2006 Q1 basis). This paper indicates a development path to this power plant that minimizes technical risks by incremental implementation of new technology.
[DOI: 10.1115/1.2982147]

1 Introduction and Objectives

Global warming has been identified as one of the most substantial environmental challenges of the 21st century. As a consequence, the energy industry is facing increasing regulations and governmental pressure regarding carbon dioxide (CO_2) emissions. This is the motivation for significant research efforts in CO_2 capture and sequestration from fossil fueled power plants [1]. In particular, many gas turbine based power plants with integrated CO_2 capture have been proposed in the literature, and an extensive comparative analysis of their performance can be found in Ref. [2]. The understanding of the costs involved remain however limited.

Technical solutions for integrated CO_2 capture at an "affordable" cost both for coal-fired and natural gas fired power plants are needed. The focus of the paper is on natural gas power plants, and the designs are carried out with the objectives of high efficiency, low capital cost, and a low quantity of carbon dioxide emitted to the atmosphere. More specifically, as in Refs. [3,4], the design goals are above 50% cycle net efficiency, below 30 $\$/t_{CO_2}$ cost of CO_2 avoided, and capture of at least 80% of the CO_2 generated. This paper reports results from a study performed in 2006 Q1 without inflation or market adjustments of the costs estimated at that time.

2 Background

Advanced power plant concepts with integrated precombustion CO_2 capture present several advantages. First, if compared with state-of-the-art atmospheric postcombustion capture, the increased CO_2 partial pressure in the generated syngas improves the driving force for CO_2 separation, thereby reducing the cost of the CO_2 separation equipment. Furthermore, some of the technical concepts that will be developed for the precombustion CO_2 capture from coal, i.e., in an integrated gasification combined cycle (IGCC) plant, can, in part, be applied to future natural gas plants as well. Additionally, fuel decarbonization is a platform allowing incremental technology improvements that provide competitive solutions to CO_2 capture, both in the near-term and long-term. Finally, the cogeneration of fuels and chemicals is possible via the precombustion pathway.

Key to precombustion capture concepts is the choice of reforming technology. Partial oxidation (POX) technology avoids requirements for expensive high temperature catalysts along with associated fuel desulphurization process steps. Furthermore, POX reactors allow for higher CH_4 conversion rates at high-pressure compared with other reforming concepts such as autothermal reformers. Only POX reforming technology is considered in this paper.

The study is based on a 400 MW scale heavy-duty gas turbine. Today, when fired with natural gas, a frame 9FB combined cycle delivers approximately 410 MW net power at around 58% low heating value (LHV) cycle net efficiency.

Both solvent-based and membrane-based CO_2 capture have been investigated for the presented cycles. An amine-based plant provides a near-term solution for CO_2 separation from pressurized syngas streams. A medium to long-term cycle improvement is presented by use of CO_2 selective separation membranes [5].

3 Novel Cycle Configurations

Two main precombustion cycles have been investigated: a base line and an advanced concept. The base line concept represents a near-term plant with little development required, while the ad-

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 1, 2008, final manuscript received May 26, 2008; published online December 23, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008 Land, Sea and Air (Gt2008), Berlin, Germany, June 9–13, 2008.

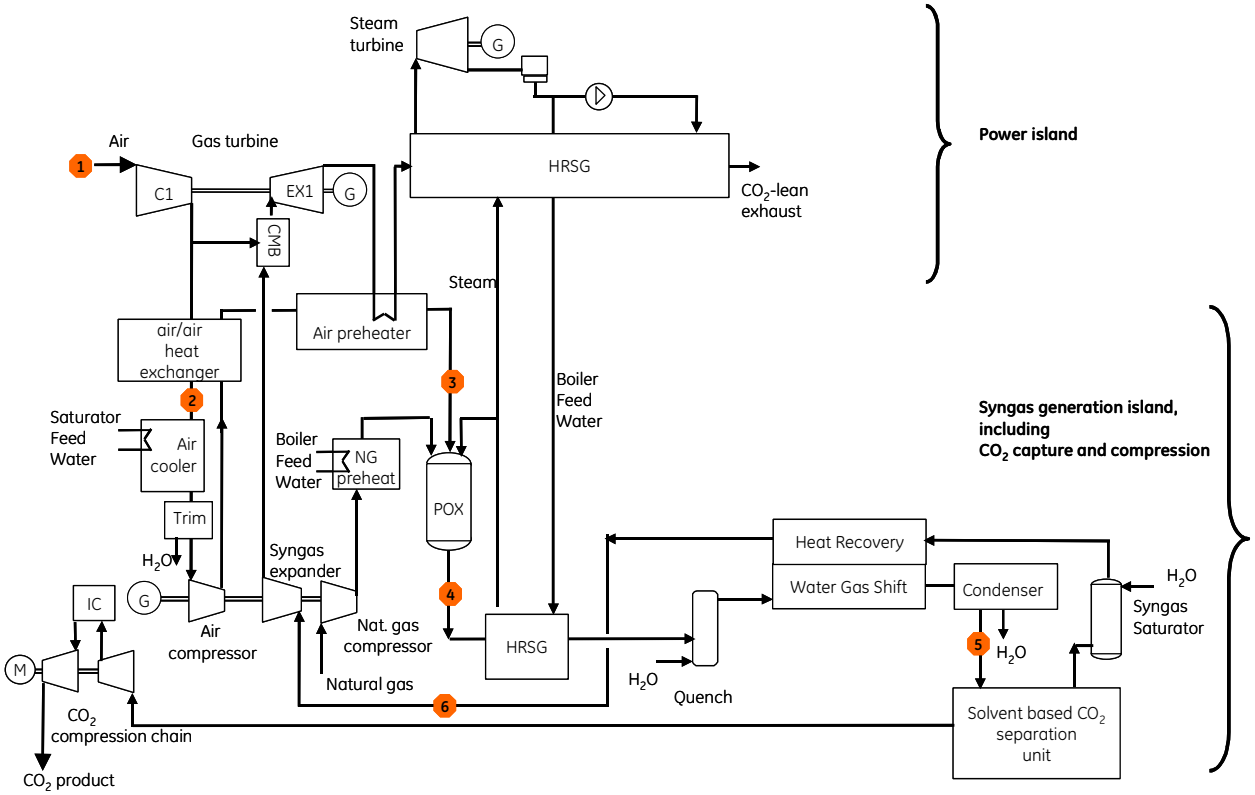


Fig. 1 Base line concept for precombustion CO_2 capture

vanced concept contains novel technology for heat recovery and CO_2 separation, and is considered longer term. By comparing the performance and cost of CO_2 avoided of these concepts, the potential benefit of an incremental addition of new technology on the cycle performance and cost of CO_2 avoided is seen.

The first cycle, denoted as the "base line" concept [6], contains mainly existing technology and is illustrated in Fig. 1. Some flow data are given in Table 1.

The base line concept consists of a high-pressure "syngas generation island," connected to a "power island" in which the syngas produced in the syngas generation island is fired in a large frame gas turbine. The exhaust heat of the gas turbine is used to produce steam for power in a steam-bottoming cycle. The syngas generation island is based on a high-pressure air-blown reformer, in which natural gas is converted to a hot syngas made essentially of N_2, CO, H_2, H_2O, and CO_2. The air feed to the POX reactor is extracted from the power island at the exit of the gas turbine compressor. This air feed is cooled down in various heat exchangers before being compressed further to reach the operating pressure of the POX reformer (around 100 bars). The pressurized air feed is then preheated before entering the POX reactor. The natural gas feed to the POX reactor is also compressed and preheated. The high exergy content of the syngas at the outlet of the POX reactor is partly recovered in a heat recovery steam generator (HRSG) producing saturated high-pressure steam that is sent to the high-pressure steam turbine of the power island. It then enters a quench reactor, which cleans, saturates, and cools the syngas, before passing to a water gas shift (WGS) reactor. There, the CO content of the syngas is converted to CO_2, which is required to allow for high CO_2 separation rates. After condensing the water content of the syngas in a condenser, the CO_2 is removed from the syngas in a solvent-based separation unit. The CO_2-lean syngas is then saturated with water in order to partly compensate for the missing CO_2 mass flow and hence to increase the power generated by the gas turbine in the power island. The saturated syngas is then preheated using some of the heat produced by the WGS reaction and finally expanded in a low temperature state-of-the-art syngas expander before being fired in the power island.

The second cycle, denoted as the "advanced" concept, contains additional elements of new technology and is illustrated in Fig. 2. Some flow data is given in Table 2. There are two main differences compared with the base line cycle. First, the syngas is generated in a novel piece of technology, the so-called syngas turbine (SynGT). Here, a new, radically compact POX reactor, under development [7], is used to fire an expander that drives an air compressor and generator, much like in an air-blown oxidizing gas turbine. In this way, the exergy content at the exit of the POX reactor is harvested directly by expanding the hot gases, and a compact simplified plant is achieved. This means that the subsequent equipment of the syngas island operates at a lower pressure than in the base line concept. The second difference to the base line concept is that the CO_2 separation is achieved using a CO_2-selective membrane.

Table 1 Flow data for the base line concept

Station	1	2	3	4	5	6
T (°C)	15	310	550	1150	40	340
P (bar)	1.013	~18	99	—	82.6	—
M (kg/s)	~640	115.0	114.4	137.2	144.6	107.9

4 Main Cycle Features

The core of the cycle performance optimization work has been done within the syngas generation island. Therefore we review here its main components and discuss technological risks.

4.1 Feed Air Heat Exchangers. The air extracted at the discharge of the 9FB compressor and sent to the syngas island needs to be cooled down and water condensed out before being compressed to the ~100 bar operating pressure of the POX reformer.

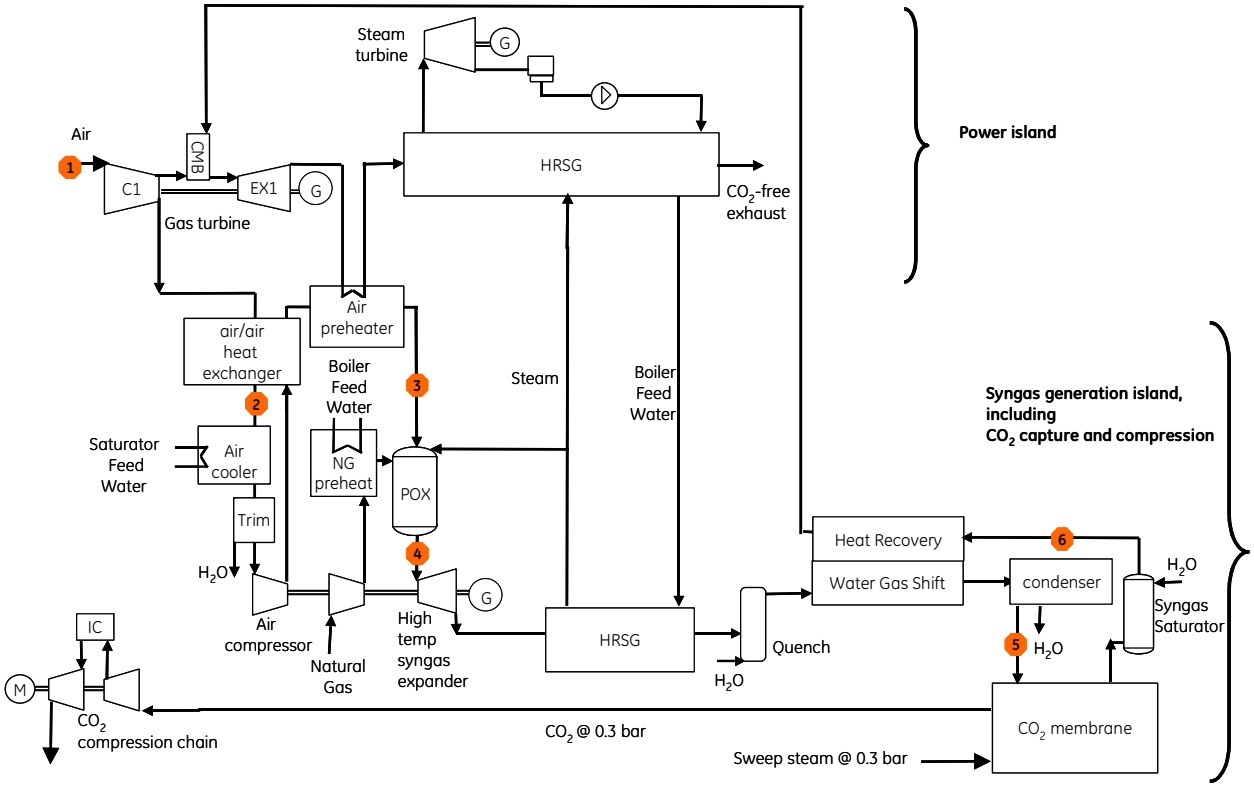


Fig. 2 Advanced concept for precombustion CO_2 capture

This is achieved using three heat exchangers successively and a water knockout drum. The first heat exchanger is assumed to be a shell and tube air/air heat exchanger that cools down the extracted air to about 310°C while preheating the high-pressure air to 375°C. The second heat exchanger is a fin tube air cooler that further cools down the extracted air to 70°C. The cooling medium is the circulating water circuit from the syngas saturator. The third heat exchanger is a trim cooler with a water knockout drum, in which water is condensed and removed.

4.2 Air and Natural Gas Preheaters. For an optimized performance of the syngas generation, the pressurized air feed to the POX reformer is preheated up to 550°C using some of the heat from the gas turbine exhaust in the power island. For safety reasons, the natural gas feed to the POX reactor is preheated using boiler feed water from the power island instead of gas turbine flue gas. To avoid coking issues, the natural gas preheat temperature is maintained below 230°C.

4.3 POX Reformer. The POX reformer is based on the well-known GE gasifier technology. A burner in a large refractory-lined vessel introduces the oxidant, steam, and fuel. The fast exothermic partial oxidation reaction occurs directly at the burner outlet in the upper section, releasing large amounts of heat. Recirculation zones and large vessel volume give a long residence time, allowing the slower endothermic steam reforming reactions to come to near-equilibrium conditions at the reactor outlet. With respect to reforming performance, the POX reaction is less pressure dependent than steam-methane reforming reaction. Therefore as opposed to autothermal reformers, the standard POX reformer can be used at very high pressures without excessive methane-slip. Operation at high-pressure allows smaller downstream units and higher driving forces for CO_2 removal. Some steam is added to the reforming process to avoid coking and improve the overall cycle performance. On the cost side, as no fuel desulphurization is required and expensive high temperature catalysts can be avoided, the POX technology is expected to have cost advantages over other reforming concepts. Also, air-blown rather than oxygen-blown POX reforming has been chosen in this study in order to avoid the large efficiency penalty caused by an air separation unit.

Regarding technology maturity and risks, some of the GE natural gas POX reformers currently operate at the same pressure and temperature conditions (80–100 bars and 1150°C) as in the concepts considered here; however, they are at a smaller scale. The extensive experience acquired on commercial large-scale oxygen-blown units could be leveraged to scale up natural gas POX reformers units. The operating costs of the POX reformer are low, with the refractory material having a replacement cycle of at least ten years. The new compact POX reformer used in the advanced concept is seen as a long-term development effort. The impact of natural gas contaminants on the size and cost of the reformer is expected to be minor for typical Norwegian natural gas applications and has been neglected at this stage.

Table 2 Flow data for the advanced concept

Station	1	2	3	4	5	6
T (°C)	15	309	550	1150	40	134
P (bar)	1.013	~18	99	—	23.8	—
M (kg/s)	~640	122.7	122.7	148.1	155.2	129.1

4.4 HRSG and Low Temperature Syngas Expander. The base line concept uses a convection-based HRSG to cool down the syngas at the outlet of the POX reactor and to produce saturated high-pressure steam for extra power generation. Saturated steam is generated to avoid high tube wall temperatures and metal dusting. This is today's technology, however it leads to a high exergy loss. A commercially available low temperature syngas expander (inlet temperature of 190°C) is used to recover the pressure in the syngas prior to combustion. Furthermore, this expander drives the air compressor and natural gas compressor.

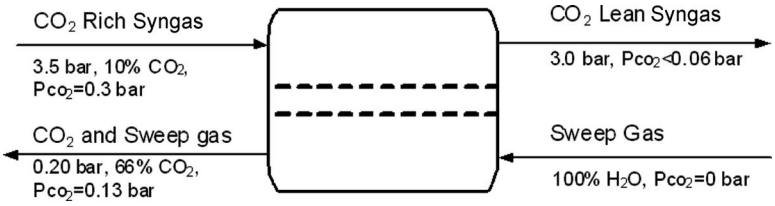


Fig. 3 Example of CO_2 separation membrane operating conditions in the advanced concept

4.5 SynGT. In the advanced concept, the SynGT comprises an air compressor, feeding a set of radically compact POX reformers (arranged in a can-annular or silo fashion) and an axial syngas expander. This module hence likens a gas turbine in principle, with the exception that the combustion is very rich instead of being lean. In combination with the downstream steam generation, the syngas island is now a combined Brayton-Rankin cycle; hence, the efficiency of the syngas process is increased. An extensive long-term development is required to realize such a module for this market. However, many aspects of modern gas turbine technology such as cooling, materials, and premixers may be transferred advantageously.

4.6 CO_2 Capture. The high-pressure level and CO_2 concentration in the proposed cycles lead to comparably high CO_2 partial pressures at the CO_2 separation unit inlet: 12.2 bars in the base line cycle and 3.6 bars in the advanced cycle. For comparison, gas turbines with atmospheric CO_2 capture typically cannot exceed a CO_2 partial pressure of 0.05 bar. The higher partial pressure of the proposed cycles increases the driving forces for CO_2 separation, which gives faster reaction kinetics and reduces the specific energy consumption of conventional solvent-based CO_2 separation systems. In addition, new types of CO_2 separation methods such as CO_2 membranes could be utilized, which have the potential to further reduce energy requirements for the CO_2 separation. As all the carbon in the fuel gas passes the CO_2 separation unit, high overall CO_2 capture rates can be achieved. Therefore, considering typical CO_2 separation efficiencies of up to 80–90% of the generated CO_2 can be captured. The CO_2 separation plant is reviewed in detail in the next section.

4.7 Stand-Alone CO_2 Compression Chain. For practical reasons and risk mitigation, it was decided to use a stand-alone CO_2-compression chain driven by an electric motor, whereas the syngas expander produces power to drive a generator as well as the natural gas and air compressors. The CO_2 compression chain is intercooled and working with 1.5 bar inlet pressure in the base line concept and with 0.3 bar (subatmospheric) in the advanced concept. This is a consequence of the different CO_2 separation technologies selected for the two cases. The final delivery pressure of the CO_2 is 100 bars.

4.8 Syngas Saturator. The syngas saturator is used in both plant concepts to increase the water content of the syngas prior to combustion. This has the effect of compensating for the reduced flow in the gas turbine due to CO_2 removal. The saturator requires a nonadiabatic saturation column with a packed-bed. Only a part of the hot water introduced to the tower is evaporated, leaving a cold exit stream to recover low temperature heat from the plant, mainly from the syngas condenser and feed air cooler.

5 CO_2 Separation Plant

Both solvent-based as well as membrane-based CO_2 capture units have been investigated. The systems have been designed to separate 80% of the overall CO_2 produced in the cycle. The CO_2 capture rate is therefore defined as

$$\text{CO}_2 \text{ capture rate } (\%) = \frac{\text{CO}_2 \text{ captured in the cycle(kg/s)}}{\text{CO}_2 \text{ produced in the cycle(kg/s)}} \times 100$$

5.1 Solvent-Based Capture Unit in Base Line Concept. A near-term solution for CO_2 separation is selected for the base line concept. This solution is based on chemical absorption using an amine-based solvent. For pressurized syngas at around 80 bars and 15 vol % CO_2, the following energy requirements have been used:

- specific thermal energy consumption:

$$1.32 \ \text{MJ}_{\text{thermal}}/\text{kg}_{\text{CO}_2}$$

- specific electrical energy consumption:

$$0.175 \ \text{MJ}_{\text{el}}/\text{kg}_{\text{CO}_2}$$

- CO_2 is supplied to the compression chain at 1.5 bars

CO_2 separation based on physical absorption has also been investigated as an alternative to the base line concept. However, the preliminary results based on an in-house model showed a large fuel slip to the CO_2 side (CH_4, H_2, and CO) due to the very high-pressure level of the feed gas (around 80 bars). This resulted in a significant penalty in the heating value of the syngas and therefore a comparably poor overall cycle performance. This alternative was not further investigated.

5.2 CO_2-Selective Membranes in Advanced Concept. A medium to long-term cycle improvement is presented in the advanced concept by use of a CO_2 separation membrane, see Fig. 3.

The flux of CO_2 through a membrane is dependent on the difference in partial pressures between the two sides. Sweep steam is used on the membrane permeate side in order to ensure a positive partial pressure differential. A novel membrane material is needed in order to reach the high selectivities required for this application. Depending on the ultimate use of the captured CO_2, additional polishing steps may need to be included.

For the performance evaluation in the advanced cycle, the following operating conditions have been selected for CO_2 separation membranes.

- CO_2 separation efficiency is 95%.
- The permeate side of the membrane is operated at a pressure of 0.3 bar absolute and the CO_2 is supplied to the compression chain at this pressure.
- 5 kg/s steam at 0.3 bar is extracted from the steam cycle and used as sweep gas in the membrane unit.

6 Syngas Firing in Gas Turbines

On the turbomachinery side, less than 20% of the compressed air is extracted at the discharge of the gas turbine compressor and is sent to the syngas generation island. This extraction is compatible with the capabilities of existing large frame machines. The mass flow of the working fluid through the gas turbine is still close to design conditions, as the loss of mass flow in the syngas due to the CO_2 separation has been partly overcome by the addi-

tion of water vapor in a syngas saturator. Off-design turbomachinery performance is nevertheless considered in the analysis to account for the slightly modified operating conditions of the gas turbine.

Compared with natural gas fueled gas turbines, syngas fueled turbines contain higher moisture in the hot gas path, which will affect durability. Increases of about 100–150°C from the previous IGCC gas turbine firing temperature level to the FB natural gas fired turbine level will require reconsideration of the combustor design as well as turbine flow path materials and coatings. The resulting syngas composition is somewhat similar to coal-based IGCC decarbonized fuels, having a Wobbe index of about 5 but with a slightly lower concentration of hydrogen. The key technology that will allow the increase in the firing temperature (and therefore efficiency) and lower NO_x is the development of a dry low NO_x syngas combustor, currently pursued at GE. The reduction or elimination of steam typically required for NO_x abatement in current high-hydrogen diffusion combustors will determine a reduction in moisture content of the hot gas path. In turn, this will allow the use of current or future materials to be employed at F and FB class conditions. Recent advances in combustor development at GE demonstrate that low emissions combustion of carbon-free syngas is achievable without steam injection by using advanced premixers [8]. Flashback and flame-holding propensity—typical challenges of hydrogen premixed combustion—can be overcome through a combination of new design introduction and fuel introduction without penalties in pressure drop across the combustor. Under these assumptions suggested by recent research performed at GE [9], it was concluded that the uprate to FB firing temperatures while maintaining low emissions on syngas combustion is achievable.

7 Performance Modeling Assumptions And Simulation Results

The performance potential of the proposed cycle was evaluated using GATECYCLE™ [10] and HYSYS® [11] as simulation platforms. The analysis was performed based on a large frame gas turbine, and an off-design turbomachinery performance has been included.

The main assumptions used for the performance analysis are given below.

- CO_2 is compressed to 100 bars, and the specific energy used for CO_2 compression is 0.32 MJ_{el}/kg_{CO_2} in the base line concept and 0.47 MJ_{el}/kg_{CO_2} in the advanced concept.
- Natural gas is supplied to the cycle at 50 bars.
- 8°C seawater is available for cooling.
- POX outlet temperature maintained at 1150°C.
- The gas turbine used in the power island is a 9FB, which has the following performance under natural gas firing: 37.9% efficiency and 279 MW net power in simple cycle, and 58% efficiency for 410 MW net power in combined cycle.

The performance of the cycles has been optimized by varying various cycle parameters, in particular, the quench inlet temperature and the steam-to-carbon ratio of the POX reactor. A sensitivity analysis showed that increasing the quench temperature is beneficial for the CO_2 capture but leads to a decrease in the cycle efficiency due to decreased high-pressure steam production. An increased steam-to-carbon ratio is beneficial to the water gas shift reaction and consequently to the amount of CO_2 capture. Therefore, by increasing the steam-to-carbon ratio in the POX reactor, a lower quench temperature can be chosen, leading to a better cycle efficiency. However, the steam-to-carbon ratio is also limited in order to keep a working fluid volume flow close to design point in the gas turbine.

It was found that at 80% overall CO_2 capture, the base line cycle configuration has the potential to reach a net efficiency of 47.5% (453 MW) and the advanced concept to reach a net efficiency of 50.7% (493 MW). This is equivalent to an efficiency penalty of 10.5% pts and 7.3% pts, respectively, including CO_2 capture and compression to 100 bars. Both cycles exhibit a substantial power boost compared with a standard natural gas fired 9FB combined cycle without CO_2 capture, reaching 40 MW for the base line concept and 80 MW for the advanced concept. Table 3 summarizes the performance results for the base line and advanced concepts.

Table 3 Cycle performance results, including CO_2 compression

	Net efficiency (LHV) (%)	CO_2 capture rate (%)	Power output (MW)
Base line concept with chemical absorption	47.5	80	455
Advanced concept with membranes	50.7	80	493

8 Cost Estimation Methodology and Cost Analysis Results

A cost analysis was carried out to evaluate the economic potential of the proposed cycles. The key element of this analysis is to estimate costs of additionally installed equipment, which is needed for CO_2 capture and compression but is not part of a traditional natural gas fired combined cycle. All cost numbers are calculated assuming the cycles and components to be mature technologies.

The economic evaluation is based on the concept of *cost of* CO_2 avoided, which is defined as

$$\text{Cost of } CO_2 \text{ avoided}(\$/t_{CO_2}) = \frac{\text{COE}_{\text{capture}} - \text{COE}_{\text{ref}}\ (\$/\text{kW h})}{\text{spec.}CO_2\ \text{emiss.}_{\text{ref}} - \text{spec.}CO_2\ \text{emiss.}_{\text{capture}}(t_{CO_2}/\text{kWh})}$$

In this equation, the term COE stands for cost of electricity and typically is subdivided into capital, fuel, and operation and maintenance costs. The calculation of cost of CO_2 avoided requires the definition of a reference combined cycle plant without CO_2 capture (index: ref) in comparison to the cycle with capture (index: capture). The reference plant is assumed to be a plant of the same type as the plant with CO_2 capture. A 410 MW 9FB combined cycle without CO_2 capture is chosen as the reference case. In the cost calculation, potential additional cost for CO_2 transport and sequestration were not included. CO_2 compression to 100 bars and CO_2 dehydration however have been considered. Equipment cost estimates were based both on internal sources, as well as on preliminary quotes from suppliers. When evaluating the cost results, it has to be kept in mind that they are rough estimates based on data from 2006 Q1 for a "typical" turnkey installation. The cost calculations account for the soot water handling and the ground water treatment. At the time of this study, the fuel cost was assumed to be 3.5 €/GJ (HHV) in compliance with earlier studies. Capital charge rates around 10% were assumed to represent financing conditions similar to those expected for state-funded infrastructure projects. It should be noted that any changes in the assumptions of the financial boundary conditions would have a substantial impact on the overall cost of CO_2 avoided level. In general, the cost assumptions and methodology used here are the same as those used in Ref. [3] so that a comparative analysis can be performed.

The calculated cost for the base line and advanced concepts are listed together with the performance results in Table 4.

Table 4 Cycle cost and performance analysis results, including CO_2 compression

	Net efficiency (LHV) (%)	Net power output (MW)	CO_2 capture rate (%)	Reference plant used for cost analysis	Cost of CO_2 avoided
Target	≥50		≥80		≤30 $/$t_{CO2}$
Base line concept with chemical absorption	47.5	455	80	9FB CC	46 $/$t_{CO2}$
Advanced concept with membranes	50.7	493	80	9FB CC	29 $/$t_{CO2}$

9 Discussion of Results

The proposed cycles with integrated precombustion CO_2 capture are able to treat all the CO_2-rich syngas at pressure. Thereby, high CO_2 capture rates can be reached and are mainly limited by the CO_2 separation efficiency of the CO_2 separation unit and the combined POX reforming and WGS conversion efficiencies. As a result, the cycle is able to capture 80–90% of the generated CO_2 and meet the CO_2 capture design criteria of at least 80% overall carbon capture.

Based on a large frame gas turbine, net efficiencies including CO_2 compression to 100 bar range between 47.5% for the base line concept with amine-based capture and 50.7% for the advanced concept with membrane-based capture and high temperature syngas expander. Contrary to the base line concept, the advanced concept meets the design target of 50% efficiency. This is the combined result of using a high temperature syngas expander (which accounts for a approximately 2% pts increase in efficiency) and a CO_2 membrane.

Both cycles exhibit an advantageous power boost compared with a standard natural gas fired 9FB combined cycle without CO_2 capture, reaching 40 MW for the base line concept and an impressive 80 MW for the advanced concept. This power boost is the result of a net power generation from the syngas island. In the base line concept, high-pressure steam is produced in the HRSG situated at the exit of the POX reformer, leading to extra power in the steam cycle. Also, the syngas saturator partly replaces the missing CO_2 mass flow by a water vapor mass flow, which limits the power loss in the gas turbine combined cycle. In the advanced concept, the power boost is even larger due to the added power from the high temperature syngas expander.

From an economic viewpoint, the advanced cycle, with 29 $/$t_{CO_2}$ meets the cost target of 30 $/$t_{CO_2}$. If the advanced POX reactor is not in use, but still including the SynGT configuration, the cost of CO_2 avoided would be 33 $/$t_{CO_2}$, which is still close to the target. The base line concept with 46 $/$t_{CO_2}$ does not reach the cost target. A detailed cost evaluation of both concepts helped to identify and quantify the "technology game changers" in use in the advanced concept: a reduction of roughly 10 $/$t_{CO_2}$ in the cost of CO_2 avoided can be attributed to the use of a CO_2 membrane, a reduction of 3 $/$t_{CO_2}$ is due to the use of the high temperature syngas expander, and another 4 $/$t_{CO_2}$ is due to the advanced POX reformer. These results illustrate how a step-by-step introduction of new technology can help to meet the cost and performance targets.

In short, the advanced precombustion cycle offers the advantage of combining a comparably high performance and promising economics. From a development point of view, the base line and the incremental introduction of new technologies, once tested and proven, is a way to mitigate risks on the pathway to a high performance advanced plant.

In this context, the following development and optimization areas can be identified:

- highly selective CO_2 membrane for operation with syngas
- high temperature syngas expander
- advanced POX reformer
- syngas combustor

10 Summary

Innovative precombustion capture cycles for natural gas have been developed and investigated. A short-term base line concept as well as a long-term advanced concept have been proposed, offering the possibility to incrementally introduce new technology and reduce risks. The cycles include a high-pressure syngas generation island based on the POX reformer technology and a power island including a CO_2-lean syngas-fired large frame gas turbine in combined cycle. The high-pressure level in the syngas generation island leads to very advantageous CO_2 partial pressures at the separation unit, as well as a reduced equipment size. The base line concept makes use of a state-of-the-art chemical absorption plant whereas the advanced concept uses CO_2 membranes. In the advanced concept, energy recovery from the generated hot syngas at the exit of the reformer has been optimized using a high temperature syngas expander, leading to a net 2% pts increase in cycle efficiency. Also a new compact and hence cost-effective POX reformer has been considered in the analysis of the advanced concept.

The cycles proposed can capture 80% or more of the generated CO_2. The base line cycle does not meet the target net efficiency of 50%; however, it offers a 40 MW power boost compared with the 410 MW generated by an equivalent standard natural gas fired large frame machine. This is due, in particular, to the generation of high-pressure steam in the syngas generation island. The advanced concept based on the same large frame machine reaches 50.7% (LHV) net efficiency and shows an even larger power boost of 80 MW. This is attributable to the high temperature syngas expander that more efficiently uses the heat released during reforming.

From an economic view point, with 33 $/$t_{CO_2}$, the advanced concept using a standard POX reformer is close to the cost of CO_2 avoided target; and with 29 $/$t_{CO_2}$, it even meets this cost target by using an advanced compact POX reformer. Compared with the short-term base line concept, the advanced concept shows a cost of CO_2 avoided reduction of about 10 $/$t_{CO_2}$ due to the use of a CO_2 membrane, another 3 $/$t_{CO_2}$ reduction thanks to the high temperature syngas expander, and another 4 $/$t_{CO_2}$ due to the advanced POX reformer. These results illustrate how a step-by-step introduction of new technology could help meet the cost and performance targets while minimizing risks.

Acknowledgment

This work has been partly funded by Statoil ASA. The authors express their appreciation to Statoil ASA and GE for the permission to publish this work.

The authors appreciate the large support by many GE individuals and would like to thank, in particular, Rich Depuy and Jatila Ranasinghe from GE Energy as well as Enrico Gori from GE Oil & Gas.

Nomenclature

P_{CO_2} = CO_2 partial pressure

Y_{CO_2} = CO_2 mole fraction
P = total pressure
T = temperature

References

[1] Steeneveldt, R., Berger, B., and Torp, T. A., 2006, "CO_2 Capture and Storage: Closing the Knowing: Doing Gap," Chem. Eng. Res. Des., **84**(9), pp. 739–763.

[2] Kvamsdal, H., Jordal, K., and Bolland, O., 2007, "A Quantitative Comparison of Gas Turbine Cycles With CO2 Capture," Energy, **32**(1), pp. 10–29.

[3] Finkenrath, M., Ursin T., Hoffmann, S., Bartlett, M., Evulet, A., Bowman, M., Lynghjem, A., and Jakobsen, J., 2007, "Performance and Cost Analysis of a Novel Gas Turbine Cycle With CO_2 Capture," ASME Paper No. GT2007-27764.

[4] Finkenrath, M., Eckstein, J., Hoffmann, S., Bartlett, M., Evulet, A., Bowman, M. J., Lynghjem, A., Jakobsen, J., and Ursin, T. P., 2006, "Advanced Gas Turbine Cycles With CO_2 Removal," *Proceedings of the 8th International Conference on Greenhouse Gas Control Technologies*, Trondheim, Norway, Jun. 19–22, Paper No. GHGT-8.

[5] Shekhawat, D., Luebke, D. R., and Pennline, H. W., 2003, "A Review of Carbon Dioxide Selective Membranes: A Topical Report," United States Department of Energy, National Energy Technology Laboratory, Report No. DOE/NETL-2003/12001.

[6] GE Patent Application No. US11/960865, "Method for High Efficiency Fuel Decarbonisation," filed Aug. 7, 2006.

[7] GE Patent Application No. US11/462867, "Systems and Methods for Power Generation with Carbon Dioxide Isolation," filed Dec. 20, 2007.

[8] Lacy, B. P., McManus, K. R., Varatharajan, B., and Shome, B., 2005, "Premixer Design for High Hydrogen Fuels," OSTI Technical Report No. 889756.

[9] Evulet, A., and Elkady, A., 2006, "Advanced Combustion Trends for CO_2 Capture-ready Plants," GE Global Research, Report No. 2006GRC610.

[10] GateCycle™ Software, www.gepower.com/prod_serv/products/plant_perf_software/en/gatecycle/index.htm

[11] www.aspen.com/hysys

Effect of Time and Temperature on Thermal Barrier Coating Failure Mode Under Oxidizing Environment

N. S. Cheruvu

K. S. Chan
ASME Fellow

Southwest Research Institute,
6220 Culebra Road,
San Antonio, TX 78238

D. W. Gandy
Electric Power Research Institute,
Charlotte, NC 28262

Thermal barrier coatings (TBCs) have been recently introduced to hot section components, such as transition pieces and the first two stages of turbine blades and vanes of advanced F, G, and H class land-based turbine engines. The TBC is typically applied on metallic-coated components. The metallic bond coat provides oxidation and/or corrosion protection. It is generally believed that the primary failure mode of TBCs is delamination and fracture of the top ceramic coating parallel to the bond coat in the proximity of the thermally grown oxide (TGO) between coatings. One of the concerns associated with the use of a TBC as a prime reliant coating is its long-term stability. The effect of long-term operation at typical land based turbine operating temperatures of below 1010°C (1850°F) of the failure mode of TBCs is unknown. Long-term isothermal tests were conducted on the thermal barrier-coated specimens at three temperatures, 1010°C (1850°F), 1038°C (1900°F), and 1066°C (1950°F), to determine the effects of long term exposure on the TBC failure location (mode). Following the isothermal testing, the samples were destructively examined to characterize the degradation of the TBC and determine the extent of TGO cracking, TGO growth, bond coat oxidation, and TBC failure location after long term exposure for up to 18,000 h. Optical microscopy and a scanning electron microscope (SEM) attached with an energy dispersive spectroscopy (EDS) system were used to study the degradation of the TBC and bond coatings. The results showed that long term isothermal exposure leads to a change in the TBC failure mode from the delamination of the TBC at the TGO/TBC interface to the internal oxidation of the bond coat and bond coat delamination. In this paper, the effect of long-term exposure on the delamination of TBC and the bond coat failure mode is discussed.
[DOI: 10.1115/1.2979747]

Introduction

The advanced turbine systems (ATS) program has played a key role in developing a successful partnership between private industries and universities. As a result, several advanced materials, modified MCrAlY bond coat and TBCs have been developed and introduced in natural gas fired G and H class turbines. The TBCs are used as a prime reliant design in these advanced turbines. The new technologies developed for G and H engines have been applied to older D, E, and F class engines to increase their performance. The TBC coatings are applied on the hot section components, such as the first two stages of turbine blades and vanes, combustor cans, and transition pieces.

TBCs are made of low thermal conductivity ceramic oxides. The widely used TBC in the industry is made of yttria stabilized zirconia (8% Y_2O_3–ZrO_2). The TBC is normally deposited using either the air plasma (APS) or electron beam physical vapor deposition (EBPVD) process over a MCrAlY or Pt-aluminide bond coated surface [1,2]. NiCoCrAlY is widely used as a bond coat for the land based turbine applications and is deposited using a low pressure plasma spray (LPPS), vacuum plasma spray (VPS), or
high velocity oxy flame (HVOF) process. Metallic coatings commonly used for turbine blades include CoCrAlY, CoNiCrAlY, NiCoCrAlYHfSi, and PWA 286. These bond coatings exhibit a duplex microstructure containing fine β-phase particles [CoAl or (CoNi) Al] in a matrix of γ. The ZrO_2 based ceramic coatings are porous and transparent to oxygen. The presence of interconnected pores in the TBC also allows for the easy ingress of oxygen from the turbine environment to the bond coat. Therefore, the formation of thermally grown aluminum Al_2O_3 oxide (TGO) at the bond coat/TBC interface during operation is inevitable [2].

The microstructure of a service-exposed TBC-coated component exhibits four different layers. These layers are (i) the TBC, (ii) the TGO between the bond coat and the TBC, (iii) the bond coat, and (iv) the superalloy substrate. The degradation of TBCs includes growth and cracking of the TGO and the formation of transient (nonaluminum) oxides with operating time. The TGO growth during operation results in constrained volume expansion that produces compressive stresses. Upon cooling, the thermal expansion mismatch between the TGO and the bond coat also leads to very high thermal compressive stresses in the TGO. The driving force for the TBC fracture or spallation is the strain energy, which depends on the TGO thickness and the TGO stress. In advanced cyclic duty engines, the TGO growth mechanism is primarily responsible for TBC spallation [3–7]. Most of the life evaluation models developed for TBCs are based on the assumption that TBC failure initiates in the TGO or in the vicinity of the TGO in the ceramic coating [8–18]. It is not well known whether this failure mode is relevant for base loaded turbines that operate at a temperature of below 1010°C (1850°F) for a long time over several thousand hours. The primary objective of this investigation was to determine the effect of long-term isothermal exposure on the TBC failure mode and bond coat oxidation.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received December 22, 2007; final manuscript received June 19, 2008; published online December 19, 2008. Review conducted by Dilip R. Ballal. Paper presented at the 2008 ASME Turbo Expo, Berlin, Germany, June 9–13, 2008.

Experimental Details

Materials and Coatings. A GTD-111 DS blade retired from a Frame 5002 engine was procured for machining test coupons from its shank section. The shank section normally operates at a much lower temperature than the blade airfoil and as a result, the material at the shank section is not expected to degrade during service. The structure and properties of the material at the shank section represent the initial, as heat-treated condition [19,20]. Compositional measurements were made on the shank section using energy dispersive X-ray spectroscopy (EDS). Typical composition of the GTD-111 test material is given in Table 1.

Cylindrical specimens 9.1 mm in diameter and 38.1 mm long (0.36 in. diameter and 1.5 in. long) were removed from the blade shank section using an electrodischarge machining process. The specimens were ground and polished to remove the recast layer. To hold and rotate the specimens during coating application, a small bolt was tack welded to one of the specimen end faces. Turbine Airfoils, Coatings, and Repairs (TACR) applied the CT102 (CoNiCrAlY) bond coat and yttria stabilized zirconia coating. A low-pressure plasma spray process (LPPS) was used to apply the CT102 coating on all specimens. The composition of the powder is given in Table 2. After the application of the CT102 coating, all specimens were given a vacuum diffusion treatment at 1121°C (2050°F) for 2 h prior to the application of the top ceramic coating by using an air plasma spray (APS) process. The chemical composition of the ceramic coating powder is given in Table 3. The as-coated specimen is shown in Fig. 1. The bolt was removed from all specimens prior to their isothermal exposure.

Testing and Evaluation

Isothermal Tests. Multiple TBC-coated GTD-111 specimens were aged in three different furnaces, which were maintained at three temperatures: 1010°C (1850°F), 1038°C (1900°F), and 1066°C (1950°F). One specimen at a time was removed at predetermined intervals for metallurgical evaluation.

Metallurgical Sample Preparation. Metallurgical mounts were prepared from all the samples removed after isothermal exposure at 1010°C (1850°F), 1038°C (1900°F), and 1066°C (1950°F) for different times. All specimens were first infilterated by epoxy to protect the TBC from spalling during the removal of a small section for metallurgical evaluation. A transverse section was removed from each of these infilterated specimens with a slow speed diamond saw. The transverse sections were then remounted in epoxy, polished, and etched using standard metallographic techniques. The polished and etched mounts were examined under optical and scanning electron microscopes for thermally grown oxide thickness, bond coat/TBC interface delamination, coating degradation, and bond coat oxidation. Chemical composition measurements were made at several locations on the TGO between the bond coat and the TBC by using EDS.

Table 1 Chemical composition of GTD-111 and IN-738 test materials (wt %)

Blade	Al	Ti	Cr	Co	Mo	Ta	W	Ni
GTD-111	3.2	5.2	14.4	9.2	2.1	4.0	3.2	Bal

Table 2 Chemical composition of the CT102 bond coat powder (wt %)

Al	Co	Cr	Ni	Y
8.0	Balance	21.0	32.0	0.5

Table 3 Chemical composition of the ceramic coating powder (wt %)

Al_2O_3	Fe_2O_3	SiO_2	TiO_2	Y_2O_3	HfO_2	ZrO_2
0.13	0.02	0.27	0.09	7.69	1.85	Balance

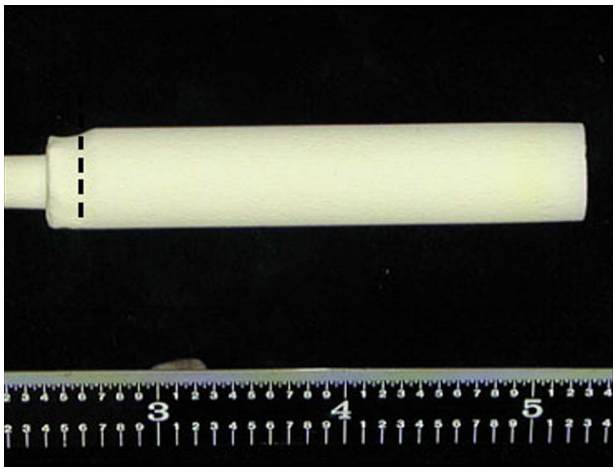


Fig. 1 An as-coated specimen

Results and Discussions

Microstructure of As-Coated Specimens. Typical microstructure of the bond and ceramic coatings in the as-deposited condition is presented in Fig. 2. The thickness of the CT102 bond coat varied from approximately 130 μm (0.005 in.) to 155 μm (0.006 in.). The bond coat was dense and exhibited a microstructure consisting of fine β-phase (NiAl) in a matrix of γ (Ni–Cr–Co solid solution). The outer surface of the bond coat exhibited an

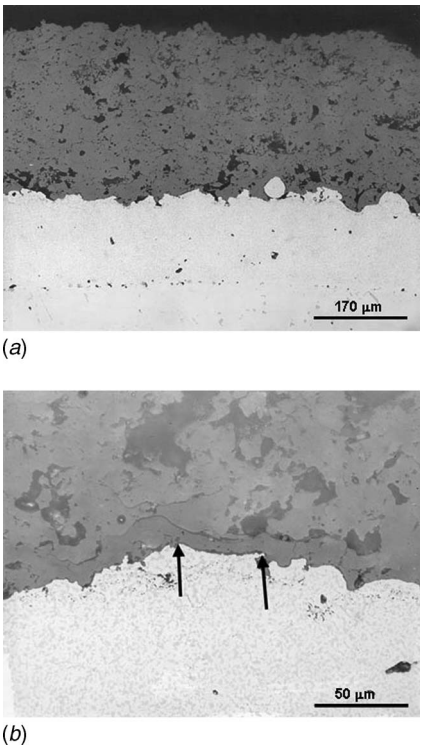


Fig. 2 Optical micrographs of the as-coated TBC on a NiCoCrAlY-coated GTD-111 specimen. The arrows point to delamination cracks at the TBC/bond coat interface.

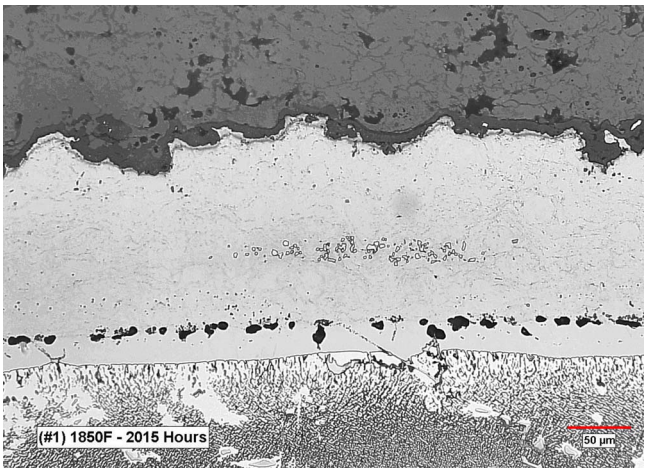


Fig. 3 Backscattered electron micrographs of mixed oxides in the TGO on the coated specimen after 2015 h of exposure at 1010°C (1850°F)

irregular surface topography, as shown in Fig. 2, which is normal for the LPPS processed coatings. However, no attempt is made to measure the surface roughness of the bond coat in this study. The rough surface is required for the adhesion of air plasma TBC. The thickness of the TBC on the specimen varied from about 250 μm to 375 μm (from 0.010 in. to 0.015 in.). Examination of the bond coat/TBC interface revealed a thin and discontinuous TGO scale in isolated areas and a few microdelamination cracks at the interface. These cracks were parallel to the interface and were about 50 μm (0.002 in.) long. Propagation of these cracks during isothermal or thermal cycling exposure could lead to TBC spallation. The microstructure of the TBC exhibited a layered structure resulting from the splattering of the powder particles during the deposition process. The pores/voids and microcracks were seen in the coating, which is normal for an APS-processed TBC [21–23].

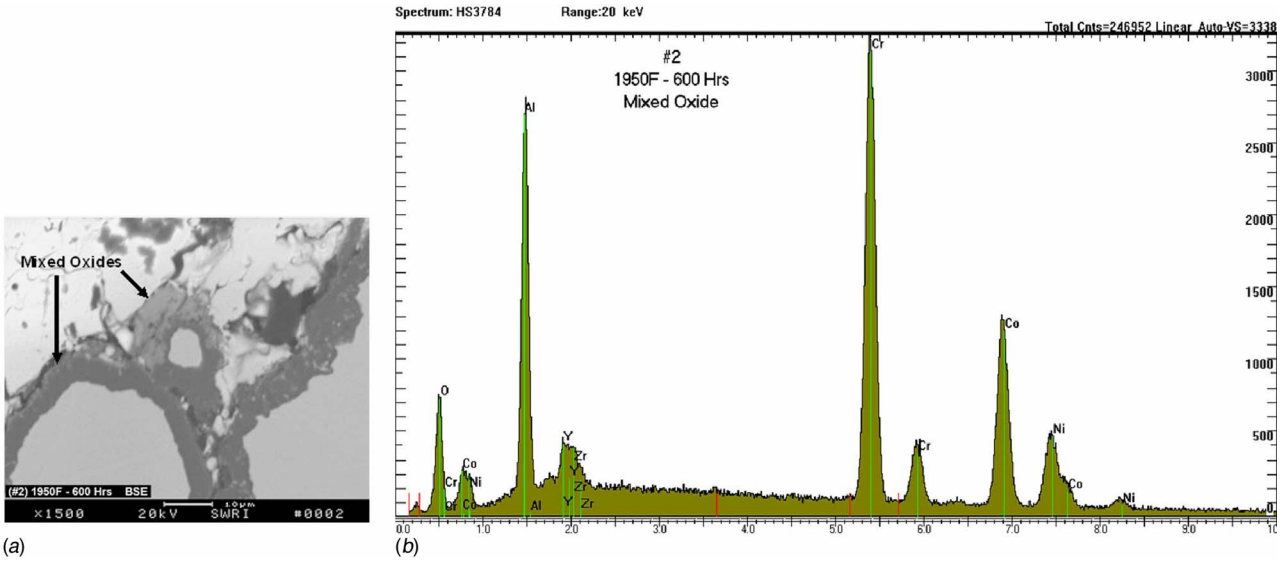
Microstructure of Exposed Specimens

Metallurgical examination of the exposed samples showed that the β-phase in the bond coat was consumed after exposure for a relatively short time at the three isothermal temperatures investigated. The β-phase in the bond coat was completely consumed in about 2000 h of exposure at the lowest temperature of 1010°C (1850°F) and 600 h exposure at the highest temperature 1066°C (1950°F). A typical microstructure of the bond coat after 2015 h of exposure at 1010°C (1850°F) is shown in Fig. 3. In addition to the depletion of the β-phase in the bond coat, a mixed oxide scale consisting of Al_2O_3 and oxides of Ni, Co, and Cr was observed between the bond coat and TBC in samples exposed for a relatively short time. For example, Fig. 4 shows a SEM back scattered electron image of the TGO on the sample after 600 h at 1066°C (1950°F). EDS measurements of the mixed oxide scale showed that it is comprised of Cr, Al, Co, and Ni oxides. It is interesting to note that the outward growth of both the Al_2O_3 and other transient oxides.

Chan and Cheruvu [24] have evaluated the effect of thermal cycling at 1010°C (1850°F) and 1066°C (1950°F) on NiCoCrAlY coating degradation [24]. In contrast to the present results, their results showed that the β-phase was present in the coating after 4500 1 h thermal cyclings between the peak temperature of 1010°C (1850°F) and the room temperature (refer to Fig. 5 in Ref. [24]). Their results further showed that the oxide scale on the NiCoCrAlY coating was free from the transient Cr or Ni oxides in the Al_2O_3 scale for up to 1500 1 h thermal cycles at 1066°C (1950°F) and 2000 1 h thermal cycles at 1010°C (1850°F) (refer to Fig. 2 in Ref. [24]).

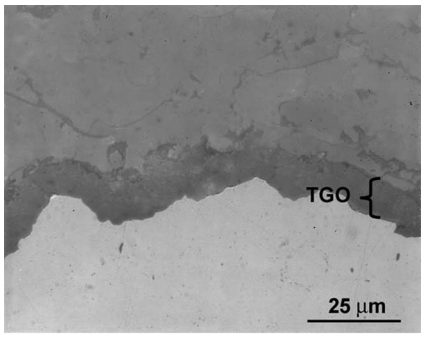
Comparison of the present results with the degradation results published earlier by the authors indicates that the bond coat underneath the TBC degrades faster than the NiCoCrAlY coating without TBC [24,25]. Consistent with these results, Stinger et al. [26,27] have reported that the TGO grew faster underneath the TBC than on the MCrAlY or on aluminide coatings without the TBC. They have postulated that the outward growth of the TGO into the TBC is responsible for the faster growth of the TGO. This suggests that the kinetics of the TGO growth is accelerated by the presence of the TBC. As the TGO grows, the β-phase in the bond coat is consumed faster.

TGO Thickness. As expected, the TGO thickness increased with the exposure time and temperature. Typical morphology of the TGO scale is shown in Figs. 5 and 6. The TGO thickness was measured on each sample at least at 30 locations, and the average values were used to determine the kinetics. These results are pre-

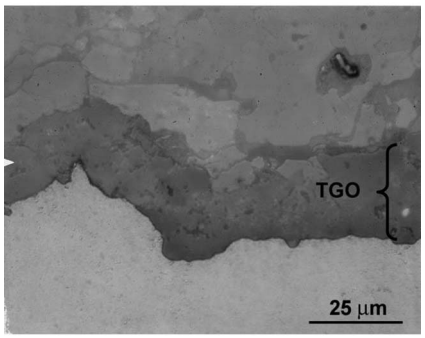


(*a*) (*b*)

Fig. 4 (*a*) SEM micrograph of mixed oxides in the TGO on the coated specimen after 600 h exposure at 1066°C (1950°F), and (*b*) EDS obtained from the mixed oxide shown by the short arrow in (*a*)

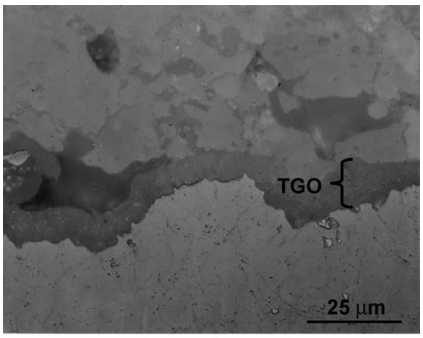


(*a*)

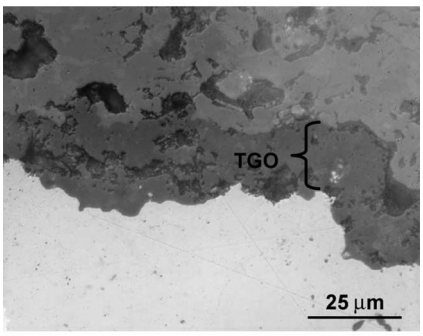


(*b*)

Fig. 5 Optical micrographs showing the variation of the TGO thickness after (*a*) 1510 h and (*b*) 2785 h of exposure at 1066°C (1950°F)



(*a*)



(*b*)

Fig. 6 Optical micrographs showing the variation of the TGO thickness on the NiCoCrAlY-coated GTD-111 specimen after (*a*) 2015 h and (*b*) 9850 h of exposure at 1010°C (1850°F)

sented elsewhere [7,25]. Typical kinetics of TGO growth results at 1010°C (1850°F) is illustrated in Fig. 7. It is clear from the results shown in Fig. 7 that the TGO growth rate significantly decreased after long-term exposure at 1010°C (1850°F). The TGO thickness on the specimens after 9850 h and 12,030h of exposure was almost comparable (18.9 μm versus 19.2 μm), as shown in Fig. 7.

TGO/Bond Coat Interface Cracking

The TGO scale exhibited delamination cracks at the TGO/TBC interface. These cracks are observed either in the TGO or in the TBC adjacent to the TGO/TBC interface. Figure 8 shows the variation of the TGO interface cracking as a function of the exposure time at 1010°C (1850°F) for the CT102 bond-coated GTD-111 specimens. A comparison of these micrographs shows that the length and the extent of cracking increase with the exposure time. On all specimens exposed to 1010°C (1850°F), the interface crack lengths were measured at ten locations on the transverse section. The average of ten cracks and the largest crack size were reported as a function of the exposure time in Fig. 9. The maximum crack length results showed a wide scatter, while the average interfacial crack length results exhibited relatively less scatter. No attempt was made to determine the standard deviation for the average interfacial crack lengths. The average interfacial crack length showed an excellent linear relationship with the exposure time, as illustrated in Fig. 9.

Internal Oxidation of Bond Coat. Metallurgical examinations of the exposed samples showed evidence of internal oxidation of the bond coat after long-term exposure. Unlike TGO growth, which occurred at the bond coat/ TBC interface, internal oxidation occurred within the bond coat. The inspection cycle at which evidence of the bond coat internal oxidation was first observed was considered as the initiation time for the oxidation of the bond coat. As the internal oxidation of the bond coat progressed (i.e., as the oxides formed within the bond coat), the TGO growth rate at the bond coat/TGO interface was found to decrease significantly, as illustrated in Fig. 7. The initiation time for the bond coat oxidation at three temperatures is presented in Table 4. It is clear from these results that the initiation time for the bond coat internal oxidation is inversely related to the exposure temperature. The extent of internal oxidation of the bond coat increased with the exposure time. A typical microstructure of the bond coat at the early stages of internal oxidation in a specimen exposed at 1010°C (1850°F) is shown in Fig. 10. Increasing the exposure temperature or time resulted in severe internal oxidation, as shown in Fig. 11.

Figure 12 shows the oxidation attack of the bond coat on the GTD-111 after exposure at 1066°C (1950°F), 1038°C (1900°F), and 1010°C (1850°F). Close examination of these micrographs indicates that internal oxidation of the bond coat and the coales-

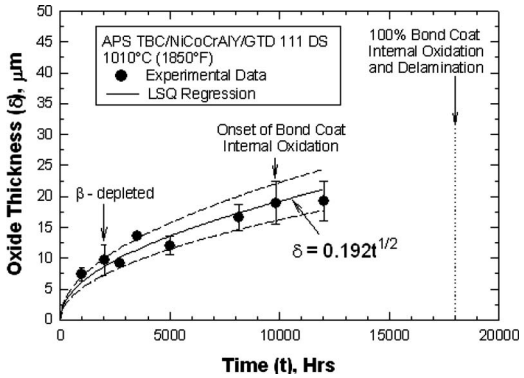
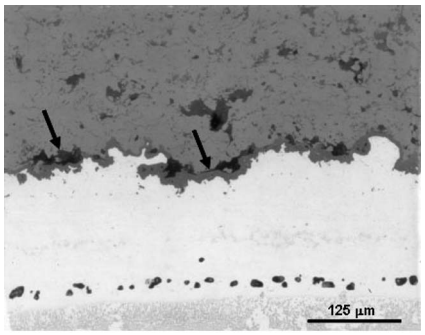
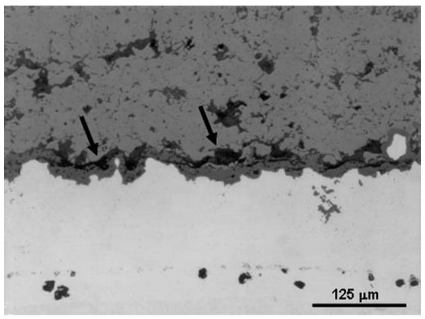


Fig. 7 Kinetics of the TGO growth at the interface of the APS TBC/NiCoCrAlY bond coat on GTD-111DS



(*a*)



(*b*)

Fig. 8 Optical micrographs of the TBC/CoNiCrAlY interface of GTD-111 specimens after (*a*) 2015 h and (*b*) 12,030 h of exposure at 1010°C (1850°F) showing the variation of the interface cracking. Arrows on the micrographs point to interface cracks.

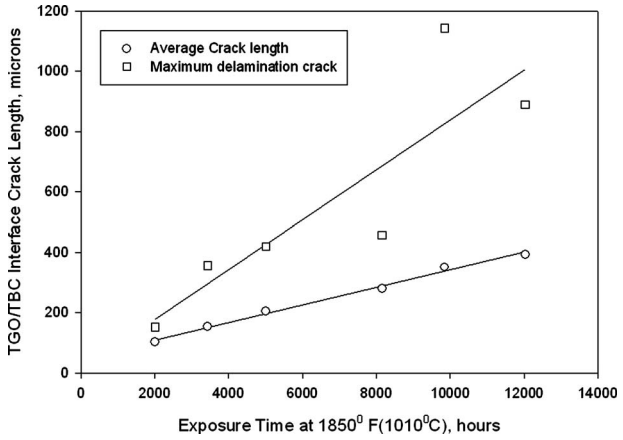


Fig. 9 Average and maximum delamination crack lengths at the TGO/TBC interface on the CoNiCrAlY-coated GTD-111 specimens as a function of the exposure time at 1010°C (1850°F)

Table 4 Initiation time for the internal oxidation of the bond coat

Bond coat/substrate	Temperature (°C(°F))	Time for internal oxidation (h)
CoNiCrAlY/GTD-111	1066 (1950)	2785
CoNiCrAlY/GTD-111	1038 (1900)	<6470
CoNiCrAlY/GTD-111	1010 (1850)	9850

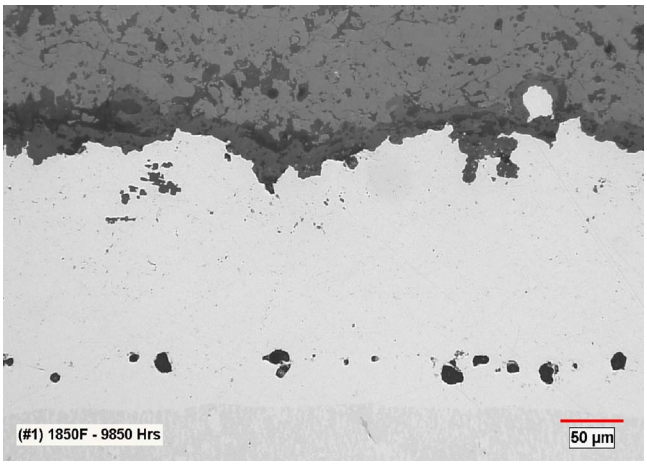


Fig. 10 Optical microstructure of the CoNiCrAlY bond coat on the GTD-111 specimens after 9850 h of exposure at 1010°C (1850°F) showing the early stages of internal oxidation

cence of the Kirkendall voids at the bond coat/substrate interface led to the delamination of the bond coat. Metallurgical examination of the exposed samples also showed that the Kirkendall voids form at the bond coat/substrate interface after a relatively short time exposure at the three temperatures investigated. Figures 13 and 14 show the extent of the interface void formation as a function of the exposure time at 1010°C (1850°) and 1066°C (1950°F), respectively. It is clear from the micrographs that the extent of voids at the bond coat/substrate interface in the samples increases with the exposure time and temperature, and the voids also coalesce with increasing exposure time. The Kirkendall voids are known to from as the result of the interdiffusion of elements between the bond coat and the substrate alloy.

These results show that the TBC failure mechanism depends on the exposure time and temperature. At relatively low temperatures and long term exposure (similar to the operating conditions of a base loaded engine), the TBC failure mechanism changes from the TGO growth controlled mechanism at the bond coat/TBC interface to internal oxidation of the bond coat and bond coat delamination. The TBC failure location also changes from the TGO/TBC interface to the bond coat/substrate interface.

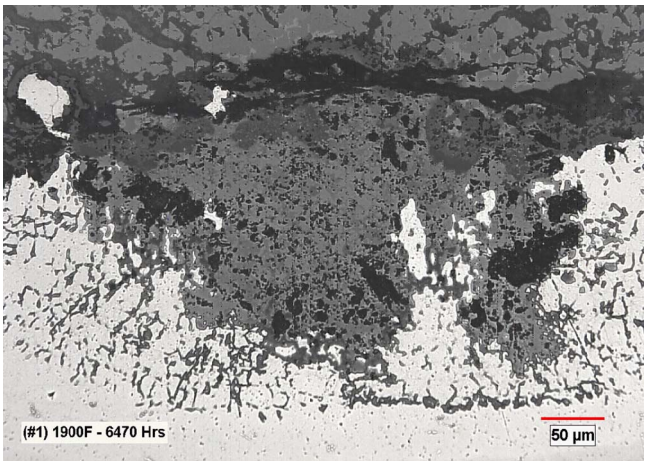
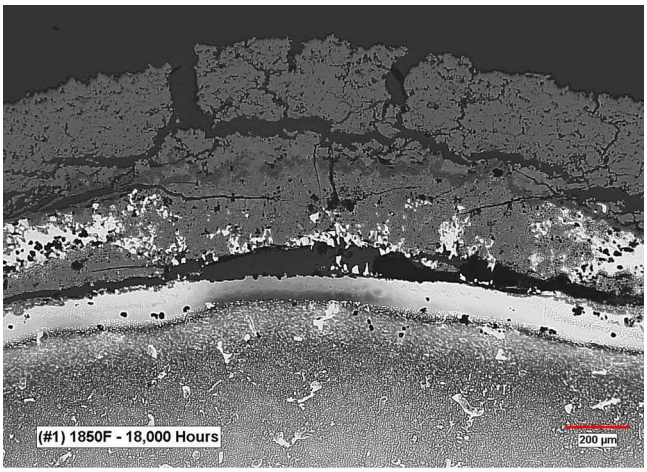
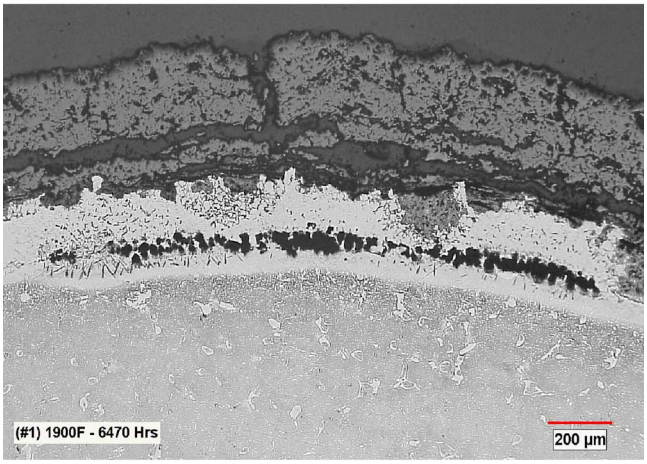


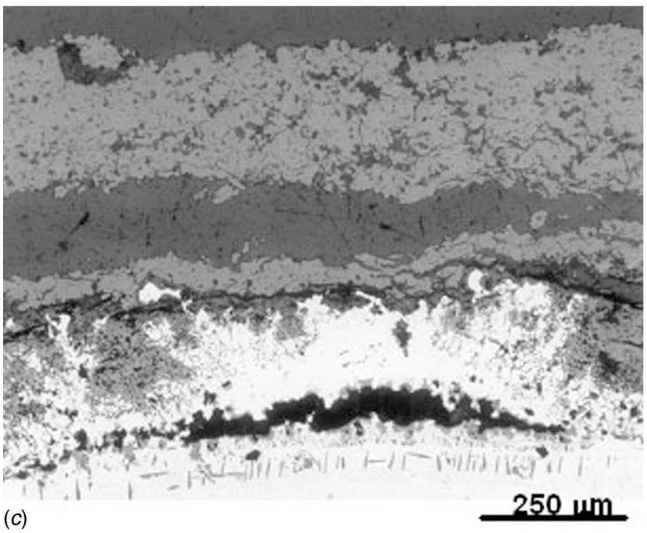
Fig. 11 Optical microstructure of the CoNiCrAlY bond coat on the GTD-111 specimens after 6470 h at 1038°C (1900°F)



(*a*)

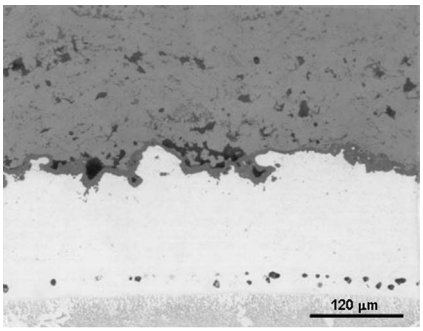


(*b*)

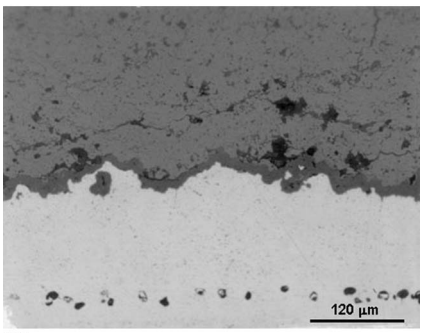


(*c*)

Fig. 12 Optical microstructure of the NiCoCrAlY bond coat on GTD-111 after (*a*) 18,000 h of exposure at 1010°C (1850°F), (*b*) 6470 h of exposure at 1038°C (1900°F), and (*c*) 2925 h of exposure at 1066°C (1950°F). Note the onset of bond delamination and void coalescence at the bond coat/substrate interface.



(*a*)



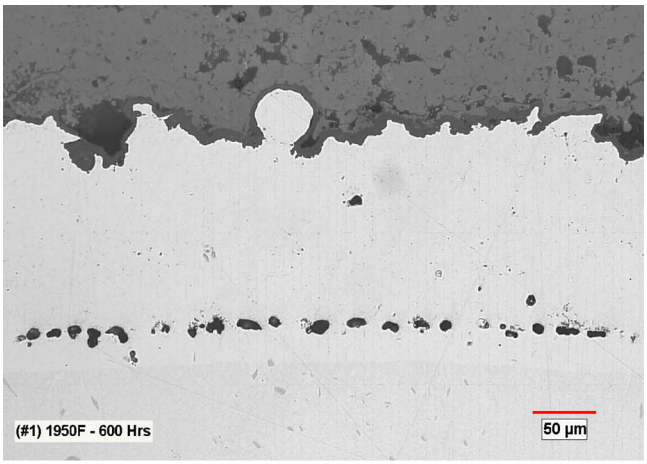
(*b*)

Fig. 13 Optical micrographs of the NiCoCrAlY bond coat/GTD-111 interface after (*a*) 2015 h and (*b*) 8155 h of exposure at 1010°C (1850°F). Note the variation of the interface voids with the exposure time.

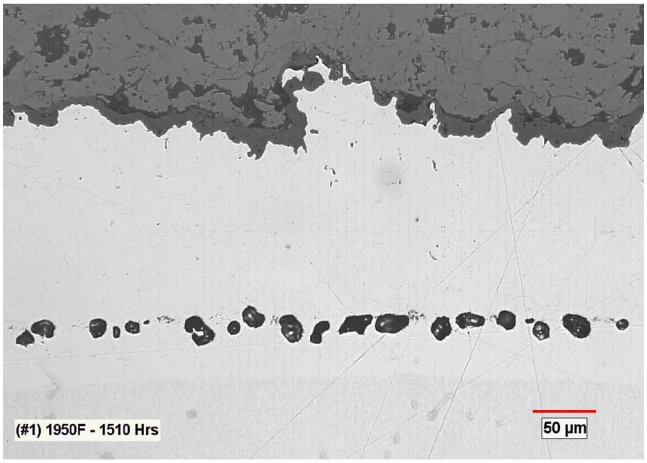
Conclusions

Long-term isothermal tests were conducted on the TBC-coated specimens at three temperatures, 1010°C (1850°F), 1038°C (1900°F), and 1066°C (1950°F) to determine the effects of long term exposure on the TBC failure location (mode). The exposed samples were destructively examined using optical microscopy and SEM attached with an energy dispersive spectroscopy (EDS) system. The following conclusions can be drawn from the results of this investigation.

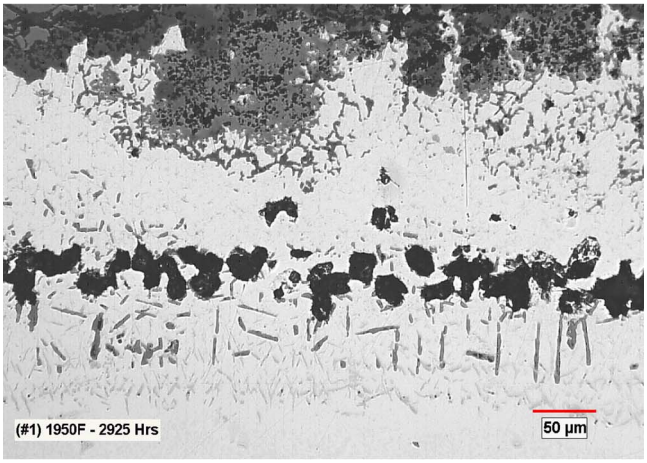
- Isothermal testing performed on TBC-CoNiCrAlY-coated GTD-111 specimens showed that the presence of the TBC accelerated the kinetics of TGO formation and bond coat degradation. The β-phase in the bond coat was completely consumed in a relatively short time after isothermal exposure.
- The presence of the TBC also accelerated the formation of mixed or transient oxides in the TGO. The TGO thickness at the bond coat/TBC interface increased with increasing exposure time.
- Delamination cracks were observed at the TGO/TBC interface. The extent and size of the cracking increased with the exposure temperature and time.
- The extent of internal oxidation of the bond coat was directly related to the exposure time and temperature.
- The extent of Kirkendall voids at the bond coat substrate also increased with increasing the exposure time and temperature, and the voids also coalesced with the exposure time.
- TBC spallation after long term exposure under isothermal conditions was associated with the oxidation and delamination of the bond coat from the substrate. Coalescence of Kirkendall voids at the interface is responsible for the delamination of the bond coat after long-term thermal exposure.



(*a*)



(*b*)



(*c*)

Fig. 14 Optical micrographs of the NiCoCrAlY bond coat/GTD-111 interface after (*a*) 600 h, (*b*) 1510 h, and (*c*) 2925 h of exposure at 1066°C (1950°F). Note the variation of the interface voids and void coalescence with the exposure time.

Acknowledgment

This work is funded by the DOE through Contract No. DE-FC626 01NT41321.

References

[1] Stecura, S., 1985, NASA Tech. Memo No. 86905, NASA, Cleveland, OH.
[2] Padture, N. P., Gell, M., and Jordan, E. H., 2002, "Thermal Barrier Coatings for Gas Turbine Engine Applications," Science, **296**, pp. 280–284.
[3] Bose, S., and Demasi-Marcin, J., 1995, "Thermal Barrier Coating Experience in Gas Turbine Engines at Pratt and Whitney," Thermal Barrier Coating Workshop, NASA Paper No. CP 3312 63.
[4] 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.
[5] Miller, R. A., 1987, "Current Status of Thermal Barrier Coatings—An Overview," Surf. Coat. Technol., **30**, pp. 1–11.
[6] Meier, S. M., Nissley, D. M., Sheffler, K. D., and Bose, S., 1992, "Thermal Barrier Coating Life Prediction Model Development," ASME J. Eng. Gas Turbines Power, **114**, pp. 259–267.
[7] Cheruvu, N. S., Chan, K. S., and Viswanathan, R., 2006, "Evaluation, Degradation, and Life Assessment of Coatings for Land-Based Combustion Turbines," Energy Materials, **1**(1), pp. 33–47.
[8] Miller, R. A., 1984, "Oxidation Based Model for Thermal Barrier Coating Life," J. Am. Ceram. Soc., **67**(8), pp. 517–521.
[9] Miller, R. A., 1987, "Progress Towards Life Modeling of Thermal Barrier Coatings for Aircraft Gas Turbine Engines," ASME Paper No. 87-ICE-18.
[10] Stangman, T. E., Liu, A., and Neumann, J., 1987, "Thermal Barrier Coating Life-Prediction Model Development," NASA Report No. CR-179648.
[11] DeMasi, J. T., Sheffler, K. D., and Ortiz, M., 1989, "Thermal Barrier Coating Life Prediction Model Development—Phase I," NASA Report No. CR-182230.
[12] Brindley, W. J., 1995, Proceedings of TBC Workshop, NASA Report No. CP 3312, pp. 189–202.
[13] Chang, G. C., Phucharoen, W., and Miller, R. A., 1987, "Thermal Expansion Mismatch and Plasticity in Thermal Barrier Coating," NASA Report No. CP-2493.
[14] Cruse, T. A., Stewart, S. E., and Ortiz, M., 1988, "Thermal Barrier Coating Life Prediction Model Development," ASME J. Eng. Gas Turbines Power, **110**, pp. 610–616.
[15] Yanar, N. M., Kim, G. M., Pettit, F. S., and Memer, G. H., 2002, Turbine Forum, Forum of Technology International Conferences on Advanced Coatings for High Temperature, Nice, France, Apr. 17–19.
[16] Meier, S. M., Nissley, D. M., and Sheffler, K. D., 1991, "Thermal Barrier Coating Life Prediction Model Development—Phase II," Report No. NASA CR-189111.
[17] Meier, S. M., Nissley, D. M., Sheffler, K. D., and Cruse, T. A., 1991, "Thermal Barrier Coating Life Prediction Model Development," ASME Paper No. 91-GT-40.
[18] Chan, K. S., Cheruvu, N. S., and Viswanathan, R., 2003, "Development of a Thermal Barrier Coating Life Model," ASME Paper No. GT2003-38171.
[19] Cheruvu, N. S., and Leverant, G. R., 1998, "Influence of Metal Temperature on Base Material and Coating Degradation of GTD-111 Buckets," *ASME International Gas Turbine and Aero Engine Congress and Exhibition*, Stockholm, Sweden, Jun., Paper No. 98-GT-5111.
[20] Cheruvu, N. S., Swaminathan, V. P., and Kinney, C. D., 1999, "Recovery of Microstructure and Mechanical Properties of Service Run GTD-111DS Buckets," ASME Paper No. 99-GT-425.
[21] Haynes, J. A., Ferber, M. K., Porter, W. D., and Rigney, E. D., 1999, "Characterization of Alumina Scales Formed During Isothermal and Cyclic Oxidation of Plasma-Sprayed TBC Systems at 1150°C," Oxid. Met., **52**, pp. 31–76.
[22] Haynes, J. A., Ferber, M. K., and Porter, W. D., 2000, "Thermal Cycling Behavior of Plasma-Sprayed Thermal Barrier Coatings With Various MCrATx Bond Coats," J. Therm. Spray Technol., **9**, pp. 38–48.
[23] Haynes, J. A., Ferber, M. K., Porter, W. D., and Rigney, E. D., 1996, "Isothermal and Cyclic Oxidation of an Air Plasma-Sprayed Thermal Barrier Coating System," ASME Paper No. 96-GT-286.
[24] Chan, K. S., and Cheruvu, N. S., 2004, "Degradation Mechanism Characterization and Remaining Life Prediction for NiCoCrAlY Coatings," ASME Paper No. GT2004-53383.
[25] Cheruvu, N. S., and Chan, K. S.,2004, "Combustion Turbine (CT) Hot Section Coating Life Management," EPRI Report No. 18.05112.
[26] Stiger, M. J., Yanar, N. M., Topping, M. G., Petitt, F. S., and Meier, G. H., 1999, "Thermal Barrier Coatings for 21st Century," Z. Metallkd., **90**, pp. 1069–1078.
[27] Steiger, M. J., Yanar, N. M., Petitt, F. S., and Meier, G. H., 1999, "Mechanism for the Failure of Electron Beam Physical Vapor DepositedThermal Barrier Coatings Induced by High Temperature Oxidation," *Elevated Temperature Coatings: Science and Technology TMS*, J. M. Hampikian and N. B. Dahotre, eds., TMS, Warrendale, PA, Feb. 1, pp. 51–62.

Co Huynh
Liping Zheng
Dipjyoti Acharya

Calnetix, Inc.,
Cerritos, CA 90703

Losses in High Speed Permanent Magnet Machines Used in Microturbine Applications

High speed permanent magnet (PM) machines are used in microturbine applications due to their compactness, robust construction, and high efficiency characteristics. These machines are integrated with the turbines and rotate at the same speeds. This paper discusses in detail the losses in high speed PM machines. A typical PM machine designed for microturbine application is presented with its detailed loss calculations. Various loss verification methods are also discussed. [DOI: 10.1115/1.2982151]

1 Introduction

Microturbines are small combustion turbines with typical outputs in the range of 20–500 kW. A typical system rotates over 40,000 rpm. One of the key enabling technologies for microturbines is the integral high speed electrical machines operating at the same speeds as the turbines, eliminating mechanical gearboxes. The result is a very compact high efficiency system that allows for ease of onsite installation. High speed permanent magnet (PM) machines are typically used in microturbine application due to their high power density and high efficiency characteristics. A good understanding of high speed PM machine characteristics, especially its losses, is critical to predict system performance and to ensure a reliable operation.

Losses in PM machines can be divided into three categories: (a) stator loss, (b) rotor eddy current loss, and (c) windage loss. The stator loss consists of copper loss and iron loss. The copper loss includes conventional I^2R loss and stray load loss due to skin effect and proximity effect. This can be calculated based on finite element analysis (FEA) or using analytical methods. The stator iron loss is divided into hysteresis loss, classical eddy current loss, and excess eddy current loss. Empirical equations or time-step transient FEA with motion can be used to calculate iron loss. The rotor loss generated by induced eddy currents in the steel shaft and permanent magnets is not significant compared with a machine's total loss. However, removing the heat from the rotor to ensure reasonable operating temperatures of its components is more difficult than removing the heat from the stator. Thus, an accurate prediction of rotor loss becomes important especially at high speed. The major causes of the rotor loss are (a) space harmonics due to the existence of stator slot opening and winding distribution and (b) time harmonics of the phase currents due to pulse width modulation (PWM). The rotor loss can be analyzed using analytical methods. However, simulations using FEA based on actual measured current waveforms or estimated current with total harmonics distortion (THD) provide a more accurate assessment. The windage loss as a result of the shearing action of the media that exists between the rotor and stator may also be significant at high speed, especially with small air gap and high cooling flow pressure in the air gap between the rotor and stator.

A spin-down test can be used to verify no-load loss, especially when the rotor's inertia is large. Testing of a back-to-back configuration can verify a machine's efficiencies and total losses at various loaded conditions. The thermal analysis model result compared with measured temperature mapping can also be used to verify machine losses.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 1, 2008; final manuscript received May 9, 2008; published online December 23, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

2 Permanent Magnet Alternator Losses

The losses in PM alternators are grouped into (a) stator loss, (b) rotor eddy current loss, and (c) windage loss.

2.1 Stator Loss. The stator loss consists of copper loss and iron loss. Copper loss is the loss due to the current going through the armature windings. The copper loss consists of I^2R loss and stray load loss. The I^2R loss is given by

$$P_{cu} = m_1 I^2 R \tag{1}$$

where m_1 is the number of phases, I is the armature current, and R is the dc armature resistance. The I^2R loss can be significant when large current flows through the conductor with large Ohmic resistance.

The stray loss comes from (a) the skin effect resulting from the same source conductors and (b) the proximity effect resulting from the field induced from adjacent conductors sharing the same slot. The skin effect is caused by electromagnetic induction in the conducting material, which opposes the currents set up by the wave E-field. The skin depth is the distance in which an electromagnetic wave entering a conducting surface is damped and reduces in amplitude by a factor of $1/e$, where e is equal to 2.71828…. The skin depth (δ) is given by

$$\delta = \sqrt{\frac{2}{\omega \mu_0 \sigma}} \tag{2}$$

where ω is the angular frequency of the current and σ is the electrical conductivity of the conducting material.

In designing the stator winding, wire strand size is selected such that skin depth is much larger than the wire radius to minimize loss due to skin effect.

Stator windings are contained inside slots. Loss due to proximity effects of conductors located in the slots of electric machines can be estimated based on the following equation [1]:

$$P_{stray} = P_{cu}(k_d - 1) \tag{3}$$

where

$$k_d = \varphi(\xi) + \left[\frac{m^2 - 1}{3} - \left(\frac{m}{2}\sin\left(\frac{\gamma}{2}\right)\right)^2\right]\psi(\xi) \tag{4}$$

$$\varphi(\xi) = \xi\frac{\sinh(2\xi) + \sin(2\xi)}{\cosh(2\xi) - \cos(2\xi)} \tag{5}$$

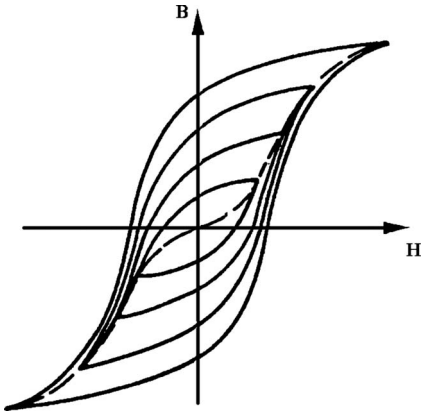


Fig. 1 Family of hysteresis loops

$$\psi(\xi) = 2(\xi)\frac{\sinh(\xi) - \sin(\xi)}{\cosh(\xi) + \cos(\xi)} \tag{6}$$

and ξ is the relative height of a conductor (the ratio of its height to the effective skin depth δ considering insulation). γ is the phase angle between the upper layer and lower layer currents for a two-layer case, m is the total number of identical conductors in layers, and kd is the average resistance coefficient, which is the ratio of effective ac resistance versus dc resistance.

The copper loss is temperature dependent, so the copper loss is calculated at the expected copper temperature. The copper I^2R loss increases when copper temperature increases due to increased winding resistance, while the copper stray load loss reduces with increased temperature. In addition to the analytical method described above, loss due to proximity and skin effects can also be simulated based on transient time-stepping FEA [2]. However, this method is very time consuming especially when multiple strands are used.

Iron loss produced in a magnetic material operating in an alternating magnetizing field is generally separated into two components: hysteresis loss and eddy current loss. Hysteresis loss is due to a form of intermolecular friction when a varying magnetic field is applied to the magnetic material. The loss per cycle is proportional to the area enclosed by the hysteresis loop on the B-H characteristics of the material. The hysteresis loss increases with the maximum magnetic field, as illustrated in Fig. 1.

The empirical formula expressing the hysteresis loss per unit volume (P_h (W/m^3)) in terms of the maximum flux density (B (T)) and frequency (f (Hz)) was developed by Steinmetz [3] as follows:

$$P_h = \eta B^n f \tag{7}$$

where η is a material constant and n is an exponent, which has typical values between 1.8 and 2.2, depending on the lamination material [4].

The term "eddy current" refers to circulating electric currents that are induced in a sheet of a conducting material when it is subjected to alternating magnetic field. These eddy currents produce power that is dissipated as heat. The eddy current loss per unit volume (P_e (W/m^3)), at frequencies which are low enough for the inductive effects to be neglected, is given by the following general equation [3]:

$$P_e = \frac{\pi^2 B^2 t^2 f^2}{\rho\beta} \tag{8}$$

where t is the thickness of the material (m), B is the peak flux density (T), ρ is the resistivity of the material (Ω m), and β is a coefficient that is related to the geometric structure.

Eddy current loss can also be divided into classical eddy current loss and excess eddy current loss for a more accurate analysis. Therefore, at a given frequency (f), the iron loss for electrical steel can be calculated from [5,6]

$$P_{\text{iron}} = k_h B^2 f + K_c (Bf)^2 + K_e (Bf)^{3/2} \tag{9}$$

where K_h, K_c, and K_e are the coefficients of hysteresis loss, classical eddy current loss, and excess eddy current loss, respectively, and B is the peak flux density. The coefficients can be calculated using the curve fitting of the iron loss data from manufacturers or from material test data.

The above equation is based on the assumption of sinusoidal excitation. When the iron core is subjected to the nonsinusoidal magnetic field, the distorted excitation, which can be represented by higher order harmonics, has to be considered [7,8]. The iron loss model considering arbitrary magnetic flux waveforms can be found in Ref. [9]. The transient time-stepping finite element method is also widely used to directly simulate the core loss [10,11].

Soft magnetic materials form the magnetic circuit in an electric machine. An ideal material would have high permeability in order to reduce the reluctance of the magnetic circuit, high saturation flux density in order to minimize the volume and weight of the iron core, and low losses. However, it is impossible to optimize all of these properties in a single material. This is because there are a large number of factors that affect magnetic properties (chemical composition, mechanical treatment, and thermal treatment are the most important), and the result is often a compromise. For example, nickel steel has low iron loss but low saturation flux density, while cobalt steel has higher saturation flux density but also higher iron loss. The iron core of a machine is made up of thin laminations in order to reduce core loss. Laminations as thin as 0.127 mm are generally used in high frequency applications in order to reduce iron loss.

Besides the laminated core, there are also two alternative materials [12]: (a) Amorphous metals (such as metglas), instead of their polycrystalline structure, have very low hysteresis and eddy current losses. Amorphous metals are produced by rapid cooling of alloys consisting of iron, nickel, and /or cobalt together with one or more of the following metalloids, which are elements or compounds exhibiting both metallic and nonmetallic properties: boron, silicon, and carbon. (b) Powder materials (such as grain-oriented electrical steels), in spite of their rather low core permeance, may be attractive for their very high frequency applications and also on account of their effective damping of vibrations.

2.2 Rotor Loss. The rotor loss generated by induced eddy current in the steel shaft and permanent magnets is not significant compared with the total machine loss. However, removing the heat from the rotor to ensure reasonable operating temperatures of its components is more difficult than removing the heat from the stator. Thus, an accurate prediction of rotor loss becomes important especially at high speed. The major causes of the rotor eddy current loss can be categorized into the following three groups: (a) no-load rotor eddy current loss caused by the existence of slots, (b) on-load rotor eddy current loss induced by the harmonics of windings' magnetomotive force (MMF), which is also called space harmonics, and (c) on-load rotor eddy current loss induced by the time harmonics of the phase currents due to PWM.

Loss due to eddy current, in general, can be expressed by

$$P = \int_V \sigma \mathbf{E}^2 dV = \int_V \mathbf{J}^2/\sigma dV \tag{10}$$

where σ is the material conductivity, $\mathbf{E}$ is the electric field, $\mathbf{J}$ is the eddy current density, and V is the volume of the material. For an accurate assessment, the rotor loss is generally simulated using a 2D FEA time-stepping transient solver with motion, in which the actual measured motor/generator current waveforms are applied. For alternators with a passive rectifier, the 2D FEA transient

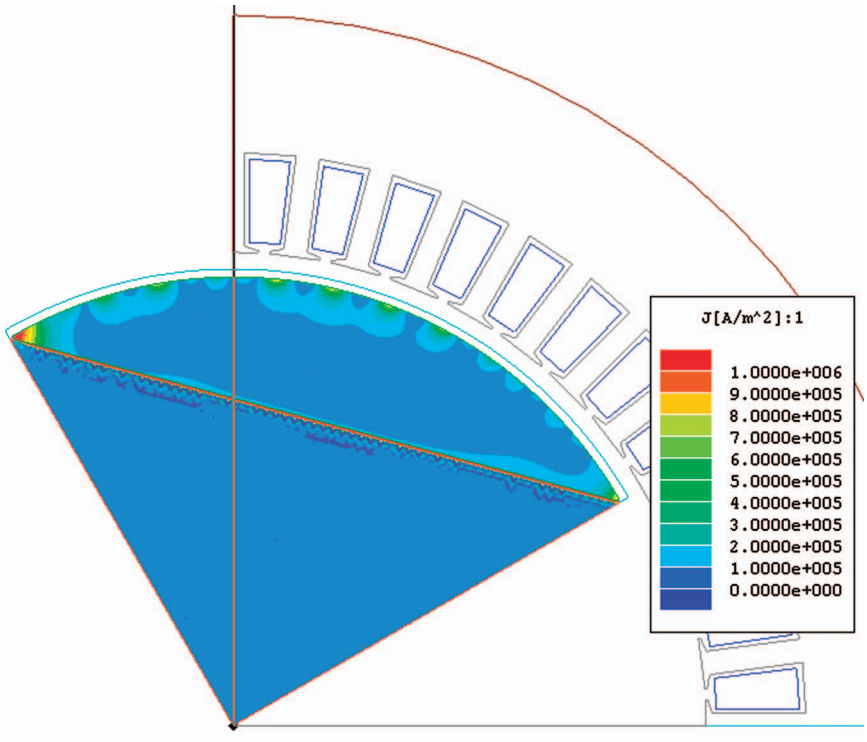


Fig. 2 Eddy currents in the rotor

solver with motion and external circuits can be used to simulate rotor loss and alternator performance. Figure 2 shows one example of eddy currents in the rotor caused by slot effect, winding space harmonics, and phase current time harmonics.

There are several methods to reduce rotor eddy current losses. Reducing the slot opening and increasing the magnetic gap between rotor and stator can reduce no-load rotor loss. Increasing the number of slots per pole and using fractional winding can reduce rotor loss caused by the space harmonics of the armature winding. Increasing the switching frequency and using external line inductors can reduce rotor loss caused by time harmonics of the phase currents. Since rotor loss caused by time harmonics is dominant in most applications, increasing the switching frequency and using external line inductance to reduce current THD is a very effective way to reduce rotor loss.

2.3 Windage Loss. Windage loss is heat generated in the fluid due to the relative motion (shearing) of the fluid that flows between the rotor and stator [13]. Windage loss, depending on various gases at various operating conditions, as used in high speed machines can be very high, contributing to overall machine inefficiency. The windage loss generation is a function of shaft rotational speed and fluid properties such as temperature, pressure, density, and temperature gradients at stator and rotor walls.

The windage loss generated in the clearance between a rotating cylinder and a stationary cylinder with homogenous laminar flow (no axial flow) can be estimated from the following system of equations [13].

Shaft rotational speed:

$$\omega = \frac{2\pi N}{60} \tag{11}$$

Reynolds number:

$$\mathrm{Re} = \omega r \frac{\rho}{\mu} \phi \tag{12}$$

Skin friction coefficient (C_d) for turbulent flow:

$$\frac{1}{\sqrt{C_d}} = 2.04 + 1.768 \ln(\mathrm{Re}\sqrt{C_d}) \tag{13}$$

Windage:

$$W = C_D \pi \rho \omega^3 r^4 \lambda \tag{14}$$

where N=rotational speed of rotor (in rpm), ρ=density of fluid (in kg/m^3), μ=kinematic viscosity of cooling media (in m^2/s), r=radius of the rotor (in m), φ=radial gap between rotor and stator (in m), and λ=length of the rotor (in m).

Theoretical relations and experimental validation taking into account the combination of axial flow and rotational flow, in the case of cooling media passing through the gap, can be included to obtain a better estimate. Also, surface roughness of the stator tooth and rotor surface affects windage loss and must be taken into account [14].

3 Typical Alternator Design and Loss Calculation

The detailed loss breakdown of a 120 kW alternator is provided as an example. The cross section of the alternator is shown in Fig. 3. Basic alternator performance parameters are shown in Table 1. The stator core is made of 0.178 mm silicon steel, and the length is 162 mm. The alternator output is connected to a passive rectifier shown in Fig. 4. A 2.86 Ω power resistor is used as an equivalent load to provide 120 kW output power.

The alternator performance was simulated using transient FEA with external circuit, and the result is summarized in Table 2. The phase current waveforms are also plotted in Fig. 5.

Based on Eq. (1), the copper I^2R loss is calculated to be 1167 W. According to Eq. (4), the resistance coefficient is 1.12, so the total copper loss is 1308 W (including 141 W of copper stray load loss). The air density at 77°C is 1.009 kg/m^3 and the viscosity is 2.075 m^2/s. According to Eqs. (11)–(14), the Reynolds number (Re) is equal to 14,179, the skin friction coefficient (C_d) is equal to 4.928×10^{-3}, and the windage loss is calculated to be 874 W.

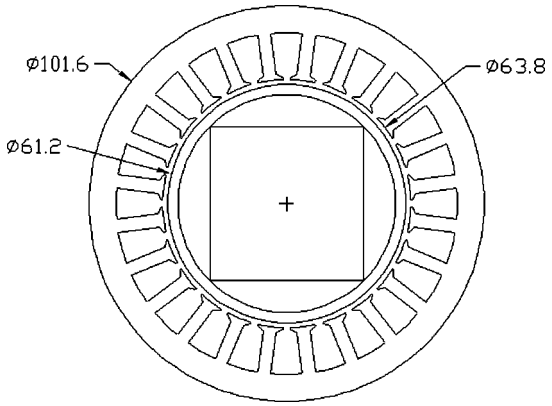


Fig. 3 The cross section of a 120 kW alternator (unit is in mm)

Based on the curve fitting of the manufacturer's data using a genetic algorithm, we get the coefficients of Eq. (9) as

$$K_h = 0.0275 \frac{W}{\text{Hz T}^2 \text{ kg}} \tag{15}$$

$$K_c = 1.83 \times 10^{-5} \frac{W}{\text{Hz}^2 \text{ T}^2 \text{ kg}} \tag{16}$$

$$K_e = 0.000277 \frac{W}{\text{Hz}^{1.5} \text{ T}^{1.5} \text{ kg}} \tag{17}$$

Table 1 Parameters of the motor/generator

Machine type	PM synchronous
Operating speed	70,000 rpm
Number of phases	3
Number of slots	24
Number of conductor /slot	6
Number of circuits	2
Winding resistance (line neutral)	16.3 mΩ @ 150°C (excluding leads)
Winding leakage inductance (per phase)	40 μH
D-axis magnetizing inductance (Lad)	39.3 μH
Q-axis magnetizing inductance (Laq)	54.6 μH
L-L Back EMF constant (V_{rms}/krpm)	8.26
Weight of core iron (kg)	2.37
Weight of teeth iron (kg)	1.20

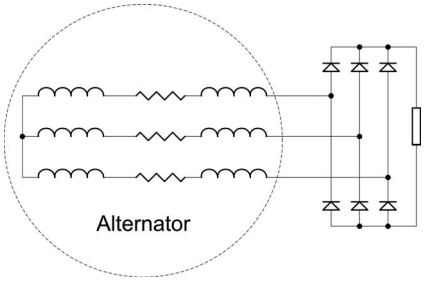


Fig. 4 Passive rectifier circuits

Table 2 Simulated performances

Output dc voltage (V)	583.7
Output dc current (A)	204.1
Output phase current (A_{rms})	154.5

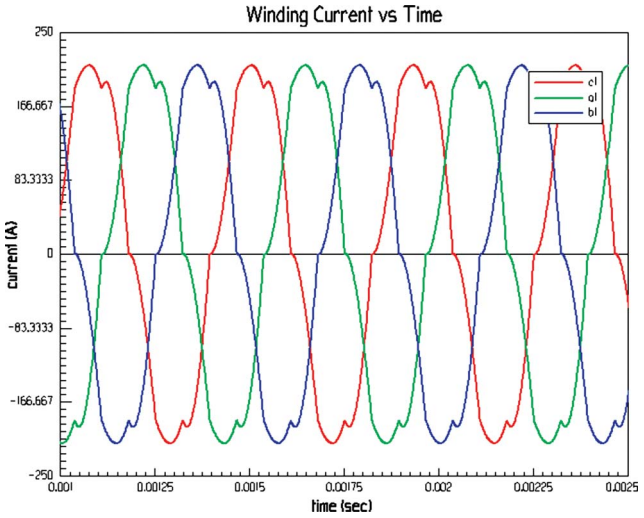


Fig. 5 Simulated current waveforms

According to FEA, the peak flux density is 1.45 T in the core and 1.17 T in the teeth. Therefore, according to Eq. (9), the core loss is calculated to be 945 W and teeth loss is computed to be 317 W. For rotor loss, the FEA simulation with a motor current waveform due to a passive rectifier was performed, and the rotor loss was found to be 221.9 W. The rotor loss due to current time harmonics is dominant because of high current harmonic distortions when using a passive rectifier. In summary, the losses of the high speed alternator at a rated load, excluding power electronics losses, are shown in Table 3.

4 Verification of Losses

Various methods can be used to verify alternator losses and operating efficiency.

4.1 Back-to-Back Test. For high speed machines, a back-to-back test configuration is most widely used to map machine performance at various load conditions from no load to full load. In this test, two identical machines are coupled together, with one operating as a motor and the other as a generator or load. Figure 6 shows a typical back-to-back test setup.

By measuring the input power to the motor and the output power from the generator, we can measure the total losses of the system that includes both units at various load conditions. Since the two machines are identical, we can calculate the total loss for each machine without introducing significant errors by compensating for small differences in winding losses due to the different operating current in each machine. For a more accurate assessment, a torque meter (if available) can also be used to directly measure the torque between the motor and generator.

Table 3 Summary of losses at a rated load

Stator loss	
Copper loss (W)	1308
Iron loss (W)	1262
Rotor loss	
Shaft loss (W)	15.9
PM loss (W)	206
Windage loss (W)	874
Total loss (W)	3666
Rated output power (W)	120,000
PM alternator efficiency (%)	97.0

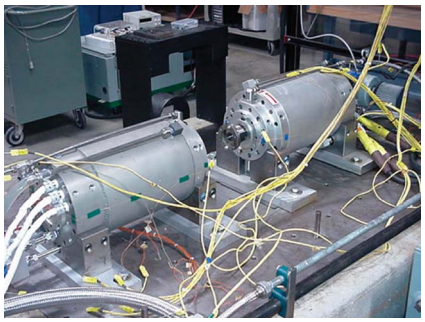


Fig. 6 Back-to-back test setup

4.2 Spin-Down Test. The relationship between the electromagnetic torque and the inertia torque is as follows:

$$T_{\text{em}} = J\frac{d\omega_r}{dt} + T_{\text{damp}} - T_{\text{mech}} \tag{18}$$

where T_{em} is the electromagnetic torque, T_{damp} is the damping torque in the direction opposite to rotation because of the friction, windage, and iron losses, T_{mech} is the externally applied mechanical torque in the direction of the rotor speed, and J is the rotor inertia.

A spin-down test is a method to measure the damping torque. When the motor is free spinning without externally applied mechanical torque and electromagnetic torque, Eq. (18) becomes

$$T_{\text{damp}} = -J\frac{d\omega_r}{dt} \tag{19}$$

Therefore, the motor losses can be calculated as

$$P_{\text{loss}} = -J \cdot \omega_r \frac{d\omega_r}{dt} \tag{20}$$

By measuring the initial speed versus time duration of the rotor from the initial speed to fully stop, the power loss can be calculated. When performing the spin-down test, any armature electrical loops should be disconnected to ensure that the motor is spinning down freely without any braking effect. If there is output power during the spin-down test, this power has to be considered in the calculation.

As an example, Fig. 7 shows the measured spin-down time versus rotor speed for a magnetic bearing supported high speed PM machine. The machine also produced output power of 140 W at 36,000 rpm, and 60 W at 8000 rpm during spin-down. The inertia of the entire rotating assembly is 0.68 kg m^2. Figure 8 shows the calculated total no-load loss versus rotor speed according to Eq. (20). The copper loss can be ignored in the calculation due to the very low output power.

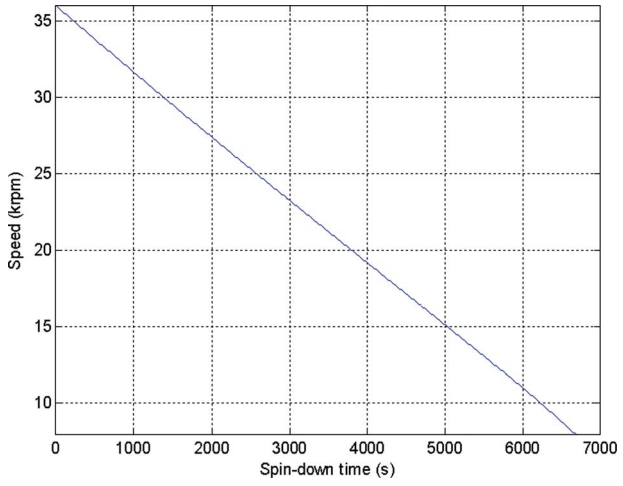


Fig. 7 Measured machine spin-down time versus rotor speed

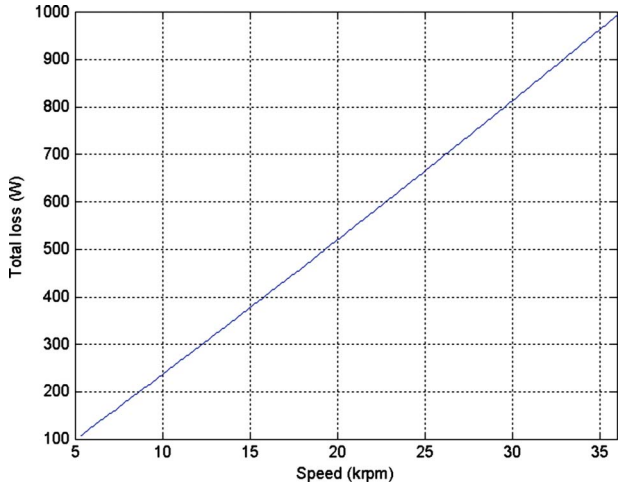


Fig. 8 Calculated no-load total loss versus speed

4.3 Temperature Mapping and Correlation With Thermal Model. In this technique, the machine is well instrumented throughout with temperature sensors such as thermocouples or thermistors. Temperature measurements are then made for various operating conditions with different loads and speeds. A finite difference thermal model of the machine is also constructed with calculated losses assigned for various components of the machine. Figure 9 shows a typical machine thermal model.

By adjusting losses in the thermal model so that predicted temperatures match the measured results for various operating conditions, we can estimate roughly different loss components of the machine. The validity of this technique depends significantly on the accuracy of the thermal model. Therefore this technique is often used as a secondary check.

5 Conclusion

High speed PM machines possess many desirable attributes that make them attractive for direct drive applications such as microturbines. In this paper, we describe the losses in high speed PM machines in detail. A typical PM machine designed for microturbine application is presented, and various verification methods are discussed. An accurate assessment of a machine's losses is critical in doing system design tradeoffs as well as in predicting system performance and ensuring a safe and reliable system operation.

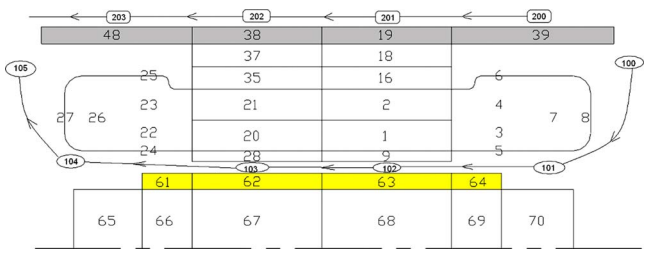


Fig. 9 A typical lumped parameter thermal model of a PM machine

References

[1] Lammeraner, J., and Stafl, M., 1966, *Eddy Currents*, CRC, Boca Raton, FL.

[2] Zheng, L., Wu, T. X., Acharya, D., Sundaram, K. B., Vaidya, J., Zhao, L., Zhou, L., Murty, K., Ham, C. H., Arakere, N., Kapat, J., and Chow, L., 2005, "Design of a Super-High Speed Permanent Magnet Synchronous Motor for Cryogenic Applications," *Proceedings of the IEEE International Electric Machines and Drives Conferences, IEMDC'05*, San Antonio, TX, May, pp. 874–881.

[3] Hamdi, E. S., 1994, *Design of Small Electrical Machines*, Wiley, New York.

[4] Mi, C., Slemon, G. R., and Bonert, R., 2003, "Modeling of Iron Losses of Permanent Magnet Synchronous Motors," IEEE Trans. Ind. Appl., **39**(3), pp. 734–742.

[5] Binesti, D., and Ducreux, J. P., 1996, "Core Losses and Efficiency of Electrical Motors Using New Magnetic Materials," IEEE Trans. Magn., **32**(5), pp. 4887–4889.

[6] Smith, A. C., and Edey, K., 1995, "Influence of Manufacturing Processes on Iron Losses," Proceedings of the IEE Electrical Machines and Drives Conference, Durham, UK, Sept., pp. 77–81.

[7] Slemon, G. R., and Liu, X., 1990, "Core Losses in Permanent Magnet Motors," IEEE Trans. Magn., **26**(5), pp. 1653–1655.

[8] Emanuel, E., 1988, "The Effect of Nonsinusoidal Excitation on Eddy Current Losses in Saturated Iron," IEEE Trans. Power Deliv., **3**(2), pp. 662–671.

[9] Roshen, W., 2007, "Iron Loss Model for Permanent Magnet Synchronous Motors," IEEE Trans. Magn., **43**(8), pp. 3428–3434.

[10] Chari, M. V. K., 1980, "Finite Element Analysis of Electrical Machinery and Devices," IEEE Trans. Magn., **16**(5), pp. 1014–1019.

[11] Shimoji, H., 2001, "Iron Loss and Magnetic Fields Analysis of Permanent Magnet Motors by Improved Finite Element Method With E&S: Model," IEEE Trans. Magn., **37**(5), pp. 3526–3529.

[12] de Jong, H. C. J., 1989, *AC Motor Design: Rotating Magnetic Fields in a Changing Environment*, Hemisphere, Chap. 2, pp. 14–15.

[13] Vrancrik, J. E., 1968, "Prediction of Windage Power Loss in Alternators," NASA-Langey, Report No. TND-4849.

[14] Becker, K. M., and Kaye, J., 1962, "Measurement of Diabetic Flowin Annulus With an Inner Rotating Cylinder," ASME J. Heat Transfer, **84**, pp. 97–105.

Misalignment in Gas Foil Journal Bearings: An Experimental Study

Samuel A. Howard
NASA Glenn Research Center,
Brook Park, OH 44135
e-mail: samuel.a.howard@nasa.gov

As gas foil journal bearings become more prevalent in production machines, such as small gas turbine propulsion systems and microturbines, system level performance issues must be identified and quantified in order to provide for successful design practices. Several examples of system level design parameters that are not fully understood in foil bearing systems are thermal management schemes, alignment requirements, balance requirements, thrust load balancing, and others. In order to address some of these deficiencies and begin to develop guidelines, this paper presents a preliminary experimental investigation of the misalignment tolerance of gas foil journal bearing systems. Using a notional gas foil bearing supported rotor and a laser-based shaft alignment system, increasing levels of misalignment are imparted to the bearing supports while monitoring temperature at the bearing edges. The amount of misalignment that induces bearing failure is identified and compared with other conventional bearing types such as cylindrical roller bearings and angular contact ball bearings. Additionally, the dynamic response of the rotor indicates that the gas foil bearing force coefficients may be affected by misalignment. [DOI: 10.1115/1.2966392]

Keywords: bearing, rotordynamics, foil bearing, gas bearing

1 Introduction

Gas foil bearings (GFBs) are currently used in several commercial applications, both terrestrial and aerospace. Aircraft air cycle machines (ACMs) and ground-based microturbines have demonstrated histories of successful long-term operation using GFBs [1]. Industrial blowers and compressors are becoming more common as well. Small aircraft propulsion engines, helicopter gas turbines, and high-speed electric motors are potential future applications.

In addition to the industrial and aeronautics applications already mentioned, of particular interest to NASA is the potential to use gas foil bearings in space nuclear power generation. One potential source of electrical power of sufficient magnitude and duration for space power is a nuclear fission based system. The system architecture would consist of a nuclear reactor heat source with the resulting thermal energy converted to electrical energy through a dynamic power conversion and heat rejection system.

A closed Brayton cycle (CBC), with an inert gas heated by the reactor driving a gas foil bearing turboalternator, is a potential power conversion scheme. The CBC for space power and propulsion is described in more details in the literature [2].

Regardless of the application, as gas foil bearings are considered for more turbomachinery applications, a better understanding of their system level characteristics is needed for successful design and integration into mainstream machinery. The system level characteristics not yet fully understood include thermal management schemes, alignment requirements, balance requirements, thrust load balancing, and others. This paper presents the results of an experimental effort to characterize gas foil bearing misalignment capabilities in the context of how precisely gas foil bearing machinery would need to be manufactured.

1.1 Technology Background. GFBs (Fig. 1) consist of an outer sleeve lined with a series of nickel-based superalloy sheet metal foils. The innermost sheet metal foil, or top foil, is smooth and constitutes the bearing inner surface against which the rotating shaft operates. The top foil is supported by a compliant structure, often made up of a layer of corrugated sheet metal foil referred to as bump foils, whose bumps behave like springs [3]. The bump foil layer gives the bearing flexibility that allows it to tolerate significant amounts of misalignment and distortion that would otherwise cause a rigid bearing to fail. In addition, microsliding between the top foil and the bump foil and the bump foil and the housing generates Coulomb damping, which can increase the dynamic stability of the rotor-bearing system [4]. Though not as common, other designs exist to achieve an elastic foundation for the compliant top foil, such as overlapping leaves, cantilevered springs, and others.

During normal operation of a foil bearing supported machine, the rotation of the rotor generates a pressurized gas film that "pushes" the top foil out radially and separates the top foil from the surface of the rotating shaft. The pressure in the gas film is proportional to the relative surface velocity between the rotor and the foil bearing top foil. Thus, the faster the rotor rotates, the higher the pressure, and the more load the bearing can support. When the rotor first begins to rotate, the top foil and the rotor surface are in contact until the speed increases to a point where the pressure in the gas film is sufficient to push the top foil away from the rotor and support its weight. Likewise, when the rotor slows down to a point where the speed is insufficient to support the rotor weight, the top foil and rotor again come in contact. Therefore, during startup and shutdown, a solid lubricant coating is used, either on the shaft surface or the foil, to reduce wear and friction [5].

The tolerance to misalignment mentioned above is as yet an unknown for GFBs. Manufacturers claim GFBs can handle large misalignment due to the flexible nature of the inner bearing surface, but there is little experimental or analytical verification of this capability. In an effort to address these concerns, the effects of misalignment are investigated experimentally in the current work.

2 Misalignment

An important issue for all turbomachinery and therefore important for gas foil bearing systems is the degree of misalignment that a rotor-bearing system can tolerate. The limits on misalignment dictate how precisely housings must be manufactured or whether bearings must be mounted in self-aligning fixtures and how much

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 28, 2008; final manuscript received April 9, 2008; published online December 19, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

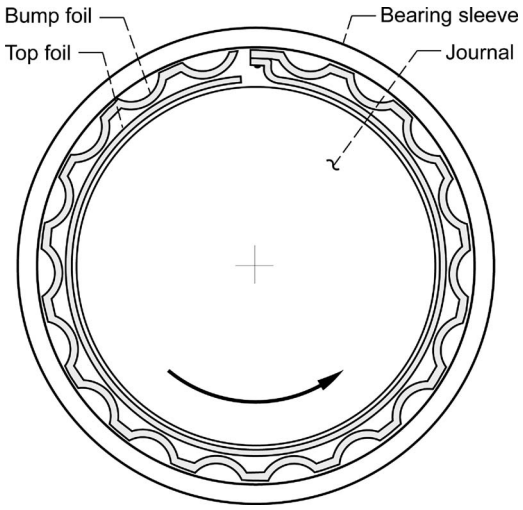


Fig. 1 Gas foil bearing schematic

thermal distortion can be tolerated. GFBs have been touted as having the ability to handle high degrees of misalignment relative to other bearing types, thus making them easier to integrate into high-speed high temperature applications. However, very little work has been reported on GFB misalignment. Carpino [6] computationally analyzed misalignment effects and found that minimum film thickness decreases with increased misalignment. Those results were for very small angular misalignments (maximum angle equal to the clearance/length radians) relative to those in the present tests (maximum angle approximately ten times larger, assuming a clearance on the order of 25 μm), yet make clear a concern associated with misalignment. The viscous losses in the bearing, which lead to power loss and heat generation, are proportional to the viscosity of the gas, and the wall shear stress of the fluid film (which varies inversely with the gas film thickness). Therefore, smaller minimum film thickness, and therefore misalignment, leads to higher power loss and greater heat generation. Previously, Gray et al. [7] reported misalignment capability of four different foil bearing configurations. They tested a "standard bearing," which is equivalent to what is referred to as a Generation I foil bearing by Dellacorte and Valco [8], a two-layered bump bearing with a second layer of bumps at the edge, a three-bump strip bearing, which has three separate bump strips axially (likely a Generation II bearing), and a standard bearing mounted in a diaphragm type flexible mounting. Their results indicate that for a load of 55.1 kN/m^2, the misalignment tolerance for the Generation I and the two-layered bearings is similar, approximately 1.0×10^{-3} rad. The three-bump strip bearing is roughly 2.6×10^{-3} rad, and the flexibly mounted Generation I bearing is roughly 7.0×10^{-3} rad. It is difficult to compare these results to the current results because the bearing sizes, speeds, and loads are different. However, that said, the magnitude of the misalignment tolerated is of the same order regardless of size, speed, and load, and more importantly, all are at least an order of magnitude higher than angular contact ball bearings.

Since angular contact ball bearings are typically used in the class of machines for which GFBs are considered as replacements, it is appropriate to compare the misalignment tolerance of the two. The amount of misalignment ball bearings can tolerate depends on the size, load, speed, and required life, but according to Zaretsky [9], a typical allowable maximum angle of misalignment for angular contact ball bearings is 3×10^{-3} rad. The current test program was instigated to quantify the level of misalignment GFBs can tolerate for comparison and to help guide future oil-free turbomachinery engineering design programs.

2.1 Experimental Setup. The rotordynamic simulator test rig at NASA Glenn Research Center (Fig. 2) [10] was used to conduct misalignment tests on two journal GFBs. The specific geometry of the individual bumps is proprietary and the reader is referred to the patent for more details on the design [3]. However, the bearings are classified as Generation III bearings as defined in Ref. [8] and the basic geometry used is listed in Table 1.

The rotordynamic simulator rig is an air turbine driven test rig

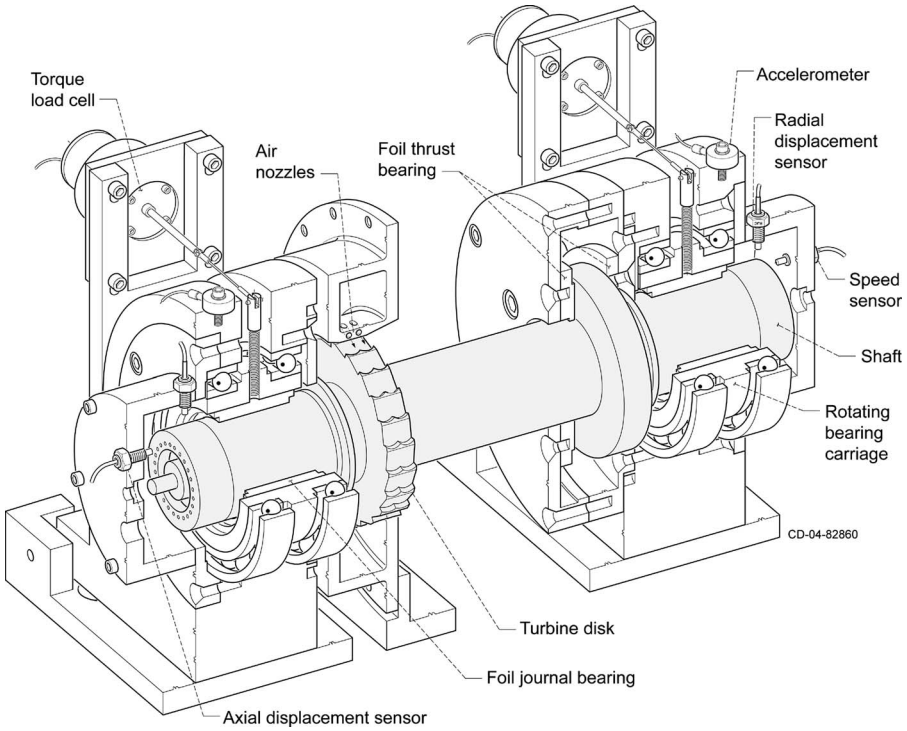


Fig. 2 Cutaway view of the rotordynamic simulator test rig

Table 1 Gas foil bearing geometry

Nominal shell outside diameter	63.5 mm
Nominal shell inside diameter	52.6 mm
Nominal shaft diameter	50.8 mm
Nominal length	40.6 mm
Number of circumferential bumps	35

that features completely oil-free operation. There are two journal GFBs at opposite ends of the rotor with two disks (a turbine and thrust bearing runner) mounted between the two journal bearings. In the configuration shown, the rotor weighs approximately 31 N (7.0 lbs) and is roughly symmetric such that each journal bearing supports half the rotor weight for a loading of 6.0 kN/m^2. The journal bearings are housed in independent structures that can be moved relative to each other in transverse and angular directions. The independent bearing supports allow the operator to impose a known misalignment on the two journal bearings using a laser-based alignment system. Each foil bearing is mounted in a ball bearing supported rotating bearing carriage, as illustrated in Fig. 2. The alignment system, shown in Fig. 3, consists of two laser heads, each having a laser emitter and sensor. One head attaches to each rotating bearing carriage. The bearing carriages rotate through 180 deg, and at three distinct angular positions, the laser alignment heads take measurements: ± 90 deg from top dead center and top dead center. From these three measurements, the lateral and angular alignments of the central axis of the rotating bearing carriages are calculated. The measurement accuracy of the laser alignment system is $\pm 1.0\%$.

For the initial series of tests described here, one bearing structure was held fixed while the other was sequentially moved a small amount laterally (0.127 ± 0.001 mm for each test) imposing a misalignment to both bearings. Two physical quantities were measured during these tests, temperature and coast-down time. The temperature was measured using type T thermocouples (open ball, 40 gauge, with a temperature range of $-100°C$–$400°C$ and $\pm 0.08\%$ accuracy) glued with thermally conductive epoxy to the underside of the bearing top foil close to the edge, ± 90 deg from top dead center, and at both ends for a total of eight thermocouples, four on each bearing (Figs. 4 and 8). Each test was run until the temperatures at a given test condition stabilized, indicating steady state operation. Data were collected at 20,000 rpm, which is a convenient speed to run the test rig in terms of vibration, noise, and time to reach steady state. The expected result was to see overall higher temperatures and higher temperature gradients between opposite sides of the bearings (T2 versus T1, for example) at higher misalignment levels. The other observed quantity, coast-down time, was used as an indirect measure of bearing torque. It was anticipated that as the level of misalignment increased, the coast-down time would decrease due to the higher bearing losses.

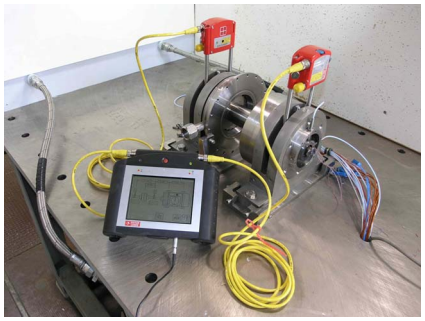


Fig. 3 Photograph of the test rig with the laser alignment system in place

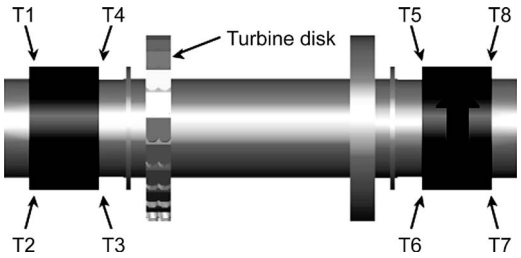


Fig. 4 Lateral misalignment method showing thermocouple locations looking down from the top. The left bearing is stationary. The right bearing moves laterally 0.127 mm for each incremental test.

In an effort to determine if the method of misalignment has a significant effect, additional tests were conducted using a different set of bearings without thermocouples. In these tests, two misalignment methods were used. The first method imposed a misalignment by holding both bearings fixed laterally and by rotating one bearing, as illustrated in Fig. 5. In the second misalignment method, the same technique was used as the previous bearings, a lateral misalignment like the previous tests illustrated in Fig. 4. A different set of bearings was used without thermocouples because the thermocouples were extremely fragile and easily damaged. Additionally, as discussed later, there was a concern that the thermocouples affected the behavior of the bearings and therefore may have influenced the results to some extent. Since there were no thermocouples on the bearings, it was not possible to determine when steady state temperature operation was reached. Therefore, each test was run for 40 min, which based on the previous tests with thermocouples is ample time to reach steady state operation. These tests were simply a go/no go type of test to determine a misalignment limit and if there was a difference between the two misalignment methods.

In addition, during this second set of tests, the displacements of the rotor in the vertical and horizontal directions at each end of the rotor were measured using eddy current displacement sensors to see if misalignment has an observable effect on the dynamics of the system.

3 Misalignment Results and Discussion

3.1 Temperature Results. Figure 6 shows the temperature data for steady operation at 20,000 rpm. In general, there is an upward trend on temperature with higher misalignment, as anticipated. The exceptions to the trend are T3 and T4. One possible cause of the downward trend in T3 and T4 is their proximity to the turbine, as seen in Fig. 4. Since T3 and T4 are adjacent to the turbine, they are affected by the temperature of the turbine outlet flow. The turbine is driven by compressed air, and as the air expands through the nozzle, it cools. At higher misalignment, more flow is required to counteract the higher torque, resulting in more turbine exhaust. The increased turbine exhaust may cool the bear-

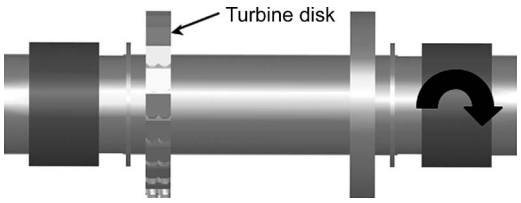


Fig. 5 Angular misalignment method. The left bearing is stationary. The right bearing moves angularly 5.0×10^{-4} rad for each incremental test.

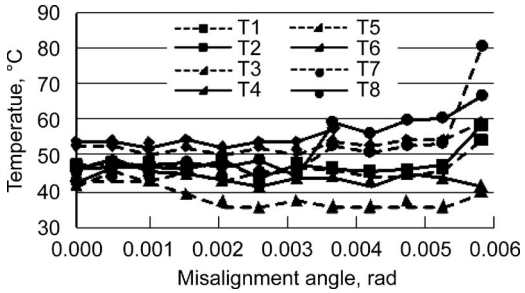


Fig. 6 Foil bearing edge temperatures at 20,000 rpm

ing in the location where T3 and T4 are mounted.

When the GFBs are misaligned, the bearings become more and more edge loaded. The thermocouples were placed on the bearing in such a way to try to see the effect of the edge-loading in the form of increased temperature. For example, as the right hand bearing in Fig. 4 is moved upward in the picture, the film thickness near Thermocouples T2, T4, T6, and T8 decreases, while at T1, T3, T5, and T7, it increases. Because thinner films are associated with higher heat generation, T2 should be hotter than T1, T4 hotter than T3, T6 hotter than T5, and T8 hotter than T7. In general, this result is observed. It should be noted that Thermocouple T6 was damaged while increasing the misalignment after the run at 3.7×10^{-3} rad of misalignment, so there are no data for T6 beyond that test. Up to that point, it was behaving as expected.

At a maximum misalignment of 5.8×10^{-3} rad $\pm0.07\times10^{-3}$ rad (1.40 mm $\pm$ 0.01 mm over a 241 mm $\pm$ 1 mm span) and 20,000 rpm, a failure was experienced and testing was stopped. GFBs can fail in several ways [11,12], but this failure was typical of the failure mode seen when load capacity is reached. The torque in the bearings increases rapidly, accompanied by an increase in temperature with no increase in speed or load. At the onset of the failure, it was observed that more turbine pressure was needed to maintain the same speed. As the failure progressed, more turbine pressure could not overcome the increase in torque, and speed decreased even with more pressure. When this occurred, the test was stopped.

In the second set of tests, it was not possible to reach failure due to the misalignment with either the angular or the translational misalignment. Thus, it was not possible to determine if the misalignment method had any affect on the rotor/bearing system. The maximum amount of misalignment capable in the test rig, 8.0×10^{-3} rad, is dictated by a clearance hole around the bolts that hold the bearing structures to the table. At 8.0×10^{-3} rad, the clearance around the bolts was reached and the bolts were contacting the structure. For future testing, modifications will be considered to enable more misalignment. Still, the preliminary result obtained here of no failure up to at least 8.0×10^{-3} rad is a notable finding in that it is a very large amount of misalignment in terms of actual hardware. This is equivalent to 8 μm/mm of misalignment. Standard precision machining and/or high thermal distortion should not cause this much misalignment. For example, in the test rig with a 241 mm bearing span, in order to induce misalignment that severe, the bearings had to be laterally misaligned by 1.93 mm. This result indicates that gas foil bearings can tolerate extreme misalignment without catastrophic failure. Additionally, there are other reasons to avoid high levels of misalignment that make it unnecessary to test further. While the bearings did not fail, the coast-down time decreases with more misalignment as seen later in Figs. 9 and 10, indicating that torque levels are increasing. Thus, power loss will increase with misalignment causing an efficiency penalty, increased thermal management requirements, and large startup power demands. Increased wear of surface coatings with misalignment is also a concern. Figures 7(*a*) and 7(*b*) show the test rig rotor before testing began and after final disassembly, respectively. Figure 8 shows two of the test bearings

Fig. 7 Photographs of the rotor before and after misalignment tests showing edge wear in the shaft coating in photo b. (*a*) Rotor before misalignment tests and (*b*) rotor after misalignment tests.

after misalignment testing. The wear at the edges of the bearing locations on the rotor and at the edges of the bearings is evident. Additionally, other components in the turbomachine cannot tolerate such extreme misalignment levels. The tip clearances in the compressor and turbine of such machines would need to increase, resulting in efficiency losses.

The results of this second set of tests give an indication of another reason for not instrumenting the second set of bearings with thermocouples. The bearings with thermocouples failed at a misalignment of 5.8×10^{-3} rad, while the uninstrumented bearings did not fail at 8.0×10^{-3} rad. The reason for this is believed to be the relative tightness of the bearings. The bearings with the

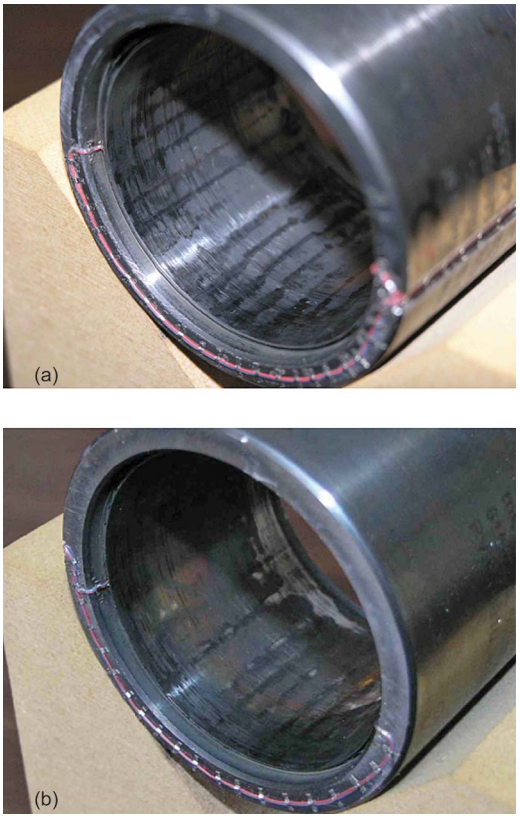


Fig. 8 Sample photographs of the test bearings after misalignment tests showing top foil wear patterns and thermocouple placement. (*a*) Turbine end bearing and (*b*) compressor end bearing.

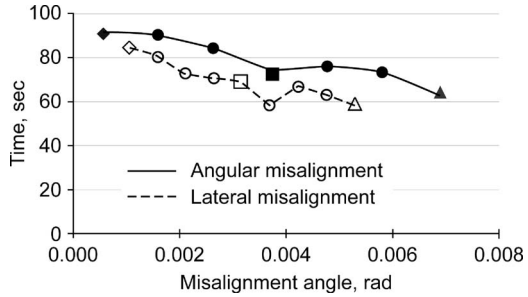


Fig. 9 Coast-down time as a function of misalignment

thermocouples were noticeably tighter than the bearings without thermocouples. In addition to manufacturing variability between the different bearings, the process of installing the thermocouples could have caused a decrease in the clearance of the bearings. Previous work has shown that decreased clearance can have the effect of decreasing a bearing's load capacity [13]. In a similar manner, a decrease in clearance, or an increase in tightness, apparently can decrease misalignment tolerance as well. This observation along with the observation that the failure mode was identical to a load capacity failure leads one to believe that load capacity and misalignment may combine in an additive fashion to contribute to bearing failure. In other words, it is probable that a perfectly aligned bearing has a certain load capacity at a given speed. Any misalignment would result in a decrease in load capacity. Similarly, a given bearing can tolerate less misalignment the more heavily it is loaded relative to its load capacity. More work is needed to determine this relationship, but it is believed to be the cause of one bearing set failing and the other not failing. In any case, the second set of bearings was not instrumented to eliminate the possibility of affecting the results.

3.2 Coast-Down Results. Figure 9 shows the coast-down time data as a function of misalignment, for both the lateral and the angular misalignment tests. The time represents how long it took to coast to a stop from a speed of 25,000 rpm. The trend shows that there is a general decrease in the time it takes to coast to a stop from a fixed speed as the misalignment increases. The decrease in coast-down time can be attributed to increased power loss or torque in the bearings. It was not known a priori how severe the wear would be at the edge of the bearings and in the shaft coating for each shutdown event. Therefore, in an effort to minimize the amount of wear to the rotor/bearing system, the coast-down test was conducted once for each misalignment condition. While the repeatability is unknown, the results do show a reasonable trend. Since the coast-down tests were only intended to verify the assertion that torque increases with misalignment, this was deemed an acceptable compromise.

Modifications are planned for the rig to enable direct torque measurements in the future. Still, quite a bit of insight can be gained by looking at the coast-down data. Figure 10 shows the same coast-down test results plotted a different way. Here, the speed versus time is shown over the entire coast-down event. In Figs. 10(*a*) and 10(*b*), three misalignment conditions are shown: small, intermediate, and large misalignment cases in each figure. Figure 10(*a*) is from the lateral misalignment test and Fig. 10(*b*) from the angular misalignment test. The shape of the data points corresponds with the shape of the data points in Fig. 9 for cross-reference. The slope of these plots is the angular acceleration of the test rotor. One can see that the rotor decelerates at about the same rate initially for all the cases but slows down quicker for the misaligned cases near the end of the coast-down event. One can also see the effect of misalignment on what is termed the lift-off speed (or touch-down speed in the case of a coast-down) of the foil bearings. The lift-off speed is the speed at which a gas film has fully developed and the rotor weight is supported on the gas

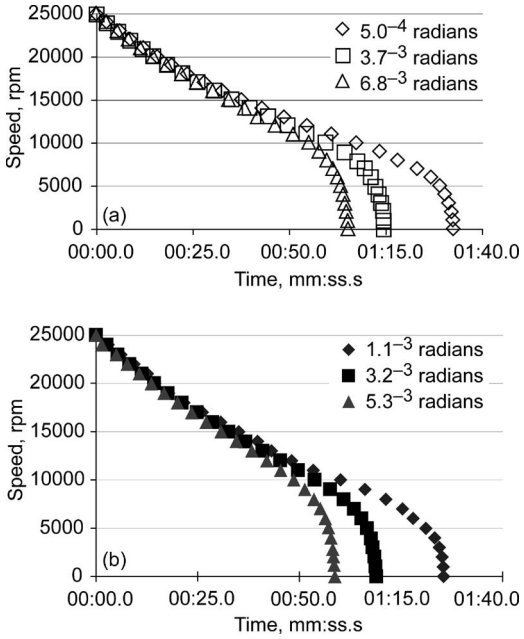


Fig. 10 Plots of the evolution of speed during coast-down for three lateral and angular misalignment conditions. (*a*) Speed versus time for coast-down at three lateral misalignment levels (*b*) speed versus time for coast-down at three angular misalignment levels.

film. It makes sense in the context of coast-down testing to think in terms of touch-down speed, but they are essentially interchangeable for the purpose of this discussion. Below the touch-down speed, there is still a metal-to-metal contact and therefore the torque is higher. Therefore, at the touch-down speed one would expect to see a significant change in the deceleration of the rotor due to a sudden increase in torque. This can, in fact, be seen in Figs. 10(*a*) and 10(*b*) and manifest itself as the knee in the curves. Additionally, as the misalignment increases in both plots, the touch-down speed increases. Thus, increased misalignment causes an increase in touch-down, or lift-off speed in much the same manner as increased load would cause an increase in lift-off speed.

3.3 Rotordynamic Results. The final observation of this series of testing is the rotordynamic behavior of the system as it changes with increased misalignment. The observations are preliminary but are noteworthy and deserving of further investigation. As the angle of misalignment increases, the results are consistent with a stiffening of the bearing. This result is not altogether surprising and complements the results of Carpino [6] mentioned earlier. Since increased misalignment leads to smaller minimum film thicknesses, which result in higher gas film pressures, it stands to reason that increased misalignment would lead to increased stiffness. The indication of this can be seen in Fig. 11 with three plots of the vertical vibration at the misaligned bearing end of the rotor as a sample. The figure shows synchronous vibration (Bode) plots of low, medium, and high misalignments from the top down: 5.0×10^{-4} rad, 3.7×10^{-3} rad, and 6.8×10^{-3} rad, respectively. One can see what looks like two potential critical speeds in the top plot at 4500 rpm and 12,500 rpm. In the middle and bottom plots, these peaks move up to approximately 5600 rpm and 13,500 rpm and 9000 rpm and 15,000 rpm, respectively. This increase in frequencies indicates the possible stiffening of the bearing one might expect. In addition, the shape, specifically the width of the peaks, changes as the misalignment increases indicating a possible change to the damping coefficients as well. Further testing is needed to verify that these peaks are critical speeds

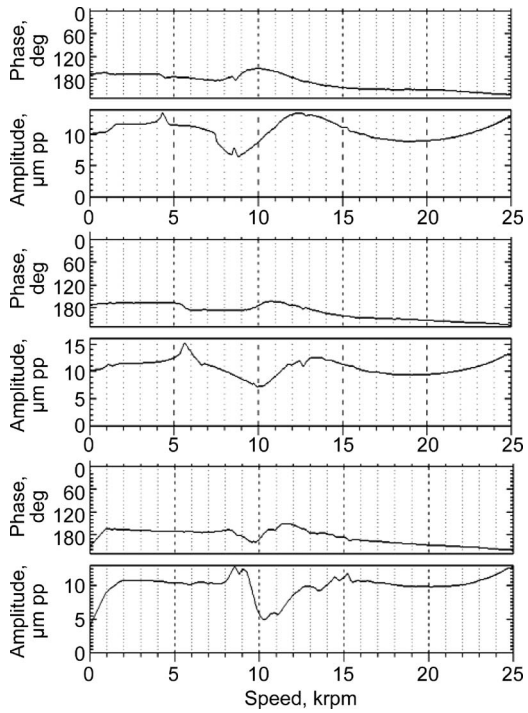


Fig. 11 Plots of the synchronous vertical vibration amplitude (Bode plot) at the angular misaligned bearing end of the rotor. The misalignments are 5.0×10^{-4} rad, 3.7×10^{-3} rad, and 6.8×10^{-3} rad, respectively from the top down.

and corresponding mode shapes of the rotor/bearing system. If they are critical speeds, the increase in frequency and width are consistent with an increase in bearing stiffness and damping.

Another rotordynamics observation relates to subsynchronous vibrations. Figure 12 shows three waterfall plots of the same vibration variable as Fig. 11 for coast-down events at the same three

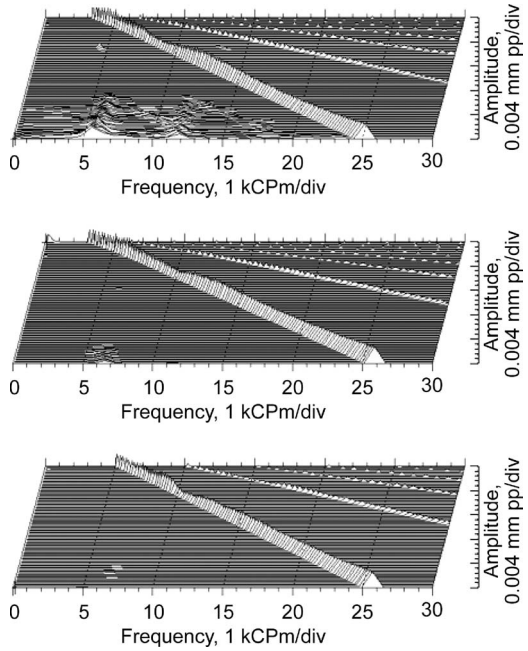


Fig. 12 Waterfall plots for increasing levels of misalignment, 1.0×10^{-3} rad, 3.7×10^{-3} rad, and 6.8×10^{-3} rad, respectively from the top down.

levels of misalignment. At the lowest misalignment, the top plot, there are subsynchronous components in the vibration that begin to show up at around 17,000 rpm. As the misalignment increases in the middle plot to 3.7×10^{-3} rad, the speed at which the subsynchronous vibrations begin to appear increases to around 22,000 rpm. At the highest misalignment, 6.8×10^{-3} rad, there are essentially no subsynchronous vibrations up to the maximum speed of 25,000 rpm. While more analysis is needed to determine the source of the subsynchronous vibrations, it is interesting to note that they occur near the frequencies of the vibratory peaks observed in Fig. 11. San Andres et al. [13] have reported similar subsynchronous vibrations with the exception that they found the subsynchronous frequency to track the run speed at a ratio of 50%, where here the frequencies seem to be fairly constant and possibly locked into the rigid body critical speeds of the rotor. This result, an increase in the speed at which subsynchronous vibrations appear, is consistent with the observed widening of peaks above as an increase in bearing damping. Again, additional testing and future analysis are needed to fully understand this outcome, but it is clear that misalignment can have an effect on system rotordynamics.

4 Conclusions

Future widespread use of gas foil bearings in high performance turbomachinery will demand an understanding of the importance of and engineering design guidelines for important system level considerations, such as thermal management schemes, alignment precision, balance requirements, thrust load balancing, and others. The present experimental results indicate that gas foil bearings are quite tolerant of high levels of misalignment, at least an order of magnitude higher than angular contact ball bearings, the technology they most often replace. However, because of other system requirements, such as a desire to keep bearing power loss and wear low, and compressor and turbine tip clearance requirements, allowable misalignment limits for a given machine are likely to be much lower than what the GFB could handle. The implication of this, from an engineering perspective, is that present design practices regarding machine precision and stackup tolerances should be sufficient for integrating GFBs into future high performance turbomachinery. The necessity of special design practices, such as flexible self-aligning bearing mounts, is not likely to be required. In addition, while the full implications are yet to be discovered, GFB misalignment appears to have an influence on system rotordynamics in terms of bearing force coefficients and the onset of subsynchronous vibrations.

Acknowledgment

The author wishes to acknowledge the support of NASA's Fundamental Aeronautics Program Office, Subsonic Rotary Wing Project.

References

[1] Heshmat, H., Walton, J. F., DellaCorte, C., and Valco, M. J., 2000, "Oil-Free Turbocharger Demonstration Paves the Way to Gas Turbine Engine Applications," ASME Paper No. 2000-GT-620.
[2] Mason, L. S., 2003, "A Power Conversion Concept for the Jupiter Icy Moons Orbiter," Paper No. AIAA-2003-6007.
[3] Heshmat, H., 2000, "High Load Capacity Compliant Foil Hydrodynamic Journal Bearing," U.S. Patent No. 6,158,893.
[4] Heshmat, H., 1994, "Advancements in the Performance of Aerodynamic Foil Journal Bearings: High Speed and Load Capability," ASME J. Tribol., **116**, pp. 287–295.
[5] DellaCorte, C., 1998, "The Evaluation of a Modified Chrome Oxide Based High Temperature Solid Lubricant Coating for Foil Gas Bearings," National Aeronautics and Space Administration, Report No. NASA/TM-1998-208660.
[6] Carpino, M., Peng, J. P., and Medvetz, L., 1994, "Misalignment in a Complete Shell Gas Foil Journal Bearing," Tribol. Trans., **37**(4), pp. 829–835.
[7] Gray, S., Heshmat, H., and Bhushan, B., 1981, "Technology Progress on Compliant Foil Air Bearing Systems for Commercial Applications," *Proceedings of the Eight International Gas Bearing Symposium*, BHRA Fluid Engineering, Cranfield, Bedford, England, April 1981, pp. 70–97, Paper No. 6.
[8] Dellacorte, C., and Valco, M. J., 2000, "Load Capacity Estimation of Foil Air

Journal Bearings for Oil-Free Turbomachinery Applications," National Aeronautics and Space Administration, Report No. NASA/TM-2000-209782.
[9] Zaretsky, E. V., 1992, *Life Factors for Rolling Bearings*, Society of Tribologists and Lubrication Engineers, Park Ridge, IL, Chap. 4, pp. 145–158.
[10] Howard, S. A., 2007, "A New High-Speed Oil-Free Turbine Engine Rotordynamic Simulator Test Rig," National Aeronautics and Space Administration, Report No. NASA/TM-2007-214489.
[11] Dykas, B. D., and Howard, S. A., 2004, "Journal Design Considerations for Turbomachine Shafts Supported on Foil Air Bearings," Tribol. Trans., **47**(4), pp. 508–516.
[12] Radil, K. C., Howard, S. A., and Dykas, B. D., 2002, "The Role of Radial Clearance on the Performance of Foil Air Bearings," National Aeronautics and Space Administration, Report No. NASA TM-2002-211705.
[13] San Andrés, L., Rubio, D., and Kim, T. H., 2007, "Rotordynamic Performance of a Rotor Supported on Bump Type Foil Gas Bearings: Experiments and Predictions," ASME J. Eng. Gas Turbines Power, **129**(3), pp. 850–857.

B. Ertas
e-mail: ertas@research.ge.com

M. Drexel
e-mail: drexel@crd.ge.com

J. Van Dam
e-mail: vandam@research.ge.com

D. Hallman
e-mail: hallman@crd.ge.com

Rotating Equipment Group,
Vibration and Dynamics Laboratory,
GE Global Research Center,
1 Research Circle,
Niskayuna, NY 12309

A General Purpose Test Facility for Evaluating Gas Lubricated Journal Bearings

The present work describes the detailed design and operational capabilities of a general purpose test facility developed to evaluate the dynamics and performance of gas lubricated journal bearings. The component level test facility was developed to serve as an initial tollgate test platform for certifying gas lubricated journal bearings into aircraft engine applications. A rotating test rig was engineered to test 70–120 mm diameter bearings at 40,000–80,000 rpm and 1200°F. The test rig described in this paper possesses design elements that enable the simultaneous application of dynamic and static load profiles of up to 1000 lb while monitoring and measuring the bearing torque. This capability allows for the characterization of several critical metrics such as bearing lift off speed characteristics, load capacity, and frequency dependent rotordynamic force coefficients. This paper discusses the functionality of the test facility and presents sample test measurements from several experiments. [DOI: 10.1115/1.2979004]

1 Introduction

The majority of present day turbomachinery uses rolling element bearings or oil lubricated journal bearings for rotor support. Although these conventional "oil-requiring" methods of supporting the rotor system have accomplished good reliability and performance, they prevent ultra high speed operation due to limitations inherent to physical stresses and power loss. Aircraft engines possess a DN speed limitation near 2.5 million due to the centrifugal loads on rolling elements, and machinery using oil-lubricated journal bearings is subject to significant power loss at high speeds. One possible solution to these speed limitations is the use of gas lubricated journal bearings, which can lift restrictions on speed and temperature, while possessing tolerable power loss at ultra high speeds. In addition to the high-speed capability of gas lubricated bearings, the removal of oil from aircraft engine turbomachinery results in significant benefits in specific fuel consumption, weight, reliability, and simplified sump designs [1]. However, realizing these benefits of gas bearings at the system level first requires component level evaluation based on bearing performance specifications that would provide safe, reliable, and efficient engine operation. The component level facility presented in this paper was designed to evaluate gas bearings based on lift off speed characteristics, load capacity, and rotordynamic force coefficients. The design of each test setup is discussed, and sample test measurements from each setup are shown.

2 Test Rig Design

The high-speed rotating component test facility is shown in Fig. 1. The table supporting the test hardware was designed with a double box support structure mounted on ball and socket rubber padded table feet. To allow for modular design and for the ease of constraining components, the table was fabricated with four *T*-slots, which are used for anchoring the bearing pedestal bases and ancillary test hardware. Two containment shields were designed based on the plastic strain energy theory [2], where the primary containment shield was mounted on four linear bearing guide blocks assembled on two rail guides on either side of the test table. The total weight of the test rig was calculated to be 5320 lb.

The side view of the main drive train and the cross section of the test rotor are shown in Fig. 2. The nonrotating components were fabricated using stainless steel 304 while the rotating assembly and bearing housing were machined from Inconel 718. A 40 kW CNC spindle with a maximum speed of 40 krpm, and a maximum torque of 10 N m drives the test rig. In line with the ac motor and the test section is a 1:2 gearbox that increases the maximum speed capability to 80 krpm. The floating end base support has the ability to slide in the *T*-slots allowing for different length test rotors and test articles. Mounted on the intermediate base supports are the bearing pedestals that support the high-speed test rotor. The main rotor-bearing system is straddle mounted with the test article located in the center of the rotor-bearing system. Although an overhung design was investigated, it posed limits on the speed and load capability compared with the straddle mounted design. Having the test article in a "suspended" configuration in the center of the test rotor allows static and dynamic forces to be applied through the housing rather than the excitation from the rotor imbalance. This feature allows the precise control and manipulation of the test article's static eccentricity, vibration amplitude, excitation frequency, and orbital shape. In addition, the straddle mounted design provides a more translational mode shape compared with the conical mode indicative of overhung systems. A disadvantage from a straddle mounted test setup is the increased time for the assembly and disassembly of the test article.

The rotor system has the provision to be hard mounted or soft mounted. The high-speed rotor-bearing system design, adopted from Refs. [3,4], uses oil–free wire mesh dampers in parallel with a double fold squirrel cage bearing support or blank inserts for the hard mounted configuration. Two sets of ABEC 7 25 mm back-to-back mounted ceramic angular contact ball bearings (the slave bearings) are used to support the test rotor, where one set is axially fixed and the other set is allowed to float axially with the use of a bearing carrier and preload spring. The floating end design is compliant to the axial growth of the rotor while the fixed end locates the rotor axially. A duplex mount configuration was chosen for an increased load capacity and longer bearing life. The back-to-back configuration also provides a larger baseline radial stiffness of the hard mounted rotor-bearing system, which contributes to the critical speed location residing above the test frequency range. In case of rotor seizure during testing a shear neck coupling

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received December 5, 2007; final manuscript June 22, 2008; published online December 19, 2008. Review conducted by Patrick S. Keogh. Paper presented at the 12th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery, Honolulu, HI, February 17–22, 2008.

