ASME
SETTING THE STANDARD

Journal of Engineering for Gas Turbines and Power

Published Quarterly by ASME

VOLUME 128 • NUMBER 2 • APRIL 2006

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

TECHNICAL PAPERS

(Contents continued on inside back cover)

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

(Contents continued)

Journal of Engineering for Gas Turbines and Power APRIL 2006

Volume 128, Number 2

Gas Turbines: Industrial and Cogeneration

Gas Turbines: Manufacturing, Materials & Metallurgy

Gas Turbines: Structures and Dynamics

Gas Turbines: Microturbines and Small Turbomachinery

Internal Combustion Engines

TECHNICAL BRIEFS

(Contents continued on facing page)

Journal of Engineering for Gas Turbines and Power APRIL 2006

Volume 128, Number 2

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

Experimental Analysis of Soot Formation in Sooting Diffusion Flame by Using Laser-Induced Emissions

Kazuhiro Hayashida
Department of Mechanical Engineering,
Kisarazu National College of Technology,
2-11-1 Kiyomidai-Higashi, Kisarazu,
Chiba 292-0041, Japan
e-mail: hayasida@m.kisarazu.ac.jp

Kenji Amagai
e-mail: amagai@me.gunma-u.ac.jp

Keiji Satoh
e-mail: d04b207@gs.eng.gunma-u.ac.jp

Masataka Arai
e-mail: arai@me.gunma-u.ac.jp

Department of Mechanical System Engineering,
Gunma University,
1-5-1 Tenjin-cho, Kiryu,
Gunma 376-8515, Japan

Two-dimensional images of OH fluorescence, polycyclic aromatic hydrocarbons (PAHs) fluorescence, and laser-induced incandescence (LII) from soot were measured in a sooting diffusion flame. To obtain an accurate OH fluorescence image, two images were taken with the laser wavelength tuned to ("on") and away from ("off") the OH absorption line. An accurate OH fluorescence image was obtained by subtracting the off-resonance image from the on-resonance image. For the PAH fluorescence and LII measurements, temporally resolved measurements were used to obtain the individual images; the LII image was obtained by detecting the LII signal after the PAH fluorescence radiation had stopped and the PAH fluorescence image was obtained by subtracting the LII image from the simultaneous image of PAH fluorescence and LII. Based on the obtained images, the relative location of OH, PAH, and soot in the flame was discussed in detail. To investigate the PAH size distribution in a sooting flame using LIF, an estimation strategy for PAH size is proposed. Emission spectra were measured at several heights in the flame using a spectrograph. Since the emission wavelength of PAH fluorescence shifts toward longer wavelengths with increasing PAH size, the main PAH components in the emission spectra could be estimated. The results suggest that PAH grows and the type of PAH changes as the soot inception region was approached. Near the soot inception region, we estimated that the PAHs, which have over 16 carbon atoms, mainly constituted the emission spectrum. [DOI: 10.1115/1.2056536]

Introduction

Soot emission from combustion devices, such as diesel engines and industrial furnaces, cause serious problems to the environment and human health. To reduce soot emission from combustion devices, a detailed knowledge about soot formation processes in flames is indispensable. Although many important studies have been carried out on soot mechanisms [1], the current understanding of the reactive and physical processes of soot is limited.

Polycyclic aromatic hydrocarbons (PAHs) have been regarded as precursors of soot particles from molecular reactants in flames and are often emitted as absorbed on the surfaces of soot particles. In previous studies, many proposals have been advanced concerning the relationship between PAHs and soot formation [2,3]. PAHs are considered to be the building blocks that undergo chemical reactions to form incipient soot [4–7]. PAHs appear just prior to soot inception and bridge the mass range between fuel molecules and soot particles [8,9]. Bittner et al. [10] showed detail distribution of the individual PAH species in a flame measured by a molecular beam mass spectrometer system. Furthermore, PAHs contribute to soot size growth by attaching to the outer surface of soot after it is formed [3,5].

Laser diagnostic techniques are able to probe combustion products nonintrusively. Laser-induced incandescence (LII) and laser-induced fluorescence (LIF) are attractive soot diagnostic techniques and can be used to obtain information about soot and PAHs, respectively. LII has been used in numerous studies for the measurement of soot volume fractions [11–15]. Recently, LII has been applied successfully to practical combustion systems [16–19]. In contrast, there are some problems in PAH measurement by LIF, and hence LIF has not been widely used for the diagnosis of soot. The problems of LIF measurement of PAHs are (1) due to the continuous broadband transition of PAHs [20], selective excitation and detection of individual PAH in a flame cannot be performed; (2) detection sensitivity differs for each PAH molecule [21]; (3) the profile of the fluorescence spectrum depends on temperature [22,23]; and (4) separation of PAH fluorescence and LII is difficult in the sooting region [24]. In recent studies, for the use of LIF in soot diagnosis, researchers have proposed techniques for detecting PAH fluorescence. Smyth et al. [25] presented two-dimensional PAH fluorescence distributions obtained from flickering diffusion flames with PAHs excited at 283.5 nm and fluorescence signals detected at 400–447 nm. In their study, to avoid the appearance of LII, the laser energy was controlled to avoid laser excitation of incandescence from soot particles. Vander Wal et al. [26] performed two-dimensional simultaneous detection of PAH fluorescence and LII in ethylene diffusion flames. To optimize both the LIF and LII intensities, they introduced temporally resolved measurements to determine reasonable detection timing. PAH and soot were excited at 266 nm and laser-induced emissions (PAH fluorescence and LII) were detected at 400, 500, and 600 nm using bandpass filters (each with a 70 nm bandwidth). They found that as the detection wavelength increased, the region of maximum fluorescence intensity approached the sooting region. They concluded that the variation of the fluorescence intensity distribution qualitatively reflected the spatial variation of the concentration of PAH size, i.e., the fluorescence centered near 400 nm was likely to be due primarily to small PAH, while the fluorescence at 500 and 600 nm were generated from larger PAH.

In the above techniques, although the relative distribution of integrated PAH fluorescence could be obtained, the concentration distribution of PAH size in a flame was difficult to determine. In order to analyze the soot formation processes, information about

Contributed by the Fuels and Combustion Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received April 12, 2004; final manuscript received February 9, 2005. Assoc. Editor: S. R. Gollahali.

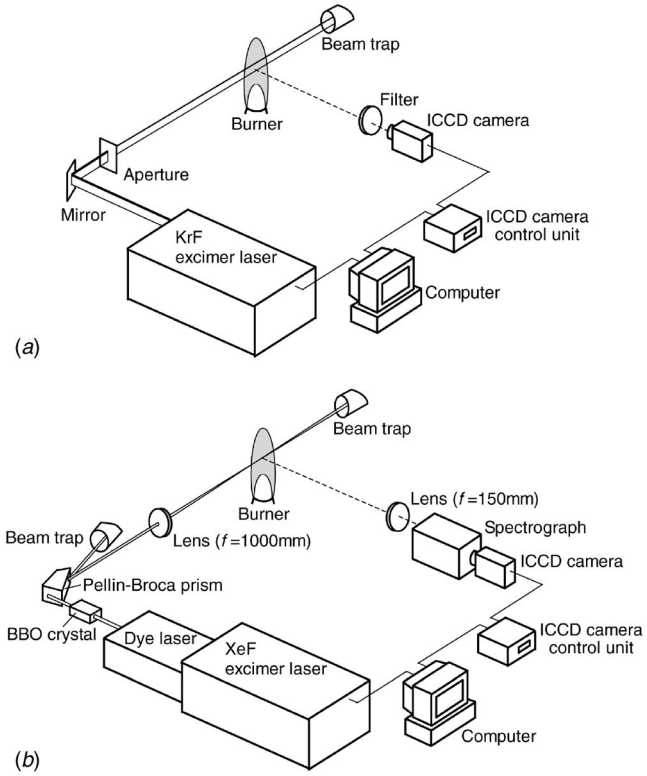


Fig. 1 Experimental apparatus

the PAH size is important. In this study, we measured the two-dimensional distributions of OH fluorescence, PAH fluorescence, and LII from soot. We clarified the detailed spatial locations of OH, PAH, and soot in a propane diffusion flame down to the overlapping region between OH, PAH, and soot. Furthermore, we propose an estimation strategy of PAH size from the profile of a PAH fluorescence spectrum. Using this strategy, we tried to estimate the PAH size distribution in a propane diffusion flame. Based on our results, we discuss the soot formation mechanisms in a propane diffusion flame.

Experimental Apparatus

The laser diagnostic system consisted of an excimer laser (Lambda Physic, LPX-150T), a dye laser (Lambda Physic, SCANMATE-2E), a spectrograph (Chromex, 250IS), and an ICCD camera (La-Vision, FlameStar II). Figure 1 shows schematics of the optical arrangements. In the two-dimensional measurements of OH fluorescence, PAH fluorescence, and LII from soot [Fig. 1(*a*)], a narrow-band KrF excimer laser (linewidth: $\sim 0.3\ \mathrm{cm}^{-1}$, pulse duration: ~ 24 ns) was used for the excitation of OH, PAH, and soot. The wavelength of this laser could be tuned in the range 247.8–248.8 nm. The laser light was formed into a light sheet ($0.5 \times 20\ \mathrm{mm}^2$) by an aperture and was introduced into a flame with a pulse energy of 100 mJ. Laser-induced emissions were detected by the ICCD camera oriented perpendicular to the laser beam direction. LIF and LII images were acquired by averaging over 1000 laser shots. Figure 1(*b*) shows the optical arrangement used in the emission spectra measurements. A XeF excimer laser was used to pump the dye laser (Rhodamine 6G), producing a beam at 580 nm (linewidth: $\sim 0.15\ \mathrm{cm}^{-1}$). A BBO crystal was used to double the dye laser output to produce 290 nm radiation. The pulse energy of the laser light was 3 mJ and the pulse duration was 20 ns. A Pellin-Broca prism was used to separate the UV laser beam (290 nm) and the visible laser beam (580 nm). The UV laser beam was focused with a spherical lens ($f=1000$ mm) and was introduced into the target flame. Laser-induced emissions were analyzed by the spectrograph with the ICCD camera. The grating of the spectrograph had 100 grooves/mm (blazed for 450 nm). A spectral bandpass of 300 nm could be captured with the ICCD camera, which had 384 (horizontal) $\times$ 288 (vertical) pixels. The spectrum within the 300 nm width was spread over the 384 rows of the ICCD camera, giving a spectral resolution of 2.4 nm using an entrance slit width of 20 μm.

The test flame used in this study was a laminar propane diffusion flame under atmospheric pressure. The test burner consists of two concentric nozzles of 6 and 56 mm in inner diameter. The fuel flowed through the inner nozzle at 0.05 L/min and the flame length was $L_f=30$ mm. To stabilize the flame, an annular airflow (temperature: 298 K) was applied from the outer nozzle. The velocity of the laminar airflow measured by a thermal anemometer (Kanomax, Aneomaster model 24-6131) was 0.5 m/s. The temperature was measured using a Pt/Pt–Rh 13% thermocouple and having a bead diameter of 300 μm. The thermocouple was kept in the flame for as short a time as possible to minimize soot accumulation on the wires. The soot deposited on the thermocouple probe was burnt off by another blue flame before each measurement. The obtained thermocouple temperatures were corrected for radiation heat losses assuming a spherical bead geometry. The correction was based on a procedure reported by Bradley et al. [27].

Experimental Methods

OH Fluorescence Measurement. The KrF excimer laser used in this study can be applied to laser-induced predissociative fluorescence (LIPF) for OH measurements [28]. LIPF has an important advantage over "normal" LIF in that the collisional quenching at atmospheric pressure is much smaller. The excitation band of OH-LIPF is the A-$X(3,0)$ band. The wavelength of the KrF excimer laser was set at 248.457 nm to excite the $P_2(8)$ transition within the A-$X(3,0)$ band. The lower state population of the $P_2(8)$ transition was insensitive to temperature variations in the vicinity of the flame front (the population changes by only 5% between 1400 K and 1700 K). OH fluorescence radiation was collected using a specially designed bandpass filter (centered at 296.5 nm, FWHM: 10 nm, maximum transmittance: 40%) that transmitted only the A-$X(3,2)$ fluorescence band near 300 nm. The ICCD camera was operated with a 50 ns gate.

In the case of a sooting flame, PAH and soot particles are also excited by the excitation laser light of OH, because the absorption bands of these substances are broad. Thus, OH fluorescence, PAH fluorescence, and LII were simultaneously detected. In order to obtain an accurate OH fluorescence distribution, we performed the following procedure similar to Puri et al. [29].

1. Two images were obtained with the laser wavelength tuned both "on" (248.457 nm) and "off" (248.469 nm) the OH $P_2(8)$ transition. The on-resonance image consisted of signals due to OH fluorescence, PAH fluorescence, and LII, while the off-resonance image consisted of only the latter two signals.
2. Since the wavelength of the off-resonance laser was only slightly different from that of the $P_2(8)$ transition, the intensity distributions of PAH fluorescence and LII were hardly changed. Thus, the OH fluorescence image could be obtained by subtracting the off-resonance image from the on-resonance image.

PAH Fluorescence and LII Measurements. In the PAH-fluorescence and LII measurements, the wavelength of the KrF excimer laser was tuned to 248.469 nm, the same wavelength as the off-resonance laser of the OH measurement. In this experiment, PAH fluorescence and LII were collected through a 248 nm cutoff filter (centered at 248 nm, FWHM: 4 nm) that eliminated

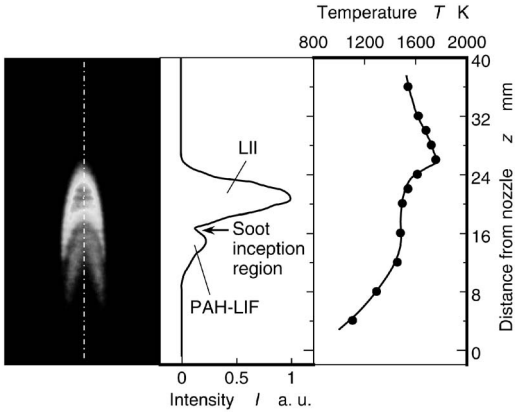


Fig. 2 Simultaneous image of PAH-LIF and LII

elastic scattering.

Figure 2 shows an example simultaneous image of PAH fluorescence and LII. The emission intensity and temperature distributions on the flame axis are also shown. PAH fluorescence appears within the fuel region and LII appears downstream from the PAH fluorescence. In this image, it can be confirmed that a low intensity region exists between PAH fluorescence and LII. The intensity profile on the flame axis clearly shows a local minimum at a distance from the nozzle of $z \approx 17$ mm in the region between the PAH fluorescence and LII. Similar results have been reported by Smyth et al. [25] and Vander Wal et al. [26]. It seems that this region is a transition region from PAH to soot and has been attributed to the presence of soot inception. By the "authors' visual observation," it was confirmed that a luminous region appeared at around $z \approx 17$ mm in the test flame. In order to investigate the soot inception region in more detail, it is necessary to separate the PAH fluorescence and LII.

Figure 3 shows the temporal variation of the Rayleigh scattering, PAH fluorescence, and LII signals. The normalized intensities are plotted. The signal intensities for each delay time were obtained by changing the gate delay of the ICCD camera, for a 5 ns gate. Intensity of Rayleigh scattering was obtained by another measurement without the flame and the 248 nm cutoff filter. From this figure it was confirmed that the rising edge of the PAH fluorescence radiation was faster than that of LII. A complete separation of PAH fluorescence and LII was difficult using only temporal-resolved measurements, even for a fast delay time. However, since the temporal characteristics of radiation are different for PAH fluorescence and LII, separate images of PAH fluorescence and LII could be derived satisfactorily using temporally resolved measurements using the following procedure.

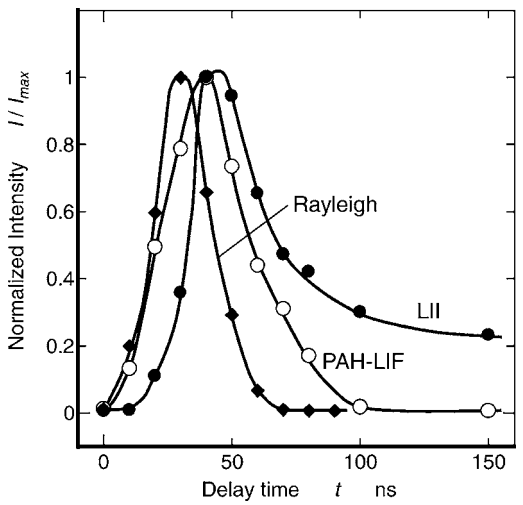


Fig. 3 Temporal variations of signal intensity (Rayleigh scattering, PAH fluorescence, and LII)

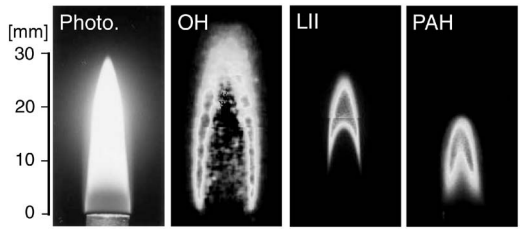


Fig. 4 Two-dimensional images of OH fluorescence, PAH fluorescence, and LII

1. Since the lifetime of LII is much longer than PAH fluorescence, detection of only the LII signal was possible for a long delay time of $t \geqslant 100$ ns.
2. The intensity of the LII image was multiplied by a certain constant so that it agreed with the LII intensity in the simultaneous image of PAH fluorescence and LII.
3. A PAH fluorescence image was derived by subtracting the corrected LII image from a simultaneous image of PAH fluorescence and LII.

In this study, the LII image was obtained at $t=100$ ns, while the simultaneous image was obtained at $t=30$ ns.

Results and Discussion

Relative Location of OH, PAH, and Soot. Figure 4 shows a flame photograph and two-dimensional images of OH fluorescence, PAH fluorescence, and LII obtained from the propane diffusion flame. Since the test flame was cylindrical shape, attenuation of the laser-induced emissions might have occurred by absorption between the measurement plane and the ICCD camera. Especially, it had been pointed out that the LII signal is attenuated by the soot absorption. In order to correct the absorption effect, and quantify the soot concentration from the LII signal, calibrations have often been performed by comparing LII signals to laser extinction measurements of soot volume fraction [11,30,31]. In this study, correction of the absorption effect was omitted, because our purpose was to obtain the relative location relationship of the species. In the OH fluorescence image, although it seems that some fluorescence is detected in the fuel region, this is an error resulting from the subtraction. It is thought that this error becomes so large that soot concentration in a flame is high, because intensity of the off-resonance image becomes strong. OH fluorescence was detected around the outer edge of the flame; the OH fluorescence image reveals that the maximum OH concentration decreases while the spatial extent of the distribution increases with increasing distance from the nozzle. LII was detected within the OH fluorescence region and PAH fluorescence was widely distributed in the fuel region. The PAH fluorescence image reveals that PAH are formed and destroyed completely inside the flame.

Figure 5 shows the axial intensity distributions of the images in Fig. 4. The temperature distribution is also shown. The PAH fluorescence distribution suggests that PAH are promptly generated at the nozzle exit, with their concentration increasing with increasing distance from the nozzle z. The peak of the PAH fluorescence intensity is located around $z=15$ mm, which coincides with the location where LII begins to appear. The PAH fluorescence intensity decreases rapidly with increasing LII intensity and the location of the disappearance of PAH fluorescence coincides with the location of the peak of LII. LII disappears around $z=27$ mm, which is the location of the peak of the OH fluorescence intensity, located slightly downstream from the location of the maximum

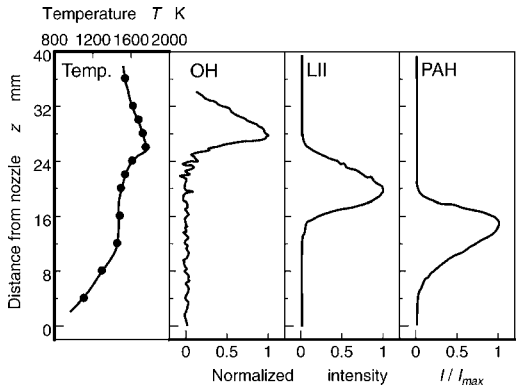


Fig. 5 Axial distributions of temperature, OH fluorescence, PAH fluorescence, and LII

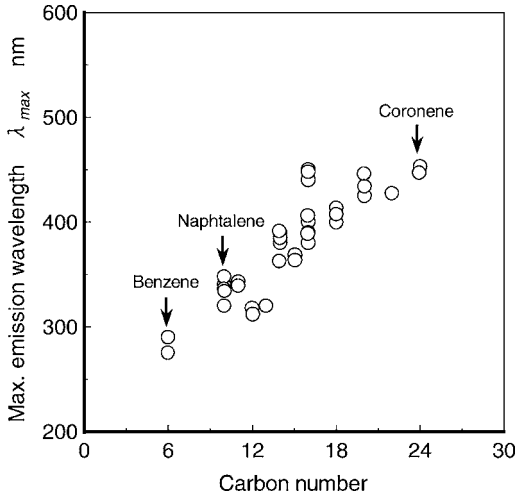


Fig. 7 Relation between peak fluorescence wavelength and PAH size

temperature. These results suggest that the PAHs were formed within the fuel region and that their concentration increased as the soot inception region was approached. Soot was formed via PAHs and was then oxidized by the oxidant (primarily OH).

Figure 6 shows selected radial profiles from the images of Fig. 4. At $z=6$ mm, LII is not observed and PAH fluorescence distribution has the peak near the flame front. LII appears at the edge of the PAH fluorescence ($z=15$ and 18 mm) and its peak coincides with the location where PAH fluorescence disappears. It is interesting that at $z=15$ mm the soot inception region indicated in Fig. 2 corresponds to the location where the profiles of PAH fluorescence and LII cross. Vander Wal et al. [26] sampled soot precursor particles in this region and investigated the particle size by transmission electron microscopy (TEM). Small particles of approximately 2 to 5 nm in size were observed.

Estimation of PAH Size. Broadband excitation and emission characteristics of PAH complicate the spectroscopic analysis of PAH in a flame. The selective excitation and detection of individual PAH in a flame is extremely difficult, and so the measurement of the concentration distribution of individual PAH had not been achieved by the LIF technique. Generally, since the energy separation between the electronic excited state and electronic ground state decreases with increasing size of PAH, the fluorescence wavelength region shifts toward longer wavelengths as the size of PAH increases [20]. Therefore, classification of PAH size might be performed using the PAH fluorescence spectrum obtained from a flame, because the peak fluorescence wavelengths reflect the size of PAH.

Figure 7 presents the relation between the peak wavelengths of PAH fluorescence and PAH size. Benzene is also indicated in this figure. The peak wavelength data were obtained from Refs. [22,23,32–37]. Individual PAHs in the vapor phase were excited by UV laser sources. Figure 7 clearly shows that the peak of PAH fluorescence sifts to longer wavelengths as the size of the PAH increases. For large PAH (carbon number >24), fluorescence data in the vapor phase was not available.

Figure 8 shows emission spectra obtained from the test flame. Measurements were performed on the flame axis. An LII spectrum obtained from the soot is also indicated in the figure for reference. The peak wavelength of the emission shifts to longer wavelengths with an increase in the distance from the nozzle z. The profile at $z=20$ mm almost agrees with the profile of LII. Variations in the spectral profile suggests that the PAH grew as the soot inception region was approached. In the spectra at $z=12-20$ mm and that of

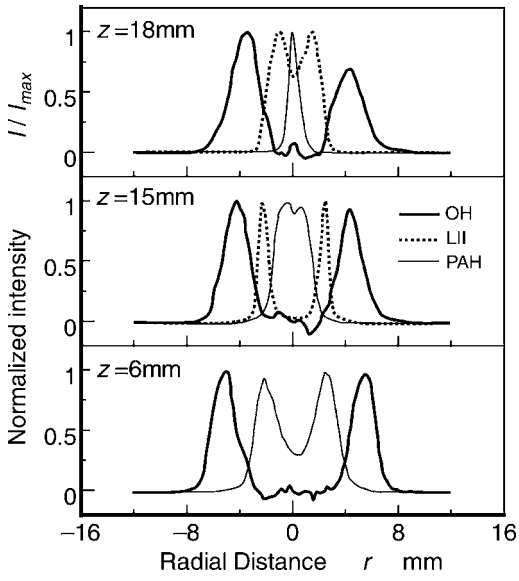


Fig. 6 Radial distributions of OH fluorescence, PAH fluorescence, and LII

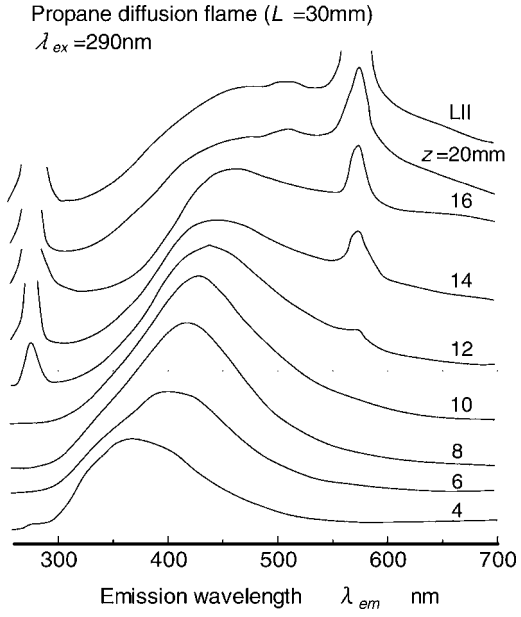


Fig. 8 Emission spectra from propane diffusion flame

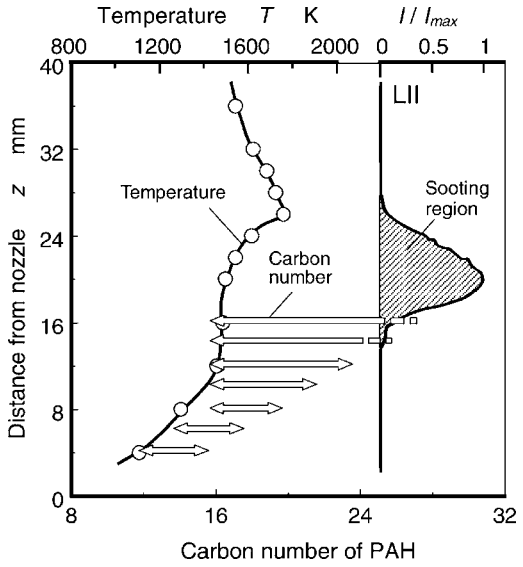


Fig. 9 Estimation of PAH size in propane diffusion flame

LII, strong Mie scattering appeared at 290 nm. The peak observed at 580 nm is the contribution from the second-order interference from the Mie scattering at 290 nm. The detection of weak Mie scattering at $z=12$ mm implies the generation of soot precursor. Therefore, it is thought that the emission spectra at $z=12$, 14, 16, and 20 mm have intensities in the longer wavelength region contributed to by LII.

Using the relation indicated in Fig. 7, we estimated the main PAH fluorescence components that constituted the spectra of Fig. 8. The estimation was performed by the following procedure.

1. The peak wavelengths of the emission spectrum in Fig. 8 were recorded.
2. The size of the PAH was estimated from the range of the plot at the obtained peak wavelength.

The estimation results are shown in Fig. 9 together with the axial distribution of the temperature and the LII intensity. The results reveal that the size of the PAH increased with increasing distance from the nozzle z. It is interesting that PAHs grew quickly with the rapid rise of temperature in the region $z<12$ mm. At $z=4$ mm, it is thought that fluorescence from PAH molecules, which contain approximately 12 to 16 carbon atoms, mainly constituted the emission spectrum in Fig. 8. Example PAH molecules corresponding to carbon numbers 12–16 are acenaphthylene ($C_{12}H_8$), acenaphthene ($C_{12}H_{10}$), fluorene ($C_{13}H_{10}$), anthracence ($C_{14}H_{10}$), phenanthrene ($C_{14}H_{10}$), pyrene ($C_{16}H_{10}$), fluoranthene ($C_{16}H_{10}$), and benzacenaphthylene ($C_{16}H_{10}$). Figure 9 suggests that the PAH grew and the type of PAH changed as the distance from the nozzle increased. This result qualitatively agrees with a previous sampling study [38]. Near the soot inception region ($z=14$ mm), it is estimated that the PAH, which have carbon atoms over 16, mainly constituted the emission spectrum. In previous studies [38,39], the relatively large PAH coronene ($C_{24}H_{12}$) was widely detected in a flame. The molecular size of coronene is 0.73 nm and the particle size of soot precursor particles is sub-nm. It would appear that precursors are formed by the polymerization and/or coagulation of large PAH.

Conclusions

LIF and LII are effective techniques for analysis of soot formation. Two-dimensional images of OH fluorescence, PAH fluorescence, and LII from soot were successfully obtained from a laminar propane diffusion flame. From the obtained images, we clarified the relative locations of OH, PAHs, and LII in the propane diffusion flame. Also, the size distribution of PAHs in the propane diffusion flame was estimated based on the relation between peak PAH fluorescence wavelengths and PAH size. The main findings of this study are as follows.

1. A two-dimensional PAH fluorescence image could be obtained by subtracting the LII image from the simultaneous image of PAH fluorescence and LII.
2. OH fluorescence was detected around the outer edge of the flame, LII was detected within the OH fluorescence region, while PAH fluorescence was widely distributed in the fuel region.
3. The peak wavelength of the emission spectra of the propane diffusion flame shifted to longer wavelengths as the soot inception region was approached.
4. Estimations of PAH size in the flame suggested that PAH size increased as the soot inception region was approached. Large PAHs, which have carbon atoms over 16, mainly existed near the soot inception region.

Acknowledgment

The authors are grateful for permission to use the flush-photometry of Gunma University, Faculty of Engineering special equipment. The authors would also like to thank T. Miyakawa for technical assistance. This study was supported by the Ministry of Education, Science, Sports and Culture, Grant-in-Aid for Scientific Research (b), 15560169, 2003 and the Program for Promoting Fundamental Transport Technology Research from Japan Railway Construction, Transport and Technology Agency.

References

[1] Bockhorn, H., 1994, *Soot Formation in Combustion: Mechanisms and Models*, Springer-Verlag, Berlin.
[2] Calcote, H. F., 1981, "Mechanisms of Soot Nucleation in Flames—A Critical Review," Combust. Flame, **42**, pp. 215–242.
[3] Richter, H., and Howard, J. B., 2000, "Formation of Polycyclic Aromatic Hydrocarbons and Their Growth to Soot—a Review of Chemical Reaction Pathways," Prog. Energy Combust. Sci., **26**, pp. 565–608.
[4] Vander Wal, R. L., 1996, "Soot Precursor Material: Visualization via Simultaneous LIF-LII and Characterization via TEM," Proc. Combust. Inst., **26**, pp. 2269–2275.
[5] Frenklach, M., 1996, "On Surface Growth Mechanism of Soot Particles," Proc. Combust. Inst., **26**, pp. 2285–2293.
[6] Melton, T. R., Inal, F., and Senkan, S. M., 2000, "The Effects of Equivalence Ratio on the Formation of Polycyclic Aromatic Hydrocarbons and Soot in Premixed Ethane Flames," Combust. Flame, **121**, pp. 671–678.
[7] Wen, Z., Yun, S., Thomson, M. J., and Lightstone, M. F., 2003, "Modeling Soot Formation in Turbulent Kerosene/Air Jet Diffusion Flames," Combust. Flame, **135**, pp. 323–340.
[8] Homann, K.-H., and Wagner, H. Gg., 1967, "Some New Aspects of the Mechanism of Carbon Formation in Premixed Flames," Proc. Combust. Inst., **11**, pp. 371–379.
[9] Fialkov, A. B., Dennebaum, J., and Homann, K.-H., 2001, "Large Molecules, Ions, Radicals, and Small Soot Particles in Fuel-Rich Hydrocarbon Flames, Part V: Positive Ions of Polycyclic Aromatic Hydrocarbons (PAH) in Low-Pressure Premixed Flames of Benzene and Oxygen," Combust. Flame, **125**, pp. 763–777.
[10] Bittner, J. D., and Howard, J. B., 1981, "Composition Profiles and Reaction Mechanisms in a Near-Sooting Premixed Benzene/Oxygen/Argon Flame," Proc. Combust. Inst., **18**, pp. 1105–1116.
[11] Quay, B., Lee, T.-W., Ni, T., and Santoro, R. J., 1994, "Spatially Resolved Measurements of Soot Volume Fraction Using Laser-Induced Incandescence," Combust. Flame, **97**, pp. 384–392.
[12] Cignoli, F., Benecchi, S., and Zizak, G., 1994, "Time Delayed Detection of Laser-Induced Incandescence for the Two-Dimensional Visualization of Soot in Flames," Appl. Opt., **33**, pp. 5778–5782.
[13] Bengtsson, P.-E., and Aldén, M., 1995, "Soot-Visualization Strategies Using Laser Techniques, Laser-Induced Fluorescence in C_2 from Laser-Vaporized Soot and Laser-Induced Soot Incandescence," Appl. Phys. B: Lasers Opt., **60**, pp. 51–59.
[14] Allouis, C., Beretta, F., and D'Alessio, A., 2003, "Sizing Soot and Micronic Carbonaceous Particle in Spray Flames Based on Time Resolved LII," Exp. Therm. Fluid Sci., **27**, pp. 455–463.
[15] Snelling, D. R., Liu, F., Smallwood, G. J., and Gülder, Ö. L., 2004, "Determination of the Soot Absorption Function and Thermal Accommodation Coefficient Using Low-Fluence LII in a Laminar Coflow Ethylene Diffusion Flame," Combust. Flame, **136**, pp. 180–190.

[16] Dec, J. E., zur Loye, A. O., and Siebers, D. L., 1991, "Soot Distribution in a D. I. Diesel Engine Using 2-D Laser-Induced Incandescence Imaging," SAE Paper No. 910224, Society of Automotive Engineers, Warrendale, PA.
[17] Geitlinger, H., Streibel, Th., Suntz, R., and Bockhorn, H., 1999, "Statistical Analysis of Soot Volume Fractions, Particle Number Densities and Particle Radii in a Turbulent Diffusion Flame," Combust. Sci. Technol., **149**, pp. 115–134.
[18] Crua, C., Kennaird, A. A., and Heikal, M. R., 2003, "Laser-Induced Incandescence Study of Diesel Soot Formation in a Rapid Compression Machine at Elevated Pressures," Combust. Flame, **135**, pp. 475–488.
[19] Pickett, L. M., and Siebers, D. L., 2004, "Soot in Diesel Fuel Jets: Effects of Ambient Temperature, Ambeint Density, and Injection Pressure," Combust. Flame, **138**, pp. 114–135.
[20] Berlman, I. B., 1965, *Handbook of Fluorescence Spectra of Aromatic Molecules*, Academic Press, New York.
[21] Petarca, L., and Marconi, F., 1989, "Fluorescence Spectra and Polycyclic Aromatic Species in a N-Heptane Diffusion Flame," Combust. Flame, **78**, pp. 308–325.
[22] Ossler, F., Metz, T., and Aldén, M., 2001, "Picosecond Laser-Induced Fluorescence from Gas-Phase Polycyclic Aromatic Hydrocarbons at Elevated Temperatures., I. Cell Measurements," Appl. Phys. B: Lasers Opt., **72**, pp. 465–478.
[23] Chi, Z., Cullum, B. M., Stokes, D. L., Mobley, J., Miller, G. H., Hajaligol, M. R., and Vo-Dinh, T., 2001, "Laser-Induced Fluorescence Studies of Polycyclic Aromatic Hydrocarbons (PAH) Vapors at High Temperatures," Spectrochim. Acta, Part A, **57**, pp. 1377–1384.
[24] Smyth, K. C., Miller, J. H., Dorfman, R. C., Mallard, W. G., and Santoro, R. J., 1985, "Soot Inception in a Methane/Air Diffusion Flame as Characterized by Detailed Species Profiles," Combust. Flame, **62**, pp. 157–181.
[25] Smyth, K. C., Shaddix, C. R., and Everest, D. A., 1997, "Aspects of Soot Dynamics as Revealed by Measurements of Broadband Fluorescence and Flame Luminosity in Flickering Diffusion Flames," Combust. Flame, **111**, pp. 185–207.
[26] Vander Wal, R. L., Jensen, K. A., and Choi, M. Y., 1997, "Simultaneous Laser-Induced Emission of Soot and Polycyclic Aromatic Hydrocarbons Within a Gas-Jet Diffusion Flame," Combust. Flame, **109**, pp. 399–414.
[27] Bradley, D., and Matthews, K. J., 1968, "Measurement of High Gas Temperatures with Fine Wire Thermocouples," J. Mech. Eng. Sci., **10**, pp. 299–305.
[28] Andresen, P., Bath, A., Gröger, W., Lülf, H. W., Meijer, G., and ter Meulen, J. J., 1982, "Laser-Induced Fluorescence with Tunable Excimer Lasers as a Possible Method for Instantaneous Temperature Field Measurements at High Pressures: Checks with an Atmospheric Flame," Appl. Opt., **27**, pp. 365–378.
[29] Puri, R., Santoro, R. J., and Smyth, K. C., 1994, "The Oxidation of Soot and Carbon Monoxide in Hydrocarbon Diffusion Flames," Combust. Flame, **97**, pp. 125–144.
[30] Shaddix, C. R., and Smyth, K. C., 1996, "Laser-Induced Incandescence Measurements of Soot Production in Steady and Flickering Methane, Propane, and Ethylene Diffusion Flames," Combust. Flame, **107**, pp. 418–452.
[31] Bryce, D. J., Ladommatos, N., and Zhao, H., 2000, "Quantitative Investigation of Soot Distribution by Laser-Induced Incandescence," Appl. Opt., **39**, pp. 5012–5022.
[32] Beretta, F., Cincotti, V., D'Alessio, A., and Menna, P., 1985, "Ultraviolet and Visible Fluorescence in the Fuel Pyrolysis Regions of Gaseous Diffusion Flames," Combust. Flame, **61**, pp. 211–218.
[33] Ware, W. R., and Cunningham, P. T., 1965, "Lifetime and Quenching of Anthracence Fluorescence in the Vapor Phase," J. Chem. Phys., **43**, pp. 3826–3831.
[34] Ohta, N., and Baba, H., 1982, "Vibronic-Level Dependence of Radiative and Nonradiative Processes in Naphthelene Vapor," J. Chem. Phys., **76**, pp. 1654–1663.
[35] Coe, D. S., and Steinfeld, J. I., 1980, "Fluorescence Excitation and Emission Spectra of Polycyclic Aromatic Hydrocarbons at Flame Temperatures," Chem. Phys. Lett., **76**, pp. 485–489.
[36] Baba, H., Nakajima, A., Aoi, M., and Chihara, K., 1971, "Fluorescence from the Second Excited Singlet State and Radiationless Processes in Pyrene Vapor," J. Chem. Phys., **55**, pp. 2433–2438.
[37] Chi, Z., Cullum, B. M., Stokes, D. L., Mobley, J., Miller, G. H., Hajaligol, M. R., and Vo-Dinh, T., 2001, "High-Temperature Vapor Detection of Polycyclic Aromatic Hydrocarbon Fluorescence," Fuel, **80**, pp. 1819–1824.
[38] Hepp, H., Siegmann, K., and Sattler, K., 1995, "New Aspects of Grows Mechanisms for Polycyclic Aromatic Hydrocarbons in Diffusion Flames," Chem. Phys. Lett., **233**, pp. 16–22.
[39] Prado, G. P., Lee, M. L., Hites, R. A., Hoult, D. P., and Howard, J. B., 1997, "Soot and Hydrocarbon Formation in a Turbulent Diffusion Flame," Proc. Combust. Inst., **16**, pp. 649–661.

Lean Blowout Limit and NO_x Production of a Premixed Sub-ppm NO_x Burner With Periodic Recirculation of Combustion Products

Jochen R. Kalb[1]
e-mail: kalb@td.mw.tum.de

Thomas Sattelmayer

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

The technological objective of this work is the development of a lean-premixed burner for natural gas. Sub-ppm NO_x emissions can be accomplished by shifting the lean blowout limit (LBO) to slightly lower adiabatic flame temperatures than the LBO of current standard burners. This can be achieved with a novel burner concept utilizing spatially periodic recirculation of combustion products: Hot combustion products are admixed to the injected premixed fresh mixture with a mass flow rate of comparable magnitude, in order to achieve self-ignition. The subsequent combustion of the diluted mixture again delivers products. A fraction of these combustion products is then admixed to the next stream of fresh mixture. This process pattern is to be continued in a cyclically closed topology, in order to achieve stable combustion of, for example, natural gas in a temperature regime of very low NO_x production. The principal ignition behavior and NO_x production characteristics of one sequence of the periodic process was modeled by an idealized adiabatic system with instantaneous admixture of partially or completely burnt combustion products to one stream of fresh reactants. With the CHEMKIN-II package, a reactor network consisting of one perfectly stirred reactor (PSR, providing ignition in the first place) and two plug flow reactors (PFR) has been used. The effect of varying burnout and the influence of the fraction of admixed flue gas has been evaluated. The simulations have been conducted with the reaction mechanism of Miller and Bowman and the GRI-Mech 3.0 mechanism. The results show that the high radical content of partially combusted products leads to a massive decrease of the time required for the formation of the radical pool. As a consequence, self-ignition times of 1 ms are achieved even at adiabatic flame temperatures of 1600 K and less, if the flue gas content is about 50–60% of the reacting flow after mixing is complete. Interestingly, the effect of radicals on ignition is strong, outweighs the temperature deficiency and thus allows stable operation at very low NO_x emissions. [DOI: 10.1115/1.2061267]

Introduction

For the reduction of the extremely temperature-sensitive NO_x production in industrial combustors, low-temperature combustion of a perfectly premixed fuel-air fresh mixture with self-ignition after admixing of hot flue gas plays an important role. The basic effect is that adiabatic mixtures of preheated fresh mixture and hot combustion products self-ignite even below the lean blowout limit (LBO) of aerodynamically stabilized burners within technically reasonable ignition delay times.

This combustion process, usually called "diluted combustion," has already been used for the reduction of NO_x emissions in low-temperature combustors. In some cases [1,2], premixed reactants are mixed with a certain amount of inert gases, which decrease the temperature rise during subsequent chemical reaction. The inert gases may be recirculated combustion products or highly preheated, oxygen-depleted air (HiTAC, [3]). Dilution can be rather high, if the combustor temperature is held above the self-ignition threshold. A volumetric combustion mode (compared to the flame front combustion mode) is generated that is often called "flameless" combustion because the visible flame front vanishes [2]. If incompletely burnt combustion products are used for the dilution of the reactants, the radical pool in the dilution gas has the additional benefit of maintaining efficient self-ignition down to lower adiabatic flame temperatures.

The Process of Periodic Recirculation of Combustion Products. The conceptual basis of the periodic mixing combustion process is the periodic mixing of partially or fully burnt products to the perfectly premixed gaseous fuel and air injected in a flow reactor. This is a steady process with spatial periodicity of admixing local reaction products from the upstream reaction zone to injected fresh mixture, subsequent self-ignition and reaction.

As an important design criterion, the flow pattern inside the reactor volume must lead to a strong interaction of the flows of local products and fresh mixture, leading to stable combustion with adiabatic flame temperatures as low as required for NO_x emissions of, for example, 1 ppm(v) (15% O_2). As shown schematically in Fig. 1, each jet of premixed fresh mixture that is injected along the circumference of the axis-symmetric combustor is being mixed with the products from the adjacent upstream reaction zone. During the mixing process, self-ignition is gradually initiated in the hot zones of the shear layers with low concentration of the reactants. This combustion reaction then continuously

[1]To whom correspondence should be addressed.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. GT2004-53410.

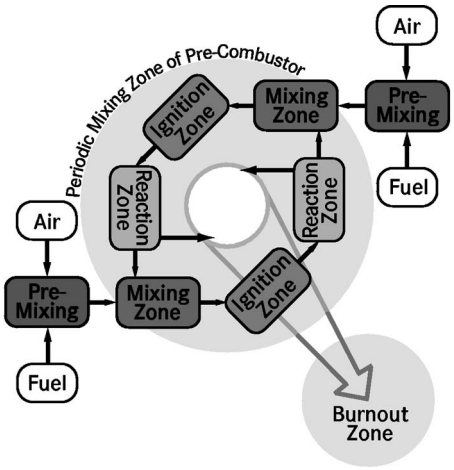


Fig. 1 Sub-ppm process with spatially periodic admixing of combustion products

converts reactants to additional products, which act in the next step as diluent after contact with the next jet of reactants. In the example depicted in Fig. 1 this effect is repeated two times per revolution.

A possible configuration of the flow reactor is a cylindrical chamber with inclined fresh gas injectors (Figs. 2 and 11). If the residence time in the cylindrical zone of circumferentially periodic mixing, ignition, and reaction is limited to the minimum required for stable operation, the gases leaving this zone (and thus the flow reactor) in axial direction will not yet be fully burnt. In these cases, a separate burnout zone is required. This zone can be established, for example, in a cylindrical extension of the reactor with a smaller cross section. The periodic processes of mixing, ignition, and reaction shown schematically in Fig. 1 are concentrated then in the precombustor (Fig. 2, front section), whereas burnout is achieved in the next stage further downstream in a burnout combustion chamber (Fig. 2, rear section).

CFD Feasibility Study. To optimize the fluid mechanics of the precombustor a CFD study using FLUENT 6.0 has been conducted. The reaction volume consisted of a three-dimensional mesh of approximately 400,000 tetrahedral cells and had adiabatic walls. The Reynolds-stress turbulence model as well as the two-step reaction mechanism for methane with six species were used with their standard parameter sets.

For the calculation of the reaction rates, the finite-rate–eddy-dissipation model, which actually uses the minimum of the rates predicted by either an Arrhenius approach or by the Magnussen EBU model [4,5], was used. Although this approach is certainly not the most sophisticated method of modeling combustion processes based on self-ignition, it has been shown previously that the EBU model is suitable for simulating (nonpremixed) diluted combustion [6,7]. The empirical constants A and B included in the EBU model have been left unchanged at their standard values $A=4.0$ and $B=0.5$. In this simulation of combustion in a self-ignition regime, the influence of constant B on the rates remains insignificant, but the heat release is proportional to A. At least in HiTAC combustion, lower rates than calculated with $A=4.0$ are observed, so that a diminished value of A was proposed [7]. On the other hand, as will be shown later in the reaction kinetics study, the elevated radical concentrations accelerate ignition in the sub-ppm NO_x combustion process.

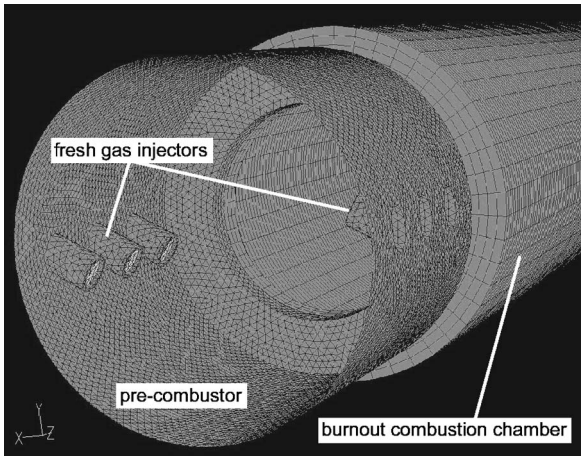


Fig. 2 Geometry of the sub-ppm NO_x Burner and the burnout combustion chamber

For the turbulent Schmidt and Prandtl number of the reacting gas mixture, the default values of $Sc_t=0.7$ and $Pr_t=0.85$ were retained [4]. In the calculation, the computational domain consisted always of the precombustor and a cylindrical burnout combustion chamber downstream. As future testing will be done at atmospheric pressure in the first phase, $p=0.1$ MPa was chosen for these simulations.

In selecting a suitable geometry, the idea of replacing aerodynamically stabilized premix burners in annular combustors of heavy-duty gas turbines by a set of precombustors played an important role. For this purpose a promising geometry for the precombustor is a cylindrical reaction volume of 0.23 m i.d. with two opposite rows of injectors (Fig. 2). With this configuration, the rotational flow is driven by the mass flow of the reactants injected into the chamber with an azimuthal velocity component. Typical velocities of 130 m/s are obtained for the mass flow of the fresh mixture at a reasonable pressure drop of the entire system. The angular momentum of the jets leads to a rotation of the bulk flow with an average velocity of ~80 m/s. The inflow velocity in the injectors of 130 m/s corresponds to methane combustion with thermal power of 150 kW (LHV) in the presented simulation. Because of the interaction of the injected jets and the recirculation flow, a system of secondary vortices is generated, which augments the mixing process considerably.

For this configuration the simulated field of total temperature (Fig. 3(*a*)) is plotted in a plane perpendicular to the burner axis penetrating the middle pair of injector ports. The premixed fresh gas is injected at a temperature of 700 K. It is then mixed with hot gases and heated until combustion begins. Although a fairly simple approach for combustion modeling is applied, the simulations are capable of predicting the very important effect of incomplete burnout ($T<T_{Ad}$) of the jets in the precombustor before they come in contact with the reactants emitted from the injector on the other side. Figure 3(*b*) displays the distribution of the CO mass fraction, which is another indicator of burnout. The maximum concentration is reached after approximately three-quarters of the flow-path length between the injector ports, and, for the admixed flue gas, an elevated value is still shown. In addition to the circular structure shown in the two plots of Fig. 3, further structures are visible near the combustor wall. They are caused by the abovementioned systems of secondary vortices between two fresh gas jets.

In summary, the simulation of the sub-ppm burner prototype shows the essential features of the process that were presented before in Fig. 1: admixing of flue gas to the fresh mixture, subsequent self-ignition, (incomplete) reaction, admixture of flue gas, and so on. Therefore, the precombustor geometry employed for assessing the proposed process via CFD modeling seems to provide a suitable configuration for practical hardware. The reason for selecting only two injector rows is the required minimum residence time for heat release, which leads to bulky designs for higher numbers of injectors.

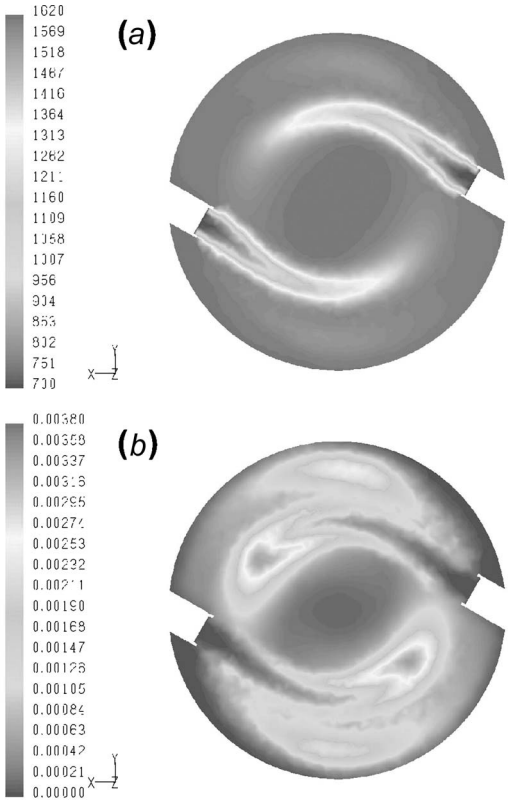


Fig. 3 (*a*) Total temperature (in Kelvin) and (*b*) CO mass fraction

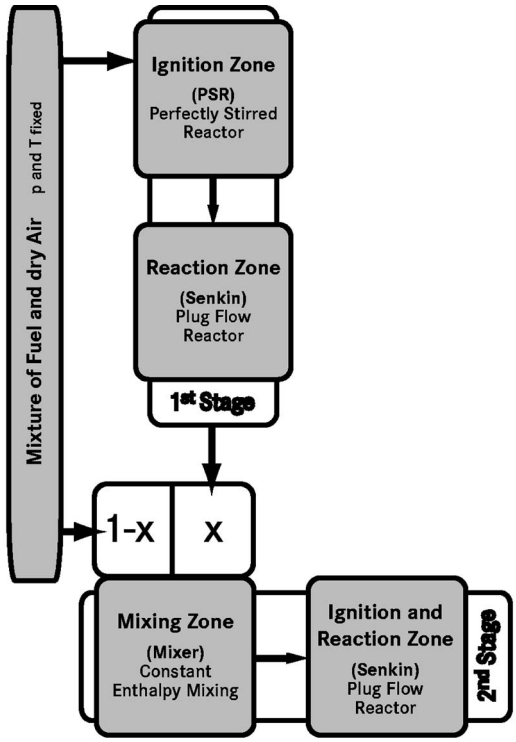


Fig. 4 Schematic diagram of the reactor network

Reaction Kinetics Study

Because including detailed chemical reaction mechanisms for fuel oxidation and nitrogen chemistry into the CFD simulation was not feasible in the current framework, it was decided to simulate the reaction kinetics of the proposed combustion process regarding its ignition performance and nitrogen oxide production characteristics separately. Therefore, a model of the process was set up using the CHEMKIN-II software package [8].

As demonstrated in the previous section, the interaction between the injected fresh gas jet and the flow of local reaction products from upstream in the CFD model is strong enough for sufficiently fast (gradual) mixing of reactants and products affecting fast (also gradual) self-ignition of the reactants. But in order to simulate the reaction kinetics of the idealized process, instantaneous perfect mixing was assumed in the applied reactor network model. Referring to Fig. 1, this means that the residence time in the mixing zone is reduced to an infinitesimally small value. Thus the time evolution of perfect reactant-product mixtures is computed for the ignition and reaction zones. The reacting mixtures are assumed to be homogeneous only in the direction transverse to the ("circular") main flow.

The Idealized Reactor Network. In order to quantify the process parameters yielding fast self-ignition and, thus, assessing the potential for NO_x reduction, the reactor network shown in Fig. 4 was employed. The purpose of the reactor network is the simulation of a single mixing process of the periodic processes in the precombustor. The partially or fully combusted mixture, which is produced by the reaction of the upstream fuel-air jet in reality, is generated in a "first stage" consisting of the perfectly stirred reactor (PSR) [9] and one plug flow reactor (PFR) [10]. The ignition and reaction after admixing of fresh reactants ("second stage") is studied in the second PFR. As in reality, both stages are supplied with reactants of the same equivalence ratio, temperature and pressure.

The First Stage. The PSR is essential, as the fresh mixture at constant preheating temperature of 700 K does not ignite in a PFR alone. In the calculations, the residence time in the PSR was reduced to the minimum required for ignition. As a consequence, the PSR acts as an artificial ignition source, which delivers combustion products with the minimum possible burnout. The temperature in the PSR is typically 150–300 K below the adiabatic flame temperature, and radical concentrations have not yet reached their maximum values. Typically, these increase further in the PFR of the first stage. This strategy leads to maximum flexibility, since the burnout and the composition of the combustion products can be influenced in a wide range by selecting the residence time in the PFR and is therefore often used for modeling perfectly premixed flames [11,12].

In order to evaluate the influence of the flue gas burnout, its temperature and its radical content on the self-ignition and NO formation after mixing, the flue gas is extracted from the first stage after predefined residence times. As OH is the most important radical for starting methane oxidation [13], its mole fraction χ_{OH} has been chosen as indicator of burnout. Five different criteria for the selection of the residence time in the PFR of the first stage were selected. These are:

- PSR exit: minimum burnout
- maximum of χ_{OH}: low burnout
- 80, 60, 40, and 20 % decay of χ_{OH} on the basis of the difference between maximum and equilibrium value
- decay of χ_{CO} to the 1.5-fold of its equilibrium value (criterion for good burnout)
- and finally equilibrium

The Second Stage. The starting point of the second stage is the instantaneous constant enthalpy perfect mixing of premixed fresh reactants and combustion products from the first stage (mixing zone, Fig. 4). The free parameter of the mixing process is the mole

fraction x of combustion products from upstream (in the following also termed flue gas) to be mixed with the fresh mixture $(1-x)$. The PFR of the second stage, modeling the ignition and reaction zone, then delivers information about the ignition delay time, which is defined as the time delay between mixing and half of the temperature rise after ignition.

Parameters. In the calculations, the worst case was studied. Pure methane is best suited for this purpose and was chosen as fuel because the higher hydrocarbons in fuels of technical interest, such as natural gas, lead to better ignition performance. The preheating temperature of the fresh mixture was 700 K. The calculations were performed for adiabatic flame temperatures from 1500 to 1900 K, with corresponding equivalence ratios ranging from 0.34 to 0.55. Calculations were made for $p=0.1$ MPa (atmospheric test conditions) as well as for 1.0 and 2.0 MPa (range of typical gas turbine combustor pressures). The fuel equivalence ratio was always adjusted, keeping the flame temperature constant in the different pressure cases.

Reaction Mechanisms. The GRI-Mech 3.0 [14] as well as the Miller and Bowman reaction mechanism [15] have been utilized for all kinetics simulations. GRI-Mech 3.0 encompasses a total of 325 reactions, 106 of which belong to the nitrogen chemistry part, whereas the Miller and Bowman mechanism has a total of 251 reactions, with 100 belonging to nitrogen chemistry. Both mechanisms take into account 53 chemical species, with 47 in common.

Ignition Performance of the Second Stage

Figures 5 and 6 give an overview of the ignition performance. A target value of 1 ms self-ignition delay time was selected in these computations. As the ignition delay time drops with increasing mole fraction of the hot combustion products, a minimum mole fraction for achieving the presumed target value can be computed. In both figures, this minimum flue gas fraction is plotted versus the adiabatic flame temperature. In Fig. 5, results are displayed for the GRI-Mech 3.0 simulations and, in Fig. 6, for the Miller and Bowman mechanism. The influence of the flue gas burnout is analyzed for the three cases of "minimum burnout" (at the PSR exit), "low burnout" (at the maximum of χ_{OH}), and "good burnout" ($\chi_{CO}=1.5\ \chi_{CO,equilibrium}$).

For GRI-Mech 3.0 and $p=0.1$ MPa (Fig. 5(a)), the interesting result obtained is that low burnout flue gas clearly has the best ignition properties. For flue gas directly from the PSR, the ignition performance is considerably worse because both flue gas temperature and radical content are lower. Nevertheless, ignition, especially for lower flame temperatures is definitely better than for good burnout flue gas, which has practically reached T_{Ad}, but has the lowest radical concentration. At higher pressures of $p=1$ or 2 MPa imposed (Figs. 5(b) and 5(c)), these differences are much weaker, but the combustion products with the maximum OH concentrations lead again to the best ignition performance.

For the Miller and Bowman reaction mechanism at $p=0.1$ MPa (Fig. 6(a)), the ignition performance of the maximum radical case is found to be remarkably better than that of the other cases, but it is worse than that derived from the corresponding simulation with GRI-Mech 3.0. The reason for this difference is that the Miller and Bowman scheme predicts considerably lower blowout temperatures of the PSR. Therefore, ignition after admixture of minimum burnout flue gas is now worse than for good burnout. The elevated radical content cannot fully compensate the temperature deficit. For the high-pressure cases of $p=1$ or 2 MPa (Figs. 6(b) and 6(c)) a clear temperature effect can be seen. Ignition becomes more effective when the burnout of the combustion products is improved. Once again, the influence of burnout on ignition performance vanishes at higher pressure.

In order to further elucidate the ignition behavior in more detail, two cases with $T_{Ad}=1600$ K and pressures $p=0.1$ and 2 MPa will be analyzed using both reaction mechanisms (Figs. 7 and 8).

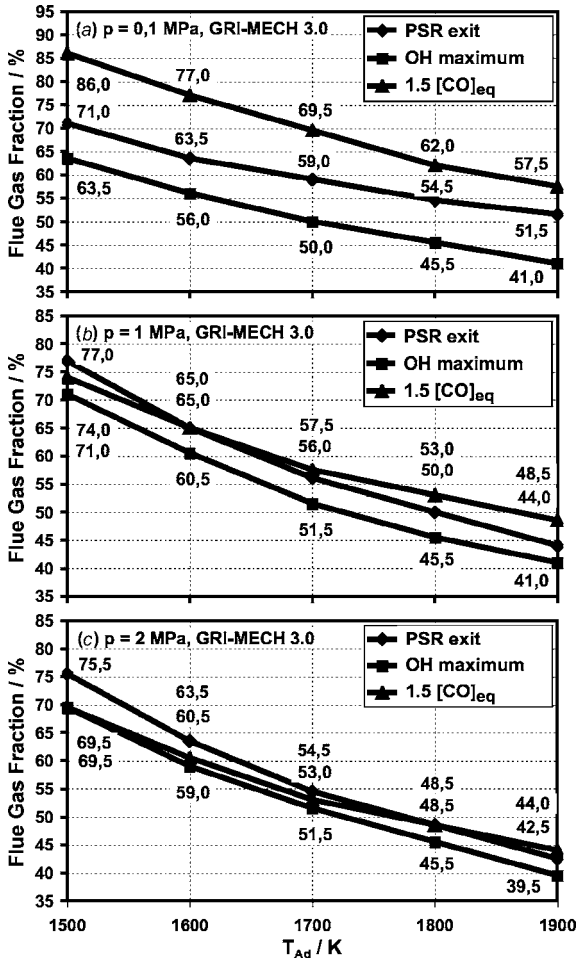


Fig. 5 Flue gas fractions required for $t_{ign}<1$ ms (GRI-Mech 3.0)

Each curve shows the ignition delay time for a constant composition and temperature of the combustion products before mixing.

With GRI-Mech 3.0, a very strong influence of the first stage's burnout on the ignition performance of the second stage is found for the atmospheric case (Fig. 7(a)). As already mentioned, best ignition performance is achieved for flue gas with maximum radical content. Interestingly, the ignition delay time becomes shorter with increasing radical concentration, although the temperature of the combustion products drops. Only the case without residence time in the PFR is an exception. Although the radical content is high, the temperature effect is becoming too strong to be fully compensated. Good burnout in the first stage practically shows no difference to complete burnout. For $p=2$ MPa (Fig. 7(b)) admixing of minimum burnout flue gas shows slowest ignition, but the burnout in the first stage has no influence any more. The maximum radical case still provides fastest ignition, but there is no essential difference to the other cases.

With the Miller and Bowman mechanism, similar results are found for the $p=0.1$ MPa case (Fig. 8(a)), although the "maximum OH" curve reveals slightly greater ignition times than the 80% curve for lower flue-gas fractions. The differences are smaller than before. For $p=2$ MPa (Fig. 8(b)) lower self-ignition times are predicted compared to the GRI-Mech 3.0 simulation, except for the minimum burnout case. Again, the difference between the different first-stage burnout cases diminishes to a large extend. Only the minimum and low burnout cases show worse ignition performance than all other cases.

In the computations, the temperature effect can be separated from the effect of the radicals due to incomplete burnout, if com-

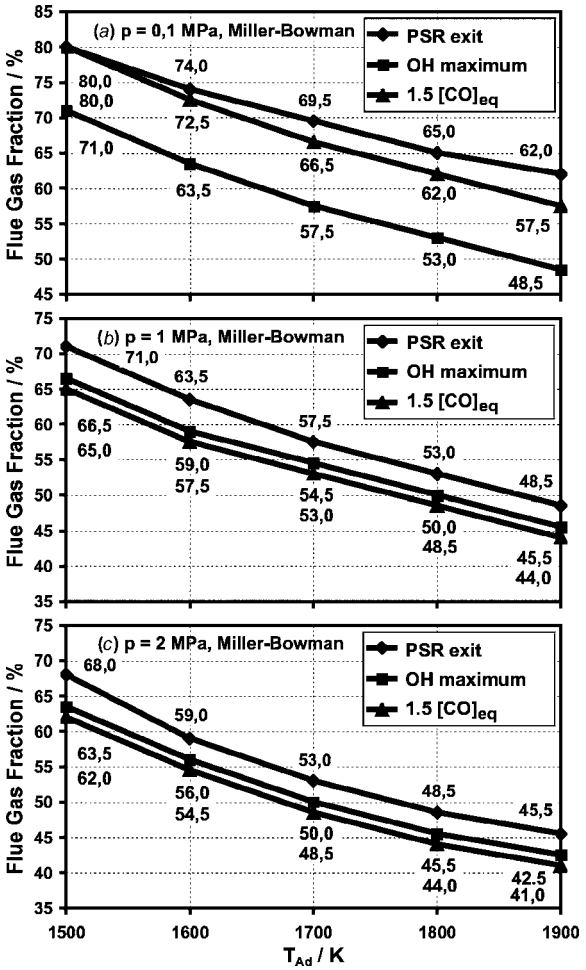


Fig. 6 Flue gas fractions required for $t_{lgn}<1$ ms (Miller and Bowman mechanism)

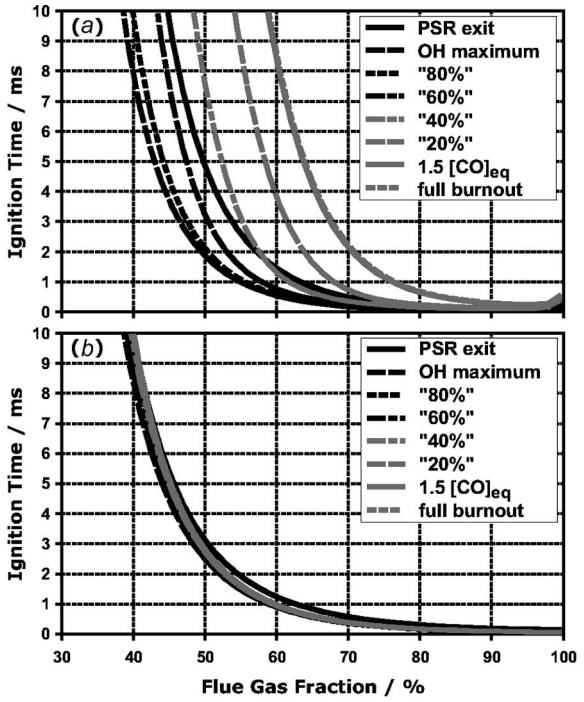


Fig. 7 t_{lgn} for various fractions of flue gas of different burnout for T_{Ad}=1600 K: (*a*) p=0.1 MPa and (*b*) p=2 MPa (GRI-Mech 3.0)

bustion products of equilibrium composition are used in combination with the temperatures of the combustion products of the eight cases shown in Figs. 7 and 8. For GRI-Mech 3.0 and $p=2$ MPa, a clear effect of the combustion product temperatures was derived from calculations with equilibrium composition combustion products (not shown here). This leads to the conclusion that the thermal effect is compensated by the ignition-accelerating effect of the radicals. The corresponding simulations on the basis of the Miller and Bowman mechanism ($p=2$ MPa), however, does not show essential differences to the previous calculation (Fig. 8(*b*)). This indicates that the strong temperature effect keeps ignition times relatively high for the cases of low and lowest burnout, but still lower values are predicted than with GRI-Mech 3.0.

Although the results obtained with the two reaction mechanisms differ slightly, both of them consistently predict that full burnout is not required in the precombustor with periodic admixing of combustion products to the injected reactants. This finding is of major technical relevance, since the precombustor volume can be strongly reduced without negative influence on the LBO. First design studies reveal that precombustors with periodic recirculation of partially burnt combustion products have the potential of approaching the size of aerodynamically stabilized premix burners.

NO_x Formation Characteristics

The analysis of the self-ignition properties of the proposed process revealed the possibility of stable combustion at low adiabatic flame temperatures, which is expected to yield very low NO_x emissions. The results will be presented subsequently. The NO formation in the first and second stages is presented separately. In order to generate NO data for a reference case, the NO emissions for the first stage are only presented for $\chi_{CO}=1.5\ \chi_{CO,equilibrium}$. This corresponds to a conventional aerodynamically stabilized perfectly premixed flame. Increasing the PSR volume allows one

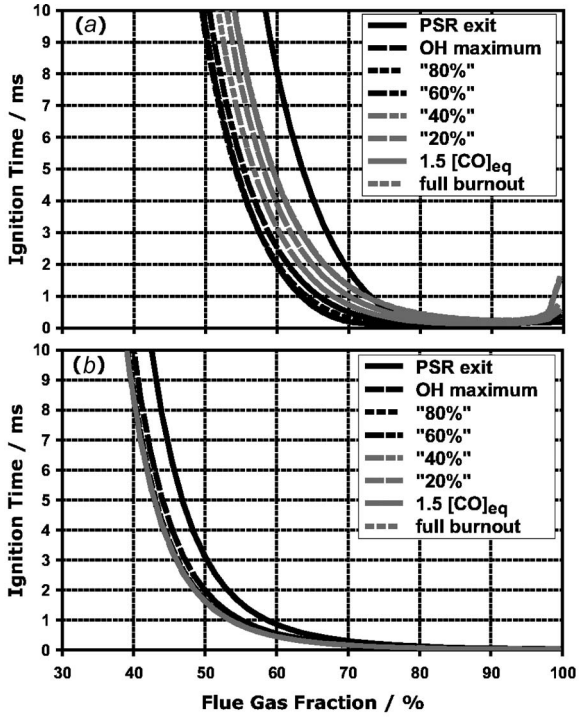


Fig. 8 t_{lgn} for various fractions of flue gas of different burnout for T_{Ad}=1600 K: (*a*) p=0.1 MPa and (*b*) p=2 MPa (Miller and Bowman mechanism)

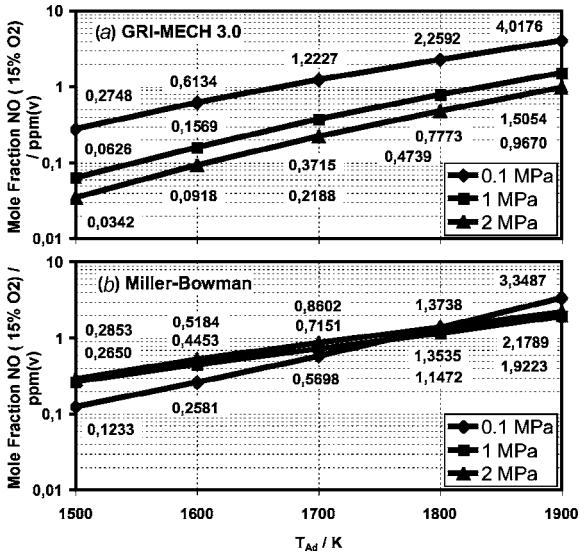


Fig. 9 NO Formation in the first stage: (*a*) GRI-Mech 3.0 and (*b*) Miller and Bowman mechanism

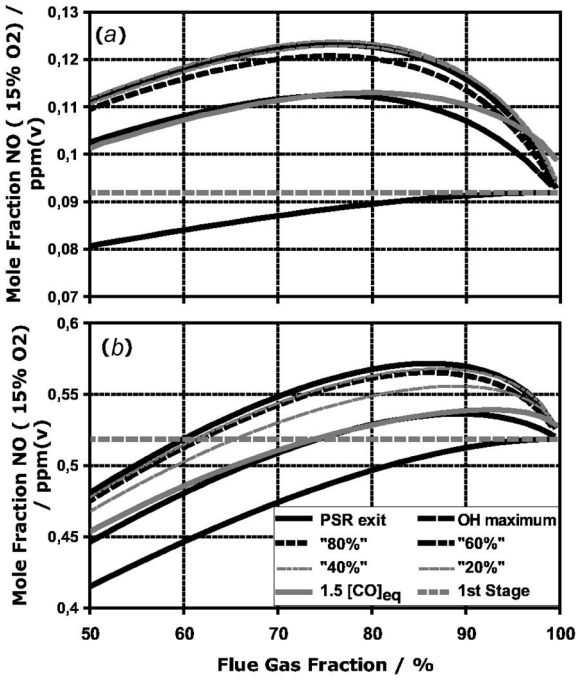


Fig. 10 NO Formation in the second stage: $T_{Ad}=1600$ K and $p=2$ MPa: (*a*) GRI-Mech 3.0 and (*b*) Miller and Bowman mechanism

to extend the calculations to temperatures below the typical LBO limit of well-mixed flames at $1650\ \mathrm{K} < T_{Ad} < 1800\ \mathrm{K}$.

In the next step the NO emissions of the sub-ppm NO_x process are studied for the identical burnout criterion. In the combustor shown in Fig. 2, carbon monoxide oxidation continues in the cylindrical burnout zone.

The First Stage. With GRI-Mech 3.0, as displayed in Fig. 9(*a*), for each of the three pressure cases the typical strong increase of the NO emission with increasing adiabatic flame temperature is predicted. Surprisingly, there is a clearly visible decrease of the NO mole fraction with increasing pressure. This effect decreases with increasing flame temperature. For the pressure of $p=2$ MPa, sub-ppm NO operation is predicted for adiabatic flame temperatures below $T_{Ad}=1900$ K. Since in experiments much higher NO emissions are typically measured for $T_{Ad}=1900$ K, there is some evidence that GRI-Mech 3.0 considerably underpredicts the NO formation at high pressures and low temperatures.

In previous investigations on the NO_x formation of methane-fired, lean-premixed combustion in high-pressure jet-stirred reactors, also using GRI-Mech 3.0 for simulation, only a relatively weak pressure dependency of χ_{NO} with decreasing χ_{NO} for increasing pressure was found, both experimentally and theoretically [16]. These findings are inconsistent with the presented results, which were obtained with the same reaction mechanism and may be subject to further investigation on the influence of the utilized reactor network model and its parameters.

For the Miller and Bowman mechanism (Fig. 9(*b*)), the pressure dependence vanishes for typical LBO temperatures of aerodynamically stabilized flames. For the low-temperature cases below $T_{Ad}=1700$ K, χ_{NO} increases with increasing pressure. The results obtained with this mechanism show better consistency with experimental findings, since sub-ppm NO operation is predicted for adiabatic flame temperature below 1750 K.

Obviously both reaction mechanisms show remarkable differences concerning their NO_x-prediction capabilities. It was noted earlier [12] that the pressure dependence of the NO_x formation at higher pressures is not very well understood, as it differs from that of low-pressure flames. Unfortunately, validation of reaction mechanisms has, in the past, very often been focused on low pressures and high temperatures. As a consequence, modeling of high-pressure diluted combustion below 1700 K is a challenge. So, great care has to be taken using available reaction schemes, when lean combustion in the low-temperature regime is simulated.

For a simple comparison of the two reaction mechanisms, a linear sensitivity analysis considering NO production was made [10,17]. One interesting finding is, for example, that for reactions, which are most important for NO formation, some characteristic changes in the relative magnitude of the sensitivity coefficients for varying pressure are different for the two mechanisms. But the origin of the discrepancy is not yet clear.

NO Formation of the Process With Periodic Recirculation of Combustion Products and Burnout Zone. At the inlet of the second stage, flue gas from the first stage, which already contains NO, is admixed to the fresh reactant mixture. From the technical point of view, the data for $T_{Ad}=1600$ K and $p=2$ MPa are of major importance for the implementation of the process in a gas turbine. In Fig. 10, χ_{NO} is plotted versus the mole fraction of the combustion products in the mixture. The plot is restricted to more than 50%, because lower values lead to long ignition delay times, which are less interesting for the technical implementation. Since extremely high recirculation rates are difficult to achieve, the results below 70% have the highest relevance for the NO emissions of the process. The corresponding single-stage NO emission is indicated as a horizontal dashed line. For combustion product fractions approaching 100%, all calculated NO emissions converge to the result for the single-stage process. Processes with nearly 100% of admixed flue gas are almost identical to the single-stage combustion process.

For GRI-Mech 3.0, NO formation drops beneath the corresponding single-stage value only in the case of low burnout of the combustion products before mixing. In all other cases, NO emissions increase moderately, but the absolute values are far below the 1 ppm limit.

Interestingly, the Miller and Bowman mechanism predicts a reduction of the NO emissions for all compositions of the flue gas in the interesting range of 50–60 % with particularly low values for the lowest burnout. The absolute values predicted are more

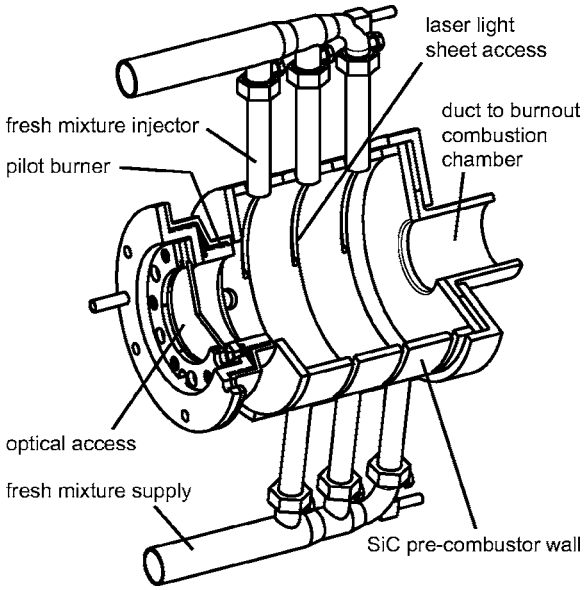


Fig. 11 Cutaway view of the precombustor

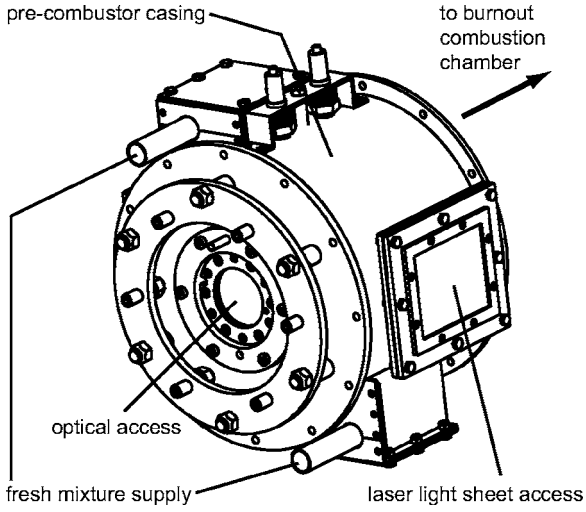


Fig. 12 Drawing of the precombustor test rig

realistic than before.

The atmospheric pressure cases (not shown here) reveal a much weaker influence of staging on the potential NO emission reduction than found for gas turbine relevant pressure levels.

Additionally to very low NO emissions, the predicted CO emissions also remain at low values. For example, with GRI-MECH 3.0 for $T_{Ad}=1600$ K, CO mole fractions below 0.6 ppm (15% O_2) for $p=2$ MPa and 2.5 ppm (15% O_2) for the atmospheric case are computed. The corresponding admixed flue gas fractions were chosen to provide fast ignition, as shown in Fig. 5. For lower adiabatic flame temperatures, CO emissions even decrease.

Conclusion

- A novel concept for the extension of the LBO of premixed flames on the basis of the periodic recirculation of partially burnt combustion products is proposed.
- The technical implementation of this process consists of a precombustor with segmented injection of the reactants and a separate burnout zone.
- CFD and reaction kinetics studies indicate that even for flame temperatures as low as 1500 K, the reaction can be stabilized in the precombustor.
- For perfectly mixed reactants, NO emissions below 1 ppm can be expected in the temperature range below 1750 K.
- Deficiencies of state-of-the-art reaction mechanisms concerning their NO prediction capabilities below 1700 K and for high pressures were detected, which limit the accuracy of the calculated NO emissions.

Outlook

For the next step of the process development, an experimental sub-ppm NO_x burner is investigated. Its configuration is essentially the same as employed in the CFD study. The walls of the cylindric reaction volume are made from sintered SiC with a maximum application temperature of 1900 K (Fig. 11). This allows thermal insulation of the burner to nearly adiabatic conditions. Optical access is provided for laser light-sheet methods in three different planes perpendicular to the burner axis. Optical analysis techniques, such as laser-induced fluorescence (LIF) and particle-image velocimetry (PIV), will be applied for the investigation of the flow field and the distribution of the reaction in the precombustor. In addition, field measurements of the temperature and the chemical composition in the precombustor will be made using conventional probes.

The test rig (Fig. 12) is fired with natural gas. For the transfer of results from atmospheric testing to engine operation, proper scaling of the ignition delay and the residence times is mandatory. In order to compensate the effect of the adverse effect of lower pressure on ignition in the tests, velocity scaling will be applied and hydrogen will be added to the fuel.

Acknowledgment

Financial support by ALSTOM Power Switzerland is gratefully acknowledged.

Nomenclature

Pr_t = turbulent Prandtl number
Sc_t = turbulent Schmidt number
T_{Ad} = adiabatic flame temperature
p = process pressure
t_{Ign} = self-ignition time
x = fraction of admixed combustion products
χ = mole fraction

References

[1] Milani, A., and Saponaro, A., 2001, "Diluted Combustion Technologies," IFRF Combustion Journal, online journal: http://www.journal.ifrf.net/Library/, 200101.
[2] Wünning, J. A., and Wünning, J. G., 1997, "Flameless Oxidation to Reduce Thermal NO-Formation," Prog. Energy Combust. Sci., **23**, pp. 81–94.
[3] Gupta, A. K., 2000, "Flame Characteristics and Challenges With High Temperature Air Combustion," ASME Paper No. IJPGC2000-15087.
[4] Fluent, 2001, *FLUENT 6.0 Documentation, User's Guide*, Fluent, Inc., Lebanon, NH.
[5] Magnussen, B. F., and Hjertager, B. H., 1976, "On Mathematical Models of Turbulent Combustion With Special Emphasis on Soot Formation and Combustion," 16th Symp. (Int.) on Combustion, Combustion Institute, Pittsburgh.
[6] Orsino, S., Weber, R., and Bollettini, U., 2001, "Numerical Simulation of Combustion of Natural Gas With High-Temperature Air," Combust. Sci. Technol., **170**, pp. 1–34.
[7] Yang, W., and Blasiak, W., 2001, "Numerical Study of Fuel Temperature Influence on Single Gas Jet Combustion in Highly Preheated and Oxygen Deficient Air," 3rd Int. Symp. on Advanced Energy Conversion Systems and Related Technologies, Nagoya.
[8] Kee, R. J., Miller, J. A., and Jefferson, T. H., 1989, "CHEMKIN: A General-Purpose, Problem-Independent, Transortable, Fortran Chemical Kinetics Code Package," Sandia Report SAND80-8003 UC-4.
[9] Glarborg, P., Kee, R. J., Grcar, J. F., and Miller, J. A., 1988, "PSR: A Fortran Program for Modeling Well-Stirred Reactors," Sandia Report SAND86-8209 UC-4.
[10] Lutz, A. E., Kee, R. J., and Miller, J. A., 1988, "SENKIN: A Fortran Program

for Predicting Homogeneous Gas Phase Chemical Kinetics With Sensitivity Analysis," Sandia Report SAND87-8248.
[11] Sattelmayer, T., Polifke, W., Winkler, D., and Döbbeling, K., 1998, "NO_x-Abatement Potential of Lean-Premixed GT Combustors," ASME J. Eng. Gas Turbines Power, **120**, pp. 48–58.
[12] Michaud, M. G., Westmoreland, P. R., and Feitelberg, A. S., 1992, "Chemical Mechanisms of NO_x Formation for Gas Turbine Conditions," *24th Symp. (Int'l.) on Combustion*, Combustion Institute, Pittsburgh, pp. 879–887.
[13] Turns, S. R., 2000, *An Introduction to Combustion: Concepts and Applications*, 2nd ed., McGraw-Hill, New York.
[14] Bowman, T., Crosley, D., Davidson, D., Eiteener, B., Frenklach, M., Gardiner, B., Golden, D. M., Hanson, B., Moriarty, N. W., Qin, Z., Smith, G. M., and Yang, H., *GRI-MECH 3.0*, http://www.me.berkeley.edu/gri_mech/
[15] Miller, J. A., and Bowman, C. T., 1989, "Mechanism and Modeling of Nitrogen Chemistry in Combustion," Prog. Energy Combust. Sci., **15**, pp. 287–338.
[16] Rutar, T., and Malte, P. C., 2001, "NO_x Formation in High-Pressure Jet-Stirred Reactors With Significance to Lean-Premixed Combustion Turbines," ASME Paper No. 2001-GT-0067.
[17] Turanyi, T., 1990, "Sensitivity Analysis of Complex Kinetic Systems: Tools and Applications," J. Math. Chem., **5**, pp. 203–248.

L. Elliott

D. B. Ingham

Department of Applied Mathematics,
University of Leeds,
Leeds LS2 9JT, UK

A. G. Kyne[1]
e-mail: fueagk@sun.leeds.ac.uk

N. S. Mera

M. Pourkashanian

Energy and Resources Research Institute,
University of Leeds,
Leeds LS2 9JT, UK

C. W. Wilson
Department of Mechanical Engineering,
University of Sheffield,
Mappin Street,
Sheffield S1 3JD, UK

A Novel Approach to Mechanism Reduction Optimization for an Aviation Fuel/Air Reaction Mechanism Using a Genetic Algorithm

This study presents a novel multiobjective genetic-algorithm approach to produce a new reduced chemical kinetic reaction mechanism to simulate aviation fuel combustion under various operating conditions. The mechanism is used to predict the flame structure of an aviation fuel/O_2/N_2 flame in both spatially homogeneous and one-dimensional premixed combustion. Complex hydrocarbon fuels, such as aviation fuel, involve large numbers of reaction steps with many species. As all the reaction rate data are not well known, there is a high degree of uncertainty in the results obtained using these large detailed reaction mechanisms. In this study a genetic algorithm approach is employed for determining new reaction rate parameters for a reduced reaction mechanism for the combustion of aviation fuel-air mixtures. The genetic algorithm employed incorporates both perfectly stirred reactor and laminar premixed flame data in the inversion process, thus producing an efficient reaction mechanism. This study provides an optimized reduced aviation fuel-air reaction scheme whose performance in predicting experimental major species profiles and ignition delay times is not only an improvement on the starting reduced mechanism but also on the full mechanism. [DOI: 10.1115/1.2131887]

Introduction

Reduction of engine development costs, improved predictions of component life or of the environmental impact of combustion processes, as well as the improved assessment of any industrial process making use of chemical kinetics is only possible with accurate chemical kinetics modeling. However, the lack of accurate and compact combustion schemes for high-order hydrocarbon fuels (e.g., aviation fuel) precludes such techniques from being utilized in many important design processes. To facilitate further development, a procedure for optimizing chemical kinetic schemes must be devised in which a reduced number of reaction steps are constructed from only the most sensitive reactions with tuned reaction rate coefficients. In developing a dependable reduced scheme, released CPU resources can be allocated to improving other aspects of the model, such as a more advanced turbulence model or a better resolved grid. A novel method for automating the optimization of reaction mechanisms of complex aviation fuels is used in this paper to produce a new set of reaction rate parameters for aviation fuel combustion.

Combustion may be modeled using a system of chemical reactions for which the rates of each reaction are known. Databases holding measurements of the reaction rate parameters (A's, β's, and E_a's) are available. These provide the logarithm of the reaction rate as a function of $1/T$, where T is temperature. However, the reaction rate coefficients reported in these databases can often be conflicting and a large variation can be observed over certain temperature ranges that cause problems to occur when calculating the concentration of species for the combustion products away from their equilibrium values.

Combustion of simple fuels, such as hydrogen, can be calculated using an eight-step reaction mechanism, for which the reaction rate data are known with confidence (see [1]). Unfortunately, for complex hydrocarbon fuels, such as aviation fuel, the mechanisms are becoming more elaborate and increasingly complex, approaching 1100 elementary reactions involving some 200 distinct species (see [2]). Hence, as all the reaction rate data are not well known, there is a high degree of uncertainty in the results obtained using these large detailed reaction mechanisms. One-dimensional diffusion flame calculations even become computationally intense at these levels, to the extent of being unable to obtain stable converged solutions, particularly under increasing degrees of strain and pressure when using inlet conditions representative of realistic combustion environments.

To date only a few studies have been performed to predict the combustion of commercial fuels, such as aviation fuel, using detailed chemical kinetics, in comparison to simpler fuels, such as hydrogen, methane, or propane. The very complexity of a fuel such as aviation fuel, comprising a mixture of n-alkanes, cycloalkanes, and aromatics, has necessarily restricted the modeling of its combustion in practical devices to the use of single-step empirical expressions or multistep global schemes to describe the chemistry. Gueret et al. [3] and Dagaut et al. [2] have experimentally investigated the oxidation of both aviation fuel and n-decane in a jet-stirred reactor (JSR) under a wide range of conditions. The aviation fuel combustion under premixed laminar flame conditions was experimentally investigated by Douté et al. [4].

Patterson et al. [5] developed the Aviation Fuel Reaction Mechanism (AFRMv1.1), a detailed chemical kinetics mechanism for aviation fuel, where, for purposes of modeling, aviation fuel is assumed to be 89% n-decane and 11% toluene. The AFRMv1.1 chemical reaction mechanism, consisting of 440 reactions be-

[1]To whom correspondence should be addressed.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. GT2004-53053.

tween 84 species, is capable of predicting the experimentally determined flame structure of Douté et al. [4] with much success using the measured temperature profile as input and is used in this study for comparison purposes.

Detailed comprehensive kinetic mechanisms are necessary to fully understand the fundamental chemistry of combustion, but are impractical for use in industrial codes because of their size. It is eminently useful to develop a moderately sized model that will predict the combustion characteristics of aviation fuel under such a wide a range of conditions as possible. Such a mechanism can be used in a network of interconnected reactor modules—combinations of plug flow, well-stirred, premixed, and nonpremixed jet reactors. Bédat et al. [6] have demonstrated a technique for integrating complex chemistry effects into direct numerical simulations while maintaining computational efficiency.

A reduced reaction mechanism for aviation fuel combustion was developed in Kyne et al. [7], and their mechanism is used in this section to test the ability of genetic algorithms (GAs) to retrieve the reaction rate coefficients. Through this section, the reduced reaction mechanism of Kyne et al. [7], comprising 167 reactions and 63 species, is referred to as the ORIGINAL mechanism. It should be noted that all reactions of sulphur and nitrogen compounds and their effect on combustion have been ignored. It is the purpose of this study to tune the reaction rate coefficients of the ORIGINAL mechanism using GAs in order to match the experimental measurements of Dagaut et al. [2], Gueret et al. [3], and Douté et al. [4].

The mechanism is also shown to satisfactorily predict species composition profiles in both spatially homogeneous and one-dimensional premixed combustion environments. The mechanism is validated for burning velocity and ignition delay time predictions by comparing the results obtained to those predicted by the detailed reaction mechanism AFRMv1.1

We note that various methods have been proposed in order to find a set of reaction rate parameters that give the best fit to a given set of experimental data. Indeed, once the problem is reformulated as an optimization problem, the objective function may be minimized or maximized by various methods. However, in the case of complex hydrocarbon fuels, the objective function is usually highly structured, having multiple ridges and valleys and exhibiting multiple local optima. For objective functions with such a complex structure, traditional gradient-based algorithms are likely to fail. Optimization methods based on the linearization of the objective function (see [8,9]) fall into the same category. A different method, namely, *the solution mapping method*, was proposed by Frenklach et al. [10] and is based on approximating the surface of the objective function by *polynomial response surfaces* generated by means of a relatively small number of computer experiments. Although efficient from the point of view of computational time, the method of *solution mapping* has a major drawback since the optimization is based not on the real values of the objective function, but on a *polynomial response surface* approximating the real landscape of the objective function. For a highly structured surface with multiple valleys and hills, the errors in approximating the object function are large and these errors are likely to affect the results of the optimization procedure. On the other hand, GAs are particularly suitable for optimizing objective functions with complex, highly structured landscapes.