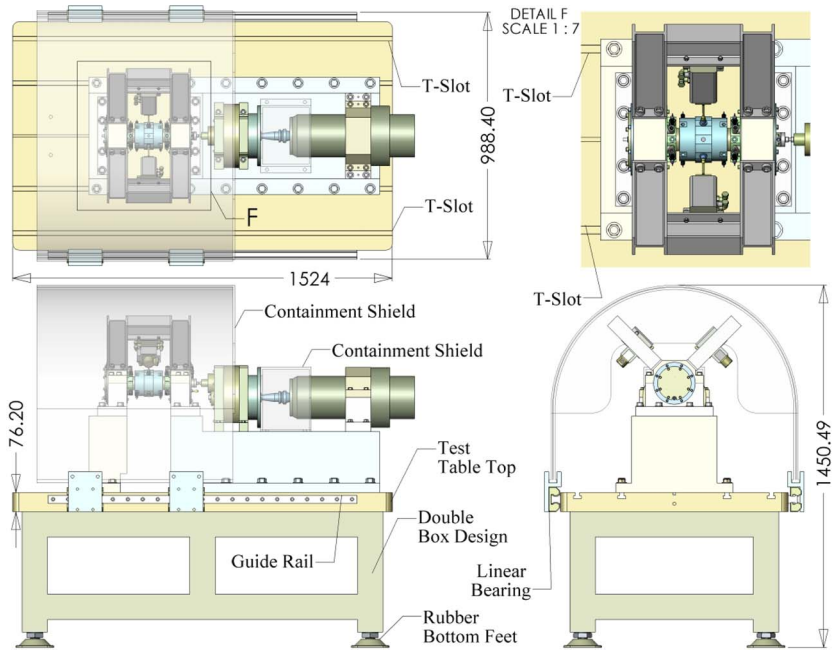


Fig. 1 Test facility layout (dimensions in millimeters)

spool piece was designed to fail at the gear tooth load rating. This would prevent damage to the electric motor drive and gearbox.

The cooling and mist lubrication system is detailed in Figs. 3 and 4. Air buffer seals were designed to provide cooling air, oil mist lubrication, and allowed vibration monitoring using two orthogonally mounted proximity probes. Pressurized air was supplied to the air-buffer seal plenum through three ports providing flow in two directions, inboard and outboard. The inboard air-buffer seal blade was designed with a smaller radial clearance to the rotor surface compared with the outboard air-buffer seal blade. This directs the majority of the cooling air through the bearing pedestal, which combines with the oil mist flow path, creating a high velocity oil mist cooling flow through the bearing cartridges while sealing the oil mist from entering the air-buffer seal plenum. The combination of oil mist and forced cooling air provides the best heat transfer characteristics when compared with oil bath or

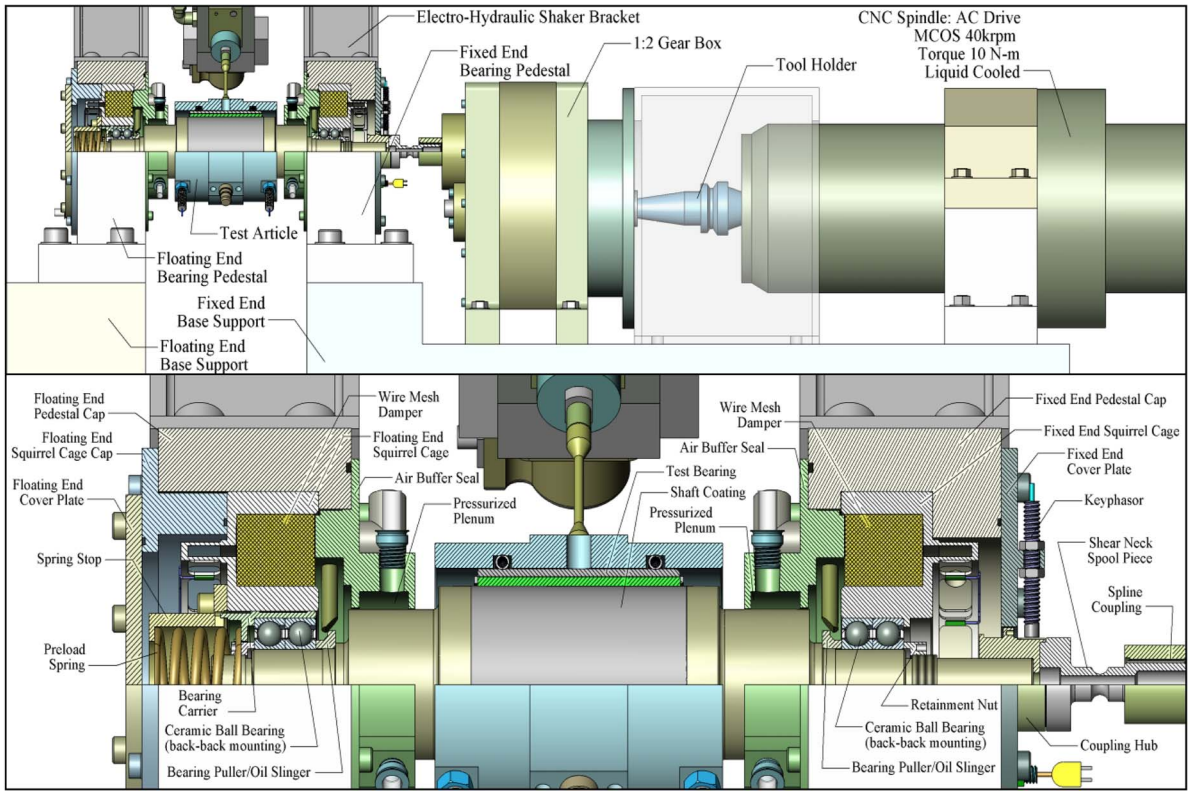


Fig. 2 Test rig cross section

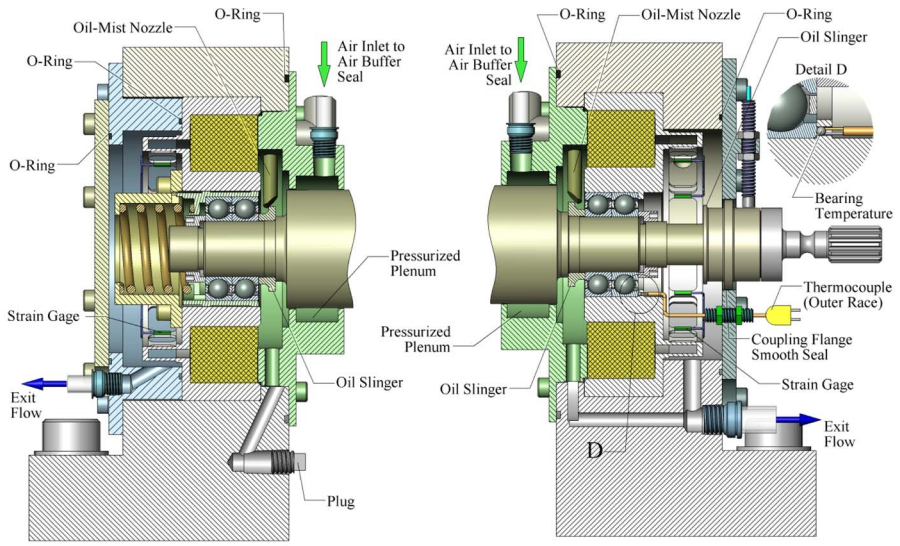


Fig. 3 Floating end and fixed end bearing pedestal cross sections

oil wicked lubrication systems, enabling higher speed and higher temperature test rig operation. A secondary function of the air-buffer seal was to provide axial compression to the wire mesh dampers. The stiffness and damping of the wire mesh dampers can be altered in situ by varying the axial compression through the air-buffer seal. To ensure sufficient heat transfer from the ball bearings a thermocouple was installed into the outer race of the fixed end bearing stack. Testing at 40,000 rpm and 1200°F test bearing temperatures materialized in the slave bearings at a maximum temperature of 140°F.

An important characteristic of the test rig operation is the baseline rotordynamics of the drive train. For longer test rotors that necessitated crossing a critical speed the soft mounted bearing configurations were used. Figure 5 shows the rotordynamic response of the fixed end bearing location for a 24-in. long rotor configuration. The waterfall plots show two critical speeds, where the first one is shown to be at 11,000 rpm and the second critical speed was observed to be at 21,000 rpm. At 40,000 rpm the steady state vibration response of the test rotor was limited to 0.45 mils (0.114 mm) p-p.

3 Bearing Lift Off Tests

One type of standard test administered on gas bearings is the lift off test. Lift off tests are performed for measuring the startup torque, lift off speed regime, running torque, and bearing touch down speeds. Figure 6 shows the test bearing housing installed in the rotating component test rig. The torque was measured using a torque arm connected to the bottom of the bearing housing, which then was assembled with a sliding fit into an aluminum block threaded onto a 0–44 N load cell. The load cell was then fastened to the head of a hex nut constrained to the support bracket cantilevered to the ground. The only constraint on the bearing housing was the torque arm, which located the housing assembly in the axial and lateral directions while permitting motion in the vertical direction of the varying dead weight bearing load.

The testing procedure involved accelerating the rotor (1000 rpm/s) from 0 rpm to a top test speed and then decelerating back to 0 rpm. An example lift off test of a hydrodynamic type gas bearing is shown in Fig. 7, which was acquired using a simultaneous dynamic data acquisition unit. The bottom graph in Fig. 7

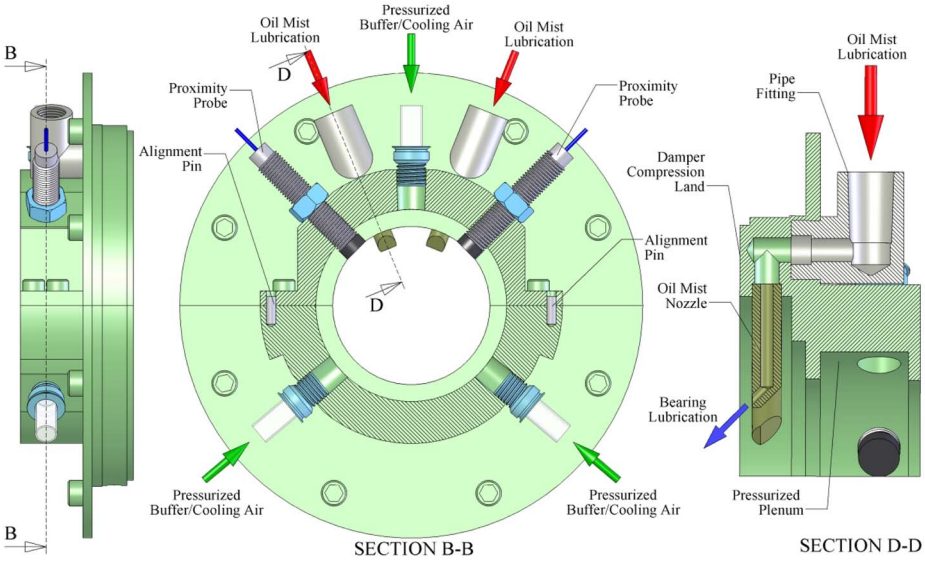


Fig. 4 Air buffer seal design and oil mist flow path

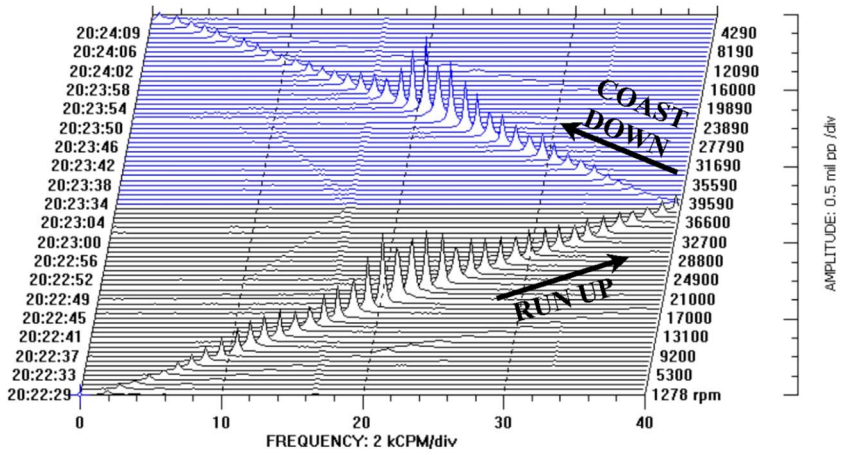


Fig. 5 Fixed-end rotor vibration: long rotor waterfall plot at 1200°F

represents the change in rotor speed as a function of time, and the top graph is the static torque measured through the load cell in Fig. 6. There are three basic regions in the torque graph for a hydrodynamic type gas bearing: (1) sliding friction, (2) a transition region, and (3) a fully developed hydrodynamic regime. The example torque signature in Fig. 7 shows sliding friction to approximately 1200 rpm and then the bearing operation enters a transition regime where the torque is dominated by boundary layer shear effects and intermittent rubbing. Above ~5000 rpm the bearing torque drops to nearly 0 in. lb, which is indicative of complete bearing lift off. In practice, it would be desired that the operating speed range and the eigenvalues of the rotor system reside above the transition region and within the fully hydrodynamic operating regime where the bearing torque is minimized and the air film is stable.

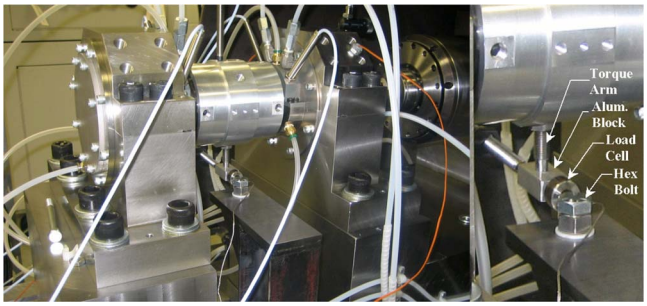


Fig. 6 Bearing housing installed in the test rig with the torque measuring device

4 Load Capacity Tests

One critical metric for evaluating gas bearings is the load capacity. The load capacity determines the maximum static loads that can safely be applied to the bearing for a given operating speed. Unlike the lift off speed tests that only require the rotor speed to traverse into the lift off speed regime, load capacity tests are conducted for higher speeds, which are often in the nominal operating speed region of the machine. In general for hydrodynamic bearings, load capacity increases with rotor speed and bearing diameter and is also dependent on the length of the bearing defining the projected area of the bearing. Past researchers have used the bearing temperature [5] or bearing torque to determine load capacity [6].

The experimental setup for the load capacity tests is shown in Fig. 8. The challenge with load capacity tests was how to remotely apply load at the test speed while monitoring an undistorted bearing torque signature. Directly applying a load to the bearing hous-

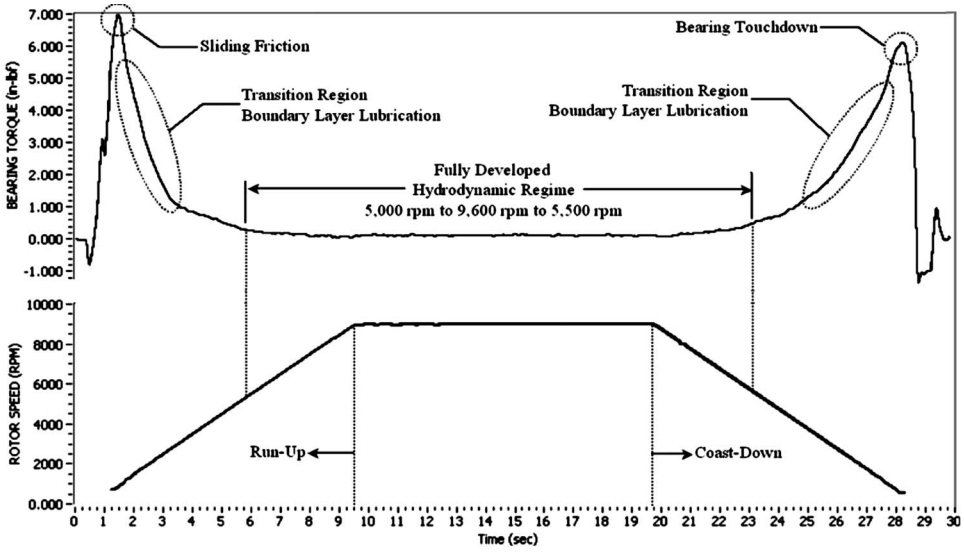


Fig. 7 Example start-up and shutdown bearing torque signature

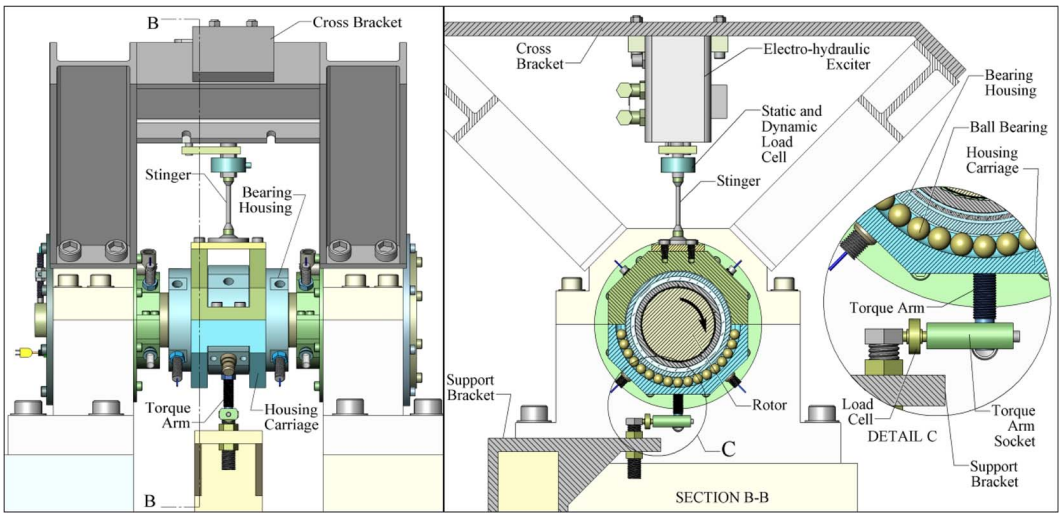


Fig. 8 Load capacity test setup

ing would have contributed to the distortion of the torque measurement, especially at higher loads. Therefore a bearing housing carriage was designed and fabricated to allow for a rotational degree of freedom to the bearing housing while applying a radial bearing load. This was achieved through two rows of ball bearings interfacing at the outer surface of the bearing housing. Figure 9 shows an example load capacity test. The testing procedure involved accelerating the rotor to the specified test speed starting with a lightly loaded bearing. After reaching the desired speed, the load was slowly increased, and the torque was monitored for any changes. A second challenge with the load capacity test was to determine when the load capacity was reached. A standard procedure for experimentally defining the load capacity for all types of gas lubricated bearings is difficult to define. However, load capacity tests using the torque measuring method in Fig. 8 clearly revealed a limiting load for stable bearing operation. As shown in Fig. 9, the torque signature possesses two different regions during the application of the bearing load. The first region begins after a bearing load of 30 lb, indicated by a steady increase in torque and a stable air film. The second region is clearly shown with an abrupt change in the trend of the torque signature. For this boundary layer lubricated region the air film is unstable and has a torque/force rate of change that is 8.5 times larger than Region 1. Using the distinction between these two regions, the bearing load capacity can be determined and for this case it was shown to be 90 lb. Repeated testing past the critical load of 90 lb, as shown in

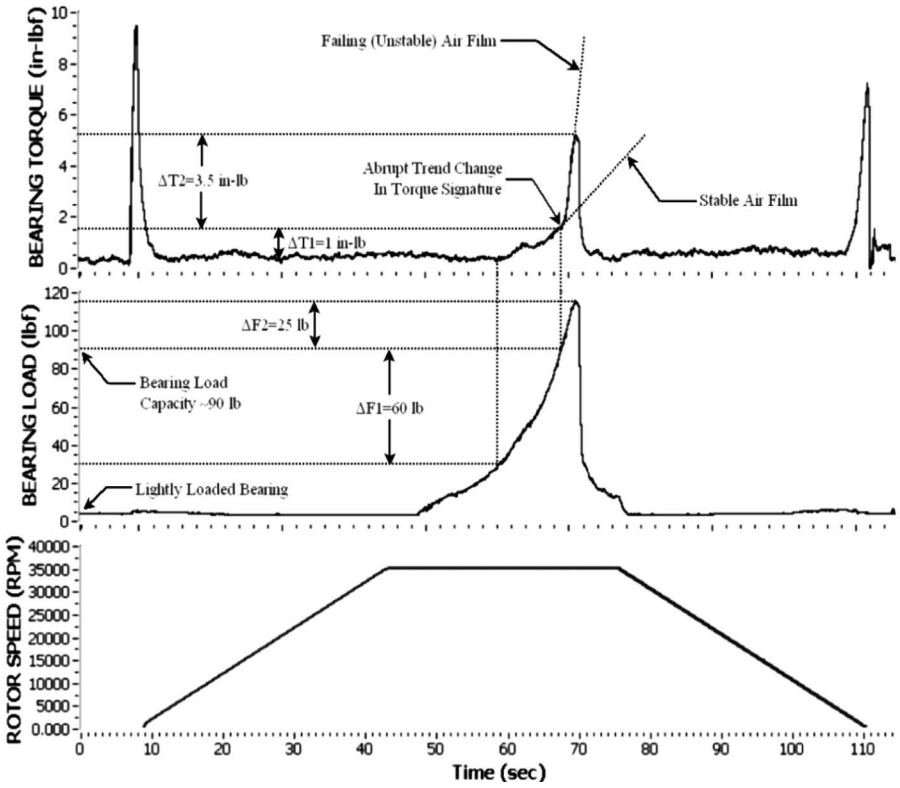


Fig. 9 Load capacity test: example output

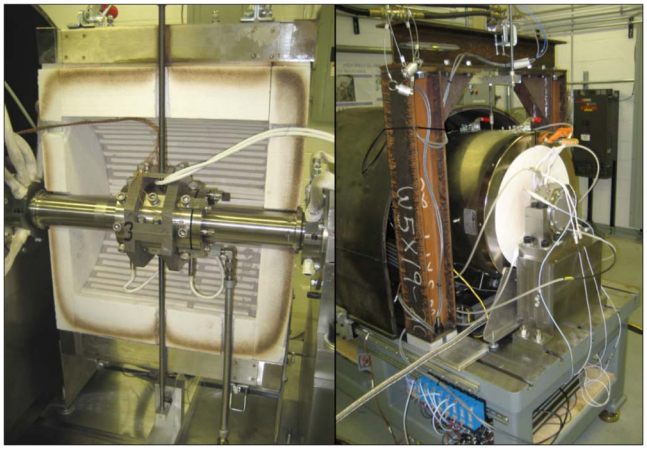


Fig. 10 High temperature test setup: long rotor with the carriage housing assembly

Fig. 9, can contribute to intermittent rubbing and can lead to bearing failure. Therefore a general rule of thumb was developed for this particular bearing design based on the critical torque value to start-up torque ratio. It was observed that the torque magnitude for the critical load capability was usually 10–15% that of the start-up torque for an unloaded bearing and was adopted as a criterion for determining the load capacity without having to traverse into the boundary flow region. After the load capacity was reached the bearing was unloaded followed by a deceleration of the test rig to 0 rpm.

Load capacity tests were ultimately conducted for varying environmental temperatures, which required using an oven. The high temperature setup is shown in Fig. 10, where the vertically split oven was installed between the floating end bearing pedestal and the fixed end bearing pedestal. The force transducers used for measuring the torque and bearing force had to be located outside the hot section. Therefore, two holes were cut on the top and bottom of the oven insulation to allow for the torque rod and the bearing load device. The carriage housing was used in this test for loading the bearing housing and was instrumented with two high temperature capacitance motion probes. The capacitance probes were used in combination with the bearing load to extract direct static stiffness values for different rotor speeds and temperatures. Since the air-buffer seals possess flow in both inboard and outboard directions, the flow of cooling air into the oven was a concern. To remedy this concern, fins were added to the long rotor design to act as deflectors and at the same time to act as cooling fins. Furthermore, the rotor was in constant rotation between temperature test points in order to prevent thermal bowing. It was observed that the rotor would bow and yield a high imbalance if the rotation was stopped between each temperature test point. Lastly, a cooling air manifold was mounted on one end of the test bearing housing to allow for the movement of air through the test bearing. This was implemented to simulate flow that would be experienced in the engine environment.

5 Rotordynamic Force Coefficients

A major area of concern toward implementing gas lubricated bearings into aircraft engines is the effect on turbomachinery rotordynamics. The majority of present day aircraft engine rotor-bearing designs incorporate squeeze film bearing support dampers in combination with squirrel cage centering springs for controlling displacements due to maneuver loads, synchronous response to imbalance while traversing through critical speeds, and nonsynchronous rotor excitation stemming from fluid forces through compressor and turbine stages. In addition, providing the necessary stiffness and damping parameters becomes imperative for peak thermodynamic efficiency, which is strongly influenced by static and dynamic clearance closures. However, with the removal of oil from the bearing system, the use of the squeeze film damper squirrel cage rotor support system is restricted. Hence, the rotordynamic characterization of the candidate gas bearing for turbine rotor support becomes paramount, especially before investing in system level engine tests.

Several past researchers have extracted force coefficients from gas lubricated journal bearings using a variety of methods [7–10]. The method discussed in this work is based on the method developed and implemented by Rouvas and Childs [11] and Childs and Hale [12]. Figure 11 shows the experimental setup for measuring asynchronous rotordynamic force coefficients. In this test configuration the electrohydraulic exciters were positioned at 45 deg locations using shaker brackets (brackets are shown in Figs. 1 and 2) that were constrained to the 45 deg bearing pedestal lands. The orthogonally mounted exciters where first used to establish the static equilibrium position of the bearing with respect to the rotor center. Next, dynamic excitations were preformed independently in the X and Y directions. Dynamic measurements for this method include relative housing motion in the X and Y directions, dynamic and static forces in the X and Y directions, and absolute housing accelerations in the X and Y directions. The excitation signal varied from single frequency excitation to pseudorandom excitation. In most cases a pseudorandom signal was engineered to produce a bearing response that contained an equivalent hous-

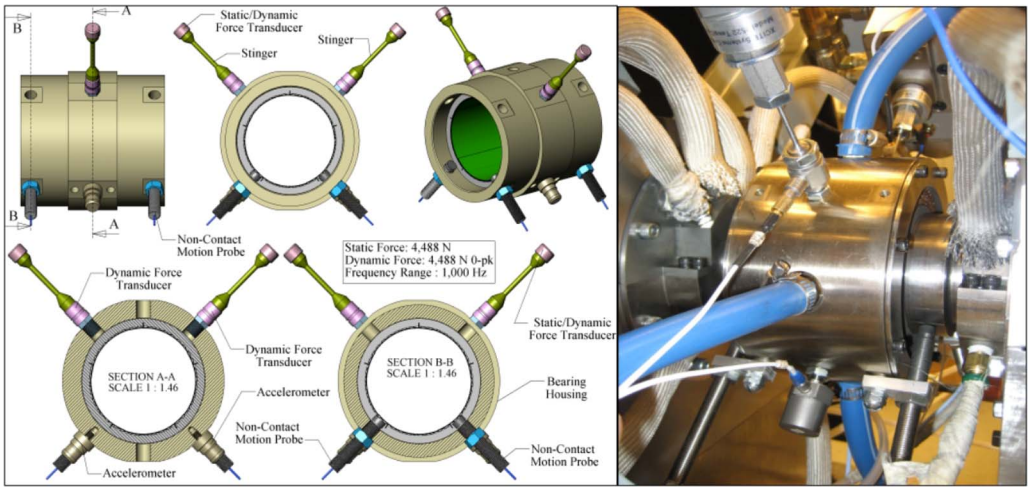


Fig. 11 Gas lubricated test bearing setup for rotordynamic characterization

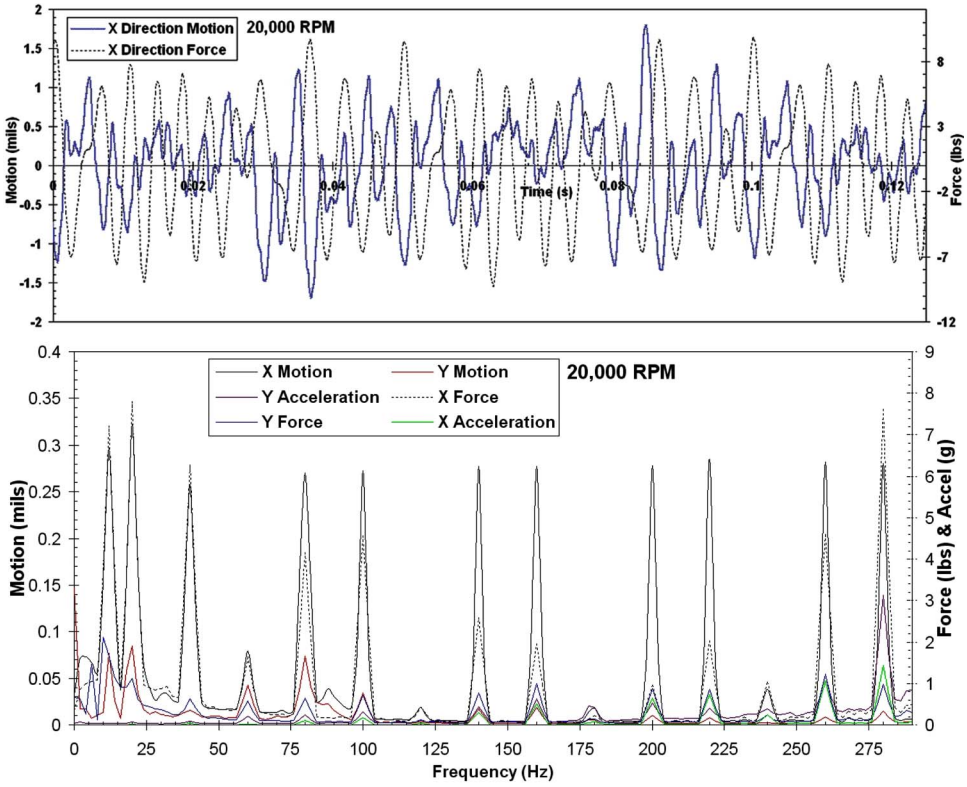


Fig. 12 Example *X* direction pseudorandom forced excitation measurements

ing motion at all frequencies. Figure 12 illustrates an example excitation in the *X* direction using a pseudorandom signal. Using these dynamic measurements the frequency dependent direct and cross-coupled stiffness and damping coefficients were calculated (Fig. 13).

6 Conclusion

The component level test facility was developed to provide an initial screening platform for candidate gas lubricated journal bearings. Selected test results presented in this paper were per-

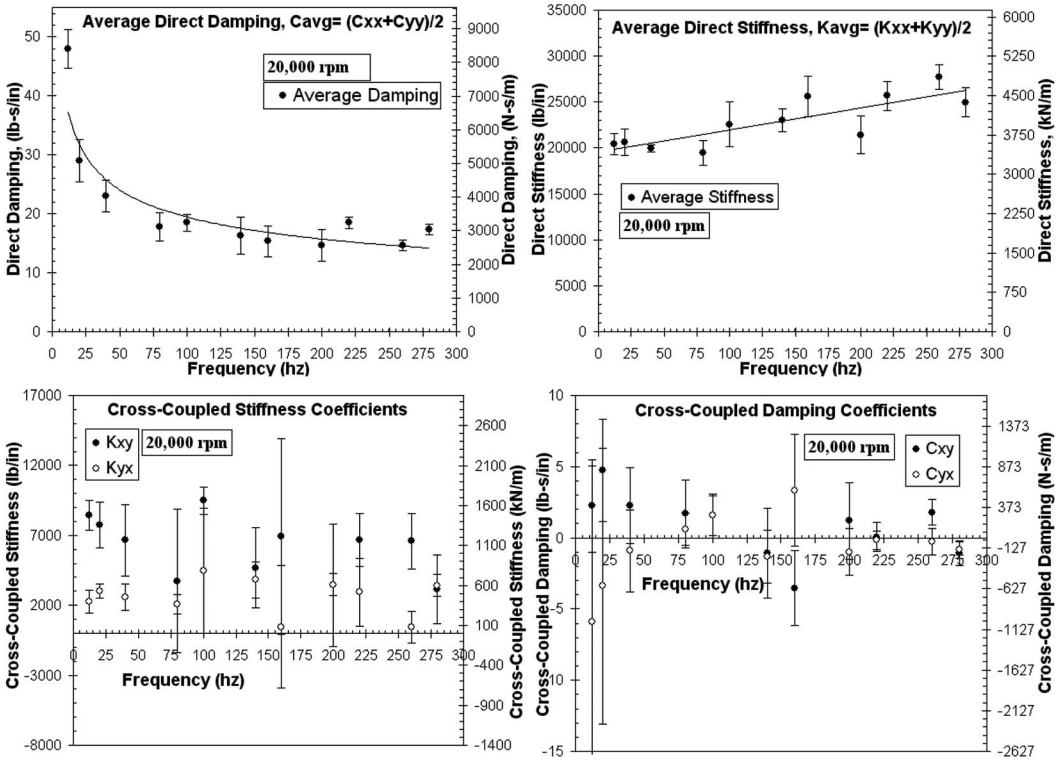


Fig. 13 Asynchronous rotordynamic force coefficients: example results

formed at 40,000 rpm and 1200°F for a variety of rotor sizes and bearing types. High fidelity torque measurements were used to identify inflection points from gas bearing torque signatures, which enabled the characterization of load capacity. The test setup also provided the ability to extract asynchronous direct and cross-coupled stiffness and damping coefficients for system level rotordynamic analysis.

References

[1] Spring, S. D., Kaminske, M., Leone, S., Drexel, M. V., Ertas, B. H., Ames, E. C., Agarwal, G., Burr, D., and Brophy, M., 2006, "Application of Compliant Foil Air Bearings for Oil Free Operation of Advanced Turboshaft Engines," *Proceedings of the American Helicopter Society 62nd Annual Forum*, Phoenix, AZ, May 8–11, Vol. III, pp. 2070–2075.

[2] Hagg, A. C., and Sankey, G. O., 1974, "The Containment of Disk Burst Fragments by Cylindrical Shells," ASME J. Eng. Power, **96**, pp. 114–123.

[3] Ertas, B. H., and Vance, J. M., 2004, "The Effect of Static and Dynamic Misalignment on Ball Bearing Radial Stiffness," J. Propul. Power, **20**(4), pp. 634–647.

[4] Ertas, B. H., Al-Khateeb, E. M., and Vance, J. M., 2002, "Rotordynamic Bearing Dampers for Cryogenic Rocket Engine Turbopumps," J. Propul. Power, **20**(4), pp. 674–682.

[5] Heshmat, H., 1994, "Advancements in the Performance of Aerodynamic Foil Journal Bearings; High Speed and Load Capability," ASME J. Tribol., **116**, pp. 287–295.

[6] Dellacorte, C., 1997, "A New Foil Air Bearing Test Rig for Use to 700°C and 70,000 rpm," NASA Technical Report No. TM-107405.

[7] Howard, S. A., Dellacorte, C., Valco, M. J., Prahl, J. M., and Hesmat, H., 2001, "Steady-State Stiffness of Foil Air Journal Bearings at Elevated Temperatures," Tribol. Trans., **44**(3), pp. 489–493.

[8] Howard, S. A., Dellacorte, C., Valco, M. J., Prahl, J. M., and Hesmat, H., 2001, "Dynamic Stiffness and Damping Characteristics of a High Temperature Air Foil Journal Bearing," Tribol. Trans., **44**(4), pp. 657–663.

[9] San Andres, L., 2006, "Hybrid Flexure Pivot-Tilting Pad Gas Bearings: Analysis and Experimental Validation," ASME J. Tribol., **128**, pp. 551–558.

[10] Lee, Y. B., Park, D. J., and Kim, H., 2006, "Numerical Analysis for Bump Foil Journal Bearing Considering Top Foil Effect and Experimental Investigation," *Proceedings of the IFToMM-Conference on Rotor Dynamics*, Vienna, Austria, Sept. 25–28, Paper No. 229.

[11] Rouvas, C., and Childs, D., 1993, "A Parameter Identification Methods for the Rotordynamic Coefficients of a High Reynolds Number Hydrostatic Bearing," ASME J. Vibr. Acoust., **115**, pp. 264–270.

[12] Childs, D. W., and Hale, K., 1994, "A Test Apparatus and Facility to Identify the Rotordynamic Coefficients of High Speed Hydrostatic Bearings," ASME J. Tribol., **116**, pp. 337–334.

Compliant Hybrid Journal Bearings Using Integral Wire Mesh Dampers

Bugra H. Ertas
Rotating Equipment Group,
Vibration and Dynamics Laboratory,
GE Global Research Center,
Niskayuna, NY 12309
e-mail: ertas@research.ge.com

The following work presents a new type of hybrid journal bearing developed for enabling oil-free operation of high performance turbomachinery. The new design integrates compliant hydrostatic-hydrodynamic partitioned bearing pads with two flexibly mounted integral wire mesh dampers. The primary aim of the new bearing configuration was to maximize the load-carrying capacity and effective damping levels while maintaining adequate compliance to misalignment and variations in rotor geometry. The concept of operation is discussed along with the description of the bearing design. Several experiments using room temperature air as the working fluid were performed that demonstrate proof of concept, which include lift-off tests, bearing load tests, and rotordynamic characterization tests. The experiments demonstrate stable operation to 40,000 rpm (2.8 $\times 10^6$ *DN) of a 2.750 in. (70 mm) diameter bearing. In addition to the experimental results, an analytical model is presented for the compliant bearing system. The aeroelastic theory couples the steady state numerical solution of the compressible Reynolds flow equation with a flexible structure possessing translational and rotational compliance. This was achieved by formulating a fluid-structure force balance for each partitioned bearing pad while maintaining a global mass flow balance through the hydrostatic restrictors and bearing lands. Example numerical results for pad pressure profile, film thickness, torque, and leakage are shown.* [DOI: 10.1115/1.2967476]

1 Introduction

Gas lubricated bearing systems offer several advantages over oil lubricated bearings such as lower power loss, less sealing requirements, extreme temperature capability, and resiliency to cavitation or ventilation [1]. Figure 1 illustrates several different gas lubricated bearing concepts. Initial gas bearing concepts [2] were designed with fixed geometry (Fig. 1(*a*)) and are categorized as self-acting or hydrodynamic bearings. Pressures generated through hydrodynamics are small when working with gas and lend to compromised load-carrying capability when compared with oil fed bearing systems. These types of bearings require tight tolerance designs and provide very little compliance with rotor growth or deformation. Increased load capacity in fixed geometry bearings can be achieved through external pressurization (Fig. 1(*b*)), but still face mechanical compliance issues and are susceptible to pneumatic hammer instabilities [3–7]. Nevertheless, bearing types as shown in Figs. 1(*a*) and 1(*b*) are used in several everyday applications such as textile spindles, hand tools, dentist drills, and gyroscopes [8]. However, when considering application into high performance turbomachinery, they have serious disadvantages such as low load capacity, alignment/misalignment challenges, and poor stability.

Improvements to fixed geometry gas bearing concepts are shown in Figs. 1(*c*) and 1(*d*). With the development of the foil bearing [9,10], high-speed and high temperature operation was enabled through the compliance of the bearing system. The bump foil and top foil provided compliance to changes in rotor geometry and misalignment making the application to ultrahigh-speed microturbomachinery possible. Additionally, advanced foil bearing designs [11,12] have reached significantly higher load capacities compared with fixed geometry bearings (Fig. 1(*a*)) and are able to sustain large excursions. However, load capability produced by the state of the art foil bearings requires very high operating speeds and foil bearings also possess transient rubbing during start-up. This fact puts limitations on these types of bearings when considering machines with large loads that operate at moderate to high speeds. Another improvement from the initial bearing concepts is the hybrid flexure pivot bearing [13] (Fig. 1(*d*)) due to the inherent high-speed stability and good load capacity. The drawback of the hybrid flexure pivot gas lubricated bearing is the very low damping. Researchers have also investigated flexibly mounted externally pressurized bearings [14–16] (Figs. 1(*e*) and 1(*f*)), in efforts to increase the stability threshold, but the designs considered still had limitations in temperature, speed, and the ability to withstand rotor deformation or growth.

This paper presents a new hybrid bearing concept developed to provide good load capacity for large operating speed margins while maintaining adequate levels of damping. The intent of the proposed bearing design has three main points: (1) compliance, (2) high load capacity at all rotor speeds, and (3) high damping.

2 Bearing Design

The bearing design is shown in Figs. 2 and 3 and details of the bearing-damper system tested in this paper are defined in Table 1 of the Appendix. There are three main components: (1) bearing/damper housing, (2) two integral wire mesh dampers (IWMDs), and (3) two end plates. The housing design in Fig. 3 is a single piece design with an outer rim and a partitioned inner rim joined together through the damper bridge located at the midspan of the bearing housing. The damper bridge uses two integral electric discharge machined (EDM) *S* springs [17] originally developed for use in combination with squeeze-film dampers as a parallel support centering spring. Most commonly used as a bearing support damper for flexure pivot bearings [18] (Fig. 12 (see the Appendix)), the integral centering springs allow radial compliance with the bearing system whereas the flexure pivot feature [18] allows rotational compliance of the bearing pads. For the bearing design in Figs. 2 and 3 the asymmetrically positioned integral

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 30, 2008; final manuscript received April 7, 2008; published online December 22, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

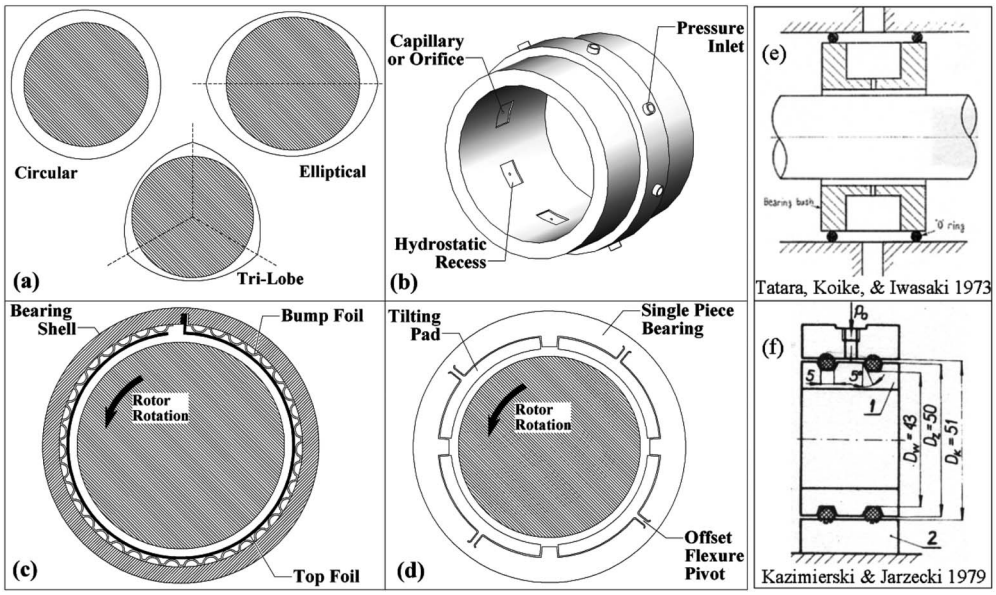


Fig. 1 Various types of gas lubricated journal bearings

springs support two functions by allowing radial motion and pad rotation. By combining the two functions of the conventional integral spring and flexure pivot design, the new configuration reduces radial space occupancy yielding a more compact design, which is advantageous when considering aerospace applications. Another important function of the integral springs is to ensure shaft-centering capability throughout the operating life cycle. Unlike the *o*-ring supported gas bearing concepts in Fig. 1, the integral centering springs provide a more reliable centering capability over time. Another unique feature of the hybrid bearing design is the asymmetric recess. The hydrostatic recess is positioned toward the leading edge of the bearing pad, which creates an off-centered pressure force. This feature helps facilitate pad rotation and strategic placement of the recess can strongly influence the pad pressure profiles at high speed. The compliant bearing system also has the ability to change clearance or set-bore diameter in response to geometry changes in the mechanical system made feasible by the partitioned inner rim. As shown in Fig. 3, the assembly between the bearing and journal is an interference fit where there exists a pad preload force on the journal surface. The pad preload force can be created by designing the set-bore clearance less than the rotor diameter or through axial compression on the dampers. Clearance is created by the pressurization of the individual pads, which can come from hydrodynamics or hydrostatics. One of the main hurdles facing gas lubricated bearing systems is the low valued damping that arises from the gas film. For recessed hybrid gas bearings the low damping stems from the low viscosity inherent to gases and the compressibility of the gas. To circumvent the low damping in the gas film, the new bearing design presented in this paper leverages existing soft-mounted bearing support concepts used in aircraft engine turbomachinery. Aircraft engine bearings are soft mounted using squeeze-film dampers [19–21] in combination with squirrel cages to control rotordynamics, even though damping in ball bearings is small. This concept works due to two main reasons: (1) The stiffness ratio between the ball bearing and squirrel cage bearing support is high and (2) the squirrel cage bearing support is mounted in parallel with a squeeze-film damper that possesses useful damping. The concept can be translated to the gas bearing system shown in Fig. 3. Figure 3 shows that as the stiffness ratio between the gas film and bearing support increases the equivalent damping of the bearing system approaches the bearing support damping. The stiffness of the gas film increases with rotor speed and in the case of hybrid bearings the stiffness of the gas film can be further enhanced through external pressurization, which maximizes the equivalent damping of

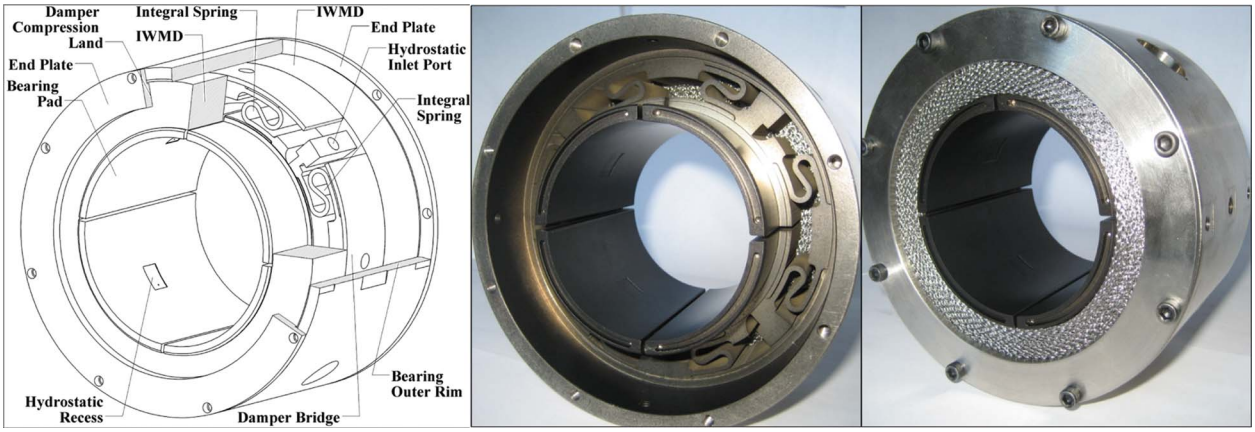


Fig. 2 Compliant hybrid journal bearing with IWMD

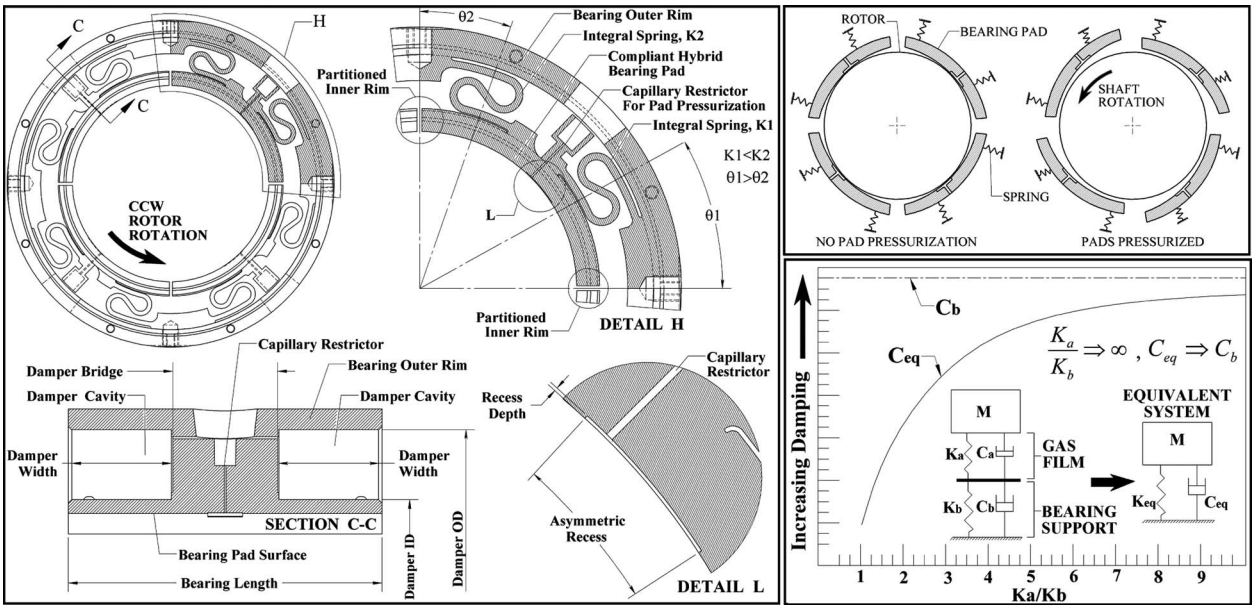


Fig. 3 Generation II style single piece bearing housing design and concept of operation

the bearing system. In addition to maximizing the equivalent damping, external pressurization enables higher load capacity.

The source of damping in the compliant hybrid bearing concept is harnessed from oil-free wire (metal) mesh dampers. Two wire mesh dampers are inserted into each side of the bearing housing in parallel with the integral housing springs and also have an interference fit between the inner diameter of the outer rim and the outer diameter surface of each partitioned bearing pad. The dampers are further constrained in the damper cavity using two end plates that can be used to vary the axial compression on the IWMD. Wire mesh bearing support dampers have mainly been restricted to laboratory level studies [22–26] with one type of known application [27]. Unlike *o*-rings, wire mesh dampers can withstand extreme temperatures and can operate in highly corrosive environments. The radial stiffness of the mesh dampers is significantly lower than the integral springs, and therefore the radial stiffness of the bearing system is primarily controlled by the integral springs. This fact allows for a linear stiffness behavior and good centering capability.

3 Bearing Stability and Modes of Operation

A major challenge with successful operation of a recessed hybrid bearing design operating with a compressible fluid was the ability to maintain stability at high speed. Past researchers have documented the usefulness of external damping in the mechanical system for suppressing pneumatic hammer [6,7]. To demonstrate the influence of damping in the bearing support, two tests were conducted with and without an IWMD. Tests were performed on a high-speed rotating test rig (Fig. 4) developed to evaluate gas bearings at the component level [28]. The test bearing was assembled at the center of the rotor where the ends of the rotor are supported by high-speed ceramic angular contact ball bearings. For this setup the hybrid bearing was constrained using a rigidly attached threaded rod through an electrohydraulic exciter providing a bearing load of 200 N. The vibration was monitored through two noncontact proximity probes measuring the relative vibration between the bearing housing and the rotor. The top waterfall plot in Fig. 4 represents the test with no bearing support dampers and shows a natural frequency at ~8 krpm. Once the rotor speed reaches ~21 krpm a subsynchronous vibration component develops at the natural frequency of the mechanical system, which shows to be significantly larger than the synchronous vibration response. The next test was performed with the IWMD, shown in the bottom waterfall plot, which demonstrates operation to five times the natural frequency to 40 krpm (2.8×10^6 DN) without subsynchronous excitation of the natural frequency.

The next set of experiments (Fig. 5) was performed to demonstrate the bearing operating modes. Important measurements in these tests were the bearing load and bearing torque. Bearing torque has been used in the past [28,29] to determine the load-carrying capacity of gas films and also has been used to determine when bearing lift-off has occurred. As torque values increase during testing, it is an indicator of higher viscous forces in the gas film generated through a decrease in film thickness. The bearing torque was directly measured from the bearing housing using a torque arm and an orthogonally mounted force transducer. Load was applied through a pulley system attached to an electrohydraulic exciter. The distortion of the torque measurement while applying the radial bearing load was minimized through the rotational degree of freedom in the pulley and spherical ball bearings.

In the simplest mode of operation, the bearing operates using hydrodynamics. This mode of operation would be applicable to situations where a surplus of pressurized gas is unavailable and the compliant bearing pads lift off the rotor surface at a corresponding rotor speed for a given load. There will exist a transient rub region before the hydrodynamics take effect and the bearing pads lift off. Therefore, this mode of operation would be most likely suited for machines with a small unit loading and high operating speeds, such as air-cycle machines and microturbine generators. Also, to maintain good bearing life, pad coatings need to be utilized and tailored to the specific application to ensure adequate number operating cycles. Test 1 demonstrating this operating mode is shown in Fig. 6 and reveals a bearing lift-off speed at 4.5 krpm. After reaching a speed of 12 krpm, a load of 100 lb (0.445 kN) was applied to the bearing without any significant changes in the torque measurement, indicating a stable air film. In practice, it would be desirable to locate rotor natural frequencies and extreme loading conditions above the lift-off speed. A variation of this mode of operation would consist of starting machine operation with hydrodynamics and then transitioning to hybrid operation. Hydrostatic pressurization of the bearing pads would begin at a certain speed and would most likely be powered through bleeding pressurized gas/liquid from the primary working fluid in turbomachinery stages. The advantage of this operating scheme compared with the previous one is the enhanced load capacity and bearing stability at higher rotor speeds. A prime ex-

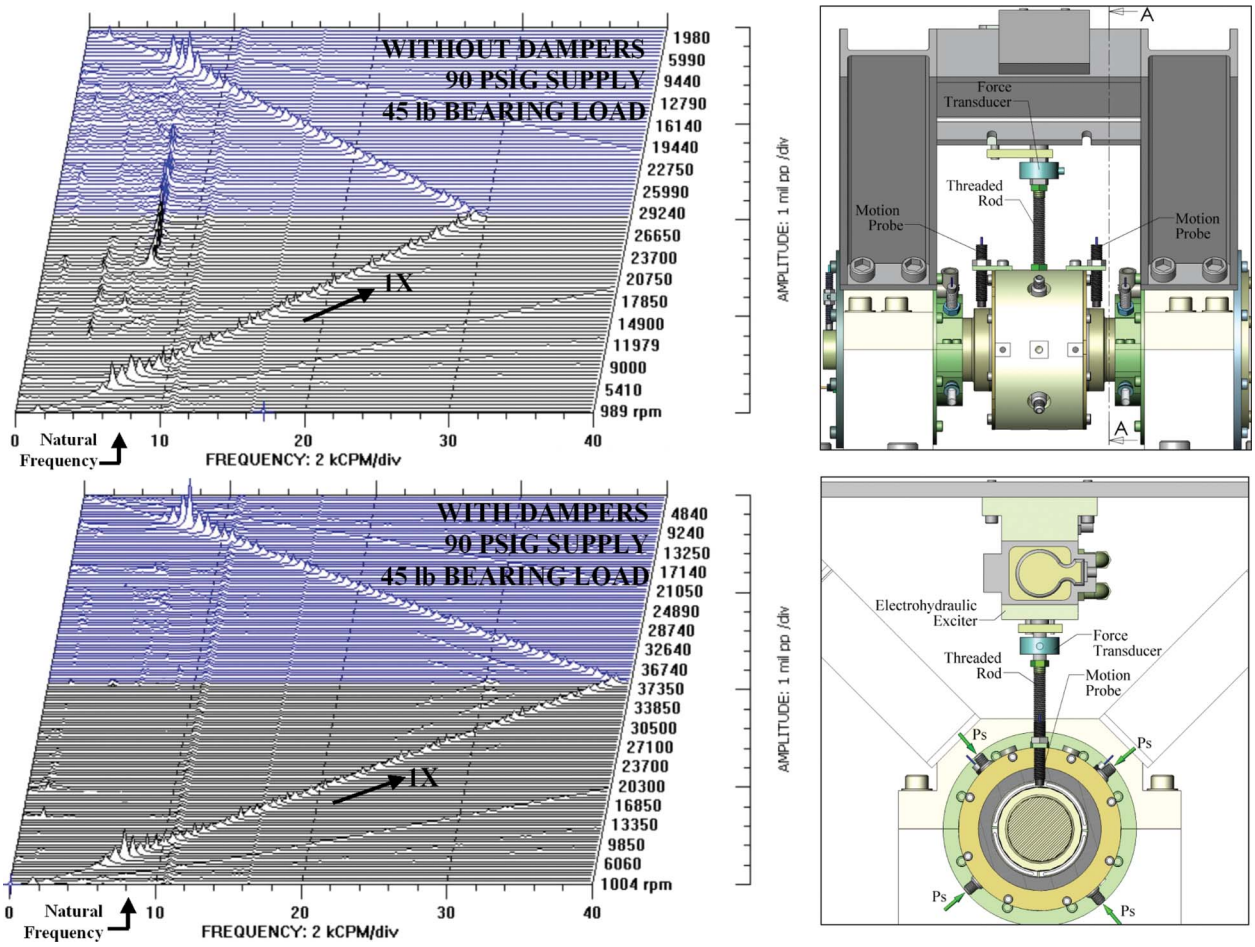


Fig. 4 Vibration response: ultrahigh-speed test to 40 krpm (6.2 bar supply pressure and 200 N bearing load)

ample for this operating mode would be helicopter gas turbine turboshaft engines or cryogenic rocket engine turbopumps due to the light unit loading during startup and the higher loading experienced at nominal operating speed ranges.

The last mode of operation is a transition from hydrostatic operation to hybrid operation. The use of this operating mode lends its advantages to machines with high unit loads where pressurized gas is available through the entire operating sequence of the machine. Having pressurized gas before rotor rotation prevents transient rubbing, reduces start-up torque, and can allow safe shutdown in emergency situations. Test 2 in Fig. 6 represents this operating mode and shows that during start-up the torque signal

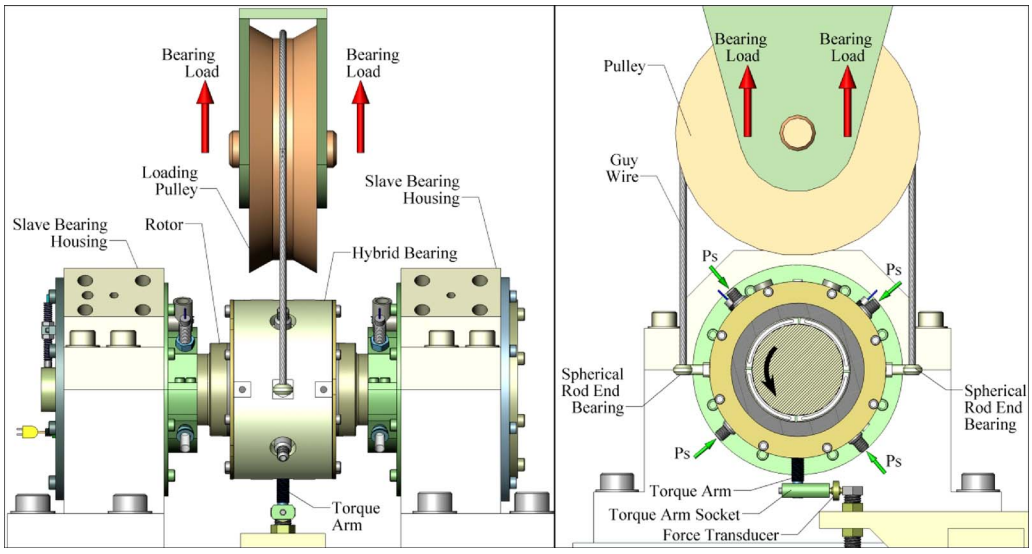


Fig. 5 Test setup for lift-off tests and load tests

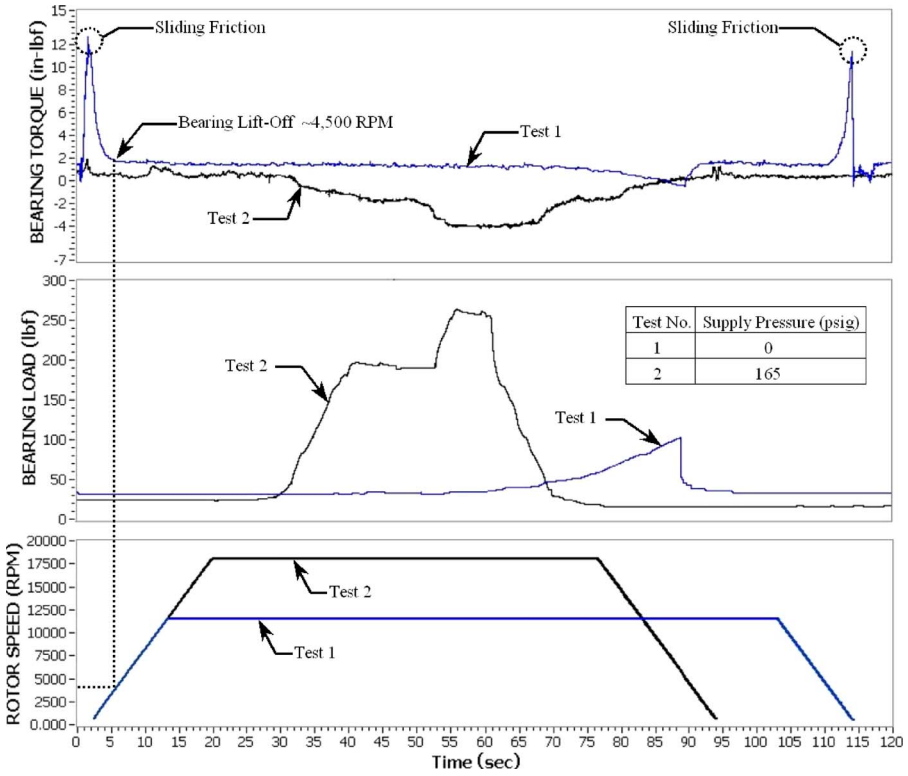


Fig. 6 Modes of operation

(without a start-up bearing torque spike) indicates a 0 rpm lift-off speed. Another advantage over the operating mode for Test 1 is the enhanced load capacity. After reaching the target test speed (18 krpm) a steady load up to 1.18 kN (265 lb) or unit load up to 230 kN/m^2 (33 psi) was applied while maintaining a stable torque signature. Note that as the load was applied for Test 2, the torque signature decreases in value. The reduction in the torque signature was most likely due to the cross-axis stiffness of the bearing, which produces an orthogonal force to the displacement. The orthogonal force components generated for the load regions of 0.89 kN (200 lb) and 1.18 kN (265 lb) were 0.16 lb (0.00071 kN) and 0.32 lb (0.00142 kN), respectively.

4 Rotordynamic Characterization

The tests presented in this section focus on the static and dynamic properties of the hybrid bearing system. Rotordynamics has historically been a concern when implementing gas bearings into rotor systems. Concerns stem from the significantly lower damping generated from a gas compared with oil. Compromised damping in bearing systems can lead to large rotor vibrations during critical speed transitions and subsynchronous excitation of rotor eigenvalues at high speeds. Therefore, a gas bearing system is sought that it is not only inherently stable but also provides useful damping to the rotating system.

The test setup is shown in Fig. 7, where the bearing housing was constrained at the center of the rotor-bearing system using two orthogonally mounted electrohydraulic exciters and stingers. The measurements included static forces, dynamic forces, static motion, dynamic motion, and acceleration both in the X and Y directions. The first test performed on the bearing was a static force-displacement test. Figure 8 represents a test conducted in the Y direction for varying supply pressures, P_S. The test involved

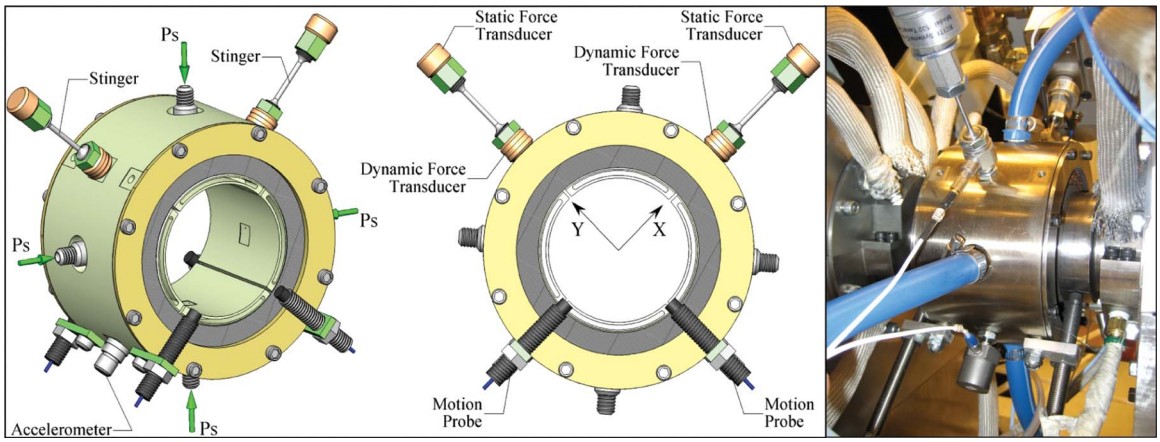


Fig. 7 Experimental setup for measuring rotordynamic force coefficients

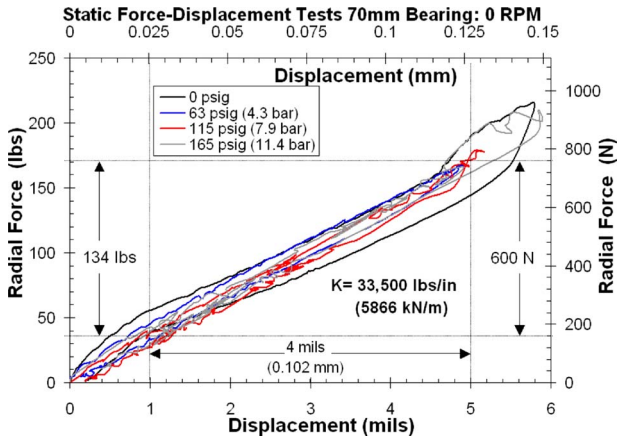


Fig. 8 Static displacement tests

loading then unloading the system in the Y direction and results clearly indicate a linear radial stiffness that has a weak dependence on the supply pressure. This result implies that the gas film is significantly stiffer than the support stiffness and therefore the stiffness of the bearing system is predominantly controlled by the integral springs.

The next set of experiments focused on determining the frequency dependent direct and cross-coupled rotordynamic force coefficients. The present test method [28] used for determining coefficients was derived from Ref. [30] and uses two independent pseudorandom forced excitations in the X and Y directions at a specific rotor speed. Equations (1)–(4) represent the direct and cross-coupled equations of motion in the complex frequency domain for the X and Y directions. The unknowns in these equations are the bearing complex impedance functions and are solved in Eqs. (5)–(8). From each impedance function, stiffness and damping coefficients can be extracted. A base line test was conducted to back out the contributions from the stingers and was ultimately subtracted from the system level force coefficients.

$$f_{xx} = (H_{xx}) \cdot D_{xx} + (H_{xy}) \cdot D_{xy} = F_{xx} - M_s A_{xx} \tag{1}$$

$$f_{yx} = (H_{xx}) \cdot D_{yx} + (H_{xy}) \cdot D_{yy} = F_{yx} - M_s A_{yx} \tag{2}$$

$$f_{yy} = (H_{yy}) \cdot D_{yy} + (H_{yx}) \cdot D_{yx} = F_{yy} - M_s A_{yy} \tag{3}$$

$$f_{xy} = (H_{yy}) \cdot D_{xy} + (H_{yx}) \cdot D_{xx} = F_{xy} - M_s A_{xy} \tag{4}$$

$$H_{xx} = \frac{f_{xx} \cdot D_{yy} - f_{yx} \cdot D_{xy}}{D_{xx} \cdot D_{yy} - D_{yx} \cdot D_{xy}} = K_{xx} + \hat{j}\Omega C_{xx} \tag{5}$$

$$H_{xy} = \frac{f_{xx} \cdot D_{yx} - f_{yx} \cdot D_{xx}}{D_{xy} \cdot D_{yx} - D_{yy} \cdot D_{xx}} = K_{xy} + \hat{j}\Omega C_{xy} \tag{6}$$

$$H_{yy} = \frac{f_{yy} \cdot D_{xx} - f_{xy} \cdot D_{yx}}{D_{yy} \cdot D_{xx} - D_{yx} \cdot D_{xy}} = K_{yy} + \hat{j}\Omega C_{yy} \tag{7}$$

$$H_{yx} = \frac{f_{yy} \cdot D_{xy} - f_{xy} \cdot D_{yy}}{D_{xy} \cdot D_{yx} - D_{yy} \cdot D_{xx}} = K_{yx} + \hat{j}\Omega C_{yx} \tag{8}$$

Two different test configurations were tested. The first setup was at 0 rpm and 0 bar supply pressure. The aim of this first test was to identify the stiffness and damping coefficients for the bearing support, namely, the integral springs and wire mesh dampers. The results for this test are shown in Fig. 9. There are four plots in Fig. 9: average direct damping, average direct stiffness, cross-coupled stiffness, and cross-coupled damping. The average direct damping shows a strong dependence on excitation frequency and agrees with past experience [26]. The direct stiffness in Fig. 9 has a weak dependence on frequency and agrees fairly well with the static stiffness tests (Fig. 8). It is important to note that both the damping and dynamic stiffness of wire mesh dampers are a strong function of vibration amplitude. The tests conducted in Fig. 9 used a pseudorandom excitation engineered to possess peak vibration amplitudes of 0.0127 mm for all frequency components. Figure 9 also indicates that the cross-coupled stiffness and damping of the bearing support are low, which is expected from a symmetric mechanical structure. The next test was conducted for 11.5 bars (165 psi (gauge)) at 18 krpm and represents the combined effect of the gas film in series with the bearing support. Results for the equivalent direct damping for rotating tests show a moderate drop in magnitude compared with the bearing support damping, but reveal a similar trend with frequency. Both tests yielded maximum damping at lower frequencies, which implies compatibility with soft-mounted rotor-bearing systems. The direct stiffness of the rotating tests was measured to be less than the direct stiffness of nonrotating tests at lower excitation frequencies and increased linearly with frequency. The cross-coupled coefficients for the rotating tests were measured to be small and at the same magnitudes of Test 1 coefficients, which is indicative of tilting pad bearing designs.

5 Theory of Steady State Operation

The following bearing analysis and model were developed to gain understanding on the fundamental operating characteristics of the compliant hybrid bearing system. It focuses on the solution of the steady state or zeroth-order governing equations for the fluid-structure model. The bearing model and code were helpful in understanding the relationship between gas film thickness values, pad pressure profiles, supply pressures, preload forces on the pad, torque, and leakage.

The general bearing model used in this analysis is shown in Fig. 10. The objective of the following analysis was to determine the equilibrium position of the rotor and bearing pads for a given bearing load and set of operating conditions. Finding the steady state equilibrium position of these components involved satisfying three main conditions: (1) mass flow balance between the hydrostatic flow and flow through the bearing lands, (2) the force balance on each partitioned bearing pad, and (3) a global force balance between the net pressure force from all pads and externally applied rotor force. Although the hybrid bearing analysis shown is unique due to the required coupling between the structure and fluid forces, it leverages existing bearing analysis methods and fundamentals [31–34]. The algorithm flowchart for the bearing analysis is shown in Fig. 11. Since the bearing is designed to have a set-bore diameter, D_P, less than the rotor diameter, D, the first step in the analysis was to guess at a set-bore diameter and recess pressure, P_r. This initializes the pad preload force, F_{PR}, which is a function of the damper and integral spring radial stiffness (Eq. (9)). The recess pressure, P_r, controls the mass flow through the hydrostatic restrictor defined in Eqs. (10) and (11) and acts as a boundary condition for the solution of the thin film flow region pressures. The rotor position relative to the housing is designated through the fixed inertial reference frame $[X, Y]$ using the rotor displacement coordinates e_x and e_Y. Next, the film thickness for each pad was calculated using Eq. (12). Assuming an isothermal ideal gas the governing compressible Reynolds flow equation for the thin film flow region can be written in the form as shown in Eq. (13) and can be converted into a nondimensional form using the relations in Eq. (14). The numerical solution of Eq. (13) was achieved using a finite difference approximation (Eqs. (15)–(18)) with the extrapolated Liebmann method [35], which is a Gauss–Seidel method applied to partial differential equations in combination with successive under-relaxation (Eq. (19)). Solution of the steady state pad pressure field leads to the calculation of pad pressure forces (Eq. (20)). After calculating the pressure forces a pad force balance equation for the X and Y directions (Eq. (22)) was generated. The free body diagram (FBD) for a pad in the bearing

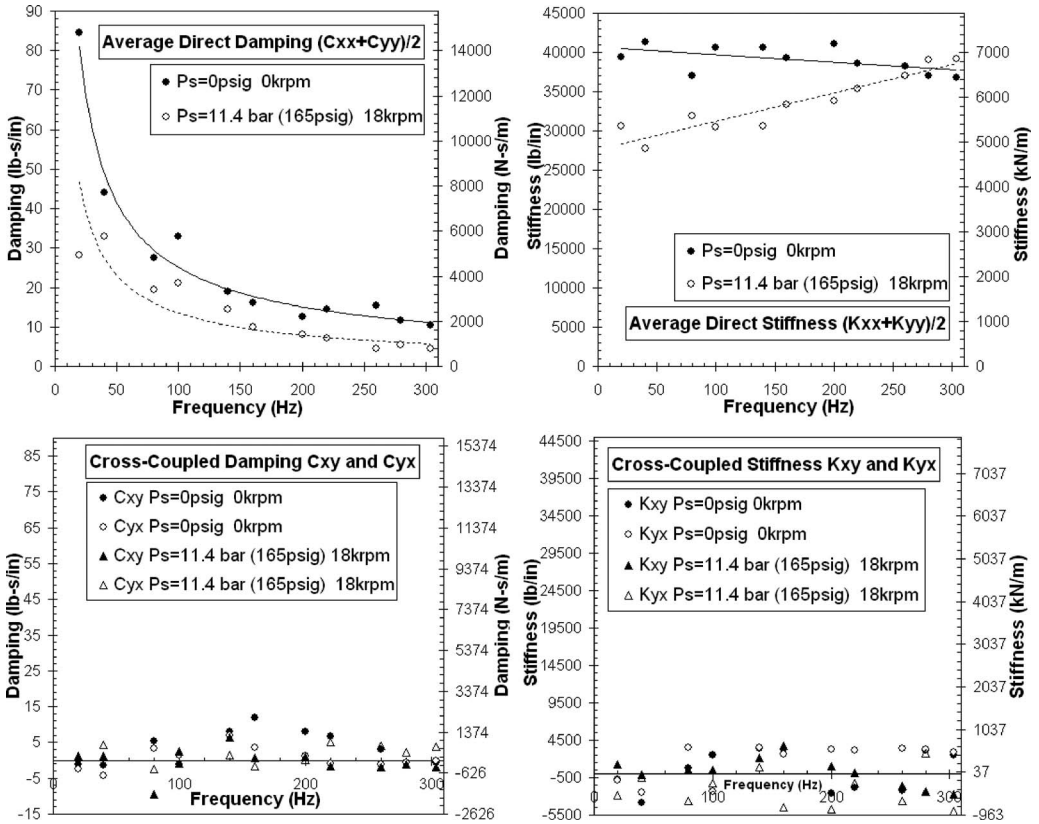


Fig. 9 Rotordynamic coefficients: 200 N *Y* direction static load and *X* and *Y* spectral vibration amplitudes: 0.0127 mm

system is shown in Fig. 10 and shows that the distributed damper force is lumped into two forces on each end of the bearing pad. All four bearing support force components shown in the FBD (Fig. 10) can be expressed using the deflection and corresponding radial stiffness (Eq. (23)). The objective of developing the force balance equations is to calculate the new position of the bearing pad. Recognizing that the pad is treated as a rigid body the deflections at points r_{S1} and r_{S2} can be written as functions of points r_0 and r_c. Ultimately the force balance equation was solved for the leading and trailing edge pad displacements yielding Δh^k and then was added to the existing film thickness $[h^k]_n$, resulting in the new film thickness $[h^k]_{n+1}$ (Eqs. (24)–(25)).

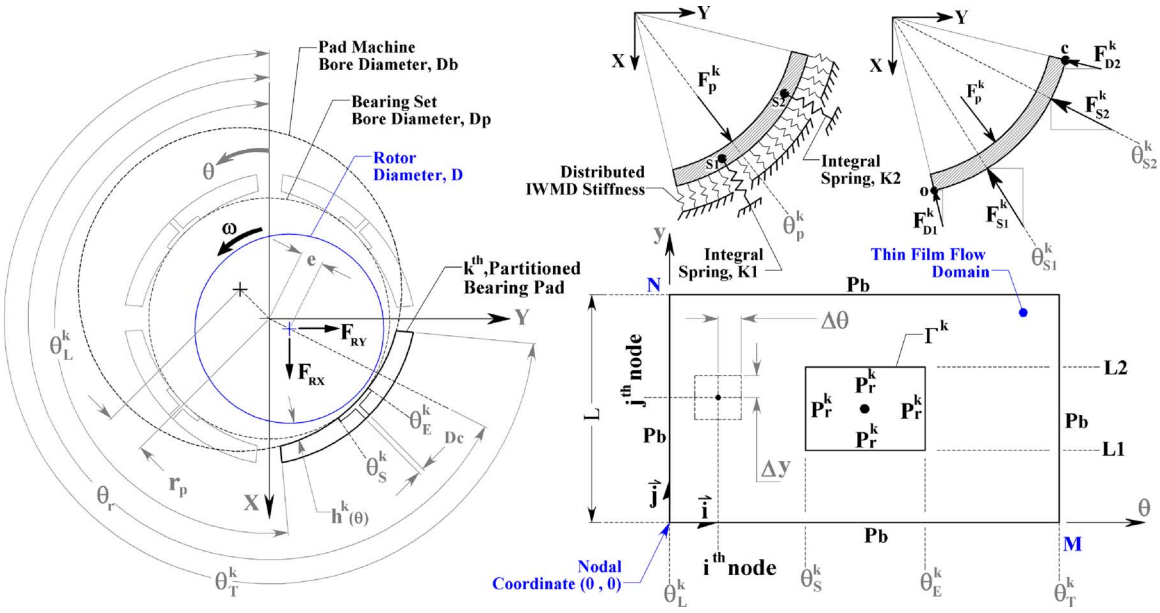


Fig. 10 Bearing model

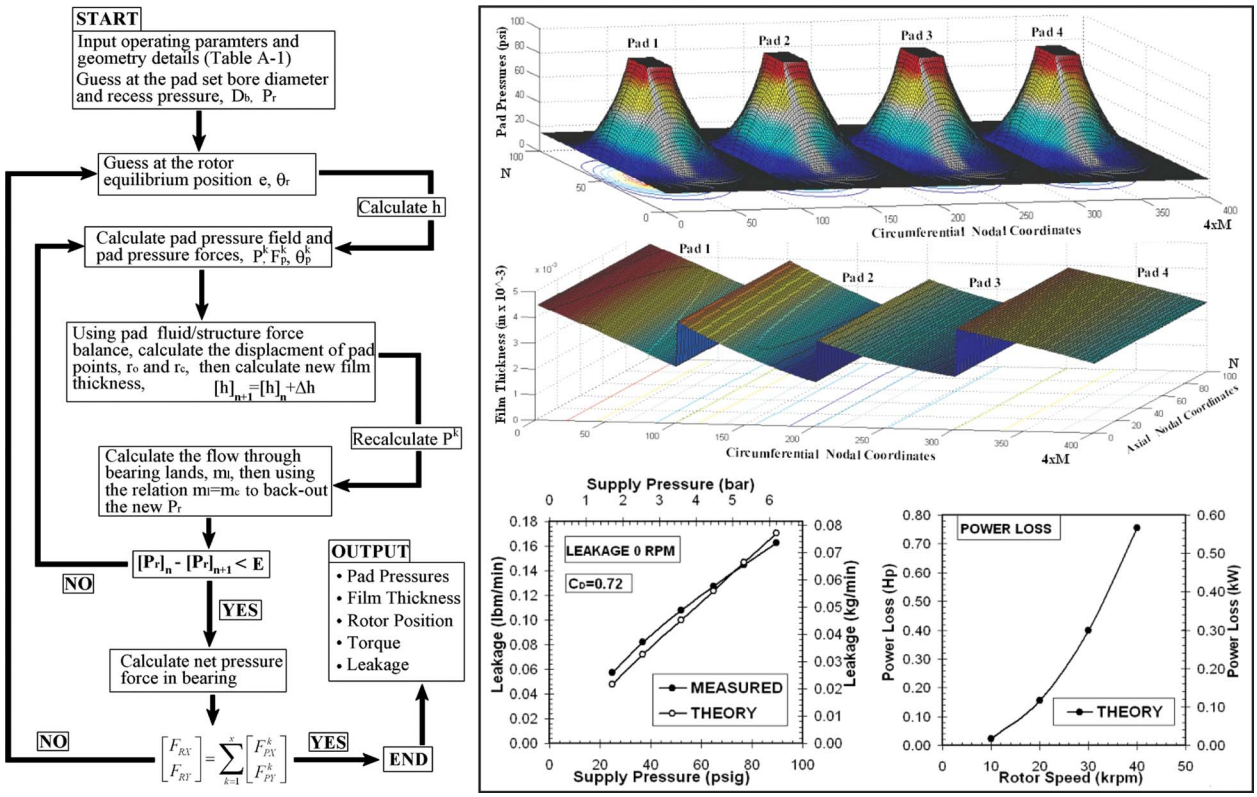


Fig. 11 Algorithm flowchart for bearing code and sample analysis results

$$F_{PR} = [F_{PR}]_D + [F_{PR}]_S = (K_D) \cdot \frac{(D - D_p)}{2} + (K_1 + K_2) \cdot \frac{(D - D_p)}{2} \tag{9}$$

$$\dot{m}_C^k = C_D \cdot \frac{\left(\frac{D_C}{2}\right)^2 \cdot \pi \cdot P_s}{\sqrt{\gamma \cdot \Re \cdot T}} \cdot \sqrt{\frac{2 \cdot \gamma^2}{\gamma - 1} \cdot \left[\left(\frac{P_r^k}{P_s}\right)^{2/\gamma} - \left(\frac{P_r^k}{P_s}\right)^{(\gamma+1)/\gamma} \right]} \tag{10}$$

$$\dot{m}_{\text{choked}} = \left(\frac{2}{\gamma + 1}\right)^{1/(\gamma-1)} \cdot \sqrt{\frac{\gamma - 1}{\gamma + 1}} \cdot \frac{\left(\frac{D_C}{2}\right)^2 \cdot \pi \cdot P_s}{\sqrt{\gamma \cdot \Re \cdot T}} \sqrt{\frac{2 \cdot \gamma^2}{\gamma - 1}} \tag{11}$$

$$h^k = C + e_X \cos(\theta) + e_Y \sin(\theta) - r_P \cos\left(\theta - \frac{(\theta_T^k - \theta_L^k)}{2}\right) \tag{12}$$

$$\frac{1}{R^2}\frac{\partial}{\partial \theta}\left((h^k)^3 P^k \frac{\partial P^k}{\partial \theta}\right) + \frac{\partial}{\partial y}\left((h^k)^3 P^k \frac{\partial P^k}{\partial y}\right) = 6\mu\omega \frac{\partial}{\partial \theta}(h^k P^k) \tag{13}$$

$$\bar{y} = \frac{y}{D}, \quad \bar{P}^k = \frac{P^k}{P_b}, \quad \bar{h}^k = \frac{h^k}{C}, \quad \Lambda = \frac{6\mu\omega}{P_b}\left(\frac{R}{C}\right)^2 \tag{14}$$

$$\left(\frac{\partial \bar{P}}{\partial \theta}\right)_{i,j} \approx \frac{\bar{P}_{i+1,j} - \bar{P}_{i-1,j}}{\theta_{i+1} - \theta_{i-1}} = \frac{\bar{P}_{i+1,j} - \bar{P}_{i-1,j}}{2\Delta\theta} \tag{15}$$

$$\left(\frac{\partial \bar{P}}{\partial \bar{y}}\right)_{i,j} \approx \frac{\bar{P}_{i,j+1} - \bar{P}_{i,j-1}}{\bar{y}_{j+1} - \bar{y}_{j-1}} = \frac{\bar{P}_{i,j+1} - \bar{P}_{i,j-1}}{2\Delta\bar{y}} \tag{16}$$

$$\left(\frac{\partial^2 \bar{P}}{\partial \theta^2}\right)_{i,j} \approx \frac{\left(\frac{\partial \bar{P}}{\partial \theta}\right)_{i+1,j} - \left(\frac{\partial \bar{P}}{\partial \theta}\right)_{i,j}}{\theta_{i+1} - \theta_i} = \frac{\bar{P}_{i+1,j} - 2\bar{P}_{i,j} + \bar{P}_{i-1,j}}{(\Delta\theta)^2} \tag{17}$$

$$\left(\frac{\partial^2 \bar{P}}{\partial \bar{y}^2}\right)_{i,j} \approx \frac{\left(\frac{\partial \bar{P}}{\partial \bar{y}}\right)_{i,j+1} - \left(\frac{\partial \bar{P}}{\partial \bar{y}}\right)_{i,j}}{\bar{y}_{j+1} - \bar{y}_j} = \frac{P_{i,j+1} - 2\bar{P}_{i,j} + \bar{P}_{i,j-1}}{(\Delta\bar{y})^2} \tag{18}$$

$$[\bar{P}_{(i,j)}]_{(n+1)} = r_f \cdot [\bar{P}_{(i,j)}]_{(n+1)} + (1 - r_f) \cdot [\bar{P}_{(i,j)}]_{(n)} \tag{19}$$

$$\begin{bmatrix} F_{PX}^k \\ F_{PY}^k \end{bmatrix} = \int_0^L \int_{\theta_L^k}^{\theta_T^k} (P^k - P_b) \begin{bmatrix} \cos\theta \\ \sin\theta \end{bmatrix} \cdot R d\theta \cdot dy \tag{20}$$

$$F_p^k = \sqrt{(F_{PX}^k)^2 + (F_{PY}^k)^2} \tag{21}$$

$$F_p^k \cdot \begin{bmatrix} \sin\theta_p^k \\ \cos\theta_p^k \end{bmatrix} = F_{D1}^k \cdot \begin{bmatrix} \sin\theta_L^k \\ \cos\theta_L^k \end{bmatrix} + F_{D2}^k \cdot \begin{bmatrix} \sin\theta_T^k \\ \cos\theta_T^k \end{bmatrix} + F_{S1}^k \cdot \begin{bmatrix} \sin\theta_{S1}^k \\ \cos\theta_{S1}^k \end{bmatrix} + F_{S2}^k \cdot \begin{bmatrix} \sin\theta_{S2}^k \\ \cos\theta_{S2}^k \end{bmatrix} \tag{22}$$

$$\begin{bmatrix} F_{D1}^k \\ F_{D2}^k \end{bmatrix} = \frac{[F_{PR}]_D}{2} + \frac{K_D}{2} \cdot \begin{bmatrix} r_o^k \\ r_c^k \end{bmatrix}, \quad \begin{bmatrix} F_{S1}^k \\ F_{S2}^k \end{bmatrix} = [F_{PR}]_S + \begin{bmatrix} K_1 \\ K_2 \end{bmatrix} \cdot \begin{bmatrix} r_{S1}^k \\ r_{S2}^k \end{bmatrix} \tag{23}$$

$$r_{S1}^k = f(r_o^k, r_c^k), \quad r_{S2}^k = f(r_o^k, r_c^k)$$

$$[h^k]_{n+1} = [h^k]_n + \Delta h^k \tag{24}$$

where

$$\Delta h^k = f(r_o^k, r_c^k). \tag{25}$$

The next step was to use the new film thickness for each bearing pad to recalculate the pad pressures. Using the new pressure profile, the mass flow from the recess into the bearing thin film flow region was calculated (Eq. (26)). Using the mass flow balance relation in Eq. (27), a new recess pressure was solved for. The iterative loop for the individual pad force balance (Eq. (22)) and mass flow balance (Eq. (27)) is executed until the recess pressures from consecutive iterations fall below a specified error value. Once the recess pressure error criterion is satisfied the global force balance between the new gas force and external rotor forces is checked (Eq. (28)). If the difference between these two force parameters is larger than the error criterion for the global force balance loop rotor position is changed and the analysis starts over. Example results for the bearing parameters in Table 1 are shown in the right side of Fig. 11. The top plot represents the pad pressure profiles and the pad film thickness values, whereas the two bottom plots show the leakage and power loss.

$$\dot{m}_l^k = \oint_{\Gamma^k} [\rho h \vec{U} \cdot \vec{n}]^k d\Gamma^k \tag{26}$$

$$\dot{m}_C^k = \dot{m}_l^k \quad \text{for } k = 1, \ldots, x \tag{27}$$

$$\begin{bmatrix} F_{RX} \\ F_{RY} \end{bmatrix} = \sum_{k=1}^{x} \begin{bmatrix} F_{PX}^k \\ F_{PY}^k \end{bmatrix} \tag{28}$$

6 Conclusions

The present work focused on a compliant hybrid journal bearing concept using oil-free wire mesh bearing support dampers. The intent of the bearing technology was to enable oil-free operation of medium to large scale rotating equipment for a wide range of operating speeds. Evaluating the bearing technology was accomplished from several viewpoints including high-speed stability, load-carrying capability, operational versatility, and rotordynamic performance. It was shown through experimental investigation that the new bearing technology was able to successfully operate in two distinct operating modes and also demonstrated stable operation at 40 krpm. The ability to externally pressurize the bearing circumvents transient rubbing issues at low rotor speeds, which enables application turbomachinery with large operating loads. The experimental results also indicated that the bearing system has significant positive damping maximized at lower frequencies and small stiffness cross-coupling. Section 5 of the paper described the fundamental bearing model based on the solution of the compressible Reynolds flow equation in combination with the local force balance equations of bearing pads and the global force balance between external rotor forces and gas forces.

Appendix

Table 1 Bearing analysis details

General bearing and gas data		Pad 1, $k=1$	
D, rotor diameter	69.85 mm	θ_E	0.977
Φ, rotor speed	3140 rad/s	θ_L	0.052
L, bearing length	74.53 mm	θ_S	0.733
F_{RX}, bearing load in X	200 N	θ_{S1}	0.419
F_{RY}, bearing load in Y	0 N	θ_{S2}	1.151
D_b, machine pad bore diameter	69.95 mm	θ_T	1.518
P_s, supply pressure	6.9 bars	Pad 2, $k=2$	
P_b, pad boundary pressure	1 bar	θ_E	2.547
γ, gas specific heat ratio	1.4	θ_L	1.622
T, gas temperature	298 K	θ_S	2.303
μ, gas viscosity	298 K	θ_{S1}	18.3×10^{-6} (N s/m^2)
P_{PR}, pad preload force	111 N	θ_{S2}	2.721
L_1, recess start coordinate	31.75 mm	θ_T	3.088
L_2, recess end coordinate	41.91 mm	Pad 3, $k=3$	
Recess depth	0.229 mm	θ_E	4.117
Capillary diameter	0.584 mm	θ_L	3.192
Capillary length	14.5 mm	θ_S	3.873
Bearing OD	136.6 mm	θ_{S1}	3.559
Damper OD	121 mm	θ_{S2}	4.291
Damper ID	80 mm	θ_T	4.658
Damper width	23 mm	Pad 4, $k=4$	
Damper mesh density	30%	θ_E	5.687
Damper material	Cu	θ_L	4.762
Wire gauge	0.254 mm	θ_S	5.443
K_D, damper stiffness per pad	700 kN/m	θ_{S1}	5.129
Integral spring stiffness $K1$	1400 kN/m	θ_{S2}	5.861
Integral spring stiffness $K2$	1400 kN/m	θ_T	6.228
Pad coating to 450 F	Teflon		
		N	100
		M	100
		rf	0.4

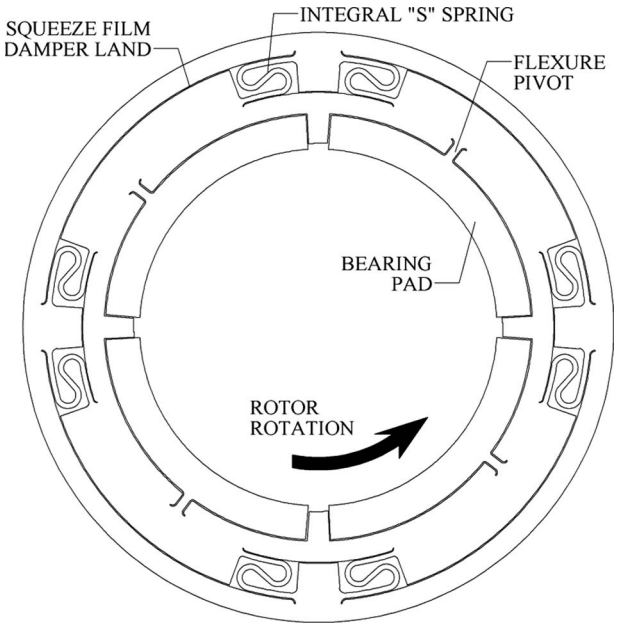


Fig. 12 Integral squeeze-film damper with flexure pivot bearing

Acknowledgment

The author would like to thank the General Electric Co. for sponsoring the present work and KMC Bearings Inc. for supplying the test bearing.

Nomenclature

A_{XX} = acceleration in X for X direction excitation, m/s^2
A_{YY} = acceleration in Y for Y direction excitation, m/s^2
A_{XY} = acceleration in Y for X direction excitation, m/s^2
A_{YX} = acceleration in X for Y direction excitation, m/s^2
C = radial clearance, m
c = point on the pad trailing edge
C_D = discharge coefficient
C_{XX} = direct damping in X direction, N s/m
C_{YY} = direct damping in Y direction, N s/m
C_{XY} = cross-coupled damping in X direction, N s/m
C_{YX} = cross-coupled damping in Y direction, N s/m
D = rotor diameter, m
D_b = pad machine bore diameter, m
D_c = capillary restrictor diameter, m
D_p = pad set-bore diameter, m
D_{XX} = motion in X direction for X direction excitation, m
D_{YY} = motion in Y direction for Y direction excitation, m
D_{XY} = motion in Y direction for X direction excitation, m
D_{YX} = motion in X direction for Y direction excitation, m
E = error between successive recess pressure iterations
e = rotor eccentricity, m
e_X = rotor eccentricity in X direction, m
e_Y = rotor eccentricity in Y direction, m
F_{D1} = damper force at the leading edge of the pad, kg m/s^2
F_{D2} = damper force at the trailing edge of the pad, kg m/s^2
F_p = center of pressure force of the pad, kg m/s^2
F_{PR} = total preload force on the pad, kg m/s^2
$[F_{PR}]_D$ = damper preload force on the pad, kg m/s^2
$[F_{PR}]_S$ = integral spring preload force on the pad, kg m/s^2
F_{S1} = integral spring force at the leading edge of the pad, kg m/s^2
F_{S2} = integral spring force at the trailing edge of the pad, kg m/s^2
F_{RX} = static rotor force in X direction, kg m/s^2
F_{RY} = static rotor force in Y direction, kg m/s^2
F_{PX} = pad pressure force in X direction, kg m/s^2
F_{PY} = pad pressure force in Y direction, kg m/s^2
F_{XX} = force in X direction for X direction excitation, N
F_{YY} = force in Y direction for Y direction excitation, N
F_{XY} = force in Y direction for X direction excitation, N
F_{YX} = force in X direction for Y direction excitation, N
H_{XX} = direct bearing impedance in X direction, N/m
H_{YY} = direct bearing impedance in Y direction, N/m
H_{XY} = cross-coupled bearing impedance in X direction, N/m
H_{YX} = cross-coupled bearing impedance in Y direction, N/m
h = film thickness, m
i = subscript, ith finite difference node in θ direction
j = subscript, jth finite difference node in y direction
$\hat{j}$ = imaginary unit, $\sqrt{-1}$
K_1 = radial stiffness of the leading edge integral spring, N/m
K_2 = radial stiffness of the trailing edge integral spring, N/m
K_D = radial stiffness of the wire mesh for a single pad, N/m
K_{XX} = direct stiffness in X direction, N/m
K_{YY} = direct stiffness in Y direction, N/m
K_{XY} = cross-coupled stiffness in X direction, N/m
K_{YX} = cross-coupled stiffness in Y direction, N/m
k = superscript, kth bearing pad
L = axial length of bearing, m
L_1 = recess axial start coordinate, m
L_2 = recess axial end coordinate, m
M = number of finite difference nodes in θ direction
M_S = system mass, kg
$\dot{m}_C$ = mass flow through hydrostatic restrictor, kg/s
$\dot{m}_l$ = mass flow through bearing lands, kg/s
N = number of finite difference nodes in y direction
n = subscript, iteration number
$\vec{n}$ = normal unit vector to recess boundary Γ
o = point on the pad leading edge
P = thin film flow pad pressure, N/m^2
P_b = pad boundary pressure, N/m^2
P_s = supply pressure, N/m^2
P_r = recess pressure, N/m^2
r_f = coefficient for successive under-relaxation
r_c = displacement at the pad trailing edge, m
r_o = displacement at the pad leading edge, m
r_{S1} = displacement at the leading edge integral spring, m
r_{S2} = displacement at the trailing edge integral spring, m
R = bearing radius, m
S_1 = point of leading edge spring location
S_2 = point of trailing edge spring location
T = gas temperature, K
$\vec{U}$ = flow velocity, m/s
X = global coordinate for fixed inertial reference frame
y = axial coordinate for pad flow field
Y = global coordinate for fixed inertial reference frame
x = number of partitioned pads
γ = ratio of specific heats of working gas
Γ = pad recess boundary
Δy = finite difference element size in axial direction, m
$\Delta\theta$ = finite difference element size in θ direction, rad
μ = gas viscosity, N s/m^2
θ = angular coordinate for pad flow field, rad
θ_E = angular coordinate for recess trailing edge, rad
θ_L = angular coordinate for pad leading edge, rad
θ_p = angular coordinate for pad center of pressure, rad

θ_r = angular coordinate of rotor displacement, rad
θ_S = angular coordinate for recess leading edge, rad
θ_{S1} = angular coordinate leading edge integral spring, rad
θ_{S2} = angular coordinate trailing edge integral spring, rad
θ_T = angular coordinate for pad trailing edge, rad
ρ = gas density, kg/m^3
Ω = bearing housing excitation frequency, rad/s
ω = angular velocity of rotor, rad/s
$\mathfrak{R}$ = specific gas constant for air, 287.05 J/(kg K)

References

[1] Hamrock, B. J., 1994, *Fundamentals of Fluid Film Lubrication*, McGraw-Hill, New York, pp. 329–330.
[2] Kingsbury, A., 1897, "Experiments With an Air-Lubricated Journal," J. Am. Soc. Nav. Eng., **9**, pp. 267–292.
[3] Lund, J. W., 1967, "A Theoretical Analysis of Whirl Instability and Pneumatic Hammer for a Rigid Rotor in Pressurized Gas Journal Bearings," ASME J. Lubr. Technol., **89**(2), pp. 154–166.
[4] Licht, L., Fuller, D. D., and Sternlicht, B., 1958, "Self-Excited Vibrations on an Air-Lubricated Thrust Bearing," Trans. ASME, **80**, pp. 411–414.
[5] Stowell, T. B., 1971, "Pneumatic Hammer in a Gas Lubricated Externally Pressurized Annular Thrust Bearing," ASME J. Lubr. Technol., **93**, pp. 498–503.
[6] Talukder, H. M., and Stowell, T. B., 2003, "Pneumatic Hammer in an Externally Pressurized Orifice-Compensated Air Journal Bearing," Tribol. Int., **36**(8), pp. 585–591.
[7] Powell, J. W., 1970, "A Review of Progress in Gas Lubrication," Review of Physics in Technology, **1**, pp. 96–129.
[8] Gross, W. A., 1962, *Gas Film Lubrication*, Wiley, New York, p. 279.
[9] Blok, H., and van Rossum, J. J., 1953, "The Foil Bearing—New Departure in Hydrodynamic Lubrication," Lubrication Engineering, **9**, p. 316.
[10] Patel, B. J., and Cameron, A., 1957, "The Foil Bearing," *Proceedings of the Conference on Lubrication and Wear*, October 1–3, 1957, Institution of Mechanical Engineers, London, p. 219.
[11] Dellacorte, C., and Valco, M. J., 2000, "Load Capacity Estimation of Foil Air Journal Bearings for Oil-Free Turbomachinery Applications," *Proceedings of the ASME/STLE Tribology Conference in Seattle*, Seattle, WA, Oct. 1–4, STLE Paper No. 00-TC-4.
[12] Heshmat, H., 1994, "Advancements in the Performance of Aerodynamic Foil Journal Bearings: High Speed and Load Capability," ASME J. Tribol., **116**, pp. 287–295.
[13] San Andres, L., 2006, "Hybrid Flexure Pivot-Tilting Pad Gas Bearings: Analysis and Experimental Validation," ASME J. Tribol., **128**, pp. 551–558.
[14] Tatara, A., Koike, H., and Iwasaki, A., 1973, "The Stability of Flexibly Supported, Externally Pressurized Gas Journal Bearings," Bull. JSME, **16**(100), pp. 1573–1579.
[15] Boffey, D., 1978, "A Study of the Stability of an Externally-Pressurized Gas-Lubricated Thrust Bearing With a Flexible Damped Support," ASME J. Lubr. Technol., **100**, pp. 364–368.
[16] Kazimierski, Z., and Jarzecki, K., 1979, "Stability Threshold of Flexibly Supported Hybrid Gas Journal Bearings," ASME J. Lubr. Technol., **101**, pp. 451–457.
[17] Ide, R. D., and Zeidan, F. Y., 1995, "Integral Squeeze Film Damper Bearings," U.S. Patent No. 5,421,655.
[18] Ide, R. D., 1991, "Hydrodynamic Bearings Having Beam Mounted Bearing Pads and Sealed Bearing Assemblies Including the Same," U.S. Patent No. 5,054,938.
[19] Della Pietra, L., and Adiletta, G., 2002, "The Squeeze Film Damper Over Four Decades of Investigations. Part I: Characteristics and Operating Features," Shock Vib. Dig., **34**(1), pp. 3–26.
[20] Adiletta, G., and Della Pietra, L., 2002, "The Squeeze Film Damper Over Four Decades of Investigations. Part II: Rotordynamic Analyses With Rigid and Flexible Rotors," Shock Vib. Dig., **34**(2), pp. 97–126.
[21] Zeidan, F. Y., San Andres, L., and Vance, J. M., 1996, "Design and Application of Squeeze Film Dampers in Rotating Machinery," *Proceedings of the 25th Turbomachinery Symposium*, Houston, TX, Sept. 17–19, Texas A&M Turbomachinery Laboratory, Houston, TX, pp. 169–188.
[22] Childs, D. W., 1978, "Space Shuttle Main Engine High-Pressure Fuel Turbopump Rotordynamic Instability Problem," J. Eng. Power, **100**, pp. 48–57.
[23] Zarzour, M., and Vance, J., 2000, "Experimental Evaluation of a Metal Mesh Bearing Damper," ASME J. Eng. Gas Turbines Power, **122**, pp. 326–329.
[24] Al-Khateeb, E. M., and Vance, J. M., 2001, "Experimental Evaluation of a Metal Mesh Bearing Damper in Parallel With a Structural Support," ASME Paper No. 2001-GT-0247.
[25] Ertas, B. H., Al-Khateeb, E. M., and Vance, J. M., 2003, "Rotordynamic Bearing Dampers for Cryogenic Rocket Engine Turbopumps," J. Propul. Power, **19**(4), pp. 674–682.
[26] Ertas, B. H., and Luo, H., 2007, "Nonlinear Dynamic Characterization of Oil-Free Wire Mesh Dampers," ASME Paper No. DETC2007/VIB-34665.
[27] Okayasu, A., Ohta, T., Azuma, T., Fujita, T., and Aoki, H., 1990, "Vibration Problems in the LE-7 Liquid Hydrogen Turbopump," *Proceedings of the 26th AIAA/SAE/ASME/ ASEE 26th Joint Propulsion Conference*, AIAA, Orlando, FL, pp. 1–5.
[28] Ertas, B. H., Drexel, M., Van Dam, J., and Hallman, D., 2008, "A General Purpose Test Facility for Evaluating Gas Lubricated Journal Bearings," *Proceedings of the 12th International Symposium on Transport Phenomena and Dynamics of Rotating Machinery Honolulu*, Honolulu, HI, Feb. 17–22, Paper No. ISROMAC12-2008-20207.
[29] Dellacorte, C., 1997, "A New Foil Air Bearing Test Rig for Use to 700°C and 70,000 rpm," NASA, Technical Memorandum No. 107405.
[30] Childs, D. W., and Hale, K., 1994, "A Test Apparatus and Facility to Identify the Rotordynamic Coefficients of High Speed Hydrostatic Bearings," ASME J. Tribol., **116**, pp. 337–334.
[31] Ghosh, M. K., and Viswanath, N. S., 1987, "Recess Volume Fluid Compressibility Effect on the Dynamic Characteristics of Multirecess Hydrostatic Journal Bearings With Journal Rotation," ASME J. Tribol., **109**, pp. 417–426.
[32] San Andres, L. A., 1992, "Analysis of Turbulent Hydrostatic Bearings With a Barotropic Cryogenic Fluid," ASME J. Tribol., **114**, pp. 755–765.
[33] San Andres, L. A., 1996, "Turbulent Flow Flexure Pivot Hybrid Bearings for Cryogenic Applications," ASME J. Tribol., **118**, pp. 190–200.
[34] San Andres, L. A., 2000, "Bulk Flow Analysis of Hybrid Thrust Bearings for Process Fluid Applications," ASME J. Tribol., **122**, pp. 170–180.
[35] Michael, W. A., 1959, "A Gas Film Lubrication Study Part II: Numerical Solution of the Reynolds Equation of Finite Slider Bearings," IBM Syst. J., pp. 256–259.

Fatigue Crack Growth Behavior Evaluation of Grainex Mar-M 247 for NASA's High Temperature High Speed Turbine Seal Test Rig

Irebert R. Delgado
e-mail: irebert.r.delgado@nasa.gov

Bruce M. Steinetz

Glenn Research Center,
National Aeronautics and Space Administration,
21000 Brookpark Road,
Cleveland, OH 44135

Clare M. Rimnac

John J. Lewandowski

Case Western Reserve University,
Cleveland, OH 44106

The fatigue crack growth behavior of Grainex Mar-M 247 is evaluated for NASA's turbine seal test facility. The facility is used to test air-to-air seals primarily for use in advanced jet engine applications. Because of extreme seal test conditions of temperature, pressure, and surface speeds, surface cracks may develop over time in the disk bolt holes. An inspection interval is developed to preclude catastrophic disk failure by using experimental fatigue crack growth data. By combining current fatigue crack growth results with previous fatigue strain-life experimental work, an inspection interval is determined for the test disk. The fatigue crack growth life of NASA disk bolt holes is found to be 367 cycles at a crack depth of 0.501 mm using a factor of 2 on life at maximum operating conditions. Combining this result with previous fatigue strain-life experimental work gives a total fatigue life of 1032 cycles at a crack depth of 0.501 mm. Eddy-current inspections are suggested starting at 665 cycles since eddy current detection thresholds are currently at 0.381 mm. Inspection intervals are recommended every 50 cycles when operated at maximum operating conditions. [DOI: 10.1115/1.2980058]

Introduction

NASA's high temperature high-speed turbine seal test rig (hereafter NASA seal rig) (Fig. 1) tests engine seals at unique combinations of temperature, pressure, and surface speed. Because of these extreme test conditions, surface cracks may develop over operational time in the bolt holes of the Grainex Mar-M 247 disk (hereafter named GXMM247), which is used as a seal runner. If allowed to grow over time, catastrophic disk failure may occur. To preclude disk failure, periodic disk bolt hole inspections are proposed using eddy-current inspection techniques.

Eddy-current inspection is a nondestructive evaluation method used to detect cracks or flaws on component surfaces and inside turbine disk bolt holes where visual inspection techniques are not practical. Eddy-current inspection of turbine disks follows a damage tolerant philosophy widely used in the aircraft industry of detecting a flaw or crack in a component prior to it growing to a critical size and causing catastrophic failure. Eddy-current inspection requires the use of a "master flaw" sample of known dimensions to compare to readings from the actual part. The flaw would be some fraction of the critical flaw size (i.e., safety factor). Ideally, this critical flaw size is determined through compact specimen tests. However, due to a lack of available material for these specimens, surface flawed fatigue crack growth (FCG) specimens were manufactured instead to obtain corresponding FCG data. Typically, the upper portion of the FCG curve on a da/dN versus ΔK plot coincides with the critical fatigue crack growth rate, which is directly related to the critical flaw size. Knowing the critical flaw size and the number of cycles to reach this flaw size, the total fatigue crack growth life of the disk bolt holes is determined.

This paper presents FCG experimental data on the GXMM247 superalloy material. Previous work by Delgado et al. [1] have provided fatigue strain-life experimental results for the material. Combining these results with the current FCG experimental data would account for the total cyclic life of the disk bolt hole surfaces. Based on these results and an additional factor of safety, a corresponding eddy-current inspection interval is recommended.

Material and Method: Experiment. Fatigue crack growth tests on the GXMM247 material were at 649°C, a load ratio of $R=0.05$, and a sinusoidal waveform at a cyclic frequency of 0.33 Hz. A surface flaw (or K_b) specimen (Fig. 2) was used to simulate the growth of a flaw within a disk bolt hole. The K_b specimen was first developed by Coles et al. [2] for evaluating turbine engine components. The K_b specimen was fabricated by electro-discharged machining (EDM) a 41 mm long by 14 mm diameter cylinder from a sacrificial GXMM247 disk (Fig. 3). The cylinders were removed at locations adjacent to the disk bolt holes and oriented perpendicular to the bolt hole axis. This oriented the machined surface flaw on the K_b specimen perpendicular to the applied cyclic stress, since cracks tend to initiate within bolt holes from circumferential and radial disk stresses during rotation [3]. For reference, the axial direction of the disk is parallel with the bolt holes. Inconel 718 cylinders were inertia-welded to both ends of the GXMM247 cylinders and the resultant ends were machined to a buttonhead shape to allow the fatigue testing machine to grip the specimen. Then, the GXMM247 portion of the specimen was machined to transition from a cylindrical shaped cross section at the Inconel 718/GXMM247 interface to a rectangular cross section at the specimen center or gauge section. Finally, a surface flaw was plunge electro-discharged machined on one side of the rectangular gauge section in the shape of half a circular disk (Fig. 4). The initial depth and width of the surface flaw were approximately half of those dimensions shown in Fig. 4 after failures outside the gauge section occurred in preliminary tests. Thus, the final surface flaw size used was 0.46 mm in depth by 0.91 mm in width. This EDM notch enlargement was necessary to encourage crack growth at the surface flaw. All FCG specimens (Fig. 5) were fabricated at Mar-Test in Cincinnati, OH.

FCG tests were performed at the NASA Glenn Research Center fatigue and fracture laboratories. A Materials Test System 810

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received January 10, 2008; final manuscript received May 28, 2008; published online December 22, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2007: Land, Sea and Air (GT2007), Montreal, Quebec, Canada, May 14–17, 2007.

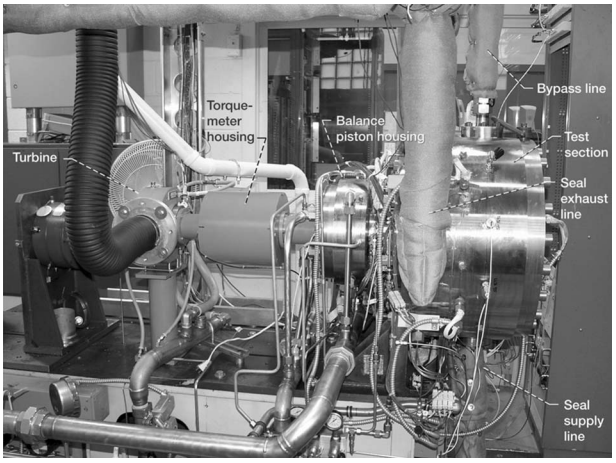


Fig. 1 NASA's turbine seal test facility

(MTS Corporation, Minneapolis, MN) was utilized (Fig. 6). The system is computer-controlled by a custom program, Material Analysis and Test Environment (MATE) which also acquires the raw FCG data [4]. Crack length computations for the K_b specimen are based on a closed-form analytical model by Gangloff and co-workers [5–7] that has been experimentally confirmed. FCG of the K_b specimen was measured using the direct current electrical potential difference (dcEPD) method. The method was initially developed by Gangloff [8] for small surface cracks in hour-glass shaped specimens. The method was modified by Vanstone and Richardson [9] for specimens having rectangular cross sections with semicircular EDM notches. FCG in a semicircular crack is related to voltage through an analytical model developed by Roe and Coffin. This is referenced in Gangloff [8]. To facilitate potential drop measurements, 0.13 mm diameter alumel wires were tack-welded to either side of the EDM notch (Fig. 7).

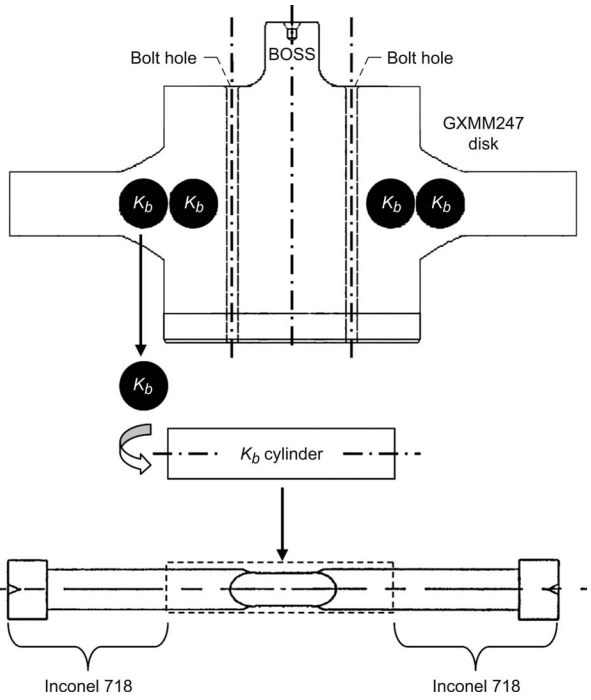


Fig. 3 Location of fatigue crack-growth K_b specimens from the sacrificial GXMM247 disk

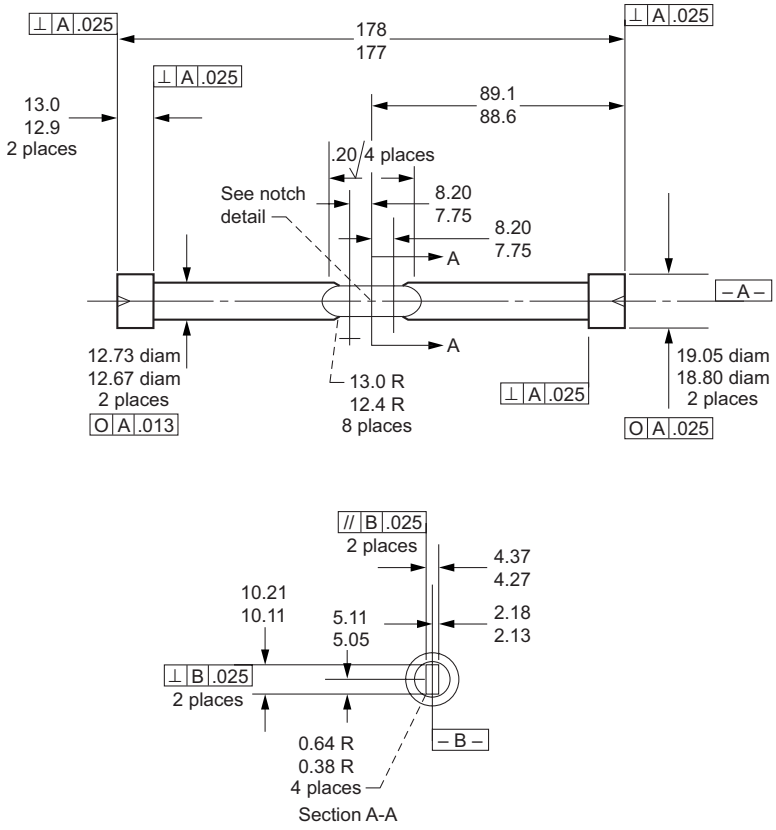


Fig. 2 Surface flaw (K_b) specimen geometry for the GXMM247 fatigue crack growth tests at 649°C. Dimensions are in millimeters.

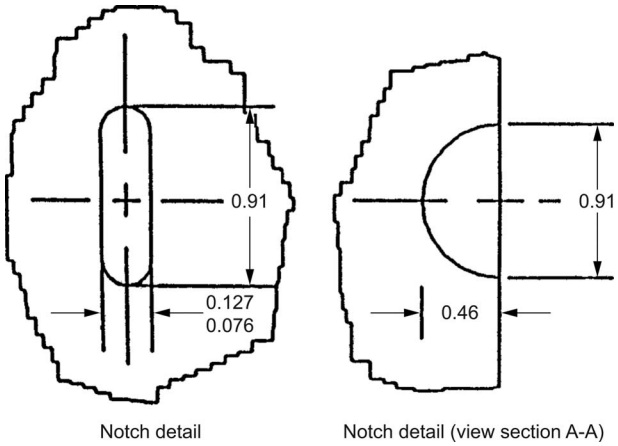


Fig. 4 Surface flaw geometry for the (K_b) specimen. Dimensions are in millimeters. See Fig. 2 for notch location on the (K_b) specimen.

Measurements of the EDM notch geometry (Fig. 8) were taken prior to installing the K_b specimen in the load frame. The measurements included the notch height, width, and depth as well as the potential probe spacing. Measurements were taken with a Nikon Measurescope 10 (Nikon, Tokyo, Japan) and a Nikon SC-102 X-Y digital readout (Nikon, Tokyo, Japan). The Measurescope has an accuracy of ± 0.0025 mm. The measurements were used as inputs into MATE to calculate for crack length, maximum load, and dcEPD voltage during the test. An assumption on the EDM crack depth was initially inputted into MATE, based on original EDM notch specifications, and was subsequently corrected during post-test calculations.

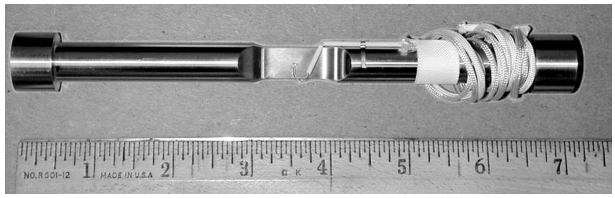


Fig. 5 K_b specimen with alumel wire attachments for potential drop crack growth measurement

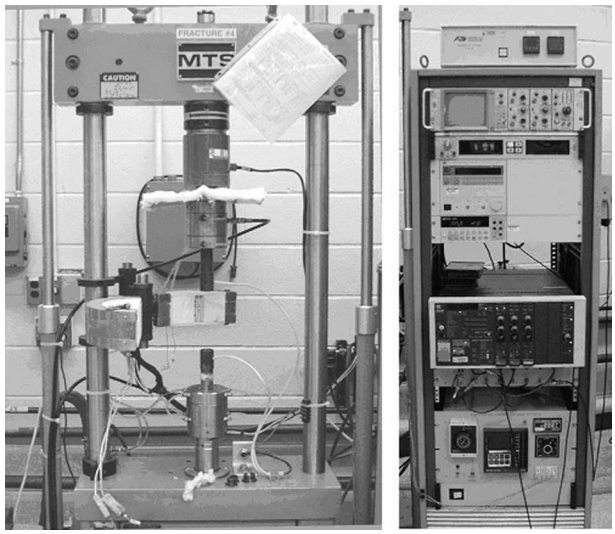


Fig. 6 Fatigue crack growth test equipment and controls at NASA's fatigue and fracture laboratories

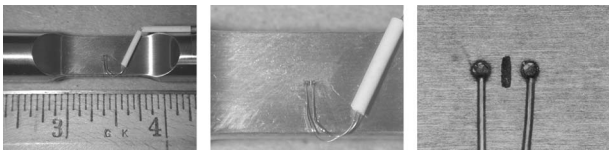


Fig. 7 Alumel wire location at the notch area of the K_b specimen

The K_b specimen was installed into the MTS load frame with 12-gauge copper leads attached to the buttonheads. A constant 10 A was applied to the leads, forming a circuit through the specimen and resulting in a potential drop across the alumel wire locations (Fig. 7), which increases with increasing crack length and depth. Baseline voltage measurements were taken to adjust for thermoelectric influences [10]. Prior to starting the precracking portion of the test, initial voltages were taken at 0 N and 17.8 kN axial loads on the K_b specimen. The specimen was precracked at room temperature with a sinusoidal waveform at 3 Hz, and $R=0.05$ ($P_{max}=25.8$kN) and grown to a depth of approximately 0.81 mm. Precracking was conducted to ensure that FCG occurred beyond the recast layer resulting from the EDM notch. Both crack length voltages and number of cycles were acquired. After precracking, reference voltage measurements (V_N) were again taken at 0 N and 17.8 kN axial loads. These voltages were used to cor-

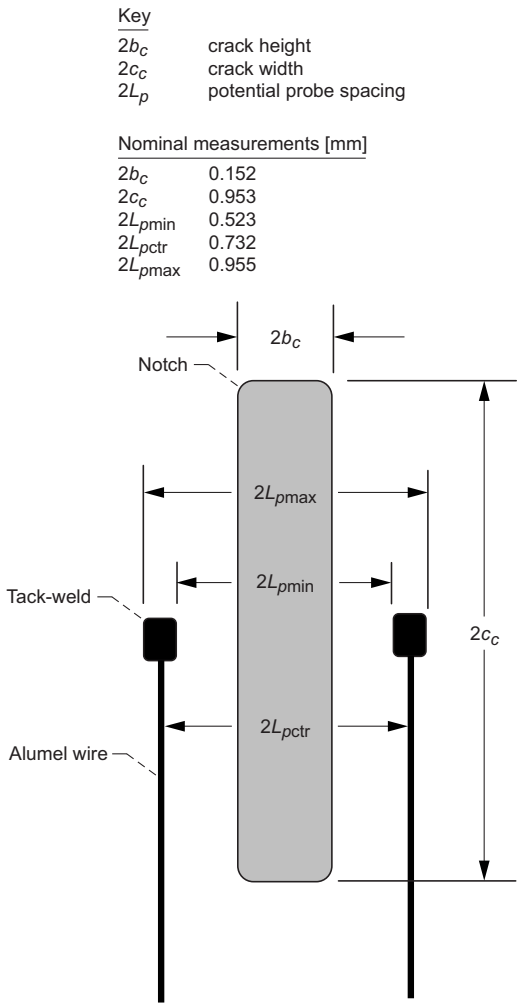


Fig. 8 Nominal fatigue crack growth pretest measurement for the K_b specimen surface flaw geometry

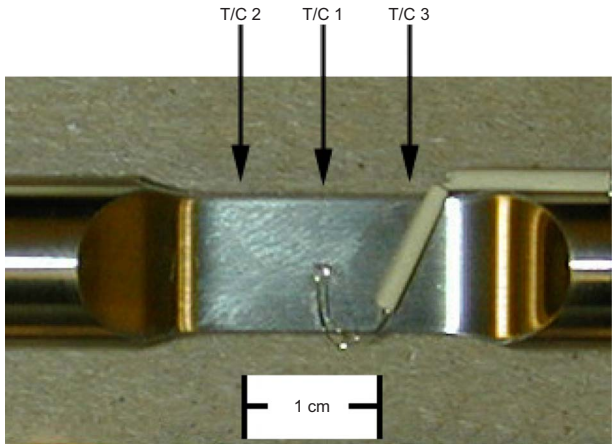


Fig. 9 Thermocouple locations on the K_b specimen gauge section indicated by the arrows. Thermocouples are not shown.

rect for voltage drift and for normalizing the acquired crack length voltages for data reduction. Reference voltages were taken prior to the precrack test and the actual crack-growth test at 649°C.

Three type-K thermocouples were tack-welded to the gauge section of the K_b specimen (Fig. 9). Thermocouple 1 (T/C 1) measured the K_b specimen test temperature and was centered in the middle of the gauge section. T/C 2 and T/C 3 controlled the specimen temperature and were spaced 6.35 mm on either side of the center thermocouple. Two half-section resistance furnaces (Fig. 6) were then positioned around the K_b specimen. The specimen was heated to 649 ± 1°C at zero load with the K_b surface temperature read-out through a digital meter. Initial voltage potentials were measured at axial loads of 0 N and 17.8 kN. An initial voltage was also taken to account for thermoelectric influences. The test was conducted with a sinusoidal waveform at 0.33 Hz, and a ratio of $R=0.05$ ($P_{max}=27.4$ kN). For each acquisition, MATE recorded the cycle count, total crack length, maximum load, corrected dcEPD voltage, number of dcEPD points acquired, thermoelectrically induced voltage, and input load ratio. Data were taken approximately every 0.0508 mm of the crack extension [11]. The crack was allowed to grow to 2.54 mm. This crack length limitation was due to the 4.32 mm gauge thickness. Experience has shown that the crack length voltage accuracy begins to drop off for crack lengths at approximately 65% of the gauge thickness [12]. Final voltages were again taken at axial loads of 0 N and 17.8 kN. After the 2.54 mm crack depth was attained, the test specimen was fractured by monotonically increasing the load to specimen failure.

Data Reduction—Fatigue Crack Growth Rate. FCG data are reduced to a Paris type relationship [13], where the FCG rate, da_c/dN, is related to the stress intensity factor, ΔK:

$$\frac{da_c}{dN} = C(\Delta K)^m \tag{1}$$

As mentioned previously, crack depth data were corrected based on post-test measurements of the fracture surface (Figs. 10 and 11). These measurements were inputted into MATE. Crack depth and length per cycle were calculated iteratively within MATE, which compares the measured voltage for a calculated crack depth to the voltage predicted by the Roe–Coffin potential solution [9] for a semi-elliptical surface notch. The Roe–Coffin solution, given by Gangloff [8], has the form

$$\frac{V}{V_N} = f(a_c, b_c, c_c, a_n, b_n, c_n, L_p) \tag{2}$$

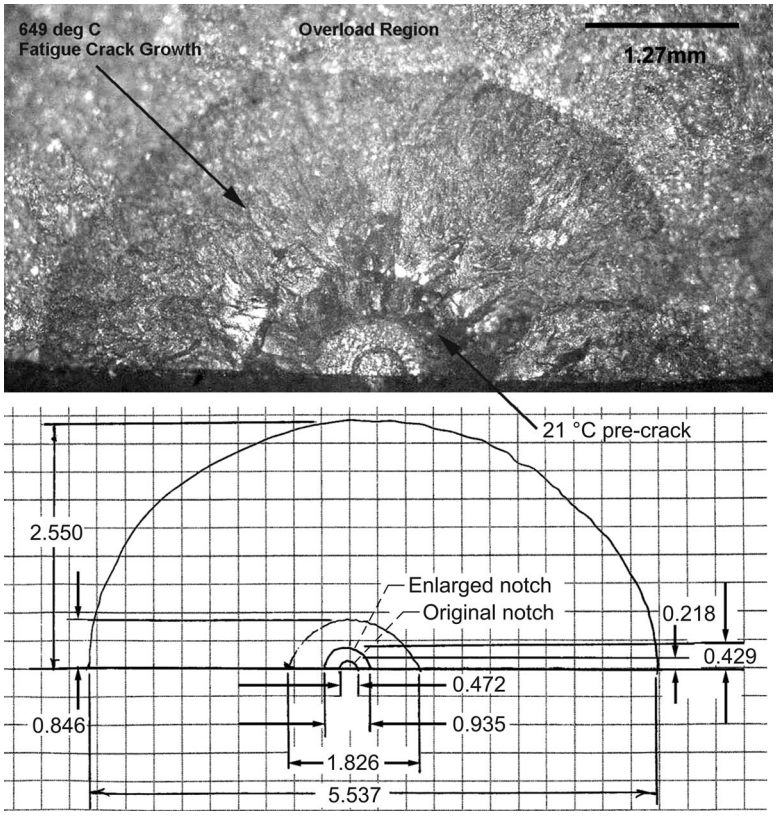


Fig. 10 Fracture surface measurements for the GXMM247 (K_b specimen 13K). Dimensions are in millimeters.

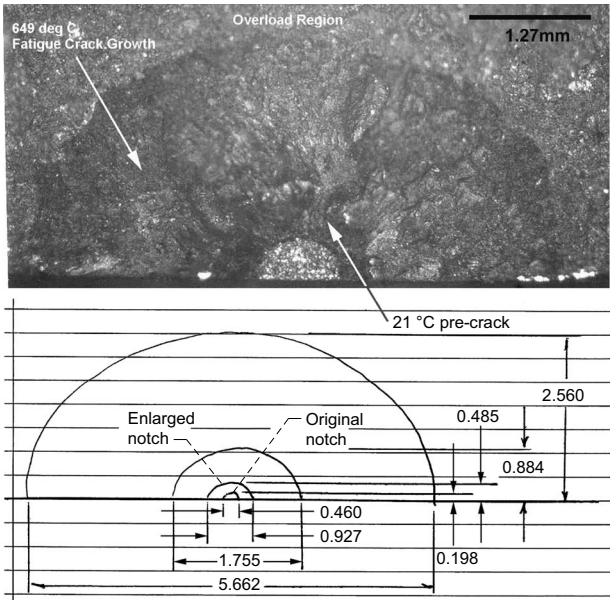


Fig. 11 Fracture surface measurements for the GXMM247 (K_b specimen 32K). Dimensions are in millimeters.

The crack depth and length were iterated within MATE such that the difference between the measured and predicted voltages of the crack depth and length were minimized. This was done by holding constant V/V_N, a_n, b_n, c_n, and L_p, and holding constant the aspect ratio, c_c/a_c, for each increment of crack depth acquired during the test. The procedure was identical for the crack length.

The fatigue crack growth rate, da_c/dN, was calculated using the incremental polynomial method, per ASTM E647 [14]. The stress intensity range, ΔK, was calculated within MATE using stress-intensity factor equations developed by Newman and Raju [13] for a semi-elliptical surface crack subjected to tensile loading

$$K_I = s\sqrt{\pi\frac{a_c}{Q}F_s\left(\frac{a_c}{c_c},\frac{a_c}{t},\frac{c_c}{b_c},\phi\right)} \tag{3}$$

The maximum stress intensity factor, $K_{\max}$, was determined by inputting the maximum recorded load at each data point into Eq. (2), which is programmed into MATE. Finally, ΔK [15] was calculated using

$$\Delta K = K_{\max}(1-R) \tag{4}$$

The resultant da_c/dN versus ΔK data were then plotted on log-log coordinates.

Data Reduction—Statistics. Linear regression analysis was conducted to evaluate the log-log transformed data

$$\log\left(\frac{da_c}{dN}\right) = \log(C) + m\log(\Delta K) \tag{5}$$

A statistical comparison between regression lines was performed to determine if the individual slopes and intercepts were statistically similar at a confidence level of 95%. If this were true, then resultant data sets from individual tests could be combined to provide a statistically stronger relationship between the fatigue crack growth rate and the stress intensity range. A test of equal variances on the error terms for each regression line was first done to validate a comparison between regression lines [16]. To test for similar slopes and intercepts, a reduced statistical model was determined, which combined data sets from successful FCG tests. The reduced statistical model was

$$Y_{ij} = \beta_0 + \beta_1 X_{ij} + \varepsilon_{ij} \tag{6}$$

A test statistic [16], F^*, was calculated and compared with the F value determined from F-tables based on a predetermined significance level (0.10) and degrees of freedom (120). F^* is defined as

$$F^* = \frac{\mathrm{SSE}(R) - \mathrm{SSE}(F)}{2} \div \frac{\mathrm{SSE}(F)}{n_1 + n_2 - 4} \tag{7}$$

For $F^* \leq F(1-a;2,n_1+n_2-4)$ a conclusion can be made that the slopes and intercepts of both regression lines are equal. Note that "a" is the desired percentile level.

FCG Behavior Considerations

Limitations on crack behavior were considered in large part due to the relatively coarse grain structure of the GXMM247 material. To validate the linear-elastic fracture mechanics (LEFM) approach, evaluations were made on the small-crack behavior, specimen geometry with respect to grain size, and plastic-zone size with respect to grain size.

Small Crack Growth Behavior. Because of the small size of the EDM notch dimensions (Fig. 4) with respect to the GXMM247 grain size of 1.6 mm [17], it was possible that FCG would occur in the small crack growth regime. Small cracks are defined as having dimensions equal to or smaller than the dimension of greatest microstructural significance, such as grain size [15]. Small cracks are characterized by higher growth rates and their ability to grow at ΔK values below the threshold stress intensity range, ΔK_{th} [18]. Small cracks may decelerate and arrest or approach a minimum in FCG rate, then accelerate, and merge with long crack growth behavior. Thus, the FCG behavior and respective crack size relative to the grain size were evaluated to determine if FCG followed small crack behavior.

Linear-Elastic Fracture Mechanics. An assessment of the validity of LEFM was made based on the specimen geometry. The surface-flawed specimen geometry was considered due to the lack of available GXMM247 material. LEFM [15] limitations were compared to specimen geometry based on

$$a_c,(t-a_c),h \geq \frac{4}{\pi}\left(\frac{K_{\max}}{\sigma_{y(0.2\%)}}\right)^2 \tag{8}$$

Plastic-Zone Size. The plastic zone size was compared with the average grain size of GXMM247. Ideally, the plastic zone should carry a number of grains within the material such that the FCG behavior is microstructurally-insensitive. The plastic zone size for plane stress conditions [15], $2r_{0\sigma}$, was estimated by

$$2r_{0\sigma} = \frac{1}{\pi}\left(\frac{K_{\max}}{\sigma_{y(0.2\%)}}\right)^2 \tag{9}$$

Fractographic Examination. Fractographic examination of fracture surfaces was performed to identify grain boundaries that may affect FCG and to determine if fatigue striations were a good indicator of cyclic stress intensity.

Fatigue crack propagation surfaces from each specimen were prepared for examination by carefully removing the gauge section with a cut-off wheel. The surface was cleaned and then mounted. Fracture surfaces were examined using a light microscope and a scanning electron microscope (SEM) under secondary electron (SE) and backscattered electron (BE) modes.

Fracture surface features were identified including initial and final EDM notches, the precrack region, low ΔK (=23.3–26.4 MPa$\sqrt{m}$) and high ΔK (=35.7–39.1 MPa$\sqrt{m}$) regions, and the overload region. SEM micrographs of the low and high ΔK regions were taken, respectively, at 1.00–1.25 mm and 2.00–2.25 mm from the EDM notch origin. A comparison was made between the crack growth rate and crack depth. The microscopic crack growth rate was determined for the low and high ΔK

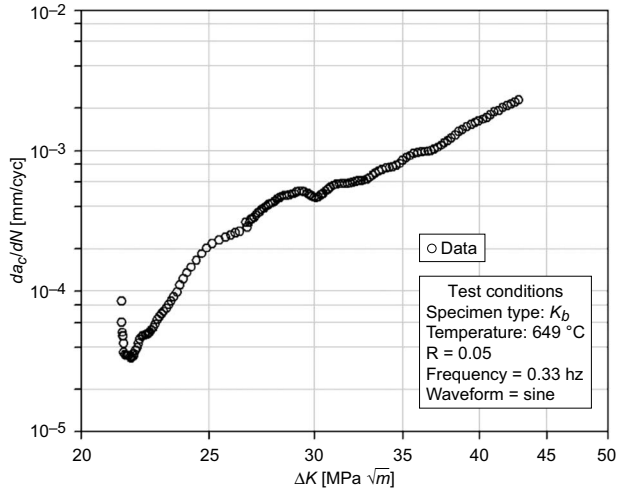


Fig. 12 Fatigue crack growth behavior for GXMM247 at 649°C (K_b specimen 13K)

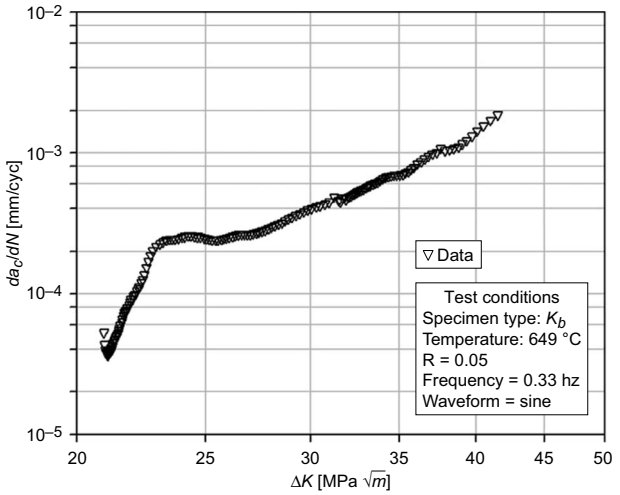


Fig. 13 Fatigue crack growth behavior for GXMM247 at 649°C (K_b specimen 32K)

regions by the measurement of striation spacings.

A comparison was made between the experimentally applied ΔK and the calculated ΔK based on striation spacings at the low and high ΔK fracture surface regions using the Bates and Clark relationship [19]

$$\text{striation spacing} = 6\left(\frac{\Delta K}{E}\right)^2 \tag{10}$$

where Young's modulus, E, was taken to be 194 GPa (see Ref. [1]). The experimental ΔK used to compare with the Bates and Clark relationship was taken from the averaged ΔK determined from the low ΔK regime (24.85 MPa$\sqrt{m}$) and high ΔK regime (37.40 MPa$\sqrt{m}$).

Results and Discussion

Test Summary. Of the original eight K_b specimens, two tests (using specimens identified as 13K and 32K) were successfully conducted after doubling the EDM notch width and depth. Four K_b specimens, previously tested, failed prematurely outside the EDM notch. The remaining two were untested. Hereafter, the results pertain only to K_b specimens 13K and 32K.

Experimental Data. The FCG behavior for the two tests specimens is shown in Figs. 12 and 13 with the coefficient, C, and exponent, m, from Eq. (1) given for each specimen in Table 1.

The deceleration and then acceleration of the da_c/dN data in the low ΔK region may be due to the crack approaching microstructural barriers such as grain boundaries or due to crack closure [18,20]. This small crack growth behavior generally transitions into long crack growth behavior at higher ΔK values [21], and this appears to have been the case in this study.

Statistics. FCG data with regressions from both data sets were analyzed for similarity. The variances between the individual regressions were found to be equivalent to a 95% confidence level. Thus a comparison between the regressions for similarity was valid. The reduced model Eq. (6) was determined with $\beta_0 = -11.20$ and $\beta_1 = 5.28$. Finally, F^* from Eq. (7) was calculated as −72.56. From standard F-tables, F (0.999, 2, 120) was 7.32. Thus $F^* \leq F$, and the slopes and intercepts were equal to within a 99.9% confidence level. The data from the two specimens were thus combined into one data set and an overall coefficient, C, and exponent, m, were determined (Fig. 14).

To be conservative in predicting the FCG life, both data sets (13K and 32K) were used with regression analysis over the entire experimentally determined range. The steeper da_c/dN versus ΔK slope in both data sets at low ΔK was thus taken into account, as well as the shallower slope at high ΔK.

FCG Behavior Considerations

Small Crack Growth Behavior. Small crack growth behavior was observed for both FCG specimens (Figs. 12 and 13). That is, the crack growth rate for both specimens was observed to initially decelerate to a minima and then to transition into the long crack growth regime where more uniform FCG behavior is expected.

In this study, the transition from small crack growth to long crack growth appeared to occur at approximately $\Delta K = 30$ MPa$\sqrt{m}$ ($da_c/dN = 4 \times 10^{-4}$ mm/cycle) for both FCG tests (Fig. 14), based on the similarity in the da_c/dN versus ΔK behavior above the transition. Consistent with this, the corresponding

Table 1 Linear regression results for the fatigue crack growth behavior of GXMM247 K_b specimens (13K and 32K) at 649°C for the entire stress intensity range tested

Specimen (K)	Intercept	Slope	R^2	Constant C	Exponent m
13K	−11.98	5.80	0.93	1.05×10^{-12}	5.80
32K	−10.53	4.82	0.90	2.94×10^{-11}	4.82

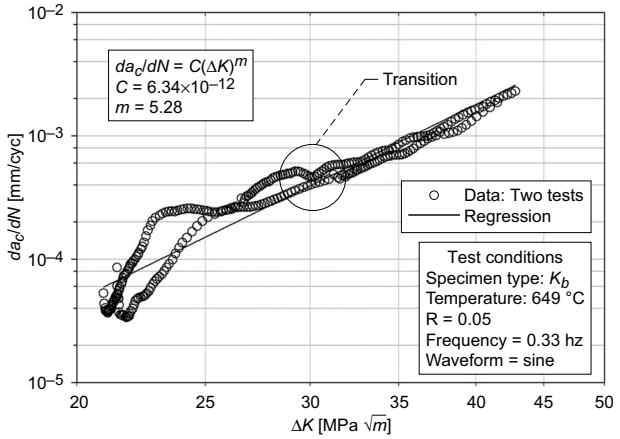


Fig. 14 Regression results using both specimen data sets for the GXMM247 fatigue crack growth K_b data

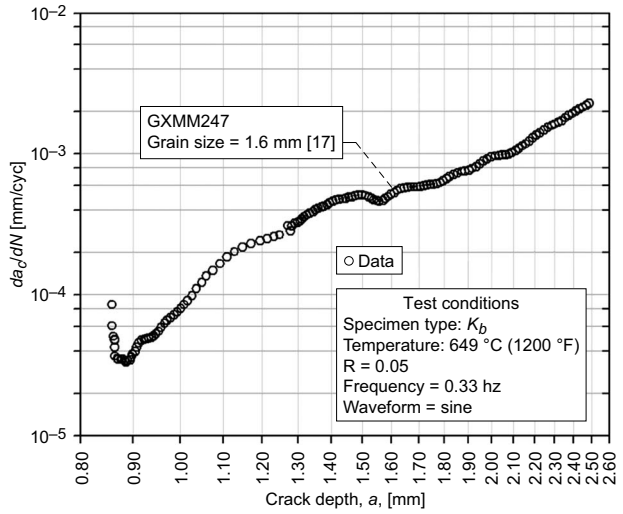


Fig. 15 Fatigue crack growth rate versus the crack depth for GXMM247 at 649°C (K_b specimen 13K)

range of crack depths at that stress intensity was between 1.55 mm and 1.65 mm (Fig. 15). This coincides with reported GXMM247 average grain sizes of 1.6 mm. In other words, once the crack grows larger than the grain size, long crack growth behavior is expected.

Linear Elastic Fracture Mechanics. Initial and final crack lengths for both FCG test specimens were compared to the K_b specimen geometry to determine LEFM applicability using Eq. (8) (Table 2). Initial crack lengths did not meet LEFM criteria for either specimen. Final crack lengths and the remaining gauge specimen thickness (i.e., $t-a_c$) also did not meet LEFM criteria.

The use of a LEFM approach for FCG tests assumes that plasticity is limited to a local region that is small compared to the specimen dimensions [15]. Predicted gross yielding of the specimen makes the use of a LEFM approach questionable. In this study, during FCG testing, the initial maximum stress level at the gauge section, calculated using the uncracked ligament, was approximately 79% of the 0.2% offset yield stress at 649°C. The final maximum stress level at the gauge section, calculated using the uncracked ligament, was nearly 118% of the 0.2% offset yield stress. Thus, the FCG behavior in these tests may have been influenced by gross yielding in the gauge region. This suggests that the FCG behavior of GXMM247 may be worse than the experimentally-obtained data. Obviously, fatigue crack propagation tests of the GXMM247 material using a specimen geometry that meets LEFM criteria would help answer this question. Because of the uncertainty in the data, a safety factor of 2 will be used to predict the FCG life. This is based on a reasonable engineering assessment.

Table 3 Initial and final plastic zone sizes for GXMM247 K_b fatigue crack growth specimens (13K and 32K) for plane stress conditions at 649°C

	Plane stress plastic zone size	
Specimen (K)	Initial (mm)	Final (mm)
13K	0.241	0.960
32K	0.230	0.902

Limitations With Plastic Zone Size. The initial and final plane stress plastic zone sizes for both specimens 13K and 32K using Eq. (9) can be found in Table 3. Ideally, the plastic zone size should carry a number of grain sizes in a material. Due to the relatively coarse grain size of the GXMM247 material, the final plane stress plastic zone size did not exceed the average grain size of 1.6 mm. Thus, the recorded da_c/dN versus ΔK behavior was likely to be microstructurally-sensitive [22]. Specifically, grain boundaries may inhibit the crack from growing causing a deceleration in the FCG rate. Alternatively a crack growing within a grain may show an acceleration in the FCG rate.

Fractographic Examination. A SEM photo of the fracture surface of specimen 13K is shown in the SE mode (Figs. 16 and 17). Fracture surface features for K_b specimen 32K were similar.

Under the BE mode, grain boundaries were observed in the fatigue crack propagation region beginning between 0.75 mm and 1.00 mm radii from the EDM notch origin (Fig. 18). Note that the 649°C FCG began at approximately 0.82 mm from the EDM notch origin.

The decreasing crack growth rate at low ΔK (Fig. 12) at a crack depth between 0.85 mm and 0.90 mm may be related to the grain boundary (Fig. 18) between 0.75 mm and 1.50 mm from the notch origin. This deceleration in crack growth rate may be due to the crack front encountering a microstructural barrier, such as a grain boundary. In fact, the minima in the FCG rate was found to correspond to a crack depth of 0.88 mm (Fig. 15) for specimen 13K. This is consistent with findings by Taylor [18] and Suresh and Ritchie [20].

The low ΔK region (Fig. 19) of specimen 13K between 1.00 mm and 1.25 mm from the EDM notch origin showed pockets of fatigue striations. The microstructural FCG rate per cycle was approximately 2.5×10^{-4} mm/cycle (Fig. 15). In contrast, the

Table 2 Comparison of initial and final crack lengths of K_b specimens (13K and 32K) with LEFM applicability at 649°C

Initial (mm)	13K	13K criteria	Pass/ fail	Initial (mm)	32K	32K criteria	Pass/ fail
a_c	0.846	0.963	Fail	a_c	0.884	0.919	Fail
$(t-a_c)$	3.472	0.963	Pass	$(t-a_c)$	3.434	0.919	Pass
h	7.976	0.963	Pass	H	7.976	0.919	Pass
Final (mm)	13K	13K criteria	Pass/ fail	Final (mm)	32K	32K criteria	Pass/ fail
a_c	2.550	3.840	Fail	a_c	2.560	3.607	Fail
$(t-a_c)$	1.768	3.840	Fail	$(t-a_c)$	1.758	3.607	Fail
h	7.976	3.840	Pass	h	7.976	3.607	Pass

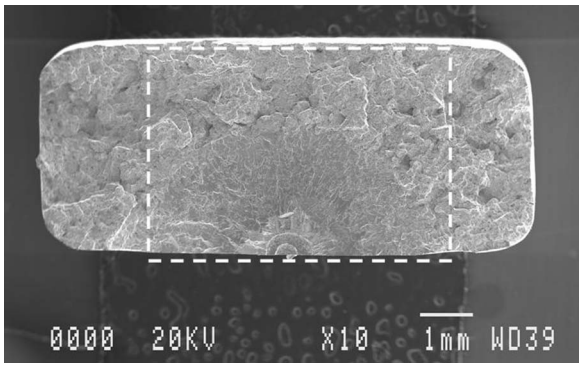


Fig. 16 GXMM247 fatigue crack growth fracture surface for K_b specimen 13K at 649°C. Dotted area is enlarged in Fig. 17.

high ΔK region (Fig. 20) shows more organized and delineated striations and the microstructural FCG rate is approximately 5×10^{-4} mm/cycle (Fig. 15). Specimen 32K showed similar

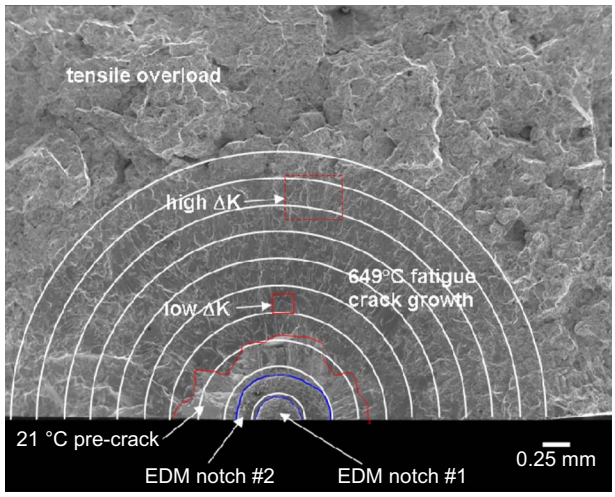


Fig. 17 Close-up of GXMM247 fatigue crack growth fracture surface features for K_b specimen 13K at 649°C. Low ΔK region: 23.3–26.4 MPa and high ΔK region: 35.7–39.1 MPa. SEM (SE mode).