The genetic-algorithm approach has been applied to a related chemical kinetics problem by Polifke et al. [11], wherein a method for the rapid extraction of the chemical kinetics rate coefficients for a simplified combustion mechanism from a given set of (detailed) chemical data is proposed. More recently, an inversion technique, based on a single-objective GA using perfectly stirred reactor (PSR) data, was developed in Harris et al. [12] in order to determine the reaction rate parameters for the combustion of a hydrogen-air mixture. In the above-mentioned study, GAs were used to determine the reaction rate parameters using measurements of the species concentrations in a perfectly stirred reactor. However, many practical combustors, such as internal combustion engines, rely on premixed flame propagation. Moreover, burner-stabilized laminar premixed flames are very often used to study chemical kinetics in a combustion environment. Such flames are effectively one-dimensional and can be made very steady, thus facilitating detailed experimental measurements of temperature and species profiles. Therefore, for laminar premixed flames it is much easier to obtain accurate experimental data to be used in the genetic algorithm inversion procedure. It was found (see [12]) that when using only PSR measurements the problem of retrieving the reaction rate coefficients has multiple solutions that give equally accurate predictions for species concentrations in PSR simulations. However, further investigations have shown that when validated against different experimental data, not all these solutions are accurate in predicting, for example, the species concentration profiles in laminar premixed flames. Hence, in order to obtain a reaction mechanism that can be used for a wide range of practical problems it is required to use an inversion procedure that incorporates multiple objectives. Another reason to use this approach is because sometimes, in practical situations, the amount of experimental data available is limited; thus, it is important to use all the available experimental data, which may come from different measurements for perfectly stirred reactors or laminar premixed flames.

Therefore the single-objective optimization technique proposed by Harris et al. [12] was further extended to a multiobjective genetic algorithm by Elliott et al. [13] in order to include into the optimization process data from premixed laminar flames. The algorithm was found to provide good results for a small-scale test problem, i.e., for the optimization of reaction rate coefficients for a hydrogen-air mixture. In this paper we use a similar approach in order to generate a new reaction mechanism for aviation fuel combustion using both perfectly stirred reactors or laminar premixed flames experimental data. However, the multiobjective optimization problem is reduced to a single-objective optimization problem.

Description of the Problem

Reaction Rate Parameters. In a combustion process, the net chemical production or destruction rate of each species results from a competition between all the chemical reactions involving that species. In this study it is assumed that each reaction proceeds according to the law of mass action and the forward rate coefficients are of the three parameter functional Arrhenius form, namely,

$$k_{fi} = A_i T^{b_i} \exp\left(-\frac{E_{a_i}}{RT}\right) \tag{1}$$

for $i=1,\ldots,N_R$, where R is the universal gas constant and there are N_R competing reactions occurring simultaneously. The rate equations (1) contain the three parameters A_i, b_i, and E_{a_i} for the ith reaction. These parameters are of paramount importance in modeling the combustion processes since small changes in the reaction rates may produce large deviations in the output species concentrations. In perfectly stirred reactors, as well as in laminar premixed flames, if the physical parameters required are specified, and the reaction rates (1) are given for every reaction involved in the combustion process, then the species concentrations of the combustion products may be calculated. It is the possibility of the determination of these parameters for each reaction, based on outlet species concentrations, that is investigated in this paper.

Perfectly Stirred Reactor (PSR) and Laminar Premixed Flames (PREMIX) Calculations. Various software packages may be used for the direct calculations to determine the output species concentrations if the reaction rates are known. In this study the perfectly stirred rector calculations are performed using the PSR FORTRAN computer program, which predicts the steady-state tem-

perature and composition of the species in a perfectly stirred reactor (see [14]), whereas the laminar premixed flame structure calculations were performed using the PREMIX code (see [15]).

We use X_k to denote the mole fraction of the kth species, where $k=1,\ldots,K$ and K represents the total number of species. The temperature and composition that exit the reactor are assumed to be the same as those in the reactor since the mixing in the reactor chamber is intense.

If PSR calculations are undertaken for N_S^{PSR} different sets of reactor conditions, then the output data that are used in the GA search procedure will consist of a set of KN_S^{PSR} mole concentrations, $(X_{jk},\ j=1,\ldots,N_S^{\mathrm{PSR}},\ k=1,\ldots,K)$, where X_{jk} represents the mole concentration of the kth species in the jth set of reactor conditions.

The laminar premixed flame structure calculations were performed using the PREMIX code for burner-stabilized flames with a known mass flow rate. If all the physical operating conditions are specified, then for a given set of reaction rates (1) the profiles of the species concentration and the burning velocity as functions of the distance from the burner's surface are calculated. If PREMIX calculations are performed for N_S^{PREMIX} different sets of operating conditions, then the output data of the code that is used by the GA in the matching process consists of a set of species concentration profiles $(Y_{jk}(x),\ j=1,\ldots,N_S^{\mathrm{PREMIX}},\ k=1,\ldots,K)$ and a set of burning velocity profiles $(V_j(x),\ j=1,\ldots,N_S^{\mathrm{PREMIX}})$, where Y_{jk} is the profile of the mole concentration along the burner for the kth species in the jth set of operating conditions and V_j is the burning velocity profile for the jth set of operating conditions.

Both PSR and PREMIX programs are not stand-alone programs, they are designed to be run in conjunction with preprocessors from the CHEMKIN library (see [16]), which handles the chemical reaction mechanism and the thermodynamic and transport properties.

It is the purpose of this paper to determine the reaction rate coefficients A's, β's, and E_a's in equation (1) given species concentrations X_{jk} in a perfectly stirred reactor and/or profiles of species concentration Y_{jk} and the burning velocity profiles V_j in laminar premixed flames.

Reformulation of the Problem as an Optimization Problem

An inverse solution procedure attempts, by calculating new reaction rate parameters that lie between predefined boundaries, to recover the profiles of the species (to within any preassigned experimental uncertainty) resulting from numerous sets of operating conditions. The inversion process aims to determine the unknown reaction rate parameters $((A_i,b_i,E_{a_i}),\ i=1,\ldots,N_R)$ that provide the best fit to a set of given data. Thus, first a set of output concentration measurements of species is simulated (or measured). If experimental measurements are taken for N_S^{PSR} different sets of PSR operating conditions and/or N_S^{PREMIX} different sets of PREMIX operating conditions, then the data will consist of a set of KN_S^{PSR} concentration measurements of species for PSR calculations and KN_S^{PREMIX} profiles of species concentrations for PREMIX calculations.

GA-based inversion procedures seek for the set of reaction rate parameters that gives the best fit to these measurements. In order to do so we consider two new objective functions that compare predicted and measured species concentrations for PSR and PREMIX simulations,

$$f_{\mathrm{PSR}}((A_i,b_i,E_i)_{i=1,N_R}) = \left\{ 10^{-8} + \sum_{j=1}^{N_S^{\mathrm{PSR}}} \sum_{k=1}^{K} W_k \frac{|X_{jk}^{\mathrm{calc}} - X_{jk}^{\mathrm{exp}}|}{X_{jk}^{\mathrm{exp}}} \right\}^{-1} \tag{2}$$

$$f_{\mathrm{PREMIX}}((A_i,b_i,E_i)_{i=1,N_R}) = \left\{ 10^{-8} + \sum_{j=1}^{N_S^{\mathrm{PREMIX}}} \left(\sum_{k=1}^{K} W_k \frac{\|Y_{jk}^{\mathrm{calc}} - Y_{jk}^{\mathrm{exp}}\|_{L^2}}{\|Y_{jk}^{\mathrm{exp}}\|_{L^2}} \right) \right\}^{-1} \tag{3}$$

where

- X_{jk}^{calc} and Y_{jk}^{calc} represent the mole concentrations of the kth species in the jth set of operating conditions using the set of reaction rate parameters $((A_i,\beta_i E_{a_i}),\ i=1,\ldots,N_R)$.
- X_{jk}^{exp} and Y_{jk}^{exp} are the corresponding experimentally measured values.
- $\|\ \|_{L^2}$ represents the L^2 norm of a function, which is numerically calculated using the trapezoidal rule.
- W_k are different weights that can be applied to each species, depending on the importance of the species. These weights are set to zero for those species for which no experimental measurements are available. The small constant 10^{-8} was introduced in order to avoid numerical overflow.

It should be noted that the fitness function (2) is a measure of the accuracy of species concentrations predictions obtained by a given reaction mechanism for PSR simulations, whereas the fitness function (3) is a measure of the accuracy in predicting species concentrations and burning velocity profiles in the case of laminar premixed flames. We note that several other fitness functions have been investigated in order to develop a fast and efficient optimization process (see [13]). For example, promising results have been obtained also if the fitness function (3) is replaced with a function that measures only the location and the amplitude of the peak in the flame speed and in the species concentration profiles. Overall, the fitness functions given by Eqs. (2) and (3) were found to outperform other possible expressions for the objective function.

In this paper, all the species for which measurements exist have been given an equal weight $W_k=1$, but one of the advantages of the expressions (2) and (3) is the flexibility in dealing with species of various importance. For example, important species can be given higher weights. Moreover, the expressions (2) and (3) allow adaptive objective functions. For example, if during the optimization process some of the most important species have already been accurately predicted, but other species that are more difficult to predict can still be improved, then the weights may be changed automatically to focus on those species and to direct the GA toward tuning the reactions involved in the production destruction of those species.

However, the fitness functions (2) and (3) can still be adapted to accommodate various types of available experimental data. For example in this study, the objective function (2) is modified in order to adapt it to the available experimental data by replacing the sum of absolute values in the expression (2) by L^2 norms of the species profiles considered as functions of the various temperatures given in the set of experimental data at various operating conditions. This was done to avoid cases when a bias arose through small levels of species' concentrations being fitted when there was a large percentage difference in the predicted and experimental point measurements. Here the difference in a continuous profile was considered to be a better judge of the correlation between measured and computed species' concentrations.

Genetic Algorithm Inversion Technique

There is a vast variety of genetic-algorithm approaches to optimization problems. GAs are, in general, a very efficient optimization technique for problems with multiple, often conflicting, objective functions. However, most GA approaches require that all the objective functions are evaluated for every individual generated during the optimization process. For the problem considered in this paper, if the number of PREMIX runs, N_S^{PREMIX}, is compa-

rable to the number of PSR runs, N_S^{PSR}, then the objective function f_{PREMIX} requires a much longer computational run time than f_{PSR}. Therefore, Elliott et al. [13] proposed a multiobjective GA for hydrogen mechanism optimization that avoids PREMIX evaluations at every generation. The algorithm is based on alternating the fitness functions that are used for selecting the survivors of the populations generated during the GA procedure. In this paper we use the experimental data of Dagaut et al. [2], Gueret et al. [3], and Douté et al. [4], which only includes species measurements for one PREMIX operating condition and 18 PSR operating conditions. Therefore, for the problems considered, the two objective functions f_{PSR} and f_{PREMIX}, are comparable from the point of view of computational time and the alternating technique proposed by Elliott et al. [13] is not required. In order to maximize the amount of information used at every generation, then the individuals (i.e., possible candidates) are evaluated using both fitness functions and the GA proceeds by optimizing the combined function

$$f = f_{PSR} f_{PREMIX} \tag{4}$$

In order to optimize this combined fitness function we consider a single-objective float number encoded GA similar to the one proposed by Michalewicz [17]. The genetic operators and the parameters used for this GA were taken as follows:

- population size $n_{pop}=50$
- number of offspring $n_{child}=60$
- uniform arithmetic crossover, crossover probability $p_c=0.65$
- tournament selection, tournament size $k=2$, tournament probability $p_t=0.8$
- nonuniform mutation, mutation probability $p_m=0.5$
- elitism, elitism parameter $n_e=2$

Various GAs, with various types of encoding, operators, and parameters may be used, but similar results are obtained. We have opted for this GA because, in general, it is known that real coded GAs perform better than binary-coded genetic algorithms for high-precision optimization problems. Moreover, the float-number encoding of the variables enable natural implementation of fine local tuning processes. Indeed, in this algorithm we use a specialized operator in order to take advantage of the floating-point representation of the solution space to increase the rate of convergence of the algorithm. The nonuniform mutation operator (see [17]) is an operator that enables fine local tuning and is defined as follows: if $x=\langle v_1, v_2, v_3, \ldots, v_k, \ldots, v_n\rangle$ is a chromosome, and the element v_k was selected for mutation (domain of v_k is $[l_k, u_k]$), the result is a vector $\langle v_1, v_2, v_3, \ldots, v_k', \ldots, v_n\rangle$ where v_k' is given by

$$v_k' = \begin{cases} v_k + \delta(i, u_k - v_k) & \text{if a random digit is 0} \\ v_k - \delta(i, v_k - l_k) & \text{if a random digit is 1} \end{cases}$$

where i is the number of generations performed and the function $\delta(i,y)$ returns a value in the range $[0,y]$ such that the probability of $\delta(i,y)$ being close to 0 increases as i increases. This property causes this operator to search the space uniformly, initially, when i is small, and, locally, at later stages. We have used the following function:

$$\delta(i,y) = \delta(i,I,y,r) = y(1 - r^{(1 - i/I)^b}) \tag{5}$$

where r is a random number in the interval [0,1], I is the maximal generation number, and b is a system parameter determining the degree of nonuniformity.

A very important aspect of the GA inversion procedure is the size of the search space that is available. For this we constrained the new sets of rate constants generated every iteration to lie between predefined boundaries that represent the uncertainty associated with the experimental findings listed in the National Institute of Standards and Technology (NIST) database (see [18]). The constraints are imposed by associating a very low fitness function to the individuals that violate these constraints.

It is worth noting that when working with accurate experimental data, or numerically simulated data, the two fitness functions (2) and (3) have a common global optimum. Therefore, the two objective functions are not conflicting and the method we have proposed can be used, although it searches for a single common optima rather than a set of nondominated solutions. It is worth noting that if the experimental data are corrupted with a large level of noise, then the two objective functions (2) and (3) may be conflicting and then Pareto-oriented techniques have to be used in order to construct the front of nondominated solutions (i.e., a set of possible optima). In this situation, other considerations have to be used to make a decision on what is the best location of the optimum.

Numerical Results. The reaction rate parameters are tuned using the experimental data of Dagaut et al. [2], Gueret et al. [3] and Douté et al. [4]. First $N_S^{PSR}=18$ different PSR operating conditions have been considered, and they correspond to the sets: (i) $p=10$ atm, $\Phi=1.5$, $\tau=0.5$ s and the initial fuel mole fraction of 10^{-3} and (ii) $p=40$ atm, $\Phi=1.0$, $\tau=2.0$ s and the initial fuel mole fraction of 2.5×10^{-4}. Calculations were performed at various temperatures from 780 K to 1150 K. For PREMIX simulation, the atmospheric pressure-rich premixed aviation fuel/O_2/N_2 flame ($\Phi=1.7$) was chosen as the basis for a comparison to the model as an experimental temperature profile is provided along with extensive mol fraction profiles. The input premixed fuel-oxidizer mixture was 2.95% aviation fuel (2.63% *n*-decane and 0.32% toluene), 28.64% O_2 and 68.4% N_2, and a mass flow rate of 0.0099 g cm^{-3} s^{-1} was employed.

A new reaction mechanism, referred to as GA-MECH, is generated using these experimental data. These mechanism is obtained by tuning the reaction rate parameters of the original mechanism, i.e., the mechanism of Kyne et al. [7]. It should be noted that the reaction rate coefficients for all 167 reactions are allowed to vary and their values are simultaneously tuned. The tuning is automatically performed by the GA by comparing the predicted species concentrations at various operating conditions to the corresponding experimental values. The reactions included in GA-MECH are the same as in the mechanism of Kyne et al. [7], but they use new reaction rate parameters. These two reaction

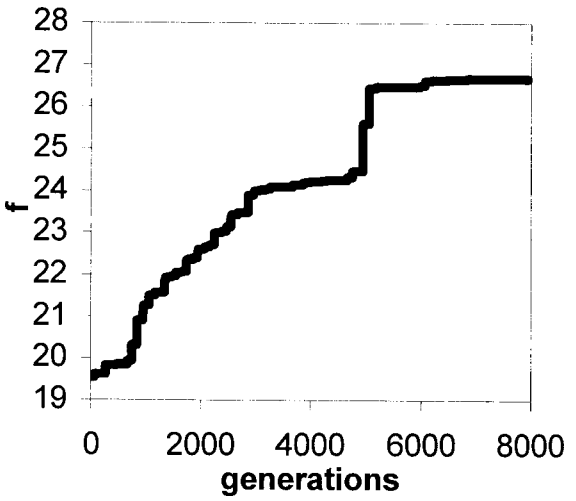


Fig. 1 The combined fitness function *f* given by Eq. (4) that measures the improvement in predicting PREMIX and PSR output species. The line depicts the evolution of the GA-MECH fitness. The ORIGINAL mechanism of Kyne et al. [7] was used as a starting point for the search.

mechanisms, ORIGINAL and GA-MECH are used to produce estimations of the output species concentrations, and these estimations are compared next to the experimental data (see [2–4]). The species concentrations predicted by the detailed mechanism AFRMv1.1 are also included for comparison purposes.

Figure 1 presents the values of the fitness function f, see Eq. (4), as a function of the number of generations. It should be noted that the fitness function f is a measure of the accuracy of species concentrations predictions obtained by a given reaction mechanism for both PSR simulations and laminar premixed flames. Therefore, it is a measure of the efficiency of the mechanism investigated. In Fig. 1, the ORIGINAL mechanism corresponds to the low fitness value obtained at the first generation, whereas the GA-MECH mechanism has a fitness value given by the value of f at the last generation. Figure 1 summarizes the degree of improvement obtained during the optimization process. In arriving at GA-MECH it was not necessary, as it sometimes is with gradient-based optimization algorithms, to restart the procedure when a local optima is reached. Instead satisfactory results were always obtained with just one run. It is interesting to note that less than 1% of the CPU time is associated with the GA. In fact, one generation involves calling the PSR code 50(number of individuals) $\times$ 18(number of PSR conditions) and the PREMIX code 50(number of individuals) $\times$ 1(number of PREMIX conditions). The results presented here are those obtained after several months of running on one processor of a 1.0 GHz workstation. However, there was no real noticeable change in the species profiles after three weeks of running even though the fitness was still rising. In order to cut down on these large optimization timescales, future work will take advantage of the fact that the GA can easily be modified to run efficiently on multiple CPU processors. In addition, preliminary research into GA clustering techniques promises a sixfold decrease in CPU runtimes without any significant loss in accuracy.

Figure 2 presents the predicted concentrations profiles obtained by PREMIX calculations at a pressure $p=1$ atm and fuel-air equivalence ratio, $\Phi=1.7$ for eight of the species for which experimental data are available in Douté et al. [4]. Douté et al. [4] have discussed the difficulties encountered in the H_2O analysis and hence have given an estimation of the accuracy of the analysis as 6% for major species ($O_2, N_2, CO, CO_2, H_2, H_2O$) and as 12% for all other species. This is in addition to a quoted accuracy of the temperature measurements of ±0.1 mm for the position and ±5% for the temperature signal. Taking this into consideration, the predictions of the GA-MECH model are observed to be quite satisfactory.

It can be seen in Fig. 2 that for all the major species (O_2, H_2O, CO, CO_2, H_2) the accuracy of predicting the species' profiles has been substantially increased by the GA-generated reaction mechanism. For these species, and also for some other minor species (see, for example, C_2H_2 and CH_4) GA-MECH clearly outperforms the other two reaction mechanisms considered in this comparison study. These two species, C_2H_2 and CH_4, are clearly underpredicted by the other two mechanisms, whereas GA-MECH brings a substantial improvement in estimating their concentrations. It is worth noting that for C_6H_6, the ORIGINAL and AFRMv1.1 are more accurate for small values of x, the distance from the surface of the burner, whereas for large values of x the GA-MECH provides more accurate results. It should also be pointed out that before any work was done in fitting the experimental species' profiles a rigorous test was made on the GA. A simulated dataset was produced using the original reaction mechanism. Then starting from a random initial population, the GA was used to fit this dataset. The success of this test convinced us that if a solution exists, then the GA will find it. Thus, the fact the GA-generated mechanism fails to accurately model the H_2, C_2H_2, and C_6H_6 profiles indicates that either the mechanism is incomplete or the two sets of experimental data are conflicting. It is interesting to note that the overall improvement in the predictions of the GA-generated reaction mechanism was undertaken with little practical implementation cost necessary. Although the actual running time was three weeks, the actual time taken in manually implementing the methodology was a matter of a few hours in updating the necessary files with additional reaction rate data where necessary.

Overall, for most of the species, GA-MECH is better in predicting experimental data in comparison to both the ORIGINAL and AFRMv1.1 mechanisms. GA-MECH clearly outperforms AFRMv1.1, although AFRMv1.1 consists of 440 reactions between 84 species, whereas GA-MECH and ORIGINAL only use 167 reactions and 63 species. Although this still represents too much detail when implemented into three-dimensional turbulent combustion models, the methodology in the optimization remains unchanged. Future work will be aimed at adapting and reducing mechanisms that can be more readily implemented into multidimensional, practical simulation applications.

Next, the GA-MECH mechanism is used to predict the flame structure of an aviation fuel/O_2/N_2 flame in a jet-stirred reactor for various inlet temperatures. Results are again compared to those obtained by the ORIGINAL and AFRMv1.1 mechanisms. Figure 3 presents the species concentration in a jet-stirred reactor calculated to the PSR program for the most important output species at several operating conditions given by $p=10$ atm, $\Phi=1.5$, $\tau=0.5$ s, and various inlet temperatures. The output species concentrations for unburned fuel ($C_{10}H_{22}+C_6H_5CH_3$) and several typical combustion products, such as O_2, CO, CO_2, H_2, CH_4, are presented. It can be seen that under similar experimental conditions the three reaction mechanisms give very similar concentration profiles for the main oxidation products.

As the experimental data are presented without addressing the accuracy of the species analysis and measurement, the overall agreement between the computational model and experiment can be considered quite good. Although not presented here, it is reported that similar results are obtained for the other operating conditions included in the optimization process, namely, $p=40$ atm, $\Phi=1.0$, $\tau=2.0$ s. Thus, for the main species, GA-MECH proves to be just as effective as the more detailed AFRMv1.1 mechanism, whereas GA-MECH is much more effective from a computational complexity and time point of view. Moreover, if some other species are investigated, then the GA-MECH outperforms the other mechanisms. Figures 3(g) and 3(h) present the concentration profiles for the output species C_6H_6 in a jet-stirred reactor under different conditions of pressure, equivalence ratio, and residence time obtained by the GA-MECH, ORIGINAL, and AFRMv1.1 reaction mechanisms. The results are again compared to experimental data. It can be seen that GA-MECH is substantially more effective in predicting the experimental C_6H_6 concentrations, even when compared to the detailed reaction mechanism AFRMv1.1.

Overall, for most of the cases considered, GA-MECH was found to be better in predicting experimental data when compared to the ORIGINAL or the AFRMv1.1 mechanisms. Substantial improvements in evaluating species' profiles were obtained by GA-MECH for both laminar premixed flames and jet-stirred reactors and represents an accurate model for the prediction of aviation fuel-air flames.

Burning Velocity and Ignition Delay Time Validations. The GA-MECH mechanism, previously validated against experimental data, is used next to predict the burning velocity of aviation fuel-air mixtures. A survey of the literature shows that measurements of the burning velocity of aviation fuel are virtually nonexistent. Therefore, the predictions of the GA-MECH mechanism are compared to those obtained using the more detailed mechanism AFRMv1.1. Predictions obtained by the ORIGINAL mechanism are also included in Fig. 4, which presents the burning velocity as a function of the distance from the surface of the burner for a rich

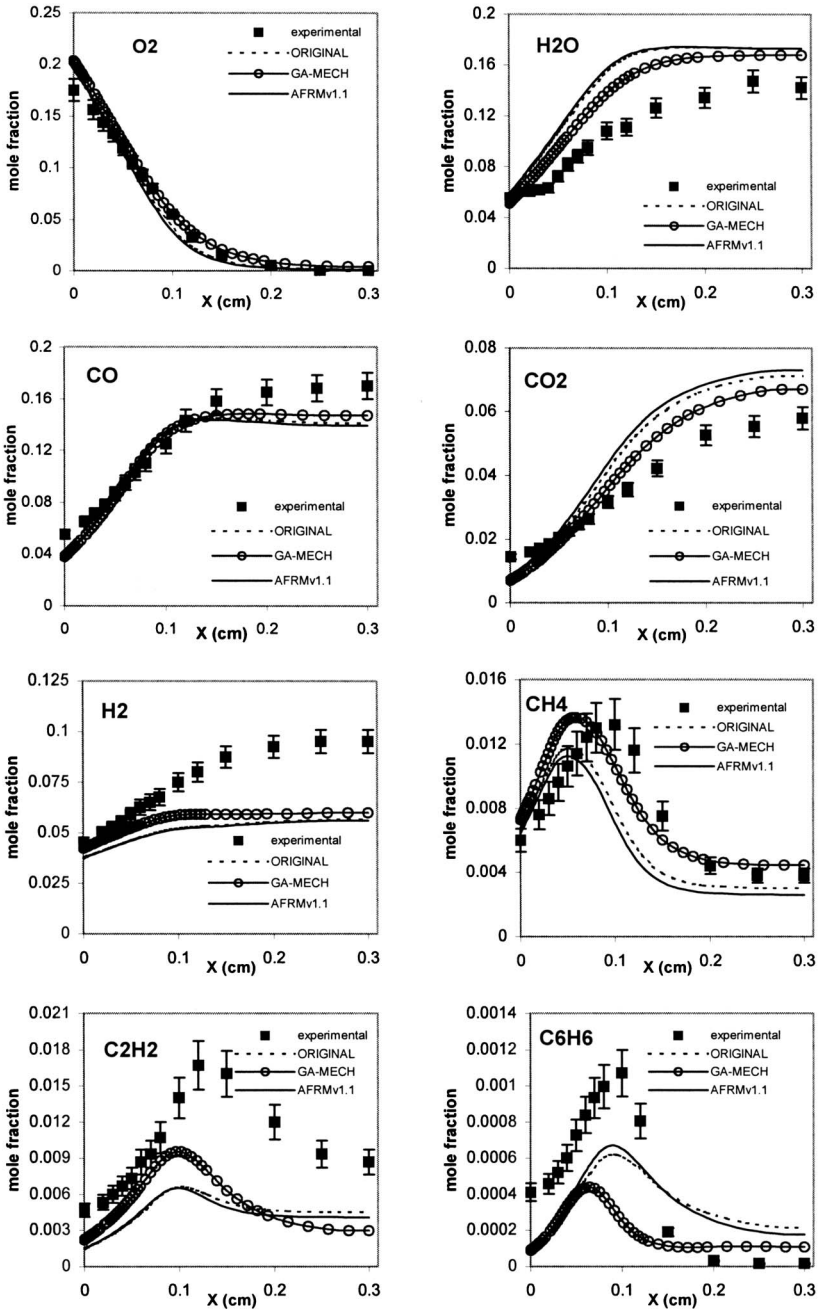


Fig. 2 Concentration profiles of various species as functions of the distance from the surface of the burner for a rich atmospheric premixed aviation fuel flame obtained by the GA-MECH (—○—), the ORIGINAL (···) mechanism of Kyne et al. [7], the AFRMv1.1 reaction mechanism (—), and the measured experimental data of Douté et al. [4] (■)

premixed aviation fuel/O_2/N_2 flame ($\Phi=1.7$). The burning velocity is plotted for various pressures, namely, $p=1$ atm, $p=2$ atm, and $p=5$ atm. It can be seen in Fig. 4 that, under similar experimental conditions, the three reaction mechanisms give similar profiles for the burning velocity, although the burning velocity calculated by the GA-MECH is slightly higher than those estimated by the ORIGINAL and AFRMv1.1 mechanisms.

In order to provide additional validation of the reaction mechanism GA-MECH generated in this study, the predicted ignition delay times are computed using the three reaction mechanisms considered: GA-MECH, ORIGINAL, and AFRMv1.1. In testing the ignition delay time, the SENKIN code of Sandia National Laboratories (see Lutz [19]) was used to predict the time-dependent chemical kinetic behavior of a homogeneous gas mixture in a closed system. For the purpose of this study, the system was considered to be adiabatic under conditions of constant volume and atmospheric pressure, with the initial mole fractions such that the air-fuel ratio is $\Phi=1.0$. Figure 5 compares the ignition delay times predicted by the reaction mechanism GA-MECH generated in this study to the ignition delay times generated by the original mechanism and the detailed AFRMv1.1 mechanism at four different initial temperatures. We note that both the GA-MECH and ORIGINAL mechanisms reproduce the same ignition delay times as the detailed mechanism AFRMv1.1. This indicates

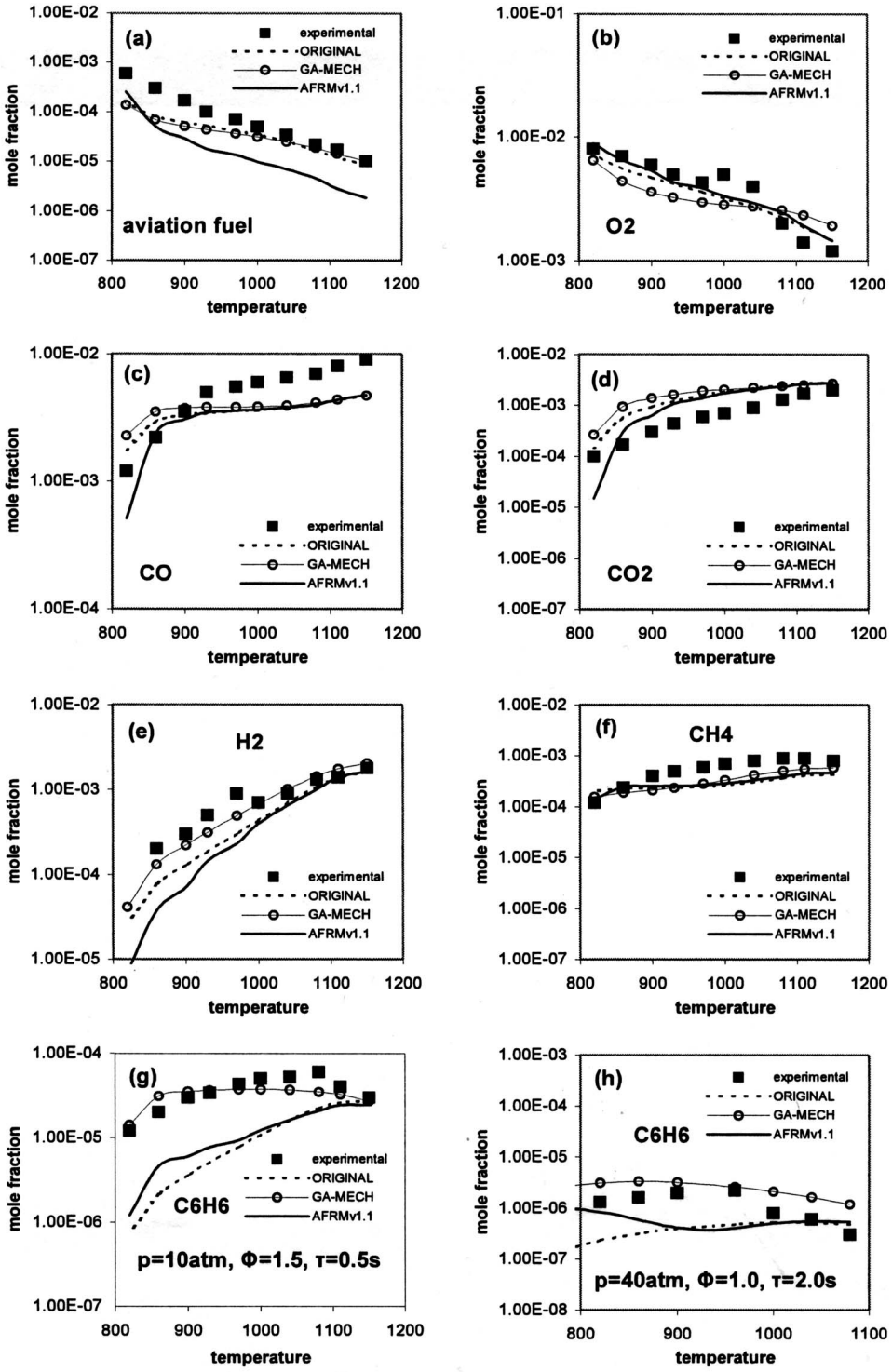


Fig. 3 Concentration profiles of (*a*) Aviation fuel; (*b*) O_2; (*c*) CO; (*d*) CO_2; (*e*) H_2; (*f*) CH_4; and (*g*) C_6H_6 obtained for various output species for aviation fuel oxidation in a jet stirred reactor at p=10 atm, residence t=0.5 s and equivalence ratio=1.5 as functions of the inlet temperatures; and (*h*) profile C_6H_6 at p=40 atm, residence t=2.0 s; and equivalence ratio=1.0. Comparisons are made between GA-MECH (—○—), the ORIGINAL (···) mechanism of Kyne et al. [7], the AFRMv1.1 reaction mechanism (—), and the measured experimental data of Dagaut et al. [2] (■).

that the GA does not just generate a mechanism whose validity only extends as far as those operating conditions included in the optimization process.

Overall, the predictions of the GA-MECH for the burning velocity and ignition delay time were found to be in good agreement with those predicted by the detailed mechanism AFRMv1.1. Moreover, it was seen in the previous section that GA-MECH outperforms the other mechanisms considered when used to match

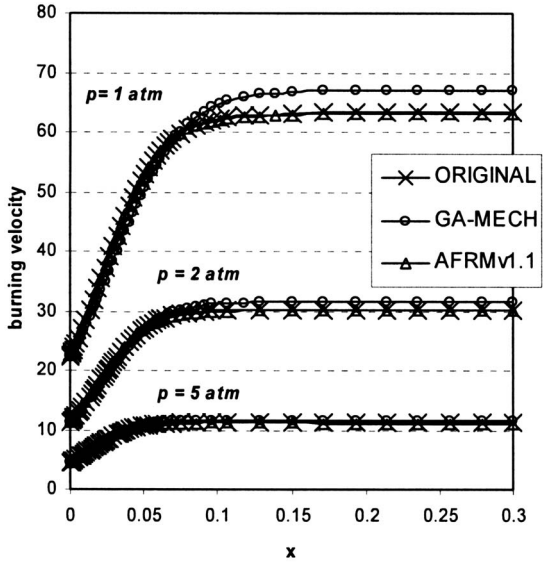


Fig. 4 Burning velocities as functions of the distance from the surface of the burner obtained by the three reaction mechanisms considered for a rich premixed aviation fuel/O_2/N_2 flame (Φ=1.7) at various pressures

the experimentally measured species profiles in both jet-stirred reactors and laminar premixed flames. Overall, the numerical results obtained in this study provide us with much confidence in the predictive capabilities of the GA-generated mechanism.

Conclusions

In this study we have built on previous reaction mechanism optimization work involving GAs by extending the complexity of the system under investigation and by developing a new multiobjective fitness function for generating a new reaction mechanism for aviation fuel-air flames. The mechanism was developed by tuning the reactions rate coefficients to match all the experimental data available for both spatially homogeneous and one-dimensional premixed combustion. The new reaction mechanism generated by the GA was found to successfully predict the flame structure of an aviation fuel/O_2/N_2 flame over a wide range of operating conditions. Moreover, the GA-generated mechanism provided estimates similar to those obtained with a much more detailed mechanism when tested for burning velocities as well as ignition delay time. The results demonstrated that that a GA-optimized reaction mechanism could be generated that outperforms a significantly more detailed reaction scheme with almost three times the number of reactions in predicting PSR and premixed flame species' profiles. The similarity in the predictions of burning velocity and ignition delay times with both optimized and nonoptimized mechanisms is reassuring because it indicates that GA-generated mechanisms should not lose their ability to predict systems outside of their optimization process. Overall, this study has proved that GAs can be used with much confidence to develop new, more effective, reaction mechanisms for the combustion of complex hydrocarbon fuels. The GA-generated mechanism is available for download from http://www.personal.leeds.ac.uk/~fuensm/project/mech.html

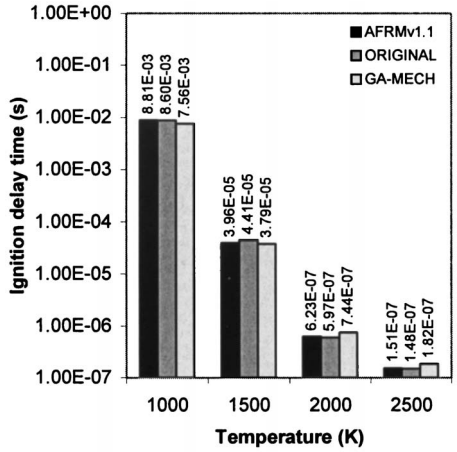


Fig. 5 The ignition delay time at various temperatures T =1000–2500 K and the air/fuel ratio Φ=1.0 as predicted by the three reaction mechanisms: GA-MECH, ORIGINAL, and AFRMv1.1

Acknowledgment

The authors would like to thank QinetiQ, the EPSRC, the Corporate Research Programme from UK MOD and the UK Department of Trade and Industry CARAD Programme for the funding and permission to publish this work.

Nomenclature

A_i = pre-exponential factor for the ith reaction
B_i = non-Arhenius parameter for the ith reaction
Ea_i = activation energy for the ith reaction
k_{fi} = forward reaction rate for the ith reaction
T = temperature
R = universal gas constant
X_{jk} = mole fraction of the kth species in the jth set of PSR conditions
K = total number of species
N_S^{PSR} = total number of PSR conditions
N_S^{PREMIX} = total number of PREMIX conditions
Y_{jk} = profile of the mole concentration above the burner for the kth species in the jth set of PREMIX operating conditions
V_j = burning velocity profile for the jth set of PREMIX operating conditions
f_{PSR} = PSR fitness function
f_{PREMIX} = PREMIX fitness function
$\| \|_{L^2}$ = L^2 norm of a function calculated using the trapezoidal rule
W_k = weighting factor applies to the kth species
ε = constant added to fitness functions to avoid numerical overflow
n_{pop} = population size
n_{child} = number of offspring
p_c = uniform arithmetic cross over probability
p_t = tournament selection probability
p_m = nonuniform mutation probability
n_e = elitism parameter
p = operating pressure
V = reactor volume
Φ = equivalence ratio
τ = residence time
Q = heat loss from reactor

Superscripts

calc = calculated values
orig = original values

References

[1] Dixon-Lewis, G., Goldsworthy, F. A., and Greenberg, J. B., 1975, "Flame Structure and Flame Reaction Kinetics. IX: Calculation of Properties of Multi-Radical Premixed Flames," Proc. R. Soc. London, Ser. A, **346**, pp. 261–275.
[2] Dagaut, P., Reuillon, M., Boetner, J.-C., and Cathonnet, M., 1994, "Kerosene

Combustion at Pressures up to 40 atm: Experimental Study and Detailed Chemical Kinetic Modelling," *Proceedings of the Combustion Institute*, Combustion Institute, Pittsburgh, Vol. **25**, pp. 919–926.
[3] Gueret, C., Cathonet, M., Boettner, J. C., and Gaillard, F., 1990, "Experimental Study and Modelling of Kerosene Oxidation in a Jet-Stirred Flow Reactor," *Proc. of 23rd Int. Symp. on Combustion*, Combustion Institute, Pittsburgh, pp. 211–216.
[4] Douté, C., Delfau, J. L., Akrich, R., and Vovelle, C., 1995, "Chemical Structure of Atmopheric Pressure Premixed n-Decane and Kerosene Flames," Combust. Sci. Technol., **106**, pp. 327–344.
[5] Patterson, P. M., Kyne, A. G., Pourkashanian, M., Williams, A., and Wilson, C. W., 2000, "Combustion of Kerosene in Counterflow Diffusion Flames," J. Propul. Power, **16**, pp. 453–460.
[6] Bédat, B., Egolfopoulos, F., and Poinsot, T., 1999, "Direct Numerical Simulation of Heat Release and NOx Formation in Turbulent Non Premixed Flames," Combust. Flame, **119**(1/2), pp. 69–83.
[7] Kyne, A. G., Patterson, P. M., Pourkashanian, M., Williams, A., and Wilson, C. J., 2001, "Prediction of Premixed Laminar Flame Structure and Burning Velocity of Aviation Fuel-Air Mixtures," *Proc. of ASME Turbo Expo2001: Land, Sea and Air*, June 4–7, New Orleans, ASME, New York.
[8] Milstein, J., 1981, "The Inverse Problem: Estimation of Kinetic Parameters," *Modelling of Chemical Reaction Systems*, K. H. Ezbert, P. Deuflhard, and W. Jager, eds., Springer, Berlin.
[9] Bock, H. G., 1981, "Numerical Treatment of Inverse Problems in Chemical Reaction Kinetics," *Modelling of Chemical Reaction Systems*, K. H. Ezbert, P. Deuflhard, and W. Jager, eds., Springer, Berlin.
[10] Frenklach, M., Wang, H., and Rabinowitz, J., 1992, "Optimisation and Analysis of Large Chemical Kinetic Mechanisms Using the Solution Mapping Method: Combustion of Methane," Prog. Energy Combust. Sci., **18**, pp. 47–73.
[11] Polifke, W., Geng, W., and Döbbeling, K., 1998, "Optimisation of Rate Coefficients for Simplified Reaction Mechanisms With Genetic Algorithms," Combust. Flame, **113** pp. 119–135.
[12] Harris, S. D., Elliott, L., Ingham, D. B., Pourkashanian, M., and Wilson, C. W., 2000, "The Optimisation of Reaction Rate Parameters for Chemical Kinetic Modelling of Combustion Using Genetic Algorithms," Comput. Methods Appl. Mech. Eng., **190**, pp. 1065–1090.
[13] Elliott, L., Ingham, D. B., Kyne, A. G., Mera, N. S., Pourkashanian, M., and Wilson, C. W., 2003, "Multi-Objective Genetic Algorithms Optimisation for Calculating the Reaction Rate Coefficients for Hydrogen Combustion," Ind. Eng. Chem. Res., **42**(6), pp. 1215–1224.
[14] Glarborg, P., Kee, R. J., Grcar, J. F., and Miller, J. A., 1988, "PSR: A FORTRAN Program for Modelling Well-Stirred Reactors," Sandia Report SAND86–8209, Sandia National Laboratories.
[15] Kee, R. J., Grcar, J. F., Smooke, M. D., and Miller, J. A., 1985, "A Fortran Program for Modelling Steady One-Dimensional Premixed Flames," Sandia Report SAND85–8240, Sandia National Laboratories.
[16] Kee, R. J., Miller, J. A., and Jefferson, T. H., 1980, "CHEMKIN: A General-Purpose, Problem-Independent, Transport Table, FORTRAN Chemical Kinetics Code Package," Sandia Report SAND80–8003, Sandia National Laboratories.
[17] Michalewicz, Z., 1996, Genetic Algorithms/Data Structures/Evolution Programs, Third Edition, Springer, Berlin.
[18] Elliott, L., Ingham, D. B., Kyne, A. G., Mera, N. S., Pourkashanian, M., and Wilson, C. W., 2003, "Incorporation of Physical Bounds on Rate Parameters for Reaction Mechanism Optimisation Using Genetic Algorithms," Combust. Sci. Technol., **175**(4), pp. 619–648.
[19] Lutz, A. E., Kee, R. J., and Miller, J. A., 1987, "SENKIN: A FORTRAN Program for Predicting Homogeneous Gas Phase Chemical Kinetics With Sensitivity Analysis," Sandia Report SAND87–8248, Sandia National Laboratories.

Experimental Study on the Role of Entropy Waves in Low-Frequency Oscillations in a RQL Combustor

J. Eckstein

E. Freitag

C. Hirsch

T. Sattelmayer

Lehrstuhl für Thermodynamik,
Technische Universität München,
85747 Garching, Germany

"Rumble" is a self-excited combustion instability, usually occurring at the start-up of aero-engines with fuel-spray atomizers at sub-idle and idle conditions, and exhibiting low limit frequencies in the range of 50 Hz to 150 Hz. Entropy waves at the (nearly) choked combustor outlet are supposed to be the key feedback mechanism for the observed self-excited pressure oscillations. The experimental study presented here aims at clarifying the role of entropy waves in the occurrence of rumble. A generic air-blast atomizer with a design prone to self-excitation has been incorporated into a thermoacoustic combustor test rig with variable outlet conditions. The thermoacoustic response of the flame was characterized by recording the OH chemiluminescence, the dynamic pressures, the dynamic temperatures, and by applying PIV. The measurements have shown the occurrence of periodic hot spots traveling with the mean flow with considerable dispersion. Measurements have been conducted with an open-ended resonance tube in order to eliminate the impact of entropy waves on the mechanism of self-excitation. The oscillation obtained, comparable in amplitude and frequency, proved that self-excitation primarily depends on convective time delays of the droplets in the primary zone and thus on the atomization characteristics of the nozzle.* [DOI: 10.1115/1.2132379]

Introduction

Liquid-fueled aero-engine combustors often are subject to an audible, self-excited combustion instability at idle and sub-idle conditions where the combustor pressures and the air temperatures at the combustor inlet are low. This instability is commonly referred to as "rumble" with typical low frequencies in the limit cycle ranging between 50 and 150 Hz.

Many scientific efforts have been undertaken to investigate the conditions and the mechanisms involved in the occurrence of combustion instabilities such as rumble. Keller [1] underlines the role of flow instabilities and subsequent fluctuations in the reaction rate as primary cause for combustion-driven oscillations. In the context of premixed systems, the work performed by Lieuwen [2] points out the role of fluctuations in the air-fuel mixture ratio and the significance of the convection times in the primary combustion zone for the occurrence and frequency of instabilities.

Pressure feedback on the combustor is obtained when temperature non-uniformities (hot spots, entropy waves) are convected through a choked combustor outlet (Marble and Candel [3]). Polifke et al. [4] have shown that entropy waves can enhance or reduce the stability with respect to self-excited thermo-acoustic oscillations.

Experimental work with fuel-spray atomizers has been conducted by Konrad et al. [5], who eliminated rumble in both single sector and full engine experiments by modifying the fuel atomizer design, particularly the fuel duct geometry, and the swirl of the generated flow field. The sensitivity of the atomizer to changes in the air flow rate for the production of unsteady heat release patterns has also been pointed out by Zhu et al. [6] based on CFD-simulations. He concludes that the limit frequency of rumble primarily depends on the convective time delays of the entropy waves from the flame to the nozzle. However, as shown by Sattelmayer [7], the combustor aerodynamics is responsible for dispersion of the entropy wave and can significantly reduce their impact as potential instability driver. His low-order model takes into account spatial dispersion of equivalence ratio and entropy waves. He comes to the conclusion that the destabilizing effect of entropy waves is bound to the low frequency domain where their dispersion is small.

To what extent entropy waves contribute to the development and sustainment of low-frequency combustion oscillations is thus not clear. In this paper we provide experimental results to clarify the role of entropy and mixing inhomogeneities as key mechanisms of the rumble instability.

Experimental Setup

For the investigations, a generic triple swirl airblast atomizer with a design typical for aero-engine systems has been mounted to a generic combustion chamber. The atomization behavior under forcing of this injector has already been investigated in Ref. [8]. Figure 1 shows a cross section of the combustor assembly with the investigation areas for PIV (small box), OH* chemiluminescence (large box), dynamic pressure ($\times$), and dynamic temperature ($\bigcirc$).

The combustion chamber is operated from near-ambient pressure conditions up to 1.6 bar static pressure (absolute). It has a square cross section and measures 90 mm $\times$ 90 mm $\times$ 280 mm, comparable to the dimensions of a single sector in a real aero-engine combustor. Exchangeable side plates allow a flexible adaptation to the requirements of the various measurement techniques applied. The primary combustion zone is optically accessible through exchangeable quartz windows. Corresponding side plates allow the application of pressure transducers and dynamic temperature probes along the whole combustor length. Ignition is achieved using a spark plug.

The generic air-blast diffusion injector/combustor system investigated is of the Rich-Quench-Lean-type. Jet A-1 is used as fuel.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisr. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria June 13–17, 2004, Paper No. GT2004-54163.

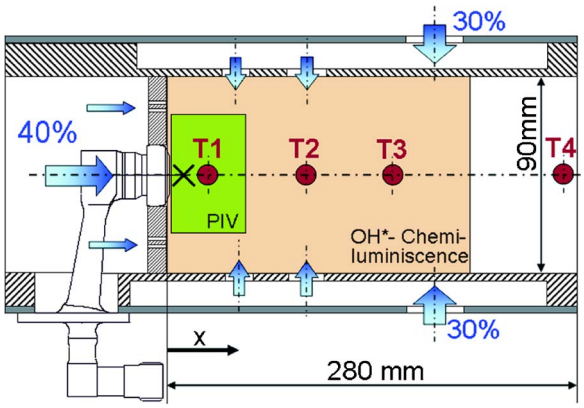


Fig. 1 Schematic representation of the generic combustor with air distribution and the measurement positions

Two rows of holes for primary and secondary air injection in the top and bottom plate of the combustion chamber provide the aerodynamic and stoichiometric operating conditions. The hole pattern and location was derived from the engine (annular) liner geometry. The hole diameters were determined from typical air distributions of such liners on the basis of CFD. The supply of the air injection is acoustically decoupled using perforated plates. Dilution, film- and liner-cooling in the downstream part of the combustion chamber are omitted. About 60% of the overall air mass flow enters the combustion chamber as admix air.

The outlet conditions of the combustion chamber are varied as follows:

1. *Venturi nozzle*: A water-cooled Venturi nozzle is used to impose choked exit conditions to the combustor. The throat measures 20 mm in diameter. For the considered operating point, the application of the Venturi led to an increase of the static combustor pressure of about 0.4 bar.
2. *Resonance tube*: An open-ended resonance tube has been designed to eliminate the impact of temperature nonuniformities on the combustor acoustics. The cylindrical tube measures 101 mm in diameter and thus corresponds to the cross-sectional area of the combustor. In order to comply with the acoustic boundary conditions imposed by the nozzle, the length of the tube has been adapted to obtain an admittance at the combustor outlet which is comparably low.

Measurement Techniques and Data Analysis. *Spatially resolved flame recordings* have been performed using an intensified High Speed camera. To detect the OH^* chemiluminescence at a wavelength of 308 nm, a DUG11-X filter was mounted on the camera, showing bandpass characteristics for wavelengths between 260 nm and 380 nm. Recordings of the self-excited flame were done at a sampling rate of 4500 Hz. Usually, a sequence of 350 camera images was saved for each operating point, covering 6 to 10 oscillation cycles of the flame. The camera was focused on the central plane of the primary zone of the combustion chamber. The viewed area in the images measures 113 mm in diameter.

Time-resolved characterization of the flame thermoacoustics included the synchronous measurement of the flame OH^* chemiluminescence, the dynamic combustor pressures, and the dynamic temperature.

For capturing the *dynamic pressure* in the combustion chamber, water-cooled pressure transducers of type PCB J106B have been used. These acceleration-compensated piezo-type transducers measure the dynamic pressure within a range of 0.7 Pa to 0.57 bar on a static head of up to 138 bar. Due to soot formation in the flame and in the dilution zone of the combustor, the pressure transducer had to be mounted upstream of the injector spray cone, 6 mm downstream of the injector outlet plane.

The OH^* *chemiluminescence* of the flame has been captured using a photomultiplier equipped with a DUG11X-filter. The photomultiplier was positioned at 300 mm radial distance from the combustion chamber with a field of view covering the whole optically accessible part (90 mm × 150 mm) as shown in Fig. 1. In general, it is not clear to what extent the OH^* chemiluminescence represents the heat release rate in the case of spray flames. However, the experience gained with the current experiment suggests that it is applicable here, since it provides data that is consistent with, e.g., the fast thermocouple measurements discussed below.

For the quantification of the *gas temperature* perturbations in the combustor, the approach described in Ref. [9] has been modified to allow the determination of the gas temperature without in situ determination of the cut-off frequency of the thermocouple in the combustor. The measurements were conducted using a water-cooled probe containing two Pt/RhPt-thermocouples with different wire diameters (0.002 in. and 0.003 in.) that are positioned at 1 mm separation on the probe. The stationary gas temperature has been reconstructed according to the approach in Ref. [9]. As shown there, the energy balance for the thermocouple implies a linearized first-order lag (P-T1) response behavior with respect to the temperature perturbation, including exchange of heat by convection and radiation. The radiative heat losses of the thermocouples to the cooler combustor walls are approximated assuming quasi-steadiness and a grey body behavior of the thermocouples. For that purpose, the wall temperature is measured at several locations with sets of thermocouples, being soldered into pocket holes in the combustor side plates. With the wall temperature T_w and the diameter of the thermocouples $d_{t1,2}$ being known, the energy balance applied to both thermocouples thus results in two equations for three unknowns, namely the two local convective heat transfer numbers $\alpha_{t1,2}$ and the gas temperature T_g. Assuming identical local gas velocities at the thermocouples $u_{t1} \equiv u_{t2}$ (which is justified with respect to their close vicinity on the probe), α_{t1}/α_{t2} becomes a function of the diameter ratio d_{t1}/d_{t2} only. Although the absolute values for α_{t1} and α_{t2} are still unknown, T_g can be calculated with good success on the basis of the two measured thermocouple signals by linear regression over their P-T1-compensated Fourier spectra.

Dynamic temperature measurements were performed along the central axis of the combustor at 20 mm, 75 mm, and 130 mm distance from the injector exit and close to the combustor outlet plane at 275 mm distance (see Fig. 1).

The acquisition of the time-resolved data is controlled by a multichannel-I/O-Board (300 kHz, 12 bit). Each channel was sampled with 10 kHz. The time-domain sequences usually contained 10,000 samples each. The relative errors of the photomultiplier and microphone measurements are estimated to be within ±5%. Also, the stationary gas temperature can be determined within this error margin. Due to the assumptions made for reconstructing the dynamic gas temperature (system linearity, constant thermocouple diameters) the maximum relative error of the amplitudes of the temperature perturbations is higher, but does not exceed ±15%. The reconstructed temperature phase however proved to be more stable. Phase errors are bound within ±10 deg.

Particle Image Velocimetry (PIV): 2D-PIV was employed to determine the phase-resolved velocity field inside the kerosene spray cone in the central longitudinal section of the primary zone. The phase-resolved measurements were triggered by the dynamic combustor pressure, using a microphone with a subsequent voltage comparitor switch. To collect detailed data at the combustor inlet, the downstream region of the injector exit was investigated, which measured 42 mm × 62 mm starting 1 mm downstream of the combustor front as shown in Fig. 1. For each phase angle, 180 single shots were averaged, leading to a statistical error below 6%. Phase errors did not exceed ±10 deg. The statistical error of the measured velocities inside the flame zone tends to increase, due to droplet burning and decreasing droplet concentrations. The higher

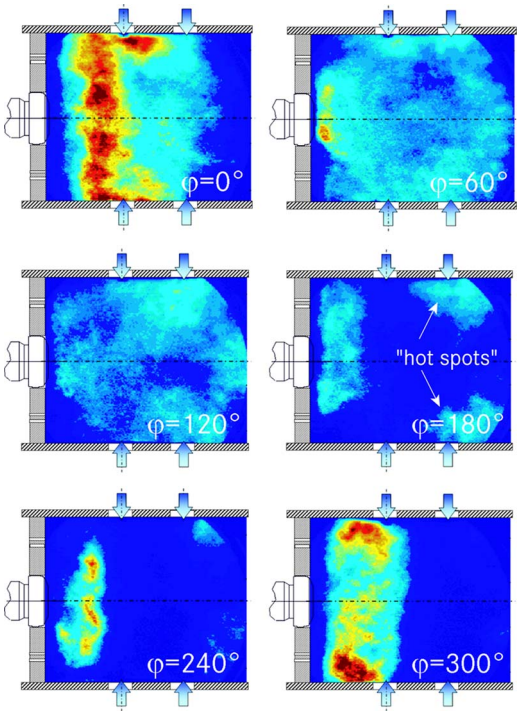


Fig. 2 High speed camera recording of the flame OH* chemiluminescence during one combustion oscillation cycle at 126 Hz

standard deviation was used to filter out misleading flame zone data during image post-processing.

Phenomenology of the Self-Excited Combustion Oscillation

In the following subsections, we describe the phenomenology of the self-excited oscillation as observed in the rig at an operating condition characterized by a relative static pressure drop over the injector of $\Delta p = 2.1\%$, no air preheating ($T_{30} = 300$ K), and an overall equivalence ratio of $\phi = 0.55$.

Venturi Nozzle With Choked Exit. Figure 2 shows one period of the self-excited combustion oscillation at 126 Hz limit frequency obtained with the Venturi nozzle mounted on the combustor exit. The images are taken from the high speed video sequence capturing the OH* chemiluminescence of the flame. The relative phase angle φ refers to the phase with maximum integral OH* radiation of the flame. The corresponding time traces of the combustor pressure together with the flame OH* chemiluminescence recorded by the photomultiplier are plotted in Fig. 3.

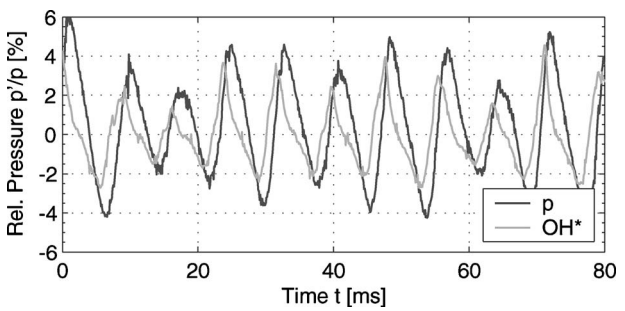


Fig. 3 Time traces of the relative combustor pressure p'/p (grey) and the normalized OH* chemiluminescence (black) obtained with the Venturi nozzle

Both Fig. 2 and Fig. 3 show a strong pulsation of the flame. High OH* intensities are mainly observed in the primary zone near the rows of admix air injection. The pressure rise inside the combustion chamber, immediately following the high reaction rates in Fig. 3, leads to periodic build-up and decay of the primary combustion zone. A Fourier analysis of the results in Fig. 3 shows that the OH* signal leads the pressure signal by 53 deg. The time traces exhibit a very stable phase relationship between the signals and a beat for both the pressure and the chemiluminescent intensity. The correlation of the modulated amplitudes is very high, i.e., oscillation cycles with high pressure amplitudes temporally coincide with highly oscillating chemiluminescence and vice versa.

The pressure amplitudes in the range of 2%–5% of the static combustor pressure partly exceed the moderate relative pressure drop of the nozzle for this operating point. As a consequence, the flow velocity from the injector decreases strongly and occasionally flame flashback towards the nozzle is observed ($\varphi = 60$ deg). In the dilution zone, i.e., downstream of secondary air injection, lower levels of OH* radiation are detected. Nevertheless, in this region with an overall lean mixture, periodic detachment of zones of locally increased reaction rate can be discerned, which originate from the primary reaction zone and which are subsequently convected towards the combustor outlet.

The oscillating combustor pressure affects the fuel spray supply to the combustion zone: Rising combustor pressure reduces the available air momentum for atomization, shifts the spray spectrum towards larger droplets and increases the convective delay times of the spray to the flame. Concerning the droplet size, the investigations in Ref. [8] have established the following quasi-steady relationship between the Sauter mean diameter of the kerosene spray d_{32} and the relative air velocity u_a for the investigated nozzle

$$\frac{d_{32}}{d_{32,0}} = \left(\frac{u_a}{u_{a,0}}\right)^{-1.6} \tag{1}$$

Following the quasi-steady approach for the pressure-velocity coupling at the injector for low frequencies, derived in Ref. [8], a relative rise of 5% of the combustor pressure with respect to the pressure drop thus leads to a 4% increase of d_{32}.

The impact of the pressure on the convective delay times of the spray from the atomizer lip to the flame tip can be seen in Fig. 4, which shows the phase-resolved spray velocities measured with PIV. The phase reference is the same as for Fig. 2. From Fig. 4, we see that the high combustor pressure for $\varphi < 90$ deg leads to a rapid decay of spray injection. The recovery of adequate spray injection velocities takes place for $\varphi > 144$ deg, when the pressure wave in the combustor is negative. The droplet peak velocities in the center plane of the combustor vary between 8 m/s for $\varphi \approx 72$ deg and 32 m/s for $\varphi \approx 288$ deg. Higher delay times for acceleration of the larger droplets increase the spread of convective droplet time delays encountered during one oscillation period. Referring to a mean flame stand-off distance of 35 mm, deduced from Fig. 2, the characteristic delay times for the smallest droplets including droplet acceleration can be estimated to 2 ms, referring to Ref. [8].

Figure 5 shows the recorded time domain data of the combustor pressure (grey) and the gas temperature 5 mm upstream of the venturi nozzle at the combustor outlet, or 40 mm upstream of the critical throat diameter (black). The reconstructed mean temperature is 1545 K and deviates by about 2% from the adiabatic flame temperature for the overall equivalence ratio set ($\phi = 0.55$, $T_{ad} \approx 1580$ K). The oscillations of pressure and temperature are nearly in phase; the temperature lags the pressure by 17 deg. The amplitude of the temperature perturbation varies considerably from cycle to cycle, which can be attributed to the influence of wave dispersion by the combustor aerodynamics [7]. The average amplitude of the temperature perturbation amounts to 31 K.

In order to investigate the impact of the combustor aerodynamics on the entropy wave produced by the flame, dynamic tempera-

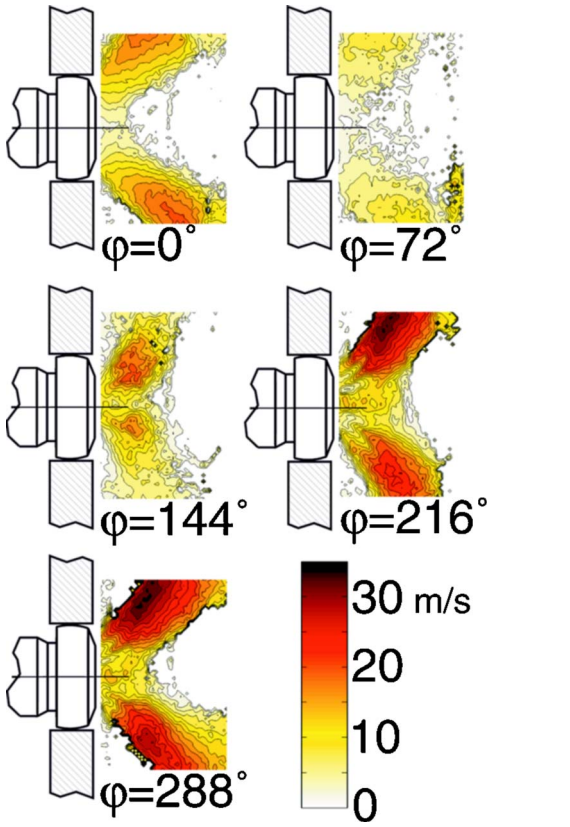


Fig. 4 Phase-resolved, averaged droplet velocities at the injector outlet measured by PIV for one oscillation cycle at 126 Hz. The velocity magnitude refers to the absolute value of the two component velocity vector $\sqrt{u^2+v^2}$ in the center plane of the combustor.

ture measurements have been performed at different axial locations of the combustor according to the positions shown in Fig. 1. In Table 1, the temperature perturbations are characterized by their Fourier amplitudes at the oscillation frequency (126 Hz). Table 1 also shows the phase difference to the synchronously measured OH^* signal at the respective frequency. The y coordinate refers to the distance from the longitudinal combustor axis. T1 and T2 correspond to thermocouple positions directly up- and downstream of the main reaction zone, which is bound to the interval $30\ \text{mm} > x_f > 50\ \text{mm}$, as seen from Fig. 2. T4 is the position 10 mm upstream of the nozzle inlet, where two radial positions have been measured denoted by T4a and T4b.

A first important conclusion can be drawn from Table 1: The signal phases obtained for T1 and T2 in the immediate vicinity of the reaction zone show that the temperature lags the OH^* chemiluminescence by about 105 deg. Taking into account minor phase delays due to convection between flame and thermocouples, a phase relationship is obtained corresponding to harmonic heat release $\dot{q}'$ and the temperature T_g' under isobaric conditions

$$\dot{q}' e^{i\omega t} = \rho c_p \frac{d(T_g' e^{i\omega t})}{dt} \Rightarrow \arg(\dot{q}') - \arg(T_g') = 90\ \text{deg} \qquad (2)$$

For the investigated diffusion flames, we now can conclude that the detected OH^* chemiluminescence is a useful measure for the local heat release oscillation of the flame, $\dot{q}'$, such that the Rayleigh criterion

$$\int_0^T \dot{q}'(t) \cdot p'(t) dt > 0 \qquad (3)$$

establishing the necessary condition for thermoacoustic self-excitation can be evaluated on the basis of, e.g., Fig. 3.

From Table 1 it can be further concluded that the dispersion of the entropy wave in the combustor is important. In the vicinity of the flame, temperature pulsations of 161 K and 102 K are obtained. While T1 feels the periodic flame movement toward the injector, the axial position of T2 is close to the second row of air injection, which explains the lower values obtained for this point. The entropy waves produced by the flame rapidly decay when traveling toward the combustor outlet. At the outlet, approximately 80% of their measured initial strength is dispersed. From the phase shift of the temperature oscillation obtained between positions T2 and T4, a convective time delay of $\tau_s = 8.3$ ms from the flame to the nozzle is obtained for the entropy wave. The resulting convective velocity comes close to the calculated bulk velocity of the gas in the combustor, which confirms the convective character of the waves.

The chemiluminescence recordings in Fig. 2 show an inhomogeneous heat release over the height of the combustion chamber with increasing intensity near the combustor walls. Also, their respective convection speed is different, such that a phase dispersion of the entropy waves should be expected at the Venturi nozzle. In order to assess the effective temperature inhomogeneity at the nozzle inlet, a measurement of the dynamic temperature has been conducted 15 mm off the centerline of the combustor, corresponding to about half the distance toward the upper rim of the convergent nozzle inlet (Positions T4b). Under the influence of the hot spots near the combustor walls, the temperature amplitude is increased by 20% with respect to the centerline value. The relative phase difference of 34 deg of the temperatures obtained between the radial positions T4a and T4b further indicates the highly dispersive character of the entropy wave convection in the combustor.

Referring to Ref. [3], the amplitude of the reflected pressure wave generated by an entropy disturbance at a choked compact nozzle can be calculated as

$$\frac{p'}{p} = -\frac{\gamma M}{2 + (\gamma - 1) M} \frac{T'}{T}. \qquad (4)$$

For the temperature fluctuations encountered at the nozzle inlet, a relative amplitude of 1.36% is obtained for the reflected pressure wave. This is sufficiently high to sustain self-excitation, but too

Table 1 Propagation of the temperature waves in the combustor

Pos.	x [mm]	y [mm]	T_g' [K]	$\Delta\varphi_{OH^*,T}$ [deg]
T1	20	0	161	104
T2	75	0	102	109
T3	120	0	47	−120
T4a	275	0	31	123
T4b	275	15	38	89

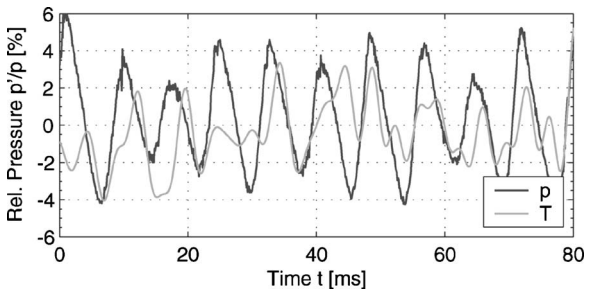


Fig. 5 Time traces of the relative combustor pressure p'/p (grey) and the gas temperature near the nozzle inlet (Position T4 in Fig. 1, black)

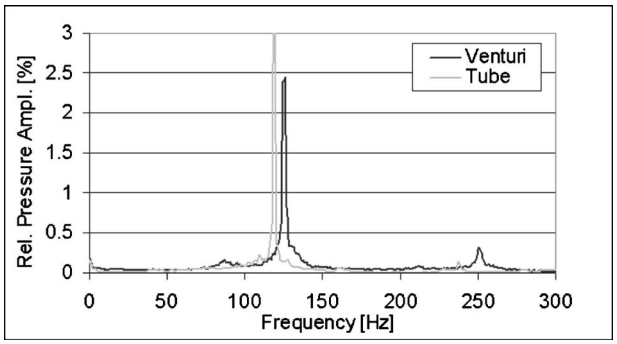


Fig. 6 Fourier spectra of the relative pressure amplitudes obtained for the operation with Venturi nozzle (black) and resonance tube (grey)

low for the level of pressure oscillation measured. Furthermore, the pressure and temperature perturbations measured near the critical outlet in Fig. 5 are nearly in phase, which disagrees with the theory predicting a phase shift of 180 deg. The role of the entropy waves on the self-excited combustion oscillation thus requires more clarification.

Resonance Tube With Open Exit. In order to eliminate the influence of entropy waves on the combustor acoustics, the Venturi at the combustor outlet was replaced by an open-ended resonance tube with a length designed to impose a comparably low outlet admittance to the combustor. Figure 6 shows the Fourier amplitudes of the obtained pressure oscillation.

The relative pressure amplitude with resonance tube (grey) rises to about 3% of the static combustor pressure, being about 0.6 percentage points higher than the oscillation obtained with the nozzle (black). The peak frequency is 118 Hz, which is 8 Hz below the frequency encountered in the choked case. The OH^* high speed recording of the pulsating flame delivers very similar images to those shown in Fig. 2.

Plotting the time traces of pressure and OH^* chemiluminescence in Fig. 7, an oscillation behavior almost identical to that obtained with the nozzle (Fig. 3) can be seen. The pressure oscillation varies less in amplitude from cycle to cycle. The heat release of the flame leads the combustor pressure by 62 deg. The Rayleigh criterion, Eq. (3), is fulfilled. High speed images of the flame (not shown here) show the same kind of flame oscillation illustrated in Fig. 2.

Interpretation of the Results

A pressure feedback by reflection of entropy waves at the nozzle can be excluded for the resonance tube. The similarity of the results thus proves that entropy waves do not play a significant role for the occurrence of the encountered low-frequency combustion instability.

Since the phase relation between pressure and heat release in both test configurations is almost the same and both pressure drop

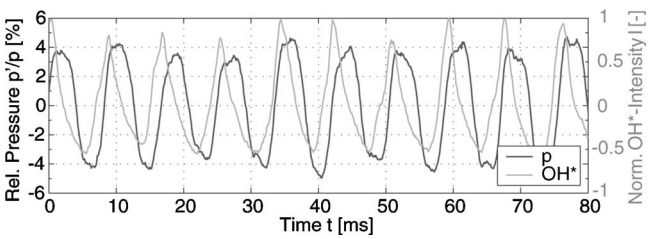


Fig. 7 Time traces of the relative combustor pressure p'/p (grey) and the normalized OH^* chemiluminescence (black) obtained with the resonance tube

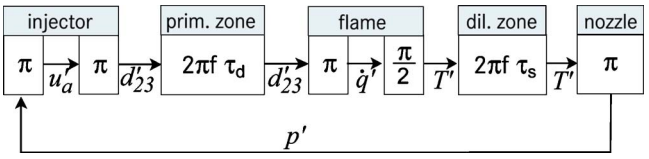


Fig. 8 Schematic representation of the feedback cycle for entropy waves with the phase relations of the quantities involved

and equivalence ratio have been maintained, we can conclude that the injector and especially the convective delay times in the primary zone are the main source for self-excitation. With the admittance of the fuel side being negligible, the fuel mass flow through the injector can be assumed constant. Consequently, as already discussed in the context of Eq. (1), the sensitivity of the atomizer to changing air momentum affects the heat release by producing a higher number of smaller droplets with correspondingly faster burn-off times. The global heat release rate will thus rise with the finer spray obtained for low pressure conditions in the combustion chamber. The phase angle between minimum pressure and maximum heat release is then a measure for the convective delay times of the smaller droplets. Referring to the measurements in Figs. 3 and 7, these delay times can be quantified as 2.80 ms (nozzle) and 2.78 ms (resonance tube). These values fit very well to the time delay estimations for smaller droplets on the basis of the PIV velocities in Fig. 4, which confirms the findings above.

For the test case with nozzle, the phase shift between temperature and pressure at the nozzle inlet is small (Fig. 5). Under those circumstances, the entropy wave destructively interferes with the self-excited combustion oscillation. This explains the beat obtained in the pressure trace of Fig. 3, which leads to a relative variation of the amplitudes by about 1.5–2 percentage points. In the Fourier spectrum for the nozzle shown in Fig. 6, a minor pressure peak at 85 Hz is found, which does not appear for the resonance tube and which is likely to be related to the influence of entropy waves.

Figure 8 shows a schematic representation of the feedback cycle for the entropy waves and the phase relationship between the quantities that influence the combustion process. Figure 8 includes the finding of this paper concerning the transfer behavior of the diffusion flame, namely the 90 deg phase relation between heat release and temperature and the fact that the heat release rate was found to increase with smaller droplet diameters. With the time delay of pressure wave propagation in the combustor being negligible and for measured τ_d=2.8 ms and τ_s=8.3 ms, the matching condition

$$\sum_i \Delta\varphi_i = n \cdot 2\pi \quad (n = 1,2,3,\ldots) \tag{5}$$

with n being a natural number is fulfilled for f_s=68 Hz. With respect to the implied model simplifications, this is a good approximation of the peak frequency at 85 Hz and shows that the oscillation of the entropy can indeed be expected in the respective frequency range of the second peak in Fig. 6. The strength of the entropy peak in the pressure spectrum is however smaller than predicted by the theory. Possible explanations are the considerable phase dispersion of the entropy waves, showing that the conditions inside the combustor are not strictly 1D, and the fact that the nozzle is water-cooled, thus not isentropic.

As shown in Fig. 9, the thermoacoustic mode at 126 Hz and its second harmonics at 252 Hz are dominating the pressure spectrum. The separate pressure peak at 85 Hz indicates that the entropy-related pressure perturbations coexist with the thermoacoustic oscillations in the combustion chamber.

The side bands found in the amplitude spectrum (Fig. 9) are a result of the non-linearity of the limit cycle. The peak frequencies of the side bands at 211 Hz and 41 Hz correspond to the combination frequencies of the entropy and the thermoacoustic oscilla-

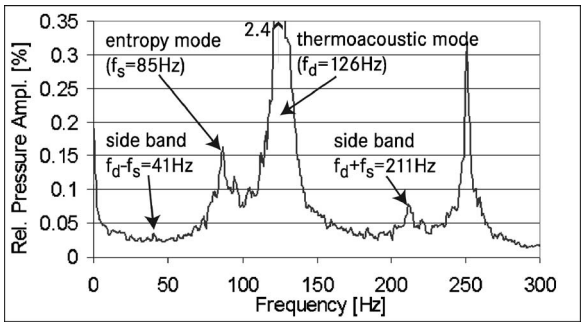


Fig. 9 Side bands of the relative pressure amplitude for the limit cycle of the self-excited oscillation with Venturi. The graph is an enlargement of Fig. 6.

tion. The weaker entropy wave, being coupled with the thermoacoustic wave by the instationary heat release of the flame, thus leads to an amplitude modulation of the dominating thermoacoustic mode which was encountered in Fig. 3.

Influence of the Operating Conditions

In order to prove if the above findings apply to a wider range of operation conditions, the limit cycle pressure was recorded for varying air mass flows and equivalence ratios. The results in Fig. 10 show that the limit frequencies are essentially determined by the absolute pressure drop of the injector Δp. Δp is related to the specific momentum flux of the atomizing air stream by the quasi-steady Bernoulli equation applied to the injector

$$\Delta p = \frac{\rho}{2} u_a^2 (1 + \zeta_{\text{inj}}) \tag{6}$$

u_a is the injection velocity. The upstream velocity is neglected in Eq. (6) due to the high area ratio between injector and supply section (1:90). Equation (6) has been proven applicable for this burner in the low-frequency regime investigated [8]. The momentum flux $\rho_a u_a^2$ is part of the Weber number, influencing the maximum stable droplet size d_{32} of the spray. Δp is thus linked with d_{32} which in turn determines the time scales for fuel transport and evaporation [8], which are the rate-controlling steps under low ambient temperatures and pressures [10,11]. The dominating influence of d_{32} and Δp on the limit cycle frequency is another indicator for the significance of the spray-related processes in the primary zone on the system acoustics. Since the wavelength of the rumble mode by far exceeds the size of the (unextended) combustor, the oscillation takes place in the combustor in a bulk mode. The time delay in the primary zone thus will influence both the instability frequency and the achievable amplification rates. The influence of the equivalence ratio is, however, low. This stays in context with the observation of soot deposition at the water-cooled exit nozzle, which indicates local quenching and partially incomplete spray combustion, becoming particularly important under conditions of low injector pressure loss.

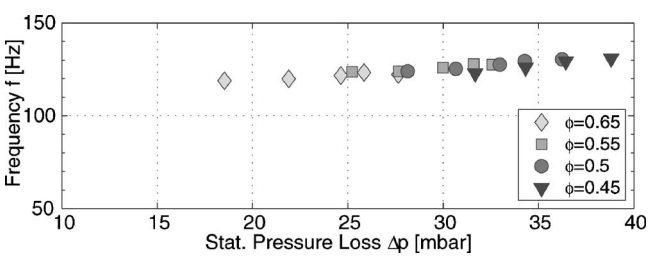


Fig. 10 Limit cycle properties of the combustion oscillations with Venturi. Limit cycle frequencies over the injector pressure loss for variable equivalence ratios.

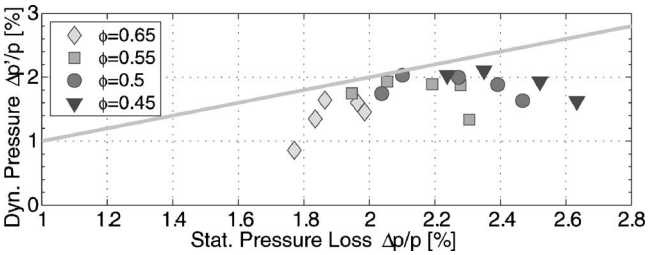


Fig. 11 Normalized pressure amplitude $\Delta p'/p$ of the self-excited oscillation over the relative static pressure drop $\Delta p/p$ across the injector

Figure 11 shows the dynamic pressure amplitudes in the limit cycle as a function of the applied static pressure loss. $\Delta p'/p$ is the effective dynamic pressure loss amplitude over the burner, which is modulated by the combustor pressure oscillations. The grey line delimits the region where dynamic pressure oscillations would exceed the static pressure loss over the injector, leading to flow reversal through the burner. The oscillation amplitudes lie below this line. The static pressure loss characteristics of the injector thus provide an upper limit for the attainable oscillation amplitudes. The measurement data indicate that the breakdown of the air injection and the deceleration of fuel supply to the flame under conditions of rising combustor pressures controls the rumble amplitudes attainable in the limit cycle.

Findings and Conclusions

The instability mechanism of a generic airblast atomizer with a design typical for aero-engine systems has been investigated under self-excitation. The experiments presented here yield the following results:

- The self-excited flame oscillation obtained is accompanied by the formation of entropy waves. These hot spots are convected with bulk velocity toward the combustor outlet. The dispersion of the entropy waves from the flame to the nozzle is high.
- Self-excited flame oscillations have also been obtained under non-choked exit conditions with a similar outlet impedance of the combustor. Entropy waves are thus not essential for the occurrence and the sustainment of the encountered low-frequency combustion oscillation.
- Unsteady heat release, which proved to be adequately represented by the OH^* chemiluminescence of the flame, primarily stems from the atomization characteristics of the airblast atomizer, its sensitivity to pressure perturbations, and thus of perturbations of the flow field in the primary combustor zone.
- The coupling between pressure and heat release rate with respect to droplet velocity measurements in the primary zone has shown that the oscillation frequency obtained is dominated by the characteristic convection time scales of the fuel from the atomizer lip to the flame, being biased toward small droplets.
- The entropy waves produced by the flame lead to the reflection of separate pressure waves at the choked nozzle. The strength of these entropy-related pressure waves is however too weak to significantly alter the dominating thermoacoustic mode of the primary zone, but results in an amplitude modulation of the combustor pressure in the limit cycle.
- The amplitude of the pressure oscillations obtained is limited by the static pressure drop over the injector.

Acknowledgment

The experimental work presented in this paper was supported by the European Commission as part of the GROWTH programme "Instability Control of Low Emissions Aero Engine Combustors" (ICLEAC) under Contract No. G4RD-CT-2000-00215. We further acknowledge Dr. Ralf von der Bank and Dr. Thomas Schilling of Rolls-Royce Deutschland Ltd & Co KG for their practical support concerning the investigated hardware.

Nomenclature

c_p = isobaric specific heat (W/kg K)
d_{32} = Sauter mean diameter (m)
f = frequency (Hz)
M = Mach number
p = pressure (Pa)
$\dot{q}$ = heat release (W/s)
T = temperature (K)
t = time (s)
u = axial velocity (m/s)
x = axial coordinate (m)
y = radial coordinate (m)
γ = ratio of specific heats
ζ_{inj} = pressure loss factor of injector
φ = phase angle (deg)
ϕ = equivalence ratio
τ = time delay (s)
ω = angular frequency (s^{-1})

Subscripts

$()_a$ = air
$()_d$ = droplet
$()_g$ = gas
$()_s$ = entropy

Superscript

$()'$ = perturbation

References

[1] Keller, J. J., 1975, "Thermoacoustic Oscillations in Combustion Chambers of Gas Turbines," AIAA J., **33**, pp. 2280–2287.
[2] Lieuwen, T., Torres, H., Johnson, C., and Zinn, B. T., 1999, "A Mechanism of Combustion Instability in Lean Premixed Gas Turbine Combustors," ASME Paper No. 99-G-3.
[3] Marble, F. E., and Candel, S. M., 1977, "Acoustic Disturbance from Gas Non-Uniformities Convected Through a Nozzle," J. Sound Vib., **55**, pp. 225–243.
[4] Polifke, W., Paschereit, C. O., and Döbbeling, K., 2001, "Constructive and Destructive Interference of Acoustic and Entropy Waves in a Premixed Combustor With a Choked Exit," Int. J. Acoust. Vib., **6**, pp. 135–146.
[5] Konrad, W., Brehm, N., Kameier, F., Freeman, C., and Day, I. J., 1998, "Combustion Instability Investigations on the BR710 Jet Engine," J. Eng. Gas Turbines Power, **120**, pp. 34–40.
[6] Zhu, M., Dowling, A. P., and Bray, K. N. C., 2000, "Self-Excited Oscillations in Combustors With Spray Atomizers," ASME Paper No. 2000-GT-0108.
[7] Sattelmayer, T., 2003, "Influence of the Combustor Aerodynamics on Combustion Instabilities From Equivalence Ratio Fluctuations," J. Eng. Gas Turbines Power, **125**, pp. 11–19.
[8] Eckstein, J., Freitag, E., Hirsch, C., Sattelmayer, T., von der Bank, R., and Schilling, T., 2005, "Forced Low-Frequency Spray Characteristics of a Generic Airblast Swirl Diffusion Burner," J. Eng. Gas Turbines Power, **127**, pp. 301–306.
[9] Heilos, A., 2002, "Spektrale Analyse der thermischen Strahlungswechselwirkung in Kohlenwasserstoffflammen," *Berichte aus der Energietechnik, Shaker*, 2002.
[10] Turns, S., 2000, "An Introduction to Combustion: Concepts and Applications," *McGraw-Hill Series in Mechanical Engineering*, 2nd ed., McGraw-Hill, New York.
[11] Sujith, R. I., Waldherr, G. A., Jagoda, J. I., and Zinn, B. T., 2000, "Experimental Investigation of the Evaporation of Droplets in Axial Acoustic Fields," J. Propul. Power, **16**, pp. 278–285.

Integration of Gas Turbines Adapted for Syngas Fuel With Cryogenic and Membrane-Based Air Separation Units: Issues to Consider for System Studies

John G. Wimer
Dale Keairns
Edward L. Parsons
John A. Ruether

National Energy Technology Laboratory,
3610 Collins Ferry Road,
Morgantown, WV 26507-0880

The purpose of this paper is to aid systems analysts in the design, modeling, and assessment of advanced, gasification-based power generation systems featuring air separation units (ASUs) integrated with gas turbines adapted for syngas fuel. First, the fundamental issues associated with operating a gas turbine on syngas will be reviewed, along with the motivations for extracting air from the turbine-compressor and/or injecting nitrogen into the turbine expander. Configurations for nitrogen-only and air-nitrogen ASU integration will be described, including the benefits and drawbacks of each. Cryogenic ASU technology will be summarized for both low-pressure and elevated-pressure applications and key design and integration issues will be identified and discussed. Finally, membrane-based ASU technology will be described and contrasted with cryogenic technology in regard to system design and integration. [DOI: 10.1115/1.2056535]

1 Introduction

Many studies [1–3] have shown that coal gasification combined cycle (CGCC) power generation systems featuring oxygen-blown gasifiers can achieve higher efficiency and reduced capital costs by integrating, to various degrees, the gas turbine with the ASU. Although it is not currently clear whether the benefits of full integration outweigh its potentially negative impacts on system availability, control and operability, especially during transient (e.g., start-up, shutdown) conditions, nearly all the concepts being considered for the next generation of ultra-efficient power plants rely on some level of integration.

Motivating the trend towards ASU integration is the opportunity for future gasification-based power plants to employ an emerging, membrane-based air separation technology that promises significant capital cost reductions[1] but must be integrated with the gas turbine to achieve acceptable cycle efficiency [6].

In recent years, the ability to operate conventional, cryogenic ASUs at elevated pressures has led to several innovations that enhance their integration with gas turbines. Consequently, elevated-pressure, cryogenic technology is currently considered to be the state-of-the-art baseline against which membrane-based technology should be compared.

Whether the ASU is cryogenic or membrane-based, integrating it within a CGCC system can help address the challenges associated with operating gas turbines originally designed for natural gas on coal-derived syngas instead.

At present, the commercial demand for gas turbines designed for combined-cycle power generation is dominated by natural gas applications. Consequently, commercially available gas turbines have been designed to attain optimal performance when fueled with natural gas, rather than other fuels such as coal-derived syngas. Due to the great expense of developing gas turbines, it is unlikely that the gas turbine industry will develop a gas turbine specifically for syngas until that market undergoes significant growth.

In the absence of gas turbines designed specifically for syngas, the only option currently available for coal applications is to adapt a natural gas-based turbine for use with syngas fuel. As will be shown, an optimal adaptation requires modifications to the turbine equipment as well as certain process changes within the overall CGCC system. This paper focuses on the latter, specifically the integration between the ASU and the syngas-adapted gas turbine.

2 Adapting a Gas Turbine for Syngas

As shown in Fig. 1, a gas turbine consists of three main elements: (1) an axial-flow air compressor, (2) a combustor that combines the compressed air with fuel, and (3) an expander that extracts work from the hot, pressurized products of combustion.

Some, but typically not all, of the pressurized air discharged from the turbine-compressor is supplied directly to the turbine-combustor. (The amount of air compressed by the turbine-compressor is significantly more than that required for stoichiometric combustion.) In conventional, air-cooled turbines, some of the compressed air is diverted for cooling various turbine components. In many applications, a portion of the discharged air is extracted for use elsewhere in the process. For example, in CGCC systems, air may be extracted and supplied to the air separation unit.

Since gas turbines are primarily designed to be fueled with natural gas, supplying them with syngas (i.e., a fuel gas synthesized via coal gasification) presents certain challenges that must be addressed. As shown by Table 1, there are substantial differences between the chemical compositions and heating values of natural gas and syngas that is produced in oxygen-blown CGCC systems. Pipeline natural gas is predominantly methane, which typically comprises over 93% of its volume. Its volumetric lower

[1] A study by Air Products and Chemicals [4] concluded that compared to a highly integrated, elevated-pressure cryogenic ASU, their membrane-based ASU would "decrease the installed specific cost ($/mTPD) of air separation equipment by 35%. Such savings decrease the installed capital cost of the overall (CGCC) facility by 7%… while the efficiency also improves by one percentage point." A similar comparison by Praxair [5] reached very similar conclusions.

Contributed by the Combustion and Fuels Division of the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received May 7, 2004; final manuscript received January 13, 2005. Assoc. Editor: P. C. Malte.

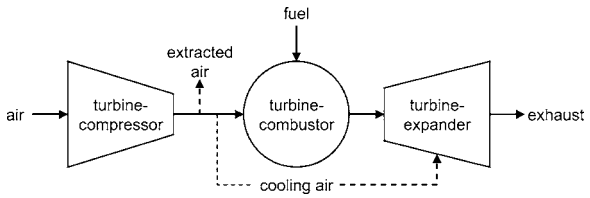


Fig. 1 Elements of a gas turbine

heating value is around 35 MJ/m^3. Syngas is primarily composed of carbon monoxide and hydrogen, but also contains a significant fraction (up to 50%) of noncombustibles, such as steam, carbon dioxide, and nitrogen. Accordingly, its heating value is much lower. The syngas streams produced at the Wabash River and Tampa Electric CGCC plants had volumetric lower heating values less than 8 MJ/m^3, or about one-fourth that of natural gas.

The differences in composition and heating value between natural gas and syngas present two main challenges for adapting a gas turbine for syngas fuel:

1. accommodating different combustion characteristics, and
2. accommodating higher volumetric- and mass-flow rates of fuel gases through the gas turbine combustor and expander.

Both of these challenges can be partially addressed by integrating the gas turbine with the air separation unit.

2.1 Accommodating Syngas Combustion Characteristics. Gas turbines intended for combined cycle applications are designed with high firing temperatures since cycle efficiency increases with the turbine firing temperature [9].

High firing temperatures, however, exacerbate the challenge of preventing excessive, localized flame temperatures in the turbine-combustors. Such "hot spots" are to be avoided for two reasons: (1) they can lead to elevated metal temperatures that reduce component life and (2) they can increase the "thermal" formation of nitrogen oxides (NO_x).

Consequently, recent designs for natural gas-fueled gas turbines have abandoned single-stage diffusion combustors in favor of lean premix[2] combustion technology in an effort to avoid hot spots in the combustor [10]. When fueled with natural gas, lean-premix gas turbines limit emissions of NO_x to 9–25 ppm (at 15% O_2) [11].

2.1.1 Use of Single-Stage Diffusion Combustors. Unfortunately, hydrogen-rich fuels such as syngas cannot be used in current, natural-gas-based designs of lean-premix combustors. Compared to methane, hydrogen's auto-ignition delay is shorter and its flame speed is faster—both by an order of magnitude [10]. Taken together, these factors produce an unacceptable risk of the combustion flame propagating upstream, or "flashing back," into the lean-premix zone.[3]

As a result, gas turbines have been adapted for syngas fuel by reverting to the single-stage diffusion combustors featured in older gas turbine designs—meaning a different approach must be employed to avoid hot spots and control NO_x formation. Avoiding hot spots is made more difficult by the higher adiabatic flame temperatures of hydrogen (2047°C) and carbon monoxide (2100°C), versus 1875°C for methane [1].

This challenge has typically been met by lowering the turbine firing temperature and/or introducing diluents, such as nitrogen from the ASU, steam raised with low-level heat in the heat recovery steam generator, or water.

2.1.2 Introduction of Diluents. Testing programs [8] at General Electric (GE) have shown that diluting the syngas to a heating value between ~4 and 5 MJ/m^3 (~100 and 125 Btu/SCF) can reduce NO_x emissions to ~10 ppmvd[4] at 15% excess O_2—even while maintaining a 1400°C (2550°F) combustor exit temperature. Dilution has been commercially demonstrated to limit NO_x emissions to approximately 15 ppmdv (at 15% excess O_2).