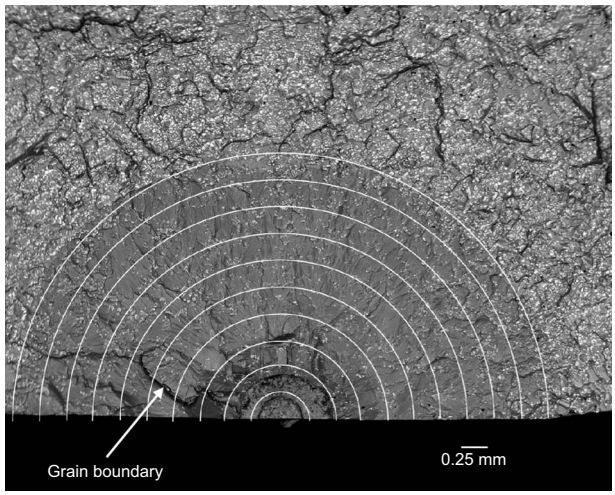


Fig. 18 GXMM247 fatigue crack growth fracture surface for K_b specimen 13K at 649°C showing the grain boundary near beginning of fatigue crack growth (SEM BE mode)

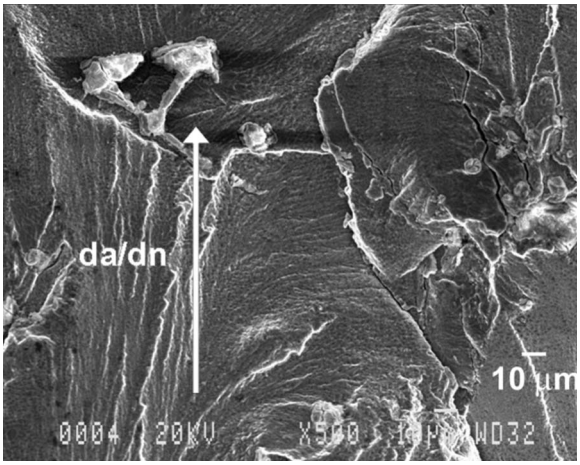


Fig. 19 Low ΔK fatigue striations of the GXMM247 fatigue crack growth fracture surface for K_b specimen 13K at 649°C (SEM SE mode, see also Fig. 17). ΔK range from 23.3 Mpa to 26.4 MPa. Arrow indicates the crack growth direction.

results.

Table 4 compares the experimentally applied ΔK and the calculated ΔK based on striation spacings taken from the low and high ΔK regions. The stress intensity ranges were overestimated by the Bates and Clark relationship, Eq. (10), by a factor of 5 for both low and high ΔK regions.

Adjusting Eq. (10) by the percent striated area of the fracture surface may result in better agreement between the experimental and calculated ΔK [19]. A rigorous survey of the fractographic

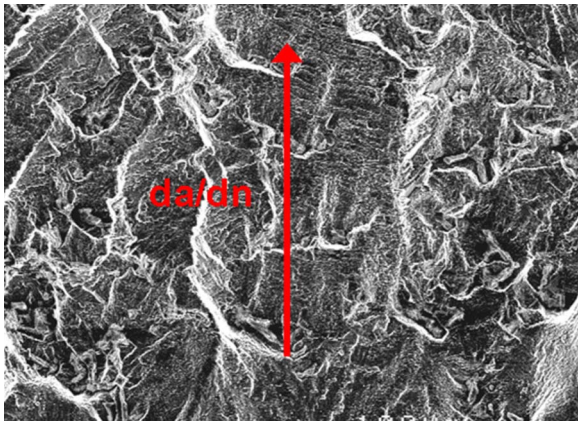


Fig. 20 High ΔK fatigue striations of the GXMM247 fatigue crack growth fracture surface for K_b specimen 13K at 649°C (SEM SE mode; see also Fig. 17). ΔK range from 35.7 Mpa to 39.1 MPa. Arrow indicates the crack growth direction.

Table 4 Comparison between the experimental ΔK and calculated ΔK from striation spacing measurements for GXMM247 at 649°C

ΔK region	ΔK (experimental)[a] (MPa$\sqrt{m}$)	Striation spacing (mm/cycle)	ΔK (Bates and Clark) [19] (MPa$\sqrt{m}$)
Low	24.85	2.5×10^{-3}	125.2
High	37.40	5.0×10^{-3}	177.1

[a]Average.

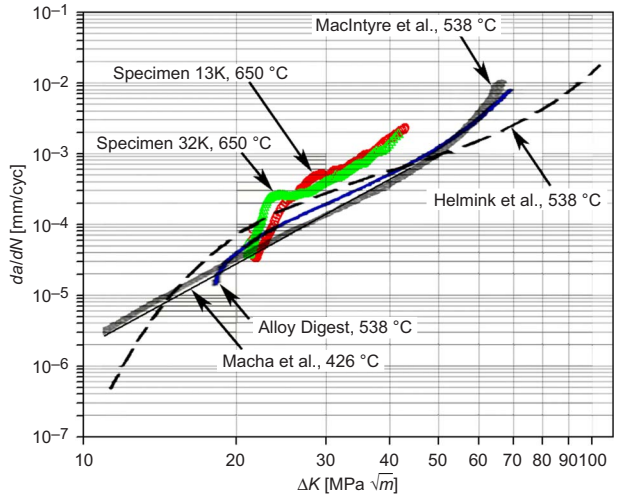


Fig. 21 Comparison of GXMM247 to literature [23–26]

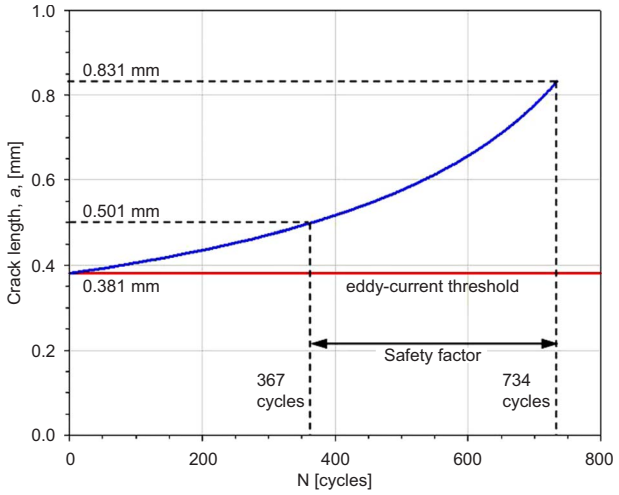


Fig. 22 GXMM247 NASA disk fatigue crack growth rate with increasing cycles at 649°C

area would be needed to determine if the Bates and Clark relationship would be a reliable predictor of cyclic stress intensity.

Comparison of Data to Literature. The regression of the combined experimental data was compared to literature at 426°C and 538°C (Fig. 21 and Table 5). The FCG rate increases with the test temperature at a constant cyclic stress-intensity. Data reported by Macha et al. [23], MacIntyre and Agarwal [24], Helmink et al. [25], and Alloy Digest [26] support this observation. The data from this study had the highest FCG rate, since tests were conducted at the highest test temperature, 649°C.

Determination of the NASA Disk Inspection Interval. As discussed previously, eddy-current inspections are proposed to examine the NASA disk bolt holes after a set number of cycles at maximum operating conditions in the NASA seal rig. A cycle is defined as a ramp up to the maximum speed and a ramp down to zero speed at maximum temperature and pressure. In general, jet engine users employ eddy current or other nondestructive inspection techniques at regular intervals to evaluate engine components for excessively long cracks. Eddy current inspection intervals are based on a critical crack length, a_{crit}, from K_{IC} tests at a specific test temperature, stress level, and detectable threshold crack size. As mentioned previously, not enough GXMM247 material was available for a proper K_{IC} test specimen. Alternatively, a surface-flawed specimen was used to characterize the da_c/dN versus ΔK behavior. By observing the transition from stable crack growth to rapid unstable crack growth behavior, a conservative estimate of the critical crack length at failure can be determined to predict the FCG life. Finally, by combining the fatigue strain-life data from previous work by Delgado et al. [1] with the current FCG data, a disk bolt hole eddy-current inspection interval can be calculated.

In determining the critical crack length to failure, a comparison of K_{IC} values was first made between the experiment and literature. Data reported by Kaufman [27] indicate that the plane-strain fracture toughness, K_{IC}, for GXMM247 is greater than 55 MPa$\sqrt{m}$ at 760°C, Fig. 21. The maximum ΔK value of 40.83 MPa$\sqrt{m}$, attained at the end of the FCG tests, can be used as a conservative estimate of K_{IC}. This is reasonable since the FCG behavior was observed to be well-behaved at the end of the test (i.e., stable crack growth rate). Assuming K_{IC} to be approximately 40 MPa$\sqrt{m}$, a_{crit} is iteratively calculated [28] using an initial crack size of 0.381 mm (the threshold level for eddy-current inspection [29]) and that the initial crack length, a_i, occurs at Cycle 1.

1. Calculate ΔK based on $\Delta K = \hat{F}\Delta\sigma\sqrt{\pi a_i}$. $\Delta\sigma$ is 699 MPa from analyses by Tong and Steinetz [30]. $\hat{F}$ is 1.12 for a half-elliptical surface crack [31].
2. Calculate da_c/dN per Eq. (1) using the values for $C(6.34 \times 10^{-12})$ and m (5.28) found for the combined data set (Fig. 14).
3. Recalculate the crack length for Cycle 2 by adding the previous crack length, a_i, to the value da_c/dN calculated in Step 2. The estimated crack growth length for the following cycle is da_c/dN.
4. Repeat Steps 1–3 until ΔK is approximately 40 MPa$\sqrt{m}$.

The resultant critical crack length is 0.831 mm at 734 cycles for $\Delta K = 40$ MPa$\sqrt{m}$ (Fig. 22). An average crack growth of 0.016 mm per 50 cycles is calculated from Cycle 0 to approximately Cycle 350. Recall that an arbitrary safety factor of 2 on the cyclic FCG life was used since FCG results indicate that LEFM conditions were not met and gross yielding was predicted at the gauge section. Thus, the calculated cyclic life of 734 cycles is reduced, by half, to 367 cycles. Also, the corresponding crack depth is reduced from 0.831 mm to a crack depth of 0.501 mm.

The total fatigue life of a material often encompasses approximately 90% of the fatigue strain-life and 10% of the FCG life [32]. Given the $R_\varepsilon = 0$ fatigue strain-life crack initiation curve at 649°C at the design strain of 0.5% from Delgado et al. [1], the mean life to crack initiation is approximately 15,000 cycles. Combining this with the predicted fatigue crack propagation cycles to failure, 734, in Fig. 22, gives a total life of 15,734 cycles. This approach suggests that, for the GXMM247 NASA disk operating at a 649°C metal temperature, crack initiation represents 95% of the cyclic life of the NASA disk while the remaining 5% is used to propagate the crack to failure.

However, considering the environment in which the NASA disk is used and the safety required for personnel and equipment, statistical and other safety factors must be used to estimate the in-

Table 5 Comparison of experimental GXMM247 test parameters with reported input parameters from available literature [23–26]

	This study	Macha et al.	MacIntyre and Agarwal	Alloy digest	Helmink
Specimen type	K_b	$C-T$	Not given	Not given	Not given
Temperature (°C)	649	426	538	538	538
Frequency (Hz)	0.33	30	0.33	Not given	Not given
Load ratio	0.05	0.1	0.05	0.05	Not given

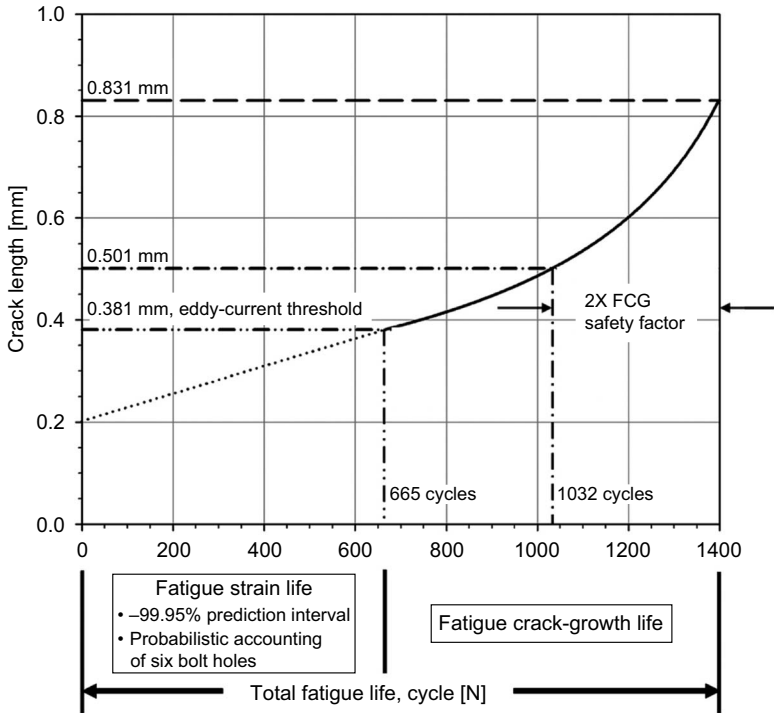


Fig. 23 Total fatigue life of GXMM247 at maximum operating conditions of the NASA turbine seal test rig

spection interval. Specifically, using the –99.95% prediction given by Delgado et al. [1], the cyclic life to crack initiation, with statistically one failure in 2000, is 1100 cycles at the design strain of 0.5% at 649°C. Accounting for the six bolt holes using the system life analysis, the resultant crack initiation life is 665 cycles [1]. Combining this with the predicted fatigue crack propagation result on a life of 367 cycles (using a factor of safety of 2) gives a total cyclic life for the GXMM247 NASA disk bolt holes of 1032 cycles at a crack length of 0.501 mm at maximum disk operating conditions. Since the eddy-current detection threshold is currently 0.381 mm, an initial NASA disk bolt hole inspection is recommended starting at approximately 665 cycles to detect crack initiation. Inspection intervals are then recommended approximately every 50 cycles thereafter to adequately monitor FCG. The NASA disk should be retired from high temperature service upon attaining either 1032 cycles or a crack depth of 0.501 mm (Fig. 23). One possible implementation plan is given in Fig. 24.

A limitation to this study was that there was a lack of sufficient material from which to make the FCG specimens; this largely contributed to the breakdown of the LEFM model. One way to examine the data is to use an elastic-plastic fracture mechanics approach. However, such an analysis was beyond the scope of this study. Also, it is acknowledged that short crack growth behavior is characterized by higher crack growth rates when compared to long crack growth behavior. Also, grain size, crystallographic orientation, and grain boundaries influence short crack growth behavior. However, modeling the short-crack growth behavior of GXMM247 was beyond the scope of this effort.

In light of these limitations, and to summarize, a number of conservative measures were taken to preclude disk failure in the test facility.

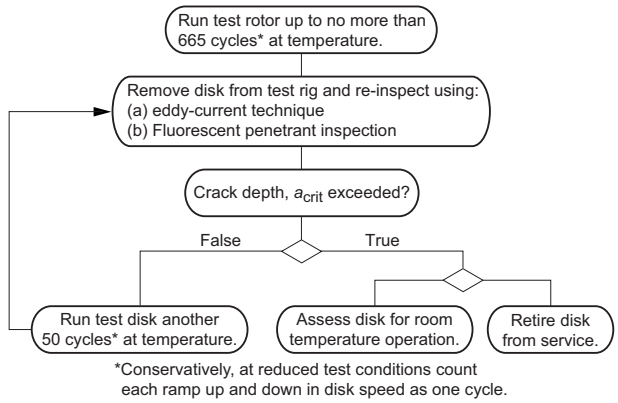


Fig. 24 Decision flow chart for implementing the eddy-current inspection interval for the GXMM247 NASA disk

1. Two FCG specimen data sets were statistically combined to provide a better characterization of the FCG behavior of GXMM247.
2. A conservative value of 40 MPa$\sqrt{m}$ was chosen for the critical crack growth rate. A comparison to available literature (Fig. 21) shows that the actual value is likely to be greater than 40 MPa$\sqrt{m}$.
3. A safety factor of 2 was placed on the cycles to failure for the disk due to the breakdown of LEFM.
4. The disk will be inspected after approximately 665 cycles based on previous strain-life analyses by Delgado et al. [1].

Summary

An eddy current inspection interval is proposed for detecting cracks in the bolt hole surfaces of the GXMM247 disk used in NASA's high temperature high-speed turbine seal test rig. Cracks may initiate and grow in the bolt holes due to cyclic stresses from seal tests at high temperature, pressure, and surface speeds. The inspection interval is based on previous fatigue strain-life data and current fatigue crack growth data resulting from tests using specimens fabricated from a sacrificial GXMM247 disk. Since not enough material was available for a proper K_{IC} test, surface-

flawed specimens were fabricated and tested to characterize the da_c/dN versus ΔK behavior of the material at maximum rig operating conditions. Data sets from successful fatigue crack growth tests were fitted to a Paris type relationship and then statistically combined.

Small crack growth behavior was observed by minima in the fatigue crack growth rate and is likely due to microstructural barriers such as grain boundaries. Also, microstructurally-sensitive behavior is likely because the final plane stress plastic zone size was predicted to be smaller than the average grain size. However, the fatigue crack growth behavior transitioned toward long crack growth behavior at higher ΔK. Also, gross yielding at the fracture surfaces indicated a more severe da_c/dN versus ΔK behavior than given by the data. Thus, LEFM criteria were not met and an arbitrary factor of 2 safety was used for determining the fatigue crack growth life. To predict the critical cycles to failure a conservative estimate of 40 MPa$\sqrt{m}$ was obtained for the critical crack growth rate after comparing the data to literature.

Using an iterative approach, critical fatigue crack propagation is predicted at 367 cycles at a crack depth of 0.501 mm using a safety factor of 2 on life. Combining this with previous fatigue strain-life experimental work gives a critical crack life of 1032 cycles at a crack length of 0.501 mm at maximum operating conditions. Eddy-current inspections are suggested starting at 665 cycles since eddy current thresholds are currently at 0.381 mm. Inspection intervals are recommended every 50 cycles at maximum operating conditions. Improvements on the fatigue crack growth life can be made using valid K_{IC} tests.

Acknowledgment

The authors wish to acknowledge Dr. Jack Telesman (NASA) and Mr. Pete Kantzos (Ohio Aerospace Institute) for their technical and experimental support.

The authors also wish to acknowledge the NASA Glenn Fatigue and Fracture Laboratories, the Army Research Laboratory and Case Western Reserve University for their support.

Nomenclature

C = fatigue crack growth coefficient (mm/cycle/(MPa$\sqrt{m}$)^m))
E = Young's modulus
$\hat{F}$ = dimensionless geometry factor
F^* = test statistic for the linear test
FCG = fatigue crack growth
F_s = boundary correction factor
GXMM247 = Grainex Mar-M 247
K_I = stress-intensity factor (Mode I) (MPa$\sqrt{m}$)
K_{IC} = Mode I critical fracture toughness (MPa$\sqrt{m}$)
$K_{\max}$ = maximum stress intensity factor (MPa$\sqrt{m}$)
L_p = potential probe spacing half-width (mm)
P or $P_{\max}$ = fatigue crack growth load (kN)
Q = shape factor
R = stress ratio
R_ε = strain ratio, $=\varepsilon_{\min}/\varepsilon_{\max}$
SSE(F) = sum squared error of the full model
SSE(R) = sum squared error of the reduced model
V = crack length voltage (V)
V = crack length voltage (V)
V_N = reference voltage (V)
X_{ij} = statistical model independent variable
Y_{ij} = statistical model predicted value
a_c = crack depth (mm)
a_i = iterative crack length (mm)
a_n = initial EDM crack depth (mm)
b_c = crack half-height (mm)
b_n = initial crack half-height (mm)
c_c = crack surface half-width (mm)
c_n = initial crack surface half-width (mm)
da_c/dN = fatigue crack growth rate (mm/cycle)
h = half-height of the K_b specimen gauge section from the EDM notch (mm)
m = fatigue crack growth exponent
n_1 = number of data points for data set 1
n_2 = number of data points for data set 2
s = remote uniform tensile stress (MPa)
t = through thickness of test specimen (mm)
ΔK = stress intensity range (MPa$\sqrt{m}$)
ΔK_{th} = threshold stress intensity range (MPa$\sqrt{m}$)
$\Delta\sigma$ = stress range (MPa)
β_0 = intercept of reduced model
β_1 = slope of reduced model
ε_{ij} = statistical model error term
ϕ = ellipse parametric angle [radians]

References

[1] Delgado, I. R., Halford, G. R., Steinetz, B. M., and Rimnac, C. M., 2005, "Strain-Life Assessment of Grainex Mar-M 247 for NASA's Turbine Seal Test Facility," ASME J. Eng. Gas Turbines Power, **127**, pp. 615–620.
[2] Coles, A., Johnson, R. E., and Popp, H. G., 1976, "Utility of Surface-Flawed Tensile Bars in Cyclic Life Studies," ASME J. Eng. Mater. Technol., **98**, pp. 305–315.
[3] Diez, A., 2000, "Crack Formation in Bolt Holes," private communication.
[4] Hartman, G. A., and Ashbaugh, N. E., 1990, "A Fracture Mechanics Test Automation System for a Basic Research Laboratory," *Applications of Automation Technology to Fatigue and Fracture Testing*, ASTM STP 1092, A. A. Braun, N. E. Ashbaugh, and F. M. Smith, eds., ASTM, Philadelphia, PA, pp. 95–110.
[5] Ashbaugh, N., and Hartman, G., 1990, "Direct Current Electric Potential Workshop," MATE Workshop, University of Dayton Research Institute.
[6] Gangloff, R. P., Slavik, D. C., Piascik, R. S., and Van Stone, R. H., 1992, "Direct Current Electrical Potential Measurement of the Growth of Small Cracks," *Small-Crack Test Methods*, ASTM STP 1149, J. M. Larsen and J. E. Allison, eds., ASTM, Philadelphia, PA, pp. 116–168.
[7] Hartman, G., 2002, "Questions on MATE Program," private communication.
[8] Gangloff, R. P., 1981, "Electric Potential Monitoring of Crack Formation and Subcritical Growth From Small Defects," Fatigue Fract. Eng. Mater. Struct., **4**(1), pp. 15–33.
[9] VanStone, R. H., and Richardson, T. L., 1985, "Potential-Drop Monitoring of Cracks in Surface-Flawed Specimens," *Automated Test Methods for Fracture and Fatigue Crack Growth*, ASTM STP 877, W. H. Cullen, R. W. Landgraf, L. R. Kaisand, and J. H. Underwood, eds., ASTM, Philadelphia, PA, pp. 148–166.
[10] Shannon, B., 2000, "Accounting for Thermoelectric Influences," private communication.
[11] Shannon, B., 2003, "FCG Data Acquisition Rate," private communication.
[12] Telesman, J., 2002, "Crack Growth Limitations," private communication.
[13] Newman, J. C., Jr., and Raju, I. S., 1983, "Stress-Intensity Factor Equations for Cracks in Three-Dimensional Finite Bodies," *Fracture Mechanics: Fourteenth Symposium—Volume I: Theory and Analysis*, ASTM STP 791, J. C. Lewis and G. Sines, eds., ASTM, Philadelphia, PA, pp. I-238–I-265.
[14] ASTM Committee E-8 on Fatigue and Fracture, 1998, "ASTM E-647–95a Standard Test Method for Measurement of Fatigue Crack Growth Rates," *Annual Book of ASTM Standards 2001*, ASTM, West Conshohocken, Vol. 3, pp. 577–613.
[15] Dowling, N. E., 1999, *Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue*, Prentice-Hall, Upper Saddle River, NJ.
[16] Neter, J., and Wasserman, W., 1974, *Applied Linear Statistical Models*, Richard D. Irwin, Homewood.
[17] Howmet Turbine Components Corporation, TB 3000: Grainex Cast Mar-M 247 Alloy, Howmet Turbine Components Corporation, Greenwich.
[18] Taylor, D., 1989, *Fatigue Thresholds*, Butterworths, London.
[19] Bates, R. C., and Clark, W. G., Jr., 1969, "Fractography and Fracture Mechanics," ASM Trans. Q., **62**(2), pp. 380–389.
[20] Suresh, S., and Ritchie, R. O., 1984, "Propagation of Short Fatigue Cracks," Int. Met. Rev., **29**(6), pp. 445–476.
[21] Taylor, D., 1985, *A Compendium of Fatigue Thresholds and Growth Rates*, Chameleon, London.
[22] Hertzberg, R. W., 1989, *Deformation and Fracture Mechanics of Engineering Materials*, 3rd ed., Wiley, New York.
[23] Macha, D. E., Cole, G. R., and Butzer, J. A., 1983, "Fine Grain, Investment-Cast Integral Turbine Wheels," *Grain Refinement in Castings and Welds*, G. J. Abbaschian and S. A. David, eds., The Metallurgical Society of AIME, New York, pp. 197–219.
[24] MacIntyre, C. A., and Agarwal, P. N., 1984, "Development of Fine Grain Cast Mar-M 247 Axial and Radial Turbine Wheels," *Advanced Aerospace Materials*

Technology, SP-597, Society of Automotive Engineers, Inc., Warrendale, pp. 35–45.
[25] Helmink, R. C., Testin, R. A., Price, A. R., Pachman, R., Erickson, G. L., Harris, K., Nesbitt, J. A., and Radavich, J. F., 2000, "Advanced Superalloys and Tailored Microstructures for Integrally Cast Turbine Wheels," *Superalloys 2000*, T. M. Pollock et al., eds., TMS, Warrendale, PA, pp. 171–179.
[26] 1995, "Microcast-X Mar-M 247," Alloy Digest.
[27] Kaufman, M., 1984, "Properties of Cast Mar-M-247 for Turbine Blisk Applications," *Superalloys 1984*, M. Gell, C. S. Kortovich, R. H. Bricknell, W. B. Kent, and J. F. Radavich, eds., Metallurgical Society of AIME, Warrendale, PA, pp. 43–52.
[28] Broek, D., 1984, *Elementary Engineering Fracture Mechanics*, 3rd ed., Kluwer Boston, Hingham.
[29] Curtis Industries, Inc., 2003, "Reference Standard Calibration for GX Mar-M 247," inspection report for reference standard, Curtis Industries, Inc., Pennsylvania.
[30] Tong, M., and Steinetz, B., 1997, "Mar-M 247 Disk Finite Element Thermal and Stress Analysis," NASA Glenn Research Center, Cleveland, OH.
[31] Dowling, N. E., 1999, *Mechanical Behavior of Materials: Engineering Methods for Deformation, Fracture, and Fatigue*, Prentice-Hall, Upper Saddle River, NJ.
[32] Kerlins, V., and Phillips, A., 1987, "Modes of Fracture," *Metals Handbook*, Vol. 12, 9th ed., ASM International, Metals Park.

Forced Response Prediction of Constrained and Unconstrained Structures Coupled Through Frictional Contacts

Ender Cigeroglu[1]
e-mail: ender@metu.edu.tr

Ning An

Chia-Hsiang Menq

Department of Mechanical Engineering,
The Ohio State University,
Columbus, OH 43210

In this paper, a forced response prediction method for the analysis of constrained and unconstrained structures coupled through frictional contacts is presented. This type of frictional contact problem arises in vibration damping of turbine blades, in which dampers and blades constitute the unconstrained and constrained structures, respectively. The model of the unconstrained/free structure includes six rigid body modes and several elastic modes, the number of which depends on the excitation frequency. In other words, the motion of the free structure is not artificially constrained. When modeling the contact surfaces between the constrained and free structure, discrete contact points along with contact stiffnesses are distributed on the friction interfaces. At each contact point, contact stiffness is determined and employed in order to take into account the effects of higher frequency modes that are omitted in the dynamic analysis. Depending on the normal force acting on the contact interfaces, quasistatic contact analysis is initially employed to determine the contact area as well as the initial preload or gap at each contact point due to the normal load. A friction model is employed to determine the three-dimensional nonlinear contact forces, and the relationship between the contact forces and the relative motion is utilized by the harmonic balance method. As the relative motion is expressed as a modal superposition, the unknown variables, and thus the resulting nonlinear algebraic equations in the harmonic balance method, are in proportion to the number of modes employed. Therefore the number of contact points used is irrelevant. The developed method is applied to a bladed-disk system with wedge dampers where the dampers constitute the unconstrained structure, and the effects of normal load on the rigid body motion of the damper are investigated. It is shown that the effect of rotational motion is significant, particularly for the in-phase vibration modes. Moreover, the effect of partial slip in the forced response analysis and the effect of the number of harmonics employed by the harmonic balance method are examined. Finally, the prediction for a test case is compared with the test data to verify the developed method.
[DOI: 10.1115/1.2940356]

1 Introduction

Mechanical systems with moving components always involve frictional contact, which appears in various applications such as turbine blades [1–10], mechanical joints [11–13], and clutches [14–16]. Due to the nonlinear nature of dry friction, dynamic analysis of structures constrained through frictional contacts is difficult. Several methods were developed in order to analyze these structures [1–16]; however, due to this difficulty, all of these methods were developed for specific cases such as shroud contact [9,10], bladed disks with wedge dampers [1–3,8], mechanical joints [11], and clutches [14,16]. Therefore, there is a need for a general approach for the analysis of structures constrained through frictional contacts.

In order to develop a general analysis method, one of the structures in the frictional contact is considered as unconstrained; therefore, it is constrained by the frictional contact and/or the geometric configuration only. This is a general case, which can also handle friction contact between constrained structures. A typical example of frictional contact between constrained and unconstrained structures can be found in bladed-disk systems with blade-to-blade dampers [1–8], where the dampers move freely in between adjacent blades. In order to establish the developed forced response analysis method, in this study, a bladed-disk system with wedge dampers is used as an illustration in the rest of the paper. Here, blade and wedge damper constitute the constrained and unconstrained structures, respectively.

Wedge damper, unconstrained structure, is a widely used friction damper, which is also referred to as underplatform damper. It has two inclined surfaces on both sides and forced against the two neighboring blades by centrifugal forces. Due to the relative motion between the wedge damper and adjacent blades, the excessive energy of the blades is dissipated through frictional contacts. In order to increase the high cycle fatigue (HCF) life of turbine blades, optimal parameter values for the wedge damper and the bladed-disk system have to be determined, which can be achieved through a forced response analysis.

Yang and Menq [1,2] developed stick-slip contact kinematics for wedge dampers under two translational degrees of freedom. Authors developed analytical stick-slip transition criteria, including the variation of normal load, in order to simulate the stick-slip motion precisely. The harmonic balance method was used to predict forced response of bladed disk with wedge dampers, and an experimental test beam was analyzed and the simulation results

[1]Corresponding author. Present address: Middle East Technical University, Ankara, Turkey.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received December 24, 2007; final manuscript received March 17, 2008; published online December 23, 2008. Review conducted by Patrick S. Keogh. Paper presented at the ASME Turbo Expo 2007: Land, Sea and Air (GT2007), Montreal, Canada, May 14–17, 2007.

were validated. A 3D wedge damper model with two-dimensional motion on the contact interface was developed by Sanliturk et al. [3]. A two-dimensional friction model for constant normal load was used, and the harmonic balance method is applied to predict forced response. A test case with two blades and a wedge damper was analyzed, and the results were compared with the simulations. In addition to wedge dampers, curved shape underplatform dampers were also studied by several researchers [4–7], and an underplatform damper with a curved and inclined surface was analyzed in Ref. [8]. All the models described above develop kinematic relations for the damper and blade, and the rigid body motion of the unconstrained structure is not modeled.

In order to determine nonlinear contact forces between two relatively moving bodies, different friction models can be utilized. In the analysis of friction damping, one-dimensional friction model was used widely [17–21]. This model is useful if the relative motion is one dimensional. It is possible to have planar motion, for which two-dimensional friction models are developed [22,23]. However, due to the interaction between two bodies, the normal load acting on contacting surfaces can vary with normal motion. Yang et al. [24] developed a one-dimensional friction model where the normal load was induced by the normal motions of the mating surfaces. Authors developed analytical transition criteria for stick-slip-separation transition and obtained analytical transition angles for simple harmonic motion. Normal load variation was also addressed by Yang and Menq [25] for three-dimensional motion (two-dimensional in-plane and one-dimensional out-of-plane motion) and by Chen and Menq [10] for three-dimensional periodic motion. Using similar criteria as developed in Ref. [24], Petrov and Ewins [26], for one-dimensional motion with normal load variation, later described an algorithm to determine transition angles numerically for periodic motion, similar to the one given in Ref. [27].

This paper presents a forced response prediction method for the analysis of constrained and unconstrained structures coupled through frictional contacts. The proposed model includes six rigid body modes and several elastic modes of the unconstrained structure; therefore, it may undergo three-dimensional translation and three-dimensional rotation, which are constrained by the friction contacts only. In the modeling of contact surfaces, discrete contact points associated with contact stiffnesses are distributed on friction interfaces. Contact stiffnesses at each contact point are determined by considering the effects of higher frequency modes, which are omitted in the dynamic analysis. The initial preload or gap at each contact pair varies with the normal force acting on the friction interface; therefore, a quasistatic contact analysis is performed initially in order to determine the contact area in addition to the initial preload or gap at each contact point due to the normal load.

In order to predict forced response of frictionally constrained structures, two-dimensional or three-dimensional friction models developed previously can be employed. However, in order to decrease computational time required for the forced response predictions, a simplified three-dimensional friction contact model based on the one-dimensional model with normal load variation [24] is proposed. In the proposed friction model, the three-dimensional relative motion on the contact surface is decomposed into two one-dimensional in-plane components and an out-of-plane component. The one-dimensional friction contact model is employed by assuming that these in-plane components are independent of each other. The harmonic balance method is used to represent the resulting nonlinear contact forces ensuing in a set of nonlinear algebraic equations. The relative motion at the contact surface is expressed by modal superposition; therefore, the number of unknowns resulting in the nonlinear equation set is only proportional to the number of modes used in the analysis. As a result, unlike receptance methods, the number of nonlinear equations is independent of the number of contact points used.

The developed method is applied to a tuned bladed-disk system with wedge dampers in order to obtain its forced response and optimal curves. In addition, the effects of normal load on the rigid body motion of the damper are studied. Specifically, the effect of the damper's rotational motion on the prediction of the forced response is analyzed. It is shown that the effect of rotational motion is significant, particularly for the in-phase vibration modes. The effect of partial slip in the forced response analysis is investigated. A blade-to-ground damper is studied in order to reveal the effects of multiple harmonics on forced response predictions. Finally, predictions for a test case are compared with available test data.

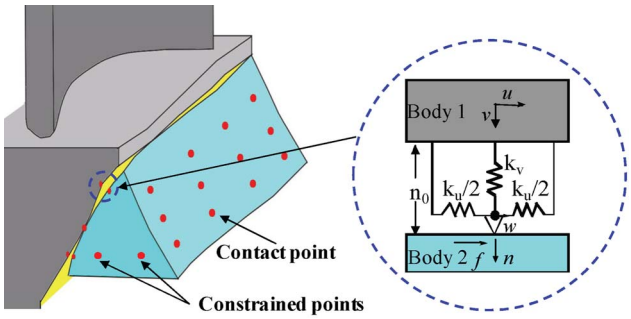


Fig. 1 Contact model for the wedge damper

2 Model for Constrained and Unconstrained Structure

The forced response prediction method for constrained and unconstrained structures coupled through frictional contacts is presented on a bladed-disk system with wedge dampers, where the blade constitutes the constrained structure, and the damper constitutes the unconstrained structure. Accordingly, the motion of the damper is constrained by the frictional contacts and the geometric configuration of the damper between adjacent blades only. Consequently, the damper undergoes three-dimensional translation and three-dimensional rotation in addition to the elastic deformation. Elastic motion of the damper is necessary if the excitation frequency and/or static forces acting on the damper are high. However, with the proposed approach, it is possible to model the damper as completely rigid or rigid in certain directions and elastic in others by using the appropriate mode shapes.

2.1 Contact Model. The interaction between the blade and the wedge damper is modeled by discrete contact points evenly distributed on the two contact surfaces of the blade and damper. At each contact pair, contact stiffnesses in the three main directions of motion are determined in order to take into account the effects of higher frequency modes, which can be represented as residual stiffnesses. It is assumed that residual stiffnesses are only present between contact pairs; hence, they are called as contact stiffnesses. The determination of contact stiffnesses will be explained in the next section.

A blade and a wedge damper in contact are given in Fig. 1, where contact points represented by the dots are shown on the left, and contact stiffnesses between a contact pair in the local coordinate system are shown on the right. The contact points on the *X-Y* plane of the damper are called constrained points, which are used to constrain the motion of the damper in the *Z* direction due to space limitations, and it should be noted that the constrained force could be at most on one of the constrained planes. These constraints are due to the physical restrictions in real gas turbine engine, where the damper can move freely under the action of contact and centrifugal forces in a volume between the adjacent blades on the disk.

The motion of the blade and the wedge damper are expressed in the blade coordinate and damper coordinate systems, respectively. The coordinate systems for the *i*th blade and wedge damper are shown in Fig. 2. The blade coordinate system is on the rotary axis

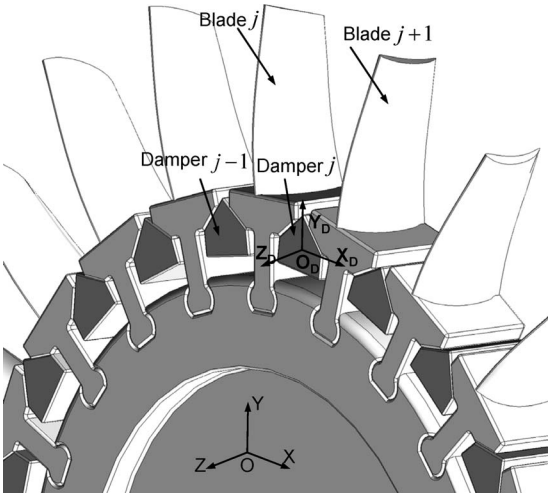


Fig. 2 Blade and damper coordinate systems

of the disk where the X and Y axes are coincident with the tangential and radial orientations of the ith blade, and the Z axis is determined by the right hand rule. The ith damper coordinate system is determined by three rotations about the blade coordinate axes. Coordinate systems for other blades and dampers can be obtained by a simple rotation about the Z axis with an amount of blade phase angle.

As shown in Fig. 3, wedge damper has four contact planes, where α and β are used to define the orientations of the left and right contact planes. Coordinate systems for the front and the back constraint planes can be obtained by 90° and −90° rotations about the X_D axis, and the coordinate systems for the right and the left contact planes can be obtained by $-(90°-\beta)$ and $90°-\alpha$ rotations about the Z_D axis. In Fig. 3 the coordinate system for the right plane is shown.

2.2 Calculation of Contact Stiffness. In the model proposed, in order to capture local deformation on the contact interface, very high frequency modes have to be included into the modal superposition approach, which is not practical in terms of calculation times. Therefore limited number of modes is used in the modal expansion process. On the other hand, higher vibration modes behave like springs at lower excitation frequencies; therefore, these omitted higher modes of the bladed-disk system can be represented by contact stiffnesses, which makes it possible to capture local deformations on the contact interface.

In the bladed-disk system with wedge dampers, dampers are initially not in contact with the blades. When the engine starts rotation due to the centrifugal effects, dampers come into contact with the neighboring blades. In order to include this effect, a normal force is assumed to act on the damper which presses it against the adjacent blades. Therefore, for a static contact analysis, where the only force acting on the bladed-disk system is the normal force, the displacement at any contact point on the blade from modal analysis can be calculated by

$$\tilde{U}_B^i = -\sum_{j=1}^{n_C} \mathbf{R}_B^{i,j\,B} F_j \tag{1}$$

$\mathbf{R}_B$ is the blade receptance matrix between contact points, ${}^B F$ is the finite element contact force vector on the damper in the blade coordinate system, and n_C is the number of contact points. The difference in blade displacements with the finite element results can be expressed in the contact plane coordinate system as

$$\Delta U_B^i = {}^C U_B^i - {}^C \tilde{U}_B^i = {}^C U_B^i + \sum_{j=1}^{n_C} {}^C \mathbf{R}_B^{i,j\,C} F_j \tag{2}$$

where ΔU_B is the vector of difference in blade displacements, and ${}^C \tilde{U}_B$ and ${}^C U_B$ are the vectors of displacements from modal superposition and finite element contact analysis considering slip and separation at the contact pairs, respectively. Superscript C on the left of each parameter denotes the contact plane coordinate system. Blade residual stiffnesses at contact points are defined as

$$\mathbf{K}_{cB}^i \Delta U_B^i = -{}^C F_i \tag{3}$$

Substituting Eq. (2) in Eq. (3), the following relation is obtained:

$$\mathbf{K}_{cB}^{i\,-1\,C} F_i = -\left(\sum_{j=1}^{n_C} {}^C \mathbf{R}_B^{i,j\,C} F_j + {}^C U_B^i \right) \tag{4}$$

from which residual stiffnesses for the blade can be calculated.

The residual stiffness equation for the damper can be obtained similarly; however, the rigid body motion of the damper has to be considered. The displacements of contact points on the damper can be given in two parts,

$$\tilde{U}_D^i = \tilde{U}_{DR}^i + \tilde{U}_{DE}^i \tag{5}$$

where $\tilde{U}_{DR}$ and $\tilde{U}_{DE}$ are the vectors of rigid body and elastic damper displacements from modal analysis, and they are given as

$$\tilde{U}_{DR}^i = \mathbf{\Psi}_i \boldsymbol{\eta} \tag{6}$$

$$\tilde{U}_{DE}^i = \sum_{j=1}^{n_C} \mathbf{R}_{DC}^{i,j\,D} F_j + \sum_{j=1}^{n_P} \mathbf{R}_{DP}^{i,j} P_j \tag{7}$$

where $\mathbf{\Psi}_i$ is the rigid body mode shape matrix for the ith contact point, $\boldsymbol{\eta}$ is the rigid body modal coefficient vector, and $\mathbf{R}_{DC}$ and $\mathbf{R}_{DP}$ are the receptance matrices between contact points and preload points, respectively. ${}^D F$ and P are the vectors of contact forces and preloads acting on the damper in the damper coordinate system. Using Eqs. (6) and (7), the difference between modal superposition and finite element displacements in the contact plane coordinate system is obtained as

$$\Delta U_D^i = {}^C U_D^i - {}^C \tilde{U}_D^i = {}^C U_D^i - \left({}^C \mathbf{\Psi}_i \boldsymbol{\eta} + \sum_{j=1}^{n_C} {}^C \mathbf{R}_{DC}^{i,j\,C} F_j + \sum_{j=1}^{n_P} {}^C \mathbf{R}_{DP}^{i,j} P_j \right) \tag{8}$$

The damper residual stiffness is defined as

$$\mathbf{K}_{cD}^i \Delta U_D^i = {}^C F_i \tag{9}$$

from which the following relation is obtained:

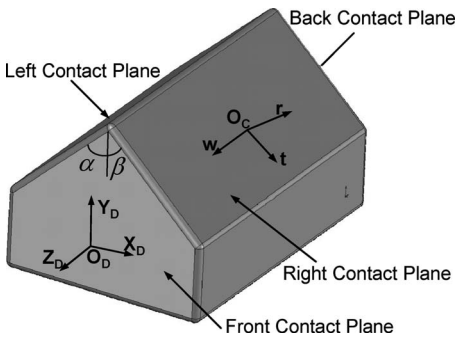


Fig. 3 Wedge damper contact planes and coordinate systems

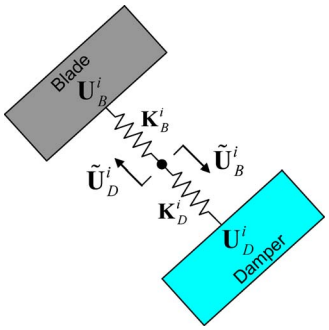


Fig. 4 Schematic view for the bounded configuration

$$\mathbf{K}_{cD}^{i\,-1}{}^{C}F_i = {}^{C}U_D^i - \left({}^{C}\boldsymbol{\Psi}_i \boldsymbol{\eta} + \sum_{j=1}^{n_C} {}^{C}\mathbf{R}_{DC}^{i,j}\,{}^{C}F_j + \sum_{j=1}^{n_P} {}^{C}\mathbf{R}_{DP}^{i,j} P_j \right) \tag{10}$$

Assuming that the blade and damper contact pairs are bounded together, as shown in Fig. 4, the residual stiffness matrix for the *i*th contact pair can be written as follows:

$$\mathbf{K}_c^{i\,-1} = \mathbf{K}_{cB}^{i\,-1} + \mathbf{K}_{cD}^{i\,-1} \tag{11}$$

where $\mathbf{K}_c^{i-1}$ is the residual stiffness matrix at the *i*th contact pair, which is referred to as contact stiffness matrix. It should be noted that in three-dimensional space there are nine unknowns in the residual stiffness matrix given in Eq. (11); hence, Eqs. (4) and (10) cannot be solved uniquely to determine these unknowns. However, if the residual stiffness matrix is assumed to be diagonal only, Eqs. (4) and (10) can be solved uniquely for the three unknowns on their diagonals. It can be concluded that the stiffnesses on the diagonal are in the three main directions of the contact plane coordinate system between blade and damper contact pairs. Since the friction model utilized is applied in the two major tangential directions, contact stiffnesses used in the developed friction model are in the three main directions; therefore, the assumption of the diagonal contact stiffness matrix is an appropriate one. Moreover, most of the friction models developed for the analysis of bladed-disk systems utilize diagonal contact stiffness matrices where off-diagonal coupling terms between tangential directions are used in the friction model given in Refs. [10,25], but to the best of the authors' knowledge, no friction model uses coupling between tangential and normal directions. Therefore, the method presented can be applied to other friction models as well.

3 Friction Contact Model and Forced Response Prediction

The relative motion at the blade-damper contact interface is three dimensional, and this relative motion is decomposed into in-plane and out-of-plane (normal) components. Furthermore, two major directions for the in-plane component of motion are determined, and the in-plane motion is approximated in these directions. Calculation of Fourier coefficients using a three-dimensional friction model as given in Ref. [10,25] is time consuming, and the main focus of this work is modeling damper motion as an unconstrained body having six degrees of freedom rigid body motion plus elastic motion. Therefore, in order to speed up the forced response calculations, these major directions are assumed to be independent of each other, and the one-dimensional friction model with normal load variation developed by Yang et al. [24] is employed. Transition criteria and analytical transition angles for harmonic motion are given in Ref. [24], and for periodic motion these criteria can be solved numerically to determine the transition angles. Therefore, the nonlinear normal force and the friction forces in major directions are obtained, which are then expressed in the contact plane coordinate system.

3.1 Initial Preload on Contact Surfaces. Depending on the engine rpm, the centrifugal force acting on the wedge dampers varies. This results in variation in the contact area and the preload/gap acting on the contact surfaces. In order to determine the initial preload/gap, a quasistatic contact analysis is performed for the given normal load. The analysis is performed as follows:

1. Initially, the contact status of all the contact pairs are assumed completely stuck.
2. The displacements of contact points and the contact forces acting on them are determined using the given contact status.
3. Using the Coulomb friction model, the contact status of each contact pair is updated.
4. The contact status of each contact pair is compared with previous contact status.
5. If the contact status is changed, go to step 2.
6. Otherwise, output the initial preload.

This analysis is an important step for the forced response calculations since the change of contact area and the preload/gap can affect the entire forced response characteristics of the blade and damper system. In order to increase the robustness of the method presented, a continuation method can be employed where the normal load is modified with a load factor between 0 and 1. Starting with a very low value, where the entire contact interface is completely stuck or very close to completely stuck, and applying the procedure given above, the contact state for this load step can be determined. The contact state for an increased load factor can be obtained by repeating the procedure given above where the previous solution is used as an initial guess at step 1. For the preload considered, the contact state and normal load acting on the contact interface will be obtained when the load factor reaches 1.

3.2 Forced Response Method. In the forced response analysis, finite element models for the blade (disk) and the damper are employed. Using receptance methods, the number of unknowns in the forced response analysis method can be decreased to the number of nonlinear (contact points) degrees of freedom multiplied by the number of harmonics. However, if the number of contact points is high, which is necessary for an accurate modeling of friction contact, this method is not suitable for forced response analysis due to large matrices involved in the solution procedure. Recently, Cigeroglu et al. [27] proposed a modal superposition method for the forced response analysis of bladed-disk systems. In this approach, the relative motion between contact surfaces is approximated by modal superposition using free mode shapes of the structure. This method is extended for multiple harmonics and is employed in this work. In the modal superposition approach, the number of unknowns involved in the solution procedure is the number of mode shapes used in the modal expansion process multiplied by the number of harmonics; therefore, it is independent of the number of contact points used. As a result of this, the modal superposition approach is suitable for accurate modeling of friction contact with more contact points or for cases when the tuned approach (cyclic symmetry) cannot be used.

The equation of motion in matrix form for a system with dry friction dampers can be written in the following form:

$$\mathbf{M} \cdot \ddot{X} + \mathbf{C} \cdot \dot{X} + \mathbf{K} \cdot X = F_e(t) + F_n(X) \tag{12}$$

where $\mathbf{M}$, $\mathbf{C}$, $\mathbf{K}$, $F_e(t)$, $F_n(X)$, and X are the mass matrix, viscous damping matrix, stiffness matrix, excitation force vector, nonlinear force vector, and displacement vector, respectively. The motion of the blade and the damper for harmonic excitation can be written in terms of its mode shapes as follows:

$$\mathbf{u}_B^j = \sum_{n=1}^{n_B} \mathbf{A}_{n,0} \boldsymbol{\phi}_j^n + \sum_{l=1}^{n_H} \mathrm{Im}\left(\sum_{n=1}^{n_B} \mathbf{A}_{n,l} \boldsymbol{\phi}_j^n e^{i(l\theta)} \right)$$

$$\mathbf{u}_D^j = \sum_{n=1}^{n_R} \mathbf{C}_{n,0}\boldsymbol{\psi}_j^n + \sum_{n=1}^{n_E} \mathbf{D}_{n,0}\boldsymbol{\Phi}_j^n + \sum_{l=1}^{n_H} \mathrm{Im}\left[\left(\sum_{n=1}^{n_R} \mathbf{C}_{n,l}\boldsymbol{\psi}_j^n + \sum_{n=1}^{n_E} \mathbf{D}_{n,l}\boldsymbol{\Phi}_j^n\right) e^{i(l\theta)}\right] \quad (13)$$

where $\boldsymbol{\phi}_j^n$, $\boldsymbol{\psi}_j^n$, and $\boldsymbol{\Phi}_j^n$ are the nth mode shape vector for the jth blade node, nth rigid body mode shape, and elastic mode shape for the jth damper node, respectively. A_n^l, C_n^l, and D_n^l are the lth harmonic modal coefficients for the blade, damper rigid body modes, and elastic modes, respectively, where $l=0$ defines the bias terms. n_B, n_R, n_E, and n_H are the number of blade modes, damper rigid body modes, damper elastic modes, and harmonics, respectively. In addition to this, i is the imaginary unit and θ is the temporal variable. The relative motion between the jth contact pair can be written in contact plane coordinates as

$$\Delta X_j = {}^C\mathbf{u}_B^j - {}^C\mathbf{u}_D^j = {}_B^C\mathbf{T}_j \cdot \mathbf{u}_B^j - {}_D^C\mathbf{T}_j \cdot \mathbf{u}_D^j \quad (14)$$

where ${}_B^C\mathbf{T}_j$ and ${}_D^C\mathbf{T}_j$ are the transfer matrices from the blade coordinate system to contact plane coordinates and from the damper coordinate system to contact plane coordinates for contact point j, respectively. Using the friction contact model and the relative motion given in Eq. (14), nonlinear contact forces can be determined in contact plane coordinates as

$${}^CF_n(\mathbf{A},\mathbf{C},\mathbf{D},\theta) \cong {}^CF_n^0(\mathbf{A},\mathbf{C},\mathbf{D}) + \sum_{l=1}^{n_H} {}^CF_{ns}^l(\mathbf{A},\mathbf{C},\mathbf{D})\sin(l\theta) + \sum_{l=1}^{n_H} {}^CF_{nc}^l(\mathbf{A},\mathbf{C},\mathbf{D})\cos(l\theta) \quad (15)$$

where ${}^CF_n^0$ is the bias term of vector of contact forces and, ${}^CF_{ns}^l$ and ${}^CF_{nc}^l$ are the vectors of sine and cosine components of the lth harmonic of contact forces, which are functions of modal coefficients $\mathbf{A}$, $\mathbf{C}$, and $\mathbf{D}$. Using the orthogonality of mode shapes and Eq. (15), Eq. (12) for a single sector can be written in the following form:

$$\boldsymbol{\Omega}_B\mathbf{A}^{\langle 0\rangle} = Q_{B_b}^0$$

$$(\boldsymbol{\Omega}_B - (l\omega)^2\mathbf{I} + i(l\omega)\mathbf{C_B})\mathbf{A}^{\langle l\rangle} = Q_e^l + Q_{B_{\mathrm{Re}}}^l + iQ_{B_{\mathrm{Im}}}^l$$

$$0 = {}_jQ_{D_b}^0 \quad (j = 1 \cdots n_R)$$

$$\boldsymbol{\Omega}_D^j\mathbf{D}_{j,0} = {}_jQ_{D_b}^0 \quad (j = n_R + 1 \cdots n_D = n_R + n_E)$$

$$(\boldsymbol{\Omega}_D - (l\omega)^2\mathbf{I} + i(l\omega)\mathbf{C_D})\mathbf{E}^{\langle l\rangle} = Q_{D_{\mathrm{Re}}}^l + iQ_{D_{\mathrm{Im}}}^l \quad (l = 1 \cdots n_H) \quad (16)$$

if mass normalized mode shapes are used. In Eq. (16), $\boldsymbol{\Omega}_B$ and $\boldsymbol{\Omega}_D$ are $n_B \times n_B$ and $n_D \times n_D$ diagonal matrices of squares of natural frequencies of the blade and the damper. $\mathbf{C_B}$ and $\mathbf{C_D}$ are modal damping matrices of the blade and the damper, and they are diagonal if the damping is proportional. $\mathbf{A}^{\langle l\rangle}$ is the vector of modal coefficients of the blade, and $\mathbf{E}^{\langle l\rangle}$ is the vector of modal coefficients of the damper for the lth harmonic, which is composed of rigid body and elastic mode shapes as $\mathbf{E}=(\mathbf{C} \quad \mathbf{D})^T$. Q_B^l and Q_D^l are the vectors of modal forces of the lth harmonic for the blade and damper, respectively, and e, b, Re, and Im stand for excitation force, bias component, real part, and imaginary part, respectively. It should be noted that the contact forces acting on the jth blade and the jth damper in contact plane coordinate system are the same in magnitude but opposite in signs. If the bladed-disk system is tuned due to the cyclic symmetry, the motion of the $(j+1)^{\mathrm{th}}$ blade can be related to the motion of the jth blade as

$$\mathbf{u}_B^{j+1}(\theta) = \mathbf{u}_B^j(\theta + \varphi), \quad \varphi = 2\pi n_{\mathrm{EO}}/n_b \quad (17)$$

where φ is the interblade phase angle, n_{EO} is the engine order, and n_b is the number of blades. Using this information, the relative displacement on the contact surfaces of the jth damper can be determined, from which contact forces acting on the jth damper can be obtained. Similar to the displacements, contact forces between the $(j-1)^{\mathrm{th}}$ damper and jth blade can be related as

$$F_n^j(\theta) = F_n^{j+1}(\theta - \varphi) \quad (18)$$

where $F_n{}^j$ is the nonlinear contact forces acting on the jth blade in the blade coordinate system. Transferring contact forces between the jth damper and the $(j-1)$th blade to the jth blade, the modal coefficients of the jth blade and damper can be determined by iteratively solving the nonlinear equation set given in Eq. (16); consequently, the blade and damper responses can be obtained from Eq. (13).

4 Numerical Results

Two different cases are analyzed in this section. In the first part, the method is applied to a tuned bladed-disk system, and forced response curves and optimal curves are presented. Moreover, the effects of normal load on the rigid body coefficients of the damper and the effects of rotational motion of the damper on the forced response results are studied. In order to show the effect of partial slip on the forced response analysis, the bladed-disk system is analyzed for different numbers of contact points. In the second part, a blade-to-ground damper system is analyzed in order to show the effects of multiharmonics on the forced response. In the results provided below, stuck case indicates that all the contact points that are initially in contact do not slip. Therefore, it is possible to have contact points, which are not in contact due to the applied normal load. These are identified in the quasistatic contact analysis.

4.1 Tuned Bladed-Disk System With Wedge Dampers. A tuned bladed-disk system with wedge dampers, composed of 65 blades and dampers, is studied in this part. Finite element models of a blade and a damper are given in Fig. 5. The point of excitation and the point where displacement is calculated are indicated by dots, as shown in the figure. In the analysis, fifth engine order excitation is considered, and ten modes for the blade and six rigid body modes for the damper are used. In the forced response calculations only the fundamental harmonic is kept due to reasons that will be explained in Sec. 4.2.

4.1.1 Forced Response and Optimal Curves. Tracking plots around the first and second modes of the blade are given in Figs. 6 and 7 for different preload cases. External force applied at the tip point is in the tangential direction for the first mode and in the radial direction for the second mode in order to excite those modes. It is observed that as the preload acting on the damper increases, the amplitude of tip point displacement decreases and the resonance frequency of the system increases. When the optimum point is reached, the amplitude is minimum, and increasing the preload further results in an increase in the vibration amplitude, which converges to the completely stuck magnitude. It is interesting to note that in Fig. 6 multiple solutions in the tracking plots exist for some preload cases, whereas in Fig. 7 none of the preload cases results in multiple solutions in the tracking plots. These multiple solutions are due to the separation of the contact surfaces. In Fig. 6, the first bending mode of the blade is excited, resulting in the separation of the contact surfaces due to the rotation of the wedge damper about the Z axis. However, in Fig. 7, since the dominant rotation is about the X axis, contact surfaces remain in contact throughout the analysis. Multiple solutions in the tracking plots are obtained by the continuation method, and they show a typical fold bifurcation, which has an unstable solution branch in between two stable solutions resulting in the jump

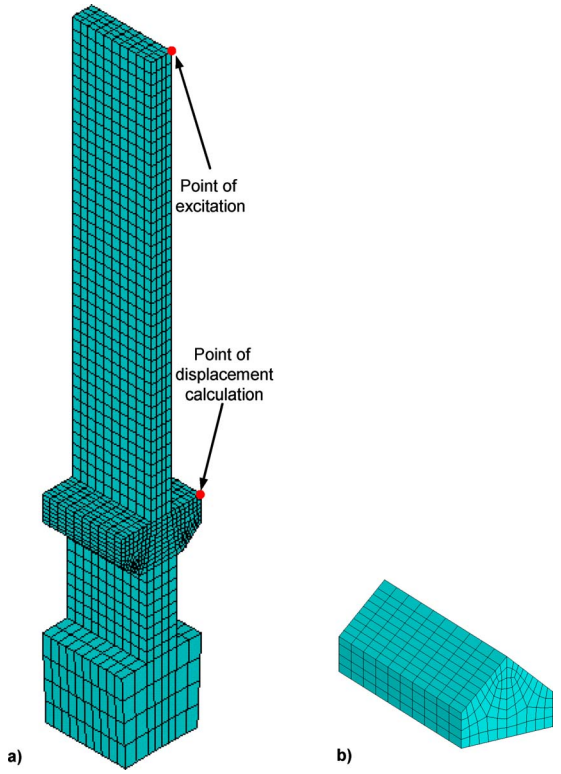


Fig. 5 Finite element models for (*a*) blade and (*b*) wedge dampers

phenomenon.

In Figs. 8 and 9, optimal curves and frequency shift curves for the first and second vibration modes of the blade are given. It is observed that the ratio of the stuck case amplitude to the optimal preload case amplitude is approximately 4.4 and 5.5 for the first and second modes of vibration, respectively. Due to the fact that there exists no separation between the wedge damper and the adjacent blades, damping at the second mode is more effective. However, since the decrease in the free response amplitude at the first mode is larger than the second mode, damper works more effectively at the first mode. The frequency shifts observed for the first and second modes are 17.2% and 3.4%, respectively.

Using the computer code (BDAMPER) developed for the analysis of bladed-disk systems with wedge dampers, higher modes of the

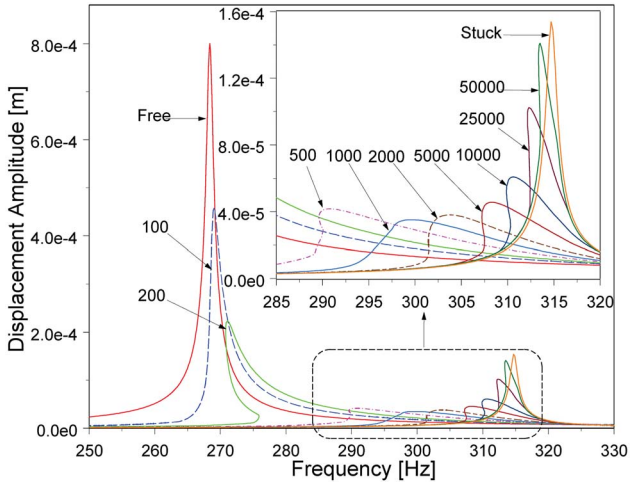


Fig. 6 Tracking plot for the first mode

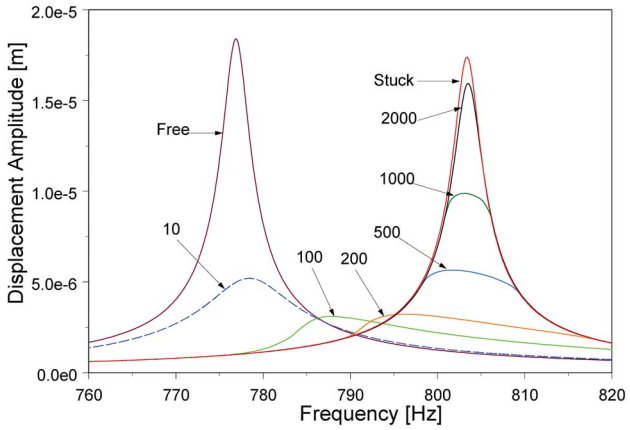


Fig. 7 Tracking plot for the second mode

blade can also be investigated. The tracking plots and optimum and frequency shift curves for the third and seventh modes of the blade are given in Figs. 10 and 11.

4.1.2 Effects of Normal Load and Excitation Frequency on Rigid Body Motion. In Figs. 12 and 13, maximum amplitudes of the bias and vibratory components of rigid body modal coefficients of the wedge damper are given for the first vibration mode, respectively. For low normal load cases, modal coefficients for translation along X and Y axes and for rotation about the Z axis

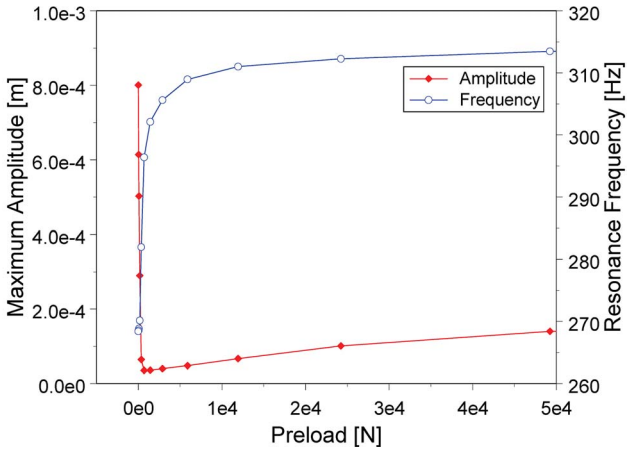


Fig. 8 Optimal and frequency shift curves: first mode

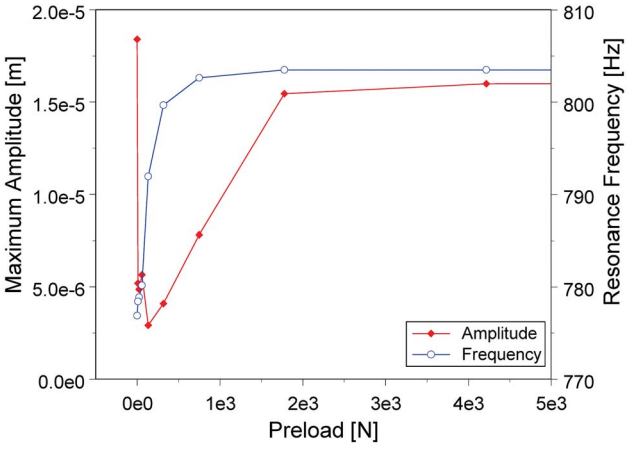


Fig. 9 Optimal and frequency shift curves: second mode

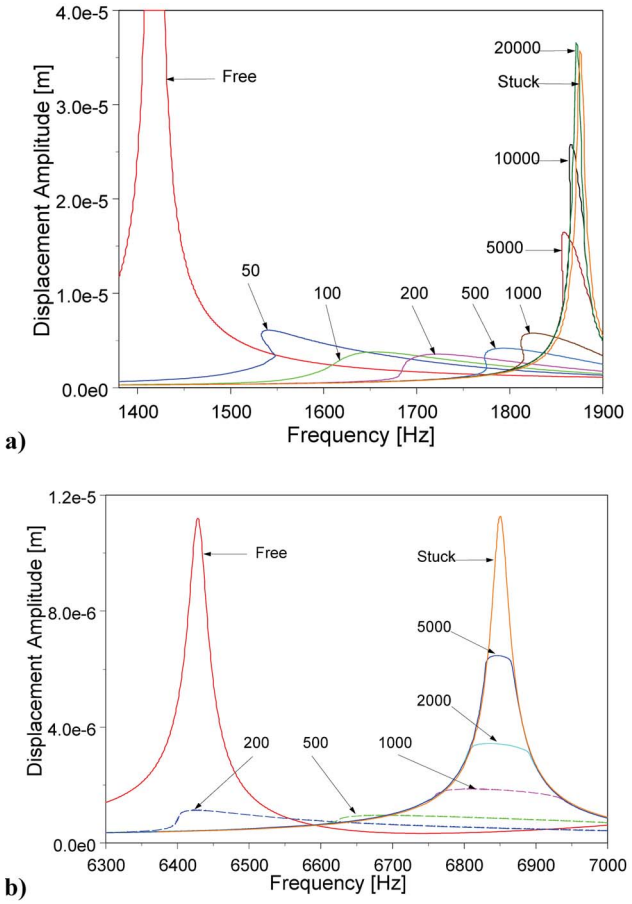


Fig. 10 Tracking plots: (*a*) third and seventh (*b*) modes

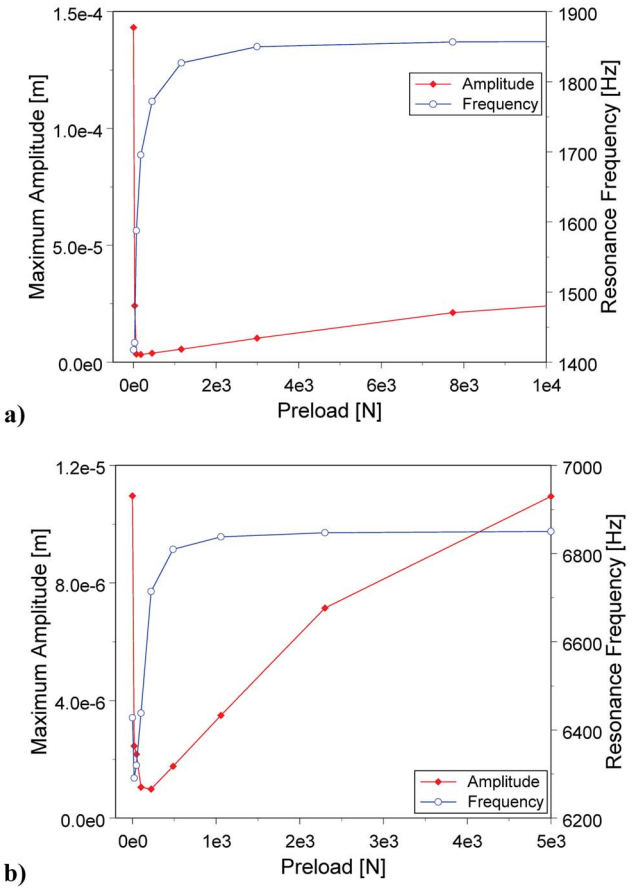


Fig. 11 Optimum and frequency shift curves: (*a*) third and seventh (*b*) mode

are the main contributions to the damper rigid body motion, and as the normal load increases, the major contribution comes from the modal coefficient for translation along the Y axis. Similarly, the major variable components of rigid body motion are translation along X and Y axes and rotation about the Z axis. For all normal load ranges, major variable components of rigid body motion are in the order of translation along X, rotation about Z, and translation along Y. It should be noted that rotation of the damper about the Z axis results in the separation of the contact surfaces, which results in multiple solutions or jumps in the tracking plots.

The bladed-disk system is analyzed using only the translational rigid body modes and translational and rotational rigid body modes of the damper at the first vibration mode of the blade. The results, including and excluding the rotational modes for different normal load cases, are compared in Fig. 14. It is observed that neglecting the rotation of the damper results in an underestimation of the maximum vibration amplitudes; in addition to this, frequency shift is overestimated for this case. It is also interesting to note that a jump phenomenon (multiple solutions) does not exist if the rotational modes of the damper are neglected. Therefore, it can be concluded that the separation of the contact surfaces is associated with the rotation of the damper.

4.1.3 Effects of Partial Slip on Forced Response. In order to observe the effects of partial slip, the forced response analysis is performed for different numbers of contact points on the left and right contact planes. Figure 15 shows the comparison of forced response results for different normal load cases for 9×9, 24×24, and 48×48 contact points. It is observed that multiple solutions in the forced response are captured better when more contact points are employed in the analysis. Moreover, using 9×9 contact points resulted in an underestimation of the vibration amplitude in the cases studied. The results for 24×24 and 48×48 contact points are closer to each other, 48×48 having the highest vibration amplitude in most of the cases. However, employing more contact nodes results in longer calculation times; therefore, optimum values for the number of contact points can be determined by comparing the forced response results.

In Fig. 16, the contact status of four sample contact points on the left and right contact planes of the wedge damper are shown

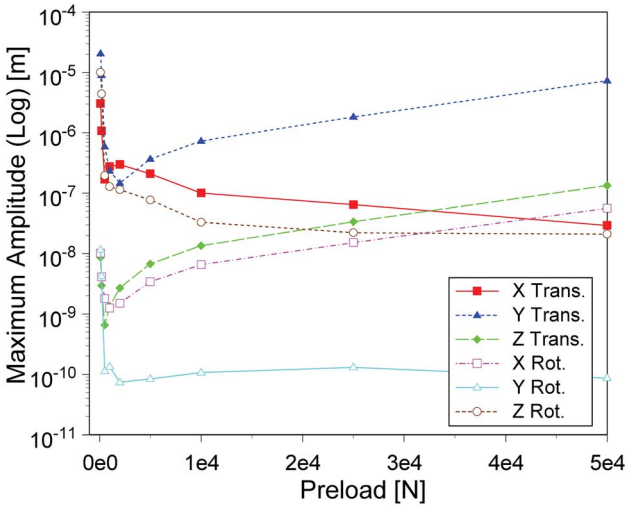


Fig. 12 Effect of normal load on the rigid body motion of damper (bias component)

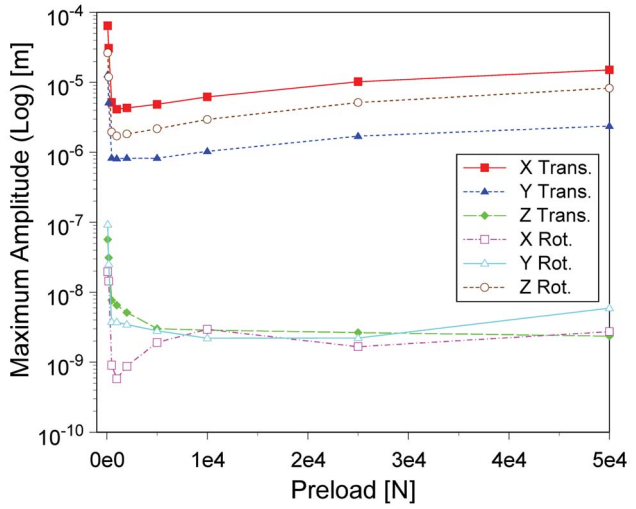


Fig. 13 Effect of normal load on the rigid body motion of damper (vibratory component)

for the normal load of 5000 N at the maximum amplitude frequency of 308.3 Hz. The length of the bar represents the periodic temporal scale. Partial slip on both contact surfaces can be clearly seen from the figure where contact points undergo different states at different times.

4.2 Blade-to-Ground Damper System

4.2.1 Effect of the Number of Harmonics on the Forced Response. In order to present the effect of the number of harmonics on the forced response of frictionally constrained structures, a blade-to-ground damper system is investigated, where the blade given in Fig. 5 is in contact with the ground from the right side, as shown in Fig. 17. This simple model is chosen in order to decrease the number of unknowns and hence decrease the computational time for the analysis as well as to control the initial preload/gap distribution, as requested.

Tracking plots for the blade-to-ground damper system are given in Fig. 18 for single-harmonic and multiple-harmonic responses. It is observed that for the cases where multiple solutions exist, the difference between the single- and multiple-harmonic solutions is

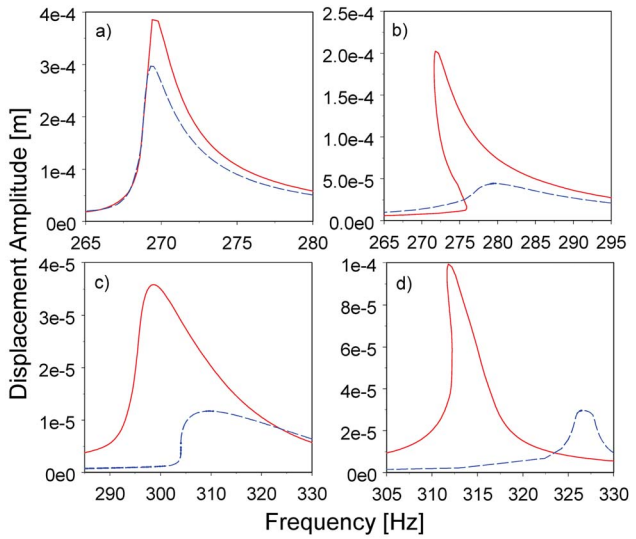


Fig. 14 Effect of rotational modes for normal load; (*a*) 100, (*b*) 200, (*c*) 1000, and (*d*) 10,000; ——, translational and rotational modes; ----------, translational modes

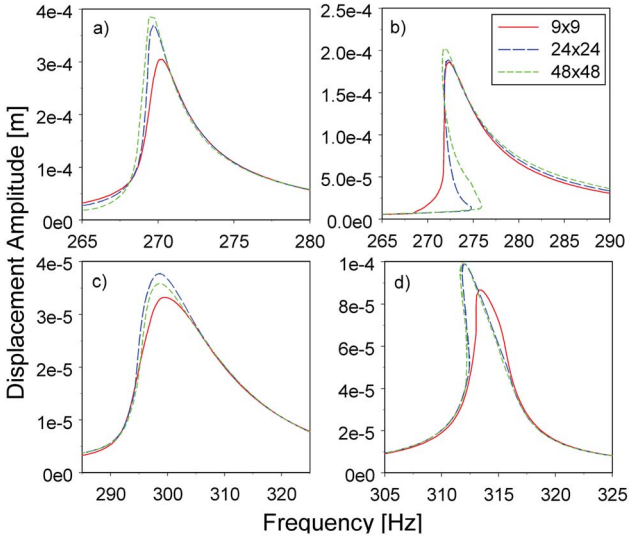


Fig. 15 Effect of the number of contact points for normal load; (*a*) 100, (*b*) 200, (*c*) 1000, and (*d*) 25,000

significant. This is due to the separation of the contact interface, and it can be concluded that, if separation occurs single-harmonic solution cannot capture the nonlinear characteristics accurately. However, for high preload cases, a single-harmonic solution cap-

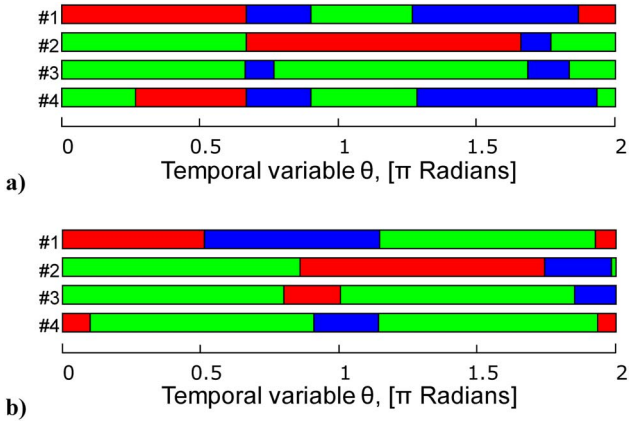


Fig. 16 Contact status of sample points on (*a*) left (*b*) right contact planes: (red) stick, (blue) slip, and (green) separation). (Color version of this figure available online only.)

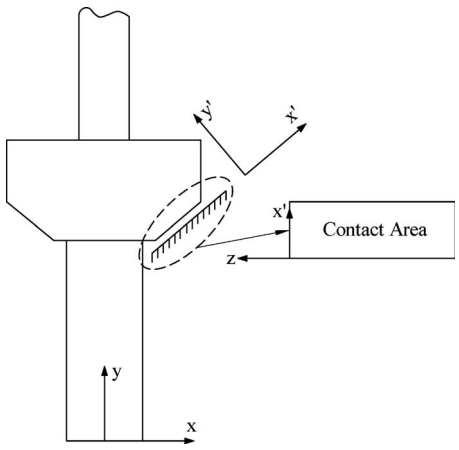


Fig. 17 Schematic for the blade to ground damper

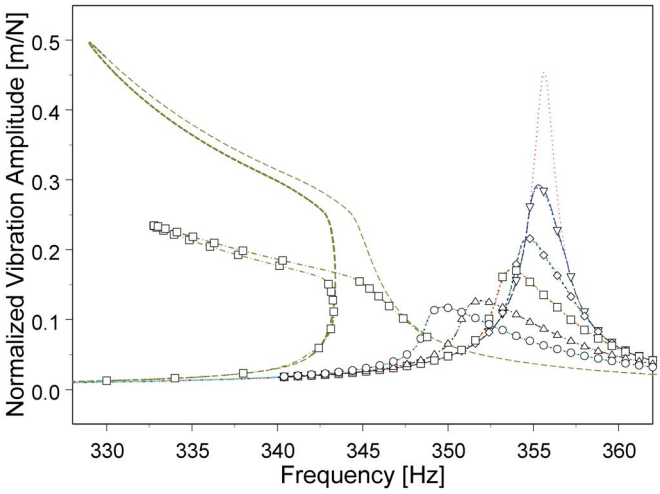


Fig. 18 Tracking plots, (·····), stuck; single-harmonic: (— — —) 1.0e6, (-------) 5.0e5, (·–·–·–) 2.5e5, (·–···–···) 1.0e5, (– – – –) 5.0e4, (– – – – –) 1.0e4; multiharmonic: (···▽···) 1.0e6, (···◇···) 5.0e5, (···□···) 2.5e5, (···△···) 1.0e5, (···○···) 5.0e4, and (- – -□- – -) 1.0e4

tures the nonlinear characteristics quite well and the difference between the single- and multiple-harmonic solutions is negligible. This result is also in agreement with the findings of Chen and Menq [10].

4.3 Comparison With the Test Case Data. In order to verify the developed method, prediction for a test blade, used by GE aircraft engines in a friction damping experiment, is compared with the test data. The schematic for the experimental setup is given in Fig. 19, where two wedge dampers are placed at each side of the test blade, and they are retained by two dummy blades without any airfoils. The normal load on the damper is adjusted by controlling the tension in the damper load wires. The test blade is excited by a pulsating air jet, where the excitation levels are controlled by air-jet supply pressure. Strain gages are placed at several locations on the blade, including the airfoil root, in order to measure the vibratory stresses.

The finite element model of the test case is given in Fig. 20. Since the developed computer code is designed for a tuned system analysis, the test blade and the dummy blades are considered as a single structure for which the interblade phase angle is zero. A comparison of the predicted frequency shift curve with the available test data is given in Fig. 21. For high damper load cases, predicted frequency shift is lower than the test data. However, overall, the predicted frequency shift curve is in good agreement with the experimental data. It should be noted that these predictions are done by using the contact stiffnesses calculated by the proposed approach, as discussed previously. In Fig. 22, the predicted normalized optimal curve and the test data are presented. It is observed that predictions and test results are in good agreement, and the optimal damper load can be predicted acceptably without any parameter tuning. The only parameter to be determined is the friction coefficient, which, in this case, was acquired experimentally. Due to the acyclic structure of the test case, complete locking of the contact surfaces is not possible, which results in mi-

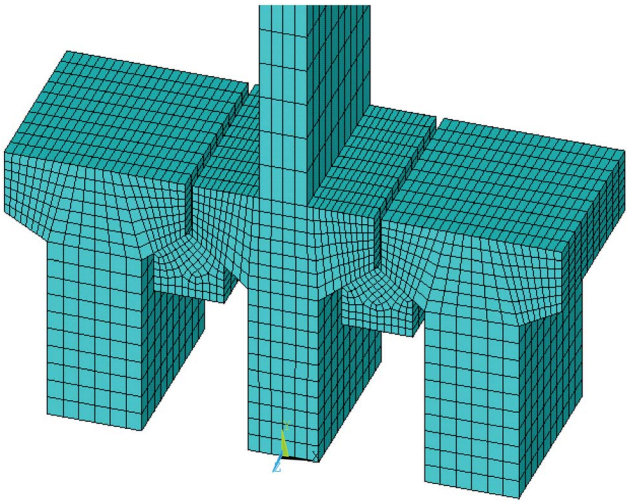


Fig. 20 Finite element model for the test case

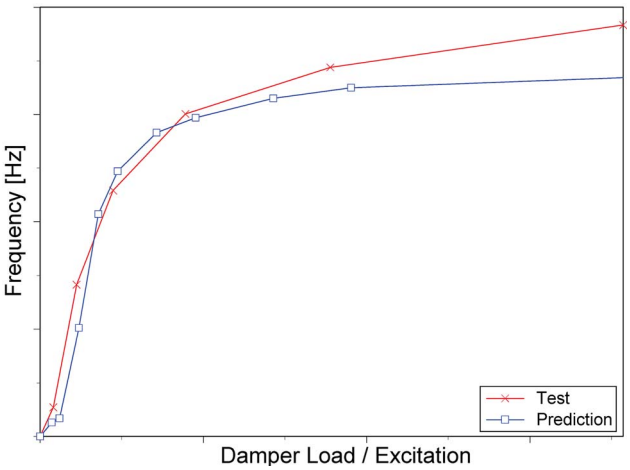


Fig. 21 Frequency shift curve

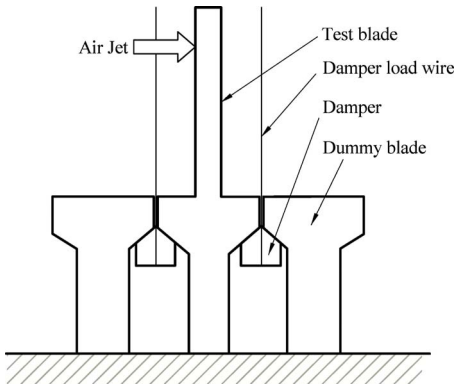


Fig. 19 Schematic view of the test case

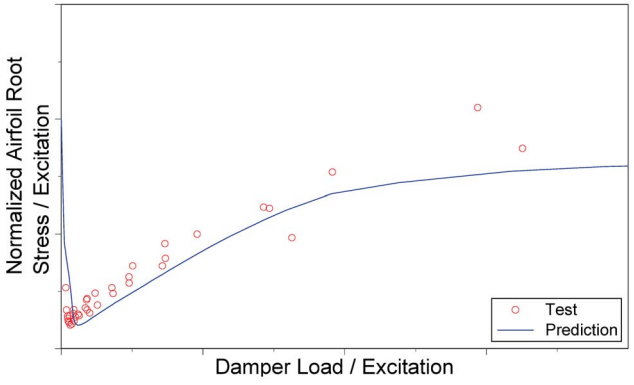


Fig. 22 Predicted normalized optimal curve and test data

croslip for those cases. This can clearly be seen from the predictions given in Fig. 22. Predictions for the test case are also performed for including more blade modes into the analysis, for which contact stiffnesses are recalculated, and it is observed that predictions obtained from both cases are similar and very close to each other. All these results are not presented here for brevity.

5 Conclusions

A forced response prediction method for the analysis of constrained and unconstrained structures coupled through frictional contacts is presented. In the developed method, the unconstrained structure is modeled as a solid body having six rigid body motions as well as elastic deformation. Discrete contact points, which are associated with contact stiffnesses in three directions of motion, are distributed on the contact surfaces. In order to determine the initial preload or gap at each contact point, a quasistatic contact analysis is performed initially for each normal load acting on the damper.

A method is proposed to calculate the contact stiffnesses used in the friction model. The suggested method is based on representing the effect of higher vibration modes by springs associated with each contact pair, which makes it possible to capture local deformations at the contact interface. Therefore, contact kinematics can be accurately estimated by using a reasonable number of mode shapes in the forced response prediction, which decreases the computational cost significantly.

For the forced response analysis, a friction model with normal load variation induced by normal motion is employed to determine the three dimensional contact forces. The harmonic balance method is employed to approximate these contact forces in order to calculate the forced response of constrained and unconstrained structures coupled through frictional contacts. Modal superposition is used to express the relative motion; therefore, the number of unknowns in the resulting nonlinear equation set is proportional to the number of modes and the number of harmonics employed in the analysis, and it is independent of the number of contact points.

The developed forced response prediction method is demonstrated on an example bladed-disk system with wedge dampers, where blades and dampers represent the constrained and unconstrained structures, respectively. A tuned bladed-disk system is studied by the method developed, and forced response results are presented. Multiple solutions (jumps) in the tracking plots for the first vibration mode are observed, which are due to the separation of the contact surfaces associated with the rotation of the damper about the Z axis.

For the first vibration mode, the major contributions to the vibratory component of motion comes from the translation along the X axis, rotation about the Z axis, and translation along the Y axis. In addition to this, the effects of rotational modes are also analyzed, and it is observed that neglecting rotational modes results in an underestimation of the vibration amplitude and an overestimation of the frequency shift. Moreover, no multiple solutions (jumps) in the tracking plots are observed if the rotational modes are neglected.

In order to analyze the effects of partial slip, forced response predictions are calculated for three different numbers of contact points. Utilizing more contact points makes it possible to capture the stick-slip-separation phenomenon of the contact surfaces more accurately, which is observed in contact status plots. The method developed can be used to obtain optimum values for the number of contact points in order to meet the accuracy and computational requirements.

The effects of multiple harmonics are also investigated on a blade-to-ground damper example. It is observed that multiple harmonics is necessary only for the case of jump, where the normal load is low; on the other hand, for high normal loads, multiple harmonics and single-harmonic solutions are approximately the same. Since friction dampers are designed to work at higher damper loads, single-harmonic solutions will be adequate for damper optimization purposes.

Finally, predictions for a test case are compared with the test data, and it is observed that simulation results and test results are in good agreement. Similar forced response predictions are obtained by increasing the number of blade modes used in the analysis, which verifies the developed forced response prediction method and contact stiffness calculation method presented. Utilizing the contact stiffnesses obtained by the proposed method, parameter tuning for contact stiffnesses is eliminated, and the only contact parameter left to be determined is the friction coefficient, which significantly simplifies the forced response prediction process.

Acknowledgment

This material is based on work supported by the Guide Consortium of the Carnegie-Mellon University, which is sponsored by the Air Force Research Laboratory under Contract No. F33615-01-C-2186. The Consortium director is Professor Jerry H. Griffin.

References

[1] Yang, B. D., and Menq, C. H., 1998, "Characterization of Contact Kinematics and Application to the Design of Wedge Dampers in Turbomachinery Blading: Part 1—Stick-Slip Contact Kinematics," ASME J. Eng. Gas Turbines Power, **120**, pp. 410–417.

[2] 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.

[3] Sanliturk, K. Y., Ewins, D. J., and Stanbridge, A. B., 2001, "Underplatform Dampers for Turbine Blades: Theoretical Modeling, Analysis, and Comparison With Experimental Data," ASME J. Eng. Gas Turbines Power, **123**, pp. 919–929.

[4] Pfeiffer, F., and Hayek, M., 1992, "Stick-Slip Motion of Turbine Blade Dampers," Philos. Trans. R. Soc. London, Ser. A, **338**, pp. 503–517.

[5] Sextro, W., Popp, K., and Walter, T., 1997, "Improved Reliability of Bladed-Disks to Friction Dampers," ASME International Gas Turbine and Aeroengine Congress and Exposition, Orlando, June.

[6] Csaba, G., 1999, "Modeling of a Microslip Friction Damper Subjected to Translation and Rotation," ASME International Gas Turbine Conference, Indianapolis, June.

[7] Jareland, M. H., 2001, "Experimental Investigation of a Platform Damper With Curved Contact Areas," ASME Design Engineering Conference, Pittsburgh.

[8] Panning, L., Sextro, W., and Popp, K., 2003, "Spatial Dynamics of Tuned and Mistuned Bladed-Disks With Cylindrical and Wedge-Shaped Friction Dampers," Int. J. Rotating Mach., **9**, pp. 219–228.

[9] Menq, C. H., Griffin, J. H., and Bielak, J., 1986, "The Forced Response of Shrouded Fan Stages," ASME J. Vib., Acoust., Stress, Reliab. Des., **108**, pp. 50–55.

[10] Chen, J. J., and Menq, C. H., 2001, "Periodic Response of Blades Having Three-Dimensional Nonlinear Shroud Constraints," ASME J. Eng. Gas Turbines Power, **123**, pp. 901–909.

[11] Ferrero, J. F., Yettou, E., Barrau, J. J., and Rivallant, S., 2004, "Analysis of a Dry Friction Problem Under Small Displacements Application to a Bolted Joint," Wear, **256**, pp. 1135–1143.

[12] Quin, D. D., and Segalman, D. J., 2005, "Using Series-Series Iwan-Type Models For Understanding Joint Dynamics," ASME J. Appl. Mech., **72**, pp. 666–673.

[13] Ouyang, H., Oldfield, M. J., and Mottershead, J. E., 2006, "Experimental and Theoretical Studies of a Bolted Joint Excited by a Torsional Dynamic Load," Int. J. Mech. Sci., **48**, 1447–1455.

[14] Padmanabhan, C., 1994, "Analysis of Periodically Excited Systems With Clearances," Ph.D. thesis, The Ohio State University.

[15] Duan, C., and Singh, R., 2005, "Transient Responses of a 2-DOF Torsional System With Nonlinear Dry Friction Under a Harmonically Varying Normal Load," J. Sound Vib., **285**, pp. 1223–1234.

[16] Duan, C., and Singh, R., 2006, "Dynamics of a 3DOF Torsional System With a Dry Friction Controlled Path," J. Sound Vib., **289**, pp. 657–688.

[17] Griffin, J. H., 1980, "Friction Damping of Resonant Stresses in Gas Turbine Engine Airfoils," ASME J. Eng. Power, **102**, 329–333.

[18] Dowell, E. H., and Schwartz, H. B., 1983, "Forced Response of a Cantilever Beam With a Dry Friction Damper Attached, Part I: Theory," J. Sound Vib., **91**, pp. 255–267.

[19] Cameron, T. M., Griffin, J. H., Kielb, R. E., and Hoosac, T. M., 1990, "An Integrated Approach for Friction Damper Design," ASME J. Vibr. Acoust., **112**, pp. 175–182.

[20] Ferri, A. A., 1996, "Friction Damping and Isolation Systems," ASME J. Vibr. Acoust., **117**(B), pp. 196–206.

[21] Cigeroglu, E., Lu, W., and Menq, C. H., 2006, "One-Dimensional Dynamic

Microslip Friction Model," J. Sound Vib., **292**, pp. 881–898.
[22] Sanliturk, K. Y., and Ewins, D. J., 1996, "Modeling Two-Dimensional Friction Contact and Its Application Using Harmonic Balance Method," J. Sound Vib., **193**, pp. 511–523.
[23] Menq, C. H., and Yang, B. D., 1998, "Non-Linear Spring Resistance and Friction Damping of Frictional Constraints Having Two-Dimensional Motion," J. Sound Vib., **217**, pp. 127–143.
[24] Yang, B. D., Chu, M. L., and Menq, C. H., 1998, "Stick-Slip-Separation Analysis and Non-Linear Stiffness and Damping Characterization of Friction Contacts Having Variable Normal Load," J. Sound Vib., **210**, pp. 461–481.
[25] Yang, B. D., and Menq, C. H., 1998, "Characterization of 3D Contact Kinematics and Prediction of Resonant Response of Structures Having 3D Frictional Constraint," J. Sound Vib., **217**, pp. 909–925.
[26] Petrov, E. P., and Ewins, D. J., 2003, "Analytical Formulation of Friction Interface Elements for Analysis of Nonlinear Multi-Harmonic Vibrations of Bladed-Disks," ASME J. Turbomach., **125**, pp. 364–371.
[27] Cigeroglu, E., and An, N., Menq, C. H., 2007, "A Microslip Friction Model With Normal Load Variation Induced by Normal Motion," Nonlinear Dyn., **50**(3), pp. 609–626.

Asymptotic Description of Maximum Mistuning Amplification of Bladed Disk Forced Response

Carlos Martel
ETSI Aeronáuticos,
Universidad Politécnica de Madrid,
28040 Madrid, Spain
e-mail: carlos.martel@upm.es

Roque Corral[1]
Industria de Turbopropulsores S.A.,
28830 Madrid, Spain
e-mail: roque.corral@itp.es

The problem of determining the maximum forced response vibration amplification that can be produced just by the addition of a small mistuning to a perfectly cyclical bladed disk still remains not completely clear. In this paper we apply a recently introduced perturbation methodology, the asymptotic mistuning model (AMM), to determine which are the key ingredients of this amplification process and to evaluate the maximum mistuning amplification factor that a given modal family with a particular distribution of tuned frequencies can exhibit. A more accurate upper bound for the maximum forced response amplification of a mistuned bladed disk is obtained from this description, and the results of the AMM are validated numerically using a simple mass-spring model.
[DOI: 10.1115/1.2968868]

1 Introduction

The term mistuning refers to the small differences among the (nominally designed to be identical) rotor blades in turbomachinery bladed disks. These small unavoidable imperfections typically result from the manufacturing process and wear, and break the cyclic symmetry of the structure, splitting the tuned eigenvalue pairs and preventing the existence of pure traveling waves (TWs) as natural vibration modes of the mistuned structure. This loss of cyclic symmetry has a dramatic effect in the forced response of the bladed disk: the resulting mistuned blade vibration patterns are no longer uniformly distributed circumferentially and can exhibit localized vibration amplitudes significantly higher than those of the tuned case. The amplification of the response levels may lead to a severe reduction in the fatigue life of the blades, and thus mistuning has a strong negative impact on safety, operability, and readiness. It is therefore very important to understand the basic mechanisms behind mistuning amplification in order to be able to produce new technological solutions that can predict and control its effect and that can result in an enhancement of overall aircraft readiness and a substantial reduction in maintenance costs.

Numerous studies have been conducted to understand mistuning effects on the dynamics of bladed disk assemblies (see, e.g., the survey paper of Slater et al. [1] and the more recent review by Castanier and Pierre [2], and the exhaustive references therein.

Despite the great research activity developed in this subject since the 1970s, there are basic questions of critical technological importance, such as the determination of the maximum vibration amplification factor of forced response due to mistuning, that still remain without a clear answer. Since the seminal work of Whitehead [3] in maximum mistuning amplification to the recent presentations by Han and and Mignolet [4] and Han et al. [5], there appears to be no solid consensus about the maximum mistuning amplification factor.

The problem of directly computing the maximum amplification factor is a formidable task because (i) mistuning breaks the cyclic symmetry of the problem, and the whole bladed disk (instead of a single sector) has to be considered, and (ii) the random nature of the mistuning requires the evaluation of the response of the system for a large number of different mistuning patterns to somehow statistically explore all possible configurations. Petrov and Ewins [6] formulated the problem of finding the mistuning pattern that gives the maximum amplification factor as an optimization problem. They were able to locate the worst mistuning pattern in a realistic bladed disk configuration that was quite CPU costly, but this direct approach did not provide much insight into the basic mechanisms of mistuning amplification.

In order to overcome the above-mentioned difficulties, several types of simplified models have been developed to reduce the required computational time. Simple lumped models with the structure modeled by 1DOF mass-spring systems (see Ref. [7] and the more recent work by Sinha [8], Kenyon and Griffin [9], and Kenyon et al.[10]) and more detailed reduced order models (ROMs) [11–14] yield good results once their coefficients are appropriately tuned but do not provide much insight into the basic physical mechanisms involved in the amplification of the forced response if their coefficients are mistuned.

Recently, a simpler model, known as the fundamental mistuning model (FMM), that reduces the set of nominal modes to a single modal family has been developed [15–17]. This method only requires us to know the natural frequencies of the tuned system, but its applicability is limited to a modal family with nearly equal frequencies (see Fig. 1).

This paper presents further reduction in the FMM for the very frequent case in which all modes of the family do not share the same frequency (see Fig. 1). This new model is derived by means of a fully consistent asymptotic analysis starting from the complete problem of the mistuned bladed disk. The small parameters in the asymptotic expansion are the small frequency corrections induced by mistuning and the small damping of the tuned modes. The resulting model, referred to hereafter as the asymptotic mistuning model (AMM), is not significantly faster than the FMM but it allows us to uncover the basic physical mechanisms of mistuning since it accounts only for the minimum subset of degrees of freedom that play a relevant role in the mistuning process. The AMM also provides a more accurate upper bound for the maximum mistuned response amplification that a forced bladed disk can experience.

The AMM has already been used by Martel et al. [18] for the

[1]Associate Professor at ETSI Aeronáuticos, Universidad Politécnica de Madrid, 28040 Madrid, Spain.

Contributed by the International Gas Turbine Institute of ASME for publication in the in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 7, 2008; final manuscript received April 9, 2008; published online December 23, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

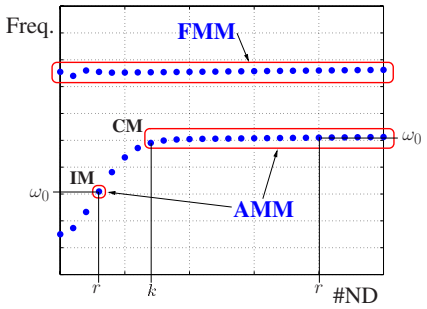


Fig. 1 Tuned natural vibration frequencies versus number of nodal diameters for a bladed disk. IM: isolated mode; CM: clustered modes.

study of the optimal intentional mistuning patterns for the stabilization of aerodynamically unstable rotors. In this paper we describe in detail the derivation of the AMM for the forced response of a general realistic mistuned bladed disk. Moreover, we apply this methodology to two frequent modal configurations not covered by the FMM: isolated and clustered modes (labeled IM and CM, respectively, in Fig. 1), obtaining the maximum amplification factor and the associated mistuning distribution and comparing the AMM results with those from the numerical simulation of a mistuned mass-spring model.

2 Asymptotic Mistuning Model

This methodology is based on the fact that the mistuned bladed disk constitutes just a small distortion of the tuned one and uses asymptotic perturbation techniques to evaluate the effect of the small mistuning in the magnitude of the forced vibration amplitude. The resulting simplified description includes the main physical mechanisms involved in the mistuning amplification process.

2.1 Tuned System. The tuned system can be regarded as the background state on top of which we compute the mistuning correction. In this section we give a rather detailed description of the forced response problem for a general tuned bladed disk. These are all well known results, but we believe that it is important to briefly collect them here to clarify the notation before proceeding with the evaluation of the corrections.

The equations of motion for the forced oscillations of a perfectly tuned bladed disk with N identical sectors take the form

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{f}(t) \tag{1}$$

where the independent bladed disk degrees of freedom (after an appropriate finite element model (FEM) discretization of the structure) are stored in $\mathbf{x}(t)$, $\mathbf{M}$ and $\mathbf{K}$ are the mass and stiffness matrices (both symmetric and positive definite), and $\mathbf{f}(t)$ accounts for the time dependent external forces acting on the blades. Note that damping has not been included in the above formulation since it is typically a small effect, and it will be added later on together with the mistuning corrections.

The natural vibration modes of the system are obtained from Eq. (1), with $\mathbf{f}(t)$ set to zero, and can be written using the following complex notation:

$$\mathbf{x} = \mathbf{X}e^{i\omega t} + \text{c.c.} \tag{2}$$

where the complex vibration mode shape $\mathbf{X}$ and the mode oscillation frequency ω (c.c. stands for complex conjugate) are given by the following symmetric generalized eigenvalue problem:

$$(\mathbf{K} - \omega^2\mathbf{M})\mathbf{X} = 0 \tag{3}$$

If we now divide the response vector as

$$\mathbf{X} = \begin{bmatrix} X_1 \\ \vdots \\ X_j \\ \vdots \\ X_N \end{bmatrix} \tag{4}$$

with the vector X_j containing the displacements of the m degrees of freedom associated with sector j, then the eigenvalue problem takes the form

$$\left(\begin{bmatrix} K & K_c & 0 & \cdots & K_c^T \\ K_c^T & K & K_c & \cdots & 0 \\ & & \ddots & \ddots & \\ K_c & 0 & \cdots & K_c^T & K \end{bmatrix} - \omega^2 \begin{bmatrix} M & M_c & 0 & \cdots & M_c^T \\ M_c^T & M & M_c & \cdots & 0 \\ & & \ddots & \ddots & \\ M_c & 0 & \cdots & M_c^T & M \end{bmatrix} \right) \times \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{5}$$

where the $m \times m$ sector stiffness and mass matrices K and M are symmetric and the coupling between adjacent sectors is represented by the coupling stiffness and mass matrices K_c and M_c.

The cyclic symmetry of the bladed disk (which is composed of a sector repeated N times and arranged periodically) is now evident in the matrices in Eq. (5) that exhibit a block circulant structure; i.e., each row can be obtained from the previous one after a wrap-around forward block shift. The associated complex eigenvectors are of the form

$$\begin{bmatrix} X_1 \\ \vdots \\ X_j \\ \vdots \\ X_N \end{bmatrix} = \begin{bmatrix} Z_k e^{i(2\pi k/N)1} \\ \vdots \\ Z_k e^{i(2\pi k/N)j} \\ \vdots \\ Z_k e^{i(2\pi k/N)N} \end{bmatrix} \quad \text{for} \quad k = 1, \ldots, N \tag{6}$$

and correspond to TWs with wavenumber k (and $\min(|k|, |N-k|)$ nodal diameters) that, taking into account Eq. (2), rotates with constant angular velocity ω_k/k. The complex vector Z_k of size m contains the mode shape details, and the only eigenvector variation from sector to sector is just a phase shift of the amount $2\pi k/N$ (the so-called "interblade phase angle").

Note that the modes $\mathbf{X}$ of the undamped problem in Eq. (2) can be taken to be real, but instead we take them to be complex (see Eq. (6)) in order to use the much simpler sector-to-sector TW formulation.

If we now insert expression (6) into Eq. (5), N decoupled eigenvalue problems of size m are obtained,

$$((K_c - \omega^2 M_c)e^{i(2\pi k/N)} + (K_c - \omega^2 M_c)^T e^{-i(2\pi k/N)} + K - \omega^2 M)Z_k = 0$$

$$\text{for} \quad k = 1, \ldots, N \tag{7}$$

which, for each k, have m real eigenvalues, $\omega_{k1}^2, \ldots, \omega_{km}^2 \geq 0$, that give the oscillation frequencies of the structure. The associated eigenvectors, $Z_{k1}, \ldots, Z_{km}$, are, in general, complex and can be normalized in such a way that the square matrix

$$P_k = [Z_{k1} | Z_{k2} \cdots | Z_{km}] \tag{8}$$

verifies

$$P_k^H (M_c e^{i(2\pi k/N)} + M_c^T e^{-i(2\pi k/N)} + M) P_k = I \tag{9}$$

$$P_k^H (K_c e^{i(2\pi k/N)} + K_c^T e^{-i(2\pi k/N)} + K) P_k = \Omega_k^2 = \begin{bmatrix} \omega_{k1}^2 & & 0 \\ & \ddots & \\ 0 & & \omega_{km}^2 \end{bmatrix}$$

$$\text{for } k = 1, \ldots, N \tag{10}$$

that is, the eigenvectors are orthonormal with respect to the mass matrix and turn into a diagonal form the sector stiffness matrix (the superindex ^{H} stands for the conjugate transpose matrix) (see, e.g., Ref. [19]).

The eigenvalue equation (Eq. (7)) remains invariant under the changes

$$k \to N-k, \quad \omega_k^2 \to \omega_{N-k}^2, \quad Z_k \to \bar{Z}_{N-k} \tag{11}$$

Thus, for every traveling wave rotating in one direction, there is also another identical one that rotates in the opposite direction with the same velocity.

In summary, the natural vibration modes of the tuned system are traveling wave modes that come in counterpropagating pairs, except for those associated with $k=0$ and $k=N/2$ (present only for N even), which have real Z_k and can be seen as nonpropagative standing waves. Property (11) implies that the plot of the natural frequencies versus the number of nodal diameters, k, is symmetric (i.e., $\omega_k=\omega_{N-k}$), and, typically, only its first half (interblade phase angle from 0 to π) is normally represented, as in Fig. 1.

We now go back to the forced problem (1) and consider a traveling wave excitation with frequency ω and engine order r,

$$\mathbf{f}(t) = \mathbf{F}e^{i\omega t} + \text{c.c.} = \begin{bmatrix} \mathbf{F}e^{i(\omega t+(2\pi r/N)1)} \\ \vdots \\ \mathbf{F}e^{i(\omega t+(2\pi r/N)j)} \\ \vdots \\ \mathbf{F}e^{i(\omega t+(2\pi r/N)N)} \end{bmatrix} + \text{c.c.} \tag{12}$$

which is of particular relevance in this context. The response of the system is much more clearly analyzed if we use the basis of the natural traveling wave modes of the system,

$$\mathbf{X} = \mathbf{P}\mathbf{A} \tag{13}$$

with

$$\mathbf{P} = \frac{1}{\sqrt{N}} \begin{bmatrix} P_1 e^{i(2\pi 1/N)1} & \cdots & P_N e^{i(2\pi N/N)1} \\ \vdots & & \vdots \\ P_1 e^{i(2\pi 1/N)j} & \cdots & P_N e^{i(2\pi N/N)j} \\ \vdots & & \vdots \\ P_1 e^{i(2\pi 1/N)N} & \cdots & P_N e^{i(2\pi N/N)N} \end{bmatrix} \quad \text{and} \quad \mathbf{A} = \sqrt{N} \begin{bmatrix} A_1 \\ \vdots \\ A_j \\ \vdots \\ A_N \end{bmatrix} \tag{14}$$

where the traveling wave modes are the columns of $\mathbf{P}$ (see Eqs. (6) and (8)) and $\mathbf{A}$ contain their amplitudes. In the TW basis, the forced problem takes now the diagonal form

$$\begin{bmatrix} \Omega_1^2 - \omega^2 I & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Omega_N^2 - \omega^2 I \end{bmatrix} \begin{bmatrix} A_1 \\ \vdots \\ A_N \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ P_r^H F \\ \vdots \\ 0 \end{bmatrix} \tag{15}$$

and its solution is very simple,

$$A_i = 0 \quad \text{for} \quad i \neq r \text{ and}$$

$$A_r = \begin{bmatrix} A_{r1} \\ \vdots \\ A_{rm} \end{bmatrix} = \begin{bmatrix} Z_{r1}^H F/(\omega_{r1}^2 - \omega^2) \\ \vdots \\ Z_{rm}^H F/(\omega_{rm}^2 - \omega^2) \end{bmatrix} \tag{16}$$

That is, the response of the tuned undamped system is a traveling wave similar to the forcing, with frequency ω and r nodal diameters, that peaks when the forcing frequency ω is equal to any of the natural frequencies $\omega_{r1}, \ldots, \omega_{rm}$ corresponding to r nodal diameters. Finally, if we also take into account a small damping d (small as compared with the vibration frequency) in the description, then the peak amplitude of each of the traveling waves becomes finite,

$$A_{rk} = Z_{rk}^H F/(\omega_{rk}^2 - \omega^2 + 2id\omega_{rk}) \quad \text{and} \quad |A_{rk}|_{\max} = |Z_{rk}^H F|/(2d\omega_{rk}) \tag{17}$$

where the damping d has been appropriately included in the above expression in order to have a temporal behavior for the unforced damped natural mode of the form $e^{(-d+i\omega_{rk})t}$ (note that we make use of the fact that the damping is small as compared with the frequency to neglect second order corrections). The forced response problem is a linear problem, and therefore we can conveniently normalize the magnitude of the maximum tuned response of a given mode to be 1 simply by setting the amplitude of the forcing to be $Z_{rk}^H F = 2id\omega_{rk}$.

2.2 Mistuning Effects. The forced response for a mistuned bladed disk is given by

$$((\mathbf{K} + \Delta\mathbf{K}) - \omega^2(\mathbf{M} + \Delta\mathbf{M}))\mathbf{X} = \mathbf{F} \tag{18}$$

which are simply the tuned Eq. (3) with the forcing (12) and the mistuning stiffness, $\Delta\mathbf{K}$, and mass, $\Delta\mathbf{M}$, matrices added. $\Delta\mathbf{K}$ and $\Delta\mathbf{M}$ are block diagonal matrices that contain the deviations from the mean stiffness K and mass M sector matrices (see Eq. (5)), that is,

$$\Delta\mathbf{K} = \begin{bmatrix} \Delta K_1 & 0 & \cdots & 0 \\ 0 & \Delta K_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & 0 & \Delta K_N \end{bmatrix}$$

$$\Delta\mathbf{M} = \begin{bmatrix} \Delta M_1 & 0 & \cdots & 0 \\ 0 & \Delta M_2 & \cdots & 0 \\ & & \ddots & \\ 0 & 0 & 0 & \Delta M_N \end{bmatrix} \tag{19}$$

$$\text{and} \quad \sum_{j=1}^{N} \Delta K_j = 0, \quad \sum_{j=1}^{N} \Delta M_j = 0 \tag{20}$$

$$\text{and} \quad \Delta K_j = \Delta K_j^T, \quad \Delta M_j = \Delta M_j^T \quad \text{for} \quad j = 1, \ldots, N \tag{21}$$

Mistuning of the sector coupling block matrices has not been included in the model above: the coupling matrices involve only a very small subset of the DOF of the sector. Thus, its mistuning is negligible as compared with that of the main sector.

As we did in the previous section, we rewrite Eq. (18) using the tuned traveling wave mode basis, Eqs. (13) and (14), to obtain

$$\left(\begin{bmatrix} \Omega_1^2 - \omega^2 I & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \Omega_N^2 - \omega^2 I \end{bmatrix} + \Delta \right) \begin{bmatrix} A_1 \\ \vdots \\ A_N \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ P_r^H F \\ \vdots \\ 0 \end{bmatrix} \tag{22}$$

The structure of the mistuning correction matrix in the traveling wave basis,

$$\Delta = \mathbf{P}^H(\Delta\mathbf{K} - \omega^2\Delta\mathbf{M})\mathbf{P} \tag{23}$$

can be more easily analyzed if we make use of the Fourier modes of the sector distribution of the mistuning mass and stiffness matrices, ΔK_k^F and ΔM_k^F,

$$\Delta K_j = \sum_{k=1}^{N} \Delta K_k^F e^{i(2\pi k/N)j} \tag{24}$$

$$\Delta M_j = \sum_{k=1}^{N} \Delta M_k^F e^{i(2\pi k/N)j} \quad \text{for} \quad j = 1, \ldots, N \tag{25}$$

which verify (see Eqs. (20) and (21))

$$\Delta K_N^F = 0, \quad \Delta M_N^F = 0, \quad \Delta K_k^F = \overline{\Delta K_{-k}^F}, \quad \Delta M_k^F = \overline{\Delta M_{-k}^F}$$

$$\Delta K_k^F = (\Delta K_k^F)^T, \quad \Delta M_k^F = (\Delta M_k^F)^T \quad \text{for} \quad k = 1, \ldots, N \tag{26}$$

After inserting the above expressions into Eq. (23), the mistuning matrix takes the form

$$\Delta = \begin{bmatrix} 0 & \Delta_{12} & \ldots & \Delta_{1N} \\ \Delta_{21} & 0 & \ldots & \Delta_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \Delta_{N1} & \Delta_{N2} & \ldots & 0 \end{bmatrix} \tag{27}$$

where each block is related to the Fourier components of the mistuning by

$$\Delta_{kj} = P_k^H (\Delta K_{k-j}^F - \omega^2 \Delta M_{k-j}^F) P_j \tag{28}$$

Note that the mistuning matrix exhibits off-diagonal blocks and, thus, mistuning couple tuned traveling waves with different numbers of nodal diameters. Moreover, according to Eq. (28), the coupling between the traveling waves with k and j nodal diameters takes place precisely through the $(k-j)$ harmonic of the mistuning.

The effect of small mistuning (and damping) in the response of the system can be computed, in first approximation, as follows. When the forcing frequency (with engine order r) is close to a resonant tuned frequency ω_0 (corresponding to a TW with r nodal diameters), we can distinguish two essentially different types of traveling wave modes:

1. Modes with frequencies that are not close to ω_0, that is, modes whose frequency ω_{kj} is such that $|\omega_0 - \omega_{kj}|$ is large as compared with the small mistuning and damping terms. The corresponding equation for this kind of modes in system (22) is of the form

$$(\omega_{kj}^2 - \omega^2) A_{kj} = \sum_{i,h=1}^{m} (\text{small terms}) A_{ih} + (\text{small forcing}) \tag{29}$$

where the coefficient in the left hand side is large as compared with the small mistuning, damping, and forcing terms in the right hand side because ω is close to ω_0 and thus away from ω_{kj}. After neglecting the small terms, the equation above yields

$$A_{kj} = 0 \tag{30}$$

That is, all traveling wave modes with frequencies well apart from ω_0 do not contribute in first approximation to the mistuning correction.

2. Modes with frequencies close to ω_0, i.e., with frequencies that are at a small distance from ω_0, of the order of the small damping and mistuning frequency corrections. We will hereafter refer to these modes as "active" modes because they are, in first approximation, the only ones that take part in the mistuned forced response near the resonant frequency ω_0. After removing the not active modes of the previous item from Eq. (22), the resulting equation for an active mode with amplitude A_a and tuned structural frequency ω_a reads

$$(\omega_a^2 - \omega^2 + 2id\omega_0) A_a + \sum_{a' \neq a} \delta_{aa'} A_{a'} = F_a \tag{31}$$

where the sum goes over the rest of the active modes and $F_a = 2id\omega_0$ for the active mode with r nodal diameters (the one that is being directly forced) and $F_a = 0$ for the rest. Note that the frequency in the damping and forcing expression can be set, in first approximation, to ω_0 for all active modes because the neglected frequency variations just produce higher order corrections. All coefficients in the above equation are small, and according to Eq. (28), the mistuning coupling coefficients can be written as

$$\delta_{aa'} = Z_a^H (\Delta K_{a-a'}^F - \omega_0^2 \Delta M_{a-a'}^F) Z_{a'} \tag{32}$$

where Z_a, $Z_{a'}$, a, and a' are, respectively, the mode shapes and the numbers of nodal diameters of the active modes A_a and $A_{a'}$, and the frequency ω has been approximated by ω_0. The mistuning coupling terms verify $\delta_{aa'} = \overline{\delta_{a'a}}$ because mistuning constitutes a purely structural conservative perturbation.

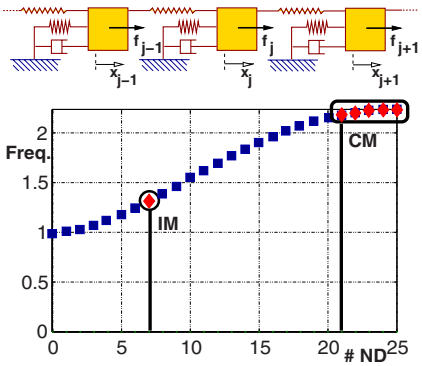


Fig. 2 Top: sketch of the simple 1DOF per sector system. Bottom: tuned frequencies versus number of nodal diameters (IM: forcing engine order 7; CM: forcing engine order 21).

In a general situation, we have as many equations like Eq. (31) as active modes, which form a simplified model for computing the mistuning effect on the forced response. We call this reduced problem AMM because the above described procedure can be regarded as an asymptotic perturbation method that gives the main (first order) effects of mistuning. The small parameters in this asymptotic expansion are the damping, the mistuning, and the small variation among the natural frequencies of the active modes.

The AMM description retains only the minimal set of traveling wave modes that are involved in the mistuning correction. The simplicity of the AMM reveals that the only TW modes that are relevant are those with frequency close to the natural frequency of the mode that is being forced (i.e., the active modes) and that the effect of mistuning is to couple these modes. Moreover, the AMM indicates precisely how this coupling among TWs takes place: it comes from the Fourier mode of the mistuning distribution with the wavenumber equal to the difference of the number of nodal diameters of the two TWs (see Eq. (32)).

3 Applications

In this section we apply the AMM to analyze the mistuning effect on the amplitude of the forced response in the two very frequent modal configurations shown in Fig. 1: IM and CM.

The results are compared with those obtained from the numerical solution of the simple 1DOF per sector system sketched in Fig. 2, which is given (in a nondimensional form) by the system of equations

$$\ddot{x}_j + \omega_a^2 (1 + \varepsilon \delta_j) x_j + \omega_c^2 (2x_j - x_{j+1} - x_{j-1}) + c\dot{x}_j = f e^{i(\omega t + 2\pi (r/N) j)} + \text{c.c.}, \quad j = 1, \ldots, N \tag{33}$$

where $\omega_a^2 = \omega_c^2 = 1$, $\varepsilon = c = .001$, δ_j is the mistuning distribution, f is scaled in order to have a tuned response equal to 1, and $N = 50$ (see the bottom plot in Fig. 2 for a representation of the tuned vibration frequencies of this system).

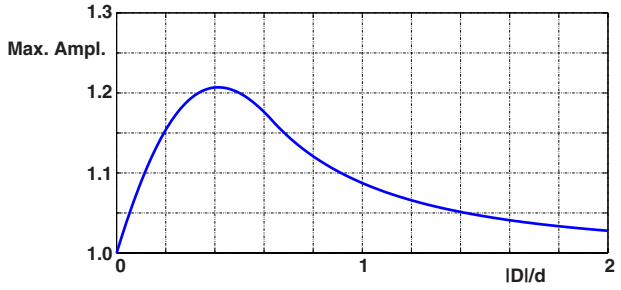


Fig. 3 Isolated mode case: maximum attainable amplitude versus mistuning amplitude $|D|$

3.1 Isolated Modes. The TW with r nodal diameters and vibration frequency ω_0 labeled IM in Fig. 1 is well apart from the rest; i.e., its distance to the rest of the frequencies is large as compared with the damping and mistuning frequency corrections. In this case, in order to take into account the effect of the mistuning in the response of the system to a forcing with engine order r and frequency close to ω_0, it is enough to consider the two counter-rotating TWs with frequency ω_0 and nodal diameters $\pm r$. In other words, these two modes are the only active modes in the AMM description, which now takes its simplest form,

$$(\omega_0^2 - \omega^2 + 2id\omega_0)A_{+r} + \bar{\delta}A_{-r} = 2id\omega_0$$

$$\delta A_{+r} + (\omega_0^2 - \omega^2 + 2id\omega_0)A_{-r} = 0 \tag{34}$$

We can further simplify this system by introducing the variable $\Delta\omega = \omega - \omega_0$ and neglecting all second order terms to obtain

$$\begin{bmatrix} -\Delta\omega + id & D \\ \bar{D} & -\Delta\omega + id \end{bmatrix} \begin{bmatrix} A_{+r} \\ A_{-r} \end{bmatrix} = \begin{bmatrix} id \\ 0 \end{bmatrix} \tag{35}$$

where the effect of mistuning is contained in the off-diagonal coefficient (see Eq. (32)),

$$D = \delta/(2\omega_0) = Z_{+r}^H(\Delta K_{2r}^F - \omega_0^2 \Delta M_{2r}^F)\overline{Z_{+r}}/(2\omega_0) \tag{36}$$

which involves only the $2r$ Fourier coefficient of the mistuning distribution.

The response of the system is, in first approximation, just a combination of the two active modes,

$$X_j = (Z_{+r}A_{+r}e^{i(2\pi r/N)j} + Z_{-r}A_{-r}e^{-i(2\pi r/N)j}) \quad \text{for} \quad j = 1, \ldots, N \tag{37}$$

and the amplitude of the displacements is proportional to

$$\text{Amplitude} \propto |A_{+r} + A_{-r}| \tag{38}$$

This amplitude is equal to 1 for the tuned case, and when mistuning is present, it is given by

$$\text{Amplitude} = \frac{|(-\Delta\omega + id) - D|}{|(-\Delta\omega + id)^2 - |D|^2|} \tag{39}$$

The maximum amplitude of the forced response for each mistuned size $|D|$ is plotted in Fig. 3. As expected, the maximum amplitude first grows and then decays when the mistuning size is increased. The global maximum attainable mistuned amplitude can be explicitly computed from the above expression to be

$$\text{Amplitude}_{\max} = \frac{1+\sqrt{2}}{2} = 1.207, \ldots \tag{40}$$

$$D_{\max} = -d(\sqrt{2} - 1)i \tag{41}$$

$$\Delta\omega_{\max} = 0 \tag{42}$$

These same results were already obtained by Macbain and Whaley [20], who used a formulation for the forced response that included just two forced oscillators with close frequencies and small damping.

In order to check the above results, several simulations of the full mass-spring system given by Eq. (33) have been performed with an engine order $r=7$ and with frequencies close to the isolated mode frequency labeled IM in Fig. 2. This situation corresponds to an isolated mode since the frequency corresponding to $r=7$ is at a large distance from the rest of the frequencies, as compared with the damping (more precisely, the distance to its closest neighbors measured in terms of the damping is given by $|\omega_{r=7} - \omega_{r=8}|/d \sim 150$ and $|\omega_{r=7} - \omega_{r=6}|/d \sim 143$).

In Fig. 4 the mistuning distribution is harmonic with the wavenumber $k=2r=14$ (see Fig. 4, middle and bottom left plots) and the precise size predicted by the AMM for the maximum amplification. The resulting maximum amplitude, 1.2066,… (see Fig. 4, top left plot), is approximately that predicted by the AMM, and as can be appreciated from the right plot in Fig. 4, the only TW modes involved in the mistuned response are those with wavenumbers 7 and −7. As predicted by the AMM, the response of the system remains almost unchanged if we add more harmonics with random amplitudes to the mistuning distribution (see Fig. 5). Finally, if we remove from the mistuning the harmonic component with wavenumber 14 and keep the others, then there is no appreciable amplification of the response and the system behaves as if tuned, generating only the directly forced TW with wavenumber 7 (see Fig. 6) and again confirming the AMM results and assumptions.

In summary, the mistuning effect in the forcing of a TW with a frequency well apart form its neighbors (i.e., with a large frequency gap as compared with the damping) can produce a maximum amplification of up to 20% approximately, which takes place for a very small mistuning size (of the order of the damping) and through the mistuning harmonic with a wavenumber equal to twice the engine order of the forcing. This particular forced response sensitivity to the mistuning harmonics was already noticed by Kenyon and Griffin [21] in their numerical simulations using a simplified ring system, and the maximum factor $(1+\sqrt{2})/2$ can be obviously guessed from some of the amplification curves presented in their paper. The systematic approach provided by the AMM goes further than those previous results, giving more precise information and allowing us to quantitatively evaluate the maximum attainable amplification factor and the associated mistuning pattern in a general bladed disk configuration.

3.2 Clustered Modes. If we now force the system with an engine order r such that the corresponding tuned TW has a frequency ω_0 on the horizontal part of the frequency distribution of the modal family (see the encircled modes labeled CM in Fig. 1), then we have more than two active modes. The AMM description, according to Eq. (31), can now be written as

$$\begin{bmatrix} d_k & & & & \\ & d_{k+1} & & D & \\ & & \ddots & & \\ & D^{\mathrm{H}} & & d_{-(k+1)} & \\ & & & & d_{-k} \end{bmatrix} \begin{bmatrix} A_k \\ A_{k+1} \\ \vdots \\ A_{-(k+1)} \\ A_{-k} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ id \\ \vdots \\ 0 \end{bmatrix} \tag{43}$$

where we have ordered the active modes consecutively in the number of nodal diameters (recall that it has to be considered only mod N). The group of active modes are those with k or more nodal diameters (see Fig. 1), and we thus have a total of $N_a = N - 2k + 1$ active modes, where N is the number of sectors of the bladed disk. The diagonal elements of the above matrix are of the form

$$d_j = \Delta\omega_j - \Delta\omega + id \tag{44}$$

where $\Delta\omega_j = \omega_j - \omega_0$ and $\Delta\omega = \omega - \omega_0$; we have again neglected higher order corrections. Making use of Eq. (32), the subdiagonal

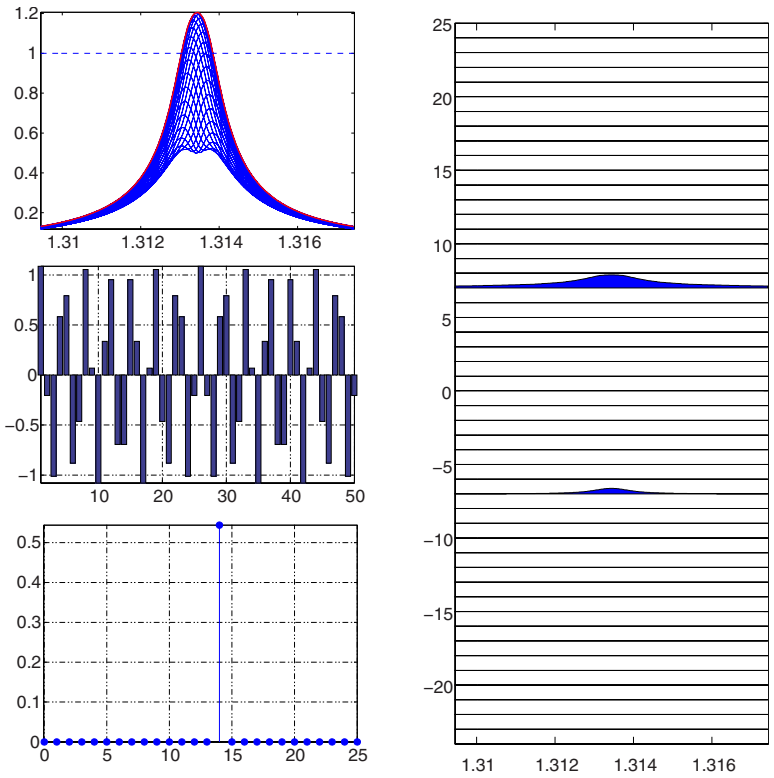


Fig. 4 Response of the system in Fig. 2 to a forcing with engine order 7 (tuned TW frequency $\omega_0=1.3135,\ldots$) and mistuning pattern composed of a single harmonic with wavenumber 14 (AMM prediction for maximum response). Top left: displacements $|x_j|$ versus forcing frequency ($\max|x_j|=1.2066,\ldots$). Middle left: mistuning distribution δ_j. Bottom left: amplitude of the Fourier modes of the mistuning distribution. Right: amplitude of the TW components of the response versus forcing frequency (wavenumber in the vertical axis).

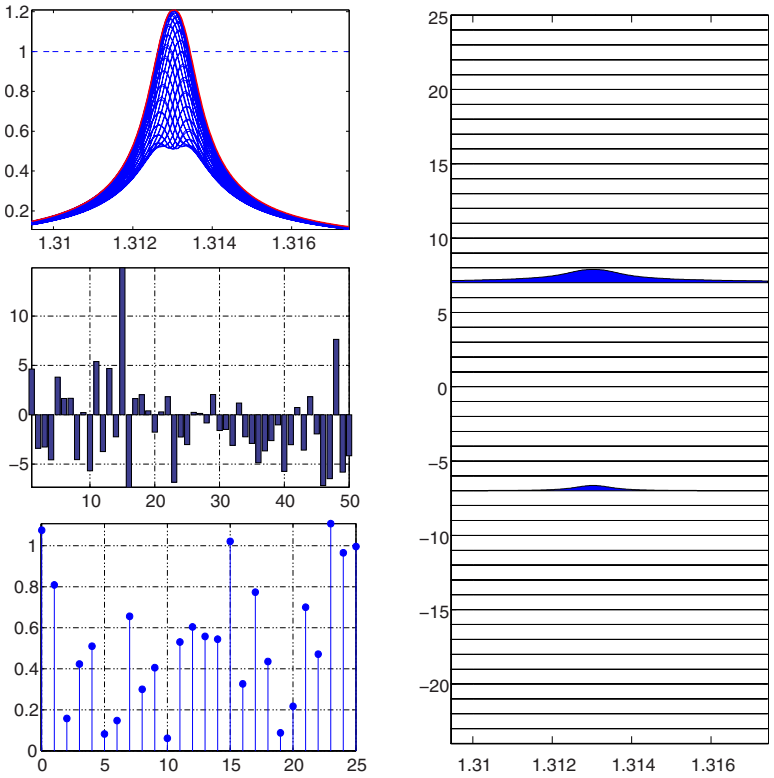


Fig. 5 Same as in Fig. 4 but with the rest of the harmonics added to the mistuning pattern with random amplitudes ($\max|x_j|=1.2085,\ldots$)

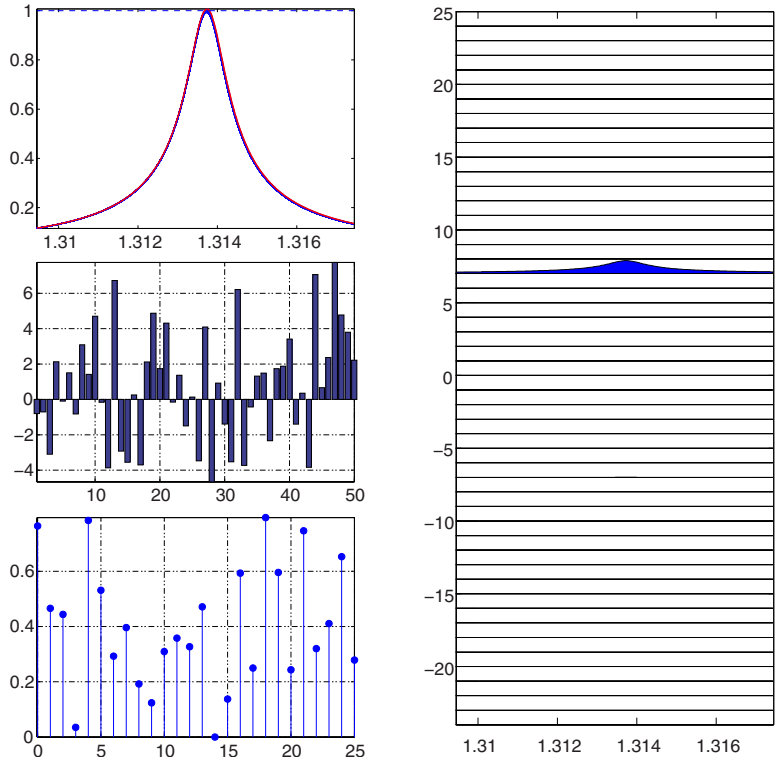


Fig. 6 Same as in Fig. 4 but for a mistuning pattern with zero harmonic with wavenumber 14 and the rest with random amplitudes ($\max|x_j| = 1.0069, \ldots$)

and superdiagonal mistuning terms (D in the matrix above) are readily seen to be

$$D_{jj'} = Z_j^H(\Delta K_{j-j'}^F - \omega_0^2 \Delta M_{j-j'}^F)Z_{j'}/(2\omega_0) \tag{45}$$

and the forcing terms are all zero except that associated with the TW with r nodal diameters.

If one takes into account that in the almost constant frequency region that we are describing, the different mode shapes are approximately equal except for the different interblade phase angle, then $Z_j = Z$ (independent of j and real valued, see Eq. (11)), and the above expression can be further simplified to

$$D_{jj'} = Z^H(\Delta K_{j-j'}^F - \omega_0^2 \Delta M_{j-j'}^F)Z/(2\omega_0) = D_{j-j'} \tag{46}$$

which depends only on the difference $(j-j')$ and allows us to rewrite the AMM as

$$\begin{bmatrix} d_k & D_1 & D_2 & \cdots & D_{N_a-1} \\ \overline{D_1} & d_{k+1} & D_1 & \cdots & D_{N_a-2} \\ & & \ddots & & \\ \overline{D_{N_a-2}} & \cdots & & d_{-(k+1)} & D_1 \\ \overline{D_{N_a-1}} & \cdots & & \overline{D_1} & d_{-k} \end{bmatrix} \begin{bmatrix} A_k \\ A_{k+1} \\ \vdots \\ A_{-(k+1)} \\ A_{-k} \end{bmatrix} = \begin{bmatrix} 0 \\ \vdots \\ id \\ \vdots \\ 0 \end{bmatrix} \tag{47}$$

Note that the only harmonics of the mistuning distribution that play a role in the mistuned forced response are $D_1, D_2, \ldots, D_{N_a-1}$, which couple different active modes.

The AMM is in this case a $N_a \times N_a$ linear system, and in order to compute the maximum mistuned forced response, we have to locate the mistuning coefficients $D_1, D_2, \ldots, D_{N_a-1}$ that maximize the amplitude of the displacement,

$$\text{Amplitude} \propto |A_k + A_{k+1} + \cdots + A_{-(k+1)} + A_{-k}| \tag{48}$$

In this case we have not been able to solve explicitly this problem, but an upper bound for the maximum response can be obtained just by using the same argument as that used by Whitehead [3,22]. If we premultiply Eq. (47) by the row vector $[\bar{A}_k, \bar{A}_{k+1}, \ldots, \bar{A}_{-(k+1)}, \bar{A}_{-k}]$ and substract the complex conjugate, we arrive at the following equation:

$$|A_k|^2 + |A_{k+1}|^2 + \cdots + |A_{-(k+1)}|^2 + |A_{-k}|^2 = \text{Re}(A_r) \tag{49}$$

which simply states that the work of the external forcing is equal to the energy dissipated by the damping. Now we can maximize the amplitude in Eq. (48) subjected only to the work constraint given by Eq. (49). This problem is completely similar to that solved by Whitehead [3,22] but with N chaged by N_a, and therefore the resulting response amplitude is at most

$$\text{Amplitude}_{\max} = \frac{1 + \sqrt{N_a}}{2} \tag{50}$$

Now, in order to locate the mistuning distribution that produces the maximum amplitude, we have to numerically solve the optimization problem given by the simplified AMM formulation, which, due to its reduced size, does not require large scale computational resources. Note that the number of variables in the optimization in the worst case when all family modes are active ($N_a = N$) is N, the number of harmonics of the mistuning distribution (which is real and with average 0) plus the forcing frequency.

The AMM formulation has been used to describe the mistuned forced response of the mass-spring system given by Eq. (33) with a forcing engine order $r=21$ and with a frequency close to the frequency of the clustered modes labeled CM in Fig. 2. We have to take into account also the symmetric TW with a negative wavenumber. Thus the cluster in Fig. 2 actually contains nine TW modes, with wavenumbers 21, 22, 23, 24, 25, −24, −23, −22, and −21.

The AMM indicates that the only relevant harmonics are those that couple the nine active modes, that is, $D_1, \ldots, D_8$ (recall that

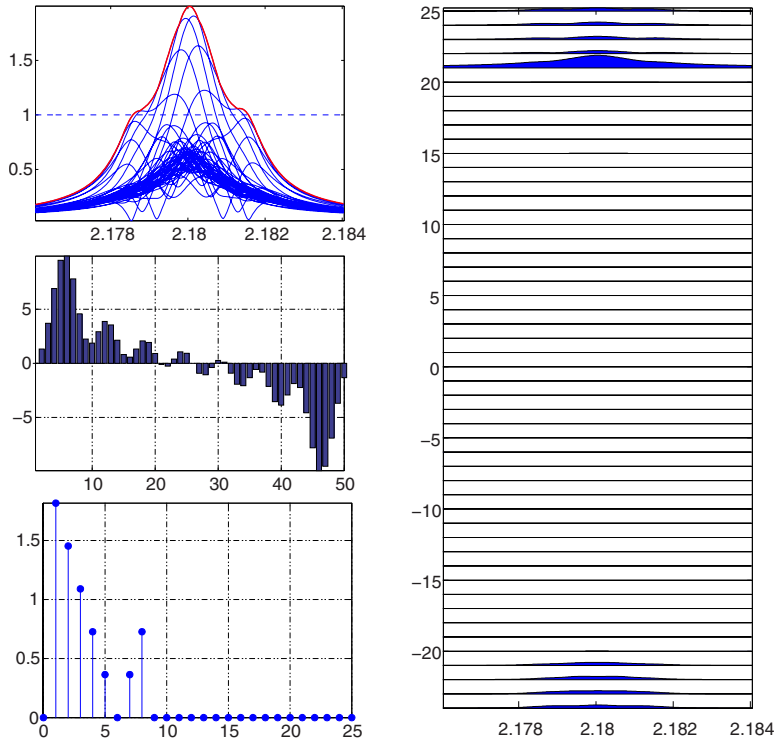


Fig. 7 Response of the system in Fig. 2 to a forcing with engine order 21 (tuned TW frequency $\omega_0=2.1830,\ldots$) and mistuning pattern composed of only harmonics with wavenumbers 1–8 (AMM prediction for a maximum response). Top left: displacements $|x_j|$ versus forcing frequency ($\max|x_j|=1.991,\ldots$). Middle left: mistuning distribution δ_j. Bottom left: amplitude of the Fourier modes of the mistuning distribution. Right: amplitude of the TW components of the response versus forcing frequency (wavenumber in the vertical axis).

$D_i=\bar{D}_{-i}$ because the mistuning is a real pattern). We have solved numerically the AMM optimization problem with a number of active modes $N_a=9$ and the corresponding tuned frequencies from the clustered modes in Fig. 2. The resulting mistuning pattern (with only the harmonics $D_1,\ldots,D_8$) has then been inserted in the full mass-spring model given by Eq. (33), and the response is shown in Fig. 7. The maximum amplitude response is $1.991,\ldots$, in close agreement with the AMM predicted value $(1+\sqrt{9})/2=2$, and the only TW modes of the response that are nonzero are those with wavenumbers $21,22,\ldots,-22,-21$ (the AMM active modes). In order to confirm that the only harmonics of the mistuning distribution that affect the response are those with wavenumbers $1,2,\ldots,8$, we add to the mistuning the rest of the harmonics with random amplitude, and, as can be seen in Fig. 8, the response of the system is practically unaltered. Finally, the response shown in Fig. 9 corresponds to removing only the mistuning harmonics with wavenumbers $1,2,\ldots,8$, and again, as predicted by the AMM, the system simply does not feel the effect of the remaining mistuning harmonics. There is almost no amplification with respect to the tuned system and no coupling among the TWs: the only TW mode of the response that is nonzero is that with wavenumber 21 (the one that is being directly excited by external forcing).

It is also interesting to mention that expression (50) can be also validated with the numerical optimization results presented by Petrov and Ewins [6] for a realistic high pressure turbine bladed disk with $N=92$ sectors. They obtained a maximum amplification factor of 5.02, and that resulting from the Whitehead [3,22] expression is $(1+\sqrt{N})/2=5.30$. From Fig. 10 of Petrov and Ewins [6], it can be estimated that for a sixth engine order excitation the number of active modes is $N_a=83$ (all modes in the family with five or more nodal diameters), and using Eq. (50) the resulting amplification factor is $(1+\sqrt{N_a})/2=5.06$, which gives a much better agreement with the full bladed disk numerical computations.

4 Concluding Remarks

We have derived a new reduced order model, the AMM, for the description of the forced response of a mistuned bladed disk. The AMM has been previously applied to the study of the stabilizing effect of intentional mistuning on the stability characteristics of aerodynamically unstable rotors [18].

The AMM captures the essential ingredients that are involved in the mechanism of mistuned response amplification and can be regarded as an extension of the FMM [15–17] that allows us to describe the dynamics associated with more general modal families (the FMM can only account for modal families with nearly identical frequencies (see Ref. [16]).

The AMM has been systematically derived from the general equations for the forced response of a mistuned rotor under the only assumption of small mistuning and damping. The AMM has been applied to describe the effect of mistuning in two frequent situations: forced response of a pair of isolated modes and forced response of a group of modes with close frequencies. In the first case the AMM can be completely analyzed, and closed form expressions have been derived for the maximum amplification factor and for its associated mistuning distribution. In the second case the problem has to be solved numerically, but the size of the AMM is much smaller than that of the original problem (N variables at most). In both cases we have found an excellent quantitative agreement between the AMM results and the numerical simulation of the forced response of the mistuned mass-spring system given by Eq. (33).

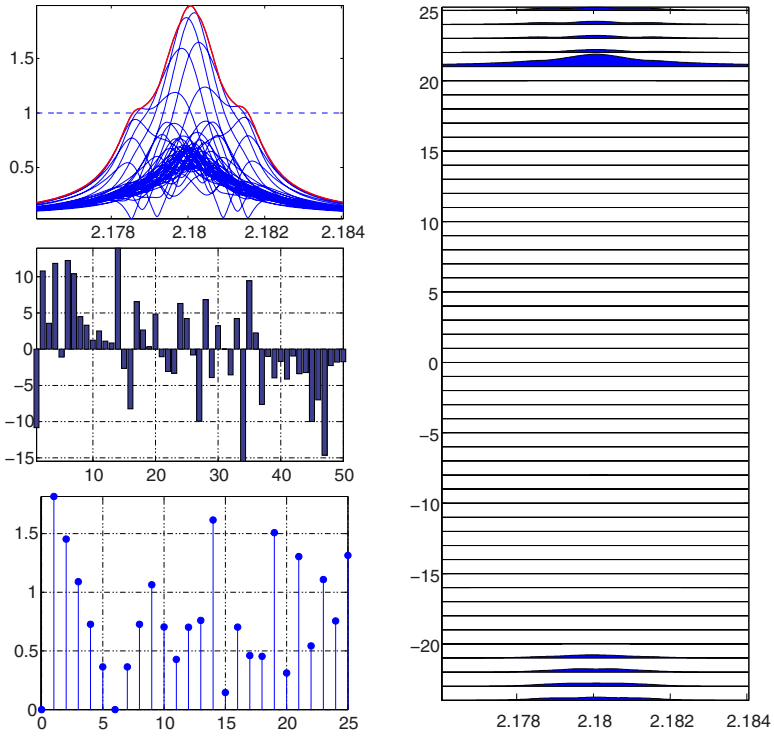


Fig. 8 Same as in Fig. 7 but with the rest of the harmonics added to the mistuning pattern with random amplitudes (max$|x_j|$=1.9851,…)

The application of the AMM allows us to draw the following final remarks about the role of mistuning in the near resonance forced response of a bladed disk:

1. The mistuned forced response of the system is composed of only those tuned TW modes that have frequencies that are close to the resonant one (the active modes).
2. The effect of the small mistuning is to couple these active TW modes. The coupling between the TWs with wavenum-

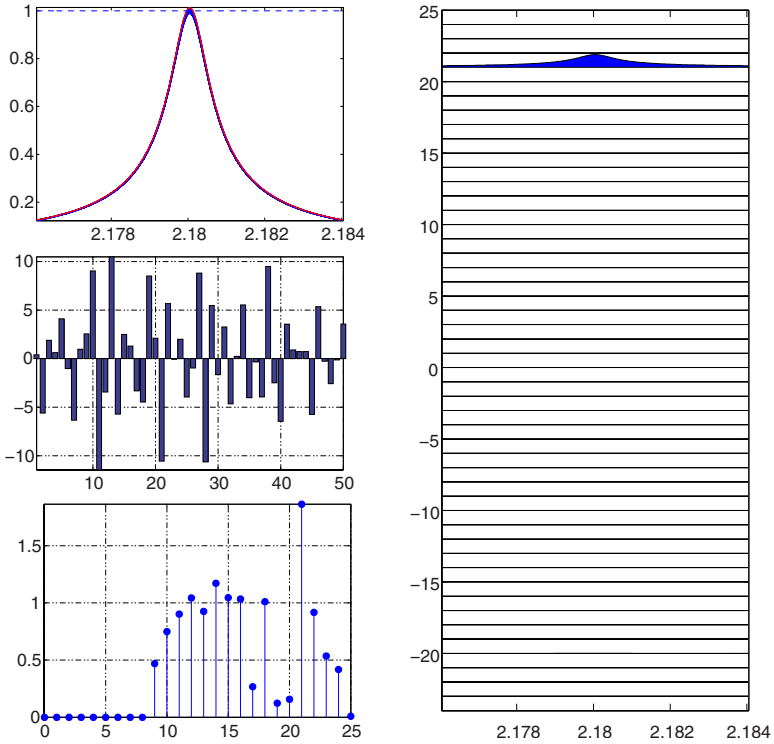


Fig. 9 Same as in Fig. 7 but for a mistuning pattern with zero harmonics with wavenumbers 1–8 and the rest with random amplitudes (max$|x_j|$ =1.0141,…)

bers k and k' takes place through the harmonic of the mistuning distribution with wavenumber $(k-k')$. The harmonics of the mistuning distribution that have an effect on the response of the system are only those that couple active modes.

3. The maximum amplification is bounded by

$$\text{Amplitude}_{\max} = \frac{1+\sqrt{N_a}}{2} \tag{51}$$

where N_a is the number of active modes. This constitutes further refinement of the upper bound obtained by Whitehead [3,22], and it is valid for modal families exhibiting a substantial variation of the natural frequencies. The typical size of the mistuning distribution that gives the maximum amplification is small, of the order of the damping, and depends on the details of the frequency distribution of the active modes.

4. The AMM requires that we assume a frequency difference among the active modes of the order of the damping or smaller and a large frequency difference between the rest of the tuned modes and the active modes, as compared with the damping. If the separation between the frequencies of the active modes and the rest of the modes is not clear but gradual (in the scale of the damping), then the maximum amplification curve does not exhibit a nice isolated maximum like that in Fig. 3, but the AMM can still be applied to determine the slope and the beginning of the amplification curve if enough modes are retained as active modes.

Acknowledgment

The authors wish to thank ITP for the permission to publish this paper and for its support during the project. The work of C.M. has also been supported by the Spanish Ministerio de Educación y Ciencia under Grant No. MTM2007-62428 and by the Universidad Politécnica de Madrid under Grant No. CCG07-UPM/000-3177. The work of R.C. has also been supported by the Spanish Ministerio de Educación y Ciencia under Grant No. TRA2006-15015/TAIR.

Nomenclature

- $\mathbf{M}$ = mass matrix
- $\mathbf{K}$ = stiffness matrix
- $\mathbf{F}$ = external forces
- $\mathbf{X}$ = DOF displacement vector
- N,m = number of sectors, number of DOFs per sector
- Z_{kj} = in-sector tuned eigenmode
- $\mathbf{P}$ = change in basis matrix from TW to displacements
- $\mathbf{A}$ = traveling wave mode amplitude vector
- $\Delta\mathbf{M}$ = mass mistuning matrix
- $\Delta\mathbf{K}$ = stiffness mistuning matrix
- r,ω = engine order and frequency of the forcing
- ω_0 = tuned natural frequency of the excited TW (r nodal diameters)
- d = damping of the tuned natural TW modes
- N_a = number of active modes
- ω_a = active mode tuned natural frequency
- Δ = mistuning matrix in the TW basis
- D,D_{ij} = mistuning coefficients in the TW basis
- ω_p,ω_c = spring frequencies of the model in Fig. 2
- ε = mistuning size of the model in Fig. 2
- δ_j = mistuning distribution of the model in Fig. 2
- c = damping of the model in Fig. 2
- f = forcing amplitude of the model in Fig. 2

References

[1] Slater, J., Minkiewicz, R., and Blair, A., 1999, "Forced Response of Bladed Disk Assemblies: A Survey," Shock Vib. Dig., **31**, pp. 17–24.
[2] Castanier, M., and Pierre, C., 2006, "Modeling and Analysis of Mistuned Bladed Disk Status and Emerging Directions," J. Propul. Power, **22**(2), pp. 384–396.
[3] Whitehead, D., 1966, "Effect of Mistuning on the Vibration of Turbomachine Blades Induced by Wakes," J. Mech. Eng. Sci., **8**(1), pp. 15–21.
[4] Han, Y., and Mignolet, M., 2007, "A Novel Perturbation-Based Approach for the Accurate Prediction of the Forced Response of Mistuned Bladed Disks," ASME Paper No. GT2007-27352.
[5] Han, Y., Xiao, B., and Mignolet, M., 2007, "Expedient Estimation of the Maximum Amplification Factor in Damped Mistuned Bladed Disks," ASME Paper No. GT2007-27353.
[6] Petrov, E., and Ewins, D., 2003, "Analysis of the Worst Mistuning Patterns in Bladed Disk Assemblies," J. Turbomach., **125**, pp. 623–631.
[7] Ewins, D., 1969, "The Effects of Detuning Upon the Forced Vibrations of Bladed-Disks," J. Sound Vib., **9**(1), pp. 65–79.
[8] Sinha, A., 1997, "Computation of the Maximum Amplitude of a Mistuned Bladed Disk Assembly via Infinity Norm," *Proceedings of the 1997 ASME International Mechanical Engineering Congress and Exposition*, ASME, New York, Vol. AD-55, pp. 427–432.
[9] Kenyon, J., and Griffin, J., 2003, "Experimental Demonstration of Maximum Mistuned Bladed Disk Forced Response," ASME J. Turbomach., **125**, pp. 673–681.
[10] Kenyon, J., Griffin, J., and Feiner, D., 2003, "Maximum Bladed Disk Forced Response From Distortion of a Structural Mode," J. Turbomach., **125**, pp. 352–363.
[11] Yang, M., and Griffin, J., 1997, "A Reduced Order Approach for the Vibration of Mistuned Bladed Disk Assemblies," ASME J. Eng. Gas Turbines Power, **119**, pp. 161–167.
[12] Castanier, M., Ottarson, G., and Pierre, C., 1997, "A Reduced Order Modeling Technique for Mistuned Bladed-Disks," J. Vibr. Acoust., **119**(3), pp. 439–447.
[13] Yang, M., and Griffin, J., 2001, "A Reduced Order Model of Mistuning Using a Subset of Nominal Modes," ASME J. Eng. Gas Turbines Power, **123**, pp. 893–900.
[14] Rivas-Guerra, A., and Mignolet, M., 2003, "Maximum Amplification of Blade Response Due to Mistuning: Localization and Mode Shape Aspects of the Worst Disks," ASME J. Turbomach., **125**, pp. 442–454.
[15] Feiner, D., and Griffin, J., 2002, "A Fundamental Model of Mistuning for a Single Family of Modes," ASME J. Turbomach., **124**, pp. 597–605.
[16] Feiner, D., and Griffin, J., 2004, "Mistuning Identification of Bladed Disks Using Fundamental Mistuning Model—Part I: Theory," ASME J. Turbomach., **126**, pp. 150–158.
[17] Feiner, D., and Griffin, J., 2004, "Mistuning Identification of Bladed Disks Using Fundamental Mistuning Model—Part II: Application," ASME J. Turbomach., **126**, pp. 159–165.
[18] Martel, C., Corral, R., and Llorens, J., 2008, "Stability Increase of Aerodynamically Unstable Rotors Using Intentional Mistuning," ASME J. Turbomach., **130**, p. 011006.
[19] Wilkinson, J. H., 1988, "The Algebraic Eigenvalue Problem," *Monographs on Numerical Analysis*, Oxford Science Publications, The Clarendon Press Oxford University Press, New York.
[20] MacBain, J., and Whaley, P., 1984, "Maximum Resonant Response Mistuned Bladed Disks," ASME J. Vib., Acoust., Stress, Reliab. Des., **106**, pp. 218–223.
[21] Kenyon, J., and Griffin, J., 2003, "Forced Response of Turbine Engine Bladed Disks and Sensitivity to Harmonic Mistuning," ASME J. Turbomach., **125**, pp. 113–120.
[22] Whitehead, D., 1998, "The Maximum Factor by Which Forced Vibration of Blades Can Increase Due to Mistuning," ASME J. Turbomach., **120**, pp. 115–119.

On the Use of Actively Controlled Auxiliary Bearings in Magnetic Bearing Systems

Iain S. Cade

M. Necip Sahinkaya

Clifford R. Burrows

Patrick S. Keogh

Department of Mechanical Engineering,
University of Bath,
Bath BA2 7AY, UK

Auxiliary bearings are used to prevent rotor/stator contact in active magnetic bearing systems. They are sacrificial components providing a physical limit on the rotor displacement. During rotor/auxiliary bearing contact significant forces normal to the contact zone may occur. Furthermore, rotor slip and rub can lead to localized frictional heating. Linear control strategies may also become ineffective or induce instability due to changes in rotordynamic characteristics during contact periods. This work considers the concept of using actively controlled auxiliary bearings in magnetic bearing systems. Auxiliary bearing controller design is focused on attenuating bearing vibration resulting from contact and reducing the contact forces. Controller optimization is based on the H_∞ norm with appropriate weighting functions applied to the error and control signals. The controller is assessed using a simulated rotor/magnetic bearing system. Comparison of the performance of an actively controlled auxiliary bearing is made with that of a resiliently mounted auxiliary bearing. Rotor drop tests, repeated contact tests, and sudden rotor unbalance resulting in trapped contact modes are considered. [DOI: 10.1115/1.2982159]

1 Introduction

Rotor/stator contact in magnetic bearings is a significant fault condition and can permanently damage the system. To prevent rotor/stator interaction it is common to incorporate sacrificial auxiliary/backup bearings. It is therefore possible, during a disturbance or fault condition, that a rotor will make contact with an auxiliary bearing. During rotor/auxiliary bearing contact significant forces may arise; these forces can also lead to significant thermal stresses due to high slip and rubbing.

Rotor/auxiliary bearing contact dynamics have received much attention in the open literature. Responses of a rigid rotor in two sliding retainer bearings have been predicted by Fumagalli et al. [1]. Rotor drop transients have been predicted using a theoretical model by Ishii and Kirk [2]. Muszynska and Goldman [3] identified chaotic rotor vibration within a rotor/bearing/stator system. Subharmonic vibrations have been identified by Ehrich [4] and Childs [5]. Experimental validation of a flexible rotor model interacting with auxiliary bearings is presented by Lawen and Flowers [6]. Efficient modeling of rotor/bearing contact conditions has been achieved using a constrained Lagrange approach [7]. Thermal stresses arising from rotor contact and rub/slip have been predicted by Keogh and Yong [8], while backward whirl problems have been discussed by Bartha [9].

Contact forces and rotor dynamic responses are dictated by physical parameters, such as radial clearance, bearing stiffness/damping, and initial contact conditions. Variations of these parameters can have significant effect in reducing contact forces or the response dynamics. However, reducing the auxiliary bearing stiffness in order to reduce contact forces may increase contact time and therefore thermal stresses. In principle, it is possible to mount the auxiliary bearing using an active system and to allow for a dynamically varying clearance gap. The contact forces may also be influenced in a controllable manner.

Direct rotor control has been achieved by Palazzolo et al. [10,11] using rolling element bearings and piezoelectric actuators. A hybrid piezoelectric/hydraulic actuator system designed to minimize rotor vibration has further been presented by Tang et al. [12]. Ulbrich et al. [13], Chavez et al. [14], and Jiang et al. [15] demonstrated a reduction in rotor/auxiliary bearing contact forces in studies using an electrically actuated auxiliary bearing. However, in this system the rotor is still radially supported by a conventional bearing.

This work considers the concept of using an active auxiliary bearing operating with an active magnetic bearing system. The effect of the active auxiliary bearing on reducing contact forces is investigated in a simulation study. Such a study is necessary to assess whether the added complexity of active auxiliary bearings outweighs the benefits of extended bearing life. If the benefits are positive, wider applications of magnetic bearings may be possible. An H_∞ control strategy, based on the work by Glover and Doyle [16], is investigated. The effectiveness of the controller is assessed in terms of rotor/bearing contact force, contact duration, and the rotor response after initial contact. Comparison of response parameters is made against those from a resiliently mounted auxiliary bearing.

2 Rotor/Auxiliary Bearing Contact Problem

A discretized rotor can be modeled in terms of a vector of lateral and rotational displacements, $\mathbf{x}_r(t)$, corresponding to different nodal positions along its shaft [17]. Corresponding disturbance and control forces acting on the rotor are denoted by $\mathbf{f}(t)$ and $\mathbf{u}(t)$, respectively. The dynamics of the rotor can be evaluated from

$$\mathbf{M}_r\ddot{\mathbf{x}}_r + (\Omega\mathbf{G}_r + \mathbf{C}_r)\dot{\mathbf{x}}_r + \mathbf{K}_r\mathbf{x}_r = \mathbf{B}_f\mathbf{f} + \mathbf{B}_u\mathbf{u} \tag{1}$$

where $\mathbf{M}_r$, $\mathbf{C}_r$, $\mathbf{K}_r$, and $\mathbf{G}_r$ are the mass, damping, stiffness, and gyroscopic matrices, respectively. Mapping matrices, $\mathbf{B}_f$ and $\mathbf{B}_u$, are used to match the dimensions of $\mathbf{f}$ and $\mathbf{u}$ to that of $\mathbf{x}_r$. When the rotor displacement is equal to the radial clearance, contact occurs between the rotor and auxiliary bearing (Fig. 1). The dynamics of the rotor are therefore governed by both the disturbance, control, and contact forces. Equation (1) can be modified to include contact forces as

$$\mathbf{M}_r\ddot{\mathbf{x}}_r + (\Omega\mathbf{G}_r + \mathbf{C}_r)\dot{\mathbf{x}}_r + \mathbf{K}_r\mathbf{x}_r = \mathbf{B}_f\mathbf{f} + \mathbf{B}_u\mathbf{u} + \mathbf{B}_c\mathbf{d}_c \tag{2}$$

where $\mathbf{d}_c$ is the contact force vector and $\mathbf{B}_c$ maps the dimension of $\mathbf{d}_c$ to that of $\mathbf{x}_r$. A generalized auxiliary bearing contact event can be modeled by specifying a function of the form $\mathbf{d}_c = \mathbf{Z}_c(\mathbf{x}_r$

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 8, 2008; final manuscript received July 24, 2008; published online December 24, 2008. Review conducted by Dilip R. Ballal.

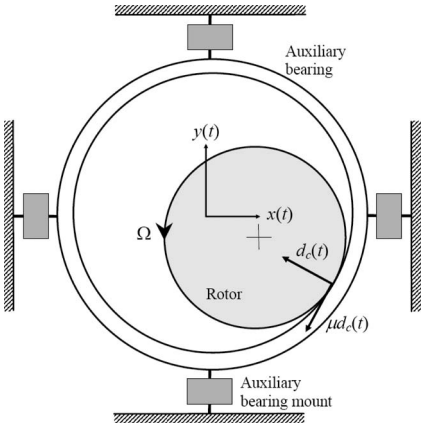


Fig. 1 Rotor/auxiliary bearing contact forces

$-\mathbf{x}_b, \dot{\mathbf{x}}_r - \dot{\mathbf{x}}_b$), relating rotor and bearing positions, $\mathbf{x}_r$ and $\mathbf{x}_b$, and velocities, $\dot{\mathbf{x}}_r$ and $\dot{\mathbf{x}}_b$, to the contact force. Equation (2) now becomes

$$\mathbf{M}_r\ddot{\mathbf{x}} + (\Omega\mathbf{G}_r + \mathbf{C}_r)\dot{\mathbf{x}}_r + \mathbf{K}_r\mathbf{x}_r - \mathbf{B}_c\mathbf{Z}_c(\mathbf{x}_r - \mathbf{x}_b, \dot{\mathbf{x}}_r - \dot{\mathbf{x}}_b) = \mathbf{B}_f\mathbf{f} + \mathbf{B}_u\mathbf{u} \tag{3}$$

Consider a single contact plane in which the rotor and auxiliary bearing displacements are (x_{ri}, y_{ri}) and (x_{bi}, y_{bi}), respectively. Contact will occur whenever

$$\delta r_i = \sqrt{(x_{ri} - x_{bi})^2 + (y_{ri} - y_{bi})^2} \geq c_{ri} \tag{4}$$

where c_{ri} is the local rotor/auxiliary bearing radial clearance. A contact force/deflection relation may be derived making the assumption of a Hertzian normal stress distribution over the contact zone. The appropriate rotor linear expression is [18]

$$\delta r_i - c_{ri} = \frac{2d_{ci}(1-\nu^2)}{\pi E l_{bi}}\left(\frac{2}{3} + \ln\frac{16R_{ri}R_{bi}}{l_{ci}^2}\right) \tag{5}$$

where d_{ci} is the normal contact force. Also, R_{ri} and R_{bi} are the rotor and inner surface bearing radii and l_{bi} is the auxiliary bearing axial length. The contact zone arclength then follows from

$$l_{ci} = 2.15\sqrt{\frac{2R_{ri}R_{bi}d_{ci}}{El_{bi}c_{ri}}} \tag{6}$$

It is also taken that the rotor and auxiliary bearing materials have similar Young's modulus, E, and Poisson ratio ν.

The contact force d_{ci} may be extracted using an appropriate numerical inverse procedure. It is also possible to consider more complex contact models, including those with damping effects, and dissimilar rotor and auxiliary bearing materials. However, the use of Eqs. (4)–(6) was considered to be appropriate for the present analysis.

The determination of x_{bi} will depend on the auxiliary bearing model and whether the active or passive case is being considered. For a passively mounted bearing, isotropic mounting stiffness k_i and damping c_i are commonly used. For an active auxiliary bearing, the actuation forces will be dependent on the chosen feedback law.

3 Rotor/Magnetic Bearing and Active Auxiliary Bearing System

For this study a rotor/magnetic bearing system (Fig. 2) consisting of two magnetic bearings with two active auxiliary bearings was simulated. The rotor (with magnetic bearing cores) is 0.6 m long with a diameter of 2 cm and a mass of 8.3 kg. The rotor is supported by two active magnetic bearings with a maximum force capability of 1.5 kN before roll-off at 185 Hz. The active magnetic

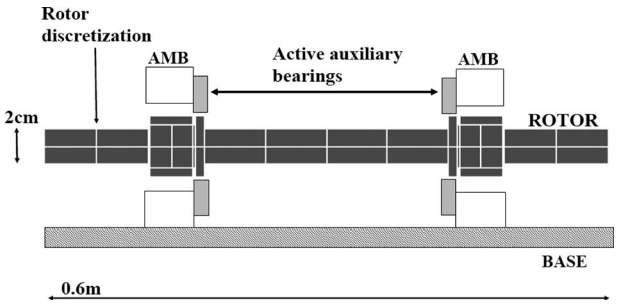


Fig. 2 Rotor/magnetic bearing/active auxiliary bearing system

bearing negative stiffness is 3.9 MN/m. Local proportional-integral-differential (PID) control was used to stabilize the rotor with controller gains $k_{Mp}=5.1$ MN/m, $k_{Mi}=0.0103$ MN s m, and $k_{Md}=0.103$ MN s/m. The rotor, under PID control, has the first four zero running speed natural frequencies occurring at 45 rad/s (mainly rigid body), 50 rad/s (mainly rigid body), 3180 rad/s (first flexible mode), and 6396 rad/s (second flexible mode) with corresponding damping ratios of 0.89, 0.92, 0.0035, and 0.071, respectively.

The active auxiliary bearings were modeled as steel bush types with a mass of 0.1 kg. The simulated auxiliary bearing actuators could provide a maximum force of 12.5 kN over a maximum operating displacement of 100 μm up to 1 kHz before roll-off. This is feasible using a piezo based actuation system. The rotor-auxiliary bearing clearance was 400 μm, and the coefficient of dry sliding friction between the rotor and bearing was set to $\mu = 0.1$. For rolling element bearings the acceleration of rollers/cage/race would need to be accounted for. In effect the friction will vary during a contact event as the race speed starts from zero and eventually attains the rolling speed of the rotor.

A recent area of controller design includes a linear matrix inequality formulation based on H_2/H_∞ norm bounds of the frequency response. The principles of formulating a state-space H_∞ controller for generalized systems are given by Glover and Doyle [16]. An H_∞ approach to controller design allows loop shaping of the input disturbances and output signals and a robust approach to both modeling error and uncertainty. An initial PID controller, with gains $k_{Ap}=1$ MN/m, $k_{Ai}=0.01$ MN m s, and $k_{Ad}=0.1$ MN m/s, was applied to the auxiliary bearing system to provide an initially stable system with feasible stiffness and damping characteristics. An H_∞ controller may then be designed to control the combined auxiliary bearing/PID system. This is a reasonable consideration when designing controllers for practical testing and allows the H_∞ controller to be switched off without causing instability.

In this study a comparison is made between an actively controlled auxiliary bearing and a passive auxiliary bearing. The passive bearing is modeled as a compliant element with stiffness and damping values of 100 MN/m and 1 kN s/m, respectively. The mass of the compliantly mounted auxiliary bearing is 0.1 kg. Figure 3 shows the static and dynamic clearance cases associated with auxiliary bearing deflection. For the static condition a circular limit applies. The active auxiliary bearing can displace; therefore the clearance gap can also vary, as shown in Fig. 3 by the shaded region.

4 Active Auxiliary Bearing Control

4.1 Auxiliary Bearing Input Disturbances. During operation the auxiliary bearing may be subject to both direct and indirect forces. Direct disturbance forces are primarily associated with rotor/bearing contact. Indirect forces may be transmitted through the bearing mounts during base motion.

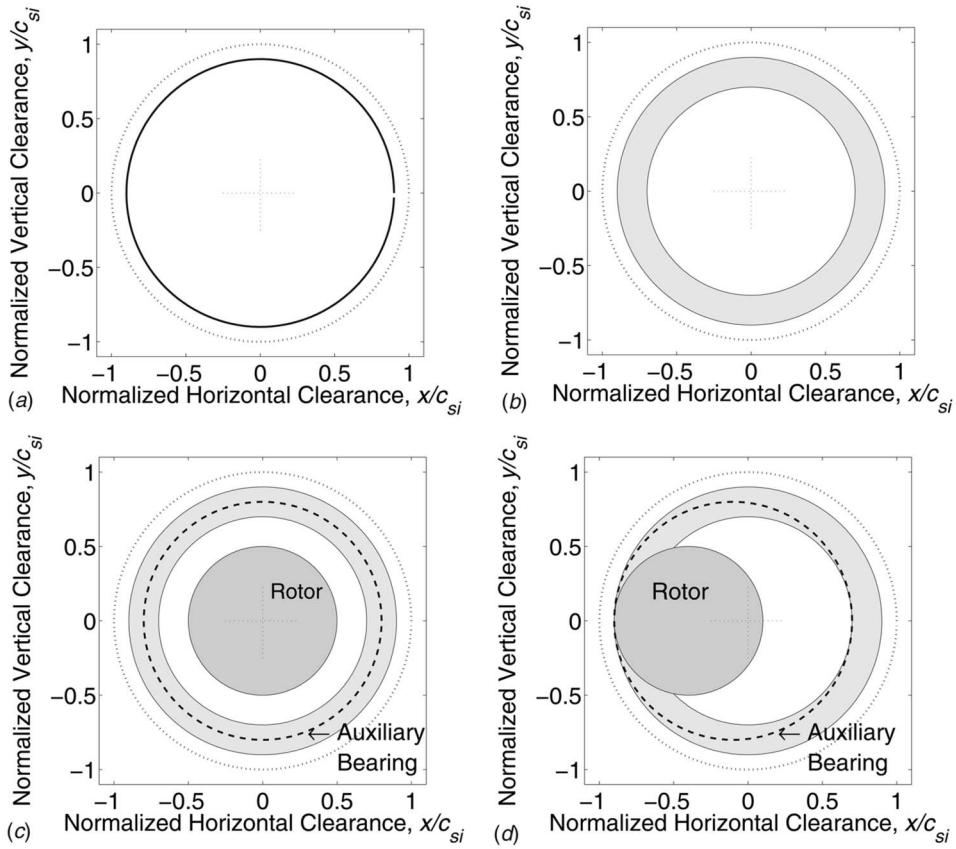


Fig. 3 (*a*) Static auxiliary bearing clearance circle and (*b*) dynamic auxiliary bearing contact region. ((*c*) and (*d*)) The active auxiliary bearing position with and without contact, respectively. Values normalized with respect to the rotor/stator clearance gap.

4.2 Closed Loop Response. The disturbance forces acting on the auxiliary bearing and corresponding equation of motion can be expressed as

$$\mathbf{F}_b = \mathbf{d}_c + \mathbf{z}_t + \mathbf{u}_b$$

$$\mathbf{M}_b\ddot{\mathbf{x}}_b + \mathbf{C}_b\dot{\mathbf{x}}_b + \mathbf{K}_b\mathbf{x}_b = \mathbf{F}_b \tag{7}$$

$\mathbf{M}_b$, $\mathbf{C}_b$, and $\mathbf{K}_b$ represent the auxiliary bearing mass, damping, and stiffness matrices, respectively. The transmitted and control forces are denoted by $\mathbf{z}_t$ and $\mathbf{u}_b$, respectively. It therefore follows that the dynamics of the auxiliary bearing can be expressed in state-space form as

$$\dot{\mathbf{q}}_b = \mathbf{A}_b\mathbf{q}_b + \mathbf{B}_c\mathbf{d}_c + \mathbf{B}_t\mathbf{z}_t + \mathbf{B}_b\mathbf{u}_b$$

$$\mathbf{y}_b = \mathbf{C}_b\mathbf{q}_b + \mathbf{D}_b\mathbf{z}_t \tag{8}$$

where $\mathbf{q}_b=[\mathbf{x}_b^T \ \ \dot{\mathbf{x}}_b^T]^T$. Here the transmitted force $\mathbf{z}_t$ relates to base motion and can be evaluated from $\mathbf{z}_t=-\mathbf{M}_b\ddot{\mathbf{Y}}$, where $\mathbf{Y}$ is a base displacement vector. $\mathbf{B}_c$, $\mathbf{B}_t$, and $\mathbf{B}_b$ are matrices used to map the dimensions of $\mathbf{d}_c$, $\mathbf{z}_t$, and $\mathbf{u}_b$ to $\mathbf{q}_b$, respectively.

Equation (8) can be transformed into Laplace domain where the $\mathcal{L}[\mathbf{q}_b(t)]=\mathbf{Q}_b(s)$, $\mathcal{L}[\mathbf{y}_b(t)]=\mathbf{Y}_b(s)$, $\mathcal{L}[\mathbf{d}_c(t)]=\mathbf{D}_c(s)$, $\mathcal{L}[\mathbf{z}_t(t)]=\mathbf{Z}_t(s)$,

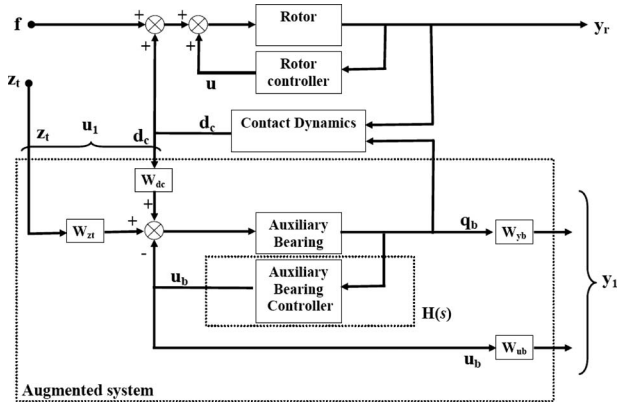


Fig. 4 Augmented rotor/auxiliary bearing system block diagram including weighting transfer functions

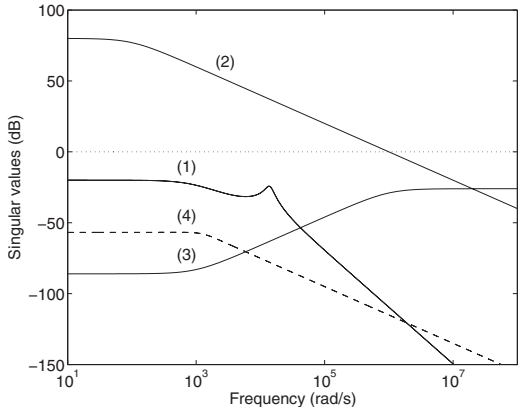


Fig. 5 (1) PID controlled auxiliary bearing singular value response, (2) $\bar{\sigma}[\mathbf{W}_{yb}(i\omega)]$, (3) $\bar{\sigma}[\mathbf{W}_{ub}(i\omega)]$, and (4) active auxiliary bearing system closed loop response

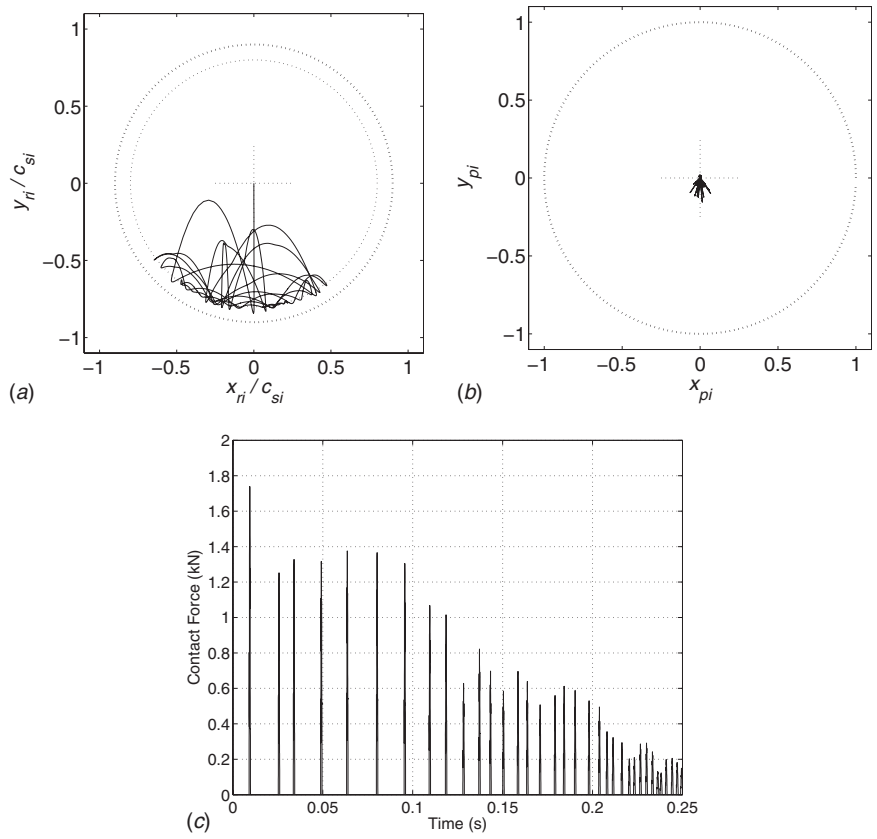


Fig. 6 (*a*) **Rotor displacements normalized with respect to rotor/stator clearance, c_{si}, for a passive auxiliary bearing.** (*b*) **Rotor deviation beyond clearance circle (x_{pi}, y_{pi}).** (*c*) **Rotor/bearing contact forces. Rotor operating speed $\Omega = 1000$ rad/s.**

and $\mathcal{L}[\mathbf{u}_b(t)] = \mathbf{U}_b(s)$. The Laplace transforms of the input disturbances, $\mathbf{D}_c(s)$ and $\mathbf{Z}_t(s)$, and control forces, $\mathbf{U}_b(s)$, can now be related to the system output $\mathbf{Y}_b(s)$ by eliminating $\mathbf{Q}_b(s)$ to yield

$$\mathbf{Y}_b(s) = \mathbf{G}_c(s)\mathbf{D}_c(s) + \mathbf{G}_t(s)\mathbf{Z}_t(s) + \mathbf{G}_b(s)\mathbf{U}_b(s) \tag{9}$$

where

$$\begin{aligned} \mathbf{G}_c(s) &= \mathbf{C}_b(s\mathbf{I} - \mathbf{A}_b)^{-1}\mathbf{B}_c \\ \mathbf{G}_t(s) &= \mathbf{C}_b(s\mathbf{I} - \mathbf{A}_b)^{-1}\mathbf{B}_t \\ \mathbf{G}_b(s) &= \mathbf{C}_b(s\mathbf{I} - \mathbf{A}_b)^{-1}\mathbf{B}_b + \mathbf{D}_b \end{aligned} \tag{10}$$

4.3 Multi-Input Multi-Output Formulation. Controller design is dependent on the control states used. For the control of an active auxiliary bearing it is important for the bearing to return to a central position after contact and to reduce the contact forces to acceptable levels, preferably zero as in contact free levitation. Practical implementation of the system relates auxiliary bearing displacement to the control force as

$$\mathbf{U}_b(s) = -\mathbf{H}_b(s)\mathbf{Q}_b(s) \tag{11}$$

where $\mathbf{H}_b(s)$ is the controller matrix transfer function. A multi-input multi-output system (Fig. 4) can be specified in the form

$$\begin{bmatrix} \mathbf{Q}_b(s) \\ \mathbf{U}_b(s) \end{bmatrix} = \begin{bmatrix} \mathbf{T}_{yz}(s) & \mathbf{T}_{yd}(s) \\ \mathbf{T}_{uz}(s) & \mathbf{T}_{ud}(s) \end{bmatrix} \begin{bmatrix} \mathbf{Z}_t(s) \\ \mathbf{D}_c(s) \end{bmatrix} \tag{12}$$

given that

$$\begin{aligned} \mathbf{T}_{yz} &= (\mathbf{I} + \mathbf{G}_c\mathbf{H}_b)^{-1}\mathbf{G}_b \\ \mathbf{T}_{yd} &= \mathbf{H}_b(\mathbf{I} + \mathbf{G}_c\mathbf{H}_b)^{-1}\mathbf{G}_t \\ \mathbf{T}_{uz} &= \mathbf{H}_b(\mathbf{I} + \mathbf{G}_b\mathbf{H}_b)^{-1}\mathbf{G}_b \\ \mathbf{T}_{ud} &= \mathbf{H}_b(\mathbf{I} + \mathbf{G}_b\mathbf{H}_b)^{-1}\mathbf{G}_t \end{aligned} \tag{13}$$

The input-output relations in Eq. (13) can be modified using weighting transfer functions to influence the control forces. This forms an augmented closed loop system (Fig. 4) of the form

$$\mathbf{W}_{\text{out}}(s)\begin{bmatrix} \mathbf{Q}_b(s) \\ \mathbf{U}_b(s) \end{bmatrix} = \mathbf{T}(s)\mathbf{W}_{\text{in}}(s)\begin{bmatrix} \mathbf{Z}_t(s) \\ \mathbf{D}_c(s) \end{bmatrix} \tag{14}$$

where appropriate diagonal forms are

$$\begin{aligned} \mathbf{W}_{\text{out}}(s) &= \begin{bmatrix} \mathbf{W}_{yb}(s) & 0 \\ 0 & \mathbf{W}_{ub}(s) \end{bmatrix} \\ \mathbf{W}_{\text{in}}(s) &= \begin{bmatrix} \mathbf{W}_{zt}(s) & 0 \\ 0 & \mathbf{W}_{dc}(s) \end{bmatrix} \end{aligned} \tag{15}$$

and

$$\mathbf{T}(s) = \begin{bmatrix} \mathbf{W}_{yb}(s)\mathbf{T}_{yz}(s)\mathbf{W}_{zt}^{-1}(s) & \mathbf{W}_{yb}(s)\mathbf{T}_{yd}(s)\mathbf{W}_{dc}^{-1}(s) \\ \mathbf{W}_{ub}(s)\mathbf{T}_{uz}(s)\mathbf{W}_{zt}^{-1}(s) & \mathbf{W}_{ub}(s)\mathbf{T}_{ud}(s)\mathbf{W}_{dc}^{-1}(s) \end{bmatrix} \tag{16}$$

The outputs from the closed loop system are the bearing states, $\mathbf{Q}_b$, and the control forces, $\mathbf{U}_b$. In the case of successful restoration of contact free levitation the control forces, together with the contact forces, will return to zero. Hence, minimization of $\mathbf{U}_b$ is a favorable control action.

The objective is to find a controller, $\mathbf{H}_b(s)$, such that the H_∞ norm of the augmented system transfer function satisfies $\|\mathbf{T}\|_\infty < 1$. If a suitable controller is found then the auxiliary bearing vibration and control forces should be attenuated in a least-

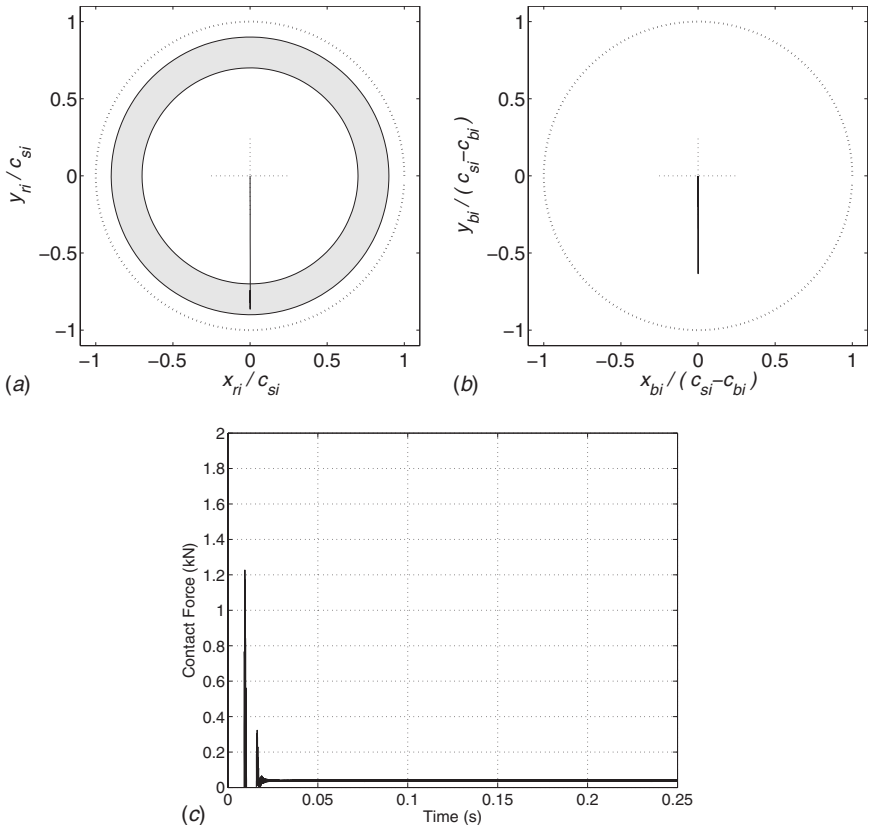


Fig. 7 (*a*) **Rotor displacements normalized with respect to rotor/stator clearance, c_{si}, for an active auxiliary bearing.** (*b*) **Active auxiliary bearing displacement $(x_{bi}, y_{bi})/(c_{si}-c_{bi})$.** (*c*) **Rotor/bearing contact forces. Rotor operating speed Ω =1000 rad/s.**

squares sense. The choice of controller, $\mathbf{H}_b(s)$, is not unique, and variations of the weighting transfer functions allow shaping of the singular value frequency response of the system, $\bar{\sigma}[\mathbf{T}(i\omega)]$. Hence the weighting functions may be used to influence the closed loop control action in the frequency domain.

4.4 $\mathbf{H}_\infty$ Controller Weightings. In order to specify suitable weighting functions it is important to consider how they effect the system dynamics. During rotor bearing contact the auxiliary bearing control forces will act in reaction to the contact force. Reducing the control force will therefore have an effect on reducing the contact force. Specifying $\mathbf{W}_{zt}=\mathbf{W}_{dc}=\mathbf{I}$ implies that there is no requirement to shape the input disturbances associated with contact forces and base motion. The control problem can now be reduced to a mixed sensitivity problem with weighting functions $\mathbf{W}_{yb}$ and $\mathbf{W}_{ub}$ shaping the auxiliary bearing position error and the control force applied to the bearing.

The onset of rotor bearing contact usually occurs suddenly and the duration may last for very short time periods. To reduce the contact forces it is reasonable to weight the control forces such that at low frequencies the control action is unhindered. At higher frequencies it is reasonable to weight the control action to reduce rotor/bearing contact forces. A suitable weighting function is of the form

$$\mathbf{W}_{ub}(s) = (10^2 + 10^{-1}s)/(10^6 + s)\mathbf{I} \tag{17}$$

A similar argument can be made for the position control of the bearing. At low frequencies it is important for the bearing to remain centralized within the clearance gap. However, at high frequencies reducing the contact forces is a priority, leading to a roll-off in the desired performance. An appropriate weighting is

$$\mathbf{W}_{yb}(s) = 10^3/(10^2 + s)\mathbf{I} \tag{18}$$

The weighting functions $\mathbf{W}_{yb}$ and $\mathbf{W}_{ub}$ and the frequency response of the PID controlled auxiliary bearing are shown in Fig. 5. The H_∞ controller, $\mathbf{H}_b(s)$, has order 10, and the closed loop frequency response of the auxiliary bearing is shown in Fig. 5.

5 Results and Discussion

Two different auxiliary bearing systems are considered.

Case 1. The rotor is supported by two magnetic bearings, each with a passive resiliently mounted auxiliary bearing.

Case 2. The rotor is supported by two magnetic bearings (same as Case 1), each operating with an actively controlled auxiliary bearing.

MATLAB and SIMULINK software was used to solve the dynamic equations of motion. A variable time step approach was undertaken using a numerical differentiation formula type solver. No time step limitations were specified.

5.1 Rotor Drop. Rotor drop was simulated by setting the magnetic bearing forces acting on the rotor to zero, giving the initial conditions $\mathbf{x}_r=0$, $\dot{\mathbf{x}}_r=0$, and $\Omega=1000$ rad/s. Gravitational forces were applied to the horizontal rotor.

Case 1. The response of the rotor and contact forces for the passive left hand side (LHS) auxiliary bearing system are shown in Fig. 6. The rotor is seen to drop and bounce within the clearance gap with increasing frequency and reducing contact forces. The first bounce contact force has a maximum value of approximately 1.75 kN. Further rotor/bearing contact events occur before the rotor eventually establishes continuous contact (not shown). During rotor/bearing contact the rotor displacement may exceed

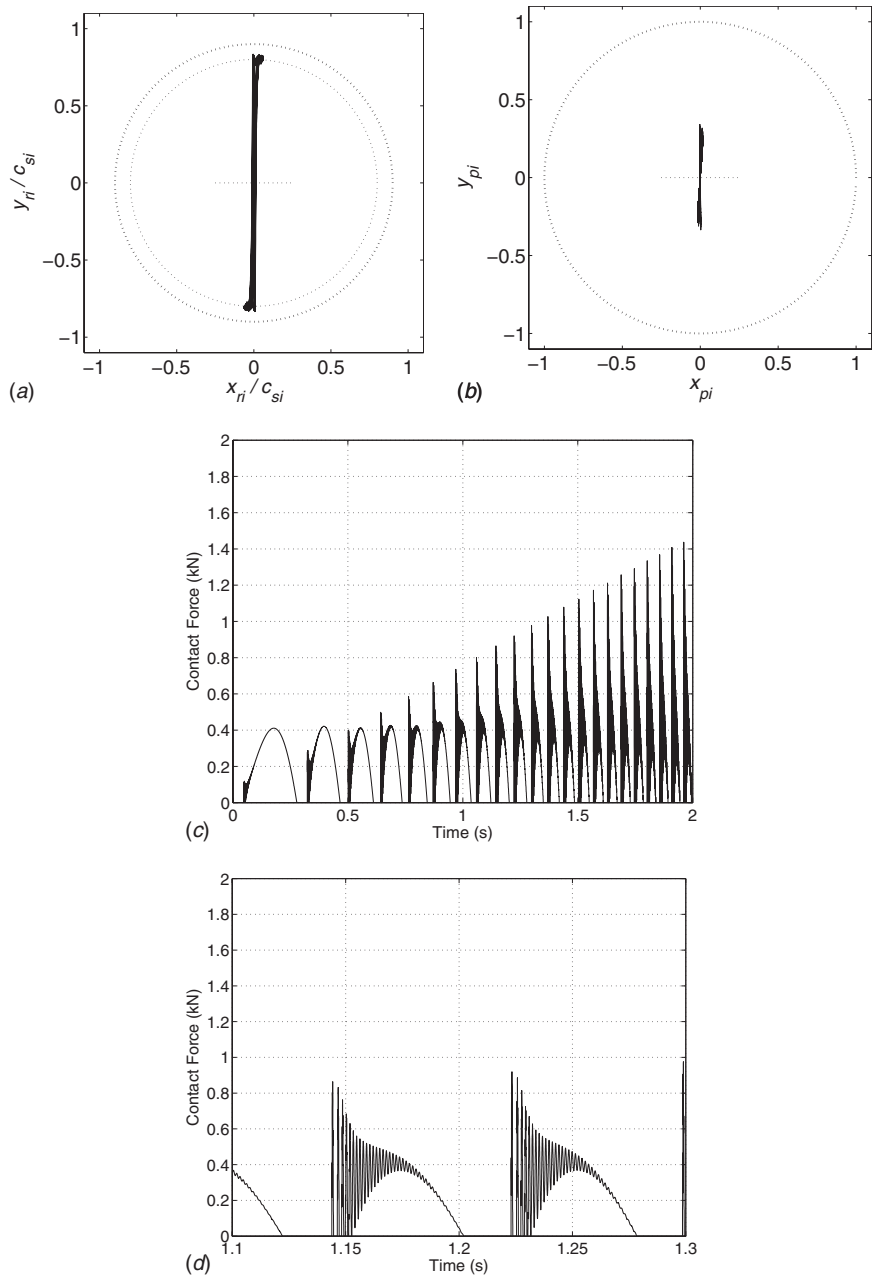


Fig. 8 (***a***) **Rotor displacements normalized with respect to rotor/stator clearance, c_{si}, for a passive auxiliary bearing.** (***b***) **Rotor deviation beyond clearance circle (x_{pi}, y_{pi}).** ((***c***) **and** (***d***)) **Rotor/bearing contact forces. Rotor operating speed Ω =500 rad/s.**

the clearance gap when using a compliant auxiliary bearing model.

The rotor may exceed the nominal clearance gap c_{bi} by $c_{si}-c_{bi}$, where c_{si} is the rotor-stator radial clearance for the magnetic bearing. Hence $\max(\delta r_i, 0)/(c_{si}-c_{bi})$ is a measure of how far the rotor can deviate. This may be decomposed into horizontal and vertical components, x_{pi} and y_{pi}, respectively. These are shown in Fig. 6(*b*) and since $\sqrt{x_{pi}^2+y_{pi}^2}<1$, rotor-stator contact does not occur.

Case 2. The rotor response and contact forces for the LHS active auxiliary bearing system are shown in Fig. 7. The first bounce contact force has a maximum value of approximately 1.25 kN. A second contact event then occurs, after which the rotor establishes continuous contact. The active auxiliary bearing is seen to reduce the rotor bouncing motion within the clearance gap. Figure 7(*b*) shows the displacement of the auxiliary bearing, x_b, normalized with respect to the maximum permissible auxiliary bearing displacement, which remains below unity.

The results show that the actively controlled auxiliary bearing is effective in

(a1) reducing the maximum contact force levels and
(b1) reducing the number of bounce events.

5.2 Unidirectional Base Motion. The response of the system to repeated rotor/bearing contacts was simulated using a chirp disturbance signal applied as a transmitted force. A 15 g disturbance acceleration was applied to the system in the vertical direction with a frequency sweep from 1 Hz to 10 Hz over 2 s. This

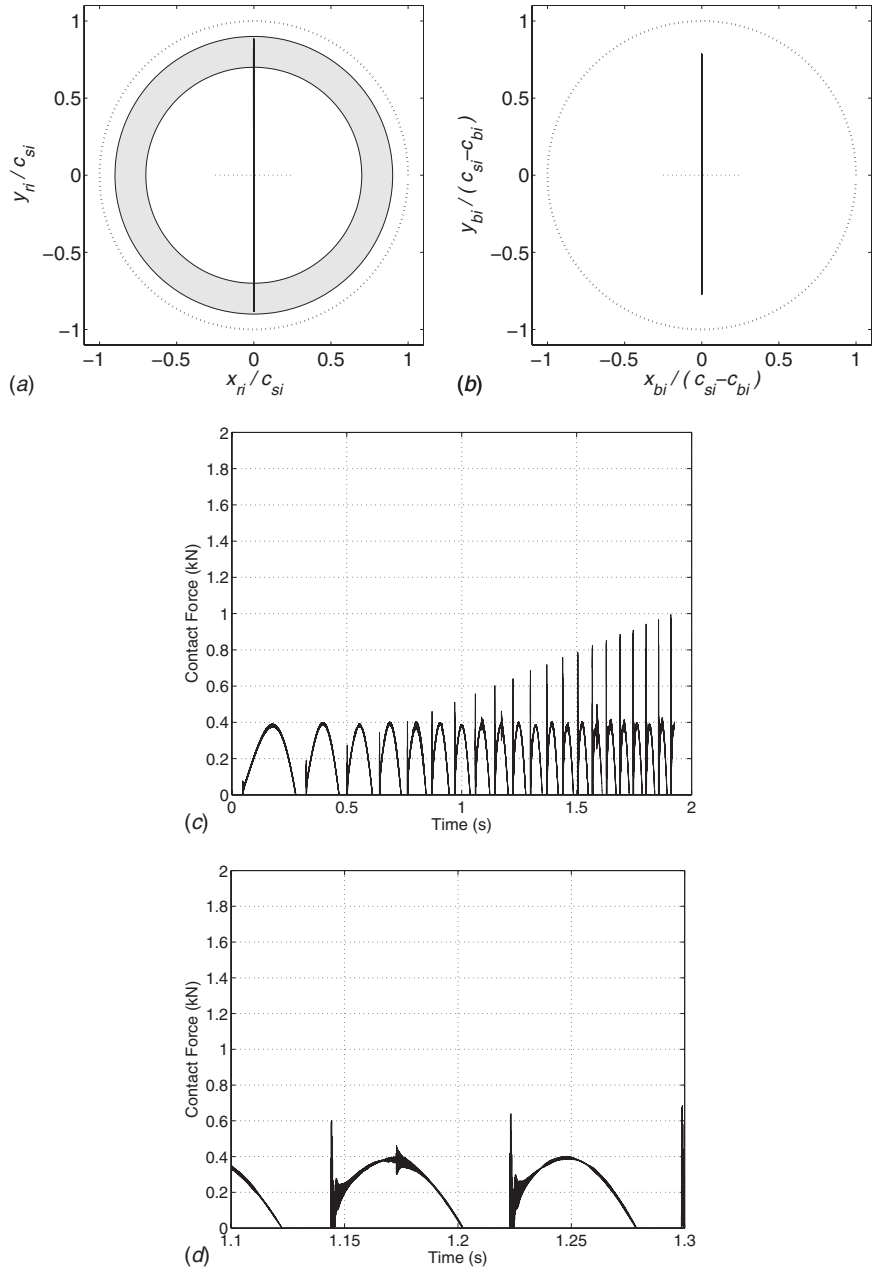


Fig. 9 (*a*) **Rotor displacements normalized with respect to rotor/stator clearance, c_{si}, for an active auxiliary bearing.** (*b*) **Active auxiliary bearing displacement $(x_{bi}, y_{bi})/(c_{si}-c_{bi})$.** ((*c*) **and** (*d*)) **Rotor/bearing contact forces. Rotor operating speed $\Omega=500$ rad/s.**

was sufficient to induce rotor contact.

Case 1. Figure 8 shows the rotor displacement, rotor penetration depth, and contact forces for the LHS passive auxiliary bearing. Figure 8(*d*) shows a zoomed view of two of the contact events shown in Fig. 8(*c*). The contact force shows initial contact events followed by continuous contact. As the sweep frequency increases, which increases the rotor vertical velocity, the contact forces increase.

Case 2. Figure 9 shows variations of rotor displacement $(x_{ri}, y_{ri})/c_{si}$, $(x_{bi}, y_{bi})/(c_{si}-c_{bi})$ and contact force, d_{ci}, for the LHS active auxiliary bearing. Again, the rotor contact forces increase with the disturbance frequency before the rotor enters continuous contact.

For this condition the actively controlled auxiliary bearing

(a2) reduces the maximum contact force levels,

(b2) reduces the number of bounce events, and

(c2) induces a small, though nonimportant, high frequency oscillation on the contact force.

The condition under (c2) could be detuned using a slightly modified controller design.

5.3 Sudden Rotor Unbalance. Sudden rotor unbalance tests were simulated by applying an unbalance disturbance vector to an initially balanced rotor. A rotor speed of $\Omega=500$ rad/s and a mass-loss eccentricity of 650 g cm were chosen.

Case 1. Figure 10 shows that a contact mode is induced with localized bounce behavior over particular repeated periods. The expanded view of one such period is shown in Fig. 10(*d*).

Case 2. Figure 11 shows the corresponding actively controlled responses. In comparison the active auxiliary bearing

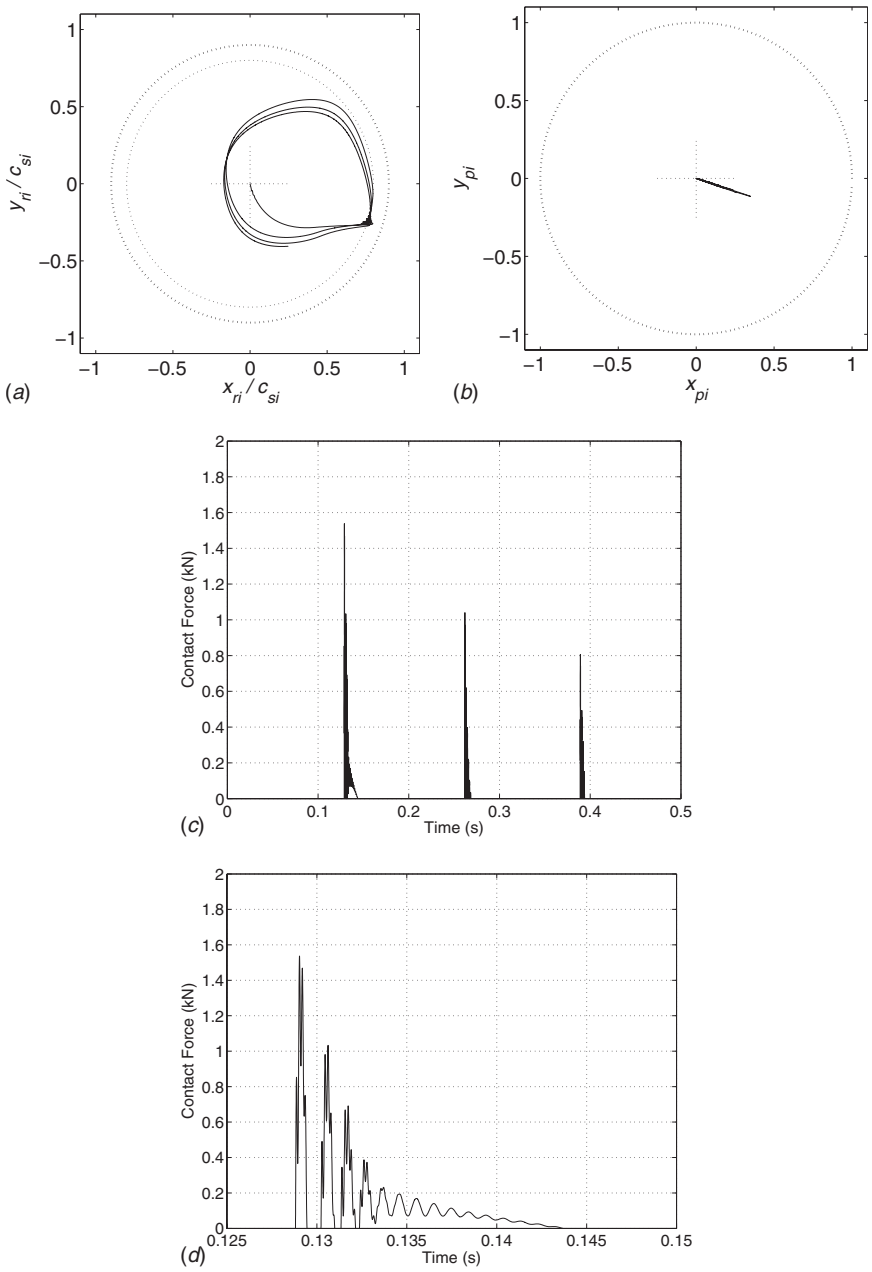


Fig. 10 (*a*) **Rotor displacements normalized with respect to rotor/stator clearance, c_{si}, for a passive auxiliary bearing.** (*b*) **Rotor deviation beyond clearance circle (x_{pi}, y_{pi}).** ((*c*) **and** (*d*)) **Rotor/bearing contact forces. Rotor operating speed $\Omega = 500$ rad/s.**

(a3) reduces maximum contact force levels and
(b3) shows high frequency control action through shorter duration contact events.

6 Conclusions

An investigation into the use of active auxiliary bearings within magnetic bearing systems has been presented. A rotor magnetic bearing system with two additional passive/active auxiliary bearings was simulated. An H_∞ approach was adopted to identify a suitable controller capable of ensuring auxiliary bearing system stability. Controller weightings were evaluated using a mixed sensitivity technique, taking account of the auxiliary bearing position and designed to reduce the contact forces during rotor/bearing contact.

Controller performance was evaluated using a simulated rotor drop, a unidirectional base motion disturbance, and a rotor mass-loss type fault. Comparison was made with a compliantly mounted passive auxiliary bearing. In all cases the active auxiliary bearing is shown to reduce the contact forces. Furthermore, the active bearing provides additional lateral control.

The results of the analysis show the potential benefits to be gained by actuating the auxiliary bearings in a magnetic bearing system. The reduction in maximum contact force levels together with contact durations should extend the operational life of a particular auxiliary bearing. Practical issues remain on the design of

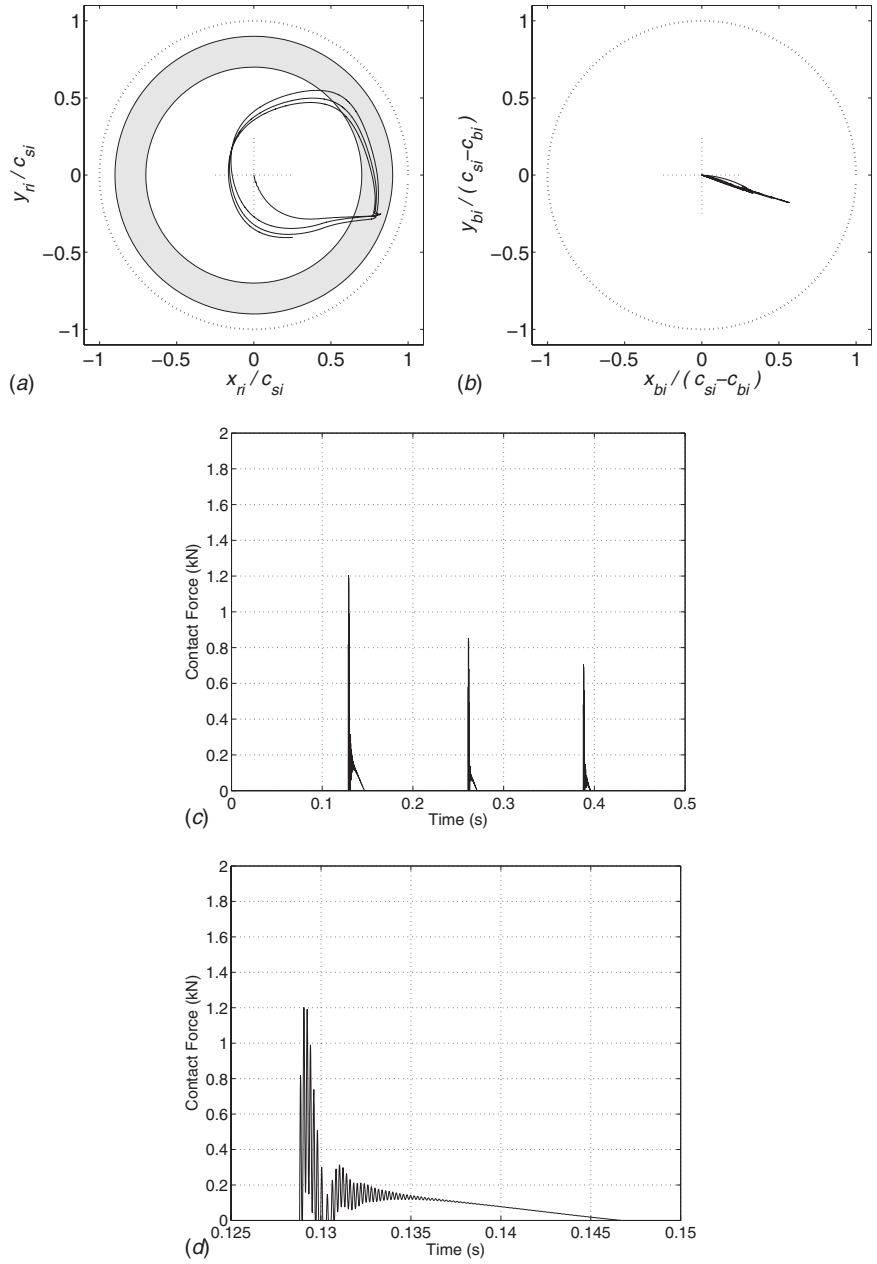


Fig. 11 (*a*) **Rotor displacements normalized with respect to rotor/stator clearance, c_{si}, for an active auxiliary bearing.** (*b*) **Active auxiliary bearing displacement $(x_{bi}, y_{bi})/(c_{si}-c_{bi})$.** ((*c*) **and** (*d*)) **Rotor/bearing contact forces. Rotor operating speed $\Omega=500$ rad/s.**

the actuation system; however, this is currently being considered using piezoelectric devices.

Acknowledgment

The authors gratefully acknowledge the funding support of the Engineering and Physical Sciences Research Council through research Grant No. EP/D031389/1.

Nomenclature

$\mathbf{A}_b$ = auxiliary bearing state matrix
$\mathbf{B}_i$ = mapping matrix
$\mathbf{C}_b$ = auxiliary bearing state-space output matrix
c_{bi} = bearing-stator radial clearance
c_i = auxiliary bearing damping
c_{ri} = rotor/auxiliary bearing radial clearance
c_{si} = rotor-stator radial clearance
$\mathbf{D}_b$ = auxiliary bearing state-space matrix
$\mathbf{D}_c$ = Laplace transform of the contact force vector
$\mathbf{d}_c$ = contact force vector
d_{ci} = localized contact force
E = Young's modulus
$\mathbf{F}_b$ = auxiliary bearing disturbance force vector
$\mathbf{f}$ = rotor disturbance force vector
$\mathbf{G}_i$ = transfer function
$\mathbf{H}_b$ = auxiliary bearing controller
$\mathbf{I}$ = identity matrix
k_i = auxiliary bearing stiffness
k_{Mp}, k_{Ap} = magnetic bearing and auxiliary bearing proportional gain

k_{Mi}, k_{Ai} = magnetic bearing and auxiliary bearing integral gain
k_{Md}, k_{Ad} = magnetic bearing and auxiliary bearing differential gain
l_{bi} = auxiliary bearing axial length
l_{ci} = contact zone arclength
$\mathbf{M}_b, \mathbf{C}_b, \mathbf{K}_b$ = auxiliary bearing mass, damping, and stiffness matrices
$\mathbf{M}_r, \mathbf{C}_r, \mathbf{K}_r, \mathbf{G}_r$ = rotor mass, damping, stiffness, and gyroscopic matrices
$\mathbf{Q}_b$ = Laplace transform of the bearing displacement vector
$\mathbf{q}_b, \mathbf{q}_r$ = auxiliary bearing state vector and rotor state vector
R_{ri}, R_{bi} = rotor and auxiliary bearing inner radii
s = Laplace transform variable
$\mathbf{T}$ = augmented plant transfer function matrix
$\mathbf{T}_{ij}$ = transfer function from i to j
$\mathbf{U}_b$ = Laplace transform of auxiliary bearing control forces
$\mathbf{U}_b$ = Laplace transform of the bearing control force vector
$\mathbf{u}$ = rotor control force vector
$\mathbf{u}_c$ = auxiliary bearing control force vector
$\mathbf{W}_{\text{in}}, \mathbf{W}_{\text{out}}, \mathbf{W}_{ij}$ = frequency dependent weighting function
x_{bi}, y_{bi} = ith node auxiliary bearing displacements
x_{pi}, y_{pi} = ith node rotor displacements beyond clearance circle
x_{ri}, y_{ri} = ith node rotor displacements
$\mathbf{x}_b, \mathbf{x}_r$ = auxiliary bearing displacement vector and rotor displacement vector
$\mathbf{Y}$ = seismic displacement vector
$\mathbf{Y}_b$ = Laplace transform of the auxiliary bearing system output vector
$\mathbf{y}_b$ = auxiliary bearing system output vector
$\mathbf{Z}_c$ = generalized contact force function
$\mathbf{Z}_t$ = Laplace transform of the transmitted force vector
$\mathbf{z}_t$ = transmitted force vector
δr_i = rotor deviation beyond clearance gap
μ = coefficient of friction
ν = Poisson ratio
Ω = rotor angular velocity
$\bar{\sigma}[\cdot]$ = singular value
$\mathcal{L}[\cdot]$ = Laplace transform
$\|\cdot\|_\infty$ = infinity norm
$[\cdot]^T$ = transpose

References

[1] Fumagalli, M., Deeyn, B., and Schweitzer, G., 1992, "Dynamics of Rigid Rotors in Retainer Bearings," *Proceedings From the Third International Symposium on Magnetic Bearings*, Alexandria, VA, July 21–31, pp. 157–166.
[2] Ishii, T., and Kirk, G., 1996, "Transient Response Technique Applied to Active Magnetic Bearing Machinery During Rotor Drop," ASME J. Vibr. Acoust., **118**, pp. 154–163.
[3] Muszynska, A., and Goldman, P., 1993, "Chaotic Vibrations of Rotor/Bearing/Stator Systems With Looseness or Rubs," *Proceedings of the ASME Nonlinear Vibrations*, Vol. 54, pp. 187–194.
[4] Ehrich, F. F., 1988, "High Order Subharmonic Response of High Speed Rotors in Bearing Clearance," ASME J. Vibr. Acoust., **10**, pp. 9–16.
[5] Childs, D., 1979, "Rub Induced Parametric Excitation in Rotors," ASME J. Mech. Des., **10**, pp. 640–644.
[6] Lawen, J., and Flowers, G., 1997, "Synchronous Dynamics of a Coupled Shaft/Bearing/Housing System With Auxiliary Support From a Clearance Bearing," ASME J. Eng. Gas Turbines Power, **119**, pp. 430–435.
[7] Sahinkaya, M. N., Abulrub, A. G., Keogh, P. S., and Burrows, C. R., 2007, "Multiple Sliding and Rolling Contact Dynamics for a Flexible Rotor/Magnetic Bearing System," IEEE/ASME Trans. Mechatron., **12**, pp. 179–189.
[8] Keogh, P., and Yong, W., 2007, "Thermal Assessment of Dynamic Rotor/Auxiliary Bearing Contact Events," ASME J. Tribol., **129**, pp. 143–152.
[9] Bartha, A., 2000, "Dry Friction Induced Backwards Whirl: Theory and Experiment," *Proceedings of the 7th International Symposium on Magnetic Bearings*, ETH Zurich, Aug. 23–25, pp. 231–236.
[10] Palazzolo, A., Lin, R., Alexander, R. M., Kascak, A., and Montauge, J., 1989, "Piezoelectric Pushers for Active Vibration Control of Rotating Machinery," ASME J. Vibr. Acoust., **111**, pp. 298–305.
[11] Palazzolo, A., Lin, R., Alexander, R. M., Kascak, A., and Montague, J., 1991, "Test and Theory for Piezoelectric Actuator-Active Vibration Control of Rotating Machinery," ASME J. Vibr. Acoust., **113**, pp. 167–175.
[12] Tang, P., Palazzolo, A., Kascak, A., and Montauge, J., 1997, "Electromechanical Modelling of Hybrid Piezohydraulic Actuator System for Active Vibration Control," ASME J. Dyn. Syst., Meas., Control, **119**, pp. 10–18.
[13] Ulbrich, H., Chavez, A., and Dhima, R., 2004, "Minimization of Contact Forces in Case of Rotor Rubbing Using an Actively Controlled Auxiliary Bearing," *Proceedings of the Tenth International Symposium on Transport Phenomena and Dynamics of Rotating Machinery*, Honolulu, HI, Mar. 7–11, pp. 31–10.
[14] Chavez, A., Ulbrich, H., and Ginzinger, L., 2006, "Reduction of Contact Forces in a Rotor-Stator-System in Case of Rubbing Through Active Auxiliary Bearing," Shock Vib., **129**, pp. 505–518.
[15] Jiang, J., Ulbrich, H., and Chavez, A., 2006, "Improvement of Rotor Performance Under Rubbing Conditions Through Active Auxiliary Bearings," Int. J. Non-Linear Mech., **41**, pp. 949–957.
[16] Glover, K., and Doyle, J. C., 1988, "State-Space Formulae for all Stabilizing Controllers that Satisfy H_∞ Norm Bound and Relations to Risk Sensitivity," Syst. Control Lett., **11**, pp. 167–172.
[17] Nelson, H., and McVaugh, J., 1976, "The Dynamics of Rotor–Bearing Systems Using Finite Elements," ASME J. Eng. Ind., **98**, pp. 593–600.
[18] Roark, R. J., and Young, W. C., 1975, *Formulas for Stress and Strain*, McGraw-Hill, New York.

Experimental Study on Impeller Blade Vibration During Resonance—Part I: Blade Vibration Due to Inlet Flow Distortion

Albert Kammerer

Reza S. Abhari

LEC, Laboratory for Energy Conversion, Department of Mechanical and Process Engineering, ETH Zurich, 8092 Zürich, Switzerland

Forming the first part of a two-part paper, the experimental approach to acquire resonant vibration data is presented here. Part II deals with the estimation of damping. During the design process of turbomachinery components, mechanical integrity has to be guaranteed with respect to high cycle fatigue of blades subject to forced response or flutter. This requires the determination of stress levels within the blade, which in turn depend on the forcing function and damping. The vast majority of experimental research in this field has been performed on axial configurations for both compressors and turbines. This experimental study aims to gain insight into forced response vibration at resonance for a radial compressor. For this purpose, a research impeller was instrumented with dynamic strain gauges and operated under resonant conditions. Modal properties were analyzed using finite element method and verified using an optical method termed electronic-speckle-pattern-correlation-interferometry. During the experiment, unsteady forces acting on the blades were generated by grid installations upstream of the impeller, which created a distorted inlet flow pattern. The associated flow properties were measured using an aerodynamic probe. The resultant pressure fluctuations on the blade surface and the corresponding frequency content were assessed using unsteady computational fluid dynamics. The response of the blades was measured for three resonant crossings, which could be distinguished by the excitation order and the natural frequency of the blades. Measurements were undertaken for a number of inlet pressure settings starting at near vacuum and then increasing. The overall results showed that the installed distortion screens generated harmonics in addition to the fundamental frequency. The resonant response of the first and the second blade mode showed that the underlying dynamics support a single-degree-of-freedom model. [DOI: 10.1115/1.2968869]

1 Introduction

The preliminary design process of compressors and turbines focuses on satisfying required cyclic parameters and essentially comprises the optimization of component aerodynamics. In many cases, the number of which is increasing, this process inevitably leads toward the added complexity of aerodynamic components or pushes well established designs toward their mechanical limits. This stage of the design process necessitates the quantification of parameters with respect to which mechanical integrity can be judged. To focus more specifically on the problem of blade high cycle fatigue, an appreciation of this problem is given in the frequently cited works by El-Aini et al. [1] and Kielb [2], which quantify the significant costs generated in this field during the design phase and operational failure.

A number of examples should be given at this point, which draw particular attention to potential problems arising from resonant vibration due to flow distortion in the inlet of compressors. Centrifugal compressors are pushed toward higher mass flow rates and pressure ratios, which, on the one hand, increase the mean stress levels and reduce the tolerable cyclic stresses and, on the other hand, demand thinner blades. On many occasions, this may in addition lead to resonance with low engine order excitation. In other applications, system integrity requires particularly short and bent inlet ducts or support struts. Either way, nonuniformities are introduced into the flow field. Increasing research and application efforts can be identified in the field of extending compressor stability through flow injection in the near tip region. This is realized using a discrete number of injectors. The problem introduced is apparent. For most applications, resonant vibration can be avoided by design modifications; however, in a number of applications, this is not the case. Whenever turbomachines are subject to continuous alteration of the operating point, i.e., turbocharging applications, industrial compression systems, or aeroengine applications, resonant vibration has to be taken into account. Bearing this in mind, computational tools have to be developed and validated experimentally.

The following short overview on publications underlines the complexity of such experiments for rotating machines and the need to perform the experiments in engine representative conditions. Extensive experimental work into the field of vibration was performed by Haupt [3] and supplemented by Jin [4] for a radial compressor. The studies aimed to understand the principles of blade excitation, quantify response amplitudes, and compare blade response within the stable and unstable operating regimes. The importance of fundamental forced response research in an engine representative environment for a fan is presented in detail in a series of publications by Manwaring et al. [5–8]. Experiments were performed to quantify the main three aspects of blade vibration, i.e., inlet flow conditions, blade pressure distribution, and

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 8, 2008; final manuscript received April 8, 2008; published online December 29, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

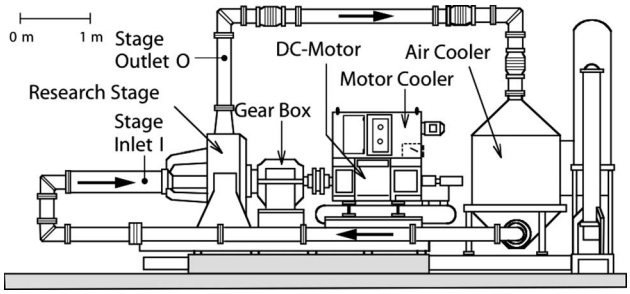


Fig. 1 Radial compressor research facility

vibratory response. The data were used to improve the application of computational tools used during the design process. A recent contribution by Dickmann et al. [9] points toward potential future design practice, employing computation of complex fluid structure interaction in centrifugal compressors.

2 Motivation and Scope of the Paper

The need for reliable data to advance the understanding of forced response in centrifugal compressors and then to develop and validate computational tools necessitates experimental techniques. This paper presents an experimental study of forced response in a centrifugal compressor. The impeller geometry is typical for turbocharging applications. Two aspects of forced response measurements are discussed in detail: (1) generation and measurement of inlet flow quantities and (2) vibratory response measurements. Vibration measurements were conducted by gradually increasing the shaft speed through resonant crossings. The inlet flow was intentionally disturbed using screens and measured using an aerodynamic probe. It will be shown that the acquired dynamic strain data corresponds well with a single-degree-of-freedom (SDOF) model for the first and the second mode. In the same way, the maximum response dependency on inlet pressure and therefore aerodynamic damping will be shown.

3 Test Facility

The test facility is a single stage centrifugal compressor, termed "RIGI" and is located in the Turbomachinery Laboratory of the Swiss Federal Institute of Technology in Zurich, Switzerland. A schematic of the test facility is given in Fig. 1. The facility operates in a closed loop, which allows the pressure to be adjusted independently of ambient conditions. In the present study, the pressure at the inlet of the impeller was varied within the range of 0.1–0.6 bar. The impeller is driven by a 440 kW dc motor with a two stage gear box. Upstream of the impeller, both the pressure and the temperature are measured and held constant to the desired levels by a control system. Downstream of the compression stage, the fluid is cooled and then discharged through a throttle prior to reentering the impeller. The throttle is used to set the mass flow rate. A standard orifice upstream of the impeller is used to measure the mass flow. In order to estimate the performance of the facility, the pressure and temperature are measured upstream and downstream of the compression stage.

4 Impeller and Main Blade Modal Properties

The impeller used in this study is a modified version of the impeller designed by Schleer [10], which has been employed in a number of experimental flow structure investigations. An in-depth description is given in Schleer [11]. With a design total pressure ratio of $p_{t,\mathrm{out}}/p_{t,\mathrm{in}}=2.8$, the impeller consists of seven pairs of main and splitter blades with an outer diameter of 400 mm. The diffuser is parallel and vaneless with an exit diameter of 580 mm and a height of 15.8 mm. Downstream of the diffuser, the flow is discharged into a toroidal collecting chamber. The maximum rotational speed is 22,000 rpm.

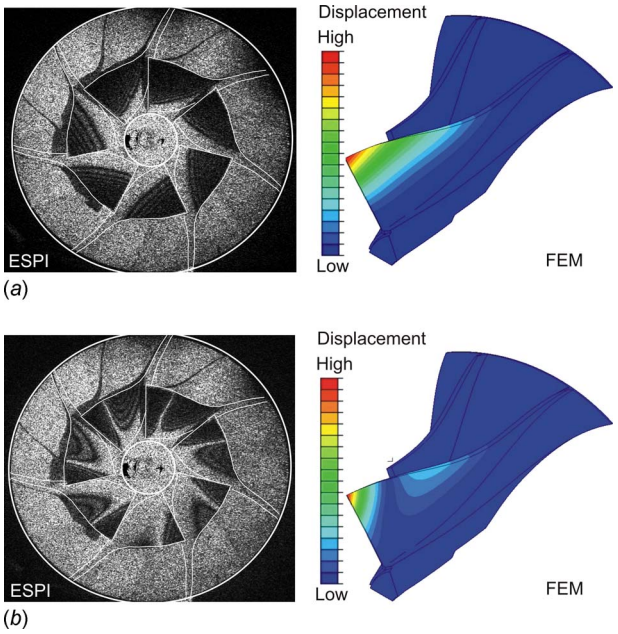


Fig. 2 Impeller main blade modal shapes for Modes 1 and 2. (*a*) Mode 1; (*b*) Mode 2.

In order to fulfill the requirements of the current study, the modified impeller design features thinner blades while all other parameters defining the geometry of the blades and the disk were retained. The redesign focused on ensuring a number of potential resonant crossings of the first and second main blade eigenfrequencies with the low order engine excitation sources. The Campbell diagram in Fig. 10 illustrates the location of the main blade eigenfrequencies, termed Modes 1–4, and a number of potential low order engine excitations. The eigenfrequencies can be seen to be well separated. Given this configuration, flow disturbances upstream of the impeller can be installed providing a specific engine order excitation and therefore forcing the main blade to vibrate at resonance. In order to illustrate the modal shapes of the main blade during resonant vibration, two studies were conducted. First of all, a modal analysis calculation was performed using finite element model (FEM) and a cyclically symmetrical model. Second, the modal shapes were optically acquired using electronic-speckle-pattern-correlation-interferometry (ESPI). The modal shapes of Modes 1 and 2 are illustrated in Fig. 2. The FEM calculation shows the local displacement for a given mode and thus the regions of local deformation during resonant vibration can be identified. The ESPI results are similarly presented. In the case of the ESPI measurements, the mode dependent deformation of the main blade is shown for the entire impeller. During these measurements, the impeller was excited by a piezoactuator mounted on the backside of the impeller. The piezo was driven by a sinusoidal excitation signal, which could be adjusted with respect to excitation amplitude and frequency. Doing so, the entire frequency range of interest was swept stepwise acquiring a single picture for a given frequency. The seven main blades can be clearly identified as exhibiting the same deformation pattern as predicted by the FEM calculations.

5 Inlet Flow Distortion

5.1 Generation of Flow Distortion. According to the Campbell diagram in Fig. 10, the upstream flow has to be distorted such that EO4–EO8 excitations are generated. This can be achieved through the employment of distortion screens upstream of the impeller. It is crucial to measure the flow upstream of the impeller in order to visualize the resultant conditions through the quantifica-

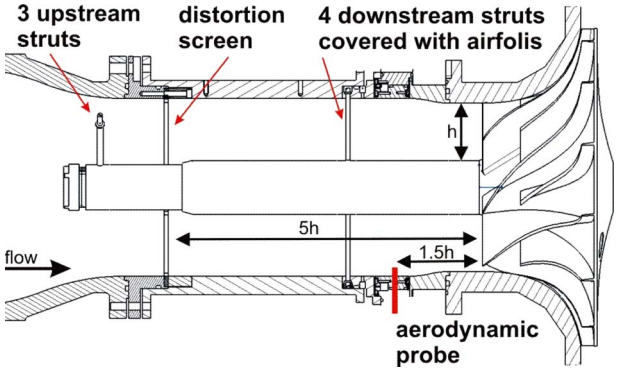


Fig. 3 Arrangement and dimensions within the inlet section

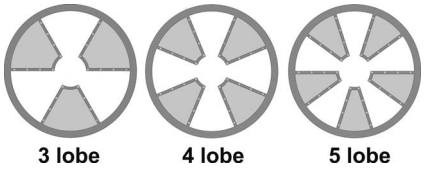


Fig. 4 Distortion screen geometry

tion of flow properties. Figure 3 illustrates the arrangement within the inlet section of the compressor. The main components are the screens that generate the flow distortion and the aerodynamic probe that measures the flow properties. Their upstream distances are 5 and 1.5 blade heights, respectively. The center of the inlet section is occupied by a cylinder covering the rotary transmitter of the strain gauge signal. The transmitter is mounted and centered using two rows of adjustable struts, i.e., three upstream and four downstream. The downstream struts are covered by symmetrical airfoils in order to avoid flow separation that would otherwise be generated across the rods. As will be shown later, the given arrangement within the inlet section creates a distinct distortion pattern.

The geometry of the installed screens is shown in Fig. 4. These are essentially made of a frame holding a wire grid with specific properties, namely, the wire thickness and the mesh width. The number of lobes determines the engine order excitation number. The grids are designed such that the area ratio between the blocked and the unblocked sections is equal for all screens. As the flow passes through the grid, losses are generated that block the fluid and force the mass flow to redistribute. As a result, two zones of different velocities and therefore total pressure are generated. The velocity difference generated by the screens depends primarily on the uniform upstream velocity facing the screen and the grid size properties. Based on these parameters, an assessment of the resultant velocities downstream of the screen can be carried out. An empirical study on the pressure drop across a grid with given wire properties was carried out by Roach [12], whereas Koo and James [13] performed measurements and calculations on the velocity distribution downstream of a blocked area for a given pressure drop. The combination of the two studies allows an assessment of the level of flow distortion due to the installed grid. To start with, Roach [12] examined the flow for a range of grid wires aiming to quantify pressure loss and turbulence properties. The pressure drop was found to correlate well using the following equation:

$$\frac{\Delta p}{q} = A\left(\frac{1}{\beta^2} - 1\right)^B \tag{1}$$

where Δp is the pressure drop over q the dynamic head, $\beta=(1-d/M)^2$ is the grid porosity, and A depends on the Reynolds number and is given in graphical form. The parameter B is equal to unity for cylindrical wires. Further correlations are given for turbulence properties, i.e., turbulence intensity, spectra, correlation functions, and length scales. These properties will not be examined in detail within this work. The study by Koo and James [13] examines the velocities downstream of a partially blocked flow for perpendicular and inclined grids using an analytical and a numerical approach. Measurements were conducted in order to verify the applicability of these models and were found to corroborate their applicability.

For a rapid assessment of the current problem, the analytical approach is very attractive, with the need to simply compute a set of equations that allow the calculation of the far upstream velocity $V_{ax,ups}$, the downstream unblocked velocity $V_{ax,unb}$, and the blocked velocity $V_{ax,blo}$. These velocities are functions of the pressure drop coefficient $\Delta p/q$ and the parameter λ. The latter represents the area split between the blocked and the unblocked part of the flow field. Figure 5(*a*) illustrates the velocity ratios for $\lambda=0.5$, i.e., the blocked and unblocked areas are of the same size. The unknown parameter for an arbitrarily chosen mesh is the pressure drop coefficient, which can be determined from the work by Roach [12].

Within the test facility, the upstream axial velocity is prescribed by the operating point setting and the associated mass flow. Therefore, the expected velocity distortion is assessed for the entire mass flow range. Figure 5(*b*) depicts the predicted velocities upstream and downstream of the grid as a function of the nondimensional mass flow rate. Here the velocities are normalized using the mean axial velocity $V_{ax,mean}$, which is a measure for the mass flow

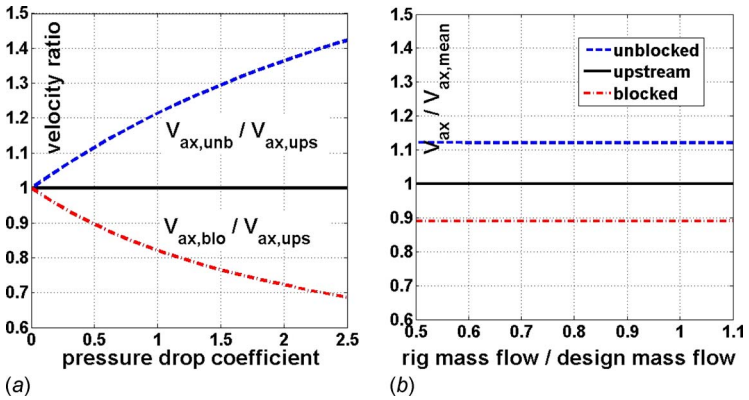


Fig. 5 Velocity distortion by Koo and James [13] and velocity ratio dependency on mass flow. (*a*) Velocity ratios for $\lambda=0.5$: (*b*) predicted velocity ratio as function of the mass flow rate.

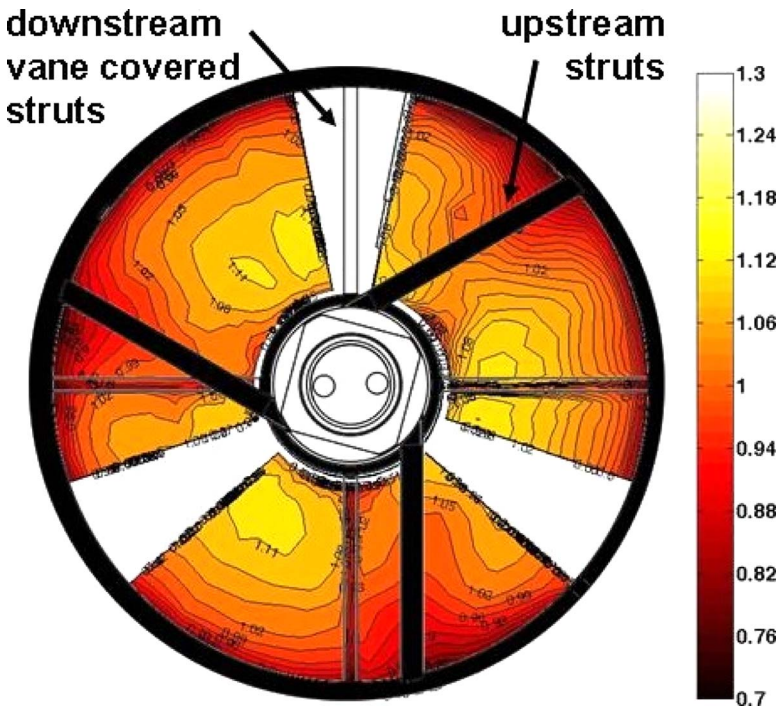


Fig. 6 Measured normalized axial velocity distribution ($V_{ax}/V_{ax,mean}$); a distortion screen is not installed. Support strut arrangement is shown.

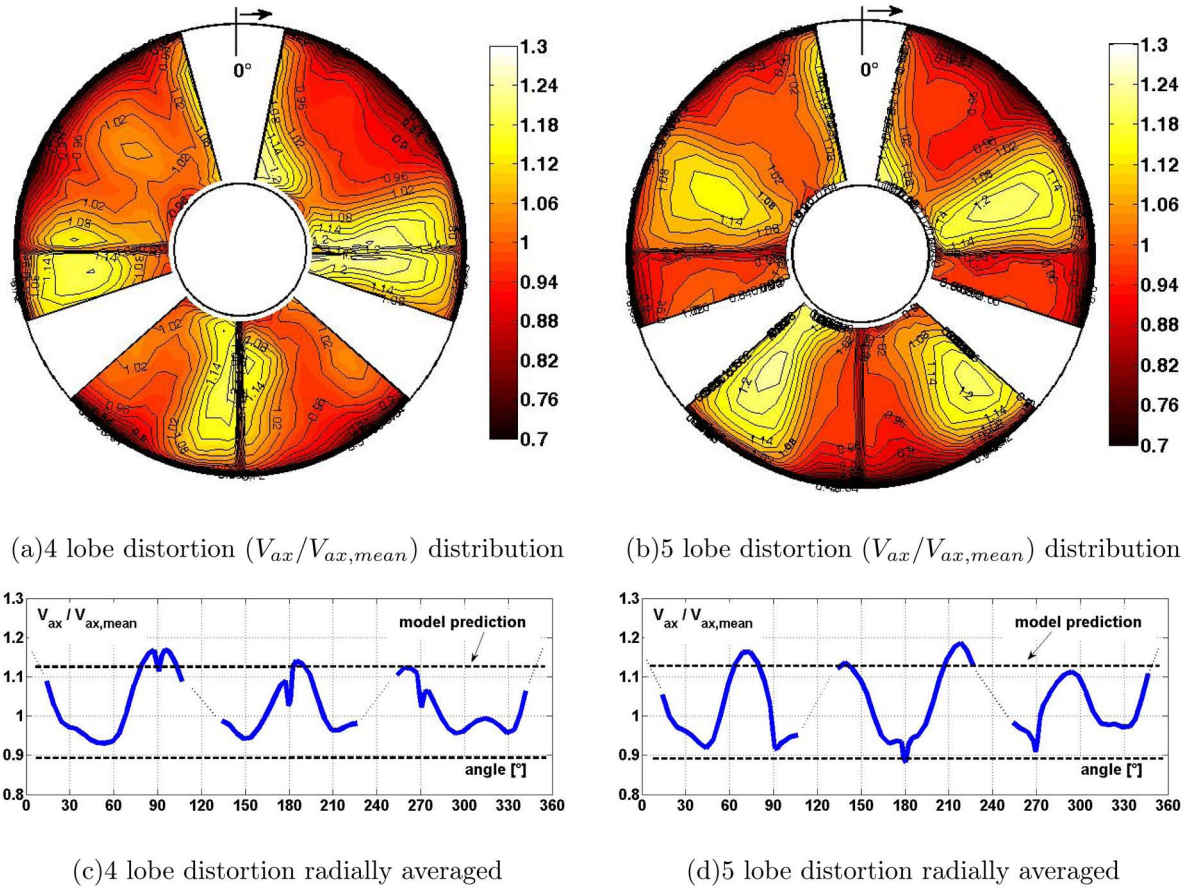
rate. It can be observed that the velocity ratios remain constant for the entire range of mass flow settings, which corresponds to a pressure loss coefficient of $\Delta p/q=0.5$.

5.2 Flow Field Measurement Upstream of the Impeller. Flow measurements upstream of the impeller were performed using a two-sensor fast response aerodynamic probe, termed FRAP. The working principals and calibration of the probe are described in detail by Pfau et al. [14]. Briefly, for each measurement point, the probe is revolved around the stem axis, simulating a virtual four-sensor probe and data are ensemble averaged using the one-per-revolution trigger in order to compute all necessary flow properties for a single point. The uncertainty within the total pressure is in the range of ± 100 Pa [14]. The steady total temperature derived from the FRAP has an uncertainty of ± 0.5 K [15].

Within the test section, the probe was mounted in a traversing mechanism enabling automated positioning of the probe head in radial and circumferential directions. Circumferentially, traversing was limited by external struts designed to support the inlet section. The white sections in Fig. 6 can be observed where the measurement data are not available. The measurable inlet area consists of three sectors each with a circumferential extent of 97.5°, thus covering 81% of the entire area. A typical measurement resolution consisted of 24 circumferential and 13 radial points with staggering applied near the walls and downstream of the struts where airfoil generated wakes were expected.

5.3 Flow Field Without Distortion Screen Installation. Initially, the flow field upstream of the impeller was measured with all necessary installations containing equipment to measure the strain. A distortion grid was not installed. Figure 6 shows the time-averaged normalized axial velocity ($V_{ax}/V_{ax,mean}$). The mean axial velocity was calculated through mass flow averaging and was found to deviate less than 2% from the axial velocity based on the measured rig performance. For better visualization of the results, all struts upstream of the measurement plane are indicated. As can be observed, the flow field is not uniform, with local deviation in an axial velocity of ≈8% from the mean value. Two regimes are clearly distinguishable, the bulk flow and the boundary layer at the tip of the cross section. Within the bulk flow near the hub, three zones of elevated velocity $V_{ax}/V_{ax,mean}\approx 1.08$ can be observed. From symmetry, it can be stated that these zones are created as a result of the three upstream struts, as also seen in Fig. 3. Downstream of the four upstream struts, the plot reveals wakes of comparatively small extent due to the symmetrical vanes covering the struts. Close to the hub of the central tube, the boundary layers were not measured due to a required minimum probe distance from the wall. At the tip, the boundary layers are clearly developed. Also within this regime, the three upstream struts exert a profound influence on the velocity distribution. The area coverage of the low momentum fluid has clearly grown. In conclusion, although the flow was not intentionally distorted using screens, a distortion pattern was measured due to the installations housing the measurement equipment. Given this flow distribution, the potential at least for EO3 and EO4 excitations was given and its effect was indeed measured as blade deflection.

5.4 Flow Field With Distortion Screen Installation. The flow fields for the four-lobe and the five-lobe screen will be presented here. In both cases, a mesh was used with an estimated pressure drop coefficient of $\Delta p/q=0.5$. For both cases, the nondimensional axial velocity is shown in Fig. 7. The effect of partially blocking the flow field can be clearly identified for the four-lobe screen in Fig. 7(*a*) and for the five-lobe screen in Fig. 7(*b*), respectively. Plots for radially mass-weighted averaging the velocity are given below in combination with the expected velocity levels as predicted by the model described previously. Here, moving along the circumference in a clockwise direction, the low and high momentum regimes can be identified. Both,, the lower limit with $V_{ax}/V_{ax,mean}=0.9$ and the upper limit $V_{ax}/V_{ax,mean}=1.13$ are slightly underpredicted with respect to the minima and maxima but are considered to be sufficiently well within prediction, bearing in mind the simplicity of the model. The four downstream struts clearly generate the local minima due to wake creation. The transition from the blocked to the unblocked area is smooth, resembling a sinusoidal function rather than a step function. This indicates that sufficient mixing takes place as the fluid passes from



(a)4 lobe distortion ($V_{ax}/V_{ax,mean}$) distribution

(b)5 lobe distortion ($V_{ax}/V_{ax,mean}$) distribution

(c)4 lobe distortion radially averaged

(d)5 lobe distortion radially averaged

Fig. 7 FRAP measured normalized velocity distribution upstream of the impeller

the screen to the measurement plane. Going into detail, the contour plots of the axial velocity reveal a number of interesting aspects. First of all, the distorted and undistorted lobes are not of the same shape and the local minima and maxima vary. For both screens, the local maximum velocities within the unblocked area amount a ratio of ≈1.2, i.e., being 20% above the mean value. Similar to the no-screen case presented previously, the upstream struts of the inlet section affect the flow field most apparently at around 280 deg. In this part of the cross section, the area occupied by the high momentum fluid is reduced, in particular, at the hub for the four-lobe screen. In the case of the five-lobe screen, the strut passes right through the high velocity fluid forcing it to decelerate. Within the blocked area, a significant amount of low momentum fluid accumulates at the tip due to the deceleration through the grid. This effect is locally attenuated by the upstream struts and the velocity ratio drops significantly below a predicted value of 0.9. Therefore, the locally confined low momentum fluid is balanced by high momentum fluid within the nonblocked regime, leading to values significantly above the prediction.

In conclusion, the screens generate the expected pattern; however, local variations are significant with the overall velocity ratio ranging between 0.82 and 1.2. Due to these local variations in the flow field, it is expected that harmonics in the excitation of the blades will be observed in addition to the fundamental excitation frequency dictated by the number of screen lobes. This conclusion is supported by unsteady CFD simulation carried out by Zemp et al. [16], applying inlet boundary conditions as measured for the five-lobe screen. Figure 8 shows the unsteady pressure difference Δp and its phase angle along the midspan of the main impeller blade. The highest fluctuation amplitudes can be observed to be present in the inducer portion of the impeller. The phase angle for this regime is constant. As expected, the fundamental excitation order EO5 generates the strongest pressure fluctuations. The fluctuation magnitude from EO3 is smaller by approximately a factor of 2. Its occurrence can be associated with the distortion pattern generated by the upstream struts as outlined previously. Its second harmonic can also be observed as EO6 excitation and is smaller by a factor of 2. Of similar magnitude is the fluctuation from EO4 excitation, which can be associated with wakes generated by vanes upstream of the impeller. Though this computational case

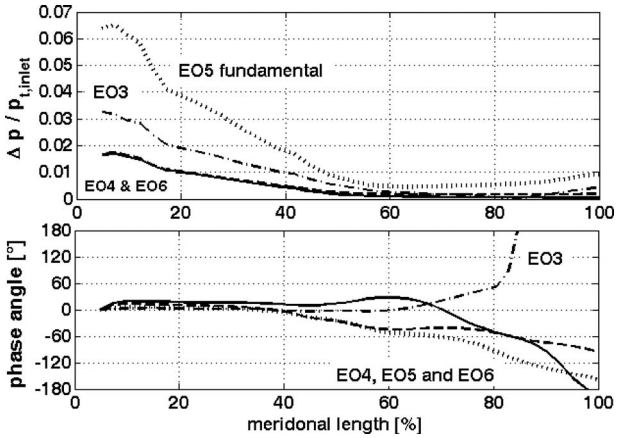


Fig. 8 Computed amplitude and phase angle of the unsteady pressure on blade surface using unsteady CFD. Measured inlet boundary conditions were applied for the five-lobe case.

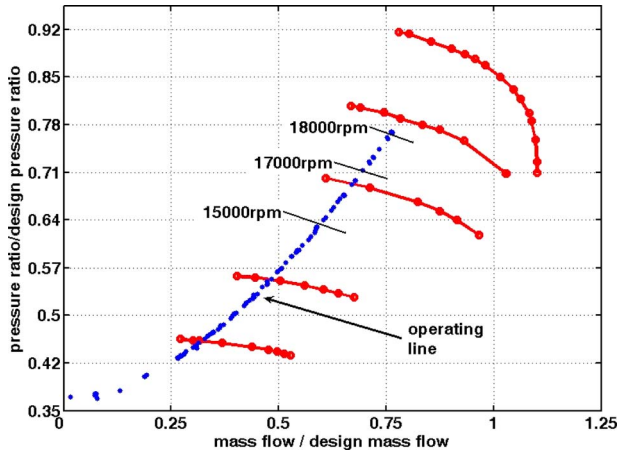


Fig. 9 Impeller performance map

study cannot be confirmed by experimental data, it helps significantly to develop the necessary appreciation of the problem.

6 Blade Strain Measurement Procedure and Modeling

6.1 Signal Acquisition. Vibratory stresses were obtained by dynamic strain gauges applied to the blade surface. Each main blade was equally instrumented with three strain gauges. The positioning of the strain gauges was optimized to capture the response of the first four main blade eigenmodes using the procedure described by Szwedowicz et al. [17]. The strain gauge signal was multiplexed such that the two blades were measured simultaneously for an ascending and a descending sweep, and then the same measurements were repeated with the next set of blades. A mercury slip ring was used to transmit data to the stationary frame of reference. Data amplification was not required due to the low noise level of the slip ring. Each data channel was scanned with a sampling rate of 200 kHz. A reference one-per-revolution signal was used to track the rotational speed.

6.2 Transient Measurement Approach. During strain measurements, stress levels due to resonant vibration were expected to be sufficiently high to potentially cause failure for a given number of cycles. Therefore, in order to overcome the risk of steady measurements, transient measurements were performed with respect to rotational speed, i.e., to sweep through the resonant speed regime with constant change in shaft speed. By doing so, the blade was effectively excited by a forcing function of constantly changing frequency, expressed by the sweep rate. In order to achieve near steady state response, the sweep rate has to be chosen sufficiently low in order to allow the structure to build up maximum response. Ewins [18] referenced the maximum linear sweep rate $S_{\max}$ prescribed by ISO standards:

$$S_{\max} = 3.6 f_{\text{res}}^2 \zeta^2 \left(\frac{\text{Hz}}{\text{s}} \right) \tag{2}$$

Using realistic values for resonant frequency and critical damping ratio for the data application yields maximum sweep rates of approximately 7 Hz/s. Figure 9 shows the compressor map and the operating line for a constant throttle setting along which blade response was measured. In a first step, sweeping was performed along the operating line up to 18,000 rpm in order to identify resonant crossings. In a second step, refined sweeping was performed with the above mentioned maximum sweep rate in order to acquire data across a resonant response upon which data analysis was subsequently performed. Sweeping was carried out for ascending and descending rotational speeds.

6.3 Modeling Transient Blade Response. In order to understand the response characteristics of a blade under transient forcing conditions, a SDOF can be used to simulate the problem. The SDOF system is described by the second-order ordinary differential equation of the form

$$m\ddot{x} + c\dot{x} + kx = F(t) \tag{3}$$

with x representing the blade displacement amplitude. The forcing function herein is given by the following equation:

$$F(t) = F_0 \cos[2\pi f(t) t + \phi_0] \tag{4}$$

where the frequency sweep $f(t)$ is modeled using the relation

$$f(t) = \frac{1}{2}\beta t + f_0 \quad \text{with} \quad \beta = \frac{f - f_0}{t - t_0} \tag{5}$$

The frequency as a function of time essentially resembles the constant impeller sweep rate during the measurements. The solution of the differential equation given above can be numerically obtained through interpolation of the excitation function and employment of the linear response of a SDOF system to step excitation. Details on this procedure can be found in Craig and Kurdila [19]. The evaluation of this approach yields the recurrence formulas

$$x_{i+1} = a_1 F_i + a_2 F_{i+1} + a_3 x_i + a_4 \dot{x}_i \tag{6}$$

$$\dot{x}_{i+1} = a_5 F_i + a_6 F_{i+1} + a_7 x_i + a_8 \dot{x}_i \tag{7}$$

where the response x_{i+1} can be calculated for a discrete time step Δt and a given forcing function. The latter is given as a vector computed using the above equations. The coefficients a_1-a_8 are functions of damping, resonant frequency, and time step and are provided for a viscous lightly damped system.

6.4 Strain Amplitude Dependency on Inlet Pressure. The inlet pressure is one of the main variation parameters in this study. Therefore, the measurement procedure described above was carried out for a number of inlet pressure settings. For the given circumstances, both the force F acting on the blade surface and the overall critical damping ratio ζ are scaled with the inlet pressure in the following way:

$$F(p) = F_{\text{ref}}\left(\frac{p}{p_{\text{ref}}} \right) \quad \zeta(p) = \zeta_M + \zeta_{A,\text{ref}}\left(\frac{p}{p_{\text{ref}}} \right) \tag{8}$$

The overall damping consists of two contributors, namely, the material ζ_M and the aerodynamic damping ζ_A of which only the latter is subject to scaling. An in-depth investigation of this problem was carried out and presented in Kammerer and Abhari [20]. For a SDOF system, the maximum response at resonance x_{res} is a function of the force and the damping. Substituting both in the equations given above, the response can be expressed as a function of inlet pressure

$$x_{\text{res}} = \frac{\frac{1}{k}F}{2\zeta} = \frac{\frac{1}{k}F_{\text{ref}}\left(\frac{p}{p_{\text{ref}}} \right)}{2\left(\zeta_M + \zeta_{A,\text{ref}}\left(\frac{p}{p_{\text{ref}}} \right) \right)} \tag{9}$$

There are essentially two conclusions to draw from this relationship. First of all, as the pressure approaches vacuum conditions, the force and the aerodynamic damping reduce to zero. Therefore, in this operating regime, material damping has a major influence on the relationship. Second, toward high pressure settings, the response amplitude is asymptotic since force and aerodynamic damping are scaled equally and mechanical damping becomes negligible.

7 Results on Strain Measurements

In the following sections, the results are presented and discussed for a number of different cases of excitation that can be distinguished depending on the distortion screen used, engine order excitation, and response mode. These are as follows.

(a) Response *without* distortion screens.

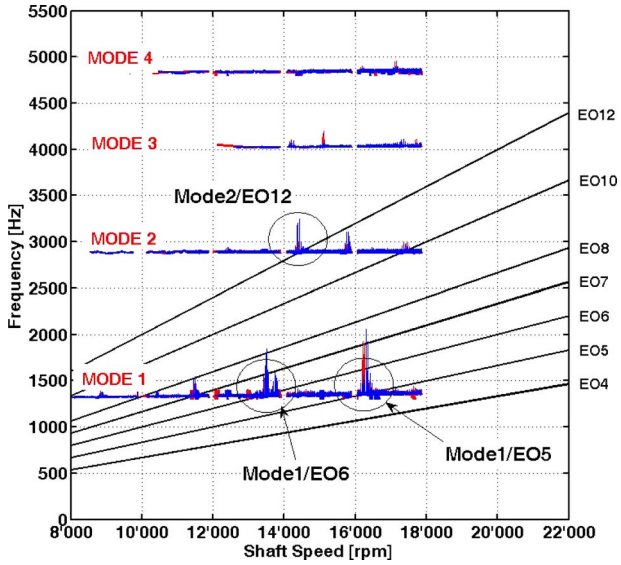


Fig. 10 Typical response of the main blade without a distortion screen installed upstream of the impeller

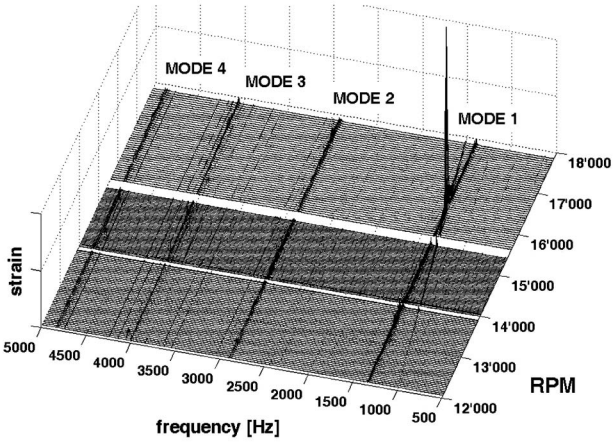


Fig. 11 Strain response spectrum for EO5 excitation using a five-lobe screen

(b) Mode1/EO5 at 16,250 rpm response is obtained from fundamental first harmonic excitation as generated by a five-lobe distortion screen. Here the response amplitude is the highest of all the cases considered since vibration is driven by the fundamental excitation frequency.

(c) Mode1/EO6 at 13,500 rpm response is obtained from the second harmonic excitation as generated by a three-lobe screen.

(d) Mode2/EO12 at 14,370 rpm response due to the third harmonic excitation as generated by a four-lobe screen.

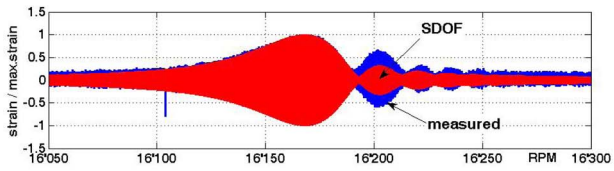
For the latter two cases, excitation is provided through harmonics, which are contained in the flow field in addition to the fundamental excitation corresponding to the screen design. The following presentation of the results focuses on the presentation of blade response in the time domain and is compared with a SDOF system model. Moreover the dependency of maximum response on inlet pressure and blade to blade variation is shown.

7.1 Strain Response Without Distortion Screens. Following the measurement procedure described previously, the main blade response was measured for an inlet configuration without a distortion screen. Figure 10 shows the Campbell diagram of a typical strain response for a maximum rotational speed of 18,000 rpm. The power spectral density of the response is plotted. Mode 1 can be identified as responding at 1350 Hz with a negligible increase in natural frequency of 1.5% over the speed range shown due to centrifugal stiffening. The response for Mode 2 occurs at 2875 Hz. The natural frequencies for Modes 3 and 4 are situated well above the 12th engine order excitation and will not be considered further.

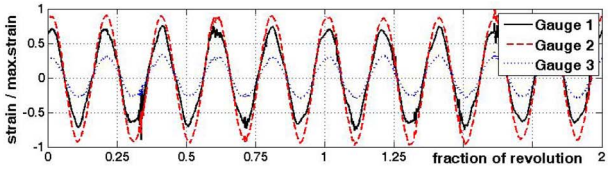
The major intention of this graph is to illustrate the potential of the given system to generate a number of resonant responses for a configuration without distortion screens. In particular, Mode 1 is excited by the 6th and the 5th engine order, whereas Mode 2 resonates with the 12th engine order. In accordance with inlet flow measurements presented previously for this configuration, the 6th and 12th engine order can be associated with harmonics introduced by upstream and downstream struts within the flow field. The strong response due to the fifth engine order is rather unexpected. Similarly, resonant responses from EO7 (Mode 1) and EO11 (Mode 2) were measured. The latter two cases prove the existence of excitation frequencies within the flow field, which might be a combination of the fundamental frequencies, i.e., EO3 and EO4. Despite the fact that the maximum strain amplitude generated by these engine orders is fractional in some cases compared with the fundamental frequencies, their very existence requires awareness of a potential contributor to high cycle fatigue. For the current research case, it has to be pointed out that second and third harmonic excitations are indeed generated within the flow field. An integer number combination of their fundamental frequencies might lead to additional excitation frequencies.

7.2 Mode1/EO5 Response. Initially, a number of resonant response properties for the data impeller are presented in detail for the crossing of Mode 1 with the fifth engine order excitation. The installed distortion screen consisted of five lobes. Figure 11 shows the acquired strain amplitude for the frequency spectrum and speed range of interest. Modes 1–4 can be clearly identified as previously presented. In addition, the excitation from the fifth engine order is sufficiently strong to remain traceable within the entire speed range before and after resonance occurs. According to Figs. 2(*a*) and 8, both the modal shape and the unsteady pressure distribution are in phase with the inducer portion of the impeller. This generally results in low aerodynamic damping and therefore relatively high strain amplitudes.

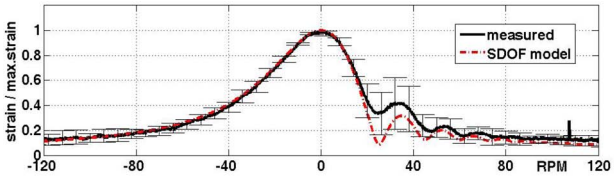
The dynamic response of a selected blade during transient measurement for a constant sweep rate and an inlet pressure setting of 0.1 bar is shown in Fig. 12(*a*). Complementary to the measured data, the calculated response is plotted using the recurrence formulas for the SDOF system introduced earlier. A crucial parameter for the computation is the critical damping ratio ζ, which was estimated from the experimental data using an amplitude-fit method. As the speed during the sweep increases, the response amplitude grows to a maximum at a rotational speed of approximately 16,170 rpm. Figure 12(*b*) shows the strain traces at resonance for all three strain gauges installed on a blade. The signals are in phase resembling a sinusoidal function. With further increase in speed, the response amplitude is modulated and its envelope is characterized by successive occurrence of lobes and nodal points. Using the Hilbert transformation, the envelope of the time signal is plotted in Figs. 12(*c*)–12(*e*) for inlet pressures of 0.1 bar, 0.2 bar, and 0.4 bar. The measured data were averaged for all available strain gauges after the Hilbert transformation. The error bars indicate the variance of the available samples used for averaging. For all three pressure cases, excellent agreement between the measurements and the SDOF dynamic model can be observed for the ascending portion of the first lobe and the resonance regime. Moreover, the sample variation is negligibly small before resonance, indicating very good agreement between the various



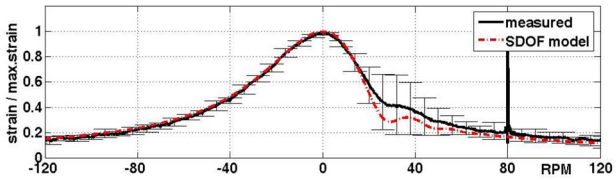
(a)Dynamic response at 0.1bar inlet pressure for main blade No. 7



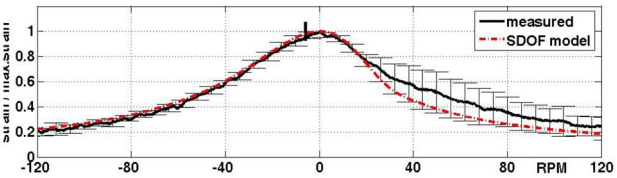
(b)Strain trace at resonance for two revolutions.



(c)Response envelope at 0.1 bar inlet pressure.



(d)Response envelope at 0.2 bar inlet pressure.



(e)Response envelope at 0.4 bar inlet pressure.

Fig. 12 Dynamic response envelope during Mode1/EO5 resonance

samples taken.

There are two parameters that influence the envelope shape, the critical damping ratio and the sweep rate. Although the latter was held constant, it should be mentioned that with an increase in sweep rate the number of lobes and nodes would increase after the maximum amplitude is achieved. The influence of the critical damping ratio on the envelope as a function of inlet pressure is apparent; as the pressure and therefore the damping increases, the main lobe widens and the second lobe diminishes. It is interesting to note an increase in variance after resonance. One main factor, in particular, is considered to cause the sample deviation for the 0.1 bar and 0.2 bar cases. Blade damping was observed to vary from blade to blade and therefore to modulate the amplitude of the dynamic response after resonance. This regime is rather sensitive to variation in damping. At 0.4 bar, these effects should diminish due to an increase in damping and therefore no comments can be made at this stage on potential causes for the rise in variance.

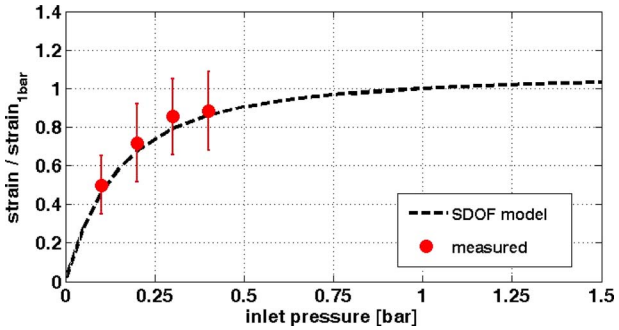


Fig. 13 Maximum response amplitude dependency on inlet pressure for Mode1/EO5

In accordance with Eq. (9), the maximum response amplitude was plotted in Fig. 13. Again, the error bars indicate the sample variation for all strain gauges considered. The SDOF curve fit was performed such that Eq. (9) was plotted using experimentally estimated damping properties, i.e., ζ_M and $\zeta_{A,\mathrm{ref}}$. The numerator $(1/k)F_{\mathrm{ref}}$ was then scaled in order to adjust for the amplitude match. The very good match between the SDOF model and the measurements essentially confirms the previous findings. Thus, for the given configuration, the blade dynamics during resonant vibration can be represented by a SDOF model.

Some light should be shed on the problem of sample variation. Figure 14 depicts the deviation of the maximum strain from the mean value used previously. The samples were acquired with strain gauges equally mounted on each blade. As can be observed, there is an offset from the mean for each blade, thus the variation is consistent and does not obey a random process. As such, the error bars are a measure of the deviation of maximum strain for the given blades. The consistent offset from the mean value could either be affected by strain gauges instrumentation, which can introduce a varying degree of damping or blade mistuning. Following the work by Whitehead [21], the maximum factor by which the amplitude of any blade could increase is $\frac{1}{2}(1+\sqrt{N})=1.8$, where $N=7$ is the number of cyclic sectors. This gives a considerable range of amplitudes within which response could be measured due to mistuning.

7.3 Mode1/EO6 Response. Resonant vibration was measured for Mode 1 using a three-lobe screen. The importance of this case lies in the fact that excitation was provided through the second harmonic of the inlet flow distortion circumferential profile. The same analysis approach was applied as in the previous case in order to compare the dynamic response with a SDOF model. Figure 15(*a*) shows the envelope of the strain measurements as the blade resonates for a 0.6 bar inlet pressure. Comparing the averaged envelope with the SDOF model, very good agreement can be observed. In contrast to the previous case presented, the plotted

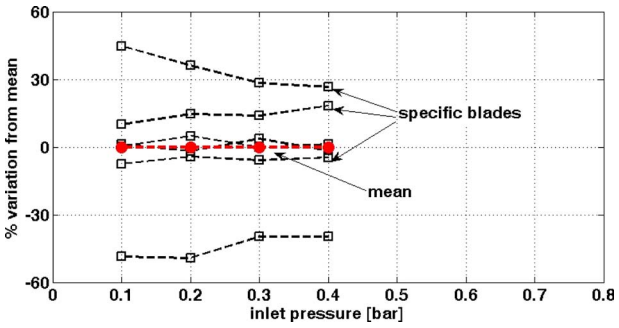
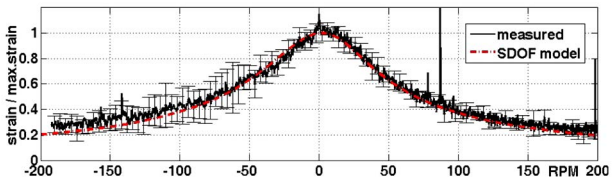
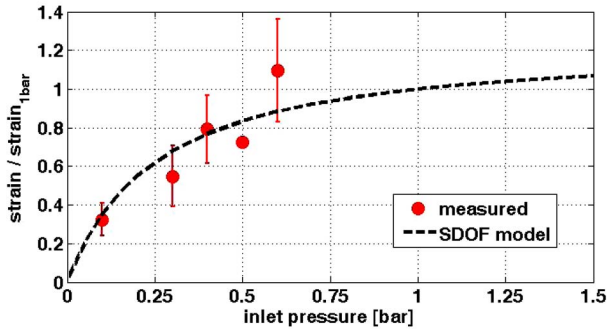


Fig. 14 Maximum strain variation from mean

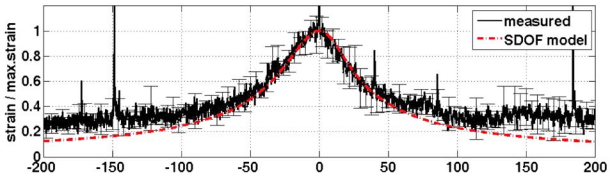


(a)Dynamic response envelope at 0.6 bar inlet pressure.

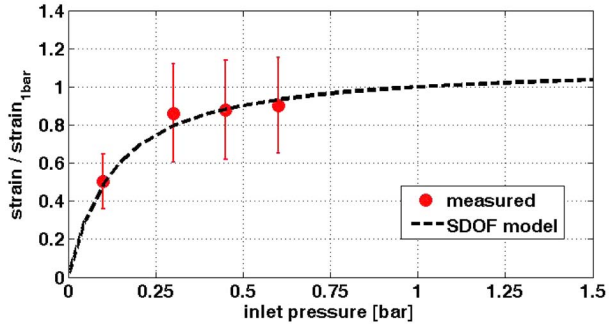


(b)Maximum response amplitude dependency on inlet pressure.

Fig. 15 Results for Mode1/EO6 case



(a)Dynamic response envelope at 0.6 bar inlet pressure.



(b)Maximum response amplitude dependency on inlet pressure.

Fig. 16 Results for Mode2/EO12 case

variance exhibits a considerable increase especially before resonance occurs. The dependency of the maximum response on the inlet pressure is given in Fig. 15(*b*). The general trend of the measured data is matched by the SDOF model. To understand the cause for the rather strong deviations of the mean value from the ideal model requires further investigations.

7.4 Mode2/EO12 Response. Resonant vibration of Mode 2 was achieved through the employment of the four-lobe screen generating the 12th engine order excitation. The second harmonic of the distortion pattern was found to be sufficiently strong to cause amplitudes approximately an order of magnitude smaller than for the Mode1/EO5 case. Figure 16(*a*) shows the averaged dynamic response for this case at a 0.6 bar inlet pressure. Also in the case of the second mode, the response under the given inlet conditions corresponds to the dynamics of a SDOF system. The variance is approximately constant for the entire speed range. The maximum response as a function of inlet pressure, Fig. 16(*b*), shows a good match for the available data points.

7.5 Strain Amplitude Comparison. A comparison between the maximum amplitudes should be given at this stage. Figure 17 compares the three cases discussed. Mode1/EO5 case shows the highest amplitudes since the excitation force in this case corresponds to the fundamental excitation frequency of the screen. As such, the velocity and therefore pressure fluctuations acting on the blade surface reach their maximum. For higher harmonics contained in the forcing functions, the amplitude decreases significantly. The maximum strain for the Mode1/EO6 is approximately 80% below the stress level experienced in the previous case. It should be pointed out that damping for the two cases is comparable. Further reduction in maximum amplitude is observed for the Mode2/EO12 case. From Part II of this paper, it is known that damping in this case is approximately double that of the Mode1 cases. The observed reduction in maximum amplitude is therefore affected by the increased damping. A statement on the amplitude of the excitation force cannot be made at this point.

8 Summary and Conclusions

Resonant response measurements were carried out for a radial compressor with design properties typical for turbocharging applications. The blade thickness of the main blade was adjusted such that resonant crossing with a number of low engine order excitations was guaranteed. FEM was used during the design procedure in order to estimate the modal properties. The final design was in agreement with experimentally estimated modal frequencies and shapes using ESPI. The impeller inlet section was equipped with a strut and tube assembly accommodating the signal transmission unit, interchangeable flow distortion screens, and a traverse mechanism, which enabled the measurements of flow properties using an aerodynamic probe. Distortion screens were designed targeting a specific flow pattern such that resonant crossings with the first and second main blade natural frequencies were possible. The velocity deficit achieved by the employment of the distortion screens was then measured and compared with the predictions. Vibration measurements were undertaken using dynamic strain gauges mounted on the blade surfaces. Vibrational response was acquired by sweeping through the resonant regime with a constant sweep rate. Finally, the results were analyzed in the time domain and then compared with the dynamic response of a SDOF model.

Traverse measurements upstream of the impeller were carried out in radial and circumferential directions for both cases, with and without distortion screens. The latter case showed the strut arrangement within the inlet section to significantly affect the flow field. Three upstream struts generated zones of elevated axial velocities within the bulk flow with local deviations of ≈8% from

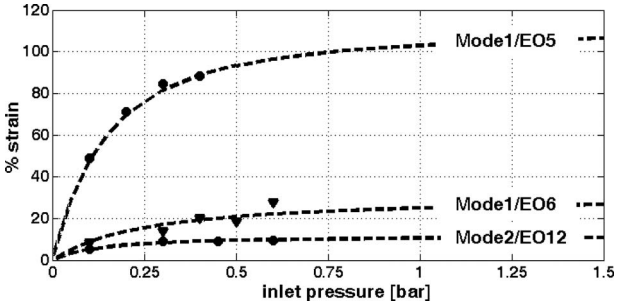


Fig. 17 Comparison of maximum strain amplitude

the mean axial velocity. At the root of these struts, the low momentum fluid accumulated. Four vane-covered downstream struts generated additional wakes. These findings show the potential of the system arrangement to generate the third and the fourth engine order excitations. This was then confirmed during subsequent strain gauge measurements, which also revealed that additional excitation frequencies were contained in the flow field and were an integer number combination of the third and fourth engine order excitations. Moreover, first mode response to the fifth engine order excitation was found to be strongest for this arrangement.

The flow field measurements using distortion screens confirmed the intended mean velocity variation upstream of the compressor. However, in both exemplified cases, the four-lobe and the five-lobe screen, the flow field was locally modulated by the strut arrangement. For a given mass flow rate, these effects gave rise to regions where low momentum fluid was concentrated and therefore the fluid was forced to redistribute and generate local maxima in axial velocity. The maximum velocities of $\approx 20\%$ above the mean level were achieved. As a result of these modulations, the flow field contained a number of excitation frequencies in addition to the intended fundamental frequency based on the number of lobes. An unsteady CFD simulation confirmed these findings using boundary conditions as measured with an aerodynamic probe.

Resonant vibration results were presented for the three cases. For the first case, Mode1/EO5, excitation was provided by the fundamental frequency as generated by the distortion screen with five lobes. In the second (Mode1/EO6) and third (Mode2/EO12) cases, excitation was provided through the second and the third harmonic of the corresponding three-lobe and four-lobe screen. The results seen in comparison with the computation of a SDOF model show very good agreement with respect to the amplitude dynamics as sweeping through the resonant regime occurred. These findings were confirmed for a number of inlet pressure settings, which altered the damping and therefore the dynamics of the blade response.

The chosen approach in this work to experimentally investigate the blade dynamics during resonance clearly revealed that a SDOF system can be used to model the problem. However, an important parameter, the overall critical damping ratio, has to be estimated. This problem is discussed in Part II of this paper. The rather complex flow generated upstream of the impeller due to mechanical obstructions revealed its potential to generate a number of excitation frequencies. Similar scenarios can be found in industrial turbocharging or pumping applications where inevitable integration of pipe bends, struts, or flow injection may introduce disturbances. For current and future applications, it is vital to understand and predict the mechanical implications of these circumstances. In this regard, computational tools are necessary in addition to the understanding of their capabilities to actually model a given problem. The data acquired in this experimental work target the provision of a basis to validate computational models. In the future, the existing data should be complemented by unsteady pressure measurements in order to gain further insight into the transition from inlet flow to unsteady forces and the resultant dynamic response.

Acknowledgment

The authors gratefully acknowledge the intellectual and financial support of their industrial partners ABB Turbo Systems and MAN TURBO AG Schweiz in this project. The contributions by Dr. Peter Sälzle and Dr. Matthias Schleer from ABB Turbo Systems and Dr. Beat Ribi and Urs Baumann from MAN TURBO AG Schweiz are greatly acknowledged. The authors also acknowledge the financial support provided by the Swiss Confederation's innovation promotion agency, CTI. Experiments on ESPI were conducted at the Swiss Materials Science & Technology Research Institute (EMPA) by Dr. Erwin Hack. The authors also thank Armin Zemp for helping to setup FRAP measurements and Cornel Reshef for his work in developing the electronic instrumentation for signal transmission and acquisition.

Nomenclature

A = factor depending on Reynolds number
B = shape factor
EO = engine order
F, F_0 = forcing function, force amplitude
M = mesh width
$S_{\max}$ = maximum sweep rate
V_{ax} = axial velocity
c = viscous damping coefficient
d = wire diameter
f = frequency
k = spring constant
m = mass
p = pressure
Δp = pressure difference
q = dynamic head
t = time
$x, \dot{x}, \ddot{x}$ = displacement, speed, acceleration
α = area ratio
β = grid porosity function or frequency gradient
ϕ_0 = phase angle
$\zeta, \zeta_M, \zeta_A,$ = critical damping ratio, material, aerodynamic
$_{\text{mean}}$ = mean velocity
$_{\text{ups}}$ = upstream of screen
$_{\text{unb}}$ = downstream of screen in the unblocked area
$_{\text{blo}}$ = downstream of screen in the blocked area
$_0$ = time independent or reference quantity
$_{\text{ref}}$ = reference quantity
$_{\text{res}}$ = at resonance
$_{\text{in,out}}$ = at inlet, outlet of the impeller

References

[1] El-Aini, Y., deLaneuville, R., Stoner, A., and Capece, V., 1997, *Proceedings of the 33rd AIAA/ASME/SAE/ASEE Joint Propulsion Conference*, Seattle, WA, Paper No. AIAA 97–3365.
[2] Kielb, R. E., 1998, *ERCOFTAC Turbomachinery Seminar and Workshop*, Aussois, France.
[3] Haupt, U., 1984, *Untersuchung des Schaufelschwingungsverhaltens Hochbelasteter Radialverdichterlaufräder* (*Fortschrittsberichte der VDI Zeitschriften*), Vol. 7, Verein Deutscher Ingenieure VDI-Verlag GmbH Düsseldorf, Hannover.
[4] Jin, D., 1990, "Untersuchung von Schaufelschwingungen und ihrer Erregungsursachen an Radialverdichtern," Ph.D. thesis, Hannover.
[5] Manwaring, S. R., and Fleeter, S., 1990, "Inlet Distortion Generated Periodic Aerodynamic Rotor Response," ASME J. Turbomach., **112**, pp. 298–307.
[6] Manwaring, S. R., and Fleeter, S., 1991, "Forcing Function Effects on Rotor Periodic Aerodynamic Response," ASME J. Turbomach., **113**, pp. 312–319.
[7] Manwaring, S. R., Rabe, D. C., Lorence, C. B., and Wadia, A. R., 1997, "Inlet Distortion Generated Forced Response of a Low-Aspect-Ratio Transonic Fan," ASME J. Turbomach., **119**, pp. 665–676.
[8] Manwaring, S. R., and Wisler, D. C., 1993, "Unsteady Aerodynamics and Gust Response in Compressors and Turbines," ASME J. Turbomach., **115**, pp. 724–740.
[9] Dickmann, H. P., Wimmel, T. S., Szwedowicz, J., Filsinger, D., and Roduner, C. H., 2006, "Unsteady Flow in a Turbocharger Centrifugal Compressor: Three-Dimensional Computational Fluid Dynamics Simulation and Numerical and Experimental Analysis of Impeller Blade Vibration," ASME J. Turbomach., **128**, pp. 455–465.
[10] Schleer, M., Mokulys, T., and Abhari, R. S., 2003, International Conference on Compressors and Their Systems, IMechE, London.
[11] Schleer, M., 2006, *Flow Structure and Stability of a Turbocharger Centrifugal Compressor*, VDI-Verlag, Düsseldorf.
[12] Roach, P. E., 1987, "The Generation of Nearly Isotropic Turbulence by Means of Grids," Int. J. Heat Fluid Flow, **8**, pp. 82–92.
[13] Koo, J. K., and James, D. F., 1973, "Fluid Flow Around and Through a Screen," J. Fluid Mech., **60**, pp. 513–538.
[14] Pfau, A., Schlienger, J., Kalfas, A. I., and Abhari, R. S., 2003, ASME Paper No. GT2003-38128.
[15] Kupferschmied, P., 1998, "Zur Methodik zeitaufgelöster Messungen mit Strömungssonden in Verdichtern und Turbinen," Ph.D. thesis, ETH Zürich, Zürich.

[16] Zemp, A., Kammerer, A., and Abhari, R. S., 2008, ASME Paper No. GT2008-50744.
[17] Szwedowicz, J., Senn, S. M., and Abhari, R. S., 2002, "Optimum Strain Gage Application to Bladed Assemblies," ASME J. Turbomach., **124**, pp. 606–613.
[18] Ewins, D. J., 2000, *Modal Testing Theory, Practice and Application*, 2nd ed., Research Studies, Baldock, Hertfordshire, England.
[19] Craig, R. R., and Kurdila, A. J., 2006, *Fundamentals of Structural Dynamics*, 2nd ed., Wiley, Hoboken, NJ.
[20] Kammerer, A., and Abhari, R. S., 2008, ASME Paper No. GT2008-50467.
[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.

Experimental Study on Impeller Blade Vibration During Resonance—Part II: Blade Damping

Albert Kammerer

Reza S. Abhari

LEC, Laboratory of Energy Conversion, Department of Mechanical and Process Engineering, ETH Zurich, 8092 Zürich, Switzerland

Forming the second part of a two-part paper, the estimation of damping is presented here. Part I discusses the experimental approach and the results on blade resonant response measurements. In the study of forced response, damping is a crucial parameter, which is measured to quantify the ability of a vibrating system to dissipate vibratory energy in response to a given excitation source. The blading of turbomachinery components is particularly prone to forced response excitation, which is one of the major causes of high cycle fatigue failure during operation. In turbocharging applications, forced response cannot be avoided due to a number of factors, i.e., change in speed, inlet bends, or obstructions in the flow field. This study aims to quantify the damping parameter for the lightly damped blades of a centrifugal compressor. The impeller geometry is typical of turbocharging applications. As a first step, the nonrotating impeller was excited using piezos, and the transfer function was derived for a number of pressure settings. Both circle-fit and curve-fit procedures were used to derive material damping. In the second step, measurements were taken in the test facility where forced response conditions were generated using distortion screens upstream of the impeller. The main blade strain response was measured by sweeping through a number of resonant points. A curve-fit procedure was applied to estimate the critical damping ratio. The contributions of material and aerodynamic dampings were derived from a linear curve-fit applied to the damping data as a function of inlet pressure. Overall, it will be shown that aerodynamic damping dominates the dissipation process for applications with an inlet pressure of 1 bar. Damping was found to depend on the throttle setting of the compressor, and where applicable computational fluid dynamics results were used to point toward the possible causes of this effect. [DOI: 10.1115/1.2968870]

1 Introduction

In turbomachinery, blades suffer from flutter or forced response depending on the type of application and the flow conditions. The prediction of vibratory response amplitudes due to these phenomena is crucial for lifetime assessment during the design phase and relies on the capability to predict the cause and magnitude of blade excitation and to assess damping mechanisms. The latter is quantified as overall damping, which is generally composed of material, mechanical, and aerodynamic dampings. The contribution of each of these damping mechanisms depends on the associated blade design properties, i.e., material, fixation, and shape. In the past, the vast majority of work and publications in the field of vibration and the quantification of damping were carried out for axial turbines and compressors. A summary of experimental approaches and problem modeling was presented by Srinivasan [1,2]. With regard to damping, research has focused on the estimation of mechanical damping as it is the sole damping property that can be intentionally influenced by the designer. Kielb and Imregun [3] outlined the contributors to damping and presented the characteristics of mechanical damping. Material damping was generally found to be minimal and could be neglected during design [4].

For radial turbomachinery, where rotors are machined from a single piece, structural damping is very low and aerodynamic damping is therefore the dominant damping mechanism. Experimental measurement of aerodynamic damping can be carried out either by cascade testing or in a rotating facility. Owing to the nature of cascade testing, this experimental approach to obtain aerodynamic damping values cannot be applied for three-dimensional radial compressor blades. However, available data on aerodynamic damping from rotating facilities are limited and are focused on axial machines. An estimation of aerodynamic damping was performed by Crawley [5] for a transonic compressor, using upstream disturbances and measuring blade response. A work on damping for a centrifugal compressor was presented by Jin et al. [6] where excitation was provided by partially blocking the diffuser. In facilities where the inlet pressure can be adjusted, aerodynamic damping is obtained by measuring the overall damping and then subtracting nonaerodynamic damping measured at vacuum conditions. Kielb and Abhari [7] used this approach to separate the contributions of mechanical and aerodynamic dampings. Their experimental approach is based on the findings described by Jeffers et al. [8]. Herein, the complexity is to provide blade excitation. During nonvacuum operation, excitation can either be generated by screens or blade rows, whereas at vacuum, excitation requires installation of mechanical actuators, i.e., piezoelectric actuators. Crucial for jet engines is the dependency of aerodynamic damping on inlet pressure, which is dependent on flight altitude. Newman [9] estimated the mode dependent aerodynamic damping in an axial transonic compressor for a number of inlet pressures. From the linear fit, structural damping was derived by extrapolating to vacuum conditions.

Parallel to experimental work, computation of flow and structure is performed in order to estimate maximum strain amplitudes

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 8, 2008; final manuscript received April 8, 2008; published online December 29, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008 Land, Sea and Air (Gt2008), Berlin, Germany, June 9–13, 2008.

within the structure. Two approaches are generally applied. Either a fully fluid structure coupled computation is performed or the two domains are treated separately. In the former case, aerodynamic damping is the result of computing damping work due to blade motion, whereas in the latter case, the procedure requires an estimate of aerodynamic damping as an input parameter. Examples of fully coupled or decoupled computation can be found in Chiang and Kielb [10], Li and He [11], Moffat and He [12], Ning et al. [13], Dickmann et al. [14], and Schmitt [15].

2 Motivation and Scope of the Paper

The work presented here aims to provide experimental data on damping estimates for a centrifugal compressor, which is made of aluminum and features thin blades. The impeller geometry is typical for turbocharging applications. The magnitude of material damping and the mode dependent aerodynamic damping as a function of inlet pressure will be discussed. Data will be compared for two operating lines, both of which are situated in the stable operating regime of the impeller.

3 Experimental Approach

The major objective of this work is to quantify the critical damping ratio and subsequently to estimate the contributions of material and aerodynamic dampings to the overall damping. Two experimental studies were carried out. First of all, aerodynamic damping measurements were performed within the test facility, and blade excitation was generated by distortion screens. Second, the impeller was bench mounted and a piezo applied to the impeller provided blade excitation. In both cases, pressure was initially set to near vacuum conditions and was then increased. For each pressure setting, damping was estimated and material damping could be estimated from extrapolating to vacuum where the contribution of aerodynamic damping was zero.

3.1 Strain Measurements With Aerodynamic Excitation. Strain gauge measurements with aerodynamic excitation were performed according to the experimental procedure presented in Part I of this paper. Distortion grids installed upstream of the impeller generated pressure fluctuations on the blade surface. The results presented in this part of this paper will be compared for the two operating lines shown in Fig. 1(*a*). The first operating line (OL1) corresponds to a realistic operating line set during operation, whereas the second operating line (OL2) is shifted toward higher mass flows. In both cases, the maximum rotational speed was 18,000 rpm. According to the Campbell diagram (Fig. 1(*b*)), the first two main blade modes could be excited by low engine order excitations. This was achieved by employing screens with three, four, and five lobes, which generate the respective frequencies for resonant response. It should be noted that the Campbell diagram shown aims to illustrate the resonant crossings, and the response amplitudes were obtained from measurements without an installed distortion screen. The results on damping estimation will be presented for the following resonant crossings: Mode1/EO5, Mode1/EO6, Mode2/EO12, and Mode2/EO10.

3.2 Strain Measurements With Piezoelectric Excitation. During strain measurement using a piezoelectric exciter, the same strain gauges were used as during measurements with aerodynamic excitation. Two piezos were mounted on the back of the impeller disk. The measurement setup required the physical separation of the wires used to supply high voltages to the piezo and the wires conducting the signal. Because of coupling between the disk and the blade, sufficient excitation was transmitted through the material in order to measure blade response. Both piezos were simultaneously driven by an amplifier with the excitation frequency provided by a function generator. Sweep excitation was performed for each mode separately within its natural frequency boundaries. For a single measurement point, sweeping was repeated and the response averaged after the windowed Fourier transformation. The entire response spectrum for a maximum frequency of 5500 Hz is shown in Fig. 2. The first four modes can be identified. Their frequencies correspond to the frequencies obtained during testing with aerodynamic excitation. It should be pointed out that the force exerted by the piezo scales linearly with the excitation frequency. Therefore, the response amplitude of Mode 1, in particular, can be observed to be relatively small in comparison to other modes. For the same reason, Mode 1's response was observed to be affected by the relative distance be-

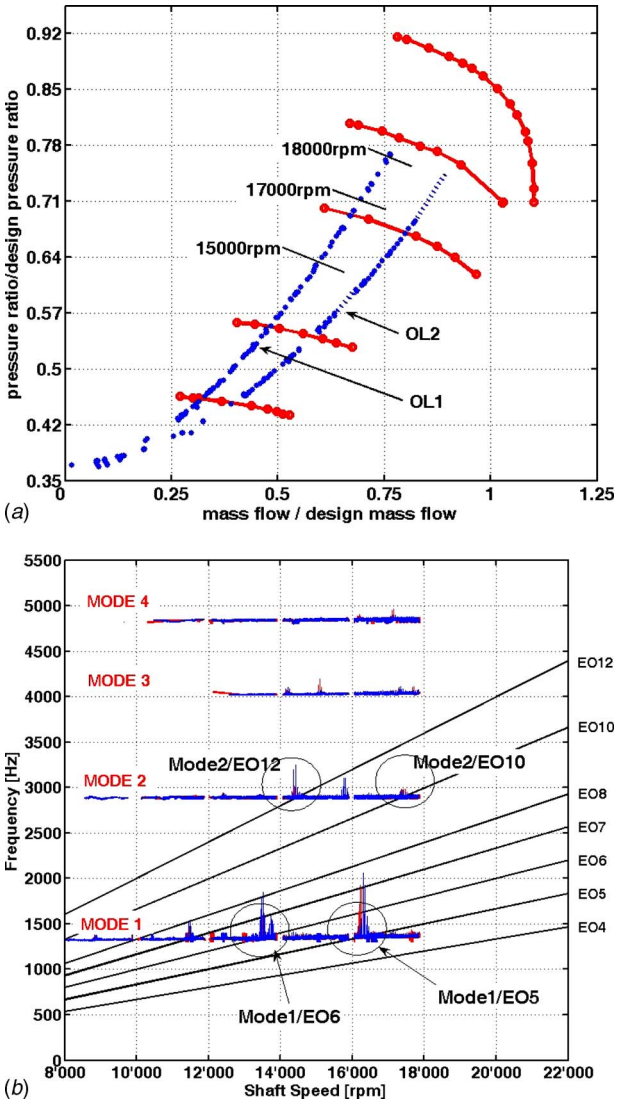


Fig. 1 Impeller performance map and Campbell diagram. (*a*) Compressor map and operating lines. (*b*) Typical response of the main blade without a distortion screen installed upstream of the impeller.

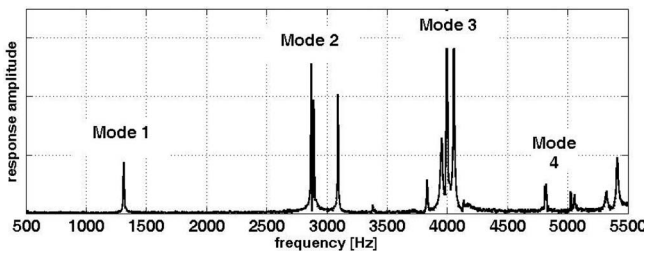


Fig. 2 Full range response spectrum with piezoelectric excitation

tween the piezo and a specific blade, i.e., blades further away from the piezo exhibited smaller amplitudes. Moreover, with an increase in pressure, damping increased and measurements of the first mode above 0.8 bar could not be used for damping estimation.

4 Damping

In general, three main contributors to damping can be identified for turbomachinery blades: material, mechanical, and aerodynamic damping. Briefly, material damping is a material property, which is measured by the energy dissipated during cyclic strain in the material. This process always has a hysteresis loop where the dissipated energy and therefore the damping depend on the amplitude of cyclic strain. Mechanical damping accounts for energy dissipated during contact friction between components. Typical applications for blades can be found at the blade root or through employment of snubbers. The material damping magnitude depends on the contact geometry and the contact pressure between the parts. Aerodynamic damping accounts for vibratory energy dissipated due to the relative motion between the blade and the fluid. The main dependency parameters are the fluid density, the blade mode shape during vibration, and the phase relation between the forcing function and the blade mode shape. The relative contribution of the three damping mechanisms strongly depends on the kind of application and design. However, in the majority of cases, material damping of turbomachinery blades is comparatively small and therefore its contribution is often neglected. Thus, mechanical and aerodynamic dampings are the main contributors to the overall damping. The former allows tuning through design measures.

For radial compressors as used in this study, mechanical damping is not present due to the lack of component friction. Damping during resonant vibration is therefore only provided through material and aerodynamic dampings. In the following, damping will be mathematically accounted for by the critical damping ratio ζ according to a vibrating SDOF system with viscous damping. In this case, the overall viscous force F_D acting on a structure is expressed by

$$F_D = -c\dot{x} = -2\zeta m\omega_n \dot{x} \quad \text{with} \quad \zeta = \frac{c}{2\omega_n m} \tag{1}$$

Accounting for contribution from mechanical and aerodynamic dampings the corresponding forces F_{DM} and F_{DA}, the total force is given as

$$F_D = F_{\mathrm{DM}} + F_{\mathrm{DA}} \tag{2}$$

$$= -2\zeta_M m\omega_n \dot{x} + F_{\mathrm{DA}} \tag{3}$$

Herein ζ_M is the critical damping ratio due to material damping. This quantity is assumed to remain constant. The damping force due to aerodynamic damping is dependent on the density, and therefore the inlet pressure was a parameter that was altered in this project. Scaling the aerodynamic damping force with inlet pressure, the following can be stated:

$$F_{\mathrm{DA}} = F_{\mathrm{DA,ref}}\left(\frac{p}{p_{\mathrm{ref}}}\right) \tag{4}$$

$$= -2\zeta_{A,\mathrm{ref}}\left(\frac{p}{p_{\mathrm{ref}}}\right) m\omega_n \dot{x} \tag{5}$$

Finally, a relationship for the overall damping can be derived as a function of inlet pressure.

$$F_D = F_{\mathrm{DM}} + F_{\mathrm{DA}} \tag{6}$$

$$-2\zeta m\omega_n \dot{x} = -2\zeta_M m\omega_n \dot{x} - 2\zeta_{A,\mathrm{ref}}\left(\frac{p}{p_{\mathrm{ref}}}\right) m\omega_n \dot{x} \tag{7}$$

$$\zeta = \zeta_M + \zeta_{A,\mathrm{ref}}\left(\frac{p}{p_{\mathrm{ref}}}\right) \tag{8}$$

For the current study, it is more convenient to apply the damping function in the following form:

$$\zeta(p) = \zeta_M + \left(\frac{d\zeta_A}{dp}\right) p \tag{9}$$

At vacuum conditions, only the material damping provides a means for vibratory energy dissipation. In an experimental approach, this property can be obtained by measuring the overall damping for a number of inlet pressure settings and then extrapolating to zero inlet pressure. For completeness, the maximum response during resonant vibration of a SDOF system is a function of inlet pressure and is expressed through

$$x = \frac{\frac{1}{k}F_{\mathrm{ref}}\left(\frac{p}{p_{\mathrm{ref}}}\right)}{2\left(\zeta_M + \zeta_{A,\mathrm{ref}}\left(\frac{p}{p_{\mathrm{ref}}}\right)\right)} \tag{10}$$

Blade response due to inlet distortion under varying inlet pressure conditions was discussed in Part I by Kammerer and Abhari [16].

5 Data Reduction

5.1 Frequency Analysis. The transient data acquisition procedure applied in this work requires the analysis of the signal with respect to both frequency and rotational speed. This is achieved through the short-time-Fourier-analysis (STFT) where the recorded signal is divided into successive windows and then Fourier transformed. Thus each window is associated with the speed range during which it was acquired. Processing the signal this way allows the identification of the change in frequency content as the shaft speed sweeps through resonance. The choice of window length and the resultant minimum frequency resolution is limited according to the uncertainty principle [17]. Owing to the nature of this problem, the window length was adjusted depending on the purpose of the analysis. The signal analysis used to perform damping estimation through amplitude fitting requires high resolution and was therefore carried out such that a frequency resolution of typically 0.2 Hz was achieved. In this case, a dense resolution of the resonant peak was required. In order to visualize the frequency content during resonant sweeping, the window length was reduced resulting in a typical resolution of 2 Hz. Attention was paid to resolving the frequency content with respect to time.

5.2 Curve-Fit Procedure. Evaluation of the critical damping ratio was performed using a curve-fit procedure, as described below. Applications of these methods were carried out by Newman [9], Jeffers [8], and Kielb and Abhari [7]. In this work, the frequency response function $Y(\omega)$ was obtained by averaging the measured signal in the frequency domain. The transfer function in the frequency domain $H(\omega)$ is defined as

$$H(\omega) = \frac{Y(\omega)}{X(\omega)} \tag{11}$$

The data to compute the excitation function $X(\omega)$ would require measurements of the unsteady pressure fluctuation on the blade surface, and these are not available. This function can be assumed to be constant. Computing the spectral density S_y from the response function and assuming the excitation spectral density to be constant, the amplitude response $|H(\omega)|^2$ of a SDOF system is given by the following function:

$$|H(\omega)|^2 = \frac{S_y}{S_x} = \frac{1/k^2}{[1-(\omega/\omega_n)^2]^2 + [2\zeta(\omega/\omega_n)]^2} \tag{12}$$

On the basis of this model, the measured data points were curve-fitted by adjusting the numerator k, the frequency ratio ω/ω_n, and the critical damping ratio ζ. In order to minimize the variance

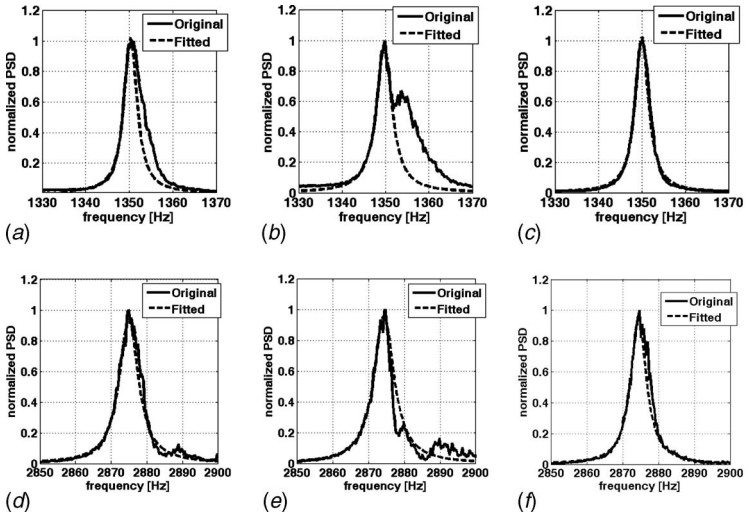


Fig. 3 Examples for amplitude fitting: ((*a*)–(*c*)) Mode 1 and ((*d*)–(*f*)) Mode 2. (*a*) Blade No. 1, (*b*) Blade No. 2, (*c*) Blade No. 3, (*d*) Blade No. 1, (*e*), Blade No. 2, and (*f*) Blade No. 3.

between the curve-fit and the data points, a least squares method was used.

A number of examples of the employment of the fitting procedure for Mode 1 are shown in Figs. 3(*a*)–3(*c*), and for Mode 2, in Figs. 3(*d*)–3(*f*). For both cases, measurements for the first three blades are compared. The data were nondimensionalized by the maximum amplitude of the spectral density. A major problem in the application of the curve-fit method in the current study is the occurrence of additional frequencies or frequency maxima within the considered bandwidth. In particular, Blade No. 2 exhibits this problem, exemplified here for the first mode. The data suggest that coupling and mistuning affect the observed results and will be illustrated later. A remedy to overcome this problem is shown for Blade Nos. 1 and 2 for both modes. Data fitting was performed with emphasis on the ascending part of the resonance peak by limiting the frequency band. In other cases, i.e., Blade No. 3, mistuning and coupling were not observed, enabling the curve-fit procedure to be applied for the entire frequency range.

5.3 Circle-Fit Procedure. In addition to the curve-fit procedure, a circle-fit procedure was implemented, aiming to verify the estimation of damping. The circle-fit method was applied to results obtained from impeller bench testing and piezoelectric excitation. The main advantage of this procedure is the synchronized acquisition of excitation and response. Based on this, the frequency transfer function $H(\omega)$ can be computed from the cross-correlation spectrum S_{xy} and the spectral density S_x:

$$H(\omega) = \frac{S_{xy}}{S_x} \tag{13}$$

The response amplitude as a function of frequency is exemplified for Mode 2 in Fig. 4(*a*). The real and imaginary components of

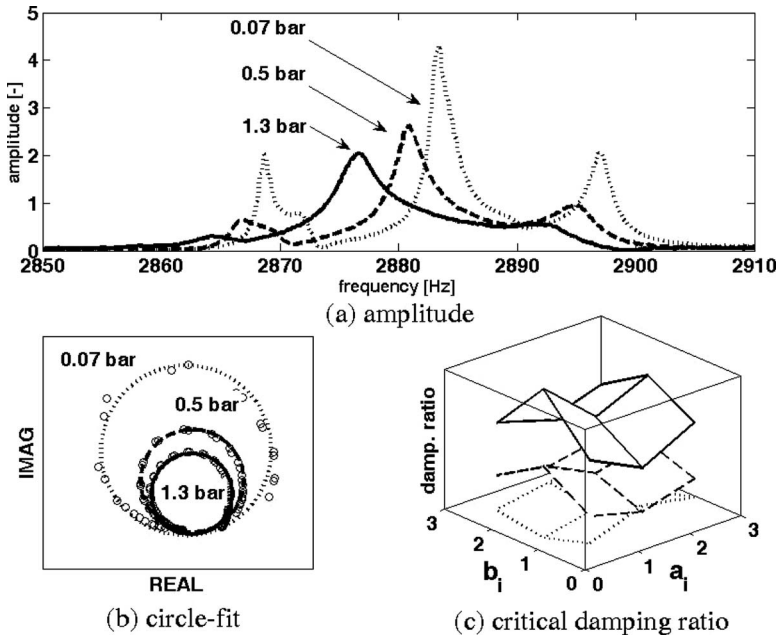


Fig. 4 Circle-fit method for damping estimation

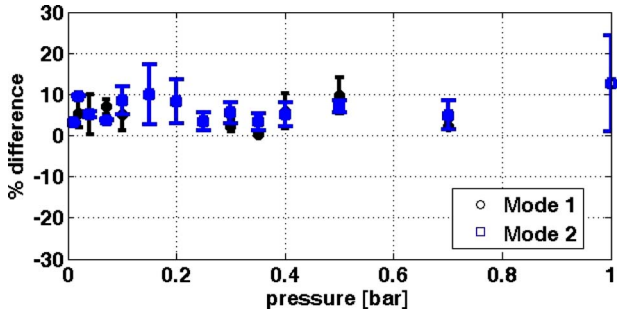


Fig. 5 Deviation in damping estimation, $\%(\zeta_{curve}-\zeta_{circle})$

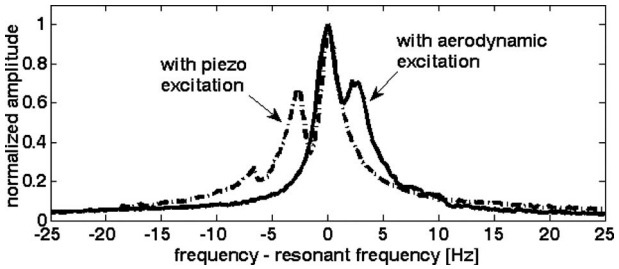


Fig. 6 Response comparison at 0.1 bar inlet pressure

the complex response function are shown in Fig. 4(*b*) for a frequency bandwidth where resonance occurs. Due to coupling between the blade sectors, three maxima can be observed in the figure of which the central with the highest amplitude was selected for damping estimation. Both graphs show the dependency of the transfer function on the inlet pressure. A circle-fit was carried out in the complex domain in order to determine the center of the data points. Following the derivations given by Ewins [18], the critical damping ratio of a SDOF system can be computed from the following equation:

$$\zeta = \frac{\omega_a^2 - \omega_b^2}{2\omega_n\left[\omega_a \tan\left(\frac{1}{2}\Theta_a\right) + \omega_b \tan\left(\frac{1}{2}\Theta_b\right)\right]} \tag{14}$$

Herein the indices a and b signify the data points before and after resonance, respectively. The angles Θ_a and Θ_b measure the angles between the resonance point and a data point a_i and b_i, respectively. Damping can therefore be calculated for a number of combinations of a_i and b_i. Figure 4(*c*) shows the estimated critical damping ratio based on this approach. Each plane represents damping for a pressure setting, and the values are averaged to obtain the final damping. The employment of this procedure revealed that a maximum of six data points from resonance was applicable. Ideally this number should be higher: however, in the current application the "carpet" plots showed significant distortion for data points further away from resonance.