If not introduced properly, diluents can degrade flame stability, potentially leading to flame blowout [1] and/or higher carbon monoxide emissions [2]. To maintain flame stability, the designer has the choice to premix the diluents with the syngas and/or inject them directly into the gas turbine combustor. GE prefers to inject diluents (whether steam and/or nitrogen) directly into the combustor for reasons associated with controllability, efficiency, and cost [8]. The optimal arrangement depends largely on the diluent in question and the composition of the syngas.

As shown in Table 2, three of the four currently operating CGCC plants use both steam and nitrogen for dilution [11].

Table 1 Comparison of natural gas and syngas produced in oxygen-blown CGCC systems

	Volumetric %		
		Syngas[b]	
Component	Natural gas[a]	Wabash River CGCC	Tampa Electric CGCC
H_2	···	24.8	27.0
CO	···	39.5	35.6
CH_4	93.1	1.5	0.1
CO_2	1.0	9.3	12.6
N_2+Ar	1.6	2.3	6.8
H_2O	···	22.7	18.7
Higher hydrocarbons	4.3	···	···
LHV;[c] MJ/m^3 (Btu/ft^3)	35 (930)	7.8 (210)	7.1 (190)
LHV, MJ/kg (Btu/lb)	48 (21,000)	9.0 (3900)	8.2 (3500)

[a]This composition is based on the mean of over 6800 samples of pipeline quality natural gas taken in 26 major metropolitan areas of the United States [7].
[b]Syngas compositions were taken from a report by GE [8].
[c]Lower Heating Values (LHV) were calculated by the author based on gas industry standard conditions: 1 atmosphere, 60°F.

2.2 Accommodating Higher Volumetric and Mass Flow Rates

2.2.1 Relationship Between Mass Flow and Output. The power output of a gas turbine is directly related to the mass flow through its expander. Due to a turbine's mechanical limitations, such as its torque limit or surge margin, there is a maximum mass flow—and power output—that the turbine can safely tolerate.

In natural gas applications, air constitutes the vast majority of the mass flow, while the natural gas fuel contributes very little. The mass flow of air that passes through the gas turbine depends on the ambient air density and is controlled volumetrically by opening and closing its inlet guide vanes.

2.2.2 Increased Mass Flows in Syngas-Fueled Turbines. The amount of fuel energy required by a particular gas turbine is based upon its simple-cycle, design-point performance when fueled with natural gas. This fuel energy requirement can be expressed as a constant ratio to the mass flow of air supplied to the turbine com-

[2]Also known as dry low-NO_x combustion (General Electric), dry-low emissions process (Rolls-Royce/Allison) and SoLoNO$_x$ process (Caterpillar/Solar Turbines).
[3]However, DOE is currently funding the development of new combustors that will enable lean-premix combustion technology to eventually be used in future syngas applications to control NO_x emissions to 3 ppmvd or less (at 15% excess O_2).
[4]Parts per million (dry volume basis).

Table 2 ASU and turbine parameters of currently operating CGCC power plants (sources: Refs. [11–14])

	Wabash River (Indiana, US)	Tampa Electric (Florida, U.S.)	Nuon Power (Buggenum, The Netherlands)	ELCOGAS (Puertollano, Spain)
Start-up year on coal	'95	'96	'94/'95	'97/'98
ASU type[a]	LP	EP	EP	EP
Absolute pressure, bar	~5	13.8	⩾10	⩾10
Oxygen purity, % volume	95%	96%	95%	85%
Gas turbine model	GE 7FA	GE 107FA	Siemens V 94.2	Siemens V 94.3
Pressure ratio	~15	~15	11.4	16.1
Nitrogen integration				
% of ASU nitrogen supplied to turbine[b]	0	~100	~100	~100
Air integration				
-% of ASU air feed supplied with extracted air	0	0	100	100
-% of compressor discharge air extracted for ASU	0	0	16	18
Diluents	H_2O	H_2O & N_2	H_2O & N_2	H_2O & N_2

[a]Low pressure (LP) or elevated pressure (EP).
[b]Typically somewhat less than 100% of the nitrogen is supplied to the turbine as a small portion is used in other plant applications, e.g., soot blowing, coal carrier gas, etc.

bustor [1]. Therefore, for a given mass flow of air supplied to the turbine combustor, a gas turbine would require the same amount of fuel *energy* regardless of whether the fuel is natural gas or syngas. However, since the volumetric and gravimetric heating values of syngas are much lower than those of natural gas (see Table 1), a much greater volume (typically four- to fivefold) and mass (typically five- to sixfold) of syngas fuel is necessary to supply the required energy to the turbine.

The introduction of certain diluents, such as steam and "nonextracted"[5] nitrogen, further inflates the net increase in flow. GE reports that the combined flow of fuel and diluents in a gas turbine adapted for syngas could be up to eight times the fuel flow of a natural gas machine [15].

Gas turbines fueled with natural gas or oil are typically designed for an expander (fuel+air) to compressor (just air) mass flow ratio of 1.02–1.03 [1] and achieve maximum efficiency when operated within this range. In contrast, the mass flow ratios of gas turbines adapted for syngas are typically well above this optimal range. For example, GE [8] reported that 1.16 is a typical mass flow ratio for gas turbines adapted for syngas fuel.

Various process solutions (see Sec. 2.2.3) can be employed to reduce the mass flow ratio of a syngas-fired gas turbine such that it is closer to the optimal, design-point value (~1.02). Although this would increase the turbine's efficiency (in terms of total enthalpy input), it would also decrease the net power output. Since coal is a relatively low-cost fuel, CGCC economics usually favor trading off lower turbine efficiency for a higher net output.

Of course, there is a limit to the extent that a gas turbine's mass flow ratio can be increased when adapting it for syngas. Mass flow is typically limited by one of the following two factors:

1. *Fault torque limit*: Unless measures are taken to limit and/or accommodate the increase in mass flow, the higher power output may exceed fault torque limits [9].
2. *Compressor surge margin*: In terms of fluidic impedance, the turbine-expander is a choked, fixed-area restriction. Consequently, apart from physical modifications, increasing the expander mass flow requires a higher pressure and/or lower temperature at the inlet[6] of the expander's first stage nozzle. Increasing the inlet pressure creates a backpressure at the turbine-compressor discharge and reduces the compressor's surge[7] margin [16], increasing the risk of unstable turbine operation. Decreasing the inlet temperature reduces the efficiency of the combined cycle.

2.2.3 Process Techniques for Accommodating Increased Flow Rates. To stay below fault torque limits and/or preserve an adequate surge margin for a gas turbine adapted for syngas-fuel, various process solutions can be employed to offset the increased mass flow through the turbine-expander:

- To the extent practical, nitrogen dilution can be replaced with syngas humidification. To achieve a given reduction in NO_x, the required mass flow of water is half the required mass flow of nitrogen [16]. (However, see Sec. 3.1 for the drawbacks of syngas humidification.)
- The turbine-compressor's inlet guide vanes can be partially closed to reduce the mass flow of air admitted into the gas turbine. However, throttling the inlet guide vanes may reduce the compressor surge margin [17]. Furthermore, if the inlet guide vanes are partially closed at full-load operation, it reduces the extent to which further throttling can be used to reduce turbine output for part load operation [17].
- Similarly, the expander's mass flow can be decreased by extracting air from the gas turbine compressor discharge and diverting it to the ASU.

Extracting air from the gas turbine compressor actually effects dual reductions in the expander's mass flow: as less air is supplied to the gas turbine combustor, less fuel is needed to satisfy its design-point ratio of fuel energy to air mass.

An integration study by SWPC [1] estimated that extracting ~6% of the air discharged from the turbine-compressor would reduce the mass flow through the expander of a nondiluted syngas-fired turbine such that it is equivalent to the mass-flow through a natural gas-fired turbine expander. Although returning diluent nitrogen from the ASU to the turbine would partially offset the reduction in mass flow achieved via extraction, the SWPC study concluded that—for systems featuring cryogenic ASUs—there is always some combination of air extraction and nitrogen injection that will achieve the design-point mass flow ratio.

Of course, all of the above measures to reduce mass flow reduce volumetric flow as well. Mitigating the increase in volumetric flow can potentially reduce the extent of equipment modifications that must be made to a gas turbine adapted for syngas fuel. Volumetric flows can also be reduced by operating the gas turbine

[5]See Sec. 3. Injected nitrogen contributes to a net increase in the turbine-expander flow rate only if it was not separated from air extracted from the turbine-compressor. (Nitrogen in extracted air is simply diverted from the turbine, separated from the air, and returned.)

[6]When the turbine flow is choked (sonic) at the first-stage nozzle, it cannot be increased by adjusting its downstream pressure.

[7]An axial compressor surges when its flow decreases enough to cause a momentary flow reversal. The flow reversal occurs when the compressor can no longer overcome its discharge pressure.

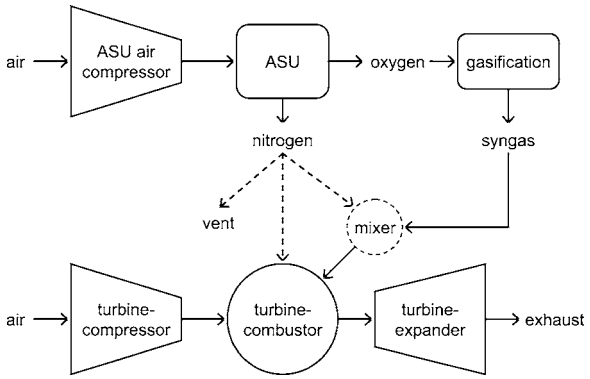


Fig. 2 Nitrogen-only integration (simplified view)

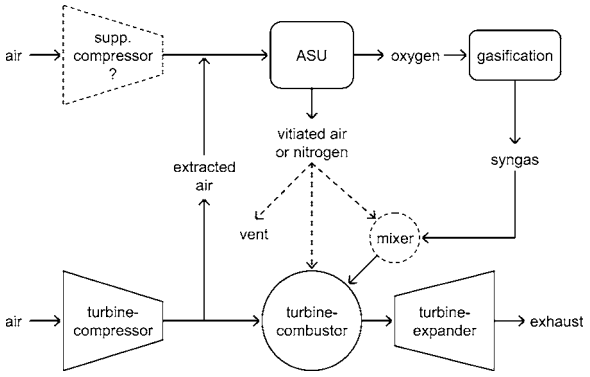


Fig. 3 Air-nitrogen integration (simplified view)

at a higher pressure and/or with a lower firing temperature.

2.2.4 Turbine Modifications to Accommodate Increased Flow Rates. To accommodate higher volumetric flows while preserving an adequate surge margin, it may be necessary to slightly[8] increase the effective area of the expander's stage-one nozzle or, in some cases, add another stage to the compressor.

To accommodate the unavoidable increase in volumetric fuel flow, gas turbines adapted for syngas fuel must be designed with a second[9] fuel handling and injection system that has larger components, e.g., gas manifolds, valves, and fuel injectors.

3 Integration Configurations

3.1 Nitrogen-Only Integration. Byproduct nitrogen from a cryogenic ASU can be used in several areas of the power generation system, most notably in the gas turbine.

As shown in Fig. 2, "nitrogen-only" integration consists of supplying nitrogen from the ASU to the gas turbine-combustor by premixing it with the syngas stream or by injecting it directly into the turbine-combustor.

The CGCC power plant in Tampa, FL features nitrogen-only integration. See Table 2.

Depending on the gas turbine design, nitrogen integration can yield several benefits.

As discussed in Sec. 2.1.2, using nitrogen as a diluent can increase the life of turbine components and reduce the formation of thermal NO_x by eliminating hot spots in the turbine combustor.

Furthermore, nitrogen dilution may be especially beneficial for gas turbines that would otherwise be derated due to low ambient air density, i.e., at high ambient air temperatures. At such conditions, more syngas can be burned to increase the power output if nitrogen dilution is used to maintain the optimal turbine firing temperature [2].

Of course, if the ASU is designed to produce residual nitrogen at high pressure, nitrogen integration would likely be required to achieve reasonable system efficiency. This is certainly the case for membrane-based ASUs, since nitrogen emerges in the nonpermeate stream at both high pressure and temperature.

As noted in Sec. 2.2.2, one drawback of nitrogen injection is that it can potentially exacerbate the various challenges (e.g., equipment modifications, surge margin, efficiency optimization) associated with increased flow rates in syngas-fired gas turbines.

For some systems, dilution may best be achieved by injecting air, carbon dioxide, or steam, rather than nitrogen, into the syngas. For example, if significant energy is required to compress the nitrogen to the injection pressure (e.g., when a low-pressure cryogenic ASU is employed), syngas humidification is often preferred over nitrogen dilution. However, one should consider the tradeoffs of using these other gases.

For example, if air is used, care must be taken that its oxygen content does not create a combustible mixture where combustion is not desired.

If steam is used it must be purified of minerals that could enhance corrosion in the turbine. This is not required for nitrogen.

Another drawback of syngas humidification is that it may require a decrease in the turbine's firing temperature. Introducing steam increases the heat transfer coefficient of the combustion gases, leading to elevated metal temperatures in the expander section [9] and a reduction in turbine life.

3.2 Air-Nitrogen Integration. As shown in Fig. 3, "air-nitrogen" integration consists of supplying the ASU with pressurized air extracted from the gas turbine compressor, followed by injection of some or all of the nitrogen/vitiated-air stream back into the gas turbine.

Both of the CGCC plants operating in Europe feature air-nitrogen integration. See Table 2.

Typically, more than enough air is discharged from the gas turbine compressor to fully supply the air required by a cryogenic or membrane-based ASU in a gasification system. Configurations are commonly referred to as "fully" or "partially" integrated depending on whether all or part of the ASU's feed air is supplied by extracted air. Likewise, the "degree of integration" is commonly defined as the percentage of ASU air feed that is met by extracting air from the gas turbine.

3.2.1 Potential Benefits of Air-Nitrogen Integration. Although air-nitrogen integration makes reliable system operation a more complex and challenging endeavor, it does offer solutions for some gas turbine issues and certain advantages with regard to system efficiency and cost.

As discussed in Sec. 2.2.3, extracting air from the gas turbine compressor reduces the flow through the expander. This reduces the extent of physical modifications required to accommodate increased volumetric flows through the turbine's expander and fuel systems. It also moves the mass flow ratio between the compressor and expander closer to its design-point value, potentially increasing the compressor's surge margin and enhancing the turbine's efficiency.

Overall system efficiency may be improved by supplying the ASU with air extracted from the turbine compressor. For a nonintegrated ASU, the air feed is typically compressed in a dedicated, intercooled, radial-flow compressor. Although the (single-stage) isentropic efficiency of a gas turbine's axial-flow compressor ($\sim$90%) is higher than that of a radial-flow compressor ($\sim$82%), it is likely that the turbine-compressor's overall efficiency will be lower because it is not intercooled. However, if the heat of compression contained in the air extracted for the ASU

[8]Substantially increasing the area of the stage-one nozzle would require a major redesign, essentially equivalent to developing a new gas turbine model.

[9]Syngas-fired gas turbines must be operated on an alternative fuel, typically natural gas, during certain off-design conditions, e.g., during startup and shutdown.

can be usefully recovered, an improvement in overall system efficiency may result. (Recovering the heat of compression is not an issue for membrane-based ASUs since they require a high-temperature air feed (Sec. 7.2.2). Heat recovery for cryogenic ASUs is less straightforward (Sec. 5.2.2).)

Furthermore, since air extraction allows the radial-flow compressor to be reduced in size—or even eliminated in the case of full integration—capital cost savings can be achieved from air-nitrogen integration of the ASU.

3.2.2 Potential Drawbacks of Air-Nitrogen Integration. Increased operational complexity is a significant drawback of air-nitrogen integration. According to the degree of integration, operation of the gas turbine, ASU, and gasifier becomes more interdependent. This decreases the flexibility of system operation and makes reliability a greater challenge. Fluctuations in the pressure and flow rate of air extracted from the turbine-compressor can cause havoc with the operation of air-integrated cryogenic ASUs. See Sec. 5.2.1. Furthermore, system start-up is much more difficult if the gas turbine, gasifier, and ASU must be brought on-line more or less in unison.

Commercial experience to date[10] with full integration was so problematic that the majority of today's designs use extracted air to supply only a portion, if any, of the air feed to a cryogenic ASU [13]. (In contrast, theoretical designs for future systems that employ a membrane-based ASU frequently feature full air-nitrogen integration.)

Another drawback of air-nitrogen integration is that the gas turbine may have to be modified to accommodate the impacts that air extraction may have on the turbine-compressor and the cooling air schemes for the combustor and expander. Although the required modifications will vary among turbine models, GE reports [9] that

- up to ~5% of the turbine-compressor discharge can typically be extracted without any modifications,
- extracting between 5% and 20%[11] is normally possible with some modifications to the compressor discharge casings, piping, and controls, and
- extracting more than 20% requires extensive modifications to the turbine.

General Electric's advanced H gas turbine may represent a notable exception to the above guidelines. Unlike previous turbine designs, which employed open-loop air cooling, the H turbine uses closed-loop steam cooling. Consequently, since none of the air from the compressor discharge must be reserved for cooling duty, more air can be extracted for supply to the ASU [18].

4 Cryogenic Air Separation

Commercialized in the early 20th century, the cryogenic air separation process uses low-temperature distillation to separate air into its product components (oxygen, nitrogen, argon). Cryogenic technology has improved significantly over the years and remains the preferred technology when high oxygen production rates are required. It is used almost exclusively for applications demanding 100 tons/day of oxygen or more [19], which is the case for oxygen-blown gasifiers.

4.1 Low-Pressure Cryogenic. Conventional cryogenic ASUs operate at low pressure, meaning that the feed air is compressed just enough to allow the nitrogen by-product to be discharged at atmospheric pressure. Depending on the oxygen purity specification and the extent of efficiency optimization, absolute air feed pressures to low-pressure (LP) cryogenic ASUs are typically 4.4 to 7.2 bar [6].

Since the boiling points of argon and oxygen are very close, separating oxygen and argon (together) from nitrogen can be accomplished in a two-column[12] distillation process, generating an oxygen product that is 95% pure and a nitrogen product that is 98% pure [19,11]. If oxygen purity greater than 95% is required, a third column would likely be required to remove argon from the oxygen product.

Product streams typically leave the ASU in gaseous form near ambient temperature and pressure. If liquid products are desired, additional refrigeration is required.

4.2 Elevated-Pressure Cryogenic. As described above, LP cryogenic ASUs have traditionally been designed to accept air at an absoute pressure—4.4 to 7.2 bar—such that product streams leave the unit at or near atmospheric pressure. Today, however, cryogenic plants can be designed to operate at much higher pressures—at or near the discharge pressures of gas turbine compressors [2]. In fact, of the four CGCC power plants currently operating, three feature elevated-pressure (EP) ASUs. See Table 2. The ability to operate at higher pressures may be partially attributed to recent improvements in manufacturing technology that increased the maximum operating pressure rating of cryogenic heat exchangers [19].

Elevated-pressure ASUs are typically favored for systems that supply the ASU with extracted air and inject all or nearly all[13] of the byproduct nitrogen into the gas turbine [3,13]. In fact, improvements in EP ASU technology have been credited with enabling nitrogen injection to become an economically viable option [15]. An elevated inlet pressure to the ASU translates to higher suction pressures for the pumps and compressors that boost its oxygen and nitrogen products to their final delivery pressures, reducing their power consumption. Accordingly, the operating pressure of an EP ASU should be adjusted to provide a good match with downstream compressors.

Another important benefit of operating at elevated pressure is that more compact equipment can be used, reducing the capital cost of the ASU.

As integrated systems begin to feature advanced gas turbines with increasingly high pressure ratios, their ASUs will be required to accept air, and deliver nitrogen, at increasingly higher pressures. The degree to which ASU operating pressure can be elevated is limited by thermodynamic and economic constraints. The differences in the volatilities of oxygen, argon, and nitrogen decrease with increasing pressure, making the separation by distillation more difficult. Consequently, if the operating pressure of a cryogenic ASU is elevated high enough, recovering the same quantity of oxygen at a given purity would require additional steps [2].

5 Integration of Cryogenic Systems

5.1 Degree of Integration (Feed Air Fraction). All or part of the air required by a cryogenic ASU can be supplied by extracting air from the discharge flow of the turbine-compressor. Sections 2.2 and 3.2 described several factors that should be considered when determining what fraction of ASU air feed should be supplied via extraction, i.e., the "degree of integration":

- effects on system start-up, reliability, and operational flexibility,

[10]Two CGCC plants in Europe experienced delays in startup and increased operating complexity as a result of employing full integration. Consequently, there is a perception that full integration is now out of favor with the gasification industry, especially when project economics value reliability over efficiency [11].

[11]GE has adapted its Model F machines for syngas such that up to 20% of the air from the compressor discharge can be extracted without affecting the cooling air system [15].

[12]In the typical two-step distillation process, the relatively impure oxygen produced in a high-pressure column is further purified in a low-pressure column.

[13]The benefit of elevated-pressure operation is diminished as a greater fraction of the byproduct nitrogen must be vented because it is not required by the gas turbine. In some cases, it may be worthwhile to utilize an EP ASU coupled with an expander to extract work from the vented nitrogen. In other cases, it may be better to simply employ an LP ASU.

- effect on turbine mass-flow balance,
- extent of gas turbine modifications required,
- comparative efficiencies of the gas turbine-compressor and the dedicated ASU compressor,
- potential for reducing the size of, or even eliminating, the dedicated ASU compressor.

In general, no more than 15%[14] of the turbine-compressor discharge flow is required to supply all of the feed air required by a cryogenic ASU in a gasification system [16]. According to the GE guidelines referenced in Sec. 3.2, extracting this much air would require some—but not "extensive"—modifications to the gas turbine.

A case study [16] by Foster Wheeler Corporation (FWC) analyzed the balance between air extraction and nitrogen dilution and its impact on system efficiency. As discussed in Sec. 2.2.2, the amount of nitrogen diluent that can be injected into the turbine-expander is constrained by the necessity to operate near the design-point value of the turbine's mass flow ratio. Any nitrogen that cannot be injected into the turbine or used elsewhere in the plant must be vented to the atmosphere. However, as more air is extracted from the turbine-compressor, more nitrogen can be injected into the turbine-expander. At some minimum degree of integration,[15] just enough air is extracted to allow the full amount of nitrogen to be injected.

Foster Wheeler concluded that this is the point at which system efficiency is maximized. At lower degrees of integration, net power is steadily reduced as the ASU's dedicated air compressor must consume more power to supply a greater portion of the ASU's air feed. At higher degrees of integration, net power is steadily reduced as the gas turbine generates less power because no more nitrogen is available to replace the mass-flow of extracted air in the turbine expander. The FWC study, which was based upon a GE 9001 FA gas turbine, determined this optimum degree of integration to be 46%. However, other gas turbines would have a different optimum value.

Air Products and Chemicals, Inc. (APCI) cited studies that concluded the optimum degree of integration typically occurs when between 25% and 50% of a cryogenic ASU's air feed is supplied by extracting air from the turbine [2]. This recommendation reflects an optimization of system cost and efficiency, while preserving the ability to start the ASU with its stand-alone air compressor [14].

Both of the CGCC plants currently operating in Europe feature full (100%) integration. See Table 2.

5.2 Conditioning the Extracted Air

5.2.1 Pressure. Many published designs of cryogenic ASUs featuring air-nitrogen integration choose to reduce the pressure of extraction air before it is introduced into the ASU. Two reasons are cited for this approach:

1. Operating at the full extraction air pressure would likely require the addition of a third distillation column to efficiently separate the oxygen at a suitable purity [18]. Gasification-based power systems with cryogenic ASUs usually specify an oxygen purity of approximately 95% by volume [2]. See Table 2. While LP cryogenic plants routinely achieve this purity with two distillation columns, EP cryogenic plants would need a third column to meet the specification if operated at full extraction air pressure. It is unlikely that the addition of a third column could be economically justified in a CGCC plant unless there was an application or market for the liquid argon byproduct that would be separated with the third column [14].
2. Expanding the extraction air to an intermediate pressure may be required to maintain stable operation of the integrated system [6,3]. During changes in ambient conditions, or when the turbine is at partial-load or off-design conditions, the pressure, and perhaps flow rate, of air extracted from the turbine-compressor will fluctuate. If the air is fed directly to a cryogenic ASU, these fluctuations would be translated to the operating pressure of its distillation columns and the suction pressures of its product pumps/compressors. Unless nontraditional controls were used, such changes in distillation pressure would affect the purity at which nitrogen and oxygen is produced [2]. Expanding the extraction air to an intermediate pressure[16] allows the ASU to operate at a steady pressure independent of potential fluctuations in extraction air pressure.

Other benefits can be derived from reducing the pressure of the extraction air. If it is expanded through a turbine that drives a compressor or an electric generator, the system's power consumption can be reduced. Furthermore, expanding the extracted air also cools it, an effect that is magnified if work is extracted during the expansion. This would reduce the duty/size of any heat exchanger required to cool the extraction air. (See Sec. 5.2.2.)

From the data in Table 2, it may be inferred that one or both of the European CGCC plants may operate their ASUs at an intermediate pressure. The pressure ratios of their turbines are reported to be 11.4 and 16.1 and they both operate their ASUs at an absolute pressure of ~10 bar.

5.2.2 Temperature. Due to the heat of compression, air extracted from the gas turbine usually has a temperature between 370 and 430°C [2] and must be cooled to ambient temperature before being introduced into a cryogenic ASU. To achieve a reasonable efficiency, heat must be recovered from the air as it is cooled.

The extracted air's high-level heat is typically recovered by transferring it to the nitrogen stream produced by the ASU. Since the nitrogen is then injected into the turbine, the transferred heat is ultimately utilized by the combined cycle process.

Various strategies have been proposed to recover the extracted air's intermediate- and low-level heat as it is cooled to ambient temperature prior to admission into the ASU. Some integrated designs choose to expand the extracted air somewhat (see Sec. 4.1.3), which further cools the air and recovers intermediate-level heat as work is extracted during the expansion. Low-level heat in the extracted air is often recovered in heat exchangers internal to the cryogenic air separation process.

5.3 Compression of Oxygen and Nitrogen Products. The pressure of oxygen emerging from a cryogenic ASU must be increased before it can be introduced into the gasifier. Likewise, before being supplied to the gas turbine, residual nitrogen must be compressed to the turbine-combustor operating pressure. To accomplish these pressure increases in an optimal manner, the nitrogen and oxygen products may be pumped in liquid form within the ASU, compressed in gaseous form downstream of the ASU, or pressurized in some combination thereof.

5.3.1 Pumping Liquid Products. Both LP and EP cryogenic ASUs can be designed such that all or part of the required oxygen and/or nitrogen product emerges from the distillation columns in liquid form. This allows the product to be pumped to a higher pressure as a liquid, reducing the need for gaseous compression, which is more energy intensive. The liquid can be pumped to its

[14]This guideline is more or less corroborated by Table 2 which shows that the two European CGCC plants fully supply their EP ASUs by extracting 16%–18% of the air discharged by their turbine-compressors. GE cites a similar range of 11%–20% air extraction for fully integrated gasification systems with cryogenic ASUs [15].

[15]I.e., the minimum degree of integration when the turbine-compressor is not throttled down. Although throttling would allow all of the ASU nitrogen to be injected at a lower degree of integration, i.e., less air extraction, FWC determined that this strategy actually decreases system efficiency.

[16]A report by FWC [16] recommended expanding the extraction air to the pressure at which air would be extracted when the gas turbine is operating at 50% load.

full delivery pressure or some intermediate pressure that provides a good match for a selected downstream compressor. These so-called "fully pumped" or "partially pumped" liquid cycles add another degree of freedom for optimizing the design of integrated systems, offering the potential to reduce—or even eliminate—stages of downstream compression.

5.3.2 Compressing Gaseous Products. Compressing product gases downstream of the ASU in a centrifugal compressor train is frequently the most economic option to achieve the required delivery pressures [19]. This can be accomplished either isothermally or adiabatically. The systems analyst should explore which approach is the most economic for a particular cycle.

Traditionally, the products of air separation have been compressed in a nearly isothermal process using many stages of intercooled centrifugal compressors, each having a pressure ratio less than 2 [2]. Although this isothermal approach minimizes the work required for compression, it requires that nearly all the heat of compression be rejected to cooling water through intercoolers.

Although adiabatic compression requires more work than isothermal compression, studies have shown that it could result in a higher overall efficiency for some cycles [2]. Since both the oxygen and nitrogen are supplied to high temperature devices in integrated systems (the gasifier and gas turbine, respectively), retaining the heat of compression in these gases could be beneficial for some systems notwithstanding the additional work required by adiabatic compression. Furthermore, the adiabatic approach would likely reduce the compressor's capital cost since intercooling equipment is not required.

6 Membrane-Based Air Separation

6.1 Description of Membrane-Based ASUs. Membrane-based air separation technology—which both APCI and Praxair are currently developing[17]—separates oxygen from air by transporting oxygen ions across a nonporous ceramic membrane. The membrane is a mixed conductor, meaning that it conducts both electrons and oxygen anions when operated at high temperature, typically 800–900°C. Given such high temperatures, oxygen anions rapidly diffuse through the membrane, driven by a chemical-potential gradient created by the difference in oxygen partial pressure across the membrane. Pure oxygen emerges from the membrane at a very low pressure (probably less than one atmosphere). Electrons flow across the membrane countercurrent to the oxygen anions, completing an internal electronic circuit.

For certain applications, a purge gas could be used to sweep away oxygen as it emerges from the permeate side of the membrane. Both APCI and Praxair have patented concepts that would use steam as the purge gas, producing a blend of oxygen and steam—a potentially attractive mixture for gasification applications. Although the absolute pressure of the purge gas could be up to 40 bar, the partial pressure of the oxygen on the permeate side of the membrane would have to be maintained near one atmosphere [5].

Presently, however, APCI and Praxair are focused on first developing and commercializing a "simpler" membrane design that does not feature a purge gas.[18] In their current designs, a downstream compressor is used to move oxygen away from the membrane and compress it to its final delivery pressure (or some intermediate pressure). The suction of the oxygen compressor maintains a low pressure on the permeate side of the membrane.

6.2 Oxygen Recovery in Membrane-Based ASUs. The fraction of oxygen that a membrane-based ASU recovers from a given flow rate of feed air can be adjusted by varying one of the following parameters (assuming that membrane composition and thickness are fixed):

1. *Feed air pressure*—Recovery increases as the pressure on the nonpermeate side of the membrane is increased. In air-nitrogen integrated systems, this pressure is usually set by the turbine-compressor discharge pressure.
2. *Membrane temperature*—Recovery increases as the membrane operating temperature is increased since oxygen diffuses through the membrane more rapidly at higher temperatures. Due to the temperature limitations of currently available materials, membrane temperatures above 900°C are not practical at present.
3. *Permeate suction pressure*—Recovery increases as the pressure on the permeate side of the membrane is decreased. Suction pressure is the most important process parameter for controlling oxygen recovery since it can be varied without affecting the operation of the integrated turbine (unlike the feed-air pressure and membrane temperature which set the conditions of the nonpermeate stream that is supplied to the turbine). The extent to which the suction pressure can be decreased is limited by the structural integrity of the membrane (leakage) and the efficiency penalty imposed by a larger pressure drop across the membrane.
4. *Membrane area*—Recovery increases as the total membrane area, i.e., the number of stages of membrane separation, is increased.

Unlike cryogenic ASUs, which routinely recover nearly 100% of the oxygen from the feed air stream, membrane-based ASUs will be designed for a much lower oxygen recovery rate (10%–75%, depending on the ASU design). Accordingly, membrane-based ASUs require a much larger volume of feed air compared to cryogenic ASUs. Their larger air feed requirement makes integration more complicated for membrane-based ASUs since the extent to which the gas turbine-compressor must be physically modified increases as the requirement for air extraction increases (see Sec. 3.2).

6.3 Product Purity in Membrane-Based ASUs

6.3.1 Oxygen Purity. Absent any leaks, membrane-based separation is 100% selective, producing highly pure (>99%) oxygen. However, membrane leakage rates have been reported [20] that correspond to an oxygen purity of 95%—the same purity that cryogenic ASUs are optimized to achieve.[19] Lower membrane leak rates would yield an oxygen stream with a lower volume of inert gases. This would allow more compact equipment to be used throughout the gasification section of the plant, which may modestly reduce its capital cost. More importantly, if a portion of the syngas is used to co-produce fuels or chemicals, an oxygen purity exceeding 95% is typically required [11].

6.3.2 Nitrogen Purity. Unlike cryogenic ASUs, membrane-based ASUs do not produce a pure stream of nitrogen. (Even with an infinitely large membrane area, some oxygen would remain in the nonpermeate stream.) Consequently, a small cryogenic ASU may be the preferred approach for producing pure nitrogen for various subsystems in the plant, e.g., coal handling and conveying [6].

7 Integration of Membrane-Based Systems

Because membrane-based ASUs are operated at high temperatures, they are far better suited than cryogenic ASUs to be thermally integrated with the combined cycle section of a CGCC plant.

[17]APCI and Praxair are developing membrane-based air separation technology under the names Ion Transport Membrane (ITM) and Oxygen Transport Membrane (OTM), respectively.

[18]Depending on future application needs and necessary technology advances, APCI and Praxair may revisit steam-purged concepts at a later time.

[19]To achieve purities >95%, a cryogenic ASU would require the addition of more distillation columns which substantially increases its cost.

7.1 Degree of Integration (Feed Air Fraction). As discussed in Sec. 6.2, the fraction of oxygen that a membrane of given composition and thickness will recover from the air feed stream is a function of three process parameters: air feed pressure, membrane temperature, and permeate suction pressure. Even if optimal values are predetermined for these process parameters, the system designer must still determine the optimal balance between total membrane area (the final parameter which determines oxygen recovery fraction) and feed air flow rate that yields the required amount of oxygen.

For integrated systems, the membrane area (and its capital cost) can be reduced by extracting more supply air from the gas turbine-compressor. However, extracting more air comes at the price of more extensive modifications to the gas turbine-compressor, and larger heat exchangers (if required) for conditioning the streams entering and leaving the ASU.

A configuration described by APCI in 2001 [3] extracted ~80% of the air discharged from the gas turbine-compressor for supply to the membrane-based ASU, but an ongoing study by APCI and SWPC indicates that somewhat less extraction may be necessary.

Since membrane-based ASUs will likely require the extraction of over half of the air discharged from the turbine-compressor, major modifications to the turbine may be required. According to the guideline described in Sec. 3.2.2, extracting more than 20% of the compressor discharge air would likely require extensive modifications to the turbine—although the H turbine may be more amenable to higher rates of extraction.

7.2 Conditioning the Extracted Air. Since the transport of oxygen ions is thermally activated and driven by the difference in oxygen partial pressures across the membrane, the air supplied to the membrane must be heated (typically to 800–900°C[20]) and compressed (typically to an absolute pressure of 7–20 bar) [21].

7.2.1 Pressure. Air Products has reported absolute operating pressures of 7–20 bar for the membrane-based ASU it is developing. This matches well with most candidate gas turbines, which have pressure ratios between 14:1 and 19:1. Since the absolute compressor discharge pressures for these machines would range between 14 and 19 bar, compression or expansion of the extracted air would not necessarily be required depending on the turbine selected and the exact operating pressure of the ASU.

It is worth noting that, unlike cryogenic ASUs, membrane-based ASUs are not substantially affected by small fluctuations in feed air pressure. Therefore, there is no need to expand the extracted air to an intermediate pressure for the express purpose of "filtering out" pressure fluctuations when the gas turbine is operated at partial load or off-design conditions.

7.2.2 Temperature. Air extracted from the gas turbine compressor discharge usually has a temperature between 370 and 430°C [2]—well below the membrane operating temperature of 800 to 900°C. To achieve the membrane operating temperature, the extracted air can be heated via heat exchange with another process stream and/or by combusting a portion of the syngas directly with the air.

For example, APCI [3] described a configuration that utilizes both heat exchange and direct combustion to heat the extraction air to the membrane's operating temperature. See Fig. 4. First, extraction air is preheated in a recuperator against the hot, nonpermeate stream exiting the membrane. Next, the preheated air is combusted with syngas to boost its temperature to the required value prior to entering the membrane. Since the membrane has a lower sulfur tolerance than the gas turbine, this configuration requires that additional sulfur be removed from the syngas that is provided to the air heater [3].

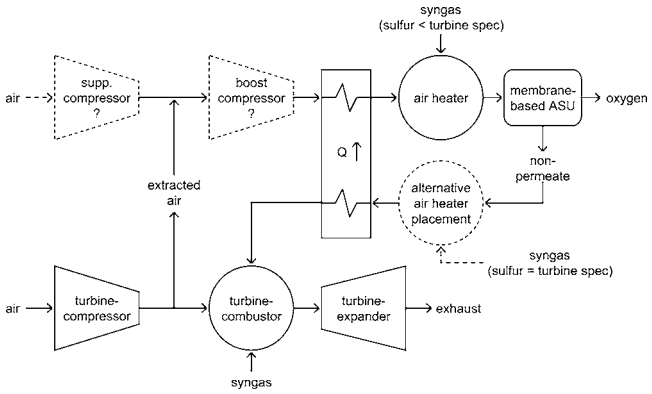


Fig. 4 Integration of a membrane-based ASU

7.3 Conditioning and Expanding the Vitiated Air Product. Since the vitiated, nonpermeate stream leaves the ASU at an elevated temperature and pressure—very close to the conditions of the ASU feed stream—its energy must be recovered to attain a reasonable system efficiency. This can be efficiently accomplished by expanding the nonpermeate stream through the gas turbine.

7.3.1 Temperature. Combined cycle units that are candidates for CGCC applications have firing temperatures (i.e., expander inlet temperatures) of 1300–1500°C,[21] well above the temperature of the nonpermeate stream as it exits the membrane-based ASU. Consequently, even if the nonpermeate stream was not cooled in a downstream recuperator (as in Fig. 4), it would still have to be combusted with syngas in order to achieve the required turbine firing temperature. The residual oxygen content of the nonpermeate stream typically exceeds what is required for this combustion.

If the nonpermeate stream is not cooled before being injected into the gas turbine combustor, it will be too hot to serve as a cooling stream for the combustor walls and other arrangements will be needed to maintain acceptable metal temperatures (perhaps around 700°C).

7.3.2 Pressure. Unlike the syngas stream, which must be supplied at a pressure that is 5 to 7 bar higher than the turbine-combustor pressure due to the pressure drop across the fuel control valves, a turbine's oxidant/diluent stream can be supplied at a pressure approximately equivalent to that of the combustor [1,22]. The operating pressure range that APCI is currently targeting, 7–20 bar (absolute), matches well with most candidate gas turbines, which have pressure ratios between 14:1 and 19:1.

A higher membrane operating pressure would be required to integrate with advanced gas turbines having pressure ratios higher than 20:1, such as the GE 7H which has a 23:1 pressure ratio [23]. In the laboratory, APCI has operated their membrane at absolute pressures up to 30 bar [14]. The alternative—compressing the hot, high-volume nonpermeate stream to the appropriate supply pressure—would likely be infeasible.

7.4 Conditioning the Oxygen Product. The conditioning required for the oxygen produced by a membrane-based ASU depends largely upon whether a purge gas is used to sweep away oxygen from the permeate side of the membrane. Depending upon the temperature and pressure of the sweep gas, additional conditioning of oxygen-containing stream may or may not be necessary.

In APCI's current design, which does not utilize a purge gas,

[20]Praxair [5] is attempting to reduce the operating temperature of the membrane-based ASU they are developing from 900 to 750°C. A lower operating temperature could reduce the capital cost of the ASU and its associated equipment.

[21]When operated on natural gas, "F" class gas turbines have firing temperatures between 1300°C and 1400°C and "G" and "H" class gas turbines have firing temperatures of 1400–1500°C.

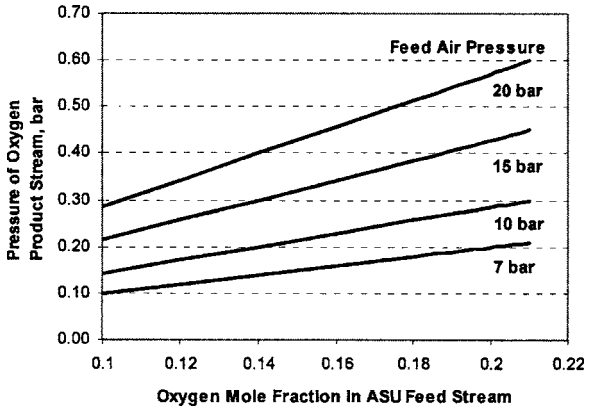


Fig. 5 Operating pressure envelope for membrane-based ASUs

the oxygen product stream emerges at a high temperature (800–900°C) but at low pressure (likely below atmospheric pressure). APCI [21] offers the following heuristic for relating the pressure of the oxygen permeate stream (P_{PERM}) to the pressure (P_{FEED}) and oxygen mole fraction (X_{O2}) of the feed stream provided to the membrane-based ASU. The heuristic is based upon both process and economic factors that will affect commercial applications of the technology:

$$P_{PERM} \approx (P_{FEED})(X_{O2})/7$$

That is, the pressure of the oxygen produced by a membrane-based ASU is approximately one-seventh of the oxygen partial pressure in the ASU feed stream.[22] In Fig. 5 this heuristic is graphed within the recommended operating ranges for the feed stream's pressure and oxygen mole fraction.

Before being compressed to the gasifier operating pressure (typically 20 bar (absolute) or higher), the oxygen must be cooled, typically via heat integration with the steam cycle or another part of the system. Unlike cryogenic ASUs, which can be designed to produce liquid oxygen that can be fully or partially pumped to its delivery pressure, the oxygen emerging from a membrane-based ASU must be compressed in gaseous form.

8 Conclusion

The differences in composition and heating value between natural gas and syngas present two main challenges for adapting a gas turbine for a CGCC system: accommodating different combustion characteristics and accommodating higher flow rates. Both of these challenges can be partially addressed by integrating the gas turbine with the air separation unit.

Conventional cryogenic ASUs can be integrated with the gas turbine in "nitrogen-only" or "air-nitrogen" configurations with only minor modifications to the gas turbine. In recent years, the ability to operate cryogenic ASUs at elevated pressures has led to several innovations that enhance their integration with gas turbines, especially in air-nitrogen configurations that inject all or nearly all of the byproduct nitrogen into the gas turbine.

Membrane-based ASUs, which are still under development, promise significant capital cost reductions for CGCC systems. Furthermore, since they are operated at high temperatures, they are better suited than cryogenic units to be thermally integrated with the combined cycle section of a CGCC plant, leading to higher system efficiency.

However, unlike cryogenic units, membrane-based ASUs require air-nitrogen integration with the gas turbine to achieve a reasonable system efficiency. Compared to nitrogen-only integration, air-nitrogen integration leads to a much greater increase in operational complexity. Furthermore, since membrane-based ASUs will likely require the extraction of over half of the air discharged from the turbine-compressor (much more than cryogenic ASUs), major modifications to the turbine may be required.

[22]Some applications may deviate from this general guideline. For example, to increase the oxygen recovery fraction, the partial pressure of oxygen on the permeate side may be decreased to less than one-seventh of that on the nonpermeate side. Praxair [5] demonstrated that the oxygen flux through the membrane they are developing has a logarithmic dependence on the oxygen partial pressure ratio across it. The oxygen flux targeted by Praxair corresponds to an oxygen partial pressure ratio of ~3.5—about half the ratio implied by the APCI heuristic cited above.

9 Disclaimer

Acknowledgment

As evidenced by the references cited, this paper has relied heavily on papers published by APCI, GE, SWPC, and Praxair. The authors also gratefully acknowledge the assistance provided by James C. Sorensen (APCI), VanEric E. Stein (APCI), and Dennis Horazak (SWPC).

References

[1] Horazak, D. A., Chen, A. G., and Pillsbury, P. W., 2000, Siemens Westinghouse Power Corporation, Gas Turbines for Coal Gasification Combined Cycle Service, Pittsburgh Coal Conference.
[2] Smith, A. R., Sorensen, J. C., and Woodward, D. W., 1996, Air Products and Chemicals, Integration of Oxygen Plants and Gas Turbines in IGCC Facilities, American Power Conference.
[3] Stein, V. E., Juwono, E, and Demetri, E. P., 2001, Air Products and Chemicals, Improving IGCC Economics through ITM Oxygen Integration, 18th International Pittsburgh Coal Conference.
[4] Allam, R. J., Foster, E. P., and Stein, V. E., 2002, "Air Products and Chemicals, Improving Gasification Economics through ITM Oxygen Integration," 5th European Gasification Conference.
[5] Prasad, R., Chen, J., van Hassel, B., Sirman, J., White, J., Apte, P., Aaron, T., and Shreiber, E., 2002, Praxair, Inc., Advances in OTM Technology for IGCC, DOE Contract No. DE-FC26-99FT40437, accessed 2/18/2004 at www.netl.doe.gov/coalpower/gasification/gas-sep/index.html.
[6] Allam, R. J., Russek, S. L., Smith, A. R., and Stein, V. E., 2000, "Air Products and Chemicals, Cryogenics and Ceramic Membranes: Current and Future Technologies for Oxygen Supply in Gasification Systems," 4th European Gasification Conference.
[7] Liss, W. H., and Thrasher, W. R., 1992, Gas Technology Institute, Variability of Natural Gas Composition in Select Major Metropolitan Areas of the United States, GRI-92/0123.
[8] Brdar, R. D., and Jones, R. M., 2000, GE Power Systems, GE IGCC Technology and Experience with Advanced Gas Turbines, GER-4207.
[9] Brooks, F. J., 2000, GE Power Systems, GE Gas Turbine Performance Characteristics, GER-3567H.
[10] Audus, H., and Jackson, A. J. B., 2000, "CO_2 Abatement by the Combustion of H_2-Rich Fuels in Gas Turbines," 5th International Conference on Greenhouse Gas Control Technologies.
[11] Ratafia-Brown, J., Manfredo, L., Hoffman, J., and Ramezan, M., 2002, National Energy Technology Laboratory, Major Environmental Aspects of Gasification-Based Power Generation Technologies, Contract No. DE-AT26-99FT20101.
[12] Hannemann, F., Koestlin, B., Zimmermann, G., Morehead, H., and Pena, F. G., 2003, Siemens, Pushing Forward IGCC Technology at Siemens, presentation to 2003 Gasification Technologies Conference.

[13] Holt, N., 2003, "Operating Experience and Improvement Opportunities for Coal-Based IGCC Plants," Mater. High. Temp. **20**(1).
[14] Stein, V. E., 2004, Air Products and Chemicals, personal communication, April 13.
[15] Todd, D. M., 2000, GE Power Systems, Gas Turbine Improvements Enhance IGCC Viability, Gasification Technologies Conference.
[16] Farina, G. L., and Bressan, L., 1999, Foster Wheeler Corporation, Improve Performance, Reduce Capital Cost by Optimizing IGCC Design, accessed 1/15/2004 at www.fwc.com.
[17] Frey, H. C., and Rubin, E. S., 1991, Carnegie Mellon University, Development and Application of a Probabilistic Evaluation Method for Advanced Process Technologies, Contract No. DE-AC21-88MC24248.
[18] Falsetti, J. S., 2000, Texaco Power and Gasification, From Coal or Oil to 550 MWe via 9H IGCC, Gasification Technologies Conference.
[19] UIG (Universal Industrial Gases, Inc.), 2003, Air Separation Technology Overview, accessed 10/2/2003 at www.uigi.com/compair.html.
[20] Prasad, R., Chen, J., Chen, H., Lane, J., White, J., Corpus, J., Schreiber, E., Spero, J., and van Hassel, B. A., 2003, Praxair, Inc., Oxygen Transport Membranes for Future IGCC Power Plants, Pittsburgh Coal Conference.
[21] Armstrong, P. A., 2002, "Air Products and Chemicals, Method for Predicting Performance of an Ion Transport Membrane Unit-Operation," Accessed 2/18/2004 at www.netl.doe.gov/coalpower/gasification/gas-sep/index.html
[22] Horazak, D. A., 2003, Siemens Westinghouse Power Corporation, personal communication, November 21.
[23] Matta, R. K., Mercer, G. D., and Tuthill, R. S., 2000, General Electric Power Systems for the 21st Century—"H" Gas Turbine Combined-Cycles. GER-3953B.

P. Dewallef
ASMA Department,
University of Liège,
1 Chemin des Chevreuils,
4000 Liège, Belgium
e-mail: p.dewallef@ulg.ac.be

C. Romessis
Laboratory of Thermal Turbomachines,
National Technical University of Athens,
P.O. Box 64069,
Athens 15710, Greece
e-mail: cristo@mail.ntua.gr

O. Léonard
ASMA Department,
University of Liège,
1 Chemin des Chevreuils,
4000 Liège, Belgium
e-mail: o.Leonard@ulg.ac.be

K. Mathioudakis
Laboratory of Thermal Turbomachines,
National Technical University of Athens,
P.O. Box 64069,
Athens 15710, Greece
e-mail: kmathiou@central.ntua.gr

Combining Classification Techniques With Kalman Filters for Aircraft Engine Diagnostics

A diagnostic method consisting of a combination of Kalman filters and Bayesian Belief Network (BBN) is presented. A soft-constrained Kalman filter uses a priori information derived by a BBN at each time step, to derive estimations of the unknown health parameters. The resulting algorithm has improved identification capability in comparison to the stand-alone Kalman filter. The paper focuses on a way of combining the information produced by the BBN with the Kalman filter. An extensive set of fault cases is used to test the method on a typical civil turbofan layout. The effectiveness of the method is thus demonstrated, and its advantages over individual constituent methods are presented.
[DOI: 10.1115/1.2056507]

Introduction

Diagnostic methods employing statistical inference can be mainly categorized into *regression methods* and *classification methods*, depending on the way information is processed. In classification problems the task is to assign an input to one of a number of discrete classes or categories. However, in regression problems the outputs represent the values of continuous variables [1]. For example, in the case of aircraft engine diagnostics, outputs of the regression algorithm is a set of numerical health parameter values, whereas for a classification algorithm each parameter is assigned to a given class ("faulty not faulty" or "low correct high").

Applying regression techniques such as Kalman filtering on gas turbine engine diagnosis [2,3] poses some stability problems when few measurements are available; health parameter estimation is unstable when a low redundancy is encountered. Indeed, for commercial aircraft engines (twin spool high bypass ratio turbofan engines), it is typical to have more than 10 parameters to estimate from only seven to nine measurements. As a result of the negative redundancy, the problem is underdetermined and the solution is not unique. Moreover, this estimation relies on uncertain observations available under the form of a noisy measurement set. Increasing the number of samples could make the redundancy positive, but because they are related to nearly the same operating points they do not represent a set of independent observations, and therefore, the estimation remains unstable. Some approaches in turbine engine diagnostics where few measurements are available can be found in [4–8], but they all assume some predefined configurations for the faults. Those predefined configurations are built into the algorithm, and it is not possible for the user to interact with them.

As stated above besides regression algorithms, classification algorithms solve the same problem in a different way and, therefore, give a different kind of results. Although less accurate, classification algorithms are more reliable and usually more stable. Moreover, some classification algorithms allow some qualitative knowledge (i.e., user experience, some events that have been observed but cannot be modeled, etc.) to be introduced into the classification rule. Those characteristics make classification algorithms very complementary to regression algorithms. Some results of classification techniques applied to turbine engine diagnostic can be found in [9,10].

In [11], a physical state variable model is combined with a neural network model to improve the diagnostic of a turbine engine by including empirical knowledge. Although the present contribution uses a different method, the final scope is the same: to decrease the proportion of false alarms and to increase the proportion of fault detection by taking advantage of two different kind of methods.

The approach presented in this contribution is based on a soft-constrained Kalman filter (SCKF) developed in [8] and a Bayesian Belief Network (BBN) developed in [9]. The principal diagnostic tool is the Kalman filter, which produces the estimations of the unknown health parameters. The estimation process incorporates information derived by the BBN—a fact that is shown to improve fault detection efficiency. A way of combining the two techniques is described in the following. For completeness, a brief description of the diagnostic problem and the structure of each of the two techniques is given first.

Problem Statement

The identification problem consists of recursively estimating a set of health parameters representing possible degradations of specific components (fan, high pressure turbine, etc.) on the basis of

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. GT2004-53541.

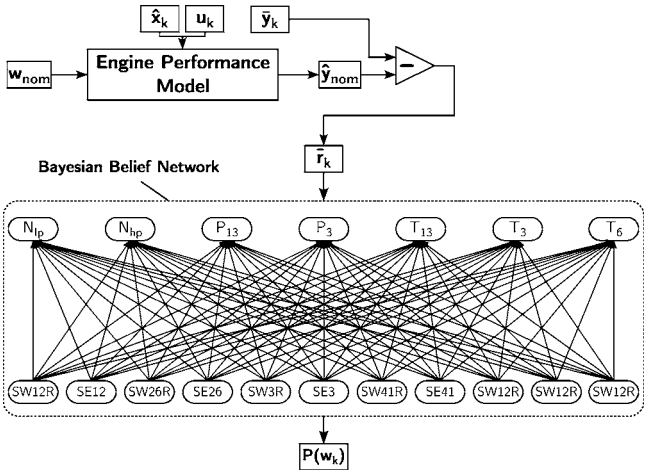


Fig. 1 Classification procedure using BBN

successive measurement samples performed on a turbine engine. A steady-state model of this engine must be made available for measurement simulation. This model is based on a set of health parameters $\mathbf{w}_k$, engine operating point $\mathbf{x}_k$, and command parameters $\mathbf{u}_k$. The set of predicted measurement $\hat{\mathbf{y}}_k$ is generated through this nonlinear model $\mathcal{G}$

$$\hat{\mathbf{y}}_k = \mathcal{G}(\mathbf{u}_k, \mathbf{x}_k, \mathbf{w}_k) \tag{1}$$

Health parameter identification is intended to solve the inverse problem of measurement prediction: inputs are the raw measurements $\bar{\mathbf{y}}_k$ and outputs are the health parameters $\hat{\mathbf{w}}_k$ to be estimated. The process starts with initial values for the health parameters $\hat{\mathbf{w}}_k$. Residuals are built as $\bar{\mathbf{r}}_k = \bar{\mathbf{y}}_k - \hat{\mathbf{y}}_k$ to compare predicted measurements $\hat{\mathbf{y}}_k$ to observed measurements $\bar{\mathbf{y}}_k$. Although classification algorithms are aimed at associating a given class to each residuals $\bar{\mathbf{r}}_k$, the purpose of regression algorithm is to fit health parameters to observed raw measurements $\bar{\mathbf{y}}_k$ and leads to drive residuals $\bar{\mathbf{r}}_k$ to zero. Regression problems are thus solved through the minimization of the following inner product, which is also known as the weighted least-squares approach where the weighting factor $\mathbf{R}_r$ is the residual covariance

$$\mathcal{J} = \bar{\mathbf{r}}_k^T \mathbf{R}_r^{-1} \bar{\mathbf{r}}_k \tag{2}$$

Classification Techniques: BBN

The classification technique is a Bayesian Belief Network (BBN), presented in [9] (to which the interested reader is referred for more details). The structure of such a network, which will be employed later in the paper, is shown in Fig. 1. It includes 18 nodes representing deviations of 11 health parameters ($\hat{\mathbf{w}}_k$), and seven measurements ($\bar{\mathbf{y}}_k$) for the test case of a turbofan engine.

The BBN is supported by an engine performance model. A set of measurement readings ($\bar{\mathbf{y}}_k$) is preprocessed together with $\hat{\mathbf{x}}_k$ and $\mathbf{u}_k$ to derive the deviation of the seven measurements from their nominal value. These deviations are presented to the BBN, from which the output probabilities $P(\mathbf{w}_k)$ are estimated. Each output node produces the probability for a health parameter to belong to a certain interval, for example, to be around the value that represents a "healthy" component or to be away from this value for a fault condition. The output is thus an indication of what the most probable values of health parameters are, and from this information the stand-alone BBN derives a fault diagnosis.

Soft-Constrained Kalman Filters

The soft-constrained Kalman filter (SCKF) has a typical structure depicted in Fig. 2. Details on the specific SCKF employed

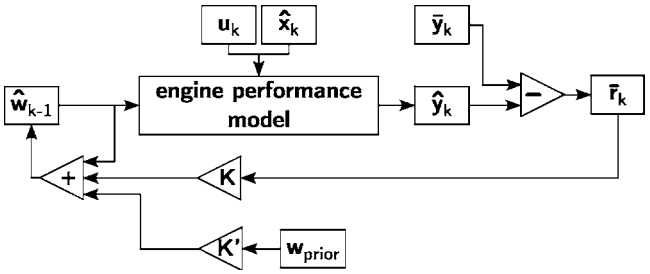


Fig. 2 Description of the soft-constrained Kalman filter

here can be found in [8], where the development of this filter is described. A Kalman filter is a regression algorithm, meaning that its outputs are numerical values of health parameters. Differences between regression methods found in the literature reside in the update rule selected to minimize $\mathcal{J}$ regarding $\hat{\mathbf{w}}_k$. Many methods found in the aircraft-engine diagnostic literature are *batch* methods where the whole set of available data is processed in one batch. In the case of non-linear engine performance model such an estimation involves the use of non-linear iterative optimization methods and the estimation is therefore called *batch iterative*. At the other hand, *recursive* procedures do not involve the batch processing of the full block of data but only a simple update of the parameters each time new data are available. Moreover, the updating scheme does not imply any optimization procedure. Generally speaking, the advantage of a recursive approach resides in the quick and simple update formula provided by those algorithms that is well suited for on-line estimation.

As outlined in Fig. 2, prior information is introduced into the identification loop. This is done through a penalizing term added to the objective function and favoring health parameters, which are in the neighborhood of predetermined values $\mathbf{w}_{\text{prior}}$. This process is called *ridge regression* [12,13]. Because this prior information is assumed to be Gaussian, it is parametrized by its mean $\mathbf{w}_{\text{prior}}$ and covariance matrix $\mathbf{D}$. Prior information on health parameters are assumed to be uncorrelated, and $\mathbf{D}$ is thus strictly diagonal. $\hat{\mathbf{w}}_{k-1}$ is updated by combining $\mathbf{w}_{\text{prior}}$ to current raw measurements $\bar{\mathbf{y}}_k$ to assess $\hat{\mathbf{w}}_k$. It is through $\mathbf{w}_{\text{prior}}$ that coupling the BBN is achieved, as explained in the next section.

In Fig. 2, the engine operating point $\mathbf{x}_k$ is assumed to be known. In this contribution it is estimated using a Kalman filter together with an additional measurement set $\bar{\mathbf{z}}_k$. The complete sequence is described in Fig. 3. A set of measurements together with the set of command parameters are preprocessed into a Kalman filter that estimates the operating point $\hat{\mathbf{x}}_k$, which is used afterward in the health-parameter estimation.

Combining Classification Techniques to Kalman Filters

The flowchart of the procedure combining the Kalman filter and BBN is summarized in Fig. 4. Both the BBN and Kalman filter use the same preprocessing through operating-point estimation. It increases the measurement prediction accuracy for both algo-

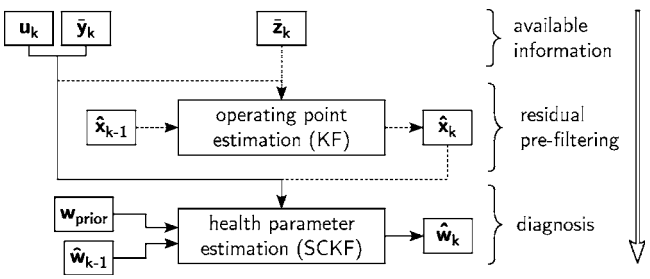


Fig. 3 Data preprocessing for operating-point estimation

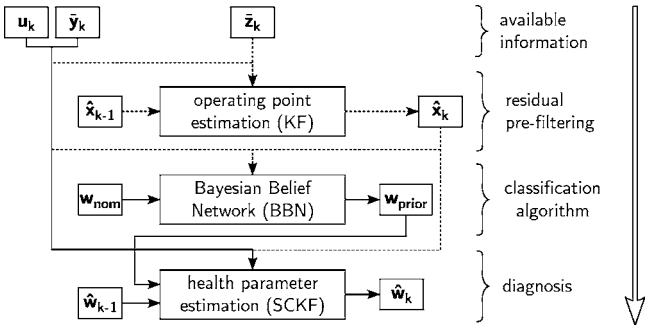


Fig. 4 Procedure followed to combine classification algorithm with Kalman filter

rithms. BBN results are fed into the SCKF as prior information to achieve the health-parameter estimation. The final outcome of the combination is the SCKF result called hereafter "results of the combination."

The communication between a classification algorithm, such as the BBN, and a Kalman filter is not straightforward because they use different probability density functions. Although prior information needed by the Kalman filter has to be Gaussian, outputs of classification algorithms are the probabilities of a measurement set to belong to a specific class and, therefore, no probability density function is attached to the results. This situation is depicted in Fig. 5. Probabilities related to outputs of the classification algorithm are assumed to be piecewise constant and must be converted into a Gaussian one to be fed into the SCKF.

The problem of Fig. 5 underlines the more general situation where qualitative knowledge is to be compared to quantitative knowledge. This difference in usually sufficient to prevent the combination between classification techniques and regression techniques. The solution proposed herein is to preserve, as much as possible, the distribution of probability given by the BBN. If the BBN diagnoses that a given parameter belongs to a specific fault case with a probability of 100%, then this must appear in the Gaussian function fed into the Kalman filter by a small covariance. Conversely, if the BBN is unable to make a diagnostic, then the probability is spread on the whole set of possible categories; this must also appear in the Gaussian function fed into the Kalman filter by an important variance. The most efficient way to translate this knowledge is to go back to the definition of the mean μ and the variance σ^2 of a given probability function p

$$\mu = \int_{-\infty}^{\infty} wp(w)dw \tag{3}$$

$$\sigma^2 = \int_{-\infty}^{\infty} (w-\mu)^2 p(w)dw \tag{4}$$

Applying relations (3) and (4) to the situation in Fig. 5 yields

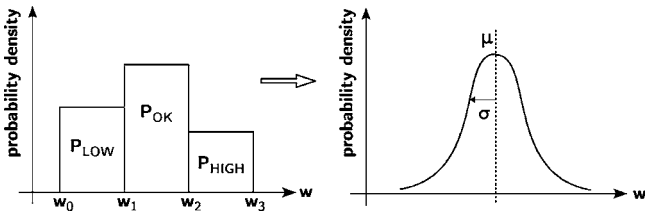


Fig. 5 Conversion of probability density from a piecewise constant function into a Gaussian function

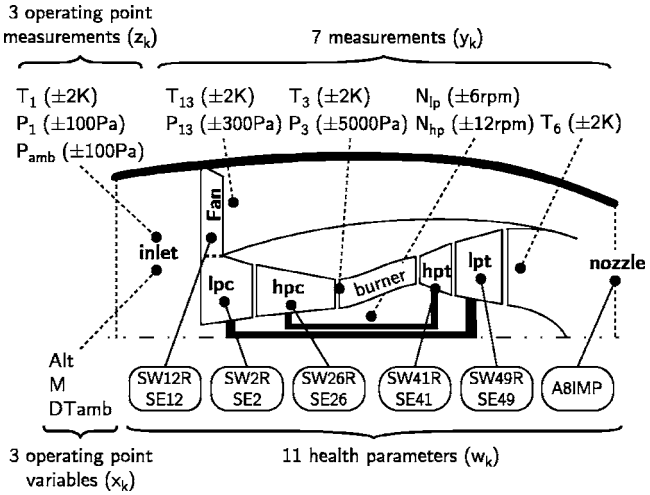


Fig. 6 Turbofan engine layout. Measurement uncertainties represent three times the standard deviation

$$\mu = \sum_{i=0}^{n_c} \int_{w_i}^{W_{i+1}} w \frac{P_i}{w_{i+1}-w_i} dw = \sum_{i=0}^{n_c} \frac{P_i}{2}(w_i + w_{i+1}) \tag{5}$$

$$\sigma^2 = \sum_{i=0}^{n_c} \int_{w_i}^{W_{i+1}} (w-\mu)^2 \frac{P_i}{w_{i+1}-w_i} dw$$

$$= \sum_{i=0}^{n_c} \frac{P_i}{3}[(w_i-\mu)^2 + (w_{i+1}-\mu)^2 + (w_i-\mu)(w_{i+1}-\mu)] \tag{6}$$

The mean and covariance needed by the Kalman filter are, respectively, built using relations (5) and (6). These are the key relations that allow the flow of information from BBN to SCKF.

Application of the Method

In order to show how the method achieves a diagnosis and "examine" its effectiveness, its application to a test case of a twin-spool mixed-flow turbofan is presented. This type of engine is representative of current-day civil aircraft propulsors. The particular engine layout chosen is shown in Fig. 6. (See Table 1.) This engine and fault cases examined has been used as a test case by several previously published diagnostic methods (e.g., [5,7–9]). It can therefore be considered as a benchmark case. The available sensors with their accuracies and the set of 11 health parameters representing some degradations of specific components (fan, high pressure turbine, etc.) are shown in Fig. 6.

An extensive set of faults, shown in Table 2, representative of possible situations expected to be encountered in practice, defined

Table 1 Detailed nomenclature of Fig. 6

ALT=flight altitude
DTAMB=ΔT from iso
M=flight Mach number
Tx=total temperature at station *x*
Px=total pressure at station *x*
SEx=efficiency degradation at component *x*
SWxR=flow capacity at component *x*
A8IMP=fouling factor of the nozzle
LPC=low-pressure compressor
HPC=high-pressure compressor
HPT=high-pressure turbine
LPT=low-pressure turbine

Table 2 Fault cases of a turbofan engine

a	−0.7% on SW2R	−0.4% on SE2	FAN, LPC
	−1% on SW12R	−0.5% on SE12	
b	−1% on SE12		
c	−1% on SW26R	−0.7% on SE26	HPC
d	−1% on SE26		
e	−1% on SW26R		
f	+1% on SW42R		HPT
g	−1% on SW42R	−1% on SE42	
h	−1% on SE42		
i	−1% on SE49		LPT
j	−1% on SW49R	−0.4% on SE49	
k	−1% on SW49R		
l	+1% on SW49R	−0.6% on SE49	
m	+1% on A8IMP		Nozzle
n	−1% on A8IMP		

in [14], is used. Faults in all individual components are included. Different types of faults are considered by involving one or more health parameters of a component.

A steady-state model of this engine is also employed [15] to simulate the measurements $\hat{\mathbf{y}}_k$ given:

1. set of command parameters $\mathbf{u}_k$
2. operating point $\hat{\mathbf{x}}_k$
3. health parameters $\hat{\mathbf{w}}_k$

Simulated measurements have been generated using the steady-state engine model during a cruise flight (ALT=10,800 m, M =0.82, DTAMB=0 K). Data sequences are generated with a duration of 5000 s with a data acquisition rate of 2 Hz. Profile of simulated faults is a steep fault with amplitude defined in Table 2 occurring at $t=50$ s followed by a slow drift of the same fault occurring at $t=2500$ s. In this way, the behavior of the method for both abrupt faults and gradual deteriorations is examined.

The time evolution of estimated health parameters will first be presented to show how the method traces their changes, abrupt or gradual. Representative test cases are chosen to demonstrate that the combined method performs better, not only when both constituents point to the same kind of answer, but also when the results of each method alone are different. Three such situations are chosen to be presented:

1. a fault on the high-pressure compressor (fault case c), which is solved by the Kalman filter, but not by the BBN
2. an LPT flow-capacity fault (case k), which is not solved by the Kalman filter but is solved by the BBN
3. an LPT fault involving two health parameters where none of the algorithms find the solution

HPC Fault: Case c. This case is dedicated to demonstrate the stability of the combination between the BBN and the Kalman filter. Both algorithms have been separately run on the same data set. The upper graph in Fig. 7 shows results of the diagnostic using the Kalman filter alone. It compares actual values of health parameter related to the high-pressure compressor (dotted lines) to identified values. Identified values are close to actual ones, showing that the identification is effective. No spreading of the fault is observed on the high-pressure turbine (not shown in the figure) nor on the fan and the low-pressure compressor. The lower graph in Fig. 7 shows similar results given by the BBN. Values in the ordinate of this graph are the corresponding mean computed by relation (5). As long as the fault is of small magnitude, namely, for ~4000 s, the BBN produces a wrong diagnosis of an LPC fault (SE2). Only for the last period of the interval, when the deviation magnitude becomes larger, is the correct component indicated.

Results of the combined method are summarized in Fig. 8 showing that the fault is still correctly located by the Kalman

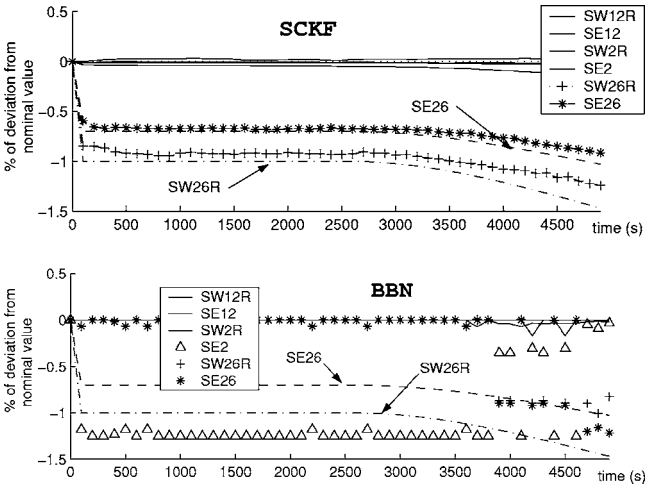


Fig. 7 Identification results of individual methods for HPC fault case c (dotted lines refer to actual values)

filter. This demonstrates that the combined algorithm is not simply a weighted mean of Kalman filter and BBN results. It seems that the SCKF estimation is robust enough not to be perturbed by BBN information.

LPT Faults: Cases k and l. Test case k involves only one health parameter: the flow capacity of the low-pressure turbine (SW49R). This case is solved by the BBN but not by the Kalman filter. This situation is summarized in the upper graph in Fig. 9, which shows results for the turbine and the nozzle using the SCKF alone. Identified health parameters SW49R, SE49, as well as SE41 are detected faulty far from the actual values (dotted lines). Conversely, the BBN is able to locate the fault. The lower

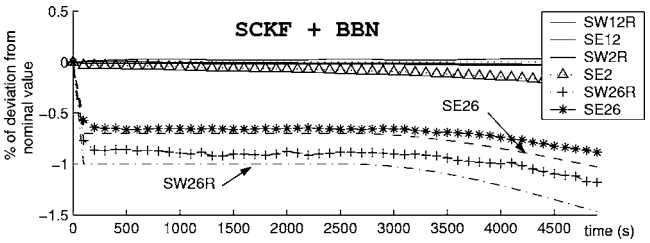


Fig. 8 Identification results of combined method for HPC fault case c (dotted lines refer to actual values)

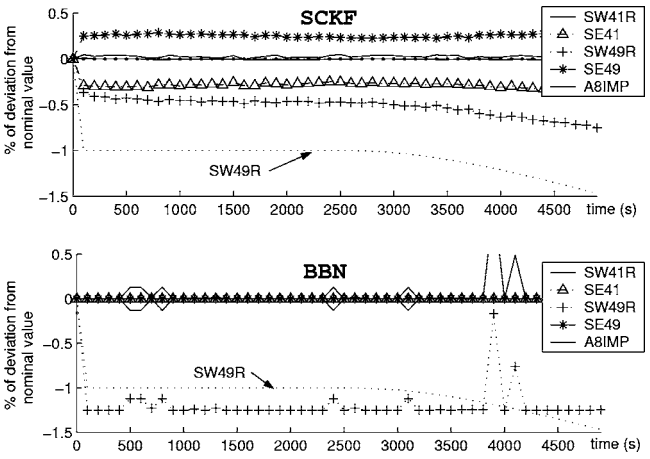


Fig. 9 Identification results of individual methods for LPT fault case k (dotted lines refer to actual values)

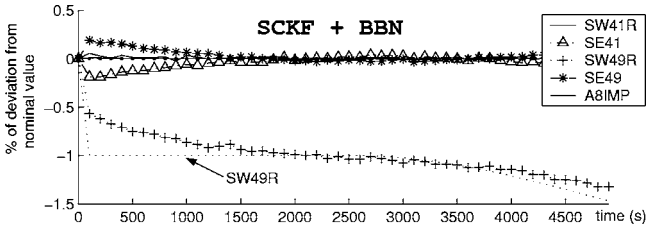


Fig. 10 Identification results of combined method for LPT fault case k (dotted lines refer to actual values)

graph in Fig. 9 indicates that the parameter SW49R is low: mean values of SW49R are around −1%, which is close to the actual value. All other parameters are assigned to their nominal value indicating that the isolation is correct.

The following results (Fig. 10) highlight the benefit of the combination also for the BBN. Identified values related to SW49R are close to the actual one, whereas the one related to SE41 and SE49 remain close to nominal values. In this case the SCKF is driven by the BBN to the correct solution.

As another illustration of this, Fig. 11 shows the results of the test case l where both SW49R and SE49 are involved in the component fault. Values derived by the SCKF alone for SW49R and SE49 remain far from the actual ones, and the difference is spread on the other parameters (SW41R and SE41). The fault is correctly located, but its magnitude is not accurately determined. The combined algorithm identification is far more accurate. The fault is not only located correctly but also accurately assessed.

LPT Fault: Case j. Because no interturbine measurement is available, case j is by far the most difficult one. This case is not solved by the BBN nor the SCKF, which is represented in Fig. 12 where values of SW49R, SE49, and SE41 identified by the SCKF remain far from actual values. Therefore, both low- and high-pressure turbines are detected as faulty. The same kind of results are obtained using the BBN where SW49R and SW41R are classified as faulty, while the fault related to SE49 is not detected.

Although identified values of SW49R and SE49 are closer to the actual ones, the fault remains poorly located (Fig. 13). SE41 is below −0.3%, and the high-pressure turbine also looks defective. This behavior is not surprising since not one of the constituent methods provides information that could be used to lead the combined method to the correct decision.

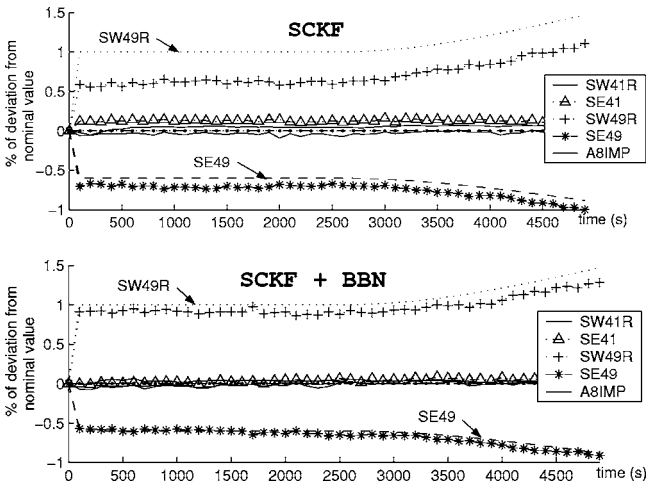


Fig. 11 Identification results of SCKF (upper figure) and combined method (lower figure) for LPT fault case l (dotted lines refer to actual values of health parameters)

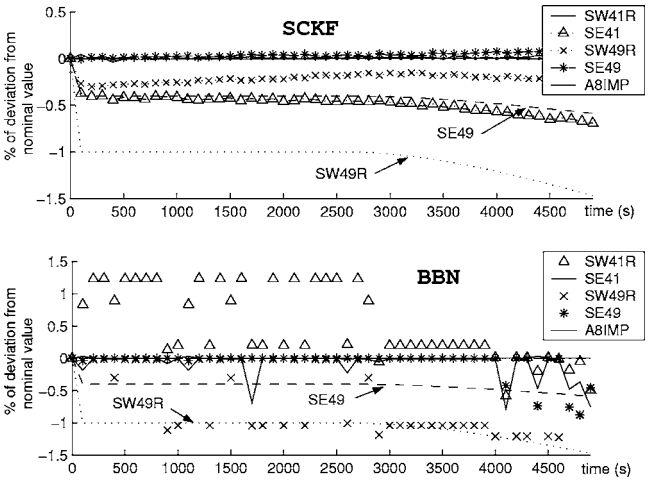


Fig. 12 Identification results of individual methods for LPT fault case j (dotted lines refer to actual values)

Diagnostic Effectiveness Overview. In order to illustrate the capabilities of the method, application results related to a number of different fault cases are shown. Table 3 compares the results of the soft-constrained Kalman filter (SCKF) alone to the combination of the SCKF with the BBN. This table shows maximum absolute value of biases $\Delta\mathbf{w}_k$ achieved by both methods after 4900 measurement samples are observed, which represents an image of the asymptotic efficiency of the method.

$$\Delta\mathbf{w}_k = 100\frac{\hat{\mathbf{w}}_k - \mathbf{w}_k^{\text{actual}}}{\mathbf{w}_k^{\text{nominal}}} \tag{7}$$

Values of $\Delta\mathbf{w}_k$ around 0.25% are considered acceptable. The main conclusion of these results is that the combination especially improves results of the Kalman filter in test cases j, k, and l related

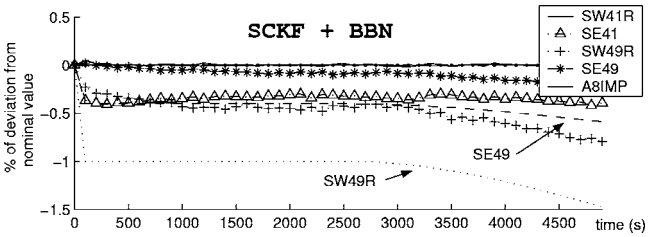


Fig. 13 Identification results of combined method for LPT fault case j (dotted lines refer to actual values)

Table 3 Comparison of identification results using soft-constrained Kalman filter with noninformative a priori (left column) and BBN a priori (right column). Results are maximum absolute values of biases defined by (7)

	SCKF alone (%)	SCKF+BBN(%)
a	0.03 on SW49R	0.05 on SW49R
b	0.09 on SE12	0.10 on SE12
c	0.07 on SW26R	0.12 on SW26R
d	0.03 on SE12	0.06 on SE2
e	0.03 on SW26R	0.04 on SW49R
f	0.02 on SE42	0.02 on SE42
g	0.17 on SW49R	0.11 on SE49
h	0.24 on SW49R	0.14 on SW49R
i	0.19 on SW49R	0.25 on SW49R
j	0.82 on SW49R	0.58 on SW49R
k	0.53 on SW49R	0.04 on SW49R
l	0.41 on SW49R	0.12 on SW49R
m	0.06 on SE41	0.05 on SE41
n	0.06 on SW2R	0.03 on SE49

Table 4 Summary of diagnosis success given by the SCKF alone, BBN alone, and combination of the SCKF with BBN a priori for the complete set of component fault detailed in Table 2

	SCKF	BBN	SCKF+BBN
a	√	√	√
b	√	√	√
c	√	-	√
d	√	√	√
e	√	√	√
f	√	√	√
g	√	√	√
h	√	√	√
i	√	√	√
j	-	-	-
k	-	√	√
l	-	√	√
m	√	√	√
n	√	√	√

to the low-pressure turbine.

Table 4 gives an overall picture of the efficiency and the gain achieved by the combination is compared to SCKF and BBN working separately. This combination is able to solve all the test cases except the case j, which is known to be difficult to identify with this set of seven measurements [7].

Discussion

The present recursive method allows one to identify component faults based on associated measurement deviations whose amplitude is around the standard deviation of measurement noise. The signal-to-noise ratio mainly affects the identification process in terms of convergence speed. The dependency of the convergence speed on the measurement noise level is illustrated in Fig. 14. Crosses represent the relative increase of time required to accumulate enough data samples as needed to converge to the asymptotic solution. The crosses compare well to the theoretical prediction (plain line), which states that the number of data samples required to obtain a given accuracy must increase as a quadratic function of the noise level. Indeed, more accurate measurements (less noise) means more adaptability (since the Kalman gain **K** is higher), an improved tracking capability, and, therefore, an increased convergence speed.

Moreover, systematic calculations have shown that the final diagnostic (asymptotic solution) is only slightly affected by the level of noise in the measurements. Beside these effects of the noise level, the nonobservability of some parameters can lead to a false diagnostic even though the noise level is strongly reduced.

With the set of seven available measurements, those results are the most meaningful that can be obtained with the combined algorithm. In order to obtain a better location of the fault, additional

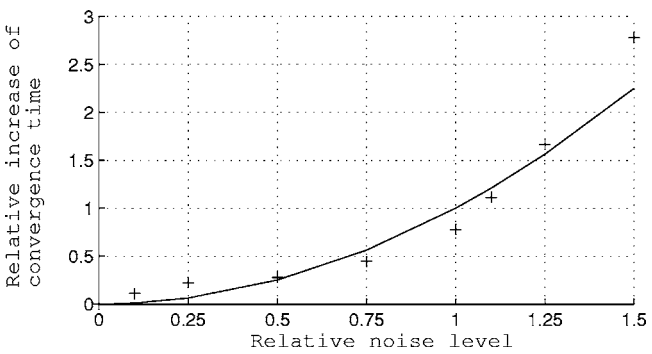


Fig. 14 Influence of the measurement noise level on the convergence speed

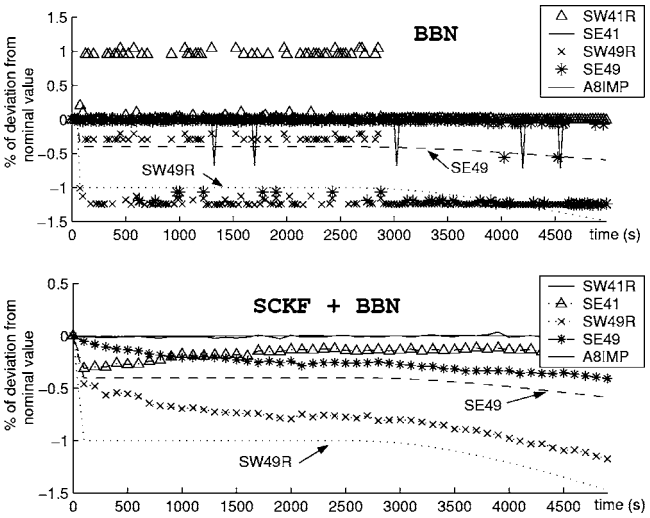


Fig. 15 Identification results of BBN alone (upper figure) and combined method (lower figure) for LPT fault case j (dotted lines refer to actual values)

knowledge must be made available. In [8], two additional measurements are considered: P26 and P49. Although this solution is ideal because it provides the best results, those two measurements may not be available. In a test bench configuration it may happen that some other measurements are available (i.e., vibrations, some measurements about the lubrication system, etc.) but are unlikely to be predicted based on the health parameters since they do not appear in the model. This underlines the weakness of the Kalman filter for which a model has to be available. Any qualitative knowledge is very difficult to include in the identification procedure.

The combined method proposed in the present paper allows the inclusion of additional information from sources other than modeled measurements (information fusion). It can thus produce results that would have been possible for the Kalman filter only with additional gas path measurements. This point is demonstrated by considering again case j, but assuming now that some additional information exists, such as historical records and related statistics. The following test case assumes that such additional knowledge favors a fault on the low-pressure turbine. The BBN is modified by tuning the a priori knowledge about the health parameters (see [9] for the detailed procedure) to make a fault on the low-pressure turbine more likely to occur.

Results using this "modified" BBN are shown in Fig. 15. A fault on SE49 is now detected in addition to the one related to SW49R. Results of the combined algorithm (lower graph in Fig. 15) are far better: SW49R as well as SE49 converge to their actual value, while SE41 is closer to its true value. Finally (after 2500 s), the detection is effective and the health parameters are accurately assessed. The diagnosis is effective and allows a more reliable decision.

Conclusions

A new methodology combining classification methods and regression methods has been developed in order to benefit of their mutual advantages. This combination has been tested on cases representative of a real-life application to underline the gain in stability and accuracy that can be achieved. The number of undetected faults are lowered and false alarms are avoided (Table 4) when compared to regression or classification methods working separately.

The estimation of the combined algorithm allows a more reli-

able detection of component faults but also achieves a better accuracy and a better fault isolation by lowering the spread of the fault on several parameters ("smearing effect").

Besides the improvements in accuracy and stability, this kind of method allows information or sensor fusion, which is a very important field of research for future works. The key advantage of combining methods is that it replaces the problem of comparing classification techniques to regression techniques by the problem of choosing which information they can share.

Acknowledgment

The paper has been produced in the framework of a collaboration between the ASMA Department of the University of Liege (Belgium) and the Thermal Turbomachinery Laboratory (LTT) of the National Technical University of Athens (Greece). This work is funded by the Walloon Region of Belgium in its Program FIRST and supported by Techspace Aero, Snecma Group.

Nomenclature

- BBN = Bayesian Belief Network
- $\mathbf{D}$ = covariance matrix of $\mathbf{w}_{\text{prior}}$
- $\mathcal{G}$ = nonlinear engine performance model
- k = time index (discrete time)
- KF = Kalman filter
- $\bar{\mathbf{r}}_k$ = residuals on raw measurements $\bar{\mathbf{r}}_k = \bar{\mathbf{y}}_k - \hat{\mathbf{y}}_k$
- $\mathbf{R}_r$ = covariance matrix of $\bar{\mathbf{r}}_k$
- SCKF = soft-constrained Kalman filter
- t = continuous time
- $\mathbf{u}_k$ = command parameters
- $\mathbf{w}_k$ = engine health parameters
- $\mathbf{w}_{\text{prior}}$ = a priori value of health parameters
- $\hat{\mathbf{w}}_k$ = estimated engine health parameters
- $\mathbf{x}_k$ = engine operating point
- $\hat{\mathbf{x}}_k$ = estimated engine operating point
- $\mathbf{y}_k$ = measurements involved in health parameter estimation
- $\bar{\mathbf{y}}_k$ = raw measurements $\mathbf{y}_k$ from data acquisition system
- $\hat{\mathbf{y}}_k$ = estimation of measurements $\mathbf{y}_k$
- $\mathbf{z}_k$ = measurements involved in operating point estimation
- $\bar{\mathbf{z}}_k$ = raw measurements $\mathbf{z}_k$ from data acquisition system

References

[1] Bishop, C. M. 1995, *Neural Networks for Pattern Recognition*, Clarendon Press, Oxford.

[2] Volponi, A., 2003, "Foundation of Gas Path Analysis (Part i and ii)," *Von Karman Institute Lecture Series: Gas Turbine Condition Monitoring and Fault Diagnosis, (2003-01)*.

[3] Provost, M. J., 2003, "Kalman Filtering Applied to Gas Turbine Analysis," *Von Karman Institute Lecture Series: Gas Turbine Condition Monitoring and Fault Diagnosis, (2003-01)*.

[4] Kobayashi, T., and Simon, D. L., 2003, "Application of a Bank of Kalman Filters for Aircraft Engine Fault Diagnostics," ASME Turbo Expo, ASME Paper No. GT2003-38550.

[5] Aretakis, N., Mathioudakis, K., and Stamatis, A., 2002, "Non-linear Engine Component Fault Diagnosis From a Limited Number of Measurements Using a Combinatorial Approach," ASME Turbo Expo, ASME Paper No. GT2002-30031.

[6] Simon, D., and Simon, D. L., 2003, "Aircraft Turbofan Engine Health Estimation Using Constrained Kalman Filtering," ASME Turbo Expo, ASME Paper No. GT2003-38584.

[7] Grodent, M., and Navez, A., 2001, "Engine Physical Diagnosis Using a Robust Parameter Estimation Method," *37th AIAA/ASME/SAE/ASEE Joint Propulsion Conference*.

[8] Dewallef, P., Mathioudakis, K., and Léonard, O., 2004, "On-Line Aircraft Engine Diagnostic Using a Soft-Constrained Kalman Filter," ASME Turbo Expo, ASME Paper No. GT2004-53539.

[9] Romessis, C., and Mathioudakis, K., 2004, "Bayesian Network Approach for Gas Path Fault Diagnosis," ASME Turbo Expo, ASME Paper No. GT2004-53801.

[10] Mathioudakis, K., 2003, "Neural Networks in Gas Turbine Fault Diagnosis," *Von Karman Institute Lecture Series: Gas Turbine Condition Monitoring and Fault Diagnosis, (2003-01)*.

[11] Brotherton, T., Volponi, A., Luppold, R., and Simon, D. L., 2003, "Estorm: Enhanced Selft Tuning On-Board Real-Time Engine Model," *2003 IEEE Aerospace Conf.*

[12] Moody, J. E., 1992, "The Effective Number of Parameters: An Analysis of Generalization and Regularization in Nonlinear Learning Systems," *Advances in Neural Information Processing System 4*, Morgan Kaufmann, San Mateo, CA.

[13] MacKay, D. J. C., 1995, "Bayesian Methods for Neural Networks: Theory and Application," Tech. Report, Cavendish Laboratory, University of Cambridge, http://wol.ra.phy.cam.ac.uk

[14] Curnock, B., 2000, Obidicote Project-Word Package 4: Steady-State Test Cases, Tech. Report DNS62433, Rolls-Royce.

[15] Stamatis, A., Mathioudakis, K., Ruiz, J., and Curnock, B., 2001, "Real-Time Engine Model Implementation for Adaptive Control and Performance Monitoring of Large Civil Turbofans," ASME Turbo Expo, ASME Paper No. GT2001-362.

Transient Analysis of Solid Oxide Fuel Cell Hybrids—Part I: Fuel Cell Models

Loredana Magistri

Francesco Trasino

Paola Costamagna

Dipartimento di Macchine Sistemi Energetici e Trasporti,
Thermochemical Power Group, DiCHEP,
Università di Genova,
Genova 16145,
Italy

The main goal of this work is the transient analysis of hybrid systems based on solid oxide fuel cells (SOFC). The work is divided into three parts: in the first, the fuel cell transient models are presented and discussed, whereas in the subsequent parts of the paper the anodic recirculation system (Part B: Ferrari, M.L., Traverso, A., Massardo, A.F., 2004, ASME Paper No. 2004-GT-53716) and the entire hybrid transient performance (Part C: Magistri, L., Ferrari, M.L., Traverso, A., Costamagna, P., Massardo, A.F., 2004, ASME Paper No. 2004-GT-53845) are investigated. In this paper the transient behavior of a solid oxide fuel cell is analyzed through the use of two different approaches: macroscopic and detailed SOFC models. Both models are presented in this paper, and their simulation results are compared to each other and to available experimental data. As a first step the transient response of the fuel cell was studied using a very detailed model in order to completely describe this phenomenon and to highlight the critical aspects. Subsequently, some modifications were made to this model to create an apt simulation tool (macroscopic fuel cell model) for the whole plant analysis. The reliability of this model was verified by comparing several transient responses to the results obtained with the detailed model. In the subsequent papers (Parts B and C), the integration of the macroscopic fuel cell model into the whole plant model will be described and the transient study of the hybrid plant will be presented.
[DOI: 10.1115/1.2056529]

Introduction

Although there have been continuous improvements in hybrid system technology, its current status still requires in-depth studies for assessing the cycle performance and understanding the safe operative limits of plants. Similarly, the transient analysis is essential to implement the control system and to define the safe operative limits.

The transient studies allow a deep investigation of the fuel cell stack and plant behavior. Dangerous operating conditions, such as steep fuel cell temperature gradient or surge problems in the air compressor, are highlighted. In this way it is possible to define the start-up and shutdown procedures, avoiding risks for the whole plant.