5.4 Comparison Between Curve-Fit and Circle-Fit Methods. A comparison between the curve-fit and the circle-fit methods was carried out based on measurements performed with piezoelectric excitation. Figure 5 quantifies the percentage of $(\zeta_{curve}-\zeta_{circle})$ as a function of pressure. Modes 1 and 2 are shown. The error bars indicate the sample variance. The mean deviation for both modes is within a range of 10%. The circle-fit method slightly underestimates the damping magnitude in comparison to the curve-fit approach. The applicability of each of the two methods was found to depend on the available data quality, which in this case was affected by the blade-to-piezodistance. On average, the blades closer to the piezo exhibited higher response amplitudes. A systematic deviation, i.e., with an increase in pressure, cannot be observed. Based on the findings, the overall agreement was considered to be good. Results shown later for the piezoelectric excitation case were obtained from averaging the estimates from both methods.

6 Results and Discussion

6.1 Blade Coupling and Mistuning. The problem of coupling and mistuning should be briefly mentioned here. The topic of coupled vibration was treated for a similar type of impeller geometry by Hasemann et al. [19] and Hagelstein et al. [20], applying holographic interferometry and finite element method (FEM). At the current stage, the level of mistuning and coupling will not be quantified: however, the following considerations should be taken into account. Mechanical coupling through the disk is present and can be judged on using FEM as well as data obtained using speckle interferometry. Blade mistuning is expected to be present due to manufacturing tolerances as well as the application of strain gauges onto the blade surfaces. Generally, for low inlet pressure, i.e., below 0.2 bar, the coupling and mistuning effects were pronounced and were observed as multiple peaks in the response spectrum of a single blade. With an increase in pressure, these peaks diminished, whereas only the blade specific response remained. Furthermore, the response spectrum was found to depend on the excitation mechanism. Figure 6 compares the response spectrum from aerodynamic and piezoelectric excitations for the first mode. In both cases, the inlet pressure was 0.1 bar. For the same resonant frequency, the appearance of shifted frequencies differs. The major difference between these two experiments is the phase angle of the applied excitation force. In the case of aerodynamic excitation, the force phase angle depends on the relative angular position between the blade and the distortion pattern. This way blade-to-blade excitation is out of phase. In the case of piezoelectric excitation, however, the experienced force is in phase for all blades: hence, they are excited simultaneously.

6.2 Damping Measurement for Piezoelectric Excitation. Damping estimation was carried out for the case with piezoelectric excitation. Both the curve- and circle-fit methods were applied. The resultant mean critical damping ratio between the two methods is shown in Fig. 7 for Modes 1 and 2. The critical damping ratio was normalized by the estimated material critical damping. The error bars indicate the sample variance of estimated damping. The major goal of this experiment was to determine material damping by extrapolating a linear curve-fit to vacuum conditions where aerodynamic damping is eliminated. For this reason, measurements were refined at near vacuum. Although in the case of Mode 1 measurements were taken above an inlet pressure of 0.4 bar, the application of the damping estimation proce-

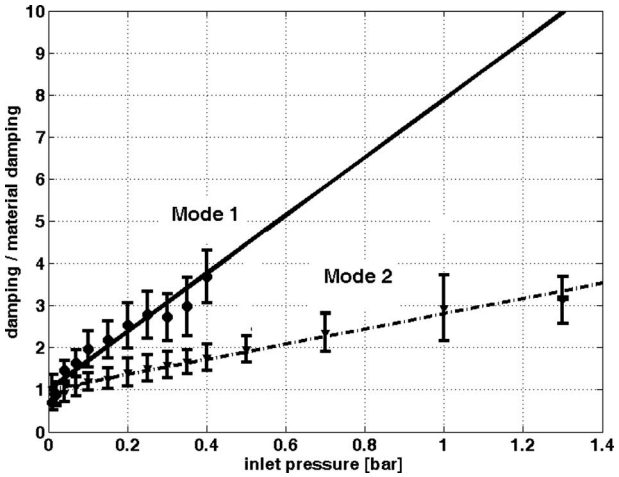


Fig. 7 Estimated damping for Modes 1 and 2 using piezoelectric excitation

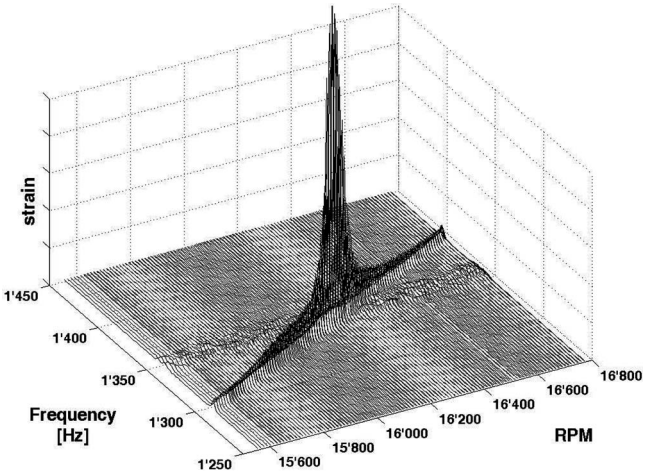


Fig. 8 Mode 1 response from EO5 excitation using a five lobe screen

dures was uncertain due to very low response amplitudes. As mentioned previously, the excitation force exerted by the piezo at low frequencies was a limiting factor. This was not the case for the second mode. This situation is also reflected by the error bars. In the case of Mode 2 as compared with Mode 1, sample variation is generally lower and deviation from the linear fit is negligible. Despite the lower data quality acquired for Mode 1, for both cases, a linear relationship between the critical damping ratio and the inlet pressure was obtained. The extrapolation of the linear fit to the vacuum yielded the critical damping ratio ζ_M due to material damping. With regard to Mode 1, material damping was found to be approximately 40% higher in comparison with Mode 2. Though this value appears to be considerable, in absolute numbers the influence is rather small considering the increase in aerodynamic damping as the inlet pressure was increased. Similar findings were reported by Srinivasan et al. [4] for a fan blade made of titanium. Here, the conclusion on material damping was that its contribution to the overall damping is negligible and should not be considered during design for resonant vibration. The same reference quantifies the influence of blade stress and temperature on material damping; however, for the given application these effects were neglected. In conclusion, in the current project employing piezoelectric excitation enabled the measurement of material damping, which is dependent on the modal shape. However, the linear fit indicates that on bench testing aerodynamic damping outweighs material damping as the pressure increases.

6.3 Mode 1 Damping Measurement. Strain measurements using aerodynamic excitation were performed for Mode 1. Two cases will be shown. In the first case, Mode1/EO5 resonance is generated due to the fundamental excitation frequency of the five lobe screen. In the second case, Mode1/EO6 resonance is generated due to the second harmonic of the three lobe screen.

Mode1/EO5. Figure 8 exemplifies the response of the main blade as resonance occurs for a 0.2 bar inlet pressure. The blade natural frequency and the excitation order can be identified. The response amplitude as a function of pressure was discussed in Part I of this paper. The dependency of damping on pressure is visualized in Fig. 9. The width of each slope is a measure of the level of damping: thus, as the inlet pressure increases from 0.1 bar to 0.4 bar, the width and therefore the damping increase. In order to estimate the critical damping ratio, the curve-fit method was applied to the response spectrum for the data taken at two distinct operating lines (see Fig. 1(*a*)). The critical damping ratio obtained was averaged across all available samples and was then normalized by the material damping estimated for Mode 1 with piezoelectric excitation. Figure 10 shows the linear fit according to Eq.

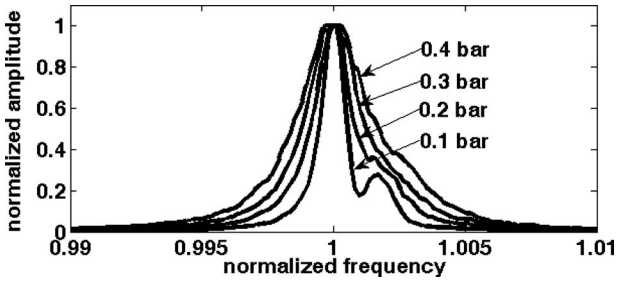


Fig. 9 Mode 1 response variation depending on inlet pressure

(9). The error bars indicate the sample variance. In the case of OL1, the variance is hardly visible in contrast to OL2.

Two main observations can be made. First of all, with a linear increase in damping, aerodynamic damping is the main contributor to the overall damping. For instance, at an inlet pressure of about 1 bar aerodynamic damping is about ten times higher than mechanical damping. Second, for vacuum conditions in the case of OL1, material damping matches the value obtained during piezoelectric excitation, whereas in the case of OL2, a slight mismatch was obtained. These observations show that employing piezoelectric excitation is a valid method to estimate material damping. Prior to the experiment, concerns were raised that the piezo would not generate amplitudes sufficiently high to allow damping comparison between the experiments since material damping depends on vibratory amplitude.

Damping comparison between the two operating lines shows that the gradient is different. In the case of the OL2, the gradient may be influenced to some degree by the curve-fit. As can be seen, the data reveal slight deviation from a linear curve-fit and variance is considerably higher than for OL1. In the authors' opinion, this variance may be introduced as a result of local modulations in aerodynamics as the mass flow varies and therefore affects the mean relative flow angle. Following the computations performed by Zemp et al. [21], the OL1 mean relative flow angle equals the blade metal angle, whereas in the case of OL2, a reduction of $\approx$7 deg was observed. Superpositioning the fluctuation introduced by the inlet distortion increases the angle deviation from the blade metal angle. Figure 11 compares the amplitude of the fundamental excitation frequency obtained from CFD calculations. Briefly, the inlet flow distribution was measured upstream of the impeller and was applied as an inlet boundary condition in an unsteady CFD calculation. The pressure fluctuations obtained at midspan were Fourier transformed and the fundamental excitation frequency was extracted from the signal. The figure shows differences in the forcing function occurring in the leading edge region of the blade. Both the amplitude and the phase are affected. CFD results

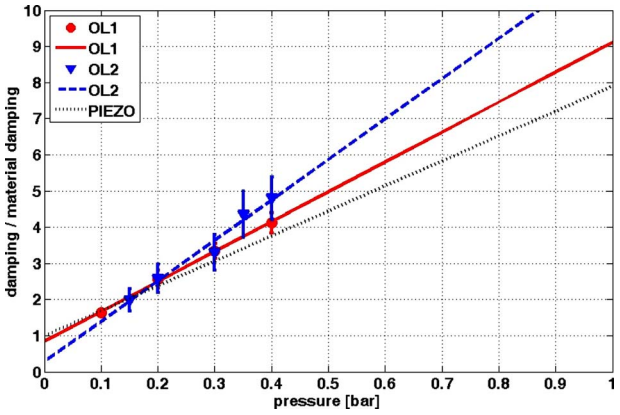


Fig. 10 Damping estimates for Mode 1 and EO5 excitation

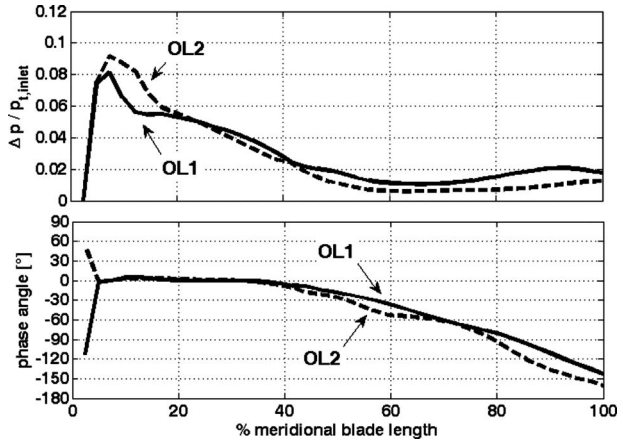


Fig. 11 First harmonic amplitude and phase for five lobe excitation obtained from CFD

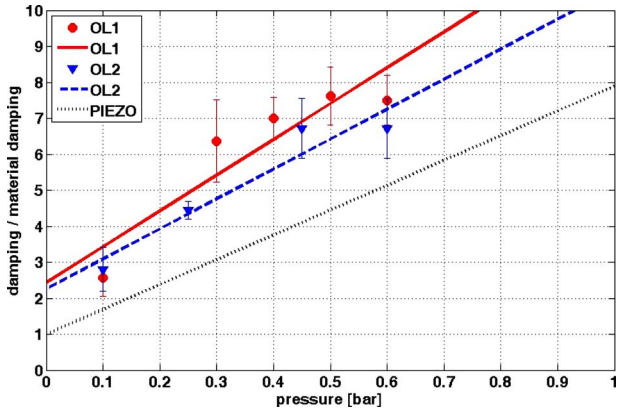


Fig. 13 Damping estimates for Mode 1 and EO6 excitation

showed that in the case of OL2, with the reduction in relative flow angle, the pressure side of the blade was affected, which leads to an increase in pressure fluctuations.

Mode1/EO6. Figure 12 exemplifies a typical response spectrum of the first mode due to EO6 excitation at 0.25 bar. The effects due to mistuning and coupling are visible. The damping ratios derived for both operating lines as a function of pressure are shown in Fig. 13. The first observation that can be made is that data points deviate more from the linear fit, and variance is higher than was the case for the data previously shown with EO5 excitation. Two factors may affect the results. First, the curve-fit procedure was found to be sensitive to the frequency bandwidth of the spectrum that was selected to perform the fit. Second, for this resonance point the impeller operates at part speed with the mean flow angle being ≈5 deg below the design incidence. This effect is in accordance with observations obtained from the previous case. Ideally, both linear fits should match material damping; however, due to data variance this is not the case. The increase in aerodynamic damping with pressure is comparable to the data obtained for Mode1/EO5, i.e., aerodynamic damping dominates as the main contributor to overall damping. Also variation between the two operating lines can be observed despite the given variance in data.

In summary, overall Mode 1 damping is dominated by aerodynamic damping. This refers to applications with an inlet pressure of approximately 1 bar. The increase in damping with the increase in pressure is comparable between the two excitation orders presented here. The magnitude of mechanical damping is in agreement between the experiments carried out with piezoelectric and aerodynamic excitations. Damping was found to depend on the operating line. In this regard, CFD simulation showed that the damping amplitude was affected by the flow field in the leading edge region. Furthermore, damping was found to be generally higher for the second harmonic excitation from the three lobe screen in comparison with fundamental excitation from the five lobe screen.

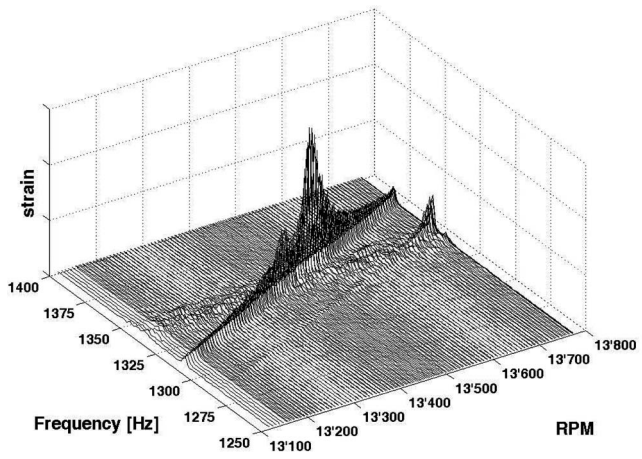


Fig. 12 Mode 1 response from EO6 excitation using a three lobe screen

6.4 Mode 2 Damping Measurement. Strain measurements using aerodynamic excitation were performed for Mode 2. Two cases will be compared. In one case, Mode2/EO12 resonance is generated due to the second harmonic excitation frequency of the 12 lobe screen. In the second case, Mode2/EO10 resonance is generated due to the second harmonic of the five lobe screen. Both cases will be presented in parallel; all results are shown in Fig. 14.

Figures 14(*a*) and 14(*b*) exemplify the response of the second mode. For both cases, mistuning and coupling are apparent due to the appearance of close frequencies at resonance. As in the cases presented earlier, damping was estimated through a curve-fit applied to the peak with the highest amplitude. Results for both cases are summarized in Fig. 14(*c*) for Mode2/EO12 and Fig. 14(*d*) for Mode2/EO10. Damping was normalized by the material damping obtained using piezoelectric excitation. The first observation to make in this regard is the substantially lower damping level due to piezoelectric excitation in comparison to damping obtained from aerodynamic excitation. Moreover, in the case of Mode2/EO12, material damping estimates between the experiments were found to agree reasonably well. The damping variance is comparatively low. In the case of Mode2/EO10, aerodynamic experiments yielded negative material damping. As shown previously, given the variance in damping estimates and the limited number of pressure settings at which data were acquired, the potential to affect the linear curve-fit was apparent. Overall, Mode 2 damping can also be seen to be dominated by aerodynamic damping.

The dependency of damping on the operating line can be identified at least for Mode2/EO12 given the low variance. Based on the same approach, in Mode2/EO10, a statement on operating line dependency cannot be made. Overall, at this stage it is comparatively difficult to reason on the cause of damping alternation. The flow around the blade is dominated by the first harmonic of the distortion screen whereas excitation is provided by the second or third harmonic, which exhibits lower amplitudes. Computation of this problem would require the accurate capture of EO12 and EO10 excitations.

In summary, Mode 2 damping is dominated by aerodynamic damping. The increase in damping was found to differ and to be higher in the Mode2/EO10 case. As far as data variance enables us to make a judgment, material damping was predicted equally

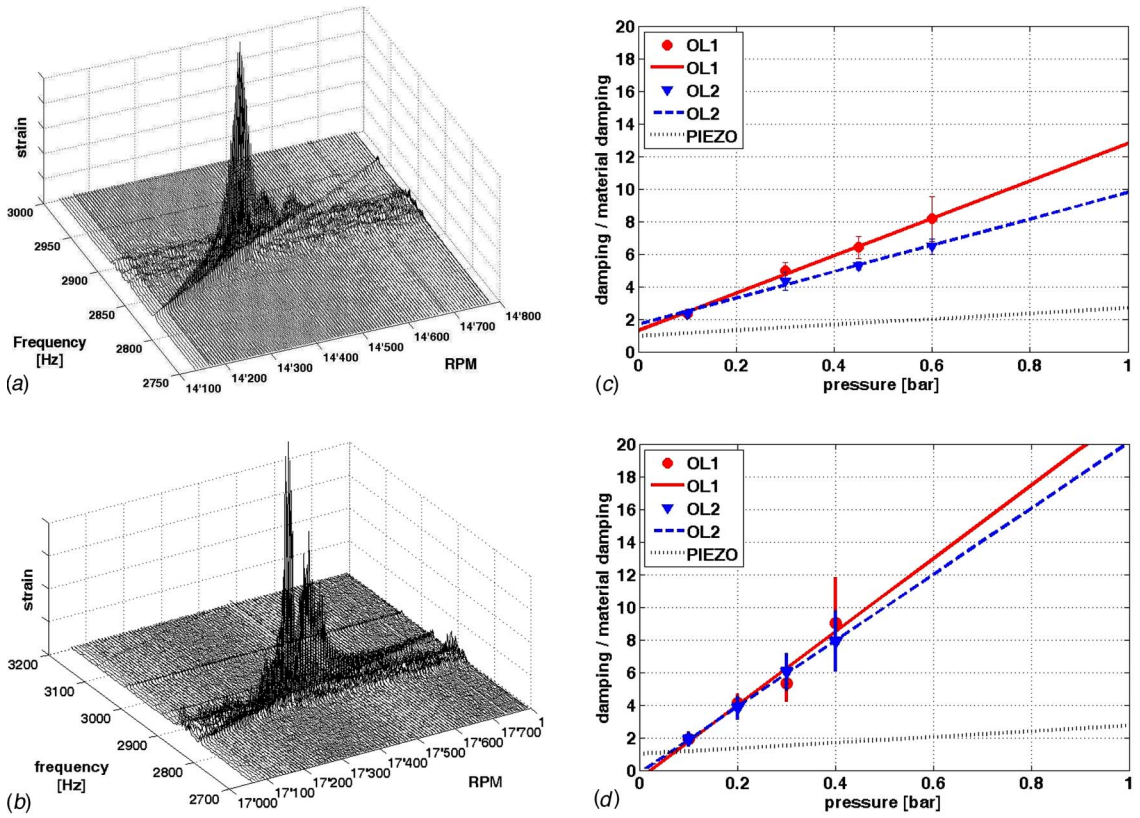


Fig. 14 Strain response and damping estimates for Mode 2: (*a*) Mode 2 response from EO12 excitation using a four lobe screen, (*b*) Mode 2 response from EO10 excitation using a five lobe screen, (*c*) damping estimates for Mode 2 and EO12 excitation, and (*d*) damping estimates for Mode 2 and EO10 excitation

when comparing piezoelectric and aerodynamic excitations. The dependency of damping on the operating line could be identified; however, an explanation cannot be given at this stage.

7 Summary and Conclusions

Damping estimation during forced response for a radial compressor was carried out on the basis of experimental data. The impeller design is typical for turbocharging applications; it features thin blades and is made of aluminum. Damping estimation was performed for the first two eigenmodes of the main blade. The experimental procedure used allowed the determination of the contributions of material and aerodynamic damping. Circle-fit and curve-fit methods were used based on the SDOF model.

In the first step, material damping was estimated using experimental data taken during piezoelectric excitation. The impeller was shaft mounted and not rotating. In this instance, piezos were mounted on the impeller disk and provided synchronous excitation while strain gauges measured the response on the blades surface. Measurements were performed for a number of pressure settings ranging from vacuum and increasing to ambient conditions. The transfer function was computed, and both the circle-fit and the curve-fit methods were applied. A comparison between the two methods showed that the estimated damping ratios deviated by approximately 10% from each other compared with the absolute value. Material damping was then calculated from a linear data fit and then extrapolation to vacuum. Material damping was found to depend on the mode shape; however, as the pressure increased the contribution of material, damping became fractional.

In the second step, the critical damping ratio was estimated based on the data taken in the rotating facility. Part I of this two-part paper describes the experimental procedure, in particular, the transient measurement approach and generation of inlet flow distortion. For both eigenmodes, Modes 1 and 2, data were acquired for two resonant conditions. A curve-fit procedure was applied in order to determine the critical damping ratio as a function of pressure, which could be curve-fitted by a linear function. The results showed that the estimation of material damping could be derived equally by the aerodynamic excitation experiment and also by piezoelectric excitation. Aerodynamic damping was found to dominate the overall damping, i.e., at a 1 bar inlet pressure aerodynamic damping is higher by a factor of 10 than material damping. The dependency of damping on the operating line was observed. In the case of Mode1/EO5, CFD results indicated that as the mass flow was varied, the unsteady pressure distribution at the leading edge changed and therefore affected the damping characteristics. In the case of Mode1/EO6, increased variance in the data was observed. Generally, the results indicated that data variance in damping measurements is affected by the mean incidence angle. In the case of Mode 2, excitation was provided by the second and the third harmonic of the inlet flow distortion. Dependency on the operating line was identified: however, this required a sufficiently low variance in data.

Acknowledgment

The authors gratefully acknowledge the intellectual and financial support of their industrial partners ABB Turbo Systems and MAN TURBO AG Schweiz in this project. They also acknowledge the financial support provided by the Swiss Confederation's innovation promotion agency, CTI.

Nomenclature

H = frequency transfer function
F = forcing function
F_D = overall damping force
F_{DM}, F_{DA} = material and aerodynamic damping force

OL1, OL2 = Operating Lines 1 and 2
S_x = spectral density of excitation function
S_y = spectral density of response function
S_{xy} = cross-correlation spectrum
X = frequency function of excitation
Y = frequency function of system response
c = viscous damping coefficient
k = spring constant
m = mass
p = pressure
Δp = pressure difference
$x, \dot{x}, \ddot{x}$ = displacement, speed and acceleration
Ω = angle
ω = frequency in radians
ω_n = natural frequency in radians
ζ = overall critical damping ratio
ζ_M, ζ_A = material and aerodynamic critical damping ratio
$_a$ = data point before resonance
$_b$ = data point after resonance
$_{\text{ref}}$ = reference quantity

References

[1] Srinivasan, A. V., 1984,"Vibrations of Bladed-Disk Assemblies—A Selected Survey," Trans. ASME, J. Vib., Acoust., Stress, Reliab. Des., **106**, pp. 165–168.
[2] Srinivasan, A. V., 1997, "Flutter and Resonant Vibration Characteristics of Engine Blades," ASME J. Turbomach., **119**, pp. 741–775.
[3] Kielb, R. E., and Imregun, M., 1999, *VKI Lecture Series 1999-05*, von Karman Institute for Fluid Dynamics, Belgium.
[4] Srinivasan, A. V., Cutts, D. G., and Sridhar, S., 1981, NASA, Technical Report No. NASA-CR-165406.
[5] Crawley, E. F., 1983, "Aerodynamic Damping Measurements in a Transonic Compressor," ASME J. Eng. Power, **105**, pp. 575–584.
[6] Jin, D., Hasemann, H., Haupt, U., and Rautenberg, M., 1991, "Schwingungen in rotierenden Maschinen: Referate der Tagung an der Universität/ Gesamthochschule Kassel," *Untersuchung der Schaufeldämpfung hochbelasteter Radialverdichterlaufräder*, Vieweg, Braunschweig, Kassel.
[7] Kielb, J. J., and Abhari, R. S., 2003, "Experimental Study of Aerodynamic and Structural Damping in a Full-Scale Rotating Turbine," ASME J. Eng. Gas Turbines Power, **125**, pp. 102–112.
[8] Jeffers, T. R., Kielb, J. J., and Abhari, R. S., 2000, ASME Paper No. 2000-GT-0359.
[9] Newman, F. A., 1988, 24th Joint Propulsion Conference AIAA, ASME, SAE, ASEE, Boston, MA, 1988, Paper No. AIAA-88-3229.
[10] Chiang, H. W. D., and Kielb, R. E., 1993, "An Analysis System for Blade Forced Response," ASME J. Turbomach., **115**, pp. 762–770.
[11] Li, H. D., and He, L., 2002, "Single-Passage Analysis of Unsteady Flows Around Vibrating Blades of a Transonic Fan Under Inlet Distortion," ASME J. Turbomach., **124**, pp. 285–292.
[12] Moffat, S., and He, L., 2003, ASME Paper No. GT2003-38640.
[13] Ning, W., Moffat, S., Li, Y., and Wells, R. G., 2003, ASME Paper No. GT2003-38642.
[14] Dickmann, H. P., Wimmel, T. S., Szwedowicz, J., Filsinger, D., and Roduner, C. H., 2006, "Unsteady Flow in a Turbocharger Centrifugal Compressor: Three-Dimensional Computational Fluid Dynamics Simulation and Numerical and Experimental Analysis of Impeller Blade Vibration," ASME J. Turbomach., **128**, pp. 455–465.
[15] Schmitt, S., 2003, "Simulation von Flattern und aerodynamischer Zwangserregung in Turbomaschinenbeschaufelungen," Ph.D. thesis, Institut für Antriebstechnik, Köln.
[16] Kammerer, A., and Abhari, R. S., 2008, ASME Paper No. GT2008-50466.
[17] Qian, S., and Chen, D., 1996, *Joint Time-Frequency Analysis Methods and Applications*, Prentice-Hall, Upper Saddle River, NJ.
[18] Ewins, D. J., 2000, *Modal Testing Theory, Practice and Application*, 2nd ed., Research Studies, Baldock, Hertfordshire, England.
[19] Hasemann, H., Hagelstein, D., and Rautenberg, M., 2000, "Coupled Vibration of Unshrouded Centrifugal Compressor Impellers. Part 1: Experimental Investigation," Int. J. Rotating Mach., **6**, pp. 101–113.
[20] Hagelstein, D., Hasemann, H., and Rautenberg, M., 2000, "Coupled Vibration of Unshrouded Centrifugal Compressor Impellers. Part 2: Computation of Vibration Behavior," Int. J. Rotating Mach., **6**, pp. 115–128.
[21] Zemp, A., Kammerer, A., and Abhari, R. S., 2008, ASME Paper No. GT2008-50744.

Method for Sensitivity Analysis of Resonance Forced Response of Bladed Disks With Nonlinear Contact Interfaces

E. P. Petrov
Mechanical Engineering Department,
Centre of Vibration Engineering,
Imperial College London,
South Kensington Campus,
London SW7 2AZ, UK
e-mail: y.petrov@imperial.ac.uk

An effective method has been developed to calculate the sensitivity of the resonance peak frequency and forced response level to variation of parameters of nonlinear friction contact interfaces and excitation. The method allows determination of the sensitivity characteristics simultaneously with the resonance peak frequency and response level calculated as a function of any parameter of interest and without significant computational expense. Capabilities of the method are demonstrated on examples of analysis of large-scale finite element models of realistic bladed disks with major types of the nonlinear contact interfaces: (i) a blisk with underplatform dampers, (ii) a bladed disk with friction damping at blade fir-tree roots, and (iii) a high-pressure bladed disk with shroud contacts. The numerical investigations show high efficiency of the method proposed.
[DOI: 10.1115/1.2969094]

1 Introduction

The resonance forced response levels and values of the resonance frequencies are of primary interests in the majority of investigations of forced response for bladed disks, which are aimed at choice of the design concept and its parameters in practical gas-turbine engines design process.

Bladed disk assemblies have customarily many contact interfaces, such as at blade-disk joints, at contact surfaces between interlocking blade shrouds, and in friction dampers. Forces occurring at these interfaces are strongly nonlinear because of friction forces, unilateral interaction forces acting along normal to a contact interface, closing and opening clearances during vibrations, etc.

In most industrial problems, the highest forced response levels that are achieved within frequency ranges of gas-turbine engine operation are of major interests because of their defining effect on high cycle fatigue failures.

The resonance frequencies and resonance forced response levels are dependent, for nonlinear structures, on excitation and on parameters of the friction and other nonlinear contact interfaces. Appropriate choices of the parameters for nonlinear contact interfaces can avoid dangerous resonance regimes by moving the resonance frequencies out of the operating frequency range and/or by decreasing the resonance amplitudes to acceptable levels.

Analysis of the forced response of bladed disks with friction contact interfaces is usually performed in the frequency domain owing to the high computational efficiency of this approach. Some examples of the frequency-domain analysis of bladed disk forced response can be found in Refs. [1–15].

In order to facilitate a choice of a set of contact interface parameters, there is a need to determine not only the resonance peak frequency and response level but also to obtain estimates on how sensitive are those to variation of design parameters and forcing levels.

The sensitivity characteristics can provide estimates for robustness, the calculated forced response levels in the presence of scatter or variability in the design parameters, which can differ from their nominal values within some ranges determined by manufacturing tolerances (see Ref. [16]). Moreover, the sensitivity characteristics allow determination of statistical characteristics of the forced response for a structure with scatters of design parameters and assessment of effects of uncertainties in the design parameter values of uncertainties of the forced response (see Ref. [17]). Furthermore, in many practical applications, there is often a need to find an optimum set of the design parameter values for a structure that gives a minimum of the forced response or satisfies other criteria formulated by the designer. The optimization problems can be efficiently solved when sensitivity characteristics are determined and, hence, sensitivity analysis is important and in many cases an unavoidable part of the optimization process (see Refs. [18–20]).

In this paper, an effective method is proposed for direct analysis of sensitivity of resonance regimes of essentially nonlinear vibrations of bladed disks to variation of contact interface parameters and levels of excitation. The method provide a capability, believed to be not known in literature before, to calculate the sensitivity of resonance frequencies and response levels while performing the direct parametric analysis of the resonance peaks, which was developed in Ref. [21]. Sensitivities of the resonance forced response to friction coefficient values, clearances, interferences, and parameters of prestressed state and loading are calculated.

The analytical derivation of special contact elements and the analytical derivation of the equation for a whole structure enable exact and extremely fast calculation of these sensitivities. All sensitivities are evaluated exactly avoiding completely the use of the finite-difference approximations. Multiharmonic balance formulation is derived to perform the resonance forced response sensitivity analysis in the frequency domain using large-scale finite element (FE) models of bladed disks and detailed modeling of the friction contact interfaces.

Numerical investigations of bladed disks with friction contact interfaces are performed to demonstrate outstanding efficiency of the approach developed.

2 Formulation of the Problem

The maximum forced response levels are achieved, as a rule, at the resonances. The resonance response levels and resonance frequencies are dependent on design parameters of a gas-turbine

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 10, 2008; final manuscript received April 14, 2008; published online December 30, 2008. Review conducted by Dilip R. Ballal. Paper presented at the ASME Turbo Expo 2008: Land, Sea and Air (GT2008), Berlin, Germany, June 9–13, 2008.

structure and on excitation levels. Therefore, development of the capability of predicting effects of design parameters and the excitation level on the resonance forced response level and on the resonance frequency can facilitate significantly the choice of the design parameters, which would allow a reduction of the resonance responses for all expected excitation levels or an avoidance of dangerous resonances by moving them out of the operating frequency ranges.

It should be noted that such capability has special importance for analysis of nonlinear structures where, in contrast to linear systems, resonance frequencies are dependent on the excitation levels and on the response levels accordingly, and an increase or decrease of the excitation level can bring the resonance peak into the operating frequency ranges.

A method developed in Ref. [21] allows, for a case of steady-state vibrations, direct calculation of the resonance frequencies ω^{res} and resonance peak response levels of a structure $\boldsymbol{a}^{\text{res}}$ as a function of design parameters and function of levels of excitation forces, i.e.,

$$\begin{aligned} \omega^{\text{res}}(\lambda) &= \omega^{\text{res}}(\boldsymbol{b}(\lambda)) \\ \boldsymbol{a}^{\text{res}}(\lambda) &= \boldsymbol{a}^{\text{res}}(\boldsymbol{b}(\lambda)) \end{aligned} \tag{1}$$

where the vector of the resonance responses $\boldsymbol{a}^{\text{res}}$ can include the resonance peak response levels selected at any set of nodes of the structural FE model, vector $b(\lambda)=\{b_1,b_2,\ldots,b_n\}^T$ comprises the parameters of interest, including the contact interface parameters such as friction coefficients, clearances, and stiffness coefficients due to contact surface roughness and also the parameters characterizing excitation forces (e.g., amplitude of a nodal force or amplitude of a modal forces), and λ is the tracing variable, which is introduced to describe variation of design parameters and excitation when operational conditions are changed (i.e., rotational speed, temperature, and excitation forces). Such variation can be simultaneous for all contact interface parameters and excitation forces (for example, when this variation is owing to variation of the rotation speed or temperature) or can be applied to a single interface parameter chosen by a designer. Moreover, the tracing variable can be used to describe the variation of excitation forces.

In this paper, a method for analysis sensitivities of the resonance frequencies and resonance peak amplitudes is developed. The method is aimed at calculation of sensitivity coefficients of these quantities with respect to a selected set of design parameters:

$$\begin{matrix} \partial\omega^{\text{res}}(\lambda)/\partial b_s, \\ \partial\boldsymbol{a}^{\text{res}}(\lambda)/\partial b_s, \end{matrix} \quad s=1,\ldots,n \tag{2}$$

where s is the parameter number, and the set of parameters with respect to which the sensitivities are calculated can include n contact interface parameters.

This calculation is performed simultaneously with the direct parametric analysis of the resonance amplitudes and frequencies illustrated by Eq. (1). The tracing variable varies within a selected range, $[\lambda^-,\lambda^+]$, for which the direct parametric analysis is performed. The key point is that the method is based on analytically derived expressions for all vectors and matrices involved in the sensitivity calculations, which allows very fast and accurate calculation without noticeable increase of the computational time.

3 Method for Sensitivity Analysis of Resonance Peak Amplitudes and Frequencies

The equation for motion for a bladed disk and, generally, for any other structure with nonlinear interfaces can be written in the following form:

$$\boldsymbol{K}\boldsymbol{q}(t)+\boldsymbol{C}\dot{\boldsymbol{q}}(t)+\boldsymbol{M}\ddot{\boldsymbol{q}}(t)+\boldsymbol{f}(\boldsymbol{q}(t),\boldsymbol{b}(\lambda))-\boldsymbol{p}(t,\boldsymbol{b}(\lambda))=\boldsymbol{0} \tag{3}$$

where $\boldsymbol{q}(t)$ is a vector of displacements; $\boldsymbol{K}$, $\boldsymbol{C}$, and $\boldsymbol{M}$ are the stiffness, viscous damping, and mass matrices used for description of linear forces. For a bladed disk rotating with speed ω, the stiffness matrix $\boldsymbol{K}$ can also include terms accounting for the rotation effects, such as geometric stiffness matrix reflecting stiffening effects of the centrifugal forces and a spin-softening-matrix describing stiffness softening due to the changing direction of the centrifugal forces under vibration. $\boldsymbol{f}(\boldsymbol{q}(t),\boldsymbol{b}(\lambda))$ is a vector of nonlinear interface forces. It is dependent on the displacements of contacting surfaces and on the parameters of the contact interfaces, $\boldsymbol{b}(\lambda)$; $\boldsymbol{p}(t,\boldsymbol{b}(\lambda))$ is a vector of periodic external excitation forces.

3.1 Frequency Domain Equation of Motion. For a search of the steady-state periodic vibration response, the time variation of displacements is represented by a restricted Fourier series:

$$\boldsymbol{q}(t)=\boldsymbol{Q}_0+\sum_{j=1}^{n}\left(\boldsymbol{Q}_j^{(c)}\cos m_j\omega t+\boldsymbol{Q}_j^{(s)}\sin m_j\omega t\right) \tag{4}$$

where $\boldsymbol{Q}_j^{(c)}$ and $\boldsymbol{Q}_j^{(s)}$ $(j=1,\ldots,n)$ are the vectors of cosine and sine harmonic coefficients for system DOFs, marked by superscripts (c) and (s) accordingly; $\boldsymbol{Q}_0$ is a vector of constant components of the displacements; m_j $(j=1,\ldots,n)$ are the specific numbers of harmonics that are kept in the displacement expansion. The multiharmonic representation of the forced response allows determination of major types of periodic vibration, which are possible for strongly nonlinear structures including (i) major resonances, (ii) superharmonic resonances, (iii) subharmonic resonances, and (iv) combination resonances. This determination can be done by an appropriate choice of values for m_j in the multiharmonic expansion of Eq. (4). For a case of search for major and superharmonic resonances, m_j can be integer numbers corresponding to those harmonics that can contribute significantly to the forced response. For a search of subharmonic and combination resonances, m_j can be appropriately selected as fractional numbers.

Substitution of Eq. (4) in Eq. (3) and application of the multiharmonic balance and condensation methods developed in Refs. [22,23] give an equation of motion in the frequency domain with respect to the harmonic coefficients of the multiharmonic expansion in the form

$$\boldsymbol{R}(\boldsymbol{Q},\omega,\lambda)=\boldsymbol{Q}+\boldsymbol{A}(\omega)(\boldsymbol{F}(\boldsymbol{Q},\boldsymbol{b}(\lambda))-\boldsymbol{P}(\boldsymbol{b}(\lambda)))=\boldsymbol{0} \tag{5}$$

where $\boldsymbol{Q}=\{\boldsymbol{Q}_0,\boldsymbol{Q}_1^{(c)},\boldsymbol{Q}_1^{(s)},\ldots,\boldsymbol{Q}_n^{(s)}\}^T$ is a vector of harmonic coefficients of displacements, $\boldsymbol{P}=\{\boldsymbol{P}_0,\boldsymbol{P}_1^{(c)},\boldsymbol{P}_1^{(s)},\ldots,\boldsymbol{P}_n^{(s)}\}^T$ is a vector of harmonic components of the excitation forces, $\boldsymbol{F}(\boldsymbol{Q},\boldsymbol{b}(\lambda))=\{\boldsymbol{F}_0,\boldsymbol{F}_1^{(c)},\boldsymbol{F}_1^{(s)},\ldots\boldsymbol{F}_n^{(s)}\}^T$ is a vector of harmonic components of nonlinear forces, and $\boldsymbol{A}(\omega)$ is a multiharmonic frequency response function (FRF) matrix of the linear part of the system, i.e.,

$$\boldsymbol{A}(\omega)=\text{diag}[\boldsymbol{A}_0,\boldsymbol{A}_1(\omega),\ldots,\boldsymbol{A}_n(\omega)] \tag{6}$$

where the FRF matrices $\boldsymbol{A}_j(\omega)$ can be generated from mode shapes and natural frequencies obtained for a linear structure when contact interactions are neglected. For each harmonic number, this matrix can be obtained in the following form:

$$\boldsymbol{A}_j(\omega)=\begin{bmatrix} \text{Re}(\tilde{\boldsymbol{A}}_j(\omega)) & \text{Im}(\tilde{\boldsymbol{A}}_j(\omega)) \\ -\text{Im}(\tilde{\boldsymbol{A}}_j(\omega)) & \text{Re}(\tilde{\boldsymbol{A}}_j(\omega)) \end{bmatrix} \tag{7}$$

$$\begin{aligned} \tilde{\boldsymbol{A}}_j &= [(1+i\eta)\boldsymbol{K}(\omega)+(m_j\omega)^2\boldsymbol{M}]^{-1} \\ &\simeq \sum_{r=1}^{N_m}\frac{\boldsymbol{\phi}_r(\omega)\boldsymbol{\phi}_r^T(\omega)}{(1-i\eta_r(\omega))(\omega_r(\omega))^2-(m_j\omega)^2} \end{aligned} \tag{8}$$

where $\omega_r(\omega)$, $\boldsymbol{\phi}_r(\omega)$, and $\eta_r(\omega)$ are the natural frequencies, the mode shapes, and the modal damping factors for the linear part of a structure; N_m is the number of modes used in Eq. (8) for the FRF matrix generation.

3.2 Equation for Determination of the Resonance Peaks. To determine resonance frequency ω^{res} together with the harmonic

coefficients of the multiharmonic representation of the resonance forced response $\boldsymbol{Q}^{\text{res}}$, Eq. (5) is complemented, as was proposed in Ref. [21], by the following condition:

$$r = \frac{1}{2}\frac{\partial}{\partial \omega}((q_{jk}^{c})^2 + (q_{jk}^{s})^2) = \boldsymbol{Q}^T \boldsymbol{I}_{jk}\frac{\partial \boldsymbol{Q}}{\partial \omega} = 0 \tag{9}$$

where q_{jk}^{c} and q_{jk}^{s} are the coefficients for cosine and sine accordingly for jth harmonic of kth DOF selected from the vector of coefficients $\boldsymbol{Q}$, and $\boldsymbol{I}_{jk}$ is a diagonal matrix, which has two units corresponding to the sine and cosine coefficients for jth harmonic of kth DOF at the main diagonal. One can see that Eq. (9) expresses the condition that the amplitude of displacement for a selected DOF, k, and harmonics, j, takes its extreme value at the resonance frequency.

The resonance peak amplitudes $\boldsymbol{Q}^{\text{res}}$ and the resonance frequency ω^{res} are then determined from the following combined equation:

$$\boldsymbol{R}^{\text{res}}(\boldsymbol{Q}, \omega, \lambda) = \begin{Bmatrix} \boldsymbol{R}(\boldsymbol{Q}, \omega, \boldsymbol{b}(\lambda)) \\ r(\boldsymbol{Q}, \omega, \boldsymbol{b}(\lambda)) \end{Bmatrix} = \boldsymbol{0} \tag{10}$$

The resonance amplitudes and resonance frequencies are calculated, as described in Ref. [21].

3.3 Calculation of the Resonance Peak Value Sensitivities. When the resonance amplitudes and the resonance frequency are found, the sensitivity of these quantities to design and the excitation parameters $\boldsymbol{b}(\lambda)$ can be calculated. The sensitivities are determined from the following equation, which is obtained by differentiating Eq. (10):

$$\begin{bmatrix} \frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Q}} & \frac{\partial \boldsymbol{R}}{\partial \omega} \\ \frac{\partial r}{\partial \boldsymbol{Q}} & \frac{\partial r}{\partial \omega} \end{bmatrix} \begin{Bmatrix} \frac{\partial \boldsymbol{Q}}{\partial \lambda} \\ \frac{\partial \omega}{\partial \lambda} \end{Bmatrix} = - \begin{Bmatrix} \frac{\partial \boldsymbol{R}}{\partial \lambda} \\ \frac{\partial r}{\partial \lambda} \end{Bmatrix} \tag{11}$$

where all derivatives are derived analytically, as shown below:

$$\frac{\partial \boldsymbol{R}}{\partial \boldsymbol{Q}} = \boldsymbol{I} + \boldsymbol{A}\frac{\partial \boldsymbol{F}}{\partial \boldsymbol{Q}}, \quad \frac{\partial \boldsymbol{R}}{\partial \omega} = \frac{\partial \boldsymbol{A}}{\partial \omega}(\boldsymbol{F} - \boldsymbol{P}) + \boldsymbol{A}\left(\frac{\partial \boldsymbol{F}}{\partial \omega} - \frac{\partial \boldsymbol{P}}{\partial \omega}\right) \tag{12}$$

$$\frac{\partial r}{\partial \boldsymbol{Q}} = \left(\boldsymbol{I}_{jk}\frac{\partial \boldsymbol{Q}^{\text{res}}}{\partial \omega}\right)^T + (\boldsymbol{Q}^{\text{res}})^T \boldsymbol{I}_{jk}\frac{\partial^2 \boldsymbol{Q}}{\partial \boldsymbol{Q}\, \partial \omega} \tag{13}$$

$$\frac{\partial r}{\partial \omega} = (\boldsymbol{Q}^{\text{res}})^T \boldsymbol{I}_{jk}\frac{\partial^2 \boldsymbol{Q}^{\text{res}}}{\partial \omega^2} \tag{14}$$

$$\frac{\partial \boldsymbol{R}}{\partial \lambda} = \boldsymbol{A}\left(\frac{\partial \boldsymbol{F}}{\partial \lambda} - \frac{\partial \boldsymbol{P}}{\partial \lambda}\right), \quad \frac{\partial r}{\partial \lambda} = -(\boldsymbol{Q}^{\text{res}})^T \boldsymbol{I}_{jk}\frac{\partial^2 \boldsymbol{Q}^{\text{res}}}{\partial \lambda\, \partial \omega} \tag{15}$$

The derivatives of matrices and vectors involved in Eqs. (12)–(15) are also derived analytically, including $\boldsymbol{F}$ and $\partial \boldsymbol{F}/\partial \boldsymbol{Q}$ (see Ref. [22]), $\partial \boldsymbol{F}/\partial \lambda$ (see Ref. [24]), and $\partial \boldsymbol{Q}^{\text{res}}/\partial \omega$, $\partial^2 \boldsymbol{Q}^{\text{res}}/\partial \boldsymbol{Q}\, \partial \omega$, $\partial^2 \boldsymbol{Q}^{\text{res}}/\partial \boldsymbol{Q}\, \partial \lambda$, $\partial^2 \boldsymbol{Q}^{\text{res}}/\partial^2 \omega$ (see Ref. [21]). All matrices and vectors involved in the resonance peak sensitivity, Eq. (11), are evaluated for the resonance peak amplitudes $\boldsymbol{Q}^{\text{res}}$ and the resonance frequency ω^{res} determined from the solution of Eq. (10).

There are following cases that usually need to be considered in the sensitivity analysis of structures.

(i) A case when the calculation of sensitivity characteristics with respect to each of the contact interface parameter b_s, $(s=1,\ldots,n)$, individually, is required. For this case, individual contact parameter values b_s are used instead of λ, in Eqs. (11)–(15), and, therefore, Eq. (11) is solved for n different vectors on right-hand side to find $\partial \boldsymbol{Q}^{\text{res}}/\partial b_s$ and $\partial \omega^{\text{res}}/\partial b_s$ for each design parameter.

(ii) A case when all parameters of interest in vector $\boldsymbol{b}(\lambda)$ are interdependent and λ is the interdependency parameter used for description of these dependencies. In practical applications, the interdependency parameter can be, for example, a rotation speed allowing for dependencies of clearances, interferences, and friction coefficients on centrifugal forces, temperatures, and gas flow variation. For this case, $\partial b_s/\partial \lambda$ are prescribed and the solution for the preceding case can be used to obtain the resonance peak response sensitivities with respect to parameter λ:

$$\frac{\partial \boldsymbol{Q}^{\text{res}}}{\partial \lambda} = \sum_{s=1}^{n} \frac{\partial \boldsymbol{Q}^{\text{res}}}{\partial b_s}\frac{\partial b_s}{\partial \lambda}, \quad \frac{\partial \omega^{\text{res}}}{\partial \lambda} = \sum_{s=1}^{n} \frac{\partial \omega^{\text{res}}}{\partial b_s}\frac{\partial b_s}{\partial \lambda} \tag{16}$$

Nevertheless, the more numerically efficient way to calculate the sensitivity with respect to the interdependency parameter is a direct evaluation of $\partial \boldsymbol{R}/\partial \lambda$ and $\partial r/\partial \lambda$, which requires the solution of Eq. (11) only once.

(iii) A case when all design parameters can be divided into groups and for each such group the dependency of parameters can be established, i.e., $\boldsymbol{b} = \{\boldsymbol{b}_1^g(\lambda_1), \boldsymbol{b}_2^g(\lambda_2), \ldots, \boldsymbol{b}_{n_g}^g(\lambda_{n_g})\}$. For this case, Eq. (11) can be applied to calculate sensitivity for each of these groups individually, assuming $\lambda = \lambda_g$, $(g=1,\ldots,n_g)$. The more numerically efficient way is, again, the direct evaluation of $\partial \boldsymbol{R}/\partial \lambda_g$ and $\partial r/\partial \lambda_g$ and the solution of Eq. (11) for each group of the parameters.

Three important advantages of the method should be highlighted.

(1) All expressions required for the resonance peak sensitivity analysis are derived analytically and, therefore, they are calculated with very high accuracy and extremely fast.

(2) The matrix on left-hand side of Eq. (11) is evaluated when the resonance peak frequency and amplitudes are calculated and, to do this, the Jacobian of Eq. (10) is calculated. Moreover, the factorization of this matrix is also made during the resonance response calculation; as a result of this prior factorization, the solution of Eq. (11), required for sensitivity calculation, does not increase noticeably the computational time, even for relatively large number of parameters with respect to which the sensitivity coefficients are determined and Eq. (11) is solved for several right-hand vectors corresponding to each design parameter of interest.

(3) The step for the tracing parameter increments λ within the analyzed range $[\lambda^-, \lambda^+]$ is a variable and defined automatically in the analysis. It is reduced when the rates of variation of the resonance frequency and response level are large, and they are increased when those are relatively small. This variation of the step size guarantees capturing all, possibly abrupt, variations of the resonance response and frequency while avoiding superfluous computations.

4 Numerical Examples of the Sensitivity Analyses

The methodology developed has been applied for analysis of different bladed disks with different types of the friction contact interfaces. Some of the investigated examples are discussed below.

4.1 A Blisk With Underplatform Dampers. A blisk of 24 blades, which is analyzed here, was manufactured for a test-rig built at Imperial College London in the framework of the EU project "Aeroelastic Design of Turbine Blades II" (ADTurbII) (see Ref. [25]). A finite element model of this blisk is shown in Fig. 1. The model of one sector of the blisk used in calculation here contains 21,555 DOFs.

A traveling wave excitation by the 19th engine-order is considered, which excites for modes with $24-19=5$ nodal diameters (NDs). Since there are no root or other contact interfaces in the blisk, apart of the underplatform dampers, the modal damping is

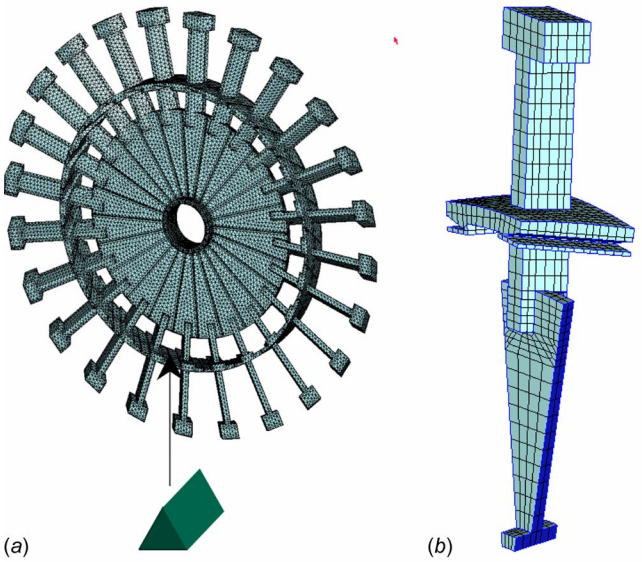


Fig. 1 A FE model of the test-rig blisk with underplatform dampers: (*a*) the whole blisk; (*b*) a sector FE model

solely due to the damping in material of the blisk and the damping loss factor is assumed low: $\eta=7.5\times10^{-5}$ and the friction damping owing to underplatform dampers (UPDs) is determined as a result of calculation.

The UPDs are modeled by rigid-body motion models developed in Ref. [26]. The parameters of these UPDs include a value of the friction coefficient at the contact interfaces of the UPDs, μ, a value of the normal loads applied at the friction interfaces N_0, which are due to the action of the centrifugal forces and proportional to the damper mass value, normal k_n and tangential k_t stiffness coefficients characterizing the elasticity of a layer of microasperities of a rough contact surface, and, for these UPD models, also the elasticity of the damper body.

The resonance frequency and resonance forced response level were calculated for this bladed disk directly as a function of parameters of the UPDs using the direct parametric analysis method. The forced response was determined at a node located at the blade tip. In Fig. 2, the results of the direct parametric analysis for the resonance peaks (shown by bold solid curve) are compared with the forced response dependencies on the rotation speed calculated for the selected values of the normal load.

The range of the normal load variation was examined, which

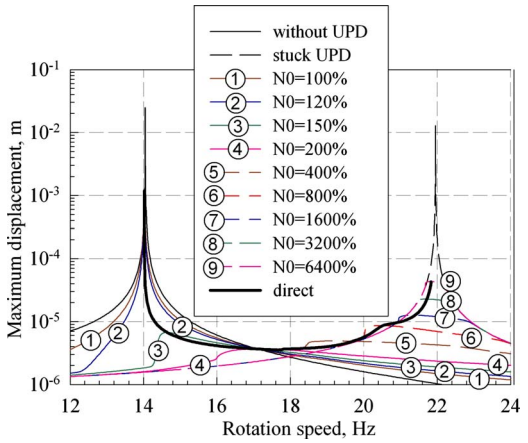


Fig. 2 Forced response calculated for different normal load values and results of the direct parametric analysis

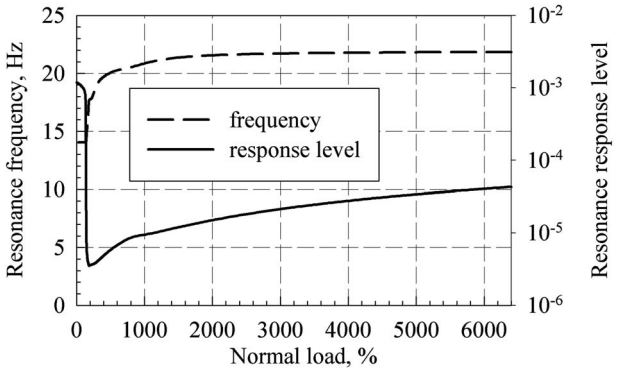


Fig. 3 Dependencies of the resonance frequency and response level on the normal load calculated

includes values from 100% of the normal load reference value to 6400%. In the direct parametric analysis, the normal load values was varied continuously and for the frequency forced response dependencies the forced response calculation were made for eight different values of the normal load selected in the range analyzed. For comparison, the forced responses of two limiting cases are displayed also: (i) the forced response of the blisk without UPDs and (ii) the forced response of the blisk with normal loads applied at the contact interfaces chosen, so large that the UPDs are always stuck. One can see that the curve obtained by direct parametric analysis follows exactly the resonance peaks calculated of all nine frequency forced response curves determined for different values of the normal load.

In Fig. 3, the rotor rotation resonance frequency and the resonance peak forced response level are shown, as they are calculated here by the direct parametric analysis, i.e., as a function of the normal load. The optimal value of the normal load can be observed directly from the plot of the resonance response level: one can see that the minimum value is achieved for value $N_0=200\%$.

The sensitivities of the resonance peak frequency and response are calculated simultaneously with the direct parametric analysis to all four UPD's parameters.

The friction contact parameters are measured in different units and their magnitudes differ significantly. Moreover, the absolute sensitivity values are dependent on the resonance forced response level and resonance frequency values, which vary significantly with the normal load variation (e.g., for the case considered here, the resonance response level and the frequency change by factor of 10^3 and by 50% accordingly when N_0 is varied from 0% to 6400%). Hence, there is some difficulty in comparing absolute values of the sensitivity coefficients. In order to allow comparison of the sensitivity to all UPD parameters, the dimensionless sensitivities are plotted here. The dimensionless sensitivity characterizes the sensitivity of relative variation of the response level (or the resonance frequency) to relative variation of a parameter. They are expressed by the following formulas:

$$s^{\%}_{ja^{\mathrm{res}}}=(b_j/a^{\mathrm{res}})\,\partial a^{\mathrm{res}}/\partial b_j,\quad s^{\%}_{j\omega^{\mathrm{res}}}=(b_j/\omega^{\mathrm{res}})\,\partial \omega^{\mathrm{res}}/\partial b_j \quad (17)$$

The dimensionless sensitivity coefficients for the resonance frequency are shown in Fig. 4 and the sensitivities for the dimensionless resonance peak response level are shown in Fig. 5. The range of the normal load variation is large and the sensitivities vary significantly within it. In order to allow more detailed display of the sensitivity variation, in each figure two plots are drawn: (i) for lower values (from 0% to 200%) and (ii) for higher values (from 200% to 6400%).

One can see that for the bladed disk, considered dimensionless sensitivities with respect to variation of the friction coefficient μ and the normal load N_0 are very close as for the resonance frequency and as for the resonance peak response. They are close over the whole range of the normal load variation, which indicates

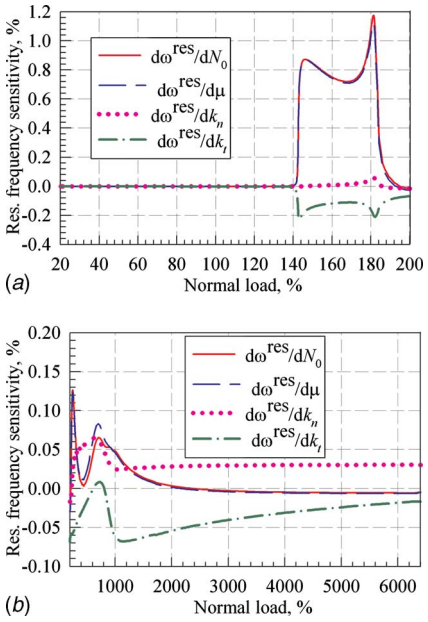


Fig. 4 Sensitivities of the resonance frequency: (*a*) for lower values of the normal load; (*b*) for higher values of the normal load

that for the considered bladed disk the increase of the friction coefficient and the increase of the damper mass will increase the resonance response and the resonance amplitude very similarly independently on the damper mass (which is proportional to the normal load) value.

The sensitivity of the resonance frequency to N_0 and μ is the largest for normal load/damper mass values within [140%, 190%] range. For this range, these UPD parameters are dominant in variation of the resonance frequency, although the sensitivity with respect to tangential stiffness coefficient k_t is noticeable and has negative sign, indicating that small k_t increase will cause decrease of the resonance frequency. For lower values of the normal load,

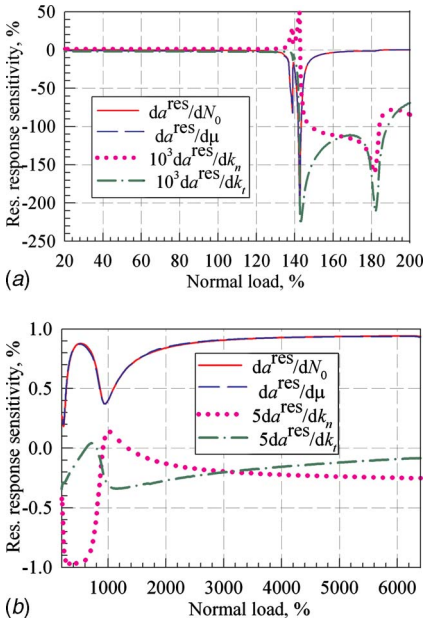


Fig. 5 Sensitivities of the resonance peak response level: (*a*) for lower values of the normal load; (*b*) for higher values of the normal load

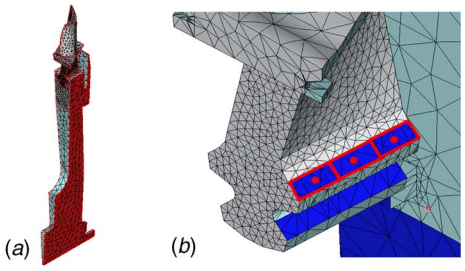


Fig. 6 A model of the bladed turbine disk with friction interfaces at blade-disk root joints: (*a*) a sector bladed disk FE model; (*b*) area friction contact elements applied contact interfaces

the resonance frequency sensitivity with respect to k_t and with respect the normal stiffness coefficient k_n are close to zero. For normal load values between 200% and 1000%, the resonance frequency sensitivities are comparable with respect to all four parameters and they change significantly over this range. For larger normal load values and, consequently, heavier damper mass, the damper stiffness properties become prevailing.

The resonance peak response is the most sensitive to variation of N_0 and μ. For a damper mass value of 144%, the dimensionless resonance response sensitivity reaches its maximum value of 200%. Therefore, the damper cannot work robustly since even a small variation of the damper parameters can change the resonance response significantly and the use of 144% should be avoided when the damper parameter choice is made. For lower normal load values, the resonance response is insensitive to damper stiffness coefficients k_t and k_n, while for higher values, these sensitivities are significant, although still smaller than sensitivity with respect to friction coefficient and the normal load.

4.2 A Turbine Bladed Disk With Friction Contacts at Blade-Disk Joints. As another example of application of the method developed, a bladed turbine disk was considered with damping generated at blade root friction contact interfaces, which are modeled by the area friction contact elements (Fig. 6).

The total number of blades in the bladed disk analyzed is 64. The damping loss factor η due to material and aerodynamic damping was assumed to be 0.001. The aerodynamic forces exciting the vibration analyzed were determined from aerodynamic calculations and are distributed over the blade airfoils. Excitation by 43rd engine-order (EO) is considered, which corresponds to a vibration mode of a tuned bladed disk with 64−43=21 nodal diameters and the frequency range corresponding to first flapwise (1F) mode is analyzed. A sector finite element model used in the analysis is shown in Fig. 6(*a*) and the number of DOFs in the sector model is 73,245. The blade root fir-tree has two lobes (see Fig. 6(*b*)). In the analysis performed, area friction contact elements were distributed over contact surfaces of first lobe (see Ref. [27]) and the lower lobes are assumed to be always fully stuck. The parameters of the area friction contact elements used in the friction modeling are as follows: (i) a value of the friction coefficient μ and (ii) a value of the normal stresses applied at the friction interfaces N_0, normal k_n, and tangential k_t, stiffness coefficients characterizing elasticity due to roughness of the contact surfaces.

In Fig. 7, the results of the direct parametric analysis for the resonance peaks (shown by bold solid curve) are compared with the forced response dependencies on the excitation frequency calculated for selected values of the gaps between contact surfaces. The forced response was determined at a node located at the blade tip. The positive gap values correspond to initial clearances and the negative values correspond to interferences that appear when centrifugal forces due to rotor rotation are applied to blades. The

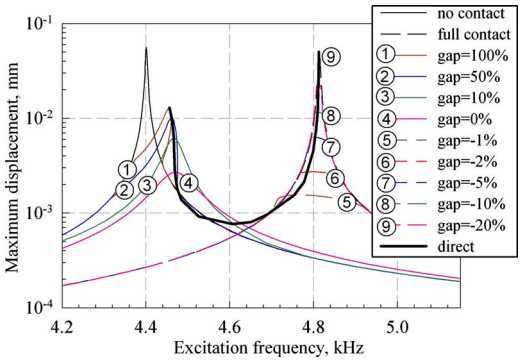


Fig. 7 Forced response calculated for different gap values and results of the direct parametric analysis

range of the gap values examined is from −100% to +100% of the reference value. For comparison, the forced responses of the two limiting cases are displayed also: (i) the forced response of the bladed disk without any contact at the first blade root lobe and (ii) the forced response calculated when there is a full contact over the first lobe and slip does not occur. Similar to previous case, one can see that the direct parametric analysis allows calculation of the resonance peaks accurately and reliably. For cases of positive gap values, the forced response exhibits stiffening behavior, since for higher amplitude levels the gaps are closed and the stiffness of the contact interfaces increases. For cases of negative gaps, there is a weak softening effect due to the opening interferences and accordingly stiffness reduction with higher amplitudes. Moreover, decrease of the interference values facilitates the friction damping (see Ref. [27]) and hence increase of the resonance amplitudes.

The dependency of the resonance frequency and response level on the gap value calculated by the direct parametric analysis is shown in Fig. 8. One can see that the resonance response level decreases with an increase in the gap value from −100% to −0.25%, where it reaches its minimum and then increases with the clearance increase. Theoretically, if the clearance can become so large that it cannot be closed by vibration, the resonance response become equal to the response of the bladed disk without the first lobe contacts.

The calculated sensitivities of the resonance frequency to parameters of the contact interfaces are plotted in Fig. 9 over the whole range of gap variation. One can see that for negative gap values, the sensitivities with respect to the friction coefficient and with respect to the normal stresses are similar, while for positive gap values, they differ significantly. For positive gap values, the sensitivity with respect to the normal stress is very close to the sensitivity with respect to the normal stiffness coefficient taken

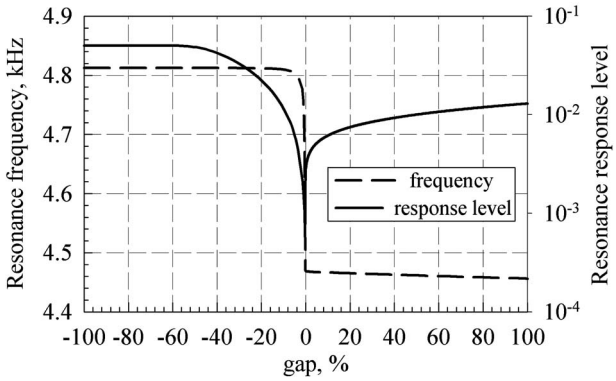


Fig. 8 Dependencies of the resonance frequency and response level on the normal load calculated

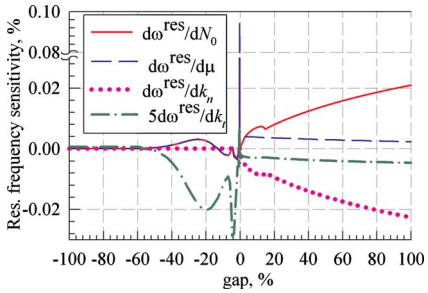


Fig. 9 Sensitivities of the resonance frequency to blade root joint parameters

with opposite signs. For negative gap values, the contact interface parameters start affecting the resonance frequency only from value −55% and in the range [−55%, 0%]. The sensitivities with respect to the friction coefficient and the tangential stiffness are solely significant.

The sensitivities of the resonance forced response level with respect to contact interface parameters are shown in Fig. 10. The response level sensitivity with respect to the friction coefficient and with respect to the normal stresses are close for negative gap values, while for positive gap values they differ significantly.

4.3 A High-Pressure Turbine Bladed Disk With Friction Contacts of Shrouds. A turbine high-pressure bladed disk shown in Fig. 11 has been also analyzed. The bladed disk comprises 92 shrouded blades and the finite single-sector model contains 162,708 degrees of freedom. The damping loss factor was assumed to be 0.003, and an excitation by fourth engine-order is

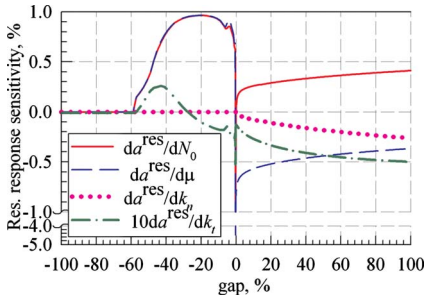


Fig. 10 Sensitivities of the resonance peak response level to blade root joint parameters

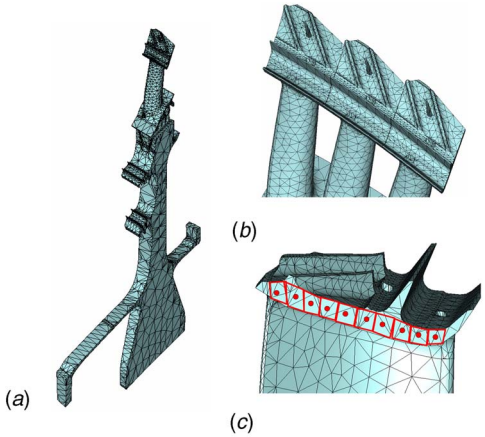


Fig. 11 A model of the high-pressure turbine disk with friction contacts of shroud: (*a*) a sector FE model; (*b*) the bladed disk shrouds; (*c*) friction contact element distributed over the contact interface

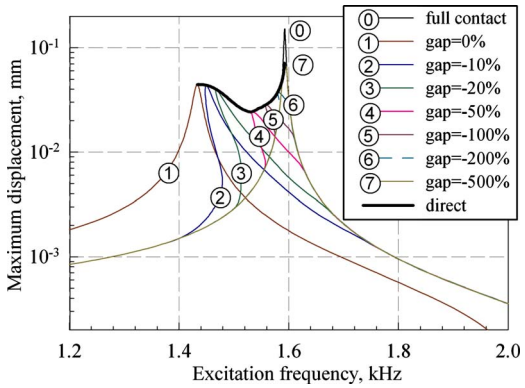


Fig. 12 Forced response calculated for different interference values and results of the direct parametric analysis

considered. The friction contact interactions of blade shrouds are modeled at each contact interface by ten friction contact elements (see Fig. 11(*c*)), which, as in two earlier considered cases, allow for unilateral nature of the force normal to the contact surface and friction forces depending on the normal load cyclic variations under vibrations. The parameters of these elements are the same as specified for the preceding case of the blade root joint interfaces.

The shroud contacts change the dynamic properties: resonance frequency, mode shapes, and resonance levels significantly. Hence two cases are examined here: (i) a case of negative gaps—interferences and (ii) a case of positive gaps—clearances. In both cases, the ranges considered for the clearance and the interference values are from 0% to 500%.

For a case of interferences, forced responses were calculated over a frequency range including the first resonance frequency. The comparison of conventional forced response calculations for different interference values and the direct parametric analysis is shown in Fig. 12.

The dependencies of the resonance frequency and the resonance response level on interference value obtained as a result of the direct parametric analysis are given in Fig. 13.

One can see that the forced response curves have softening characteristics, which is due to the reduced stiffness when amplitudes become large enough to open interferences. The reduction of the resonance response level is achieved owing to the energy dissipation at shroud contact interfaces and changes of the shroud contact stiffness. The resonance response has its minimum at the interference 45% of some reference value.

The sensitivity of the resonance frequency and resonance response to shroud contact interface parameters are shown in Figs. 14 and 15, respectively.

One can see that the parameter providing highest sensitivities is the normal stiffness coefficient k_n, which is dominant over the whole range of the interference variation. The sensitivity with respect to this parameter reaches a maximum in its absolute value at 32% of the interference value. At this interference value, the sensitivities with respect to N_0, μ, and k_t have also their extreme values. One can note that, in contrast to both previous bladed disks considered in this paper, the sensitivity with respect to the normal stresses and with respect to the friction coefficient is significantly different over the whole range of the interferences.

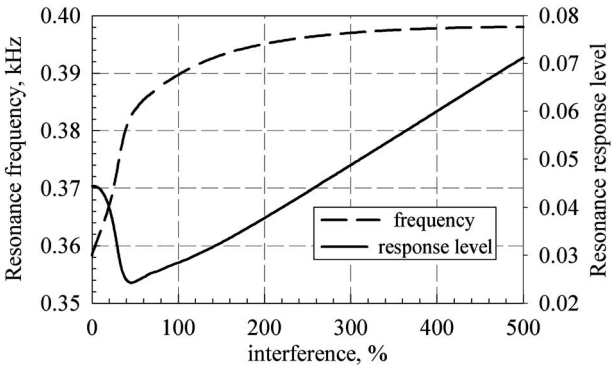


Fig. 13 Dependencies of the resonance frequency and response level on the normal load: a case of interferences

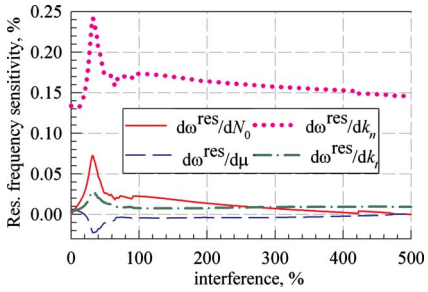


Fig. 14 Sensitivities of the resonance frequency to shroud friction contact parameters: a case of interferences

For a case of clearances, the frequency range analyzed includes two resonance peaks. The comparison of results obtained by the conventional forced response analysis performed for selected clearance values with the results of direct parametric analysis is shown in Fig. 16.

For both resonances, the direct parametric analysis curve goes through points of the maximum forced response level. The maximum response level is not necessarily achieved at the tip point of the forced response curve, as for a case of first resonance considered here, where the response levels are lower than at the points of maximum satisfying the condition given by Eq. (9). In the vicinity of first resonance peak, the forced response curves exhibit notice-

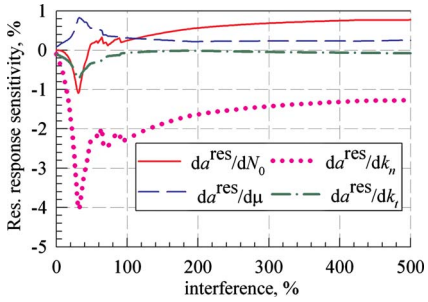


Fig. 15 Sensitivities of the resonance peak response level to shroud friction contact parameters: a case of interferences

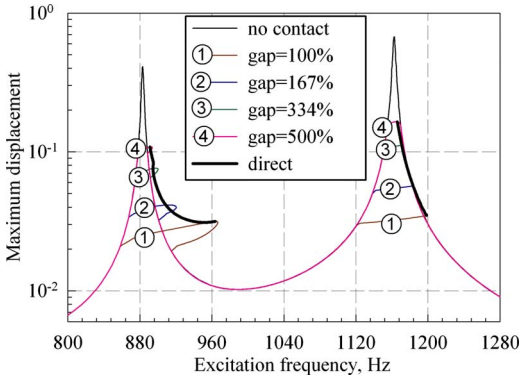


Fig. 16 Forced response calculated for different clearance values and results of the direct parametric analysis

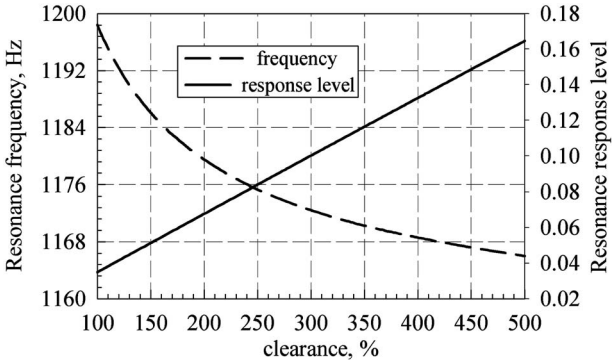


Fig. 17 Dependencies of the resonance frequency and response level on the normal load: a case of clearances

able stiffening behavior due to the closing at higher response levels, the clearances between the shrouds. In the vicinity of second resonance peak, stiffening can also be observed, but major source of the resonance response reduction is due to the dissipation of energy at the friction contact interfaces. Calculated dependencies of the resonance frequency and the resonance response level on the clearance value are plotted in Fig. 17 for second resonance peak.

Sensitivities of the resonance frequency and the resonance forced response level are shown in Figs. 18 and 19 for a range of clearance value variation from 100% to 500%.

It is evident that the dimensionless stiffness coefficients of the resonance frequency and the resonance response level calculated with respect to the clearance and with respect to the normal of the contact surface are similar over the whole range of the clearance value analyzed here. Resonance frequency sensitivity with respect to the friction coefficient is significant and negative, which indicates that increase of the friction coefficient can reduce the resonance frequency. The effects of the clearance and the normal stiffness on the resonance frequency are opposite to those of the friction coefficients: Their increase increases the resonance frequency. For the considered range of the clearance values, the sensitivity of the resonance response with respect to the clearance and the normal stiffness are negative and almost constant. The sensitivities of the resonance frequency and response coefficients with respect to variation of the tangential stiffness of the rough contact surface are very small.

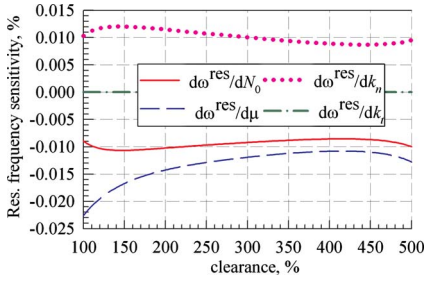


Fig. 18 Sensitivities of the resonance frequency to shroud friction contact parameters: a case of clearances

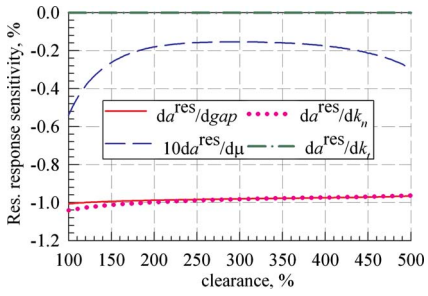


Fig. 19 Sensitivities of the resonance peak response level to shroud friction contact parameters: a case of clearances

5 Conclusions

For the first time, an effective method has been developed to calculate the sensitivity of the resonance frequency and resonance steady-state forced response level to variation of parameters of nonlinear friction contact interfaces and excitation.

The method allows determination of the sensitivity coefficients simultaneously with the direct parametric analysis of the resonance regimes, i.e., simultaneously with the resonance frequency and resonance peak response level calculated as a function of any contact interface parameter or as a function of the excitation level.

The method does not require additional computational expense of any significance compared with that needed for the direct parametric analysis alone.

The resonance peak sensitivities are calculated in frequency domain using multiharmonic representation of the forced response. Analytical expressions have been derived for determination of the resonance peak sensitivities, which include analytical derivatives of the multiharmonic balance equation of motion and of the resonance conditions.

Capabilities of the method are demonstrated on examples of analysis of large-scale finite element models of realistic bladed disks with major types of the nonlinear contact interfaces: (i) a blisk with underplatform dampers, (ii) a bladed disk with friction damping at blade fir-tree roots, and (iii) a high-pressure bladed disk with shroud contacts.

The numerical investigations show high efficiency of the method proposed in determination of resonance response and frequency sensitivities to parameters of friction contact interfaces. The sensitivity characteristics give important information about susceptibility of the resonance regimes to small parameter variations and can be used in analyses of statistic characteristics, robustness, and effects of uncertainties of the design parameters on forced response levels (see Refs. [16,17]).

Acknowledgment

The author is grateful to Rolls-Royce plc. for providing the financial support for this project and for giving permission to publish this work.

References

[1] Cameron, T. M., and Griffin, J. H., 1989, "An Alternating Frequency/Time Domain Method for Calculating Steady Response of Nonlinear Dynamic Systems," Trans. ASME, J. Appl. Mech., **56**, pp. 149–154.
[2] Sanliturk, K. Y., Imregun, M., and Ewins, D. J., 1997, "Harmonic Balance Vibration Analysis of Turbine Blades With Friction Dampers," Trans. ASME, J. Vib. Acoust., **119**, pp. 96–103.
[3] Yang, B.-D., and Menq, C.-H., 1997, "Modelling of Friction Contact and Its Application to the Design of Shroud Contact," Trans. ASME: J. Eng. Gas Turbines Power, **119**, pp. 958–963.
[4] Pierre, C., Ferri, A. A., and Dowell, E. H., 1985, "Multi-Harmonic Analysis of Dry Friction Damped Systems Using an Incremental Harmonic Balance Method," Trans. ASME, J. Appl. Mech., **52**, pp. 958–964.
[5] Wang, J. H., and Chen, W. K., 1993, "Investigation of the Vibration of a Blade With Friction Damper by HBM," Trans. ASME: J. Eng. Gas Turbines Power, **115**, pp. 294–299.
[6] Berthillier, M., Dupont, C., Mondal, R., and Barrau, R. R., 1998, "Blades Forced Response Analysis With Friction Dampers," Trans. ASME, J. Vib. Acoust., **120**, pp. 468–474.
[7] Chen, J. J., and Menq, C. H., 2001, "Prediction of Periodic Response of Blades Having 3D Nonlinear Shroud Constraints," Trans. ASME: J. Eng. Gas Turbines Power, **123**, pp. 901–909.
[8] Sanliturk, K. Y., Ewins, D. J., and Stanbridge, A. B., 2001, "Underplatform Dampers for Turbine Blades: Theoretical Modelling, Analysis and Comparison With Experimental Data," Trans. ASME: J. Eng. Gas Turbines Power, **123**, pp. 919–929.
[9] Csaba, G., 1998, "Forced Response Analysis in Time and Frequency Domains of a Tuned Bladed Disk With Friction Dampers," J. Sound Vib., **214**(3), pp. 395–412.

[10] Charleux, D., Gibert, C., Thouverez, F., and Dupeux, J., 2006, "Numerical and Experimental Study of Friction Damping in Blade Attachments of Rotating Bladed Disks," Int. J. Rotating Mach., **2006**, pp. 1–13.
[11] Laxalde, D., Thouverez, F., Sinou, J.-J., and Lombard, J.-P., 2007, "Qualitative Analysis of Forced Response of Blisks With Friction Ring Dampers," Eur. J. Mech. A/Solids. **26**, 676–687.
[12] Cigeroglu, E., An, N., and Menq, C.-H., 2007, "Wedge Damper Modeling and Forced Response Prediction of Frictionally Constrained Blades," ASME Paper No. GT2007-27963.
[13] Zucca, S., Borrajo, J., and Gola, M. M., 2006, "Forced Response of Bladed Disks in Cyclic Symmetry With Underplatform Dampers," ASME Paper No. GT2006-90785.
[14] Yang, B.-D., and Menq, C.-H., 1998, "Characterization of Contact Kinematics and Application to Design of Wedge Dampers in Turbomachinery Blading," Trans. ASME: J. Eng. Gas Turbines Power, **120**, pp. 410–423.
[15] Petrov, E. P., and Ewins, D. J., 2004, "State-of-the-Art Dynamic Analysis for Nonlinear Gas Turbine Structures," J. Aerosp. Eng., **218**(G3), pp. 199–211.
[16] Petrov, E. P., 2007, "Analysis of Robustness for Forced Response of Nonlinear Structures With Uncertainties in Parameters of Joints," *Proceedings of the 36th International Congress on Noise Control Engineering*, Istanbul, Turkey, Aug. 28–31.
[17] Petrov, E. P., 2007, "A Sensitivity-Based Method for Direct Stochastic Analysis of Nonlinear Forced Response for Bladed Discs With Friction Interfaces," ASME Paper No. GT2007-27981.
[18] Haug, E. J, Choi, K. K., and Komkov, V., 1986, *Design Sensitivity Analysis of Structural Systems*, Academic, Orlando, FL.
[19] Haftka, R. T., and Adelman, H. M., 1989, "Recent Developments in Structural Sensitivity Analysis," Struct. Optim., **1**, pp. 137–151.
[20] Keulen, F., Haftka, R., and Qu, X.-Y., 2000, "Review of Options for Structural Dynamics Optimization," *Proceedings of the 45th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics and Materials Conference and Exhibit*, Atlanta, USA, Apr. 3–6, Paper No. AIAA2000-1351.
[21] Petrov, E. P., 2007, "Direct Parametric Analysis of Resonance Regimes for Nonlinear Vibrations of Bladed Discs," ASME J. Turbomach., **129**, pp. 495–502.
[22] Petrov, E. P., and Ewins, D. J., 2003, "Analytical Formulation of Friction Interface Elements for Analysis of Nonlinear Multi-Harmonic Vibrations of Bladed Discs," ASME J. Turbomach., **125**, pp. 364–371.
[23] Petrov, E. P., 2004, "A Method for Use of Cyclic Symmetry Properties in Analysis of Nonlinear Multiharmonic Vibrations of Bladed Discs," ASME J. Turbomach., **126**, pp. 175–183.
[24] Petrov, E. P., 2004, "Method for Direct Parametric Analysis of Nonlinear Forced Response of Bladed Discs With Friction Contact Interfaces," ASME J. Turbomach., **126**, pp. 654–662.
[25] Elliott, R., Green, J. S., and Seinturier, E., 2005, "Aeroelastic Design of Turbine Blades–ADTurB II Overview," *Proceedings of the Sixth European Turbomachinery Conference*, Lille, France, Mar. 7–11, Paper No. AMP-105-01/62.
[26] Petrov, E. P., and Ewins, D. J., 2007, "Advanced Modelling of Underplatform Friction Dampers for Analysis of Bladed Disc Vibration," ASME J. Turbomach., **129**, pp. 143–150.
[27] Petrov, E. P., and Ewins, D. J., 2006, "Effects of Damping and Varying Contact Area at Blade-Disc Joints in Forced Response Analysis of Bladed Disc Assemblies," ASME J. Turbomach., **128**, pp. 403–410.

Further Development of a Smoke Sensor for Diesel Engines

D. Gould

D. P. Gardiner

M. LaViolette

W. D. Allan
e-mail: billy.allan@rmc.ca

Department of Mechanical Engineering,
Royal Military College of Canada,
Kingston, ON K7K 7B4, Canada

This paper describes experimental research aimed at developing an on-board smoke sensor for diesel engines. The sensor element was similar to a conventional spark plug. Electrical heating of the insulator was used to prevent carbon fouling from the diesel soot. The sensing element created sparks within the exhaust pipe and changes in smoke levels were detected through analysis of the voltage levels of the sparks. The system was tested in a heavy duty diesel engine equipped with exhaust gas recirculation (EGR) and compared with reference measurements of the filter smoke number (FSN). The experiments showed good sensitivity to step changes in smoke levels (accomplished by varying EGR levels) at smoke levels below 0.5 FSN. However, the sensor suffered from temperature induced signal drift and was unstable under some circumstances. The use of a spark plug with a smaller electrode tip diameter improved the signal stability. It is proposed that measurement and control of the electrode temperature will be necessary to control the signal drift. [DOI: 10.1115/1.2978995]

Keywords: diesel, smoke, soot, particulate, emissions, sensor, diagnostics, control

1 Introduction

It has been shown that an on-board diesel smoke sensor could be useful for control and diagnostic applications [1,2]. These include closed loop EGR control, transient fuel control, and on-board diagnostics for exhaust aftertreatment systems. Laboratory instruments capable of real time smoke measurements exist [3,4] but a rugged inexpensive sensor suitable for direct installation in a diesel exhaust pipe (like the exhaust gas oxygen sensor of a gasoline engine) is not currently available. Approaches that would require a sampling system or optical access to the exhaust have been considered to be unsuitable for vehicle applications [2].

A number of researchers have studied approaches to smoke sensing based on the electrical properties of the soot particles [5–10]. These approaches involve sensors that insert electrodes into the exhaust flow. In one type of sensor, the electrodes are used to detect the naturally occurring electrical charge of the soot particles, and no voltage is applied to the electrodes by the sensor circuit [5–7,10]. A more recent development is a type of sensor where a high voltage is applied between a pair of electrodes [8,9]. The electrical conductivity of the carbonaceous particles causes current flow across the electrode gap.

The approach described in this paper also uses a sensor with electrodes. However, in this case, the voltage between the electrodes is so high that a spark is created. During the period following breakdown when spark current flows across the gap through the spark, the spark sustaining voltage is monitored. Changes in smoke levels are detected based on changes to the spark voltage caused by the soot particles.

This concept was introduced in an earlier publication [11]. The original design goals were taken from the 2000 DOE Workshop on future sensor requirements [2] and included the ability to defect when engine-out smoke levels exceeded 2 Bosch smoke units with a response time of about 1 s. As described further in this paper, improvements in the smoke emissions of diesel engines have made it necessary to measure much lower smoke levels. The present paper describes the results of ongoing work to improve the sensitivity to low smoke levels and to understand and overcome problems that affected the consistency of the measurement.

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received December 1, 2006; final manuscript received May 5, 2008; published online December 18, 2008. Assoc. Editor: Margaret Wooldridge. Paper presented at the 2005 Fall Conference of the ASME Internal Combustion Engine Division (ICEF2005), Ottawa, Ontario, Canada, September 11–14, 2005.

2 Background

Figure 1 shows the voltage waveform of an actual spark with the phases of the spark identified based on the definitions provided by Maly [12]. When the ignition system is energized, the voltage applied to the spark plug increases until the gas between the electrodes ionizes and forms a spark. This leads to an abrupt drop in voltage, termed the breakdown phase. Following breakdown, the spark is sustained at voltage levels that are relatively low compared with the peak voltage needed to create the spark. Measurements of this spark sustaining voltage are used to determine smoke levels in the current approach.

As shown in Fig. 1, the sustaining voltage of a spark can be at two different levels depending on whether conditions lead to an arc or a glow type of discharge. The difference in voltage levels between the arc phase and the glow phase is due primarily to the lower cathode fall voltage of the arc phase. This lower cathode fall voltage exists because an arc discharge achieves electron emission from tiny molten hot spots on the cathode surface [12].

Glow discharges have a cold cathode electron liberation mechanism, which requires a higher voltage to sustain the spark [12]. Low current sparks ($<$50 mA) are predominantly glow discharges because the current density needed to create and maintain arc hot spots is unavailable. However, brief periods of arc activity can be observed during the first few microseconds of the spark when the highest current levels are experienced.

It was shown in earlier work [11] that the presence of carbon at the cathode could increase the duration of the arc phase and lower the minimum current threshold for an arc phase to occur. This behavior offered a means of detecting the carbonaceous soot particles in diesel exhaust by determining whether a spark was in arc phase or glow phase at a given elapsed time after breakdown.

An initial version of the sensor was developed and tested [11], and results from this work are shown in Fig. 2. These tests showed a reasonable correlation between the sensor signal and the filter smoke number (FSN) in the range of 1–3. However, the results were inconsistent from test to test, and there was poor sensitivity at low smoke levels.

In engine tests using current diesel engine technology with high pressure common rail injection, Stumpf et al. [13] measured FSN values below 0.5 for all of the operating conditions tested except

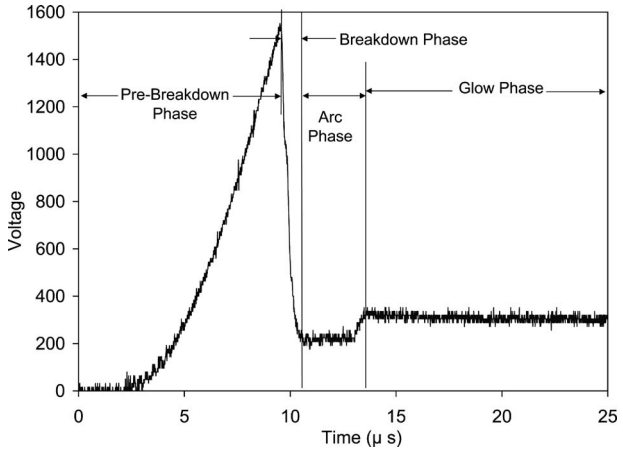


Fig. 1 Spark voltage waveform

those with high EGR rates [13]. Furthermore, this study showed that increasing the injection pressure could reduce soot emissions such that FSN values below 0.1 could be achieved over operating conditions ranging from idle to 1.2 MPa brake mean effective pressure (BMEP). Thus an on-board smoke sensor will require high sensitivity in order to provide useful diagnostic information for current and future diesel engines. The remainder of this paper describes new results using an improved apparatus and focuses on measurements of smoke levels below 1 FSN.

Muntean [14] demonstrated that there is good correlation between the smoke number measurements and the dry particulate mass concentration in the exhaust gas. The correlation with the total particulate mass depends on the soluble organic fraction adsorbed in the solid particles. Muntean stated that "given the predominance of the dry portion of the particulates in heavy duty engines, it can provide a good surrogate indicator of the total particulate mass of the engine." Thus, the authors believe that the trends indicated by the reference FSN measurements in the current study would reflect variations in the particulate mass emissions of the test engine.

3 Experimental Details

3.1 Test Engine. The test engine was a Steyr model WD615.98 9.7 l 6-cylinder heavy duty direct injection diesel mounted on a water brake engine dynamometer. The engine was originally turbocharged and intercooled in its vehicle application (a 10 ton heavy logistic military vehicle) and had a compression ratio of 15.5:1. For the tests described in this paper, the turbocharger was removed (converting it to a naturally aspirated engine) so that a simple EGR system (a valve controlled duct between the intake and exhaust pipes) could be used to vary exhaust smoke levels over a wide enough range at a fixed speed and load. The EGR was used only to increase the exhaust smoke levels, and the actual EGR rates were not measured.

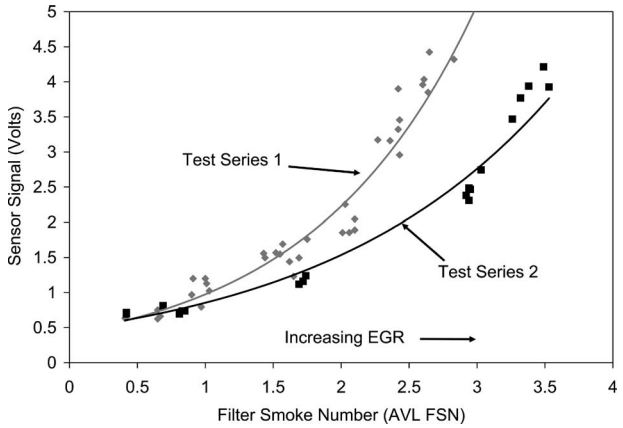


Fig. 2 Correlation between first generation smoke sensor signal and filter smoke number (1000 rpm, 3.25 bar BMEP, variable EGR). Adapted from Allan et al. [11]

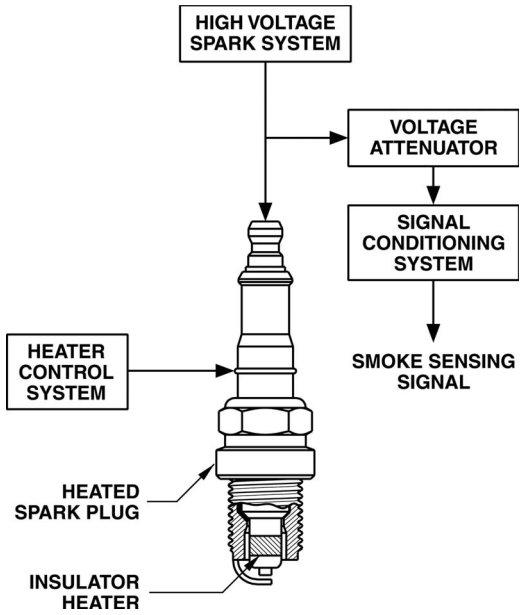


Fig. 3 Smoke sensor system

3.2 Test Conditions. The engine was tested at 1000 rpm and 150 N m torque (1.9 bar BMEP). This test condition was selected because with the simple EGR system, it was possible to vary the exhaust smoke levels from a FSN below 0.15 to FSN values above 1. During the EGR sweeps, the engine was operated at a fixed fuel rack setting as the changes in EGR levels had little impact on the engine speed and torque. The authors recognize that the engine setup was not representative of current diesel technology. The approach used for the present work simply provided a means of exposing the developmental sensor to the desired range of smoke levels for early proof-of-concept evaluations.

3.3 Reference Measurements. Reference smoke measurements were obtained using an AVL 415S smoke meter. This device performs automatic smoke measurements based on the filter paper method [15]. As described in Ref. [15], the filter paper method uses measurements of the diffuse reflection of the blackened filter paper to determine the FSN. Unlike older manual and automated piston pump approaches, it uses continuous sampling through a diaphragm pump. The sampled volume can be varied in relation to the soot concentration, and the volume used to calculate the smoke values is corrected to standard pressure and temperature conditions of 1 bar and 25°C, respectively.

The output of this smoke meter is expressed as a FSN from 0 to 10. FSN measurements are similar to the Bosch smoke number measurements obtained from a manual pump device except that the manual pump device does not account for the pressure and temperature of the sampled volume [15]. During the remainder of this paper, the reference smoke measurements will be presented as FSN values.

3.4 Smoke Sensor. The smoke sensor system is depicted in Fig. 3. The latest version of the smoke sensor is based on a prototype electrically heated spark plug provided by a spark plug manufacturer. The heating element also serves as a temperature sensor so that the insulator temperature may be held constant at

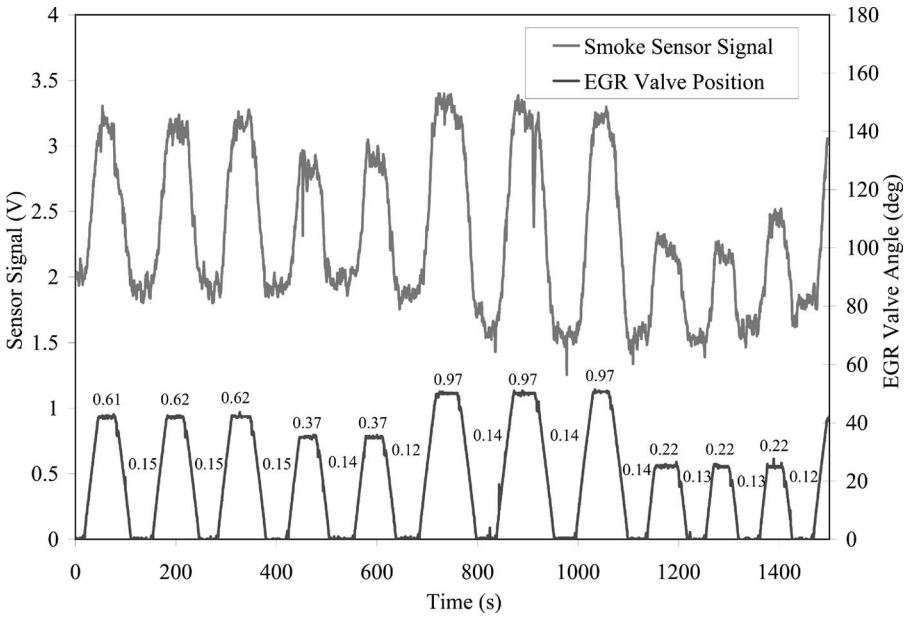


Fig. 4 Smoke sensor signal at varying EGR levels (heated spark plug) (numerical values are FSN values for steady EGR valve positions)

different exhaust temperatures. This arrangement has proven to eliminate carbon-fouling problems while requiring only modest heater power levels (typically <25 W). Further details regarding the design of the heated spark plug are available in Ref. [16].

The spark plug was fired by a programmable current ignition system, which allowed the spark current to be adjusted so that few arc discharges were produced when no soot was present. The spark voltage was attenuated by a 1000:1 probe. The signal conditioning system compared the voltage of each spark to a known threshold and determined whether the discharge was in arc mode or glow mode at a selected time after breakdown. The smoke sensor signal (0–5 V) corresponded to the percentage of sparks that were in arc mode.

In comparison with the work described in Ref. [11], sensitivity to low smoke levels was improved through higher fidelity measurements of the spark voltage and by focusing on measurements very early after the breakdown phase of the spark. Note that the timing of the breakdown phase is depicted in Fig. 1.

4 Results and Discussion

4.1 Signal Response to Variations in Smoke Levels. Figure 4 shows an example of the response of the sensor to variations in smoke levels caused by changes in EGR flows. For each EGR cycle, the EGR valve was opened from the fully closed position to a target EGR level and then returned to the fully closed position. The closed EGR valve position resulted in smoke levels between 0.12 FSN and 0.15 FSN while the different open positions provided smoke levels between 0.22 FSN and 0.97 FSN.

It can be seen that the sensor exhibited a substantial change in signal level to each up or down step in EGR including those which generated differences in smoke levels as low as 0.1 FSN. However, the relationship between the sensor signal and the actual FSN value was inconsistent. For example, the signal level for the closed EGR value setting varied between 1.5 V and 2 V. The filter smoke measurements without EGR varied little (from 0.12 FSN to 0.15 FSN), and the lowest signal values from the smoke sensor did not correspond to the lowest FSN measurements. In other words, the smoke sensor responded to changes in smoke levels but had an inconsistent baseline or zero level. We have used the term "signal drift" to refer to this problem. As shown in Fig. 5, the resulting correlation between the smoke sensor signal and actual FSN values exhibited a substantial amount of scatter.

Figure 6 examines the response of the smoke sensor in terms of the change in signal (Δ sensor voltage) occurring during a change in the EGR valve position versus the accompanying change in filter smoke number (Δ FSN). The "delta plot" depicts what the correlation might be if the baseline drift of the smoke sensor could be controlled. As shown in Fig. 6, plotting the data in this manner removed much but not all of the scatter. Evidently drift in the baseline signal was not solely responsible for imperfections in the correlation. Thus, variations or drift in the effective gain or span of the signal and/or underlying inconsistencies in the sensor's response at a given smoke level also existed.

4.2 Signal Drift and Instability. The origins of the signal drift problem may be explained by examining the temperature sensitivity of the sensor. Figure 7 shows the response of the sensor to changes in operating temperature induced by varying the power supplied to the spark plug's electric heater. The resistance of the heating element (which has a positive temperature coefficient of resistance) was measured to provide a relative indication of the sensor temperature. It can be seen that increasing the temperature

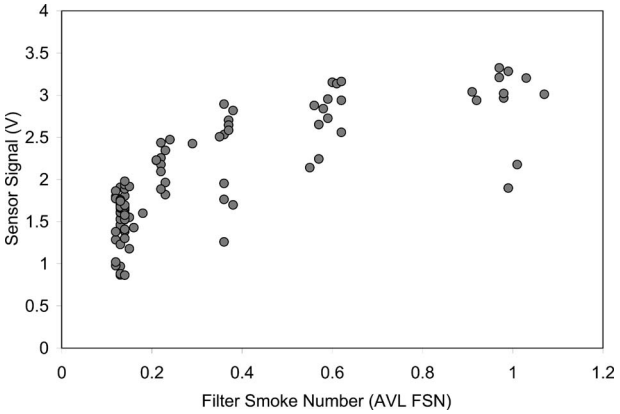


Fig. 5 Correlation between heated spark plug smoke sensor signal and filter smoke number (1000 rpm, 1.9 bar BMEP, variable EGR)

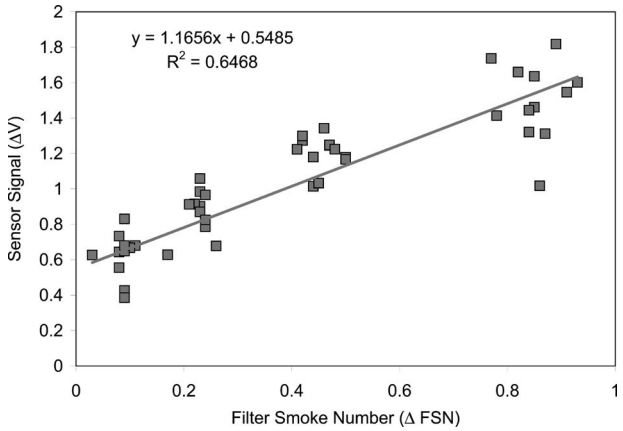


Fig. 6 Correlation between change in heated spark plug sensor signal and change in filter smoke number (1000 rpm, 1.9 bar BMEP, variable EGR)

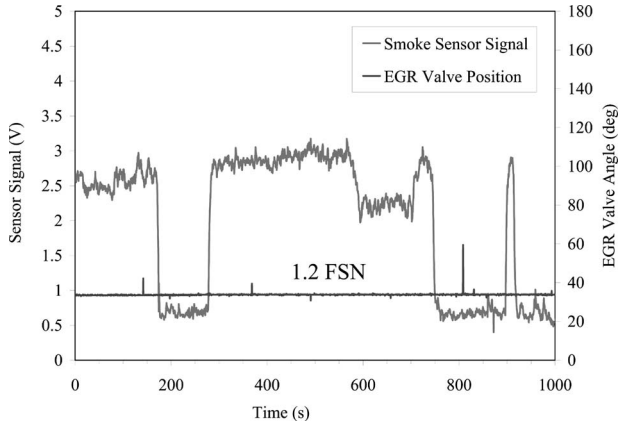


Fig. 9 Short term instability behavior (numerical values are FSN values for steady EGR valve positions)

reduced the sensor signal and vice versa. Thus the sensor was sensitive to temperature as well as smoke. This was problematic because increasing the EGR levels to increase the smoke also raised the temperature of the exhaust gas and exhaust pipe; consequently the sensor temperature rose as the EGR was increased.

Figure 8 shows the behavior of the sensor signal during two consecutive EGR cycles. It can be seen that the signal increased abruptly when the EGR valve was opened and the FSN increased from 0.12 to 0.52. However, the signal drifted downward during the period the EGR valve was open and the sensor was warming up. When the EGR valve was closed (and the exhaust temperature fell), the initial signal drop was followed by upward drift in the sensor signal as the sensor cooled off.

Note that the sensor signal levels at the beginning and end of the plot were almost identical despite the fact that the smoke levels were 0.12 FSN and 0.61 FSN, respectively. This shows that the temperature sensitivity could offset the smoke sensitivity completely under some circumstances. Maintaining a constant heater temperature (which determines the insulator temperature) did not eliminate the temperature sensitivity problem. It is likely that the temperature of the center electrode tip must be controlled and temperature sensing within the electrode will be required.

Another problem that was encountered can be described as "short-term instability." A particularly severe example of this is shown in Fig. 9. It can be seen that abrupt changes in the signal level could occur during operation at a constant EGR level despite a constant FSN.

Examination of the spark plug electrodes (Fig. 10) indicated that the spark location could change during operation, as the outer edge of the 1.85 mm diameter center electrode had a polished

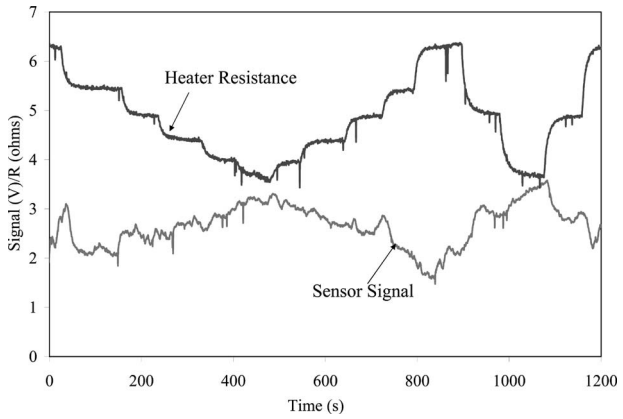


Fig. 7 Effect of electric heating on sensor signal

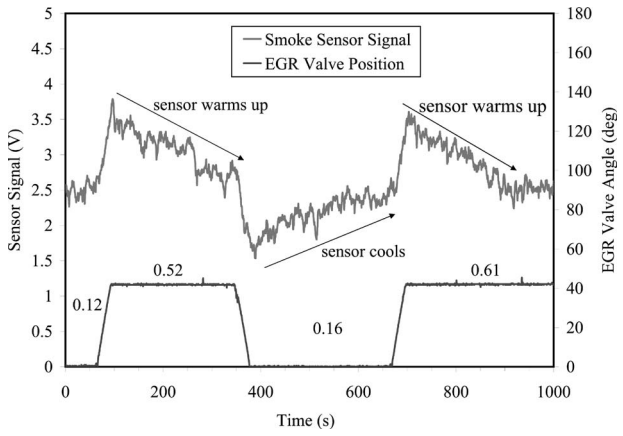


Fig. 8 Signal drift during EGR cycle (numerical values are FSN values for steady EGR valve positions)

Fig. 10 Arcing pattern on heated spark plug

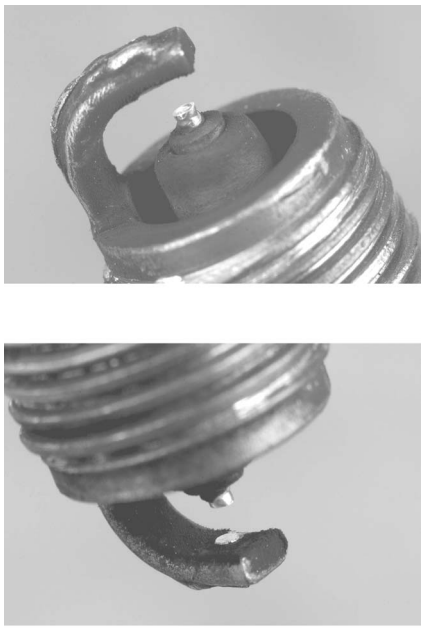


Fig. 11 Arcing pattern on platinum pin spark plug

appearance due to arcing. (Note the clean appearance of the insulator nose due to the electric heating.) It was hypothesized that the spark might encounter variations in soot particle buildup as it moved across the relatively large surface area of the electrode, and this could cause abrupt changes in the signal.

In order to investigate this possibility, a spark plug with a much smaller "active" electrode area was tested. This spark plug had a 1.0 mm diameter platinum pins on the center and ground electrode so the spark remained focused over a smaller area. The platinum spark plug was not electrically heated but proved to be suitable for brief tests at low smoke levels where serious insulator fouling could be temporarily avoided. As shown in Fig. 11, the spark remained confined to the smaller area of the platinum pins. Note also that the fouled appearance of the insulator in contrast with the electrically heated spark plug in Fig. 10.

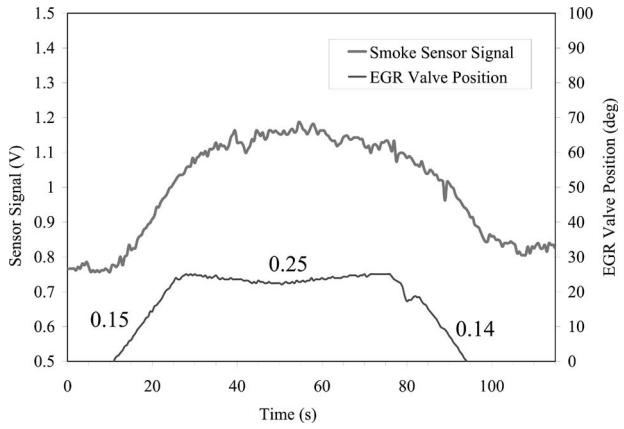


Fig. 13 Smoke sensor response to 0.1 FSN change in smoke (numerical values are FSN values for steady EGR valve positions)

Results with the platinum spark plug are shown in Fig. 12. Short-term instability problems were markedly reduced. As shown in Fig. 13, good signal response was achieved for variations in smoke levels as low as 0.1 FSN. It should be emphasized that this improved performance was achieved simply because the electrode had a smaller surface area. The use of platinum provides a means of achieving acceptable durability with a small diameter electrode.

The correlation between the sensor signal (using the unheated platinum spark plug) and the filter smoke number is shown in Fig. 14. In Fig. 15 the data are plotted in the same manner as in Fig. 6, showing the relationship between the change in smoke sensor signal and the change in filter smoke number during changes in the EGR valve position. From these figures, it is apparent that further work is needed to achieve a more consistent relationship between the sensor signal and actual smoke levels.

Newer spark plug designs with even smaller precious metal electrode pins (platinum and iridium) are now available. Future work on the smoke sensor concept will employ these designs to

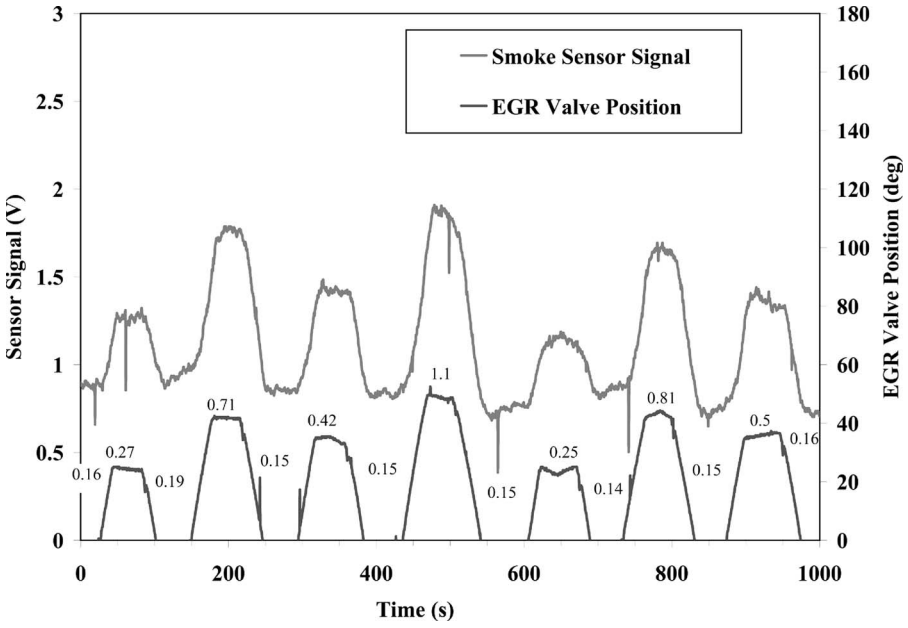


Fig. 12 Smoke sensor signal at varying EGR levels (unheated platinum spark plug) (numerical values are FSN values for steady EGR valve positions)

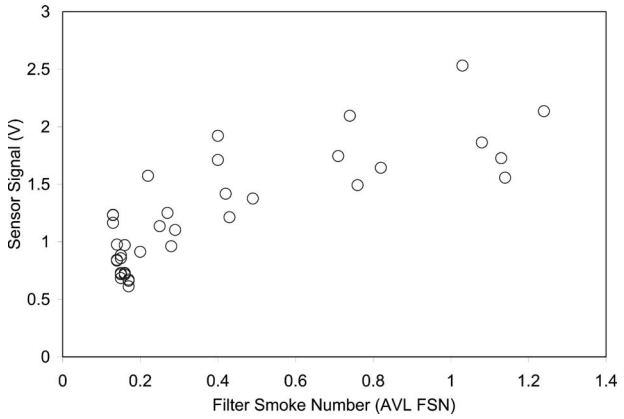


Fig. 14 Correlation between unheated platinum spark plug smoke sensor signal and filter smoke number (1000 rpm, 1.9 bar BMEP variable EGR)

eliminate short-term instability along with electrode temperature sensing and active heater control to overcome temperature-related signal drift problems.

5 Summary

Further development of a diesel smoke sensor has been carried out using an electrically heated spark plug as the exhaust probe and analysis of the spark voltage waveforms to detect changes in smoke levels. It has been demonstrated that this concept offers good sensitivity to step changes in smoke levels (accomplished by varying EGR levels) at smoke levels below 0.5 FSN. However, the signal drifts as the exhaust gas and exhaust pipe (where the sensor is mounted) warm up or cool down. Maintaining a constant heater temperature (which determines the insulator temperature) cannot eliminate this temperature sensitivity. It is likely that the temperature of the center electrode tip must be controlled and temperature sensing within the electrode will be required.

The sensor has also exhibited short-term instability problems. This is believed to be related to the movement of the spark between different locations on the electrode surface. Preliminary tests using an unheated spark plug (which is very vulnerable to fouling) with a smaller electrode tip diameter have shown a marked improvement in short-term signal stability.

It is recognized that there are numerous other variables that may have the potential to impact the sensitivity of the smoke measurements such as differences in particulate characteristics (due to the factors such as fuel composition, engine design, and operating conditions), flow conditions at the spark gap, and the composition of the exhaust gas. Further work will be required to address these and other issues.

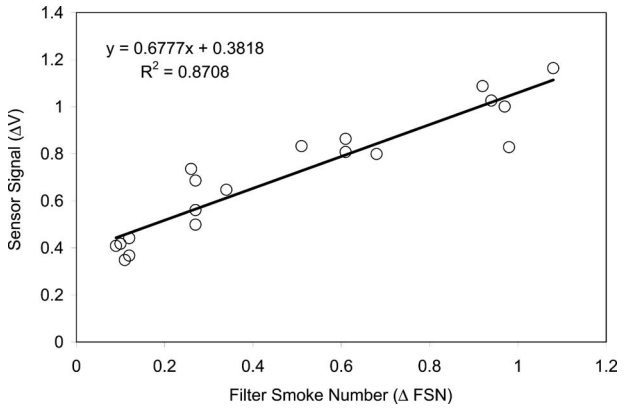


Fig. 15 Correlation between change in unheated platinum spark plug sensor signal and change in filter smoke number

Acknowledgment

The financial support of Canada's Program of Energy Research and Development (PERD) is gratefully acknowledged. The authors also wish to thank Dr. Ron Patrick of ECM: Engine Control and Monitoring for providing the heated spark plug. The assistance of Stephen Guy and Dan Faux with the experimental setup is gratefully acknowledged.

References

[1] Zhao, F., and Asmus, T., 2000, "Diesel Closed Loop Control via Smoke Sensor," presented at the Department of Energy Sensor Workshop, Berkeley, CA, Jan. 25–26.
[2] Anon., 2000, "Particulate Sensors: Requirements, Recommendations/Issues," Minutes of the Department of Energy Sensor Workshop, Berkeley, CA, Jan. 25–26.
[3] Schrami, S., Heimgartner, C., Will, S., and Leipertz, A., 2000, "Application of a New Soot Sensor for Exhaust Emission Control Based on Time Resolved Laser Induced Incandescence (TIRE-LII)," Society of Automotive Engineers, SAE Paper No. 2000-01-2864.
[4] Halsch, C., Beck, H. A., and Niessner, R., 2004, "A Photoacoustic Sensor System for Time Resolved Quantification of Diesel Soot Emissions," Society of Automotive Engineers, SAE Paper No. 2004-01-0968.
[5] Schweimer, G. W., 1986, "Ion Probe in the Exhaust Manifold of Diesel Engines," Society of Automotive Engineers, SAE Paper No. 860012.
[6] Collings, N., Baker, N., and Wolber, W. G., 1986, "Real-Time Smoke Sensor for Diesel Engines," Society of Automotive Engineers, SAE Paper No. 860157.
[7] Hong, G., Collings, N., and Baker, N. J., 1987, "Diesel Smoke Transient Control Using a Real-Time Smoke Sensor," Society of Automotive Engineers, SAE Paper No. 871629.
[8] Warey, A., Hendrix, B., and Hall, M., 2004, "A New Sensor for On-Board Detection of Particulate Carbon Mass Emissions from Engines," Society of Automotive Engineers, SAE Paper No. 2004-01-2906.
[9] Warey, A., and Hall, M. J., 2005, "Performance Characteristics of a New On-Board Engine Exhaust Particulate Matter Sensor," Powertrain and Fluid Systems Conference and Exhibition, San Antonio, TX, Oct. 24–27, SAE Paper No. 2005-01-3792.
[10] Kittelson, D., Ma, H., Rhodes, M., and Krafthefer, B., 2004, "Particle Sensor for Diesel Combustion Monitoring," 2004 Diesel Engine Emissions Reduction (DEER) Conference, Aug. 29–Sep. 2.
[11] Allan, W. D. E., Freeman, R. D., Pucher, G. R., Faux, D., Bardon, M. F., and Gardiner, D. P., 2003, "Development of a Smoke Sensor for Diesel Engines," SAE Powertrain Congress, Pittsburgh, PA, Oct. 27–30, SAE Paper No. 03-01-3084.
[12] Maly, R., 1984, "Spark Ignition: Its Physics and Effects on the Internal Combustion Process," *Fuel Economy in Road Vehicles Powered by Spark Ignition Engines*, C. J. Hilliard, ed., Plenum, New York.
[13] Stumpf, M., Merkel, S., Eckert, P., Wagner, U., Velji, A., Spicher, U., and Moehler, O., 2007, "Investigation of Soot Concentration and Particle Size Distribution on a Single Cylinder Diesel Engine," *Proceedings of the ASME Internal Combustion Engine Division Fall Technical Conference, ICEF 2007*, Charleston, SC, Oct. 14–17, Paper No. ICEF2007-1709.
[14] Muntean, G. G., 1999, "A Theoretical Model for the Correlation of Smoke Number to Dry Particulate Concentration in Diesel Exhaust," SAE Paper No. 1999-01-0515.
[15] Anon., 2001, "Measurements of Smoke Values With the Filter Paper Method," AVL List GMBH, AVL Engine Instrumentation Application Note AT1007E.
[16] Suzuki, T., and Ito, Y., 2000, "Heater Equipped Spark Plug," U.S. Patent No. 6,060,821.

Study of Fuel Temperature Effects on Fuel Injection, Combustion, and Emissions of Direct-Injection Diesel Engines

Gong Chen
Department of Mechanical Engineering,
Gannon University.
Erie, PA 16541

The influence of inlet liquid fuel temperature on direct-injection diesel engines can be noticeable and significant. The work in this paper investigates the effects of inlet fuel temperature on fuel-injection in-cylinder combustion, and output performance and emissions of medium-speed diesel engines. An enhanced understanding and simplified modeling of the variations in the main fuel-injection parameters affected by inlet fuel temperature are developed. The study indicates that the main injection parameters affected include the injection timing at the injector end relative to the injection-pump actuation timing, the fuel-injection rate, the fuel-injection duration, and the injection spray atomization. The primary fuel temperature effects on the injection parameters are from the fuel bulk modulus of elasticity and the density with the fuel viscosity less significant as the injector-nozzle flow is usually in a turbulent region. The developed models are able to predict the changes in the injection parameters versus the inlet fuel temperature. As the inlet fuel temperature increases, the nozzle fuel-injection-start timing is predicted to be relatively retarded, the injection rate is reduced, and the needle-lift duration is prolonged from the baseline. The variation trends of the engine outputs and emissions versus fuel temperature are analyzed by considering its consequent effect on in-cylinder combustion processes. It is predicted that raising fuel temperature would result in an increase in each of CO, HC, PM, and smoke emissions, and in a decrease in NO_x, and may adversely affect the fuel efficiency for a general type of diesel engine at a full-load condition. The experimental results of the outputs and emissions from testing a medium-speed four-stroke diesel engine agreed with the trends analytically predicted. The understanding and models can be applied to compression-ignition direct-injection liquid fuel engines in general. [DOI: 10.1115/1.3019006]

1 Introduction

The inlet liquid fuel properties and conditions influence the combustion processes of compression-ignition direct-injection diesel engines. Some previous work and results from studying the effects of the fuel properties and conditions on diesel engine fuel injection and combustion were presented by Hsu and co-worker [1,2] and Maeda et al. [3]. Particularly, an investigation of the fuel temperature effects on fuel-injection behavior and engine combustion was conducted and some of the results from the investigation were available and reported by Hsu [1].

Inlet fuel temperature (T_f) herewith studied in this paper refers to the temperature of the liquid fuel at the inlet to the engine fuel-injection pump, as shown in Fig. 1. For an engine without controlling or modulating fuel temperature in operation, T_f is largely related to the ambient temperature.

For high-power medium-speed diesel engines, which usually utilize the configuration of a quiescent-type combustion chamber and direct fuel injection, a timely mixing of fuel and air in the engine cylinders to achieve an efficient combustion depends heavily on the fuel-injection process, in which injection pressure is highly important. Thus, this type of engine invariably requires relatively high injection pressure, and the effects of inlet fuel temperature through fuel injection on in-cylinder combustion of this type of engines are more noticeable and significant.

An appropriate understanding and determination of the fuel temperature effects on engine processes, performances, and output parameters are important in the evaluation of engine design and engine capabilities. Some analyses and experimental results for the effects of engine inlet fuel temperature on fuel injection and combustion have been reported, such as those by Hsu [1]. The following issues need to be studied and further investigated: (a) appropriate and simple models, which can be used to describe and predict the variations in injection parameters over the inlet fuel temperature, (b) understanding of the variations of in-cylinder combustion parameters versus inlet fuel temperature, and (c) analytical and engine experimental investigation of the variation trends in the engine output performances and emissions over inlet fuel temperature.

The study in this paper is inspired to address those issues and to investigate the effects of inlet fuel temperature on those engine parameters of interest with diesel engines. An enhanced fundamental understanding and simplified modeling of the variations in main fuel-injection parameters affected by inlet fuel temperature are developed. The variation trends of the engine performance and emissions versus fuel temperature are analyzed by considering its consequent effect on in-cylinder combustion processes. The predicted variation trends are compared with experimental results from testing a medium-speed four-stroke diesel engine.

The fuel temperature range under this study is a range during which the liquid fuels are well in their liquid phase. While the resulting model equations appear in the main sections of the text, the derivation and development of the equations for modeling the fuel-injection parameters affected by inlet fuel temperature in this work are described in more detail in the Appendix.

Manuscript received October 24, 2007; final manuscript received August 23, 2008; published online December 18, 2008. Review conducted by Margaret Wooldridge. Paper presented at the ASME International Combustion Engine Division 2003 Fall Technical Conference (ICEF2003), Erie, PA, Sept. 7–10, 2003.

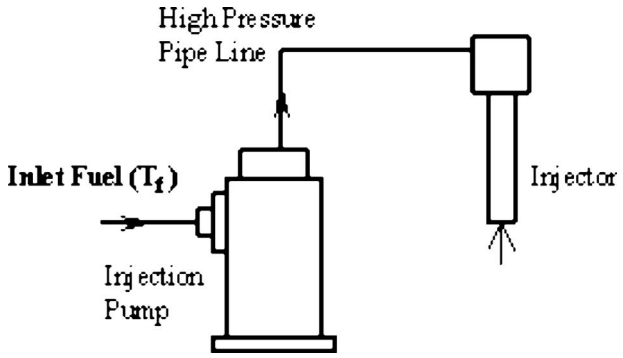


Fig. 1 Indication of engine inlet fuel temperature

2 Influences on Fuel Properties

The output performance and exhaust emissions of diesel engines are essentially related to the in-cylinder combustion processes. The effects of inlet fuel temperature on the engine in-cylinder combustion are mainly through its influence on fuel injection. The fuel properties, which are affected by fuel temperature and have relative-significant effects on fuel injection, need to be studied, and their variations need to be modeled.

Fuel properties affected. The physical properties of petroleum fuel, such as diesel, vary as its temperature changes, as previously described by Burman and DeLuca [4], Schmidt [5], and Hsu and co-worker [1,2]. Changes in fuel discharge and fuel injection into the cylinder due to varying the engine inlet fuel temperature ((T_f) mainly result from changing the following fuel properties: the fuel bulk modulus of elasticity (K_f), the fuel density (ρ_f), and the fuel viscosity (μ_f or ν_f). As T_f changes, these fuel properties vary. It can be estimated that the effects of changing the fuel viscosity on the fuel-injection process due to varying fuel temperatures are less significant than the effects of changing K_f and ρ_f, as mentioned by Hsu [1]. This is further analyzed later in this paper. In general, the modulus of elasticity and the density of a particular liquid fuel can be expressed as functions of fuel pressure and temperature as

$$K_f = f_1(p_f, T_f) \quad \text{and} \quad \rho_f = f_2(p_f, T_f) \tag{1}$$

Using the values of K_f and ρ_f of typical petroleum fuels from Ref. [4], Fig. 2 depicts the estimated trends in the variation in K_f at various fuel pressures, and ρ_f and ν_f versus fuel-inlet temperature T_f. The trends of variation are presented in Fig. 2 by the normal-

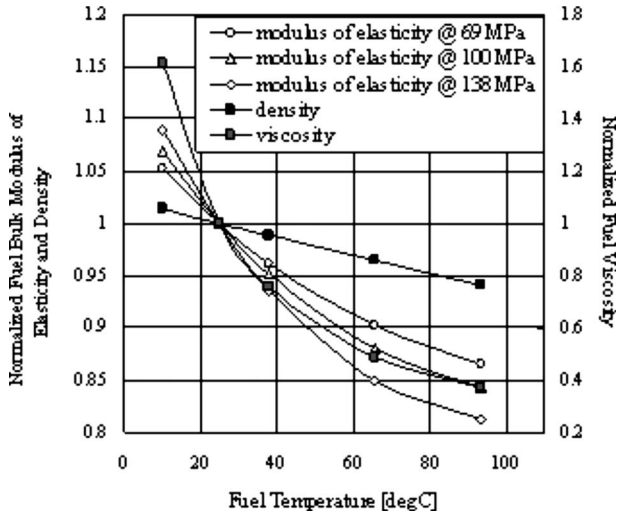


Fig. 2 Temperature effects on modulus of elasticity, density, and viscosity of liquid fuel

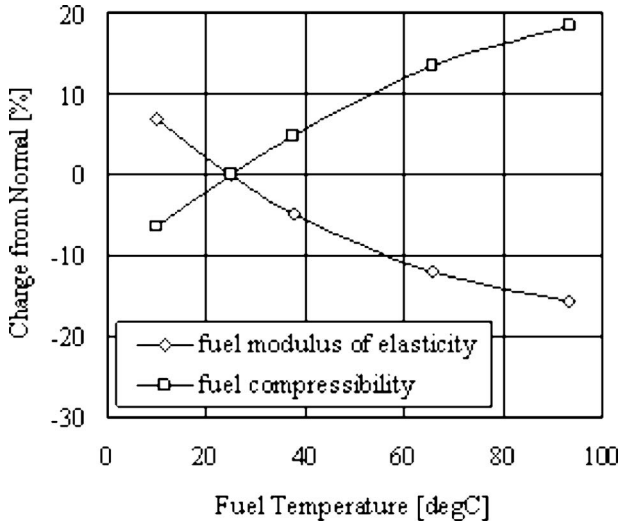


Fig. 3 Fuel compressibility versus fuel temperature (p_{f0} =100 MPa)

ized fuel properties, which are defined as the ratios of the fuel properties at an arbitrary fuel temperature to those properties at a fuel standard temperature. For instance, the trends of the normalized fuel properties show that as T_f increases, ρ_f decreases, and K_f is decreased.

Effects on fuel acoustic velocity and compressibility. The acoustic velocity (u_p) and compressibility (β_f) of a liquid fuel are related to the modulus of elasticity and the density of the fuel, as shown in Refs. [1,4]. Based on the definition of the acoustic velocity and compressibility, the variations in u_p and β_f are derived in the Appendix to give the equations in connection to K_f and ρ_f as

$$\frac{\Delta u_p}{u_{p0}} = \sqrt{\frac{K_f}{K_{f0}}\frac{\rho_{f0}}{\rho_f}} - 1 \tag{2}$$

and

$$\frac{\Delta \beta_f}{\beta_{f0}} = \frac{K_{f0}}{K_f} - 1 \tag{3}$$

where the subscript 0 denotes a reference baseline state. The variation in the fuel compressibility over T_f is estimated based on these relations and shown in Fig. 3, where the "Change from Normal" represents the variation of those indicated fuel properties versus T_f from their normal values at a baseline T_f. The trend of the fuel acoustic velocity varying versus T_f is estimated and plotted in Fig. 4. As T_f increases, the modulus of elasticity decreases, and thereby, the acoustic velocity becomes lower and the compressibility is higher.

3 Modeling Main Effects on Fuel Injection

As the inlet fuel temperature varies, the effects on the engine fuel injection are mainly due to the changes in the fuel acoustic velocity, the fuel compressibility, which is directly related to the modulus of elasticity, and the fuel density. Then, modeling the major fuel-injection parameters affected by fuel temperature is considered.

Start-of-injector-needle-lift timing. The start of fuel-injection timing at the fuel injector-nozzle end in its corresponding crank angle (CA) position, relative to the actuation timing at the injection-pump end, mainly depends on the following two hydraulic lag factors: the time interval required for fuel pressure of the pump-to-injector line to be built up to reach level (p_i) for starting to lift the injector-nozzle needle, and the hydraulic delay due to

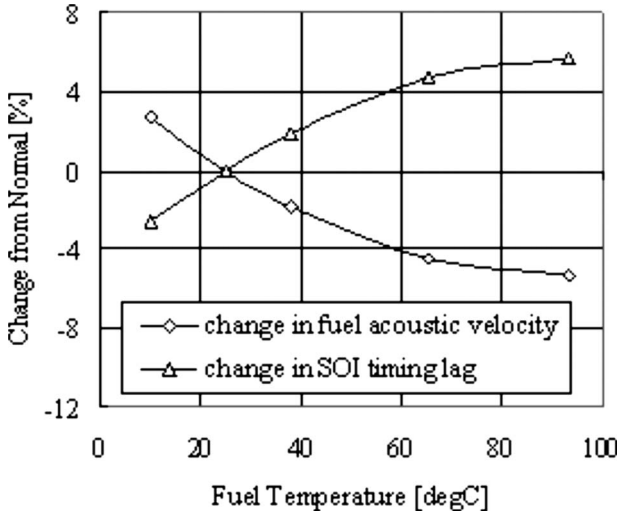


Fig. 4 Predicted fuel temperature effects on fuel acoustic velocity and SOI-timing lag (p_{j0}=100 MPa)

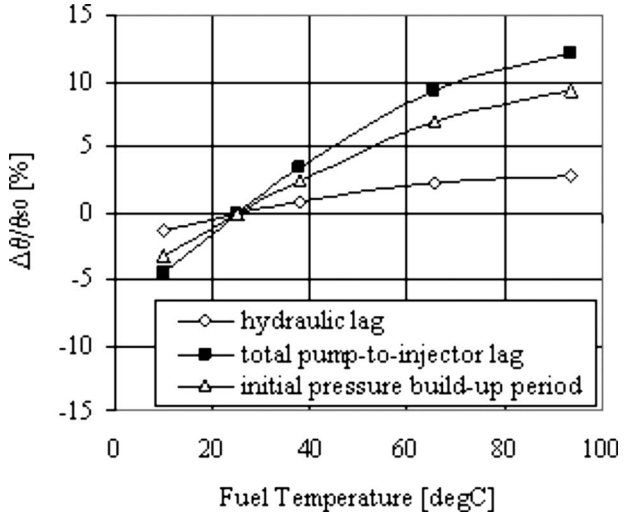


Fig. 5 Predicted fuel temperature effect on start of injection timing (p_{j0}=100 MPa, $\theta_{i0}/\theta_{s0}=\theta_{j0}/\theta_{s0}=0.5$)

the pressure wave propagating from the pump end to the injector end.

For the initial pressure build-up period (θ_{si}) in the degree of crank angle, its variation from a baseline state (θ_{si0}), due to varying the properties of fuel, can be evaluated using the relevant equation from the Appendix as

$$\frac{\Delta\theta_{si}}{\theta_{si0}} = \frac{K_{f0}}{K_f} - 1 \tag{4}$$

Thus according to Eq. (4), a decreased K_f caused by increasing fuel temperature T_f will make the initial pressure build-up interval longer.

A reduced acoustic velocity from increasing fuel temperature causes the hydraulic lag between the fuel-injection-pump end and the injector-nozzle end to be longer, changing both the start-of-injector-needle-lift timing and the end-of-lift timing at the injector nozzle, though the start-of-actuation timing and end-of-actuation timing at the pump end remain unchanged. The variation in this part ($\Delta\theta_{sj}$) of pump-to-injector hydraulic lag compared with the corresponding baseline lag (θ_{sj0}) can be estimated by a simplified relation, as derived in the Appendix and given here as

$$\frac{\Delta\theta_{sj}}{\theta_{sj0}} = \sqrt{\frac{K_{f0}}{K_f}\frac{\rho_f}{\rho_{f0}}} - 1 \tag{5}$$

The effect of T_f on θ_{sj} based on this prediction is plotted in Fig. 4. As T_f increases, the hydraulic lag increases, and thus the injector start-of-injection (SOI) timing becomes more retarded, relative to the baseline pump actuation timing, which is not changed.

Then, the total variation ($\Delta\theta_s$) of the timing lag (θ_s) of pump-to-injector actuation combines the two factors above and can be written as

$$\frac{\Delta\theta_s}{\theta_{s0}} = \frac{\theta_{i0}}{\theta_{s0}}\left(\frac{K_{f0}}{K_f} - 1\right) + \frac{\theta_{j0}}{\theta_{s0}}\left(\sqrt{\frac{K_{f0}}{K_f}\frac{\rho_f}{\rho_{f0}}} - 1\right) \tag{6}$$

where it is needed that $\theta_{i0}/\theta_{s0}+\theta_{j0}/\theta_{s0}=1$. Using Eq. (6) with values of the timing fractions θ_{i0}/θ_{s0} and θ_{j0}/θ_{s0} identified for a particular engine fuel-injection system, the total variation due to varying fuel temperature can be determined. Figure 5 illustrates the results of $\Delta\theta_{si}/\theta_{si}$, $\Delta\theta_{sj}/\theta_{sj}$, and $\Delta\theta_s/\theta_{s0}$ versus T_f in an example that θ_{i0} and θ_{j0} are equally distributed.

Fuel injection line pressure. A higher fuel compressibility due to a smaller modulus of elasticity due to increasing T_f would result in a lowered fuel-injection line pressure. The variation in the average pressure can be estimated based on K_f and ρ_f, which is shown in more detail in the Appendix, to yield

$$\frac{\Delta p_j}{p_{j0}} = \left(\frac{K_f}{K_{f0}}\right)^n - 1 \tag{7}$$

where $n \leq 1$ is needed to correct the effect of K_f to a practical fuel-injection case in which fuel is compressed not solely in a closed volume. Correlating with general experimental measurements, $n=0.5$ is reasonably used in the study of this paper. Then, the estimated change in the averaged injection line pressure versus fuel temperature can thus be graphically illustrated in Fig. 6. It is shown that as T_f increases, the injection pressure is lower.

Fuel injection rate. The lowered injection pressure due to increasing T_f reduces the rate of fuel injection (m_{fr}, kg/s, or kg/°CA) into the cylinder. The variation of the averaged fuel-injection rate is related to K_f and ρ_f by the following form, as derived and described in more details in the Appendix.

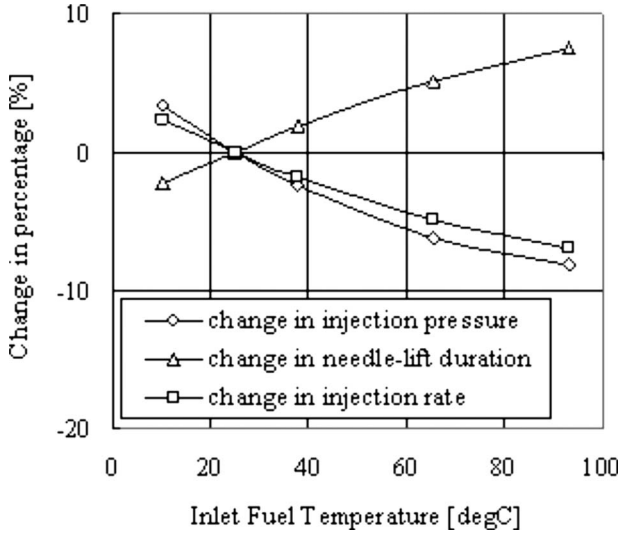


Fig. 6 Predicted changes in line pressure, injection rate, and injector needle-lift duration versus inlet fuel temperature (p_{j0} =100 MPa)

Table 1 Predicted relative changes in fuel injection parameters ($\Delta T_f = +20°C$)

Parameter	$\Delta K_f/K_{f0}$	$\Delta\rho_f/\rho_{f0}$	$\Delta\beta_f/\beta_{f0}$	$\Delta u_p/u_{p0}$	$\Delta p_j/p_j$	$\Delta m_f/m_{fr0}$	$\Delta\theta_s/\theta_{s0}$	$\Delta\theta_d/\theta_{d0}$
Variation	−8%	−2%	+9%	−3%	−3.5%	−3%	+5%	+2.7%

$$\frac{\Delta\dot{m}_{fr}}{m_{fr0}} = \sqrt{\left(\frac{K_f}{K_{f0}}\right)^n \frac{\rho_f}{\rho_{f0}}} - 1 \tag{8}$$

The estimated change in the average injection rate versus fuel temperature is presented in Fig. 6. It is shown that as T_f increases, the fuel-injection rate decreases.

Injector needle-lift duration. As T_f changes, the fuel-injection duration will be affected by the fuel-injection rate and the amount of fuel (m_f in mass) injected per injection, which may change as well over T_f. Using the injector needle-lift duration (θ_d in °CA) to represent the injection duration, the effect of T_f on θ_d is derived in the Appendix as

$$\frac{\Delta\theta_d}{\theta_{d0}} = \frac{m_f}{m_{f0}}\sqrt{\left(\frac{K_{f0}}{K_f}\right)^n \frac{\rho_{f0}}{\rho_f}} - 1 \tag{9}$$

The reduced injection rate, due to increasing T_f, prolongs the fuel-injection duration for an unchanged amount of fuel injected per injection (m_f). This implies a combined effect of K_f and ρ_f. As T_f increases, the lowered injection pressure due to the lowered K_f (Eq. (7)) reduces the volumetric injection rate, and the decrease in the fuel density causes the volume of an unchanged amount of fuel injection to increase, both of which prolong the fuel-injection duration. All of this will result in a retardation in the end-of-injection timing, and thereby an increase in fuel late burning and a decrease in the engine fuel efficiency, and likely an increase in m_f for the engine operating at the same speed and shaft load. If the change in m_f is negligible, the variation in θ_d through K_f and ρ_f becomes

$$\frac{\Delta\theta_d}{\theta_{d0}} = \sqrt{\left(\frac{K_{f0}}{K_f}\right)^n \frac{\rho_{f0}}{\rho_f}} - 1 \tag{10}$$

Based on this relation, the variation trend in θ_d due to changing T_f, without taking into account the consequent change in fuel consumption, is calculated and is illustrated in Fig. 6.

Influence of fuel viscosity on nozzle fuel flow. The viscosity of a liquid fuel varies as T_f changes. Consider that the main effect of the fuel viscosity on the fuel-injection system is on the fuel flow through the injector-nozzle hole of a length L_n and a diameter D_n. The flow velocity u_j at the nozzle exit and the pressure drop Δp_v due to viscous and additional losses through the nozzle hole can be roughly estimated through

$$u_j = C_D\sqrt{\frac{2\Delta p_n}{\rho_f}} \quad \text{m/s} \tag{11}$$

and

$$\Delta p_v = \left(f_v\frac{L_n}{D_n} + \Sigma K\right)\frac{\rho_f u_j^2}{2} \quad \text{MPa} \tag{12}$$

where C_D is the nozzle discharge coefficient accounting for the nozzle loss, Δp_n is the pressure drop across the nozzle, f_v is the flow friction coefficient, and ΣK is the sum of the nozzle entrance and exit loss coefficients. For a typical medium-speed engine fuel-injection system, the nozzle flow velocity is usually in the order of 10^2 m/s and the nozzle-hole diameter is in the order of 10^{-1} mm. The nozzle flow of a common diesel fuel, with the Reynolds number (Re) sufficiently high (see Appendix), is usually in a turbulent region. Thus, the coefficient f_v and C_D remain slightly changed, and it is reasonable to assume the fuel viscosity effect on the nozzle flow due to a change in fuel temperature is less significant, compared with those from K_f and ρ_f.

Fuel atomization regarding D_{SM}. With the injector-nozzle geometry and engine operating condition being fixed while changing inlet fuel temperature, the fuel spray atomization and droplet size are mainly affected by the pressure drop across the nozzle, which is directly related to the fuel modulus of elasticity and influenced by the viscosity and surface tension of fuel as well. Due to the relatively high injection pressure of present-day medium-speed diesel engine, a change in the Sauter mean diameter (D_{SM}) of the droplet size distribution versus the injection pressure may become insignificant, as seen from the results presented by Heywood [6] and by Hiroyasu and Kadota [7].

Using their correlations with the injection pressure-fuel temperature relation derived in the present paper, the estimated change in D_{SM} would be a decrease of about 1% as fuel temperature is decreased by 25°C. The combined effect of the fuel viscosity and surface tension on D_{SM} is more significant, as described by Heywood [6], Hiroyasu et al., [8] and Gambhir et al. [9]. Using the results from Refs. [5,8,9], it is estimated that an increase of 25°C in fuel temperature would result in a decrease in D_{SM} by about 20%. The change is largely due to that on the fuel viscosity.

Nevertheless, when fuel temperature changes, the change in the injection rate and duration, and in turn, the rate of in-cylinder combustion and heat release from varying the injection pressure is more significant in affecting the in-cylinder combustion heat release. This was experimentally demonstrated by Hsu [1] through an in-cylinder combustion and heat release analysis. Therefore, the fuel viscosity effects on fuel injection and combustion, regarding changes in fuel spray atomization due to varying fuel temperature, are also less significant as the nozzle flow is in the turbulent region. If the nozzle flow is not in the turbulent region, the effect of viscosity would become more significant, relative to those due to K_f and ρ_f.

Predicted effects of inlet fuel temperature. For the purpose of illustration of using the developed analytical models of this work, Table 1 predicts changes in the relevant injection parameters along with an increase of 20°C in the engine inlet fuel temperature that is initially at 30°C. Some of the results are then compared with those available and comparable parameters measured by Hsu [1] from an engine experiment. It is shown from this preliminary comparison that the model-predicted variations in the injection pressure, the start-of-needle-lift timing, and the needle-lift duration agree with the engine experimental results in the trends. For instance, the model estimated a decrease of 3.5% in the injection pressure and an increase of 2.7% in the needle-lift duration, while the measurements showed a decrease of 2.9% in the pressure and an increase of 2.6% in the duration, based on an equivalence of a 20°C increase in fuel temperature.

Consider also a hypothetical direct-injection diesel engine operating at 1000 rpm has a nominal injection pressure of 100 MPa, a nominal injection duration of 30°CA at an initial level of T_f, and 0.8 m of the pump-nozzle fuel line length. It is estimated using the developed model that increasing T_f from 25°C to 50°C would approximately result in the injector SOI timing more retarded by about 0.3°CA, and an increase of 1.1°CA in the injection duration.

4 Influences on Combustion and Engine Output Performance and Emissions

The engine overall output performance and exhaust emissions are essentially related to the in-cylinder combustion processes.

Table 2 Predicted trends in engine outputs and emissions versus inlet fuel temperature (medium-speed diesel engine, full load)

	m_a	T_{max}	P_{max}	T_e	m_f	$D\text{NO}_x$	$B\text{NO}_x$	BCO	PM	Smoke
$T_f\uparrow$[a]	$\uparrow$[a]	$\downarrow$[b]	$\downarrow$[b]	$\uparrow$[a]	$\uparrow$[a]	$\downarrow$[b]	$\downarrow$[b]	$\uparrow$[a]	$\uparrow$[a]	$\uparrow$[a]

[a]$\uparrow$: increase.
[b]$\downarrow$: decrease.

The in-cylinder global combustion parameters under consideration in the investigation included peak combustion temperature T_{max}, peak firing pressure p_{max}, the amount of air charging or air-to-fuel ratio a/f, cylinder exhaust temperature T_e, and the rate of fuel-combustion heat release. Consider and understand that the changes in fuel-injection properties due to varying inlet fuel temperature affect the in-cylinder combustion and the engine output performance and emissions. Then, predicting general trends of the effects of the changes in the fuel-injection properties on the combustion, output performance, and emission parameters was achieved in the analytical study based on the understanding and consideration.

With the effects of inlet fuel temperature T_f on the fuel-injection parameters, studied and discussed in Sec. 3, a change in T_f would cause all of the following fuel-injection parameters to vary: the nozzle fuel-injection-start timing, the injection pressure, the injection rate, and the injection needle-lift duration. In the case that T_f varies with all of the other engine input parameters including the injection-pump actuation timing remaining constant, the actual position/timing of the combustion duration will change to be either retarded or advanced, depending on the direction of T_f's variation. Thus, the general trends of changes in those in-cylinder combustion, output performance, and emission parameters due to varying T_f can be predicted simply from the changes in the injection and combustion properties. The predicted variations in the in-cylinder combustion parameters and engine outputs and emissions are summarized in Table 2, for a full-load operation of a general turbocharged medium-speed diesel engine.

The method and results of predicting the trends shown in Table 2 of the engine output performance and emission parameters are further discussed herewith for medium-speed diesel engines based on the understanding and consideration achieved and described earlier in this paper. As T_f increases, for instance, the delayed nozzle fuel-injection-start timing, the reduced injection rate, and the prolonged injection needle-lift duration together will result in an increase in late burning combustion, which will produce the trends in the variation in the outputs and emissions shown in Table 2. Both the peak cylinder temperature T_{max} and peak cylinder pressure p_{max} will decrease as T_f increases, while the cylinder exhaust temperature T_e will be higher. In general, an increase in T_f will be consequently accompanied with a decrease in both the NO_x volume concentration $D\text{NO}_x$, and the brake specific $B\text{NO}_x$. Due to the increased late burning, which causes more incomplete combustion of the fuel-air mixture toward the exhaust valve opening, all of the other exhaust emissions including carbon monoxide (CO), unburned hydrocarbon (HC), exhaust particulate emissions (PMs), and smoke will become higher. The increase in the engine cylinder air charge rate m_a in the full-load condition is due to the increased late burning, which increases the cylinder exhaust thermal energy driving the turbocharger.

The variation in the engine fuel efficiency depends on the nominal fuel-injection timing position. For a high-power medium-speed engine operating at full load, the start-of-injection timing is usually set for the peak firing pressure close to but not exceeding a maximum allowable structural limit, and increasing T_f may adversely affect the fuel efficiency. Under partial load conditions where if the injection timing is set to achieve the optimal indicated fuel efficiency at a nominal T_f condition, a change in T_f to increase or decrease both may also adversely affect the fuel efficiency.

5 Experimental Study

The investigation in this paper includes results of an engine experiment for measuring and verifying the fuel temperature effects on in-cylinder combustion and engine outputs and emissions as well. The experimental investigation was conducted by testing a GE-7FDL locomotive diesel engine. With the specifications of its series shown in more detail by Chen et al. [10], the engine is a turbocharged 16-cylinder four-stroke direct-injection diesel engine, generates 3356 kW of nominal rated power at the engine rated speed of 1050 rpm.

The engine fuel efficiency was determined in this test by measuring the amount of fuel consumed by the engine during a test interval. The exhaust emissions were measured by the equipment as follows: NO_x by a chemiluminescence analyzer, CO by a nondispersive infrared analyzer, HC by a heated flame ionization analyzer, PM by a minivariable split particulate tunnel, and smoke by a full-flow opacity meter.

As fuel-inlet temperature changed while testing the engine, all of the other inlet conditions and operating input parameters remained constant. The results of the in-cylinder combustion global parameters of the engine operating under full-load condition at two different inlet fuel temperatures were obtained and were plotted in Fig. 7, showing the trends of variations. Those parameters included the amount of air charge, engine averaged peak firing pressure, and the engine fuel efficiency. The variation trends of those parameters affected by varying the inlet fuel temperature indicate that, as T_f increased, the cylinder peak firing pressure decreased and the fuel efficiency dropped, indicating an increase in late burning at the full-load condition. The variation in a/f ratio was negligibly small in this case, and the amount of air charge

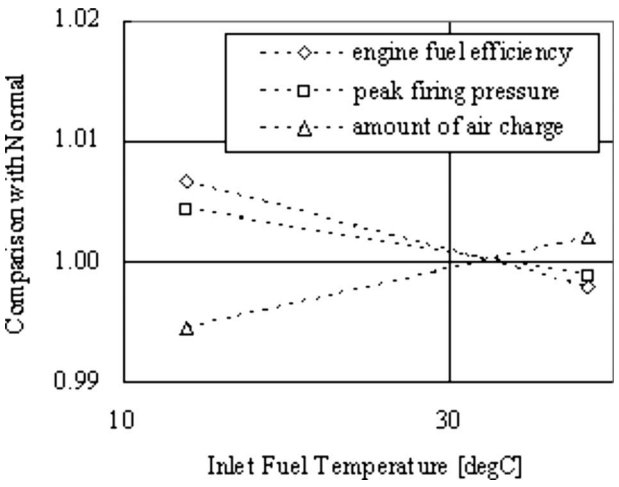


Fig. 7 Variation of engine outputs versus inlet fuel temperature (full load)

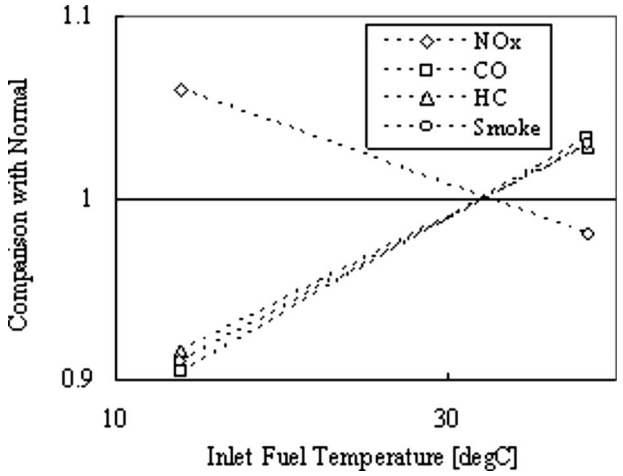


Fig. 8 Engine emissions versus inlet fuel temperature (full load)

into the cylinder m_a slightly increased as the inlet fuel temperature became higher. This also indicates the increases in the late combustion and exhaust thermal energy, which drives the turbocharger as T_f increases, as predicted by the preceding analytical study.

The results of the exhaust emissions from the engine operating in full-load condition at these two different levels of fuel temperature were also obtained and are plotted in Fig. 8 for verification. The trends of variation in the exhaust emissions versus the inlet fuel temperature show that, as T_f increased, the NO_x emissions decreased while the rest of the emissions including CO, HC, and smoke increased. This also agrees with the preceding analytical prediction in this paper—the fuel-injection behavior is affected by fuel-inlet temperature—the consequent changes in fuel injection due to varying inlet fuel temperatures affect the in-cylinder combustion and the engine output performance and emissions.

6 Conclusions

For high-power medium-speed direct-injection engines, the effects of inlet fuel temperature are noticeable and relatively significant. An appropriate understanding and determination of the fuel temperature effects are needed. The effects of inlet fuel temperature on fuel injection, combustion, and engine output performances and emissions of the medium-speed diesel engines have been analytically and experimentally investigated in this paper. The investigation results indicate the following.

(1) As the inlet fuel temperature changes, the main fuel properties varying and affecting fuel-injection and in-cylinder combustion are the bulk modulus of elasticity and the density of liquid fuel, relative to the fuel viscosity, which is less significant as the nozzle flow is generally in a turbulent region. Primarily, the affected fuel-injection parameters include the nozzle fuel start-of-injection timing, the fuel end-of-injection timing, the fuel-injection line pressure and the injection rate, the fuel-injection duration, and fuel spray atomization. The simplified analytical models developed in this paper are able to predict the variations of the injection parameters versus fuel temperature. As inlet fuel temperature increases, it is predicted that the nozzle fuel-injection-start timing is retarded due to an increase in the total hydraulic lag, the fuel-injection rate drops, and the injector needle-lift duration is prolonged from its baseline. The change in the mean spray droplet size is not significant as a function of K_f and ρ_f as opposed to the fuel viscosity μ_f or ν_f. However, the change in the injection rate and duration due to K_f and ρ_f are relatively significant in affecting engine combustion and emissions as the inlet fuel temperature changes and the nozzle flow is in a turbulent region. The model results agree with the available experimental results in predicting trends.

(2) The engine combustion, output performance, and emissions versus the inlet fuel temperature have been analytically studied for the trends of variation in the global in-cylinder combustion parameters including peak cycle temperature, peak cylinder pressure, air-to-fuel ratio, cylinder exhaust gas temperature, and the rate of combustion heat release. The combustion process, output performance, and emissions of a medium-speed diesel engine were also experimentally investigated. The analytical prediction and experimental results show an agreement in the trends and indicate that increasing inlet fuel temperature would result in an increase in CO, HC, PM, and smoke emissions and in a decrease in NO_x. Increasing T_f may adversely affect the fuel efficiency for a general medium-speed diesel engine at a full-load condition.

(3) Reducing the absolute changes in the actual start-of-injection timing and end-of-injection timing caused by the fuel temperature effects is a design goal. The results of this investigation indicate that this goal can be achieved if the distance between the injection pump and injector nozzle of the fuel-injection system is short as long as the volume of fuel contained in the high-pressure line is sufficient for the injection capacity required.

(4) With the enhanced understanding and analyses, the effects of fuel temperature on engine combustion, output performance, and emissions can be taken into account in the engine design and evaluation. Comparable engine performance and emissions along with inlet fuel temperature in the determination of engine capabilities and performances can be obtained. Due to the similar nature of the injection and combustion systems, the understanding and modeling developed can be applied to liquid fuel compression-ignition direct-injection engines in general.

Acknowledgment

The author wishes to thank General Electric Co. for permission to use relevant engine experimental results in this paper.

Nomenclature

a/f = air-to-fuel mass ratio
BCO = brake exhaust carbon monoxide, g/h
$B\mathrm{NO_x}$ = brake exhaust nitrogen oxide, g/h
C_D = injector-nozzle flow discharge coefficient
D_n = injector-nozzle-hole diameter, m
D_{SM} = Sauter mean diameter of spray droplets, ϵm
$D\mathrm{NO_x}$ = volumetric concentration of NO_x, ppm
f_v = viscous friction coefficient
K_f = bulk modulus of elasticity of liquid fuel, MPa
L_n = injector-nozzle hole length, m
L_p = fuel system pump-to-nozzle distance, m
m_a = amount of air charge into cylinder, kg
m_f = fuel injected per injection, kg
m_{fr} = fuel-injection rate, kg/°CA
$\mathrm{NO_x}$ = exhaust nitrogen oxide
p_i = line pressure to start lifting the injector, MPa
p_j = fuel-injection line pressure, MPa
$p_{\max}$ = cylinder peak firing pressure, MPa
p_f = pressure of fuel, MPa
PM = exhaust particulate emissions, g/h
RPM = engine operation speed, rpm
T_f = engine inlet fuel temperature, °C
$T_{\max}$ = cylinder peak combustion temperature, °C
u_j = fuel-injection velocity at nozzle exit, m/s
u_p = acoustic velocity within liquid fuel, m/s
β_f = liquid fuel compressibility, 1/MPa

θ_d = fuel-injection needle-lift duration, °CA
θ_s = total lag of injector actuation relative to that of the injection pump $\theta_{si}+\theta_{sj}$, °CA
θ_{si} = lag of injector actuation relative to that of the injection pump due to pressure wave propagation °CA
θ_{sj} = lag of injector actuation relative to that of the injection pump due to initial pressure build-up °CA
ρ_f = density of liquid fuel, kg/m^3
μ_f = viscosity of liquid fuel, $N\,s/m^2$
Δ = change from a normal reference state
Δp_n = pressure drop across injector nozzle, MPa
Δp_{ne} = effective pressure drop across injector nozzle, MPa
Δp_v = nozzle-loss pressure drop, MPa

Appendix

The analyses and modeling results for predicting the variations of the fuel-injection parameters in studying the fuel temperature effects in this paper are described in more detail as follows.

Variation in liquid fuel pressure-wave propagation velocity. The acoustic velocity of a liquid fuel can be determined by the following equation:

$$u_p = \sqrt{\frac{dp_f}{d\rho_f}} = \sqrt{-\frac{1}{\rho_f^2}\frac{dp_f}{d\frac{1}{\rho_f}}} = \sqrt{\frac{K_f}{\rho_f}} \quad \text{m/s} \tag{A1}$$

Therefore, a relative change in the acoustic velocity due to changing K_f and ρ_f as a result of changing T_f is derived as

$$\frac{\Delta u_p}{u_{p0}} = \frac{\sqrt{\frac{K_f}{\rho_f}} - \sqrt{\frac{K_{f0}}{\rho_{f0}}}}{\sqrt{\frac{K_{f0}}{\rho_{f0}}}} = \sqrt{\frac{K_f}{K_{f0}}\frac{\rho_{f0}}{\rho_f}} - 1 \tag{A2}$$

Variation in liquid fuel compressibility. Since the compressibility of a liquid fuel is related to the fuel bulk modulus of elasticity by:

$$\beta_f \sim \frac{1}{K_f} \quad 1/\text{MPa} \tag{A3}$$

A relative change in the compressibility versus K_f can be written as

$$\frac{\Delta\beta_f}{\beta_{f0}} = \frac{\frac{1}{K_f} - \frac{1}{K_{f0}}}{\frac{1}{K_{f0}}} = \frac{K_{f0}}{K_f} - 1 \tag{A4}$$

Variation in pump-to-injector actuation lag. The hydraulic lag between the actuation of the fuel-injection pump and the start-of-injection of the injector nozzle in a direct-injection engine can be calculated by

$$\theta_s = \frac{L_p}{u_p}\frac{360 \times \text{RPM}}{60} = \frac{6 \times L_p \times \text{RPM}}{\sqrt{\frac{K_f}{\rho_f}}} \quad ^\circ\text{CA} \tag{A5}$$

where L_p is the fuel system pump-to-nozzle length and RPM is the speed of engine operation in rpm. Then, a relative change in the lag of the pump actuation to the start-of-injection at the injector-nozzle end θ_s is derived as

$$\frac{\Delta\theta_s}{\theta_{s0}} = \frac{\sqrt{\frac{\rho_f}{K_f}} - \sqrt{\frac{\rho_{f0}}{K_{f0}}}}{\sqrt{\frac{\rho_{f0}}{K_{f0}}}} = \sqrt{\frac{K_{f0}}{K_f}\frac{\rho_f}{\rho_{f0}}} - 1 \tag{A6}$$

Variation in interval for injector-nozzle-actuation pressure build-up. In general, a pressure (p_i) at a preset level is needed to be built up in the pump-to-injector line to open the injector nozzle for injection. The line pressure during this initial pressure build-up period depends on the fuel modulus of elasticity (K_f), the nominal line total volume (V_{f0}), and the pump-plunger displacement volume (ΔV_f) as $p_i = K_f \Delta V_f / V_{f0}$. Since ΔV_f is a function of the pump displacement in cam or crank angle (θ_i) as $\Delta V_f = U_i \theta_i$, where U_i is the plunger displacement per unit degree of cam or crank angle and can simply be assumed a constant in that initial period, then θ_i, the initial pressure build-up interval, can be determined from

$$\theta_i = \frac{p_i V_{f0}}{K_f U_i} \tag{A7}$$

The variation in θ_i from a baseline θ_{i0} of a fuel-injection system with the fuel properties changeable now yields

$$\frac{\Delta\theta_i}{\theta_i} = \frac{K_{f0}}{K_f} - 1 \tag{A8}$$

Variation of average fuel-injection pressure. The fuel-injection line pressure p_j is a function of the fuel modulus of elasticity K_f or of the fuel compressibility β_f. For an engine fuel-injection system with a fuel compressed volume (ΔV_f) due to the pump plunger displacement in the high-pressure-fuel space between the pump and the injector nozzle, p_j can be expressed by the following relation:

$$p_j \sim \left(K_f \frac{\Delta V_f}{V_{f0}}\right) \quad \text{MPa} \tag{A9}$$

A relative change in p_j due to changing K_f as a result of changing T_f is then resulted from

$$\frac{\Delta p_j}{p_{j0}} = \left(\frac{K_f}{K_{f0}}\right)^n - 1 \tag{A10}$$

where $n \leq 1$ needs to be used to correct the effect of K_f shown in Eq. (A4) to a practical fuel-injection case in which fuel is compressed not solely in a closed volume.

Injection flow velocity and variation of pressure drop across the nozzle. The average fuel discharge velocity at the nozzle exit can be estimated with the nozzle discharge coefficient C_D accounting for the nozzle flow loss by

$$u_j = C_D \sqrt{\frac{2\Delta p_n}{\rho_f}} \quad \text{m/s} \tag{A11}$$

where $\Delta p_n = p_j - p_c$ is the pressure drop across the injector nozzle between the injection pressure p_j and the cylinder content pressure p_c.

The effective pressure drop Δp_{ne} here can be used to estimate the pressure drop accounting for accelerating the fuel flow across the injector nozzle. The Δp_{ne} is the pressure drop obtained by subtracting the nozzle viscous friction pressure drop Δp_v from the total pressure drop Δp_n across the nozzle. Then the Δp_v, and the relationship between the total and the effective pressure drop can be written, respectively, as

$$\Delta p_v = \left(f_v \frac{L_n}{D_n} + \sum K\right)\frac{\rho_f u_j^2}{2} \quad \text{MPa}$$

and

$$\Delta p_n = \Delta p_{ne} + \Delta p_v \approx \frac{1}{2}\rho_f u_j^2 \left(1 + f_v \frac{L_n}{D_n} + \sum K\right) \quad \text{MPa} \tag{A12}$$

The nozzle flow Reynolds number at the injection nozzle exit is

$$R_e = \frac{\rho_f u_j D_n}{\mu_f} = \frac{C_D\sqrt{2\rho_f \Delta p_n} D_n}{\mu_f} = \frac{\sqrt{2\rho_f \Delta p_{ne}} D_n}{\mu_f} \tag{A13}$$

For a typical direct-injection diesel engine fuel-injection system, the fuel flow velocity u_j at the nozzle exit is usually in the order of 10^2 m/s and the nozzle-hole diameter D_n is in the order of 10^{-1} mm, which results in R_e in the order of 10^4. For example, assuming the nozzle effective pressure drop $\Delta p_{ne}=70$ MPa and $D_n=0.3$ mm of a hypothetical direct-injection diesel engine at a fuel temperature of 25 °C would give $\mathrm{Re} \approx 4.0 \times 10^4$. This means the fuel flow at the nozzle exit is in a turbulent region, and thereby the friction coefficient f_v and the discharge coefficient C_D remain about unchanged as T_f varies.

Then, a relative variation in the effective injection pressure drop can be estimated by a derived simplified form as

$$\frac{\Delta(\Delta p_{ne})}{\Delta p_{ne0}} \approx \frac{\Delta(\Delta p_n)}{\Delta p_{n0}} \approx \frac{\rho_f}{\rho_{f0}}\left(\frac{u_j}{u_{j0}}\right)^2 = \left(\frac{K_f}{K_{f0}}\right)^n - 1 \tag{A14}$$

which means both the pressure drop Δp_n and Δp_{ne} are mainly affected by the fuel bulk modulus of elasticity, concerning a fuel temperature change.

Variation in the average fuel-injection rate. Then the averaged fuel-injection rate (in kg/ °CA) is

$$m_{\mathrm{fr}} = \frac{u_j A_n \rho_f}{\dfrac{360 \times \mathrm{RPM}}{60}} \quad \text{kg/ °CA} \tag{A15}$$

where A_n is the nozzle-hole fuel flow area. For the slightly changed flow discharge coefficient in the turbulent flow of the nozzle, described in the preceding analysis, a relative change in the injection rate consequently becomes

$$\frac{\Delta \dot{m}_{\mathrm{fr}}}{m_{\mathrm{fr0}}} = \frac{u_j \rho_f - u_{j0}\rho_{f0}}{u_{j0}\rho_{f0}} = \sqrt{\left(\frac{K_f}{K_{f0}}\right)^n \frac{\rho_f}{\rho_{f0}}} - 1 \tag{A16}$$

Variation in nozzle needle-lift duration. For the amount of fuel mass m_f per injection, the average nozzle needle-lift duration can be calculated from

$$\theta_d = \frac{m_f}{\rho_f A_n u_j}\frac{360 \times \mathrm{RPM}}{60} \quad \text{°CA} \tag{A17}$$

Thus, a relative change in the needle-lift duration can be derived, which gives

$$\frac{\Delta\theta_d}{\theta_{d0}} = \frac{\dfrac{m_f}{\rho_f u_j} - \dfrac{m_{f0}}{\rho_{f0}u_{j0}}}{\dfrac{m_{f0}}{\rho_{f0}u_{j0}}} = \frac{m_f}{m_{f0}}\frac{\rho_{f0}u_{j0}}{\rho_f u_j} - 1 \tag{A18}$$

If the change in m_f is negligible, substituting Eqs. (A11) and (A14) into Eq. (A18) results in

$$\frac{\Delta\theta_d}{\theta_{d0}} = \frac{\sqrt{\dfrac{\rho_f}{K_f}} - \sqrt{\dfrac{\rho_{f0}}{K_{f0}}}}{\sqrt{\dfrac{\rho_{f0}}{K_{f0}}}} = \sqrt{\left(\frac{K_{f0}}{K_f}\right)^n \frac{\rho_{f0}}{\rho_f}} - 1 \tag{A19}$$

By including the change in m_f in evaluating the variation, the equation becomes

$$\frac{\Delta\theta_d}{\theta_{d0}} = \frac{m_f}{m_{f0}}\sqrt{\left(\frac{K_{f0}}{K_f}\right)^n \frac{\rho_{f0}}{\rho_f}} - 1 \tag{A20}$$

References

[1] Hsu, B. D., 2002, *Practical Diesel Engine Combustion Analysis*, Society of Automotive Engineers, Warrendale, PA.

[2] Hsu, B.D., and Hoffman, J.G., 1985, "The Effect of Diesel Fuel Properties on the Combustion of a Medium Speed Diesel Engine," ASME Paper No. 85-DGP-14.

[3] Maeda, Y., Murayama, T., and Fukazawa, S., 1975, "Influence of Fuel Properties on the Injection Characteristics of Fuel Injection System in Diesel Engines," Bull. JSME, **18**(118).

[4] Burman, P. G., and DeLuca, F., 1962, *Fuel Injection and Controls*, Simmons-Boardman, New York.

[5] Schmidt, P. F., 1985, *Fuel Oil Manual*, Industrial, New York.

[6] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.

[7] Hiroyasu, H., and Kadota, T., 1974, "Fuel Droplet Size Distribution in Diesel Combustion Chamber," SAE Trans., **83**.

[8] Hiroyasu, H., Arai, M., and Tabata, M., 1989, "Empirical Equations for the Sauter Mean Diameter of a Diesel Spray," SAE Paper No. 890464.

[9] Gambhir, G. S., Abata, D. L., and Michalek, D. J., 2003, "An Experimental and Computational Investigation on the Pre-Heating of Fuel to Improve Cold Starting of Diesel Engines," ASME Paper No. ICES2003–662.

[10] Chen, G., Flynn, P. L., Gallagher, S. M., and Dillen, E. R., 2003, "Development of the Low-Emission GE-7FDL High-Power Medium-Speed Locomotive Diesel Engine," ASME J. Eng. Gas Turbines Power, **125**, pp. 505–512.

Real-Time Self-Learning Optimization of Diesel Engine Calibration

Andreas A. Malikopoulos
e-mail: amaliko@umich.edu

Dennis N. Assanis
e-mail: assanis@umich.edu

Panos Y. Papalambros
e-mail: pyp@umich.edu

Department of Mechanical Engineering,
University of Michigan,
Ann Arbor, MI 48109

Compression ignition engine technologies have been advanced in the past decade to provide superior fuel economy and high performance. These technologies offer increased opportunities for optimizing engine calibration. Current engine calibration methods rely on deriving static tabular relationships between a set of steady-state operating points and the corresponding values of the controllable variables. While the engine is running, these values are being interpolated for each engine operating point to coordinate optimal performance criteria, e.g., fuel economy, emissions, and acceleration. These methods, however, are not efficient in capturing transient engine operation designated by common driving habits, e.g., stop-and-go driving, rapid acceleration, and braking. An alternative approach was developed recently, which makes the engine an autonomous intelligent system, namely, one capable of learning its optimal calibration for both steady-state and transient operating points in real time. Through this approach, while the engine is running the vehicle, it progressively perceives the driver's driving style and eventually learns to operate in a manner that optimizes specified performance criteria. The major challenge to this approach is problem dimensionality when more than one controllable variable is considered. In this paper, we address this problem by proposing a decentralized learning control scheme. The scheme is evaluated through simulation of a diesel engine model, which learns the values of injection timing and variable geometry turbocharging vane position that optimize fuel economy and pollutant emissions over a segment of the FTP-75 driving cycle. [DOI: 10.1115/1.3019331]

1 Introduction

Advanced compression ignition engine technologies, such as fuel injection systems, variable geometry turbocharging (VGT), exhaust gas recirculation (EGR), and variable valve actuation (VVA), have alleviated the traditional disadvantages of diesel engines, and have facilitated their use in the passenger vehicle market. These technologies provide an increased number of engine controllable variables that can be used for engine calibration to optimize one or more engine performance criteria, e.g., fuel economy, pollutant emissions, and engine acceleration. Current engine calibration methods generate a static tabular relationship between the values of the controllable variables and the corresponding steady-state engine operating points to achieve optimal performance with respect to the specified criteria. This relationship is incorporated into the electronic control unit (ECU) that aims to maintain performance optimality. While the engine is running, values in the tabular relationships are being interpolated to provide the values of the controllable variables for each operating point.

These methods, however, seldom guarantee optimal engine calibration for the entire operating domain, especially during transient operation [1]. The latter often constitutes the largest segment of engine operation compared with the steady-state one. Fuel consumption and emissions during transient operation are extremely complicated and are highly dependent on engine calibration [2,3]. Research efforts in addressing transient operation have focused on simulation-based methods to derive calibration maps for transients of particular driving cycles. However, prespecifying the entire transient engine operation as imposed by different driving cycles and deriving the optimal values of the controllable variables associated with transient operating points are not possible in practice, thus preventing a priori optimal calibration.

To address these issues, an alternative approach was implemented recently, which treats the engine as a controlled stochastic system and the engine operation as a Markov decision process (MDP) [4]. Engine calibration is formulated as a centralized decision-making problem under uncertainty. The predictive optimal stochastic control algorithm (POSCA) [5] was developed, allowing the engine to learn the values of the controllable variables in real time. While the engine is running the vehicle, it progressively perceives the driver's driving style and eventually learns to operate in a manner that optimizes specified performance criteria. Consequently, optimal calibration is achieved for steady-state and transient engine operating points designated by the driver's driving style. The engine's ability to learn its optimum calibration is not limited, however, to a particular driving style. The engine can learn to operate optimally for different drivers if they indicate their identity before starting the vehicle. The engine can then adjust its operation to be optimal for a particular driver based on what it has learned in the past regarding his/her driving style.

A major challenge to this centralized decision-making approach is the increase in the problem's dimensionality when more than one controllable variable (decision maker) is considered. Decentralized decision making requiring limited information is a highly desirable feature in this situation. It is necessary when complete information among decision makers, which is required in centralized decision making, is impractical due to the increase in the problem's dimensionality. Mathematical learning theory has been developed in systems to address the modeling and control aspects of sequential decision making under uncertainty [6–8]. Learning automata have been applied to network routing in which decentralization is attractive and large uncertainties are present [9,10]. The resulting system performance has demonstrated that decentralized learning schemes can be successful, while the problem's dimensionality remains tractable.

Manuscript received March 18, 2008; final manuscript received October 5, 2008; published online December 19, 2008. Review conducted by Christopher J. Rutland. Paper presented at the 2007 Fall Conference of the ASME Internal Combustion Engine Division (ICEF2007), Oct. 14–17, 2007, Charleston, SC.

The problem of decentralized control for a class of large scale interconnected dynamic systems in continuous time domain was studied by Wu [11]. In this offline approach, it is assumed that the considered systems are linear time varying, and the interconnections between each subsystem are unknown. Szer and Charpillet [12] proposed a model-free distributed reinforcement learning (RL) algorithm that utilizes communication to improve learning among the decision makers in a Markov decision process formalism. Scherrer and Charpillet [13] developed a general iterative heuristic approach in which at each decision epoch the focus is on a subgroup of decision makers, and their policies, given the rest of the decision makers, have fixed plans. Beynier and Mouaddib [14] introduced the notion of expected opportunity cost to better assess the influence of a local decision of an agent on the others. An iterative version of the algorithm was implemented to incrementally improve the policies of agents leading to higher quality solutions in some settings. Yagan and Chen-Khong [15] implemented a model-free coordinated reinforcement learning for decentralized optimal control, assuming that each decision maker can partially observe the state condition. This decentralized scheme is suited for partially observable Markov decision processes. Shen et al. [16] developed a decentralized Markov game model to estimate the belief among the decision makers. In the proposed model, the model-free Q-learning algorithm was employed to adjust dynamically the payoff function of each player. Wheeler and Narenda [17] employed a game-theoretic approach and developed a decentralized learning control scheme in finite Markov chains with unknown transition probabilities and costs. In this scheme, the decision makers demonstrate a myopic behavior; namely, they are unaware of the surrounding world. In attempting to improve his/her performance, each decision maker selects a control action, observes the corresponding cost associated with the occupied state, and then updates the action. Although many of these algorithms address the decentralized learning problem, their use of the accumulated data acquired over the learning process is inefficient, and they require a significant amount of experience to achieve acceptable performance. This requirement arises due to the formation of these algorithms in deriving control policies without learning the system models en route. In addition, the requirement of real-time derivation of the values of the engine controllable variables imposes an additional computational burden in implementing such control schemes.

This paper proposes a decentralized learning control scheme that can utilize efficiently the accumulated data acquired from the engine output to achieve acceptable performance in a real-time implementation. The proposed scheme differs from Wheeler and Narenda's in the sense that the decision makers (engine controllable variables) do not demonstrate a myopic behavior explicitly. On the contrary, a random hierarchy among them is assumed, based on which each one observes the control actions of the other.

The remainder of this paper is organized as follows: In Sec. 2, the mathematical framework of modeling the engine operation as a MDP is reviewed. The decentralized learning control scheme is introduced in Sec. 3. The effectiveness of the method is demonstrated in Sec. 4 through simulation of a diesel engine calibration with respect to the injection timing and VGT vane position over a segment of the FTP-75 driving cycle. Conclusions are drawn in Sec. 5.

2 Modeling Engine Operation as a Markov Decision Process

In implementing self-learning optimization for engine calibration in real time, the engine is treated as a controlled stochastic system, and engine operation is modeled as a MDP. The engine performance criteria, e.g., fuel economy, emissions, and engine acceleration performance, are considered controlled random functions. The objective is to select the optimal control policy (optimum values of the controllable variables) for the sequences of engine operating point transitions, corresponding to the driver's driving style, that optimize one or more engine performance criteria (random functions). The problem of engine calibration is thus formulated as a centralized sequential decision-making problem under uncertainty.

The MDP [18] provides the mathematical framework for modeling such problems [19]. It comprises a decision maker (controller), states (engine operating points), actions (controllable variables), the transition probability matrix (driver), the transition reward matrix (engine performance indices), and optimization criteria (e.g., maximizing fuel economy, minimizing pollutant emissions, and maximizing engine acceleration). In this framework, the controller (decision maker) is faced with the problem of influencing engine operation over time by selecting optimal actions.

Following the exposition in Ref. [5], a discrete-time stochastic controlled MDP is defined as the tuple,

$$s_k = \{\boldsymbol{S}, \boldsymbol{A}, \mathbf{P}(\cdot,\cdot), \mathbf{R}(\cdot,\cdot)\} \tag{1}$$

where $\boldsymbol{S}=\{1,2,\ldots,N\}$, $N \in \mathbb{N}$ denotes a finite state space, $\boldsymbol{A} = \cup_{s_k \in \boldsymbol{S}} A(s_k)$ stands for a finite action space, $\mathbf{P}(\cdot,\cdot)$ is the transition probability matrix, and $\mathbf{R}(\cdot,\cdot)$ is the transition reward matrix. The decision-making process occurs at each of a sequence of decision epochs $k=0,1,2,\ldots,M$, $M \in \mathbb{N}$. At each epoch, the decision maker observes a system's state, $s_k=i \in \boldsymbol{S}$, and executes an action, $a_k \in A(s_k)$, from the feasible set of actions, $A(s_k) \subseteq \boldsymbol{A}$, at this state. At the next epoch, the system transits to the state $s_k = j \in \boldsymbol{S}$ imposed by the conditional probabilities $p(s_{k+1}=j|s_k = i,a_k)$, designated by the transition probability matrix $\mathbf{P}(\cdot,\cdot)$. The conditional probabilities of $\mathbf{P}(\cdot,\cdot)$, $p:\boldsymbol{S}\times\boldsymbol{A}\rightarrow[0,1]$, satisfy the constraint

$$\sum_{j=1}^{N} p(s_{k+1}=j|s_k=i,a_k) = 1 \tag{2}$$

Following this state transition, the decision maker receives a reward associated with the action a_k, $R(s_{k+1}=j|s_k=i,a_k)$, where $R:\boldsymbol{S}\times\boldsymbol{A}\rightarrow\mathbb{R}$, as imposed by the transition reward matrix $\mathbf{R}(\cdot,\cdot)$. The states of a MDP possess the Markov property, stating that the conditional probability distribution of future states of the process depends only on the current state and not on any past states; i.e., it is conditionally independent of the past states (the path of the process) given the present state. Mathematically, the Markov property states that

$$p(s_{k+1}|s_k,s_{k-1},\ldots,s_0) = p(s_{k+1}|s_k) \tag{3}$$

The solution to a MDP can be expressed as a policy $\pi = \{\mu_0,\mu_1,\ldots,\mu_M\}$, which provides the action to be executed for a given state, regardless of prior history; μ_κ is a function mapping states s_k into actions $a_k=\mu_\kappa(s_k)$, such that $\mu_\kappa(s_k) \in A(s_k)$. Such policies are addressed as admissible. Consequently, for any initial state at decision epoch $k=0$, s_0, and for any finite sequence of epochs $k=0,1,2,\ldots,M$, $M \in \mathbb{N}$, the expected accumulated value of the rewards of the decision maker is given by

$$J^{\pi}(s_0) = \underset{\substack{s_k \in \boldsymbol{S} \\ a_k \in A(s_k)}}{E} \left\{ R_M(s_M) + \sum_{k=0}^{M-1} R_k(s_{k+1}=j|s_k=i,a_k) \right\}$$
$$\forall i,j \in \boldsymbol{S}, \quad \forall\, a_k \in A(s_k) \tag{4}$$

where $R_M(s_M)$ is the reward at the final state. In the finite-horizon context, the decision maker should maximize the accumulated value for the next M decision epochs; more precisely, an optimal policy π^* is one that maximizes the overall expected accumulated value of the rewards,

$$J^{\pi^*}(s_0) = \max_{\pi \in \boldsymbol{A}} J^{\pi}(s_0) \tag{5}$$

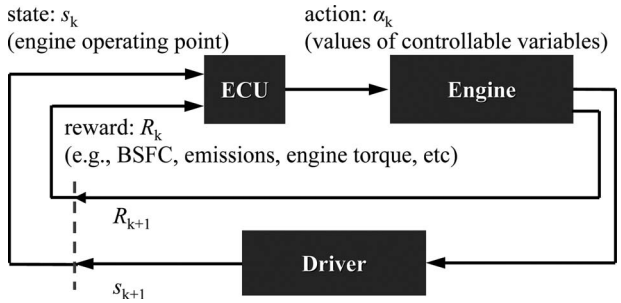


Fig. 1 Learning process during the interaction between the engine and the driver

Consequently, the optimal policy $\pi^*=\{\mu_0^*,\mu_1^*,\ldots,\mu_M^*\}$ sequence is given by

$$\pi^* = \arg\max_{\pi\in A} J^{\pi^*}(s_0) \tag{6}$$

Dynamic programming (DP) has been widely used for many years as the principal method for solving Markov decision problems [20]. However, DP algorithms require the realization of the transition probability matrix, $\mathbf{P}(\cdot,\cdot)$, and the transition reward matrix, $\mathbf{R}(\cdot,\cdot)$. For complex systems, e.g., an internal combustion engine, with large state space, these matrices can be either impractical or impossible to compute. Alternative approaches for solving Markov decision problems have been developed in the field of RL [21,22]. RL has aimed to provide simulation-based algorithms for learning control policies of complex systems, where exact modeling is infeasible or expensive [23]. In this framework, the system interacts with its environment in real time and obtains information, enabling it to improve its future performance by means of rewards associated with the control actions taken. This interaction allows the system to learn in real time the course of action (control policy) that optimizes the rewards. The majority of RL algorithms are founded on dynamic programming. They utilize evaluation functions attempting to successively approximate Eq. (4). These evaluation functions assign to each state the total reward expected to accumulate over time starting from a given state when a policy π is employed. However, in learning engineering systems in which the initial state is not fixed, recursive updates of the evaluation functions to approximate Eq. (4) would demand significant amount of time to achieve the desired system performance [24].

For the engine calibration problem built upon the MDP theoretical framework, the POSCA is employed [5]. POSCA is intended to derive the control policy (values of the engine controllable variables) for any initial state (engine operating point). In applying this algorithm to more than one controllable variable, however, limitations arise due to the requirement for the algorithm to account for all combinations of the controllable variables in a single set of a finite action space $\boldsymbol{A}$. To overcome this problem, the decentralized learning control scheme is implemented.

3 Decentralized Learning Control Scheme

While the engine is running the vehicle and interacting with the driver, the probability of engine operating point transitions designates the elements of the transition probability matrix, $\mathbf{P}(\cdot,\cdot)$. The desired engine performance criteria, e.g., fuel economy and pollutant emissions, are represented by the elements of the transition reward matrix, $\mathbf{R}(\cdot,\cdot)$. Through this interaction, the driver introduces a state $s_k \in \boldsymbol{S}$ (engine operating point) to the engine's ECU, and on that basis the ECU selects an action, $a_k \in A(s_k)$ (combination of values of the controllable variables). As a consequence of its action, the ECU receives a numerical reward, $R_{k+1} \in \mathbb{R}$, and the engine transits to a new state, $s_{k+1} \in \boldsymbol{S}$, as illustrated in Fig. 1. POSCA aims to compute the control policy (values of the controllable variables) for the sequence of engine operating transitions based on the conditional probabilities of the matrix, $\mathbf{P}(\cdot,\cdot)$. During this process, however, when two or more controllable variables are considered, the combinations of their values can grow intractably, resulting in a huge feasible action space $\boldsymbol{A}=\cup_{s_k\in S}A(s_k)$.

The decentralized learning control scheme proposed in this paper establishes a learning process that enables the derivation of the values of the controllable variables to occur in parallel phases [25]. A random hierarchy among them is assumed, based on which each one observes the control actions of the other. In particular, POSCA is employed to derive the control actions of the first member in the hierarchy of decision makers with respect to the sequence of state transitions. At the same time, the algorithm is engaged separately to derive the control actions of the second member in the hierarchy of decision makers with respect to the policy as learned from the first one. Similarly, the algorithm is employed to derive the control actions of the third decision maker with respect to the second one, and so forth.

For instance, in implementing a diesel engine calibration with respect to the injection timing, α, and VGT vane position, β, a feasible set of values, $\boldsymbol{A}$ and $\boldsymbol{B}$, for each controllable variable is defined. The decentralized learning enables the engine to implement two different mappings in parallel. In the first, injection timing is mapped to the states as a result of the correspondence of the driver's driving style to particular engine operating points, i.e., $\boldsymbol{S}\times\boldsymbol{A}\rightarrow\mathbb{R}$. In the second, VGT is mapped to the injection timing, i.e., $\boldsymbol{A}\times\boldsymbol{B}\rightarrow\mathbb{R}$. The learning algorithm utilizes these two mappings to derive the control policies, $\pi_\alpha^* \in \boldsymbol{A}$ and $\pi_\beta^* \in \mathbf{B}$ (values of injection timing and VGT), for the driver's driving style as expressed by the incidence in which particular states or particular sequences of states arise.

The decentralized learning process of the engine transpires at each decision epoch k in conjunction with the injection timing $a_k \in \boldsymbol{A}$ taken for each state $s_k=i \in \boldsymbol{S}$ and the VGT vane position $\beta_k \in \boldsymbol{B}$ for each $a_k \in \boldsymbol{A}$. At the early epochs, and until full exploration of the feasible sets $\boldsymbol{A}$ and $\boldsymbol{B}$ occurs, the mapping from states to probabilities of selecting a particular value of injection timing $a_k \in \boldsymbol{A}$ and the mapping from $a_k \in \boldsymbol{A}$ to probabilities of selecting VGT $\beta_k \in \boldsymbol{B}$ are constant; namely, the values of each controllable variable are selected randomly with the same probability,

$$p(a_k|s_k) = \frac{1}{|\boldsymbol{A}|}, \quad \forall\, a_k \in \boldsymbol{A}, \quad \forall\, s_k \in \boldsymbol{S} \tag{7}$$

and

$$p(\beta_k|a_k) = \frac{1}{|\boldsymbol{B}|}, \quad \forall\, a_k \in \boldsymbol{A}, \quad \forall\, \beta_k \in \boldsymbol{B} \tag{8}$$

Exploration of the entire feasible set for each variable is important to evade suboptimal solutions. POSCA is thus used after the exploration phase to realize the policies π_α^*, and π_β^* by means of the expected values of the rewards, $V(s_{k+1}|s_k,a_k)$ and $V(a_{k+1}|a_k,\beta_k)$, generated by the mappings $\boldsymbol{S}\times\boldsymbol{A}\rightarrow\mathbb{R}$ and $\boldsymbol{A}\times\boldsymbol{B}\rightarrow\mathbb{R}$, respectively. The expected values of the rewards are defined to be

$$\begin{aligned} V(s_{k+1}=j|s_k=i,a_k) :=\ & p(s_{k+1}=j|s_k=i,a_k)\cdot R(s_{k+1}=j|s_k=i,a_k) \\ & + \max_{a_{k+1}\in A}\left[\sum_{l=1}^{N} p(s_{k+2}=l|s_{k+1}=j,a_{k+1})\right. \\ & \left.\cdot R(s_{k+1}=l|s_k=j,a_{k+1})\right], \\ & i,j=1,2,\ldots,N, \quad N=|\boldsymbol{S}| \end{aligned} \tag{9}$$

and

$$V(a_{k+1}=m|a_k=n,\beta_k) := p(a_{k+1}=m|a_k=n,\beta_k)\cdot R(a_{k+1}=m|a_k=n,\beta_k) + \max_{\beta_{k+1}\in \boldsymbol{B}}\left[\sum_{p=1}^{\Lambda} p(a_{k+2}=p|a_{k+1}=m,\beta_{k+1}) \cdot R(a_{k+2}=p|a_{k+1}=m,\beta_{k+1})\right], \quad m,n=1,2,\ldots,\Lambda, \quad \Lambda=|\boldsymbol{A}| \tag{10}$$

In deriving the control policies of the injection timing and VGT in self-learning calibration, which is treated in a stochastic framework, all uncertain quantities are described by probability distributions. The control policies π^*_α, and π^*_β are computed by utilizing the max-min control approach, whereby the worst possible values of the uncertain quantities within the given set are assumed to occur. This essentially ensures that the control policies will result in at least a minimum overall reward value. Consequently, at state $s_k=i$, POSCA computes the control policy π^*_α in terms of the values of injection timing α as

$$\pi^*_\alpha(s_k) = \arg\max_{\bar{\mu}_k(s_k)\in A(s_k)} \min_{s_{k+1}\in \boldsymbol{S}} [V(s_{k+1}=j|s_k=i,a_k)], \quad \forall\, i,j \in \boldsymbol{S} \tag{11}$$

For the control policy π^*_α, POSCA computes the control policy π^*_β of the values of the VGT vane position β as

$$\pi^*_\beta(a_k) = \arg\max_{\beta_k(a_k)\in \boldsymbol{B}(a_k)} \min_{a_{k+1}\in \boldsymbol{A}} [V(a_{k+1}=m|a_k=n,\beta_k)], \quad \forall\, m,n \in \boldsymbol{A} \tag{12}$$

Employing this decentralized learning control scheme, derivation of the values of more than one controllable variable can be achieved, while the problem's dimensionality remains tractable.

4 Real-Time Self-Learning Injection Timing and VGT in a Diesel Engine

The decentralized learning control scheme introduced in the previous section is applied here to a four-cylinder, 1.9 l turbocharged diesel engine. The objective is to find the injection timing and VGT vane position while the engine is running the vehicle that maximize the engine brake torque. Consequently, the controller's inputs are the engine operating points and brake torque, while the outputs are the injection timing and VGT vane position. Injection timing is an important controllable variable in the combustion process and affects performance and emissions [26]. The major objective of injection timing is to initiate fuel injection at the crank angle, resulting in the maximum brake torque (MBT). It designates the ignition delay defined to be the crank angle between the start of injection (SOI) and the start of combustion (SOC). The VGT technology was originally considered to increase engine brake torque at tip-ins and to reduce turbo-lag. VGT has a system of movable guide vanes located on the turbine stator. By adjusting the guide vanes, the exhaust gas energy to the turbocharger can be regulated, and thus the compressor mass airflow and exhaust manifold pressure can be controlled.

The software package ENDYNA THEMOS CRTD by TESIS [27] suitable for real-time simulation of diesel engines is employed. The software utilizes thermodynamic models of the gas path and is well suited for testing and development of ECU. In the example, the existing static correlation involving injection timing and VGT is bypassed to incorporate the decentralized learning control scheme and is used as a baseline comparison. The engine models with the baseline and self-learning calibration are run repeatedly over the same driving style represented by a segment of the FTP-75 driving cycle, illustrated in Fig. 2. Every run over this driving style constitutes one complete simulation. Before initiating the first simulation of the engine model, the elements of the transition probability and reward matrix are assigned the value of zero. That is, the engine at the beginning has no knowledge regarding the particular driving style and the values of the rewards associated with the controllable variables (injection timing and VGT).

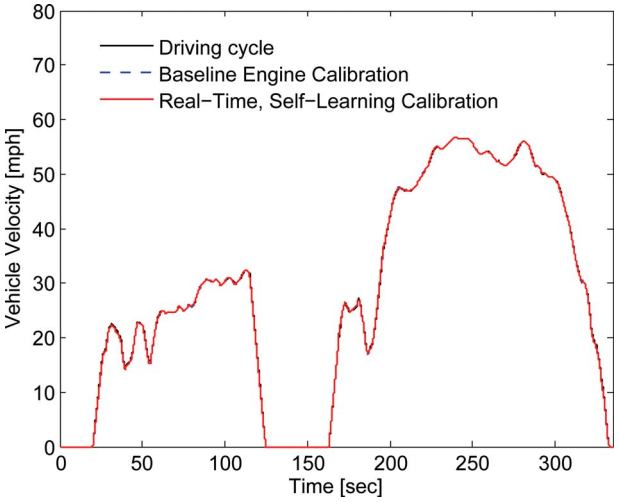


Fig. 2 Segment of the FTP-75 driving cycle

After completing the learning process, the decentralized control scheme specified the values of the injection timing and VGT vane position. The vehicle with the self-learning calibration was able to follow the segment of the driving cycle requiring lower gas-pedal position rates for the same engine speed, as illustrated in Figs. 3–5. The implication is that the derived policy of injection timing and VGT resulted in higher engine torque compared with the baseline calibration. The injection timing (before top dead center (BTDC)) for both vehicles is illustrated in Figs. 6 and 7. While the baseline calibration interpolates values of the injection timing of steady-state operating points, the injection timing derived by the decentralized scheme corresponded to the engine operating point transitions imposed by the driver's driving style, and thus, self-learning calibration was able to capture transient engine operation. Drivability issues that may be raised in such a noisy injection timing response could be addressed by tightening the allowable space of two successive control actions (injection timing values). Lower gas-pedal position rates resulted in reducing the fuel mass injection duration, shown in Fig. 8, and consequently, less fuel mass was injected into the cylinders, as illustrated in Fig. 9 (in

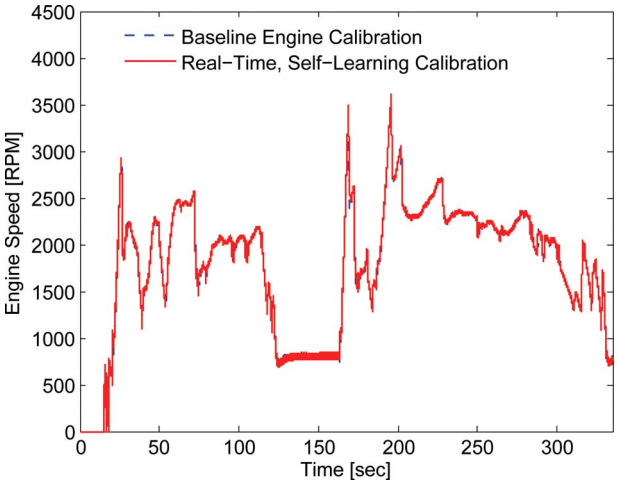


Fig. 3 Engine speed

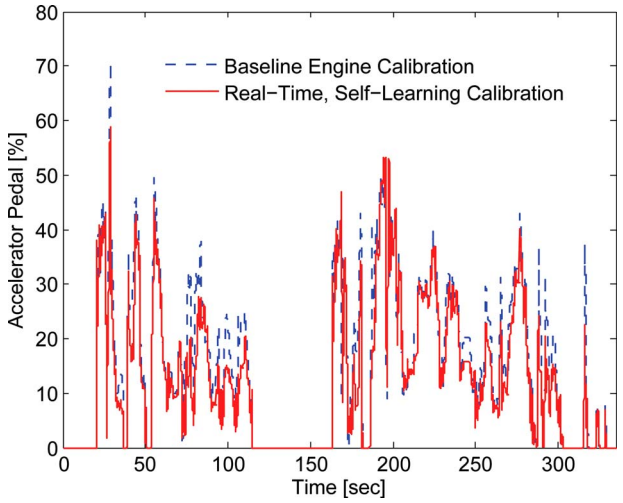


Fig. 4 Gas-pedal position rate representing a driver's driving style

zoom-in for clarity). In the decentralized learning of the engine, the injection timing was mapped to the engine operating points

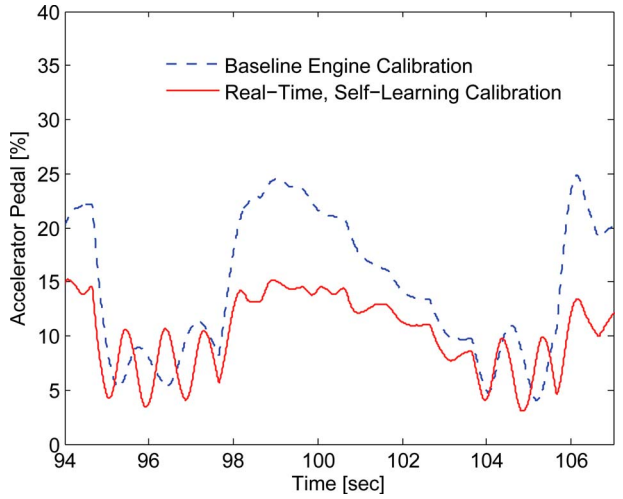


Fig. 5 Gas-pedal position rate representing a driver's driving style (zoom-in)

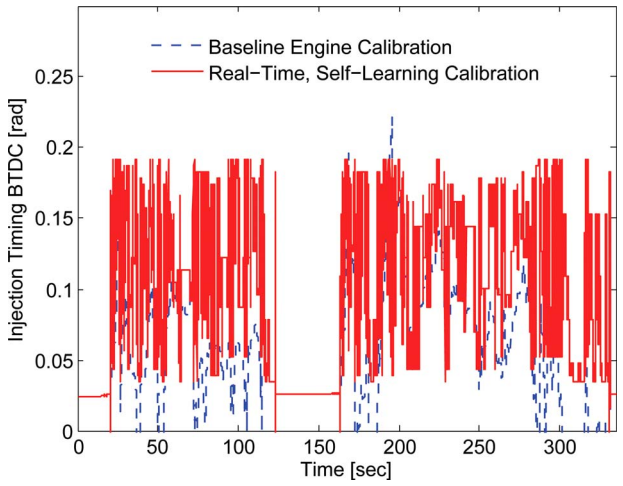


Fig. 6 Injection timing

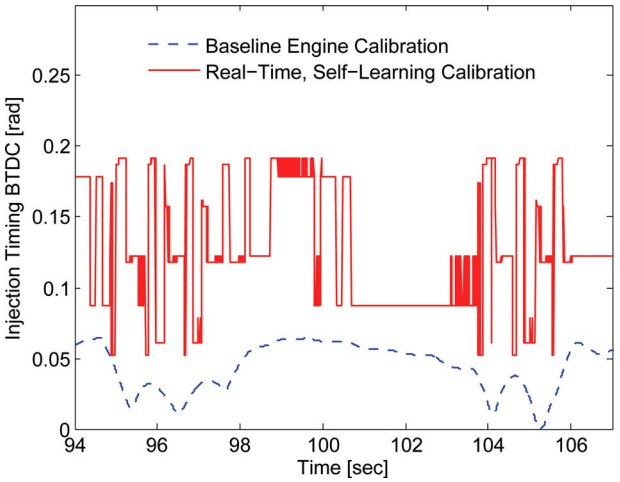


Fig. 7 Injection timing (zoom-in)

(states), while the VGT vane position was mapped to the injection timing. The derived VGT policy is illustrated in Figs. 10 and 11. Having the engine operate at the maximum brake torque, a 9.3% overall improvement of fuel economy was accomplished, as illustrated in Fig. 12, compared with the baseline calibration.

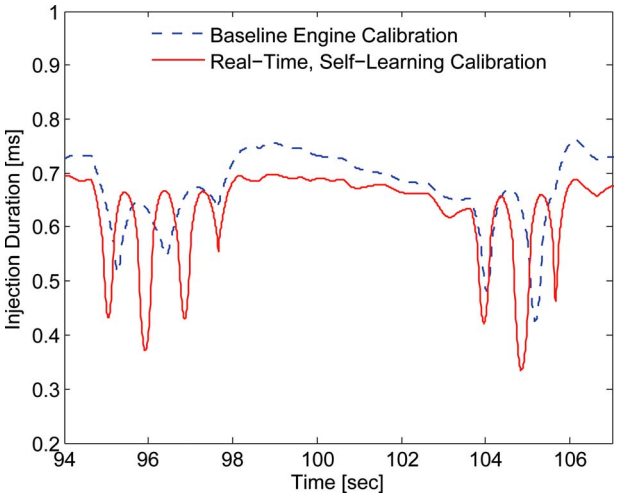


Fig. 8 Fuel mass injection duration (zoom-in)

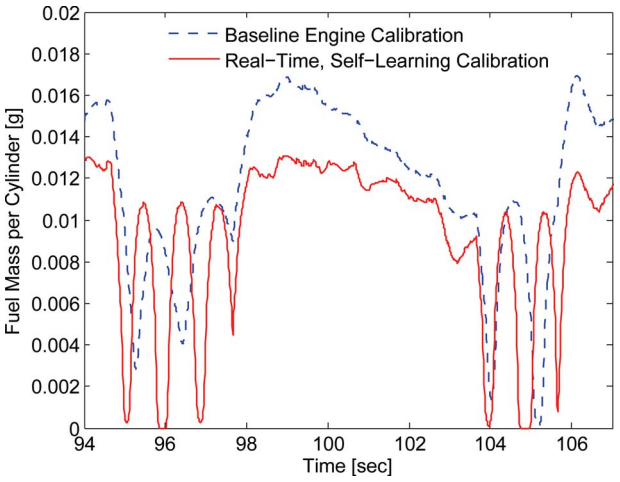


Fig. 9 Fuel mass injected per cylinder (zoom-in)

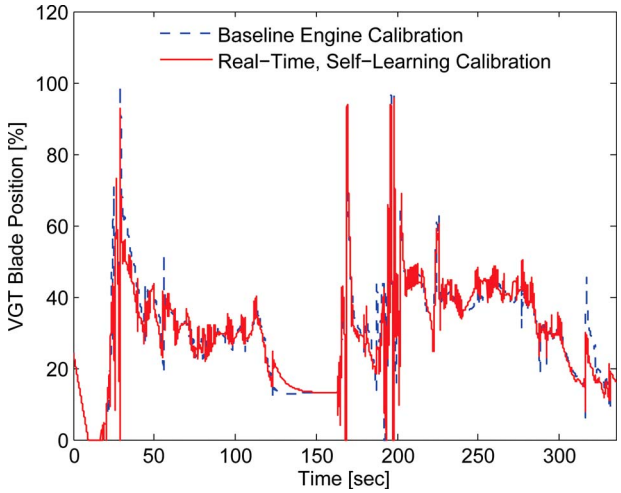


Fig. 10 VGT vane position

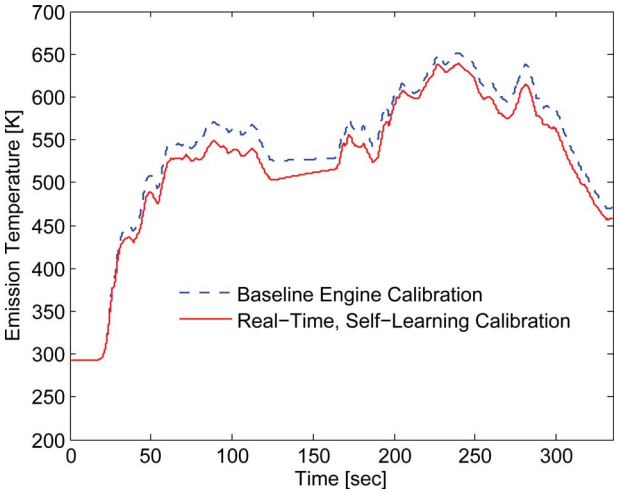


Fig. 13 Emission temperature in the exhaust manifold

At constant engine speed and constant fuel mass per cycle, for a given injection duration, and at fixed brake mean effective pressure (BMEP), if the injection timing is advanced from the MBT, then brake specific fuel consumption (BSFC) is decreased and NO_x emissions are increased. In our case, however, fuel mass per cycle, injection duration, and BMEP all vary, and thus, this behavior is altered. Since the injection duration is decreased, ignition delays decrease as well. Less time is available for fuel-air mixing, resulting in a less intense premixed burn and a lower premixed spike in the heat release rate curve. Cylinder bulk gas temperatures are therefore decreased, and so is NO_x production, as illustrated in Figs. 13 and 14.

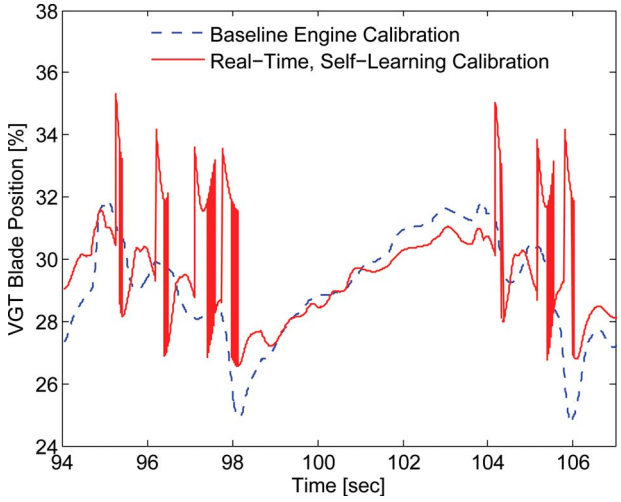


Fig. 11 VGT vane position (zoom-in)

5 Concluding Remarks

This paper presented a decentralized learning control scheme that addresses the problem dimensionality in the centralized decision-making approach as employed in making the engine into an autonomous intelligent system. In this scheme, a learning process is established, which enables the derivation of the values of the controllable variables to occur in parallel phases. The values for more than one controllable variable can thus be determined while keeping the problem's dimensionality tractable. The example presented an application of this scheme to a diesel engine with respect to injection timing and VGT vane position. The engine was able to realize the values of injection timing and VGT for a driving style represented by a segment of the FTP-75 driving cycle. Future research should validate this method for more than two controllable variables and the implications for the required learning time.

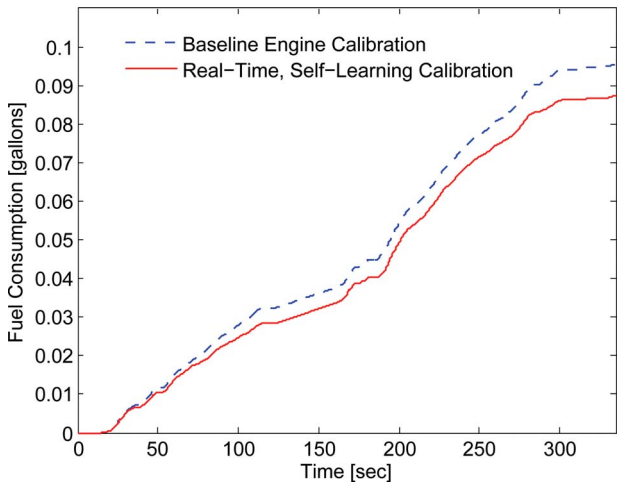


Fig. 12 Fuel consumption for the driving cycle

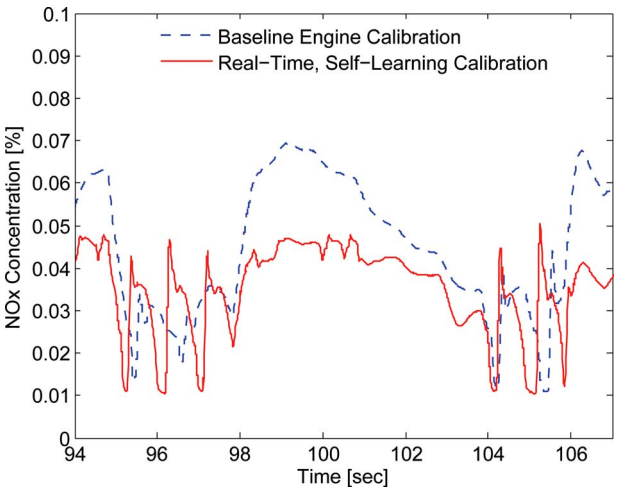


Fig. 14 NO_x concentration of emissions (zoom-in)

The decentralized control scheme, in conjunction with POSCA, can guarantee optimal calibration for steady-state and transient engine operating points designated by the driver's driving style. The ultimate goal of this approach is to fully exploit the engine's given technology in terms of the optimum specified performance criteria, e.g., fuel economy, pollutant emissions, and engine acceleration, that can be achieved. It aims to provide an answer to the following question: "For an engine with a given technology, what are the optimum performance criteria that a driver can get with respect to his/her driving habits?"

The long-term potential benefits of this approach are substantial. True fuel economy of vehicles will be increased while meeting emission standard regulations; drivers will be able to evaluate their driving behavior and learn how to improve fuel economy and reduce emissions by modifying it. This capability can be also especially appealing in engines utilized in hybrid-electric powertrain systems when real-time optimization of the power management is considered.

Acknowledgment

This research was partially supported by the Automotive Research Center (ARC), a U.S. Army Center of Excellence in Modeling and Simulation of Ground Vehicles at the University of Michigan. The engine simulation package ENDYNA THEMOS CRTD was provided by TESIS DYNAware GmbH. This support is gratefully acknowledged.

References

[1] Atkinson, C., and Mott, G., 2005, "Dynamic Model-Based Calibration Optimization: An Introduction and Application to Diesel Engines," SAE World Congress, Detroit, MI, Apr. 11–14, Paper No. SAE 2005-01-0026.

[2] Samulski, M. J., and Jackson, C. C., 1998, "Effects of Steady-State and Transient Operation on Exhaust Emissions From Nonroad and Highway Diesel Engines," SAE Transactions-Journal of Engines, Vol. 107.

[3] Green, R. M., 2000, "Measuring the Cylinder-to-Cylinder EGR Distribution in the Intake of a Diesel Engine During Transient Operation," SAE Transactions-Journal of Engines, Vol. 109.

[4] Malikopoulos, A. A., 2008, "Real-Time, Self-Learning Identification and Stochastic Optimal Control of Advanced Powertrain Systems," Ph.D. thesis, Department of Mechanical Engineering, University of Michigan, Ann Arbor.

[5] Malikopoulos, A. A., Papalambros, P. Y., and Assanis, D. N., 2007, "A Learning Algorithm for Optimal Internal Combustion Engine Calibration in Real Time," *Proceedings of the ASME 2007 International Design Engineering Technical Conferences and Computers and Information in Engineering Conference*, Las Vegas, NV, Sept. 4–7.

[6] Bush, R. R., and Mosteler, F., 1958, *Stochastic Models for Learning*, Wiley, New York.

[7] Atkinson, R. C., Bower, G. H., and Crothers, E. J., 1965, *An Introduction To Mathematical Learning Theory*, Wiley, New York.

[8] Tsypkin, Y. Z., 1971, *Adaptation and Learning in Automatic Systems*, Academic, New York.

[9] Narendra, K. S., and Wheeler, R. M., Jr., 1983, "N-Player Sequential Stochastic Game With Identical Payoffs," IEEE Trans. Syst. Man Cybern., **13**, pp. 1154–1158.

[10] Srikantakumar, P. R., and Narendra, K. S., 1982, "A Learning Model for Routing in Telephone Networks," SIAM J. Control Optim., **20**, pp. 34–57.

[11] Wu, H., 2007, "Decentralized Iterative Learning Control for a Class of Large Scale Interconnected Dynamical Systems," J. Math. Anal. Appl., **327**, pp. 233–245.

[12] Szer, D., and Charpillet, F., 2004, "Improving Coordination With Communication in Multi-Agent Reinforcement Learning," *Proceedings of the 16th IEEE International Conference on Tools With Artificial Intelligence, ICTAI 2004*, Boca Raton, FL, pp. 436–440.

[13] Scherrer, B., and Charpillet, F., 2002, "Cooperative Co-Learning: A Model-Based Approach for Solving Multi-Agent Reinforcement Problems," *Proceedings of the 14th IEEE International Conference on Tools With Artificial Intelligence*, Washington, DC, pp. 463–468.

[14] Beynier, A., and Mouaddib, A.-I., 2006, "An Iterative Algorithm for Solving Constrained Decentralized Markov Decision Processes," *Proceedings of the 21st National Conference on Artificial Intelligence and the 18th Innovative Applications of Artificial Intelligence Conference*, Boston, MA, Paper No. AAAI-06/IAAI-06, pp. 1089–1094.

[15] Yagan, D., and Chen-Khong, T., 2007, "Coordinated Reinforcement Learning for Decentralized Optimal Control," *Proceedings of the 2007 First IEEE International Symposium on Approximate Dynamic Programming and Reinforcement Learning*, Honolulu, HI (IEEE Catalog No. 07EX1572), pp. 296–302.

[16] Shen, D., Chen, G., Cruz, J. B., Jr., Kwan, C., and Kruger, M., 2007, "An Adaptive Markov Game Model for Threat Intent Inference," *Proceedings of the IEEE Aerospace Conference*, Big Sky, MT, pp. 1–13.

[17] Wheeler, R., and Narenda, K., 1986, "Decentralized Learning in Finite Markov Chains," IEEE Trans. Autom. Control, **31**(6), pp. 519–526.

[18] Puterman, M. L., 2005, *Markov Decision Processes: Discrete Stochastic Dynamic Programming*, 2nd revised ed., Wiley-Interscience, New York.

[19] Sennott, L. I., 1998, *Stochastic Dynamic Programming and the Control of Queueing Systems*, 1st ed., Wiley-Interscience, New York.

[20] Bertsekas, D. P., and Shreve, S. E., 2007, *Stochastic Optimal Control: The Discrete-Time Case*, 1st ed., Athena Scientific, Nashua, NH.

[21] Bertsekas, D. P., and Tsitsiklis, J. N., 1996, *Neuro-Dynamic Programming* (Optimization and Neural Computation Series 3), 1st ed., Athena Scientific, Nashua, NH.

[22] Sutton, R. S., and Barto, A. G., 1998, *Reinforcement Learning: An Introduction (Adaptive Computation and Machine Learning)*, MIT, Cambridge, MA.

[23] Borkar, V. S., 2000, "A Learning Algorithm for Discrete-Time Stochastic Control," Probability in the Engineering and Information Science, **14**, pp. 243–258.

[24] Malikopoulos, A. A., Papalambros, P. Y., and Assanis, D. N., 2007, "A State-Space Representation Model and Learning Algorithm for Real-Time Decision-Making Under Uncertainty," *Proceedings of the 2007 ASME International Mechanical Engineering Congress and Exposition*, Seattle, WA, Nov. 11–15.

[25] Malikopoulos, A. A., Assanis, D. N., and Papalambros, P. Y., 2007, "Real-Time, Self-Learning Optimization of Diesel Engine Calibration," *Proceedings of the 2007 Fall Technical Conference of the ASME Internal Combustion Engine Division*, Charleston, SC, Oct. 14–17.

[26] Heywood, J., 1988, *Internal Combustion Engine Fundamentals*, 1st ed., McGraw-Hill, New York.

[27] TESIS, http://www.tesis.de/en/

Numerical-Experimental Study and Solutions to Reduce the Dwell-Time Threshold for Fusion-Free Consecutive Injections in a Multijet Solenoid-Type CR System[1]

Andrea E. Catania
Alessandro Ferrari
Ezio Spessa

IC Engines Advanced Laboratory,
Politecnico di Torino,
Corso Duca degli Abruzzi 24,
10129 Torino, Italy

In "multijet" common rail (CR) diesel injection systems, when two consecutive injection current pulses approach each other, a merging of the two injections into a single one can occur. Such an "injection fusion" causes an undesired excessive amount of injected fuel, worsening both fuel consumption and particulate emissions. In order to avoid this phenomenon, lower limits to the dwell-time values are introduced in the control unit maps by a conservatively overestimated threshold, which reduces the flexibility of multiple-injection management. The injection fusion occurrence is mainly related to the time delay between the electrical signal to the solenoid and the nozzle opening and closure. The dwell-time fusion threshold was found to strongly decrease particularly with the nozzle closure delay. A functional dependence of the nozzle opening and closure delays on the solenoid energizing time and nominal rail pressure was experimentally assessed, and the injection temporal duration was correlated to the energizing time and rail pressure. A multijet CR injection-system mathematical model that was previously developed, including thermodynamics of liquids, fluid dynamics, mechanics of subsystems, and electromagnetism equations, was applied to better understand the cause and effect relationships for nozzle opening and closure delays. In particular, numerical results on the time histories of delivery- and control-chamber pressures, pilot- and needle-valve lifts, and mass flow rates through Z and A holes were obtained and analyzed to highlight the dependence of nozzle opening and closure delays on injector geometric features, physical variables, and valve dynamics. The nozzle closure delay was shown to strongly depend on the needle dynamics. Parametric numerical tests were carried out to identify configurations useful for minimizing the nozzle closure delay. Based on the results of these tests, a modified version of a commercial electroinjector was built, so as to achieve effectively lower nozzle closure delays and very close sequential injections without any fusion between them. [DOI: 10.1115/1.2938394]

Introduction

Diesel share among passenger cars has been continuously increasing over the past years mainly because of the fuel economy and "fun-to-drive" benefits of modern direct injection (DI) diesel engines. In fact, most of turbocharged DI diesel engines that are equipped with sophisticated electronically controlled high-pressure fuel injection systems (FISs) feature superior fun-to-drive aspects with respect to naturally aspirated gasoline engines with identical maximum rated power because of their favorable torque distribution, combined with good noise, harshness, and vibration (NHV). At the same time, engine manufacturers are increasingly challenged to find solutions to reduce the emissions without compromising the performance, reliability, and fuel economy, making the diesel engine an obvious choice in a variety of applications.

As in the recent past, the injection system will continue to be a key component in the diesel evolution toward innovative combustion systems while retaining both excellent fun to drive and fuel economy. More flexible multiple injections per cycle, increased injection pressure, and variable-orifice nozzles will be means for additional significant improvements in conventional diesel combustion. With regard to multiple-injection strategies, as is well known, the common rail (CR) already features the largest flexibility in the injection timing and pressure among diesel FISs. Currently, multijet CRs can enable up to five events per cycle, including pilot injection, triple main injection, and afterinjection.

The pilot injection consists of a small fuel amount (in the order of 1–3 mm^3 of diesel fuel) injected just before the main injections in order to increase the in-cylinder temperature and pressure. This limits the premixed combustion phase, yielding smooth heat release and reduced NO_x emissions.

Afterinjection consists of a small amount of fuel injected after the main-combustion end in order to reduce particulate matter (PM) emission by soot oxidation and, more recently, to promote regeneration in diesel particulate filters, as well as to generate hydrocarbons to purge a lean NO_x catalyst. As shown in Ref. [1],

[1]ICES2006 best paper award.

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 14, 2007; final manuscript received April 10, 2008; published online December 22, 2008. Review conducted by Christopher J. Rutland. Paper presented at the 2006 Spring Conference of the ASME Internal Combustion Engine Division (ICES2006), Aachen, Germany, May 7–10, 2006.

combining pilot and aftertinjections can give a significant improvement in NO_x and soot emissions without virtually any indicated mean effective pressute (IMEP) deterioration. It is also conceivable to split the aftertinjection, with the purpose of reducing wall wetting and prevent oil dilution [2].

With regard to main injection, the most advanced strategies currently enable up to three main events. Main-injection shots are also conventionally referred to as "preinjection," "main injection," and "postinjection." The combined use of exhaust gas recirculation (EGR), retarded injection timing, and injection splitting strategies, with carefully controlled fuel amounts in each pulse and dwell time (DT) between consecutive shots, can give rise to significant improvements in the NO_x and PM trade-off in both heavy-duty and small-displacement diesel engines [3–5].

In the past few years, novel diesel combustion concepts, such as the low-temperature combustion (LTC) of highly premixed charges, are gaining interest in the engine community for their potential of drastic containment in engine-out PM and NO_x emissions. Such LTC processes can be grouped in the wider homogeneous charge compression ignition (HCCI) concept, which they approximate at different levels [6–10].

The engine fuel handling apparatus should be well suitable to manage multiple injections, so as to achieve thorough fuel-air mixing and avoid wall wetting [11,12]. More specifically, it was assessed [13] that for early injections, short pulse widths should be employed to give better evaporation and mixing rate, whereas for late injections, short DTs result in less stratification of fuel. Furthermore, the injection-rate distributions, the mass ratio among pulses, and the number of pulses were shown to be critical parameters in the HCCI injection mode [14]. Finally, it is worth recalling that conventional diesel combustion processes are still required at high loads. Hence, the engine FIS should be well suitable for both conditions of HCCI and enhanced conventional diesel combustion.

All these premises support an approach to the design of the fuel handling apparatuses in the perspective of manufacturing injection systems that are capable of operating repeatable close injections of small and precise fuel amounts. Although the next generation of CR FISs applying the piezotechnology, which has recently been cast in production, is likely to achieve these objectives, the potential of conventional solenoid CR systems is not yet fully exploited. Only recently [15,16], the hydraulic dynamic response of a solenoid CR has thoroughly been investigated in the case of multiple injections for different values of energizing time (ET), nominal rail pressure, and DT. It was ascertained that the injection-system geometrical layout has a great influence on the pressure oscillations that are triggered by the injector-nozzle closure [15]. The amplitude and frequency of these pressure waves, as well as multiple-injection performance and the fluid-dynamic interaction between injectors of the same apparatus, are strongly dependent on the geometrical features of the injector inlet pipe and rail. Hence, any system component can be properly designed to assure prompt response with minimum disturbances [16,17], allowing a first extension of the DT range in comparison to current limits [17].

The present work addresses the issue of further reducing the lower limit in the DT range. Specific attention was paid to the merging of two consecutive injections, which can occur when the corresponding current pulses approach each other so that an undesired excessive injected-fuel amount is produced. The cause and effect relationship for injection fusion occurrence was investigated through an integrated numerical-experimental approach. From both specifically oriented experimentation on the bench [15] and parametric tests carried out by the novel NAIS code [18,19], a modified version of a production CR solenoid injector was realized to accomplish closer sequential injections without fusion.

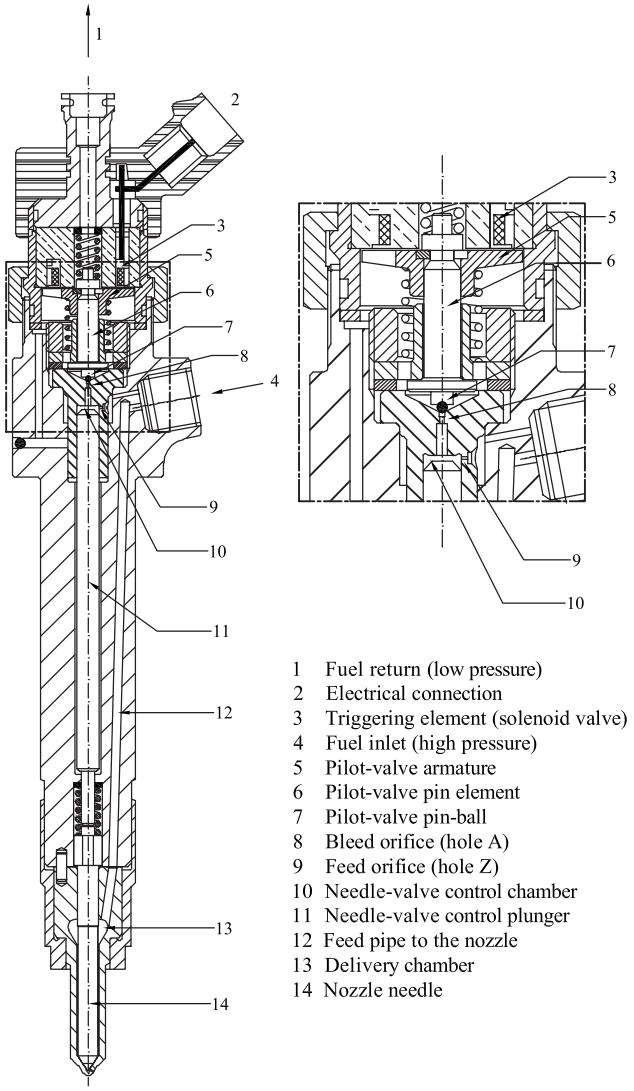


Fig. 1 Electroinjector

Experimental Facility

The experimental study was carried out by the Moehwald-Bosch MEP2000/CA4000 test bench described in Ref. [15], which was instrumented for a fluid-dynamic characterization of the high-pressure electronically controlled FIS. The facility has maximum shaft power and speed of 35 kW and 6100 rpm, respectively, with the possibility of simulating fast injection-system transient operations and choosing test parameters, so as to simulate real engine operating conditions. An injection-rate meter (EVI) was used to take injection flow-rate traces by recording the pressure signal produced by each injection in a test loop. The measured pressure time histories are converted into fuel mass distributions using the method developed in Ref. [15]. The oil injected mass was also gauged by an EMI2 device. Two thermopiezotransducers were applied to take temperature and pressure traces in the rail and at the injector inlet. Additional details on the experimental facilities and measuring equipment are given in Ref. [15].

A new generation solenoid-type CR FIS for a 1900 cc JTD engine was taken for study. It was controlled by a development electronic unit capable of managing up to five injections per engine cycle. A three-piston pump of the Radialjet Bosch type, having a displacement of 550 mm^3/rev, was used to supply the accumulator with the ISO 4113 oil, which simulates the diesel fuel. The electroinjector (Fig. 1) was equipped with a minisac nozzle. A

description of main electroinjector components is detailed in Ref. [19]. The system had a rail with a volume of 20 cm^3 and injector inlet pipes of 125 mm and 2.4 mm in length and internal diameter, respectively.

Common Rail Mathematical Model

Thermofluid Dynamics Model. The schematic of the hydraulic system was basically made up of chamber elements, i.e., volumes with all the characteristic dimensions of the same order, and pipe elements, i.e., volumes with one dimension prevailing over the others. In the chambers, pressure and temperature take uniform values that depend only on time, according to a lumped mass model. In the pipes, the thermofluid dynamic properties are allowed to vary with both time and axial coordinate, following a 1D approach [19]. The chamber mathematical model and the pipe-flow equations are detailed in Refs. [18,19]. Because of impulsive phenomena giving rise to high-frequency pressure oscillations, the wall shear stress was simulated by a frequency-dependent friction model [20,21], which was modified to take the turbulence of flow into account [19]. For predicting the temperature variations due to the liquid fuel compressibility, which can play a significant role in high-pressure injection-system simulation [22], the energy equation was applied. The effects of both fluid/wall heat transfer and viscous power dissipation were considered through a bulk polytropiclike specific heat, depending on the thermodynamic fluid-flow evolution [19].

Analytical expressions of ISO 4113 oil physical properties (such as isothermal bulk elasticity modulus and sound speed, thermal expansivity, specific heat at constant pressure, and specific heat ratio, as well as kinematic viscosity) were derived on the basis of experimental data [18,22].

Mechanical and Electromagnetic Submodel. The needle valve and the control plunger were modeled as two-degree-of-freedom mechanical systems whose behavior is described by two independent Newton dynamic equilibrium equations, which include fluid-dynamic, mechanical, as well as electromagnetic forces.

The pilot valve is made up of two mobile elements, i.e., the stem, or pin, and the armature (Fig. 1). These are integral during the solenoid excitation-induced valve lift. However, as the pin head reaches its stroke-limiter seat, the armature is allowed to swing with respect to the pin to damp the counterstroke of the pin on the basement and thus avoid oscillations of this element. In accordance with a mass-damper-spring model, both pin and armature were treated like harmonic oscillators.

The electromagnetic behavior of the solenoid was predicted by a simple and effective lumped parameter approach, which is detailed in Ref. [19]. From the recorded solenoid current and voltage time histories, the magnetic flux was obtained as a function of time by the integration of the generalized Ohm's law, and applying the Hopkinson law, the electromagnetic force was derived.

Results and Discussion

Fluctuations in the Injected Volume. When multiple injections happen, the pressure oscillations triggered by a specific pulse can influence the next-shot injected volume. As a result, an uncontrolled fluctuation of the volume introduced in the engine cylinder takes place as DT between two consecutive injection pulses is varied. Such a phenomenon was thoroughly studied in Refs. [15,16] and is briefly recalled in Figs. 2–4 for a double injection made up of a main pulse followed by a postinjection. Figure 2 reports the experimental total injected volume as a function of DT between the main-injection end and the postinjection start current signals at fixed ET of each pulse (i.e., $ET_{main}=1000$ μs, $ET_{post}=400$ μs) for the two different rail nominal pressure values $p_{rail}=1000$ bar and 1250 bar.

Figures 3 and 4 show the time histories of the inlet pressure

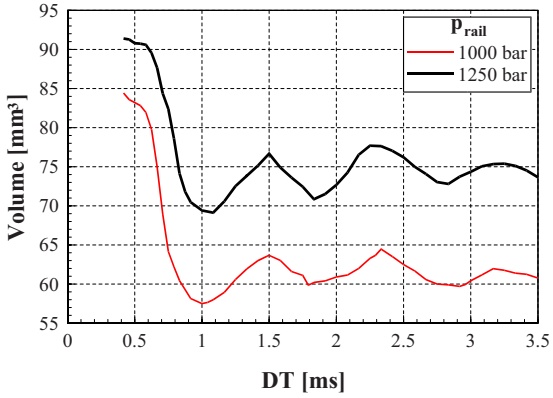


Fig. 2 Total injected-volume fluctuations

$p_{inj,in}$ to the injector (upper plot), the delivery-chamber pressure p_{dc} (middle plot), and the injected flow rate (lower plot) at DT $=1500$ μs (Fig. 3) and DT$=835$ μs (Fig. 4) for $p_{rail}=1000$ bar. Circle symbols refer to the experimental data, whereas the solid lines represent computed time histories. The abscissa $t-t_0$ stands for the time elapsed from an arbitrarily taken reference instant t_0. For DT$=1500$ μs (Fig. 3), the start of postinjection (SOPI) occurs around a maximum of the delivery-chamber pressure p_{dc}. Higher p_{dc} levels at postinjection start cause earlier and higher needle lifts, so that longer injection durations occur [15], corresponding to higher injected volumes (DT$=1500$ μs in Fig. 2). For DT $=835$ μs (Fig. 4), SOPI takes place during a decreasing phase in the temporal history of the delivery-chamber pressure (more specifically, SOPI occurs around the minimum that p_{dc} time history

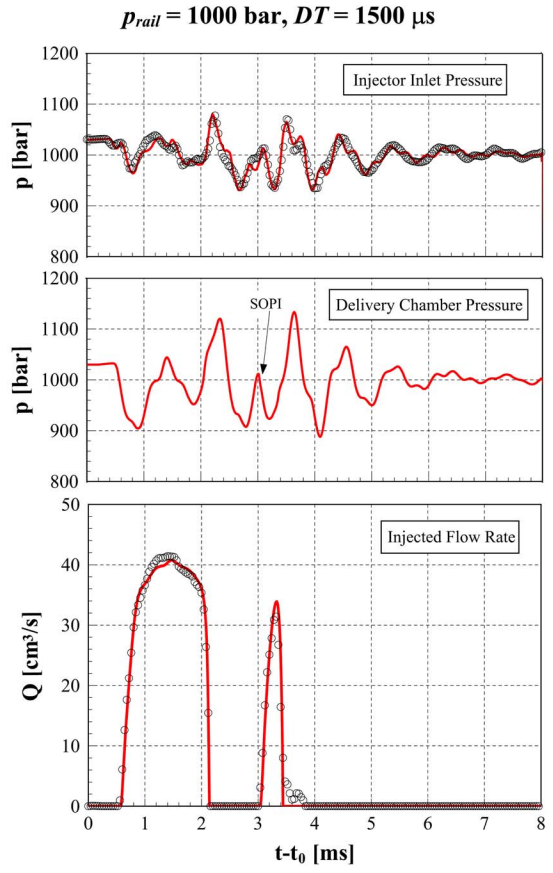


Fig. 3 $p_{inj,in}$, p_{dc}, and Q_{inj} for a high SOPI p_{dc} level

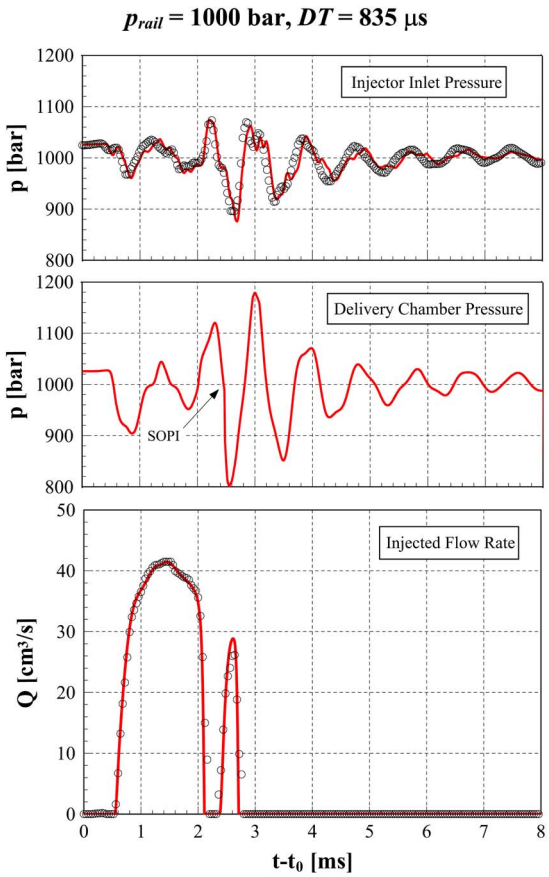


Fig. 4 $p_{inj,in}$, p_{dc}, and Q_{inj} for a low SOPI p_{dc} level

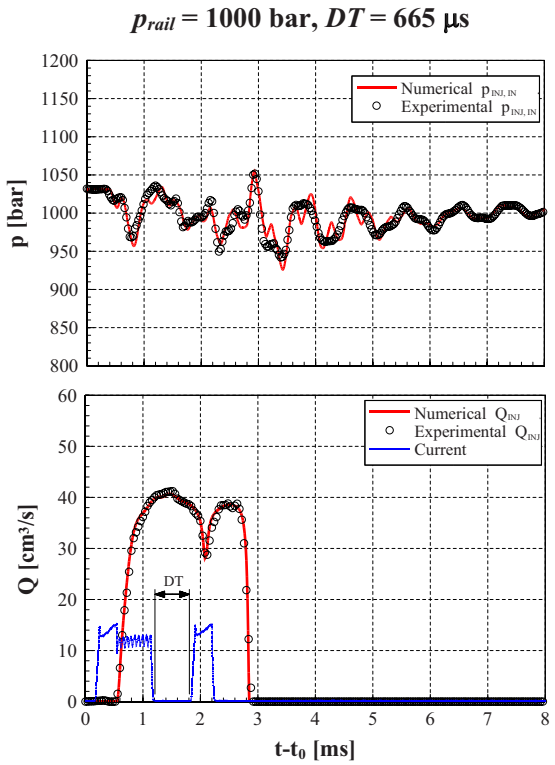


Fig. 5 $p_{inj,in}$, Q_{inj}, and EC at a main-post injection fusion

would present in the absence of any postinjection event). Therefore, a retarded and lower needle lift happens with a consequent shorter duration of the jet pulse, which in this case corresponds to a total injected volume close to a local minimum (as can be seen in Fig. 2 at DT=850 μs). It is worth recalling that the fluctuation frequencies of the injected volume for DT $\geqslant$ 835 μs (Fig. 2) are the same as those of the free oscillations of pressure, which are shown in Figs. 3 and 4 after the postinjection event [15]. It should be pointed out that CR layouts featuring short injector inlet pipes with large internal diameters reduce the amplitude of these oscillations [16,17].

Injection Fusion. The sudden increase in injected volume for DT lower than $\approx$850 μs (Fig. 2) shows the occurrence of a hydraulic instability that is referred to as "injection fusion" in this paper. It takes place when energizing current profiles of two consecutive injections are approached until the corresponding flow rates merge into a single injection event. An example of injection fusion is given in Fig. 5 for a double-pulse occurrence as is described in the previous subsection. More specifically, Fig. 5 reports the time histories of the injector inlet pressure (upper plot), injected flow rate (lower plot), and qualitative energizing current (EC) profiles (thin line in the lower plot) at DT=665 μs, for p_{rail}=1000 bar, ET_{main}=1000 μs, and ET_{post}=400 μs. The circle symbols in Fig. 5 refer to experimental data, whereas solid lines represent numerically predicted time histories. Although the current profiles of the consecutive pulses are distinct, the main injection and postinjection rates show a merging. As mentioned, the fusion of consecutive injections can cause an undesirable excessive amount of injected fuel, which yields a worsening in both pollutant emissions and fuel consumption. With the objective to avoid injection fusion, lower thresholds to the DT values are usually introduced in the engine control unit maps, limiting the flexible management of multiple injections and CR system capability of performing a larger number of injection shots. As will be shown, the reason of the injection fusion essentially resides in the temporal delay occurring between the start and stop of electrical signals to the solenoid and the needle lifts at both valve opening and closure. Therefore, this subsection will be aimed at a better understanding of the cause and effect relationships for nozzle opening delay (NOD) and nozzle closure delay (NCD) by the above-mentioned integrated numerical-experimental approach.

Figure 6 shows the predicted time histories (solid line) of the injector inlet pressure (upper plot) and of the injected flow rate (lower plot) in comparison to the experimental distributions (circle symbols in the plots) for a single injection shot occurring at p_{rail}=1000 bar and ET=1000 μs. The measured energizing current profile (thin line in the lower plot) is also represented in scale. In Fig. 6, a predicted NOD $\approx$ 0.35 ms is apparent between the start of current-signal rising ($t-t_0 \approx 0.1$ ms) and the instant at which the fuel injection begins ($t-t_0 \approx 0.45$ ms). Similarly, NCD $\approx$ 1 ms can be inferred between the instant of the electric current shutoff ($t-t_0 \approx 1.1$ ms) and the time at which the end of injection occurs ($t-t_0 \approx 2.1$ ms). With NCD > NOD, the interval of time over which the fuel is injected, i.e., the injection temporal length (ITL) (in Fig. 6) is larger than ET and is given by

$$\mathrm{ITL} = \mathrm{ET} + \mathrm{NCD} - \mathrm{NOD} \tag{1}$$

The cause-effect relations in the NCD and NOD occurrences can be investigated with the support of the computed quantities reported in Figs. 7–9 for the same p_{rail}=1000 bar and ET =1000 μs values indicated above Fig. 6. Figure 7 plots the delivery-chamber p_{dc} (thin solid line with solid dots) and control-chamber p_{cc} (thick solid line) pressure time distributions, as well as the pilot valve l_{pv} (dashed line) and needle lift l_n (dashed line with solid diamonds) time histories. Figure 8 shows the computed time history of the injector-needle velocity, whereas Fig. 9 plots the computed volumetric flow rates through the Z and A holes (Fig. 1, and schematic of the control chamber in Fig. 35 in the Appendix) as functions of time. Due to the delay in the response

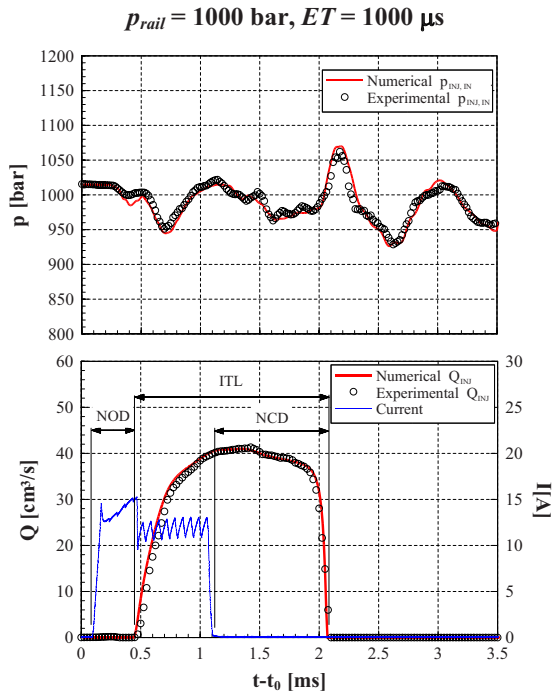


Fig. 6 Single-injection $p_{inj,in}$, Q_{inj}, EC, and characteristic time intervals

of electric circuit and of system fluid mechanics to solenoid current pulse, the predicted pilot-valve opening instant ($t-t_0 \approx 0.13$ ms, in Fig. 7) exhibits a small shift in relation to the start of the measured current signal ($t-t_0 \approx 0.1$ ms, in Fig. 6). Then, the pin element (6 in Fig. 1) of the pilot valve moves upward and

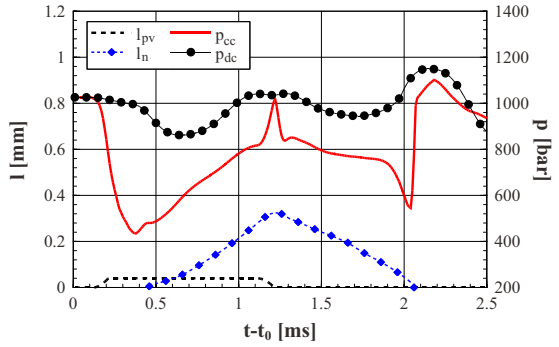


Fig. 7 Predicted lifts (l_{pv}, l_n) and pressures (p_{cc}, p_{dc})

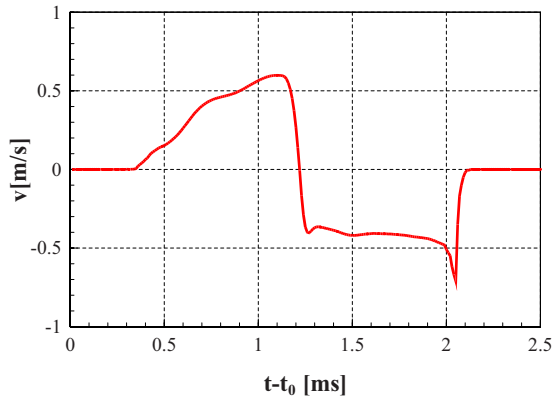


Fig. 8 Predicted needle velocity

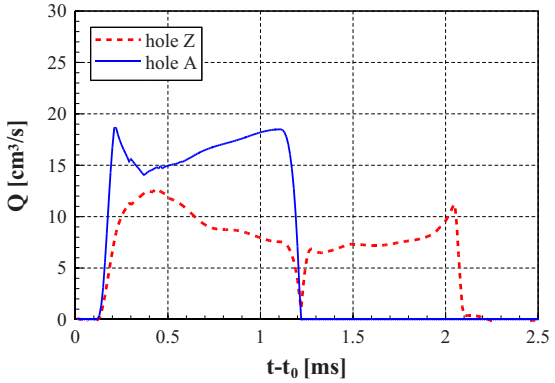


Fig. 9 Predicted flow rates through *A* and *Z* holes

quickly approaches its stroke end (i.e., ≈40 μm, in Fig. 7). At that time, the calculated p_{cc} has started to decrease. This can be explained, taking account of the results in Fig. 9, by the following mass-balance equation:

$$\frac{dp_{cc}}{dt} = \frac{a^2}{V_{cc}}(\dot{m}_Z - \dot{m}_A) \tag{2}$$

where V_{cc} is the control-chamber volume, a is the sound speed, $\dot{m}_Z$ indicates the flow rate entering the control chamber through the Z hole, and $\dot{m}_A$ is the flow rate that leaves the control chamber through the A hole. Details about the way Eq. (2) can be derived from the mass conservation equation of the control chamber and about the way $\dot{m}_Z$ and $\dot{m}_A$ are calculated can be found in the Appendix. Soon after pilot-valve opening, Fig. 9 shows that the fuel flow rate from the control chamber to the reservoir through the A hole exceeds that from the fuel inlet to the control chamber through the Z hole. In fact, the pressure drop through the A hole is larger than that through the Z hole since the pressure downstream from the A hole remains at values lower than ≈3 bar (nearly the reservoir pressure level), which is much lower than the pressure p_{cc} in the control chamber, p_{cc} being the same order of magnitude as the pressure upstream from the Z hole, that is, close to p_{rail}. In addition, the A hole has a larger diameter than the Z hole. Thus, Eq. (2) justifies the early p_{cc} decreasing stage that is shown in Fig. 7.

As can be observed from needle-lift (l_n in Fig. 7) and needle-velocity (Fig. 8) patterns, the injector needle proceeds to move upward later with respect to the time at which the pin element of the pilot valve reaches its stroke end. In fact, p_{cc} must be low enough (≈440 bar in Fig. 7) so that the resulting pressure force on the control plunger (11 in Fig. 1) and the nozzle needle (14 in Fig. 1) overcomes the spring preload acting on the needle and so brings upward both mobile elements. It has to be pointed out that the needle lift time history was plotted with reference to the lift value at which the nozzle opens and fuel injection starts ($t-t_0 \approx 0.45$ ms, in the lower plot of Fig. 6). However, the needle velocity (Fig. 8) is greater than zero at $t-t_0 \approx 0.35$ ms, i.e., well before fuel injection starts ($t-t_0 \approx 0.45$ ms). It can be speculated that during this time interval, the needle takes ≈0.10 ms to regain axial nozzle deformation. Thus, the force reacting to the needle compression of the nozzle, which contributes to lifting the needle, gradually reduces, so that the needle velocity (Fig. 8) remains low. As a consequence, during the time interval between the start of needle movement and the beginning of fuel injection, the tiny reduction of control-chamber volume caused by needle motion plays a role in slightly raising p_{cc}, according to the following, and more general, continuity equation:

$$\frac{dp_{cc}}{dt} = \frac{a^2}{V_{cc}}\left[(\dot{m}_Z - \dot{m}_A) + \rho_{cc} S_{cp} \frac{dl_n}{dt}\right] \quad (3)$$

where ρ_{cc} is the fuel density in the control chamber and S_{cp} is the control-plunger cross-section area (Appendix).

Therefore, it can be concluded that the NOD indicated in Fig. 6 ($\approx$0.35 ms) is made up of two periods: the time interval required to partly discharge the control chamber ($\approx$0.25 ms), so as to start the needle motion, and the time interval required by needle lift to regain the nozzle axial deformation ($\approx$0.10 ms).

However, the minimum ET that is required to induce the nozzle opening is shorter than NOD and, in theory, equals only the first contribution ($\approx$0.25 ms), i.e., the time period that is necessary to start the needle movement with enough acceleration to complete the nozzle axial-deformation recovery. For any ET shorter than this threshold value (i.e., $\approx$250 μs), the pressure level reached in the control chamber at the end of the electric pulse is not so low as to yield nozzle opening, giving rise only to a partial recovery of axial nozzle deformation and to unsteady flow phenomena in the electroinjector. It should be pointed out that increasing the diameter of the *A* hole, so as to reduce the time required to empty the control chamber, could be an effective means of reducing the values of both NOD and ET thresholds. However, a too large *A*-hole diameter would increase the risk of mechanical separation of control plunger and nozzle needle during the injection phase, as is detailed further on. In addition, NOD reduction can be obtained by using materials with higher hardness for the nozzle. This can reduce the initial axial nozzle deformation, thus decreasing the time necessary to regain it.

The pressure level p_{dc} in the delivery chamber (Fig. 7) remains virtually equal to the rail pressure until the nozzle starts opening. Actually, a slight p_{dc} reduction (for $t-t_0<0.45$ ms, in Fig. 7) is induced by the propagation of the rarefaction wave that arises in the control chamber because of the pilot-valve opening.

After the injection flow rate has started, the needle velocity (Fig. 8), p_{cc} (Fig. 7), and l_n (Fig. 7) undergo an increase. Actually, although the flow rates through the *A* hole continues to exceed the flow rate through the *Z* hole, as Fig. 9 shows, the reduction of the control-chamber volume, caused by the control piston lift, overcomes this effect, determining a raise in the pressure, as can be derived from Eq. (3). The increase in p_{cc} produces a counterpressure for the needle motion, thus maintaining a relatively low needle velocity (lower than $\approx$0.6 m/s in Fig. 8). To that end, the diameter of the *A* hole is a critical parameter. If the *A* hole is too large, the needle lift does not produce an adequate increase in p_{cc}, so that the needle velocity can strongly rise, and separation between nozzle needle and control plunger is likely to occur when such a mobile element unit reaches its mechanical stroke end.

With regard to the pressure in the delivery chamber, Fig. 7 shows that as soon as the needle begins the opening phase (for $t-t_0>0.45$ ms), p_{dc} significantly decreases as a result of the fuel delivery. This decrease in p_{dc} draws the fuel from the injector-drilled pipe toward the nozzle and, consequently, the flow rate to the delivery chamber can grow until it exceeds the injected flow rate because the section of the injector-drilled pipe is larger than the restricted nozzle flow area. Then, p_{dc} begins to increase ($t-t_0\approx0.70$ ms), approaching ($t-t_0\approx1.05$ ms) the measured value before the injection start.

When the current signal goes back to zero ($t-t_0\approx1.1$ ms, in Fig. 6), the pilot valve starts to close. As can be inferred from the computed l_{pv} pattern in Fig. 7, the pilot valve starts opening at $t-t_0\approx0.13$ ms and closes at $t-t_0\approx1.20$ ms; i.e., it stays open for a time interval that is approximately equal to ET. As soon as this valve closes, a water-hammer phenomenon occurs, giving rise to a sudden pressure increase in the control chamber (p_{cc} at $t-t_0\approx1.2$ ms, in Fig. 7). In fact, with reference to the right hand side of Eq. (3), $\dot{m}_A$ becomes equal to zero and the other terms contribute both to a pressure increase inside the control chamber, leading to the water-hammer pressure peak. This slows down the needle, stopping it at the maximum lift of $\approx$0.32 mm (Fig. 7), which is lower than the lift corresponding to the needle mechanical stroke end. The duration of this phase ($\approx$0.12 ms) can be easily estimated as the period between the instant of current shutoff ($t-t_0\approx1.1$ ms, in Fig. 6) and the time at which the velocity of the needle becomes zero ($t-t_0\approx1.22$ ms, in Fig. 8); i.e., the needle stops at its maximum lift.

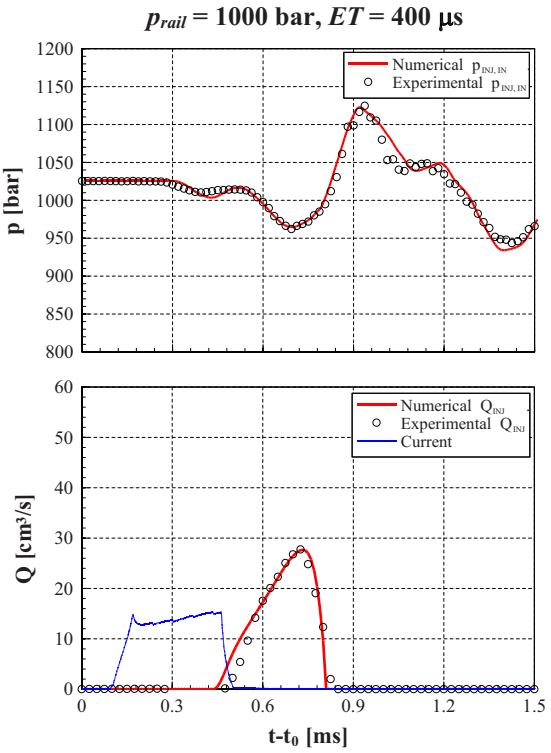


Fig. 10 Numerical versus experimental $p_{inj,in}$ and Q_{inj}

It is worth observing that the intensity of p_{cc} peak generally increases with ET. In fact, as ET is raised, the control-chamber volume at the time of pilot-valve closure decreases, with a consequent increase in the p_{cc} peak. However, the intensity of this peak can be high also for low ETs, as shown in Figs. 10–12 for a single injection with ET=400 μs and the same p_{rail}, as in Figs. 6–9. In fact, the pilot-valve closure occurs at $t-t_0\approx0.64$ ms (Fig. 11) when the flow rate through the *Z* hole is still elevated (Fig. 12).

In relation to the dependence of the p_{cc} peak on p_{rail}, Figs. 13–15 report the results obtained for a single injection pulse with the same ET=1000 μs as in the case of Figs. 6–9 but with a higher nominal rail pressure (p_{rail}=1250 bar). It can be seen that the intensity of the p_{cc} peak caused by the pilot-valve closure at

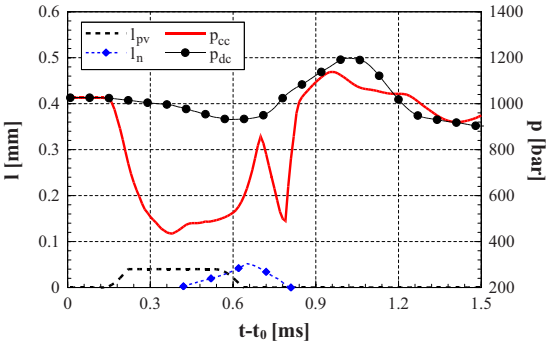


Fig. 11 Predicted lifts (l_{pv}, l_n) and pressures (p_{cc}, p_{dc})

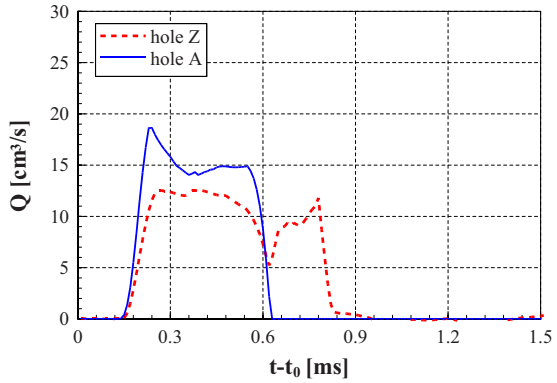


Fig. 12 Predicted flow rates through *A* and *Z* holes

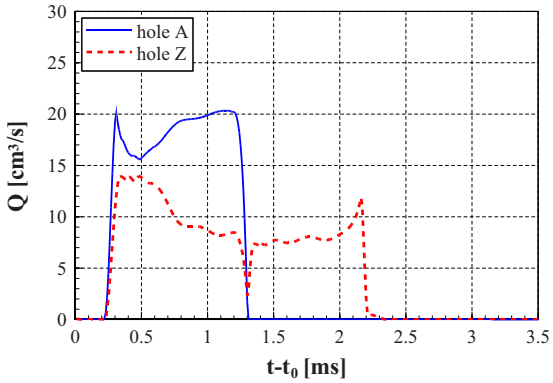


Fig. 15 Predicted flow rates through *A* and *Z* holes

p_{rail}=1250 bar (Fig. 14) is higher than that at p_{rail}=1000 bar (Fig. 7). This can be ascribed to the increased maximum needle lift and to the larger $\dot{m}_A$ and $\dot{m}_Z$ values at p_{rail}=1250 bar (Fig. 15 versus Fig. 9).

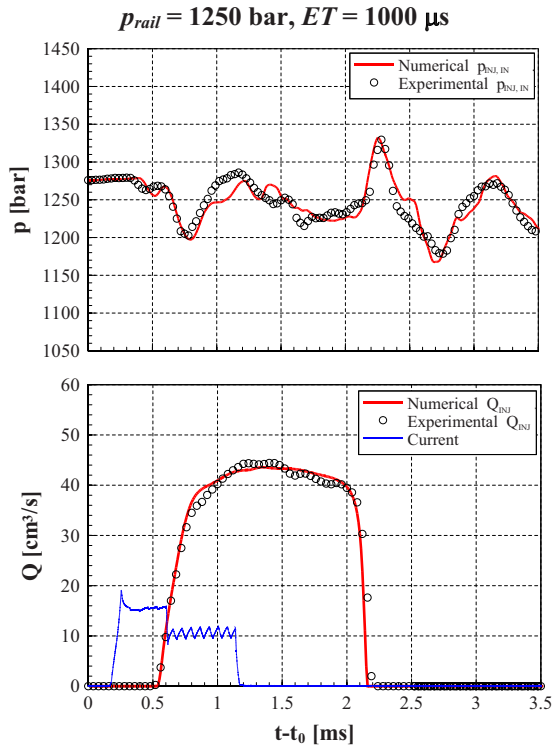


Fig. 13 Numerical versus experimental $p_{inj,in}$ and Q_{inj}

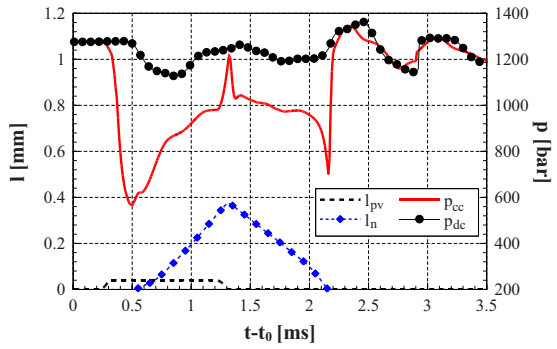


Fig. 14 Predicted lifts (l_{pv}, l_n) and pressures (p_{cc}, p_{dc})

After the needle stops at its maximum lift, the water hammer due to the pilot-valve closure further decelerates the needle, thus changing the direction of its motion, enabling injector closure. During this phase ($t-t_0>1.25$ ms, in Fig. 8) the needle absolute velocity remains at low values. As the needle lift decreases, p_{cc} reduces because of the control-chamber volume increase. When the needle approaches its seat ($t-t_0>1.9$ ms, in Fig. 7), p_{dc} rises due to the flow passage restriction there, which cannot deliver the flow rate coming from the nozzle drilled passage. The p_{dc} growth is useful in reducing the needle velocity, thus avoiding any possible nozzle reopening. Consistent with the injection-rate pattern (Fig. 6), l_n plot (Fig. 7), and needle-velocity distribution (Fig. 8), injector closure occurs at $t-t_0\approx2.06$ ms. It is worth pointing out that the absolute needle velocity is higher than zero, beyond the period $t-t_0\approx2.06$ ms, because of the nozzle axial deformation.

The water hammer due to the end of injection is evident for $t-t_0>2.1$ ms in the distributions of both p_{dc} (Fig. 7) and $p_{inj,in}$ (Fig. 6). On the other hand, the p_{cc} peak at $t-t_0\approx2.20$ ms (Fig. 7) is caused by the fuel entering the control chamber through the *Z* hole (Fig. 9), which cannot be compensated by an increase in the control-chamber volume. Therefore, it can be summarized that the NCD value shown in Fig. 6 ($\approx$1 ms), for a single injection pulse, is the sum of two contributions: the time interval required by the needle to invert its motion (deceleration phase $\approx$0.15 ms, in Fig. 8) soon after the electric input power to the solenoid is shut off and the time interval necessary to definitely close the nozzle (closure phase $\approx$0.85 ms, in Fig. 7).

When two consecutive injections take place, the lower DT limit to avoid injection fusion, i.e., injection fusion threshold (IFT), depends on the NCD value of the first-injection pulse and on the NOD value of the subsequent injection. Since NCD and NOD depend on ET and p_{rail}, IFT must be a function of these.

NOD Dependence on ET and p_{rail}. Figure 16 reports experimental values of NOD as functions of ET for a single injection pulse at different nominal values of rail pressure (p_{rail}=750 bar, 1000 bar, and 1250 bar). NOD was evaluated as the difference between the time instant corresponding to the pressure rise in the EVI profile and that corresponding to the switch on of the electrical impulse to the CR injector solenoid.

For p_{rail}=1000 bar, NOD results in a slight increase from 380 μs to 400 μs as ET varies from 400 μs to 1500 μs. Such increase can be ascribed to the fact that longer injections produce higher temperatures in the electroinjector at constant p_{rail}. As the fluid temperature increases, the sound speed decreases. Furthermore, the fuel density decreases, thus reducing $\dot{m}_Z-\dot{m}_A$. Based on Eqs. (2) and (3), this results in a longer time required for the control-chamber discharge, which, in turn, causes a longer NOD.

NOD was shown to slightly decrease when higher p_{rail} were used at fixed ET. In fact, higher p_{rail} implies higher p_{cc} at pilot-

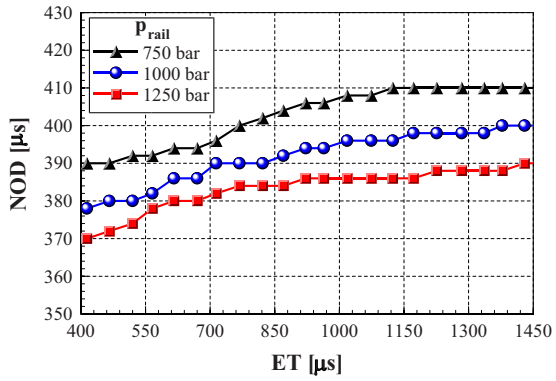


Fig. 16 NOD dependence on ET and p_{rail}

valve opening, thus resulting in a larger flow rate through the A hole during the discharge of the control chamber. Such effect, which is beneficial to NOD reduction, is mitigated by the presence of a larger drop in p_{cc}, as can be inferred from a comparison between Fig. 14 (p_{rail}=1250 bar and $\Delta p_{cc} \approx 710$ bar) and Fig. 7 (p_{rail}=1000 bar and $\Delta p_{cc} \approx 590$ bar).

NCD Dependence on ET and p_{rail}. Figure 17 shows measured NCD values as a function of ET for a single injection at different rail pressures (p_{rail}=750 bar, 1000 bar, and 1250 bar). NCD was measured as the time interval between the shutoff of the electrical impulse to the solenoid and the end of injection.

At each p_{rail}, NCD has a virtually linear increasing trend in a wide ET range. In particular, Fig. 17 shows that NCD$\approx$ET+b (where b is a constant value, basically dependent on injector configuration and on nominal rail pressure), up to a constant NCD level that is reached at a discriminating value of ET. Such a level decreases as p_{rail} increases. In fact, the larger ET is, the longer is the distance covered by the needle during its opening lift and, therefore, the longer is the time employed by the needle to come back to its nozzle closure position (after the shutoff of the current pulse to the solenoid). However, when ET is long enough to let the needle reach its mechanical stroke end (i.e., 0.43 mm for the considered electroinjector), a further increase in ET cannot raise NCD more.

It is worth pointing out that although NCD increases with ET, the time interval corresponding to the deceleration phase of the needle is roughly constant ($\approx$100–150 μs). Actually, as ET grows, both needle velocity and p_{cc} increase at the start of the pilot-valve closure; the higher p_{cc} is the larger is the counterforce on the control plunger, with a stronger damping effect on the needle speed. As ET increases and NCD linearly grows with it, the needle deceleration period reduces to a progressively smaller percentage of NCD.

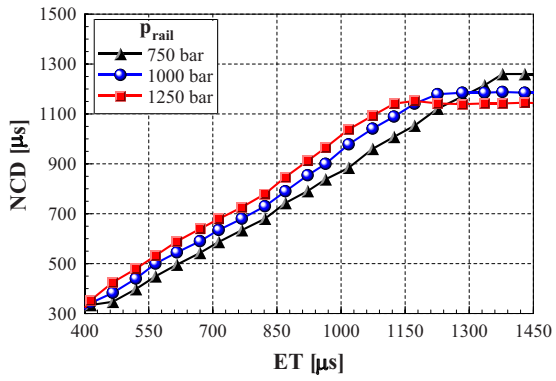


Fig. 17 NCD dependence on ET and p_{rail}

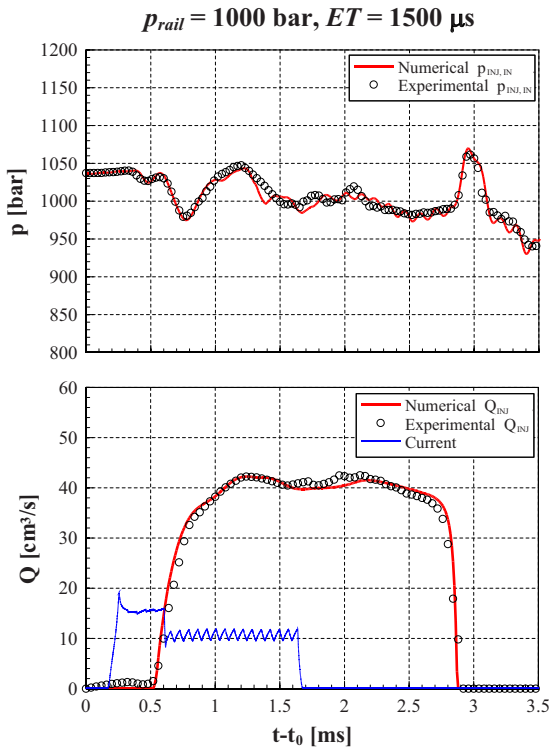


Fig. 18 Numerical versus experimental $p_{inj,in}$ and Q_{inj}

With reference to the dependence of NCD on p_{rail} at fixed ET, Fig. 17 shows that NCD becomes larger as p_{rail} increases. Actually, a higher p_{rail} increases the difference between p_{dc} and p_{cc} soon after the pilot-valve opening and at the initial stage of the needle lift, as is supported by the results of Figs. 7 and 14. Consequently, a greater acceleration of the needle occurs and the l_n peak value increases with p_{rail}. As the needle moves upward, the predicted difference between p_{dc} and p_{cc} gradually tends to the same value for all p_{rail} levels, whereas it increases again with p_{rail} during the nozzle closure. Thus, larger damping forces act on the needle, contrasting its downward acceleration during nozzle closure. The combined effect of the higher l_n peak and larger damping forces results in longer NCDs as p_{rail} increases.

However, when ET is long enough to let the needle reach its mechanical stroke end (i.e., ET greater than the above-mentioned discriminating value), NCD decreases as the p_{rail} level increases. In order to clarify this behavior, the operating conditions in which the needle reaches its stroke end are examined in Figs. 18–23 for ET=1500 μs. More specifically, Figs. 18 and 21 show the time histories of the injector inlet pressure (upper plots), injected flow

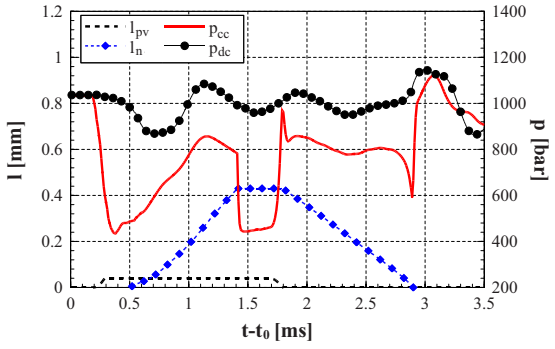


Fig. 19 Predicted lifts (l_{pv}, l_n) and pressures (p_{cc}, p_{dc})

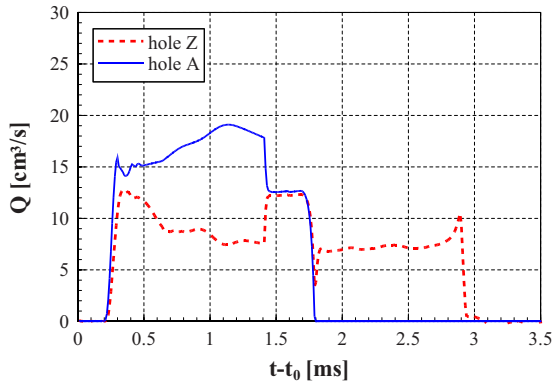


Fig. 20 Predicted flow rates through *A* and *Z* holes

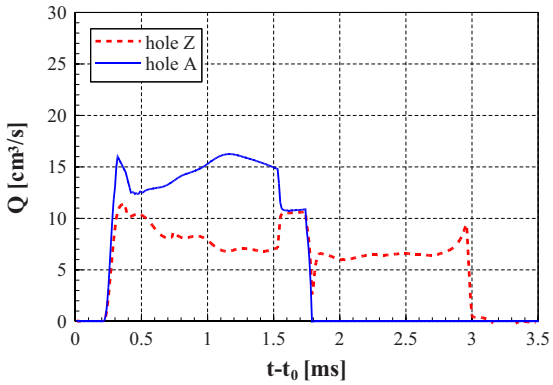


Fig. 23 Predicted flow rates through *A* and *Z* holes

rate (lower plots), and injection current profiles (thin lines in the lower plots) at $p_{rail}=1000$ bar (Fig. 18) and 750 bar (Fig. 21). Symbols refer to experimental data and solid lines represent numerically predicted time histories. Figures 19, 20, 22, and 23 report the time histories of l_{pv}, l_n, p_{cc}, and p_{dc} (Figs. 19 and 22) and those of volumetric flow rates through *Z* and *A* holes (Figs. 20 and 23) for tests at $p_{rail}=1000$ bar (Figs. 19 and 20) and 750 bar (Figs. 22 and 23) with ET=1500 μs. By comparison of Figs. 19 and 22, it can be inferred that the water hammer produced by pilot-valve closure is higher for $p_{rail}=1000$ bar (p_{cc} increases of about 530 bar at $t-t_0\approx1.75$ ms, in Fig. 19) than for $p_{rail}=750$ bar (p_{cc} increases of about 400 bar at $t-t_0\approx1.75$ ms, in Fig. 22). Such results are consistent with the flow-rate values through the *A* hole, which are higher for $p_{rail}=1000$ bar (Fig. 20) than for $p_{rail}=750$ bar (Fig. 23) during pilot-valve opening. Therefore, a higher p_{rail} corresponds to a greater impulse to the needle from its stroke end downward, resulting in a larger needle acceleration and thus in a shorter NCD. This is the predominant effect in relation to the whole needle-closure dynamics, as was assessed by the numerical model.

Finally, it is interesting to compare, for the same nominal rail pressure value ($p_{rail}=1000$ bar), the two cases: The needle reaches its mechanical stroke end (Figs. 18–20, ET=1500 μs); the needle stops at a lower lift (Figs. 6–9, ET=1000 μs). When the needle reaches its stroke end ($t-t_0\approx1.4$ ms, in Fig. 19), p_{cc} undergoes an abrupt decrease because the volume of the control chamber cannot reduce any further and, consequently, the flow rate going out through the *A* hole is subjected to a sudden reduction before the pilot-valve closure (Fig. 20, $t-t_0\approx1.4$ ms). Then, the water hammer produced by the pilot-valve closure is much more significant for the case in which the mechanical end of the stroke is reached by the needle (p_{cc} increase of about 530 bar at $t-t_0\approx1.75$ ms, in Fig. 19) with respect to the opposite case (p_{cc} increase of about 200 bar at $t-t_0\approx1.15$ ms, in Fig. 7).

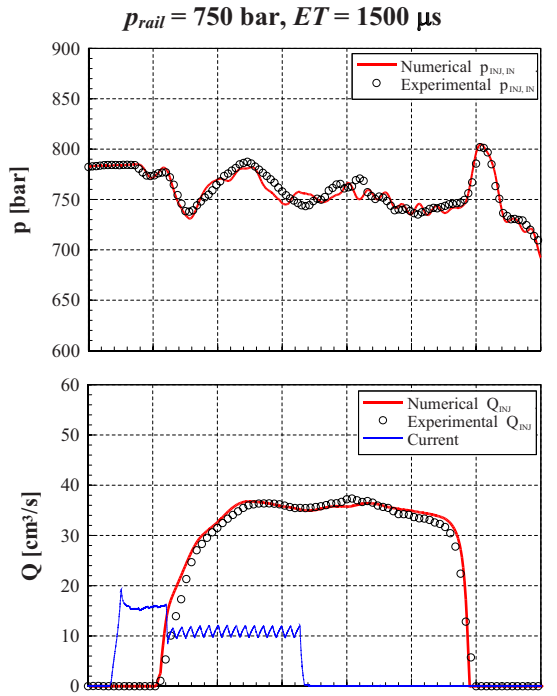


Fig. 21 Numerical versus experimental $p_{inj,in}$ and Q_{inj}

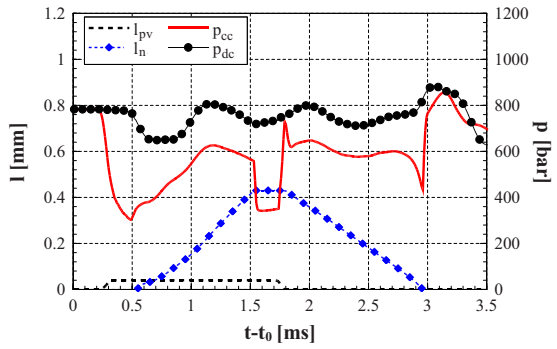


Fig. 22 Predicted lifts (l_{pv}, l_n) and pressures (p_{cc}, p_{dc})

ITL Dependence on ET and p_{rail}. Figure 24 reports the ITL experimental values given by Eq. (1) as a function of ET, for a

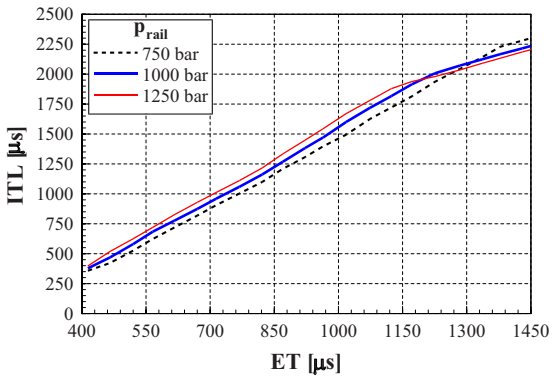


Fig. 24 Experimental ITL dependence on ET and p_{rail}

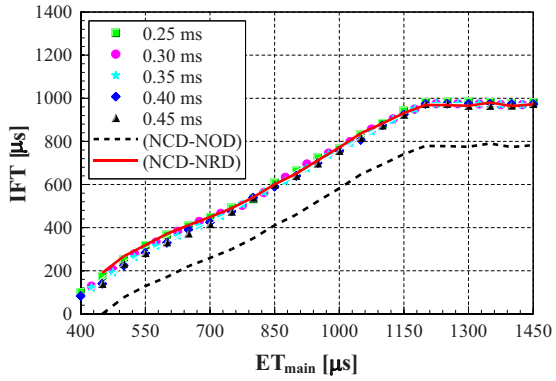


Fig. 25 IFT predictive correlations versus experimental data

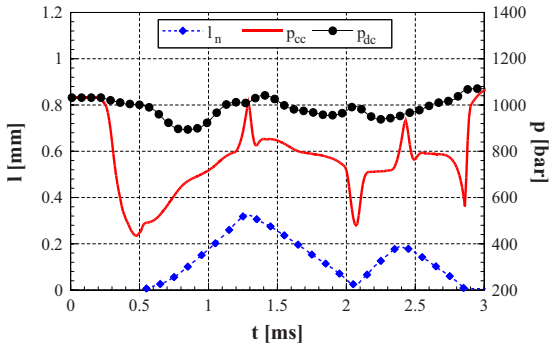


Fig. 26 Predicted l_n, p_{cc}, and p_{dc} time histories for a double injection with flow-rate fusion: p_{rail}=1000 bar, ET_{main}=1000 μs, ET_{post}=400 μs, DT=665 μs

single injection, at different values of the rail pressure (namely, p_{rail}=750 bar, 1000 bar, and 1250 bar). For the injection system under investigation, ITL is shorter than ET when ET<430 μs, as can be inferred from Figs. 16 and 17, showing that NCD is lower than NOD when ET<430 μs.

At fixed p_{rail}, ITL shows a linear increasing trend versus ET, keeping the same slope until the needle reaches its mechanical stroke end. Therefore, this part of the diagram can be interpolated by the following linear correlation:

$$\mathrm{ITL} = 2\mathrm{ET} - b_1 \tag{4}$$

Actually, in this range, NOD is virtually independent of ET, and NCD increases linearly with ET. According to the results in Fig. 24, b_1 can be expressed as a function of p_{rail},

$$b_1 = |q| \cdot p_{\mathrm{rail}}^{-|r|} \tag{5}$$

q and r are quantities that depend on the nozzle configuration.

Once the needle has reached its mechanical stroke end, the slopes of the trend lines in Fig. 24 change, in accordance with the linear correlation,

$$\mathrm{ITL} = \mathrm{ET} + b_2 \tag{6}$$

In fact, both NOD and NCD are virtually constant and b_2 can take the dependence on p_{rail} into account by a relation of the type

$$b_2 = \sum_{i=0}^{2} z_i p_{\mathrm{rail}}^{i} \tag{7}$$

in which the coefficient z_i changes with the nozzle configuration and the maximum needle lift. It is interesting to observe that the ITL dependence on p_{rail} is almost the same as that of NCD, NOD being dependent on p_{rail} to a much smaller extent than NCD.

Injection Fusion Threshold. The IFT can be defined as the minimum value of DT that does not cause any merging between two consecutive injections. It can be measured by progressively making the second injection pulse closer to the first one until the two corresponding injection rates merge into a single one.

In Fig. 25, the IFT of a double main injection–postinjection is plotted as a function of ET_{main} for p_{rail}=1000 bar. The symbols refer to the experimental values of IFT, obtained for different values of ET_{post}, which were selected in a typical automotive application range, i.e., 250–450 μs, as reported in the figure label. The solid and dashed lines plot the outcomes of two different correlations for IFT, as follows. It can be deduced from experimental results (symbols, in Fig. 25) that IFT does not depend on ET_{post} and shows a similar pattern to that of NCD at the same nominal rail pressure (Fig. 17). Therefore, at fixed p_{rail}, IFT only depends on the ET of first injection.

The dashed line in Fig. 25 is obtained by evaluating IFT as

$$\mathrm{IFT} = \mathrm{NCD}_{\mathrm{main}} - \mathrm{NOD}_{\mathrm{post}} \tag{8}$$

where NCD_{main} is varied as a function of ET_{main} according to the results reported in Fig. 17 for p_{rail}=1000 bar, whereas a constant average value of ≈390 μs was assessed for NOD_{post} based on the results corresponding to p_{rail}=1000 bar in Fig. 16. It is evident that Eq. (8) significantly underestimates the experimental IFT values. This suggests that when injection fusion of two pulses occurs, the time delay required to reopen the nozzle for the second pulse, called nozzle reopening delay (NRD), should be lower than the NOD measured for a single pulse under the same operation parameters due to different system dynamic responses.

Figure 26 shows the computed l_n, p_{cc}, and p_{dc} time histories for a double injection in which a flow-rate fusion occurs (p_{rail}=1000 bar, ET_{main}=1000 μs, ET_{post}=400 μs, DT=665 μs). For this case, the corresponding measured and computed values of the injector inlet pressure and injection rate time histories were already reported in Fig. 5 and discussed in addition to the measured time distributions of the solenoid input current. The same quantities for a single injection with the same p_{rail} of 1000 bar and ET of 400 μs have also been reported in Figs. 10 and 11 and discussed. It can be easily observed that when the postinjection electric current pulse is given to the solenoid ($t-t_0$≈1.8 ms in Fig. 5), the needle is still open (l_n at $t-t_0$≈1.8 ms in Fig. 26) and the level of p_{cc} (solid line in Fig. 26) is lower than that measured at the start of the current pulse for the single injection in Fig. 11 ($t-t_0$ ≈0.10 ms). Therefore, a shorter time is required for the partial control-chamber discharge, and no axial deformation of the nozzle must be regained, thus supporting a substantial reduction of NRD with respect to NOD. Based on several numerical calculations at different p_{rail} and ET, NRD was shown to be almost constant and approximately equal to half the mean NOD value.

The solid line in Fig. 25 is obtained by evaluating IFT as

$$\mathrm{IFT} = \mathrm{NCD}_{\mathrm{main}} - \mathrm{NRD}_{\mathrm{post}} \tag{9}$$

where NCD_{main} is evaluated as a function of ET_{main}, according to the results reported in Fig. 17 for p_{rail}=1000 bar, whereas a constant average value of ≈200 μs was consistently chosen for NRD_{post}. The agreement of IFT evaluated with Eq. (9) and the experimental data is excellent, substantiating the validity of the approach used in the present investigation.

Figure 27 shows measured values of IFT versus ET_{main} for different values of p_{rail}. It can be inferred that ET_{main} strongly influences IFT, whereas, according to the results in Figs. 16 and 17, p_{rail} plays a minor role in the evaluation of IFT.

Numerical Tests for IFT Reduction. In this section, the effectiveness of simple injector-design modifications for IFT reduction was explored. In particular, for instance, as before, a main injection and postinjection double pulse (p_{rail}=1250 bar, ET_{main}

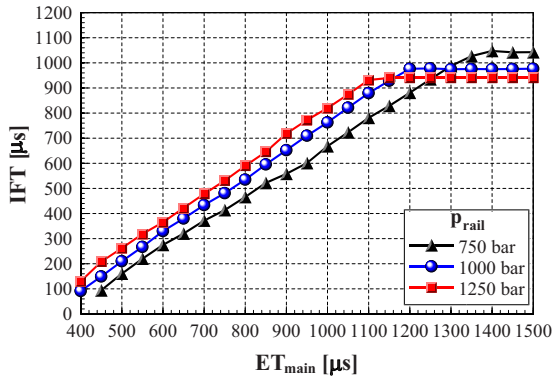


Fig. 27 IFT dependence on ET and p_{rail}

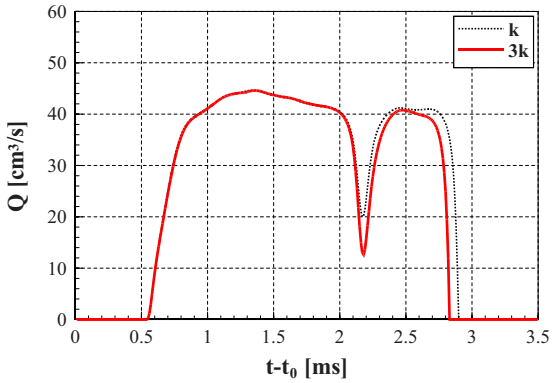


Fig. 30 Needle-spring stiffness effect on IFT

=1000 μs, and ET$_{post}$=400 μs) was studied, considering the production injector as a reference, or base one. DT was chosen so that injection fusion occurs (DT=750 μs < IFT ≈ 820 μs, Fig. 27).

Figure 28 reports the experimental (circles) and predicted (solid line) injection-rate patterns, with current pulse signals (thin line), as reference conditions for the IFT reduction parametric analysis. The numerical simulation was applied to study the effects of the following design variables on both the injected flow-rate time history and IFT: the overall mass of needle and control piston (Fig. 29), the needle-spring stiffness (Fig. 30), and preload (Fig. 31), and the needle-plunger mechanical stroke end (Fig. 32). The captions in Figs. 29–32 indicate the variation of each parameter with respect to the original value of the base injector configuration (namely, M, k, F, and L). For all the examined cases, the computed injection rate of the baseline injector is reported by dotted lines in the figures, for comparison.

With regard to the influences of the overall mass of the control plunger and needle (Fig. 29), as well as of the needle-spring stiffness (Fig. 30), no significant effects on NOD and NCD were found even though the reference mass (M) was halved ($0.5M$) and the reference stiffness (k) was tripled ($3k$).

An increase in the nominal needle-spring preload (Fig. 31) reduces NCD$_{main}$, thus lowering IFT. More specifically, if the preload is doubled ($2F$) with respect to the reference value (F), injection fusion disappears. A further increase in the needle-spring preload would further decrease IFT but would also give rise to a lower flow rate during the initial phase of the injection main pulse. In order to meet mass-injection capabilities, it can be considered both to enlarge the diameter of the A hole and to increase the

p_{rail} = 1250 bar, DT = 750 μs

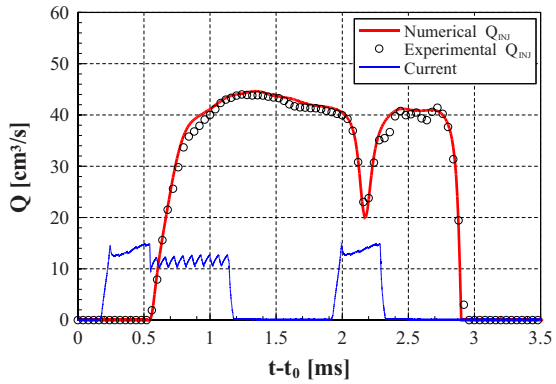


Fig. 28 Reference conditions for IFT reduction study: ET$_{main}$ =1000 μs, ET$_{post}$=400 μs

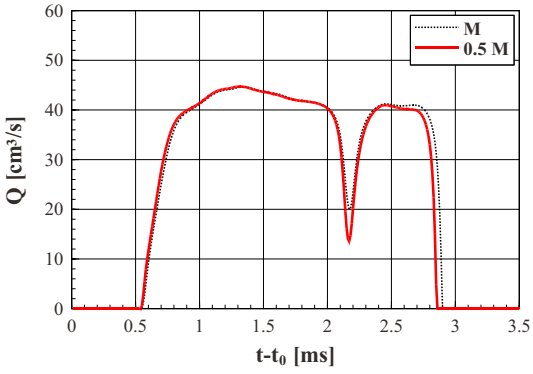


Fig. 29 Needle and control-piston mass effect on IFT

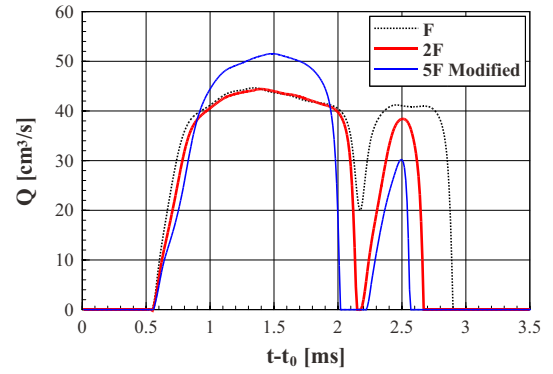


Fig. 31 Needle-spring preload effect on IFT

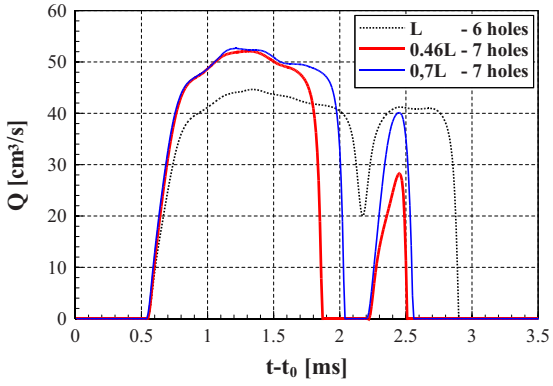


Fig. 32 Needle mechanical stroke-end effect on IFT

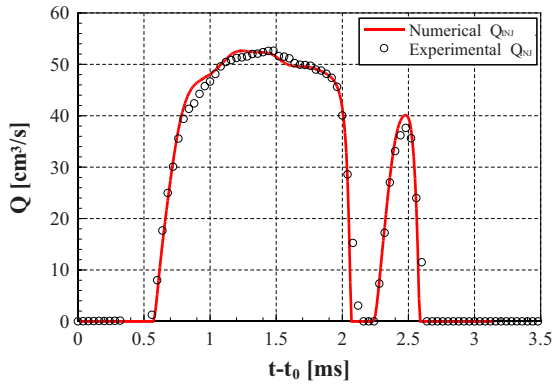


Fig. 33 Numerical versus experimental flow-rate time patterns for the modified injector: p_{rail}=1250 bar, ET_{main}=1000 μs, ET_{post}=400 μs, DT=750 μs

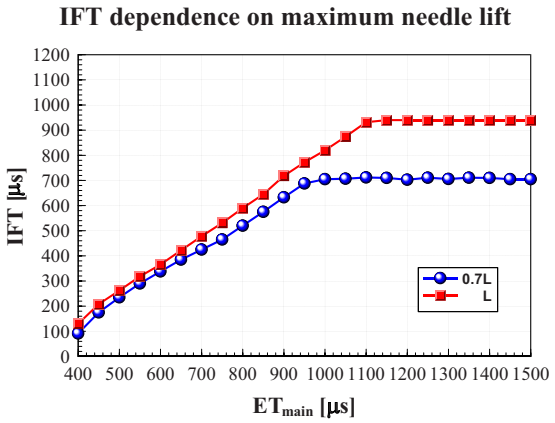


Fig. 34 IFT versus ET_{main} at p_{rail}=1250 bar for two needle mechanical stroke-end limited maximum lifts

nozzle-hole number. Hence, the injection rate that was obtained with a larger A-hole diameter and with a seven-hole nozzle (instead of the six-hole base nozzle) is reported in Fig. 31 for a preload of 5F (Modified in the legend). This substantiates a further reduction of IFT.

Figure 32 presents another simple and effective solution for reducing IFT, without any substantial design modification of the electroinjector. The different injection-rate patterns in the figure were computed by modifying the mechanical stroke-end values of the injector (L, 0.46L, and 0.7L, where L=0.43 mm is the value of the mechanical stroke end for the reference injector). In effect, this value can be easily reduced by replacing the spacer between the control plunger and the nozzle needle (Fig. 1) with a thicker one. The results of Fig. 32 show that shortening the length (L) of the needle mechanically limited stroke can effectively decrease IFT (a reduction of about 400 μs was achieved). In fact, a lower maximum needle lift reduces NCD (as per comments to Fig. 17). It is worth pointing out that such a significant reduction in IFT occurs at large ET_{main} because the peak value of the needle lift tends to increase with the ET.

However, a lower stroke-end confined maximum value of l_n reduces the maximum fuel amount that can be injected. Therefore, a nozzle configuration with seven holes (same diameter as the base one, i.e., d_h=0.14 mm) was taken for numerical tests in order to preserve the fuel injection capability. It should also be stated that nozzle configurations with seven holes, d_h=0.14 mm, are already present in commercial CR systems.

In order to experimentally assess the numerical predictions, a modified version of the multijet electroinjector, featuring a reduced maximum lift (0.7L) and a nozzle with seven holes (d_h =0.14 mm), was built and tested. For such injector configuration, and at the same working conditions of Figs. 28–32, Fig. 33 shows the experimental flow-rate time distributions of the main injection and postinjection and their numerical prediction (this being the same as in Fig. 32 for the case of 0.7L). As can be seen, the accuracy of the model in predicting the system injection performance, when changes in the base injector setup are made, is quite satisfactory. This further supports the present investigation achievements. It is worth noting that the amount of fuel delivered by the injector with reduced maximum lift is higher than the injected quantity by the standard one due to the presence of a larger nozzle-hole number in the modified injector. Hence, it can be expected that a further reduction in the needle stroke-end confined lift should maintain the injection capability of the seven hole injector closer to the reference injector one.

Figure 34 shows the significant experimental IFT reduction that was achieved by decreasing the injector maximum lift from 0.43 mm to 0.3 mm (0.7L). Improved results can be obtained by reducing the needle lift, for example, to 0.2 mm ($\approx$0.5L).

Conclusion

The merging of two consecutive injections from solenoid CR FIS, which can take place when the related current pulses approach each other, was successfully investigated using an integrated experimental-numerical procedure to analyze the cause-effect relationships and solutions for IFT reduction.

The main achievements are as follows:

- The NOD is made up of the time required to partially empty the control chamber to start the needle motion and the time spent by the needle to move upward so as to regain any nozzle axial deformation (the minimum ET that is required to yield needle movement, being shorter than NOD and thus equal only to the first contribution). NOD values are shown to be almost independent of the ET and nominal rail pressure.
- The NCD is also the sum of two contributions: the time required by the needle to invert its motion (needle deceleration phase) once the electric input power to the solenoid is shut off and the period that is necessary to definitively close the nozzle (nozzle closure phase). At constant p_{rail}, NCD is shown to increase linearly with ET, until a constant NCD level is reached for ET larger than a specific value depending on p_{rail}. For a given ET value, the NCD linear function of ET rises with p_{rail}, different from the NCD constant function of ET.
- The ITL, i.e., the temporal interval during which the fuel is injected, is different from ET and is evaluated as ITL=ET +NCD−NOD. Based on the previously described dependence of NOD and NCD on p_{rail} and ET, a correlation was proposed to estimate ITL for different operating conditions.
- When two consecutive injections are fulfilled, the lower limit to the DT value for avoiding injection fusion, i.e., the IFT, is dependent both on the first-injection NCD and on the subsequent-injection NOD. This latter was shown to be shorter than the NOD of a single injection having the same ET and p_{rail} of the second pulse. It is called NRD and is related to IFT by $IFT=NCD_{main}-NRD_{post}$, namely, for a main-post double injection.
- Finally, simple injector modifications were analyzed to find design solutions for IFT reduction. It was shown that either lowering the stroke-end limited needle lift or increasing the needle spring preload can be an effective means of reducing IFT. The former solution tends to lower the maximum fuel amount that can be injected. Hence, it should be combined with a larger nozzle-hole number. The other solution also produces a reduction of the injected flow rate, which can be compensated by nozzles with a larger hole number and by increasing the diameter of the A passage. However, this lat-

ter is a rather critical parameter influencing the injector dynamics. Thus, the reduction of the stroke-end limited needle lift seems to be the best solution to reduce IFT. It can be easily realized by replacing the spacer between the control plunger and the injector needle with a thicker one.

- A modified version of the commercial multijet electroinjector was realized by a plunger-needle spacer thicker than the standard one, so as to reduce the maximum needle lift to 0.3 mm, and by a nozzle with seven holes of the same base diameter, instead of six. It yielded an effective reduction of both NCD and IFT, allowing closer and, thus, more injections without fusion. Better results can be expected by reducing the maximum needle lift to 0.2 mm.

Acknowledgment

Financial support to this research was provided by the Fiat Research Center, FA-GM Powertrain, and MUR (Ministry of University and Research) under projects COFIN04–COFIN06. The authors would like to thank S. Canale, G. Bonetto, F. Guglielmo, and E. Rigon of the Fiat Research Center for their invaluable technical support.

Nomenclature

a	=	sound speed
b, b_1, b_2	=	coefficients
d	=	diameter, internal pipe diameter
F	=	injector-spring preload
k	=	injector-spring stiffness
l	=	lift
L	=	maximum allowable needle lift
$\dot{m}$	=	mass flow rate
M	=	overall mass of injector needle and control plunger
NO_x	=	nitrogen oxides
p	=	pressure
q	=	coefficient
Q	=	volumetric flow rate
r	=	coefficient
S	=	surface
t	=	time
V	=	volume
z	=	coefficient
μ	=	discharge coefficient
ρ	=	density

Subscripts

0	=	reference value
av	=	average
A	=	A hole
cc	=	control chamber
cp	=	control plunger
dc	=	delivery chamber
h	=	nozzle holes
inj	=	injected
inj, in	=	injector inlet
main	=	main injection
n	=	needle
post	=	post injection
pv	=	pilot valve
rail	=	at the rail
Z	=	Z hole

Appendix: Control Chamber Continuity Equation

The control-chamber continuity equation can be written by considering an isentropic flow evolution [18] as follows:

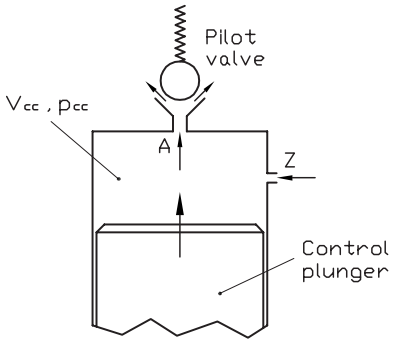


Fig. 35 Control chamber

$$\dot{m}_A + \frac{V_{cc}}{a^2}\frac{dp_{cc}}{dt} = \dot{m}_Z + \rho_{cc}S_{cp}\frac{dl_n}{dt} \tag{A1}$$

where ρ_{cc} is the fuel density in the control chamber, V_{cc} is the control-chamber volume, S_{cp} is the surface area of the control-plunger cross section, a is the liquid isentropic sound speed, whereas $\dot{m}_z$ and $\dot{m}_A$ represent the flow rates through Z and A holes of the injector, respectively (refer to the schematic of the control chamber in Fig. 35). Since the fluid motion through Z and A holes is characterized by high values of the Reynolds number, the related flow rates can be evaluated by means of Torricelli's formula,

$$\dot{m}_Z = \mu_Z\rho_{av_Z}A_Z\sqrt{\frac{2|\Delta p_Z|}{\rho_{av_Z}}}, \quad \dot{m}_A = \mu_A\rho_{av_A}A_A\sqrt{\frac{2|\Delta p_A|}{\rho_{av_A}}} \tag{A2}$$

where ρ_{av} is an average density value in between the densities upstream and downstream from the considered hole and Δp is the pressure drop through the considered hole.

The last term to the right hand side in Eq. (A1) takes account of the decrease in the control-chamber volume due to needle lift. Therefore, until the injector needle is at rest, Eq. (A1) reduces to

$$\frac{dp_{cc}}{dt} = \frac{a^2}{V_{cc}}(\dot{m}_Z - \dot{m}_A) \tag{A3}$$

References

[1] Park, C., Kook, S., and Bae, C., 2004, "Effects of Multiple Injections in a HSDI Diesel Engine Equipped With Common Rail Injection System," SAE Paper No. 2004-01-0127.

[2] Pfahl, U., Hirtler, W., and Cartus, T., 2003, "Passenger Car Investigations of NO_x-Adsorber and DPF Combination to Fulfill Future Diesel Emission Limits," SAE Paper No. 2003-01-0043.

[3] Nehmer, D. A., and Reitz, R. D., 1994, "Measurement of the Effect of Injection Rate and Split Injections on Diesel Engine Soot and NO_x Emissions," SAE Paper No. 940668.

[4] Pierpont, D. A., Montgomery, D. T., and Reitz, R. D., 1995, "Reducing Particulate and NO_x Using Multiple Injections and EGR in a D. I. Diesel," SAE Paper No. 950217.

[5] Chen, S. K., 2000, "Simultaneous Reduction of NO_x and Particulate Emissions by Using Multiple Injections in a Small Diesel Engine," SAE Paper No. 2000-01-3084.

[6] U.S. Department of Energy, 2001, "Homogeneous Charge Compression Ignition (HCCI) Technology," a report to the U.S. Congress.