In order to achieve these results reliable simulation models are required. All the components of the plant (such as fuel cell, gas turbine, ejector, etc.) must be studied in as much detail as possible without neglecting the requirements of the system simulations in terms of flexibility, computational time, and geometrical input data requested. Moreover, the whole model should be easily adaptable to different plant configurations.

This paper deals with fuel cell transient simulation; here the fuel cell model and the time response are described and analyzed. The fuel cell has been considered as a stand-alone unit in order to understand its behavior without taking into account the influence of the other components of the plant. The integration of the fuel cell model and system model and the study of the transient behavior of the whole plant is the object of Part II [1] and Part III [2] of this work.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. 2004-GT-53842.

Solid Oxide Fuel Cell Models (SOFC)

In this paper the cell geometry taken into consideration is tubular, with a design similar to the one proposed in [3]; a simplified scheme of the fuel cell and of some surrounding facilities, such as the reformer, the preheater, etc., is reported in [4]. It is important to point out that the dimensions of the SOFC under consideration are typical of the precommercialization phase currently in progress (for example, the cell length is 1.5 m). The cell is considered to be coupled to a reforming unit, which feeds the SOFC with a mixture of H_2, CH_4, CO, CO_2, H_2O. As a consequence, chemical reactions, such as shifting and methane reforming, occur in the anodic compartment of the SOFC.

Two cell models have been developed and are presented; since the equations of the simplified model are easily explained by considering them as a simplification of the detailed model, the latter model will be presented first. Even if this model addresses a single cell, we can still use it to draw conclusions about the simulation of the overall stack to be included in the hybrid systems (HS). However, the passage from the single cell to the stack size cannot simply be made by assuming that all the cells of the stack show a behavior identical to the single cell operated under adiabatic conditions. Indeed, the effects of heat radiation between contiguous cells in the stack, which tends to smooth the temperature profile along each cell, cannot be neglected. The behavior of the whole stack has been analyzed by considering a series of single cells undergoing the same dynamics. The radiation effects have been taken into account in a simplified way. As the details of the internal geometry of a tubular SOFC stack are not always available for academic work, radiation effects have been considered by increasing the thermal conductivities of the SOFC materials.

With the aim of including the transient fuel cell model in a whole hybrid plant transient model [2], the stack has been simulated in the simplest way compatible with good accuracy of the results. Starting from a very detailed model for a stand-alone single cell (single cell model), described in depth in [4], some changes have been made in order to study the stack behavior

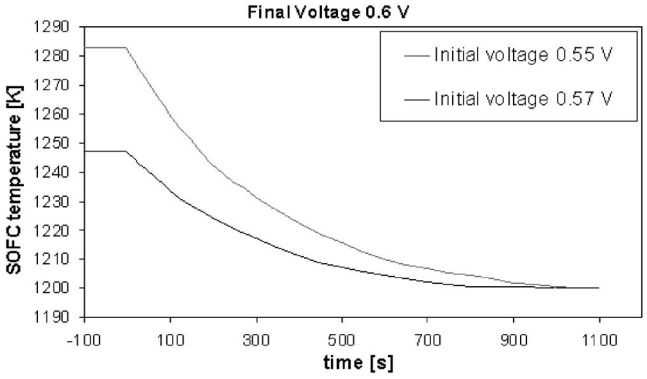


Fig. 1 Temperature profile due to instantaneous voltage increase obtained with single cell model

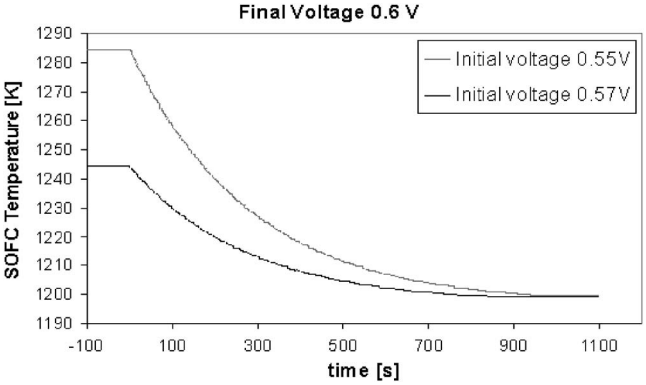


Fig. 3 Temperature profile due to instantaneous voltage increase obtained with stack model

(stack model), reduce the computational time, and simplify the integration with the plant code that includes the dynamic simulation of the gas turbine, anodic ejector, combustion chamber, etc. As a first step the results of the simplified dynamic fuel cell model have been compared to the ones obtained with the previous model as depicted in Fig. 1–4.

Single-Cell and Stack-Model Analogies. The main hypotheses of both models are:

- The cell is adiabatic.
- The cell voltage is uniform.
- All the chemical reactions within the anodic stream are at equilibrium.
- The electrochemical reaction of H_2 is taken into consideration while the electrochemical reaction of CO is neglected.

The cell models include the following:

- *Electrochemical performance*: the current-voltage behavior of the stack is evaluated by subtracting the overall voltage losses from the thermodynamic potential (i.e., the voltage of the stack under conditions of thermodynamic reversibility, Eq. (1) (Eqs. (2)–(6))). In our model, we simulate both ohmic losses (Eqs. (2)–(4)) and activation losses due to sluggish electrode kinetics (Eqs. (5)–(6)). In this way, Eqs. (7)–(10) allow the evaluation of the electrical power supplied by the cell and of the electrochemical reaction rate r.
- *Equilibrium of reforming and shifting chemical reactions,* which is made through Eqs. (11) and (12). These equations allow the calculation of the reaction rates r of reforming and shifting.
- Mass balances of anodic and cathodic flows (Eq. (13)), taking into account that molar flow rate variations are due to the reaction rates r of both chemical and electrochemical reactions.
- *Energy balances of gaseous flows,* including the anodic gas (Eq. (15)), the air flowing within the quartz tube (Eq. (16)) and the cathodic gas (air) flowing into the annular space between the quartz tube and the cathode (Eq. (17)). The equations include the convective exchanges due to mass transfer with the solid (only for the anodic and cathodic gas flow), and the heat exchanges between the gases and the solid surfaces (either the inner tube or the electrochemical cell itself). The evaluation of the heat exchange coefficient is made on the basis of the Nusselt number, assuming that the flow regimen is laminar.
- *Energy balance of the tube (Eq. (18)) and of the solid PEN* (positive(P)-electrolyte(E)-negative(N)) structure (Eq. (19)), including forced convection between the solid and gases, in-plane conduction, and, in Eq. (19), the dissipative heat sources due to chemical and electrochemical reactions. The thermal balance of the solid results in a one-dimensional Fourier problem with distributed heat sources.

Input data are geometric characteristics, operating conditions, inlet flow conditions, and gas and material properties.

Single-Cell and Stack-Model Differences. In the single cell model the evaluation of the electrical performance is made locally along the cell length: indeed, in order to integrate the model, the cell is divided into a number of calculation points, and the evaluation of the electrical performance is repeated in each of them. The differential equations of the model have been integrated using a relaxation method for the Fourier problem and the charge balance, combined with a finite difference method to solve the other

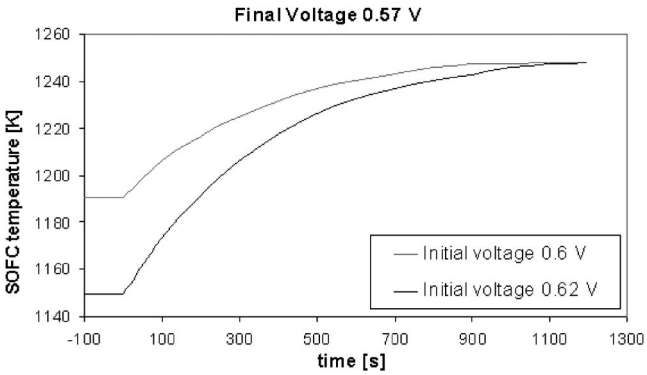


Fig. 2 Temperature profile due to instantaneous voltage reduction obtained with single cell model

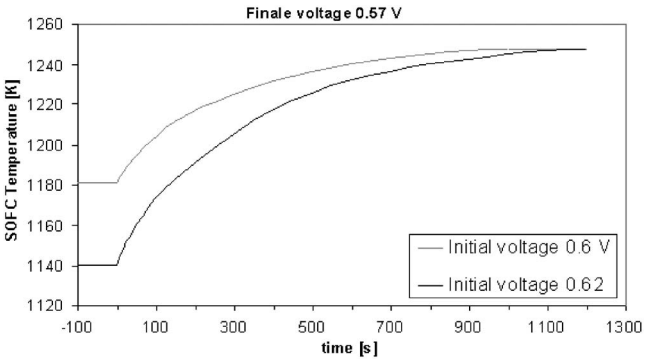


Fig. 4 Temperature profile due to instantaneous voltage reduction obtained with stack model

Table 1 Models equations

	DETAILED AND SIMPLIFIED	
Electrochemical Reaction	$V_{nr}=\frac{-\Delta G}{2F}=\frac{-\Delta G^{*}}{2F}+\frac{R_{t}T}{2F}\ln\frac{p_{H_2}\cdot(p_{air})^{\gamma}}{p_{H_2O}}$ (1)	
	$\rho_{a}=A_{a}\exp\left(-\frac{B_{a}}{T}\right)$ (2) $\rho_{c}=A_{c}\exp\left(-\frac{B_{c}}{T}\right)$ (3) $\rho_{e}=A_{e}\exp\left(-\frac{B_{e}}{T}\right)$ (4)	
	$\frac{1}{\Omega_{max}}=D_{1}\frac{2F}{R_{t}T}\left(\frac{p_{H_2}}{p^{*}}\right)^{m_{1}}e^{-(\alpha_{1}\cdot nF)}$ (5) $\frac{1}{\Omega_{max}}=D_{2}\frac{4F}{R_{t}T}\left(\frac{p_{air}}{p^{*}}\right)^{m_{2}}e^{-(\alpha_{2}\cdot nF)}$ (6)	
	$\Omega_{int}=\sum\Omega_{max}+\sum\Omega_{min}+\Omega_{ohm}$ (7) $V_{a}=\Omega_{int}I$ (8) $V_{u}=V_{int}-V$ (9) $W_{e}=VI$ (10)	
Equilibrium Chemical Reactions	$K_{eq,ref}=\frac{p_{H_2}^{3}p_{CO}}{p_{CH_4}p_{H_2O}^{2}}$ reforming (11) $K_{eq,shift}=\frac{p_{CO_2}p_{H_2}}{p_{CO}p_{H_2O}}$ shifting (12)	
	DETAILED MODEL	SIMPLIFIED MODEL
Mass Balance	$\frac{dN}{dx}=S\sum_{r}\nu_{i,r}r_{r}$ (13)	$N_{i}^{out}=N_{i}^{in}+SI\sum_{r}\nu_{i,r}r_{r}$ (14)
Energy Balance	$\frac{1}{S_{a}}\left[\sum N_{i}\cdot c_{p,i}\cdot\frac{dT_{a}}{dx}+\sum c_{p,i}\cdot\frac{dN_{i}}{dx}(T_{a}-T_{i})\right]+\frac{h_{a}}{b_{a}}(T_{a}-T_{e})=0$ (15)	
	$\frac{1}{S_{ae}}\sum N_{i}\cdot c_{p,i}\cdot\frac{dT_{ae}}{dx}+\frac{h_{a-ae}}{b_{ae}}(T_{ae}-T_{a-ae})=0$ (16)	
	$\frac{1}{S_{a}}\left[\sum N_{i}\cdot c_{p,i}\cdot\frac{dT_{a}}{dx}+\sum c_{p,i}\cdot\frac{dN_{i}}{dx}(T_{a}-T_{i})\right]+\frac{h_{a}}{b_{a}}(T_{a}-T_{e})+\frac{h_{a-ae}}{b_{a}}(T_{a}-T_{a-ae})=0$ (17)	$T^{out}=\frac{\sum N_{i}^{in}C_{p,i}T^{in}+SI\sum r_{r}(-\Delta H_{r})-W_{e}}{\sum N_{i}^{out}C_{p,i}}$ (20)
	$\frac{1}{s_{a-ae}}\cdot\left[h_{a-ae}(T_{ae}-T_{a-ae})+h_{a-ae}(T_{a}-T_{a-ae})-k_{a-ae}\frac{d^{2}T_{a-ae}}{dx^{2}}\right]=0$ (18)	
	$\frac{1}{s_{e}}\cdot\left[h_{e}(T_{a}-T_{e})+h_{e}(T_{ae}-T_{e})+Wi+\frac{1}{mI}\sum H_{i}\frac{dF_{i}}{dx}\right]-k_{e}\frac{d^{2}T_{e}}{dx^{2}}=0$ (19)	

equations.

In the stack model, the electrical current-voltage performance is evaluated following the same procedure described in the previous section, with the difference that, in the present case, the evaluation of the electrical performance is made only once, based on the average values of the physical-chemical variables. In the simplified model, the other equations are in the form of macroscopic balances, i.e., finite equations, in a form that is the integration of the differential equations between the inlet and outlet of the fuel cell. Indeed, these equations simply express a balance between inlet and outlet flows of mass and energy, and only allow the evaluation of the average values of the physical-chemical variables within the electrochemical reactor.

As with the detailed model, the simplified model is also integrated through a relaxation method for the evaluation of the cell performance at the design point. This kind of approach is also useful for the investigation of the system transients because of its temporal step conception that helps to trace the evolution of the system until steady-state conditions are reached using

$$\frac{d(T_{\text{solid}})_i}{dt}=\frac{(q_{\text{solid}})_i}{(C_{\nu\_\text{solid}})_i} \tag{1}$$

A complete presentation of the model equations is given in [4] and in Table 1.

Assessment of Stack Model Results

The stack model can be used to generate steady-state characteristic curves, such as current-voltage and current-power curves, with different air and fuel flows at different operating temperatures. The steady-state results are in good agreement with literature data [5].

Comparison of Single-Cell and Stack-Model Results. Figures 1 and 3 show the temperature profile due to the instantaneous step (increment) of the operating voltage obtained with the two models (single-cell and stack models). Similar results are shown in Figs. 2 and 4 where an instantaneous reduction of the voltage is considered.

In practice it is possible to reach similar conditions applying a potentiostatic control strategy [6]. In both cases the final state is reached starting from different initial conditions in terms of voltage and current density and, consequently, cell temperature. The relaxation time, temperature profile, and type of transient obtained with each model are similar in each case considered and, as the figures show, the relaxation time is independent of the particular magnitude of the current density step [7]. Obviously the values are not exactly the same because of the nonlinearity of the equation systems that lead the thermal and electrochemical phenomena.

In fact, the stack code, as already said, is a single node model (lumped volume) and the differences in the evaluation of the temperature profile come from using the single-node approach instead of using the multinode one of the single-cell model. The nonlinearity of the system implies different average values of some parameters (such as temperature) when evaluated with a different numbers of calculation nodes [8]. However, in this case, the difference between the results obtained with the two models are negligible, especially considering that the final aim of the stack model is a definition of the hybrid plant control strategy [1,2].

Solid Oxide Fuel Cell Transient Behavior

The analysis of the SOFC transient behavior has been carried out with the previously described stack model, complete with reformer and postcombustor models. In an SOFC, the thermal transient involves larger time constants compared to the electrochemical and electrical time response as analyzed in [6]. In fact, the electrochemical response is nearly instantaneous, and the electrical transient, due to the variation in reactant concentration along the cell, is normally a fraction of a second (cell length/fuel velocity) [6]. This great difference between the three time constants allows the use of a lumped volume model for the dynamic analysis of the SOFC, thus, leading only to the evaluation of the thermal time scale.

In all the cases considered in Figs. 5–16 the composition of the inlet fuel has been kept constant and the fuel supply system response has been considered instantaneous. Two different kinds of transients have been investigated: decrease of fuel flow rate (Figs. 5–10) and increase of fuel flow rate (Figs. 11–16).

The modification of the fuel flow rate is related to a simultaneous modification of the electrical load. In this way it has been

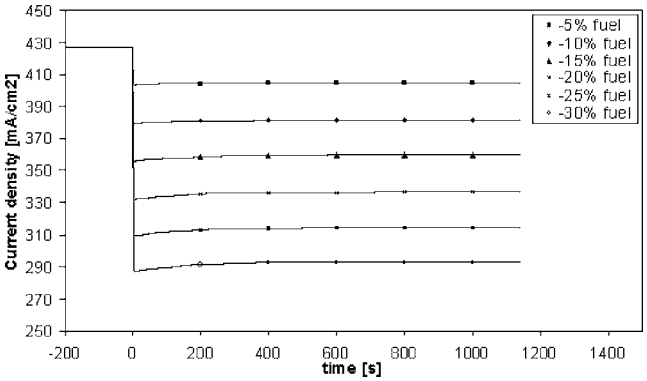


Fig. 5 Current density distribution due to instantaneous fuel flow-rate decrease

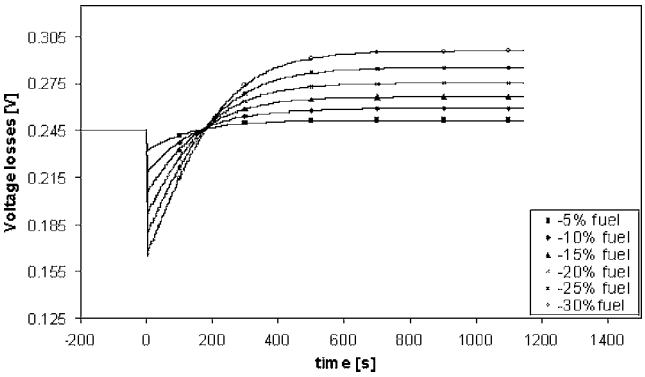


Fig. 8 Overpotential profiles due to instantaneous fuel flow-rate decrease

considered as a "load follow device," where the current density is supposed to be controlled by the electrical conditioning system, while the fuel flow rate is scheduled in order to keep the utilization factor constant. This approach has been considered in order to understand the transient phenomena in the stand-alone stack; different control strategies are under consideration where the stack is part of the whole hybrid system, as considered in Part III [2] of this work.

Electrical Load Decrease. Figures 5–10 show the transient behavior of different stack parameters, such as current density (Fig. 5), stack temperature (Fig. 6), total electrical resistance (Fig. 7), voltage losses (Fig. 8), theoretical cell voltage (Fig. 9), and real voltage (Fig. 10). As shown in Fig. 5, the variation of the current density, led by the electrochemical response, is almost instantaneous, whereas the thermal transient (Fig. 6) appears longer. In both cases the fuel variation affects only the final condition reached by the stack, not the transient characteristic time.

Figure 7 shows the strong influence of the thermal transient on the total stack resistance; in fact, a decrease in fuel lowers the stack temperature and as a consequence increases the electrical resistance. The opposite variation of electrical resistance and current density leads to the voltage losses behavior as shown in Fig. 8. After a fuel-current drop the operating voltage shows a steep increase due to the instantaneous fall of the polarizations. This drop of the polarization losses results in the beginning of the transient where the current density has been decreased but the temperature of the stack is still high, due to the thermal inertia of

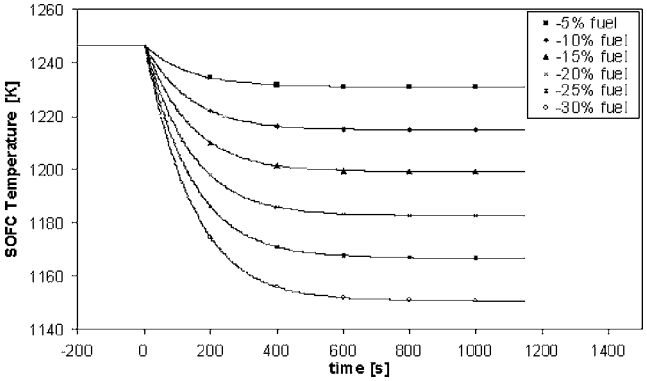


Fig. 6 Temperature profiles due to instantaneous fuel flow-rate decrease

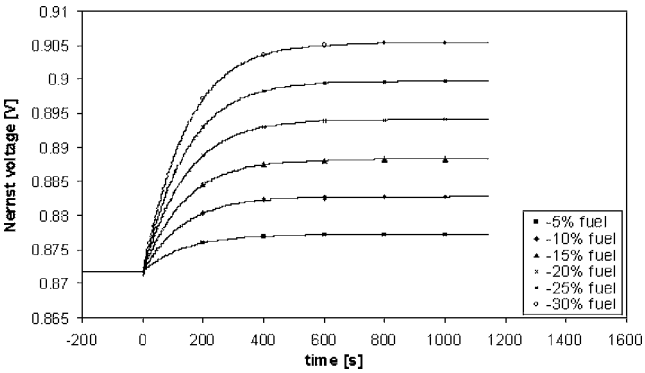


Fig. 9 Nernst voltage behavior due to instantaneous fuel flow-rate decrease

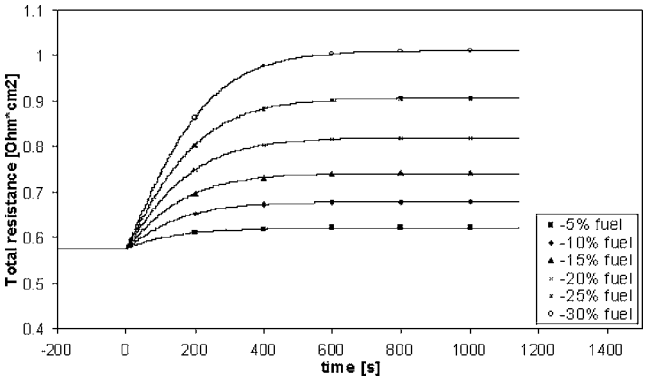


Fig. 7 Total resistance behavior due to an instantaneous fuel flow-rate decrease

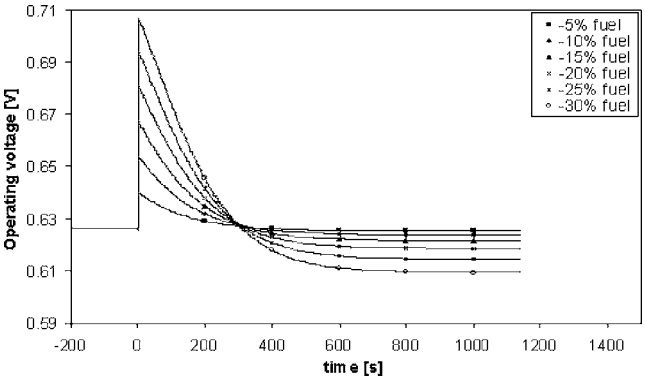


Fig. 10 Operating voltage profiles due to instantaneous fuel flow-rate decrease

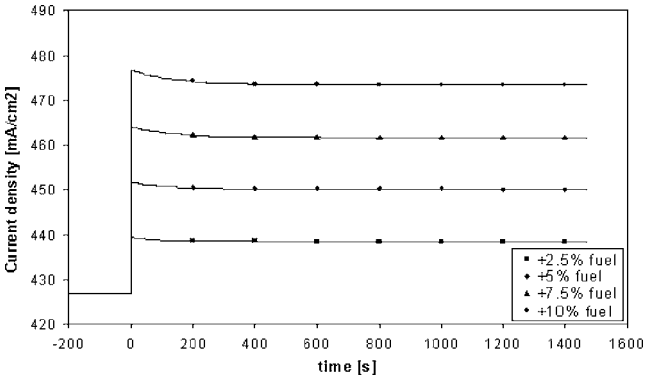


Fig. 11 Current density distribution due to instantaneous fuel flow-rate increase

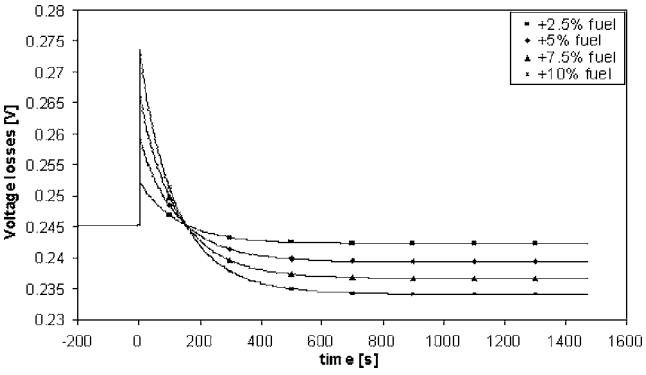


Fig. 14 Overpotential profiles due to instantaneous fuel flow-rate increase

the stack.

After this first drop the effect of the lowered stack temperature prevails over the positive effect of the decreased current on the polarization, resulting in an increase in the overpotential (Fig. 8) and a consequent reduction in the operating voltage (Fig. 10); obviously the larger the initial current decrease the larger the initial voltage spike.

In all the cases studied here, the relaxation time (defined as the time necessary to recover 90% of the dynamic parameter drop resulting from the difference $P(t\rightarrow\infty)-P(t=+0)$) is not dependent on the initial current density-fuel step.

Electrical Load Increase. In a similar way Figs. 11–16 show the effect of a load (i.e., fuel) increase on the stack transients. The fuel variation was limited to +10% to take into account constraints on the distribution of the cell temperature. In this case the behavior of the different stack parameters is the reverse of that obtained for a load decrease, but nonetheless, the characteristic time is not dependent on the initial current density-fuel step and is in fair agreement with that obtained in the previous case.

Heat Capacity Influence. The relaxation time depends on the fuel cell material properties, mass, and configuration of the solid structure [7]. A comparison of Figs. 10 and 17 shows the influence of the heat capacity K(kJ/K) of the stack on the relaxation time, the initial conditions, and current-density step being the same in both figures and K in Fig. 17 being 75% of the value in Fig. 10. The resulting effect is a shortening of the relaxation time, whereas an increase in K would lead to its increase.

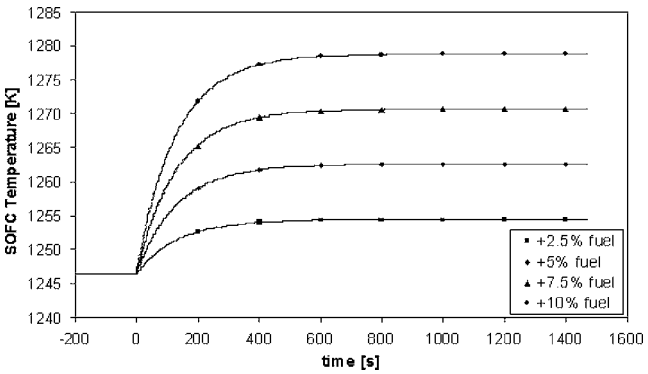


Fig. 12 Temperature profiles due to instantaneous fuel flow-rate increase

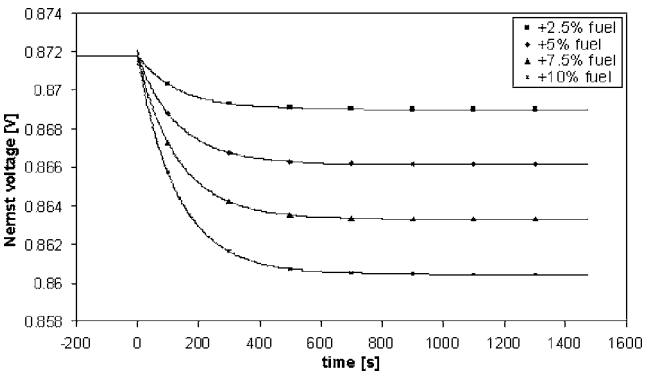


Fig. 15 Nernst voltage behavior due to instantaneous fuel flow-rate increase

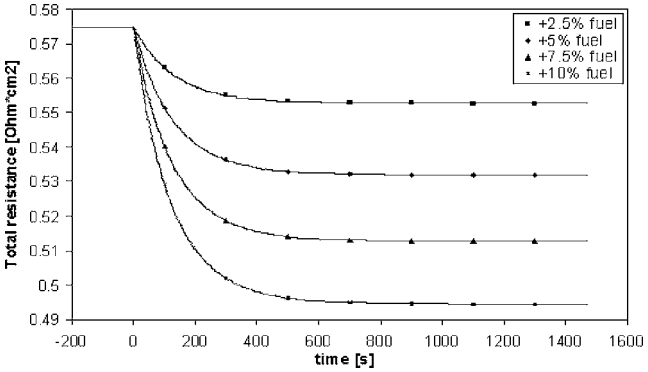


Fig. 13 Total resistance behavior due to an instantaneous fuel flow-rate increase

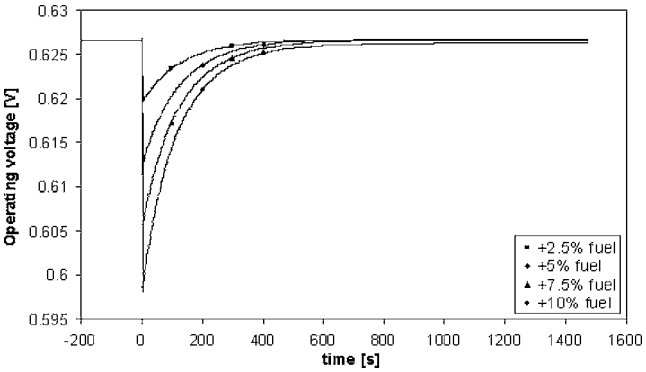


Fig. 16 Operating voltage profiles due to instantaneous fuel flow-rate increase

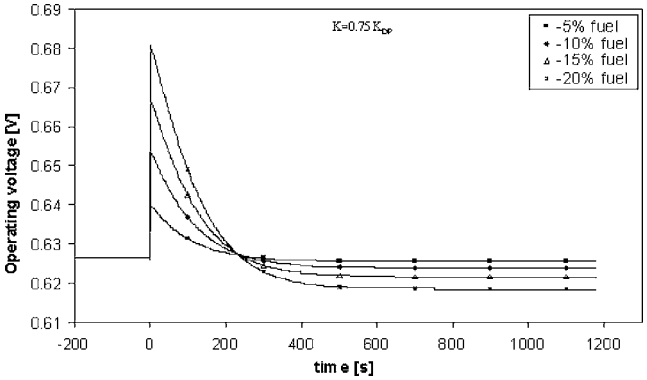


Fig. 17 Operating voltage profiles due to instantaneous fuel flow-rate decrease

Conclusions

The main goal of this work is the transient analysis of a pressurized hybrid system based on solid oxide fuel cell technology for further system studies. The work is divided in three parts: Part A (this paper) is devoted to cell model presentation; in Part II [1], the interaction of the stack and anodic recirculation system is carefully investigated; in Part III [2], the whole hybrid system transient response, including the turbomachinery influence, is analyzed.

In this part, two models for stack transient investigation have been presented and used for validation. One of the main goals is the possibility of developing an apt model to be used in the whole system analysis with minimum time calculation requests and a good description of the transient behavior.

Different load follow transients of the stack have been presented just to show the capability of the model that will also be used in Parts II and III [1,2]. The influence of several parameters has been studied and discussed.

Acknowledgment

This work has been partially sponsored by the project MIUR-FISR 2000, national coordinator A. F. Massardo.

Nomenclature

A, B = coefficients in Eqs. (2)–(4), Ω m, K
B = hydraulic diameter, m
C_p = molar specific heat, J mol^{-1} K^{-1}
D_1, D_2 = coefficients, A
d = cell diameter, m
E = activation energy, J mol^{-1}
F = Faraday constant, C mol^{-1}
ΔG = Gibbs free energy change, J mol^{-1}
ΔG° = standard Gibbs free energy change, [J mol^{-1}]
H = enthalpy, J mol^{-1}
h = heat transfer coefficient, W m^{-2} K^{-1}
I = electrical current, A
l = cell length, [m]
N = molar flow rate, mol s^{-1}
$K_{p,\mathrm{ref}}$, $K_{p,\mathrm{shift}}$ = equilibrium constants, Pa^2, –
k = thermal conductivity of the solid, W m^{-1} K^{-1}
p = partial pressure, Pa
p° = reference pressure, Pa
P = overall pressure, Pa
r = reaction rate, mol m^{-3} s^{-1}
R_g = gas constant, J mol^{-1} K^{-1}
S = cross section, m^2
s = thickness, m
T = temperature, K
V_{oc} = thermodynamic electrical potential, V
V_p = loss of electrical potential, V
V = electrical potential of stack, V
W_{el} = electrical power, W
x = cell longitudinal coordinate, -

Greek Letters

ν = stoichiometric coefficient, -
Ω = electrical resistance, Ω
ρ = electrical resistivity, Ω^{-1}

Subscripts

a, c, el, = anode, cathode, electrolyte
air = air in the quartz tube
i = chemical component
k = chemical reaction
pol = polarization
s = solid structure
tube = quartz tube
ohm = ohmic

References

[1] Ferrari, M. L., Traverso, A., Massardo, A. F., 2004, "Transient Analysis of Solid Oxide Fuel Cell Hybrids, Part B: Anode Recirculation Model," ASME Paper No. 2004-GT-53716.
[2] Magistri, L., Ferrari, M. L., Traverso, A., Costamagna, P., and Massardo, A. F., 2004, "Transient Analysis of Solid Oxide Fuel Cell Hybrids, Part C: Whole Cycle Model," ASME Paper No. 2004-GT-53845.
[3] Singhal, S. C., 2000, "Advances in Solid Oxide Fuel Cell Technology," Solid State Ionics, **135**, pp. 305–316.
[4] Magistri, L., Bozzo, R., Costamagna, P., and Massardo, A. F., 2002, "Simplified Versus Detailed SOFC Reactor Models and Influence on the Simulation of the Design Point Performance of Hybrid System," ASME Paper No. 2002-GT-30653.
[5] Singhal, S. C., 1997, "Recent progress in Tubular Solid Oxide Fuel Cell Technology," *Proc. of 5th International Symposium on Solid Oxide Fuel Cells*, Electrochemical Society, Penningtin, NJ, **97-18**, pp. 37–50.
[6] Haynes, C., 2002, "Simulating Process Settings for Unslaved SOFC Response to Increases in Load Demand," J. Power Sources, **109**, pp. 365–376.
[7] Achenbach, E., 1995, "Response of a Solid Oxide Fuel Cell to Load Change," J. Power Sources, **57**, pp. 105–109.
[8] Gemmen, R. S., Liese, E., Rivera, J. G., Jabbari, F., and Brower, J., "Development of Dynamic Modeling Tools for Solid Oxide and Molten Carbonate Hybrid Fuel Cell Gas Turbine Systems," ASME Paper No. 2000-GT-554.

Dynamic Simulation of Carbonate Fuel Cell-Gas Turbine Hybrid Systems

Rory A. Roberts

Jack Brouwer

National Fuel Cell Research Center,
University of California,
Irvine, CA

Eric Liese

Randall S. Gemmen

National Energy Technology Laboratory,
Morgantown, WV

Hybrid fuel cell/gas turbine systems provide an efficient means of producing electricity from fossil fuels with ultra low emissions. However, there are many significant challenges involved in integrating the fuel cell with the gas turbine and other components of this type of system. The fuel cell and the gas turbine must maintain efficient operation and electricity production while protecting equipment during perturbations that may occur when the system is connected to the utility grid or in stand-alone mode. This paper presents recent dynamic simulation results from two laboratories focused on developing tools to aid in the design and dynamic analyses of hybrid fuel cell systems. The simulation results present the response of a carbonate fuel cell/gas turbine, or molten carbonate fuel cell/gas turbine, (MCFC/GT) hybrid system to a load demand perturbation. Initial results suggest that creative control strategies will be needed to ensure a flexible system with wide turndown and robust dynamic operation. [DOI: 10.1115/1.1852565]

1 Introduction

Fuel cell systems provide promise for more efficient and environmentally friendly energy production. Certain classes of ''high temperature'' fuel cells are capable of using hydrocarbon fuel as long as the system contains a reformer that can process the fuel to create a hydrogen/carbon-monoxide rich mixture. This capability allows fuel cells to use renewable (biomass) as well as fossil fuels. For large-scale high temperature fuel cells, one can also use the high quality exhaust heat for various purposes such as thermal energy for industrial processes, chemical recuperation (reformation) using the fuel supply, or cogeneration of electricity via a gas or steam turbine, as is the case with a hybrid fuel cell/gas turbine (FC/GT) cycle.

In FC/GT hybrid systems, it is very important to properly match and integrate the fuel cell with the gas turbine portions of the cycle. Too little mass flow in the cathode could lead to overheating of the fuel cell. At the same time, too much mass flow could lower the turbine inlet temperature (TIT) of the gas turbine leading to efficiency decreases, performance degradation, and even gas turbine shutdown. A complete shutdown of the turbine can be damaging to the fuel cell and other system components. In the case of a carbonate fuel cell (or molten carbonate fuel cell, MCFC), the operating temperature is generally desired to be around 650°C. This operating temperature is good for reformation and processing of the original hydrocarbon fuel, but at the same time this temperature is too low for a typical gas turbine TIT. As a result, MCFC/GT hybrid systems generally consist of a fuel cell operating with sufficient excess fuel followed by oxidation of the excess fuel emitted from the anode so as to raise the turbine inlet temperature. This approach can produce very high TIT but at the expense of the overall system fuel-to-electrical efficiency. The fuel cell is the most efficient part of the power plant. So, utilizing as much fuel as possible in the fuel cell (without hindering the performance of the fuel cell) while concurrently maintaining good gas turbine performance is the goal.

The MCFC can handle higher fuel utilizations than what is needed to maintain an acceptable TIT. As a result, there is some design parameters that should allow control even under dynamic conditions. One additional parameter that can be manipulated in the design of an MCFC/GT hybrid system is the mass flow rate through the compressor and therefore the fuel cell. However, when the mass flow is lowered such operation could lead to over heating of the MCFC. The MCFC requires that the heat being generated within the stack be carried off by the anode and cathode gas in order to maintain an optimal temperature. On the other hand, one could design the fuel cell module to directly use a portion of this heat by incorporating internal reformation in the MCFC stack design.

Internal reformation involves endothermic reactions that can absorb the heat directly where it is being generated by the electrochemical reactions. Use of internal reformation provides an additional means of cooling the MCFC, and results in decreased cathode mass flow cooling requirements for the MCFC. As a result, internal reformation increases system efficiency by reducing thermal loss and increasing TIT, and it reduces cost through reduction of heat exchanger requirements [1]. In addition, this approach should provide a more robust hybrid performance envelope. However, it also produces a quicker responding cooling effect due to the immediate proximity of the endothermic reforming action to the stack. Such quick response times need to be carefully considered in the design of the control system, and in the operation of the hybrid system.

Once these overall design decisions are set, control strategies must be employed that can maintain the proper temperatures throughout the MCFC/GT power plant. This needs to be done in order to achieve efficient, safe, and reliable operation during perturbations (e.g., load changes) that may be imposed on the system. The dynamic simulation capabilities developed at National Fuel Cell Research Center (NFCRC) and National Energy Technology Laboratory (NETL) and presented herein are useful for: (1) determining design requirements, (2) analyzing dynamic response, and (3) developing control strategies for MCFC/GT hybrid systems.

Work has been done by others to find the optimal design of different types of fuel cell-gas turbine hybrids. To analyze the performance of hybrid system design, research groups have developed steady-state models to analyze the design and off-design performance of various fuel cell-gas turbine hybrid systems [2,3] approach.

This paper presents the development and comparison of two dynamic MCFC/GT models. The two models have similar features in their thermodynamic approach to simulating a MCFC/GT

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. 2004-GT-53653.

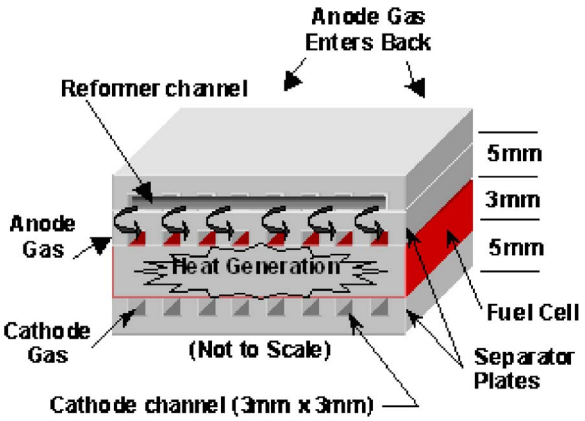


Fig. 1 MCFC diagram

hybrid. The NETL model uses C++ in combination with the ProTRAX software package. The NFCRC is constructed in SimulinkTM software package. The respective solution strategies for determining dynamic hybrid system performance vary according to the software packages used by the parties.

The details of the MCFC and gas turbine simulation approach have been presented and discussed in previous papers [4–6]. Suffice it to note that identical assumptions are made in formulating the set of governing equations for the fuel cell models, which includes discretized solution of the Nernst equation, all major electrochemical losses (polarizations), mass conservation, energy conservation and heat transfer processes. Similar solution strategies are employed for all components, but, slightly different compressor and turbine maps are used, and a significantly different approach to simulating heat exchangers is employed by the two parties. Each of these models, however, retains an appropriate dependance upon heat exchanger design and operating conditions through nondimensional numbers (e.g., Reynolds, Nusselt, Prandtl).

2 Carbonate Fuel Cell Model

Each of the MCFC models is constructed to simulate the fundamental operation of a MCFC similar to that currently manufactured by FuelCell Energy, Inc. Results from NFCRC and NETL fuel cell models are compared under the same operating conditions, with analysis based on a 1.08 m^2, single cell, co-flow, planar fuel cell with prescribed anode and cathode inputs.

A partially reformed anode gas and cathode gas mixture representative of that which characterizes the FuelCell Energy, Inc. system was used for comparison of the fuel cell models [1]. The anode gas was modeled entering a reformation channel (part of the separator plate) in counter-flow with the internal cathode and anode gas flows. Figure 1 illustrates this process. The anode gas is preheated and pre-reformed in the reformation channel. It was found that by creating this counter flow in-stack reformer and having the reformation begin where the anode and cathode gases exit, the temperature profile in the fuel cell is more uniform. If the prereformation channel were not present in the geometrical configuration of the model then most of the methane would reform near the entrance of the MCFC anode compartment. This would cause a dramatic cooling effect and create a large temperature gradient through the cell. This large gradient would hinder the electrochemical performance, lead to mechanical stresses, and possibly cause premature deterioration of the cell materials.

2.1 Modeling of the Internal Reformation. The capability of simulating internal reformation of the type described above is a major advancement of the model that is presented first in this paper. The internal reformation model is broken into two concurrent steps, steam reformation of methane and water gas shift.

$$CH_4 + H_2O \rightarrow CO + 3H_2 \tag{1}$$

$$CO + H_2O \rightarrow CO_2 + H_2 \tag{2}$$

The rate of the steam reformation step is determined by a dynamic Arrhenius rate expression and the water gas shift is assumed to be in equilibrium. The reformation model uses rates that are consistent with the use of typical nickel-based catalysts. The equilibrium constant for Eq. (1) is determined from the following curve fit [7],

$$K_{P_{\text{ref}}} = 1.198E + 17 * e^{(-26{,}830)/T} \tag{3}$$

To convert this from partial pressure parameters to concentration parameters the following equation is used

$$K_{\text{ref}} = \frac{K_{P_{\text{ref}}}}{(R_u * T)^2} \tag{4}$$

The pre-exponential term of Eq. (3) [7] is scaled by the catalytic surface area as follows:

$$K_o = A_{\text{surf}} K'' \tag{5}$$

The forward reformation reaction rate constant is determined by the following expression:

$$k_{\text{forward}} = K_o * (R_u * T)^{(a+b)} * e^{(-E_a)/R_u T} \tag{6}$$

where

$$\text{Ea} = \text{activation energy } (23{,}000 \text{ [kJ/kmol]}),$$

$$a = 1,$$

and

$$b = -1.25.$$

The forward reaction rate for the reformation is:

$$R_{\text{Forward}} = C^a_{\text{CH}_4} * C^b_{\text{H}_2\text{O}} * k_{\text{forward}} \tag{7}$$

In the current internal reformation model the rate of the reverse reformation reaction (i.e., reverse rate of reaction (1)) is determined through the equilibrium constant as:

$$k_{\text{reverse}} = \frac{k_{\text{forward}}}{K_{\text{ref}}} \tag{8}$$

The reverse reaction rate is determined from the following

$$R_{\text{reverse}} = C_{\text{CO}} * C^3_{\text{H}_2} * k_{\text{reverse}} \tag{9}$$

The result of water gas shift chemistry, which is assumed to be in equilibrium, is determined by calculating the equilibrium constant using Eq. (10) and solving the quadratic in Eq. (11) [7]

$$K_p = 1.767E - 2 * e^{4400/T} \tag{10}$$

$$(1 - K_P)y^2 + [C_{\text{H}_2} + C_{\text{CO}_2} + K_P * (C_{\text{CO}} + C_{\text{H}_2\text{O}})]y + (C_{\text{H}_2} * C_{\text{CO}_2} - K_P * C_{\text{CO}} * C_{\text{H}_2\text{O}}) = 0 \tag{11}$$

Where y is the net change in moles, which can be used in the following equations.

$$\begin{aligned} \text{CH}_2 &= \text{CH}_2 \text{ in} + y \\ \text{CCO}_2 &= \text{CCO}_2 \text{ in} + y \\ \text{CCO} &= \text{CCO in} - y \\ \text{CH}_2\text{O} &= \text{CH}_2\text{O in} - y \end{aligned} \tag{12}$$

The CO is assumed to be consumed/created in the fuel cell by only water gas shift chemistry. Direct electrochemical oxidation of CO and hydrocarbons is neglected, which has been shown experimentally to be a reasonable assumption [8].

Table 1 Fuel cell parameters for comparison

Parameter	Unit	Value
Number of Channels		216
Cell Active Area	m^2	1.08
	Anode Specification	
Inlet Temperature	K	923
Inlet Pressure	Pa	104,425.2
Exit Pressure	Pa	104,425.2
Channel Width	m	0.0031
Channel Height	m	0.0013
Inlet CH_4	mole frac	0.2798
Inlet CO	mole frac	0.005
Inlet CO_2	mole frac	0.0346
Inlet H_2	mole frac	0.1168
Inlet H_2O	mole frac	0.5662
	Cathode Specification	
Inlet Temperature	K	923
Inlet Pressure	Pa	104,425.2
Exit Pressure	Pa	104,425.2
Inlet CO_2	mole frac	0.1553
Inlet H_2O	mole frac	0.1553
Inlet N_2	mole frac	0.559
Inlet O_2	mole frac	0.1294
Channel Width	m	0.0031
Channel Height	m	0.0032
Exchange Current Density	amp/m^2	50
Diffusion Limiting Current Density	amp/m^2	4000
Transfer Coefficient		0.75
	Cell Specification	
Thickness	m	0.01
Heat Capacity	J/kg K	800
Density	kg/m^3	1500
Net Resistance	ohm m^2	-6.667×10^{-7} (T−273) $+4.7833\times10^{-4}$/Acell
	Separator Specification	
Thickness	m	0.001
Heat Capacity	J/kg K	611
Density	kg/m^3	7900

2.2 Fuel Cell Model Comparisons. The MCFC geometry and dynamic response of a laboratory scale MCFC model were presented in a previous paper [6] along with the dynamic equations used in each of the models. The same MCFC models are used in the comparison presented in this paper. The most significant change in the models is the addition of internal reformation simulation capabilities.

Table 1 presents the parameters used in the current model. The two models were simulated with partially reformed gas entering the MCFC anode. The values are compared to ensure that predicted performance matches closely. The comparison of steady state results are presented in Table 2. The voltages of the models were held constant at 0.7 V. The models compare very well with regard to predictions of exit species concentrations, current production, and overall fuel and oxidant utilization.

3 MCFC/GT Hybrid Model

The MCFC/GT models have similar features and have the FuelCell Energy, Inc. configuration of a Direct FuelCell/Turbine® (DFC/T®) sub-MW system. A diagram of the system is presented in Fig. 2. The fuel heater is slightly different than the DFC/T power plant presented by FuelCell Energy, Inc. [9].

3.1 Gas Turbine Model. The small gas turbine (or microturbine) simulated in this paper contains a single stage radial compressor and expander connected on the same shaft to a generator. Generic radial turbine and compressor maps are used to establish compressor and turbine performance over a wide range of operation. The NFCRC gas turbine model has a pressure ratio of 3:1. The design parameters for the gas turbine are provided in Table 3. NETL's gas turbine model is slightly different. At steady state, the compressor pressure ratio is 4:1. The turbine inlet temperature is 785°C. The design parameters and the turbine and compressor maps used by NETL are based upon measured performance of a small gas turbine engine.

Table 2 Steady-state results for the MCFC with internal reformation

Parameter	Unit	NFCRC	NETL
	Anode		
Outlet			
Outlet Temperature	K	923	923
Outlet CH_4	mole frac	0.0007	0
Outlet CO	mole frac	0.00133	0
Outlet CO_2	mole frac	0.499	0.4922
Outlet H_2	mole frac	0.1195	0.1417
Outlet H_2O	mole frac	0.3794	0.366
	Cathode		
Outlet			
Outlet Temperature	K	923	923
Outlet CO_2	mole frac	0.0739	0.0629
Outlet H_2O	mole frac	0.18	0.1833
Outlet N_2	mole frac	0.648	0.66
Outlet O_2	mole frac	0.09695	0.0925
	Cell		
Voltage	V	0.7	0.7
Current	A	1216	1212
Fuel Utilization		75.10%	74.37%
Oxygen Utilization		35.34%	40.70%

The dynamic response of the gas turbine engine models is based upon solution of a dynamic expression that governs shaft speed, which balances compressor and generator loads and gas turbine inertia with power produced in the turbine. The compressor plenum volume (note, the NETL model does not contain a plenum volume) provides for mass storage in the system. The mass flow presented in Table 4 is higher than what today's microturbines can provide, but is not unrealistic to expect from emerging turbo-machinery designed for hybrid systems applications. With these design parameters, the micro-turbine produces about 10% of the total hybrid system power at the design point.

3.2 1D Heat Exchanger Model. NETL's heat exchanger models simulate a state-of-the-art primary surface type heat exchanger following McDonald [10]. It is modeled in one dimension with 3 nodes. The NFCRC heat exchanger models assume a flat plate design, which are simulated using first principals (i.e., concurrent solution of the momentum, convection and conduction equations in a shell and tube geometry as a function of heat exchanger length). It resolves one dimension of the heat exchangers in discrete lumped parameter cells. As a basis for comparison, both models are assumed to have the same thermal capacitance. The following assumptions for heat exchanger masses were made: the high temperature heat exchanger has a mass of 400 kg, the steam generator and fuel heater section has a mass of 250 kg, and the low temperature heat exchanger has a mass of 300 kg.

3.3 Catalytic Oxidizer Model. The catalytic oxidizer model of both NETL and NFCRC completely oxidizes all of the remaining fuel from the fuel cell anode gas stream. The adiabatic flame temperature is calculated for the combined anode/gas turbine exhaust gas mixture. The catalytic oxidizer is assumed to have no mass, therefore there is no thermal capacitance associated with the catalytic oxidizer in either model.

4 MCFC/GT Hybrid Dynamic Simulation Results

The NETL and NFCRC integrated hybrid systems models were each run with the same set of inlet conditions and operating parameters. This was done to establish that both models produce similar steady-state results under the same conditions. The steady-state results from both the NETL and NFCRC hybrid system models for these conditions are presented in Table 4. The majority of

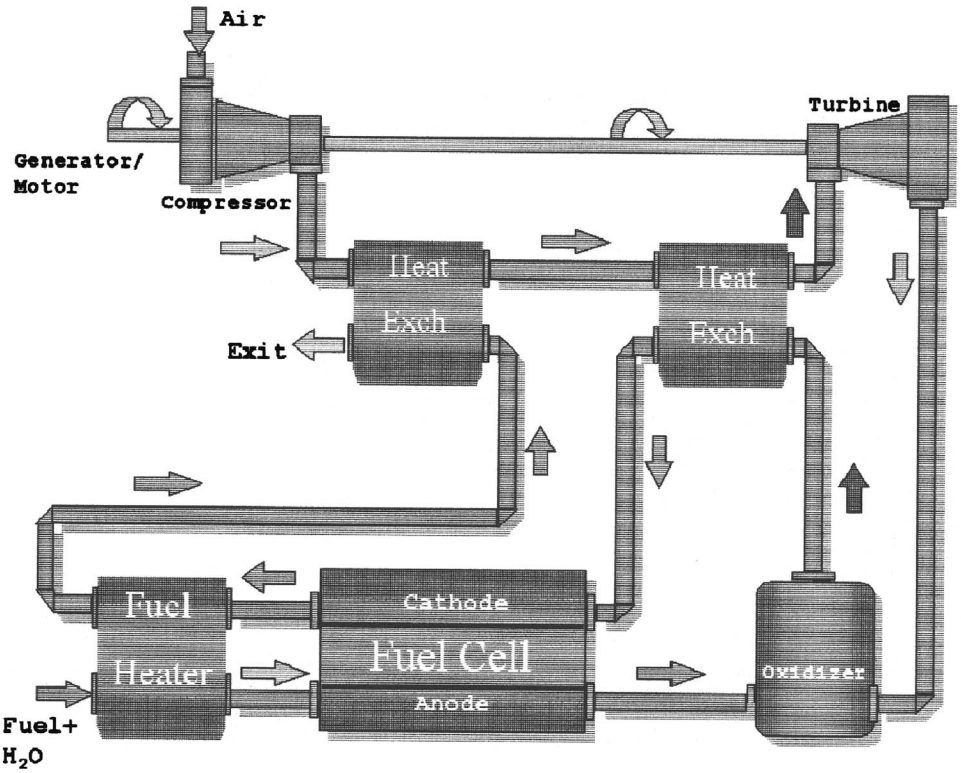


Fig. 2 MCFC/GT hybrid power plant schematic

the steady state performance comparisons presented in Table 4 suggests a relatively similar predicted performance for the hybrid system. There is a disparity between the cathode inlet temperature and the MCFC stack power predictions of the two models, which is probably due to the differences in the heat exchanger simulations.

Once the steady-state results were verified to be similar, each of the dynamic models was subjected to the same perturbation. The perturbation was a drop in overall load demanded from the fuel cell, which was accomplished by applying an increase in fuel cell operating voltage of 0.01 V from 0.76 V to 0.77 V.

Table 3 Gas turbine parameters

Design Spec	Value
Design Mass Flow Rate (kg/s)	1.33
Design Turbine Inlet Pressure for NFCRC (kPa)	304
Design Turbine Inlet Pressure for NETL (kPa)	405
Design Turbine Inlet Temperature (K)	1050
Design Compressor Inlet Pressure (kPa)	101.325
Compressor Impeller Radius (m)	0.055
Design Compressor Inlet Temperature (K)	298
Design Speed (RPM)	65,000
Diffuser Expansion Ratio	1.4
Plenum Volume (NFCRC only) (m^3)	0.8

Table 4 Steady-state hybrid results

Description	NFCRC	NETL
Catalytic Exhaust Temp (K)	1129.44	1106.724
Cathode Inlet Temp (K)	856.45	847.77
Gas Turbine Power (kW)	136.35	127.32
Fuel Cell Power (kW)	856.95	869.287
Efficiency	56.20%	56.50%
Current (single cell) (A)	1025.06	1039
Voltage (single cell) (V)	0.76	0.76

4.1 Hybrid Dynamic "Open Loop" Response. This section of the paper presents the predicted open loop system dynamics, since MCFC voltage was the manipulated variable and no other system operating parameter was controlled in these simulations. The fuel flow rate was held constant. Figures 3–5 present several hybrid system responses produced by each model during the voltage perturbation described above. When the voltage is increased, both the current of the MCFC and efficiency drop as expected. Figure 3 presents the current change and the change of the overall efficiency for each of the hybrid systems simulations (that of NETL and that of NFCRC). A difference in the current predictions is presented, which leads to a slightly higher efficiency for the NETL hybrid model (not significant). Before the perturbation, the current in both models is different, partly because of the different temperatures in the MCFC component of each model and different gas compositions through the MCFC stack that result from differences in internal reforming. The NFCRC fuel cell ends up operating at a lower temperature. This increases the internal resistance of the fuel cell, which decreases the current density of the fuel cell as well.

Figure 4 presents the catalytic oxidizer and cathode inlet temperature open loop response for the same fuel cell load perturbation. Note that the time scales associated with the dynamic response of the catalytic oxidizer and cathode inlet temperatures predicted by the two models compare well. In addition, the magnitude of the predicted final catalytic oxidizer temperature of the two models compares well. The difference in the magnitude of catalytic oxidizer temperature is about 30 K before the perturbation. This is due to more fuel being available in the anode off gas in the NFCRC model because of the lower current production. After the perturbation, the catalytic oxidizers reach the same temperatures because the MCFC currents are the same after the perturbation, which is seen in Fig. 3.

The increase in catalytic oxidizer temperature could be detrimental to the heat exchanger and the increase in cathode inlet temperature as seen in Fig. 4 could overheat the MCFC. This

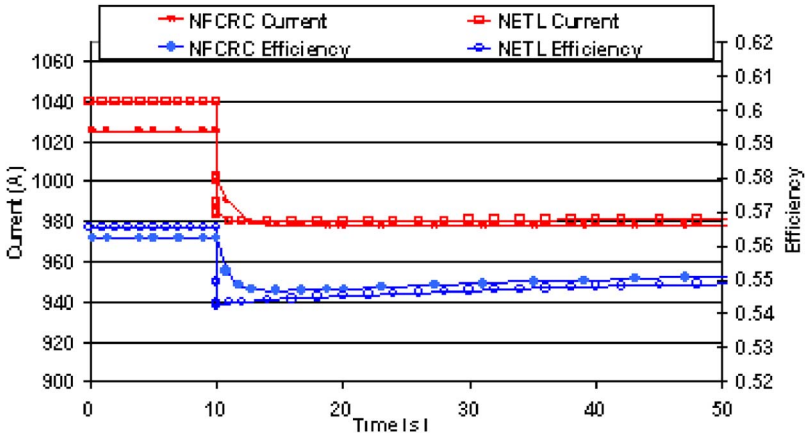


Fig. 3 1.08 m2 fuel cell response and overall hybrid efficiency during an increase of fuel cell voltage 0.76V→0.77V for an open loop system

increase in cathode inlet temperature could possibly be controlled by decreasing the amount fuel flow entering the MCFC or by increasing the air mass flow through the MCFC.

Figure 5 presents the change in the total MCFC/GT hybrid system power and the MCFC and gas turbine power for both hybrid system models. For the open loop system, both hybrid models predict that the MCFC power drops followed by a rise in gas turbine power due to a rise in catalytic oxidizer temperature. Although the magnitudes of power predicted by the respective models are not the same, the time scales associated with the dynamic response of the hybrid system are well matched. The fuel cell fairly rapidly drops in power output (order of 10 s), followed by a gas turbine response that endures longer than the 50 s plotted in Fig. 5. Since the heat exchangers are modeled with small masses in this case, the gas turbine power rises fairly quickly and the total power of the system does recover slightly due to the increase in gas turbine power.

4.2 Hybrid System Control Strategies. In order to minimize the thermal impact on the fuel cell, it is desirable to maintain a fairly constant inlet temperature to the cathode. This can be done in various ways. One method adjusts the turbine speed to adjust airflow to the oxidizer, changing the temperature in the oxidizer and ultimately the temperature at the cathode inlet. The generator load is adjusted in order to maintain the desired turbine shaft speed. The controller would likely use the catalytic oxidizer temperature as the controlled parameter since response of the cathode inlet temperature would be impacted by the thermal mass of the high temperature heat exchanger. However, the present case presents a turbine that starts out at 100% speed. Since increases in turbine speed are not allowed at this point, the oxidizer temperature cannot be controlled by increasing turbine speed. As a result the NETL model allows additional adjustment of fuel flow to maintain the original fuel utilization. A decrease in load demand to the fuel cell will require a decrease in fuel flow in order to maintain the desired fuel utilization. The overall result should lower the oxidizer temperature to a value below its setpoint, which can be followed by further adjustments to turbine speed to increase the oxidizer temperature back to the setpoint.

For the NFCRC model, a different control strategy is implemented. A proportional integral (PI) controller is used to control the fuel flow rate in order to maintain the catalytic oxidizer temperature. Controlling the catalytic oxidizer temperature by adjusting the amount of fuel that enters the catalytic oxidizer can be accomplished by adjusting the total fuel flow entering the MCFC stack. For the case of a drop in load demand in the power plant, lowering the fuel flow rate will increase the fuel utilization in the MCFC stack and in turn lower the amount of fuel entering the catalytic oxidizer. In order to maintain fairly constant power from

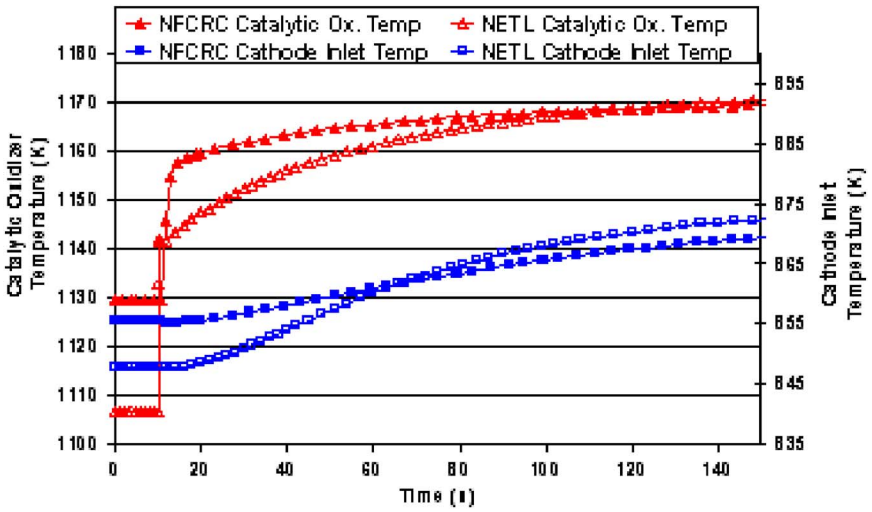


Fig. 4 Catalytic oxidizer and cathode inlet temperature during an increase of fuel cell voltage 0.76V→0.77V for an open loop system

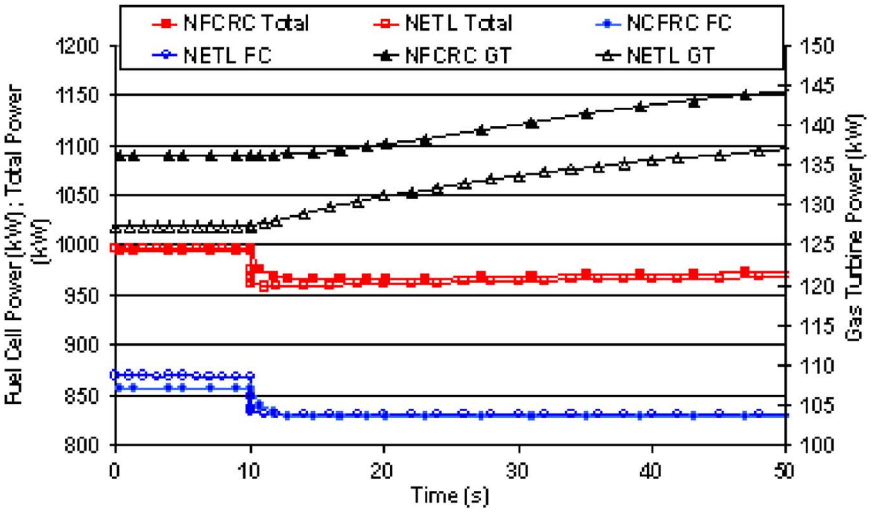


Fig. 5 Power of fuel cell and gas turbine during an increase of fuel cell voltage 0.76V→0.77V for an open loop system

the gas turbine, the temperature of the catalytic oxidizer should remain approximately constant, and this can be achieved with such a control strategy. This should in turn maintain fairly constant mass flow through the gas turbine while the fuel utilization in the fuel cell is left to float. This control strategy may alter the MCFC performance, but it allows control of the catalytic oxidizer temperature and gas turbine performance parameters.

The goal of both of these control strategies is to maintain an adequate cathode flow and inlet temperature, a net power production from the gas turbine, while maintaining MCFC temperature near an optimal operating temperature. It is generally accepted that the operating temperature of a MCFC is 650°C. There is a balance between the MCFC performance and cell stack and component degradation that makes this a desirable temperature. A study done by Au [11] shows that if the operating temperature of the MCFC changes from 600 to 700°C then the overall CHP power plant efficiency is not much affected.

4.3 Hybrid Dynamic "Closed Loop" Response. The predicted dynamic response to a drop in power demand of the two hybrid MCFC/GT system models, each with a control loop applied as described above to maintain catalytic oxidizer temperature, is presented in this section.

Figures 6–8 show the response of the MCFC/GT power plant to a drop in load demand with the fuel flow rate into the MCFC controlled to maintain the catalytic oxidizer at a reference temperature. Figure 6 shows the initial drop in current produced by the MCFC stack due to the change in operating voltage. Given the voltage increase, the fuel utilization decreases due to the lower current. Each of the models predicts these overall impacts.

In order to maintain constant fuel utilization, the NETL model initially reduces the fuel flow in response to the perturbation. Because the fuel valve stroke time that is used is very small (1 ms) there is an almost immediate reduction in the fuel flow from the proportional action of the PID controller on the fuel valve. Thus the catalytic oxidizer temperature presented in Fig. 7 also shows a sharp reduction after the initial sharp increase, and this temperature is coincidentally approximately the original oxidizer temperature. After this the fuel controller continues to reduce the fuel flow in order to bring the fuel utilization back to its original value and this is why the oxidizer temperature continues to decrease. However, the turbine speed control begins to act by reducing the turbine speed to reduce airflow until the catalytic oxidizer temperature is back to its set point. Note that the NETL cathode inlet

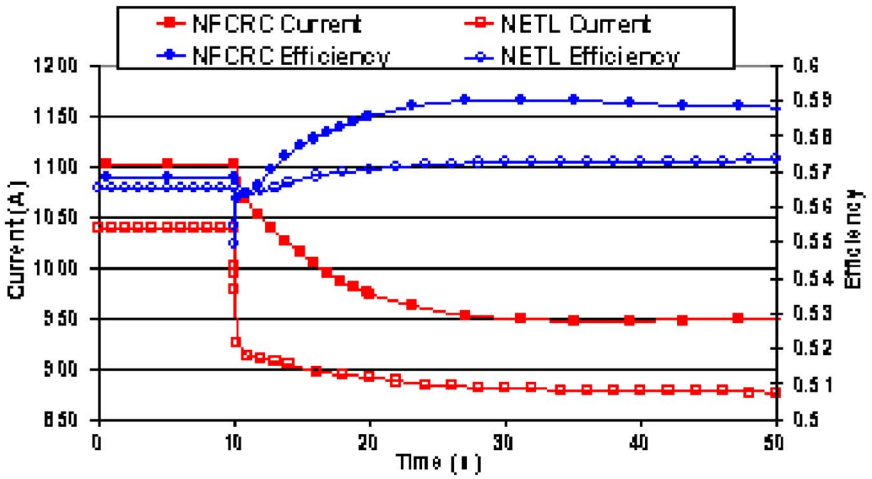


Fig. 6 1.08 m2 fuel cell response and overall hybrid efficiency during an increase of fuel cell voltage 0.76V→0.77V for a closed system

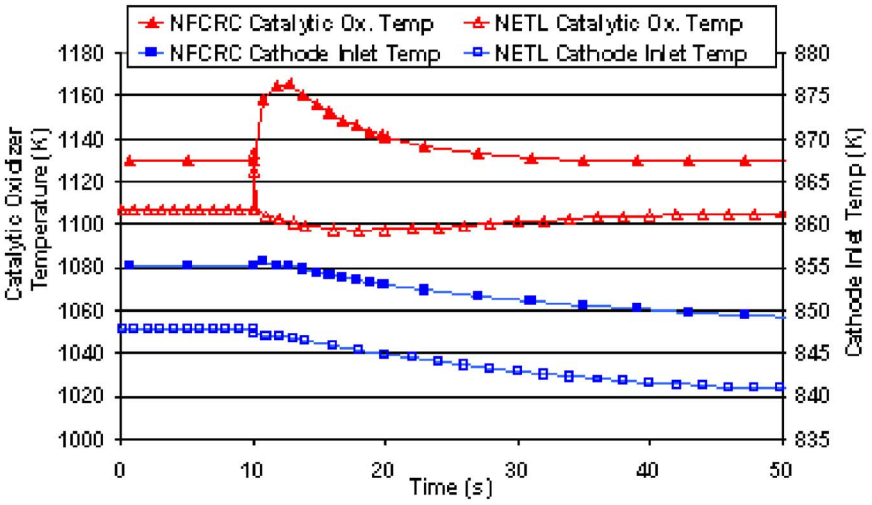


Fig. 7 Catalytic oxidizer and cathode inlet temperature response during an increase of fuel cell voltage 0.76V→0.77V for a closed system

temperature does decrease slightly and so the oxidizer temperature set point would have to be increased to reach the original setpoint for cathode inlet temperature.

The NFCRC model, which incorporated a simple PI controller for overall fuel flow, based on the difference between the catalytic oxidizer temperature and its setpoint exhibits a different dynamic response. Each of the two control strategies requires approximately 35 s to bring the catalytic oxidizer temperature back to its setpoint. However, the simple control strategy used in the NFCRC model allows the catalytic oxidizer temperature to rise more than 30° and linger at higher temperatures for a longer period of time than the NETL control strategy. The NFCRC controller parameters lead to an over-damped catalytic oxidizer temperature response, suggesting that modifications to the proportional and integral gains on the controller could improve the dynamic response.

The results of Fig. 6 show that the predicted fuel cell current and overall hybrid efficiency of the two hybrid system models are different before the perturbation as discussed earlier. The MCFC stack currents do approach the same values, but the NFCRC MCFC approaches this state more slowly due to the slower fuel flow rate controlled response. The gas turbine power for each model presented in Fig. 8, reflects the differences in the two control strategies. The NETL gas turbine power decreases due to the drop in load on the generator allowing the shaft speed to increase and the air flow rate to increase. NFCRC gas turbine power shows a small temporary increase due to the increase of the catalytic oxidizer temperature and TIT. This increase in gas turbine power and the decrease in fuel flow rate allows the NFCRC hybrid system to reach higher efficiencies as seen in Fig. 6.

Even though the control strategy for the NFCRC MCFC/GT hybrid system leads to a more efficient operating condition, the corresponding sharp rise in catalytic oxidizer temperature for any length of time may not be acceptable. This temperature could melt the catalyst. One may desire the quicker response that the NETL model provides over a more efficient strategy in order to extend the life of this system. At the same time, if the masses of the heat exchangers are much larger, then the controller response time may be sufficient.

As mentioned earlier, the heat exchangers that are modeled in each hybrid system are different. When combined with the differ-

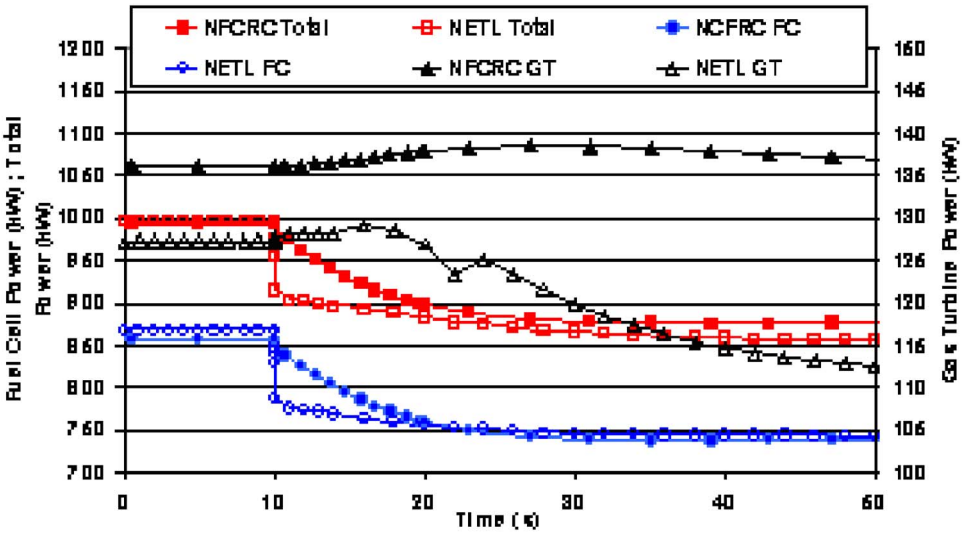


Fig. 8 Power of fuel cell and gas turbine during an increase of fuel cell voltage 0.76V→0.77V for a closed system

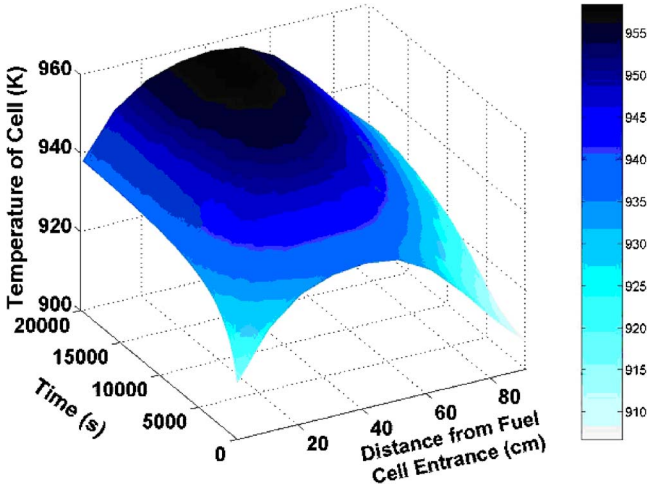


Fig. 9 Temperature profile in the MCFC during the open loop voltage perturbation

ent control strategies that are applied, the heat exchangers create the observed predicted performance differences of the two hybrid system models.

4.4 Rise in Cathode Inlet Temperature. As pointed out earlier, maintaining proper temperature control with the various components and gas streams throughout the hybrid system is very important. One critical temperature that was focused on was the cathode inlet temperature. Controlling this temperature is very important in operating the MCFC stack at a safe and desirable temperature. For the open loop dynamic response for the MCFC/GT hybrid system it was shown in Fig. 4 that if the MCFC stack experiences a 0.01 V change or a 5% change in power demand that the cathode inlet temperature can rise 30°C. Figure 9 presents the temperature profile in the MCFC as a function of time and distance from the entrance of the fuel cell gas streams. Figure 9 shows that the entire MCFC temperature profile rises around 30°C as well. This results in peak internal temperatures that approach dangerous levels of 955 K (682°C). This high temperature is extreme enough to degrade the performance and the life of the MCFC. And this result arises from only a small perturbation in power demand on the system.

5 Summary and Conclusions

The comparison of two different MCFC/GT models developed and applied in two different simulation environments resulted in good agreement. The models were able to predict the same trends and responses and significantly showed similar time scales associated with the dynamic response of hybrid fuel cell gas turbine system components and operating conditions. Steady state predictions of MCFC performance demonstrated similar predictive capabilities for the two fuel cell models. A MCFC/GT cycle of the type designed by FuelCell Energy, Inc. was presented and simulated. The simulated open-loop uncontrolled dynamic response of this MCFC/GT system to a fuel cell load perturbation was presented. The dynamic responses of the two hybrid system models compare well. Closed-loop dynamic responses of the hybrid system to fuel cell load perturbations using two different control strategies are also presented.

Several differences between the two hybrid system models are noted, due in large part to the differences in the specific heat exchanger and gas turbine models used. The gas turbine models used different compressor and turbine maps and were set at slightly different design conditions, while entirely different heat exchanger technology was assumed in the respective models. The results point to important precautions for both modeling and designing an MCFC/GT system. The performance of subcomponents such as heat exchangers and control algorithms can dictate success or failure for these systems. The correct heat exchanger for the application can determine whether or not a system will properly operate.

The results suggest that control strategies that are implemented in hybrid MCFC/GT systems need to be robust yet flexible enough to allow the system operate under various conditions. The simple control strategies implemented in this paper were capable of controlling the hybrid system during a simple fuel cell load perturbation. In the case of more complex perturbations or operating conditions, such as those extant with being connected to the grid, more sophisticated control schemes should be considered. Nonetheless, the current model dynamic predictions suggest that control of these complex hybrid systems is quite feasible.

Hybrid fuel cell systems must be designed to be able to recover and maintain safe and efficient performance at all times. The models developed and applied in the current paper can be valuable tools applied to study a large number of complex perturbations and control strategies to gain confidence that the hybrid system (components+controls) that one desires will function properly.

Nomenclature

A_{surf} = surface area
Ea = activation energy
K_j = equilibrium constant for concentrations
K_{pi} = equilibrium constant for partial pressures
k_i = rate constants from the Arrhenius equation
K_o = Pre-exponential for the Arrhenius equation
K'' = Pre-exponential per unit area
R_i = Dynamic reaction rates (kmol/s) for reaction i
C_j = Concentrations of species j
y = Change in kmoles for water–gas shift reaction

References

[1] Lukas, M. D., Lee, K., Y., and Ghezel-Ayagh, H., 1999, "Development of a Stack Simulation Model for Control Study on Direct Reforming Molten Carbonate Fuel Cell Power Plant," IEEE Trans. Energy Conversion, PE-468-EC-0-01-1999.
[2] Costamagna, P., Magistri, L., and Massardo, A. F., 2000, "Design and Part-Load Performance of a Hybrid System Based on a Solid Oxide Fuel Cell Reactor and a Micro Gas Turbine," J. Power Sources, **96**, pp. 352–368.
[3] Rao, A. D., and Samuelsen, G. S., 2002, "Analysis Strategies for Tubual Solid Oxide Fuel Cell Based Hybrid," J. Eng. Gas Turbines Power, **124**, pp. 503–509.
[4] Gemmen, R. S., Liese, E., Rivera, J. G., Jabbari, F., and Brouwer, J., 2000, "Development of Dynamic Modeling Tools for Solid Oxide and Molten Carbonate Hybrid Fuel Cell Gas Turbine Systems," 2000 Proc. ASME Turbo Expo, Munich, Germany, paper no. 2000-GT-0554.
[5] Liese, E. A., Gemmen, R. S., 2002, "Dynamic Modeling Results of a 1 MW Molten Carbonate Fuel Cell/Gas Turbine Power System," 2002 Proc. ASME Turbo Expo, Amsterdam, The Netherlands, paper no. GT-2002-30110.
[6] Roberts, R. A., Brouwer, J., Gemmen, R., and Liese, E., 2003, "Inter-Laboratory Dynamic Modeling of a Carbonate Fuel Cell for Hybrid Application," 2003 Proc. ASME Turbo Expo, Atlanta, Georgia, paper no. GT-2003-38774.
[7] Hou, K. and Hughes, R., 2000, "The Kinetics of Methane Steam Reforming Over a Ni/a-Al2O Catalyst," Chem. Eng. J., **82**, pp. 311–328.
[8] Weber, A., Bastain, S., Muller, A. C., Herbstriitt, D., Ivers-Tiffee, E., 2002, "Oxidation of H_2, CO, and Methane in SOFCs With Ni/YSZ-Cermet Anodes," Soil Sci. Soc. Am. J., **152–153**, pp. 543–550.
[9] Ghezel-Ayagh, H., Daly, J. M., and Wang, Z.-H., 2003, "Advances in Direct Fuel Cell/Gas Turbine Power Plants," 2003 Proc. ASME Turbo Expo, Atlanta, Georgia, paper no. GT2003-38941.
[10] McDonald, C. F., 2000, "Low Cost Recouperator Concepts for Microturbine Applications," 2000 Proc. ASME Turbo Expo, Munich, Germany, paper no. 2000-GT-167.
[11] Au, S. F., McPhail, S. J., Woudstra, N., and Hemmes, K., 2003, "The Influence of Operating Temperature on the Efficiency of a Heat and Power Fuel Cell Plant," J. Power Sources, **122**, pp. 37–46.

Modeling and Simulation of a Gas Turbine Engine for Power Generation

Qusai Z. Al-Hamdan
Air Force Engineering Branch,
Royal Jordanian Air Force,
Amman, Jordan
e-mail: qusai@go.com.jo

Munzer S. Y. Ebaid
Head of Scientific Research Department,
King Abdullah II Design and Development Bureau (KADDB),
Amman, Jordan

The gas turbine engine is a complex assembly of a variety of components that are designed on the basis of aerothermodynamic laws. The design and operation theories of these individual components are complicated. The complexity of aerothermodynamic analysis makes it impossible to mathematically solve the optimization equations involved in various gas turbine cycles. When gas turbine engines were designed during the last century, the need to evaluate the engines performance at both design point and off design conditions became apparent. Manufacturers and designers of gas turbine engines became aware that some tools were needed to predict the performance of gas turbine engines especially at off design conditions where its performance was significantly affected by the load and the operating conditions. Also it was expected that these tools would help in predicting the performance of individual components, such as compressors, turbines, combustion chambers, etc. At the early stage of gas turbine developments, experimental tests of prototypes of either the whole engine or its main components were the only method available to determine the performance of either the engine or of the components. However, this procedure was not only costly, but also time consuming. Therefore, mathematical modelling using computational techniques were considered to be the most economical solution. The first part of this paper presents a discussion about the gas turbine modeling approach. The second part includes the gas turbine component matching between the compressor and the turbine which can be met by superimposing the turbine performance characteristics on the compressor performance characteristics with suitable transformation of the coordinates. The last part includes the gas turbine computer simulation program and its philosophy. The computer program presented in the current work basically satisfies the matching conditions analytically between the various gas turbine components to produce the equilibrium running line. The computer program used to determine the following: the operating range (envelope) and running line of the matched components, the proximity of the operating points to the compressor surge line, and the proximity of the operating points at the allowable maximum turbine inlet temperature. Most importantly, it can be concluded from the output whether the gas turbine engine is operating in a region of adequate compressor and turbine efficiency. Matching technique proposed in the current work used to develop a computer simulation program, which can be served as a valuable tool for investigating the performance of the gas turbine at off-design conditions. Also, this investigation can help in designing an efficient control system for the gas turbine engine of a particular application including being a part of power generation plant. [DOI: 10.1115/1.2061287]

1 Modeling of Gas Turbine Components

Aero-derivative and industrial gas turbine engines are used for a variety of applications, such as electrical power generation, driving pumps, compressors on gas and liquid fuels, etc. The engine configuration may vary to suit the application. The common configurations are a single-, twin-, or triple-spool construction or a single-stage or multistage construction. In this study only the gas turbines used for electrical power generation are considered.

A gas turbine engine essentially consists of the following component parts: (i) intake, (ii) compressor(s), (iii) combustion chamber(s), (iv) turbine(s), and (v) engine auxiliaries, such as fuel pump, lubrication pump, electrical power supply, starting gear, and control system. A block diagram of the gas turbine engine showing these components is given in Fig. 1.

Overall performance of the complete gas turbine engine is mainly determind by the main components i, ii, iii, and iv. The mathematical model for each component was created using physical laws or emphrical data when available. The thermodynamic properties of combustion gases and air at various stages throughout the gas turbine are calculated by considering variation of temperature or instead universal gas constant (R) can be used since (R) is temperature independent.

Tables containing the values of the specific heats against temperature variation have been published in many references, such as Chappel and Cockshutt [1]. In the present work, to compute the values of specific heats at constant pressure and various temperatures for air and combustion gases, data from the tables were fitted with polynomial curves to obtain Eqs. (1)–(5). These equations provide details of the polynomials. Here, T_a and T_g refer to the average temperatures during the compression and expansion processes in the compressor and turbine, respectively.

For air at the low-temperature range of 200–800 K,

$$C_{\mathrm{Pa}} = 1.0189 \times 10^{3} - 0.13784 T_a + 1.9843 \times 10^{-4} T_a^2 + 4.2399 \times 10^{-7} T_a^3 - 3.7632 \times 10^{-10} T_a^4 \quad (1)$$

For air at the high-temperature range of 800–2200 K,

Contributed by the Gas Turbines Electric Power of the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received February 22, 2004; final manuscript received April 27, 2005. Review conducted by S. van der Linden.

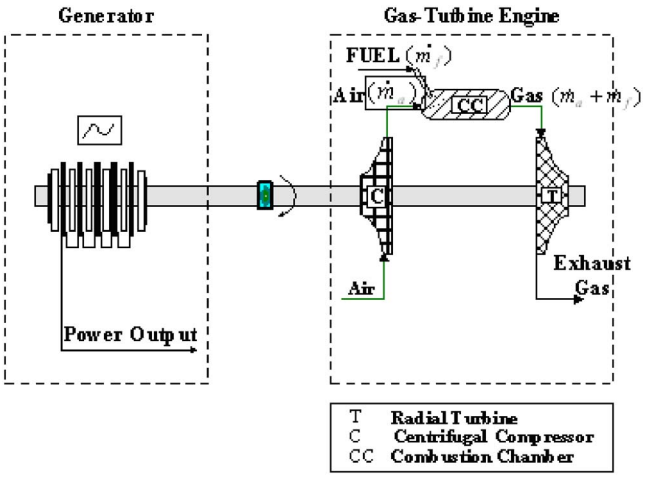


Fig. 1 Schematic diagram of simple gas turbine engine

$$C_{\mathrm{Pa}} = 7.9865 \times 10^2 + 0.5339T_a - 2.2882 \times 10^{-4}T_a^2 + 3.7421 \times 10^{-8}T_a^3 \quad (2)$$

For specific heats of products of combustion,

$$C_{Pg} = C_{\mathrm{Pa}} + [f/(1+f)]B_T \quad (3)$$

where B_T at the low-temperature range of 200–800 K,

$$B_T = -3.59494 \times 10^2 + 4.5164T_g + 2.8116 \times 10^{-3}T_g^2 - 2.1709 \times 10^{-5}T_g^3 + 2.8689 \times 10^{-8}T_g^4 - 1.2263 \times 10^{-11}T_g^5 \quad (4)$$

and B_T at high-temperature range of 800–2200 K,

$$B_T = 1.0888 \times 10^3 - 0.1416T_g + 1.916 \times 10^{-3}T_g^2 - 1.2401 \times 10^{-6}T_g^3 + 3.0669 \times 10^{-10}T_g^4 - 2.6117 \times 10^{-14}T_g^5 \quad (5)$$

1.1 Compressor Modeling and Analysis. The performance of a compressor is fully described by a number of dimensionless parameters. The dimensionless parameters would be the same for every system of units. The dimensionless parameters are shown in Table 1.

Compressor performance, sometimes called compressor map, is usually represented by overall performance characteristics as shown in Fig. 2. These maps are in general, obtained experimentally but sometimes they can be predicted with reasonable accuracy using geometric properties of the components, i.e., intake, impeller, diffuser, and casing [2–4].

Mathematically, the compressor performance is described using the dimensionless parameters as given below

$$\frac{\tau_c}{d_2^2 P_{o1}} = \frac{1}{2\pi}\frac{1}{\eta_c}\left(\frac{d_2 N}{\sqrt{C_{\mathrm{Pa}}T_{o1}}}\right)^{-1}\frac{\dot{m}_a\sqrt{C_{\mathrm{Pa}}T_{o1}}}{d_2^2 P_{o1}}\left[\left(\frac{P_{o2}}{P_{o1}}\right)^{(\gamma_a-1)/\gamma_a} - 1\right] \quad (6)$$

$$\frac{\tau_c}{\delta} = f\left(\eta_c, \frac{N}{\sqrt{\theta}}, \frac{\dot{m}_a\sqrt{\theta}}{\delta}, \frac{P_{o2}}{P_{o1}}\right) \quad (7)$$

Equation (6) is in complete dimensionless form, and Eq. (7) is the general form.

The compression power W_c is given by

$$W_c = d_2^2 P_{o1}\sqrt{C_{\mathrm{Pa}}T_{o1}}\frac{\dot{m}_a\sqrt{C_{\mathrm{Pa}}T_{o1}}}{d_2^2 P_{o1}}\frac{1}{\eta_c}\left[\left(\frac{P_{o2}}{P_{o1}}\right)^{(\gamma_a-1)/\gamma_a} - 1\right] \quad (8)$$

Using the compressor characteristics, if any two dimensionless parameters are known then the rest of the parameters can be determined easily.

Table 1 Dimensionless compressor parameters

PARAMETER	MEANINGS
$\left(\frac{\dot{m}_a\sqrt{C_{pa}T_{01}}}{d_2^2 P_{01}}\right), \left(\frac{\dot{m}_a\sqrt{\theta}}{\delta}\right)$	Compressor dimensionless mass flow parameters
$\left(\frac{d_2 N}{\sqrt{C_{pa}T_{01}}}\right), \left(\frac{N}{\sqrt{\theta}}\right)$	Compressor dimensionless speed parameters
$\left(\frac{\tau_c}{d_2^2 P_{01}}\right), \left(\frac{\tau_c}{\delta}\right)$	Compressor dimensionless torque parameters
$\left(\frac{P_{02}}{P_{01}}\right)$	Compressor dimensionless pressure ratio parameters
$\left[\frac{(P_{02}/P_{01})^{\gamma_a-1/\gamma_a}}{(T_{02}/T_{01}-1)}\right]$	Isentropic compressor efficiency
$\left(\theta = \frac{T_{01}}{T_{0\_ref}}\right)$	Dimensionless temperature parameter
$\left(\theta = \frac{P_{01}}{P_{0\_ref}}\right)$	Dimensionless pressure parameter

The final stagnation temperature in the compression process T_{o2} can be calculated from the following equation:

$$T_{o2} = T_{o1} + \frac{T_{o1}}{\eta_c}\left[\left(\frac{P_{o2}}{P_{o1}}\right)^{(\gamma_a-1)/\gamma_a} - 1\right] \quad (9)$$

In order to solve Eqs. (6)–(9) the needed input data can be obtained from the compressor performance map. This may require interpolation and will be dealt with in the section dealing with computer simulation.

1.2 Combustion Chamber Modeling and Analysis. The combustion chamber performance is normally given in terms of combustion efficiency η_{cc} and the factor for the loss of stagnation pressure ξ_{cc}. Using these parameters the fuel/air ratio F and the stagnation pressure at the exit of the combustion chamber P_{o3} can be determined from Eq. (10) and (11). However, there are many good programs that provide a more accurate determination of F, such as the NASA chemical equilibrium program of Gordon and McBride [5].

$$F = \frac{1}{\dfrac{\eta_{cc}(LCV)}{C_{Pg}(T_{o3} - T_{o2})} - 1} \quad (10)$$

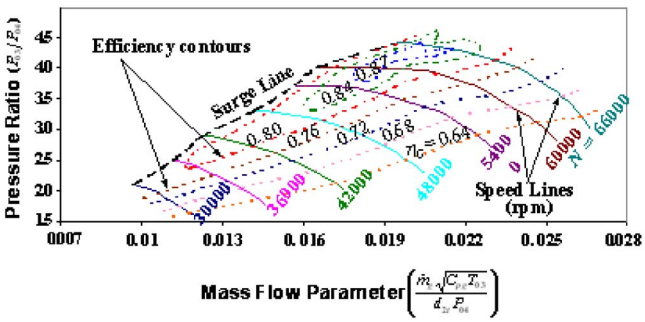


Fig. 2 General compressor characteristics (map)

Table 2 Dimensionless turbine parameters

PARAMETER	MEANINGS
$\left(\frac{\dot{m}_g\sqrt{C_{pg}T_{03}}}{d_2^2P_{03}}\right),\left(\frac{\dot{m}_g\sqrt{\theta}}{\delta}\right)$	Turbine dimensionless mass flow parameters
$\left(\frac{d_2N}{\sqrt{C_{pg}T_{01}}}\right),\left(\frac{N}{\sqrt{\theta}}\right)$	Turbine dimensionless speed parameters
$\left(\frac{\tau_c}{d_2^2P_{03}}\right),\left(\frac{\tau_c}{\delta}\right)$	Turbine dimensionless torque parameters
$\left(\frac{P_{03}}{P_{04}}\right)$	Turbine dimensionless pressure ratio parameter
$\left[\frac{(P_{03}/P_{04})^{\gamma_a-1/\gamma_a}}{(T_{03}/T_{04}-1)}\right]$	Isentropic Turbine efficiency
$\left(\theta=\frac{T_{03}}{T_{0\_ref}}\right)$	Dimensionless temperature parameter
$\left(\delta=\frac{P_{03}}{P_{0\_ref}}\right)$	Dimensionless pressure parameter

$$P_{o3}=(1-\xi_{cc})P_{o2} \tag{11}$$

1.3 Turbine Modeling and Analysis. The performance characteristics of a turbine, such as those of a compressor, are fully described by a number of dimensionless parameters. These parameters and their corresponding meanings are tabulated in Table 2 [6–11].

The turbine performance is represented by overall performance characteristics, also known as turbine map as shown in Fig. 3. This map is in general obtained experimentally but it can also be predicted with reasonable accuracy by using geometric properties and on the basis of previous experience [2–4].

In the current work, turbine cooling has not been accounted for and will be taken into consideration in future work. Using the turbine characteristics, if any two dimensionless parameters are known then the rest of the parameters can be determined easily. Mathematically, the turbine performance is described using the dimensionless parameters as follows:

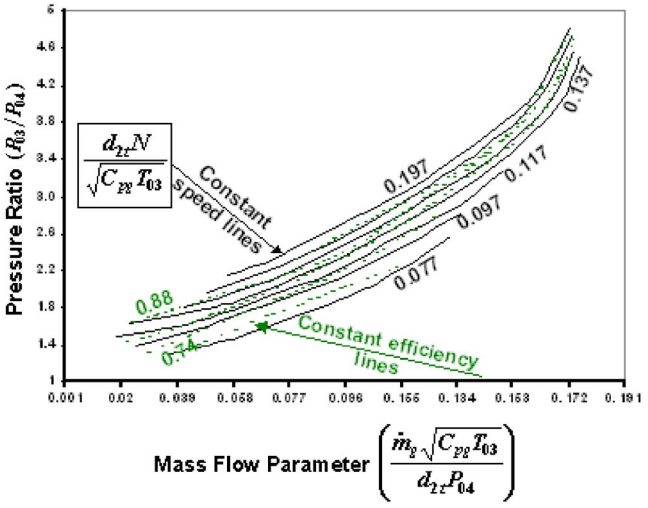


Fig. 3 Turbine Characteristics (Map)

$$\frac{\tau_t}{d_2^2P_{o3}}=\frac{1}{2\pi}\eta_t\left(\frac{d_2N}{\sqrt{C_{Pg}T_{o3}}}\right)^{-1}\frac{\dot{m}_g\sqrt{C_{Pg}T_{o3}}}{d_2^2P_{o3}}\left[1-\left(\frac{P_{o4}}{P_{o3}}\right)^{(\gamma_g-1)/\gamma_g}\right] \tag{12}$$

$$\frac{\tau_t}{\delta}=f\left(\eta_t,\frac{N}{\sqrt{\theta}},\frac{\dot{m}_g\sqrt{\theta}}{\delta},\frac{P_{o3}}{P_{o4}}\right) \tag{13}$$

Equation (12) is in complete dimensionless form, and Eq. (13) is the general form.

The expansion power W_t and the final stagnation temperature T_{o4} in the expansion process are calculated using Eqs. (14) and (15), respectively.

$$W_t=d_2^2P_{o3}\sqrt{C_{Pg}T_{o3}}\frac{\dot{m}_g\sqrt{C_{Pg}T_{o3}}}{d_2^2P_{o3}}\eta_t\left[1-\left(\frac{P_{o4}}{P_{o3}}\right)^{(\gamma_g-1)/\gamma_g}\right] \tag{14}$$

$$T_{o4}=T_{o3}-T_{o3}\eta_t\left[1-\left(\frac{P_{o4}}{P_{o3}}\right)^{(\gamma_g-1)/\gamma_g}\right] \tag{15}$$

2 Component Matching

Considering a simple gas turbine used for electrical power generation application schematically shown in Fig. 1. The performance of the compressor and the turbine are known by their characteristics maps as shown in Figs. 2 and 3. In this gas turbine engine, the component's matching should meet the following conditions:

i. The compressor shaft speed equals the turbine shaft speed,
$$N_c=N_t=N$$

ii. The gas mass flow through turbine is of the sum the air mass flow through compressor and the fuel mass flow,
$$\dot{m}_g=\dot{m}_a+\dot{m}_f$$

iii. Assuming that the pressure loss in the combustion chamber is a constant small percentage (ξ_{cc}) of the combustion chamber inlet pressure,
$$P_{o3}=(1-\xi_{cc})P_{o2}$$

iv. Assuming that the pressure loss in the compressor Inlet is a constant small percentage (ξ_c) of the atmospheric pressure.
$$P_{o4}=(1-\xi_c)P_{o2}\cong P_{o1}$$

v. Power flows, also, in balance.

It should be noted that the second condition is subject to modification in that it is common practice to bleed air from the compressor at various stations to provide cooling air for bearings and turbine blade cooling. Quiet often it is sufficiently accurate to assume that the bleed air equals the fuel flow, and therefore, the mass flow is the same throughout the compressor and the turbine, i.e., $\dot{m}_g=\dot{m}_a=\dot{m}$.

The steady-state or equilibrium operation of this gas turbine engine can be achieved by the matching of its compressor and turbine. Matching the compressor and the turbine can be done by superimposing the turbine performance map on the compressor map while meeting the components matching conditions. This matching procedure is schematically shown in Fig. 1.

Superimposing the turbine performance map on the compressor map can be achieved by applying Refs. [2–4] or by applying the new approach of making both map's axes (the abscissa and the ordinate) being identical. The main difficulty here is that of temperatures: T_{o1} for the compressor and T_{o3} for the turbine. The problem was solved by introducing a new dimensionless matching parameter $[\dot{m}N/d_{2c}P_{o1}]$ as presented in Sec. 2.1.

2.1 The Compressor. The abscissa of the compressor characteristics, i.e., the mass flow parameter is, $\dot{m}_a\sqrt{C_{\text{pa}}T_{o1}}/d_{2c}^2P_{o1}$. The matching parameter $[\dot{m}N/d_{2c}P_{o1}]$ was obtained by multiplying the mass flow parameter with the dimensionless speed parameter $d_{2c}N/\sqrt{C_{\text{pa}}T_{o1}}$ as follows:

$$\left[\frac{\dot{m}_a\sqrt{C_{\text{pa}}T_{o1}}}{d_{2c}^2P_{o1}}\right]\times\left[\frac{d_{2c}N}{\sqrt{C_{\text{pa}}T_{o1}}}\right]=\left[\frac{\dot{m}N}{d_{2c}P_{o1}}\right] \tag{16}$$

The ordinate, i.e., the pressure ratio parameter P_{o2}/P_{o1}, remains unchanged. Once these transformations had been made, the compressor characteristics map was plotted again where P_{o2}/P_{o1} as the ordinate and $[\dot{m}N/d_{2c}P_{o1}]$ as the abscissa.

2.2 The Turbine. The abscissa of the turbine characteristics, i.e., the mass flow parameter is $\dot{m}_g\sqrt{C_{pg}T_{o3}}/d_2^2P_{o3}$. The matching parameter $[\dot{m}_gN/d_{2c}P_{o4}]$ was obtained by multiplying this parameter with the dimensionless speed parameter, turbine pressure ratio, and the ratio of the turbine rotor diameter to compressor impeller diameter as follows:

$$\left[\frac{\dot{m}_g\sqrt{C_{pg}T_{o3}}}{d_{2t}^2P_{o3}}\right]\times\left[\frac{d_{2t}N}{\sqrt{C_{pg}T_{o3}}}\right]\times\left[\frac{P_{o3}}{P_{o4}}\right]\times\left[\frac{d_{2t}}{d_{2c}}\right]=\left[\frac{\dot{m}_gN}{d_{2c}P_{o4}}\right] \tag{17}$$

To satisfy the compressor-turbine matching conditions specified previously, i.e., $\dot{m}_g=\dot{m}_a=\dot{m}$ and $P_{o4}\cong P_{o1}$. Then the matching parameter of the turbine is equal to the matching parameter of the compressor, i.e.,

$$\frac{\dot{m}_gN}{d_{2c}P_{o4}}=\frac{\dot{m}_aN}{d_{2c}P_{o1}}$$

For the turbine pressure ratio parameter, the ordinate axis of the turbine characteristics map $[P_{03}/P_{04}]$ for matching is developed into

$$\left[\frac{P_{o3}}{P_{o4}}\right]\times\left[\frac{1}{1-\xi_{cc}}\right]\times\left[\frac{P_{o4}}{P_{o1}}\right]=\left[\frac{P_{o2}}{P_{o1}}\right] \tag{13'}$$

Note that $P_{o3}=(1-\xi_{cc})P_{o2}$ and $P_{o4}\cong P_{o1}$.

Once these transformations had been made, the turbine characteristics map was plotted again in terms of these new parameters using Eqs. (17) and (18).

3 Gas Turbine Computer Simulation

A computer program for simulating a gas turbine engine would basically satisfy matching conditions analytically between the various components to produce the equilibrium running line. Representing this line either in the form of *lookup tables* or an *equation* is known as modeling and solving that equation with the help of a computer is computer simulation such that all energy and mass balances, all equations of state of working substances, and the performance characteristics of all components are satisfied.

Testing of the gas turbine engine is expensive and time consuming. Therefore, simulation can be an economic and fast tool for predicting its performance. The simulation of the gas turbine engine can be one of the following:

1. simulation at the design stage where no real gas turbine engine to meet the design specifications yet exists
2. simulation at the application stage where engine is already constructed
3. simulation at the application stage where the generation of design data for additional or auxiliary equipment, such as Lube oil requirements, blow off valve requirements, and limitations for transformers, etc., are needed
4. simulation for performance extrapolation of existing plant to meet higher output requirements

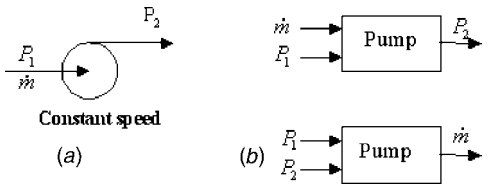


Fig. 5 (*a*) Centrifugal pump in fluid-flow diagram, and (*b*) Possible information-flow blocks representing pump

This study concentrates on the second purpose. The performance of the gas turbine plant at off-design conditions would always be of interest where the gas turbine simulation program for this purpose. This simulation may pin-point the cause of operating problems or shows how the effectiveness of the gas turbine plant may be improved.

Essentially, transient operation of a gas turbine plant is more difficult than the steady-state operation of a gas turbine plant. That field of study was considered to be outside the scope of present paper.

3.1 Information-Flow Diagrams. For system simulation, fluid-flow and energy-flow diagrams are standard engineering tools. An equally useful tool is the information-flow diagram, for example, a block diagram of a control system is an information-flow diagram wherein a block signifies that an output can be calculated when the input is known. A centrifugal pump might appear in a fluid-flow diagram, such as shown in Fig. 5(*a*), while in the information-flow diagram the blocks shown in Fig. 5(*b*). These figures represent functions or expressions that permit calculation of the outlet pressure for one block and the flow rate for the other. A block, as in Fig. 5(*b*), is called a transition function and may be an equation or may be tabular data to which interpolation would be applicable.

Figure 5 shows only one component. To illustrate how these individual blocks can build the information-flow diagram for a gas turbine plant, consider the simple gas turbine cycle in Fig. 1 shown earlier. The components in this cycle are the compressor, the combustion chamber and the turbine.

The information-flow diagram is arranged in Fig. 6 in a manner that might be used if the net power output W_{net} was to be calculated for the system with a given rate of fuel mass flow rate, i.e., heat input at the combustion chamber. Further input information includes the ambient conditions T_{o1}, P_{o1} and rotational speed N.

The compressor block diagram signifies that when the rotational speed N_c, inlet pressure P_{o1}, inlet temperature T_{o1}, and airflow rate $\dot{m}_a$ are specified, the outlet pressure P_{o2} and the compressor efficiency η_c can be determined from the compressor characteristics map as shown previously in Fig. 2 furthermore, power W_c required by the compressor and outlet temperature T_{o2} can be calculated from Eqs. (8) and (9), respectively.

The combustion chamber block diagram signifies that when the fuel flow rate $\dot{m}_f$, inlet temperature T_{o2}, and inlet pressure P_{o2} are specified, the outlet pressure P_{o3} and the outlet temperature T_{o3} can be calculated from Eqs. (10) and (11), respectively.

The turbine block diagram signifies that when the rotational speed N_t, inlet pressure P_{o3}, inlet temperature T_{o3}, and gas flow rate $\dot{m}_g$ are specified, the outlet pressure P_{o4} and the turbine efficiency η_t can be determined from the turbine characteristics map as shown previously in Fig. 3. Furthermore, power W_t delivered by the turbine and outlet temperature T_{o4} can be calculated from Eqs. (14) and (15), respectively.

Sometimes the arrangement of the system permits a direct numerical calculation for the first component of the system using input information. The output information for this first component is all that is needed to calculate the output information of the next component and so on to the final component of the system whose output is the output information of the system. Such a system

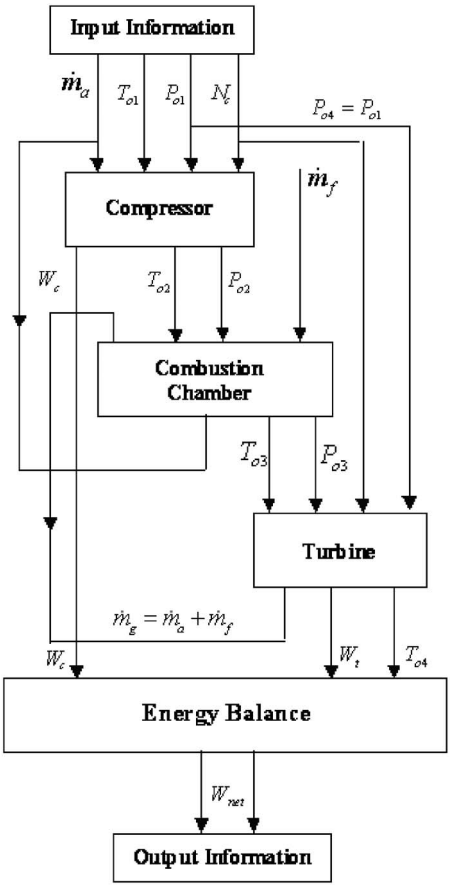


Fig. 6 Information flow diagram of a simple gas turbine engine

simulation consists of sequential calculations.

Before starting the design of the computer program for simulation purposes, it was necessary to identify its main features. Those features are summarized as follows:

i. The computer simulation should allow the user to simulate components individually or as a complete plant.
ii. The simulation program should be modeled for the linking with another program to finally simulate the CPP plant. The outputs of this computer program should contain the needed parameters to start simulating the steam power plant.
iii. The simulation program should be modular so that various modules may be assembled to represent different gas turbine plant configurations.
iv. The simulation program should be user friendly and written in such manner that data can be transferred from one module to another easily and efficiently.

To produce a running line, analytically, the computer simulation program will use the components mathematical models and the components characteristics. That running line will be essential to compute the various gas turbine performance parameters.

In order to use these maps in a computer program it was necessary to have them in a special form. This form can take one of the following shapes:

i. deriving an equation to describe the performance of the component and solving this equation to calculate the performance parameters for any selected point on the performance map
ii. storing the compressor characteristics in look-up tables and then using an interpolation or extrapolation technique to determine the values of the performance parameters for any selected point on the performance map
iii. using a neural network technique to teach the computer the pattern of the compressor characteristics then computing the performance parameters for any selected point on the performance map

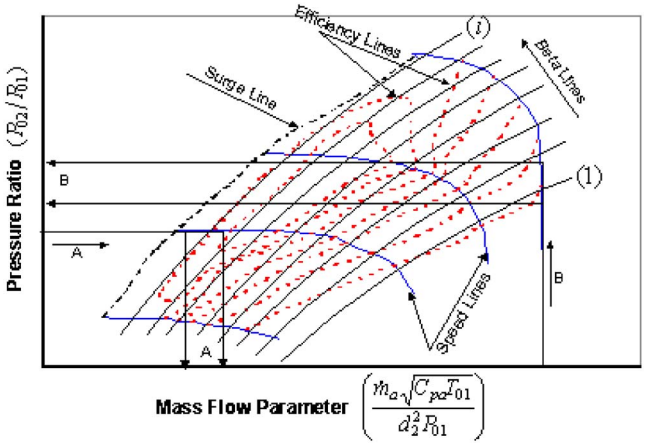


Fig. 7 Problems with reading compressor maps

In the present work, the second option was chosen as it produced a more reliable estimate of the parameters at any point on the component characteristics map.

3.2 Representation of Compressor and Turbine Maps in the Computer Program. The compressor or turbine characteristics in the standard format as shown in Figs. 2 and 3 cannot be used directly in the computer program; they require some processing to convert the maps into a good numerical representation. There are many problems and difficulties associated with the numerical representation process.

3.3 Representation of Compressor Maps in the Computer Program. It is not possible to read the compressor map parameters with given speed N_{Dim} and pressure ratio P_{o2}/P_{o1}, as there might be two values for the mass flow parameter $\dot{m}_{\mathrm{Dim}}$ at given pressure ratio (see point A in Fig. 7).

It is also not possible to determine the efficiency η_c from the compressor map with given speed N_{Dim} and mass flow parameter $\dot{m}_{\mathrm{Dim}}$, because at some parts of the compressor map the speed lines can be vertical. Hence, there might be two values for the pressure ratio P_{o2}/P_{o1} at a single value of mass flow parameter $\dot{m}_{\mathrm{Dim}}$ (see point B in Fig. 7).

Introducing a new coordinates, here called the BETA (β) [4] lines, was the solution for this problem. This allowed an independent map reading using the shape of the parameter lines with the β line and speed parameter N_{Dim}.

The auxiliary coordinates (β lines) can be selected arbitrarily with only two conditions. First, there are no intersections between the β lines within the range of interest, and second, the β lines are equally spaced. The β lines will have any numbers of lines with each line has a parameter number starting from 1. The β lines can be a parabolic lines or straight lines (straight lines are special parabolic lines).

Three-dimensional look-up tables (Tables 3–5) were created to represent the compressor characteristics. These tables represented the mass flow parameter $\dot{m}_{\mathrm{Dim}}$ versus rotational speed parameter N_{Dim}, the pressure ratio P_{o2}/P_{o1} versus rotational speed parameter N_{Dim} and the compressor efficiency η_c versus rotational speed parameter N_{Dim}.

If the values of any parameter with β line parameter are specified, the program searches and picks the other two parameters

Table 3 Mass flow parameter versus rotational speed parameter

Speed Parameter (N_{Dim})	BETA Line (β_{line}) →				
	β_{line_1}	β_{line_2}	β_{line_3}	β_{line_4}	β_{line_i}
N_{Dim_1}					
N_{Dim_2}		Equivalent Mass Flow Parameter ($\dot{m}_{Dim}$) →			
N_{Dim_3}					
N_{Dim_4}					
N_{Dim_i}					

Table 6 Pressure ratio versus mass flow parameter of the compressor surge line

Pressure Ratio (P_{02}/P_{01}) ↓	Mass Flow Parameter ($\dot{m}_{Dim}$) →

from these look-up tables. The other problem with the numerical representation of the compressor map is the surge line. One or both of the following can achieve the solution to this problem:

i. deriving an equation to describe surge line and solving this equation to check if the points are beyond the surge line
ii. storing the surge line in look-up table and then use an interpolation or extrapolation technique to check if the points are beyond the surge line

Either one of the two solutions can be efficient. In this program the second method was used for maintaining the consistency of the whole program (Table 6) shows the pressure ratio P_{o2}/P_{o1} versus mass flow parameter $\dot{m}_{\mathrm{Dim}}$ of the surge line.

A linear interpolation technique was used to estimate the parameters values lying at intermediate points. This method of interpolation is followed in most books of numerical analysis.

3.4 Representation of Turbine Maps in the Computer Program. As in the compressor case, the β lines must be introduced to the turbine maps to solve the problem of converting the maps into tabulated data. This can be seen clear in Fig. 3, where at the chocking condition and same speed parameter, more than one pressure ratio results from the same dimensionless mass flow parameter.

Three-dimensional look-up tables, as in the compressor case, have been developed and used to represent the turbine characteristics. These tables have the same forms as those of the compressor characteristics shown in Table 6. In these tables, linear interpolation technique was also used to compute the values lying at intermediate points.

Table 4 Pressure ratio versus rotational speed parameter

Speed Parameter (N_{Dim})	BETA Line (β_{line}) →				
	β_{line_1}	β_{line_2}	β_{line_3}	β_{line_4}	β_{line_i}
N_{Dim_1}					
N_{Dim_2}		Equivalent Pressure Ratio Parameter (P_{02}/P_{01}) →			
N_{Dim_3}					
N_{Dim_4}					
N_{Dim_i}					

Table 5 Compressor efficiency versus rotational speed parameter

Speed Parameter (N_{Dim})	BETA Line (β_{line}) →				
	β_{line_1}	β_{line_2}	β_{line_3}	β_{line_4}	β_{line_i}
N_{Dim_1}					
N_{Dim_2}		Equivalent Compressor Efficiency Parameter (η_c) →			
N_{Dim_3}					
N_{Dim_4}					
N_{Dim_i}					

3.5 Computer Simulation Program. The computer simulation program uses the components models based on either mathematical equations or performance characteristics to achieve matching between the various components in the gas turbine plant. This matching produces the engine equilibrium running line. The equilibrium running line can be used to calculate the different gas turbine performance parameters.

The principal advantages of gas turbine simulation program would be as follows

1. The computer simulation program can help in investigating the effects of the components performance characteristics on the performance of the complete engine. This investigation can be carried out at the design stage without bearing the cost of manufacturing and testing an expensive prototype.
2. The conceptual designs of the engine can be studied and the choice of particular concept can be made to suit the specified operational requirements.
3. The matching of the components can be explored for the design, off-design, and transient conditions.
4. The simulation program can serve as a valuable tool for investigating the performance of the gas turbine at off-design conditions. This investigation can help in designing an efficient control system for the gas turbine engine for a particular application, such as being a part of the combined power and power (CPP) plant.

The information flow diagram for the simple gas turbine cycle shown in Fig. 6 was used to create a computer simulation program. The flow chart for the program logic is shown in Fig. 8. The program is also suitable for dealing with the simulation of other configurations. But for the sake of brevity the flow charts for each configuration have been omitted.

4 Results and Discussion of Modeling, Matching, and Simulation

The output of the new methodology presented in this work is illustrated graphically in Figs. 9 and 10, which show complete typical performance characteristics of a centrifugal compressor and complete typical performance characteristics of a radial turbine [3], respectively.

In order to match the turbine with the compressor, Figs. 9 and 10 have to be reproduced by introducing the matching parameter

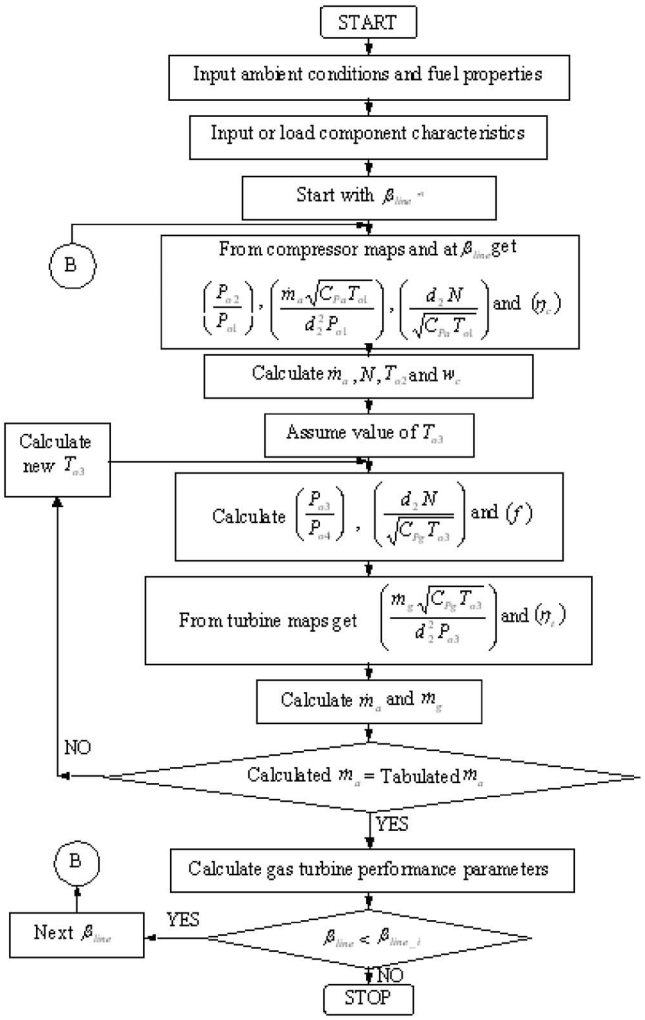


Fig. 8 Computer simulation flowchart

$(\dot{m}N/d_{2c}P_{o1})$. The transformation is shown in Figs. 11 and 12. For the compressor it is worth noting that the constant speed lines were shifted apart, nevertheless the trends stay the same. For the turbine, the trend of the constant speed lines has changed. The reason is because the turbine inlet temperature T_{o3} is not constant along any constant speed line while for the compressor case; the compressor inlet temperature T_{o1} is constant.

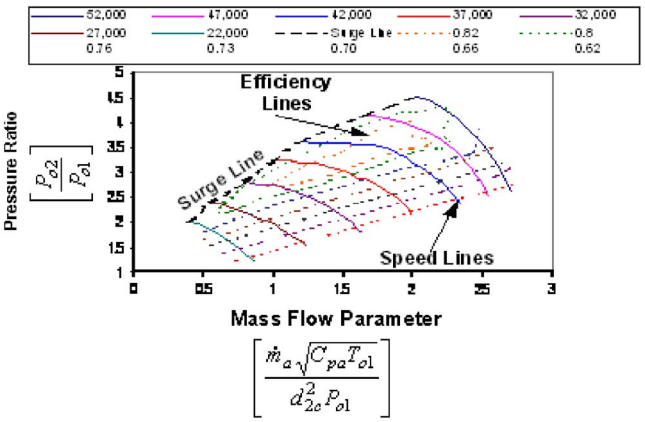


Fig. 9 Complete performance characteristics of a centrifugal compressor [3]

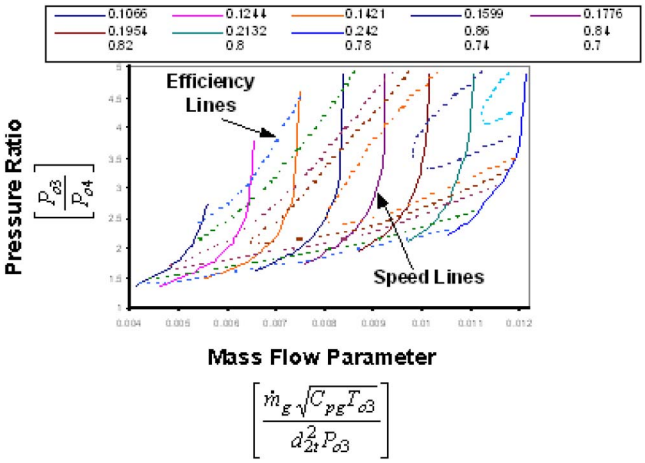


Fig. 10 Complete performance characteristics of a radial turbine [3]

Superimposing Fig. 11 on Fig. 12 produces the complete matching characteristics of the gas turbine performance as depicted in Fig. 13. At any point within the matching range the following parameters can be computed as given in Table 7.

The turbine inlet temperature T_{o3} lines of 650 and 1400 K were computed and drawn in Fig. 13. It can be seen that the changes of constant T_{o3} lines at various pressure ratios are linear and showed

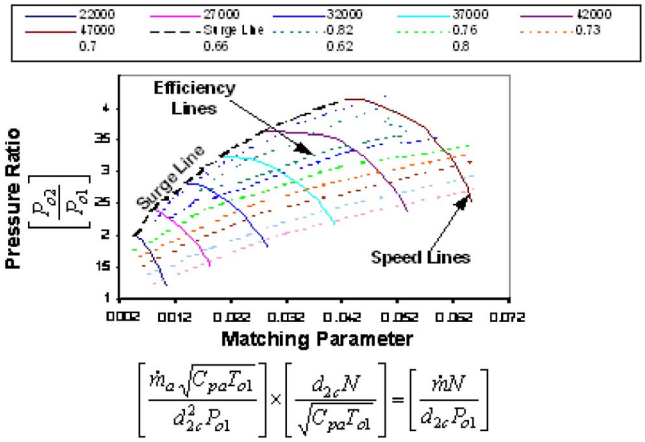


Fig. 11 Transformed performance characteristics of centrifugal compressor

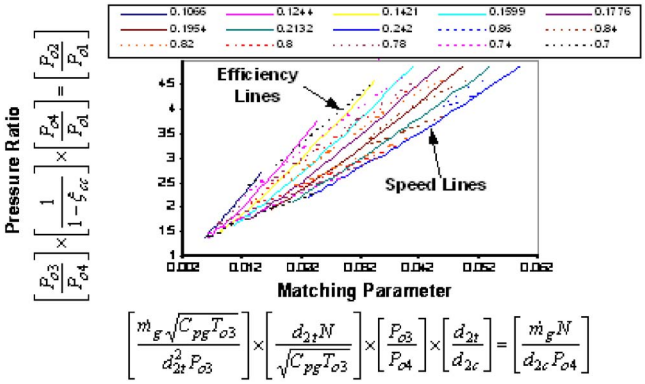


Fig. 12 Transformed performance characteristics of radial turbine

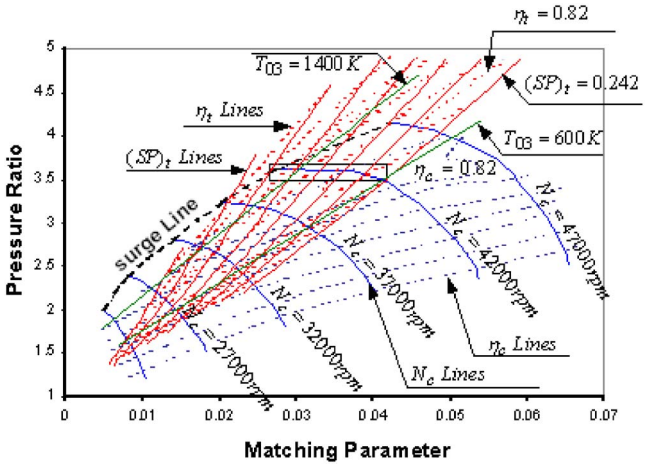


Fig. 13 Complete matching characteristics of the gas turbine performance

divergence at higher values of speed and pressure ratios. The area between these two lines represents the accepted working range for the gas turbine engine.

For a power generation driven by a gas turbine engine, let us consider any running line, for example, a speed of 42,000 rpm. For clearity, part of this running line enclosed by a rectangular box in Fig. 13 was enlarged to produce Fig. 14. Based on the graphical analysis, the above parameters can be calculated within the specified working range. The output results are given in Tables 8 and 9.

Using the results in Tables 8 and 9, the relationship between thermal efficiency and specific fuel consumption with the net power output are drawn in Fig. 15. It shows that the maximum thermal efficiency of 17.57% is attained, which corresponds to a net power output of 228 kW and minimum specific fuel consumption of 0.3327 kg/kW.hr.

Note that the low thermal efficiency is due to the fact that this selected gas turbine engine has a pressure ratio (r) of 4 and turbine inlet temperature ($To3$) of 1000 K.

Figure 16 shows that the gas mass flow rate is decreasing with an increasing net power output. At the same time the gas exhaust temperature is increasing under the same condition. This can be explained because the turbine is considered a constant volumetric flow component. Increasing the turbine work output at constant speed can be achieved by raising the turbine inlet temperature T_{o3}. In order to accommodate the same amount of volumetric mass flow at this higher T_{o3}, the mass flow rate must decrease. Consequently, the gas exhaust temperature increases.

Figure 17 shows the variation of temperatures T_{o2}, T_{o3}, and T_{o4} with the net power output. It can be seen that the variation of T_{o2} is fairly small because T_{o2} depends on the compressor pressure ratio and within the working range of the constant speed of 42,000, the pressure ratio variation is small (see Fig. 14).

Figure 18 shows the relationship between the turbine and compressor torque with net power output of the gas turbine engine at

Table 7 Gas turbine parameters

Turbine inlet temperature (T_{o3})	Gas turbine thermal efficiency (η_{gt})
Compressor power (W_c)	Net output torque (τ_{net})
Turbine power (W_t)	Turbine torque (τ_t)
Net power output (W_{net})	Specific fuel consumption (SFC)
Compressor torque (τ_c)	

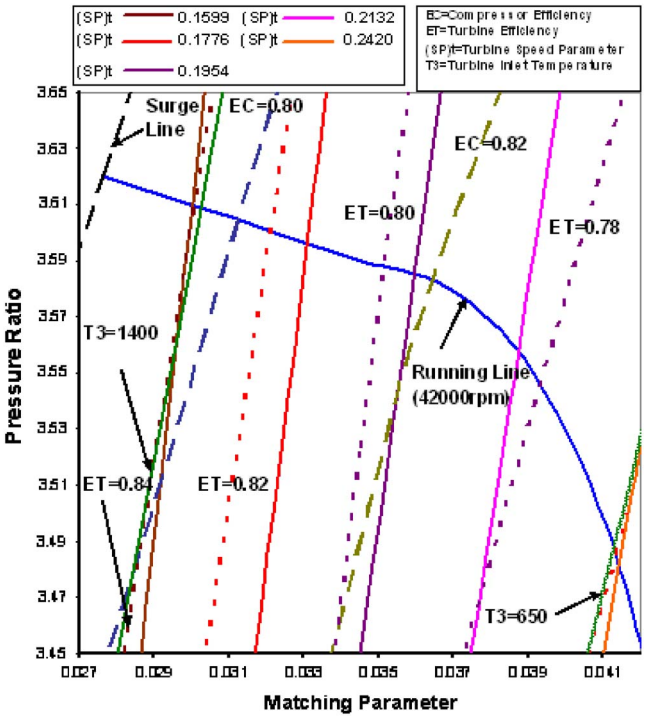


Fig. 14 Matching characteristics at running line of 42,000 rpm

a constant speed of 42,000 rpm. It can be noted that the variation of torque and net power output is linear and increases at a constant rate.

5 Closure

Modeling, matching, and simulation of a gas turbine engine for power generation has been presented. A computer program for simulating a gas turbine engine has been developed that can satisfy the necessary matching conditions analytically and, thus, achieve matching between the various components in order to produce the equilibrium running line. Representing the data for this line either in the form of lookup tables or an equation is known as modeling; solving that equation with the help of a computer is computer simulation. Thus, modeling and simulation together satisfy all energy and mass balances, all equations of state of the working fluids, and the performance characteristics of all components.

Table 8 Calculated parameters within the specified working range of 42,000 rpm

#	Matching Parameter	r	Turbine Speed	η_c
1	0.0413	3.483	0.242	83
2	0.0387	3.558	0.2132	83.2
3	0.0359	3.585	0.1954	81.8
4	0.0331	3.597	0.1776	80.8
5	0.03	3.61	0.1599	79.5
#	W_t	W_{net}	Fuel/Air Ratio	$\dot{m}_f$
1	296.79	28.55	0.00611	0.01095
2	374.59	118.74	0.01143	0.01921
3	415.78	172.66	0.01596	0.02488
4	455.72	228.08	0.02205	0.03169
5	493.48	283.08	0.03036	0.03955

Table 9 Calculated parameters within the specified working range of 42,000 rpm

#	η_t	T_{o2}	T_{o3}	T_{o4}	W_c
1	82	436.82	651.97	508.69	268.24
2	84.5	439.48	840	647.02	255.85
3	84.5	443.17	1000	769.1	243.12
4	82.8	445.58	1210.52	936.02	227.64
5	80	448.7	1493.34	1165.4	210.4
#	η_{gt}	SFC	τ_c	τ_t	τ_{net}
1	6.37	0.7699	203.39	225.04	21.65
2	15.1	0.3464	194	284.04	90.04
3	16.95	0.3327	184.35	315.27	130.92
4	17.57	0.3481	172.61	345.55	172.94
5	17.48	0.3861	159.54	374.19	214.65

The results of component matching, modeling, and simulation are presented in this paper and lead to the following concluding remarks:

i. Matching conditions between the compressor and the turbine may be met by superimposing the turbine performance characteristics on the compressor performance characteristics with suitable transformation of the coordinates.

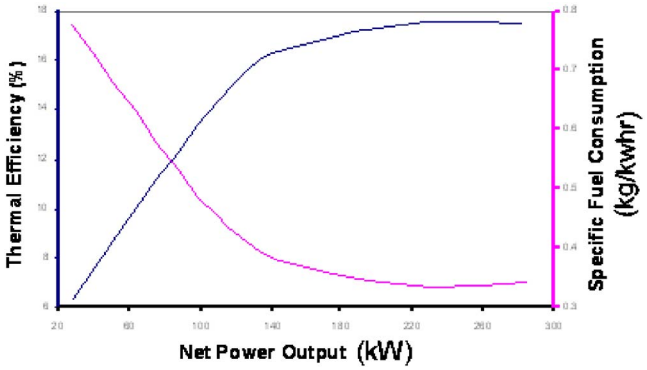


Fig. 15 Variation of thermal efficiency and specific fuel consumption at 42,000 rpm

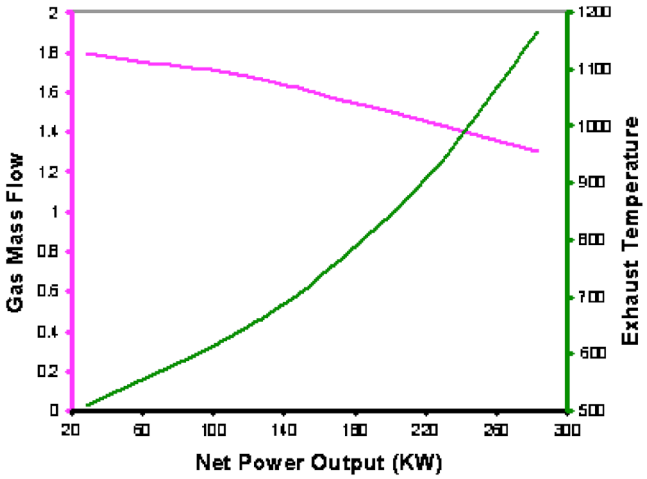


Fig. 16 Variation of exhaust flow and temperature at 42,000 rpm

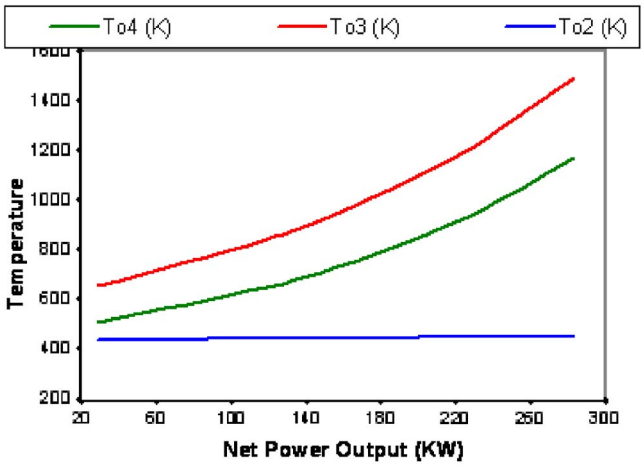


Fig. 17 Variation of various temperatures at 42,000 rpm

ii. Matching technique can be used to determine the following:

- the operating range (envelope) and running line of the matched components
- the proximity of the operating points to the compressor surge line
- the maximum operating point at the maximum turbine inlet temperature (T_{o3})
- most importantly, (from the figures) whether the gas turbine engine is operating in a region of adequate compressor and turbine efficiencies

iii. A computer program has been written for simulating a gas turbine engine. This program basically satisfies the matching conditions analytically between the various components to produce the equilibrium running line. Hence, it can serve as a very useful tool for simulating gas turbine engines. The principal advantages of the gas turbine simulation program are summarized as follows:

- The computer simulation program can help in investigating the effects of the components' performance characteristics on the performance of the complete engine. This investigation can be carried out at the design stage without bearing the cost of manufacturing and testing an expensive prototype.

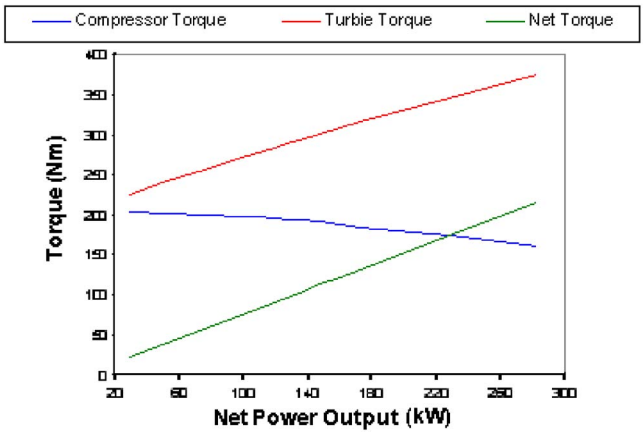


Fig. 18 Variation of torque values at 42,000 rpm

- The conceptual designs of the engine can be studied, and the choice of a particular concept can be made to suit the specified operational requirements.
- The matching of the components can be explored for the design, off-design, and transient conditions.
- The simulation program can serve as a valuable tool for investigating the performance of the gas turbine at off-design conditions. This investigation can help in designing an efficient control system for the gas turbine engine of a particular application, including being identical to the combined power and power (CPP) plant.

Acknowledgment

The authors are especially thankful to Professor F.S Bhinder for his guidance and valuable contribution. The authors would also like to thank the Royal Jordanian Air Force and King Abdullah II Design & Development Bureau for their endless support and encouragement, and the University of Hertfordshire for its cooperation and help throughout this research program.

Nomenclature

C_P = specific heat at constant pressure
C_V = specific heat at constant volume
γ = ratio of specific heats
$\dot{m}$ = mass flow rate
w, W = specific work output, work output
P = pressure
T = temperature
r = pressure ratio
η = efficiency
LCV = lower calorific value
τ = torque
F = fuel-to-air ratio, function
ξ = pressure loss in combustion chamber
d = diameter

Subscripts

1,2,3 = state points in the cycles
gt = gas turbine
s = isentropic
o = stagnation
g = gas
a = air
c = compressor
t = turbine
cc = combustion chamber

Superscript

$\cdot$ = Rate

References

[1] Chappel, M. S., and Cockshutt, E. P., 1974, *Gas Turbine Cycle Calculations: Thermodynamic Data Tables for Air and Combustion Products for Three Systems of Units*, NRC No.14300, Ottawa.
[2] Cohen, H., Rogers, G. F. C., and Saravanamuttoo, H. I. H., 1996, *Gas Turbine Theory*, 4th Edition, Longman, London.
[3] Ainley, D. G., and Mathieson, G. C. R., 1957, *A Method of Performance Estimation for Axial-Flow Turbines*, Aeronautical Research Council.
[4] Kurzke, J., 1996, "How to Get Component Maps for Aircraft Gas Turbine Performance Calculations," ASME paper 96-GT-164.
[5] Gordon, S., and McBride, B., 1994, *Computer Program for Calculation of Complex Chemical Equilibrium Compositions and Applications* NASA Reference Publication 1311, Vols. I and II, Lewis Research Center, Cleveland, Ohio.
[6] Walsh, P., and Fletcher, P., 1998, "Gas Turbine Performance," 1st Edition, Blackwell Science Ltd., London, pp. 175–185, 200, 215.
[7] Bhinder, F. S., and Mango, O. I. K., 1995, "A Parametric Study of the Combined Power and Power (CPP) Plant for Generating Electricity," ASME Turbo Cogeneration. Vienna.
[8] Horlock, J. H., 1997, "Aero-Engine Derivative Gas Turbines for Power Generation: Thermodynamic and Economic Perspectives," ASME J. Eng. Gas Turbines Power, **119**(1), pp. 119–123.
[9] Horlock, J. H., 1995, "Combined Power Plants: Past, Present and Future," ASME J. Eng. Gas Turbines Power, **117**(4), pp. 608–616.
[10] Cerri, G., 1987, "Parametric Analysis of Combined Gas-Steam Cycles," ASME J. Eng. Gas Turbines Power, **109**(1), pp. 46–55.
[11] Harman, R. T. C., 1981, *Gas Turbine Engineering*, First Edition, Macmillan Press Ltd., London, UK.

E. Kakaras
Associate Professor
e-mail: ekak@central.ntua.gr

A. Doukelis
Research Associate
e-mail: adoukel@central.ntua.gr

Laboratory of Steam Boilers and Thermal Plants,
National Technical University of Athens,
Iroon Polytechniou 9,
Athens 15780, Greece

A. Prelipceanu
Process Engineer
e-mail: alexander.prelipceanu@linde-le.com

S. Karellas
Research Associate
e-mail: karellas@ltk.mw.tum.de

Lehrstuhl für Thermische Kraftanlagen,
Technische Universität München,
85747 Garching, Germany

Inlet Air Cooling Methods for Gas Turbine Based Power Plants

Background: Power generation from gas turbines is penalized by a substantial power output loss with increased ambient temperature. By cooling down the gas turbine intake air, the power output penalty can be mitigated. Method of Approach: The purpose of this paper is to review the state of the art in applications for reducing the gas turbine intake air temperature and examine the merits from integration of the different air-cooling methods in gas-turbine-based power plants. Three different intake air-cooling, methods (evaporative cooling, refrigeration cooling, and evaporative cooling of precompressed air) have been applied in two combined cycle power plants and two gas turbine plants. The calculations were performed on a yearly basis of operation, taking into account the time-varying climatic conditions. The economics from integration of the different cooling systems were calculated and compared. Results: The results have demonstrated that the highest incremental electricity generation is realized by absorption intake air-cooling. In terms of the economic performance of the investment, the evaporative cooler has the lowest total cost of incremental electricity generation and the lowest payback period (PB). Concerning the cooling method of pre-compressed air, the results show a significant gain in capacity, but the total cost of incremental electricity generation in this case is the highest. Conclusions: Because of the much higher capacity gain by an absorption chiller system, the evaporative cooler and the absorption chiller system may both be selected for boosting the performance of gas-turbine-based power plants, depending on the prevailing requirements of the plant operator. [DOI: 10.1115/1.2131888]

Introduction

Gas turbines have gained widespread acceptance in the power generation, mechanical drive, and gas transmission markets. Their compactness, high power-to-weight ratio, and ease of installation have made them a popular prime mover. Improvements in hot section materials, cooling technologies, and aerodynamics have allowed increases in firing temperatures. Consequently, thermal efficiencies are currently very attractive, with simple cycle efficiencies ranging between 32 and 42 % and combined cycle efficiencies reaching the 60% mark. Yet, gas turbine output and efficiency is a strong function of the ambient air temperature. Depending on the gas turbine type, power output is reduced by a percentage between 5% and more than 10% of the ISO-rated power output (15 °C) for every 10 K increase in ambient air temperature. At the same time, the specific heat consumption increases by a percentage between 1.5% and more than 4% [1]. The highest losses in gas turbine power output usually coincide with periods of high electricity demand. This poses a significant problem for turbine operators, especially in the rapidly deregulating power generation sector where the structure of supply agreements and the dynamics of an open market mean that power producers are paid significantly more for power generated during high demand periods. Even if this is not the case, during high-demand periods power margins may have been exhausted, and in such cases, augmenting power can have a notable impact on the system's reliability, as well as on the decision for new power plant installations. By the addition of an air-cooling system at the compressor intake, the inlet air can be conditioned to lower temperatures than ambient, thus improving plant performance at high ambient temperatures.

The purpose of this paper is to review the state of the art in applications for reducing the gas turbine intake air temperature in gas-turbine-based power plants. Three different cooling, methods for reducing the intake air temperature are examined: evaporative cooling, refrigeration cooling, and evaporative cooling of precompressed air. Two combined cycle (CC) power plants and two gas turbine plants owned and operated by the Public Power Corporation of Greece (PPC) have been selected as test cases for analysis of the merits from integration of the different gas turbine intake air-cooling methods. Following the discussion on the different cooling methods, the results of the effect of integration of the inlet air-cooling systems under investigation on the operation of the selected power plants are presented and discussed. The calculations are performed on a yearly basis of operation and the time-varying climatic conditions are taken into account. Finally, the economics from integration of the different cooling systems are calculated and compared.

Gas Turbine Intake Air-Cooling Methods

The performance deterioration of gas turbines in terms of power output and efficiency at high ambient air temperatures, coinciding with time periods of high electricity demand, implies that the integration of an intake air-cooling system could be an important consideration to boost the performance of the plant. Two different techniques are mainly available for intake air-cooling in gas-turbine-based power plants: evaporative cooling and refrigeration cooling. A third innovative approach not yet implemented is evaporative cooling of precompressed air. Thermal energy storage systems are not discussed in this paper.

Evaporative Cooling. Evaporative cooling is based on the evaporation of water in the intake air of the gas turbine. As water evaporates, the latent heat of evaporation is absorbed from the surrounding air. As a result, the air is cooled during the process. In the limiting case, the air leaves the cooler at a saturated state. The evaporative cooling process is, essentially, identical to the adiabatic saturation process, since the heat transfer between the air stream and the surroundings is usually negligible. Therefore, the evaporative cooling process follows a line of constant enthalpy on the psychrometric chart.

Traditional evaporative coolers that use media for evaporation

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received July 15, 2004; final manuscript received September 23, 2005. Review conducted by M. Boyce.

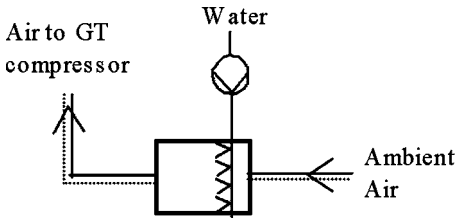


Fig. 1 Schematic layout of direct inlet fogging of air

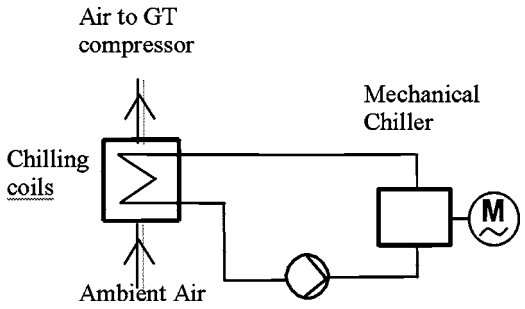


Fig. 2 Schematic layout of inlet aircooling with mechanical chiller and chilling coils

of the water have been widely used in the gas turbine industry over the years, especially in hot arid areas. In this type of cooler, water is distributed over media blocks (which are made of fibrous corrugated material [2]) positioned at the gas turbine intake. The airflow through the blocks evaporates the water. The effectiveness of these coolers, which as defined as the ratio of the temperature difference between inlet and exit of the cooler and the temperature difference if the air left the cooler at a saturated state, is, typically, 80–85%, meaning that the wet bulb temperature can never be attained. In addition, the increased pressure loss caused by the media blocks reduces the power output gain. Water quality requirements, however, can be less stringent than those required for direct fog-cooling systems.

Direct inlet fogging is a method of cooling where demineralized water is converted into a fog by means of high-pressure nozzles operating at 70–200 bar. This fog then provides cooling when it evaporates in the air inlet duct of the gas turbine. This technique, in contrast to the traditional evaporative coolers, allows effectiveness close to 100% to be attained. An elaborate discussion on the aspects of direct inlet fogging can be found in Ref. [3]. A schematic layout of direct inlet fogging for a gas-turbine-based power plant is shown in Fig. 1.

Evaporative cooling is the most commonly employed system for inlet air-cooling because of the simplicity of components and installation and the low equipment cost [4]. Unfortunately, the gain from the use of an evaporative cooler is limited by the difference between the dry and wet bulb temperatures, and it is high for dry ambient conditions, whereas the gain for wet ambient conditions is low.

In order to further increase power augmentation, additional fog nozzles can be installed near the compressor intake. The water droplets evaporate mainly inside the compressor as the air is heated up during compression. These systems are called high-fogging or wet-compression systems. About 6% power gain can be achieved for 1% water mass flow relative to compressor air mass flow [5].

Refrigeration Cooling. Refrigeration cooling is used when cooling below the wet bulb temperature is desired. Both mechanical and absorption refrigeration systems are available for gas turbine intake air-cooling. In a mechanical refrigeration system, the refrigerant vapor is compressed by means of a centrifugal, screw, or reciprocating compressor. Centrifugal compressors are typically used for large systems in excess of 1000 tons and are usually driven by an electric motor. Mechanical refrigeration has high auxiliary power consumption for the compressor driver. This is also a drawback during the peak hours of operation, when electricity demand reaches its maximum. Against it, absorption chillers exploit thermal energy from the turbine exhaust gases. Absorption chiller systems typically employ lithium-bromide (Li-Br) and water, with the Li-Br solution being the absorber and the water acting as the refrigerant. Such systems can typically produce chilled water of 6–7 °C. Recent innovative designs are capable of producing chilled water at temperatures as low as 2 °C [6]. Absorption systems can be designed to be either single or double effect. A single-effect system will have a coefficient of performance (COP) of 0.7–0.8 and a double-effect unit a COP of 1.2–1.3, also depending on the quality of the heat source. Part-load performance of absorption systems is relatively good and efficiency does not drop off at part load as it does with mechanical refrigeration systems. Absorption systems also have lower operating and maintenance costs than mechanical refrigeration systems and very low auxiliary power consumption [7]. The refrigerant produced by the mechanical or absorption chillers is typically circulated to chilling coils at the gas turbine inlet, which cool down the intake air to 3–4 K higher than the refrigerant temperature. A schematic layout of this type of inlet air-cooling system for a gas-turbine-based power plant is shown in Fig. 2. A more innovative approach is the use of a direct-contact air-cooler, which allows a smaller temperature difference between chilled water and cooled air (of the order of 1–2 K), as well as a lower inlet pressure drop with respect to the chilling coils [6]. A schematic layout of this type of inlet air-cooling system utilizing an absorption chiller is shown in Fig. 3.

Evaporative Cooling of Precompressed Air. In this gas turbine inlet air-cooling method, an electrically driven fan is used for precompression of the inlet air supplied to the compressor of the gas turbine, resulting in a temperature rise of the intake air. The precompressed air is then cooled by an evaporative cooler. Evaporative cooling of precompressed air is quite effective because of the higher potential of evaporative cooling for the hotter precompressed airstream. This method can be applied to compensate for the reduction in the air mass flow as a consequence of high ambient temperatures. This can be achieved by controlling the pressure rise of the fan, so that the air mass flow at the inlet of the compressor is maintained largely constant, regardless of ambient conditions. Using this method, the gas turbine can be operated over a long period of time at its optimum operating point with regard to power output and efficiency. Also, more heat can be recovered from the exhaust gases compared to the noncooled case. This means that, for a cogeneration system, steam production can be held constant to meet the required process steam demand. The maximum possible airflow rate is limited by the exhaust condition for the turbine and flow conditions through the compressor (surge

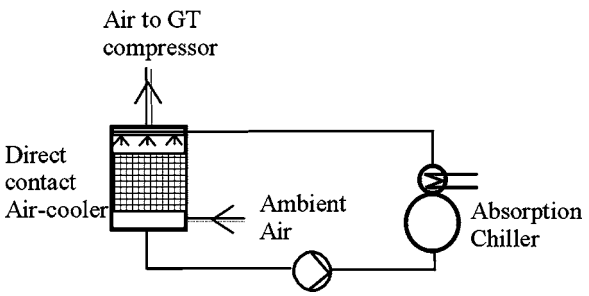


Fig. 3 Schematic layout of inlet airooling with absorption chiller and direct-contact air-cooler

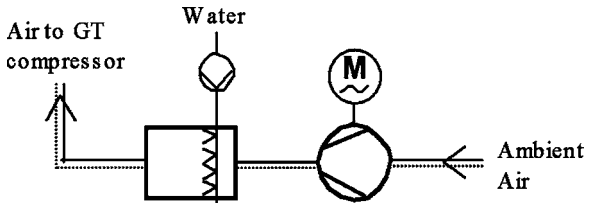


Fig. 4 Schematic layout of evaporative cooling of precompressed air

limit). A description of the modeling of this system has been presented by Nixdorf et al. [8]. A schematic layout of evaporative cooling of precompressed air is shown in Fig. 4.

Gas-Turbine-Based Power Plant Test Cases

In order to examine the merits from integration of the different intake air-cooling methods described above for gas-turbine-based power plants, two combined cycle (CC) power plants and two gas turbine plants owned and operated by PPC have been selected as test cases for analysis. Lavrio IV CC power plant, installed close to Athens, has a gross capacity of 560 MW_{el} and a net efficiency of 50.8% at ISO conditions and consists of three 119 MW_{el} European Gas Turbines SA Model BG-9171E gas turbines, three double-pressure heat recovery steam generators, and one 203 MW_{el} steam turbine. Lavrio V CC will be installed during the several years and will be a 380-400 MW_{el} single-shaft combined cycle. The simple cycle gas turbines selected as test cases are one 21.3 MW_{el} ABB GT10 and one 31.5 MW_{el}LM2500+ installed in the islands of Rhodes and Crete, respectively.

The cycles have been simulated using the thermodynamic cycle calculation software ENBIPRO (ENergie-BIllanz-PROgram) [9]. ENBIPRO is a powerful tool for heat and mass balance solving of complex thermodynamic circuits, calculation of efficiency, and exergetic and exergoeconomic analysis of power plants. The software code models all pieces of equipment that usually appear in power plant installations for both full-load and part-load operation and can accurately calculate all thermodynamic properties (temperature, pressure, enthalpy) at each node of the thermodynamic circuit, power consumption of each component, flue gas composition, etc. The code has proven its validity by accurately simulating a large number of power plants and through comparison of the results with other commercial software [10]. Table 1 presents an output file concerning power output and efficiency calculations for Lavrio V.

The effect of ambient temperature variation on the power output and efficiency of each block has been accurately simulated with ENBIPRO by using the manufacturer data of the blocks that were provided by PPC. For the case of the Lavrio V CC, which will be installed during the next several years, manufacturer data for the 385 MW_{el} GE S109FA single-shaft combined cycle have been used in the current study. The monthly full-load operating hours of the power plants have been estimated as shown in Table 2, using operating data provided by PPC. PPC has also provided mean monthly meteorological data (relative temperature and humidity) from the sites of installation. For the Lavrio combined cycles, typical meteorological year data with hourly values were available. Figure 5 presents the mean monthly dry and wet bulb temperatures for the power plant site in Crete. The relative humidity during the summer months is about 50%. Mean monthly values in the other plant sites are quite similar.

As has been discussed previously, the integration of three different types of gas turbine intake air-cooling systems has been examined: evaporative cooling, refrigeration cooling, and evaporative cooling of precompressed air. In the calculations of direct inlet fogging, 95% effectiveness of the coolers is assumed and the intake air temperature is calculated from the wet bulb temperature derived from the values of ambient temperature and corresponding relative humidity and the cooler effectiveness. Power consumption of the pumps utilized by the system has been taken into account.

Table 1 Sample ENBIPRO power and efficiency output file

Leistungen der einzelnen Arbeitsmaschinen	
Compressor: 0102	−241.113 MW
Gas Turbine: 0105	502.108 MW
Steam Turbine: 0208	59.581 MW
Elec. Drive Pump: 0212	0.000 MW
Elec. Drive Pump: 0602	−0.018 MW
Elec. Drive Pump: 0702	−0.411 MW
Steam Turbine: 0714	50.303 MW
Elec. Drive Pump: 0802	−1.095 MW
Steam Turbine: 0814	28.695 MW
Final Results	
Gas-Turbine/Air-Compressor Block(s)	
Total available power (Pgt - Pcmp):	260.995 MW
Compress.-turbine transm. gear losses (99.0%):	−2.610 MW
GT-generator electrical losses (99.0%):	−2.584 MW
Net Elect. power production from Gas Turbine:	255.801 MW
Net efficiency ratio of gas turbine:	37.226%
Steam-Turbine Block(s)	
Total Available Power (Pst):	138.579 MW
Turbine-Generator transm. gear losses (99.0%):	−1.386 MW
ST-Generator electrical losses (99.0%):	−1.372 MW
Net El. Power Production from Steam-Turbine:	135.821 MW
Net Efficiency Ratio of Steam-Turbine:	19.766%
Auxilliary Equipment (pumps, fans, etc.)	
Calc. power consumption (without losses):	−1.524 MW
Electromechanical losses (99.0%):	−0.015 MW
Total power consumption (with losses):	−1.539 MW
Complete Installation	
Total net electric power production:	390.084 MW
Combusted fuel total heating power:	687.149 MW
Installation net electric efficiency ratio:	56.77%

Concerning the refrigeration cooling method, both mechanical and absorption chiller have been studied. In each power plant test case, the chiller size is selected so that the mean monthly compressor intake air temperature can be maintained at a value of 5 °C throughout the warmest month of the year. In this way, maximum additional electricity generation is achieved during the hot summer months, where electricity demand peaks are observed. The chiller sizes used in the calculations are summarized in Table 3. The same chiller size as for the LM 2500+ is required for the smaller capacity ABB GT 10, since it is a less efficient engine with approximately the same airflow rate.

Table 2 Full-load operating hours of the blaocks

	Operating hours	
Months	**Lavrio IV, V**	**ABB GT10, LM2500+**
January	200	0
February	200	0
March	300	0
April	400	0
May	590	0
June	700	700
July	700	700
August	700	700
September	600	700
October	500	0
November	300	0
December	200	0
Total	**5390**	**2800**

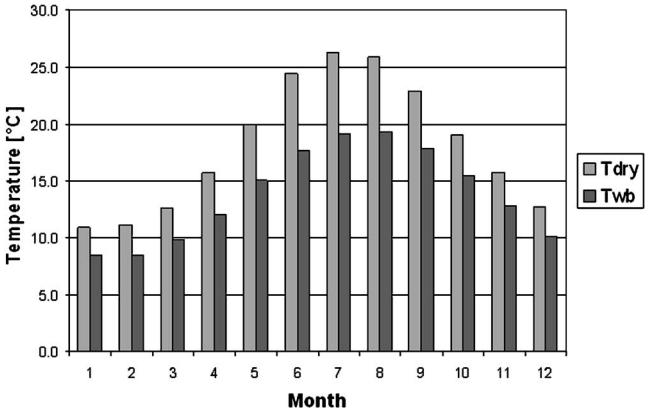


Fig. 5 Monthly dry and wet bulb temperatures for the gas turbine power plant site in Crete

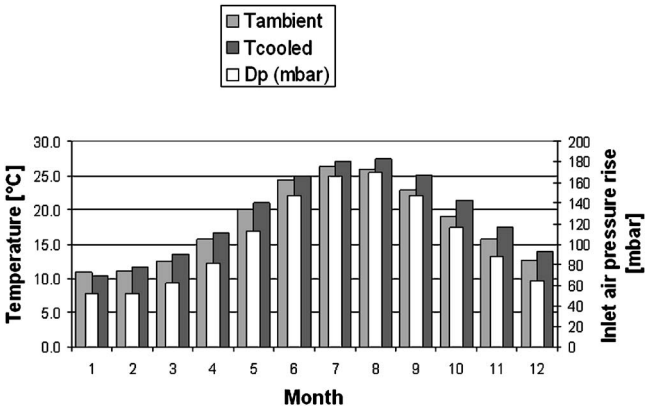


Fig. 6 Ambient and gas turbine inlet air temperature and inlet air pressure rise for the method of evaporative cooling of precompressed air, gas turbine power plant in Crete

The absorption chilling system examined in the current study is an innovative system consisting of a direct-contact air-cooler and a specially designed Li-Br absorption chiller, capable of lowering the intake air temperature to 5 °C. The latter has been analyzed in detail by Bies et al. [1] and Kakaras et al. [11], whereas alternative scenarios of usage of the system for various power plants, representative of different countries and different climatic conditions, have been presented in detail by Kakaras et al. [12], demonstrating its potential. This air-cooling system has been successfully operated within a research and demonstration project [13].

Variation of the COP of mechanical chillers with cooling load was calculated on the basis of data found in Ref. [7], whereas manufacturer data were available on heat consumption and the variation of COP versus load for the absorption chillers. In the combined cycle test cases, where flue gas exit temperature is low, absorption chillers are operated on steam of appropriate thermodynamic parameters extracted from the steam cycle of the power plant. In the simple cycle gas turbine test cases, flue gas from the turbine exhaust produces low-pressure steam in an additional flue gas heat exchanger, which is then used for the operation of the absorption chillers. Manufacturer data on the auxiliary consumption of the chillers were also available. The gas turbine inlet pressure drop resulting from the direct contact air-cooler is assumed 4 mbar and from the chilling coils used for the mechanical chillers as 10 mbar and have been taken into account for the estimation of the capacity gain from integration of the inlet air-cooling system.

For simulation of the method of evaporative cooling of precompressed air, the required pressure rise from the fan was determined in dependency to ambient temperature and relative humidity, targeting an air mass flow equal to the mass flow corresponding to an ambient temperature of 5 °C. The net power has been calculated by subtracting the power required for air precompression. Figure 6 presents the ambient and gas turbine inlet air temperature and the inlet air pressure rise, as calculated for the gas turbine power plant in Crete.

Table 3 Chiller sizes for the test cases

Power plant	Design cooling load [MW]
Lavrio IV CC	3×16
Lavrio V CC	25
Rhodes ABB GT 10	3
Crete LM 2500+	3

Results From Integration of the Inlet Air-Cooling Systems in the Power Plants

The operation of the plants with integration of the intake air-cooling systems under consideration has been evaluated on a yearly basis of operation, using the mean monthly values of temperature and relative humidity available for the sites of installation and the monthly full-load operating hours of the power plants, as depicted in Table 2.

Tables 4–7 present the monthly and annual increase in capacity from integration of the different cooling systems in the four power plants under study, whereas Table 8 presents the mean annual efficiency for different intake air-cooling options. Integration of evaporative cooling of precompressed air has only been examined for the simple cycle gas turbines, to keep the size of the required fans in acceptable limits. In all the examined cases, the absorption intake air-cooling system demonstrates the highest gain in capacity, as well as an increase in efficiency, except for the Lavrio IV case, where there is a slight decrease. According to the results, the gain in capacity when using evaporative cooling of precompressed air is much higher than simple evaporative cooling. The latter results in the lowest gain from the three air-cooling methods. Nevertheless, the efficiency penalty in the case of evaporative cooling of precompressed air is remarkable.

Economic Performance of the Inlet Air-Cooling Systems

From the plant operator's point of view, an important factor for deciding which system should be selected for intake air-cooling,

Table 4 Capacity gain (in %, with respect to noncooled operation) from integration of the different intake air-cooling options in Lavrio IV CC

Month	Evaporative cooling	Absorption chiller	Compression chiller
January	0.89	1.43	0.65
February	0.90	1.19	0.41
March	1.19	2.32	1.56
April	1.71	3.85	3.10
May	2.24	5.52	4.74
June	3.31	7.86	7.08
July	3.97	8.91	8.12
August	3.48	8.84	8.05
September	2.85	7.39	6.61
October	1.63	5.09	4.32
November	1.07	3.22	2.47
December	0.78	1.90	1.15
Mean annual	**2.43 (13,226 kW)**	**5.91 (32,171 kW)**	**5.13 (27,953 kW)**

Table 5 Capacity gain (in percent, with respect to noncooled operation) from integration of the different intake air-cooling options in Lavrio V CC

Month	Evaporative cooling	Absorption chiller	Compression chiller
January	0.41	0.26	–0.52
February	0.38	0.12	–0.66
March	0.73	0.83	0.07
April	1.42	2.14	1.39
May	2.28	3.96	3.18
June	4.03	7.07	6.27
July	5.08	8.58	7.78
August	4.56	8.65	7.83
September	3.43	6.46	5.67
October	1.66	3.52	2.75
November	0.84	1.57	0.82
December	0.45	0.54	–0.22
Mean annual	**2.75 (10,297 kW)**	**4.86 (18,225 kW)**	**4.08 (15,282 kW)**

apart from the gain in electricity generation that has been presented in the previous section, is the economic performance of the investment. In order to assess the economic performance of the three gas turbine inlet air-cooling systems under consideration, the basic economic indicators of the investments, namely, the payback period (PB) and the total cost of incremental electricity generation

Table 6 Capacity gain (in percent, with respect to noncooled operation) from integration of the different intake air-cooling options in the ABB GT10

Month	Evap. cooling	Absorption chiller	Compression chiller	Evaporative cooling of precompressed air
January	-	-	-	-
February	-	-	-	-
March	-	-	-	-
April	-	-	-	-
May	-	-	-	-
June	4.30	9.43	5.98	5.10
July	4.91	10.96	7.15	5.84
August	4.45	10.59	6.74	5.41
September	3.08	8.21	4.75	3.94
October	-	-	-	-
November	-	-	-	-
December	-	-	-	-
Mean annual	**4.18 (852 kW)**	**9.80 (1995 kW)**	**6.15 (1253 kW)**	**5.07 (1033 kW)**

Table 7 Capacity gain (in %, with respect ot noncooled operation) from integration of the different intake air-cooling methods in the LM2500+

Month	Evap. cooling	Absorption chiller	Compression chiller	Evaporative cooling of precompressed air
January	-	-	-	-
February	-	-	-	-
March	-	-	-	-
April	-	-	-	-
May	-	-	-	-
June	7.52	16.94	13.74	13.49
July	8.72	19.35	15.94	15.14
August	7.89	18.67	15.22	14.42
September	5.35	14.72	11.53	11.34
October	-	-	-	-
November	-	-	-	-
December	-	-	-	-
Mean annual	**7.37 (2088 kW)**	**17.42 (4939 kW)**	**14.11 (3999 kW)**	**13.60 (3855 kW)**

Table 8 Mean annual efficiency resulting from integration of the different intake air-cooling methods

	Lavrio IV	Lavrio V	ABB GT10	LM-2500+
Without cooling	49.70	55.82	32.49	36.62
Evaporative cooling	49.67	56.14	33.03	37.23
Absorption chiller	49.46	55.97	33.26	37.73
Compression chiller	48.88	55.26	32.43	36.99
Evaporative cooling of precompressed air			31.65	32.05

(ICOE) have been calculated. ICOE is calculated by dividing the difference in total electricity generation costs after and before integration of the gas turbine inlet air-cooling systems with the incremental electricity generation (MWh) due to integration of the cooling systems. PB is then obtained from the equation PB $=[\text{Investment}/\text{IEG}(\text{POE}-\text{ICOE})]$. The mean fuel price and mean electricity selling price (POE) have been provided by the plants operator. The specific costs for each cooling system used in the calculations are presented in Table 9, as derived from a market survey.

Table 10 presents the results of the economic assessment of the integration of the different intake air-cooling systems in the power plants under investigation in terms of mean annual gain in capacity, total cost of incremental electricity generation and PB of investment.

In terms of the total cost of incremental electricity generation, the results demonstrate that the evaporative cooler has the lowest cost, followed by the absorption chiller. The highest electricity generation cost arises from integration of the evaporative cooling system of precompressed air, due to the high-efficiency penalty related to operation of the air precompressor. The absorption cool-

Table 9 Specific costs for each cooling method in € (euro) per kilowatt cooling load (respectively, kilograms per second air mass flow for the fan)

Cooling method	Specific costs
Evaporative cooling	60 [€/kW_C]
Compression chiller	75 [€/kW_C]
Absorption chiller	250 [€/kW_C]
Fan for air precompression	4500 [€/kg/s]

Table 10 Economics of the integration of the intake air-cooling systems in Chania CCPP (mean electricity selling price =88 euro/MWh)

Cooling system	Capacity gain (KW)	ICOE (Euros/MWh)	PB (Years)
Lavrio IV evaporative cooling	13,226	35.76	0.2
Lavrio IV absorption chiller	32,171	37.92	1.4
Lavrio IV compression chiller	27,953	46.96	0.6
Lavrio V evaporative cooling	10,297	24.39	0.1
Lavrio V absorption chiller	18,225	29.20	1.1
Lavrio V compression chiller	15,282	39.03	0.5
ABB GT10 evaporative cooling	852	58.92	0.5
ABB GT10 absorption chiller	1995	74.06	9.6
ABB GT10 compression chiller	1253	103.05	-
ABB GT10 evaporative cooling of precompressed air	1033	155.08	-
LM-2500+ evaporative cooling	2088	67.62	0.3
LM-2500+ absorption chiller	4939	71.01	3.2
LM-2500+ compression chiller	3999	81.58	3.1
LM-2500+ evaporative cooling of precompressed air	3855	194.43	-

ing system presents the highest capacity gain, but also the highest investment costs and, thus, PBs are much higher than evaporative systems and mechanical chillers.

In deciding among the alternative options for gas turbine intake air-cooling, both the economics of the investment and the capacity gain have to be considered from the plant operator. In the above-presented results, evaporative cooling is better in terms of total cost of incremental electricity generation and PB. If economics were the only consideration, the evaporative cooler would be the selected option. Yet, the absorption chiller system yields more than double the gain in mean annual capacity. This is an important parameter for a power producer in countries like Greece, since the summer period is when the electricity peak demand occurs and the additional capacity boost is strongly needed. It should also be mentioned here that all calculations have been conducted assuming that the electricity tariff is uniform throughout the year, as is currently the case in Greece. If the electricity tariff were higher for peak-demand periods, as may be the case in the future, the economics of the investment could prove more favorable for the absorption chiller. It can, therefore, be concluded that selection between the evaporative cooler and the absorption chiller system for boosting the performance of gas-turbine-based power plants merits close attention.

Conclusions

In the current study, three different cooling methods for reducing the intake air temperature in gas-turbine-based power plants have been examined: evaporative cooling, refrigeration cooling, and evaporative cooling of precompressed air. Two combined cycle power plants and two gas turbine plants owned and operated by the Public Power Corporation of Greece (PPC) have been selected as test cases for analysis of the merits from integration of the different gas turbine intake air-cooling methods. Following the discussion on the different cooling methods, the results of the effect of integration of the inlet air-cooling systems under investigation on the operation of the selected power plants were presented and discussed. The calculations were performed on a yearly basis of operation and the time-varying climatic conditions were taken into account. Finally, the economics from integration of the different cooling systems were calculated and compared.

The results have demonstrated that the highest incremental electricity generation is realised by absorption intake air-cooling. In terms of the economic performance of the investment, the evaporative cooler has the lowest total cost of incremental electricity generation and the lowest PB. Concerning the cooling method of precompressed air, the results show a significant gain in capacity, but the total cost of incremental electricity generation in this case is the highest. Because of the much higher capacity gain by an absorption chiller system, the evaporative cooler and the absorption chiller system may both be selected for boosting the performance of gas-turbine-based power plants, depending on the prevailing requirements of the plant operator.

Acknowledgment

The current work has been conducted within the framework of the EU project EI/1105/98, "The cooled gas turbine—a demonstration of innovative gas turbine technology," and the assistance of EU is gratefully acknowledged.

Nomenclature

- CC = combined cycle
- COP = coefficient of performance
- ICOE = total cost of incremental electricity generation
- IEG = Incremental Electricity Generation
- Li-Br = lithium-bromide
- PB = Payback period
- POE = Mean electricity selling price
- PPC = Public Power Corporation of Greece
- $T_{ambient}$ = ambient air temperature
- T_{cooled} = gas turbine inlet air temperature

References

[1] Bies, D., Johanntgen, U., and Scharfe, J., 1999, "Optimised Cooling of the Compressor Intake Air-A New Way for the Improvement of Power and Efficiency in Gas Turbine Plants," *Proc. of International Gas Turbine Congress*, Kobe, Japan, pp. 429–436.

[2] Johnson, R. S., 1988, "The Theory and Operation of Evaporative Coolers for Industrial Gas Turbine Installations," ASME Paper No. 88-GT-41.

[3] Meher-Homji, C. B., 1999, "Gas Turbine Power Augmentation by Fogging of Inlet Air," *Proc. of 28th Turbomachinery Symposium*, Turbomachinery Laboratory, College Station, TX, pp. 92–113.

[4] Kraneis, W., 2000, "The Increased Importance of Evaporative Coolers for Gas Turbine and Combined-Cycle Power Plants," *VGB PowerTech 2000*, VGB, Essen, Vd. 9, pp. 36–40.

[5] Lecheler, S., and Hoffmann, J., 2003, "The Power of Water in Gas Turbines: ALSTOM'S Experience With Air Inlet Cooling," *Proc. of PowerGen Latin America*, Sao Paulo, PennWell, Tulsa, OK.

[6] Bies, D., and Prelipceanu, A., 2003, "Optimised Cooling of the Intake Air by Absorption Refrigeration Technology - An Innovative Technology to Improve Gas Turbine Performance," *Proc. of 1 International Conference on Industrial Gas Turbine Technologies*, European Communities, Brussels, pp. 460–470.

[7] Petchers, N., 2003, *Combined Heating, Cooling & Power Handbook: Technologies and Applications, An Integrated Approach to Energy Resource Optimisation*, Fairmont Press, Lilburn, GA.

[8] Nixdorf, M., Prelipceanu, A., and Hein, D., 2002, "Thermo-Economic Analysis of Inlet Air Conditioning Method of a Cogeneration Gas Turbine Plant," ASME Paper No. GT-2002-30561.

[9] Stamatelopoulos, G. N., 1996, "Berechnung und Optimierung von Kraftwerkskreisläufen," VDI-Fortschrittsberichte Reihe 6, Nr. 340, Braunchweig.

[10] Institut für Wärme und Brennstofftechnik, TU Braunschweig, http://www.wbt.ing.tu-bs.de.

[11] Kakaras, E., Doukelis, A., and Karellas, S., 2004, "Compressor Intake Air Cooling in Gas Turbine Plants," Energy–The International J., **29**(12-15).

[12] Kakaras, E., Doukelis, A., and Scharfe, J., 2001, "Applications of Gas Turbine Plants With Cooled Compressor Intake Air," ASME Paper No. 2001-GT-110.

[13] European Commission, 2003, "The Cooled Gas Turbine-A Demonstration of Innovative Gas turbine Technology," Internal Technical Report, EI/1105/98 Thermie B, project funded by the European Commission.

Gas Turbine Engine Durability Impacts of High Fuel-Air Ratio Combustors: Near Wall Reaction Effects on Film-Cooled Backward-Facing Step Heat Transfer

David W. Milanes
Graduate Research Assistant

Daniel R. Kirk
Postdoctoral Associate

Krzysztof J. Fidkowski
Graduate Research Assistant

Ian A. Waitz
Professor and Deputy Head

Gas Turbine Laboratory,
Department of Aeronautics and Astronautics,
Massachusetts Institute of Technology,
Cambridge, MA 02139

As commercial and military aircraft engines approach higher total temperatures and increasing overall fuel-to-air ratios, the potential for significant chemical reactions to occur downstream of the combustor is increased. This may take place when partially reacted species leave the combustor and encounter film-cooled surfaces. One common feature on turbine endwalls is a step between various engine components and seals. Such step features produce recirculating flows which when in the vicinity of film-cooled surfaces may lead to particularly severe reaction zones due to long fluid residence times. The objective of this paper is to study and quantify the surface heat transfer implications of such reacting regions. A shock tube experiment was employed to generate short duration, high temperature (1000–2800 K) and pressure (6 atm) flows over a film-cooled backward-facing step. The test article contained two sets of 35 deg film cooling holes located downstream of a step. The film-cooling holes could be supplied with different gases, one side using air and the other nitrogen allowing for simultaneous testing of reacting and inert cooling gases. A mixture of ethylene and argon provided a fuel-rich free stream that reacted with the air film resulting in near wall reactions. The relative increase in surface heat flux due to near wall reactions was investigated over a range of fuel levels, momentum blowing ratios (0.5–2.0), and Damköhler numbers (ratio of characteristic flow time to chemical time) from near zero to 30. The experimental results show that for conditions relevant for future engine technology, adiabatic flame temperatures can be approached along the wall downstream of the step leading to potentially significant increases in surface heat flux. A computational study was also performed to investigate the effects of cooling-jet blowing ratio on chemical reactions behind the film-cooled step. The blowing ratio was found to be an important parameter governing the flow structure behind the backward-facing step, and controlling the characteristics of chemical-reactions by altering the local equivalence ratio. [DOI: 10.1115/1.2056532]

1 Introduction

As specific thrust requirements of aircraft engines increase, higher turbine inlet temperatures will be employed. Higher fuel-air ratios, therefore, will be needed to meet future aircraft performance requirements. This leads to the potential for residual fuel present in the freestream to react with turbine film-cooling flows. These reactions would be promoted by the long fluid residence times that characterize the recirculation region behind backward-facing steps, such as the lip of a turbine endwall. This high fuel-air ratio phenomenon can be divided into three phases:

1. exhaust of energetic species from the combustor
2. migration of the energetic species downstream into the turbine
3. mixing of the energetic species with turbine film-cooling flows

Phase One: Species Exhaust. Increasing fuel-air ratios can lead to energetic species exiting the combustor as shown by Lukachko et al. [1]. Both incomplete mixing and insufficient time to complete reactions can lead to the exhaust of low-molecular-weight hydrocarbon species. The higher temperatures associated with high fuel-air ratios promote the production of reactive dissociation products, such as CO and OH (even under conditions of thermodynamic equilibrium and complete mixing). All of these reactive species are potential by-products of high-fuel-air-ratio combustion and fuel for possible downstream reactions.

Phase Two: Downstream Migration. After leaving the combustor, the reactive species migrate through the turbine where they may continue to mix with leaner regions also present in the free stream. As the reactive species and oxidizer mix, chemical reactions are initiated. Lukachko et al. [1] found the characteristic width of the reactive species streak to be an important parameter governing the degree to which chemical reactions are completed during the migration from the combustor through the turbine. Reactive streaks exiting the combustor with widths greater than 10–20% of the radial passage height are large enough to prevent full mixing and reaction before the streak reaches film-cooled surfaces in the turbine.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. 2004-GT-53259.

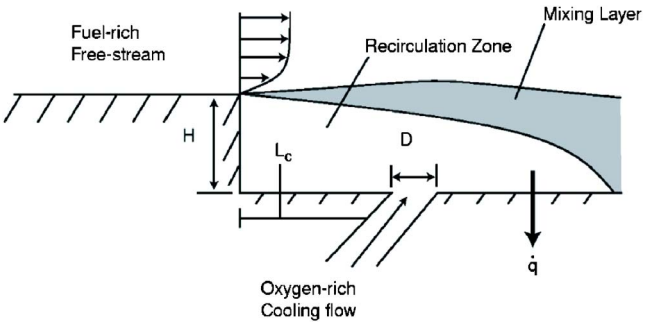


Fig. 1 Backward-facing step geometry

Phase Three: Near-Wall Mixing. As reactive streaks pass through the turbine, mixing between the fuel-rich streaks and oxidizer-rich turbine film-cooling flows can occur. Given the right turbine fluid dynamic and thermodynamic conditions, the time required for chemical reactions to complete may be less than the time during which the fuel and oxidizer streaks are in contact. Thus, the possibility of near-wall chemical reactions arises leading to elevated heat flux to turbine components.

High free-stream pressure and temperature increase the chemical reaction rate and, therefore, reduce the characteristic chemical reaction time, increasing the potential for reactions. If the length of time during which the fuel and oxidizer streaks are in contact is extended, then chemical reactions are also promoted. The recirculation zone behind a backward-facing step is characterized by long fluid residence times. These long fluid residence times provide an increased contact time for any fuel and oxidizer streaks that enter the recirculation zone. Thus, a backward-facing step can promote chemical reactions.

2 Backward Facing Step Overview

Figure 1 shows a typical backward-facing step geometry. The two flow regions of greatest interest for assessing the potential for reaction are the mixing layer and the recirculation zone. Basic details of the flow structure and heat transfer characteristics behind the backward facing step are presented in the following subsections. More details on the flow structure associated with a backward facing step can be found in Ref. [2–4].

The Reynolds analogy does not hold in the recirculation region of backward-facing step flow [5]. Therefore, experiments or computational simulations are required to determine convective heat transfer coefficients downstream of backward-facing steps. Research such as that of Vogel and Eaton [4] shows that the convective heat transfer coefficient increases with downstream distance from the step face until a maximum is reached at the reattachment point—the point at which the separated free-stream flow reattaches to the wall downstream of the step face. This value is typically on the order of 1.2 times the convection coefficient just upstream of the step face [5].

2.1 Film-Cooled Backward-Facing Step Flow Structures. Reference [6] discusses the flow structures behind a film-cooled backward-facing step. If the momentum blowing ratio I is sufficiently small, then the cooling jet does not have enough momentum to alter the structure of the recirculation zone. As I increases, the cooling jet has enough momentum to alter the recirculating flow, but does not have sufficient momentum to pass through the shear layer that forms the border between the recirculation zone and the free stream. If I is increased even further, then the cooling jet has sufficient momentum to pass through both the recirculating fluid near the wall and the shear layer.

As shown by Harinaldi et al. [7], the influence of a cooling jet on the structure of the recirculation zone not only depends on the momentum blowing ratio, but also on the distance from the step face to the cooling hole. As discussed in Sec. 2.2, this distance is captured in the normalized cooling distance parameter.

Table 1 Governing Nondimensional Parameters For Step Reacting Flow Study

$$\mathrm{Re}_H = \frac{\rho u_\infty H}{\mu} \text{ step height Reynolds number} \quad (1)$$

$$\mathrm{ER} = \frac{W_2}{W_1} \text{ expansion ratio} \quad (2)$$

$$\delta_H^* = \frac{\delta}{H} \text{ normalized boundary layer thickness} \quad (3)$$

$$I = \frac{\rho_c u_c^2}{\rho_\infty u_\infty^2} \text{ momentum blowing ratio} \quad (4)$$

$$D^* = \frac{H}{D} \text{ normalized step height} \quad (5)$$

$$L_c^* = \frac{H}{L_c} \text{ normalized cooling distance} \quad (6)$$

$$Q_s = \frac{q_{\mathrm{hot}} - q_{\mathrm{cold}}}{q_{\max} - q_{\mathrm{cold}}} \text{ scaled heat flux} \quad (7)$$

$$\mathrm{Da} = \frac{\tau_{\mathrm{flow}}}{\tau_{\mathrm{chem}}} \text{ Damköhler number} \quad (8)$$

$$H^* = \frac{\Delta H}{H_\infty} \text{ heat-release potential} \quad (9)$$

$$B = \frac{\rho_c u_c}{\rho_\infty u_\infty} \text{ mass blowing ratio} \quad (10)$$

2.2 Governing Nondimensional Parameters and Data Scaling. A set of nondimensional parameters that govern reacting film-cooled backward-facing step flow has been formulated. A complete parameter set is derived by superposition of three sets of backward-facing step governing flow parameters. The first set, developed by Adams et al. [2], governs uncooled, nonreacting backward-facing step flow. Coolant injection effects are accounted for by adding a second set of parameters proposed by Harinaldi et al. [7]. Finally, effects because of near-wall reactions are captured through the nondimensional parameters developed by Kirk et al. [8] for reacting film-cooled flows. Table 1 summarizes the important nondimensional parameters used in this paper to perform the step-reacting flow studies. More details on these parameters may be found in Refs. [1,6,8]. Discussed below are the parameters that are unique to this investigation and not commonly found in the literature concerning backward-facing step fluid mechanics and heat transfer.

2.2.1 Film-Cooled Reacting Flat Plate Parameters. Kirk et al. [8] proposed a set of nondimensional parameters that govern reacting film-cooled flat-plate heat transfer. These parameters are summarized below:

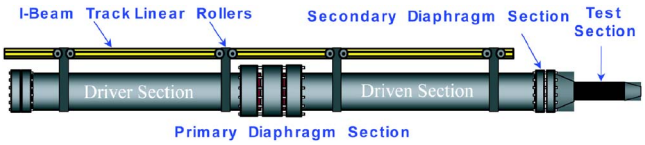


Fig. 2 Shock tube on linear rollers

Scaled heat flux Qs. This dimensionless parameter measures the augmentation in wall heat flux due to chemical reactions. The scaled heat flux *Qs* can range from 0% (if $q_{hot}=q_{cold}$ and no reactions occur), to 100% (if $q_{hot}=q_{max}$), and the reactions proceed until the adiabatic flame temperature is achieved. The scaled heat flux is a function of other dimensionless parameters, including mass blowing ratio, heat release potential, and Damköhler number.

Damköhler number. The Damköhler number Da compares the characteristic flow time to the characteristic chemical time and can be written as

$$\mathrm{Da} = \frac{\tau_{\mathrm{flow}}}{\tau_{\mathrm{chem}}} = \frac{L/u_\infty}{\tau_{\mathrm{chem}}} = \frac{X_r/M\sqrt{\gamma RT_\infty}}{\tau_{\mathrm{chem}}} \tag{11}$$

If this ratio is larger than unity, then there is enough time to permit chemical reactions. When Da is less than unity, chemical reactions will not have enough time to complete. For this study the characteristic length scale is selected to be the step reattachment distance—typically 6–8 step heights downstream of the step face for backward-facing step flow.

Heat-release potential. The energy content of the free-stream flow is measured by the heat-release potential H^*:

$$H^* = \frac{\Delta H}{H_{f\infty}} \approx \frac{T_{\mathrm{ad}} - T_{f\infty}}{T_{f\infty}} \tag{12}$$

Thus, H^* is the ratio of the potential increase in enthalpy of the free stream due to reaction relative to the free-stream enthalpy. If C_p is constant, then H^* can be approximately expressed as the ratio of the maximum total temperature change to the free-stream total temperature. The maximum temperature change due to chemical reaction occurs when the reaction proceeds until the adiabatic flame temperature is achieved. More details on the formulation and application of the non-dimensional heat release potential can be found in Ref. [1,6,8].

3 Transient Shock Tube Testing and Experimental Apparatus

To simulate conditions in which near wall reactions may occur a simplified experiment was devised and implemented. The purpose of this section is to discuss the experimental apparatus, relevant time scales, and data acquisition.

3.1 Shock Tube Experiment and Facility Overview. The shock tube provides a flexible, cost-effective facility for exploring a wide range of flow-chemistry interactions at the high temperatures associated with modern combustors and turbines. Sufficient time must be allowed for a quasi-steady flow to develop over the film-cooled flat plate so that relevant heat flux measurements may be obtained. However, the short run times are advantageous because the test specimen is exposed to high-temperature flow for only a fraction of a second. The gas composition and test gas pressure are regulated to yield different free-stream stagnation temperatures and pressures. Similar transient techniques have been used to make non-reactive heat transfer measurements [9] and reactive measurements over a film-cooled flat plate [8].