[7] Foster, D. E., and Najt, P. M., 1983, "Compression-Ignited Homogeneous Charge Combustion," SAE Paper No. 830264.

[8] Gray, A. W., and Ryan, T. W., 1997, "Homogeneous Charge Compression Ignition (HCCI) of Diesel Fuel," SAE Paper No. 971676.

[9] Stanglmaier, R. H., and Roberts, C. E., 1999, "Homogeneous Charge Compression Ignition (HCCI): Benefits, Compromises, and Future Engine Applications," SAE Paper No. 1999-01-3682.

[10] Ando, H., and Kuwahara, K., 2001, "A Keynote on Future Combustion Engines," SAE Paper No. 2001-01-0248.

[11] Helmantel, A., and Denbratt, I., 2004, "HCCI Operation of a Passenger Car Common Rail DI Diesel Engine With Early Injection of Conventional Diesel Fuel," SAE Paper No. 2004-01-0935.

[12] Buchwald, R., Brauer, M., Blechstein, A., Sommer, A., and Kahrstedt, J., 2004, "Adaption of Injection System Parameters to Homogeneous Diesel Combustion," SAE Paper No. 2004-01-0936.

[13] Su, W., Zhang, X., Lin, T., Pei, Y., and Zhao, H., 2004, "Study of Pulse Spray,

Heat Release, Emissions and Efficiencies in a Compound Diesel HCCI Combustion Engine," ASME ICED Fall Conference, Paper No. ICEF2004-927.
[14] Su, W., Wang, H., and Liu, B., 2005, "Injection Model Modulation for HCCI Diesel Combustion," SAE Paper No. 2005-01-0117.
[15] Catania, A. E., Ferrari, A., Manno, M., and Spessa, E., 2005, "Experimental Investigation of Dynamic Effects on Multiple-Injection Common Rail System Performance," ASME J. Eng. Gas Turbines Power, **130**(3), p. 032806.
[16] Catania, A. E., Ferrari, A., and Manno, M., 2005, "Parametric Analysis of Layout Effects on Common Rail Multiple-Injections," ASME ICED Fall Conference, Paper No. ICEF2005-1288.
[17] Baratta, M., Catania, A. E., and Ferrari, A., 2006, "Hydraulic Circuit Design Rules to Eliminate the Dependence of the Injected Fuel on Dwell Time in Multijet Common Rail Systems," ASME J. Fluids Eng., in press.
[18] Catania A. E., Ferrari, A., Manno, M., and Spessa, E., 2006, "A Comprehensive Thermodynamic Approach to Acoustic Cavitation Simulation in High-Pressure Injection Systems by a Conservative Homogeneous Barotropic-Flow Model," ASME J. Eng. Gas Turbines Power, **128**(2), pp. 434–445.
[19] Catania, A. E., Ferrari, A., and Manno, M., 2005, "Development and Application of a Complete Common-Rail Injection System Mathematical Model for Hydrodynamic Analysis and Diagnostics," ASME J. Eng. Gas Turbines Power, **130**(6), p. 062809.
[20] Zielke, W., 1968, "Frequency-Dependent Friction in Transient Pipe Flow," ASME J. Basic Eng., **90**, pp. 109–115.
[21] Kagawa, T., Lee, I. Y., Kitagawa, A., and Takenaka, T., 1983, "High Speed and Accurate Computing Method of Frequency-Dependent Friction in Laminar Pipe Flow for Characteristics Method," Trans. Jpn. Soc. Mech. Eng., Ser. B, **49**, pp. 2638–2644.
[22] Catania, A. E., Ferrari, A., and Spessa, E., 2008, "Temperature Variations in the Simulation of High-Pressure Injection-System Transient Flows Under Cavitation," Int. J. Heat Mass Transfer, **51**, pp. 2090–2107.

Modeling Pressure Fluctuation With Cross Flow in a Tight-Lattice Rod Bundle

Weizhong Zhang[1]
e-mail: zhang.weizhong@jaea.go.jp

Hiroyuki Yoshida
e-mail: yoshida.hiroyuki@jaea.go.jp

Kazuyuki Takase
e-mail: takase.kazuyuki@jaea.go.jp

Thermal and Fluid Engineering Group,
Japan Atomic Energy Agency,
2-4 Shirakata,
Tokai, Naka,
Ibaraki 319-1195, Japan

To explore the mechanism of differential pressure fluctuation inducing cross flow between subchannels in the tight-lattice rod bundle, an evaluation method is presented, which permits the prediction in detail of the unsteady differential pressure fluctuation behavior between subchannels. The instantaneous fluctuation of differential pressure between two subchannels in gas-liquid slug flow regime is deemed as a result of the intermittent nature of slug flow in each subchannel. The method is based on the detailed numerical simulation result of two-phase flow that pressure drop occurs mainly in the liquid slug region and it is, however, negligibly small in the bubble region. The instantaneous fluctuation of differential pressure between two subchannels is associated with pressure gradient in the liquid slug for each channel. In addition to a hydrostatic gradient, acceleration and frictional gradients are taken into account to predict pressure gradient in the liquid slug. This method used in conjunction with the numerical simulation code works satisfactorily to reproduce numerical simulation results for instantaneous fluctuation of differential pressure between two modeled subchannels. It is shown that the static head, acceleration, and frictional pressure drops in the liquid slug are main contributions to the fluctuation of differential pressure between subchannels. [DOI: 10.1115/1.3032392]

Keywords: tight lattice, interface tracking, cross flow, mixing, differential pressure

1 Introduction

An innovative flexible fuel cycle water reactor (FLWR) is under development in Japan Atomic Energy Agency (JAEA). To effectively utilize uranium and plutonium resources, the reactor is required to achieve a high conversion ratio (more than 1.0) and a high discharge burn-up with the plutonium mixed oxide (MOX) fuel [1]. Such a high conversion ratio can be attained by reducing water fraction and consequently lessening moderation of neutrons in the core. Therefore, a tight hexagonal lattice arrangement with the smallest gap clearance of some 1 mm between adjacent fuel rods and high void fraction more than 70% as average are adopted in the design stage of the FLWR core. This makes the cooling situation in the reactor core severe and brings up the question of thermal-hydraulic feasibility of the core. In view of this, JAEA has started a research and development project to investigate thermal-hydraulic performance in tight-lattice rod bundles of the FLWR in collaboration with power companies, reactor suppliers, and universities since 2002 [2].

As known, many constitutive equations are used in the existing subchannel analysis codes. However, their applicability to the tight-lattice arrangement of the FLWR core may be questionable. New experimental data may be needed to validate these equations or to newly develop them for the FLWR. However, in consideration of the large scale of the FLWR core as well as the tight arrangement of fuel rods, it may be economically infeasible to perform full mock-up experiments to investigate all thermal-hydraulic characteristics in the design stage. This leads to a design method called "design-by-analysis" [3]. This method makes the best use of numerical approaches in design. By combining three-field subchannel analysis codes with newly developed constitutive equations and using state-of-the-art two-phase fluid computational dynamics, improved numerical approaches are supposed to be able to replace high cost experiments partially, to benefit from a shorter development period, and to facilitate a wide spectrum of design optimization [4].

Accurate prediction of cross flow rates between subchannels is required for the evaluation of boiling transition in the FLWR core. With use of a statistical method (correlation analysis), the previous study [5] has been revealed that there exists strong correlation between gas/liquid cross flow rates (or mixing coefficients) and differential pressure, and cross flow results mainly from the differential pressure between subchannels. Therefore, as the driving force of cross flow, differential pressure should be evaluated with accuracy for the prediction of cross flow rates. In view of this, this study aims at exploring the mechanism of differential pressure fluctuation between subchannels based on the results obtained by the detailed two-phase-flow numerical simulation. Since cross flow occurs significantly in the slug flow [6], we limit this study to the slug flow regime. In this flow regime, the instantaneous fluctuation of differential pressure between two subchannels may be deemed as a result of the intermittent nature of slug pattern and evaluated by the difference of pressure profile in each subchannel for the tight-lattice rod bundle, a convenient modeling strategy is to consider a unit cell consisting of one Taylor bubble with its surrounding liquid film, plus one subsequent liquid slug in a subchannel. The instantaneous fluctuation of differential pressure between two subchannels is associated with pressure gradient in the liquid slug for each channel. In addition to a hydrostatic gradient, acceleration and frictional gradients are taken into account to predict pressure gradient in the liquid slug. Finally, it will be shown that the way of evaluation can reproduce numerical simulation results of instantaneous differential pressure fluctuation between subchannels.

2 Numerical Simulation

2.1 Numerical Simulation Using Advanced Interface-Tracking Method. A two-phase flow simulation code (TPFIT) with advanced interface tracking method has been developed at JAEA

[1]Corresponding author.

Manuscript received July 22, 2008; final manuscript received August 18, 2008; published online December 30, 2008. Review conducted by Dilip R. Ballal. Paper presented at the Sixteenth International Conference on Nuclear Engineering (ICONE16), May 12–15, 2008, Orlando FL.

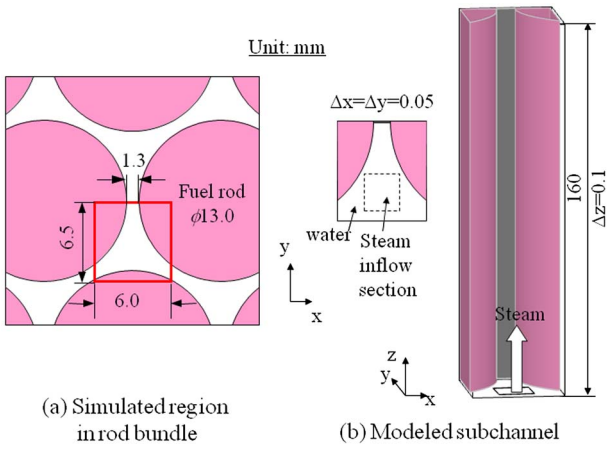


Fig. 1 Modeled single subchannel without cross flow

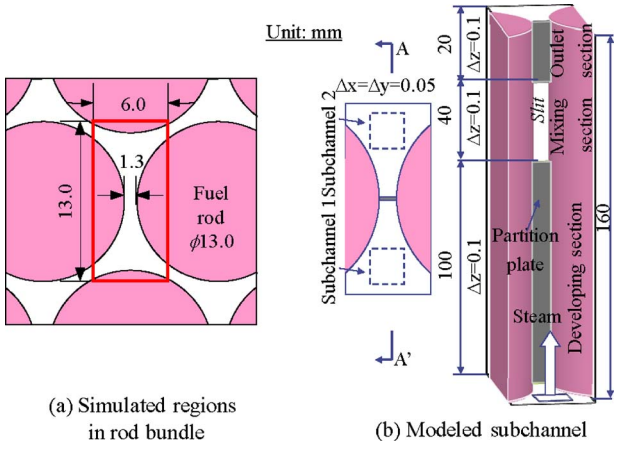


Fig. 2 Modeled two subchannels with cross flow

[3]. The one-fluid model based on the method of volume of fluid (VOF) [7] for the description of two-phase flow was adopted in the TPFIT code. Basis equations consist of mass, momentum, and energy conservative equations for compressible flow, together with transport equations for volume fraction of gas and liquid phases. The momentum equation is solved by a cubic interpolated pseudoparticle (CIP) method [8]. The energy equation is used to obtain the Poisson equation of the static pressure. Temperature is estimated with fluid property routine by the static pressure and local density of both phases. The incomplete lower upper decomposition and conjugate gradient squared method (ILUCGS) is used to solve the Poisson equation of the static pressure. The staggered grid in the Cartesian coordinate system is used as the discretization scheme. Surface tension in the momentum equation is estimated by the continuum surface force (CSF) model [9]. The volume fractions of gas and liquid are calculated by an improved interface tracking method [6] to reduce the numerical diffusion of volume fraction. For more details of numerical methods, please refer to Ref. [6].

2.2 Modeled Subchannels. In this study, the mechanism of pressure drop in a single subchannel without the effect of cross flow will be examined first. On this basis, the model for evaluation of pressure drop in a single subchannel will be built and then be applied to evaluate differential pressure between two subchannels. For the tight-lattice rod bundle the differential pressure may be obtained by subtracting the predictions of pressure profile in each subchannel. The simulated regions in the rod bundle of the FLWR and the modeled subchannels for the numerical simulation are shown in Figs. 1 and 2. The diameter of fuel rods is 13.0 mm. The smallest gap spacing between two adjacent fuel rods is 1.3 mm or 1.0 mm. In the case of the gap spacing of 1.3 mm, the length and width of the simulated region are 13.0 mm and 6.0 mm, respectively. In the case of 1.0 mm, they are 12.2 mm and 5.6 mm, respectively. Figure 2(*b*) describes the modeled two subchannels, subchannel 1 and subchannel 2. The two subchannels are separated by a partition plate with a thickness of 0.2 mm; in the upper part of which, there is a slit with a height of 40 mm and a width of 1.3 or 1.0 mm corresponding to the gap spacing of fuel rods. Water and steam flow into the subchannels through the bottom. Gas and liquid flows pass through three sections along the axial direction, i.e., developing, mixing, and outlet sections. The dimensions are shown in Figs. 1 and 2.

2.3 Simulation Results. The TPFIT code was used to simulate the adiabatic two-phase flow in the modeled subchannels. The grid interval of x and y directions is set to 0.05 mm, i.e., $\Delta x = \Delta y = 0.05$ mm, as shown in Figs. 1 and 2. Δz is set to 0.1 mm. Grid numbers are $120 \times 260 \times 1600$ and $112 \times 244 \times 1600$ for the gap clearances of 1.3 and 1.0, respectively. Saturated water and steam at 7.2 MPa are used as working fluids. All boundary conditions along the flow direction are assigned to be nonslip, and the outlet pressure is fixed to 7.2 MPa. Initial values of void fraction are set to 0.0 in all regions. From the bottom of the simulated region, steam and liquid are injected at constant velocities with values tabulated in Table 1. The flow areas of steam and liquid in the inlet are 5.76 mm^2 and 16.89 mm^2, respectively, in the case of the gap spacing of 1.3 mm, and 4.62 mm^2 and 13.69 mm^2 in the case of the gap spacing of 1.0 mm. The total mass flux is set to 400 kg/m^2 s. The outlet void fraction ranges from 0.24 to 0.44, flow pattern from the bubbly-slug to the slug-churn. Eight cases were simulated. Two cases (A-1 and A-2) simulate the two-phase flow in a single channel without cross flow. Six cases (B-1–B-6) simulate the cross flow through a slit between two subchannels.

Table 1 Simulated case

Case name	Channel type	Gap spacing (mm)	Length of mixing section (mm)	Outlet void fraction (–)	Gas inlet velocity (m/s)	Liquid inlet velocity (m/s)	Flow pattern
A-1	Single channel			0.24	1.00	0.70	bubbly-slug
A-2		1.3		0.35	1.95	0.69	slug
B-1	Two subchannels			0.24	1.00	0.70	bubbly-slug
B-2				0.35	1.95	0.69	slug
B-3		1.3		0.44	2.83	0.67	slug-churn
B-4				0.24	1.00	0.70	bubbly-slug
B-5				0.34	1.95	0.69	slug
B-6		1.0	40	0.43	2.83	0.67	slug-churn

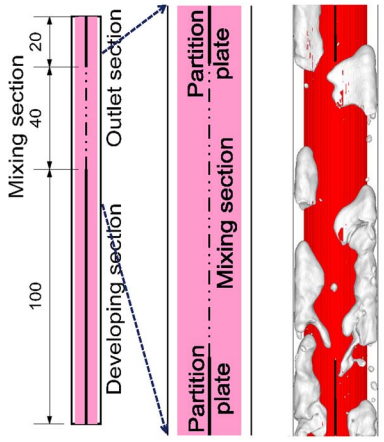


Fig. 3 Simulation results: an example of slug behavior

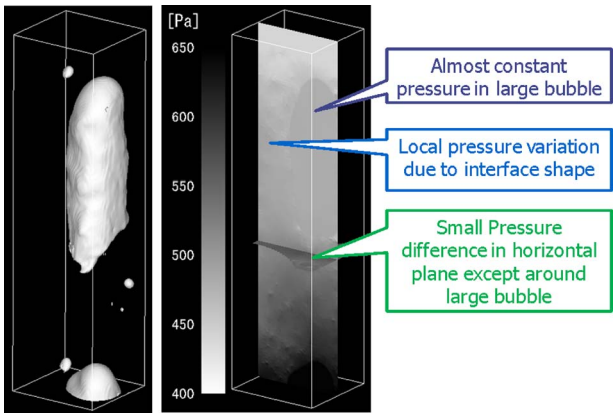


Fig. 5 Detailed pressure distribution in a slug unit

The time step used is 0.5 ms, and the time length of simulations is 1.0 s. Flow parameters, tabulated in Table 1, are the same for two subchannels. Three cases (B-1–B-3) are for the gap spacing of 1.3 mm and three (B-4–B-6) for that of 1.0 mm. For the sake of comparison three cases for the 1.0 mm gap spacing have the same flow parameters as those for the 1.3 mm gap spacing. The character and figure in the case names (see Table 1) stand for the channel type and the case number, respectively. An example of the simulation results observed from section A-A′ indicated in Fig. 2(*b*) is shown in Fig. 3. The liquid-gas interface is contoured at the mesh where the local void fraction is equal to 0.5. White parts represent steam and other parts are full of water. Figure 4 shows the cross-sectional-averaged pressure (solid line) and void fraction (dash line) distributions in a subchannel along the flow direction. The bottom abscissa denotes the value of pressure relative to the outlet of the subchannel where the pressure is fixed to 7.2 MPa as a boundary condition in the numerical simulation. From Fig. 4, it can be observed that the pressure distribution corresponds to that of void fraction in a subchannel, and in the region where the void fraction is low, a high pressure loss occurs. In order to explore the mechanisms of pressure drop, a slug unit is focused on and the pressure distribution in it is shown in detail in Fig. 5. It can be observed that in the large bubble zone, the gray scale changes little and the pressure keeps almost constant, and in the liquid film zone around the bubble, the local pressure varies due to interface shape. It goes without saying that the pressure in the bubble is slightly larger than that in the liquid film due to surface tension. In addition, pressure difference in the horizontal plane is small except around the large bubble. However, below the bubble, it can be observed that the gray scale changes a lot and thus a large pressure gradient exists there. This is reasonable because the flow after the bubble is highly disturbed and vortices may occur there. Moreover, the liquid in the slug may move faster than that in the liquid film around the bubble, acting as a scoop, picking up all the slow moving liquid in the film ahead and accelerating it to slug velocity. By this mechanism, the fast moving liquid builds its volume and becomes a slug [10]. Figure 6 shows the cross-sectional-averaged pressure and void fraction distribution. It is clear that in the bubble region, pressure drop is almost zero, significant pressure drop occurs in the liquid slug, and the largest pressure gradient takes place under the bottom of the bubble.

3 Hydrodynamic Model

In what follows, an approximate model for prediction of differential pressure between subchannels will be presented, following the existing slug flow models for a single channel [10,11]. Based on the above-mentioned results of the numerical simulation, a typical pressure profile through a slug unit at an instant in time is illustrated in Fig. 7. A sharp rise in pressure takes place across the slug front zone. Besides wall friction and gravity, a force is necessary to accelerate the slow moving liquid in the film ahead of the slug to the velocity of the slug. There follows a linear change in pressure due to gradient shear and static head term in the slug core, and almost equivalent to that which would take place in the full pipe flow with no slip between the distributed gas and liquid.

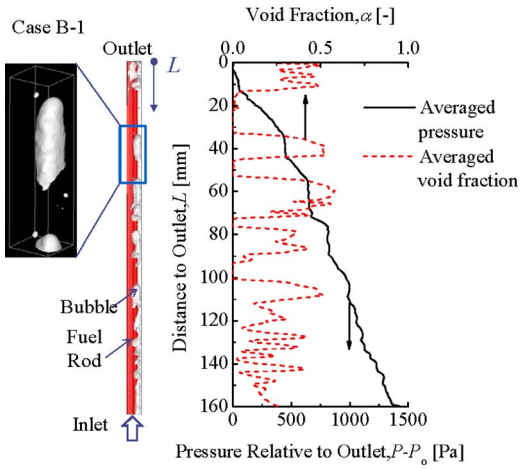


Fig. 4 Axial cross-sectional-averaged pressure and void fraction distributions

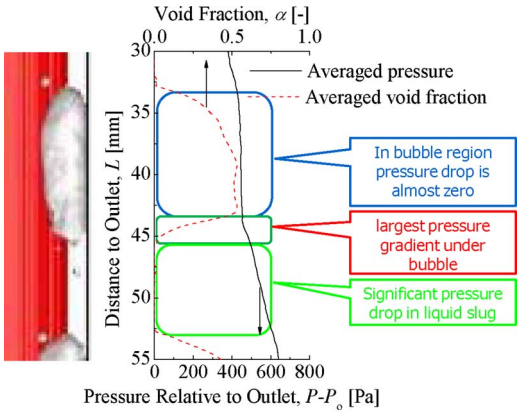


Fig. 6 Cross-sectional-averaged pressure and void distributions in a slug unit

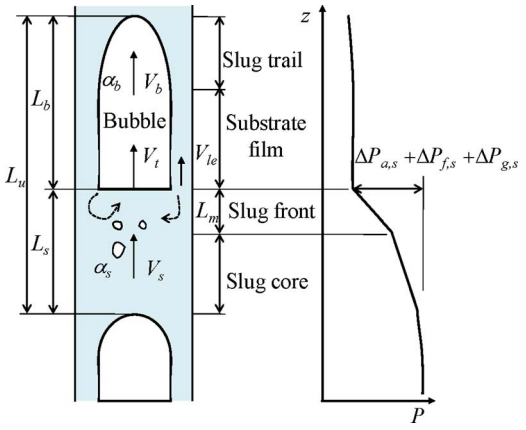


Fig. 7 Schematic of model for evaluation of pressure drop

3.1 Pressure Drop Across a Bubble Zone. Based on the simulation results, the pressure almost keeps constant in this zone. Therefore, pressure drop across a bubble zone is assumed to be zero as follows:

$$\frac{dP_b}{dz} = 0 \tag{1}$$

3.2 Pressure Drop Across a Slug. There are three contributions to the pressure drop across a slug. The first, dP_a, is the pressure drop that results from the acceleration of the slow moving liquid film to the slug velocity in the slug front zone. The second, dP_f, is the pressure drop required to overcome the wall shear. The third, dP_g, is the static head pressure drop. The total pressure gradient across a liquid slug is thus given by

$$\frac{dP_s}{dz} = \frac{dP_{a,s}}{dz} + \frac{dP_{f,s}}{dz} + \frac{dP_{g,s}}{dz} \tag{2}$$

3.2.1 Acceleration Contribution. A slug that has stabilized in length can be considered as a body receiving and losing mass at equal rates. The velocity of the liquid in the film just before pickup is lower than that in the slug, and a force is therefore necessary to accelerate this liquid to the slug velocity. This force manifests itself as a pressure drop. If the pressure along the liquid film in the bubble zone can be assumed to be essentially constant, this force can be evaluated by the sum of gravity and wall shear stress in the liquid film zone [11]. Here the following equations are employed to evaluate the acceleration pressure drop for vertical flow:

$$\Delta P_a = \frac{W_{le}}{A}(V_s - V_{le}) \tag{3}$$

where W_{le} is the rate at which mass is picked up by the liquid slug, which can be obtained from the following equation:

$$W_{le} = \rho_L \cdot A \cdot (1 - \alpha_{le}) \cdot (V_t - V_{le}) \tag{4}$$

where α_{le} and V_{le} are the void fraction and mean velocity of the liquid film in front of the liquid slug. V_t is the propagation velocity of the slug unit or average translational velocity of the nose of the slug given by [11]

$$V_t = 1.2V_s + 0.35\sqrt{gD_b} \tag{5}$$

Here, D_b is the bubble diameter, which may be approximately estimated from the average void fraction of the bubble zone, α_b.

$$D_b = \sqrt{A \cdot \alpha_b} \tag{6}$$

where

$$\alpha_b = \frac{1}{L_b}\int_0^{L_b} \alpha_l dx \tag{7}$$

Here α_l is the local void fraction in the bubble zone. In Eq. (3), the mean velocity of fluid in the slug, V_s, is given as

$$V_s = \frac{1}{A}\left(\frac{W_L}{\rho_L} + \frac{W_G}{\rho_G}\right) \tag{8}$$

The pressure drop due to acceleration takes place in the slug front, which penetrates a distance into the body of the slug. The depth of penetration of the liquid film into the slug appears to depend on the relative velocity between slug and film and may be obtained from the following equation [10]:

$$L_{\mathrm{sf}} = 0.3 \cdot \frac{\rho_L (V_s - V_{le})^2}{2} \tag{9}$$

And the acceleration gradient in the slug front is given by

$$\frac{dP_{a,s}}{dz} = \frac{\Delta P_a}{L_{\mathrm{sf}}} \tag{10}$$

3.2.2 Frictional Contribution. The frictional pressure drop takes place when the liquid slug moves along the channel wall. For the calculation of this term, the similarity analysis for the two-phase frictional pressure drop developed by Dukler et al. [12] is applied. Within the liquid slug the bubble size is usually small and thus the flow can be deemed as the homogeneous one with negligible two-phase slip. Under this condition, the frictional pressure drop can be calculated as follows:

$$\frac{dP_{f,s}}{dz} = \frac{2f_s}{D_h}[\rho_L(1 - \alpha_s) + \rho_G \alpha_s] \cdot V_s^2 \tag{11}$$

For "nonslip" conditions, f_s could be correlated as a unique function of Re_s,

$$f_s = 0.079/\mathrm{Re}_s^{0.25} \tag{12}$$

Here the Reynolds number Re_s is defined in the following manner:

$$\mathrm{Re}_s = D_h V_s \frac{\rho_L(1 - \alpha_s) + \rho_G \alpha_s}{\mu_L(1 - \alpha_s) + \mu_G \alpha_s} \tag{13}$$

When the void fraction α_s is equal to 0, this correlation is identical to the ones for the single phase flow.

3.2.3 Static Head Pressure. The static head term for the slug zone can be obtained from the following simple equation:

$$\frac{dP_{g,s}}{dz} = [\rho_L(1 - \alpha_s) + \rho_G \alpha_s] \cdot g \tag{14}$$

where α_s is the cross-sectional-averaged void fraction in the slug zone.

3.3 Axial Pressure Drop in Question. The axial relative pressure in question, as shown in Fig. 8, could be obtained through the integration of differential pressure as follows:

$$P = \int_0^L \frac{dP}{dz} dz \tag{15}$$

where

$$\frac{dP}{dz} = \begin{cases} \dfrac{dP_s}{dz} & \text{in slug zone} \\ \dfrac{dP_b}{dz} & \text{in bubble zone} \end{cases} \tag{16}$$

Here dP_s/dz and dP_b/dz can be calculated from Eqs. (1) and (2), respectively. Δz is the mesh size of the z direction in the numerical simulation.

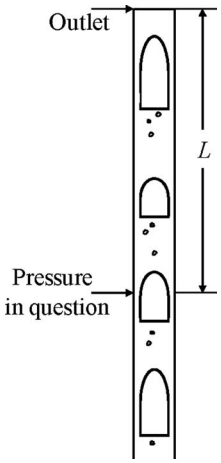


Fig. 8 Calculation of pressure in question

3.4 Local Variables. The above equations are incomplete and some parameters are still unknown, such as the local void fraction in the bubble zone α_l, the mean velocity of the liquid film in front of the liquid slug, the bubble length L_b as well as that of the slug unit L_u, and instantaneous bubble distribution. These parameters may be predicted by assuming the idealized bubble shape and setting the initial bubble distribution along the channel. We leave them for future study. In this study we aim at exploring the mechanism of differential pressure fluctuation inducing cross flow as a first step; therefore numerical simulation results are used to evaluate these unknown parameters for simplicity. In addition, a criterion to reduce the detailed three-dimensional simulation results is introduced to determine locations and lengths of bubble zones (or slug zones). The regions, where the cross-sectional-averaged void fraction is larger than 0.3, are deemed as bubble zones. Other regions are regarded as slug zones.

4 Results and Discussion

In what follows, the evaluation method for the pressure drop shown in this study will be first compared with the simulations results of the pressure drop distribution in a single subchannel without cross flow as well as that in a subchannel with cross flow. After that, it will be used to evaluate the differential pressure fluctuation between subchannels.

4.1 Evaluation of Model With Simulation Results for Single Subchannel. Case A-1 simulates the two-phase flow in a single subchannel without cross flow. Figure 9 shows the evaluation of the model with its simulation results at the time of 0.8100 s. The ordinate denotes the pressure value relative to the value at the outlet of the channel, and the abscissa denotes the distance from the outlet. The solid black line represents the numerical simulation results, and the solid gray line represents the prediction of the model. From the figure, it can be seen that the agreement is satisfactory between the numerical simulation results and the model, except for the channel inlet where the formation of the slug unit is still immature. The static head term, acceleration term, and frictional term are also shown in the figure. For this case, it can be seen that the static head term accounts for a large proportion of the total pressure drop in the slug.

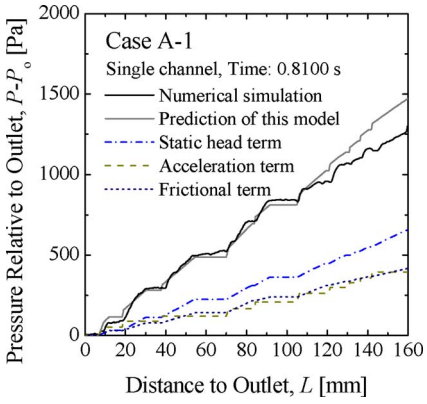


Fig. 9 Evaluation of model with simulation results for single subchannel without cross flow

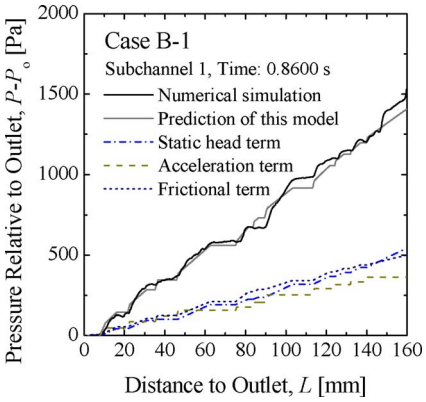


Fig. 10 Evaluation of model with simulation results for subchannels with cross flow

4.2 Evaluation of Model With Simulation Results for Subchannels. In this section, the model will be evaluated with the axial pressure profile obtained from the numerical simulation for a subchannel with cross flow. Figure 10 shows the evaluation results for case B-1 at an instant with a time of 0.8600 s. From the figure, it can be seen that the prediction of the model is generally in agreement with the simulation results, especially for mixing section (abscissa from 20 mm to 60 mm). This means that the model may be applicable to subchannels. Furthermore, we can see that the three contributions to the pressure drop have almost similar weights in the total pressure drop.

4.3 Evaluation of Model With Simulation for Differential Pressure. From the numerical simulation results, differential pressure may be defined as the difference of the cross-sectional-averaged pressure distribution in each subchannel. In this study, the way to evaluate the cross-sectional-averaged pressure distribution in a single subchannel has been developed. For the tight-lattice rod bundle, it is assumed that the cross flow has minor effect on the pressure distribution in each subchannel. Thus we may evaluate the pressure distribution for each subchannel by the developed model first, and then the difference of them may be defined as the differential pressure predicted by this model. In this way, the prediction of the differential pressure by the model is compared with that obtained from the numerical simulation, as shown in Fig. 11. It can be seen that generally the prediction of this model may reproduce the results of numerical simulation.

5 Concluding Remarks

A simple method is presented in this study for the prediction of differential pressures between subchannels based on the existing model for the slug flow in a single channel. The instantaneous fluctuation of differential pressure between the two subchannels is deemed as a result of the intermittent nature of the slug pattern in a subchannel. A unit cell consisting of one Taylor bubble and its surrounding liquid film, plus one subsequent liquid slug in a subchannel, is considered in this model. The instantaneous fluctuation of differential pressure between the two subchannels is associated

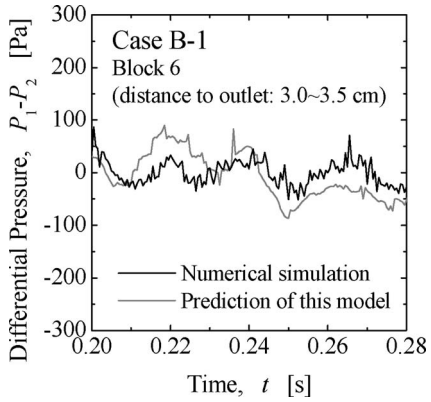


Fig. 11 Evaluation of fluctuation of differential pressure between subchannels

with the pressure gradient in the liquid slug for each channel. In addition to a hydrostatic gradient, acceleration and frictional gradients are taken into account to predict the pressure gradient in the liquid slug. This model may be used to evaluate the instantaneous fluctuation of differential pressure between subchannels provided that the void fraction distribution and bubble shape are given. It is shown that the static head, acceleration, and frictional pressure drops in the liquid slug are main contributions to the fluctuation of differential pressure between subchannels.

Acknowledgment

This research was conducted using a supercomputer of the Japan Atomic Energy Agency.

Nomenclature

A = cross-sectional flow area of the channel, m^2
G = mass flux of the subchannel, kg/m^2 s
f = friction coefficient
g = gravitational acceleration, m/s^2
D_h = hydraulic diameter, m
i,j,k = grid numbers
L = length, m
P = pressure, Pa
Re = Reynolds number
V = velocity, m/s
V_{le} = liquid velocity at the end of the liquid film, m/s
V_s = mean velocity of the fluid in the slug, m/s
V_t = translational velocity at the slug nose, m/s
W = mass flux, kg/s
W_{le} = liquid pickup and shedding rate, kg/s
W_L = total liquid mass flux in the channel, kg/s
W_G = total gas mass flux in the channel, kg/s
x,y,z = coordinates, m
α = void fraction of the steam
μ = dynamic viscosity, N s/m^2
ρ = density, kg/m^3

Subscripts

a = acceleration
b = bubble
f = frictional
g = gravitational
G = gas
l = liquid film zone
le = end of the liquid film
L = liquid
s = slug
sf = slug front
t = total
u = unit

References

[1] Okubo, T., Iwamura, T., Takeda, R., Yamauchi, T., and Okada, H., 2003, "Design Study on Reduced-Moderation Water Reactor (RMWR) Core for Plutonium Multi-Recycling," *Proceedings of GENES4/ANP2003*, Kyoto, Japan, Paper No. 1145.

[2] Ohnuki, A., Takase, K., Kureta, M., Yoshida, H., Tamai, H., Liu, W., Nakatsuka, T., and Akimoto, H., 2005, "Advances of Study on Thermal/Hydraulic Performance in Tight-Lattice Rod Bundles for Reduced-Moderation Water Reactors," *Proceedings of 13th International Conference on Nuclear Engineering*, Beijing, China, Paper No. ICONE13-50100.

[3] Yoshida, H., Nagayoshi, T., Ose, Y., Takase, K., and Akimoto, H., 2004, "Investigation of Water-Vapor Two-Phase Flow Characteristics in a Tight-Lattice Core by Large-Scale Numerical Simulation (I), Development of a Direct Analysis Procedure on Two-Phase Flow With an Advanced Interface Tracking Method," Trans. Atomic Energy Soc. Japan, **3**, pp. 233–241.

[4] Ninokata, H., 2006, "Development of the Advanced Subchannel Analysis Code-Generalized Boiling Transition Model for a Wide Variety of Fuel Bundle Geometry," *Proceedings of 5th Korea-Japan Symposium on Nuclear Thermal Hydraulics and Safety*, Jeju, Korea, Paper No. NTHAS5-PL02.

[5] Zhang, W., Yoshida, H., Ose, Y., Ohnuki, A., Akimoto, H., Hotta, A., and Fujimura, K., 2008, "Numerical Investigation of Cross Flow Phenomena in a Tight-Lattice Rod Bundle Using Advanced Interface Tracking Method," J. Power Energy Syst., **2**, pp. 456–466.

[6] Yoshida, H., Kureta, M., Kitou, K., Takase, K., and Akimoto, H., 2003, "Numerical Simulation of Void Drift Using Interface Tracking Method," 10th International Topical Meeting on Nuclear Reactor Thermal Hydraulics, Seoul, Korea, Paper No. E00208.

[7] Harlow, F. H., and Welch, J. E., 1965, "Numerical Calculation of Time-Dependent Viscous Incompressible Flow of Fluid With Free Surface," Phys. Fluids, **8**, pp. 2182–2189.

[8] Yabe, T., and Aoki, T., 1991, "A Universal Solver for Hyperbolic Equations by Cubic-Polynomial Interpolation," Comput. Phys. Commun., **66**, pp. 233–242.

[9] Brackbill, J. U., Kothe, D. B., and Zemach, C., 1992, "A Continuum Method for Modeling Surface Tension," J. Comput. Phys., **100**, pp. 335–354.

[10] Dukler, A. E., and Hubbard, M. G., 1975, "A Model for Gas-Liquid Slug Flow in Horizontal and Near Horizontal Tubes," Ind. Eng. Chem. Fundam., **14**, pp. 337–347.

[11] Taitel, Y., and Barnea, D., 1990, "Two-Phase Slug Flow," Adv. Heat Transfer, **20**, pp. 83–132.

[12] Dukler, A. E., Wicks, M., and Cleveland, R. G., 1964, "Frictional Pressure Drop in Two-Phase Flow—B. An Approach Through Similarity Analysis," AIChE J., **10**, pp. 44–51.

Development of an Innovative Plate-Type SG for Sodium Cooled Fast Reactor: Concept and Test Fabrication

Yoshihisa Nishi
e-mail: y-nishi@criepi.denken.or.jp

Izumi Kinoshita
e-mail: kinosita@criepi.denken.or.jp

Central Research Institute of Electric Power Industry (CRIEPI),
2-11-1 Iwado Kita, Komae-shi,
Tokyo 201-8511, Japan

The concept of an innovative plate-type steam generator (SG) for a fast reactor, fabricated using hot isostatic pressing (HIP), was presented. The heat-transfer plate, which is assembled with rectangular tubes and fabricated using HIP, is surrounded by a leak-detection layer. The optimum form for the detection layer was determined by crack extension analysis. The concept can be applied in both pool-type and loop-type liquid-metal cooled fast reactors (LMFRs). In this report, the fabrication technique is described as applied to a loop-type LMFR. Optimum HIP conditions of 1423 K (1150°C), 1200 kgf/cm^2(117 MPa), and 3 h for modified 9Cr–1Mo steel were determined from HIP tests, tensile tests, and structural inspection. Nickel-type solder (BNi-5) and gold-type solder (BAu-4) were examined as potential joining materials to laminate the heat-transfer plates. Tensile test comparisons showed that BAu-4 was superior, so it was used. No problems were observed in the fabrication of a partial model of a SG (HIP of the rectangular tubes, brazing the plates, welding the header, etc.)
[DOI: 10.1115/1.3032417]

1 Introduction

In the sodium cooled LMFR, a secondary cooling system and countermeasures for sodium-water reaction accidents were prepared as mitigation systems, with a large impact on LMFR construction costs. A drastic reduction in LMFR construction costs is made possible if these systems are assumed to be unnecessary and a compact SG is developed.

One new concept is to include a third layer between the sodium side and the steam/water side to make the listed countermeasures unnecessary.

If a crack in a tube wall on one side is detected while it remains within the layer, the plant can be stopped before a sodium/water reaction accident can occur. The leak-detection layer might safely replace the equipment currently required to cope with a potential sodium-water reaction in the cooling system. However, a double-wall-tube type SG may present a problem because of the time delay for the leak detection due to a long detection path.

A plate-fin type heat exchanger, usually used for air-conditioning systems (Fig. 1), is known to be compact and efficient. Its leak-detection time might be improved by the inclusion of a leak-detection layer between the cooling flow plates that would securely separate the fluids in the plates with a simple and compact structure that can shorten the detection path.

Reasons why a conventional plate-fin type heat exchanger cannot be used as is for a LMFR are that finned section is not well inspectable at the time of fabrication and that the flow stagnation occurs in narrow section where corrugated plate is in contact with flat plate and thus causes possibility of deposition of corrosive substances in the water/steam flow.

Therefore, a heat exchanger of the plate-fin type without the corrugated plate as a fin was developed using the HIP method (Fig. 2) [1–4].

Manuscript received July 22, 2008; final manuscript received September 16, 2008; published online December 30, 2008. Review conducted by Dilip R. Ballal.

2 The Concept of the Plate Type SG for the LMFR

2.1 For a Pool-Type LMFR. Figure 3 shows a conceptual image of the SG. Plates, with individual channels to carry the feed water in and the steam out, set up at right angles in the hot plenum when it is used in a pool-type reactor. Primary sodium flows down between the plates. Two heat-transfer tube structural concepts (vertical and horizontal tube types) inside the plate are shown in Fig. 4.

2.2 For a Loop-Type LMFR. When applying to the loop-type LMFR (Fig. 5), the structural concept inside a heat-transfer plate becomes simpler. A structure like that in Fig. 1 is thought to be promising when used in a loop-type reactor, as shown in Fig. 5.

The leak-detection layer is constructed by brazing, which produces a bond strength lower than the fracture toughness of the material. Therefore, it can be expected that a crack that reaches the surface will extend along the interface and reach a leak-detection channel before the crack reaches the opposite cooling flow path (Fig. 6). However, to enable detection even if the crack does not extend along the interface, the optimum detection channel pitch was determined by crack extension analysis (Fig. 7). In this report, test fabrication was conducted on the assumption of applying to the loop-type LMFR.

3 Hip Condition

The trial manufacture for the selection of HIP conditions used the modified 9Cr–1Mo steel that is a candidate for fast reactor SG material. The plate for the trial manufacture was assembled from four rectangular heat-transfer pipes, upper and lower flat panels, and square prisms in both sides (Fig. 8). The HIP pressure was constant (1200 kgf/cm^2 (117MPa)), and the HIP temperature was varied from 1173 K (900°C) to 1423 K (1150°C) at intervals of 50 K. Three specimens for the tensile test were set in the HIP furnace in addition to the trial manufacture heat-transfer plate, and the tensile strength of the HIP interface was evaluated after the HIP process.

Figure 9 shows the diffusion bonding process [5], in which bonding progresses via plastic deformation and diffusion, so that

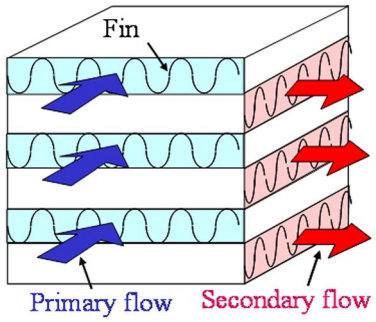


Fig. 1 A plate-fin type heat exchanger

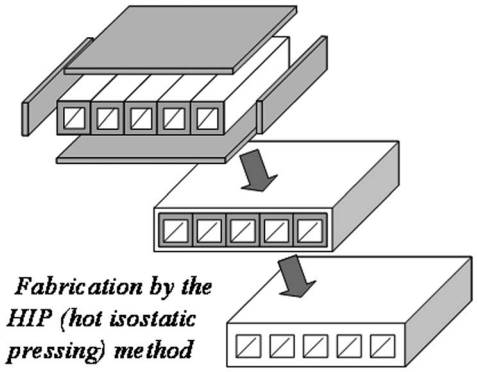


Fig. 2 Production of the plate using HIP

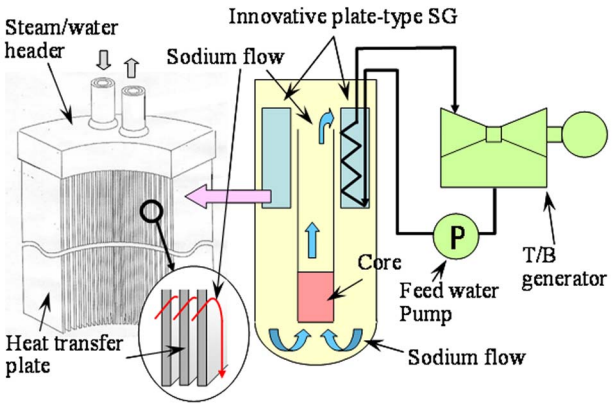


Fig. 3 Conceptual image of the innovative SG for a pool-type reactor

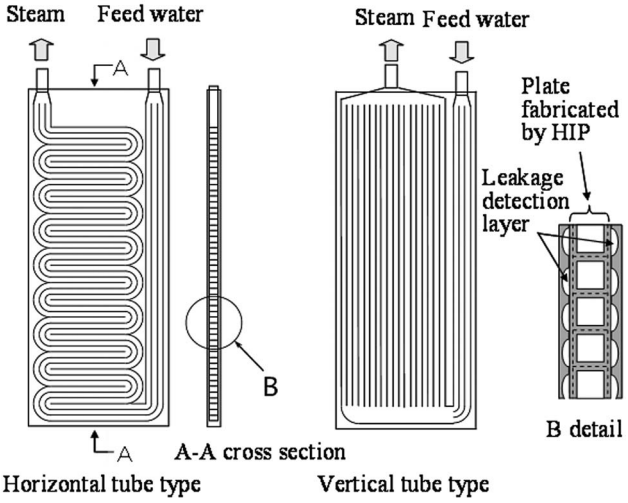
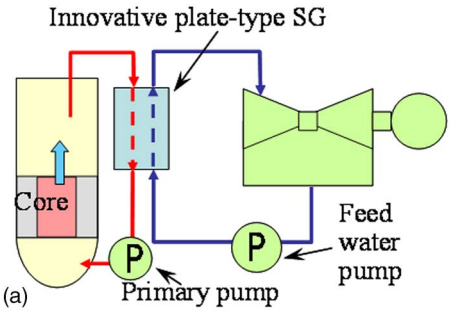


Fig. 4 Heat-transfer tube structural concepts in side the heat-transfer plate for a pool-type LMFR



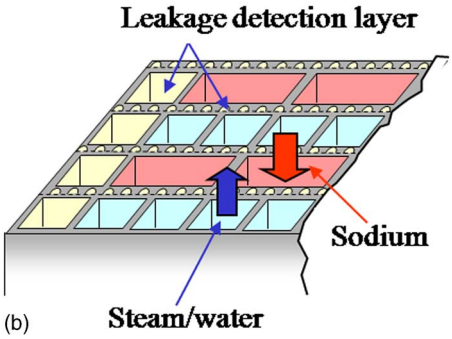


Fig. 5 (*a*) Conceptual image of the innovative SG for a loop-type reactor—cooling system of the loop-type LMFR; (*b*) conceptual image of the innovative SG for a loop-type reactor—schematic of the SG for a loop-type LMFR

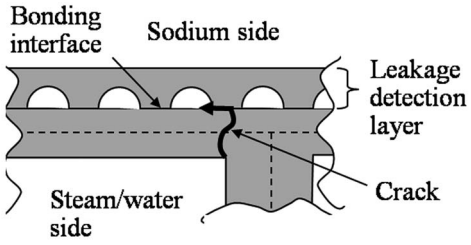


Fig. 6 Leak detection layer

eventually all traces of the junction plane disappears. The void near the bonding interface decreases, the grains of both layers grow and cross the bonding trace, and recrystallization occurs. However, when the HIP bonding temperature is too high, there is a possibility of strength degradation and embrittlement because of the coarsening of the grain.

Figure 10 shows microphotographs of the progress of construction. It can be seen that the trace line of the bonding interface

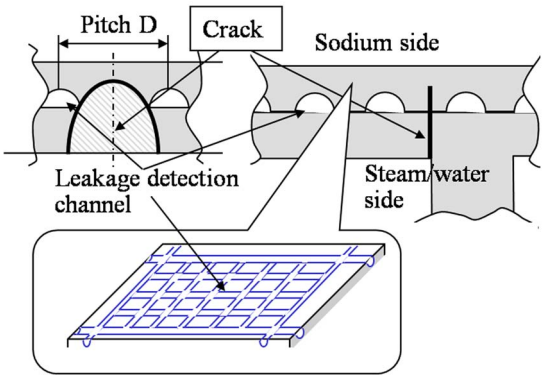


Fig. 7 Pitch of the leak detection groove

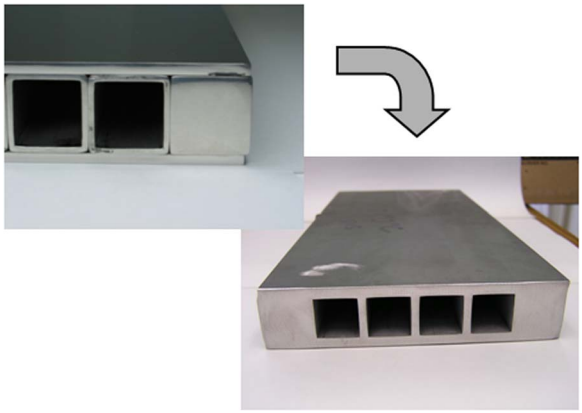


Fig. 8 Photo of the trial fabrication using HIP

thins with increasing temperature and completely disappears at 1423 K (1150°C). No remarkable coarsening of grain structure was observed. Table 1 shows the postannealing results of the tensile testing of the specimen produced by the optimal HIP process at 1423 K (1150°C). To laminate the heat-transfer plate made with HIP, brazing was required, so the thermal effects of brazing should be considered. This table also shows tensile test results that reveal two kinds of thermal effect. The temperatures were determined by the characteristics of the candidate brazing material.

All tensile strengths were within the range of the standard for the materials, and toughness was greater than the standard value. High bonding strength at the HIP-formed interface was confirmed.

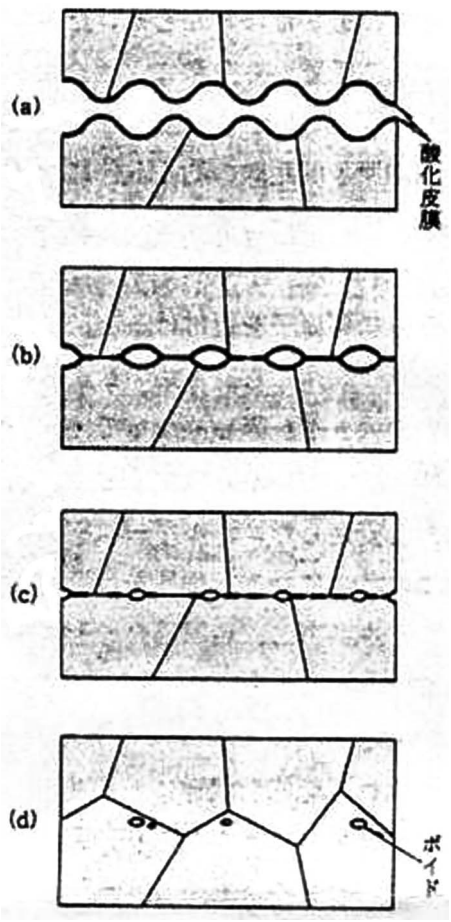


Fig. 9 Image of the diffusion bonding process

4 Brazing Conditions

A postbrazing tensile test was performed on modified 9Cr–1Mo steel for the selection of an optimal brazing material. After the heat-transfer plates are laminated by brazing, it is necessary to weld inlet and outlet headers for each of the fluids onto each of appropriate entrance and exit positions, respectively. Therefore, an interference test with welding was performed.

BAu-4 (82Au–18Ni) and BNi-5 (71Ni–19Cr–10Si) were selected as candidate brazing alloys. BAu-4 has a low melting temperature (1223 K (950°C)) and comparatively high strength. However, BNi-5 is also comparatively high strength and is often used in nuclear power plants because it contains no boron, but the melting temperature of BNi-5 is about 1373 K (1100°C), and it is a matter of concern that this is close to the selected HIP temperature.

The results of the tensile test are shown in Table 2. The lowest tensile strength was 329 MPa with BAu-4 brazing, which was 56% of the strength of the base material, while with BNi-5 brazing the bond was just 17% as strong as the base material. A fracture surface of BAu-4 proved to be rough while the interface of BNi-5 smooth. That confirmed good diffusion bonding of BAu-4 brazing alloy. It was judged that BAu-4 is preferable for its tensile strength.

An interference test was performed by welding on the perimeter of a plate, which has some brazed poles on the plate. The distances between the welding and the brazing were 20 mm, 10 mm, 5 mm, and 0 mm (Fig. 11). The required distance in the brazing and the weld was evaluated from the process of the welding and a color check. Figure 12 shows the test plate after welding. An excellent fillet was shown by the color check and the good condition of the weld was confirmed by visual observation for the BAu-4 solder. However, the brazing alloy was immersed in the weld and blisters and dents were seen in the 0 mm distance case. These defects were not seen where the weld and the brazing were 5 mm or more apart. With BNi-5, no welding defect was found even at 0 mm. It was found that welding with BNi-5 was trouble-free even when the brazing alloy was immersed in the weld.

From the viewpoint of the separation of the welded and brazed parts, it was judged that BNi-5 is preferable. However, in this study, BAu-4 was selected because of its excellent brazing qualities and high bonding strength, on the assumption that the separation will be 5 mm or more in practice.

5 Thermal Stress Evaluation and Sizing of the Leak-Detection Channel

Whether a crack that progresses to the bonding interface extends along the interface or goes straight across the layer containing the other fluid is determined by the stress distribution around the crack, the balance between the stress intensity and the fracture toughness, etc. Therefore, the channel pitch was evaluated by crack extension analysis to allow the certain detection of leakage on one side before the crack penetrates to the other side, even if it goes straight, as shown in Fig. 7.

This SG is assumed to be for use in a medium or small LMFR. To clarify the boundary conditions of the thermal-stress analysis, the thermal design was performed on the assumption of use in a 4S reactor [6]. Figure 13 shows the boundary conditions of the analysis. ABAQUS was used for the thermal stress analysis of the SG. Figure 14 shows the node division for modeling the heat-transfer plate. Three rectangular heat-transfer water/steam-side tubes and the corresponding sodium-side tubes to the water/steam-side tubes were modeled.

To evaluate the uncertainty over the analysis, validation analysis using a detail dividing for the part on which the stress concentrated was performed. As a result, the difference of the maximum stress value is lower than 1%.

The distribution of the σxx stress at A (low-temperature side) and D (high-temperature side) positions are shown in Fig. 15. It

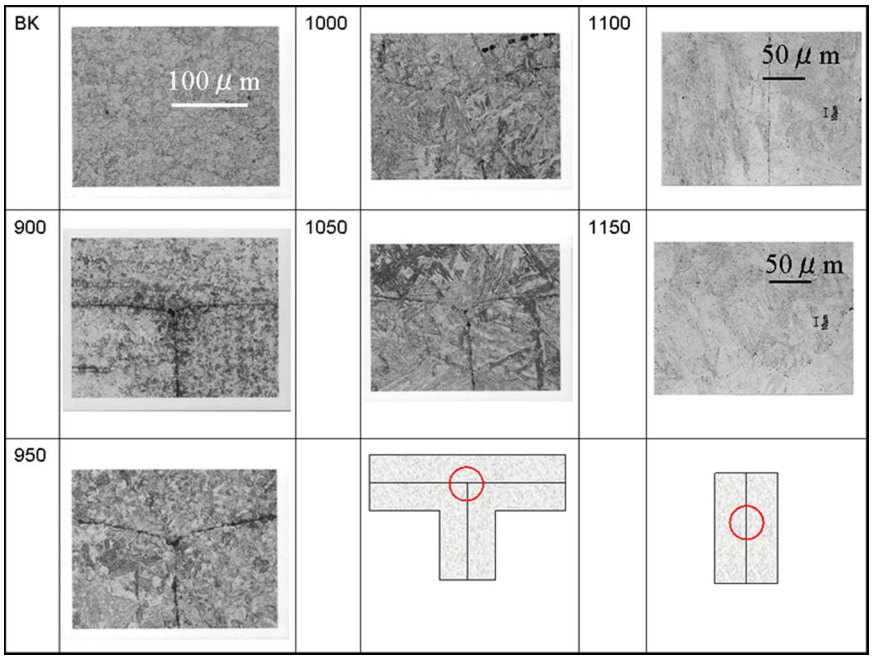


Fig. 10 Microphotos near the interface (parameter: temperature (°C))

Table 1 Results of the tensile test (HIP)

Condition	Strength (N/mm^2)	Expansion (%)	Break position
HIP (1423 K (1150°C), 1200 kgf/cm^2, 3 h)	668	24	Base
	662	21	Base
	659	20	Base
HIP (1423 K (1150°C), 1200 kgf/cm^2, 3 h)+thermal shock (1253 K (980°C), 20 min)	661	20	Base
	640	20	Base
	675	22	Base
HIP (1423 K (1150°C), 1200 kgf/cm^2, 3 h)+thermal shock (1453 K (1160°C), 20 min)	725	19	Base
	724	19	Base
	715	18	Base
Modified 9Cr–1Mo	585–760	18<	

was found that the tensile stress was predominant on the low-temperature side and compressive stress on the high-temperature side in the figure. The reason is that the SG contracts on the low-temperature side, and this contraction is constrained by the high-temperature side.

Figure 16 shows the Von Mises stress over the entire model. The maximum value was 427.2 MPa, and the temperature of the corresponding position was 584 K (311°C) or less. The thermal structure feasibility of the SG at the steady state condition was confirmed by the mechanical properties of modified 9Cr–1Mo steel. The perpendicular stress component on the heat-transfer plate was σyy, and the maximum value was 272.9 MPa. The

Table 2 Results of the tensile test (brazing)

Brazing material	Strength (N/mm^2)	Break position
BAu-4	481	Around the brazed interface
	340	
	329	
BNi-5	165	Fusion line (surface is smooth)
	159	
	102	
Modified 9Cr–1Mo	585–760	

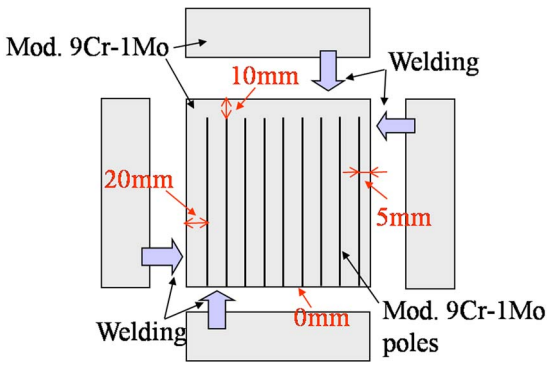


Fig. 11 Component for welding marginal test

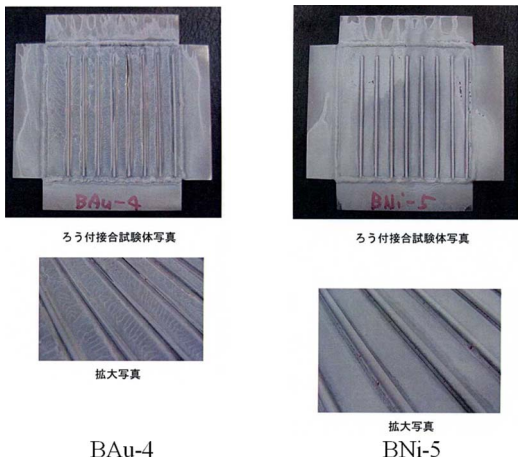


Fig. 12 Photograph of the test plate after welding

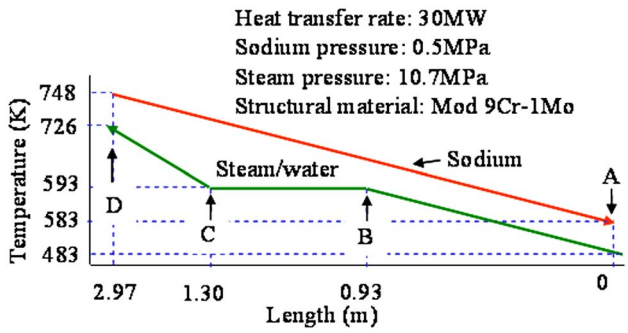


Fig. 13 Boundary condition

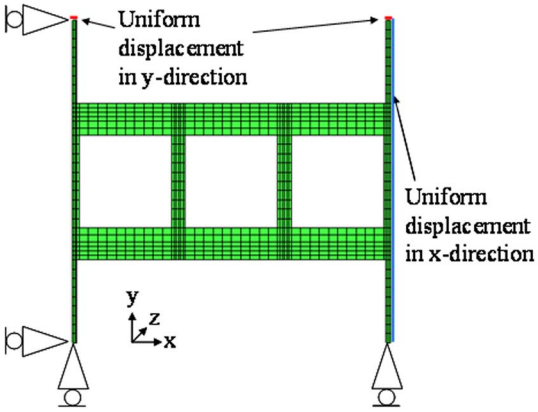


Fig. 14 Node deviation (*X-Y* cross section)

durability of the bonding surface was confirmed from data on the bonding strength of the BAu-4 brazing alloy.

Figure 17 shows the distribution of the maximum stress along three cross cut lines (No. 2, No. 1, No. 3) shown in Fig. 15. The membrane stress and bending stress can be obtained by linear interpolation from this distribution. Fatigue crack growth analyses were performed under the following conditions to consider the startup and stopping of the plant: The peak stress was assumed to be the maximum value obtained by the steady state analysis and the minimum stress was assumed to be zero. The Paris rule was used for crack extension, and the equation the Newman–Raju proposed for flat plate; equation based on JSME maintenance standard and two equations proposed for T-joint were used for stress intensity factor. The channel pitch required so that a straight crack width extends to the channels as it reaches 80% of the distance to the opposite flow path was evaluated.

It was found that the equation of Newman–Raju led to the most conservative result. Leakage can be detected before penetration if there is a leak-detection channel pitch of 11 mm or less, so a leak-detection channel pitch of 10 mm was selected.

6 Test Fabrication

Figure 18 shows the heat-transfer plate made under the selected HIP conditions. Chamfering was performed on the upper and lower surfaces of the plate after HIP, and grooving for leak-detection was applied afterwards.

Hot pressing was used to braze the plate. Hot pressing is a diffusion joining method that uses huge unidirectional pressure in a high-temperature environment. The pressure was changed gradually in preliminary pressure testing and the buckling point was evaluated. The brazing of the plate with hot pressing was done with a pressure of 1.0 kgf/mm^2 after evaluating the results of the preliminary test. Figure 19 shows the finished test body. The proposed SG also requires the attachment of headers for sodium, water/steam, and the leak-detection gas. In the test fabrication, headers were formed from pipes cut lengthways into half cylinders and TIG-welded in place. (Fig. 20). No problems were observed in the header welding process. It was confirmed that the assembling of heat-transfer plates and the welding of the header proceeded well, without crushing the channels in the leak-detection layer, as shown by a ventilation test.

Thus the construction of this innovative plate SG for a loop-type FBR was demonstrated by the test fabrication of a partial model.

7 SG Sizing

To clarify the effectiveness of this innovative plate-type SG, a size comparison was made with a double-wall tube SG. The SG of the 4S reactor (designed in FY 2003) [6] developed by CRIEPI, Tokyo, and the Toshiba Corporation, Tokyo, was selected as the

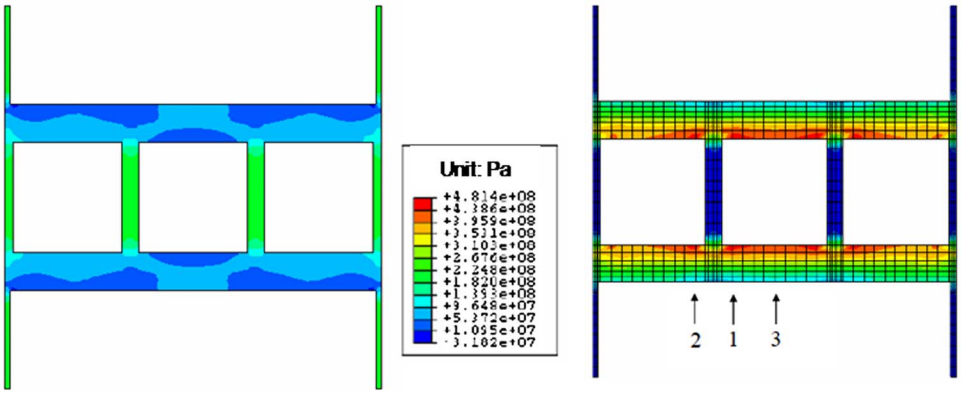


Fig. 15 σxx stress (left side: position D; right side: position A)

Fig. 16 Distribution of the Von Mises stress

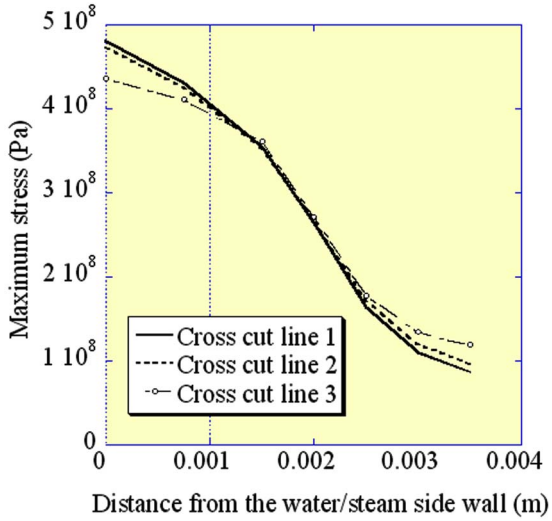


Fig. 17 Distribution of the maximum stress along three cross cut lines at the A position

Fig. 18 The layer of the rectangular tube with leak detection grooves

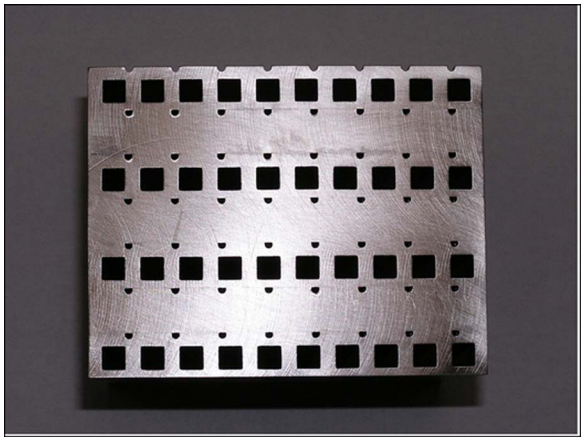


Fig. 19 Test body after brazing and hot pressing

standard. In the sizing of the innovative plate-type SG, equation of Hoe et al. [7] for sodium-side heat convection, Dittus–Boelter's equation [8] for subcooled and superheated regions on the water/steam side, and the equation of Schrock–Grossman [8] for the nucleate boiling/forced convective evaporation region were used. These equations are usually used for a shell and tube type heat exchanger with the circular tube of LMFR. However, because the information of the correlation in a rectangular tube was few, these equations were used in this study.

Figure 21 shows the SG size obtained from the calculation. This innovative plate-type SG appears to require only about 60% of the volume of a double-wall tube SG, although a porosity of 70% is assumed to account for the leak-detection layer.

8 Conclusion

A concept for an innovative plate-type heat exchanger fabricated using HIP is presented as an SG for a LMFR. The concept can be applied in both pool-type and loop-type LMFRs. This

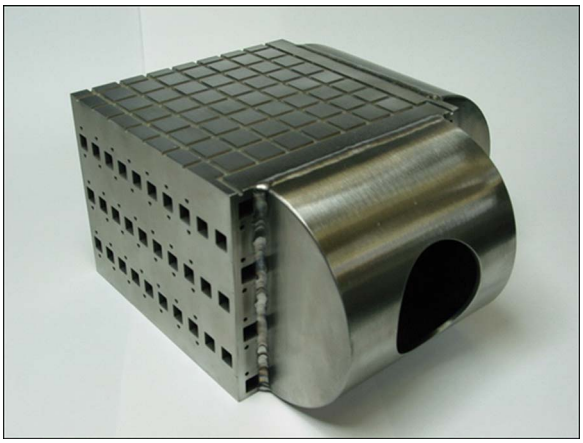


Fig. 20 Partial model of SG

	Helical coil Double wall tube SG	Innovative Plate-type SG
SG cross section	1.16m(diameter)	1mx1m
Heat transfer height	5.54m	3.35m
Porosity of the leakage detection layer	~0%	70%
Volume	5.85m^3	3.36m^3

Heat transfer rate : 30MW
Sodium outlet/inlet temp.
748/583K (475 /310 °C)
Steam pressure
10.7MPa
Steam/water temp.
726/483K (453 / 210 °C)

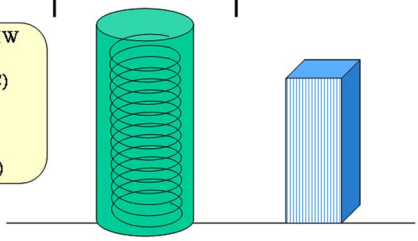


Fig. 21 Results of the size comparison

study focused on the fabrication of the innovative plate-type SG for a loop-type LMFR, and the following conclusions were obtained.

1. Microscopic inspection and tensile tests pointed to optimal conditions of 1423 K (1150°C), 1200 kgf/cm^2, and 3 h for modified 9Cr–1Mo steel.
2. Two candidates were examined as brazing materials to laminate the plates, and BAu-4 was selected on the basis of its excellent joining performance, with the proviso that a margin of 5 mm or more should separate it from welds.
3. Thermal stress analysis and crack extension evaluation were performed on the plate-type SG, showing that the stress in the steady state condition was within the allowable range, and that early leak detection (before penetration to the other channel) is possible with a 10 mm pitch for the leak-detection channels.
4. Hot pressing was used to braze the heat-transfer plate, and the test fabrication of a partial model of the innovative plate-type SG went well. No problem was observed in the welding of headers.

Thus, the practicality of this fabrication method for the innovative plate-type SG has been demonstrated.

References

[1] Kinoshita, I., Nishi, Y., Ueda, N., Takemoto, M., and Maekawa, I., 2002, "Development of Plate Type Heat Exchanger for Steam Generator of Fast Reactors; (1) Concept of SG Structure," J11, *Proceedings of the 2002 Annual Meeting of the Atomic Energy Society of Japan*, Kobe, Japan, Vol. II, p. 500.

[2] Nishi, Y., and Kinoshita, I., 2004, "Development of Plate Type Heat Exchanger for Steam Generator of Fast Reactors; (2) Test Fabrication by HIP," C20, *Proceedings of the 2004 Annual Meeting of the Atomic Energy Society of Japan*, Kyoto, Japan, Vol. II, p. 246.

[3] Nishi, Y., and Kinoshita, I., 2005, "Development of Plate Type Heat Exchanger for Steam Generator of Fast Reactors; (3) Fabrication of Small-Scale Model," 3511, *Proceedings of the Mechanical Engineering Congress 2005 Japan* (MECJ-05), Chofu, Japan, Vol. 3, p. 181.

[4] Nishi, Y., and Kinoshita, I., 2007, "Development of Plate Type Heat Exchanger for Steam Generator of Fast Reactors; (4) Evaluation of the Crack Propagation," 0619, *Proceedings of the Mechanical Engineering Congress 2007 Japan (MECJ-07)*, Osaka, Japan, Vol. 3, p. 159.

[5] Suzumura, A. , and Seto, S., 1993, "The Latest Bonding Technology and the Application," JSME, Nikkan Kogyo Shinbun, Ltd., p. 36.

[6] Ueda, N., Kinoshita, I., Nishi, Y., Minato, A., Yokoyama, T., and Nishiguchi, Y., 2002, "Current Design Status of Sodium Cooled Super-Safe, Small and Simple Reactor," *Proceedings of the Tenth International Conference on Nuclear Engineering* (ICONE10), Arlington, MA, Paper No. ICONE10-22353,

[7] Hoe, R. J., Dropkin, D., and Dwyer, O. E., 1958, "Heat-Transfer Rates to Crossflowing Mercury in a Staggered Tube Bank—II," Trans. ASME, Vol. **79**, pp. 899–907.

[8] Nishikawa, K., and Fujita, Y., 1982, *Heat Transfer*, Rikogakusha, Tokyo, Japan, pp. 171–243.

M. Sharabi

W. Ambrosini
e-mail: walter.ambrosini@ing.unipi.it

Dipartimento di Ingegneria Meccanica Nucleare
e della Produzione,
Università di Pisa,
Via Diotisalvi 2,
56126 Pisa, Italy

S. He
School of Engineering,
University of Aberdeen,
Aberdeen AB24 3UE, UK
e-mail: s.he@abdn.ac.uk

Pei-Xue Jiang
e-mail: jiangpx@mail.tsinghua.edu.cn

Chen-Ru Zhao

Department of Thermal Engineering,
Tsinghua University,
Beijing 100084, China

Transient Three-Dimensional Stability Analysis of Supercritical Water Reactor Rod Bundle Subchannels by a Computatonal Fluid Dynamics Code

The paper describes the application of computational fluid dynamics (CFD) in simulating density wave oscillations in triangular and square pitch rod bundles. The FLUENT code is used for this purpose, addressing typical conditions proposed for supercritical water reactor (SCWR) conceptual design. The RELAP5 code and an in-house 1D linear stability code are also adopted to compare the results for instability thresholds obtained by different techniques. Transient analyses are performed both by the CFD code and RELAP5, with increasing heating rates and constant pressure drop across the channel, up to the occurrence of unstable behavior. The obtained results confirm that the density wave mechanism is similar in rod bundle and in axisymmetric configurations.
[DOI: 10.1115/1.3032437]

1 Introduction

Water at supercritical pressure was used for cooling engineering systems for many decades to increase the thermal efficiency of energy conversion plants. The possibility of using supercritical water as a coolant in nuclear reactors was also considered in the 1950s and 1960s by Westinghouse and General Electric. However, the first conceptual designs were not economically competitive involving a large reactor volume and a complex system design [1]. More recently, there has been a renewed interest in reactors cooled by supercritical water, stimulated by pioneering works at the University of Tokyo started in 1989; this concept acquired worldwide interest and was selected as one of the six Generation IV nuclear energy systems [2].

Fluids at supercritical pressures show dramatic changes in density and transport properties as they cross the pseudocritical temperature, at which a maximum in specific heat at constant pressure is observed. As in boiling water reactors, supercritical water reactors (SCWRs) employ a once-through forced convection cooling system, where the whole coolant supplied to the core by feed water pumps is directly fed to the turbine. The coolant enters the core at a temperature below the pseudocritical one and changes continuously from low temperature high density fluid to low density gaslike fluid at the core outlet [3–6]. In this process, local pressure drops at the inlet and at the outlet, as well as distributed friction, play a considerable role in determining delays between flow perturbations and their effects in terms of total pressure drop across the channel. These delays are at the root of the well-known dynamic instability mechanism identified under the name of "density waves," which occurs in boiling channels and is predicted to occur also in heated channels containing fluids at supercritical pressures, as suggested in early studies by Zuber [7].

Investigations were carried out to study the stability problem in heated channels with supercritical water using linear and nonlinear codes. A linear stability analysis in the frequency domain was carried out by Yi et al. [8] based on single-channel thermal-hydraulics model. Their results showed that the stability of SCWR can be achieved at normal operation and by proper orifice design at channel inlet. Zhao et al. [9] made a nondimensional analysis of the conservation equations in a three-region model assuming a pseudo-two-phase system. The results were represented in stability maps in terms of pseudosubcooling number and expansion number and showing a large margin of stability for the adopted SCWR reference design. Different forms of dimensionless numbers used to define the stability thresholds were proposed in other works [10–12].

The dimensionless numbers developed by Ambrosini and Sharabi [12] (transpseudocritical number and subpseudocritical number) are extrapolations to the supercritical case of the classical dimensionless numbers used in the analysis of stability in boiling channels. These formulations were found to have some generality in the sense that stability thresholds expressed using them do not depend on system pressure or on the working fluid at a considerable extent. Similarities and differences between instabilities in boiling channels and those predicted for heated channels with supercritical fluids were pointed out in Ref. [13], and parametric studies showing the effect of power distribution, inlet and outlet throttling, and hydraulic diameter can be found in Ref. [14].

In a recent work by Sharabi et al. [15], the first stability computational fluid dynamic (CFD) calculations were performed for an axisymmetric circular channel geometry, making use of both wall functions and low-Reynolds number models for dealing with turbulence. In that work, the heated channel was brought to unstable conditions by progressively increasing the heating power, while keeping constant the pressure drop across it.

The work presented in this paper is a further step in this line of research, involving calculations performed with the FLUENT CFD code [16], for 3D slices of fuel bundle subchannels with both triangular and square pitches. The technique adopted to detect

Manuscript received July 29, 2008; final manuscript received July 31, 2008; published online January 5, 2009. Review conducted by Dilip R. Ballal.

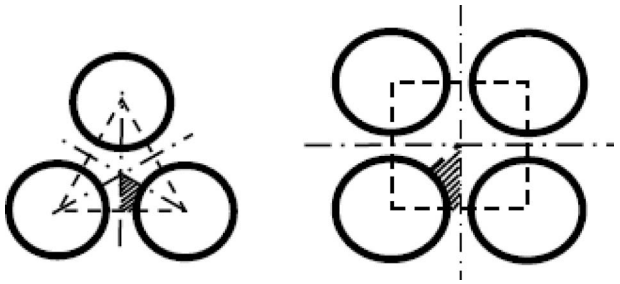


Fig. 1 Subchannels of triangular and square pitch assemblies in SCWRs

unstable behavior is the same as mentioned above and assessed in previous analyses with 1D and CFD codes, involving running transient calculations with increasing heating power in time, up to the occurrence of oscillations. The transient 3D nature of the calculations required the use of efficient computational techniques and appropriate spatial discretization. The obtained results from the CFD model are also compared with those predicted by an in-house linear code based on one-dimensional discretization of the governing equations (see, e.g., Ref. [17]) and the RELAP5/MOD3.3 [18].

2 Physical System

Most of the conceptual designs of SCWRs make use of a thermal neutron spectrum and employ water rods inside the assemblies to provide sufficient neutron moderation, to compensate for the density decrease from inlet to outlet, as mentioned earlier. The fuel assemblies can either be of a square or triangular lattice with water rods. On the other hand, in a supercritical water fast reactor (SWFR) core [19], fuel rods are tightly arranged inside fuel assemblies in a triangular lattice. In this study, both the triangular lattice of the SWFR and the square lattice of the thermal core are investigated; however, water rods are not considered for the sake of simplicity. The geometries addressed are shown in Fig. 1.

Thanks to symmetry, only 1/6 of the triangular pitch subchannel and 1/8 of the square pitch one are considered. In both cases, the system pressure is 25 MPa and water enters the channel at a temperature of 280°C. The following data are adopted from the reference designs [19,20].

- For the square pitch subchannel,
  - rod diameter: 10.2 mm
  - pitch: 11.2 mm
  - active height: 4.2 m
- For the triangular pitch subchannel,
  - rod diameter: 7.6 mm
  - pitch: 8.664 mm
  - active height: 3 m

3 Computational Methods

The FLUENT CFD code [16] is used in the calculations performed in this work. The standard k-ε turbulence model is adopted in the analysis. The wall function approach is used to bridge the viscosity-affected near-wall region with the bulk flow. This approach is preferred in this study to the low-Reynolds number turbulence models, since the latter require employing very fine meshes close to the wall. Although low-Reynolds number models are able to predict the phenomenon of heat transfer deterioration, which is expected to occur at relatively large heat to mass flux (q''/G) ratios [15], such calculations are very time consuming for this unsteady three-dimensional problem. However, as reported in Ref. [15] for stability analysis in circular channels, the predictions of the onset of instability using the two options of turbulence

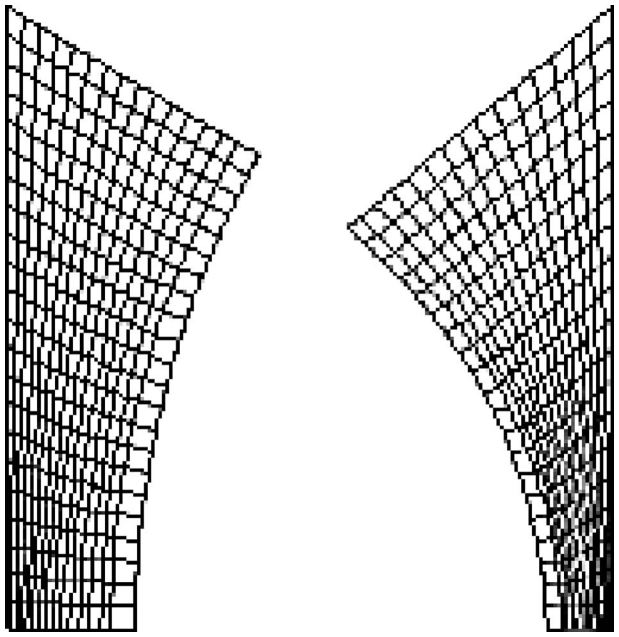


Fig. 2 View of the grid adopted for triangular and square pitch assemblies

modeling was similar. FLUENT offers two choices for wall functions: the standard wall functions and the nonequilibrium wall functions. The latter is suggested for use in complex flows where the assumption of the local equilibrium between turbulent production and dissipation is relaxed [21]. This second option is chosen in the present study.

The two domains appearing as the dashed areas in Fig. 1 are discretized by structured meshes, as shown in Fig. 2. About 300 grid nodes in the transversal directions and 150 in the axial direction were used in the meshing. For all the transient calculations, the value of y^+ is checked to be always in the range between 30 and 500, where the selected wall functions are applicable.

The adopted boundary conditions include the no-slip at the solid walls (the curved sides), symmetry boundary conditions at the other three sides shown in Fig. 2, an inlet fluid temperature of 280°C, and an inlet turbulence intensity of 7%. A constant pressure drop is assigned across the whole channel; local pressure drops at the inlet and the outlet are not considered in the simulations. Though localized pressure drops at the inlet increase the system stability while localized pressure at the outlet decreases it [15], these effects are not considered for the sake of simplifying the addressed problem.

The first order upwind scheme is used for discretizing the convective terms in the governing equations. This is suitable for better numerical stability, compared with higher order schemes, for such complicated three-dimensional conditions with sharply changing fluid properties and transient flow. The PISO algorithm [22] is used, as recommended for transient calculations; the body force weighted discretization was also selected, as suggested for buoyant flows.

Water property tables at a pressure of 25 MPa in the temperature range of interest were generated using the NIST package [23]. A more refined detail is used in the vicinity of the pseudocritical temperature, since the properties change considerably in a very narrow region around this temperature and accurate values of fluid properties are important.

The analysis starts with a steady-state calculation, whose results are used as initial conditions for the transient evolution. The heat flux is then increased linearly at a rate of 100 W/m^2 every 0.1 s starting from 650 kW/m^2, and the overall pressure drop across the channel is kept constant at the obtained steady-state values,

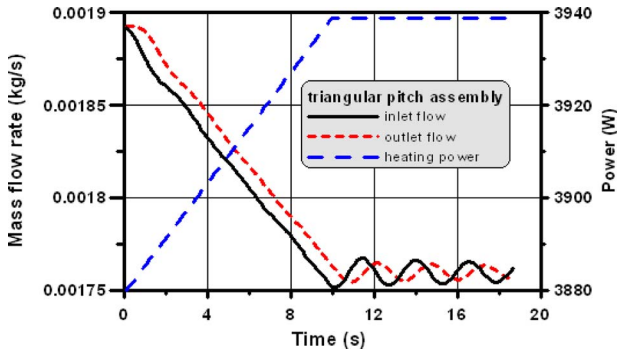


Fig. 3 Transient system response for the triangular pitch assembly obtained by stopping the power increase at 10 s

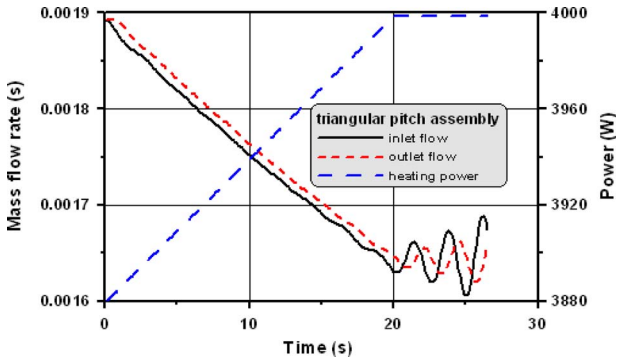


Fig. 4 Transient system response for the triangular pitch assembly obtained by stopping the power increase at 20 s

being 30 kPa and 60 kPa for the square and the triangular pitch assemblies, respectively. These values were adjusted to give flow rate conditions close to the onset of oscillations, thereby saving running time. The time step is kept constant at 0.1 s, and the convergence criteria are chosen to be 6×10^{-5} for the residual of the continuity equation and 10^{-7} for the residuals of all the other equations.

Mesh insensitivity tests were made for the steady-state calculations demonstrating that the results for two extreme cases (one for the lowest mass flow rate at inlet and another one for the highest mass flow rate) are independent of the adopted mesh.

The threshold of instability is evaluated in dimensionless form making use of the subpseudocritical number N_{SPC} and the transpseudocritical number N_{TPC} defined as [12]

$$N_{SPC} = [h_{pc}(p) - h_{in}]\frac{\beta_{pc}(p)}{C_{p,pc}(p)} \tag{1}$$

$$N_{TPC} = \frac{\dot{Q}}{W}\frac{\beta_{pc}(p)}{C_{p,pc}(p)} \tag{2}$$

where the subscript pc refers to properties calculated at the pseudocritical temperature corresponding to the operating pressure, and $\dot{Q}$ and W are the heating power and the channel inlet mass flow rate, respectively. The subpseudocritical number N_{SPC} is constant in all the calculations, being 1.54, corresponding to the selected inlet temperature.

4 Results and Discussions

In Fig. 3, the results of a transient calculation for the triangular pitch assembly are presented. The heating power is increased linearly and, as a consequence, the flow rate decreases due to the increase in friction and acceleration losses resulting from fluid expansion. After 10 s, the heating power is stabilized and kept constant at the reached level. Decaying oscillations with a phase lag between inlet and outlet flows are observed, indicating that the system is still stable in these conditions. However, it is expected that the system is close to the stability threshold, as it can be noted by the small damping of the oscillations, leading to a value of the decay ratio close to unity.

This is confirmed by the results presented in Fig. 4, obtained in a similar analysis but at a higher heating power value. In this case, after the leveling off of the power, the system oscillates with increasing amplitude showing the presence of amplified oscillations. This suggests that flow oscillations start at some time between 10 s and 20 s. In a further calculation (Fig. 5), the power is allowed to increase continuously up to 30 s showing oscillations in counter phase at the inlet and outlet sections. Similar trends are shown also for the square pitch assembly, as illustrated in Fig. 6.

The threshold for the onset of unstable behavior is then identified in terms of the value of N_{TPC} using Eq. (2). As shown in Fig. 7, during the calculation performed with continuous increase in power up to 30 s, the N_{TPC} increases with time up to the point where it starts to oscillate because of flow instability. It can be seen that the triangular pitch assembly has a quite similar behavior, possibly slightly more stable compared with the square pitch assembly. This may be attributed to the distributed pressure drops, which are larger for the more tight triangular pitch assembly (i.e., it has a stabilizing effect). The values of the transpseudocritical number at the threshold of instability are approximately $N_{TPC} \approx 3.8$ and $N_{TPC} \approx 3.6$ for the triangular and square pitches, respectively.

Stability analyses of the two subchannels were also performed making use of the RELAP5/MOD3.3 system code. The calculations are based on the hydraulic diameters of the subchannels (being 3.29 mm and 5.46 mm for the triangular and square pitch assemblies, respectively). Figure 8 illustrates the obtained results in

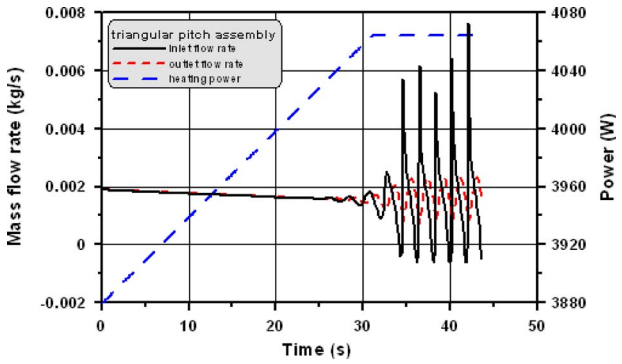


Fig. 5 Transient system response for the triangular pitch assembly obtained by stopping the power increase at 30 s

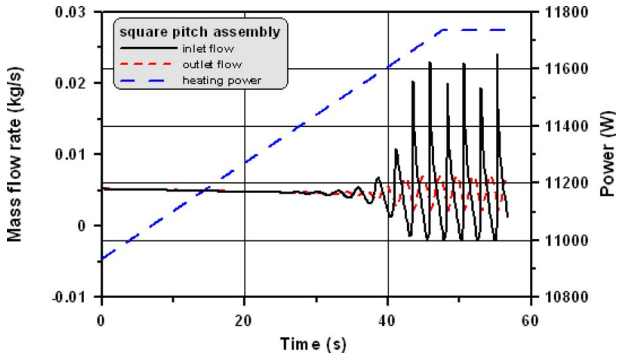


Fig. 6 Transient system response for the square pitch assembly obtained by stopping the power increase at 30 s

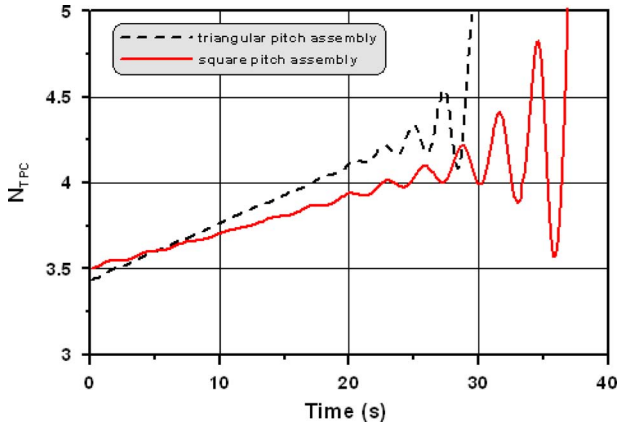


Fig. 7 Onset of oscillations in terms of N_{TPC} for the triangular and square pitch assemblies

terms of N_{TPC}, being about 3.4 and 3.3 at the threshold of instability for the triangular and square pitch assembly, respectively.

Stability maps obtained by the in-house linear stability code for both subchannels are shown in Figs. 9 and 10. Again, a slightly more stable behavior for the triangular pitch assembly is observed with N_{TPC} around 3.2 for N_{SPC} roughly equal to 1.54.

As it is seen from the results predicted by the three codes, notwithstanding the considerable diversity of their models, predictions by the one-dimensional models are in rather acceptable agreement with those of the three-dimensional CFD models, regarding the threshold for the onset of oscillations. Obviously, the analysis by CFD provides more detailed information on the phenomena.

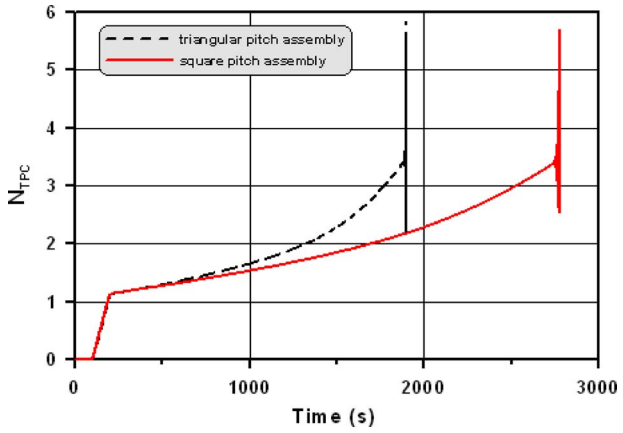


Fig. 8 Onset of instability based on N_{TPC} as predicted by the RELAP/MOD3.3 for the triangular and square pitch assemblies

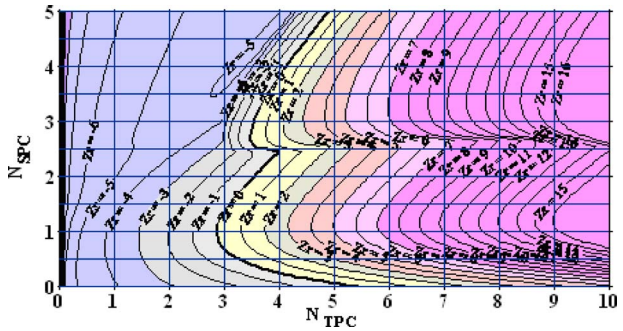


Fig. 9 Stability map for the triangular pitch assembly ($\Lambda=8.0$, Fr=0.15, and maximum Courant number=0.9)

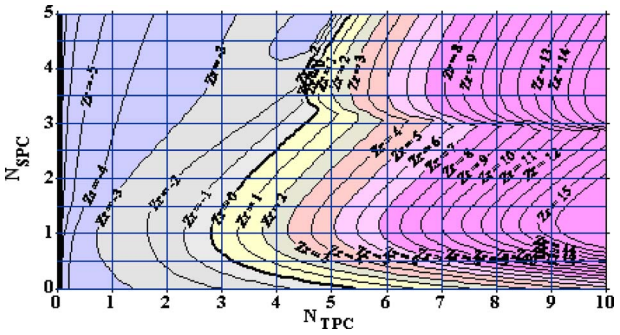


Fig. 10 Stability map for the square pitch assembly ($\Lambda=6.7$, Fr=0.03, and maximum Courant number=0.9)

The axial distributions of the flow rate computed by the CFD models at subsequent positive and negative swings during oscillations are shown in Figs. 11 and 12, for the triangular and square pitch assemblies, respectively. The distributions have similar characteristics as those in boiling channels obtained by one-dimensional nonlinear codes (see, e.g., Ref. [13]).

Concerning the fluid and wall temperatures during oscillations, Figs. 13 and 14 show the contours obtained at the same times selected for showing the flow rate profiles along the channel axis in Figs. 11 and 12. For the triangular pitch assembly, it is seen that

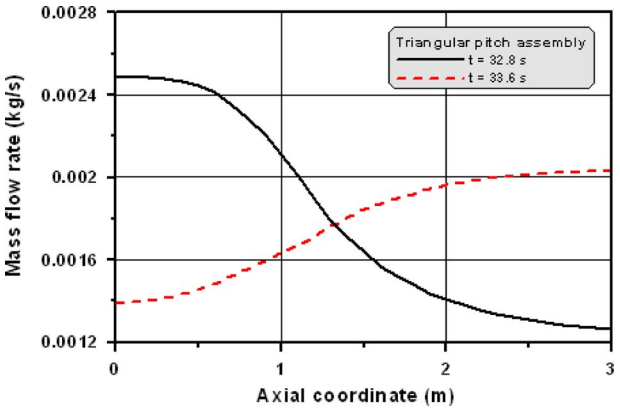


Fig. 11 Mass flow rate distribution along the axial coordinate for the triangular pitch assembly

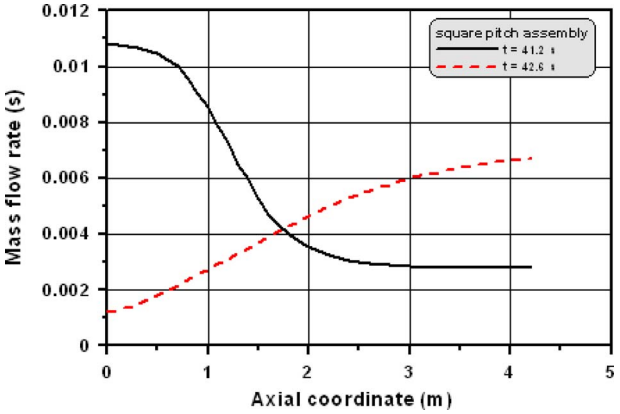


Fig. 12 Mass flow rate distribution along the axial coordinate for the square pitch assembly

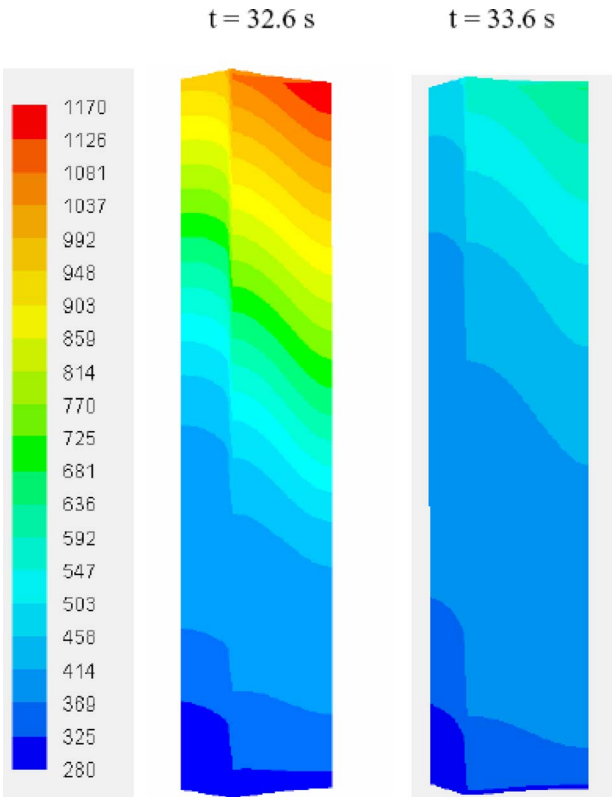


Fig. 13 Temperature contours during oscillations for the triangular pitch assembly

at time 32.6 s the cladding temperature increases to very high levels in the upper half of the subchannel as the flow rate decreases; then, in the subsequent swing, as the flow rate increases again, the cladding temperature decreases back to lower levels. A

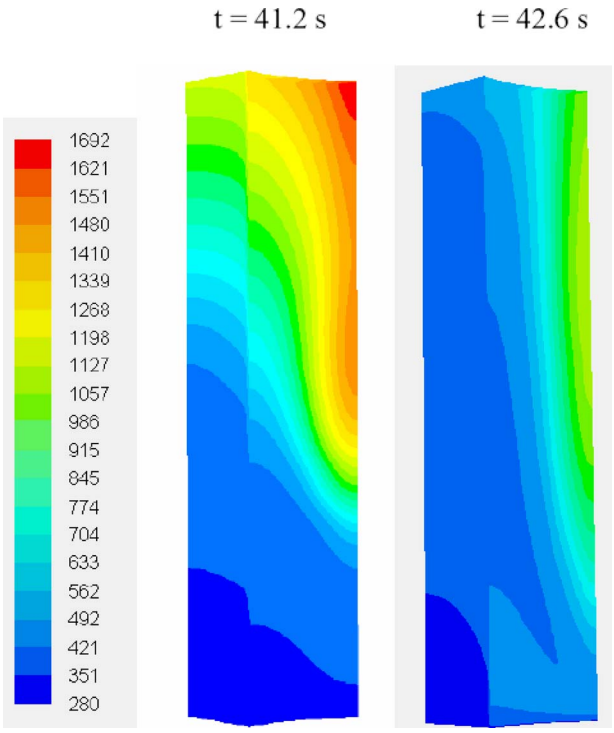


Fig. 14 Temperature contours during oscillations for the square pitch assembly

similar behavior is observed also for the square pitch assembly but with larger temperature gradients along the periphery of the rods; this is mainly due to the nonuniform distribution of coolant in the cross section of the square pitch assembly (see Figs. 1 and 2). These observations on the greater transversal gradients of temperature on the cladding surface in the case of square pitch assembly are in agreement with the results of previous steady-state calculations by Yang et al. [24], who proposed a remedy for this problem by incorporating a spacer structure to make the flow more uniform in the cross-sectional area. However, this idea needs further analysis before considered effective.

In a previous work [15], using low-Reynolds number turbulence models, it was found that the heat transfer deterioration phenomenon can occur before the start of the oscillations, and it has an effect on both the characteristic of the oscillations (amplitudes and periods) and the distribution of wall temperature. Hence, it is expected that during the oscillations the wall temperatures can even exceed those predicted by the wall function models adopted here.

5 Conclusions

The results presented in this paper represent a further step in the analysis of heated channels cooled by supercritical water using CFD models. Whereas the previous work was mainly devoted to circular tubes [15], rod bundle geometry was addressed here, although inlet and outlet throttling were neglected. The use of wall functions to represent the near-wall region may constitute an additional limitation of the analysis performed in the present study.

Though for the above reasons, the obtained data are not directly applicable to the actual core conditions, nevertheless they provide additional information on stability mechanisms in triangular and square pitch rod bundles. In particular, it is shown that density wave oscillations in these geometries have similar characteristics as in circular channels. Moreover, the reasonable match between the stability thresholds evaluated by CFD and 1D models suggests that the density wave mechanism is depicted nearly in the same way by both models.

Nomenclature

Roman Letters

C_p = specific heat (J/(kg K))
D = diameter (m)
f = friction factor
Fr = Froude number, $w_{in}^2/(gL)$
g = gravity (m/s^2)
G = mass velocity ($kg/(m^2\ s)$)
h = specific enthalpy (J/kg)
L = subchannel length (m)
N_{SPC} = subpseudocritical number
N_{TPC} = transpseudocritical number
p = pressure (Pa)
q'' = heat flux (W/m^2)
$\dot{Q}$ = heating power (W)
w = fluid velocity (m/s)
W = mass flow rate (kg/s)
Zr = dimensionless amplification exponent

Greek Letters

β = isobaric thermal expansion coefficient (1/K)
Λ = friction parameter=$fL/(2D)$

Subscripts

in = inlet
pc = value calculated at the pseudocritical temperature

References

[1] Squarer, D., Schulenberg, T., Struwe, D., Oka, Y., Bittermann, D., Aksan, N., Maraczy, C., Kyrki-Rajamäki, R., Souyri, A., and Dumaz, P., 2003, "High Performance Light Water Reactor," Nucl. Eng. Des., **221**(1–3), pp. 167–180.

[2] Yamada, K., Ookawa, M., Asanuma, Y., Sakurai, S., Kitou, K., and Oka, Y., 2007, "Recent Activities and Future Plan of Thermal-Spectrum SCWR Development in Japan," Third International Symposium on SCWR—Design and Technology, Shanghai, China.

[3] Dobashi, K., Oka, Y., and Koshizuka, S., 1998: "Conceptual Design of a High Temperature Power Reactor Cooled and Moderated by Supercritical Light Water," *Proceedings of the Sixth International Conference on Nuclear Engineering, ICONE6*, ASME, New York.

[4] Heusener, G., Mueller, U., and Squarer, D., 2000, "High Performance Light Water Reactor (HPLWR)," Nucl. Europe Worldscan, **XX**(1/2), pp. 59–60.

[5] Heusener, G., Mueller, U., Schulenberg, T., and Squarer, D., 2000, "European Development Program for a High Performance Light Water Reactor (HPLWR)," SCR-2000, The University of Tokyo, Tokyo, Japan.

[6] Oka, Y., 2000, "Review of High Temperature Water and Steam Cooled Reactor Concepts" *Proceedings of the First International Symposium on Supercritical Water-Cooled Reactors*, Design and Technology, SCR-2000, The University of Tokyo, Tokyo, Japan.

[7] Zuber, N., 1966, "An Analysis of Thermally Induced Flow Oscillations in the Near-Critical and Super-Critical Thermodynamic Region," Research and Development Center, General Electric Company, Schenectady, NY, Report No. NASA-CR-80609.

[8] Yi, T. T., Koshizuka, S., and Oka, Y., 2004, "A Linear Stability Analysis of Supercritical Water Reactors, (I) Thermal-Hydraulic Stability," J. Nucl. Sci. Technol., **41**(12), pp. 1166–1175.

[9] Zhao, J., Saha, P., and Kazimi, M., 2005, "Stability of Supercritical Water-Cooled Reactor During Steady-State and Sliding Pressure Start-Up," 11th International Topical Meeting on Nuclear Reactor Thermal-Hydraulics (NURETH-11), Popes' Palace Conference Center, Avignon, France, Paper No. 106.

[10] Chatoorgoon, V., Voodi, A., and Fraser, D., 2005, "The Stability Boundary for Supercritical Flow in Natural Convection Loops, Part I: H_2O Studies," Nucl. Eng. Des., **235**, pp. 2570–2580.

[11] Ortega, T. G., Class, A., Lahey, R. T., and Schulenberg, T., 2006, "Stability Analysis of a Uniformly Heated Channel With Supercritical Water," 14th International Conference on Nuclear Engineering (ICONE 14–89733), Miami, FL.

[12] Ambrosini, W., and Sharabi, M. B., 2006, "Dimensionless Parameters in Stability Analysis of Heated Channels With Fluids at Supercritical Pressures," 14th International Conference on Nuclear Engineering (ICONE 14–89862), Miami, FL.

[13] Ambrosini, W., 2007, "On the Analogies in the Dynamic Behaviour of Heated Channels With Boiling and Supercritical Fluids," Nucl. Eng. Des., **237**(11), pp. 1164–1174.

[14] Ambrosini, W., and Sharabi, M. B., 2007, "Assessment of Stability Maps for Heated Channels With Supercritical Fluids Versus the Predictions of a System Code," Nuclear Eng. and Technol., **3**(5), pp. 591–682.

[15] Sharabi, M. B., Ambrosini, W., and He, S., 2008, "Prediction of Unstable Behaviour in a Heated Channel With Water at Supercritical Pressure by CFD Models," Ann. Nucl. Energy, **35**(5), pp. 767–782.

[16] Fluent Inc., 2004, "Fluent User Guide."

[17] Ambrosini, W., Di Marco, P., and Susanek, A., 1999, "Prediction of the Boiling Channel Stability by a Finite-Difference Numerical Method," Second International Symposium On Two-Phase Flow Modelling and Experimentation, Pisa, Italy.

[18] SCIENTECH Inc., 1999, "RELAP5/MOD3.3 Code Manual, Code Structure, System Models and Solution Methods," The Thermal Hydraulics Group, Idaho, Vol. I, June.

[19] Yoo, J., Ishiwatari, Y., Oka, Y., and Liu, J., 2006, "Conceptual Design of Compact Supercritical Water-Cooled Fast Reactor With Thermal Hydraulic Coupling," Ann. Nucl. Energy, **33**, pp. 945–956.

[20] Yamaji, A., Oka, Y., and Koshizuka, S., 2001, "Conceptual Core Design of a 1000 MWe Supercritical Pressure Light Water Cooled and Moderated Reactor," ANS/HPS Student Conference, A&M University, Texas.

[21] Kim, S. E., and Choudhury, D., 1995, "A Near-Wall Treatment Using Wall Functions Sensitized to Pressure Gradient," *Separated and Complex Flows*, ASME, New York, Vol. FED-217, pp. 273–280.

[22] Issa, R. I., 1986, "Solution of Implicitly Discretized Fluid Flow Equations by Operator-Splitting," J. Comput. Phys., **62**, pp. 40–65.

[23] Lemmon, E. W., McLinden, M. O., and Hurber, M. L., eds., 2002, *NIST Reference Fluid Thermodynamic and Transport Properties—REFPPRO*, Aug., NIST Standard Reference Database 23 (Software and Source), V. 7.0, U.S. Department of Commerce.

[24] Yang, J., Oka, Y., Ishiwatari, Y., Liu, J., and Yoo, J., 2007, "Numerical Investigation of Heat Transfer in Upward Flows of Supercritical Water in Circular Tubes and Tight Fuel Rod Bundles," Nucl. Eng. Des., **237**, pp. 420–430.

Weld Material Investigations of a WWER-440 Reactor Pressure Vessel: Results From the First Trepan Taken From the Former Greifswald NPP

Udo Rindelhardt
e-mail: u.rindelhardt@fzd.de

Hans-Werner Viehrig
e-mail: h.w.viehrig@fzd.de

Joerg Konheiser
e-mail: j.konheiser@fzd.de

Jan Schuhknecht
e-mail: j.schuhknecht@fzd.de

Forschungszentrum Dresden-Rossendorf (FZD),
PF 510119,
D-01324 Dresden, Germany

Between 1973 and 1990 four units of the Russian nuclear power plants type WWER-440/230 were operated in Greifswald (former East Germany). Material probes from the pressure vessels were gained in the frame of the ongoing decommissioning procedure. The investigations of this material started with material from the circumferential core weld of unit 1. First, this paper presents results of the reactor pressure vessel (RPV) fluence calculations depending on different loading schemes and on the axial weld position based on the Monte Carlo code TRAMO. The results show that the use of the dummy assemblies reduces the flux by a factor of 2–5 depending on the azimuthal position. The circumferential core weld (SN0.1.4) received a fluence of 2.4×10^{19} neutrons/cm^2 at the inner surface; it decreases to 0.8×10^{19} neutrons/cm^2 at the outer surface. The material investigations were done using a trepan from the circumferential core weld. The reference temperature T_0 was calculated with the measured fracture toughness values, K_{Jc}, at brittle failure of the specimen. The K_{Jc} values show a remarkable scatter. The highest T_0 was about 50°C at a distance of 22 mm from the inner surface of the weld. The Charpy transition temperature TT_{41J} estimated with results of subsized specimens after the recovery annealing was confirmed by the testing of standard Charpy V-notch specimens. The VERLIFE procedure prepared for the integrity assessment of WWER RPV was applied on the measured results. The VERLIFE lower bound curve indexed with the Structural Integrity Assessment Procedures for European Industry (SINTAP) reference temperature, RTT_0^{SINTAP}, envelops the K_{Jc} values. Therefore for a conservative integrity assessment the fracture toughness curve indexed with a RT representing the brittle fraction of a data set of measured K_{Jc} values has to be applied. [DOI: 10.1115/1.3032461]

1 Introduction

The integrity of the reactor pressure vessel (RPV) in nuclear power plants (NPPs) is nowadays checked by surveillance probes, which are usually inserted in the RPV at the beginning of the plant operation. The lead factor of this probes is typically clearly higher than 1, so that the neutron fluence of these probes reaches the designed RPV end of life fluence after a comparable short time (few cycles). However, it should be noted that the flux and the temperature at the surveillance positions mostly differ from the real RPV conditions.

Therefore, the investigation of original RPV material after the end of a NPP operation is useful to check and to validate this approach. Such an opportunity is given now by investigating material from the former Greifswald NPP (WWER-440/230; start of four units between 1973 and 1978; shutdown of all units in 1990). The well documented different irradiation/annealing states of the four Greifswald RPVs allow a very comprehensive approach to study embrittlement phenomena. This approach includes the following:

- neutron and gamma fluence calculations
- Nb based experimental dosimetry
- a comprehensive material investigation program

The procedure to get the trepans from the RPV had to be included in the general decommissioning course; therefore the first trepans (diameter of 120 mm) were gained in the late autumn of 2005 from unit 1 of this NPP. The axial positions of the trepans related to the main features of the RPV are shown in Fig. 1. Details of the trepanning procedure are given in Ref. [1]; first investigation results were presented in Ref. [2].

This paper presents results of the RPV fluence depending on different loading schemes and on the axial position based on Monte Carlo calculations. Further new results of material investigations based on the master curve (MC) concept are given.

2 Neutron Fluence and Core Loading

The high neutron flux in the RPV is a well known problem for the Russian WWER-440-type NPP. During the development and design of the first reactors, the neutron embrittlement of the steel was considered to be negligible, therefore no surveillance assemblies were originally foreseen. The operation of the first units showed however [3], that a fast embrittlement process was going on. Two causes contribute to this phenomenon. On one hand, the WWER-440 RPV is subject to a much higher neutron flux than comparable western reactor types due to its compact construction. On the other hand, steel impurities (especial copper and phosphor) accelerate the embrittlement of the circumferential welds.

The second mechanism could be avoided in later versions of this reactor (WWER-440/213), and in case of the older units a flux reduction through the use of dummy assemblies (fuel elements

Manuscript received July 29, 2008; final manuscript received August 6, 2008; published online January 5, 2009. Review conducted by Dilip R. Ballal. Paper presented at the Sixteenth International Conference on Nuclear Engineering (ICONE 16), May 12–15, 2008, Orlando, FL.

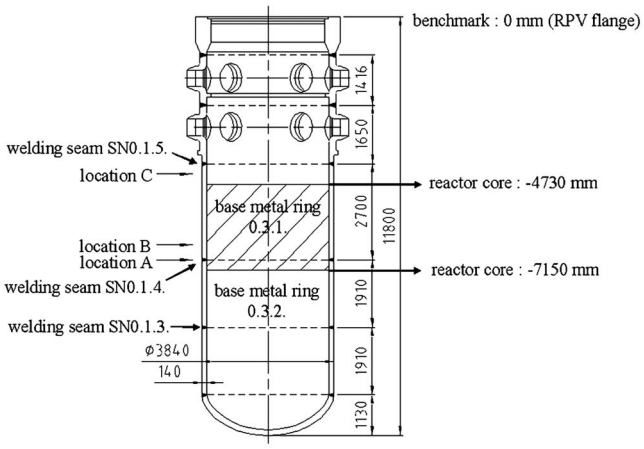


Fig. 1 WWER-440 reactor with trepan positions

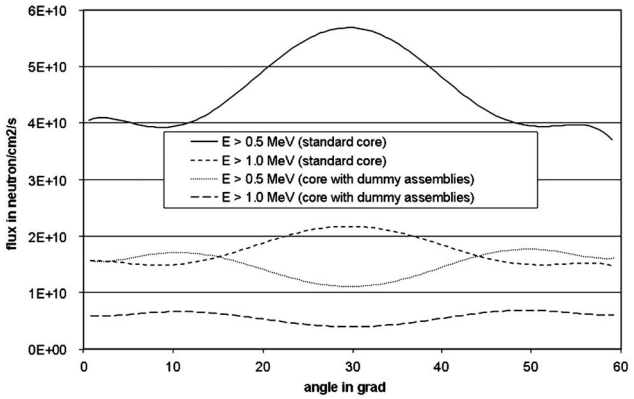


Fig. 3 Flux distribution $E>0.5$ MeV and 1.0 MeV outside of the RPV (110 cm above the core bottom)

without fuel) in the outer rows was reached. Additionally an annealing procedure was applied around the weld with the highest neutron fluence.

The influence of dummy assemblies on the fluence of the Greifswald unit 1 RPV was calculated using the program TRAMO, which was earlier described in detail [2]. This NPP was operated with the original complete fuel loading up to the 11th cycle. Beginning with the 12th cycle, the outer row of fuel assemblies was replaced by steel shielding assemblies (see Fig. 2) to reduce the flux. Nevertheless, after the 13th cycle an annealing procedure had to be applied at the weld SN0.1.4 of the RPV. This weld is located about 30 cm above the core bottom (Fig. 1). Two additional cycles were operated after the annealing procedure.

In Fig. 2 the 60 deg symmetry sector of the core used for the calculations is shown. The positions of the dummy assemblies are marked. The fuel pins were replaced by seven steel cylinders with a diameter of 48 mm in the dummy assemblies. During cycle 14 only the four central dummy assemblies were used; at the outer two positions fuel assemblies with high burn-up were inserted.

The neutron flux of the RPV generally depends on the azimuthal position because of the radial asymmetry of the core. Figure 3 shows the flux at the outer surface of the RPV for the standard core loading and a core loading with dummy assemblies.

The axial position corresponds to the flux maximum. The flux was determined to be 5.5×10^{10} neutrons/s cm^2 and was found at the central position of 30 deg (see Fig. 2). The values for the edge positions (0 deg and 60 deg, respectively) are approximately 30% lower. The flux was reduced by inserting the dummy assemblies by a factor of 2–5. The flux maximum was moved by inserting dummy assemblies into the azimuthal position of 10 deg and 50 deg, respectively, in the last cycles. The azimuthal difference increases now up to 50%.

Figure 4 shows the development of the fast fluences (energy limits $E>1.0$ MeV and $E>0.5$ MeV) at the inner surface of the RPV in the flux maximum. The value of 0.5×10^{20} neutrons/cm^2 was already reached at unit 1 after 11 cycles. For the originally planned lifetime of 30 years the fluence can be estimated to 1.5×10^{20} neutrons/cm^2 with a mean fluence contribution per cycle of 4.5×10^{18} neutrons/cm^2. Due to the loading changes after the 11th cycle, the mean fluence per cycle could be reduced to 7.7×10^{17} neutrons/cm^2. Those are only 16% of the original value, the total fluence after 30 years would be expected to be only 0.7×10^{20} neutrons/cm^2.

The maximal allowed neutron fluence is documented in Russian nuclear regulatory documents [4]. The reference value is the fluence for energies $E>0.5$ MeV. The values depend on the composition of the steel and differ therefore between all units. The "limiting fluence" values for the first WWER-440/230 units are given between 0.5×10^{20} neutrons/cm^2 and 2.0×10^{20} neutrons/cm^2.

In Fig. 5 the reduction of the fast neutron fluence inside the RPV wall is shown for the axial position approximately 70 cm above the core bottom. For both indicated azimuthal positions the fluence values are given for $E>0.5$ MeV and $E>1.0$ MeV, respectively. The neutron fluence at the position of 0 deg/60 deg is

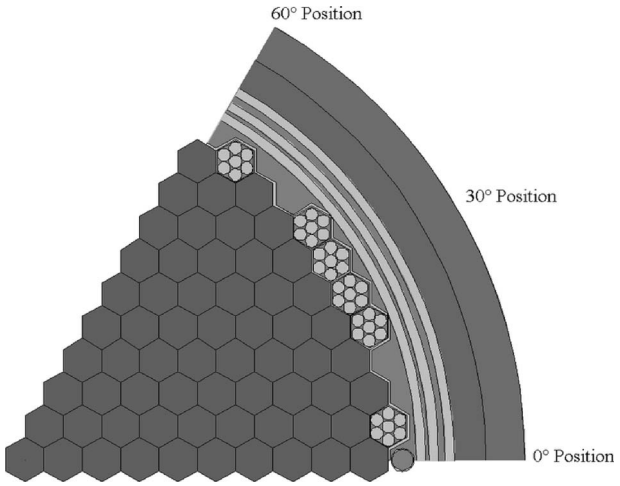


Fig. 2 Horizontal sectional drawing of the WWER-440 core with positions of dummy assemblies

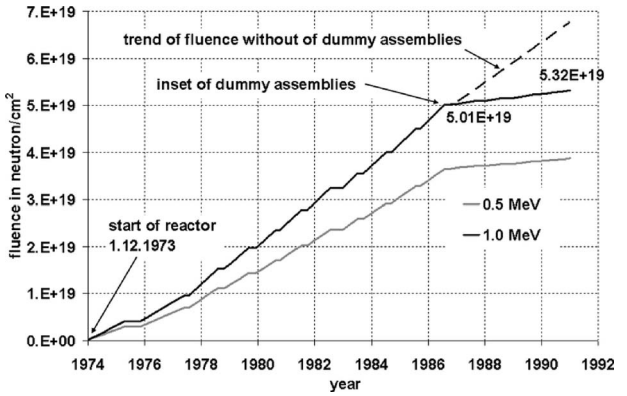


Fig. 4 Development of the fluences $E>1.0$ MeV and $E>0.5$ MeV at the inside of RPV in the flux maximum

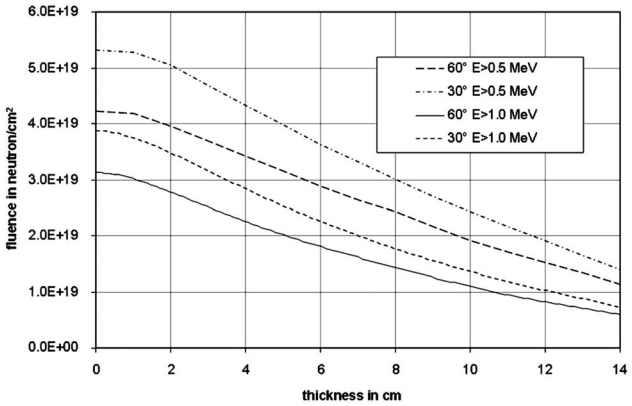


Fig. 5 The fluences $E>1.0$ MeV and the fluences $E>0.5$ MeV in an axial position 70 cm over the core bottom

only 20% smaller compared with the position of 30 deg, which is due to the influence of the dummy assemblies in the last four cycles.

It should be mentioned that the attenuation of the gamma fluence due to the dummy assemblies is generally smaller. Although the calculations show that the fission gammas are decreased by using the dummy assemblies up to 90%, the total gamma fluxes were reduced only to 30% due to the additional production of gammas by neutron reactions in the dummy assemblies.

3 Neutron Fluences in Different Welds

It is known that for the WWER-440 RPV the circumferential welds are particularly sensitive to embrittlement phenomena. Therefore fluence calculations were performed for all three welds indicated in Fig. 1. In Fig. 6 the fluences through the RPV wall are given for these welds.

The weld SN0.1.4 is located 30 cm above the core bottom and receives the highest fluences of all welds. The calculation scheme and the results for this weld were described in detail in Ref. [2], so that only short information will be given here.

The fluences show the expected decrease from the inner to the outer surface of the RPV. The azimuthal differences between 0 deg/60 deg and 30 deg correspond to the above described flux maximum position. The maximum fluences were determined to be 4.0×10^{19} neutrons/cm^2 for $E>0.5$ MeV and to be 2.4×10^{19} neutrons/cm^2 for $E>1.0$ MeV. The fluences of weld SN0.1.4 are approximately 20% smaller as the maximum fluences found for the position 110 cm above the core bottom (see Fig. 1). The maximum gamma fluence was found to be 1.7×10^{20} gamma/cm^2 for $E>1$ MeV.

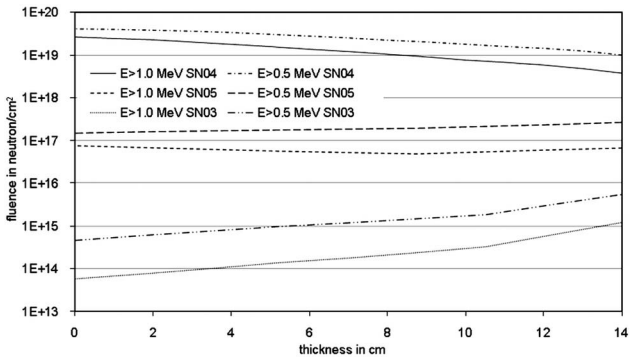


Fig. 6 The fluences $E>0.5$ MeV and $E>1.0$ MeV at the welds SN0.1.3, SN0.1.4, and SN0.1.5

The weld SN0.1.5 is located 58 cm above the core near the inlet nozzle. Therefore this weld underlies considerably stronger stress loads than the weld SN0.1.4 at certain accidents. Such states with high material stresses occur when cold emergency water is fed into the reactor during an accident.

In the Russian nuclear regulatory documents [4] generally a maximum fluence of 1.0×10^{18} neutrons/cm^2 is accepted as uncritical for RPV components. If this limit is exceeded, special proofs and verifications are requested.

The maximum neutron fluence was found to be 2.6×10^{17} neutrons/cm^2 ($E>0.5$ MeV). The neutron fluence is approximately two orders of magnitude smaller than the fluence at weld SN0.1.4. Surprisingly, the fluence was higher at the outer surface than at the inner surface of the RPV. This can be explained on one hand due to the air cavity outside of the RPV. The neutrons leaving the core in the horizontal direction are scattered by the peripheral components along the air cavity in the axial direction and enhance the flux at the outer surface near the weld SN0.1.5. On the other hand, the neutrons on the direct way are stronger moderated and absorbed into the water and internals within the RPV. For the fluence $E>1.0$ MeV a similar behavior was found; the gradient over the wall was even smaller. Here the influence of the direct (uncollided) fission neutrons becomes noticeable at the inner surface. The azimuthal fluence differences at weld SN0.1.5 were smaller compared with the differences at weld SN0.1.4 (maximal 20%).

The gamma fluences at the inner surface of weld SN0.1.5 with the maximum value ($E>1.0$ MeV) of 3.3×10^{18} gamma/cm^2 are larger than at the outer surface. This is due to the higher production of gammas from the thermal neutron capture reaction within the RPV (internals and water). The contribution of the fission gammas is comparatively small ($<10\%$).

In the case of weld SN0.1.5 the above given fluence limit of 1×10^{18} neutrons/cm^2 would not be reached within the planned lifetime. However, this value could be exceeded in case of lifetime extension of WWER-440 power plants.

The weld SN0.1.3 is located approximately 160 cm below the core. The fluence values were calculated, in this case over the whole RPV sector, because the azimuthal core asymmetry can be neglected in this position. On the other hand, the influence of the inner fuel assemblies was taken into account. Neutrons, which are generated in this core region, can remarkably contribute to the fluence. Therefore in these calculations all sources in the lower core region were considered. A statistical error of up to 10% was accepted for the integral neutron values.

The calculation showed that the neutron fluence is approximately three to four orders of magnitude smaller compared with the fluence of weld SN0.1.4. A fluence of 1.0×10^{16} neutrons/cm^2 ($E>0.5$ MeV) was found. Material from the weld SN0.1.3 can be regarded as unirradiated.

Opposite to the fluence behavior at weld SN0.1.5, the fluence gradient over the RPV wall for the weld SN0.1.3 was found to be negative for all integral neutron fluences greater 0.1 MeV. Caused by the large distance to the core, no direct fission neutrons contribute to the fluence. The air cavity outside of the RPV plays a dominant role in neutron transport. It should be mentioned that the thermal fluence shows an opposite gradient.

The gamma fluence (>1 MeV) at the inner surface of weld SN0.1.3 was calculated to be 3.0×10^{16} gamma/cm^2. The gradient over the wall is positive. Similar to the neutron fluence the gamma fluence for weld SN0.1.3 was found to be three to four orders of magnitude smaller than the gamma fluence for weld SN0.1.4.

4 Material and Specimens

The trepaning procedure was described in Ref. [1]. The investigated trepan 1-1 represent the irradiation/annealing/irradiation condition of the welding seam SN0.1.4. The welding seam

Table 1 Chemical composition of trepan 1-1 (mass %)

No. disk	Protocol[b]	1-1.1	1-1.3	1-1.12
Location[a]		8	22	94
C	0.05			
Si	0.47			
Mn	1.22	1.06	0.97	0.93
Cr	1.48	1.49	1.35	1.23
Ni	0.23	0.22	0.19	0.22
Mo	0.41	0.4	0.43	0.4
V	0.16	0.14	0.14	0.09
P	0.037	0.038	0.03	0.028
Cu	0.103	0.125	0.141	0.141

[a]Distance of the disk center from the inner surface in millimeters.
[b]Manufacturing protocol of the RPV unit 1—welding seam SN0.1.4.

SN0.1.4 is a multilayer submerged weld. The welding seams of the WWER-440 RPVs consist of a welding root welded with an unalloyed wire Sv-08A and the filling material welded with the alloyed wire Sv-10KhMFT. Table 1 contains the chemical composition of the weld metal within trepan 1-1 in different thickness locations, and the specifications of the manufacturing protocol [5]. The chemical compositions measured at the three locations of the welding seam belong to the alloyed filling material and generally agree with the information in the manufacturing protocol. The copper and phosphorus contents are within the range, as specified in the manufacturing guidelines of the WWER-440 generation 1 (model 230), but both are clearly higher than in the specification of next generation (model 213) with maximum allowed P and Cu contents of 0.01% and 0.1%, respectively [6,7].

The trepan 1-1 was cut in into disks with a thickness of 10 mm using a wire traveling electro-erosion discharging machine (EDM). From one disk ten Charpy size SE(B) specimens were machined. Figure 7 shows the trepan, and Fig. 8 exemplifies the cutting scheme of the disk 1-1.1. The location of the welding seam within trepan 1-1 was examined by metallographic means and is schematically depicted in Fig. 7. The welding root is located within a distance of about 60–80 mm relative to the inner RPV wall.

The Charpy size SE(B) specimens were precracked ($a/W=0.5$) and 20% side grooved. As shown in Fig. 8, the orientation of the SE(B) specimen is TS (specimen axis axial and crack growth direction across the vessel wall) according to ASTM E399.

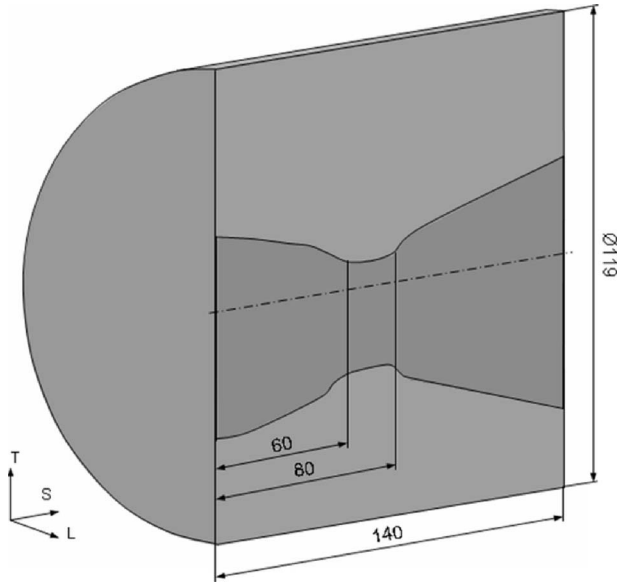


Fig. 7 Trepan 1-1 with the location of the welding seam

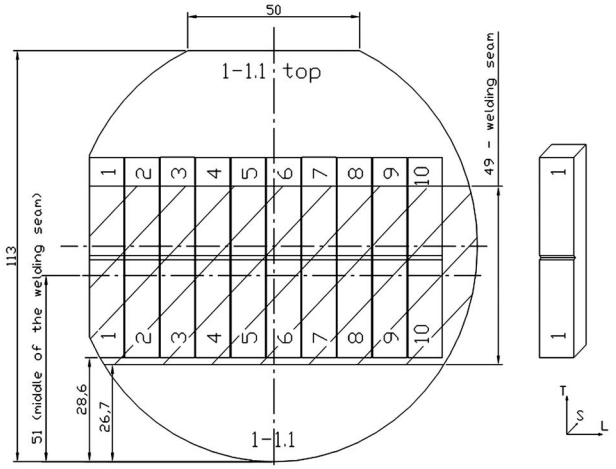


Fig. 8 Cutting scheme of disk 1-1.1, as an example

The TS orientation is in correspondence with the surveillance specimens in Russian WWER-440/213 reactors. From the broken halves of the tested SE(B) specimens Charpy V-notch specimens according to DIN EN 10045-1 (1991) "Metallic Materials: Charpy Impact Test; Part 1: Test Method" and EN ISO 14556 (2000) "Steel—Charpy V-Notch Pendulum Impact Test—Instrumented Test Method" were manufactured by the reconstitution technique.

5 Testing and Evaluation

SE(B) specimens from seven disks with thickness locations shown in Table 2 were tested and evaluated according to ASTM E1921-05 "Standard Test Method for Determination of Reference Temperature, T_0, for Ferritic Steels in the Transition Range." The precracked and side-grooved Charpy size SE(B) specimens were monotonously loaded until they failed by cleavage instability. Standard MC reference temperatures T_0 were evaluated with the measured J integral based cleavage fracture toughness values, K_{Jc}, applying the multitemperature procedure of ASTM E1921-05. In addition, the Structural Integrity Assessment Procedures for European Industry (SINTAP) containing a modification of the MC analysis were used for the evaluation of the measured K_{Jc}- values. The SINTAP lower tail analysis contains three steps and guides

Table 2 Location of the investigated disks within trepan 1-1, neutron fluences, and MC test results according to ASTM E1921-05 and Charpy V parameter

Code disk	1-1.1	1-1.3	1-1.6	1-1.8	1-1.9	1-1.11	1-1.12
Location[a]	8.3	21.9	42.3	59.6	69.9	83.5	93.8
Neutron fluence ($E>1$ MeV) in 10^{18} neutrons/cm^2							
Before[b]	24.23	21.47	16.89	13.43	11.63	9.56	8.21
After[b]	0.71	0.64	0.51	0.42	0.36	0.3	0.26
Total	24.93	22.11	17.4	13.85	11.99	9.86	8.47
ASTM E1921—05							
T_0 (°C)	10.3	49.1	−5	−4.5	−40.7	−28.4	19.8
σT_0 (K)	6.4	6.3	6.4	6	6.4	6	6.3
SINTAP (°C)							
T_0^{SINTAP}	32.5	49.1	−5	−4.5	−13.6	18.6	40.9
Charpy V							
TT_{41J} (°C)	51.4	32.6					20.1
USE (J)	130.8	153					120.3

[a]Distance of the disk center from the inner surface in millimeters.
[b]Annealing.

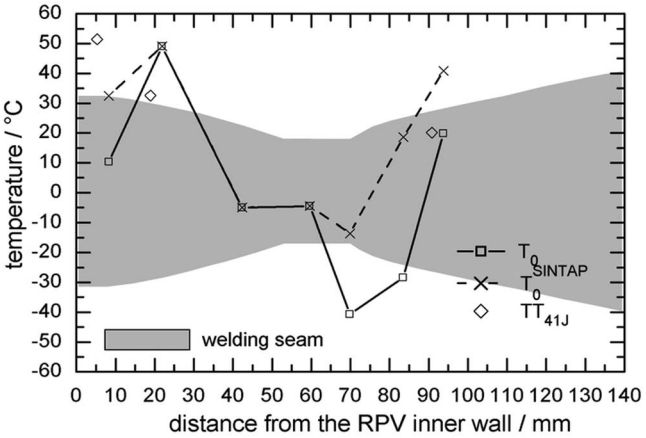


Fig. 9 Course of the reference temperatures T_0 and T_0^{SINTAP} through the welding seam SN0.1.4

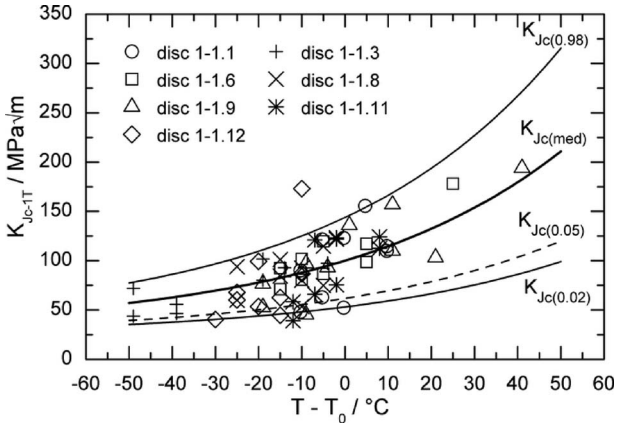


Fig. 10 K_{Jc} values measured on Charpy size SE(B) specimens, adjusted to 1T (25.4 mm) specimen size, versus the test temperature normalized to T_0 of the individual disks, and master curves

the user toward the most appropriate estimate of the reference temperature, T_0^{SINTAP}, describing the population having the lower toughness [8,9].

Instrumented Charpy V-notch impact tests on reconstituted specimens were performed according to DIN EN 10045–1 (1991). The impact energy, lateral expansion and fracture appearance temperature curves were fitted by the tanh approach. Charpy V parameters as transition temperatures and the upper shelf energy (USE) were evaluated.

6 Results and Discussion

Table 2 summarizes the MC and Charpy V test results of the investigated disks of trepan 1-1. The table also contains the location of the disks within the trepan and the calculated neutron fluences in the center of the disks. The test results comprise the reference temperatures evaluated according to ASTM E1921–05 (T_0) and the SINTAP procedure (T_0^{SINTAP}), as well as the Charpy V transition temperature at an impact energy of 41 J (TT_{41J}) and USE. T_0 and TT_{41J} data presented in Table 2 and depicted in Fig. 9 vary through the thickness of the trepan 1-1 and, thus, the welding seam. Along the wall thickness, T_0 shows a wavelike behavior. After an initial increase in T_0 from 10°C at the inner surface to 49°C at a 22 mm distance from it, T_0 again decreases to −41°C at a distance of 70 mm, finally increasing again to 20°C at 94 mm from the center. The lowest T_0 values were measured on SE(B) specimens from disks 1-1.9 and 1-1.12 located in the root region of the welding seam. Figure 10 shows the K_{Jc} values versus the test temperature normalized to T_0 of the individual disks. The K_{Jc} values generally follow the course of the MC, though the scatter is large. Nevertheless, the K_{Jc} values are close to or above the 2% fracture probability line. However, more than 5% of the K_{Jc} data fall below the 5% fracture probability lower bound ($K_{Jc(0.05)1T}$) curve. This strongly indicates that the material is not fully homogeneous, which is not unusual for the investigated multilayer weld metal.

The SINTAP MC evaluation enables conservative lower bound type fracture toughness estimates also for inhomogeneous materials. As shown in Table 2, the SINTAP MC evaluation gives reference temperatures T_0^{SINTAP} for disks 1-1.1, 1-1.9, 1-1.11, and 1-1.12, clearly higher than the standard T_0. The course of T_0^{SINTAP} through the welding seam SN0.1.4 in Fig. 9 also shows the lowest values in the root region and an increase from the inner surface to a 22 mm distance within the filling layers. The filling layer (disk 1-1.12) beyond the root has a T_0^{SINTAP} value comparable with disk 1-1.3.

The Charpy V transition temperature TT_{41J} (Table 2) cannot be correlated with the T_0 because the location of the notch root in the Charpy V-notch specimen is different from the crack tip in the SE(B) specimen. The structure of the investigated multilayer welding seam is different within short distances caused by the applied welding technology. There are no Charpy V data available from the weld metal SN0.1.4 in the initial condition. The evaluated TT_{41J} of 51°C is close to the so called "critical embrittlement temperature" for the initial condition, T_{K0}, of welding seam SN0.1.4 reported in the manufacturing protocol as 46°C [5]. The evaluated upper shelf energy is on the level expected for WWER-440 weld metal [5,6]. The estimated critical embrittlement temperature in the irradiated condition, T_K, before the annealing in 1988 was reported as 186°C [5,10]. To determine T_K after recovery annealing of the RPV 6 mm thick compact specimen, samples were cut from the inner wall of the welding seam SN0.1.4. by means of an EDM machine in 1988 [6]. Subsized impact specimens KLST ($3\times3\times27$ mm^3) according to DIN EN 10045–1 (1991) were machined from those compact specimens. The orientation of the KLST specimens was TL (specimen axis axial and crack growth direction circumferential). Using the KLST specimens a ductile-to-brittle transition temperature at an impact energy of 1.9 J, $TT_{1.9\ J}$, of −35°C was determined for the inner surface of the welding seam SN0.1.4 [6]. $TT_{1.9\ J}$ of −35°C can be converted into a Charpy V-notch transition temperature as follows:

$$TT_{41J} = TT_{1.9J} + 65\ \text{K} = -35\,^\circ\text{C} + 65\ \text{K} = 30\,^\circ\text{C} \qquad (1)$$

The scatter of this correlation, ±35 K, is rather high [11] and therefore the uncertainty of the applied conversion is likewise. The TT_{41J} close to the inner RPV wall is 51°C (Table 2). Taking into account the difference in the orientation of the specimens and the re-irradiation of two cycles, the TT_{41J} estimated with KLST specimens after the recovery annealing is realistic. In this case the re-irradiation causes an increase in TT_{41J} by 21 K.

As shown in Table 2 and Fig. 9, the highest value of the MC T_0 was not determined at the surface layer. There is a difference of about 40 K between the locations 8 mm and 22 mm distant from the inner surface of the welding seam SN0.1.4. In this case the T_K values determined with samples from the inner surface do not represent the most conservative value.

The MC approach and the associated reference temperature, T_0, as defined in the test standard ASTM E1921, rapidly moves from the research laboratories to application in integrity assessment of components and structures [12,13]. It is more naturally suited to probabilistic analyses because it defines both a mean transition toughness value and a distribution around that value. It is reasonable to expect that in the future the determination of nuclear power plant operating limits will be based on master curve meth-

Table 3 Location of the investigated disks within trepan 1-1, neutron fluences, and reference temperatures evaluated according to the VERLIFE procedure

Code disk	1-1.1	1-1.3	1-1.6	1-1.8	1-1.9	1-1.11	1-1.12
Location[a]	8.3	21.9	42.3	59.6	69.9	83.5	93.8
	Neutron fluence ($E>1$ MeV) (10^{18} neutrons/cm^2)						
Before[b]	24.23	21.47	16.89	13.43	11.63	9.56	8.21
After[b]	0.71	0.64	0.51	0.42	0.36	0.3	0.26
Total	24.93	22.11	17.4	13.85	11.99	9.86	8.47
	VERLIFE						
RTT_0 (°C)	27.5	66.3	12.2	12.6	−23.4	−11	37
RTT_0^{SINTAP} (°C)	49.7	66.3	12.2	22	3.6	35.7	62.3

[a]Distance of the disk center from the inner surface in millimeters.
[b]Annealing.

ods. For WWER reactors the "Unified Procedure for Lifetime Assessment of Components and Piping in WWER NPPs—VERLIFE" defines a reference temperature, RTT_0, used in the integrity assessment of WWER reactors as

$$RTT_0 = T_0 + \sigma \tag{2}$$

where T_0 is the reference temperature according to ASTM E1921–05, σ is a margin $\sigma=\sqrt{\sigma_1^2+\delta T_M^2}$, σ_1 is the standard deviation according to ASTM E1921–05, and δT_M considers the scatter in the materials (if this value is not available the application of the following values is suggested: $\delta T_M=10\,^\circ$C for the base material and $\delta T_M=16\,^\circ$C for weld metals).

The VERLIFE procedure suggests the following RTT_0 indexed lower bound curve for the WWER base and weld metal:

$$K_{JC}^{5\%}(T) = \min\{25.2 + 36.6 \cdot e^{[0.019(T-RTT_0)]}; 200\} \quad \text{in} \quad \text{MPa}\sqrt{\text{m}} \tag{3}$$

Equation (3) agrees with the standard MC for 5% fracture probability in ASTM E1921-05. Table 3 summarizes the reference temperatures, RTT_0, evaluated according to the VERLIFE procedure for the investigated disks of trepan 1-1. The RTT_0 were evaluated using the T_0 and T_0^{SINTAP} according to ASTM E1921-05 and the SINTAP MC procedure, respectively. Figure 11 shows the K_{Jc} values versus the test temperature normalized to the individual T_0 of the different thickness locations. The K_{Jc} values adjusted to a specimen thickness of 1T (25.4 mm) are not completely enveloped by the VERLIFE WWER lower bound fracture toughness

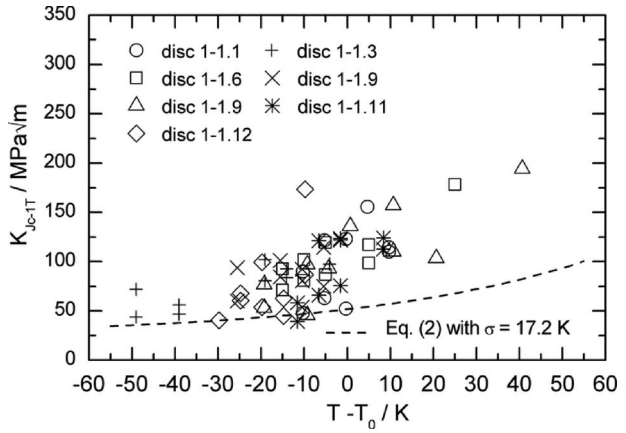


Fig. 11 K_{Jc} values measured on Charpy size SE(B) specimens, adjusted to 1T (25.4 mm) specimen size versus the test temperature normalized to RTT_0 of the different disks and VERLIFE lower bound curve

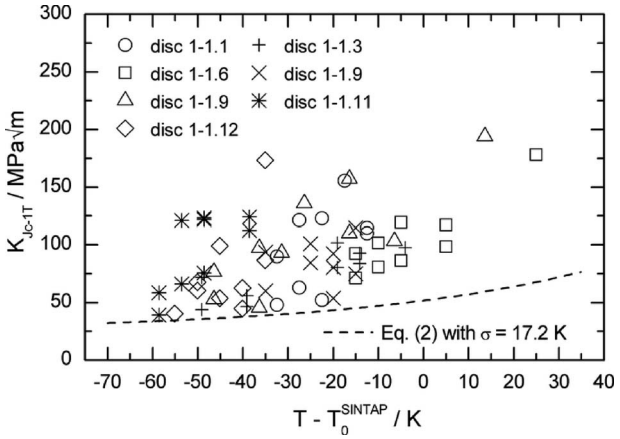


Fig. 12 K_{Jc} values measured on Charpy size SE(B) specimens, adjusted to 1T (25.4 mm) specimen size versus the test temperature normalized to RTT_0^{SINTAP} of the different disks and VERLIFE lower bound curve

curve. The application of RTT_0^{SINTAP} as the index temperature ensures that all measured K_{Jc} values are now enveloped by the VERLIFE lower bound fracture toughness curve (Fig. 12).

7 Summary and Conclusion

RPV neutron fluence calculations were carried out for different loading schemes and the welds SN0.1.3, SN0.1.4, and SN0.1.5 of the NPP Greifswald unit 1 using the Monte Carlo code TRAMO. With exception of the weld SN03 the statistic errors of the integral values were even smaller 1% (for the SN03 10%). The results show that the use of the dummy assemblies reduces the flux by a factor of 2–5 depending on the azimuthal position. The fluence increase is reduced to 1/6 at the position of the maximum fluence.

The neutron fluence at the different circumferential welds is closely related to their distance to the core. The circumferential core weld (SN0.1.4) received a fluence of 2.4×10^{19} neutrons/cm^2 at the inner surface; it decreases to 0.8×10^{19} neutrons/cm^2 at the outer surface. The neutron fluences at the other welds are 2 and 4 orders of magnitude, respectively, smaller according to their distances from the core. It should be mentioned that in these cases the fluence gradient can be even negative through the wall.

At first the trepan of the core welding seam was investigated by MC and Charpy V-notch testing. Specimens from seven locations through the thickness of the welding seam were tested according to ASTM E1921-05. The reference temperature T_0 was calculated with the measured fracture toughness values, K_{Jc}, at brittle failure of the specimen. Generally the K_{Jc} values measured on precracked and side-grooved Charpy size SE(B) specimens of the investigated weld metal follows the course of the master curve. The K_{Jc} values show a remarkable scatter. More values than expected lie below the 5% fractile. In addition the MC SINTAP procedure was applied to determine T_0^{SINTAP} of the brittle fraction of the data set. There are remarkable differences between T_0 and T_0^{SINTAP} indicating a macroscopic inhomogeneous weld metal. The highest T_0 was about 50 °C at a distance of 22 mm from the inner surface of the weld. It is 40 K higher compared with T_0 at the inner surface. This is important for the assessment of ductile-to-brittle temperatures measured with subsized Charpy specimens made of weld metal from the inner RPV wall. This material does not represent the most conservative condition. Nevertheless, the Charpy transition temperature TT_{41J} estimated from results of subsized specimens after the recovery annealing was confirmed by the testing of standard Charpy V-notch specimens.

The VERLIFE procedure prepared for the integrity assessment of WWER RPV was applied on the measured results. It enables

the determination of a reference temperature, RTT_0 to index a lower bound fracture toughness curve. This curve agrees with the MC 5% fractile, as specified in ASTM E1921-05. The measured K_{Jc} values are not enveloped by this lower bound curve. However, the VERLIFE lower bound curve indexed with the SINTAP reference temperature RTT_0^{SINTAP} envelops the K_{Jc} values. Therefore for a conservative integrity assessment the fracture toughness curve indexed with a RTT_0^{SINTAP} representing the brittle fraction of a data set of measured K_{Jc} values has to be applied.

Nomenclature

- ASTM = American Society for Testing Materials
- EDM = electro-erosive discharging machine
- KLST = subsized impact specimens ($3 \times 3 \times 27$ mm^3)
- K_{Jc} = J-integral base fracture toughness at the onset of cleavage failure of the specimen
- PVR = pressurized water reactor
- RTT_0 = reference temperature of the VERLIFE lower bound fracture toughness curve based on the ASTM E1921 T_0
- RTT_0^{SINTAP} = reference temperature of the VERLIFE lower bound fracture toughness curve based on T_0^{SINTAP}
- SE(B) = single edge bend specimen
- SN0.1x = number of RPV weld
- T_K = critical temperature of brittleness according to the Russian code
- T_0 = reference temperature according to the Master Curve approach in ASTM E1921
- T_0^{SINTAP} = reference temperature evaluated according to the MC extension in SINTAP
- TT_{41J} = Charpy V transition temperature related to an impact energy of 41 J
- $TT_{1.9\ J}$ = transition temperature related to an impact energy of 1.9 J determined with KLST specimens
- VERLIFE = Unified Procedure for Lifetime Assessment of Components and Piping in WWER
- WWER = Russian water moderated, water operated energy reactor (PVR type)
- 1T = 1 in. (25.4 mm)
- σ = margin $\sigma = \sqrt{\sigma_1^2 + \delta T_M^2}$
- σ_1 = standard deviation according to ASTM E1921-05
- δT_M = considers the scatter in the materials

References

[1] Konheiser, J., Rindelhardt, U., Viehrig, H.-W., Böhmer, B., and Gleisberg, B., 2006, "Pressure Vessel Investigations of the Former Greifswald NPP: Fluence Calculations and Nb Based Fluence Measurements," ICONE14/FEDSM2006 Proceedings on DVD, Contribution ICONE 14-89578.

[2] Rindelhardt, U., Viehrig, H. W., Konheiser, J., Noack, K., and Gleisberg, B., 2007, "RPV Material Investigations of the Former Greifswald NPP," ICONE 15 Proceedings on DVD, Contribution ICONE15-10335.

[3] Ahlstrand, R., Hietanen, O., Juntunen, T., McNiven, U., Nurkkala, P., Rajamäki, P., Snellman, J., and Vuorenmaa, A., 1991, "Identifying Life-Limiting Factors at the Loviisa Power Plant and Management of the Ageing Process," International Conference on Nuclear Plant Life Extension, Upgrading, Repair, Refurbishment, Uprating and Ageing, Berlin, 2–4 December, *Proceedings in Nuclear Engineering International*, pp. 343–354.

[4] PNAE G-7-002-86, 1989, "Strength Calculation Norms for Components and Pipelines of Nuclear Power Installations," Energoatomizdat, USSR Gosatomenergonadzor, Moscow, p. 525, in Russian.