The shock tube used in this experiment consists of a 7.3 m driven section and 8.4 m driver, as shown in Fig. 2, both constructed from 30 cm dia stainless-steel pipe. Free-stream pressure

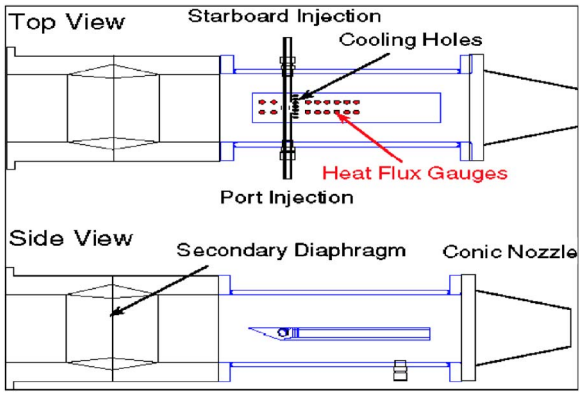


Fig. 3 Test section with flat-plate test article

was measured by four pressure transducers located in the driven section, which were also used to measure shock speed and infer total temperature. Argon was used as the test gas because of its high specific heat ratio, allowing for test temperatures up to 2800 K. Furthermore, it is inert and will not react with the seeded fuel. Additional details concerning shock tubes and the MIT shock tube facility are found in Ref. [10].

Figure 3 shows the top and side views of the test section attached to the driven end of the shock tube. A shock wave, generated upstream where the driver and driven sections meet, travels down the driven section producing a high-enthalpy test gas. When it reaches the end of the driven tube, it ruptures a secondary diaphragm initiating flow over the test article. The area of the conic nozzle downstream of the test article is selected to achieve a desired free-stream Mach number.

3.2 Test Time Considerations. Calculations were performed to ensure that there would be ample time for the starting of flow through the test section and sufficient boundary-layer development over the plate during the test. Table 1 summarizes the duration of the test, flow speed, and flow time over the range of temperatures studied. Furthermore, it was shown that there is sufficient time for the boundary layer to reach a quasi-steady state [10,11].

A primary objective of the experiments was to probe a range of Damköhler numbers, as defined above, representative of conditions in current and future gas turbine technology. Ethylene, C_2H_4, was chosen as the fuel because of its relatively short ignition time and ability to satisfy the shock-matching condition [10].

Chemical ignition times for ethylene were estimated over the range of temperatures used in the experiment. Table 2 presents the time to 95% completion, which was taken as ten times the chemical ignition time [1]. Examining the ratio of the flow to chemical time indicates that the chemical times are much shorter than the available test time. Although the shock tunnel tests are short in duration, they are sufficiently long for full heat release of the fuel to take place.

Using the definitions in Eq. (8) and ethylene as the fuel, Damköhler numbers that range from 0 to 30 can be examined using the facility. In the experiment the Damköhler number was controlled through changing the freestream temperature. (In the

Table 2 Test time-scale summary, ethylene fuel, $\Phi=1.0$, $P=6$ atm

T(K)	τ_{test}	u_{flow}	τ_{flow}	$C_2H_4\tau_{\mathrm{chem95}}$	Da (ms)
1000	40	176	0.14	30	0.005
1900	22	243	0.11	0.08	1.4
2800	8	295	0.086	3e-3	28.7

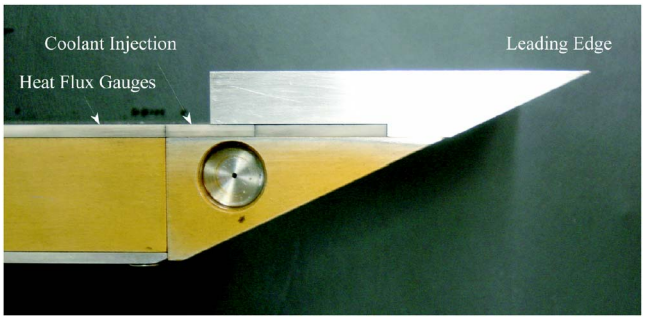


Fig. 4 Side view of step model

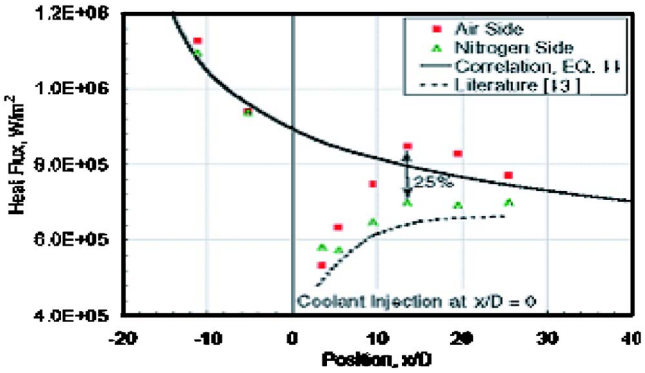


Fig. 5 Augmented flat-plate heat flux due to local reaction, $B=1.0$, $H^*=0.18$, Da=20

numerator, T_∞ sets the free-stream velocity since the test section Mach number is fixed at 0.3. In the denominator, for constant pressure and $\Phi=1.0$, the chemical time is only a function of temperature.)

The range of nondimensional fuel enthalpy H^* studied in this investigation was 0.005–0.80. In the experiments the value of H^* increased with decreasing Damköhler number because the fuel mass fraction of the free stream was kept constant for each case. The results of the reacting flow tests will be shown in Sec. 4.3.

3.3 Test Article Description. Test articles were designed to withstand the temperature and pressure conditions inside the test section of the shock tunnel, to replicate film-cooling flow conditions, and to allow high-frequency heat flux measurements. Figure 4 shows the test plate used in the reacting film-cooling experiments. The plate is 7.62×25.4 cm with a leading-edge angle of 25 deg. The leading edge of the test plate was made of aluminum. The surface of the flat plate was covered with a 0.254 cm thick sheet of MACOR; the low thermal conductivity of MACOR facilitated surface heat flux measurements. A spanwise row of eight 0.254 cm dia cooling jets are located at a distance of 5.59 cm from the plate leading edge. The center-to-center cooling hole spacing is 0.762 cm. The cooling jets, inclined at an angle of 35 deg, are grouped into two sets of four, with the plate's line of symmetry dividing the two groups. Different gases can be supplied to each group allowing for simultaneous testing of reacting and inert cooling gases. High-frequency temperature gages located behind the cooling holes enable rapid sampling of the plate surface temperature.

An existing film-cooled flat plate was modified by attaching a step piece to the leading edge as shown in Fig. 4. The step height is 1.02 cm, and the distance from the step face to the center of the cooling holes is 0.51 cm. These values corresponded to $L_c^*=2$ and $D^*=4$.

Capturing temperature measurements during the short duration tests required the use of high-frequency response instrumentation. Platinum RTD gages were used in conjunction with an ADTEK MMADCAP data acquisition system to provide plate surface measurements. The gages are made of 0.32 cm MACOR cylinders with a platinum resistance element painted on the tip of the cylinder. The thickness of the resistance element is small enough so that the gage provides negligible conduction heat transfer resistance. Therefore, the gage temperature directly indicates the plate surface temperature. Using the ACQ processing method described in Ref. [6] and in more detail by Vidal [12], the temperature versus time profile can be integrated to determine surface heat flux if the material properties of the surface are known. The sampling rate was 100 kHz.

3.4 Experimental Uncertainty. The two largest sources of uncertainty in the reacting step experiments are the free-stream temperature T_∞, and the convective heat transfer coefficient h. Uncertainty in h only affects the calculation of scaled heat flux, but uncertainty in the free-stream temperature propagates through the calculation of quantities, such as Da and Re_H. The free-stream temperature is known to within 10% using ideal shock theory for each measured shock speed, as detailed by Kirk et al. [8]. The convection coefficient is taken from [5] with an estimated uncertainty of 10% due to the presence of a film-cooling jet. Other sources of uncertainty are the location of the reattachment point and the value of measured heat flux. The latter, as Kirk et al. [8] suggests, is known to within 2%. The former is located at the point of highest heat flux [4] and is known to within half the distances between consecutive heat-flux gages. These errors are shown using 95% confidence intervals on the figures that include experimentally derived quantities.

4 Test Description and Experimental Results

This section presents the experimental results obtained from the backward-facing step experiments for cases with and without blowing, as well as with and without reaction taking place.

4.1 Backward-Facing Step Validation Tests. A series of tests were performed using the flat plate with blowing and no step to qualify the facility and the experiment. These results are detailed in Ref. [8]. Figure 5 shows a sample of the influence of local reaction on measured surface heat flux on the flat plate with $B=1.0$, $H^*=0.18$, and Da=20.

Examining Fig. 5, differences of up to 25% exist between the reactive and nonreactive sides downstream of the film-cooling holes, whereas the agreement between measured heat flux on the upstream gages is within several percent. The plot indicates the influence of the reaction extends downstream to the last gage location at $x/D=25$. The solid line shows the empirical turbulent boundary layer correlation, which is given below [11].

$$q=h(T_\infty-T_w)=\rho C_p \mathrm{St} M\sqrt{\gamma RT_\infty}(T_\infty-T_w) \tag{13}$$

The agreement of the nitrogen (inert) cooled side to the literature [13] predicted decrease in surface heat flux, which is shown by the dash lined is within 15%. To ensure that this behavior is attributable to local reactions, the gas injection sides were switched and the experiment repeated with similar results.

Next a series of tests was conducted using the step geometry with no blowing to assess how well the heat-flux measurements would match data available in the literature. These results are shown in Fig. 6.

Figure 6 shows a plot of Stanton number, normalized by the maximum Stanton number value measured downstream of the step face, versus nondimensional reattachment distance X^*. Note that the streamwise distance and reattachment distance used in calculating X^* are measured from the step face. The data measured from the shock tunnel runs are compared to the experiment of Vogel and Eaton [4]. As seen in the Fig. 6, the shock tunnel data matches Vogel and Eaton's measurements closely in both magnitude and trend.

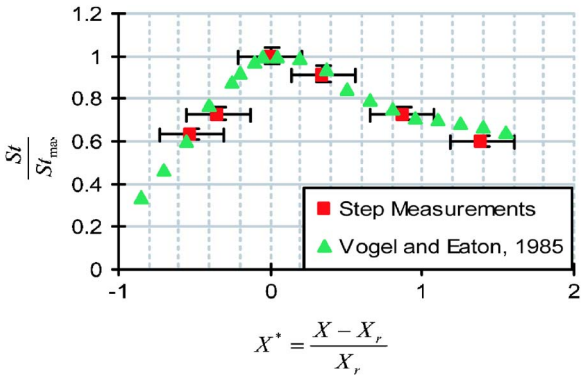


Fig. 6 Nondimensional backward-facing step heat-flux comparison, no blowing (B=0)

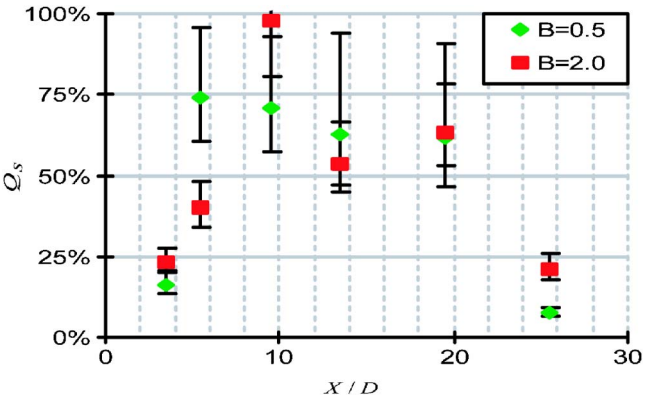


Fig. 9 Comparison of step reaction experiments at H^*=0.3 at B=0.5 and B=2.0

4.2 Step Tests With Blowing. Figure 7 shows the downstream Stanton number profile behind a film-cooled backward-facing step with various blowing ratios. The nondimensional distance X/D is measured relative to the cooling hole location. The heat flux profiles of the B=0.5 and B=0.0 cases are similar in trend and magnitude. For B=2.0, however, the Stanton number is relatively constant after reaching a peak around 10 X/D.

4.3 Summary of Reacting Flow Results. Figure 8 shows the dimensional heat flux measurements for Da=10, B=0.5, H^*=0.3 for both the flat plate and D^*=4, L_c^*=2 step. This figure shows a 17% increase in heat flux was measured at a downstream location of 10 cooling hole diameters. Because the nonreacting step heat flux is considerably higher than the flat plate heat flux for the same nondimensional conditions, the 17% heat flux increase due to chemical reactions could pose a greater problem for turbine end-wall durability. Note however, the percentage differences due to reaction shown in Fig. 8 cannot be directly applied to a turbine situation because of the very different temperatures and pressures (in particular the cold wall) used in the experiments. To assess the impact in a turbine environment it is necessary to begin with the scaled heat flux Q_s and apply it for the specific conditions encountered as shown in Ref. [8].

Figure 9 shows the downstream scaled heat flux profile for Da=10 and H^*=0.3 for B=0.5 and 2.0. A more rapid increase is seen for B=0.5, but a higher peak value is achieved for B=2.0. Here it can be seen that heat fluxes approaching those associated with the adiabatic flame temperature (Q_s=100%) occur at 5–15 dia downstream of the step.

The downstream profile of the maximum scaled heat flux is shown in Fig. 10 for B=0.5 and Da=10. Different values of heat release potential, ranging from 0.24 to 0.32, are shown in Fig. 10. The heat flux shown is the maximum value for each gage location measured during the steady-state test window. Kirk [8] found that scaled heat flux is nearly independent of heat release potential for a reacting film-cooled flat plate. Figure 10 shows that this is not the case for reacting film-cooled backward-facing steps.

The trends for the different H^* values are similar: a rapid rise in scaled heat flux followed by a relatively constant region, and then a decrease as mixing with the free stream reduces the temperature. This behavior is similar to the scaled heat flux trend observed for a lifted jet on a flat plate, [8].

The variation of scaled heat flux with Damköhler number is shown in Fig. 11. Both step and flat plate data are presented for H^*=0.24–0.87. The flat plate data show the expected trend of H^* independence, and the attached blowing ratio results in higher Q_s as discussed in Ref. [8]. Relative to the film-cooled flat plate at similar conditions, the step geometry promotes relatively higher heat flux at realtively lower Da.

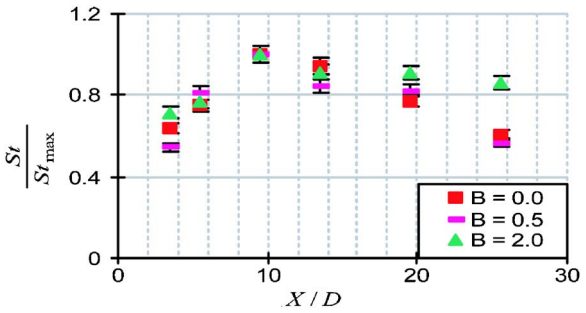


Fig. 7 Comparison of step results with and without blowing

5 Numerical Simulations

The trends in the experimental data are more complex than those for the film-cooled flat plate consistent with the larger number of important nondimensional parameters. The objective of the

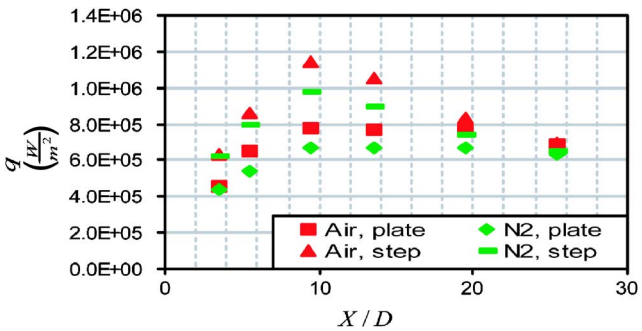


Fig. 8 Comparison between reacting flow tests for a flat plate and step at B=0.5 and H^*=0.3

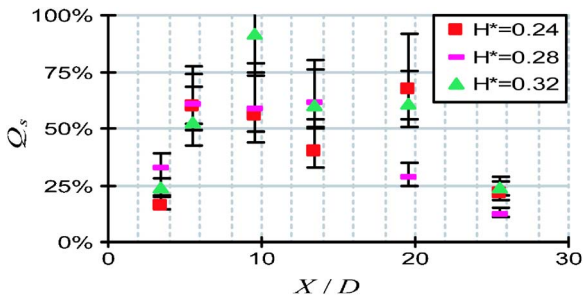


Fig. 10 Comparison of step reaction experiments at B=0.5, H^*=0.24, 0.28, and 0.32

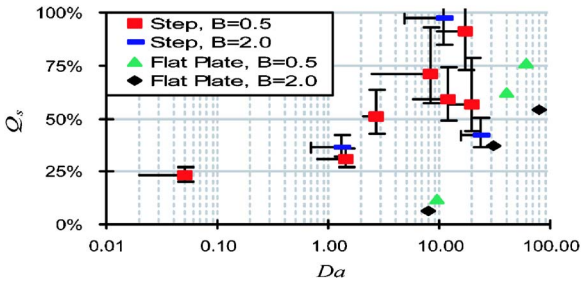


Fig. 11 Step *Qs* versus Da, step versus Flat Plate

numerical simulations presented here was to provide a means to visualize the general behavior of a film-cooled backward-facing step reacting flow field to enable the experimental results to be explained. This included assessing differences in the behavior for different blowing ratios, Damköhler numbers, and free-stream fuel levels.

A commercially available solver, FLUENT, was used to model the flow field. A tetrahedral grid was created for a single 35 deg film-cooling hole located 0.5 H downstream of a backward-facing step. The cooling hole diameter was fixed at 0.25 H. The grid domain extended 10 H upstream of the step face and 30 H downstream of the step face. The simulation was performed using a second-order discretization scheme. Turbulence away from the walls was modeled using the "Realizable" $k-\varepsilon$ equations, which have been found to better predict flows with strong adverse pressure gradients and separation than the standard $k-\varepsilon$ equations [14]. Near the wall, FLUENT's standard wall function model was implemented which encompasses law of the wall scaling and near-wall turbulence modeling [8]. Convergence was determined based on overall mass and energy imbalance within the domain of $<0.01\%$.

A simple one-step irreversible reaction model was employed to model the chemistry as summarized in [8]

$$a + \tfrac{1}{2}b \rightarrow c + \text{heat} \tag{14}$$

$$\tau_{\text{chem}} \approx \frac{(\text{fuel})_o}{\text{Ae}\dfrac{-E_a}{R_u T_o}(\text{fuel})^a(\text{O}_2)^b} \tag{15}$$

The free-stream flow was nitrogen seeded with hydrogen a, and the cooling flow was oxygen b. The Da number sweep was performed through varying the preexponential factor A.

5.1 Numerical Results. Figure 12(a) shows the total temperature contours for the respective nondimensional parameters specified with adiabatic wall thermal boundary conditions. The primary view in each subfigure shows the symmetry plane passing through $z/D=0$, and the top view in each subfigure is a cross section of the upstream and downstream walls. In each case, the free-stream temperature was 1200 K, the coolant temperature was 300 K, and the free-stream velocity was 100 m/s.

The effects of changing the blowing ratio are demonstrated in Fig. 12. For $B=0.1$, reactions are concentrated near the step face, resulting in significant temperature rise within the recirculation region and near the wall. As B increases, reactions within the recirculation zone are suppressed and moved into the mixing layer. In the highest blowing ratio simulated, the cooling jet penetrated the mixing layer to provide a flow blockage resulting in the entrainment of some free-stream flow in the region between the cooling jet and step face. This entrained free-stream flow, which is seeded with fuel, mixes with flow entrained from the cooling jet and reacts, resulting in the total temperature increase seen in Fig. 12(e).

Figure 13 shows the local equivalence ratio for the same fuel level and wall thermal boundary conditions as in Fig. 12, $H^*=0.3$ and an adiabatic wall, respectively (the cooling hole is displayed in light shading for scale reference). In Fig. 13, however, the Damköhler number was lowered to a value of zero thereby removing the chemical reactions. Figure 13 shows that as the blowing ratio increases the local equivalence ratio within the recirculation zone becomes leaner. The equivalence ratio drops to zero as the blowing ratio increases making the recirculation zone too lean to sustain combustion. The region of near-unity equivalence ratio shifts to the mixing layer as B increases. This is consistent with the shift of the reacting region to the mixing layer as blowing ratio increases as noted in Fig. 12.

As the blowing ratio increases, the coolant injection velocity

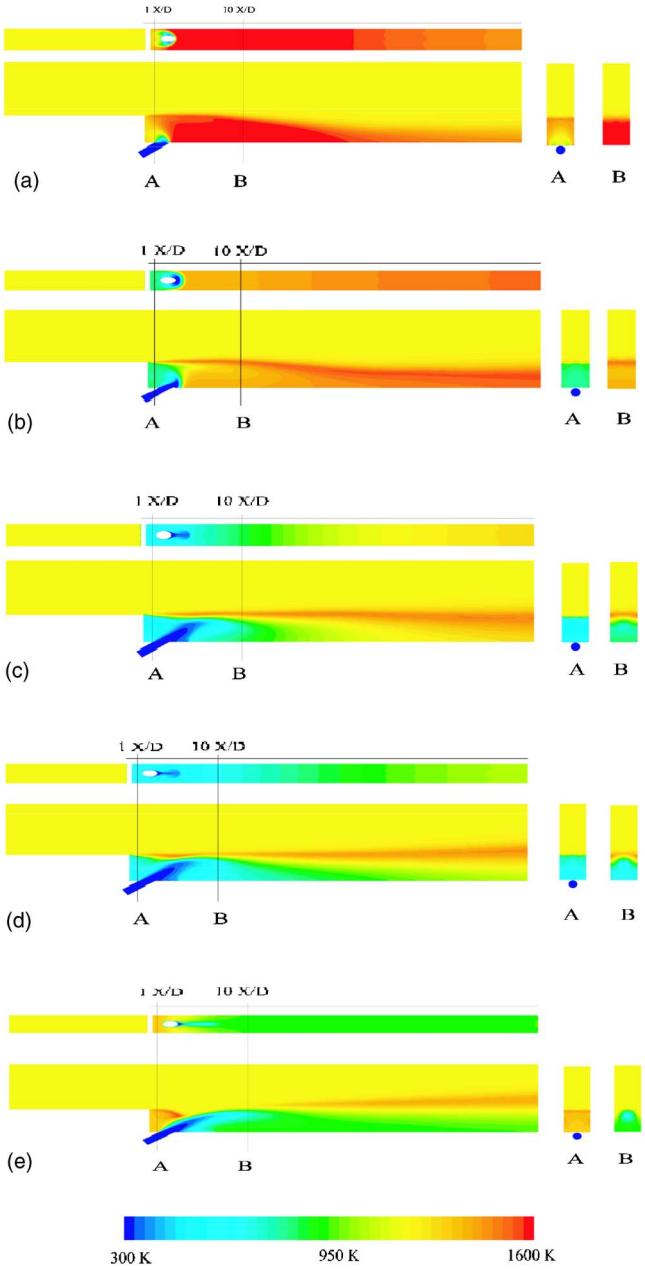


Fig. 12 (a) Temperature contours $B=0.1$, $I=0.002$, $H^*=0.3$, Da $=12$, adiabatic wall, (b) Temperature contours $B=0.5$, $I=0.055$, $H^*=0.3$, Da$=12$, adiabatic wall; (c) Temperature contours $B=2.0$, $I=0.875$, $H^*=0.3$, Da$=12$, adiabatic wall; (d) Temperature contours $B=4.0$, $I=3.50$, $H^*=0.3$, Da$=12$, adiabatic wall; and (e) Temperature contours $B=10.0$, $I=21.9$, $H^*=0.3$, Da$=12$, adiabatic wall

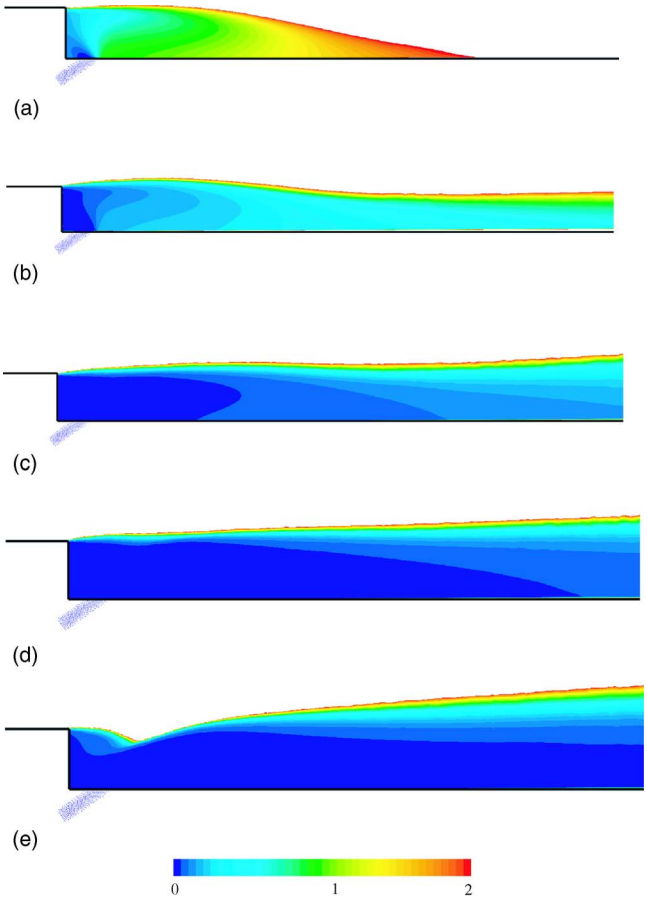


Fig. 13 (a) Equivalence ratio contours B=0.1, I=0.002, H^*=0.3, Da=0; (b) Equivalence ratio contours B=0.5, I=0.055, H^*=0.3, Da=0; (c) Equivalence ratio contours B=2.0, I=0.875, H^*=0.3, Da=0; (d) Equivalence ratio contours B=4.0, I=3.50, H^*=0.3, Da=0; and (e) Equivalence ratio contours B=10.0, I=21.9, H^*=0.3, Da=0

increases given the same coolant at the same thermodynamic state. This increases the effective downstream velocity at the edge of the recirculation zone that borders the mixing layer. As detailed in Ref. [6], this reduces the effective normal velocity across the shear layer by reducing the streamwise velocity difference between the higher-temperature free-stream flow and the lower-temperature recirculation zone flow. Because turbulent mixing is proportional to the effective normal velocity across the shear layer, mixing from the free stream to the recirculation zone decreases with increasing blowing ratio. This reduces the average temperature and fuel content of the flow entrained into the recirculation zone at the reattachment point.

The qualitative trends shown in the numerical simulations are consistent with the trends in blowing ratio seen in Fig. 9 for the experimental data. Further work is ongoing to pursue more quantitative comparisons and to assess other parametric dependencies. The simulations also suggest that, for blowing ratios of 4 or greater, it is possible to mitigate and delay the impacts of local reactions on surface heat transfer by moving the reactions to the shear layer and downstream.

Summary and Conclusions

As the equivalence ratios within modern combustors increase, the potential for reactive species to interact at film-cooled locations must be taken into account during the design of gas turbine engines. This paper presented experimental measurements and numerical simulations to explore the impacts that secondary reactions have on a film-cooled, backward-facing step. In summary:

- Under certain operating conditions, augmentations in surface heat flux can be produced due to chemical reactions in the film-cooling layer.
- The experimental data indicate that the scaled heat flux Q_s increases with increasing Damköhler number to a value of around 80–100% at Da=30 for B=0.5 and B=2.0. Relative to a cooled flat plate, the step promotes greater surface heat flux at lower Da.
- The blowing ratio controls the flow structure behind the backward-facing step and, therefore, controls the characteristics of reactions downstream of the film-cooled backward-facing step by altering the local equivalence ratio.

Future work will examine variations in step geometry, including step height, location of cooling holes relative to the step, as well as cooling hole spacing and arrayed geometries.

Acknowledgment

The authors would like to thank Professor Edward Greitzer and Stephen Lukachko of MIT for their guidance and the MIT Department of Aeronautics and Astronautics for funding this research.

Nomenclature

A = preexponential factor
B = mass blowing ratio, $\rho_c u_c / \rho_\infty u_\infty$
C_p = constant pressure specific heat J/kg K
Da = Damköhler number, $\tau_{\text{flow}} / \tau_{\text{chem}}$
D = cooling holes diameter, m
D^* = normalized step height, H/D
E_a = activation energy
ER = expansion ration, W_2/W_1
h = heat transfer coefficient, W/m^2 K
H = step height, m
H^* = heat release potential, $(T_{ad}-T_{t\infty})/T_{t\infty}$
I = momentum blowing ratio, $\rho_c u_c^2 / \rho_\infty u_\infty^2$
k = thermal conductivity, W/m K
L = length scale, m
L_c^* = normalized cooling distance, H/L_c
M = Mach number, u_∞/a_∞
P = static pressure, N/m^2
Pr = Prandtl number, $C_p \mu / k$
q = heat flux, W/m^2
Q_s = scaled heat flux ratio, $(q_{\text{hot}}-q_{\text{cold}})/(q_{\max}-q_{\text{cold}})$
R = specific gas constant (J/kgK)
R_u = universal gas constant J/kmol K
Re_H = step height Reynolds number, $\rho u_\infty H / \mu$
St = Stanton number, $h/u_\infty \rho Cp$
T = temperature, K
W = channel height, m
X = streamwise distance, m
X^* = nondimensional reattachment distance
X_r = reattachment distance m
y = vertical (normal to surface) distance, m
z = lateral distance, m

Greek

γ = ratio of specific heats
δ = boundary layer thickness, m
δ_H^* = normalized boundary layer thickness, δ/H
κ = thermal diffusivity, m^2/s
μ = viscosity, $N\,s/m^2$
ρ = density, kg/m^3
τ = characteristic time, s
Φ = equivalence ratio

Subscripts

ad = adiabatic
c = coolant
chem = chemical time to 95% completion
cold = nonreactive heat flux
d = driving
f = Film
hot = reactive heat flux
max = maximum heat flux
r = reattachment point
t = total or stagnation quantities
test = test time
w = wall
∞ = freestream

References

[1] Lukachko, S. P., Kirk, D. R., and Waitz, I. A., 2003, "Turbine Durability Impacts of High Fuel-Air Ratio Combustors, Part 1: Potential for Intra-Turbine Oxidation of Partially-Reacted Fuel," ASME J. Eng. Gas Turbines Power, **125**(3), pp. 742–750.

[2] Adams, E., Johnston, J., and Eaton, J., "Experiments on the Structure of Turbulent Reattaching Flow," Department of Mechanical Engineering, Stanford University, Report MD-43.

[3] Eaton, J., and Johnston, J., 1980, "Turbulent Flow Reattachment: An Experimental Study of the Flow and Structure behind a Backward-Facing Step," Department of Mechanical Engineering, Stanford University, Report MD-39.

[4] Vogel, J., and Eaton, J., 1985, "Combined Heat Transfer and Fluid Dynamic Measurements Downstream of a Backward-Facing Step," ASME J. Heat Transfer, **107**, pp. 922–929.

[5] Aung, W., and Goldstein, R., 1972, "Heat Transfer in Turbulent Separated Flow Downstream of a Rearward-Facing Step," Isr. J. Technol., **10**, pp. 35–41.

[6] Milanes, D. W. 2003, "Near Wall Reaction Effects on Film-Cooled Backward-Facing Step Heat Transfer," Master's thesis, Massachusetts Institute of Technology.

[7] Harinaldi, Ueda, T., and Mizomoto, M., 2001, "Effect of Slot Gas Injection to the Flow Field and Coherent Structure Characteristics of a Backstep Flow," Int. J. Heat Mass Transfer, **44**, pp. 2711–2726.

[8] Kirk, D. R., Guenette, G. R., Lukachko, S. P., and Waitz, I. A., 2003, "Turbine Durability Impacts of High Fuel-Air Ratio Combustors—Part II: Near Wall Reaction Effects on a Film-Cooled Flat Plate and Application to Gas Turbine Heat Transfer," ASME J. Eng. Gas Turbines Power, **125**(3), pp. 751–759.

[9] Kercher, D. M., "Short Duration Heat Transfer Studies at High Free-Stream Temperatures," ASME Paper No. 82-GT-129.

[10] Kirk, D. R., 1999, "Aeroacoustic Measurement and Analysis of Transient Hot Supersonic Nozzle Flows," Master's thesis, Massachusetts Institute of Technology.

[11] Keyes, W. M., and Crawford, M. E. 1980, *Convective Heat and Mass Transfer*, McGraw-Hill, New York.

[12] Vidal, R. J., 1956, "Model Instrumentation Techniques for Heat Transfer and Force Measurements in a Hypersonic Shock Tunnel," Cornell Aeronautical Laboratory, Inc. Contract No. AF33(616)-2387, Report No. AD917-A-1.

[13] Rohsenow, W., ed., 1985, *Handbook of Heat Transfer Applications*, 2nd Edition, McGraw-Hill, New York.

[14] Shih, T. H., Liou, W. W., Shabbir, A., and Zhu, J. 1995, "A New $k-\varepsilon$ Eddy-Viscosity Model for High Reynolds Number Turbulent Flows—Model Development and Validation." Comput. Fluids, **24**(3), 227–238.

Parametric Analysis of Combined Cycles Equipped With Inlet Fogging

R. Bhargava
Universal Ensco, Inc.,
1811 Bering Drive,
Houston, TX 77057

M. Bianchi

F. Melino

A. Peretto

DIEM–University of Bologna,
Viale Risorgimento 2,
40136 Bologna, Italy

In recent years, deregulation in the power generation market worldwide combined with significant variation in fuel prices and a need for flexibility in terms of power augmentation specially during periods of high electricity demand (summer months or noon to 6:00 p.m.) has forced electric utilities, cogenerators and independent power producers to explore new power generation enhancement technologies. In the last five to ten years, inlet fogging approach has shown more promising results to recover lost power output due to increased ambient temperature compared to the other available power enhancement techniques. This paper presents the first systematic study on the effects of both inlet evaporative and overspray fogging on a wide range of combined cycle power plants utilizing gas turbines available from the major gas turbine manufacturers worldwide. A brief discussion on the thermodynamic considerations of inlet and overspray fogging including the effect of droplet dimension is also presented. Based on the analyzed systems, the results show that high pressure inlet fogging influences performance of a combined cycle power plant using an aero-derivative gas turbine differently than with an advanced technology or a traditional gas turbine. Possible reasons for the observed differences are discussed. [DOI: 10.1115/1.1765122]

Introduction

Ambient temperature has a strong influence on the gas turbine (GT) performance, with power output dropping by about 0.54% to 0.90% for every 1°C of temperature rise, [1]. Similarly, combined cycle power plant (CCPP) performance is also influenced by an increase in the ambient temperature because of decrease in the GT performance and efficiency of the cooling system (depending on the cooling system type) associated with a condensing steam turbine (ST).

One way to prevent the loss in performance of CCPPs, caused by high ambient temperatures, is to cool air at the inlet of gas turbine compressor. There are several methods to cool air at the compressor inlet such as continuous cooling systems (in which direct refrigeration, electrically driven chillers or absorption chillers are used), conventional evaporative systems (that use wetted honeycomb media for water evaporation) and thermal energy storage (which provides to build a cold reserve during nonpeak hours). Another technique, known as fogging, consists of cooling air at the inlet of compressor by spraying demineralized water converted into a fog by atomizing nozzles operating at high pressure. Evaporation of fog droplets cools the compressor inlet air, increasing its density and results in increased mass flow rate through the compressor. In case of inlet evaporative fogging, a saturated air flow at the wet bulb temperature is obtained and the power boost is mainly associated with increased air flow through the compressor. If the amount of injected water is more than that required for air to saturate, overspray fogging is realized. The effect of overspray fogging (also known as wet compression) is to further increase gas turbine power output. With sufficient amount of overspray, compressor work input decreases resulting in additional power boost.

The choice among all the available inlet cooling techniques is not simple because many parameters such as air temperature, relative humidity, air flow to GT output ratio, and number of hours per day for needed power enhancement are to be considered, [2].

It is important to note that various inlet cooling techniques, available in the market today, have certain limitations such as modifications in the plant lay out combined with increased investment and operating cost (with continuous cooling, or thermal energy storage approach) and low evaporation efficiency (with conventional media type evaporative cooling systems). Similarly, inlet fogging approach may not be advantageous at locations where shortage of a water source exists or if the ambient conditions are very humid.

Recently, various options (inlet fogging, steam injection, heat recovery steam generator (HRSG) with supplementary firing and their combinations) to augment power for combined cycle plants during peak demand periods were evaluated by Tawney et al. [3]. Their study revealed that under given economic conditions and operating dispositions, inlet fogging combined with a supplementary fired HRSG showed comparatively higher return on investment for combined cycle applications. Jones and Jacobs [4] presented a detailed review of available options for enhancing combined cycle performance. Their study also included an economic assessment of performance enhancement alternatives by considering a combined cycle plant consisting of two GE PG7241(FA) gas turbines, two unfired three-pressure level HRSGs, and one GE D11 reheat condensing steam turbine with a wet cooling tower. Among various alternatives examined, inlet fogging was found least sensitive to variations in the economic parameters considered due to its insignificant impact on non-peak period plant performance combined with low initial investment and modest gain in incremental peak-period power generation.

A comprehensive parametric study on the effects of inlet fogging (both evaporative and overspray) on the various existing gas turbines has been recently presented by Bhargava and Meher-Homji [5]. Classifying a large set of gas turbines from major gas turbine manufacturers worldwide into three categories, namely, traditional (TIT$<$1200°C), aero-derivative and advanced technology (TIT$>$1200°C), their study showed higher performance improvement due to fogging for aero-derivative machines where TIT (turbine inlet temperature) is the temperature at inlet of the first stage rotor.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 2002; final manuscript received March 2003. Assoc. Editor: H. R. Simmons. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Atlanta, GA, June 16–19, 2003, Paper No. 2003-GT-38187.

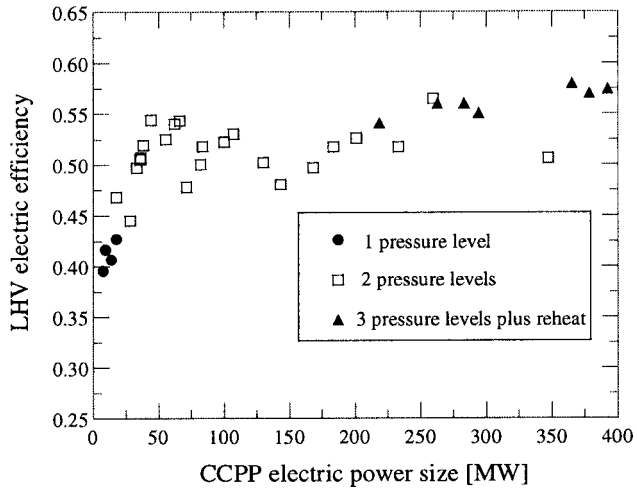


Fig. 1 LHV electric efficiency versus electric power size for combined cycle systems [6]

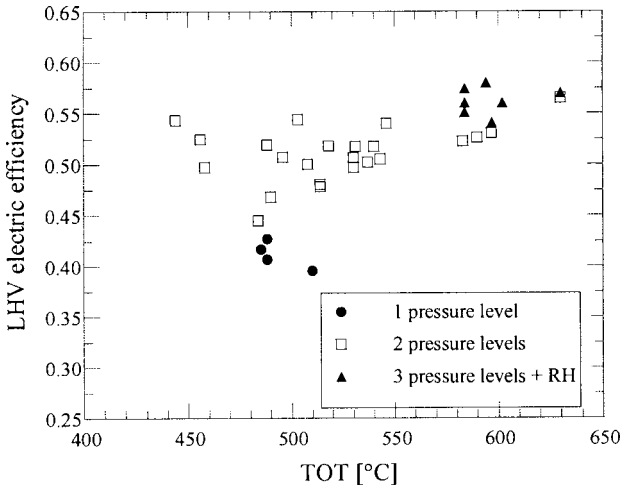


Fig. 2 LHV electric efficiency versus gas turbine exhaust gas temperature for combined cycle systems [6]

In this paper, results of a parametric study on the effects of inlet evaporative and overspray fogging applied on CCPPs are discussed.

Combined Cycle Power Plants

The current state of the art of CCPP performance is shown in Fig. 1, [6], which represents LHV (lower heating value) electric efficiency (η_{el}) versus electric power size (P_{el}) in a range up to 400 MW_{el} for different configurations of combined cycle systems.

CCPPs usually come in different configurations depending on the electric power size as is evident from Fig. 1. It could be seen from Fig. 1 that up to 25 MW_{el} the most widespread CCPP configuration consists of a single-pressure level HRSG. From about 25 to 250 MW_{el}, CCPP with two-pressure levels HRSG or, sometimes, CCPP with three-pressure levels HRSG are usually adopted. For electrical power output greater than 250 MW_{el}, the most developed configuration consists of a three pressure levels HRSG combined with reheat. It must be noted that number of pressure levels in a HRSG is also a function of gas turbine outlet gas temperature (TOT) as is evident from Fig. 2.

For the present study, 15 gas turbines with power output ranging from 5 MW to 260 MW (see Table 1 for main performance characteristics) employed in the CCPPs have been selected. In the present study, three most widely used combined cycle configurations discussed above, namely, one and two-pressure levels HRSG and three-pressure levels HRSG with reheat are simulated using a commercially available computer program, [7].

The main design features of the adopted CCPP configurations are described below and the key design parameters for bottoming cycle are reported in Table 2:

- a surface condenser using wet cooling tower with mechanical draft (the cooling water for the condenser is cooled by mixing with air in a wet cooling tower);
- integral deaerator;
- the steam supplied from HRSG to the steam turbine is superheated at all pressure levels;
- economizers are placed in parallel configuration (for two and three-pressure levels HRSG);
- parallel disposition of superheater and reheater (for three-pressure levels HRSG).

Table 1 Performance for selected gas turbines[1]

GT model	$\eta_{el,GT}$	TIT [°C]	β	$P_{el,GT}$ [MW]	TOT [°C]	M_{air} [kg/s]
Taurus 60	0.297	1093	12.0	5.31	518	21.6
Titan 130	0.329	1121	15.8	13.3	497	49.0
FT8	0.372	1160	20.0	24.7	463	82.6
RB211/6562	0.356	1227	20.8	27.9	495	92.1
GTX100	0.366	1288	20.0	42.3	553	118.5
W251 B11/12	0.324	1149	15.3	48.3	519	170.4
GT8C2	0.338	1177	17.6	55.3	514	191.9
V64.3	0.348	1310	15.8	65.9	588	186.2
MS 6001 FA	0.337	1288	14.8	69.2	597	202.2
W501 D5A	0.344	1177	14.2	117.9	536	374.4
MS 9001 EC	0.348	1204	14.2	168.8	561	510.2
V94.2A	0.354	1260	14.0	177.1	575	504.7
W 501 G	0.386	1427	19.2	245.5	599	544.9
V94.3A	0.381	1316	17.0	259.9	592	635.9
MS 9001 FA	0.369	1327	15.8	254.8	606	641.6

[1] values obtained considering the back pressure due to the HRSG

Table 2 Bottoming cycle design parameters

Pressure levels	1	2	3
Steam turbine inlet pressure	30 bar	15 bar 50 bar	3 bar 10 bar 120 bar
Reheat pressure	-	-	50
condenser pressure	0.05 bar		
Deaerating pressure	1.2 bar		
Minimum stack temperature	110 °C		
Minimum ΔT of pinch point	10 °C		
Minimum ΔT of approach point	20 °C		
Minimum ΔT of sub cooling	5 °C		
Steam turbine isentropic efficiency	0.860		
Mechanical efficiency	0.980		
Electric generator efficiency	0.985		

The overall performance parameters of CCPPs with selected gas turbines corresponding to the ISO (international standard organization) ambient conditions are provided in Table 3.

The Influence of Ambient Conditions on CCPP Performance. The variation of ambient conditions (T, p, and RH) influences the GT and bottoming cycle performance and, consequently the overall performance of a CCPP. At a non-ISO ambient condition, the GT operates at off-design point and consequently the thermal energy discharged to the HRSG varies (because of variation of exhaust gas mass flow rate and temperature). Also, the bottoming cycle performance is influenced because the condenser now operates under off-design conditions and its performance is linked to the condenser cooling system adopted. With an increase in ambient temperature, cooling water temperature increases which subsequently increases the condenser pressure. As a consequence, the steam turbine electric power output reduces.

The changes in performance parameters for different CCPP configurations, utilizing selected gas turbines as listed in Table 1, at the selected base ambient condition compared to the ISO condition are summarized in Table 4. The selected base ambient condition is characterized by an ambient temperature, pressure and relative humidity of 40°C, 1.013 bar, and 40% RH, respectively. This base ambient condition thereinafter will be referred as the hot day case (HDC). The values of ambient conditions for HDC are selected mainly because inlet fogging is found more effective at high ambient temperature and low humidity.

It should be mentioned that RH only slightly influences the CCPP performance in comparison to the effect of ambient temperature. Nevertheless, as it will be shown below, RH has an important role in fogging employment and strongly influences its benefits. Moreover, as far as the ambient pressure is concerned, its influence on CCPP performance is negligible in the fogging performance variation.

A review of the results presented in Table 4 highlights the following:

- compared to the ISO condition, HDC condition reduces GT mass flow rate (by about 10% to 16%) and increases turbine outlet temperature (by about 20 to 29°C);
- loss in GT power output ranges from 13 to 25% at the HDC condition compared to the ISO condition; these results are consistent with the previous study made on GT performances under hot ambient conditions and in simple cycle operation [5];
- at HDC condition, compared to the ISO condition, steam turbine (ST) power output drops by about 8 to 13%. The observed

Table 3 Overall performance of selected combined cycle power plants at ISO conditions

GT model	Pressure levels	$\eta_{el,CC}$	$P_{el,CC}$ [MW]	ε_{HRSG}
Taurus 60	1	0.4423	7.92	0.7455
Titan 130	1	0.4633	18.72	0.7178
FT8	2	0.4990	33.15	0.7098
RB211/6562	2	0.4957	38.80	0.7417
GTX100	2	0.5202	60.21	0.7881
W251 B11/12	2	0.4750	70.93	0.7619
GT8C2	2	0.4899	80.30	0.7575
V64.3A	2	0.5149	97.57	0.8115
MS 6001 FA	2	0.5095	104.48	0.8196
W501 D5A	2	0.5009	171.44	0.7754
MS 9001 EC	3	0.5273	255.86	0.8223
V94.2A	3	0.5349	267.45	0.8304
W 501 G	3	0.5510	350.03	0.8373
V94.3A	3	0.5561	378.86	0.8336
MS 9001 FA	3	0.5500	380.32	0.8393

Table 4 A comparative performance of combined cycle power plants at ISO and hot day case ambient conditions

GT model	Solar Turbines TAURUS 60			Solar Turbines TITAN 130			Pratt & Withney FT8			Rolls Royce RB 211/6562			Alstom Power GTX 100		
Case	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]
Pressure Levels	1			1			2			2			2		
CCPP Electrical Power Output [MW]	7.92	6.80	-16.43	18.72	15.72	-19.06	33.15	27.36	-21.18	38.80	31.92	-21.54	60.21	51.46	-17.00
CCPP Heat Rate [kJ/kWh]	8140	8384	2.91	7754	8041	3.57	7215	7448	3.13	7262	7501	3.19	6920	7065	2.05
CCPP LHV Electrical Efficiency	0.4423	0.4294	-3.00	0.4633	0.4477	-3.48	0.4990	0.4833	-3.25	0.4957	0.4799	-3.29	0.5202	0.5096	-2.08
GT Electrical Power Output [MW]	5.31	4.41	-20.59	13.27	10.76	-23.35	24.72	19.70	-25.48	27.86	22.23	-25.33	42.33	35.58	-18.98
GT LHV Efficiency	0.2969	0.2774	-7.02	0.3291	0.3063	-7.45	0.3720	0.3480	-6.91	0.3559	0.3342	-6.51	0.3658	0.3523	-3.83
GT Mass Flow Rate [kg/s]	21.92	19.55	-12.13	49.81	44.43	-12.09	83.97	72.19	-16.31	93.63	81.48	-14.92	120.81	105.42	-14.60
GT Outlet Temperature [°C]	518	543	25 °C	497	519	22 °C	463	492	29 °C	495	516.0	21 °C	553	579	26 °C
ST Electrical Power Output [MW]	2.60	2.39	-8.78	5.45	4.97	-9.76	8.43	7.66	-10.10	10.94	9.70	-12.84	17.87	15.88	-12.55
High Pressure Steam Mass Flow [kg/s]	2.76	2.63	-4.63	5.78	5.48	-5.48	7.68	7.52	-2.11	9.82	9.34	-5.16	15.52	14.88	-4.30
Low Pressure Steam Mass Flow [kg/s]	-	-	-	-	-	-	1.45	1.03	-40.70	1.37	0.98	-40.68	1.22	0.78	-56.61
HRSG efficiency	0.7455	0.7833	4.82	0.7178	0.7516	4.49	0.7098	0.7544	5.91	0.7417	0.7794	4.84	0.7881	0.8237	4.33

GT model	SIEMENS WESTINGHOUSE W 251 B11/12			Alstom Power GT8C2			Ansaldo Energia V64.3A			GE MS 6001 FA			SIEMENS WESTINGHOUSE W 501 D5A		
Case	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]
Pressure Levels	2			2			2			2			2		
CCPP Electrical Power Output [MW]	70.93	60.67	-16.92	80.30	69.41	-15.68	97.57	85.35	-14.32	104.48	89.85	-16.29	171.44	147.60	-16.15
CCPP Heat Rate [kJ/kWh]	7578	7839	3.33	7349	7558	2.77	6992	7189	2.74	7066	7318	3.44	7187	7385	2.68
CCPP LHV Electrical Efficiency	0.4750	0.4593	-3.42	0.4899	0.4763	-2.86	0.5149	0.5008	-2.82	0.5095	0.4919	-3.58	0.5009	0.4875	-2.75
GT Electrical Power Output [MW]	48.30	40.01	-20.72	55.33	46.63	-18.66	65.92	56.72	-16.23	69.18	57.93	-19.41	117.88	98.79	-19.33
GT LHV Efficiency	0.3235	0.3029	-6.80	0.3376	0.3200	-5.49	0.3479	0.3327	-4.55	0.3374	0.3172	-6.36	0.3444	0.3263	-5.56
GT Mass Flow Rate [kg/s]	173.38	156.44	-10.83	195.18	177.01	-10.26	189.99	173.11	-9.75	206.30	186.15	-10.82	381.24	342.75	-11.23
GT Outlet Temperature [°C]	519	540	21 °C	514	534	20 °C	588	609	21 °C	597	622	25 °C	536	559	23 °C
ST Electrical Power Output [MW]	22.63	20.66	-9.56	24.97	22.78	-9.59	31.65	28.64	-10.52	35.30	31.92	-10.62	53.55	48.81	-9.72
High Pressure Steam Mass Flow [kg/s]	19.89	19.41	-2.47	21.96	21.46	-2.33	27.59	27.10	-1.81	30.87	30.29	-1.91	46.38	45.02	-3.02
Low Pressure Steam Mass Flow [kg/s]	2.20	1.60	-37.75	2.56	1.90	-35.00	1.29	0.80	-62.14	1.21	0.67	-81.32	4.34	3.04	-42.64
HRSG efficiency	0.7619	0.7940	4.04	0.7575	0.7895	4.05	0.8133	0.8436	3.59	0.8196	0.8522	3.83	0.7754	0.8073	3.95

GT model	NUOVO PIGNONE GE MS 9001 EC			SIEMENS WESTINGHOUSE V94.2A			SIEMENS WESTINGHOUSE W 501 G			Ansaldo Energia V94.3A			NUOVO PIGNONE GE MS 9001 FA		
Case	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]	ISO	HDC	Δ [%]
Pressure Levels	3 + RH			3 + RH			3 + RH			3 + RH			3 + RH		
CCPP Electrical Power Output [MW]	255.9	217.9	-17.42	267.4	230.8	-15.87	350.025	313.802	-11.54	378.862	332.244	-14.03	380.318	321.009	-18.48
CCPP Heat Rate [kJ/kWh]	6827	7045	3.09	6730	6926	2.83	6533	6633	1.51	6474	6603	1.95	6546	6720	2.59
CCPP LHV Electrical Efficiency	0.5273	0.5110	-3.19	0.5349	0.5198	-2.90	0.5510	0.5428	-1.51	0.5561	0.5452	-2.00	0.5500	0.5357	-2.67
GT Electrical Power Output [MW]	168.75	138.39	-21.94	177.14	147.88	-19.79	245.46	217.12	-13.05	259.85	223.63	-16.20	254.81	209.157	-21.83
GT LHV Efficiency	0.3478	0.3245	-7.16	0.3543	0.3330	-6.39	0.3864	0.3755	-2.90	0.3814	0.3670	-3.94	0.3685	0.3490	-5.56
GT Mass Flow Rate [kg/s]	519.90	466.52	-11.44	514.69	464.47	-10.81	557.59	504.15	-10.60	649.51	586.58	-10.73	655.42	575.67	-13.85
GT Outlet Temperature [°C]	561	585	24 °C	575	601	26 °C	599	627	28 °C	592	615	23 °C	606	631	25 °C
ST Electrical Power Output [MW]	87.11	79.51	-9.56	90.31	82.95	-8.87	104.57	96.69	-8.15	119.01	108.62	-9.57	125.50	111.85	-12.21
High Pressure Steam Mass Flow [kg/s]	57.41	57.2	-0.37	60.68	60.68	0.00	72.5	72.8	0.41	81.61	80.8	-1.00	87.22	84.21	-3.57
Medium Press Steam Mass Flow [kg/s]	14.4	11.99	-20.10	12.91	10.7	-20.65	11.82	9.612	-22.97	14.63	12.17	-20.21	13.01	10.34	-25.82
Low Pressure Steam Mass Flow [kg/s]	2.399	1.826	-31.38	2.274	1.731	-31.37	2.305	1.74	-32.47	2.746	2.097	-30.95	2.638	1.927	-36.90
HRSG efficiency	0.8223	0.8587	4.24	0.8304	0.8645	3.95	0.8373	0.8705	3.81	0.8336	0.8678	3.94	0.8393	0.8748	4.06

reduction in ST power output is due to the reduction of the GT thermal energy discharged and to the increase in condensing pressure. Nevertheless, small increase in the gas turbine outlet temperature and HRSG efficiency (calculated with respect to ambient temperature) compensates the ST power output reduction;

- the cooling system of the condensing steam turbine is strongly influenced by ambient temperature and relative humidity. In particular, the temperature rise of the cooling water (due to the hot ambient temperature) increases the condenser pressure from 0.05 bar to about 0.1 bar;
- net electric power output of CCPPs evaluated decreases by about 11 to 22%. The greater fraction of this decrease is associated with GT performance loss since the losses for bottoming cycle are comparatively smaller as is evident from Table 4.

Inlet Fogging and Overspray Models—Thermodynamic Considerations

Considering an adiabatic process, the thermal energy loss from a nonsaturated air stream ($\dot{Q}_{\text{air}}$) with the injection of water may be expressed as the sum of energy, per unit time, consumed for water evaporation ($\dot{Q}_{ev}$) and energy stored in the water droplets ($\dot{Q}_d$) not yet evaporated:

$$\dot{Q}_{\text{air}} = \dot{Q}_{ev} + \dot{Q}_d . \tag{1}$$

Considering that $\dot{m}$, T, and c_p represent mass flow rate, temperature and constant pressure specific heat, respectively, and using the subscripts a and ac to distinguish the air before and after cooling (see Fig. 3) $\dot{Q}_{\text{air}}$ becomes

$$\dot{Q}_{\text{air}} = c_{pa}\dot{m}_a T_a - c_{\text{pac}}\dot{m}_{ac} T_{ac} . \tag{2}$$

Equation (2) is valid if the air temperature T_{ac} is greater than the wet bulb temperature T_{wb} ($T_{ac} \geqslant T_{wb}$). Moreover, the energy consumed for water evaporation $\dot{Q}_{ev}$ may be expressed as

$$\dot{Q}_{ev} = \dot{m}_{ev}(\Delta h_{\text{lat}} + \Delta h_{sh}) \tag{3}$$

where $\dot{m}_{ev}$ is the evaporated water mass flow rate, Δh_{lat} the latent heat and Δh_{sh} energy spent to heat the evaporated water from the evaporation temperature to T_{ac}.

Also, Δh_{lat} can be evaluated, expressing the droplet temperature T_d in °C, with the following equation:

$$\Delta h_{\text{lat}} = 2501.7 - 2.3704 T_d \ \ [\text{kJ/kg}]. \tag{4}$$

Neglecting $\dot{Q}_d$ (that means to assume T_d constant during the whole evaporation process) and the difference between $\dot{m}_a$ and $\dot{m}_{ac}$ (due to the water evaporation) and assuming the average value of the air specific heat (c_p), Eq. (1) reduces to

$$\dot{m}_a c_p (T_a - T_{ac}) = \dot{m}_{ev}(\Delta h_{\text{lat}} + \Delta h_{sh}). \tag{5}$$

Equation (5) may be employed to evaluated air cooling due to water injection and evaporation and may be utilized for the inlet evaporative and overspray fogging analysis. As was the case for Eq. (2), Eq. (5) is valid if $T_{ac} \geqslant T_{wb}$.

In the case of inlet evaporative fogging and with reference to Fig. 3, a complete evaporation of the entire water mass flow rate injected ($\dot{m}_{\text{inj}}$) with the spray system before the compressor inlet occurs, implying that $\dot{m}_{\text{inj}} - \dot{m}_{ev} = 0$. Whereas, in the overspray fogging case, liquid water enters inside the compressor (i.e., $\dot{m}_{\text{inj}} - \dot{m}_{ev} > 0$).

It should be noted that the total amount of water mass flow rate ($\dot{m}_w$) consumed is greater than the water quantity ($\dot{m}_{\text{inj}}$) used for inlet evaporative or overspray fogging because of water losses (leakage) due to deposit of the droplets against the silencer, the inlet air trash screen and the duct walls. To account for water losses, a spray system efficiency (η_s) is introduced and then the total water consumption can be expressed as

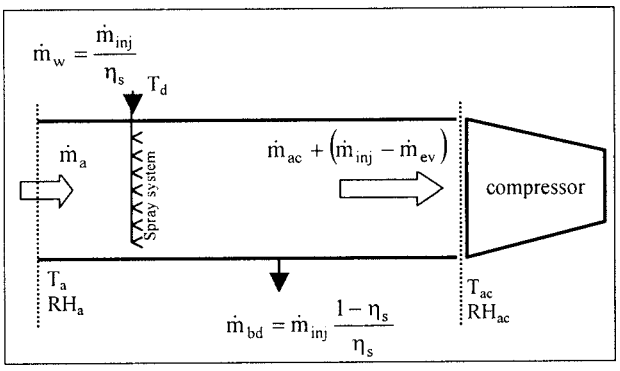


Fig. 3 A schematic showing nomenclatures at the compressor inlet with fogging

$$\dot{m}_w = \frac{\dot{m}_{\text{inj}}}{\eta_s} . \tag{6}$$

The expression of blowdown mass flow rate $\dot{m}_{bd}$ is reported in Fig. 3 as a function of spray system efficiency and injected mass flow rate.

Inlet Fogging. With the above-mentioned hypothesis (such as, adiabatic process, droplet temperature constant during the evaporation) and the air becoming saturated at the compressor inlet ($\text{RH}_{ac} = 100\%$ and $T_{ac} = T_{wb}$) for the evaporative fogging case, air temperature decreases at the compressor inlet (from T_a to T_{wb}, as shown in Fig. 4) and the evaporated water mass flow rate ($\dot{m}_{ev}$) can be calculated using Eq. (5) and the psychrometric chart of Fig. 4.

A decrease in the compressor inlet air temperature, due to air saturation (from point a to point wb of Fig. 4), increases the corrected speed $\text{CS} = n/\sqrt{T}$, the pressure ratio (β) and the corrected flow ($\text{CF} = m\sqrt{T}/p$) because the compressor operation point moves toward the design point (ISO condition) as shown in Fig. 5.

Overspray Fogging. In the case of overspray fogging, the amount of water injected in the air stream is more than that required for the air to saturate at a given ambient condition. Thus, the saturated air enters inside the compressor together with liquid water.

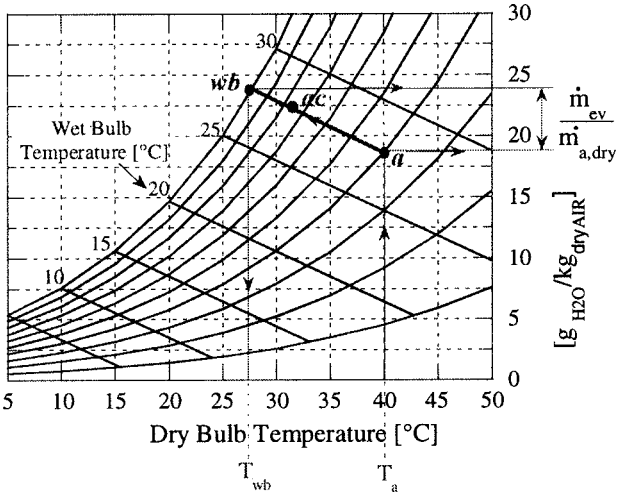


Fig. 4 Psychrometric chart (@ $p = 1.01325$ bar)

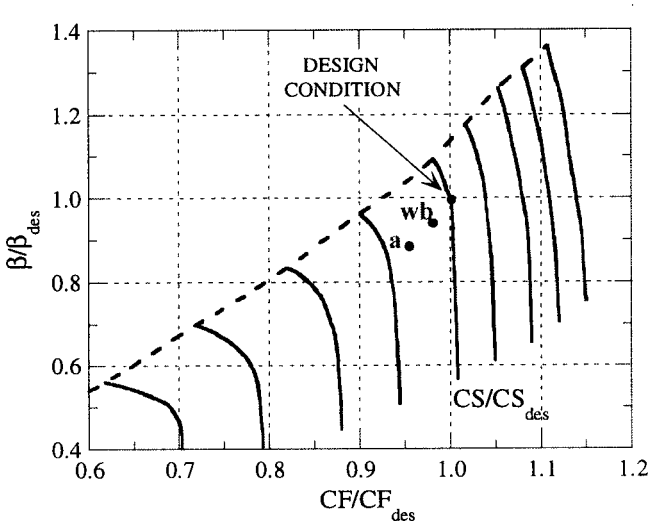


Fig. 5 A typical compressor performance map

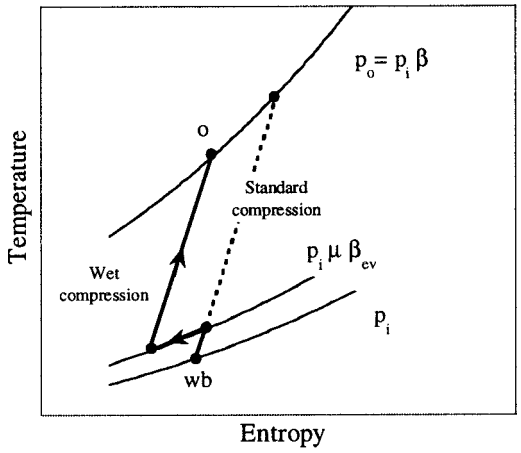


Fig. 7 Temperature-entropy diagram for a compression process in presence of overspray fogging

As air temperature increases inside the compressor due to compression effect, it reduces the air relative humidity below the saturation point and, consequently, resulting in additional evaporation of water droplets.

In the overspray model presented here, it has been assumed that the whole overspray water evaporates simultaneously and immediately when air temperature inside the compressor is sufficient to evaporate all the water. Obviously, the pressure ($p_{ev}=p_i\beta_{ev}$) at which evaporation occurs depends on the values of compressor polytropic efficiency (η_p) and the amount of overspray (ϕ) defined as the ratio between the evaporated mass of overspray water ($\dot{m}_{ev,os}$) and saturated air mass flow rate at the compressor inlet ($\dot{m}_{ac}$):

$$\phi=\frac{\dot{m}_{ev,os}}{\dot{m}_{ac}}. \tag{7}$$

In Fig. 6, for the three different values of compressor polytropic efficiency, the value of ϕ and the increase of temperature during compression (starting from saturated air at 28°C and pressure equal to 1.01325 bar), are presented versus the evaporation pressure ratio, β_{ev}.

Figure 6 shows that for a typical compressor polytropic efficiency of 0.90 and limiting the value of ϕ at 0.02, all the water mass flow rate evaporates for β_{ev} values around 2. This overspray fogging model does not take into account the effects of number and size of water droplets which have significant influence on the evaporation time. Nevertheless, it is possible to take into account the droplet number and its dimension by introducing a "delay coefficient" μ that increases the evaporation pressure value ($p_{ev}=p_i\mu\beta_{ev}$). This delay coefficient is a function of the droplet dimension and, in particular:

- $\mu=1$ means that no evaporation delay is considered (usually, this approximation is correct when the size of droplets is less than 10 μm);
- $\mu>1$ means that an evaporation delay is considered. Typically μ ranges from 1 to 1.4 depending on the droplet dimension, [7].

It should be noted that, the compressor discharge temperature (T_o) depends on the p_{ev} value (and, therefore, also on μ) and, with reference to Fig. 7, it may be expressed as

$$T_o=T_{wb}\beta^{(k-1)/(k\eta_p)}-\phi\frac{\Delta h_{\text{lat}}}{c_{pa}}\left(\frac{\beta}{\mu\beta_{ev}}\right)^{(k-1)/(k\eta_p)} \tag{8}$$

where k is the average value of specific heat ratio and the value of η_p is assumed constant.

Equation (8) shows that by increasing μ, the compressor discharge temperature and specific work increase. A more detailed discussion on the derivation of Eq. (8) is presented in the Appendix.

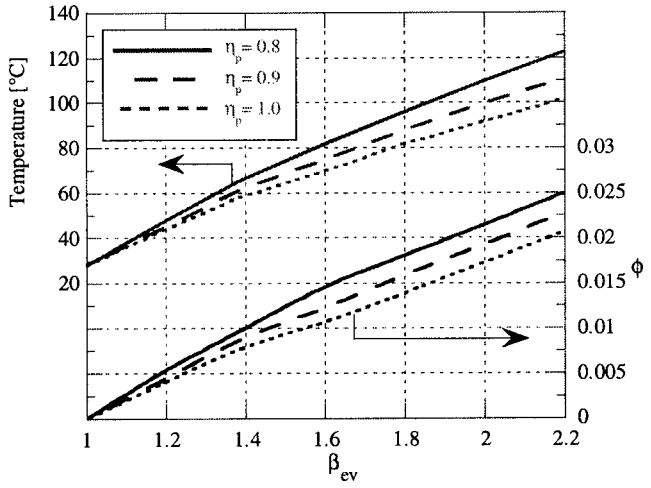


Fig. 6 Air temperature and evaporated water mass through compressor

CCPP Performance With Inlet Evaporative and Overspray Fogging

To study the effects of high pressure inlet fogging on CCPPs performance parameters, inlet evaporative and overspray fogging conditions were analyzed for various CCPPs using gas turbines listed in Table 1. In the case of overspray fogging, an additional water mass flow rate ranging from 0.5% to 2% of saturated air flow rate was considered. The various results are presented as change (percent change or just difference) in performance parameters with respect to the HDC condition. This approach allows comparison of various systems operating under same base ambient conditions. In addition, results for ISO condition are also included in each figure. This presentation allows to view changes in a performance parameter both with respect to ISO and HDC conditions for various fogging cases examined.

Net combined cycle plant output, comparing inlet fogging case to HDC, shows an increase in a range of 6% to 10% for various systems as is evident from Fig. 8. This observed change implies a

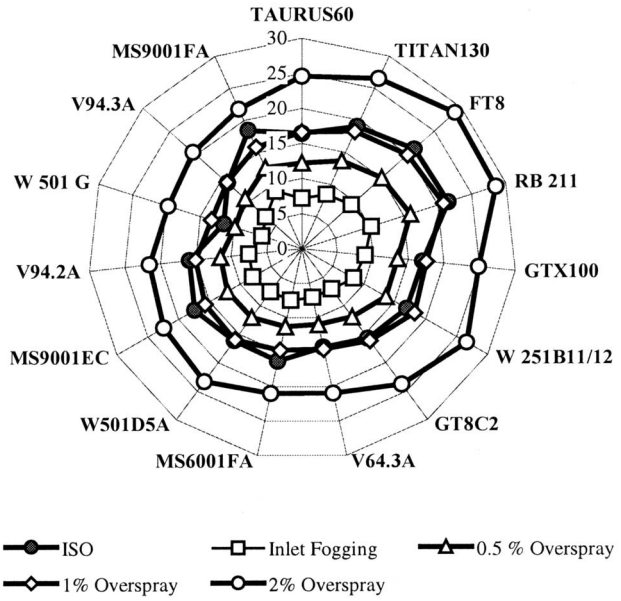


Fig. 8 Percent power output change of CCPPs @ generator (reference case—HDC)

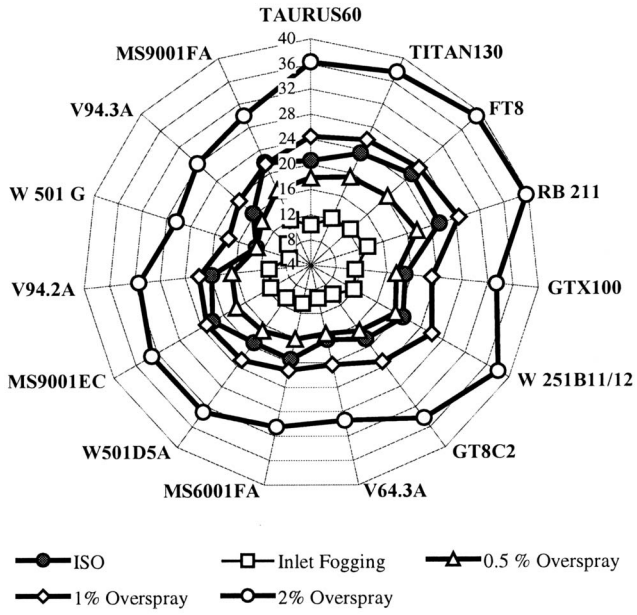


Fig. 9 Percent power output change of gas turbines @ generator (reference case—HDC)

gain of about 0.50% to 0.75% in net plant output for every 1°C of inlet air cooling. It may be noted that for the assumed base ambient condition (HDC), the maximum amount of cooling achievable is 12°C. As anticipated and evident from Fig. 8, percent increase in net plant output becomes higher with the application of overspray fogging for each CCPP. In particular, with 1% overspray fogging, all the lost power due to high ambient temperature (HDC condition) is recovered for various machines examined in this study with few exceptions (such as MS9001FA, MS6001FA and FT8) as shown in Fig. 8. Whereas, with 2% of overspray a net gain in electrical power output compared to the ISO case could be achieved for various systems examined (Fig. 8). Based on the limited number of machines examined, results suggest that a higher amount (approximately, 4–9%) of boost in net power output can be achieved for a CCPP with an aero-derivative gas turbine (such as RB211 and FT8) than a CCPP with an advanced technology gas turbine (for example, MS9001FA, MS6001FA, W501G, etc) with 2% overspray fogging (Fig. 8). Furthermore, a CCPP using a traditional gas turbine (such as W251 B11/12, TITAN 130, MS9001EC, etc.) shows higher (approximately, 2–7%) gain in net power output with 2% overspray fogging compared to a CCPP using an advanced technology gas turbine.

The observed differences in the amount of power boost between an aero-derivative and advanced technology or traditional gas turbines can be attributed, among other factors, to the fact that the rate of loss in power output due to increased ambient temperature is comparatively higher for an aero-derivative gas turbine, [5]. The other factors could include differences in compressor performance characteristics and flexibility of changing rotational speeds for aero-derivative machines compared to the other two categories of gas turbines.

A further examination of the data shows that differences in power boost for three categories of gas turbines change (increase) with the increased amount of overspray. It is important to mention that, based on the field performance data, the commercial program used tend to overestimate effects of fogging. Despite of differences in the absolute values of certain performance parameters, the study presented does show trends of performance gains due to fogging for CCPPs.

The observed increase in net plant output can mainly be attributed to GT performance enhancement as shown in Fig. 9. Whereas, the effect of fogging on steam turbine power output (Fig. 10) is small (maximum value, 6%), compared to the GT power output boost, and especially no significant change seems to exist between ST power boost and the type of applied fogging strategy. This observation could be explained considering that drop in ST performance is principally caused by ambient temperature and relative humidity effects on the condenser cooling system and marginally by GT.

Comparing GT power output between inlet fogging case to HDC, an increase of about 8–12% power boost (equivalent to

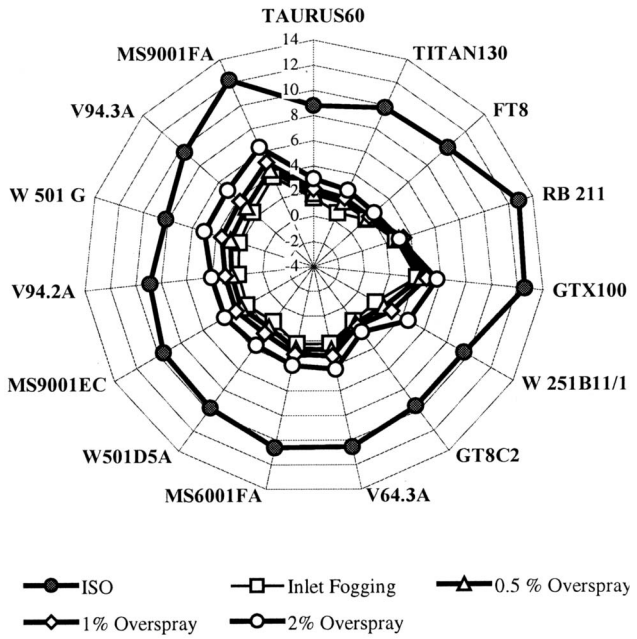


Fig. 10 Percent power output change of steam turbines @ generator (reference case—HDC)

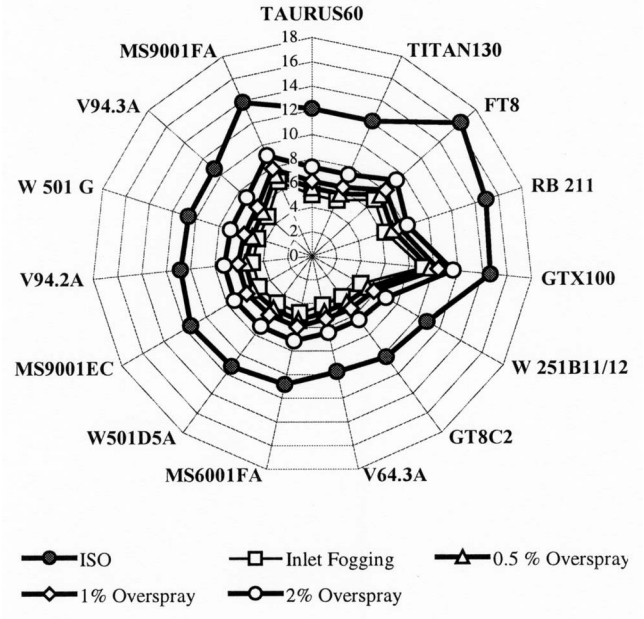


Fig. 11 Percent air mass flow rate change of gas turbines (reference case—HDC)

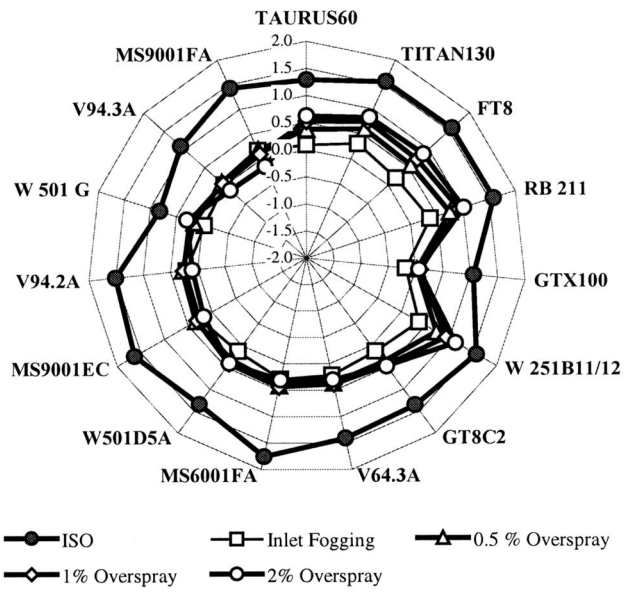


Fig. 13 Percetange points change of LHV CCPP efficiency (reference case—HDC)

approximately, 0.7–0.8% power boost for every 1°C of cooling of air at compressor inlet) can be observed (Fig. 9). Furthermore, more power output gain is obtained with the use of overspray. This amount of power boost grows with the increase in injected overspray water mass flow rate (Fig. 9).

Among the other parameters that influence CCPPs performance, GT mass flow rate and turbine outlet temperature are presented, in Figs. 11 and 12, respectively. Inlet evaporative and overspray fogging increase total mass flow rate through the compressor but obtained values are less than the ISO case (Fig. 11). This could be explained considering that air mass flow rate (for constant compressor speed) depends only on air density (and so by air temperature and relative humidity) at the compressor inlet. Air temperature at the compressor inlet is always 28°C and relative humidity 100% for different fogging conditions; this implies that air density is lower compared to the ISO condition. The total air mass flow rate through a compressor increases as the fogging condition changes from inlet evaporative fogging to 2% overspray case and is mainly due to the increased amount of injected water. A closer look at the data shows that increase in total air mass flow rate due to fogging (evaporative or overspray fogging) is higher, by approximately 2–3% for aero-derivative gas turbines (Fig. 11) compared to the other gas turbines with the exception of two gas turbines (GTX100 and MS9001FA).

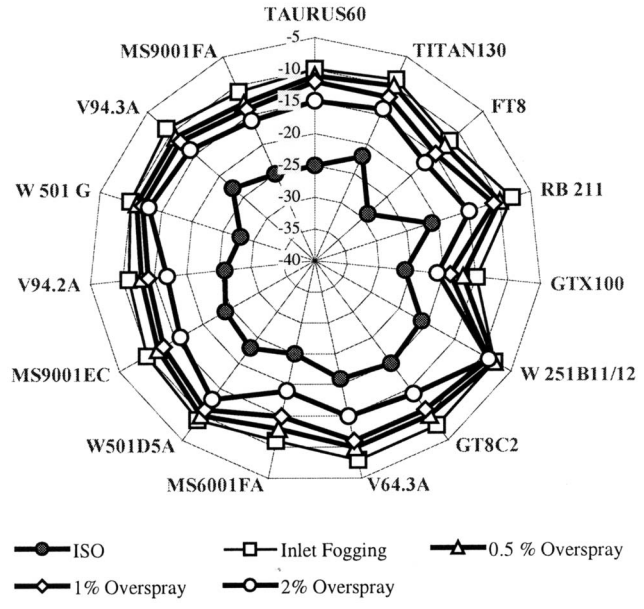


Fig. 12 A comparative change in gas turbine outlet temperature, TOT [°C] (reference case—HDC)

It is not surprising to note that the value of TOT increases (by 20 to 30°C) with the increase in ambient temperature (i.e., for the HDC condition) for various systems examined as shown in Fig. 12. Inlet and overspray fogging applications reduce TOT value but does not reach its ISO value. It could be observed that the difference in TOT value compared to the ISO condition reduces with the increased amount of overspray fogging as shown in Fig. 12 mainly due to the inter-cooling effect.

The change, expressed in percentage point, in the plant LHV electric efficiency compared to the HDC condition is found small (less than one percentage point) due to different fogging conditions as shown in Fig. 13. A slightly higher increase in efficiency for aero-derivative gas turbine compared to other gas turbines is also evident in Fig. 13. Furthermore, efficiency loss for various CCPPs due to change in ambient condition to the HDC condition could not be recovered even with 2% overspray fogging.

It is quite evident from Fig. 14 that a CCPP with an advanced technology gas turbine requires considerably more amount of injected water per unit percent power boost compared to the other gas turbine based systems. The main reason being that the injected water amount is a percentage fraction of the inlet air mass flow rate. Typically, an advanced technology gas turbine has a high air mass flow rate as is evident from Table 1. Thus, to compare injected water flow rates for different combined cycle systems, it would be more appropriate to normalize water flow rate by inlet air flow rate and percent power boost. A comparison of such a parameter does show that somewhat lower amount (less than 1

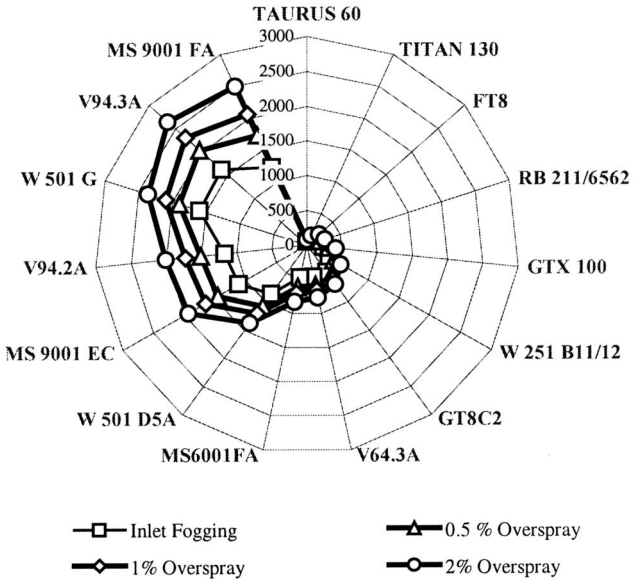


Fig. 14 Injected water consumption per unit percent net CCPP output change [(liters/hr)/% power boost] (reference case—HDC)

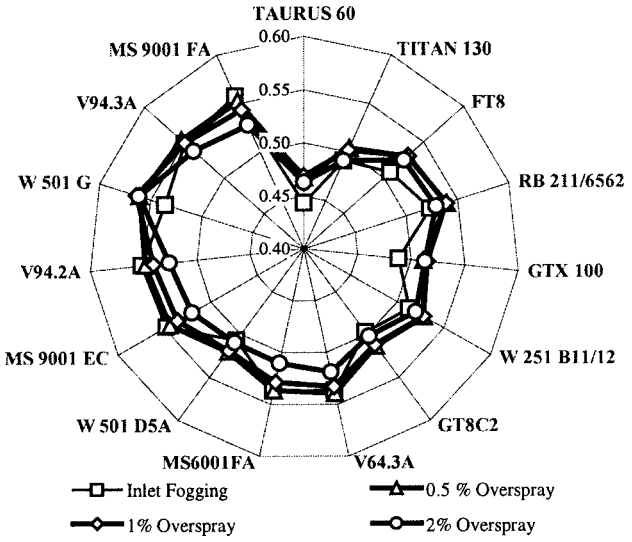


Fig. 15 Incremental efficiency for CCPP due to fogging (reference case—HDC)

liter/hr per unit of air mass flow rate for 1% of power boost) of injected water is required for an aero-derivative based CCPP in comparison to the other types of gas turbines.

Incremental efficiency, defined as the ratio of incremental power divided by additional fuel consumption due to fogging, [8], is indicative of fuel efficiency associated with the fogging technology. The trend of incremental efficiency values shows that gas turbines having high efficiency without fogging also have high incremental fuel efficiency in presence of fogging when used in combined cycle applications as is evident in Fig. 15. It is interesting to note that for many advanced technology gas turbine based CCPPs, value of incremental fuel efficiency decreased with the increase of overspray amount. On the contrary, incremental fuel efficiency always increased with increased amount of overspray for aero-derivative and traditional gas turbine based CCPPs. Also, the gain in incremental efficiency, compared to the base case (HDC) is approximately 6% (max.).

Concluding Remarks

Based on the comprehensive analysis of different configurations of combined cycle power plants with varying net power output (8 MW–380 MW) in presence of both evaporative and overspray fogging, reported in this paper, the following concluding remarks can be summarized:

- High pressure fogging is effective for CCPPs in recovering the lost power output due to high ambient temperatures. The amount of power boost increases with the increased amount of overspray fogging.
- The gain in net power output due to fogging, both inlet evaporative and overspray fogging, is mainly associated with the performance enhancement in the gas turbine section of a CCPP.
- The amount of net power boost is found higher for a CCPP with traditional or aero-derivative machines (such as, W251B11/12, FT8, Titan-130) compared to a CCPP using an advanced technology gas turbine (for example, W501G, V94.3A, MS9001FA).
- The amount of injected water, normalized by saturated air flow rate at the compressor inlet and percent net power boost, is smaller (by approximately, 1 liter/hr per unit flow rate and per unit % net power boost) for a CCPP using an aero-derivative gas turbine than a CCPP using other two categories of gas turbines. These results imply that an aero-derivative gas turbine based combined cycle plant will have lower operating cost with fogging compared to a CCPP using an advanced technology or a traditional gas turbine.
- Inspite of the fact that a CCPP using an advanced technology gas turbine showed smaller power boost due to fogging, its incremental fuel efficiency is found higher compared to the systems with other two types of gas turbines.

Nomenclature

$\dot{m}$ = mass flow rate (kg/s)
c_p = constant pressure specific heat (kJ/kg−K)
μ = delay coefficient
k = specific heat ratio
n = compressor rotational speed (rpm)
p = pressure (bar)
P_{el} = electric power output (MW)
Q = thermal power (kJ)
T = temperature (°C)
Δ = change in a performance parameter with respect to the hot day case
Δh = enthalpy difference (kJ/kg)
β = compressor pressure ratio
ϕ = amount of overspray (as a fraction of saturated air mass flow rate)
$\varepsilon_{\mathrm{HRSG}}$ = $(\mathrm{TOT}-T_{\mathrm{stack}})/(\mathrm{TOT}-T_a)$, HRSG effectiveness
η_{el} = LHV electric efficiency
η_p = polytropic efficiency
η_s = spray system efficiency

Subscripts and Superscripts

a = air at the compressor inlet before cooling
ac = air at the compressor inlet after cooling
air = air
bd = blowdown
d = droplet
db = dry bulb
des = design
el = electric
ev = evaporated
i = compressor inlet
inj = injected

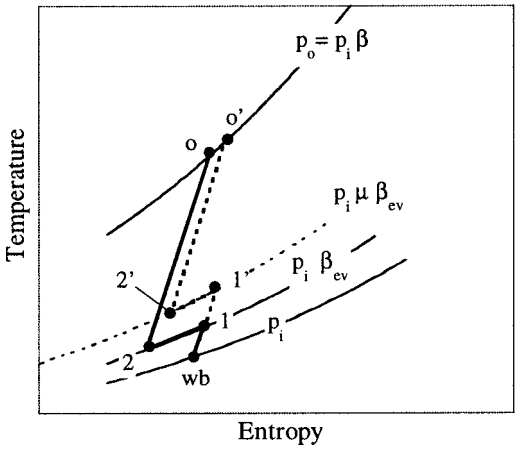


Fig. A1 Temperature-entropy diagram for overspray compression process considering different droplet diameters

lat = latent
o = compressor outlet
os = overspray
s = spray
sh = superheated
stack = flue gas at the HRSG outlet
w = water
wb = wet bulb

Acronyms

CCPP = combined cycle power plant
CF = corrected flow
CS = corrected speed
GT = gas turbine
HDC = hot day case
HR = heat rate
HRSG = heat recovery steam generator
ISO = international standard organization
LHV = low heating value
RH = relative humidity
ST = steam turbine
TIT = gas turbine inlet temperature
TOT = gas turbine outlet temperature

Appendix

Thermodynamic Consideration About Overspray. Droplet diameter influences the evaporation time and, consequently, the evaporation pressure ratio and the compressor discharge temperature. In particular, as droplet diameter increases, evaporation pressure ratio and compressor outlet temperature increase.

In Fig. A1, two compression processes (both from p_i to p_o) with different evaporation pressure p_{ev}, are presented. The compression process with dotted line takes into account effect of droplet size (or delay coefficient). Defining β the total pressure ratio ($\beta = p_o/p_i$) and β_{ev} the evaporation pressure ratio ($\beta_{ev} = p_{ev}/p_i$), and assuming k and η_p constant, [8,9], it is possible to express the temperature T_1 as

$$T_1 = T_{wb} \cdot \beta_{ev}^{(k-1)/(k\eta_p)}. \tag{A1}$$

Furthermore, indicating $T_1 - T_2 = \Delta T$, for the thermal equilibrium (Eq. (5) in which Δh_{sh} is neglected), values

$$\dot{m}_{ac} \cdot c_{pa} \Delta T = \dot{m}_{ev,os} \cdot \Delta h_{\text{lat}} \tag{A2}$$

$$\Delta T = \frac{\dot{m}_{ev,os}}{\dot{m}_{ac}} \cdot \frac{\Delta h_{\text{lat}}}{c_{pa}} = \phi \cdot \frac{\Delta h_{\text{lat}}}{c_{pa}}. \tag{A3}$$

Moreover, the compressor discharge temperature T_o can be expressed as

$$T_o = T_2 \cdot \left(\frac{\beta}{\beta_{ev}}\right)^{(k-1)/(k\eta_p)} = (T_1 - \Delta T) \cdot \left(\frac{\beta}{\beta_{ev}}\right)^{(k-1)/(k\eta_p)} \tag{A4}$$

introducing Eq. (*A*3) and (A1), Eq. (*A*4) reduces to

$$T_o = T_{wb} \cdot \beta^{(k-1)/(k\eta_p)} - \Delta T \cdot \left(\frac{\beta}{\beta_{ev}}\right)^{(k-1)/(k\eta_p)}. \tag{A5}$$

If now droplet dimension is taken into account, the water evaporation will be for a pressure ratio ($\mu\beta_{ev}$) greater than β_{ev} ($\mu \geqslant 1$ is here indicated as "delay coefficient").

It should be noted that, if no variation occurs in the injected water mass flow rate,

$$T_{1'} - T_{2'} = T_1 - T_2 = \Delta T \tag{A6}$$

and Eq. (*A*5) becomes

$$T_o = T_{wb} \beta^{(k-1)/(k\eta_p)} - \phi \frac{\Delta h_{\text{lat}}}{c_{pa}} \left(\frac{\beta}{\mu\beta_{ev}}\right)^{(k-1)/(k\eta_p)}. \tag{A7}$$

Equation (*A*7) is the same as Eq. (8) discussed in the main text.

It should be highlighted that the compressor discharge temperature ($T_{o'}$) considering the droplet dimension ($\mu > 1$) is higher than that (T_o) obtained without evaporation delay ($\mu = 1$):

$$T_{o'} > T_o. \tag{A8}$$

References

[1] Chaker, M., Meher-Homji, C. B., and Mee, III, T., 2002, "Inlet Fogging Of Gas Turbine Engines—Part A: Fog Droplet Thermodynamics, Heat Transfer And Practical Considerations," ASME Paper No. GT-2002-30562.
[2] Jolly, S., Nitzken, J., and Shepherd, D., 1998, "Evaluation Of Combustion Turbine Inlet Air System," Presented at the Power-Gen Asia, New Delhi, India, Sept. 29–Oct. 1.
[3] Tawney, R., Pearson, C., and Brown, M., 2001, "Options to Maximize Power Output for Merchant Plants in Combined Cycle Applications," ASME Paper No. 2001-GT-0409.
[4] Jones, C., and Jacobs, III, J. A., 2000, "Economic and Technical Considerations for Combined Cycle Performance Enhancement Options," GE Power Systems, GER-4200.
[5] Bhargava, R., and Meher-Homij, C. B., 2002, "Parametric Analysis of Existing Gas Turbines With Inlet Evaporative and Overspray Fogging," ASME Paper No. GT-2002-30560.
[6] *Gas Turbine World 2001–2002 Handbook*, 2002, Pequot Publishing Inc., Fairfield, CT.
[7] *Thermoflex 5.2*, Thermoflow 8, Release 1, Thermoflow, Inc., Sudbury, MA.
[8] Utamura, M., Kuwahara, T., Murata, H., and Horii, N., 1999, "Effects of Intensive Evaporative Cooling on Performance Characteristics of Land-Based Gas Turbine," Joint Power Generation Conference, PWR-Vol. 34, ASME, New York.
[9] Horlock, J. H., 2001, "Compressor Performance With Water Injection," ASME Paper No. 2001-GT-0343.

Optimal Design of Gas Turbine Cogeneration Plants in Consideration of Discreteness of Equipment Capabilities

Ryohei Yokoyama[1]
Associate Professor
Fellow ASME
e-mail: yokoyama@ese.me.osakafu-u.ac.jp

Koichi Ito
Professor
Fellow ASME

Department of Mechanical Engineering,
Osaka Prefecture University,
1-1 Gakuen-cho,
Sakai, Osaka 599-8531, Japan

To attain the highest economic and energy-saving characteristics of gas turbine cogeneration plants, it is necessary to rationally determine capacities and numbers of gas turbines and auxiliary equipment in consideration of their operational strategies corresponding to seasonal and hourly variations in energy demands. Some optimization approaches based on the mixed-integer linear programing have been proposed to this design problem. However, equipment capacities have been treated as continuous variables, and correspondingly, performance characteristics and capital costs of equipment have been assumed to be continuous functions with respect to their capacities. This is because if equipment capacities are treated discretely, the number of integer variables increases drastically and the problem becomes too difficult to solve. As a result, the treatment of equipment capacities as continuous variables causes discrepancies between existing and optimized values of capacities and expresses the dependence of performance characteristics and capital costs on capacities with worse approximations. In this paper, an optimal design method is proposed in consideration of discreteness of equipment capacities. A formulation for keeping the number of integer variables as small as possible is presented to solve the optimal design problem easily. This method is applied to the design of a gas turbine cogeneration plant, and its validity and effectiveness are clarified. [DOI: 10.1115/1.2131889]

Introduction

Because of their potential of high economic and energy-saving characteristics by efficient utilization of exhaust heat, cogeneration plants have been installed increasingly into industries, districts, and buildings in recent years. To utilize the potential and attain the highest economic and energy-saving characteristics, design and operation are important issues. For example, it is important to rationally determine their structures by selecting energy-producing and -conversion equipment from many alternative ones so that they match energy demand requirements. It is also important to rationally determine capacities and numbers of selected equipment in consideration of their operational strategies, such as on/off status of operation and load allocation corresponding to seasonal and hourly variations in energy demands.

One of the ways to rationally determine the aforementioned design and operation items of cogeneration plants is to use combinatorial optimization approaches, which are based on the mathematical programing, such as the mixed-integer linear programing (MILP) [1–5] and the mixed-integer nonlinear programing [6], as well as the meta heuristics, such as the simulated annealing [7] and the genetic algorithm [8]. For example, in the approach based on the MILP, the selection, numbers, and on/off status of operation of equipment are expressed by integer variables, and the capacities and load allocation of equipment by continuous ones. However, the optimal design problem has often been treated in consideration of single-period operation [1], or multiperiod operation for a small number of periods [2], to avoid excessive difficulty of the problem. This is because the number of integer variables increases with those of equipment and periods, and it becomes difficult to obtain the optimal solution in a practical computation time using the conventional solution algorithm for the MILP, which combines the branch and bound method with the simplex one. Recently, some efforts have been made to treat the optimal design problem in consideration of multiperiod operation for a larger number of periods [3–5]. Nevertheless, equipment capacities have still been treated as continuous variables, and correspondingly, performance characteristics and capital costs of equipment have been assumed to be continuous functions with respect to their capacities. This is because if equipment capacities are treated discretely, then the number of integer variables increases drastically and the problem becomes too difficult to solve. As a result, the treatment of equipment capacities as continuous variables causes discrepancies between existing and optimized values of capacities and expresses the dependence of performance characteristics and capital costs on capacities with worse approximations.

In this paper, an optimal design method is proposed in consideration of discreteness of equipment capacities. A formulation for keeping the number of integer variables as small as possible is presented to solve the optimal design problem easily. This method is applied to the design of a gas turbine cogeneration plant, and its validity and effectiveness are investigated.

Optimal Design Method

Summary of Optimal Design Problem. To consider seasonal and hourly variations in energy demands, a typical year is divided into multiple periods and energy demands are estimated for each period. As shown in Fig. 1, a superstructure for an energy supply plant is created to match energy demand requirements. The superstructure is composed of all the units of equipment considered as candidates for selection, and a real structure is created by selecting some units of equipment from the candidates. Furthermore,

[1]To whom correspondence should be addressed.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. GT2004-53678.

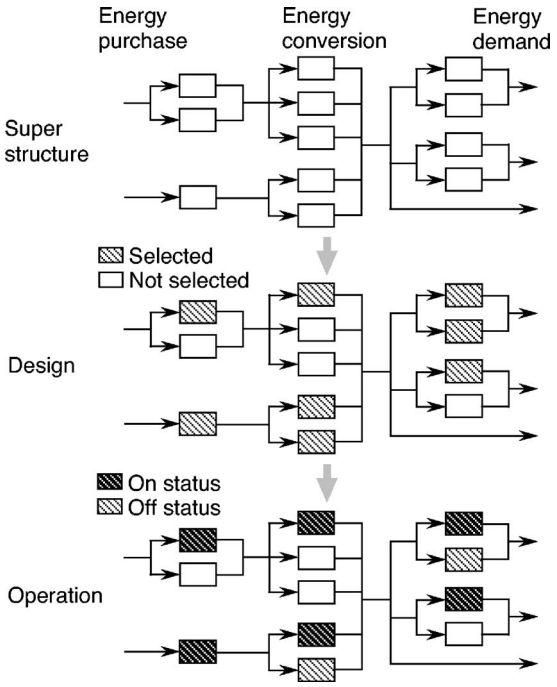


Fig. 1 Concept of superstructure

some units of equipment are operated to satisfy energy demands for each period. The selection, capacities, and numbers of equipment are considered as design variables, and the on/off status of operation and load allocation of equipment as operation ones. The hierarchical relationship between the design and operation variables is shown in Fig. 2. The selection and capacities are expressed by binary variables, the numbers and on/off status of operation by integer variables, and the load allocation by continuous variables.

As fundamental constraints, performance characteristics of equipment and energy balance relationships are considered. If necessary, other constraints, such as relationships between maximum contract demands and consumption of purchased energy and operational restrictions are considered. As the objective function to be minimized, the annual total cost is adopted, typically, and is evaluated as the sum of annual capital cost of equipment and annual operational cost of purchased energy. These constraints and objective function are expressed as functions with respect to the design and operation variables.

In the following, an optimal design problem is formulated for the energy supply plant with a simple superstructure shown in Fig. 3. The formulation can easily be extended to energy supply plants with complex superstructures.

Selection, Capacities, and Numbers of Equipment. The energy supply plant is composed of I blocks, each of which corresponds to a type of equipment. The capacity of the ith type of equipment is selected from its J_i candidates. In addition, the number of the ith type and the jth capacity of equipment is determined within its maximum N_{ij}. The selection and number of the ith type and the jth capacity of equipment are designated by the binary variable γ_{ij} and the integer variable η_{ij}, respectively. By these definitions, the following equations are obtained:

$$\left.\begin{aligned} &\eta_{ij}/N_{ij} \leqslant \gamma_{ij} \leqslant \eta_{ij} \ (j=1,2,\cdots,J_i) \\ &\sum_{j=1}^{J_i} \gamma_{ij} \leqslant 1 \\ &\gamma_{ij} \in \{0,1\} \ (j=1,2,\cdots,J_i) \\ &\eta_{ij} \in \{0,1,\cdots,N_{ij}\} \ (j=1,2,\cdots,J_i) \end{aligned}\right\} \quad (i=1,2,\cdots,I) \tag{1}$$

Here, it is assumed that multiple units with the same capacity can be selected for a type of equipment. To select multiple units with different capacities for a type of equipment, multiple blocks for the type of equipment should be included in the plant.

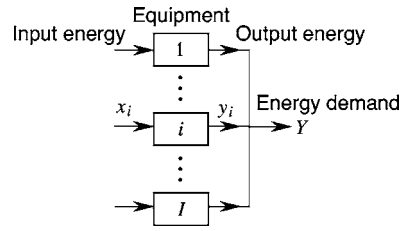


Fig. 3 Energy supply plant with simple superstructure

Performance Characteristics of Equipment. For simplicity, the case of $N_{ij}=1$ is first considered. A relationship between the flow rates of input and output energy is shown in Fig. 4, as performance characteristics of the ith type of equipment. Here, the discontinuity of the relationship due to the on/off status of operation is expressed by a binary variable, and the relationship for the on status of operation is approximated by a linear equation as follows:

$$\left.\begin{aligned} &y_i(k) = P_i x_i(k) + Q_i \delta_i(k) \\ &\underline{X}_i \delta_i(k) \leqslant x_i(k) \leqslant \bar{X}_i \delta_i(k) \\ &\delta_i(k) \in \{0,1\} \end{aligned}\right\} \quad (i=1,2,\cdots,I;\ k=1,2,\cdots,k) \tag{2}$$

where δ_i is the binary variable for the on/off status of operation, x_i and y_i are the continuous variables for the flow rates of input and output energy, respectively, and P_i, Q_i, $\underline{X}_i$, and $\bar{X}_i$ are the performance characteristic values (i.e., P_i and Q_i are the slope and intercept, respectively, of the linear relationship between the flow

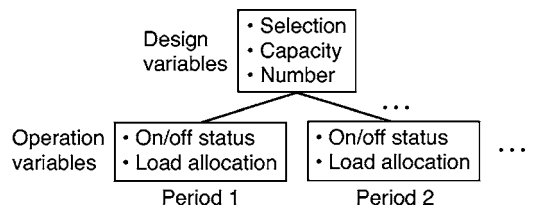


Fig. 2 Hierarchical relationship between design and operation variables

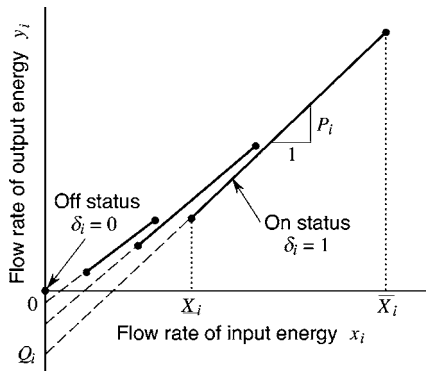


Fig. 4 Modeling of performance characteristics of equipment

rates of input and output energy, and $\underline{X}_i$ and $\bar{X}_i$ are the lower and upper limits, respectively, for the flow rate of input energy). The argument k is the index for periods, and K is the number of periods. The first equation in Eq. (2) expresses the flow rate of output energy as a function with respect to that of input energy for the on status and makes the flow rate of output energy zero for the off status. The second equation in Eq. (2) makes the flow rate of input energy within its lower and upper limits for the on status and zero for the off status.

The values of P_i, Q_i, $\underline{X}_i$, and $\bar{X}_i$ depend on the selected capacity and are expressed as follows:

$$\left.\begin{aligned} P_i &= \sum_{j=1}^{J_i} p_{ij}\gamma_{ij} \\ Q_i &= \sum_{j=1}^{J_i} q_{ij}\gamma_{ij} \\ \underline{X}_i &= \sum_{j=1}^{J_i} \underline{x}_{ij}\gamma_{ij} \\ \bar{X}_i &= \sum_{j=1}^{J_i} \bar{x}_{ij}\gamma_{ij} \end{aligned}\right\} \quad (i = 1,2,\cdots,I) \tag{3}$$

where p_{ij}, q_{ij}, $\underline{x}_{ij}$, and $\bar{x}_{ij}$ are the performance characteristic values of the ith type and the jth capacity of equipment (i.e., p_{ij} and q_{ij} are the slope and intercept, respectively, of the linear relationship between the flow rates of input and output energy, and $\underline{x}_{ij}$ and $\bar{x}_{ij}$ are the lower and upper limits, respectively, for the flow rate of input energy). In addition, the on status of operation is feasible only if a capacity is selected, and the following equation is obtained:

$$\delta_i(k) \leqslant \sum_{j=1}^{J_i} \gamma_{ij} \quad (i = 1,2,\cdots,I;\ k = 1,2,\cdots,K) \tag{4}$$

Therefore, Eq. (2) results in

$$\left.\begin{aligned} & y_i(k) = \sum_{j=1}^{J_i} p_{ij}\gamma_{ij}x_i(k) + \sum_{j=1}^{J_i} q_{ij}\gamma_{ij}\delta_i(k) \\ & \sum_{j=1}^{J_i} \underline{x}_{ij}\gamma_{ij}\delta_i(k) \leqslant x_i(k) \leqslant \sum_{j=1}^{J_i} \bar{x}_{ij}\gamma_{ij}\delta_i(k) \\ & \delta_i(k) \leqslant \sum_{j=1}^{J_i} \gamma_{ij} \\ & \delta_i(k) \in \{0,1\} \end{aligned}\right\} \quad (i = 1,2,\cdots,I;\ k = 1,2,\cdots,K) \tag{5}$$

Since δ_i is common to all the capacities, this formulation keeps the number of binary variables as small as possible, which makes the computation time as short as possible.

Equations (2)–(5) are formulated in the case of $N_{ij}=1$. However, these equations can be extended for the performance characteristics of multiple units of equipment with the same capacity in the case of $N_{ij}>1$.

As a result, Eq. (5) can be extended as follows:

$$\left.\begin{aligned} & y_i(k) = \sum_{j=1}^{J_i} p_{ij}\gamma_{ij}x_i(k) + \sum_{j=1}^{J_i} q_{ij}\gamma_{ij}\delta_i(k) \\ & \sum_{j=1}^{J_i} \underline{x}_{ij}\gamma_{ij}\delta_i(k) \leqslant x_i(k) \leqslant \sum_{j=1}^{J_i} \bar{x}_{ij}\gamma_{ij}\delta_i(k) \\ & \delta_i(k) \leqslant \sum_{j=1}^{J_i} \eta_{ij} \\ & \delta_i(k) \in \{0,1,\cdots,\max_{1\leqslant j\leqslant J_i} N_{ij}\} \end{aligned}\right\} \quad (i = 1,2,\cdots,I;\ k = 1,2,\cdots,K) \tag{6}$$

where δ_i is the integer variable for the number of equipment at the on status of operation. Here, it is assumed that δ_i units of equipment are operated at the same load level and the sums of the flow rates of input and output energy are expressed by x_i and y_i, respectively. This assumption is validated if the simple performance characteristics expressed by Eq. (2) are used. The third equation in Eq. (6) means that the number of equipment at the on status of operation may not be larger than that selected. Since δ_i is common to all the capacities, $\max_{1\leqslant j\leqslant J_i} N_{ij}$ is used as its maximum in the fourth equation in Eq. (6).

Capital Costs of Equipment. The capital cost of each unit of equipment depends on its capacity and performance characteristics. The capital cost of the ith type of equipment C_i is calculated as follows:

$$C_i = \sum_{j=1}^{J_i} c_{ij}\eta_{ij} \quad (i = 1,2,\cdots,I) \tag{7}$$

where c_{ij} is the capital cost of the ith type and the jth capacity of equipment.

Objective Function and Energy Balance Relationship. As mentioned previously, the annual total cost is adopted as the objective function z to be minimized and is expressed by

$$\begin{aligned} z &= \sum_{i=1}^{I} \left(RC_i + \varphi_i \sum_{k=1}^{K} T(k)x_i(k) \right) \\ &= \sum_{i=1}^{I} \left(R\sum_{j=1}^{J_i} c_{ij}\eta_{ij} + \varphi_i \sum_{k=1}^{K} T(k)x_i(k) \right) \end{aligned} \tag{8}$$

where R is the capital recovery factor, φ_i is the unit cost for energy charge of the input energy consumed by the ith type of equipment, and T is the duration per year of each period.

As the energy balance relationship, the following equation is considered:

$$\sum_{i=1}^{I} y_i(k) = Y(k) \quad (k = 1,2,\cdots,K) \tag{9}$$

where Y is the energy demand for each period.

Solution Method. The aforementioned formulation leads to the following optimal design problem:

minimize z of Eq. (8)
with respect to γ_{ij}, η_{ij}, $\delta_i(k)$, $x_i(k)$, and $y_i(k)$
$(i=1,2,\cdots,I;\ j=1,2,\cdots,J_i;\ k=1,2,\cdots,K)$
subject to Eqs. (1), (6), and (9).

To reformulate this optimal design problem as a MILP problem, the nonlinear terms due to the products of γ_{ij} and x_i, and γ_{ij} and δ_i in Eq. (6) are replaced by the continuous variables $\xi_{ij}(k)$ and $\zeta_{ij}(k)$, respectively, as follows:

$$\xi_{ij}(k) = \gamma_{ij}x_i(k)$$

$$(i = 1,2,\cdots,I;\ j = 1,2,\cdots,J_i;\ k = 1,2,\cdots,K) \tag{10}$$

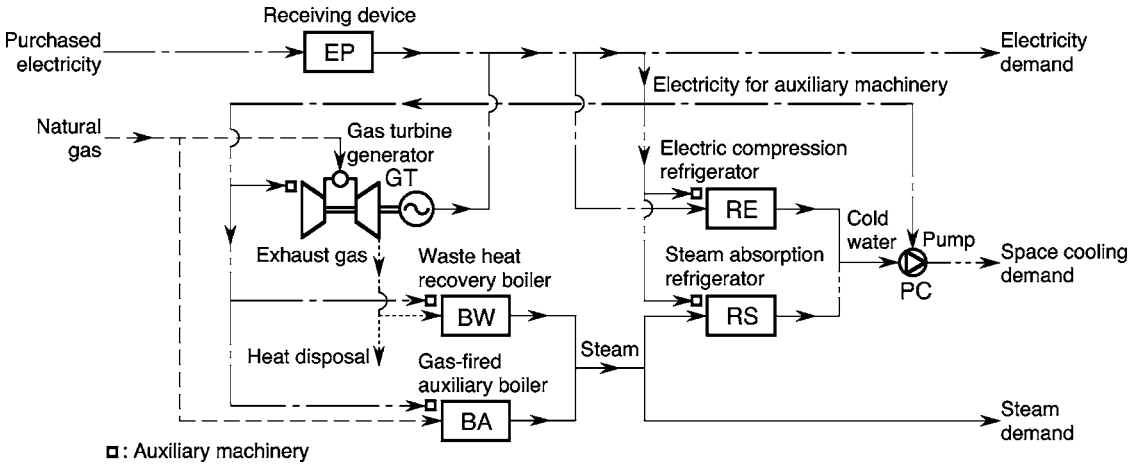


Fig. 5 Superstructure of gas turbine cogeneration plant

$$\zeta_{ij}(k) = \gamma_{ij}\delta_i(k)$$

$$(i = 1,2,\cdots,I;\ j = 1,2,\cdots,J_i;\ k = 1,2,\cdots,K) \tag{11}$$

In addition, the following constraints are employed:

$$\left.\begin{aligned} &\underline{x}_i(k)\gamma_{ij} \leqslant \xi_{ij}(k) \leqslant \tilde{x}_i(k)\gamma_{ij} \\ &x_i(k) + \tilde{x}_i(k)(\gamma_{ij} - 1) \leqslant \xi_{ij}(k) \leqslant x_i(k) \end{aligned}\right\}$$

$$(i = 1,2,\cdots,I;\ j = 1,2,\cdots,J_i;\ k = 1,2,\cdots,K) \tag{12}$$

$$\left.\begin{aligned} &\underline{\delta}_i(k)\gamma_{ij} \leqslant \zeta_{ij}(k) \leqslant \tilde{\delta}_i(k)\gamma_{ij} \\ &\delta_i(k) + \tilde{\delta}_i(k)(\gamma_{ij} - 1) \leqslant \zeta_{ij}(k) \leqslant \delta_i(k) \end{aligned}\right\}$$

$$(i = 1,2,\cdots,I;\ j = 1,2,\cdots,J_i;\ k = 1,2,\cdots,K) \tag{13}$$

where $\underline{(\,)}$ and $\tilde{(\,)}$ are lower and upper bounds, respectively, and these values can be set as follows:

$$\left.\begin{aligned} &\underline{x}_i(k) = 0 \\ &\tilde{x}_i(k) = \max_{1\leqslant j\leqslant J_i} N_{ij}\bar{x}_{ij} \end{aligned}\right\} \quad (i = 1,2,\cdots,I;\ k = 1,2,\cdots,K) \tag{14}$$

$$\left.\begin{aligned} &\underline{\delta}_i(k) = 0 \\ &\tilde{\delta}_i(k) = \max_{1\leqslant j\leqslant J_i} N_{ij} \end{aligned}\right\} \quad (i = 1,2,\cdots,I;\ k = 1,2,\cdots,K) \tag{15}$$

The constraints of Eqs. (12) and (13) are because, for example, Eq. (12) means that if $\gamma_{ij}=0$, then $\xi_{ij}(k)=0$, or else if $\gamma_{ij}=1$, then $\xi_{ij}(k)=x_i(k)$, which makes Eq. (10) valid, indirectly. This procedure can linearize the nonlinear terms without any approximations and transform the optimal design problem into a MILP one.

The reformulated MILP problem can be solved using the commercial solver GAMS/CPLEX [9].

Numerical Study

Plant Structure. Figure 5 shows the superstructure of a gas turbine cogeneration plant investigated as an example. This plant is composed of gas turbine generators, waste heat recovery boilers, gas-fired auxiliary boilers, electric compression refrigerators, steam absorption refrigerators, a receiving device for purchasing electricity, and pumps for supplying cold water.

Electricity is supplied to users by operating the gas turbine generators and purchasing electricity from an outside electric power company. Electricity is also used to drive the electric compression refrigerators, pumps, and other auxiliary machinery in the plant. Exhaust heat generated from the gas turbines is recovered in the form of steam by the waste heat recovery boilers, and is used for heat supply. An excess of exhaust heat is disposed of to the environment in the form of exhaust gas. A shortage of steam is supplemented by the gas-fired auxiliary boilers. Cold water for space cooling is supplied by the electric compression and steam absorption refrigerators. Steam is used for space heating and hot water supply.

Input Data and Conditions. A district that is supplied with energy is investigated as an example. It is composed of hotels and office buildings, and its total floor area is 383.7×10^3 m^2. Energy demands are estimated by setting three representative days per year (i.e., summer, winter, and midseason days) and dividing each representative day into six sampling times each of which has 4 hr. Peak energy demands in summer and winter are also considered. As a result, the number of periods is $K=20$. Table 1 shows the definition of periods.

It is assumed that the gas turbine generators and waste heat recovery boilers are selected together as cogeneration units. It is also assumed that a unit of the receiving device is selected, and its capacity is considered as a continuous variable. The pumps are common to all the possible structures, and only their power consumption is considered. As a result, the number of types of equipment is $I=4$. The number of capacities is set at $J_1=10$ for the gas turbine cogeneration units, and those are set at $J_i=4$ ($i=2, 3, 4$) for the gas-fired auxiliary boilers, and electric compression and steam absorption refrigerators, respectively. The maximum numbers of selected equipment are set at $N_{ij}=4$ ($i=1,2,\cdots,4$; $j=1,2,\cdots,4$) for all types and capacities of equipment. Table 2 shows the capacities and performance characteristics of candidates of equipment for selection. Although the performance characteristic values are shown only at the rated load status, their changes at the part-load status are also taken into account. As an example, Figs. 6(a) and 6(b) show the approximated values of all the candidates, and the comparison of the cataloged and approximated values of some candidates, respectively, as the performance characteristics of the gas turbine cogeneration units.

As mentioned previously, the annual total cost is adopted as the

Table 1 Definition of periods

Period k	Representative day	Duration T hr/yr
1–6	Summer	368
7–12	Winter	484
13–18	Midseason	608
19	Summer peak	0
20	Winter peak	0

Table 2 Capacities and performance characteristics of candidates of equipment for selection

Equipment	Capacity/performance[a]	Candidate			
	#	1	2	3	4
	Max. power output (MW)	1.29	1.60	2.00	2.40
	Max. steam output (MW)	5.69	3.34	4.10	4.57
	Power-generating efficiency	0.140	0.173	0.169	0.179
	Heat recovery efficiency	0.617	0.362	0.347	0.341
	#	5	6	7	8
Gas turbine cogeneration unit	Max. power output (MW)	2.93	3.50	3.54	4.36
	Max. steam output (MW)	6.44	6.97	6.89	8.92
	Power-generating efficiency	0.256	0.271	0.273	0.273
	Heat recovery efficiency	0.563	0.540	0.531	0.559
	#	9	10		
	Max. power output (MW)	5.23	5.32		
	Max. steam output (MW)	8.91	9.05		
	Power-generating efficiency	0.301	0.306		
	Heat recovery efficiency	0.513	0.521		
Gas-fired auxiliary boiler	#	1	2	3	4
	Max. steam output (MW)	5.24	6.55	7.86	9.82
	Thermal efficiency	0.92	0.92	0.92	0.92
Electric compression refrigerator	#	1	2	3	4
	Max. cooling output (MW)	2.82	3.52	4.22	5.28
	Coefficient of performance	4.57	4.73	4.76	5.04
Steam absorption refrigerator	#	1	2	3	4
	Max. cooling output (MW)	3.46	5.18	6.91	8.64
	Coefficient of performance	1.20	1.20	1.20	1.20

[a]At rated load status

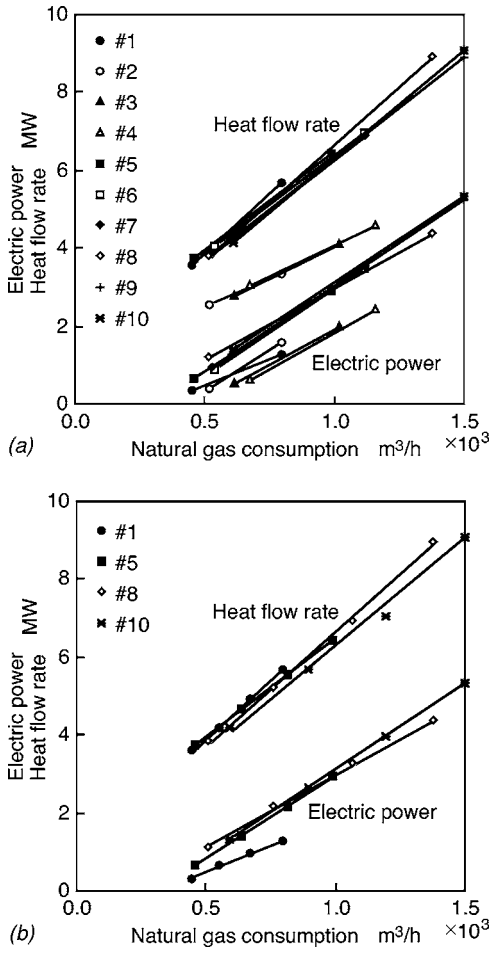


Fig. 6 Performance characteristics of gas turbine cogeneration units: (*a*) approximated values of all candidates and (*b*) comparison of cataloged and approximated values of some candidates

objective function from the economic viewpoint. Table 3 shows the capital unit costs of equipment, and the unit costs for demand and energy charges of electricity and natural gas purchased from outside utility companies. Here, the demand charges of electricity and natural gas in proportion to their maximum contract demands are included in the annual total cost. In evaluating the annual total cost, the capital recovery factor is set at 0.964 with the interest rate 0.05 and the life of equipment 15 yr. In addition, the annual primary energy consumption is also evaluated from an energy-savings viewpoint, with the thermal efficiency of purchased electricity 0.388 and the lower heating value of natural gas 11.57 kWh/m^3.

Results and Discussion. The optimization calculation is conducted for the following three plants:

1. Plant A: gas turbine cogeneration plant without constraint
2. Plant B: gas turbine cogeneration plant with the constraint of no disposal of exhaust heat
3. Plant C: conventional plant without gas turbine cogeneration unit

Plants A and B are the gas turbine cogeneration plants to be investigated. Plant C is a conventional plant as a reference to evaluate the economic and energy-saving characteristics of the cogeneration plants. It may be interesting to analyze the trade-off relationship between the economic and energy-saving characteristics by the multiobjective optimization with the annual total cost and primary energy consumption as the objective functions [10]. Here, in place of the multiobjective optimization, the single-objective optimization is conducted with the annual total cost as the objective function, and how the constraint of no disposal of exhaust heat affects the design and operation as well as the economic and energy-saving characteristics is investigated. For this purpose, plant A has no constraint, and plant B has the constraint.

Table 4 shows the optimal values of capacities and numbers of equipment, and maximum contract demands of electricity and natural gas for plants A to C. In plant A, to utilize the potential of cost saving by cogeneration, three units of the largest gas turbine cogeneration unit are selected. According to Table 2 and Fig. 6, this unit has the highest power-generating efficiency at the rated load status and keeps a relatively high power-generating efficiency, even at part-load status, although it does not have the

Table 3 Capital unit costs of equipment, and unit costs for demand and energy charges of utilities

Equipment/utility		Unit cost
Gas turbine generator		230.0×10^3 yen/kW
Waste heat recovery boiler		9.6×10^3 yen/kW
Gas-fired auxiliary boiler		6.6×10^3 yen/kW
Electric compression refrigerator		34.4×10^3 yen/kW
Steam absorption refrigerator		30.1×10^3 yen/kW
Receiving device		56.3×10^3 yen/kW
Electricity	Demand charge	1740 yen/(kW month)
	Eenergy charge	10.77 yen/kWh (summer) 9.79 yen/kWh (other seasons)
Natural gas	Demand charge	2033 yen/(m^3/hr month)
	Energy charge	30.88 yen/m^3

highest exhaust heat recovery efficiency. This is because a large amount of exhaust heat is disposed of as shown below and, consequently, the unit does not necessarily need a high exhaust heat recovery efficiency. In relation to this result, a unit of the smallest gas-fired auxiliary boiler is selected, and the capacity of the receiving device is relatively small. To utilize exhaust heat efficiently, four units of the largest steam absorption refrigerator and no electric compression refrigerator are selected. In plant B, the constraint of no disposal of exhaust heat decreases the number of the largest gas turbine cogeneration units from three to two. Although the capacity and number of the gas-fired auxiliary boiler does not change, the capacity of the receiving device becomes relatively large. Since the amount of exhaust heat decreases, the capacity of the steam absorption refrigerators also decreases, and a unit of the largest electric compression refrigerator is selected. In plant C, three units of a larger gas-fired auxiliary boiler are selected, and the capacity of the receiving device becomes large. The capacities and numbers of the steam absorption refrigerators and electric compression refrigerator are the same as those in plant B.

Table 5 shows the annual total cost and its items, and the annual primary energy consumption for plants A to C. The annual amount of exhaust heat disposed of is also included. In plant A, the capital cost increases while the demand and energy charges decrease as compared to those in plant C. As a result, the total cost decreases by 14.8%. This decrease may be attractive from an economic viewpoint; however, a large amount of exhaust heat is disposed of, and consequently, the primary energy consumption decreases only by 2.4%. This decrease may not be reasonable from an energy-saving viewpoint. In plant B, the total cost decreases by 9.6% while the primary energy consumption decreases by 10.3% as compared to those in plant C. This result indicates that the design and operation of cogeneration plants may change their economic and energy-saving characteristics significantly.

Figures 7 and 8 show the optimal operational strategies for electricity, steam, and cold water supply for plants A and B, respectively. In plant A, the gas turbine cogeneration units are operated at the rated or high part-load status. As a result, a large amount of exhaust heat is disposed of in most of the periods. This is because the power-generating efficiency of the gas turbine cogeneration units decreases drastically at low part-load status. In plant B, the gas turbine cogeneration units are operated at low part-load status or stopped not to dispose of exhaust heat in the periods with low thermal energy demands.

Conclusions

An optimal design method to determine the selection, capacities, and numbers of equipment of energy supply plants has been

Table 4 Optimal values of capacities and numbers of equipment, and maximum contract demands of utilities

Plant	Equipment/utility	Candidate	Number	Capacity
A	Gas turbine cogeneration unit	10	3	15.96 MW[a]
	Gas-fired auxiliary boiler	1	1	5.24 MW
	Electric compression refrigerator	-	-	0.00 MW
	Steam absorption refrigerator	4	4	34.56 MW
	Receiving device	-	-	7.57 MW
	Electricity	-	-	7.57 MW
	Natural gas	-	-	4.53×10^3 m^3/hr
B	Gas turbine cogeneration unit	10	2	10.64 MW[a]
	Gas-fired auxiliary boiler	1	1	5.24 MW
	Electric compression refrigerator	4	1	5.28 MW
	Steam absorption refrigerator	3	4	27.64 MW
	Receiving device	-	-	13.66 MW
	Electricity	-	-	13.66 MW
	Natural gas	-	-	3.47×10^3 m^3/hr
C	Gas turbine cogeneration unit	-	-	0.00 MW[a]
	Gas-fired auxiliary boiler	3	3	23.58 MW
	Electric compression refrigerator	4	1	5.28 MW
	Steam absorption refrigerator	3	4	27.64 MW
	Receiving device	-	-	24.15 MW
	Electricity	-	-	24.15 MW
	Natural gas	-	-	2.17×10^3 m^3/hr

[a]Maximum power output

Table 5 Annual costs and primary energy consumption

Plant	Cost/energy	Value
A	Capital cost	523.3×10^6 yen/yr
	Demand charge	268.5×10^6 yen/yr
	Energy charge	728.6×10^6 yen/yr
	Total cost	1520.4×10^6 yen/yr
	Heat disposal	62.2 GWh/yr
	Primary energy consumption	262.9 GWh/yr
B	Capital cost	427.6×10^6 yen/yr
	Demand charge	369.9×10^6 yen/yr
	Energy charge	816.0×10^6 yen/yr
	Total cost	1613.5×10^6 yen/yr
	Heat disposal	0.0 GWh/yr
	Primary energy consumption	241.7 GWh/yr
C	Capital cost	243.6×10^6 yen/yr
	Demand charge	557.2×10^6 yen/yr
	Energy charge	983.1×10^6 yen/yr
	Total cost	1783.9×10^6 yen/yr
	Heat disposal	0.0 GWh/yr
	Primary energy consumption	269.5 GWh/yr

proposed in consideration of discreteness of equipment capacities. An optimal design problem has been formulated as a mixed-integer linear programing one, and a formulation for keeping the number of integer variables as small as possible has been presented to solve the problem easily. This method has been applied to the design of a gas turbine cogeneration plant. Through a numerical study, it has turned out that the proposed method can rationally determine the selection, capacities, and numbers of equipment. It has also turned out that the method can also clarify the influence of constraints on the design and operation of the gas turbine cogeneration plant.

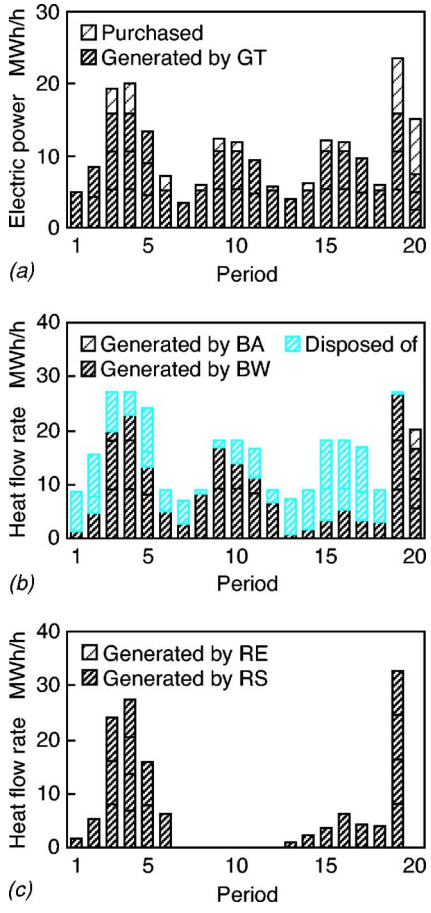


Fig. 7 Optimal operational strategies for plant A: (*a*) electricity supply, (*b*) steam supply, and (*c*) cold water supply

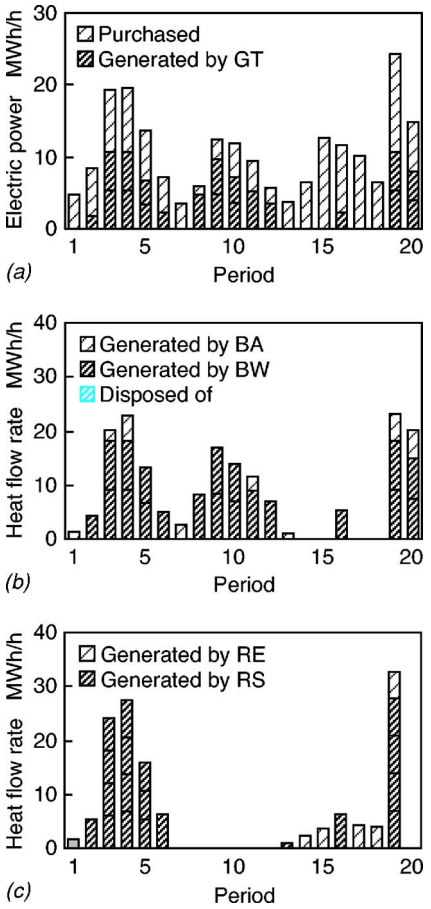


Fig. 8 Optimal operational strategies for plant B: (*a*) electricity supply, (*b*) steam supply, and (*c*) cold water supply

Nomenclature

C,c = capital cost of equipment
I = number of types of equipment
J = number of capacities of equipment
K = number of periods
k = index for periods
N = maximum number of selected equipment
P,p = slope of linear relationship between flow rates of input and output energy of equipment
Q,q = intercept of linear relationship between flow rates of input and output energy of equipment
R = capital recovery factor
T = duration per year of period
$\underline{X},\underline{x}$ = lower limit for flow rate of input energy of equipment
$\bar{X},\bar{x}$ = upper limit for flow rate of input energy of equipment
x = flow rate of input energy
Y = energy demand
y = flow rate of output energy
z = annual total cost
γ = selection of equipment
δ = number of equipment at on status of operation
ζ = product of γ and δ

η = number of selected equipment
ξ = product of γ and x
φ = unit cost for energy charge of input energy
$\underset{\sim}{(\,)}$ = lower bound
$\widetilde{(\,)}$ = upper bound

Subscripts

i = index for types of equipment
j = index for capacities of equipment

References

[1] Papoulias, S. A., and Grossmann, I. E., 1983, "A Structural Optimization Approach in Process Synthesis—I: Utility Systems," Comput. Chem. Eng., **7**(6), pp. 695–706.

[2] Horii, S., Ito, K., Pak, P. S., and Suzuki, Y., 1987, "Optimal Planning of Gas Turbine Co-Generation Plants Based on Mixed-Integer Linear Programming," Int. J. Energy Res., **11**(4), pp. 507–518.

[3] Iyer, R. R., and Grossmann, I. E., 1998, "Synthesis and Operational Planning of Utility Systems for Multiperiod Operation," Comput. Chem. Eng., **22**(7-8), pp. 979–993.

[4] Yokoyama, R., Hasegawa, Y., and Ito, K., 2000, "Structural Optimization of Energy Supply Systems by a Decomposition Method for Mixed-Integer Linear Programming," Trans. JSME, Ser. C, **66**(652), pp. 4016–4023 (in Japanese).

[5] Yokoyama, R., Hasegawa, Y., and Ito, K., 2002, "A MILP Decomposition Approach to Large Scale Optimization in Structural Design of Energy Supply Systems," Energy Convers. Manage., **43**(6), pp. 771–790.

[6] Adjiman, C. S., Schweiger, C. A., and Floudas, C. A., 1998, "Mixed-Integer Nonlinear Optimization in Process Synthesis," *Handbook of Combinatorial Optimization*, Kluwer, Dordrecht, Vol. 1, pp. 1–76.

[7] Painton, L. A., and Diwekar, U. M., 1994, "Synthesizing Optimal Design Configurations for a Brayton Cycle Power Plant," Comput. Chem. Eng., **18**(5), pp. 369–381.

[8] Fujita, K., Akagi, S., Hirokawa, N., and Yoshida, K., 1998, "Optimal Planning Method of Energy Plant Configurations Based on a Genetic Algorithm," Trans. JSME, Ser. C, **64**(617), pp. 354–361 (in Japanese).

[9] Brooke, A., Kendrick, D., Meeraus, A., and Raman, R., 1998, *GAMS a User's Guide*, GAMS Development Corp., Washington, DC.

[10] Yokoyama, R., and Ito, K., 1995, "Multi-Objective Optimization in Unit Sizing of a Gas Turbine Cogeneration Plant," ASME J. Eng. Gas Turbines Power, **117**(1), pp. 53–59.

Gas Turbine Compressor Washing: Historical Developments, Trends and Main Design Parameters for Online Systems

Friederike C. Mund

Pericles Pilidis

Department of Power, Propulsion & Aerospace Engineering,
School of Engineering,
Cranfield University,
Cranfield, Bedfordshire MK43 0AL, UK

By being exposed to atmospheric conditions gas turbines are inevitably subjected to sources of fouling. The resulting degradation can be partially recovered by cleaning the compressor. Based on open literature and patents, the different approaches leading to the most advanced method of compressor online washing have been compiled. The origins of online washing and the development trends over the decades are outlined, and the current systems are categorized. The introduction of system categories has been justified by a field survey. Additionally, the main design parameters of online washing systems are summarized. [DOI: 10.1115/1.2132378]

Introduction

The degradation caused by the adherence of particles on the compressor airfoil and annulus surfaces is conventionally defined as fouling. The buildup of material changes the shape of the airfoils, changes the incidence angle of the following airfoils, increases the surface roughness, and decreases the throat area of the compressor [1]. With a reduced efficiency, pressure ratio, and stall margin, the power output is limited at constant firing temperature. Typically, 70–85% of the overall performance loss during operation can be attributed to compressor fouling [2]. It can be partially recovered by compressor cleaning, and suggested measures for a compressor cleaning procedure are a reduced mass flow by up to 3% [3,4], a reduced power output by 3%, or a reduction of the overall pressure ratio by 2% [5].

Methods of Compressor Cleaning

To recover performance losses due to fouling, compressors are cleaned to remove the deposited particles. Several approaches have been taken over the years to wash fouled compressors, and these will now be described here. A summary of advantages and disadvantages for each method is given in Table 1.

The most obvious method to wash a compressor is the manual procedure. The engine has to be disassembled and cleaned manually using brushes and detergent. Simply cleaning the IGVs by hand will often bring significant benefits, but it is laborious, requires a shutdown and cooling of the engine, and is rarely needed [6]. However, Leusden et al. [7] and Abdelrazik and Cheney [8] stated that the level of power recovery cannot be achieved by any other cleaning method. The authors also consider a proposed method to insert a washing hose through the blade pitches up to the eighth compressor stage as a manual method; one "patiently" has to insert the hose [9].

To avoid the manual labor, abrasive so-called grit-blasting methods were developed. The deposits were removed by injecting charcoal, rice, nutshells, or synthetic resin particles into the airstream of the operating engine. The procedure was undertaken during engine operation because high air speeds were needed to accelerate the particles to achieve the necessary impact force [10]. Satisfactory cleaning results were reported in extreme cold environments with this simple and fast method without downtime. However, the removal of oily deposits was not entirely satisfactory since the last stages remained contaminated and care had to be taken not to contaminate seals and the internal passages of the secondary air systems of the engine [11]. Despite the disadvantages, it was widely used for aero engines [12]. General Electric claimed to overcome the disadvantages of abrasive cleaning in a patent for high-bypass turbofan applications by using materials of 70% carbon content [13]. An increased roughness due to the abrasion was avoided, and with remaining blade smoothness the compressor capacity was preserved. It was further claimed that no remainder of abrasives would cause the clogging of cooling holes.

In the 1980s, concerns increased regarding erosion due to the impact of particles. Especially with the introduction of protective blade coatings, damage of the first stages evoked further damage downstream [12]. For older engines, the use of ground spent catalysts, rice hulls, and nut shells may still have been satisfactory, but for state-of-the-art engines, this practice disappeared due to the potential damage [6,14]. A milder method to wash off the deposits involves liquid injection in the form of demineralized water with and without the addition of detergents. This was found to be more effective than the abrasive cleaning methods [15] and has become the leading applied method. It can be clustered into two different processes, offline and online.

The soak or crank wash, also referred to as offline compressor wash, requires a shutdown of the engine and is performed unfired, with the starter motor turning the engine [16]. The rotational speed is therefore on the order of 20% [17–19] to a maximum of 30% [8] of the normal operational speed, resulting in a reduced airflow through the engine. The downtime increases for larger, heavier engines (V93: 8 h [20], MS7001E: 6 h [21]). It accumulates from cooling the engine, spraying the cleaning fluid into the engine, running it at reduced speed, and rinsing the engine at higher speeds. Furthermore, the intake needs to be cleaned, and while starting the turbine, the compressor dries. Even though it is now the leading applied method for aero engines [17,22], early concerns addressed the dispersive characteristics of the injected fluid. According to Langford [13], significant fractions of the blade contamination in aero-engine applications consisted of engine material resulting from the rubbing parts. Langford therefore

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received September 23, 2004; final manuscript received July 27, 2005. Review conducted by R. Yokoyama.

Table 1 Advantages and disadvantages of compressor cleaning methods

Method	Advantages	Disadvantages
Manual cleaning (brushes, washing agent)	Very effective	Shut down of engine Laborious
Grit blasting (charcoal, rice, nut shells, synthetic resin particles)	Simple and fast No engine downtime Effective in cold environments	Less effective at rear stages and for oily deposits Clogging of internal cooling passages Erosion Increased surfaces roughness Damage of blade coatings
Soak, crank, offline washing (demineralized water, washing agents)	Very effective	Shutdown of engine
Fired, online washing (demineralized water, washing agents)	No interference with load profile Extends intervals of crank washes	Less effective Cannot replace offline washing (complementary)

considered the liquid solvents to chemically attack not only the aluminum contaminants, but also engine parts made of the same material, causing chemical erosion.

The efficiency of crank washes is very high, and the power recovery is close to the original level or the level reached after a major overhaul [15]. However, the downtime results in lost revenue for the owner. Aero-derivative engines, being lightweight with lower metal mass, may cool down within 1.5–3 h reducing the downtime significantly [6]. The engine has to be cooled to avoid thermal stresses of the components [16,23]. Initially, the fluid was only preheated to allow for an earlier injection and the higher fluid temperature did not increase the washing efficiency [24]. Recent numerical investigations also indicated that the droplet temperature adjusts close to airflow temperature before entering the compressor, independent of initial injection temperature. It was therefore suggested that preheating does not affect the droplet temperature in the compressor [25]. However, Fielder [16] reported improved washing efficiency with preheated fluids in marine applications. Because of the required cooldown, crank washes are often performed concurrently with other maintenance work on the gas turbine.

A further method to clean compressors using liquid injection is the so-called online or fired washing. It is considered to be the most advanced method because it is performed fired, generally under full load with no or little reduction of the unit's capacity and speed [3,26]. An official definition of online washing can be found in an instruction sheet of the Association of Power and Heat Generating Utilities [27]: "At rated (nominal) shaft speed, generally demineralized water is injected into the engine upstream of the first compressor stage." Being considered to be not as efficient as offline cleaning, frequencies of daily to once per month and durations between 5 and 30 min were recommended, complementary to offline washes. This requires a permanently installed washing device, but offers the opportunity of a more frequent washing scheme during engine operation. For cleaning with fluids, demineralized water, solvents, and surfactants are most commonly used [19] for the injection from a spray ring upstream of the compressor. Modern cleaning agents consist of surfactants to reduce surface tension and solubilize and disperse dirt, soaps or degreasers to enhance wetting abilities, high-temperature carriers to keep the cleaning agent in solution, inhibitors to reduce redepositing, and chelates against calcium or magnesium scales [28]. The washing device is provided by the original equipment manufacturer (OEM) or may be installed or replaced by a retrofit washing device from a supplier.

Generally, it is stated that online and offline washes are complementary. Frequent online washing helps to decrease the fouling rate, but over longer periods a gradual performance loss will still occur. Offline washing is more effective and must therefore be used to recover these losses [26]. Using online procedures between crank washes extends their interval [6]. But extending the intervals between washes increases the risk of larger portions of insoluble compounds being washed off at the front end of the compressor and, consequently, contaminating the blade rows downstream [29]. The equipment for online and offline washing is, in most cases, not identical. For online washing, generally, more nozzles and of different types are used and nozzles for offline washing are fitted in closer proximity to the compressor inlet face [7,10,29,30]. The offline procedure being carried out at much lower rotational speeds allows for higher injected fluid rates and larger droplet distributions without putting the components at risk [7].

Considerations toward economical aspects of compressor washing can be found, to a greater extent, in the literature. A comprehensive review of compressor-washing techniques that summarizes online and offline cleaning benefits and penalties from extensive field data can be found in Stalder [29]. But with varying ambient conditions, atmospheric pollutants, and different types of engines, it is difficult to generalize on a standard compressor-washing scheme. The most beneficial combination of online and offline washes is often a case of trial and error and is tailored to an individual plant. Saravanamuttoo [31] indicated in a comment that it does not appear likely that a universal cleaning procedure can be obtained, but discussions of online and offline procedures have since been controversial [32]. Based on field experience, individual criteria of when and how the washing should be performed and what type cleaning fluid would be appropriate were identified by Haub and Hauhe [21] and Stalder [29].

Historical Perspective of Compressor Online Washing Developments

Patents and publications are good indicators to the state of the art. In particular, they can give insight into obstacles and solutions where a technology is applied. Based on these sources of open literature, the historical developments of online washing are presented for the last four decades. As online and offline washing methods are complementary, the development of the two systems is closely related. The proceedings in the field of offline washing, which provide interesting approaches and new ideas that potentially could improve online washing, are therefore included. Considering the definition of online washing with respect to the rotational speed of the shaft, a gradual movement from offline to online washing started in the early 1970s. The rotational speed of the engine during the spray injection was rising from crank speed to higher levels, and the washing was performed at higher and higher loads. However, the early literature often did not provide the necessary details to distinguish between offline and online washing. In such cases, it was not mentioned whether or not the engine was fired or whether the engine speed or load was given. Further details in the text about cooldown or soaking procedures were taken into account and are stated in context.

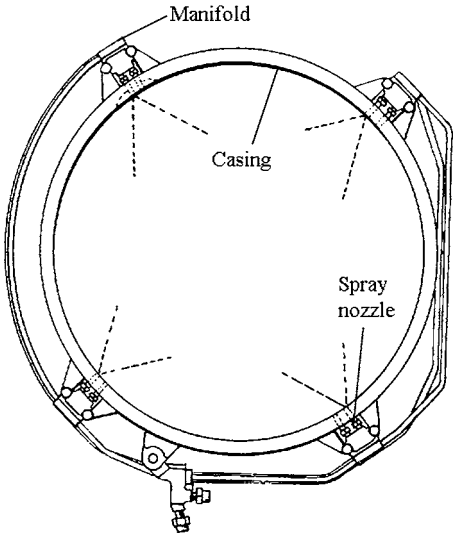


Fig. 1 Removable spray manifold by Freid and Tapparo [33]

1970s. In the 1970s, the first publications became available addressing compressor washing. The predominant development issue was that spray systems should not disturb the airflow pattern and cause compressor performance loss due to inlet distortion. A patent filed in 1970 mentioned cleaning by spraying fluid into the intake airstream of a compressor [33]. The system was a removable manifold assembly including four nozzles, and the location of the nozzles were specified such that no interference with the airflow took place. The position of the nozzles was claimed to be ideal when not protruding into the airflow (see Fig. 1). This indicated that the washing procedure was designed for engine operation; however, the rotational speed was not mentioned. A uniform spray throughout the intake was suggested, generally injected transverse to the direction of the airflow.

Mansson [34] stated that removed dirt from the low-pressure compressor redeposited on the blades of the high-pressure compressor and offered a solution for the problem of washing later stages within the compressor. Although the injection of water into the airflow was only described for the startup of the engine at low rotational speed and airflow, Mansson suggested spray injectors within the connecting or interstage duct between two industrial compressors (see Fig. 2). Hydraulically operated nozzles were moved through apertures in the casing into the internal flow path to inject water parallel into the airstream. A shower of finely dispersed water was injected at about midradius of the duct. After the washing process, the spray nozzles were automatically withdrawn from the duct so as to not introduce turbulence into the air and effect the operation of the high-pressure compressor. This idea has not become a general applicaion; however, interstage injection has recently become of greater interest and is discussed in the section "Recent Development Trends."

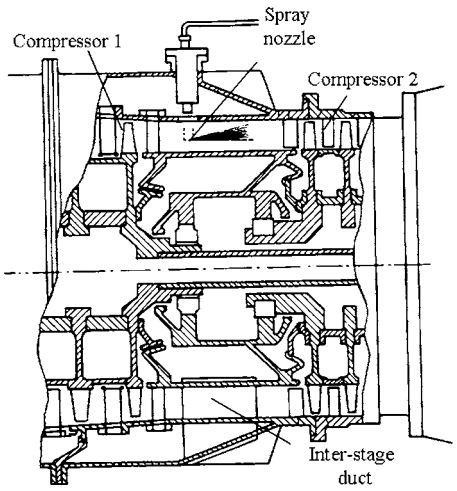


Fig. 2 Inter-compressor injection nozzle by Mansson [34]

But as for all newly introduced methods early in their development status, putting online washing into practice went along with negative experiences as well. Elser [35] discussed the cleaning procedure for an aero-derivative engine specified by the OEM. Even though it was possible to perform the wash during normal scheduled engine operation, the power recovery was not satisfactory. The specification for the wash was based on operating hours rather than power loss, and the accuracy of the performance measurements was too low to show the gained power after each wash. A slow degradation of the engine of 3.9% of the power output over 1375 operating hours was observed, although the washes were performed every two days as recommended. Based on these experiences and the high cost of the cleaning fluid, Elser [36] later suggested an offline soak-wash scheme and reported the successful and cheaper replacement of all washes during operation with regular soak washes. Based on field experience, the OEM shortly after presented the fired washing procedure only for continuous-duty power plants or intermittent-duty engines operating in sandy or desert condition [5]. The recommended washing frequency of 6–8 days was then based on a reduced pressure ratio of 2% at given rotational speed. A mixture of kerosene, water, and alcohol, depending on ambient temperature, was to be injected into the compressor just above the rotational speed where the bleed valves closed. The engine was then running well away from full load, at about half of the design air mass flow. Strictly speaking, this method cannot be considered online because of the reduced shaft speed. However, it was initiated in the 1970s when washing procedures moved toward spray injection at higher loads.

1980s. Within the next decade, washing procedures were established without interference of the gas turbine load profile. The development turned toward the properties of the injected sprays. Attempts were made to describe effective spray patterns qualitatively, and chemical properties of the washing agents were optimized. A description of a fired cleaning procedure of a large power plant was first reported by Adams and Schmitt-Wittrock [20]; but still at the time, a slight reduction of the power output was required during wash cycles. The underlying reason being a reduced stall margin of the compressor during the washing procedure, the requirement for power reduction is a function of the compressor design and surge margin. The washing procedure by an OEM for an industrial heavy-duty gas turbine suggested the injection at full speed but reduced load, where a shutdown of the combined-cycle power plant was not necessary [4]. McDermott explicitly mentioned the full- or near-full-load operating condition in his patent of a method and apparatus for cleaning gas turbine engines, which was filed in 1986 [37], and claimed best results when the engine was operating at full speed and the compressor at maximum speed. In 1989, the term "online washing" was attributed, by Thames et al. [6], to not interfere with continuous gas turbine operations. They also stated that online washing relies on the principle of keeping an engine clean rather than allowing the engine to degrade significantly and acknowledged economical benefits of extended time intervals between offline washes via online washing.

It was generally observed that spray characteristics were important for achieving a good cleaning effect and to avoid component damage, but the spray patterns were only described qualitatively. A first hint toward blade stresses and subsequent failure due to the impingement of large droplets from "bulky" wash sprays was found in a patented offline washing apparatus in 1980 [38]. McDermott [37] claimed for his proposed washing system that the formation of spray sheets that were centrifuged toward the casing

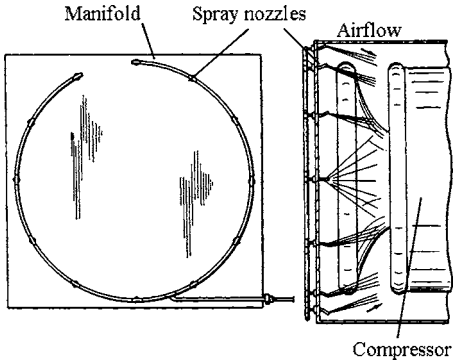


Fig. 3 Spray nozzle manifold (left) and installation opposite of the compressor (right) suggested by McDermott [37]

of the compressor was avoided by injecting the fluid across the airstream rather than parallel to it. The installation of the spray manifold was suggested opposite the compressor or on the bell-mouth side, as illustrated in Figs. 3 and 4, respectively. McDermott attributed a proper dispersion of the fluid to the injection of flat-fan spray patterns into turbulent zones of relatively low airspeed in front of the engine (see Fig. 4) and claimed that a fog of fine droplets was formed such that the droplets were too small to be strongly affected by the centrifugal forces within the compressor. Becker and Bohn [4] also reported that to avoid blade erosion a droplet mist has to be ejected over the entire blading cross section. However, the system presented for a heavy-duty gas turbine also used two additional nozzles with a low jet dispersion to vigorously spray through the inlet guide vanes onto the suction side of the slowly passing rotor blades. This approach was an early response to the conflicting design variable of the droplet size for effective cleaning systems. Small droplets would avoid erosion because of their lower impact and would tend to follow the airflow. But larger droplets were able to propagate further into the compressor to clean the stages downstream and to penetrate the boundary layer to reach the blade surface.

During the late 1980s, demineralized water, water-detergent mixtures, and water-petroleum solvent mixtures were the fluids applied for compressor washes and the chemical industry became active in the field. Thames et al. [6] investigated washing agents, procedures, and schemes based on a survey of gas turbine users, manufacturers, and cleaning agent suppliers. The issue of a possible redeposition of contaminants from the front to the rear of the compressor was acknowledged. It was considered to be only problematic for heavily fouled compressors and attributed to the vaporization of fluid with rising compressor temperature, as it was not often found in borescope inspections.

The need for cleaning fluids with higher boiling points, which penetrate the compressor further, and lower freezing points, so they could be used and stored in cold conditions, was addressed with a patent from Woodson et al. [39]. Furthermore, a patent of a universal cleaning solution for online and offline washing that was fully combustible without the formation of detrimental byproducts and biodegradable when included in waste water indicated the rising environmental concern [40]. Gas turbine users evaluated advantages and disadvantages of the cleaning solutions in field tests. The selection of the appropriate cleaning fluid depended on the matter of contamination, and therefore the investigation of quantity and type of compressor deposits were suggested; preheating of the cleaning fluid also proved beneficial [3]. Test facilities used fouled blades to investigate the cleaning efficiency of different washing solutions experimentally. A test of eight different fluids proved the environmentally sound ones less efficient than the solvent-based solutions. The newly introduced corrosion inhibitors worked satisfactorily [41]. The chemical development of the cleaning solutions is ongoing, and a fairly recent idea to rinse the engine with an anti-static liquid after on an offline wash to avoid further buildup of debris was patented by an aero-engine manufacturer [42].

1990s. During the 1990s, the risk of spray nozzles loosening due to vibration and corrosion was of rising concern. To avoid parts entering the compressor during operation, adaptations of the injection locations and safer nozzle fittings were introduced. Gen-

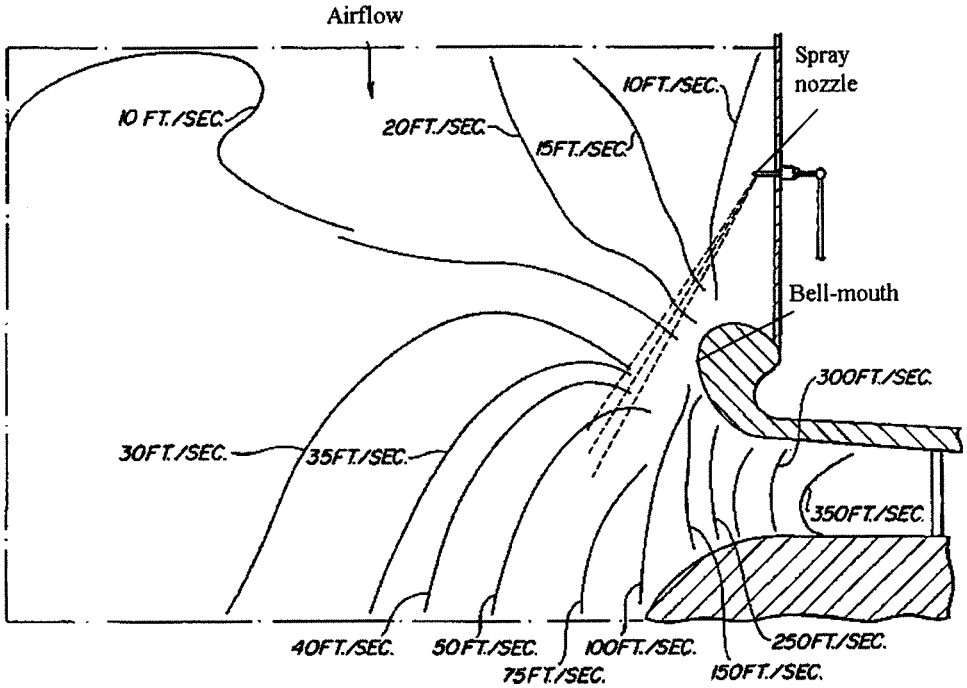


Fig. 4 Spray injection from the bell-mouth side into regions of low airflow velocity and high turbulence [37]

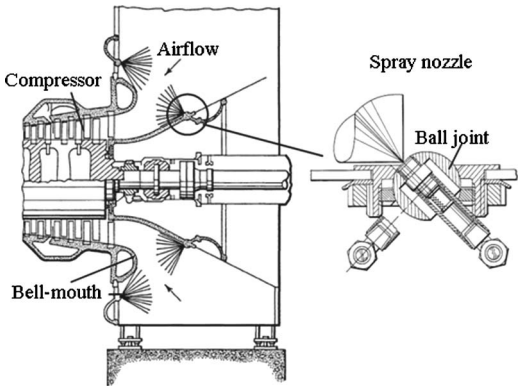


Fig. 5 Washing device and nozzle detail by Kolev and Robben [43], side view

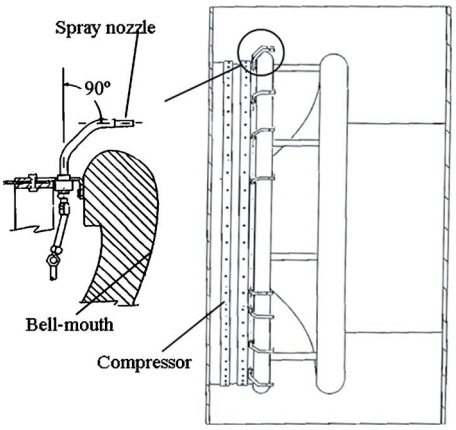


Fig. 6 Injection system with fittings from outside the bell mouth [21]

erally, it was perceived that the further away the spray nozzles were located from the compressor, the less likelihood of a loose nozzle entering the compressor. Therefore, installations opposite of the compressor were favorable. Kolev and Robben [43] patented a nozzle that used a ball joint (see Fig. 5 (right)) that was claimed to be more secure. As illustrated in Fig. 5 (left), the location of the spray nozzles was at a significant distance from the compressor inlet.

However, the main development issues were the optimization of the fluid distribution at the compressor inlet and also the quantitative investigations toward the appropriate droplet size. It was known that the impingement of large droplets from the washing spray could cause blade erosion, damage the coatings, or induce vibration. However, after a nozzle and wash fluid were fitted, the droplet size distribution was only adjustable via the operating pressure or temperature. By providing a spherical fitting for the cone spray nozzle, a readjustment of the injection direction in three dimensions, inducing a different dispersion through the airflow, was possible. As illustrated in Fig. 5, the protruding parts into the flow passage were kept at minimum and adjustments were possible even during engine operation. The patent was followed shortly thereafter by an evaluation of the washing system installed in a power plant in The Netherlands [44]. Additional measurements of the fuel supply and ambient conditions facilitated the detection of an average power recovery due to a single online wash of 1%. Substantial economic benefits were shown due to the efficient combination of online and offline washes.

An adaptation of the bell-mouth fitted system of McDermott [37] was found for an industrial application [21] where the nozzles were fitted from behind the bell mouth lip rather than inside the bell mouth to reduce the likelihood of ingesting a loose nozzle (see Fig. 6). This was also supported by the lower air velocities in this location. A further improved system was installed in a different power plant where the nozzles were fitted through the sidewall of the bell mouth with more nozzles fitted above the shaft cone than underneath. The authors stated that a better mixing of the solution and the airflow was achieved.

Up until the early 1990s, the issue of finding the right droplet size to enable the droplets to follow the airflow but avoid damage to the blades was addressed qualitatively, but not quantitatively. Naval applications started quantifying the droplet diameters in practice. Even though the statistical nature of droplet size distributions can be characterized by specific size parameters, such as the Sauter mean diameter, the data available to the authors lacked the statistical definition of the size distribution and referred to droplet diameter ranges. Depending on the gas turbine and the washing system, droplet sizes in the range of 125–200 μm and 80–200 μm were reported by Patterson and Spring [45]. A significantly higher mean droplet diameter of 800 μm was published for the online wash system of a different naval application [46]. In both cases, flat-fan spray nozzles were used at a relatively low-pressure level. The differences of droplet diameters may also be related to the different types of contaminants in the marine environment. Where salt deposits were reported to be removable with water washes, severe buildup of baked soot on the front stages, originating from other engine exhausts, were found to be only removable with offline washes [8].

It was just two years later that a washing method for aero engines was patented that claimed to produce droplet sizes that followed the airflow similarly to the trajectories of contaminants (see Fig. 7, left). A high operating pressure between 5 and 8 MPa was suggested to achieve droplet sizes between 120 and 250 μm and a high injection velocity of 100–126 m/s to overcome the centrifugal effect of the rotating blade rows. Depending on the existence of inlet guide vanes, different pressure levels, flow rates, and droplet sizes were suggested [22]. The author also described an adaptation of the system for industrial engines, which is depicted in Fig. 7 (right), and addressed common washing practice for aviation and marine applications [19]. Potential problems of the to-date cleaning concepts were summarized. The large quantities of liquid were attributed to causing blade stress. The waste of detergent and extensive preparation before the washes led to complex installations and procedures. To overcome this, the presented new cleaning method was claimed to use up to 90% less fluid. It was stated that particle size and speed were crucial factors to decide whether the droplet was peripherally centrifuged or was able to follow the airflow into the engine. As the droplet size also had a major influence on the impact force when impinging on the blades, a high operating pressure was suggested with a low flow.

From naval applications, it was reported that online washes were not effective in cleaning the rear stages [14]. The lack of cleaning, especially toward the rear of the compressor and the root of the rotor blades, was attributed to large droplets centrifuged outward and small droplets not being able to penetrate the bound-

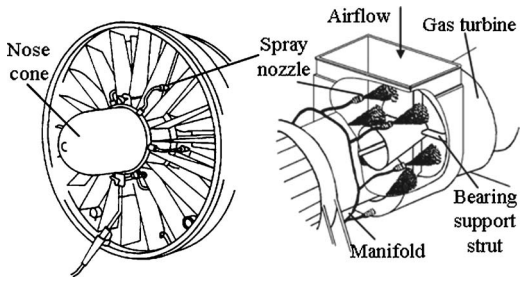


Fig. 7 Washing device for aero-engines [22] and adaptation for a stationary application [19]

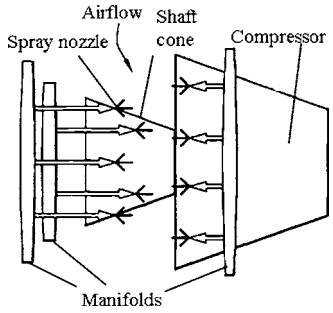


Fig. 8 Wet compression and compressor washing system located upstream of the compressor by Trewin and Carberg [49]

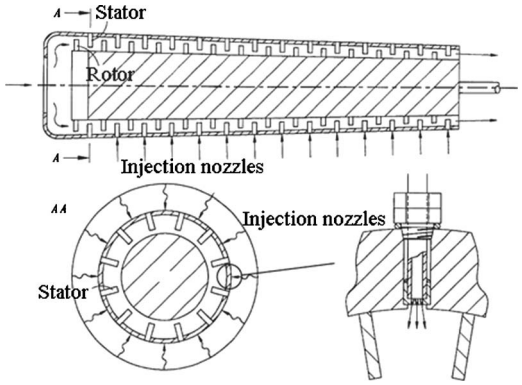


Fig. 9 Wet compression and compressor washing system with injection nozzles throughout the compressor by Ingistov [50]

ary layer of the blades [47]. A concept of different pressure levels to inject different droplet sizes for industrial gas turbines was suggested [47]. Providing initially small droplets in the range of 80–120 μm to clean the first stages, the system was claimed to adjust the right pressure to generate droplet sizes between 130 and 170 μm to clean the rear stages. The range of required droplet sizes was specified between 50 and 500 μm depending on the compressor and the cleaning fluid used.

Because of the large variety of engines, intake configurations, and washing systems, and the difficulties to undertake experiments, a generally applicable washing procedure, including the influencing design variables, has not yet been found. Finding the balance between efficient cleaning and avoiding erosion is still a learning process. General Electric faced erosion problems of the zero-stage rotor blades for the Frame 9FA class engine fleet and suggested a discontinuation of the online washing procedures after 100 h of online water wash [48].

Recent Development Trends. The latest published developments were focused on the flexible usage of the installations or the combined use of air and fluid to manipulate fluid distributions. An engine manufacturer followed up using different droplet sizes for different purposes and patented a double function system for compressor cleaning and power augmentation through wet compression in 2002. The conventional online cleaning arrangement was additionally connected to an atomizing air supply. For wet compression the air supply was switched on to generate smaller droplets required for evaporation, whereas for compressor washing the system was operated only with fluid [49]. As illustrated in Fig. 8, the system consisted of three nozzle manifolds, two located at different distances opposite of the compressor and one located at the bell-mouth side of the compressor. The washing system itself was operated at a single low-pressure level. Interestingly, a power generation company followed the combined use approach with a patent of a wet compression and cleaning system, where the spray nozzles were fitted between the stators within the blade rows [50]. This idea picked up on the early interstage injection by Mansson [34]. Figure 9 illustrates a meridional view of the gas path, a cross section through the first stationary blade row, and the nozzle detail of the system. Jeffs [51] also reported the development of a new type of air-assisted nozzles for a washing system for large gas turbines. To allow small droplets to penetrate the airflow, the spray is shielded by two high-velocity flat-profile air jets.

Further developments include the earlier mentioned injection of a second anti-static fluid to reduce the rate of fouling [42] and the suggestion of using sound waves of a fixed or variable frequency of $<$500 Hz to clean the internal surfaces [52].

Categories of Compressor Online Washing Systems

Based on historical developments, different features of the washing systems were observed. Nozzles were configured such that they either target the compressor inlet directly from close proximity spraying rather parallel to the airflow direction or they were located further upstream, spraying transversely to the airflow, not necessarily into the direction of the compressor. The systems were either operated at low pressures, below 1 MPa, or at significantly higher pressures around 5 MPa and above. The amount of liquid used per wash varied greatly and was thought to relate to the size or the airflow of the engine. In some cases, it was possible to derive the ratio of air-to-fluid mass flow via a given injection time. With the engine airflow derived from manufacturers' specifications, an initial dataset from open literature was compiled (see Table 2). No clear relationship of the air-fluid ratio or the amount of liquid to the engine size or airflow was found, and the need of a broader data set became evident. As the open literature provided insufficient data, in particular, for the airflow to fluid ratios, a survey was undertaken to fill this gap. Based on Internet sources, gas turbine users in Europe, gas turbine manufacturers, and washing system suppliers were contacted through e-mail, fax, or letter. The response provided the missing link to assign the systems presented in the historical perspective to qualitative system categories using the operating pressure and the fluid rate of the system. As illustrated in Fig. 10, the current systems were classified into categories ranging from lower (LP) to high levels of operating pressures (HP) and from low and medium to high air-fluid ratio (LF, MF, HF). For simplicity, the categories were abbreviated as, for example, high-pressure low-flow system (HPLF).

It was observed that for LF systems, the nozzles were generally installed within the inlet plenum at significant distance to the com-

Table 2 Washing parameters for different engines available in open literature

Engine	Power (MW)	Airflow (kg/s)	Liquid per wash (l)	Fluid rate (l/s)	Air-fluid ratio	Operating pressure (MPa)
TF40B	13	80	11 [46]	0.091	135	—
Avon	14	75	265 [5]	0.482	162	0.2
LM2500	25	68	76 [6]	0.367	189	—
RB211-24G	29.5	92	90–180 [53]	0.500	184	7
GT 11 D5	66	200 (est.)	172 [44]	0.190	1053	—
V94.3A	260	620	300 [54]	0.833	787	4–6

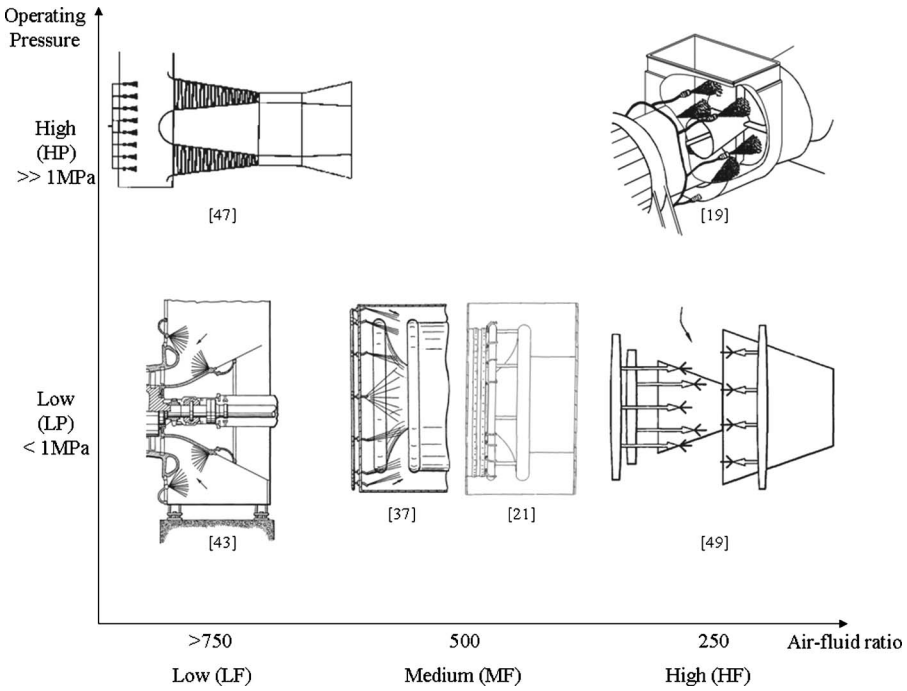


Fig. 10 Qualitative categories of online washing systems

pressor, whereas for HF systems, the injectors were located in close proximity to the compressor inlet. HP systems generally used only one ring of nozzles installed either opposite of the compressor or at the bell mouth. With the exception of MF systems, LP systems used several nozzle rings, combining the injection from opposite with the injection from the bell mouth. These principals are sketched in Fig. 11.

Indications were found that the systems using high-fluid injection rates (HF) use a smaller number of nozzles than the systems using a lower amount of fluid. Asplund [19] stated that the advantage of the high operating pressure was that a lower number of nozzles was required, reducing installation and maintenance costs. References were found for smaller engines up to 60 MW output where less than nine nozzles were used for the online installation [19,53]. The number of nozzles for offline equipment is generally defined by the area to be wetted, usually the area between two support struts [29], and may range from five to nine. Asplunds HPHF nozzle configuration, therefore, has strong parallels to offline systems. The LPHF system suggested by Trewin and Carberg [49] is a dual-function system, where the nozzles are used for online washing and fogging, which may indicate a design compromise. Several sites using the current online system of the patent applicant reported around nine online nozzles for gas turbines in the power range of 120–260 MW. However, retrofitted systems were also found using twice the number of nozzles. Significantly higher numbers of nozzles were found for LF systems. The HPLF system by Hayward et al. [47] defined the number of nozzles required via the complete coverage of the first stage. For current systems, this was generally found to be in excess of three times the number of struts. For the LP systems, a similar trend was observed. Jeffs [55] reported a ratio of 1:4 for the number of offline and online nozzles for the LPLF system by Kolev and Robben [43]. Stalder [29] and Leusden et al. [7] added that a relatively high number of nozzles were required to ensure a better fluid distribution.

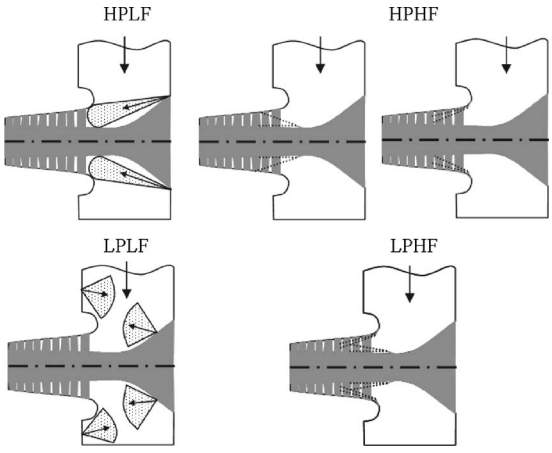


Fig. 11 Principal nozzle locations for basic online washing system categories

The air-fluid ratio is based on the gas turbine inlet mass flow and therefore related to a specific engine. To prove the suggested categories, each system had to be investigated over a range of engine mass flows covering the most relevant engine types. Open literature often only provided ranges of fluid flow rates that applied for a range of engine sizes, and information from gas turbine manufacturers and washing system suppliers was of crucial importance. But even fluid flow rates were restricted information, and thus, only a few data sets were made available. Because of the confidentiality concerns, the inquiries to the gas turbine users were reduced to the details on mass flow of the washing fluid, operating pressure, and number and arrangement of nozzles. Even though the investigated data set can only represent a small picture of the broad range of present systems, it allowed a first quantitative analysis of the current systems.

It was found that the best way to present the results was the air-fluid ratio versus the power output of the gas turbine (simple cycle) where a higher air-fluid ratio indicated a lower fluid injection rate per air mass flow through the engine. The results are illustrated in Fig. 12. Each marker represents the air-fluid ratio of a washing system for a certain gas turbine. Recommendations of suppliers and engine manufacturers are shown using large markers. Feedback from gas turbine users (labeled with "GTU") using the equivalent system or literature referencing the equivalent sys-

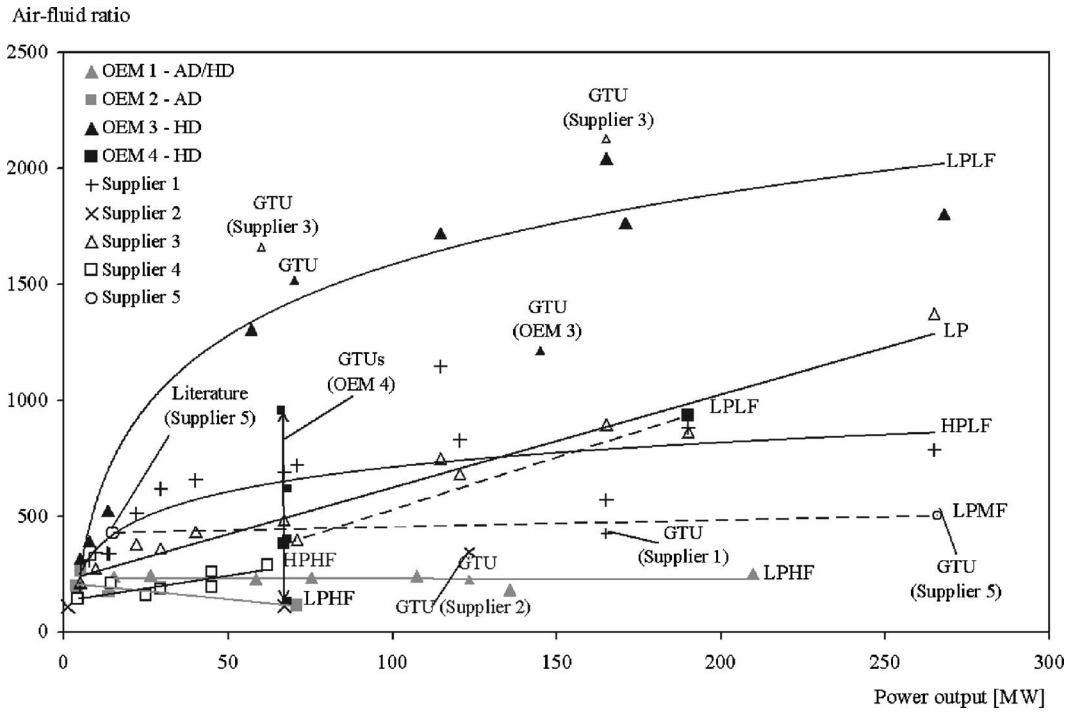


Fig. 12 Air-fluid ratio versus engine power output for specific state-of-the-art washing systems based on field data and washing recommendations

tem are using the same but smaller markers. To visualize the behavior of the injected fluid rate over the broad engine range for a specific washing system, trend lines were produced, including manufacturers' recommendations and feedback from gas turbine users. According to the system categories discussed before, the trend lines were also labeled with the suggested abbreviations.

Despite a high scatter and limited data, different trends were observed for the air-fluid ratio versus the engine power output:

- constant air-fluid ratio over the whole engine range
- linear scaling of air-fluid ratio with increasing engine power output
- exponential relationship of air-fluid ratio and engine power output

These trends only become apparent for larger engine sizes; as for the range of small engines with a power output below 30 MW, the air-fluid ratio does not differentiate.

A constant air-fluid ratio was recommended in particular, by aero-derivative (AD) engine manufacturers (OEM 1 and OEM 2) who also recommended the lowest air to fluid ratio. This may originate from washing regimes for aero engines using large amounts of water [19,22]. Both trend lines represent LPHF systems. For OEM 1, it should be noted that the data from the higher end of the power output refer to engines from their heavy-duty (HD) range. For the LP system suggested by Haub and Hauhe [21], only one additional response was received from a gas turbine user. The system was used with almost the same medium flow rate for a small and heavy-duty gas turbine. Even though only two data sets were available, the large difference in engine size indicated a constant air-fluid ratio over the whole engine range for the system (LPMF, supplier 5).

Several systems showed a relatively linear increase of air-fluid ratio for increasing engine power output. This was found for the high-pressure system presented by Asplund [19] for gas turbines of smaller power output (supplier 4). Despite the vendor's philosophy of using high pressures and low amounts of fluid, the system falls into the HPHF category when compared to the other systems on the market. Depending on the gradient, a linear behavior made the assignment to a HF or LF system more difficult because, for smaller engine sizes the air-fluid ratio was low, and for large engine sizes the air-fluid ratio was high. This was, in particular, found for recommendations of system supplier 3. It was also observed that on two occasions the system was used by gas turbine users at significantly different air-fluid ratios (see GTU (supplier 3)). In these cases, the washing recommendations of the gas turbine manufacturer were more restrictive in terms of injected fluid rate. This can be found when retrofits were installed and washing recommendations of the system supplier and the gas turbine manufacturer were not in the same range. Certainly guarantee limits for the gas turbine have to be appreciated and may result in a compromised washing scheme. The recommendations of a recently developed system by an engine manufacturer (OEM 4) also showed a linear increase of air-fluid ratio with rising engine size, even though only two engine examples were available. For this system, several gas turbine users gave feedback for the smaller engine example and a wide spread of injected fluid rate was reported, well away from the engine manufacturers recommendations (GTUs of OEM 4). In most cases, the injection rate was lower than recommended injection rate by the OEM. Only in one case was a higher injection rate reported. The user justified this by shortening the washing procedure. This LP system is claimed to use a high number of nozzles to create a homogeneous field of droplets [7] and can be classified as a LPLF system for larger gas turbines.

Two washing systems showed an exponential trend of the air-fluid ratio. The recommendations from OEM 3 combined with user feedback and literature [44] resulted in a relatively steep exponential behavior (LPLF). For larger turbines, the amount of injected water was low and the associated washing system supplier found that droplets were deflected by the strong airflows. This was recently improved by the development of air-shielded nozzles, where two flat air shields protect the fluid spray cone and allow the droplets to penetrate further [51]. An exponential relationship was also found for a high-pressure system (HPLF) from system supplier 1 [47]. The level of air-fluid ratio was smaller, but the system still uses significantly less fluid per airflow than MF or

Table 3 Main influencing parameters on washing efficiency

Air flow	Droplet
Inlet mass flow	Droplet size
	Fluid density
Velocity flow field	Injection velocity (magnitude, direction, and location)

HF systems. Interestingly, a gas turbine user reported the application of a higher injection rate than recommended by the system supplier or the gas turbine manufacturer (GTU supplier 1). Recommendations of supplier 2 using a LP system were defined as a function of the gas turbine mass flow. A decreasing fluid rate for air mass flows up to 400 kg/s was followed by a constant rate for larger engines [10]. This resulted in an exponential function of the air-fluid rate when applied to some example engines. However, the slope was relatively low, and for engines with a higher power output than 50 MW, the air-fluid ratio remains constant. Even though one data set was received from a gas turbine user for a medium-size engine of 123 MW output injecting less than recommended fluid (GTU (supplier 2)), the system used a relatively high fluid rate over the whole engine range and falls into the LPHF category.

Generally, it has to be stated that the data scatter was relatively high and the large variety of engines contributed additionally. Different makes have different airflow-to-power output ratios, and variations in air fluid ratio occur using the same washing system for different engine makes of the same power output.

Main Design Variables of Compressor Washing Systems

In this section, the information collected on the state-of-the-art online washing systems is broken down to discuss the main requirements and design parameters that affect the efficiency of the washing system. Several authors have suggested requirements for effective compressor washing; however, the criteria were qualitatively and relatively broad. Becker and Bohn [4] and Harris and Calabrese [46] claimed that a complete wetting of the first blade row from hub to tip is essential. McDermott [37] suggested a continuous covering of cleaning fluid over the whole surface area in front of the engine for effective cleaning results, and Hayward et al. [47] recommended the installation of "sufficient nozzles to give a 360° coverage of the first stage blading." A practical perspective was recently given by Lambart et al. [28] as "the ideal spray pattern for online compressor washing covers the compressor inlet plane completely without overlapping onto the plenum wall, bell mouth or shaft cone to avoid drainage and waste of fluid." These guidelines aim to clean the first blade row from hub to tip most effectively. In power generation applications, these blades are generally stationary (the inlet guide vanes). The centrifugal effect within the following rotating blade rows forces droplets toward the blade tip, and it was found that progressing downstream the blades were left untreated toward the shaft of the engine [47]. A compromised approach to reduce this problem was reported by a gas turbine user, targeting the spray toward one-third of the distance from the shaft to the casing instead of the midspan area [56] to leave a greater spray margin close to the shaft. To achieve a complete coverage of the first stage, the spray droplets need to penetrate the accelerating airflow through the intake bend. The droplet trajectories are based on the force balance between the airflow and the droplets, and the resulting main influencing parameters for compressor washing efficiency are summarized in Table 3.

In general, a washing system is designed for a specific operating condition. For relative changes of air flow or fluid flow to the design condition, the force balance is affected and the droplet trajectories change. If the airflow or air velocity is reduced, then droplets can more easily penetrate the airflow. If the droplet size or density is reduced, then the droplets are stronger deflected by the airflow. The inlet total mass flow of the gas turbine depends on the engine size in the first place, but also on the ambient temperature, the installation altitude, the power setting, and the condition in terms of degradation of the engine. Kacprzynski et al. [57] presented field data of a naval application that indicated a better efficiency recovery for water washing at higher power settings. For the investigated twin-spool engine, the inlet mass flow depended directly on the power setting. The result indicated that the washing system was designed for application at high power settings. Even though two extremely different operating conditions were addressed, Leusden et al. [7] also stated that different spray nozzles were used for online and offline cleaning due to the different operating conditions (and subsequently inlet mass flows) of the gas turbine. As the velocity of the airflow within the intake duct varies in magnitude and direction, areas of low airflow velocity are easier to penetrate by droplets than areas of high airflow velocities. The velocity vectors of the airflow may work favorably or adversely for the droplet trajectory.

The main contributing factors for the droplet force are the droplet mass and injection velocity. Smaller droplets tend to follow the flow, and larger droplets behave in a more ballistic manner. This implies better traveling capabilities for larger droplets, but the higher impact force may put the engine hardware at risk and a sensible compromise has to be found. With higher densities, droplets become heavier for the same droplet size. However, the range of density changes for compressor washing fluids stays within certain boundaries. Compared to demineralized water, cleaning solutions with 10% higher density (at 20°C) were found [40] and for solvent solutions lower densities of up to 5% were reported [39,46]. A higher droplet velocity allows a better penetration of the airflow. However, the injection system only provides in initial injection force for the droplet and the following trajectory is strongly influenced by the velocity field of the airflow.

As discussed earlier, very different concepts exist of how to inject the spray. The injection nozzles may directly target the compressor or, if a mixing of the spray with the airflow is desired, the nozzles may spray into the inlet plenum just upstream of the converging annulus. The injection direction can vary from being parallel to the airflow to almost transverse direction. The transverse direction is mainly employed for increased spray dispersion [37] as the breakup of a droplet occurs above a certain shear stress level between the droplet and the airflow. Spray patterns involve a certain range of injection directions. Generally, flat-fan-shaped spray patterns were found [28,37,38,45,46], but also cone-shaped spray patterns were used [43]. Increasing surrounding airflow results in a reduction of the spray angle, and a spray angle of 90 deg injected into still air may reduce to 60 deg when injected into the engine at full load conditions [37]. Given these implications, the main scope of the injection system is the transfer of droplets through the intake system to the compressor inlet. Once entered, the compressor, droplet size, and speed decide if the droplet is centrifuged toward the casing and also determine the impact force when a droplet impinges onto a blade [19] or if a droplet is deflected by the boundary layer [28]. Therefore, the injection system has to generate the appropriate droplet size and velocity spectrum to assure the transfer of droplets toward the compressor inlet face.

Within the compressor, characteristics of the fluid become of greater importance. Scheper et al. [2] stated that "all blading surfaces must become thoroughly wet" and that "a sufficient flow rate" has to be used "to counteract the considerable drying effect generated by the rotor rotation." Even though these early requirements were addressing offline washing procedures, the most important characteristics of the washing solution for online washing were known as temperature sustainability, and wetting efficiency.