[5] Böhmer, B., Böhmert, J., Müller, G., Rindelhardt, U., and Utke, H., 1999, "Embrittlement Studies of the Reactor Pressure Vessel of the Greifswald-440 Reactors," TACIS Service DG IA, Brussels, Belgium, Technical Report No. NUCRUS96601.

[6] Davies, L. M., 1997, "A Comparison of Western and Eastern Nuclear Reactor Pressure Vessel Steels," European Commission, Luxembourg, AMES Report No. 10.

[7] Brumovský, M., Valo, M., Kryukov, A., Gillemot, F., and Debarberis, L., 2005, *Guidelines for Prediction of Irradiation Embrittlement of Operating WWER-440 Reactor Pressure Vessels*, IAEA, Vienna.

[8] Wallin, K., Nevasmaa, P., Laukkanen, A., and Planman, T., 2004, "Master Curve Analysis of Inhomogeneous Ferritic Steels," Eng. Fract. Mech., **71**(16–17), pp. 2329–2346.

[9] Pisarski, H. G., and Wallin, K., 2000, "The SINTAP Fracture Toughness Estimation Procedure," Eng. Fract. Mech., **67**(6), pp. 613–624.

[10] Ahlstrand, R., Klausnitzer, E. N., Langer, D., Leitz, Ch., Pastor, D., and Valo, M., 1993, "Evaluation of the Recovery Annealing of the Reactor Pressure Vessel of NPP Nord (Greifswald) Units 1 and 2 by Means of Subsize Impact Specimens," *Radiation Embrittlement of Nuclear Reactor Pressure Vessel Steels: An International Review*, L. E. Steel, ed., American Society for Testing and Materials, Philadelphia, Vol. 4, pp. 321–343.

[11] Klausnitzer, E., Kristof, H., and Leistner, R., 1985, "Assessment of Toughness Behaviour of Low Alloy Steels by Sub-Size Impact Specimens," Transactions of the 8th International Conference of Structural Mechanics in Reactor Technology (SMIRT), Vol. G, pp. 33–37.

[12] Kirk, M., Lott, R., Server, W. L., Hardies, R., and Rosinsky, S., 2000, "Bias and Precision of T_0 Values Determined Using ASTM Standard E1921-97 for Nuclear Pressure Vessel Steels," *Effects of Radiation on Materials: 19th International Symposium*, M. L. Hamilton, A. S. Kumar, S. T. Rosinsky, and M. L. Grossbeck, eds., American Society for Testing and Materials, West Conshohocken, PA, Paper No. ASTM STP 1366, pp. 142–161.

[13] Brumovský, M., Faidy, C., Karzov, G., KiSig, K., Kryukov, A., Lapena, J., Lott, R., Lyssakov, V. N., Nanstad, R., Planman, T., Rosinski, S., Serrano, M., Server, W. L., Servini, F., and Viehrig, H.-W., 2005, "IAEA Guidelines for Application of the Master Curve Approach to Reactor Pressure Vessel Integrity in Nuclear Power Plants," IAEA-Technical Reports No. 429.

Air/Water Counter-Current Flow Experiments in a Model of the Hot Leg of a Pressurized Water Reactor

Christophe Vallée
Institute of Safety Research,
Forschungszentrum Dresden-Rossendorf e.V.,
P.O. Box 51 01 19,
01314 Dresden, Germany
e-mail: c.vallee@fzd.de

Deendarlianto
Department of Mechanical and Industrial Engineering,
Faculty of Engineering,
Gadjah Mada University,
Jalan Grafika No. 2,
Yogyakarta 55281, Indonesia

Matthias Beyer

Dirk Lucas

Helmar Carl

Institute of Safety Research,
Forschungszentrum Dresden-Rossendorf e.V.,
P.O. Box 51 01 19,
01314 Dresden, Germany

Different scenarios of small break loss of coolant accident for pressurized water reactors (PWRs) lead to the reflux-condenser mode in which steam enters the hot leg from the reactor pressure vessel (RPV) and condenses in the steam generator. A limitation of the condensate backflow toward the RPV by the steam flowing in counter current could affect the core cooling and must be prevented. The simulation of counter-current flow limitation conditions, which is dominated by 3D effects, requires the use of a computational fluid dynamics (CFD) approach. These numerical methods are not yet mature, so dedicated experimental data are needed for validation purposes. In order to investigate the two-phase flow behavior in a complex reactor-typical geometry and to supply suitable data for CFD code validation, the "hot leg model" was built at Forschungszentrum Dresden-Rossendorf (FZD). This setup is devoted to optical measurement techniques, and therefore, a flat test-section design was chosen with a width of 50 mm. The test section outlines represent the hot leg of a German Konvoi PWR at a scale of 1:3 (i.e., 250 mm channel height). The test section is mounted between two separators, one simulating the RPV and the other is connected to the steam generator inlet chamber. The hot leg model is operated under pressure equilibrium in the pressure vessel of the TOPFLOW facility of FZD. The air/water experiments presented in this article focus on the flow structure observed in the region of the riser and of the steam generator inlet chamber at room temperature and pressures up to 3 bar. The performed high-speed observations show the evolution of the stratified interface and the distribution of the two-phase mixture (droplets and bubbles). The counter-current flow limitation was quantified using the variation in the water levels measured in the separators. A confrontation with the images indicates that the initiation of flooding coincides with the reversal of the flow in the horizontal part of the hot leg. Afterward, bigger waves are generated, which develop to slugs. Furthermore, the flooding points obtained from the experiments were compared with empirical correlations available in literature. A good overall agreement was obtained, while the zero penetration was found at lower values of the gaseous Wallis parameter compared with previous work. This deviation can be attributed to the rectangular cross section of the hot leg model. [DOI: 10.1115/1.3043816]

1 Introduction

In the event of a loss-of-coolant-accident (LOCA) in a pressurized water reactor (PWR), emergency strategies have to be mapped out in order to guarantee a safe removal of heat from the reactor core, also in case of component breakdown. During a small break LOCA with failure of the high pressure emergency core cooling (ECC) system and of the main feed pumps, a natural circulation starts in the primary circuit. This allows the heat removal, also if steam is generated in the reactor core due to the depressurization of the primary circuit. But if further the water level in the reactor pressure vessel (RPV) falls below the hot leg nozzle, only steam will flow to the steam generator. Therefore the natural circulation breaks down and switches to the reflux-condenser mode.

In the reflux-condenser mode, the steam coming from the RPV condenses in the vertical U-tubes of the steam generator. In each half of the steam generator, the condensate flows down the tube in which it has been formed. Therefore, about one-half of the condensate flows as usual over the pump to the downcomer, whereas the other half flows over the hot leg back to the upper plenum. In the hot leg, the condensate has to flow in counter current to the steam.

The horizontal stratified counter-current flow of condensate and steam is only stable for a certain range of flow rates. If the steam flow increases too much, the condensate is clogged in the hot leg. This is the beginning of the counter-current flow limitation (CCFL): The liquid is carried over by the steam and partially entrained in the opposite direction to the steam generator. As a consequence, the hot leg and steam generator (SG) are flood, which decreases the water level in the RPV and reduces the core cooling. In the case of an additional increase in the steam flow, the condensate is completely blocked and the cooling of the reactor core from the hot leg is impossible.

For the validation and optimization of accident management strategies, such transient scenarios are reproduced in dedicated facilities or rather simulated. The use of one-dimensional system codes is state of the art, but these programs are not able to predict important flow conditions with the required accuracy and spatial resolution. In particular, the CCFL conditions are dominated by 3D effects, which requires the use of a computational fluid dynamics (CFD) approach. However, the actual CFD codes applied to two-phase flows do not meet the high level of confidence

Manuscript received July 29, 2008; final manuscript received July 30, 2008; published online January 5, 2009. Review conducted by Dilip R. Ballal. Paper presented at the 16th International Conference on Nuclear Engineering (ICONE16), Orlando, FL, May 12–15, 2008.

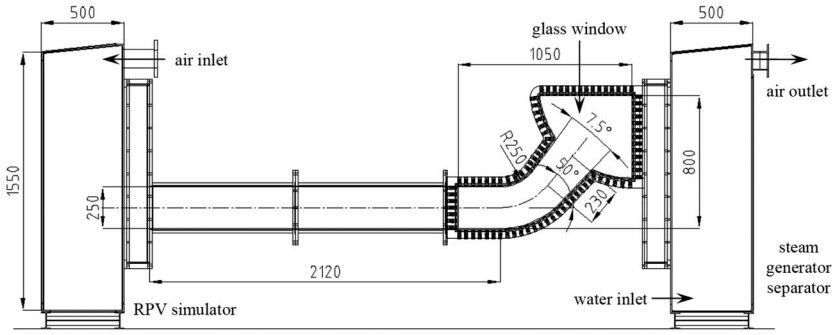


Fig. 1 Schematic view of the hot leg model test section (dimension in millimeters)

needed in the field of nuclear reactor safety. Therefore, further theoretical model development as well as code validation are necessary.

In order to support these tasks, air/water counter-current flow experiments were performed in a model of the hot leg of a PWR. This "hot leg model" is installed at the Transient twO Phase FLOW (TOPFLOW) facility of Forschungszentrum Dresden-Rossendorf (FZD). The experiments were performed before and around the onset of flooding. The measured global parameters such as water levels and pressure drop will be analyzed in order to characterize the flow. Furthermore, the comparison with detailed visual observations will be used to explain this behavior. Furthermore, the effect of the liquid mass flow rate and of the system pressure will be evaluated. Finally, the onset of flooding as revealed by the data will be discussed in comparison to the proposal of other investigators.

2 The Hot Leg Model

The test section of the hot leg model is schematically shown in Fig. 1. The main components consist of the test section itself, the reactor pressure vessel simulator located at the lower end of the horizontal channel, and the SG separator connected to the SG inlet chamber. The test section reproduces the hot leg of a PWR from the German *Konvoi* type at a scale of 1:3. In order to provide optimal observation possibilities, the test section is not composed of pipes like in the original power plant, it is a 50 mm thick channel representing a cut through the vertical midplane of the hot leg and of the steam generator inlet chamber. Consequently, the test section is composed of a horizontal rectangular channel, a bend that connects it to an upward inclined and expended channel, and a quarter of a circle representing the steam generator inlet chamber. The horizontal part of the test section is 2.12 m long and has a rectangular cross section of 0.05×0.25 m^2. The riser is 0.23 m long, has an inclination of 50 deg to the horizontal plane, and an expansion angle of 7.5 deg. Moreover, the SG and RPV simulators are identical vessels with $0.8 \times 0.5 \times 1.55$ m^3 ($D \times W \times H$) cubic shape.

In order to visualize the flow over big size windows at high pressures, the hot leg model is operated under pressure equilibrium. This condition is realized in the pressure vessel of the TOPFLOW facility of FZD (Fig. 2), where the test section is installed in. The air outlet of the condenser is permanently connected to the inside atmosphere of the vessel. A compressor system allows to increase the air pressure in the chamber to a maximum of 5 MPa. The test section is equipped with glass side walls in the bended region of the hot leg and of the steam generator inlet chamber to allow visual observation, as shown in Fig. 1. The flow behavior was recorded by a high-speed video camera at frequencies of 60–100 Hz and a shutter speed of 1/1000 s.

Furthermore, global parameters were measured: The water levels in both separators were determined by the measurement of the differential pressure between the top and the bottom of the vessels with differential pressure transducers. A vortex meter was used to measure the injected water mass flow rate. The injected air mass

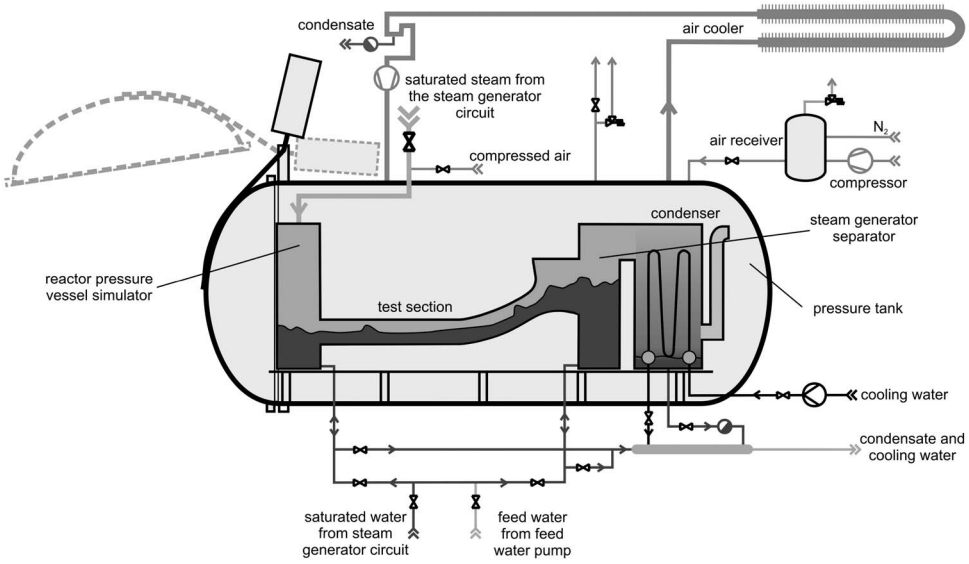


Fig. 2 Schematic view of the experimental apparatus

flow rate was measured and controlled using a thermal mass flow meter. The down-flowing water mass flow rate (discharge flow) was determined by calculating the increase in water level of the RPV simulator. The water level inside the SG separator was measured to check the mass balance. The temperatures both of the air and water were measured by thermocouples. The differential pressure between the SG and RPV simulators was measured by a differential pressure transducer. The signals of these global parameters were transmitted to a personal computer via a data acquisition system running at 1 Hz.

3 Experiments and Results

3.1 Experimental Procedure. The two-phase flow experiments were performed with air and water at room temperature (around 20°C) and system pressures of 1.5 bar and 3.0 bar. The water was injected from the feed water pump in the SG separator, from where it can flow through the test section to the RPV simulator. The air was injected in the RPV simulator and flowed through the test section in counter current to the water flow to the SG separator. From there, the air is released over the condenser to the inner atmosphere of the pressure vessel. The onset of flooding was obtained by a stepwise increase in the air flow rate with small increments under a constant water flow rate. The onset of flooding was defined as the limiting point of stability of the counter-current flow indicated by the maximum air flow rate at which the down-flowing water flow rate is equal to the inlet water flow rate. This method was used by previous investigators such as by Zabaras and Dukler [1] and Deendarlianto et al. [2]. The experiments were carried on until the point of zero-liquid penetration appears when the down-flowing water flow rate was equal to zero. During the experiments, the water mass flow rate was varied between 0.1 kg/s and 0.9 kg/s and the air mass flow rate was varied between 0.16 kg/s and 0.31 kg/s.

3.2 Flow Behavior Observed in the Riser of the Hot Leg Model at a System Pressure of 3.0 bar

3.2.1 Flow Behavior at a Water Flow Rate of 0.3 kg/s. Figure 3 shows the water levels measured in the SG and RPV separators, the pressure difference between the two vessels, and the injected air mass flow rate. During this flooding experiment performed at a system pressure of 3.0 bar, a constant water mass flow rate of 0.3 kg/s was injected. The water levels inside both separators are shown in the upper graph. The pressure difference between the vessels and the injected air mass flow rate is represented in the lower graph of this figure. From the slopes of the curve of the water level in the RPV simulator shown in Fig. 3, the experiment can be divided into three regions.

(1) In the first region (Fig. 3: Region I), the water level in the RPV simulator increases with the increase in air mass flow rate; meanwhile the water level in the SG separator is almost constant. This means that all the injected water flows from the SG separator to the RPV simulator. In this region, it is also found that the pressure difference between the vessels is still low and slightly increases with the air mass flow rate. We define this region as the stable counter-current flow. The flow behavior can be described from the high-speed camera images shown in Fig. 4. It should be noticed that the flow pattern on the inclined plane of the riser is a supercritical stratified flow for $t<61.0$ s. Supercritical flow means that the local Froude number of the liquid film is larger than unity. In the geometry of the hot leg, the supercritical flow condition is due to gravity acceleration. In the horizontal part of the hot leg, the supercritical flow changes to subcritical flow, and a hydraulic jump as a transition from supercritical to subcritical flow is observed near the bended region (Fig. 4(*a*)). Furthermore, the air/water interface in the inclined riser is stable, indicating that the water flow is not disturbed by the air stream.

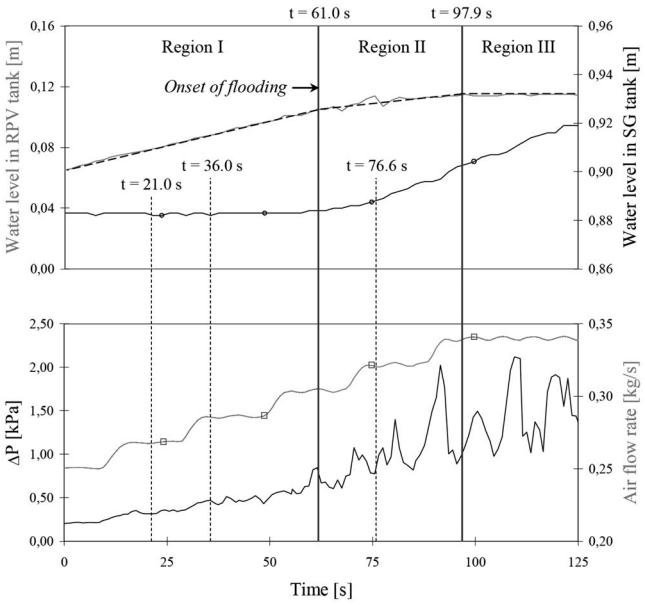


Fig. 3 Variation in the water levels in the RPV simulator (top diagram, top gray curve) and in the SG separator (top diagram, bottom black curve with points) of the air mass flow rate (bottom diagram, top gray curve with squares) and of the pressure drop over the test section (bottom diagram, bottom black curve) measured at a water mass flow rate of 0.3 kg/s and a pressure of 3.0 bar

(2) At an injected air mass flow rate of 0.30 kg/s (Fig. 3: $t=61.0$ s), a limitation of the discharge water flow is detected. Here the slope of the curve of the water level in the RPV simulator and SG separator begins to decrease and increase, respectively. This means that a part of the water injected in the SG separator does not flow to the side of the RPV simulator. This point is defined as the onset of flooding, and subsequent Region II as the partial delivery region. Around the onset of flooding, the pressure difference between the vessels begins to present higher fluctuations, as shown in the lower graph of Fig. 3. Visual observation indicates that at this point, the air/water interface becomes wavier and a large amplitude wave grows with droplet entrainment from its crest. This phenomenon was captured by the camera and is shown in Fig. 4(*b*). Due to the waves, the free cross section available for the air flow decreases. Therefore the air is accelerated above the wave; this starts to blow up the liquid slug, finally breaking up the slug into small droplets. With further increase in the air mass flow rate, the liquid slugs reduce and sometimes block the whole cross section for the air flow in the test section. Consequently, the pressure drop over the test section increases at these flow conditions ($\dot{m}_G=0.32$ kg/s), which is revealed by the pressure difference measured between the vessels ($t=76.6$ s in Fig. 3).

(3) With further increase in the air mass flow rate up to 0.34 kg/s (Fig. 3: $t=97.9$ s), the injected water mass flow rate of 0.3 kg/s is hindered to flow to the RPV simulator and the water level measured there shows a plateau (Region III). Therefore, the water remains completely in the test section and in the SG separator, where the water level rises. This region corresponds to the zero-liquid penetration. The visual observations indicate that large amplitude rolling waves are formed near the bend and block the cross section of the rectangular duct. Furthermore, large two-phase mixing regions were observed with droplet detachment at the wave crest and bubble entrainment in the bended region

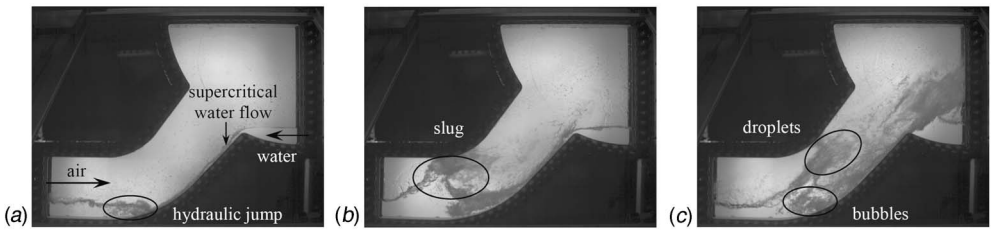


Fig. 4 Flow behavior during the counter-current flow of air and water at a water flow rate of 0.3 kg/s and a pressure of 3.0 bar. (*a*) $\dot{m}_G$=0.27 kg/s; *t*=21.00 s; (*b*) $\dot{m}_G$=0.32 kg/s; *t*=76.59 s; and (*c*) $\dot{m}_G$=0.34 kg/s; *t*=97.92 s.

due to the rolling flow pattern (see Fig. 4(*c*)).

3.2.2 Flow Behavior at a Water Flow Rate of 0.9 kg/s. Another experiment performed with a higher water mass flow rate of 0.9 kg/s is presented. The measured global parameters are shown in Fig. 5, which indicate also three main flow behaviors.

(1) In Region I of stable counter-current flow, the water level in the RPV simulator increases and the water level in the SG separator keeps constant. The pressure difference between the vessels is low and stable. From visual observations, it is noted that before the inception of flooding (t=78.0 s), the flow pattern is also a supercritical stratified flow, but no hydraulic jump is detected in the bended region, as shown in Fig. 6(*a*).
(2) At an injected air mass flow rate of 0.30 kg/s (Fig. 5: $t>78.0$ s), a limitation of the counter-current flow is detected, as marked as the onset of flooding in Fig. 5. Close inspection of the figure reveals that the behavior of water levels changes a little with increasing water mass flow rate. At the initiation of flooding, the slope of the water level in the RPV simulator begins to decrease. Meanwhile, the water level in the SG separator remains almost constant for about 5.5 s. This phenomenon is marked as A in Fig. 5 and means that a part of the liquid does not flow to either of the tanks. During this period of time, the high-speed camera images indicate that the water flow begins to be partially reversed. Therefore, at the initial stage of flooding, the flow in the horizontal part of the hot leg becomes subcritical and bigger waves are observed (Figs. 6(*b*)–6(*d*)). This causes an accumulation of water in the test section, which explains the decrease in the slope of the water level measured in the RPV simulator. Furthermore, no liquid is transported to the steam generator inlet chamber, neither from waves nor from droplets. Consequently, the water level in the SG separator remains relatively constant. In Region II, the behavior of the pressure drop in the test section is found to be similar to that at lower water mass flow rate ($\dot{m}_L$=0.3 kg/s), while the absolute maximum pressure difference measured increases with higher water flow rate.
(3) The zero-liquid penetration is reached with further increase

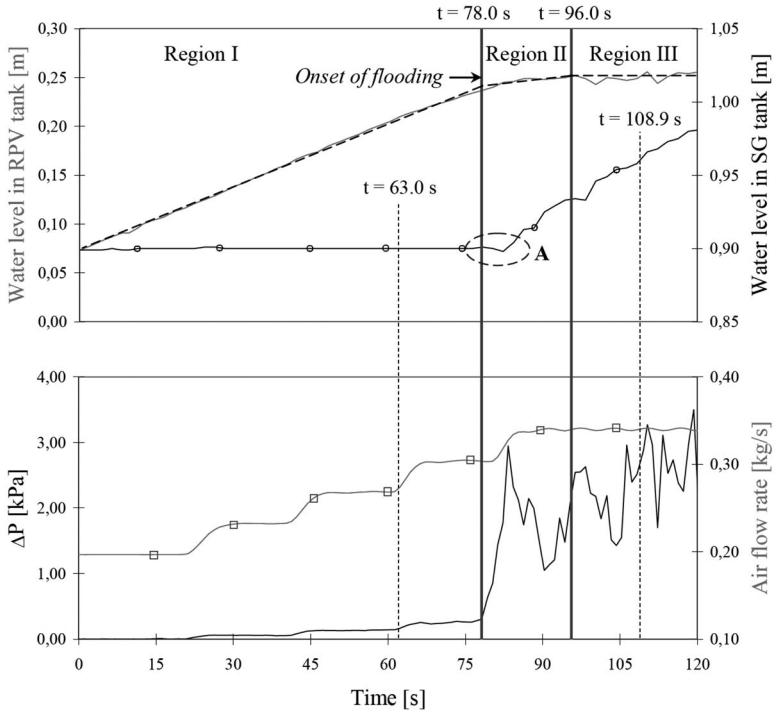


Fig. 5 Variation in the water levels in the SG separator (top diagram, bottom black curve with points) and in the RPV simulator (top diagram, top gray curve) of the pressure difference between the vessels (bottom diagram, bottom black curve) and of the air mass flow rate (bottom diagram, top gray curve with squares) measured at a water mass flow rate of 0.9 kg/s and a pressure of 3.0 bar

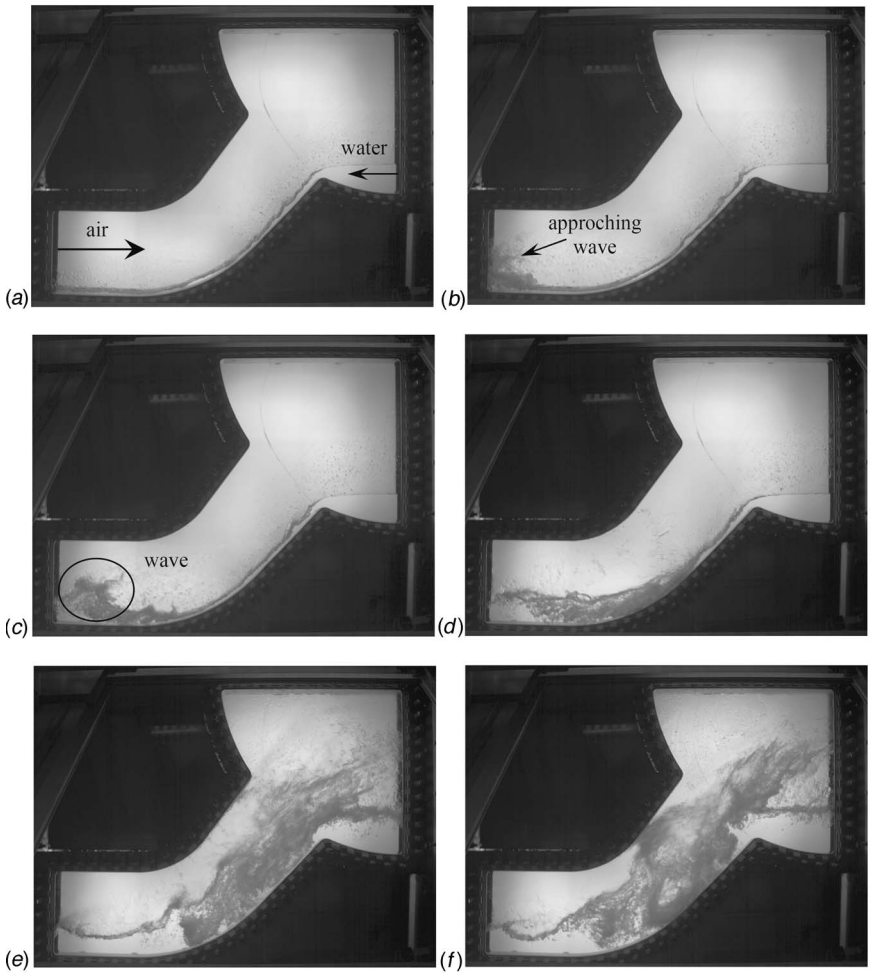


Fig. 6 Flow behavior observed during the counter-current flow experiment at a water flow rate of 0.9 kg/s and a pressure of 3.0 bar. (*a*) Before the onset of flooding ($\dot{m}_G$=0.27 kg/s; *t*=63.00 s); (*b*) at the onset of flooding ($\dot{m}_G$=0.30 kg/s; *t*=78.00 s); (*c*) $\dot{m}_G$=0.30 kg/s; *t*=78.60 s; (*d*) $\dot{m}_G$=0.30 kg/s; *t*=79.10 s; (*e*) $\dot{m}_G$=0.34 kg/s; *t*=96.00 s; and (*f*) $\dot{m}_G$=0.34 kg/s; *t*=108.90 s.

in the air mass flow rate up to 0.34 kg/s (Fig. 5: t=96.0 s). The visual observations (Figs. 6(*e*) and 6(*f*)) indicate that in Region III, the flow behavior is similar to that at lower water mass flow rate ($\dot{m}_L$=0.3 kg/s) with highly mixed zones.

3.2.3 General Flow Behavior. The general flow behavior in the hot leg model during counter-current flow limitation can be described from the experimental analysis presented in Secs. 3.2.1 and 3.2.2. Before the onset of flooding, a hydraulic jump is observed close to the bend at low liquid mass flow rate, whereas at higher water mass flow rate, no hydraulic jump appears in this region. However, at the onset of flooding, water always accumulates in the horizontal part of the channel due to a flow reversal at the interface. The high air velocity generates waves at the water surface with spume at the crest. A hydraulic jump appears at the transition between the down-flowing water from the SG separator and the accumulated water. Afterward, bigger waves develop to slugs, which induce the increase and the fluctuations of the pressure drop along the test section observed in Figs. 3 and 5.

These observations are similar to those obtained by Siddiqui et al. [3] and Wongwises [4] who examined the flow behavior during flooding in a horizontal pipe connected to a bend. This indicates that the flow behavior in hot leg typical geometries is not significantly affected by the cross-sectional geometry.

3.3 Influence of the System Pressure on the CCFL. In order to investigate the influence of the pressure on the counter-current flow limitation, similar experiments were performed at the same water flow rates and a lower system pressure of 1.5 bar. From the measured global parameters, each run was divided into regions according to the examples shown in Sec. 3.2. In each region, the flow rates of both fluids were evaluated for the arrangement of a flooding diagram.

For a meaningful comparison of these experiments, the effect of the pressure on the physical properties (i.e., air density) must be considered. Therefore, the dimensionless superficial velocity J_i^* (or Wallis parameter) is used to plot the flooding diagram. For the phase i, this is given by

$$J_i^* = j_i \cdot \sqrt{\frac{\rho_i}{g \cdot D \cdot (\rho_L - \rho_G)}} \tag{1}$$

where j is the superficial velocity, ρ is the density of the fluid, g is the gravitational acceleration, and D is the inner pipe diameter.

The Wallis parameter is convenient for all types of comparisons because this is a nondimensional parameter. However, it contains a length term D (Eq. (1)), which was commonly defined for pipes and not for channels with rectangular cross section. Hihara et al. [5], who examined the slugging of a counter-current gas/liquid flow in a horizontal rectangular channel, substituted this length

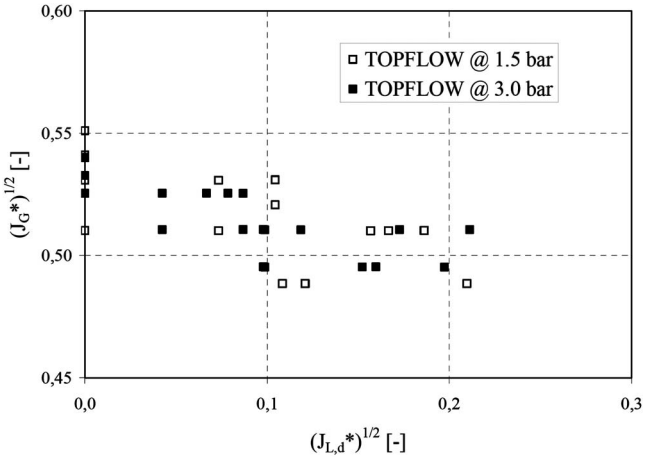


Fig. 7 Flooding diagram obtained for the hot leg model

term by the duct height H. Furthermore, this way is supported by the experimental results of Zapke and Kröger [6], who investigated counter-current flows in inclined and vertical rectangular ducts. From experiments in channels with different rectangular cross sections, they concluded that the flooding gas velocity depends only on the height of the channel and not on its width (see Fig. 7(*a*) in their paper). Therefore, in this work the Wallis parameter is calculated with the length term $H=0.25$ m.

The liquid Wallis parameter was calculated from the discharge water flow rate obtained for each run with the method presented in Sec. 3.2. In each region in which the runs were divided, the discharge water flow rate was evaluated from the gradient of the water level increase in the RPV simulator. In regions were CCFL was observed, the square root of the Wallis parameter was plotted for the gas versus the liquid, leading to the flooding diagram (Fig. 7).

First, the decreasing trend of the experimental points shows that the air flow rate required for CCFL decreases with increasing water flow rate. This trend was usually observed in horizontal, inclined, and also vertical conduit experiments as reported in literature by Wallis [7], Hewitt and Wallis [8], Zabaras and Dukler [1], Celata et al. [9], Choi and No [10], Zapke and Kröger [6], and Deendarlianto et al. [2].

Moreover, Fig. 7 demonstrates that the flooding is not affected by the system pressure when using the Wallis parameter as coordinates. The high-speed camera images confirm this statement. At 1.5 bar, the onset of flooding coincides with the reversal of the flow in the horizontal part of the hot leg due to high air velocities as observed at 3.0 bar (Sec. 3.2). After the flow reversal, liquid slugs are generated and develop near the bend.

4 Comparison of the Present CCFL Data With Other Studies

In this section, a comparison of the present CCFL data with different correlations for pipes available in literature is presented.

4.1 General Formulation of Flooding Correlations. Over the past decades, the prediction of CCFL was realized with flooding correlations deduced from experimental data. The most frequently used formulation of flooding correlations was given by Wallis [11] with the following expression:

$$(J_G^*)^{1/2} + m \cdot (J_L^*)^{1/2} = C \tag{2}$$

where J_i^* is the Wallis parameter of the phase i according to Eq. (1). C and m are constants determined from experimental results. Wallis derived the above correlation from the study of surface waves in vertical pipes, in which the onset of waves was associated with the occurrence of instabilities caused by flooding. In the case of the vertical conduits, Wallis [7] reported that the values of both constants m and C are equal to unity.

4.2 Correlations Obtained for Hot Leg Typical Geometry. In the past, several researchers studied the flooding in hot leg typical geometries constituted of a horizontal part connected to an inclined riser. They proposed special correlations for this geometry, which should be compared with our data.

Richter et al. [12] performed experiments in a simulator of the hot leg of a PWR with air and water. The inner pipe diameter and the length of the horizontal part were 0.203 m and 0.914 m, respectively. They published experimental flooding data in the region of the highest possible air mass flow or, consequently, to the lowest possible water mass flow. They correlated the obtained flooding data with the following equation:

$$(J_G^*)^{1/2} + (J_L^*)^{1/2} = 0.7 \tag{3}$$

Ardron and Banerjee [13] presented a correlation on the basis of a flooding model, assuming that the initiation of flooding coincides with the formation of slug flow in the horizontal pipe close to the bend. Since this model is very complex, they proposed the following empirical correlation:

$$(J_G^*)^{1/2} = 1.444 - 0.004 \cdot \lambda - \cosh(\lambda^{0.057} \cdot (Fl_p^*)^{-0.020} \cdot (J_L^*)^{0.7/2}) \tag{4}$$

where

$$\lambda = \frac{L_H \cdot (\mathrm{Re}^*)^{-0.2}}{D}$$

$$\mathrm{Re}^* = \frac{D}{\upsilon_G} \cdot \sqrt{g \cdot D \cdot \frac{\rho_L - \rho_G}{\rho_G}}$$

$$Fl_p^* = \frac{\upsilon_G}{\upsilon_L} \cdot \sqrt{\frac{\rho_G}{\rho_L}} \tag{5}$$

Here L_H is the length of the horizontal part and υ is the kinematic viscosity.

Ohnuki et al. [14] performed an experimental study on counter-current two-phase flows in a horizontal pipe connected to an inclined riser. They also proposed an empirical correlation to predict the onset of flooding by using the Wallis parameter, in which the effects of the length to the diameter of the horizontal pipe and the length of the inclined riser I are considered:

$$(J_G^*)^{1/2} + 0.75 \cdot (J_L^*)^{1/2} = \ln\left\{\left(\frac{L_H}{D}\right) \cdot \left(\frac{1}{I}\right)\right\}^{-0.066} + 0.88 \tag{6}$$

Later on, Lopez-De-Bertodano [15] also proposed a correlation based on the assumption that the initiation of flooding in the hot leg PWR coincides with the formation of slug flow:

$$(J_G^*)^{1/2} + 0.798 \cdot (J_L^*)^{1/2} = 0.619 \tag{7}$$

Furthermore, correlations were presented by Kang et al. [16] and Kim and No [17] to predict the CCFL in a simulated hot leg. Kang et al. obtained the following flooding correlation from their experimental work:

$$(J_G^*)^{1/2} + 0.397 \cdot (J_L^*)^{1/2} = 0.603 - 0.00234 \cdot \left(\frac{L_H}{D}\right) \tag{8}$$

Kim and No [17] proposed a correlation by regression through 356 experimental points of different researchers:

$$(J_G^*)^{1/2} + 0.614 \cdot (J_L^*)^{1/2} = 0.635 - 0.00254 \cdot \left(\frac{L_H}{D}\right) \tag{9}$$

Recently, Navarro [18] carried out experiments with air and water in small scale hot legs. The experimental work covered the effect of geometrical parameters on flooding, such as the inner pipe diameter, the ratio of length to diameter, the length, as well as

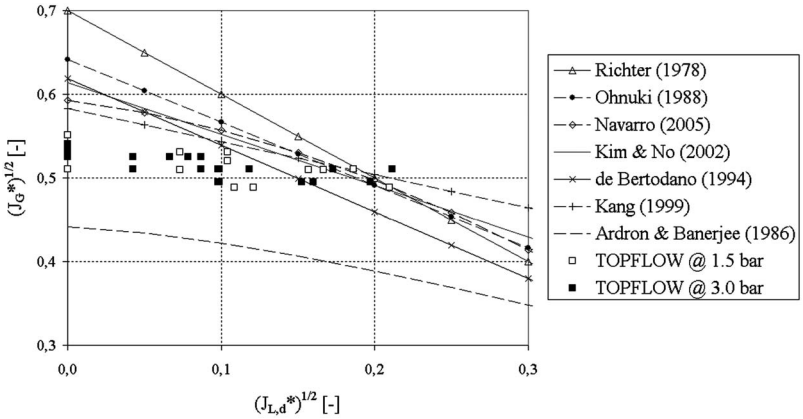


Fig. 8 Comparison of the present data with different CCFL correlations obtained for hot leg typical geometries

the inclination of the riser. As a result, he proposed an experimental correlation to predict the CCFL in the hot leg of a PWR:

$$(J_G^*)^{1/2} + 0.2452 \cdot (J_L^*)^{1/2} = 0.5963 - 1.17 \cdot J_L^* \qquad (10)$$

4.3 Comparison With This Work. The present data obtained at the hot leg model were compared with the empirical correlations mentioned in Sec. 4.2. The correlations including geometry parameters were calculated for the hot leg model according to the dimensions indicated in Fig. 1. As shown in the flooding diagram (Fig. 8), the present data come closer together with the correlations reported by Ohnuki et al., Lopez-De-Bertodano, Kang et al., and Kim and No, especially at higher liquid discharge flow rates. The trend indicated by the experimental points is similar to those found by the previous investigators, but the slope is lower. Therefore, the zero-liquid penetration (i.e., $(J_{L,D}^*)^{1/2}=0$) was obtained at lower values of the gaseous Wallis parameter.

One main difference could explain this deviation: All the flooding correlations found in literature are based on pipe experiments, whereas the hot leg model has a rectangular cross section. Therefore, the deviation revealed in Fig. 8 is probably caused by these geometrical differences.

5 Conclusions

The counter-current flow limitation was investigated experimentally in a horizontal rectangular channel connected to an inclined riser as a model of the hot leg of a pressurized water reactor. Counter-current flow limitation, or the onset of flooding, was found by analyzing the water levels measured in the separators. A confrontation with the high-speed observation images indicates that the initiation of flooding coincides with the reversal of the flow in the horizontal part of the hot leg due to high air velocities. After the flow reversal, the formation of liquid slugs was observed, which develop near the bend. This flow behavior around the onset of flooding is not significantly influenced by the water flow rate or the system pressure.

Furthermore, the CCFL data were compared with empirical correlations for analog geometries available in literature. This comparison shows that the Wallis parameter can be applied to rectangular cross sections by using the channel height as length, instead of the diameter. The obtained flooding curve is similar to those reported by other investigators, but its slope and the gaseous Wallis parameter for zero penetration are lower. This is attributed to the rectangular cross-sectional geometry of the hot leg model.

Nevertheless, this analysis of the data shows that it is suitable for CFD code validation. Image processing methods will be applied in order to extract local quantitative information from the high-speed camera images. This will be used for comparison with CFD simulations. Furthermore, similar experiments are planned with steam and water at saturation in order to investigate the phenomenon at reactor-typical boundary conditions.

Nomenclature

C = constant
D = inner pipe diameter (m)
g = gravitational acceleration (m/s^2)
H = channel height (m)
I = length of the inclined riser (m)
j = superficial velocity (m/s)
J^* = Wallis parameter
L_H = length of the horizontal part (m)
m = constant
$\dot{m}$ = mass flow rate (kg/s)
P = pressure (Pa)
t = time (s)
υ = kinematic viscosity (m^2/s)
ρ = fluid density (kg/m^3)

Acknowledgment

This work is carried out in the frame of a current research project funded by the German Federal Ministry of Economics and Labor, Project No. 150 1329.

The authors would like to thank Horst-Michael Prasser for initiating the project as well as the TOPFLOW team for their work on the test facility and the preparation of the experiments, namely, Heiko Pietruske, Peter Schütz, Klaus Lindner, Heiko Rußig, Marko Tamme, and Steffen Weichelt.

References

[1] Zabaras, G. J., and Dukler, A. E., 1988, "Counter-Current Gas-Liquid Annular Flow, Including the Flooding State," AIChE J., **34**(3), pp. 389–396.
[2] Deendarlianto, Ousaka, A., Kariyasaki, A., and Fukano, T., 2005, "Investigation of Liquid Film Behaviour at the Onset of Flooding During Adiabatic Countercurrent Air-Water Two-Phase Flow in an Inclined Pipe," Nucl. Eng. Des., **235**, pp. 2281–2294.
[3] Siddiqui, H., Banerjee, S., and Ardron, K. H., 1986, "Flooding in an Elbow Between a Vertical and a Horizontal or Near-Horizontal Pipe," Int. J. Multiphase Flow, **12**, pp. 531–541.
[4] Wongwises, S., 1996, "Flooding in a Horizontal Pipe With Bend," Int. J. Multiphase Flow, **22**, pp. 195–201.
[5] Hihara, E., Soejima, H., and Saito, T., 1985, "Slugging of Countercurrent Gas-Liquid Flow in a Horizontal Channel," Nippon Kikai Gakkai Ronbunshu (B)/Trans. Jpn. Soc. Mech. Eng., Ser. B, **51**, pp. 394–397.
[6] Zapke, A., and Kröger, D. G., 2000, "Countercurrent Gas-Liquid Flow in Inclined and Vertical Ducts—I: Flow Patterns, Pressure Drop Characteristics and Flooding," Int. J. Multiphase Flow, **26**, pp. 1439–1455.
[7] Wallis, G. B., 1969, *One-Dimensional Two-Phase Flow*, McGraw-Hill, New York.

[8] Hewitt, G. F., and Wallis, G. B., 1966, "Flooding and Associated Phenomena in Falling Film Flow in a Vertical Tube," UKAEA Report No. AERE R-4022.
[9] Celata, G. P., Cumo, M., Farello, G. E., and Setaro, T., 1991, "Hysteresis Effect in Flooding," Int. J. Multiphase Flow, **17**, pp. 283–289.
[10] Choi, K. Y., and No, C. H., 1995, "Experimental Studies of Flooding in Nearly Horizontal Pipes," Int. J. Multiphase Flow, **21**, pp. 419–436.
[11] Wallis, G. B., 1961, "Flooding Velocities for Air Water in Vertical Tubes," UKAEA Report No. AAEW-R 123.
[12] Richter H. J., Wallis, G. B., Carter, K. H., and Murphy, S. L., 1978, "Deentrainment and Counter-current Air-Water Flow in a Model PWR Hot Leg," U.S. Nuclear Regulatory Commission Report No. NRC-0193-9.
[13] Ardron, K. H., and Banerjee, S., 1986, "Flooding in an Elbow Between a Vertical and a Horizontal or Near-Horizontal Pipe. Part II: Theory," Int. J. Multiphase Flow, **12**, pp. 543–558.
[14] Ohnuki, A., Adachi, H., and Murao, Y., 1988, "Scale Effects on Countercurrent Gas-Liquid Flow in a Horizontal Tube Connected to an Inclined Riser," Nucl. Eng. Des., **107**, pp. 283–294.
[15] Lopez-De-Bertodano, M., 1994, "Counter-Current Gas-Liquid Flow in a Pressurized Water Reactor Hot Leg," Nucl. Sci. Eng., **117**, pp. 126–133.
[16] Kang, S. K., Chu, I.-C., No, H. C., Chun, M.-H., and Sung, C.-H., 1999, "Air-Water Counter-Current Flow Limitation in a Horizontal Pipe Connected to an Inclined Riser," Journal of Korean Nuclear Society, **31**, pp. 548–560.
[17] Kim, H. Y., and No, H. C., 2002, "Assessment of RELAP5/MOD3.2.2γ Against Flooding Database in Horizontal-to-Inclined Pipes," Ann. Nucl. Energy, **29**, pp. 835–850.
[18] Navarro, M. A., 2005, "Study of Countercurrent Flow Limitation in a Horizontal Pipe Connected to an Inclined One," Nucl. Eng. Des., **235**, pp. 1139–1148.

Effect of CANDU Bundle-Geometry Variation on Dryout Power

Laurence K. H. Leung
Atomic Energy of Canada Limited,
Chalk River,
ON, K0J 1J0, Canada

Dryout powers have been evaluated at selected inlet-flow conditions for two proposed designs of Canada deuterium uranium, CANDU® (a registered trademark of Atomic Energy of Canada Limited (AECL)) bundles and compared with those of the 37-element and CANDU Flexible, CANFLEX® (a registered trademark of AECL and Korea Atomic Energy Research Institutes (KAERI)) bundles. These proposed designs consist of a large center element (18 mm for one design and 20 mm for the other) and three rings of elements of 11.5 mm in outer diameter. The critical heat flux for each bundle design has been predicted using the correlation derived with Freon data obtained from the corresponding full-scale bundle test. An improvement in dryout power has been shown for the proposed design having a 20 mm center element with a radial power profile corresponding to the natural-uranium fuel as compared with other bundles, particularly the natural-uranium 37-element bundle, with a symmetric-cosine axial power profile. The dryout power improvement is further enhanced for the upstream-skewed axial power profile. [DOI: 10.1115/1.3043821]

1 Introduction

The AECL has been improving the CANDU fuel design for better dryout power and safety characteristics and lower void reactivity than the current natural-uranium 37-element fuel. The introduction of the CANDU flexible (CANFLEX) fuel has led to improved dryout power and safety characteristics. Full-scale bundle experiments have shown improvements in dryout power [1] and the reduction in linear element rating has lowered the fuel temperature for the CANFLEX bundle in uncrept and crept channels as compared with the 37-element bundle [2]. In addition, a low-void-reactivity fuel (LVRF) bundle has been developed, with burnable neutron absorber in the center element to reduce the void reactivity [3].

Further optimization of the CANFLEX fuel has been performed to achieve additional reduction in void reactivity and improvement in dryout power [4]. The amount of burnable neutron absorber has been increased in the center element and sizes of the inner-ring and center elements have been changed. The inner-ring element of the CANFLEX bundle has been reduced from 13.5 mm to 11.5 mm, and the center element has been increased from 13.5 mm to either 18 mm or 20 mm. These bundles are referred to as the CANFLEX-variant bundles (or simply the variant bundles from here on).

AECL has been using chlorofluorocompounds to study thermalhydraulic characteristics of CANDU fuel bundles since the 1970s. These compounds (or referred to as modeling fluids) have a much lower boiling point and latent heat of vaporization than water and hence can significantly reduce the cost and risk of boiling heat-transfer experiments. The reduction in pressure and power requirements allows the flexibility in testing and modification to the bundle simulator in support of new fuel designs. Thermalhydraulic parameters obtained with these modeling fluids are transformed into water-equivalent values based on the fluid-to-fluid modeling parameters to provide (or supplement) information in the analysis of a similar water-cooled system. Several series of Refrigerant-134a-cooled full-scale bundle experiments have been carried out to examine the impact of element-size variation on dryout power and critical heat flux (CHF). Correlations for CHF were derived with these data and have been applied to quantify the variation in dryout power for these proposed designs.

Dryout powers have been predicted using the corresponding CHF correlation for these variant bundles and compared with those of CANFLEX and 37-element bundles. The objective of this paper is to present the analysis result for the impact of CANDU bundle-geometry variation on critical heat flux and dryout power.

2 Experiments

Dryout-power experiments were performed with the test string installed inside the vertical test station of the MR-3 Freon loop at Chalk River Laboratories (CRL). Figure 1 shows the schematic of the loop and Fig. 2 illustrates the vertical test station and pressure-tap locations. A flow tube was inserted inside the test station to house the electric-heated simulator. The inside diameter of the flow tube was 103.45 mm.

The simulator has the overall configuration of a CANFLEX bundle, except that the center-element diameter has been increased and the diameter of the inner-ring elements has been reduced to 11.5 mm. Two sizes of center element (one of 18 mm and the other of 20 mm) were tested simulating two different designs. Figure 3 compares the geometry of the proposed designs to those of the 37-element and CANFLEX bundles. Element diameters are summarized in Table 1 for these bundle designs. Bundle junction simulators and sizes and configurations of most appendages (including AECL patented heat-transfer enhancing "buttons") were maintained the same as those of the CANFLEX bundle. Spacer heights on the center element and the inner-ring elements were adjusted to fit gaps between neighboring elements. The simulator was kept eccentric inside the flow tube simulating the gravity effect on the bundle in a horizontal channel using stainless-steel springs. These springs were welded on five bearing pads at the simulated top position of the bundle.

The simulator was heated electrically with a uniform axial power profile and had an overall heated length of 4.33 m (0.481 m in each bundle segment). It exhibited a nonuniform radial power profile simulating the fresh natural-uranium (NU) fuel. The nonuniform radial power profile was achieved using Inconel tubes of different wall thicknesses in the three rings (tubes in the same ring

Manuscript received August 11, 2008; final manuscript received August 14, 2008; published online January 6, 2009. Review conducted by Dilip R. Ballal. Paper presented at the 16th International Conference on Nuclear Engineering (ICONE16), Orlando, FL, May 12–15, 2008.

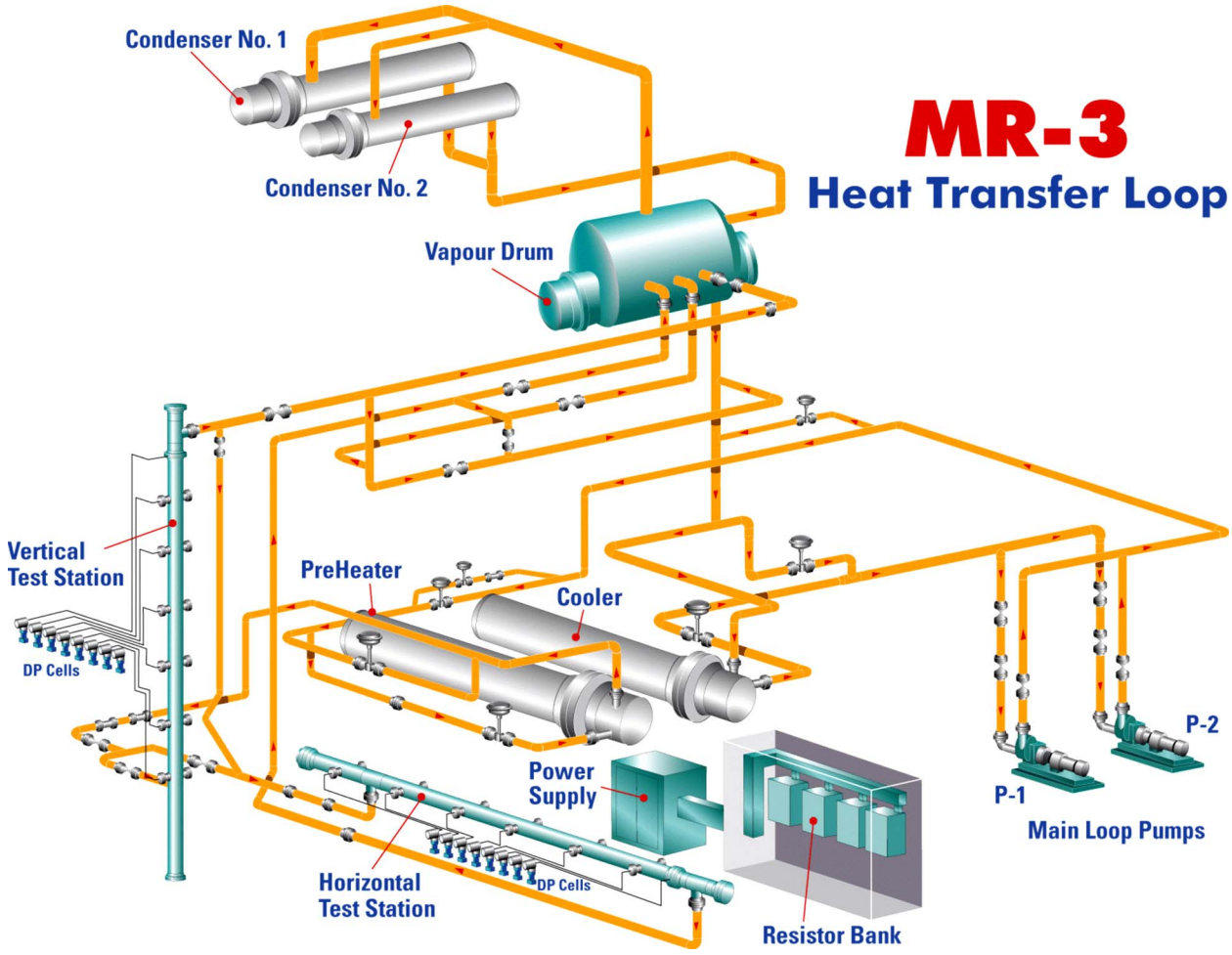


Fig. 1 Schematic of MR-3 heat transfer loop

have the same wall thickness). An external resistor bank was connected to various rings of elements providing the flexibility to change the radial power profile of the simulator. Radial power profiles corresponding to the fresh NU fuel and slightly enriched uranium (SEU) fuel at various burnups were tested.

These experiments covered mainly high pressure and high-flow conditions in Refrigerant-134a flow: outlet pressures of 1.84 MPa and 2.11 MPa; mass-flow rates of 12 kg s^{-1}, 15 kg s^{-1}, and 19 kg s^{-1}; and inlet-fluid temperature varied from 42°C to 59°C. Based on the fluid-to-fluid transformation factors [5], the water-equivalent values of outlet pressures are 11 MPa and 12.5 MPa; mass-flow rates are 17 kg s^{-1}, 21 kg s^{-1}, and 26 kg s^{-1}; and inlet-fluid temperatures are from 260°C to 300°C.

In each test run, the prescribed flow conditions (i.e., outlet pressure, mass-flow rate, and inlet-fluid temperature) were established. The power applied to the bundle simulator was increased gradually. Thermocouples installed inside all elements were rotated and moved axially to monitor inner-surface temperatures at each power. Dryout has been considered to occur when any one of these thermocouples exhibited a spike in temperature (of about 5°C). The corresponding power is identified as the dryout power for that test run.

3 Comparison of Experimental Dryout Powers and CHF Values

Dryout-power measurements obtained from these experiments were analyzed and compared against those of various bundle designs of the NU radial power profile. The local to average power ratios are 1.099, 0.899, 0.953, and 0.658 for outer-ring, middle ring, inner-ring, and center elements of the Variant-20 bundle. Those for elements of the Variant-18 and CANFLEX bundles are similar. Figure 4 compares dryout power measurements of the CANFLEX and the variant bundles at the outlet pressure of 1.84 MPa. The dryout power decreases with increasing inlet-fluid temperature and increases with increasing mass-flow rate. Dryout-power measurements of the Variant-18 bundle (i.e., variant bundle with an 18 mm center element) are the lowest among these bundles at the same flow conditions. Those of the Variant-20 bundle (i.e., variant bundle with a 20 mm center element) are the highest at the mass-flow rate of 18.9 kg s^{-1} but are about the same as dryout powers of the CANFLEX bundle at the mass-flow rate of 12.2 kg s^{-1}. Similar trends in dryout power measurements have been observed at the outlet pressure of 2.11 MPa.

CHF values have been established from dryout power measurements (i.e., CHF=(dryout power)/(heated area)) and examined against critical qualities (cross-sectional average values established from heat balance). Figure 5 compares CHF values of the CANFLEX and variant bundles at the outlet pressure of 1.84 MPa. The CHF decreases with increasing critical quality and increases with increasing mass-flow rate. As observed previously for dryout power measurements, CHF values of the Variant-18 bundle are the lowest among these bundles at the same local conditions. Those of the Variant-20 bundle, however, are the highest for both mass-flow rates. The increase in CHF for the Variant-20 bundle, as compared with the CANFLEX bundle, is attributed to the slight variation in heated area (which is larger for the CANFLEX bundle than the Variant-20 bundle). Similar CHF trends have been observed at the outlet pressure of 2.11 MPa.

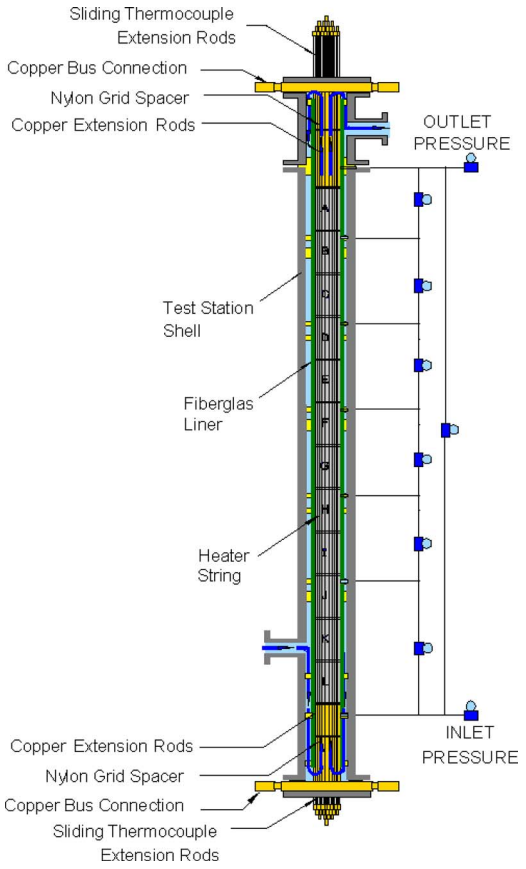


Fig. 2 Instrumentation setup at the vertical test station

4 Assessment of Dryout Powers for Various Bundle Designs in Reactor Fuel Channels

Dryout-power characteristics of various bundle designs in CANDU fuel channels (of 103.9 mm inside diameter) have been assessed at conditions of interest. The dryout power is established by comparing local (or boiling-length-average) heat flux against predicted CHF at local cross-sectional average conditions along the bundle string in the fuel channel. Local pressures are predicted using the pressure-drop model [6]. This assessment covers the following light-water conditions: outlet pressures of 10.5 MPa, 11.5 MPa, and 12.5 MPa; mass-flow rates of 17 kg s^{-1}, 21 kg s^{-1}, 26 kg s^{-1}, and 28 kg s^{-1}; and inlet-fluid temperatures of 260°C, 275°C, and 290°C.

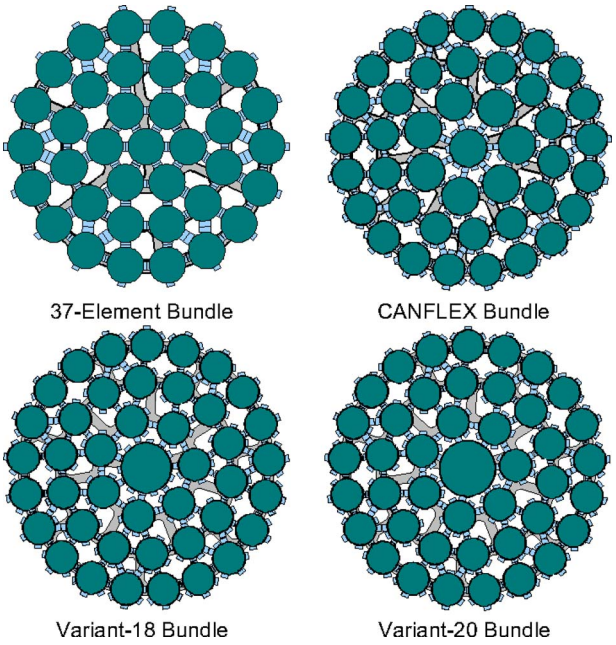


Fig. 3 Comparison of bundle geometry

Table 1 Summary of element diameters in bundle designs

Bundle	No. of element	Element diameter (mm) Center	Inner ring	Middle ring	Outer ring
37-Element	37	13.06	13.06	13.06	13.06
CANFLEX	43	13.5	13.5	11.5	11.5
Variant-18	43	18	11.5	11.5	11.5
Variant-20	43	20	11.5	11.5	11.5

The CHF correlation has been derived using experimental water-equivalent CHF values obtained with the corresponding design having a radial power profile of the NU fuel. It is expressed in the general form as

$$\mathrm{CHF} = (a_1 P^{a_2} G^{a_3} + a_4 P^{a_5} G^{a_6} x_{cr}) \times 1000 \tag{1}$$

where CHF is in kW m^{-2}, P is the local pressure in MPa, G is the local mass flux in Mg m^{-2} s^{-1}, and x_{cr} is the critical quality (all corresponding to cross-sectional average values). Water-equivalent values are transformed from Freon CHF values using fluid-to-fluid modeling factors [5]. The pressure is modeled using the density ratio, that is,

$$\left(\frac{\rho_f}{\rho_g}\right)_{\text{water}} = \left(\frac{\rho_f}{\rho_g}\right)_{\text{Freon}} \tag{2}$$

where ρ_f and ρ_g are the saturated liquid and saturated vapor densities in kg m^{-3}, respectively. The mass flux is modeled using the Weber number, that is,

$$\left(G\left(\frac{D_{\text{hy}}}{\sigma\ \rho_f}\right)^{0.5}\right)_{\text{water}} = \left(G\left(\frac{D_{\text{hy}}}{\sigma\ \rho_f}\right)^{0.5}\right)_{\text{Freon}} \tag{3}$$

where A_f is the flow area in m^2, G is the mass flux in kg m^{-2} s^{-1}, D_{hy} is the hydraulic-equivalent diameter of the bundle in meters, and σ is the surface tension in N m^{-1}. No specific modeling criteria are required for the critical quality, that is,

$$(x_{\text{cr}})_{\text{water}} = (x_{\text{cr}})_{\text{Freon}} \tag{4}$$

The CHF is modeled using the boiling number, that is,

$$\mathrm{Bo} = \left(\frac{\mathrm{CHF}}{G\ H_{fg}}\right)_{\text{water}} = \left(\frac{\mathrm{CHF}}{G\ H_{fg}}\right)_{\text{Freon}} \tag{5}$$

where H_{fg} is the latent heat of vaporization in kJ kg^{-1}. Table 2 lists optimized coefficients in the CHF correlation for each bundle design.

A separate set of experimental CHF data has been applied in establishing the CHF correlation for the 37-element bundle. Kalra and Ahmad [7] obtained these data in the same test facility using a 6 m long 37-element bundle with Refrigerant-12 flow. These data covered a wide range of flow conditions but only up to pressures of 11 MPa. Optimized coefficients for the CHF correlation are also listed in Table. 2.

Two separate axial power profiles along the bundle string have been implemented in the assessment (see Fig. 6). The symmetric-cosine profile corresponds to the eight-bundle-shift refueling scheme, while the upstream-skewed profile is relevant to either the four-bundle-shift or the two-bundle-shift refueling scheme [8]. The boiling-length-average heat-flux approach is applied to account for the effect of axial power profile on CHF. This approach

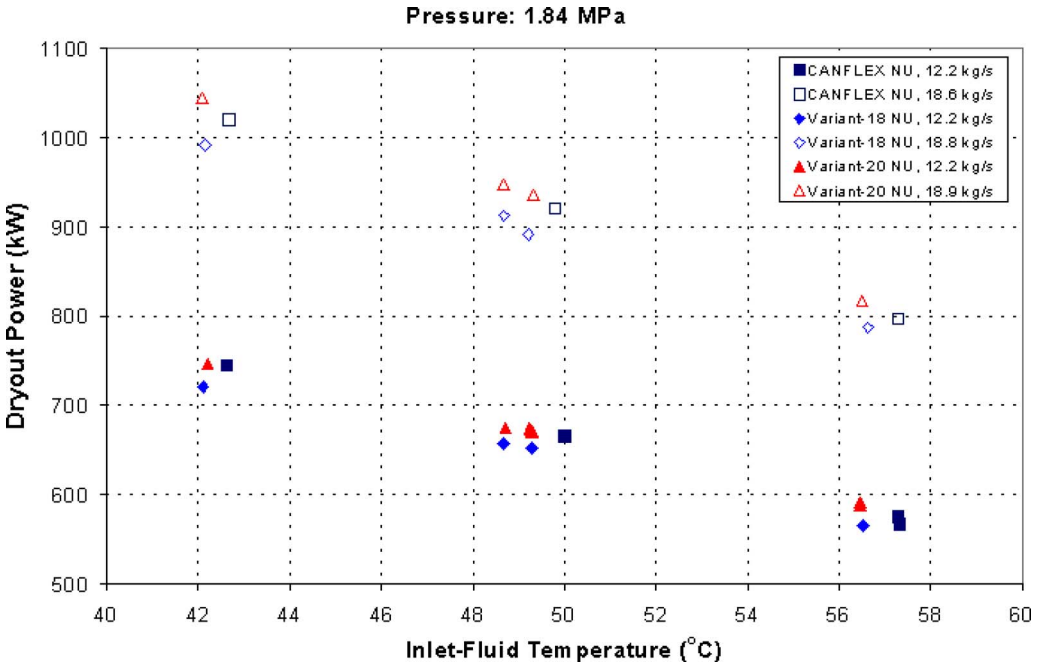


Fig. 4 Comparison of dryout power measurements

has been shown to consolidate CHF values for uniform and non-uniform axial power profiles in a variety of heated channels at the same local flow conditions [9].

4.1 NU Radial Power Profiles and Symmetric-Cosine Axial Power Profile. Figure 7 illustrates the predicted variation in dryout power with mass-flow rates for various bundle strings of the NU radial power profile having a symmetric-cosine axial power profile at the outlet pressure of 11.5 MPa and inlet-fluid temperature of 275°C. Predicted dryout powers for the 37-element bundle are systematically lower than those of the CANFLEX bundle. This is consistent with the comparison shown previously for these two designs [10]. The improvement in dryout power for the CANFLEX bundle is mainly attributed to the introduction of the additional nonload bearing appendages (i.e., "buttons"). All CANFLEX-type bundles exhibit better dryout power characteristics than the 37-element bundle. Predicted dryout powers for the Variant-20 bundle are about the same as those for the CANFLEX bundle, and those for the Variant-18 bundle are lower.

Dryout-power improvements, as compared with the 37-element bundle, for CANFLEX-type bundle designs are presented in Fig. 8 for the assessed range of flow conditions. As illustrated above, the improvement for the Variant-18 bundle is the smallest, while those for the CANFLEX and Variant-20 bundles are about the same. Overall, the outlet pressure has the strongest impact on the

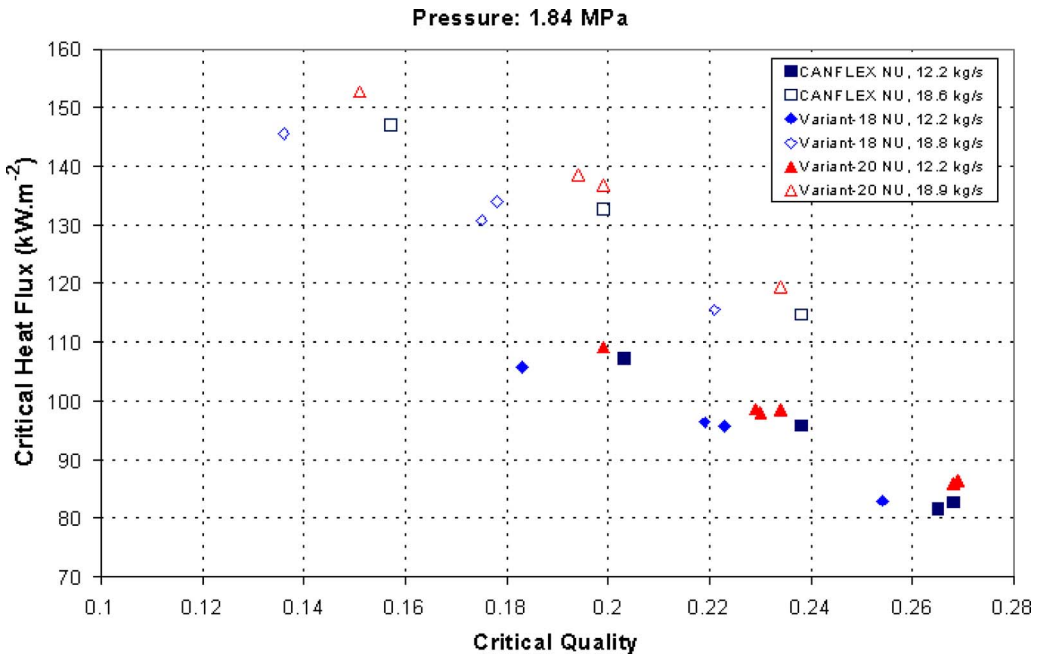


Fig. 5 Comparison of experimental CHF values

Table 2 Coefficients in CHF correlation

	37-Element	CANFLEX	Variant-18	Variant-20
a_1	12.0791	9.1330	1.4269	1.9598
a_2	−1.2439	−0.9068	−0.1857	−0.2931
a_3	0.6635	0.5118	0.5348	0.5682
a_4	−95.321	−156.93	−0.8586	−1.2990
a_5	−2.0928	−1.8239	0.2102	−0.0319
a_6	0.9083	0.5243	0.6134	0.8126

dryout power improvement. On average, the improvement increases from about 10% at the pressure of 10.5 MPa to 19% at the pressure of 12.5 MPa. As indicated previously, the CHF database for the 37-element bundle covers pressures of up to 11 MPa only. The observed improvement at the pressure of 12.5 MPa may be affected by the extrapolation of the CHF correlation for the 37-element bundle. The impact of inlet-fluid temperature on the dryout power improvement is also noticeable. The improvement decreases with increasing fluid temperature with an average variation of about 4%. The effect of mass-flow rate on the improvement is relatively small (within ±1%) and hence has not been identified separately.

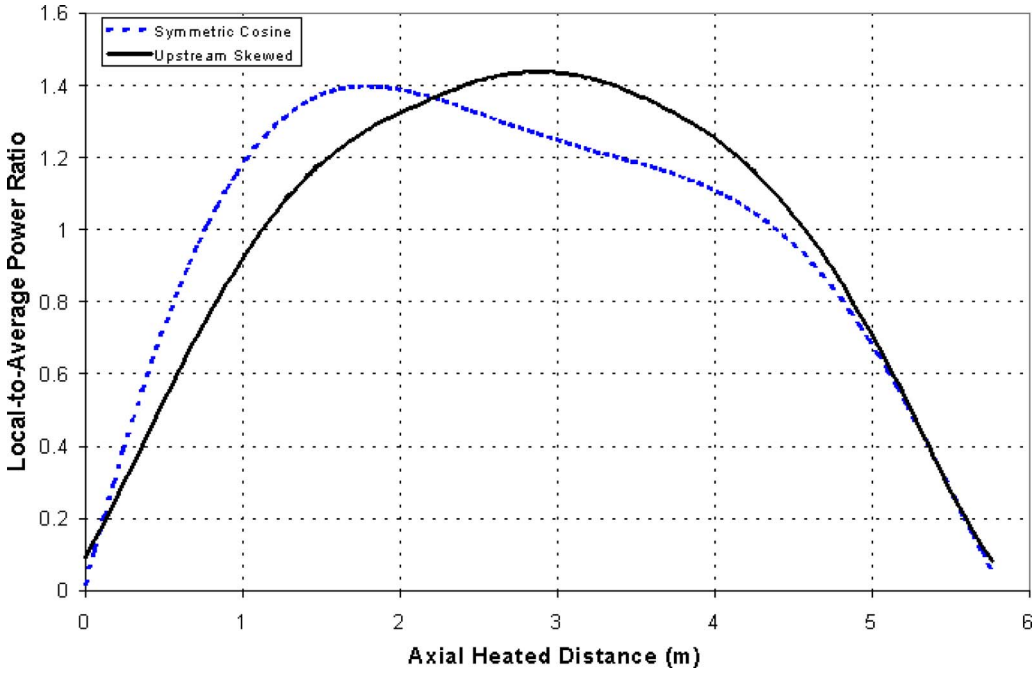


Fig. 6 Axial power profiles covered in the assessment

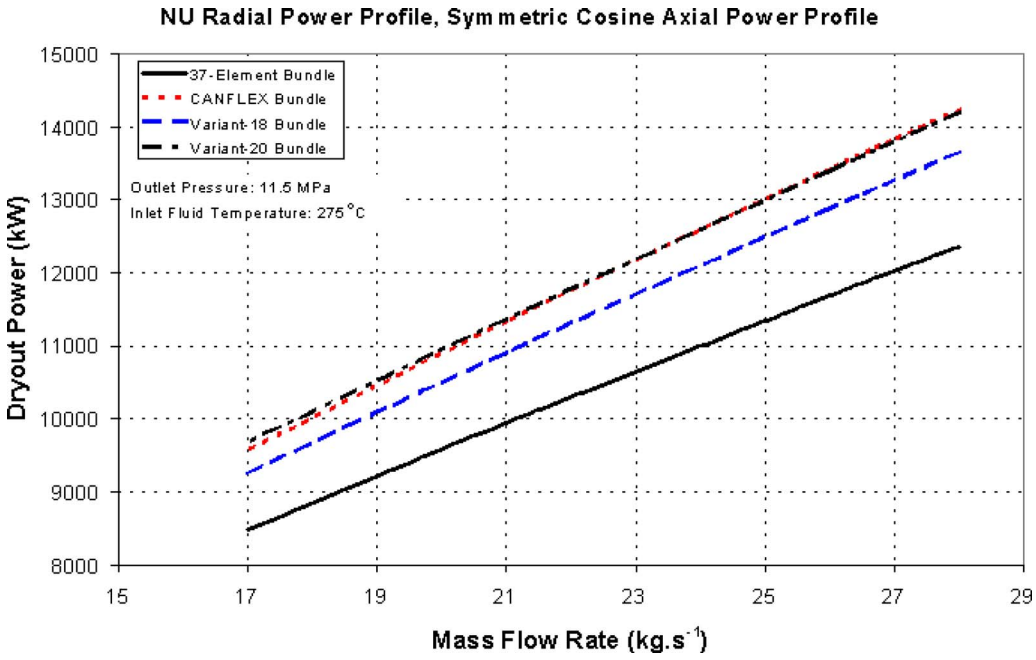


Fig. 7 Dryout power predictions for bundle strings of NU radial power profile and symmetric cosine axial power profile

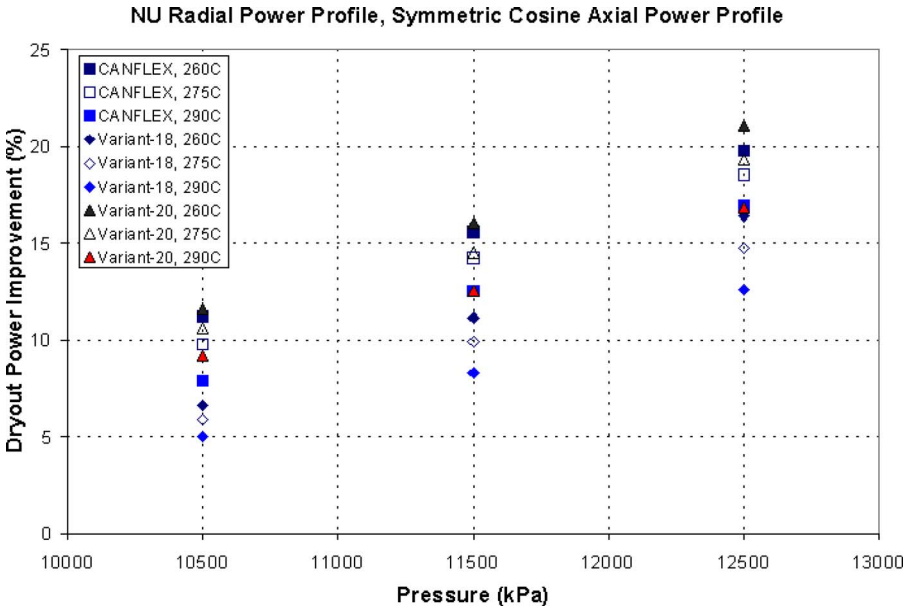


Fig. 8 Dryout power improvements for bundle strings of NU radial power profile and symmetric cosine axial power profile

4.2 NU Radial Power Profiles and Upstream-Skewed-Cosine Axial Power Profile. The introduction of SEU fuel may require a change in refueling scheme from the current eight-bundle shift in the C6 reactor to either four-bundle or two-bundle shift. This change affects the axial power profile along the bundle string in the fuel channel (as illustrated in Fig. 6) and consequently impacting the dryout power. An improvement in dryout power has been predicted for the upstream-skewed axial power profile due to the shifting of the peak power location to upstream locations with low coolant enthalpy and consequently reducing local heat fluxes at downstream locations where dryout is likely to occur [8].

Figure 9 compares predicted dryout powers for variant bundle strings with the NU radial power profile and an upstream-skewed axial power profile to those for CANFLEX and 37-element bundles with the NU radial power profile and a symmetric-cosine axial power profile. The change in axial power profile has led to further improvements in dryout power for variant bundle strings, compared to the 37-element bundle with the symmetric-cosine axial power profile. Dryout powers for the Variant-20 bundle with an upstream-skewed axial power profile are higher than those for the CANFLEX bundle with a symmetric-cosine profile. Similarly, the dryout power increases for the Variant-18 bundle with an upstream-skewed axial power profile and is comparable to those for the CANFLEX bundle with a symmetric-cosine profile.

Figure 10 illustrates dryout power improvements for variant bundle strings of the NU radial power profile and upstream-skewed axial power profile, compared with the 37-element bundle of the symmetric-cosine axial power profile. As mentioned above, a further improvement in dryout power has been observed for the

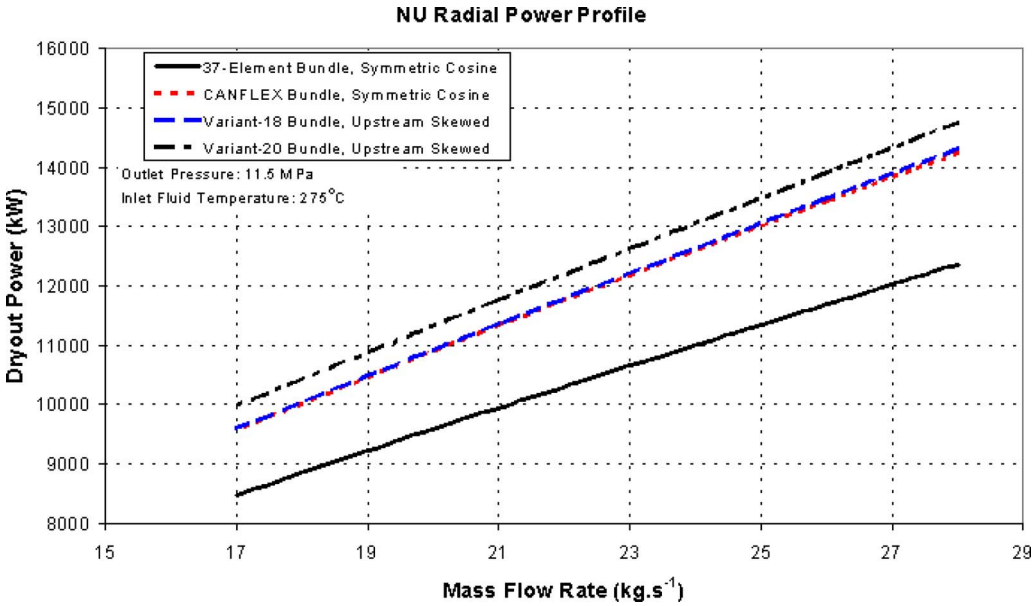


Fig. 9 Dryout power predictions for variant bundle strings of NU radial power profile and upstream-skewed axial power profile

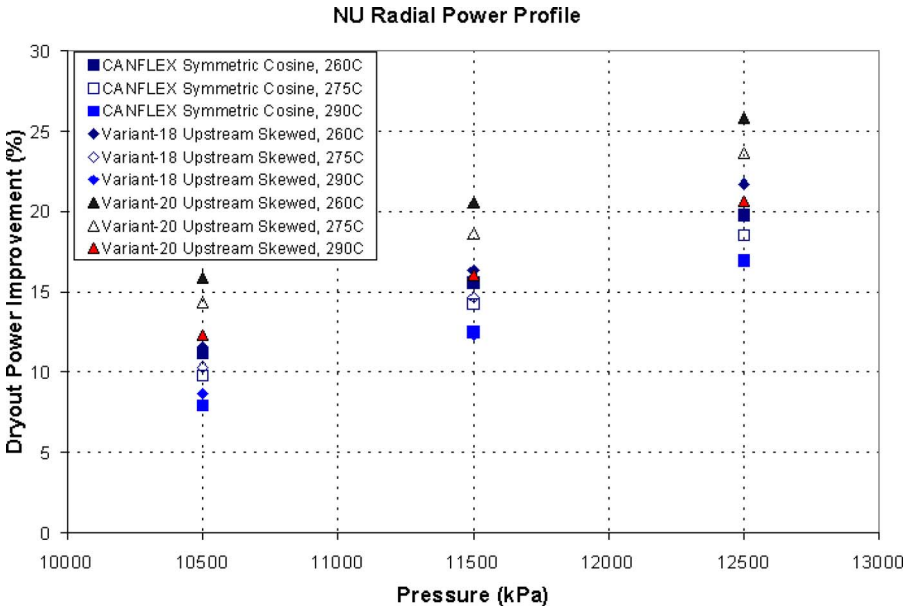


Fig. 10 Dryout power improvements for variant bundle strings of NU radial power profile having upstream-skewed axial power profile

Variant-20 bundle with the upstream-skewed axial power profile over the current range of assessed flow conditions. The dryout power improvement varies from about 14% at the pressure of 10.5 MPa to 23% at the pressure of 12.5 MPa. Therefore, the change in axial power profile has led to further improvement in dryout power by about 4% for the Variant-20 bundle.

The dryout power improvement for the Variant-18 bundle with the upstream-skewed axial power profile, on average, varies from 10% to 19% at pressures of 10.5 MPa and 12.5 MPa, respectively, compared with the 37-element bundle with the symmetric-cosine axial power profile. This range is similar to the dryout power improvement observed for the CANFLEX bundle with the symmetric-cosine axial power profile.

4.3 SEU Radial Power Profile and Upstream-Skewed-Cosine Axial Power Profile. The use of SEU fuel increases the burnup compared with the NU fuel. This has led to a relatively large variation in radial power profile at various locations along the bundle string in the fuel channel, impacting the CHF characteristic and dryout power performance. The variation in radial power profile depends on levels of enrichment and discharge burnup. One option being considered utilizes 2.4% SEU fuel in elements of various rings in the bundle with an average discharge burnup of about 20,000 MWd/TU. The center element contains only burnable neutron absorber in a Zircaloy matrix. It generates a small amount of power due to gamma heating despite the lack of uranium fuel. This power has not been modeled in the current assessment. The variation in radial power profile for the corresponding bundle string at several locations of the fuel channel is shown in Fig. 11. In addition, the radial power profile for the CANFLEX fresh NU fuel bundle is shown in the same figure for comparison.

A steep radial power profile is observed at the first bundle location (i.e., Bundle 1) of the string, which contains the fresh fuel. The radial power profile gradually flattens, with the local relative heat flux decreases at outer-ring elements and increases at middle ring and inner-ring elements, along the string. It approaches the fresh NU fuel profile of the CANFLEX bundle at the last bundle (i.e., Bundle 12). The CHF for the 2.4% SEU bundle string is anticipated to be lower than that for the CANFLEX fresh NU fuel bundle due to the consistently higher local relative heat flux at the outer-ring elements (where dryout has been observed in the full-scale bundle test [1]). This would have an impact to the dryout power prediction.

The effect of radial power profile on CHF is predicted using the methodology described previously for the CANFLEX bundle [11]. It accounts for the deviation of the local relative heat flux at each element ring from the optimum profile, which corresponds to the simultaneous dryout at one or more elements in all the rings. A bundle exhibiting the optimum profile represents the most balanced-fuel configuration leading to the best CHF characteristic. CHF ratios with respect to the fresh NU fuel are calculated using the local radial power profile at various bundle locations and are shown in Fig. 12. These ratios are applied together with the CHF correlation for the fresh NU fuel (e.g., Eq. (1)) to determine local CHF values at various SEU bundle locations.

The dryout power performance for variant bundles with 2.4% SEU fuel in element rings is better than that for the 37-element bundle (see Fig. 13). However, the improvement has been reduced as compared with the NU fuel (see Fig. 9) due to the anticipated effect of the radial power profile. Nevertheless, the dryout power for the Variant-20 bundle of 2.4% SEU radial power profile remains higher than that for the CANFLEX bundle of NU radial power profile. Similarly, the dryout power for the Variant-18 bundle has been reduced and is lower than that for the CANFLEX bundle.

Figure 14 shows dryout power improvements for Variant bundles of 2.4% SEU fuel having an upstream-skewed-cosine axial power profile. The average improvement varies from 11% to 20% at pressures of 10.5 MPa and 12.5 MPa, respectively, for the Variant-20 bundle, and from 6% to 16% for the Variant-18 bundle. Overall, the improvement is about 1% higher than that observed for variant bundles of NU fuel having a symmetric-cosine axial power profile shown in Fig. 8. Therefore, the dryout power improvement from changing the axial power profile has compensated the loss from the introduction of the 2.4% SEU fuel.

5 Conclusions and Final Remarks

Dryout-power measurements for several bundle designs have been obtained using axially uniform-heated simulators cooled with Refrigerant-134a flow. Those for the variant bundle having a 20 mm center element are generally the highest as compared with

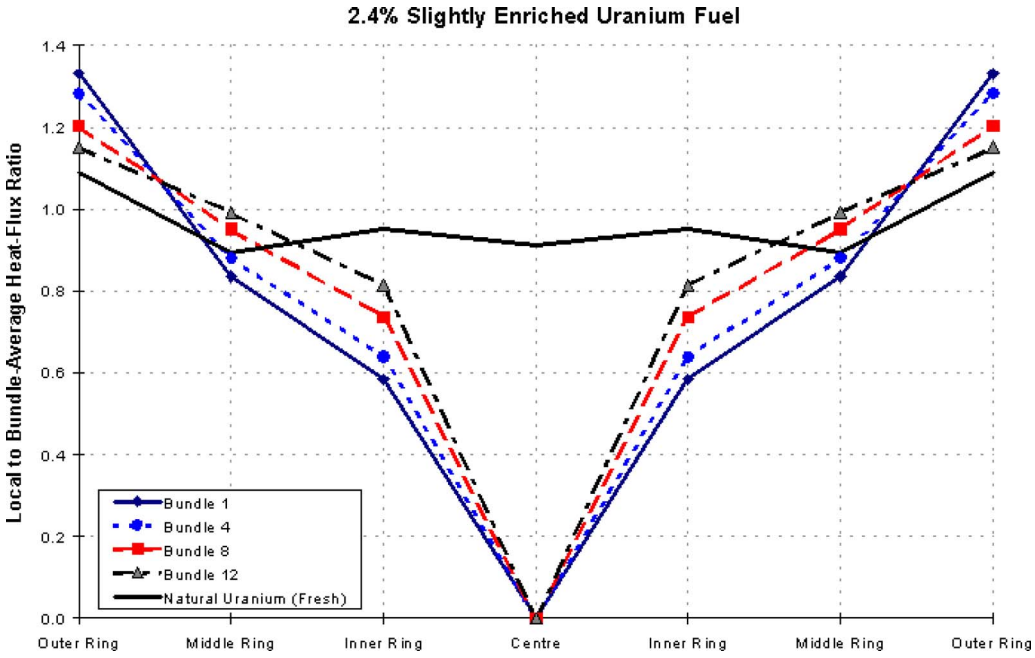


Fig. 11 Variation in radial power profile at locations of 2.4% SEU bundle string

37-element and other CANFLEX-type bundles. CHF values have been evaluated using dryout power measurements and transformed into water-equivalent values, which were used to develop a CHF correlation for each bundle design. A separate CHF correlation for the 37-element bundle has been derived using the previous experimental data obtained with Refrigerant-12 flow.

The dryout power performance has been assessed for various bundle designs with radial power profiles corresponding to NU and 2.4% SEU fuels, and symmetric-cosine and upstream-skewed-cosine axial power profiles at high pressure and high-flow conditions. The variant bundle having a 20 mm center element exhibits generally the best dryout power performance as compared with 37-element and CANFLEX-type bundles. The improvement in dryout power increases (by about 6%) for the upstream-skewed-cosine axial power profile as compared with the symmetric-cosine shape. The introduction of a 2.4% SEU fuel in element rings has led to relatively large variation in the radial power profile at various locations of the bundle string in the fuel channel. Consequently, the dryout power improvement has been reduced by about 3%.

Flow conditions covered in the Freon bundle experiments are relatively limited (mainly at high flows and high pressures). Additional dryout power measurements are required to extend CHF correlations for reactor analyses of postulated accident scenarios.

The current assessment focuses only on the impact of element-diameter variation on dryout power. Additional geometry optimization (e.g., varying the pitch-circle diameter of the element ring) has been examined using the ASSERT subchannel code to enhance

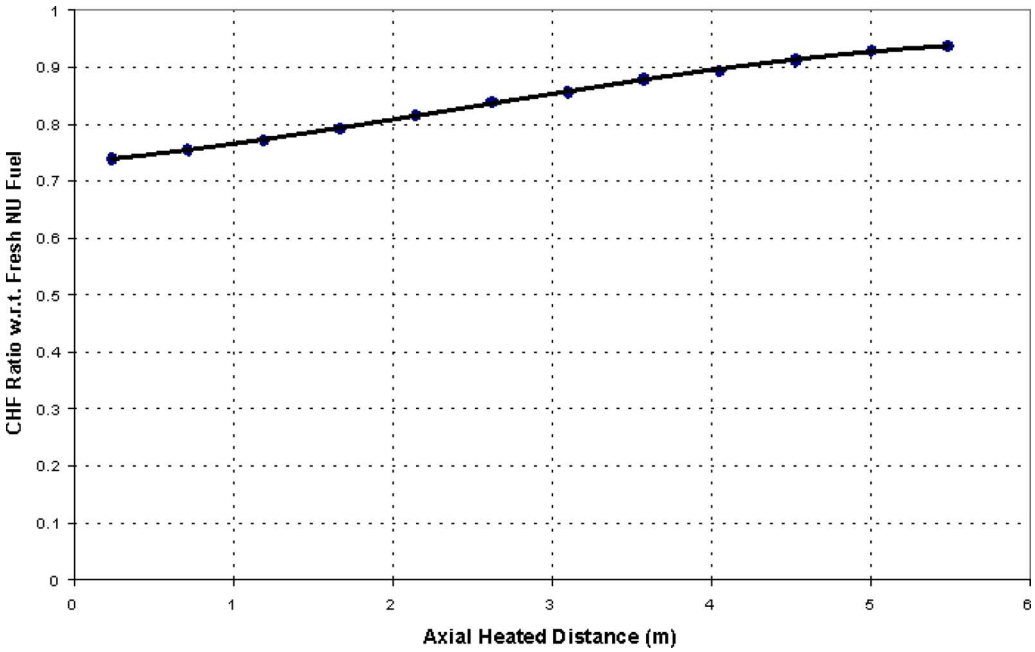


Fig. 12 Predicted CHF ratios with respect to fresh NU fuel at locations of 2.4% SEU bundle string

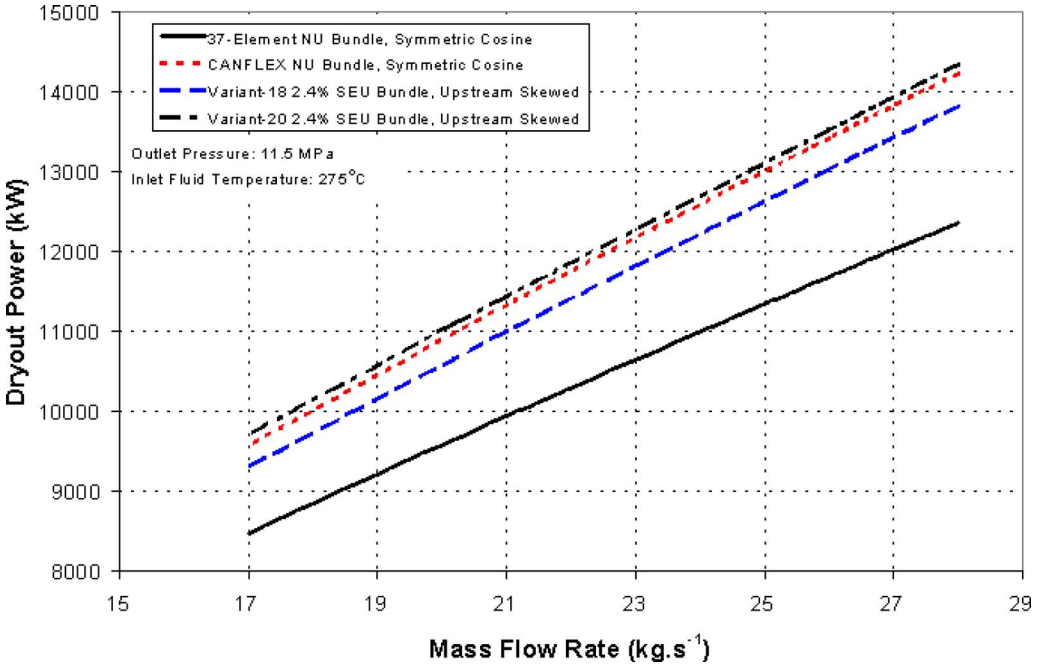


Fig. 13 Dryout power predictions for variant bundle strings of 2.4% SEU radial power profiles and upstream-skewed axial power profile

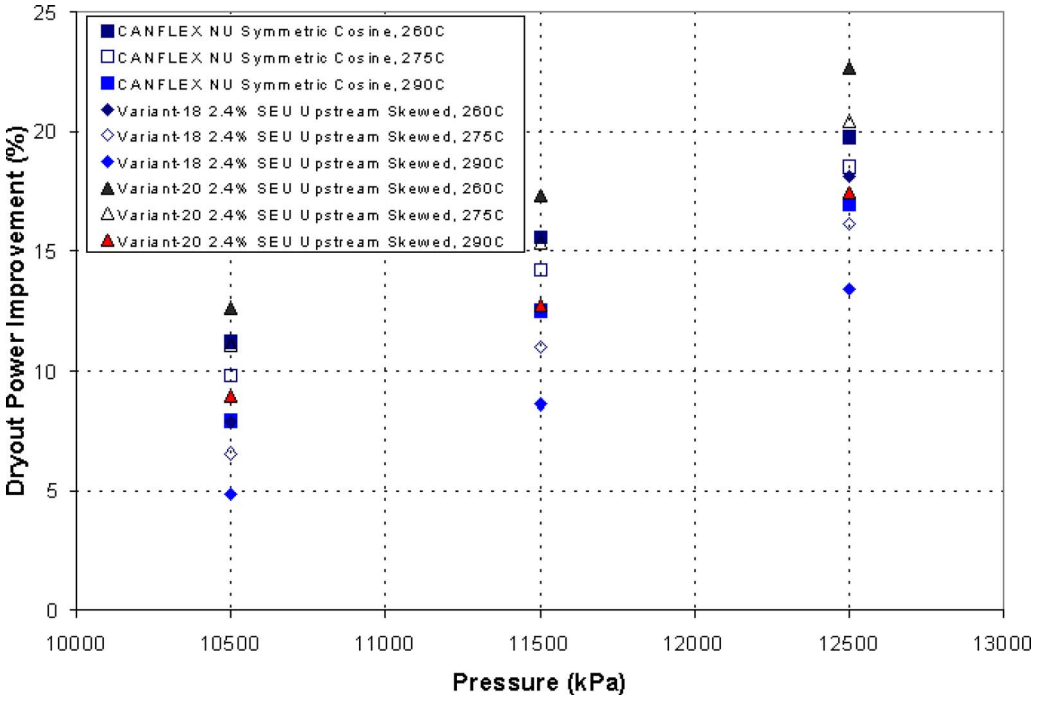


Fig. 14 Dryout power improvements for variant bundle strings of 2.4% SEU radial power profile having upstream-skewed axial power profile

the dryout power performance [4]. This would lead to an improved design of the variant bundle.

Acknowledgment

The author would like to express his appreciation to D.E. Bullock and H. Rummens for performing the CHF experiments and E.N. Onder for analyzing the experimental data and for commenting on the manuscript.

Nomenclature

A_f = flow area (m^2)
D_{hy} = hydraulic-equivalent diameter (m)
CHF = critical heat flux (kW m^{-2})
H_{fg} = latent heat of vaporization kJ kg^{-1}
P = local pressure (MPa)
G = local mass flux (Mg m^{-2} s^{-1})
x_{cr} = critical quality

ρ_f = saturated liquid density (kg m^{-3})
ρ_g = saturated vapor density (kg m^{-3})
σ = surface tension N m^{-1}

References

[1] Leung, L. K. H., Jun, J. S., Dimmick, G. R., Bullock, D. E., Inch, W. W. R., and Suk, H. C., 2001, "Dryout Power of a CANFLEX Bundle String With Raised Bearing Pads," *Proceedings of the Seventh International Conference on CANDU Fuel*, Kingston, Canada, Sept. 23–27.

[2] Lane, A. D., Boczar, P. G., Carter, T. J., Groeneveld, D. C., and Sejnoha, R., 1990, "The Program to Develop and Demonstrate the CANFLEX Bundle, Its Additional Capabilities and Supporting Technologies," *Proceedings of the 11th Annual Canadian Nuclear Society Conference*, Toronto, Canada, Jun. 3–6.

[3] Lau, J. H., Doria, F. J., Chow, H., Leung, L. K. H., Palleck, S., Sim, K. S., He, Z., and Purdy, P., 2004, "Design and Qualification of the Bruce CANFLEX—Low Void Reactivity Fuel (LVRF)—An Overview," *Proceedings of the 25th Annual Canadian Nuclear Society Conference*, Toronto, Canada, Jun. 6–9.

[4] Rao, Y., and Leung, L. K. H., 2007, "Thermalhydraulics Performance Optimization of CANDU Fuel Using ASSERT Subchannel Code," *Proceedings of the 2007 International Congress on Advances in Nuclear Power Plants (ICAPP2007)*, Nice, France, May 13–18, Paper No. 7549.

[5] Leung, L. K. H., and Groeneveld, D. C., 2000, "Fluid-to-Fluid Modelling of Critical Heat Flux in 37-Element Bundles," *Proceedings of the 21st Nuclear Simulation Symposium*, Ottawa, ON, Sept. 24–26.

[6] Snoek, C. W., and Leung, L. K. H., 1989, "A Model for Predicting Diabatic Pressure Drops in Multi-Element Fuel Channels," Nucl. Eng. Des., **110**(3), pp. 299–312.

[7] Kalra, S. P., and Ahmad, S. Y., 1979, "Critical Heat Flux Measurements in a Vertical Multi-Element Segmented Cluster," *Proceedings of the Fluid Flow and Heat Transfer Over Rod or Tube Bundles*, ASME, New York, Dec. 2–7.

[8] Leung, L. K. H., Rudzinski, K. F., and Chan, P. S., 2000, "The Dryout-Power Improvement of CANFLEX SEU Bundles in CANDU Reactors," *Proceedings of the 2000 Annual Canadian Nuclear Society Conference*, Toronto, ON, Jun. 11–14.

[9] Leung, L. K. H. and Groeneveld, D. C., 2006, "Critical Heat Flux in Axially Non-Uniform-Heated Channels," *Proceedings of the 13th International Heat Transfer Conference*, Sydney, Australia, Aug. 13–18.

[10] Leung, L. K. H., Jun, J. S., Dimmick, G. R., Bullock, D. E., Inch, W. W. R., and Suk, H. C., 2001, "Dryout Power of a CANFLEX Bundle String With Raised Bearing Pads," *Proceedings of the Seventh International Conference on CANDU Fuel*, Kingston, Canada, Sept. 23–27.

[11] Leung, L. K. H., 2002, "A Prediction Method for the Effect of Radial Heat-Flux Distribution on Critical Heat Flux In CANFLEX Bundles," *Proceedings of the 22nd CNS Nuclear Simulation Symposium*, Ottawa, ON, Nov. 3–5.

Akira Yamaguchi[1]
e-mail: yamaguchi@nucl.eng.osaka-u.ac.jp

Takashi Takata
e-mail: takata_t@nucl.eng.osaka-u.ac.jp

Graduate School of Engineering,
Osaka University,
2-1, Yamada, Suita,
Osaka 565-0871, Japan

Hiroyuki Ohshima
Japan Atomic Energy Agency,
4002, Narita, O-arai,
Ibaraki 311-1393, Japan
e-mail: ohshima.hiroyuki@jaea.go.jp

Yoshitaka Kohara
Graduate School of Engineering,
Osaka University,
2-1, Yamada, Suita,
Osaka 565-0871, Japan
e-mail: kohara_y@qe.see.eng.osaka-u.ac.jp

Yoshihiro Deguchi
Mitsubishi Heavy Industry,
5-717-1, Fukahori-machi,
Nagasaki 851-0392, Japan
e-mail: yoshihiro_deguchi@mhi.co.jp

Numerical Prediction and Optimization of Depressurized Sodium-Water Reaction Experiment With Counterflow Diffusion Flame

Sodium-water reaction (SWR) is a design basis accident of a sodium-cooled fast reactor (SFR). A breach of the heat transfer tube in a steam generator results in contact of liquid sodium with water. Typical phenomenon is that the pressurized water blows off, vaporizes, and mixes with the liquid sodium. It is necessary to quantify the SWR phenomena in the safety evaluation of the SFR system. In this paper, a new computer program has been developed and the SWR in a counterflow diffusion flame is studied by a numerical simulation and an experiment. The experiment is designed based on the numerical simulation so that the stable reaction flame is maintained for a long time and physical and chemical quantities are measured. From the comparison of the analysis and the experiment, there exist discrepancies that may be caused by the assumptions of the chemical reaction. Hence, a new experiment is proposed to enhance the measurement accuracy and to investigate the reason of the disagreement. The authors propose a depressurized experiment and show the preliminary result of the experiment. It is found that a stable chemical reaction flame is formed. With the depressurization, it is expected that the flame location can be controlled and the reaction region becomes thicker because of decrease in the reactant gas density. [DOI: 10.1115/1.3043822]

1 Introduction

1.1 Background and Purpose. In a sodium-cooled fast reactor (SFR), liquid sodium is used as the heat transfer fluid to carry the energy from the reactor core to the steam generation (SG) system. The liquid sodium has an excellent heat transport capability and a large safety margin to the boiling point (1153 K) at the atmospheric pressure. On the other hand, it has chemical reactivity in contact with water vapor. One of the design basis accidents of the SFR is the water leakage into the liquid sodium through a heat transfer tube of SGs, i.e., a sodium-water reaction (SWR).

Consequently, coupled phenomena of the chemical reaction and the thermal-hydraulics of sodium and water vapor are of importance from the safety viewpoint. Large-scale experiment series of SWR were performed in Japan [1]. However, the SWR phenomena are generally complex and the experimental measurement technology is not well matured. Therefore, a numerical simulation is used to investigate the coupled phenomena and local quantities such as mass concentrations, flow velocities, and temperatures.

According to Takata et al. [2], two chemical reaction types are considered, i.e., a surface reaction and a gas-phase reaction. In the initial phase of the SG tube failure accident, the temperature is well below the sodium boiling point and little sodium vapor exists. Therefore, the water vapor and the liquid sodium react at the gas-liquid interface. It is the surface reaction that occurs at the liquid sodium surface. Subsequently, the liquid sodium is heated up by the exothermic reaction. At this stage, sodium vaporizes and the sodium vapor and the water vapor encounter. It results in the gas-phase reaction. In general, most of the sodium remains a liquid phase because of a large amount of the liquid sodium inventory.

This paper deals with a numerical simulation and a SWR experiment in the gas phase. It is advantageous to study the gas-phase reaction to understand the characteristics of the SWR. In the gas-phase reaction, one can calculate the spatial distributions of the reactants and product including airborne particulates as well as a temperature by solving conservation equations. For validating the numerical method, a counterflow diffusion flame experiment was performed [3]. Numerically obtained spatial distributions of quantities are compared with the experimental measurement. According to the comparison, discrepancies are found between the experiment and the analysis. Here a new experiment is proposed to enhance the measurement accuracy of the chemical and thermal quantities, that is, an experiment in depressurized conditions.

1.2 Sketch of SWR. The pressure of the water side is approximately 17 MPa and that of the sodium side is 0.2 MPa. The temperatures in the sodium side are 625 K and 743 K for the cold leg and the hot leg, respectively, of the heat transport system. The sodium is in liquid phase because the boiling temperature is greater than the system temperature by 500 K. A breach of a heat transfer tube in the SG results in contact of the liquid sodium with the water. Therefore, sudden influx of the water and steam into the liquid sodium takes place. The water changes into the vapor phase by sudden depressurization in quite a short time. Typical phenomenon is that the pressurized water and steam blow off and mix with the liquid sodium where many heat transfer tubes exist. At

[1]Corresponding author.

Manuscript received August 12, 2008; final manuscript received August 22, 2008; published online January 6, 2009. Review conducted by Dilip R. Ballal. Paper presented at the 16th International Conference on Nuclear Engineering (ICONE16), Orlando, FL, May 12–15, 2008.

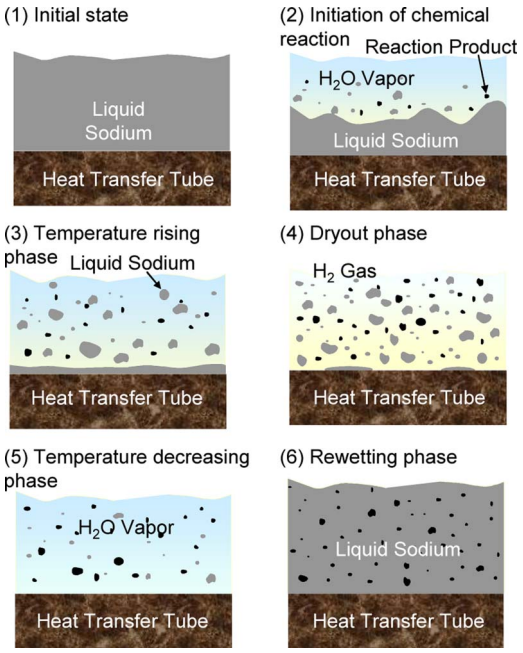


Fig. 1 Sketch of the sodium-water reaction phases

the interface of the liquid sodium and the water vapor, the SWR takes place and reaction heat and reaction products, i.e., sodium hydroxide and hydrogen, mostly are generated.

A sketch of the reaction zone around the tube surface is shown in Fig. 1. At the interface of the bulk sodium and the water vapor jet, the relative velocity is large. Hence it is reasonable to consider that liquid sodium droplets are entrained into the water vapor flow. In the initial phase, the tube is surrounded by the liquid sodium as shown in Fig. 1(1). When the water vapor flows into the liquid sodium and reaches the tube surface, the water vapor pushes the liquid sodium aside. But liquid film remains on the tube surface (Fig. 1(2)). In this phase, the SWR takes place and the region is heated up and sodium hydroxide is generated at the liquid sodium surface or in the gas phase as aerosols. As the reaction proceeds and the temperature increases, the liquid film becomes thinner (Fig. 1(3)) and comes to the dryout of the tube surface (Fig. 1(4)). In this situation, the sodium runs out and the SWR goes down.

In safety consideration, what we are interested in is the tube temperature in the temperature rising phase and the dryout phase. As mentioned above, the reaction is dominant near the interface of the water vapor and the bulk liquid sodium. As the reaction proceeds, the region is covered with the reaction product hydrogen gas (see Fig. 1(4)). It is a mechanism that mitigates the temperature increase. Subsequently, the water vapor replaces the hydrogen and the temperature begins to decrease if the water leakage continues (Fig. 1(5)). There may be still a small amount of liquid droplets of sodium since the ejecting water vapor entrains the liquid sodium droplets and carry them to the zone. The SWR continues, but it is mild and the temperature transient is benign. After the water leakage ceases, liquid sodium flows in again and the tubes are cooled down (Fig. 1(6)). At this stage, the SWR process is terminated.

2 Computational Model

In the computer program, Navier–Stokes equations and chemical reaction equations are solved interactively. In addition, a dynamic equation of airborne particulates is coupled with the basic equations of thermal-hydraulics. A source of the particulates is the chemical reaction products, i.e., sodium hydroxide and sodium oxide. Details of the computational models are given in Ref. [3]. In this section the authors describe the governing equations and the computational modeling briefly.

2.1 Governing Equations. Governing equations for thermal-hydraulics are described. Taking the Reynolds average of the conservation equations, one obtains time-averaged equations as follows:

For mass conservation equation,

$$\frac{\partial \rho}{\partial t} + \mathrm{div}\ \rho \bar{\mathbf{u}} = \sum_k S_k \tag{1}$$

For momentum conservation equation,

$$\frac{D\bar{u}_i}{Dt} = -\frac{1}{\rho}\mathrm{grad}\ \bar{p} + \mathrm{div}(\nu\ \mathrm{grad}\ \bar{u}_i - \overline{u_i' u_j'}) + g_i - \frac{\bar{u}_i}{\rho}\sum_k S_k \tag{2}$$

For conservation equation of chemical species,

$$\frac{D\rho Y_k}{Dt} = \mathrm{div}(\rho D_k\ \mathrm{grad}\ Y_k) + S_k \tag{3}$$

For energy conservation equation,

$$\frac{D(\rho C_p \bar{T})}{Dt} = \mathrm{div}(\lambda\ \mathrm{grad}\ \bar{T} - \rho C_p \overline{u_i' T'}) + F + Q_R \tag{4}$$

The explanation of the variables is given in the Nomenclature. $\bar{\cdot}$ and $\cdot'$ indicate the time average and fluctuating parts of a quantity, e.g., $u_i = \bar{u}_i + u_i'$. $S_k = \partial(\rho Y_k)/\partial t$ is the production term of species k by the chemical reaction.

Assuming the gradient-diffusion hypothesis, the Reynolds stress $\overline{u_i' u_j'}$ and the turbulent heat flux $\overline{u_i' T'}$ are defined using the turbulent viscosity and the turbulent thermal diffusivity. The equations for the turbulence energy and the energy dissipation ratio are based on those for the low-Reynolds number two-equation model proposed by Yang and Shih [4], which is a variation of Jones and Launder's model [5] and describes the near-wall effect. A zero-equation model is used with regard to the temperature field. According to the recommendation by Nagano and Tsuji [6], Wassel–Catton equation is used for the turbulent Prandtl number. The correlations and the parameters appearing in the turbulence model are given by Yang and Shih [4] and Nagano and Tsuji [6].

2.2 Aerosol Dynamics. For evaluating the aerosol behavior, the evolution of the size distribution is calculated by solving the following general dynamics equation:

$$\frac{\partial n(v)}{\partial t} + \mathrm{div}\left(\frac{n(v)}{C_u(v)}\mathbf{u}\right) + \mathrm{div}(n(v)\mathbf{U}) = \frac{1}{2}\int_0^v n(v')n(v-v')\beta(v', v-v')dv' - n(v)\times\int_0^\infty n(v')\beta(v, v')dv' + S(v) \tag{5}$$

Aerosols are transferred in a thermal-hydraulic field by convection, diffusion, thermophoresis, diffusiophoresis, and gravity fallout. In the transfer processes of the aerosol mass, coagulation and sedimentation are major mechanisms of the change in the aerosol size distribution. The Brownian movement, the velocity difference of two aerosol particles, and the turbulence are the mechanisms for the coagulation. Gelbard and Seinfeld [7] gave details on the modeling of the coagulation process, which is used in the present model.

The second term in the left hand side of Eq. (5) is the advection term due to the gas velocity. In this equation, $C_u(v)$ is the Cunningham correction factor to explain slip effect for smaller diameter particles than the mean free path of the gas. The third term in

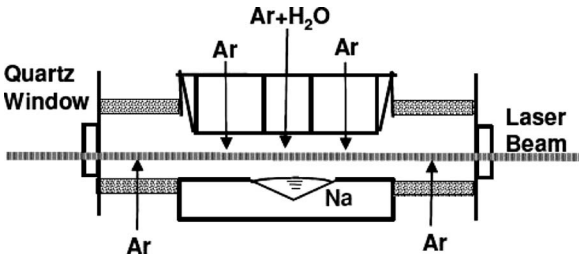


Fig. 2 Sodium-water counterflow diffusion flame experiment

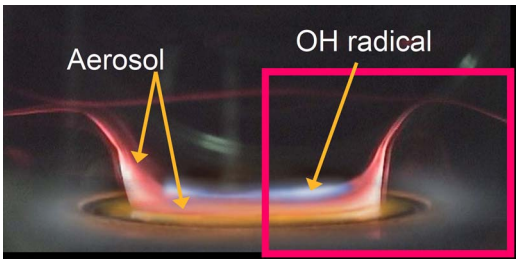


Fig. 3 Visualization of sodium-water reaction

the left side of Eq. (5) represents the mass transfer caused by applied forces to the aerosols. **U** is the terminal velocity vector of a particle, which is expressed as

$$\mathbf{U} = \mathbf{U}_{\text{dif}} + \mathbf{U}_g + \mathbf{U}_{tp} + \mathbf{U}_{dp} \tag{6}$$

where the velocities with subscripts dif, *g*, *tp*, and *dp* are contributions from the Brownian diffusion, the gravity, the thermophoresis, and the diffusiophoresis forces, respectively.

The aerosol source term is obtained based on the reaction rate evaluated by the chemical reaction model. One assumes that the statistical distribution of the primary aerosol diameter follows a lognormal distribution. The volume fraction of the aerosols is little and the interaction of the aerosols and the air flow can be neglected in the gas dynamics calculation.

2.3 Radiation Heat Transfer. We assume one-dimensional radiation heat transfer because we are dealing with the pool surface phenomena. Using the gas temperature T_i in the *i*th computational cell numbered from the pool surface in the vertical direction, one obtains the following equations regarding the volumetric heat generation rate Q_i of an individual cell:

$$q_i = Q_i L_i = \varepsilon_i G_p \prod_{j=1}^{i-1} (1-\varepsilon_j) + \varepsilon_i \sum_{j \neq i} \left\{ \varepsilon_j \sigma T_j^4 \prod_{k=i+1}^{j-1} (1-\varepsilon_k) \right\} - 2\varepsilon_i \sigma T_i^4 \tag{7}$$

$$G_p = \varepsilon_p \sigma T_p^4 + (1-\varepsilon_p)\sigma \sum_{j=1}^{n} \varepsilon_j T_j^4 \prod_{k=1}^{j-1} (1-\varepsilon_k) \tag{8}$$

Here L_i and T_i are the cell size and the local gas temperature, respectively, in the *i*th computational cell. G_p in the first term of the right hand side of Eq. (7) is the leaving flux from the pool surface and is given by Eq. (8). The pool surface emissivity is influenced by the surface conditions and ε_p=0.65 is assumed according to the measurement by Hashiguchi et al. [8].

The total gas mixture emissivity including aerosols is evaluated by Modak's [9] radiative property model and is applied to each computational cell.

$$\varepsilon_i = \varepsilon_a + \varepsilon_{\text{gas}} - \varepsilon_a \varepsilon_{\text{gas}} \tag{9}$$

where ε_{gas} and ε_a are the gas and the aerosol emissivity, respectively. ε_{gas} is calculated from the emissivity of the individual nontransparent gas with a correction for the overlapping of absorption bands from different gases. It can be assumed that the emissivity is equal to the absorptivity for aerosols. ε_a is evaluated according to Felske and Tien [10] as a function of the local value of gas temperature, the mass density, and the volume fraction of the aerosols.

3 SWR Experiment and Code Validation

3.1 SWR in Counterflow Diffusion Flame. To understand the SWR phenomena and to validate the numerical simulation method, a counterflow laminar combustion experiment of SWR was performed by Yamaguchi et al. [3]. Figure 2 shows the schematics of the sodium-water counterflow experiment.

The outline of the experiment is as follows. A liquid sodium pool is heated up to 800 K (less than the boiling point by 253 K) so that the sodium evaporates slowly. Through the nozzle above the liquid pool surface, water vapor diluted to 0.7 vol % by argon gas flows out toward the sodium pool surface. A strainer is installed at the exit of the mixture gas flow to provide flat velocity distribution. Hence a counterflow diffusion flame is formed between the pool surface and the flow nozzle. The reaction zone exists in between the sodium pool and the strainer. It is important to keep the water vapor density at a very low value for visualizing the flame zone. Otherwise, we cannot see the experimental cell inside because it is filled with white smoke (reaction product aerosol of sodium hydroxide).

Figure 3 is the photo of the reaction zone as well as the aerosol mass concentration taken in the experiment. The sodium pool diameter is 0.03 m. In the experiment, the existence of OH radical is measured by the laser-induced fluorescence (LIF) technique. Bright region above the pool shows the OH radical fluorescence intensity. In the reaction, the water molecule decomposes first of all to produce OH radical, which bonds with the Na molecule. Therefore, the chemical reaction takes place in the region where OH fluorescence intensity is high. The region is located 2 mm from the liquid sodium pool surface. The thickness of the reaction zone is approximately 1 mm.

The aerosol concentration is visualized by the scattering of the laser beam from the aerosol particles. We can see the red light emission from the levitating aerosols. The aerosol existing area is like a "crown" shape. This state is maintained for a long time duration and is very stable. Figure 4 shows the numerical results of the aerosol mass concentration. The numerical simulation is performed in two-dimensional geometry assuming axis symmetry.

In Figs. 3 and 4, we can see that the aerosol is dense above the liquid sodium pool. They move upward at the pool edge and are removed from the reaction zone. The flow field is in a mixed condition of forced and natural convections. Then the flow field and aerosol distributions are influenced by the temperature distribution as well. Also we can see that the stable reaction flame is established. Since the saturated vapor pressure of sodium is very low, the flame location tends to adhere to the pool surface. It is important to lift the flame off the pool surface for measurements with an enhanced accuracy.

However, trial-and-error process by a series of experiments is not practical to find out the best solution. Thus, the experimental condition (the geometry of the cell, boundary condition of flow

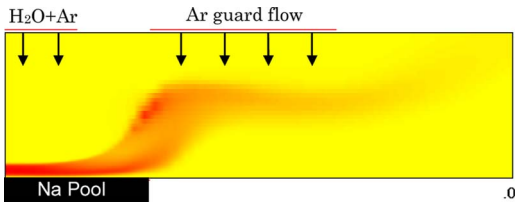
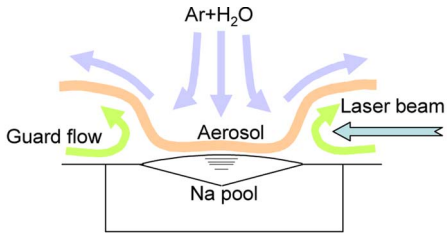
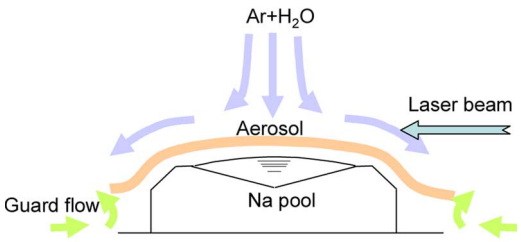


Fig. 4 Aerosol mass concentration



(a) Original geometry of the sodium pool



(b) Modified geometry of the sodium pool

Fig. 5 Modification of the sodium pool: (*a*) original geometry of the sodium pool and (*b*) modified geometry of the sodium pool

velocities, and temperatures) is optimized based on numerical simulations. To lift the flame up, double guard flows at different radial coordinates are used as shown in Fig. 2. One is an inner downward argon flow surrounding the H_2O+Ar flow. The inner guard flow is used to confine the water vapor in the above-pool region. Another is an outer upward argon flow located at the peripheral of the experiment cell. The outer guard flow plays an important role of lifting up the reaction flame.

With this process, the flow and the temperature fields surrounding the flame are adjusted so that the flame zone is detached away from the pool surface and the reaction products are removed from the reaction zone. The combination of the two guard flows is found to be appropriate to establish the optimal flow field for continuous chemical reaction and clear visualization and measurement. From the comparison of the numerical simulation and the experimental observation, the performance of the present simulation is satisfactorily accurate and the optimization procedure of the flow and temperature and the experimental design works well. The dilution of water vapor and two guard flows enable sufficiently long duration time of the reaction. Accordingly one can extend the reaction duration time as long as 1 h and achieve steady state reaction flame, and no change in the liquid sodium level is observed during 5 min measurement. It is long enough for the laser measurement and visualization of the thermal-hydraulic field.

3.2 Modification of Sodium Pool Geometry. It is noted from Fig. 3 that the reaction zone is closed to the pool surface and a dense aerosol region exists above the sodium pool peripheral. The laser beam is irradiated in the lateral direction permeating through the quartz window as shown in Fig. 2. Therefore there is a concern that the thick aerosol region may deteriorate the transparency or visibility from the window. Thus a modification of the geometry of the edge of the sodium pool is considered.

Figure 5(*a*) shows the original geometry and schematic flow pattern as well. A qualitative explanation of the flow field is given in the following. The temperature of the outer guard flow is kept lower than the inner guard flow and the upward flow velocity is small. Therefore the outer guard flow direction turns to horizontal toward the pool center. Then it turns upward flow at the peripheral of the sodium pool. The upward guard flow is useful to remove the reaction product aerosol to the upper right of the figure and to make the flame visible. However, the aerosol region is stretched

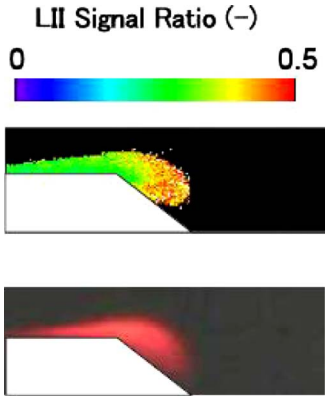


Fig. 6 Photo of the laser scattering by aerosols (bottom) and relative size of the aerosols measured by LII (top)

upward at the pool peripheral, which may deteriorate the side view. Therefore, the edge of the pool is cut off as shown in Fig. 5(*b*) and the flow pattern is changed by adjusting the water vapor flow and the guard flow conditions. We expect the aerosol distribution as in Fig. 5(*b*) and better visibility from the window.

The bottom of Fig. 6 is the photo of the aerosols in the experiment with the modified geometry in Fig. 5(*b*). The water vapor concentration is 1.6 vol % and $Ar+H_2O$ flow velocity is 0.12 m/s. It is seen that the upward stretching of the aerosol at the pool peripheral is suppressed and is similar to what the authors expected in advance. The numerical simulation results of the aerosol density for the same condition as the experiment are shown in Fig. 7. It can be said that the aerosol distribution is in agreement with the experiment although quantitative comparisons are necessary in the future.

The laser-induced incandescence (LII) measurement result is shown in the top of Fig. 6. The LII signal intensity was measured twice at 200 ns and 300 ns after the laser irradiation. The ratio of the two LII signals is shown in Fig. 6 (top). The aerosol particle temperature rises by the laser irradiation followed by the cooling down as time goes. Hence the ratio of the two LII signal intensities lies between zero and unity. Large particles with high thermal capacity are cooled down slowly and the LII signal ratio is close to unity. On the other hand, small particles with a little mass heat capacity are cooled down fast and the LII signal ratio approaches to zero. Therefore, the LII signal ratio corresponds to the particle diameter. Although quantitative measurement is not performed yet, we understand that the particle grows and the size becomes larger in the pool peripheral from Fig. 6.

In the numerical simulation, the diameter of the incipient aerosol generated by the chemical reaction is assumed to be 0.5 μm. The aerosol diameter is divided into groups from 0.5 μm to 100 μm in lognormal scale. A small aerosol is subjected to the forces dominant in a small scale such as thermophoresis and diffusion [11]. Being driven by the forces, they move upward or downward before they coagulate to grow larger. Large aerosols follow the macroscopic phenomena such as a gas flow.

Figure 8 shows the LIF measurement result and the OH radical

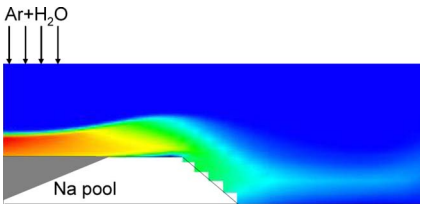


Fig. 7 Mass concentration of aerosols

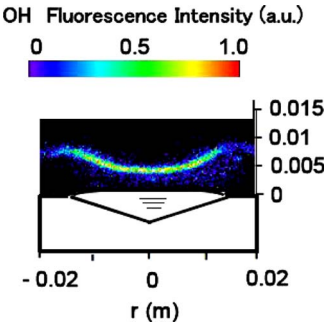


Fig. 8 Concentration of the OH radical measured by LIF

concentration in arbitrary unit. We see that the reaction zone exists at 4 mm above the pool surface at the pool center. If the existence of the OH radical corresponds to the reaction region, the reaction region lies above the dense aerosols. It is noted that the reaction zone is axisymmetric, which justifies the assumption in the numerical simulation.

Numerical results of the mass concentration distribution of the sodium and the water vapor along the vertical axis at the pool center are shown in Fig. 9 to investigate the chemical reaction and the mass concentration in the reaction zone. The solid lines indicate the computational results and the symbols indicate the experimental measurements. From the simulation results, the reaction flame locates at 2 mm distant from the pool surface, where their mass concentrations approach zero. No sodium vapor exists beyond the combustion flame and no water vapor below it as an infinite reaction rate or a flame sheet concept is assumed.

On the contrary, in the measurement the sodium vapor and the water vapor are almost consumed and disappear at 3 mm from the pool surface. There is a discrepancy between the experiment and the numerical simulation in the flame location. From Fig. 8, the OH radical locates at 4 mm above the pool, which is more consistent with the experimental results in Fig. 9. Further investigation is needed regarding the assumption of chemical reaction rate constant, the reaction process, and/or the reaction products.

3.3 Proposal of Depressurized Experiment. To investigate the reason of the discrepancies between the analysis and the experiment mentioned in Sec. 3.2, the authors propose a depressurized experiment. We expect that the flame moves upward and the reaction region becomes thicker because of the decrease in the reactant gas density in the depressurized conditions. Therefore, the thermal-hydraulic and chemical quantities can be measured with higher accuracy.

The flame location is influenced by the relationship of the water and sodium mass fluxes. To lift up the flame location above the pool surface, one needs to either increase the sodium mass flux and/or decrease the water vapor mass flux. With the depressurization, the sodium evaporation is enhanced and the sodium mass flux increases without ascending the pool temperature by improving the capacity of the heating unit. Current pool temperature condition (800 K) is consistent with the actual situation in which the design temperature ranges from 625 K to 743 K. It is easy to control the water vapor mass flux by changing the approaching flow velocity and the water vapor concentration. Another advantage is that the gas density is decreased. The depressurization will decrease the frequency of water and sodium molecular collision. It will enlarge the reaction region.

Table 1 Computational conditions

	Pressure (kPa)	Velocity (m/s)	Water vapor (vol %)
Case 0	100	0.12	2.7
Case 1	25	0.12	2.7
Case 2	25	0.06	5.4

The authors performed sensitivity analyses to decide the most appropriate pressure and the experimental conditions. The minimum pressure is decided to be 25 kPa considering the practicability of the experiment. Two analyses are shown in comparison with the ambient pressure (100 kPa) case designated as Case 0. In Case 0 and Case 1, the water vapor concentration is assumed to be 2.7 vol % and the flow velocity is 0.12 m/s. They are changed to 5.4 vol % and 0.06 m/s in Case 2. In the three cases, the water vapor mass flux is the same at the nozzle. The computational conditions are summarized in Table 1.

Figure 10 shows the sodium and water vapor densities as a function of the distance from the pool surface. The flame location lies around 3 mm from the pool surface in Case 1 and Case 2 while the flame location is below 1 mm in Case 0. Figure 11 shows the chemical reaction product concentration. Below the reaction flame is sodium-rich and Na_2O is the dominant reaction product. On the other hand, above the flame is water-vapor-rich and the NaOH is the dominant constituent. In the experiment, the constituent will be measured, which gives important information and insight as well as the validation data on the chemical reaction model.

We have little information on the reaction rate constant and fundamental reactions. In the present simulation, chemical equivalence model is used that assumes infinite reaction rate constant. Therefore, with this modeling, the thickness of the reaction zone is infinitely small. The expansion of the reaction zone is determined by the reaction rate constant and mobility of the reactants.

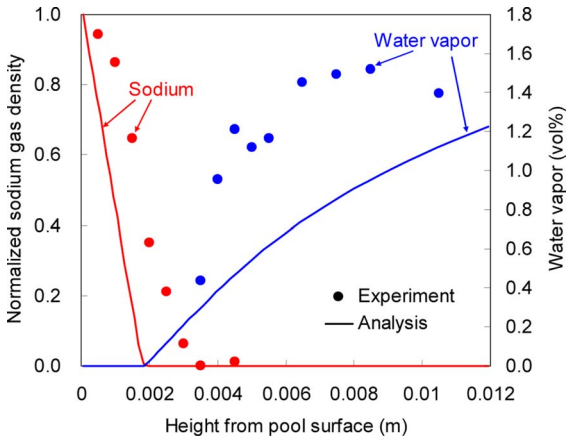


Fig. 9 Sodium and water vapor mass concentrations

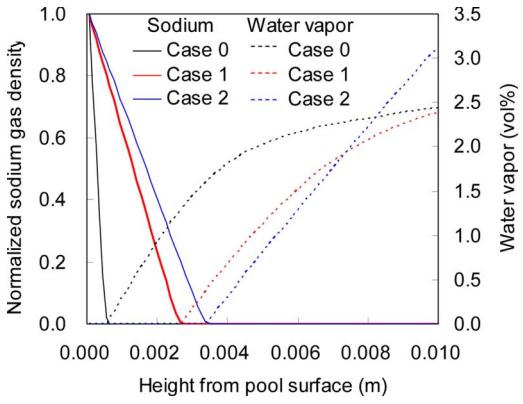


Fig. 10 Velocity vector above the pool surface

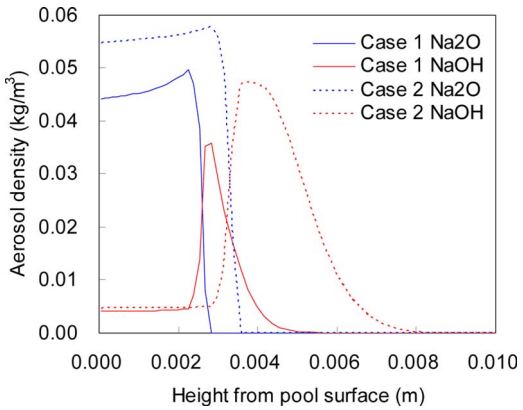


Fig. 11 Instantaneous temperature distribution

If we observe the experimental results carefully and interpret them in comparison with the numerical simulation, we may estimate the rate constant.

Figure 12 shows the aerosol mass density contour calculation. It is clear that the depressurization of the experimental cell lifts up the reaction flame, which facilitates the laser measurement from the side window.

It is noteworthy that the reaction zone where the OH radical exists is located above the aerosol (see Figs. 3, 6, and 8). Also, the existence of the OH radical has been proven. An assumption that we employ is that the chemical reaction rate is infinite, i.e., chemical equilibrium. As in Fig. 9, the water vapor and sodium vapor disappear at a flame plane with no thickness in the numerical simulation. On the other hand, the thickness of the OH radical seems to be an order of 1 mm as in Fig. 8. Further investigation is necessary by measuring the reaction zone thickness and investigating the relationship of the thickness and flow velocity or mass flux. With the information, the characteristic time of the convection flow and the chemical reaction will be quantified.

Although the depressurized experiment is currently ongoing, a preliminary result is shown in Fig. 13. The experimental condition is almost the same as Case 2 and the pressure is 24 kPa. In the photograph, we see the reacting region where the laser beam is scattered by the reaction products. The flame location is distant

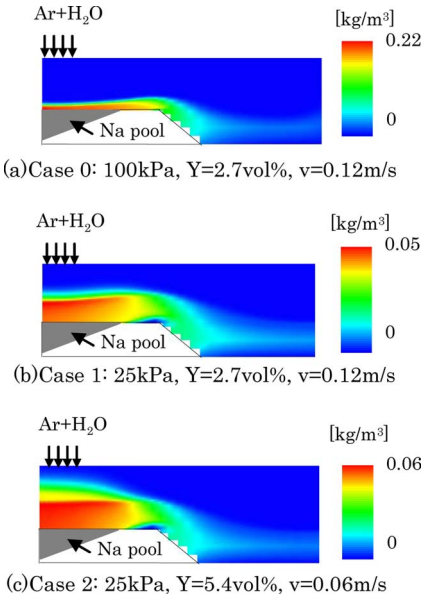


Fig. 12 Aerosol mass density for three cases

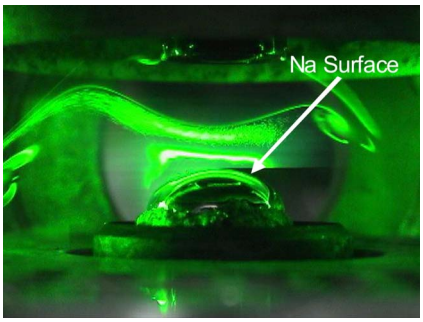


Fig. 13 Preliminary experimental result

from the sodium surface by approximately 5 mm. In the future, the authors expect quantitative measurement to be performed and the sodium-water reaction to be understood.

4 Conclusions

In this paper, a new computer program has been developed for the gas-phase reaction of the SWR. A counterflow diffusion flame is studied by a numerical simulation and an experiment. The experimental conditions are decided with the numerical simulation so that the stable reaction flame is maintained for a long time and physical and chemical quantities are measured. From the comparison of the analysis and the experiment, there exist discrepancies of the mass concentration distributions and the location of the flame. It may be caused by the assumptions of the chemical reaction model. Therefore, a new experiment is proposed to enhance the measurement accuracy and to investigate and explain the reason of the disagreement. The authors propose a depressurized experiment. We expect that the flame moves upward and the reaction region becomes thicker because of the decrease in the reactant gas density in depressurized conditions. It is expected based on the simulation that the experiment cell is depressurized to 25 kPa, the stable and thick reaction flame is generated, and the optic measurement is facilitated. Although the SWR is a complex phenomenon, the numerical simulation is useful and helpful to design an experiment to perform basic experiment. At the same time, information obtained from the experiment is useful to establish the modeling for the safety analysis code.

Nomenclature

C_p = specific heat
C_u = Cunningham correction factor
D_k = diffusion coefficient
F = heat transferred by gas diffusion
g_i = ith component of the gravity vector
k = identifier of chemical species
$n(v)$ = aerosol number density with its volume v
p = pressure
Q_R = radiation heat transfer per unit volume
S_k = mass production per unit volume
S = aerosol production rate
T = temperature
t = time
$\mathbf{u}$ = velocity vector
$\mathbf{U}$ = terminal velocity vector of the particle
u_i = ith component of the velocity vector
Y_k = molar mass density
β = aerosol coagulation rate
ε = emissivity
λ = thermal conductivity
v = kinetic viscosity
ρ = density
σ = Stephan–Boltzmann constant

References

[1] Tanabe, H., and Wachi, E., 1990, "Review on Steam Generator Tube Failure Propagation Study in Japan," *Proceedings of the IAEA/IWGFR Specialists' Meeting on Steam Generator Failure and Failure Propagation*, Aix-en Provence, France, pp. 33–35.

[2] Takata, T., Yamaguchi, A., Fukuzawa, K., and Matsubara, K., 2005, "Numerical Methodology of Sodium-Water Reaction With Multi-Phase Flow Analysis," Nucl. Sci. Eng., **150**, pp. 221–236.

[3] Yamaguchi, A., Takata, T., Ohshima, H., and Suda, K., 2006, "Sodium-Water Reaction and Thermal Hydraulics or Gas-Liquid Interface: Numerical Interpretation of Experimental Observation," *Proceedings of the 14th International Conference on Nuclear Engineering*, Miami, FL, Jul. 17–20, Paper No. ICONE14–89615.

[4] Yang, Z., and Shih, T. H., 1993, "New Time Scale Based k-e Model for Nearwall Turbulence," AIAA J., **31**, pp. 1191–1198.

[5] Jones, W. P., and Launder, B. E., 1972, "Low-Reynolds Number Phenomena With Two-Equation Model of Turbulence," Int. J. Heat Mass Transfer, **15**, pp. 301–314.

[6] Nagano, Y., and Tsuji, T., 1995, *Analysis of Turbulent Flows*, Computational Fluid Dynamics Series 3, H. Miyata, ed., University of Tokyo Press, Tokyo, p. 236.

[7] Gelbard, F., and Seinfeld, J. H., 1980, "Simulation of Multicomponent Aerosol Dynamics," J. Colloid Interface Sci., **78**(2), pp. 485–501.

[8] Hashiguchi, I., 1978, preprints 1978 Annual Meeting, Atomic Energy Society of Japan, Tokyo, Paper No. A55, in Japanese.

[9] Modak, A. T., 1979, "Radiation From Products of Combustion," Fire Research, **1**, pp. 339–361.

[10] Felske, J. D., and Tien, C. L., 1973, "Calculation of the Emissivity of Luminous Flames," Combust. Sci. Technol., **7**, pp. 25–31.

[11] Yamaguchi, A., and Tajima, Y., 2003, "Numerical Simulation of Non-Premixed Diffusion Flame and Reaction Product Aerosol Behavior in Liquid Metal Pool Combustion," J. Nucl. Sci. Technol., **40**(2), pp. 93–103.

On Moore's Law and Its Application to Spark Ignition Engine Technology

Marc LaViolette
Department of Mechanical Engineering,
Royal Military College of Canada,
Kingston, ON, K7K 7B4, Canada
e-mail: laviolette-m@rmc.ca

Moore's law relates how the integration of semiconductors has progressed in time. This research shows that the exponential trend shown in the electronics manufacturing industry can have applications elsewhere. This study shows that the internal combustion engine followed the same trend for over 70 years. Though not the most used engine variable, engine power density shows the same trends for engines as transistor density does for microchips. This now mature technology has ended its period of rapid growth. However, the present day engine trends can show how Moore's law can be extended to include the slower growth of long established technologies. Because exponential growth cannot go on forever, the extension Moore's law requires that the logistic function be used. The new function also allows one to predict a theoretical value for maximum power density.
[DOI: 10.1115/1.2978996]

1 Introduction

In 1965 Moore [1], a cofounder of Intel™, published a seminal paper on the future of integrated electronics. By observing that the complexity for minimum component costs has *increased* at a rate of roughly a factor of 2 per year, he was able to predict that the growth of transistors in integrated circuits would be 30 years hence. These ideas were deemed so important that the exponential relation describing the growth now bears his name. The spark ignition (SI) internal combustion engine is now a mature technology but in its infancy the growth was startling. Otto's first production engine was a 7.9 L, which could only produce 3.28 kW. By today's standard, that power is farcical. Did the engine industry follow Moore's law? Does it have the same growth today as in its infancy? Can we predict what it will be in ten years?

2 Performance Parameter

In illustrating the growth of a technology with time, one must choose a variable that is applicable to all engines in all eras. Power is not a good comparator because displacement has always been used, even by Otto, to scale up the power production. Mean effective pressure (MEP) is the common variable used presently. It is a fair comparator because almost all engines have approximately the same maximum engine speed and inlet air density. This is not true for early engines whose maximum speed was below 200 rpm for supercharged engines. Therefore power density, Ω, or power per unit displacement, was chosen as a representative parameter. Specifically, brake power can be written as [2]

$$\Omega = \frac{\dot{W}_B}{V_d} = \eta_{\text{mec}}\eta_{\text{th}}\eta_{\text{vol}}\frac{\rho_a}{t_{\text{cycle}}}f \ \text{HV} \tag{1}$$

Ω allows even extremely different designs from different eras to be compared. All major variables appear, though not always in a straightforward manner. Though the octane number of a fuel is not explicit to the formulation, its effects are taken into account. The higher compression ratio obtainable through the use of higher octane numbered fuels means a higher cycle thermal efficiency, η_{th}. The cycle efficiency also takes into account heat losses. The engine speed and number of strokes per cycle are hidden in the time to complete a cycle, t_{cycle}. It seems more logical to include the supercharging effects in air density, ρ_a, instead of volumetric efficiency, η_{vol}. The same has been done for any ram effects, or intake and exhaust tuning that increases the airflow.

3 Power Density Versus Time

Figure 1 shows how the power density has varied over time. The plot comprises approximately 900 spark ignition engines from all industries and applications (see Refs. 3–13), e.g., aircraft, automobile, Grand Prix racing, production vehicles, etc. For a given year there is a broad range of power densities. However, there is a trend within a given year. All high performance engines, those having small life expectancies, have the largest power density.

For the first 60 years, the increase seen in Fig. 1 is exponential with the power density doubling approximately every eight years. However, from the 1940s onwards, the improvement slows and seems to want to reach an asymptote. This upper asymptote is not surprising as there are physical limits to all the parameters given in the right hand side of Eq. (1). Hence there is a maximum value of power density.

The trend shown in Fig. 1 is one normally found in biology or economics where there is resource limited exponential growth. The function used to describe this growth is called the logistic function.

4 Logistic Function

Since this function is unfamiliar to most engineers, a brief description by Strang [14] is given here. This function was first used by Verhulst in Refs. 15 and 16 to describe population growth where the growth rate is proportional to the population but with a competition term, which shows that the resources get consumed as the square of the population. The first order differential function is of the form

$$\frac{dx}{dt} = ax - bx^2 \quad \text{whose solution is} \quad x = \frac{a}{b + e^{-a(t-t_0)}\left(\frac{a}{x_0} - b\right)} \tag{2}$$

Usually, the asymptote can be determined only when the maximum growth rate has been achieved. It is impossible to determine the asymptote from early data because, for small values of t, Eq. (2) behaves like

$$e^{a(t-t_0)} \tag{3}$$

This is the form Moore used to predict integrated circuit growth.

5 Growth of Power Density

Equation (2) was used to plot curves in Fig. 1. The values of the constants a, b, and x_0 for all three curves are shown in Table 1. To determine these constants, a least-squares fit minimizing the relative error was used to determine a, which was then held constant for all three curves. The b and x_0 of all three curves were then adjusted to give curves that reflected the three life expectancies: short, average, and long.

Any prediction is only as good as the data used for the predic-

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received December 5, 2006; final manuscript received May 9, 2008; published online December 18, 2008. Review conducted by Christopher J. Rutland. Paper presented at the 2005 Fall Conference of the ASME Internal Combustion Engine Division (ICEF2005), Ottawa, ON, Canada, September 11–14, 2005.

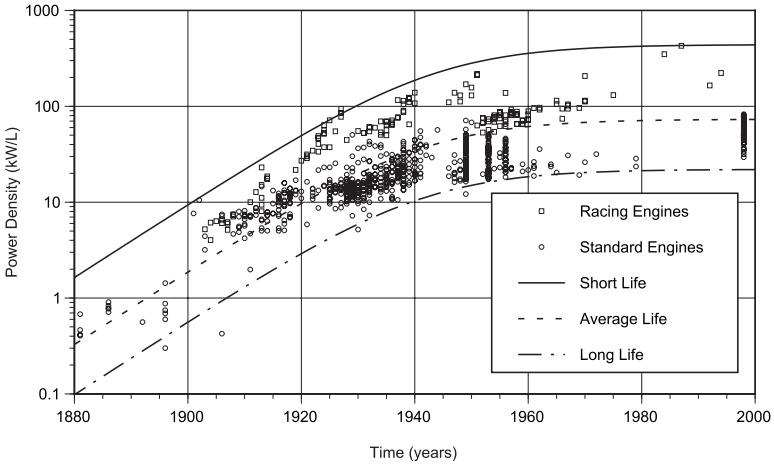


Fig. 1 Evolution of the power density of SI engines

tion. With this *caveat* in mind, one can calculate that the maximum improvement of Ω happened in the early 1940s. The maximum growth rate for average life was 1.62 kW/L year. For short life expectancies it was 9.69 kW/L year. The trend tends to suggest that maximum engine power densities will be $\Omega_\infty = 440$ kW/L for racing engines. For the longer life expectancy of automotive applications the maximum will be $\Omega_\infty = 73.3$ kW/L. The present day for average life growth rate is small 0.0237 kW/L year.The outlook for improvement seems bleak.

The present economic incentive seems to be a better use of this highly optimized solution for generating power. The development of hybrid vehicles points in this direction. The engine itself is not being improved upon but the way in which it is used is being strictly controlled.

6 Conclusion

The initial growth of the internal combustion engine was phenomenal. The power density doubled approximately every eight years showing the early exponential growth expected from a novel technology. The growth rate reached a maximum in the 1940s at which time Moore's law was no longer applicable. However, the logistic function allows the trend to be matched even after the growth is no longer exponential. The present growth rate has dwindled to almost a hundredth of the maximal value, and we can expect very small improvements in engine power density in the future. Engine development will likely continue with a focus on fuel economy without sacrificing power density. Because this technology is mature, its growth history can serve as a model to other technologies still in their infancy.

Nomenclature

a = logistic equation constant, 1/year
b = logistic equation constant, (L/kW year)
f = fuel to air mass ratio, kg of fuel/kg of air
t = time, year

Table 1 Logistic function constants for the curves in Fig. 1

Life expectancy	a (1/year)	b (L/kW year)	Ω_0 (kW/L)	t_0 (year)
Short	0.08805	2.00×10^{-4}	1.5	1879
Average	0.08805	1.20×10^{-3}	0.30	1879
Long	0.08805	4.00×10^{-3}	0.09	1879

HV = heating value of the fuel, MJ/kg
t_{cycle} = time to complete a cycle, s
V_d = displacement, L
$\dot{W}_B$ = brake power, kW
η = efficiency
ρ_a = inlet air density, kg/m^3
Ω = power density, kW/L

Subscripts

0 = initial value
∞ = asymptotic value
mec = mechanical
vol = volumetric
th = thermal

There is no SI unit for year. The largest SI time unit is the day. The year given is reckoned using the Gregorian calendar.

References

[1] Moore, G. E., 1965, "Cramming More Components Onto Integrated Circuits," Electronics, **38**(8), pp. 114–117.
[2] Lumley, J. L., 1999, *Engines: An Introduction*, Cambridge University Press, Cambridge, UK.
[3] Hobbs, L. S., 1971, *The Wright Brothers' Engines and Their Design* (Smithsonian Annals of Flight Vol. 5), Smithsonian Institution, Washington, DC.
[4] Burls, G. A., 1924, *Cost of Power Production by Internal Combustion Engines*, Blackie, Old Bailey, London.
[5] Smith, H. H., 1986, *Aircraft Piston Engines* (McGraw-Hill Series in Aviation), Sunflower University Press, Manhattan, KS.
[6] Ludvigsen, K., 2000, *Classic Grand Prix Cars*, Sutton, Stroud, UK.
[7] Ludvigsen, K., 2001, *Classic Racing Engines*, Haynes North America, Newbury Park, CA.
[8] Wilkinson, P. H., 1949, *Aircraft Engines of the World*, Paul H. Wilkinson, New York.
[9] Wilkinson, P. H., 1953, *Aircraft Engines of the World*, Paul H. Wilkinson, New York.
[10] Wilkinson, P. H., 1956, *Aircraft Engines of the World*, Paul H. Wilkinson, New York.
[11] Carpenter, R. C., and Diederichs, M. E., 1908, *Internal Combustion Engines*, Van Nostrand, New York.
[12] Clerk, S. D., and Burls, G., 1913, *The Gas, Petrol, and Oil Engine*, Vols. 1 and 2, 2nd. ed., Wiley, New York.
[13] Witz, A., 1923, *Moteurs à Gaz, à Essence et à Pétrole*, Vol. 1, 5th ed., Albin Michel, Rue Huyghens, Paris.
[14] Strang, G., 1986, *Introduction to Applied Mathematics*, Cambridge University Press, Cambridge, MA.
[15] Verhulst, P., 1845, "Recherches Mathématiques sur la loi d'Accroissement de la Population," *Nouveau Mémoire de l'Académie Royale des Sciences et Belles-Lettres de Bruxelles*, Vol. 18, Éditions de l'Académie, Bruxelle, pp. 1–41.
[16] Verhulst, P., 1847, "Deuxième Mémoire sur la loi d'Accroissement de la Population," *Mémoires de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique*, Vol. 20, Éditions de l'Académie, Bruxelle, pp. 1–32.