For online washing it is vital that the cleaning fluid stays in solution. To prevent evaporation due to the rising temperatures within the compressor, with exit temperatures being well above water evaporation temperature [55], the evaporation point of the cleaning fluid is increased by chemical treatment. Modern clean-

ing agents contain high-temperature carriers retaining the cleaning fluid in solution [28], and liquids are on the market that withstand up to 400°C [19]. However, in borescope and visual inspections it was found that droplets may traverse up to the fifth or sixth stage and then vaporize, and the residue was centrifuged along the compressor casing [28,30]. For a large single-shaft gas turbine, it was stated that all fluid was evaporated by stage 8 [50] and poor penetration into the latter stages was also observed for naval applications [45]. The mechanism of compressor cleaning relies on an efficient wetting of the blade surface to take advantage of surfactants, and dispersing organic and inorganic soils. A splashing of droplets over the surface is not desirable, as a certain contact time of the fluid on the blade surface has to be allowed. The wetting speed depends on the surface tension, drop dimensions, velocity, and viscosity of the washing agent [3]. Modern washing solutions appreciate this effect with ingredients to reduce the surface tension of the droplets and therefore increase the wetting speed [28]. Above a certain level of liquid film on the blades, droplets and debris wash off and power recovers [29]. Because of an improved ability to remove oily debris, the water-detergent mixture may be preheated up to the boiling point before being injected into the compressor. But a side-effect of this procedure is that the margin until evaporation point reduces significantly [3].

References

[1] Diakunchak, I. S., 1992, "Performance Deterioration in Industrial Gas Turbines," ASME J. Eng. Gas Turbines Power, **114**, pp. 161–168.

[2] Scheper, G. W., Mayoral, A. J., and Hipp, E. J., 1978, "Maintaining Gas Turbine Compressors for High Efficiency," Power Eng., **82**(8), pp. 54–57.

[3] Mezheritsky, A. D., and Sudarev, A. V., 1990, "The Mechanism of Fouling and the Cleaning Technique in Application to Flow Parts of the Power Generation Plant Compressors," ASME Paper No. 90-GT-103.

[4] Becker, B., and Bohn, D., 1984, "Operating Experience With Compressors of Large Heavy-Duty Gas Turbines," ASME Paper No. 84-GT-133.

[5] Bagshaw, K. W., 1974, "Maintaining Cleanliness in Axial Compressors," *Gas Turbine Operation and Maintenance Symposium 1974*, National Research Council (eds.), Toronto, Canada, pp. 247–264.

[6] Thames, J. M., Stegmaier, J. W., and Ford, J. J., Jr., 1989, "On-Line Compressor Washing Practices and Benefits," ASME Paper No. 89-GT-91.

[7] Leusden, C. P., Sorgenfrey, C., and Duemmel, L., 2004, "Performance Benefits Using Siemens Advanced Compressor Cleaning System," ASME J. Eng. Gas Turbines Power, **126**, pp. 763–769.

[8] Abdelrazik, A., and Cheney, P., 1991, "Compressor Cleaning Effectiveness for Marine Gas Turbines," ASME Paper No. 91-GT-11.

[9] Butler, J. J., 2002, "Inside Out Gas Turbine Compressor Cleaning Method," U.S. Patent No. 6,394,108 B1.

[10] Faddegon, C. J., 1999, "Effective Compressor Cleaning as a Result of Scientific Testing and Field Experience" (private communication).

[11] Kulle, V., 1974, "Gas Turbine Blade Cleaning at Alberta Gas Trunk Line (AGTL)," *Gas Turbine Operation and Maintenance Symposium 1974*, National Research Council (eds.), Toronto, Canada, pp. 265–271.

[12] Brittain, D., 1983, "Cleaning Gas Turbine Compressors," Aircr. Eng., **55**(1), pp. 15–17.

[13] Langford, J. R. F., 1977, "Contamination Removal Method," U.S. Patent No. 4,065,322.

[14] Millsaps, K. T., Baker, J., and Patterson, J. S., 2004, "Detection and Localization of Fouling in a Gas Turbine Compressor From Aerodynamic Measurements," ASME Paper No. GT2004–54173.

[15] Boyce, M. P., Bowman, J. C., Meher-Homji, C. B., and Focke, A. B., 1985, "Energy Efficient Operation of Gas Turbine Compressor Sets," ASME Paper No. 85-GT-70.

[16] Fielder, J., 2003, "Evaluation of Zero Compressor Wash Routine in RN Service," ASME Paper No. GT2003–38887.

[17] Testman, F. L., Jr., 1999, "Engine Wash Recovery System," U.S. Patent No. 5,899,217.

[18] Anon., 1999, *LM2500 Military Marine Gas Turbine Installation Design Manual*, GE Marine Engines, Cincinnati.

[19] Asplund, P., 1997, "Gas Turbine Cleaning Upgrade (Compressor Wash)," ASME Paper No. 97-AA-135.

[20] Adams, J., and Schmitt-Wittrock, P., 1981, "Optimierung der Reinigungsintervalle von Gasturbinenverdichtern," Brennst.-Waerme-Kraft, **33**(1), pp. 25–31.

[21] Haub, G. L., and Hauhe, W. E., Jr., 1990, "Field Evaluation of On-Line Compressor Cleaning in Heavy Duty Industrial Gas Turbines," ASME Paper No. 90-GT-107.

[22] Asplund, P., 1996, "A Method of Washing Objects, Such as Turbine Compressors," World Intellectual Property Organization, Patent Cooperation Treaty, WO96/40453.

[23] Chum, K. Y., 2002, "Effectiveness of Online and Offline Compressor Water Wash," *9FA Users' Conference*, Nov. 11.

[24] Stalder, J.-P., 2004, Comment to the Paper Presentation: Engdar, U., Orbay, R., Genrup, M., and Klingmann, J., 2004, "Investigation of the Two-Phase Flow Field of the GTX100 Compressor Inlet During Off-Line Washing," ASME Paper No. GT2004–53141.

[25] Engdar, U., Orbay, R., Genrup, M., and Klingmann, J., 2004, "Investigation of the Two-Phase Flow Field of the GTX100 Compressor Inlet During Off-Line Washing," ASME Paper No. GT2004–53141.

[26] Margolis, H., 1991, "U.S. Navy On-Line Compressor Washing of Marine Gas Turbines," ASME Paper No. 91-GT-309.

[27] VGB, 1991, "Reinigungsverfahren für Verdichter-und Turbinenbeschaufelungen an Gasturbinen im offenen Prozess," *VGB Technische Vereinigung der Grosskraftwerk-betreiber e.V.*, VGB Merkblatt 106, 2nd Edition, VGB-Kraftwerkstechnik GmbH, Essen, Germany, pp. 10–11.

[28] Lambart, P., Gordon, R., and Burnett, M., 2003, "Developments in On-Line Gas Turbine Compressor Cleaning," *Institution of Diesel and Gas Turbine Engineers, 2nd Gas Turbine Conference 2003*, Milton Keynes, UK, pp. 136–142.

[29] Stalder, J.-P., 2001, "Gas Turbine Compressor Washing State of the Art: Field Experiences," ASME J. Eng. Gas Turbines Power, **123**, pp. 363–370.

[30] Mast, M., and Bohrenkämper, G., 2001, "Modernisation and Upgrading Products for Siemens V94.2 Gas Turbines," *10th Meeting of V94.2 Gas Turbine Users*, May 7–8, Rosignano Solvay, Italy.

[31] Saravanamuttoo, H. I. H., 1974, "Comments or Questions and Speakers' Replies," *Gas Turbine Operation and Maintenance Symposium 1974*, National Research Council (eds.), Toronto, Canada, p. 305.

[32] Meher-Homji, C. B., and Bromley, A., 2004, "Gas Turbine Axial Compressor Fouling and Washing," *33rd Turbomachinery Symposium*, Texas A&M University, pp. 163–191.

[33] Freid, W. B., and Tapparo, D. J., 1971, "Wash Manifold," U.S. Patent No. 3,623,668.

[34] Mansson, M., 1977, "Washing Apparatus for a Compound Compressor," U.S. Patent No. 4,046,155.

[35] Elser, W., 1973, "Betriebserfahrungen beim Reinigen der Verdichter von Rolls-Royce Strahltriebwerken," Brennst.-Waerme-Kraft, **25**(9), pp. 347–348.

[36] Elser, W., 1975, "Betriebserfahrungen beim Reinigen der Verdichter von Rolls-Royce-Strahltriebwerken," Brennst.-Waerme-Kraft, **27**(1), pp. 19–21.

[37] McDermott, P., 1991, "Method and Apparatus for Cleaning a Gas Turbine Engine," U.S. Patent No. 5,011,540.

[38] Hornak, S. S., Lowdermilk, R. S., and Miller, R. A., 1980, "Removable Wash Spray Apparatus for Gas Turbine Engine," U.S. Patent 4,196,020.

[39] Woodson, J. B., Cooper, L. A., White, H. M., and Fischer, G. C., 1989, "Cleaning Gas Turbine Compressors," U.S. Patent 4,808,235.

[40] Kaes, G., 1991, "Method of a Cleaning Agent for the Cleaning of Compressors Especially Gas Turbines," U.S. Patent No. 5,076,855.

[41] Kolkman, H. J., 1992, "Performance of Gas Turbine Compressor Cleaners," ASME Paper No. 92-GT–360.

[42] Ackerman, J. F., Stowell, W. R., and Johnson, R. A., 2003, "Methods and Apparatus for Washing Gas Turbine Engines," U.S. Patent No. 6,630,198B2.

[43] Kolev, S., and Robben, R., 1993, "Injection Device for the On-Line Wet Cleaning of Compressors," U.S. Patent No. 5,193,976.

[44] Stalder, J.-P., and van Oosten, P., 1994, "Compressor Washing Maintains Plant Performance and Reduces Cost of Energy Production," ASME Paper No. 94-GT–436.

[45] Patterson, J. S., and Spring, S. K., 1992, "On-Line Detergent Washing: Reducing the Environmental Effects on the LCAC Gas Turbine Engines," ASME Paper No. 92-GT-269.

[46] Harris, H., and Calabrese, M., 1994, "On-Line Detergent Fluid Evaluation on a TF40B Gas Turbine Engine," ASME Paper No. 94-GT-452.

[47] Hayward, J., Winson, G., and Raatrae, A., 1999, "Wash System for Gas Turbine Compressors," European Patent Application No. EP0933502A2.

[48] GE Systems, 2001, "Compressor Rotor R-0 Blade Erosion," Technical Information Letter, TIL 1303–1R1, GE Energy Services.

[49] Trewin, R. R., and Carberg, W. G., 2002, "Combined System for Cooling by Water Injection and Washing of a Gas Turbine Compressor," European Patent Application No. EP1205640A2.

[50] Ingistov, S., 2002, "Method an Apparatus for Increasing the Efficiency of a Multi-Stage Compressor," U.S. Patent No. 6,398,518B1.

[51] Jeffs, E., 2003, "Innovative On-Line Wash Nozzles for Large Turbines," Turbomach. Int., **44**(3), reprint.

[52] Fetescu, M., 1998, "Method and Apparatus for Keeping Clean and/or Cleaning Gas Turbine Using Externally Generated Sound," U.S. Patent No. 5,758,486.

[53] Syverud, E., Bakken, L. E., Langnes, K., and Bjørnås, F., 2003, "Gas Turbine Operation Offshore; On-Line Compressor Wash at Peak Load," ASME Paper No. GT-2003–38071.

[54] Harland, S., 2002, Didcot B power plant visit (private communication).

[55] Jeffs, E., 1992, "Compressor Washing on Line for Large Gas Turbines," Turbomach. Int., **33**(5), pp. 49–51.

[56] Fairhurst, W., 2003 (private communication).

[57] Kacprzynski, G. J., Caguiat, D. E., Gumina, M., Galie, T. R., Roemer, M. J., and McGroarty, J. J., 2001, "A Prognostic Modelling Approach for Predicting Recurring Maintenance for Shipboard Propulsion Systems," ASME Paper No. 2001-GT-0218.

Properties, Weldability, and Applications of Modern Wrought Heat-Resistant Alloys for Aerospace and Power Generation Industries

M. D. Rowe

V. R. Ishwar

D. L. Klarstrom

Haynes International, Inc.,
1020 W. Park Avenue,
Kokomo, IN 46904-9013

Alloy selection and alloy design both require consideration of an array of material attributes, including in-service properties, weldability, and fabricability. Critical properties of modern heat-resistant alloys for gas turbine applications include high-temperature strength, thermal stability, oxidation resistance, and fatigue resistance. In this paper, the properties of 12 solid-solution-strengthened and six age-hardenable heat-resistant alloys are compared. Weldability is an important attribute and can be a major limiting factor in the use of certain alloys. Weldability test methods are discussed, and the resistance of alloys to solidification cracking and strain-age cracking is compared. The use of weldability testing in the development of modern heat-resistant alloys is discussed with several examples cited. Finally, alloy selection for gas turbine components is outlined, taking into account both alloy properties and fabricability. [DOI: 10.1115/1.2056527]

Introduction

High-performance heat-resistant alloys are critical materials of construction in modern gas turbines and other components in the power-generation and aerospace industries. The chemical compositions of selected alloys of importance are provided in Table 1. Modern alloys are usually developed with a certain key application in mind and must possess an array of favorable properties to be selected for a given application. Alloy development is a process of compromise. For example, addition of niobium to an alloy may favorably increase strength while harming oxidation resistance and weldability. Successful alloy development involves optimizing certain key properties for the selected application while minimizing any negative side effects. Alloys that become accepted in a wide range of applications are those that excel in more than one property, such as strength and oxidation resistance, while still offering good weldability, fabricability, and the capability to be manufactured at a reasonable cost. This paper discusses the properties of modern heat-resistant alloys, then covers methods of ranking the weldability and fabricability of alloys. Finally, alloy selection for gas turbine applications is discussed, taking into account both service capability and fabricability of the alloys.

Properties of Wrought Heat-Resistant Alloys

A variety of cast and wrought nickel-based, cobalt-based, and iron-based heat-resistant alloys are used in flying and industrial gas turbine components. Although turbine blades, vanes, etc., use significant quantities of cast alloys, hot gas path components, such as combustors, transitions, flame holders, etc., predominantly use wrought superalloys. In this paper, only the properties, weldability, and applications of wrought superalloys for gas turbine applications are discussed.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. GT2004-54309.

Because of the demanding combination of high temperatures and a combustion environment in gas turbine engines, materials used in these applications need to have suitable high-temperature mechanical properties and environmental resistance characteristics. Some of the important high-temperature properties commonly required for the alloys used in these applications include high-temperature creep strength, thermal fatigue resistance, and thermal stability. The environmental resistance properties include oxidation, nitridation, and in certain cases, sulfidation or carburization. Table 2 summarizes the effects of the common alloying elements in heat-resistant alloys. The combination of these elements defines the resultant mechanical properties at elevated temperatures, the environmental resistance capabilities, and fabricability of the alloy. These resultant properties, in turn, determine the applications for which an alloy is suitable.

The heat-resistant alloys can be divided into two general categories: solid-solution-strengthened alloys and age-hardenable alloys. Solid-solution strengthened alloys use elements such as molybdenum or tungsten in solid solution to provide strengthening. In addition, they use the principle of in-service carbide precipitation, usually in the form of chromium-rich $M_{23}C_6$ carbides, to enhance their resistance to high-temperature creep. They cannot be hardened using aging heat treatments. Age-hardenable alloys use precipitation hardening as the mechanism for strengthening.

Solid-Solution-Strengthened Alloys. One of the early cobalt-based superalloys developed for gas turbine applications was HAYNES® 25 alloy, also known as L-605 alloy. This alloy possesses good oxidation resistance and mechanical strength at temperatures up to about 980 °C. However, because of concerns over the availability of the strategic element cobalt in the 1940s, iron-based superalloys, such as MULTIMET® alloy, were developed. This alloy, also known as N-155 alloy, was easier to fabricate than alloy 25, but it had properties close to alloy 25 in terms of high-temperature strength and oxidation resistance. MULTIMET alloy is used in afterburner parts, such as tail pipes, tail cones, etc, in aircraft engines. A modern day improvement over MULTIMET alloy is HAYNES 556 alloy. It has significantly better oxidation resistance than MULTIMET alloy in combustion gas environments.

Table 1 Nominal chemical compositions of heat-resistant alloys

Solid-solution-strengthened heat-resistant alloys (wt %)

Alloy	UNS No.	Ni	Co	Fe	Cr	Mo	W	Mn	Si	Al	Ti	C	B	Others
W	N10004	63a	2.5[a]	6	5	24	1[a]	1k	1k	–	–	0.12[a]	–	V-0.6[a]
X	N06002	47a	1.5	18	22	9	0.6	1k	1k	0.5k	0.15[a]	0.10	0.008[a]	Nb-0.5[a]
HR-160®	N12160	37a	29	2[a]	28	1[a]	1[a]	0.5	2.75	0.4k	0.5	0.05	–	Nb-1[a], Cu-0.5[a]
230®	N06230	57a	5[a]	3[a]	22	2	14	0.5	0.4	0.3	0.1[a]	0.10	0.015k	La-0.02, Cu-0.5[a]
617	N06617	54a	12.5	1	22	9	–	–	–	1.2	0.3	0.07	–	–
625	N06625	62a	1[a]	5k	21	9	–	0.5k	0.5k	0.4k	0.4k	0.10k	–	Nb+Ta-3.7, Cu-0.5[a]
25	R30605	10	51a	3k	20	1[a]	15	1.5	0.4k	–	–	0.10	–	–
188	R30188	22	39a	3[a]	22	–	14	1.25k	0.35	–	–	0.10	0.015k	La-0.03
MULTIMET®	R30155	20	20	30a	21	3	2.5	1.5	1k	–	–	0.12	–	Nb+Ta-1.0, N-0.15
556™	R30556	20	18	31a	22	3	2.5	1	0.4	0.2	–	0.1	0.02[a]	Ta-0.6, Zr-0.02, La-0.02, N-0.20
HR-120®	N08120	37	3[a]	33a	25	1[a]	0.5[a]	0.7	0.6	0.1	0.2k	0.05	0.004	Nb-0.7, B-0.004, N-0.20, Cu-0.5[a]
214™	N07214	75a	–	3	16	0.5[a]	0.5[a]	0.5k	0.2[a]	4.5	0.5[a]	0.04	0.01k	Zr-0.1k, Y-0.01

Age-hardenable heat-resistant alloys

Alloy	UNS No.	Ni	Co	Fe	Cr	Mo	W	Mn	Si	Al	Ti	C	B	Others
Waspaloy	N07001	58a	13.5	2[a]	19	4.3	–	0.1[a]	0.15[a]	1.5	3	0.08	0.006	Zr-0.05
R-41	N07041	52a	11	5k	19	10	–	0.1[a]	0.5[a]	1.5	3.1	0.09	0.006	Zr-0.07[a]
263	N07263	52a	20	0.7[a]	20	6	–	0.4	0.2	0.6k	2.4k	0.06	0.005[a]	Zr-0.04[a]
242™	N10242	65a	2.5[a]	2k	8	25	–	0.8k	0.8k	0.5k	–	0.03[a]	0.006k	Cu-0.5[a]
718	N07718	52a	1[a]	19	18	3	–	0.35k	0.35[a]	0.5	0.9	0.05	0.004	Nb+Ta-5.0
X-750	N07750	70b	1[a]	8	16	–	–	0.35k	0.35k	0.8	2.5	0.08k	–	Nb+Ta-1.0, Cu-0.5[a]

[a]maximum (not intentionally added), k maximum (intentionally added), a balance, b minimum HASTELLOY®, HAYNES®, HR-120®, HR-160®, 230®, 556™, MULTIMET®, 214™, and 242™ are trademarks of Haynes International, Inc.

Because of increasing requirements to reduce the dependence on strategic elements, a nickel-based superalloy, HASTELLOY® X alloy, was developed in the 1950s. This alloy was initially used as a combustor material in the Pratt and Whitney Aircraft's JT3D engine for the Boeing 707 airplane. The applications for this alloy steadily increased over the years, and even today alloy X is used extensively in both flying and industrial gas turbine engines.

With further demands for higher-temperature capabilities, particularly for military engines, and less emphasis on concerns regarding strategic elements, such as cobalt, HAYNES 188 alloy was developed in the late 1960s. This cobalt-based superalloy is superior to 25 alloy in terms of both oxidation resistance and thermal stability. It is used in the engines for the F-14, F-15, and F-16 fighter aircraft and in a number of commercial aircraft engines.

During the cobalt crisis in the late 1970s and early 1980s, the need for lower cobalt-containing superalloys again became important. This led to the development of alloys such as 617 and HAYNES 230 alloys. Although these alloys have found extensive use in both flying and industrial gas turbines, as will be discussed later in this paper, 230 alloy is superior to 617 alloy in terms of high-temperature environmental resistance (oxidation and nitridation resistance), as well as thermal fatigue and low-cycle fatigue resistance.

Table 2 Elemental effects in heat-resistant alloys

Element	Effect
Nickel	High temperature strength, stability
Cobalt	High temperature strength
Chromium	Oxidation resistance
Molybdenum	High temperature strength
Tungsten	High temperature strength
Iron	Reduces cost
Silicon	Oxidation and sulfidation resistance
Carbon	High temperature strength and grain size control
Boron	High temperature creep strength
Manganese	Oxidation resistance
Lanthanum	Oxidation resistance
Ytrium	Oxidation resistance
Aluminum	Deoxidation (0.2–0.3%), gamma prime formation (0.6–2%)
Titanium	Gamma prime formation (2–3%)

Age-Hardenable Alloys. Age-hardenable alloys are well known for possessing high strength at intermediate temperatures where a precipitation hardening reaction, such as γ' $(Ni_3(Al,Ti))$, γ'' (Ni_3Nb), or other similar mechanism, is active. Table 1 provides a listing of the common age-hardenable alloys. HAYNES 263 alloy is often used for combustors and transition components in gas turbine engines. HAYNES 718 alloy is used extensively for compressor components. HAYNES R-41 alloy, also known as René-41 alloy, and Waspaloy alloy possess the highest strength characteristics at temperatures of 700–900 °C, and they are used to provide support to components in the hot gas path. At temperatures greater than about 900 °C, precipitation hardening is no longer effective, and their strength decreases dramatically.

Mechanical Properties. The typical tensile properties of several solid-solution-strengthened and age-hardenable superalloys are given in Table 3. Stress rupture properties are given in Table 4 and are illustrated in Fig. 1. The age-hardenable alloys have excellent strength levels in the temperature range of 650–870 °C. Up to 650 °C, 718 alloy has the highest strength, and is widely used up to this temperature. Although R-41 and Waspaloy alloys have much higher strength than solid solution strengthened alloys, such as 25, 188, and 230, fabrication problems in terms of weldability limit the use of these two alloys. This subject is discussed in greater depth in subsequent sections on welding. Among the solid-solution-strengthened alloys, the cobalt-based alloys 25 and 188 offer excellent stress-rupture properties; however, they are limited by thermal stability and relatively high cost. Of the nickel-based solid-solution-strengthened alloys, 230 alloy offers relatively high-stress rupture properties in the temperature range from 760 to 980 °C.

Oxidation Resistance. Alloys developed for gas-turbine and high-temperature applications depend mainly on chromium for oxidation resistance. Typical chromium contents range from 16 to 25% by weight. Small additions of elements such as manganese and silicon promote the formation of protective spinel oxides. Rare-earth elements, such as lanthanum, have been successfully utilized in alloys such as 188, 556, and 230 to further improve the oxidation resistance by enhancing their resistance to oxide spallation. The small, but potent addition of lanthanum to these three alloys dramatically improves oxidation resistance, as indicated in

Table 3 Typical tensile properties of common heat-resistant alloys in sheet form

(A) Yield Strength

Temp. (°C)	Alloy X (MPa)	Alloy 25 (MPa)	Alloy 188 (MPa)	Alloy 230 (MPa)	Alloy 625 (MPa)	Alloy 263[a] (MPa)	Alloy 718[a] (MPa)	Alloy R-41[a] (MPa)	Waspaloy Alloy[a] (MPa)
20	380	475	465	390	485	635	1175	815	910
650	245	330	275	270	370	525	995	760	770
760	235	285	270	285	340	515	750	750	770
870	195	250	250	225	205	220	420	575	415
980	90	125	130	120	80	87	90	275	135

(B) Ultimate Tensile Strength, MPa

Temp. °C	Alloy X	Alloy 25	Alloy 188	Alloy 230	Alloy 625	Alloy 263[a]	Alloy 718[a]	Alloy R-41[a]	Waspaloy Alloy[a]
20	765	1005	945	860	900	1050	1395	1250	1335
650	580	745	710	670	755	835	1155	1110	1195
760	465	640	620	585	595	725	855	1040	970
870	310	415	415	400	345	350	305	705	545
980	177	235	245	225	165	170	145	335	220

(C) Tensile Elongation, %

Temp. °C	Alloy X	Alloy 25	Alloy 188	Alloy 230	Alloy 625	Alloy 263[a]	Alloy 718[a]	Alloy R-41[a]	Waspaloy Alloy[a]
20	44	51	53	48	49	36	22	22	27
650	54	60	60	55	56	36	20	14	21
760	53	42	63	46	53	33	10	17	12
870	59	45	64	59	54	55	28	11	12
980	66	36	59	71	44	72	37	11	19

[a]Properties in age-hardened condition

Table 5. HAYNES 214 alloy differs from the other alloys in that it forms a protective aluminum oxide layer because of its 4.5% aluminum content. Excellent adherence of this oxide scale is promoted through the addition of a small amount of yttrium.

As mentioned earlier, between the cobalt-based alloys 25 and 188, the newer 188 alloy has much better oxidation resistance. Similarly, 556 alloy has much better resistance than the older iron-based MULTIMET alloy. Among the nickel-based alloys, 230 and 214 alloys offer the best oxidation resistance. Compared to the X alloy, 230 alloy offers two to three times better oxidation resistance. In the case of 214 alloy, the excellent oxidation resistance is attributed to the formation of protective aluminum oxide scale. However, the alloy is susceptible to strain-age cracking during welding, as discussed later in this paper. Therefore the applications for 214 alloy are mainly thin sheet or nonstructural components that demand excellent oxidation resistance at high temperatures.

Thermal Stability. Thermal stability is an important consideration for selecting heat-resistant alloys because it has a significant impact on thermal fatigue and repairability. Thermal stability reflects the resistance of a material to degradation in ductility after long thermal exposures. Residual room-temperature tensile elongation following such exposure is a common measurement of thermal stability. The thermal stability of several common alloys is compared in Table 6 and Fig. 2. Alloys 188, 25, X, and 625 suffer significant degradation after exposure to high temperatures. The cobalt alloys 188 and 25 form a deleterious Co_2W-type laves phase, while X forms σ and μ phases. In the case of 625 alloy, the degradation is related to the formation of Ni_3Cb phase. Alloy 230 does not form any of these deleterious phases, and the limited loss of ductility that does occur is related to the precipitation of $M_{23}C_6$-type carbides, which form in the other alloys as well.

The thermal stability of an alloy is important for long term service capability in components such as combustors. Because of thermal exposure and thermal cycling, loss of ductility due to exposure will result in reduced thermomechanical fatigue life. Table 7 shows the effect of exposure to 760 °C for 1000 h on the low-cycle fatigue behavior of several alloys at 427 °C. The data are plotted in Fig. 3. All of the alloys suffer some reduction in the low cycle fatigue life after exposure at 760 °C; however, 230 alloy is least affected by aging. Although the data reflect only an exposure time of 1000 h, longer exposure periods are expected to result in even greater degradation of properties for those alloys that are affected.

Table 4 Typical stress rupture properties of common heat-resistant alloys in sheet form

Stress to Produce Rupture in 1000 Hours, MPa

Temp. °C	Alloy X	Alloy 25	Alloy 188	Alloy 230	Alloy 625	Alloy 263[a]	Alloy 718[a]	Alloy R-41[a]	Waspaloy Alloy[a]
650		395		250	380	440	600		550
705		260		170	220	310	295	465	395
760	105	180	160	125	125	185	97	295	240
870	43	83	65	57	29	47		89	68
980	17	28	17	18	12				

[a]Properties in age-hardened condition
Some data based on Larson-Miller extrapolation

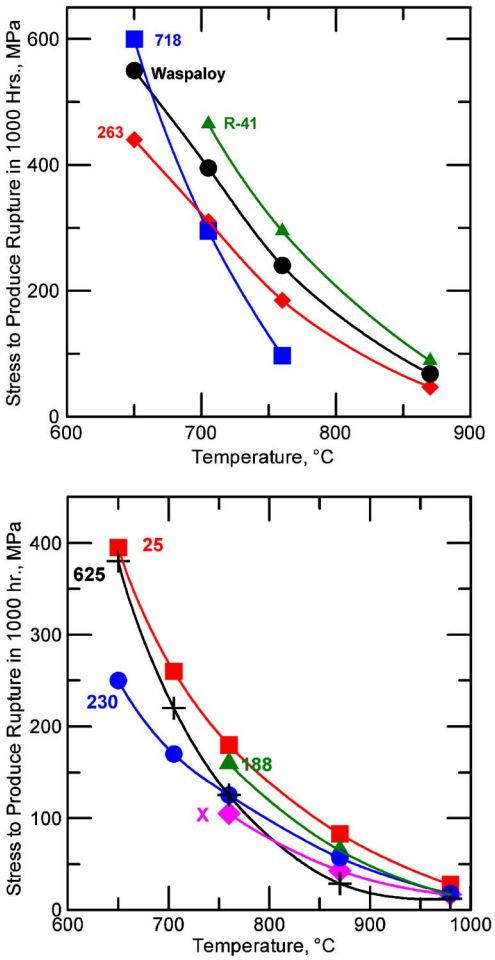


Fig. 1 Stress rupture properties of solid-solution-strengthened (bottom) and age-hardenable (top) alloys. Alloy 625 is included in both plots for reference.

Table 6 Thermal stability of several heat-resistant alloys: Room temperature tensile elongation after thermal exposure at the indicated temperature[a]

Alloy	No Exposure (%)	650 °C (%)	760 °C (%)	870 °C (%)
230	50	38	35	36
X	57	19	19	30
625	46	18	13	26
25	51	12[b]		5[c]
188	56	29	11	22

[a]All alloys, with the exception of 25 alloy, were exposed for 8000 hr at the indicated temperatures.
[b]Exposed for 2500 hr
[c]Exposed for 1000 hr

Weldability of Wrought Heat-Resistant Alloys

Many gas-turbine components rely on welding for their fabrication; therefore, weldability must be taken into account in alloy selection and design. Over the years, many alloys have been developed that excel in one or two major attributes, such as high-temperature strength or oxidation resistance, only to find that another attribute, such as weldability, severely limits use of the alloy. As discussed previously, R-41 alloy is among the highest strength wrought alloys available in the temperature range from about 700 to 900 °C. Although the alloy has been successfully used for many years in certain applications, widespread use of the alloy is limited by its susceptibility to strain-age cracking during heat treatment after welding and fabrication. Cracking is also a problem during manufacture of the alloy in wrought forms, which substantially increases its cost. This type of cracking is illustrated in Fig. 4. Alloy 263 is often selected over R-41 alloy, despite its much lower strength, owing to its superior fabricability. In modern alloy development it is recognized that, in addition to offering excellent service performance capabilities, an alloy needs additional attributes such as ease of welding and ease of manufacture in wrought forms in order to be successful in a wide range of applications and to be made available at a reasonable cost. Two of the most common problems encountered in the welding of superalloys are strain-age cracking and solidification cracking. Tests have been developed over the years to provide a measure of alloy

Table 5 Dynamic oxidation resistance of several heat-resistant alloys[a]

	980°C/1000 hr		1095°C/1000 hr	
	Metal loss (μm)	Average metal affected (μm)	Metal loss (μm)	Average metal affected (μm)
Alloy				
Cobalt-base				
25	157	211	>635[b]	>635[b]
188	28	89	191	249
Iron-base				
MULTIMET	300	366	1247[d]	1367[d]
556	43	124	221	274
Nickel-base				
X	69	142	229	328
625	124	180	>787	>787
617	69	249	315	>610[c]
230	20	71	56	132
214	10	25	13	30

[a]Results were generated using standard burner rig high-velocity oxidation tests. Samples were exposed in a rotating carousel to combustion products of No. 1 and No. 2 fuel oil burned at an air:fuel ratio of about 50:1. Gas velocity was 0.3 Mach. Samples were automatically removed from the gas stream every 30 min, and fan cooled to under 260 °C. After exposure, samples were evaluated metallographically to determine the extent of metal loss (metal thinning), and the average depth of internal penetration. Average metal affected is the sum of metal loss and average depth of internal penetration.
[b]Consumed in 165 hr
[c]Internal penetration through entire thickness
[d]Extrapolated from 225 hr

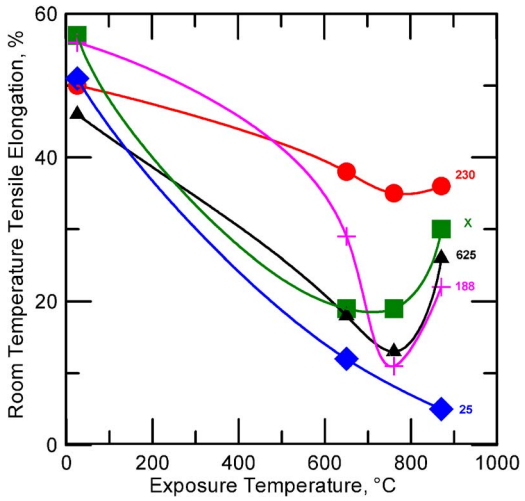


Fig. 2 Thermal stability of solid-solution-strengthened alloys (room temperature tensile elongation after aging for 1000 hr at various temperatures)

cracking resistance for use in alloy selection and design. The test methods and alloy rankings are described in the following section.

Strain-Age Cracking. Strain-age cracking is a type of postweld heat-treatment cracking that can occur in age-hardenable alloys. This type of cracking occurs in the solid state and typically follows an intergranular fracture path. Cracking can be associated with the heat affected zone, weld metal, or parent metal. Alloys containing titanium and aluminum that are strengthened by precipitation of the γ' phase are especially susceptible. These alloys exhibit low ductility during the initial stages of precipitation strengthening. The susceptibility of alloys to cracking generally increases with increasing atomic percent aluminum plus titanium. Alloys R-41 and Waspaloy are susceptible to this type of cracking, but are successfully welded in many applications, especially in sheet-metal thickness. Alloys with a higher aluminum and titanium content than these two alloys are not commonly available as wrought materials and are difficult to weld without cracking. Alloys containing columbium, such as 718, are more resistant to cracking than alloys strengthened with only titanium and aluminum.

Test methods used for strain-age cracking fall into two general categories: restrained weldment tests and mechanical property tests. In a restrained weldment test, a weldment is made under restrained conditions to simulate an actual fabrication, then the assembly is heat treated and inspected for cracking. Mechanical property tests measure some mechanical property that is assumed to be predictive of strain-age cracking behavior. Mechanical property tests are quantitative and relatively economical, but it is not always clear whether or not they accurately predict cracking behavior in real fabrications.

The restrained weldment category of test includes the restrained circular patch test, as well as simpler types of restrained plate tests. Early investigators used the restrained circular patch test to evaluate the resistance of alloys to strain-age cracking [3,4]. In this test, a circular disk with a circular cutout in the center is welded into a thick restraining ring along the outer perimeter. A circular insert is then welded into the hole in the restrained disk, thus, creating a heavily restrained weldment. The assembly is subjected to a postweld heat treatment and inspected for cracking.

Table 7 Comparison of low-cycle fatigue behavior at 427 °C for annealed material versus aged material: total strain range of 1.0%, $R=-1.0$, $F=0.33$ HZ, cycles to failure

Alloy	Annealed	Aged (760 °C, 1000 hr)
617	5096	3072
X (Ref. [1])	7750	4889
625 (Ref. [2])	8454	5851
230 (Ref. [1])	8742	7033
188 (Ref. [1])	12,647	6970

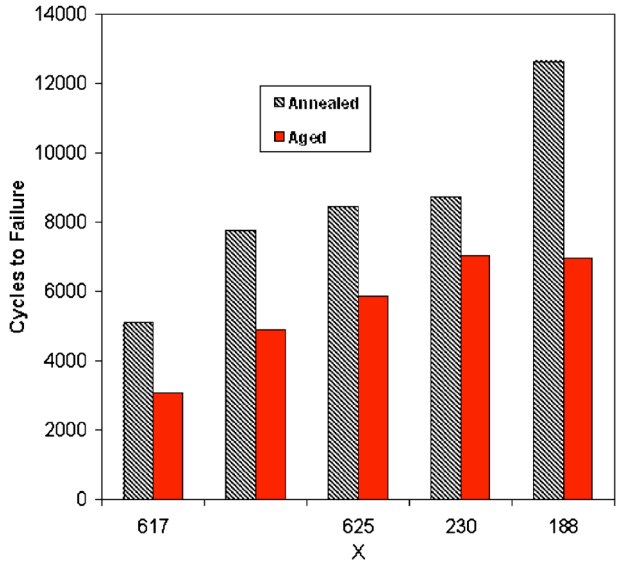


Fig. 3 Comparison of low-cycle fatigue behavior at 427 °C for annealed material versus material aged at 760 °C for 1000 hr. (Total strain range of 1.0%, $R=-1.0$, $F=0.33$ HZ).

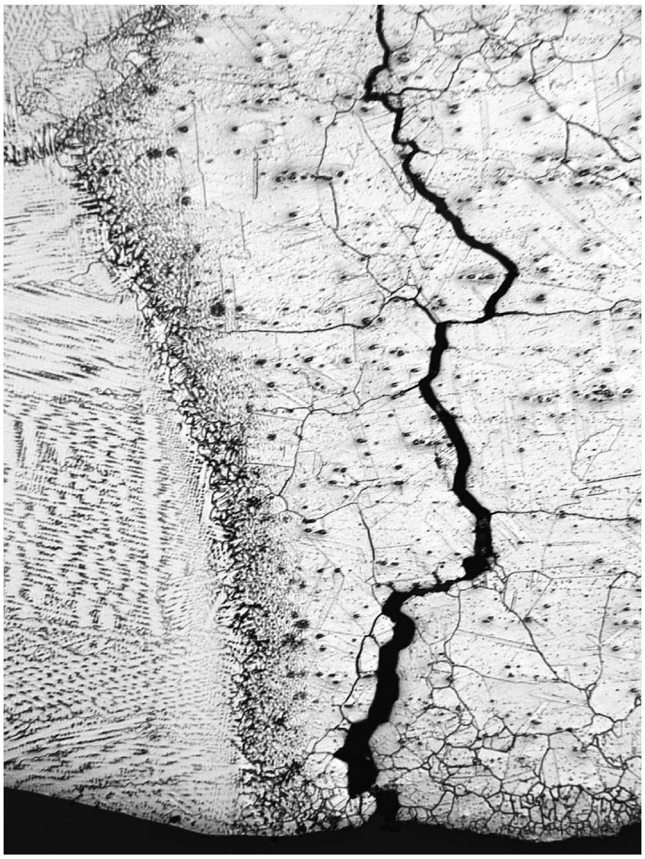


Fig. 4 Strain-age cracking of a simulated repair weld in R-41 alloy

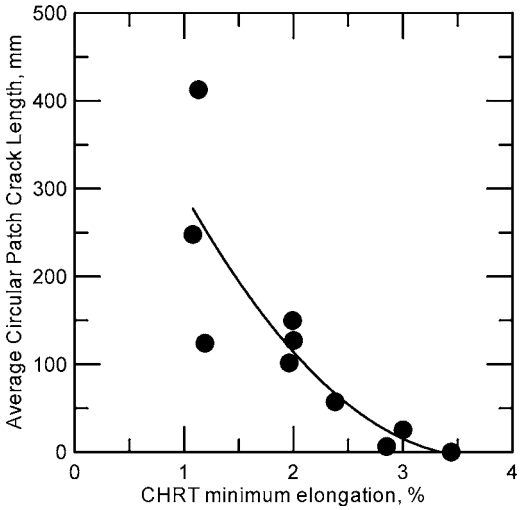


Fig. 5 Correlation between restrained circular patch test results and controlled heating rate test (CHRT) minimum elongation [8]

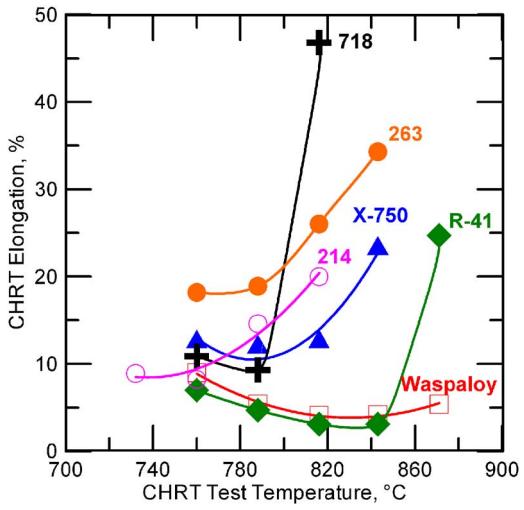


Fig. 6 Ranking of the susceptibility of commercial alloys to strain-age cracking by the controlled heating rate test (CHRT)

The circular patch test has the advantage of employing an actual restrained weldment, which is similar to a real fabrication. However, the cost in terms of labor and material for this type of test is relatively high and results are semi-quantitative at best. Semi-quantitative comparisons can be made by counting the number of cracks or the total crack length, but the test is only reproducible if the parts are precisely machined and joined with a highly controlled, automated welding process [5]. Repeatable results are difficult to obtain owing to the many variables involved that are difficult to control, such as bead shape and level of restraint. Restrained weldment tests are useful to demonstrate that an alloy can be welded, but it has many disadvantages as a ranking test.

Several investigators have used the Gleeble thermomechanical testing apparatus to simulate strain-age cracking. The Gleeble offers extreme flexibility in the thermal cycles and mechanical strains that can be applied, which has led to a wide variety of test methods and specimen geometries. Recent work is underway on development of Gleeble test methods at both Pratt and Whitney and The Ohio State University [6,7]. Gleeble test methods offer great potential because of the flexibility of the apparatus to simulate weld thermal cycles, heat-treatment thermal cycles, and apply varying degrees of restraint. However, the majority of Gleeble test methods have been the subject of a single paper, therefore the documentation is limited, and they have not been correlated to the behavior of actual restrained weldments.

One of the most thoroughly documented and economical test methods appears to be the controlled heating rate test (CHRT) developed at Rocketdyne in the late 1960s [8]. This test involves heating a sheet tensile specimen at a rate of 14–17 °C/min to a given test temperature, then pulling the specimen to failure and measuring the elongation. The test is performed over a range of temperatures from about 700 to 900 °C, and the minimum ductility over that range is the measure of cracking susceptibility for a given alloy or heat. The test is as economical as a standard tensile test and can be performed on the same equipment. Fawley and Prager evaluated several different mechanical tests and found that the CHRT gave the best correlation to the results of highly controlled and automated restrained circular patch tests, as shown in Fig. 5. The authors went on to use the CHRT to rank the cracking susceptibility of several heats of R-41 alloy and evaluate the effects of heat treatments and test atmosphere on cracking susceptibility. Testing in an inert atmosphere or vacuum was found to significantly improve resistance to cracking. The most important parameter determining the resistance of various R-41 alloy heats to cracking was found to be grain size; cracking susceptibility increases with larger grain size. The effectiveness of controlling grain size and heat treating in an inert atmosphere to control strain-age cracking was later confirmed in production at Rocketdyne.

Ranking Susceptibility of Alloys to Strain-Age Cracking. The controlled heating rate test can be used to rank the susceptibility of commercial alloys to strain-age cracking, as shown in Fig. 6. The susceptibility of alloys generally increases with the total content of γ'-forming elements, aluminum, titanium, niobium, and tantalum, on an atomic percent basis. Alloys that rely on niobium for strengthening, such as alloy 718, recover their ductility more quickly as temperature increases than alloys that are strengthened with aluminum and titanium alone. Alloys that perform well in the CHRT also tend to possess favorable hot working characteristics, because they have good ductility over a range of temperatures. This is helpful in manufacturing alloys at an affordable price.

Solidification Cracking. Solidification cracking occurs when solidification strains exceed the ductility of the solidifying weld metal due to the formation of low-melting-point liquid films along solidification grain boundaries. This type of cracking, also known as hot cracking, frequently occurs along the weld centerline. Solidification cracking can be caused by contamination of the weld fusion zone with sulfur or low-melting-point metals, but some alloys may crack in the absence of contamination. Alloys vary in their susceptibility to solidification cracking depending on their chemical composition. Heat-resistant alloys often contain elements, such as columbium, silicon, boron, or zirconium, that can promote solidification cracking.

Solidification Cracking Test Methods. The varestraint test is perhaps the most common test method used to rank the susceptibility of alloys to solidification cracking. In the test, a bending strain is applied to a sample while an autogenous gas tungsten arc weld is in progress [9]. Cracking is induced in the partially solidified metal at the trailing edge of the weld pool. The level of applied strain can be varied to tailor the severity of the test to alloys that are more or less susceptible to cracking. The amount of cracking is then quantified as the total crack length or maximum crack distance allowing a ranking of materials with regard to cracking resistance. The minimum level of augmented strain necessary to produce cracking is another indicator of cracking susceptibility; a greater critical strain to produce cracking indicates greater resistance to cracking. The most commonly used ranking parameter is the maximum crack distance because it is believed to

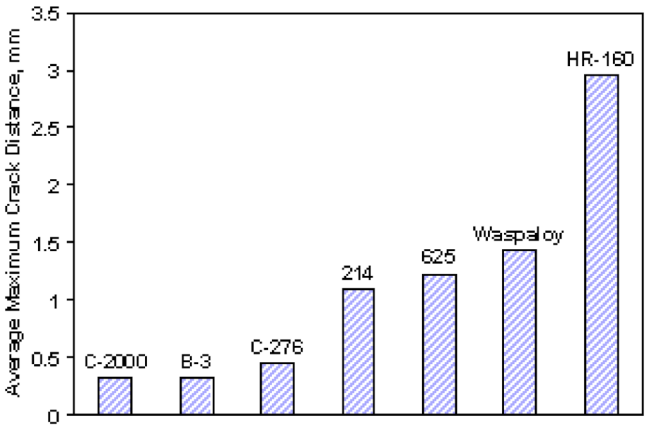


Fig. 7 Ranking of the susceptibility of commercial alloys to solidification cracking by the varestraint test

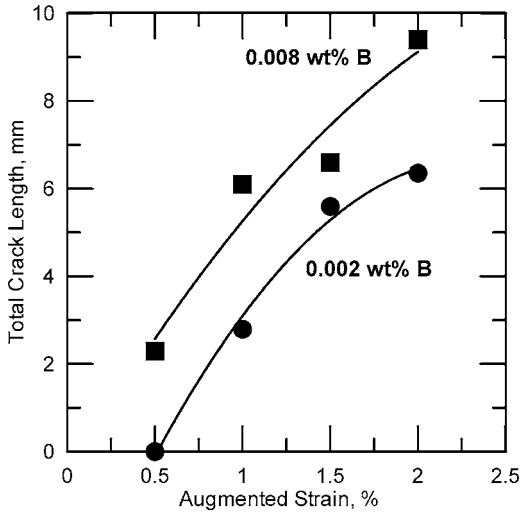


Fig. 8 Influence of boron content on solidification cracking resistance of 230 alloy, as indicated by the varestraint test

correlate to the solidification temperature range of the alloy. Additional research would be of interest to show which varestraint test ranking parameter best correlates to cracking behavior in actual weldments. In addition to solidification cracking, the longitudinal varestraint test can also detect susceptibility to some types of heat-affected zone cracking.

The sigmajig test [10] was developed more recently than the varestraint test, and offers a more quantitative index of cracking susceptibility. In the test, stress is applied to a 50×50 mm sheet, transverse to the welding direction, then an autogenous weld is made with the gas tungsten arc welding (GTAW) process. The test is repeated with stepwise increases in the preapplied stress until a critical stress level to produce cracking is determined. The critical stress level varies from one alloy to another; higher critical stress indicates greater resistance to cracking.

When the varestraint and sigmajig tests are both used on the same set of alloys, the two tests do not always provide exactly the same results. Maroef et al. used the two tests to investigate the influence of silicon and iron on the solidification cracking resistance of HR-160 alloy [11]. Both tests indicated the predominant detrimental effect of silicon. However, the varestraint test indicated a slight detrimental effect of iron, whereas the sigmajig test indicated a slight beneficial influence of iron. Future research may shed more light on the relationship between weldability tests and cracking behavior in actual fabrications.

Elements that are known to increase the susceptibility of nickel-based alloys to solidification cracking include silicon, boron, zirconium, niobium, sulfur, and phosphorus. The detrimental influence of sulfur and phosphorus is well known, and these elements are controlled to low levels in modern alloys. However, elements such as boron, zirconium, silicon, and niobium must be tolerated in many alloys owing to their beneficial influence on properties other than weldability.

A comparison of several commercial alloys is given in Fig. 7. The corrosion resistant alloys HASTELLOY C-2000®, C-276, and B-3®, which possess excellent resistance to solidification cracking, are included for comparison. The heat-resistant alloys in Fig. 7 have greater susceptibility to solidification cracking than the corrosion resistant alloys owing to the addition of boron and zirconium in 214 and Waspaloy alloys, niobium in 625 alloy, and silicon in HR-160 alloy. These alloying additions have a somewhat detrimental effect on weldability, but are necessary to ensure the service capability of the alloys.

Use of Weldability Testing in Modern Alloy Development. Both varestraint testing and restrained plate weldment testing were used to develop a weldable filler metal for 230 alloy. It was found that boron, added to the base metal to improve creep rupture strength, increased susceptibility of the alloy to solidification cracking in the weld metal, as shown in Fig. 8 [12]. Thus, a boron-free weld metal was developed. Ernst showed correlation between varestraint test results and restrained plate weldment test results [13].

In the development of 214 alloy it was necessary to strike a compromise between two weldability issues: strain-age cracking and solidification cracking. Boron and zirconium additions improve resistance of the alloy to strain-age cracking, as shown in Fig. 9. However, boron and zirconium increase susceptibility of the alloy to both solidification cracking and HAZ liquation cracking [14]. Weldability testing was used to determine the least addition of boron and zirconium that would be effective in imparting resistance to strain-age cracking. In addition, a boron- and zirconium-free filler metal was developed to avoid solidification cracking. Varestraint testing has also been used to demonstrate that nitrogen additions to the shielding gas improve the weldability of 214 alloy by binding zirconium and boron in the form of nitrides [15]. This improvement was later confirmed in production welding.

Alloys 718 and 625 depend on niobium for strengthening. These alloys are known to be moderately susceptible to solidifi-

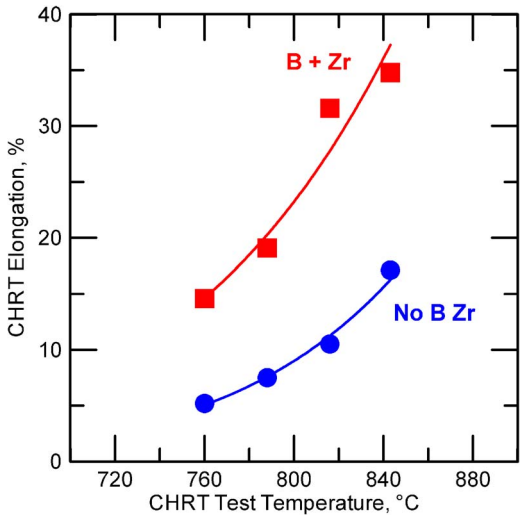


Fig. 9 Effect of boron and zirconium additions on the strain-age cracking resistance of 214 alloy, as measured by the controlled heating rate test

cation cracking due to formation of niobium-rich eutectic-type products on solidification. Varestraint testing has been used to show that carbon has an important beneficial influence on solidification cracking resistance in niobium-bearing alloys by promoting formation of niobium carbide rather than laves phase as the terminal solidification product [16]. Although the compositions of 718 and 625 are well established, this example illustrates how weldability test methods reveal information that can be used to develop alloys with improved weldability.

Selection and Applications of Heat-Resistant Alloys

Multiple factors must be taken into account when selecting alloys for demanding applications, including the in-service properties of an alloy as well as weldability and fabricability. One of the most important applications of heat-resistant alloys is in gas turbines, such as auxiliary power units (APUs), aircraft engines, and industrial gas turbines. The selection of the alloys used in these applications depends on the conditions of temperature, stresses, and environment (oxidation, sulfidation, and nitridation).

Wrought superalloys are used for many applications in aircraft and industrial gas turbines, including combustors, fuel nozzles, manifolds, igniters, transitions, casings, seal rings, exhaust, and noise-suppression components. Combustor alloys are selected for high-temperature strength, fatigue resistance, and resistance to oxidation, along with ease of forming and welding. Fabrication of combustors requires alloys in the form of sheet and plate. The commonly used combustor alloys include 188, 230, X, 263, 617, and 625. Combustors most commonly fail by fatigue resulting from thermal cycling of the combustor and buckling. Because of increasing requirements for higher thrust in aerospace engines and higher efficiency in industrial engines, the combustion temperatures continue to be increased. Consequently, alloys such as 188 and 230 are increasingly replacing the less capable alloys, such as X.

Materials of construction for fuel nozzles, manifolds, and igniters must possess high-temperature strength and environmental resistance. Required product forms are billet, bar, and casting. Alloys 25, 188, and X are commonly used for these components. Cobalt-based alloys are preferred if the combustion environment contains sulfur, as with the use of fuel oil, due to their combination of high-temperature strength, oxidation, and sulfidation resistance.

Transition ducts carry the hot combustion products from the combustors into the high-pressure turbine section. Fatigue and buckling are common failure modes for transitions, thus high-temperature mechanical strength and fatigue resistance are necessary for these components. Alloys in the form of plate and sheet are required. The alloys used for transitions include X, 263, 617, and 230.

Alloys for casings are selected based on their combination of high-temperature tensile and creep strength, fatigue resistance, and oxidation resistance. Alloys 718 and Waspaloy are commonly used for construction of casings.

Materials for seal ring construction require high-temperature strength and ductility for the purpose of ballistic containment. Furthermore, a low coefficient of thermal expansion is critical. Alloy 242 is age-hardenable by a novel mechanism of strengthening by formation of $Ni_2(Mo,Cr)$ long-range-ordered domains. This mechanism also provides 242 alloy with good ductility and resistance to cracking. In combination with a low coefficient of thermal expansion, these properties have led to the selection of 242 alloy for seal ring applications.

Exhaust components require materials with high strength at moderate temperatures, oxidation resistance, and good welding and forming characteristics. X and 625 alloys are commonly used for exhaust components.

Conclusions

Alloy selection depends on the specific service requirements demanded by the application as well as the complexity of fabrication of the component. The lowest cost alloy that meets these requirements will be selected. Many alloys are available that excel in one property or another, but may fall short in attributes such as weldability or thermal stability. In the future, successful alloys will need to offer not only excellent properties, such as strength and oxidation resistance, but at the same time possess excellent weldability and fabricability. This can only be accomplished by using weldability test methods to screen experimental alloys at early stages of development. Alloys such 230 and 188 excel in their combination of high temperature strength and fabricability. However, there are increasing demands from gas turbine designers and manufacturers for alloys that surpass these alloys in terms of strength, yet retain ease of fabricability. Research devoted to development of alloys that meet these demanding requirements is ongoing.

References

[1] Klarstrom, D. L., and Lai, G. Y., 1988, "Effects of Aging on the Low Cycle Fatigue Behavior of Three Solid-Solution-Strengthened Superalloys," *Superalloys 1988*, S. Reichman et al. (eds.), TMS, pp. 585–594.

[2] Srivastava, S. K., 1996, Haynes International Technical Report No. 13866.

[3] Lepkowski, W. J., Monroe, R. E., and Rieppel, P. J., 1960, "Studies on Repair Welding Age-Hardenable Nickel-Base Alloys," Weld. J. (Miami, FL, U. S.), **39**(9), pp. 392s–400s.

[4] Weiss, S., Hughes, W. P., and Macke, H. J., 1962, "Welding Evaluation of High Temperature Sheet Materials by Restrained Patch Testing," Weld. J. (Miami, FL, U. S.), **41**(1), pp. 17s–22s.

[5] Thompson, G., Nunez, S., and Prager, M., 1968, "Practical Solutions to Strain-Age Cracking of Rene 41, Weld. J. (Miami, FL, U. S.), **47**(7), pp. 299s–313s.

[6] Lin, W., 2000, "A Methodology for Quantifying Postweld Heat Treatment Cracking Susceptibility," abstract, 81st Annual AWS Convention, p. 158.

[7] Norton, S., and Lippold, J. C., 2001, "Development of a Gleeble Based Test for Postweld Heat Treatment Cracking Susceptibility," abstract, 82nd Annual AWS Convention, p. 129–130.

[8] Fawley, R. W., Prager, M., Carlton, J. B., and Sines, G., 1970, "Recent Studies of Cracking During Postwelding Heat Treatment of Nickel-Base Alloys," WRC Bulletin No. 150, Welding Research Council, NY.

[9] Savage, W. F., and Lundin, C. D., 1965, "The Varestraint Test," Weld. J. (Miami, FL, U. S.), **44**(10), 433s–442s.

[10] Goodwin, G. M., 1987, "Development of a New Hot-Cracking Test–The Sigmajig," Weld. J. (Miami, FL, U. S.), **66**(2), 33s–38s.

[11] Maroef, S., Rowe, M. D., and Edwards, G. R., 2003, "The Effect of Silicon and Iron on the Weldability of Ni-Co-Cr-Si HR-160 Alloy," *Proc. of 6th Int. Trends in Welding Research Conf.*, April 15–19, 2002, Pine Mountain, GA, ASM International, Metals Park, OH.

[12] Tawancy, H. M., 1983, Haynes International Technical Report No. 11191.

[13] Ernst, S. C., 1994, "Weldability Studies of HAYNES® 230 Alloy," Weld. J. (Miami, FL, U. S.), **73**(4): 80s–89s.

[14] Cieslack, J., Stephens, J. J., and Carr, M. J., 1988, "A Study of the Weldability and Weld Related Microstructure of Cabot Alloy 214," Metall. Trans. A, **19A**(3), pp. 657–667.

[15] Lai, G. Y., Herchenroeder, R. B., and Patriarca, C. R., 1994, "Shielding Gas to Reduce Weld Hot Cracking," U.S. Patent 5,306,358.

[16] DuPont, J. N., Robino, C. V., and Marder, A. R., 1998, "Solidification and Weldability of Nb-Bearing Superalloys," Weld. J. (Miami, FL, U. S.), **77**(12), pp. 417s–431s.

Contact Dynamic Response With Misalignment in a Flexible Rotor/Magnetic Bearing System

P. S. Keogh
Department of Mechanical Engineering,
Faculty of Engineering and Design,
University of Bath,
Bath BA2 7AY, UK

M. O. T. Cole
Department of Mechanical Engineering,
Faculty of Engineering and Design,
University of Bath,
Bath BA2 7AY, UK
and
Department of Mechanical Engineering,
Chiang Mai University,
Chiang Mai 50200, Thailand

This paper investigates the vibration characteristics of rotor displacement signals in a magnetic bearing system under conditions when rotor contact with auxiliary bearings is possible. Since these signals may be used for feedback control, it is necessary to determine how they may affect the ability of the controller to regain rotor levitation. An experimental system is used to demonstrate the sensitivity of the rotor nonlinear dynamic behavior to unbalance, which is sufficient to cause contact during rotor run-up through rigid-body and flexural mode critical speeds. Complex rotor dynamics may involve contact with more than one auxiliary bearing or bush. Application of appropriate rotating forces to the rotor through a magnetic bearing is also shown to induce similar contact dynamics. Thus, an alternative procedure for assessing the nonlinear rotor dynamic behavior is established with the potential for identification of appropriate control forces. The contact dynamics are also considered in the presence of auxiliary bearing misalignment. Misalignment may arise through physical translation of a housing or through steady-state offset errors in sensor measurements. A misalignment of 50% of the nominal radial clearance is applied at an auxiliary bearing. Various contact modes are evident as the rotor is run up in speed. During rundown, different contact dynamics may be encountered and the level of such hysteresis is assessed. [DOI: 10.1115/1.2056530]

Introduction

To provide a secondary support and prevent rotor-stator lamination contact in magnetic bearing systems, it is necessary to install auxiliary bearings or bushes having appropriate radial clearance at a number of axial locations along the rotor. The dynamics induced when the rotor makes contact with an auxiliary bearing may have adverse consequences even if the magnetic bearings are still fully functional. It is important to be able to explain such behavior in order that controllers may be designed to restore the rotor to a noncontacting condition. For the well-known case of backward whirl [1], contact forces are likely to be significantly greater than the load capacity of a magnetic bearing. This means that once backward whirl has become established, it would not be possible to apply control forces of appropriate magnitude to restore position control. However, for other contact modes, restoration of position control without rundown is a feasible option. The basic problem for control design is that two or more rotor vibration states may exist for the same unbalance condition, namely, a single noncontacting orbit state and one or more contact mode states. Hence a controller designed for unbalance compensation alone could be inappropriate for a contact mode, resulting in even harder contact.

There have been many studies on the interaction of a rotor with a stator component. If the rub is light in the sense that contact forces do not directly influence the rotor dynamics, frictional heating of the rotor may lead to thermal bending [2]. In most studies involving more significant contact forces, magnetic bearing characteristics are not usually considered and the rotor dynamics are concerned with motion in a clearance circle. However, these papers reveal the nature and range of nonlinear responses. Analysis of synchronous rub behavior originates from the works of Johnson [3] and Black [4]. Subsynchronous response was examined by Childs [5], Muszynska [6], and Ehrich [7]. This led to the study of chaotic response [8,9]. A useful experimental study was undertaken by Wu and Flowers [10] to establish regimes of rub and bouncing modes. Further work by Lawen and Flowers [11,12] and Flowers et al. [13] focuses on the synchronous dynamics associated with rotor-auxiliary bearing contact. The papers of Fumagalli et al. [14] and Fumagalli and Schweitzer [15] are noted for the measurement of vibration and contact forces associated with high-speed rotors. In contrast, Kirk et al. [16], Kirk, [17], and Schmied and Pradetto [18] are notable for the experimental procedures involving drop tests with larger industrial rotors.

In recent years, a number of researchers have developed analytical procedures to investigate the contact problem. Isaksson [19] examined unbalance response and stability, Markert and Wegener [20], Markert et al. [21], and Wegener and Markert [22] studied the influence of different contact models in single disk and multidisk rotor systems. With regard to the contact problem involving a rotor levitated by a magnetic bearing, Ecker [23,24] considered the problem for a single mass rotor and determined different regimes of dynamic behavior. Cuesta et al. [25,26] performed experimental tests and developed an analytical model to identify nonlinearity in the orbit behavior. The present authors have also developed modeling and experimental methods to restore position control of a rotor from a state of contact with an auxiliary bearing [27–29]. It is acknowledged that other auxiliary bearing designs have shown potential. Chen et al. [30] provide details on a novel zero-clearance design. The foil bearing of Swanson and Heshmat [31] was also shown to function under a significant range of fault conditions.

This paper considers a flexible rotor/magnetic bearing system, which contains several critical speeds in the running speed range corresponding to modes involving rotor flexure. It cannot be simplified to a Jeffcott rotor model since the rigid-body and flexural modes are not well separated in frequency. The primary objective is to develop experimental procedures to investigate the rotor dynamics induced by contact with one or more auxiliary bearings and/or bushes. It is proposed that the magnetic bearings may be used to induce contact by applying appropriate circular forces.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. GT2004-53818.

Misalignment is also investigated, which may arise in practice through physical translation of a housing or through steady-state offset errors in sensor measurement signals.

Preliminary Rotor Dynamics

The rotor dynamics may be represented using finite element analysis, resulting in a matrix equation of the form

$$\mathbf{M}\ddot{\mathbf{x}} + \mathbf{C}(\Omega)\dot{\mathbf{x}} + \mathbf{K}\mathbf{x} = \mathbf{f}_w + \mathbf{f}_u + \mathbf{B}_c\mathbf{f}_c + \mathbf{B}_u\mathbf{u} \tag{1}$$

For a beam element formulation, the rotor displacement vector $\mathbf{x}$ defines translation and pitch at each element node in two orthogonal planes through the axis of rotation. The damping matrix in Eq. (1) is speed dependent through gyroscopic effects. Unbalance forces are generally possible at each element node, whereas contact forces from auxiliary bearings and magnetic bearing forces apply at specific element nodes. Hence the matrices $\mathbf{B}_c$ and $\mathbf{B}_u$ are required to increase the dimensions of $\mathbf{f}_c$ and $\mathbf{u}$ to that of $\mathbf{x}$.

If displacement transducer signals are used for feedback to the magnetic bearings the measured states are generally a subset of $\mathbf{x}$ according to

$$\mathbf{x}_m = \mathbf{C}_m\mathbf{x} \tag{2}$$

To stabilize the negative stiffness characteristic and provide a centralizing effect on the rotor a proportional-integral-derivative (PID) law may be used. The control force vector is written as

$$\mathbf{u} = -\mathbf{K}_P\mathbf{C}_m\mathbf{x} - \mathbf{K}_D\mathbf{C}_m\dot{\mathbf{x}} - \mathbf{K}_I\int_0^t \mathbf{x}_m(\tau)d\tau + \mathbf{u}_r \tag{3}$$

where $\mathbf{u}_r$ is a residual magnetic bearing force vector, either from another control law or from applied test signals. It is taken in Eq. (3) that any inherent negative magnetic bearing stiffness has been absorbed into the rotor stiffness matrix. Combining Eqs. (1) and (3) now gives

$$\mathbf{M}\ddot{\mathbf{x}} + [\mathbf{C}(\Omega) + \mathbf{B}_u\mathbf{K}_D\mathbf{C}_m]\dot{\mathbf{x}} + (\mathbf{K} + \mathbf{B}_u\mathbf{K}_P\mathbf{C}_m)\mathbf{x} + \left(\mathbf{B}_u\mathbf{K}_I\int_0^t \mathbf{x}_m(\tau)d\tau - \mathbf{f}_w\right) = \mathbf{f}_u + \mathbf{B}_c\mathbf{f}_c + \mathbf{B}_u\mathbf{u}_r \tag{4}$$

If the rotor-magnetic bearing system is operating normally without contact with any auxiliary bearing then $\mathbf{f}_c = \mathbf{0}$. Control forces in $\mathbf{u}_r$ may then be used to achieve desired performance objectives, for example, on vibration and transmitted force levels. In cases of power loss and rotor drop [14–18], $\mathbf{f}_c \neq \mathbf{0}$, and all control terms in Eq. (4) need to be deleted. The concern of the present paper is the operation of the system when control terms remain and contact with an auxiliary bearing is possible. This may arise, for example, due to appreciable unbalance over a speed range, a momentary malfunction, or a disturbance input, such as base motion. The problem of restoring position control from a contacting to a noncontacting condition has received little attention. This is a nontrivial problem since a controller designed to attenuate rotor vibration in a noncontacting condition may increase contact force levels if applied to the rotor in a contacting condition. This has been demonstrated in [27–29], for controllers that utilize the complex synchronous components of measured vibration signals. The problem arises due to large phase shifts induced in the synchronous components, which may cause the magnetic bearings to apply inappropriate control forces. Likewise, modification of PID gain settings may be ineffective and lead to magnetic bearing saturation problems.

The design of controllers that are able to restore position control of a rotor from a contacting condition requires an understanding of the influence of the contact dynamics on measurement signals that are used for feedback. The options are as follows:

(a) undertake detailed analytical and numerical modeling to establish characteristics of expected measurement signals

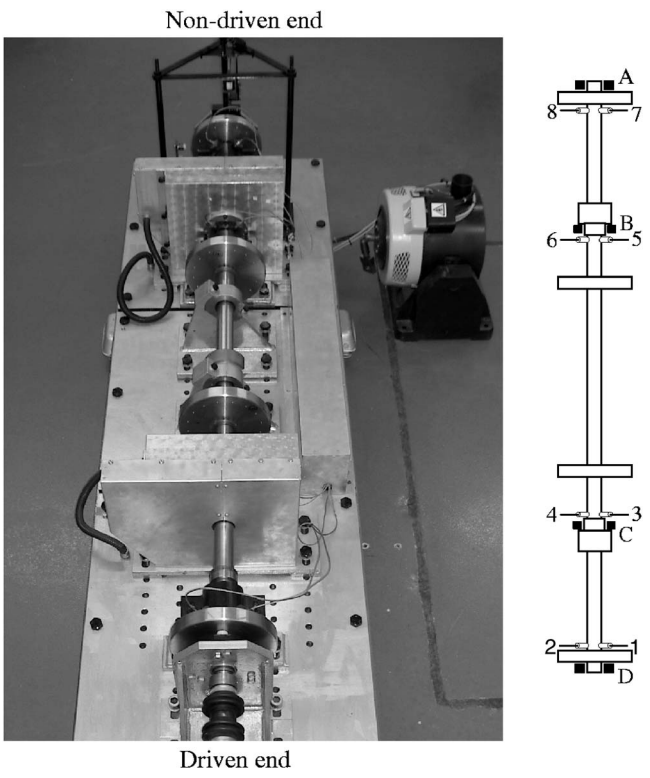


Fig. 1 Rotor/magnetic bearing facility. Auxiliary bearings (A–D), sensors (1–8) at ±45 deg to vertical.

(b) apply different unbalance distributions to the rotor and assess the measured signals in an identification process

(c) use the magnetic bearings to apply forces that are capable of inducing contact with an auxiliary bearing

Although many studies report modeling as described in option (a), the implications for controller design are not covered. Furthermore, almost all modeling is focused on a system with only one auxiliary bearing. For option (b), the authors have reported the use of applied unbalance in [27] to establish control action for restoration of rotor position without contact. However, a problem with the approach of option (b) is that the rotor must be physically run up through a speed range with sufficiently high levels of rotor unbalance to induce contact. If a rotor prototype is available, this may be feasible, though high contact forces and slip speeds may result in some degradation of the rotor/auxiliary bearings and, hence, the identification process.

The option (c) approach was proposed by the authors in [29] on a theoretical basis. The technique was used successfully for the purpose of robust controller design, but no experimental validation of this identification method was given. The orbit induced by a particular unbalance condition at speed Ω may be reproduced using circular forces $[\mathbf{u}_r = \mathrm{Re}(\mathbf{U}_r e^{i\Omega t})]$ applied by the magnetic bearings at the same frequency. An advantage is that the rotor may be run at a much lower speed $\omega \ll \Omega$, such that uncertain unbalance distributions become insignificant ($\mathbf{f}_u \approx \mathbf{0}$). The assumption here is that the contact forces are not strongly dependent on the slip speeds. Since there is the potential of developing an inverse procedure to establish appropriate control actions, the intention is to demonstrate experimentally that applying appropriate magnetic bearing forces may reproduce contact dynamics induced by unbalance.

Figure 1 shows the experimental rotor/magnetic bearing system. The rotor is a stainless-steel shaft, of length 2 m and radius 25 mm, on which are mounted four equally spaced disks. Each disk has a mass of 10 kg and radius 10 cm. Also mounted on the

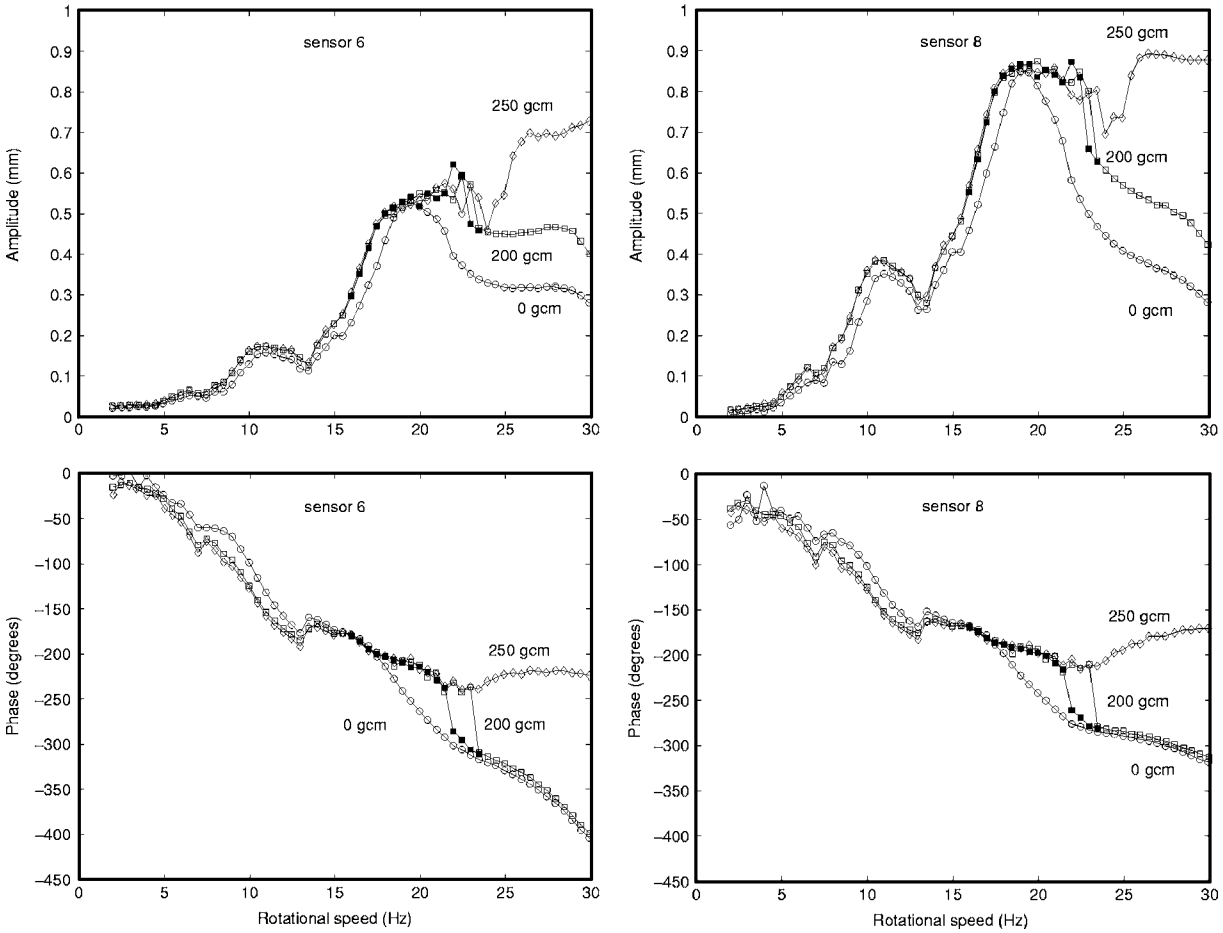


Fig. 2 Synchronous amplitudes and phases extracted from measured signals: Unbalance condition 0 gcm (○), 200 gcm on run-up (□) and rundown (■), 250 gcm (◇). 0% misalignment.

shaft are lamination stacks for two similar radial magnetic bearings capable of applying 1700 N over a bandwidth of 100 Hz. A finite element analysis indicated a first free-free flexural mode of ~27 Hz. The magnetic bearing proportional gains were set to give a nominal rigid-body lateral natural frequency of 15 Hz, while the derivative gains resulted in ~30% of critical damping. The total rotor mass was ~100 kg, giving a residual control force capacity of ~1200 N in each magnetic bearing. Auxiliary rolling element bearings (B and C) of nominal radial clearance 0.7 mm are integral to each magnetic bearing. Additionally, two greased, split bronze bushes (A and D) are located at the ends of the rotor and close to the two outer disks. These have a nominal radial clearance of 0.9 mm. The rotor was driven by a motor through a universal coupling. Rotor displacements could be measured using four eddy current displacement transducer pairs located just inboard of the end disks ((1,2) and (7,8)) and adjacent to the magnetic bearings ((3,4) and (5,6)). The control axes of each magnetic bearing were arranged at ±45 deg to the vertical, as was the case with the transducer pairs. The complete base frame of the system was mounted on isolators to the floor.

Unbalance-Induced Contact

An initial unbalance condition of the rotor was established such that ~24 Hz the largest orbit was at the nondriven end disk. Also the rotor could be run up in speed to 30 Hz without contact with any of the auxiliary bearings or bushes. Further unbalance masses could then be added at a fixed circumferential location on the nondriven end disk to induce contact between the rotor and nondriven end bush. Initially, the nondriven end bush (A) was aligned to be concentric with the stationary rotor (0% misalignment), when levitated under PID control. The integral feedback gain, which has the effect of centralizing the rotor within the magnetic bearings, was then set to zero. This did not change the centralized position of the rotor since static levitation forces had been established. However, it was found in trial runs of the rotor involving contact that a nonzero integral gain led to integrator wind-up, which resulted in the rotor being pulled hard onto the auxiliary bearings in a nonrepeatable manner. With the integral gain set to zero, this undesired dynamic response was eliminated.

Figure 2 shows the results of different unbalance conditions as the rotor is run up, then down, through the speed range. Specifically, synchronous components are extracted from a measurement signal $y(t)$ over the kth period according to

$$r_k e^{i\phi_k} = \frac{1}{T}\int_{(k-1)T}^{kT} y(t)e^{-i\Omega t}dt \tag{5}$$

where the phase is referred to a fixed point on the shaft. For noncircular orbits the synchronous components will depend on the sensor orientation. Significant transient behavior was allowed to decay before the synchronous components were extracted. Three cases are considered with additional unbalances on the nondriven end disk of 0, 200, and 250 gcm, relative to the initial indeterminate unbalance distribution. The synchronous amplitudes and phases correspond with sensor 6 at the nondriven end magnetic bearing and sensor 8 at the nondriven end disk, both arranged at 45 deg (clockwise) to the vertical when viewed from the nondriven rotor end. The 0 gcm responses did not involve contact with any auxiliary bearing, though only marginally so given that

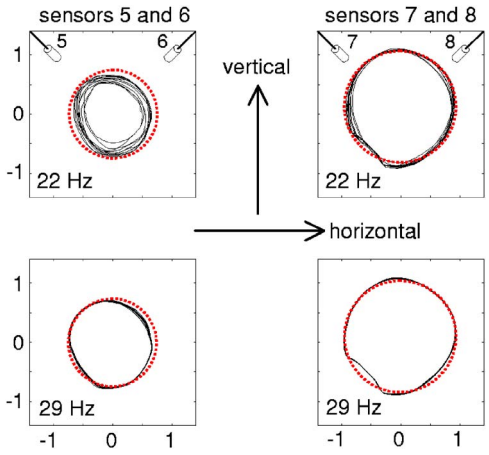


Fig. 3 Measured orbits with 0% misalignment condition and 250 gcm unbalance. Orbit dimensions are in millimeters and rotation is in the anti-clockwise sense within nominal clearance circles (· · · · · · · · ·).

the radial clearance close to sensor 8 is 0.9 mm. The phase plots for 0 gcm indicate significant rotor flexural response for speeds above 17 Hz. The 200 and 250 gcm responses are similar and slightly larger than those for 0 gcm at speeds up to 17 Hz. However, as the speed was increased further, audible contacts were apparent. For sensor 8, there is a limiting value of ~0.9 mm on the synchronous amplitudes and the synchronous phases are seen to deviate significantly from the 0 gcm case, indicating a nonlinear rotor dynamic response with contact. For 200 gcm, contact behavior ceased around 24 Hz, with the rotor dynamic response around 40% higher that the 0 gcm response through to 30 Hz. As the speed was then run down slowly from 30 Hz, the 200 gcm response remained without contact to ~22 Hz. Contacts then persisted until 17 Hz. Although small, the level of hysteresis in the response between run-up and rundown is clear.

The sensitivity of the response to unbalance is demonstrated with the relatively small change from 200 to 250 gcm. In Fig. 2, there is little to distinguish between synchronous components up to 17 Hz. However, the 250 gcm responses involved audible contact through to 30 Hz, both on run-up and rundown. Also, contacts occur both at the nondriven end disk and magnetic bearing, as evident from the amplitudes at sensors 6 and 8 above 26 Hz. Figure 3 shows the nature of measured orbits with 250 gcm unbalance. These are shown with the nominal clearance circles of bearings A and B. Note that the axial locations of the auxiliary bearings are slightly different from the sensor pair planes (Fig. 1). The orbits are forward whirl in the anti-clockwise sense. The bounce in the contact mode at sensors 7 and 8 is noted in the lower left quadrant, possibly the result of small errors in alignment.

To consider misalignment influences, the nondriven end bush (A) was shifted horizontally by 50% of the radial clearance. Using similar procedures to the 0% misalignment tests, Fig. 4 shows synchronous responses for two cases of unbalance. With the 200 gcm unbalance the rotor dynamic response was without aux-

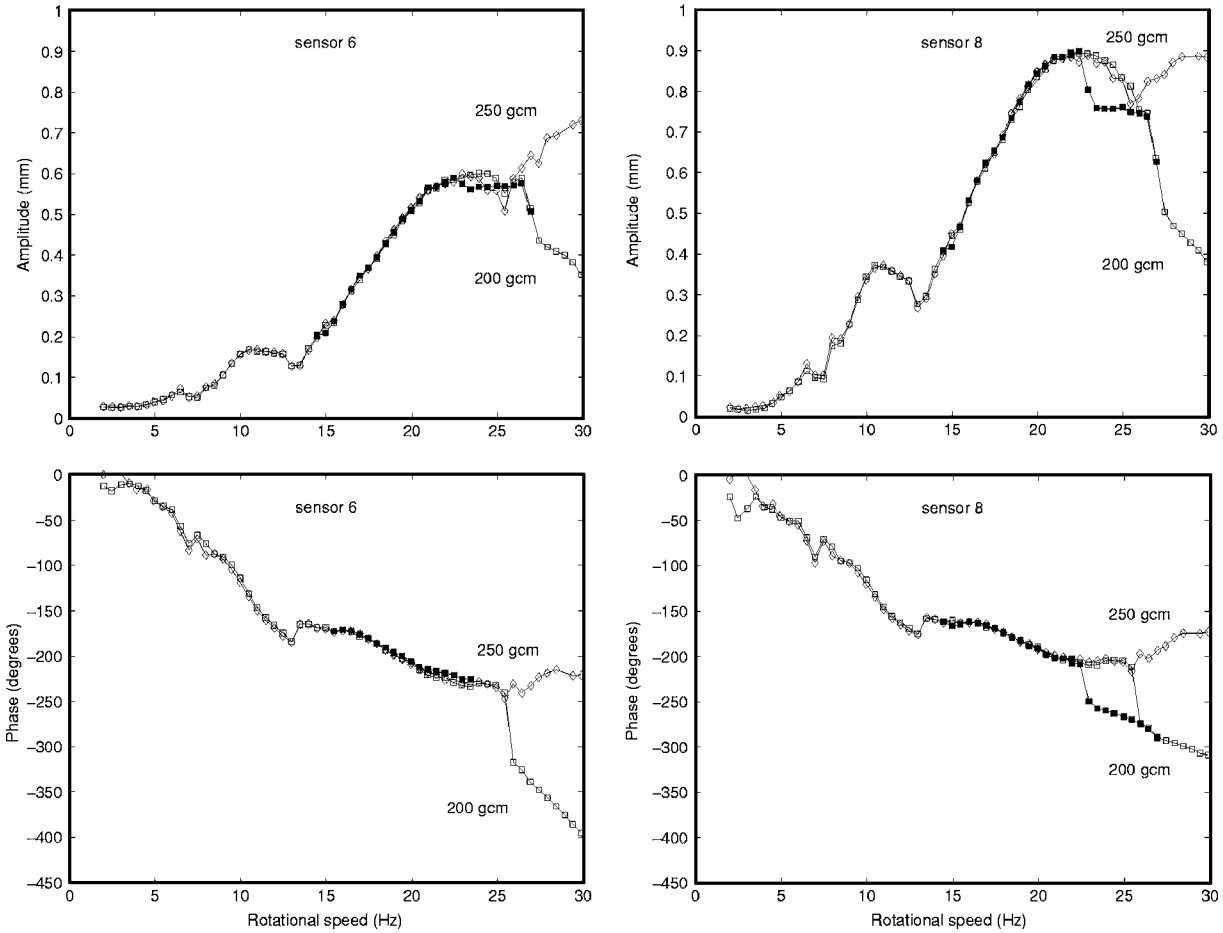


Fig. 4 Synchronous amplitudes and phases extracted from measured signals: Unbalance condition 200 gcm on run-up (□) and rundown (■), 250 gcm (◇). 50% misalignment.

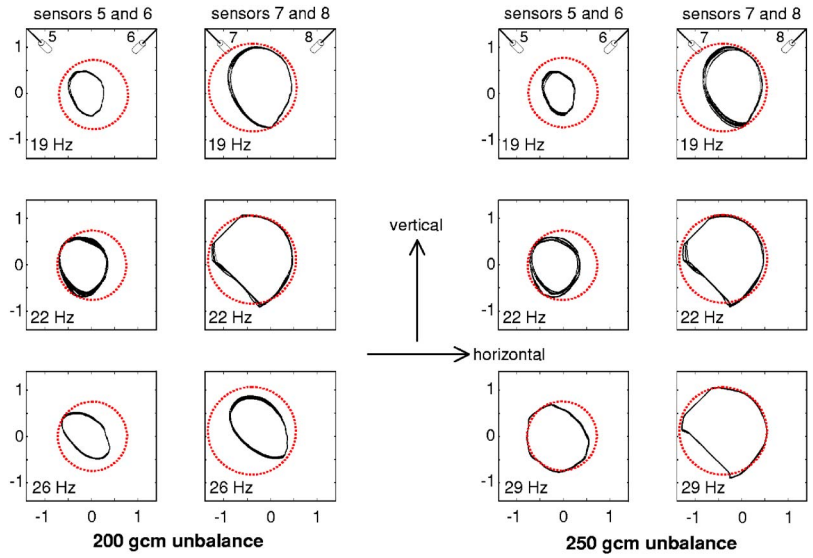


Fig. 5 Measured orbits with 50% misalignment of nondriven end bush to the left of the origin. Orbit dimensions are in millimeters and rotation is in an anticlockwise sense within nominal clearance circles (· · · · · · · ·).

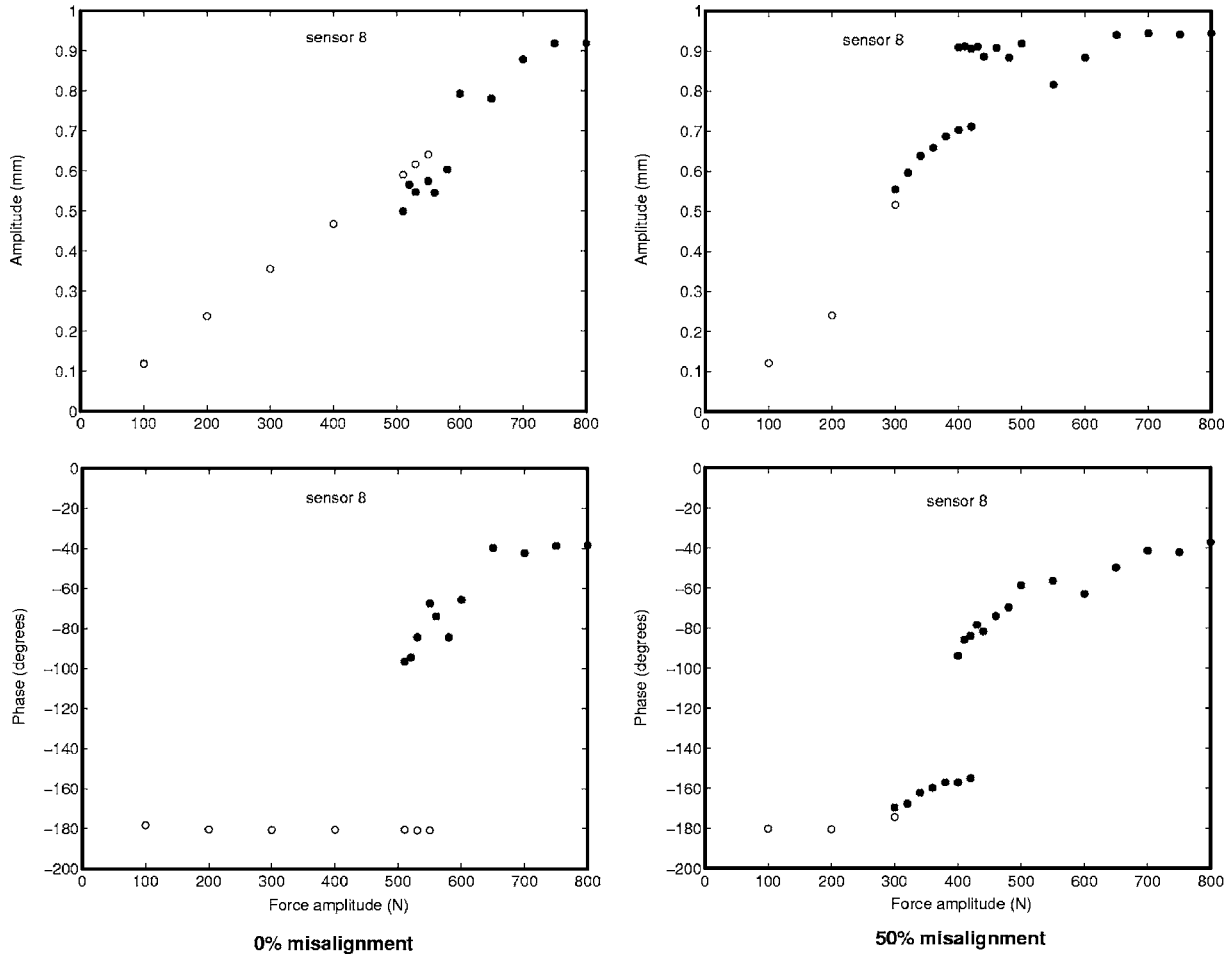


Fig. 6 Synchronous amplitudes and phases extracted from measured signals at circular forcing frequency of 24 Hz and rotor speed 2 Hz. No rotor contact is indicated by ○, and rotor contact by ●.

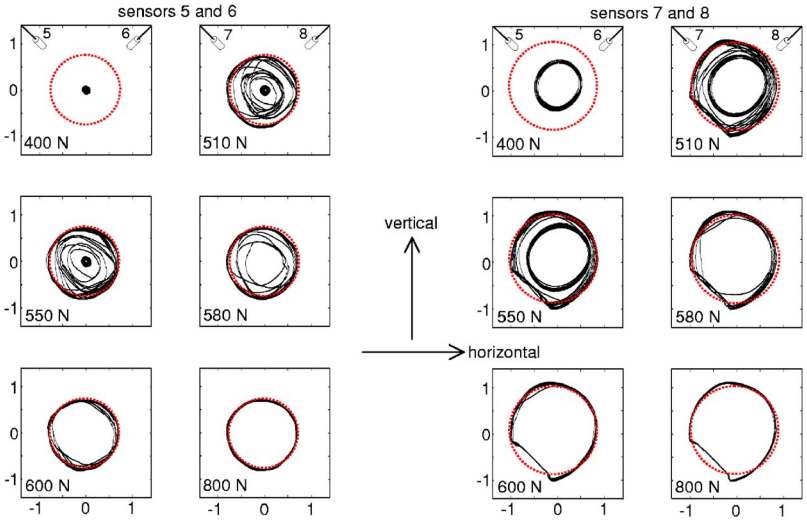


Fig. 7 Measured orbits with 0% misalignment condition under circular forcing from nondriven end magnetic bearing at 24 Hz with rotor speed 2 Hz. Orbit dimensions are in millimeters within nominal clearance circles (· · ·).

iliary bearing contact up to 16 Hz. Audible contacts were apparent from 16 to 26 Hz, which the ceased up to 30 Hz. The audible contacts coincided with nonlinear behavior, ending with the sharp drop in phases at 26 Hz for both sensors 6 and 8. As the rotor was run down from 30 Hz, audible contacts became evident just over 22 Hz through to 16 Hz. The nonlinear behavior on the rundown begins with the sharp rise in the phase around 22 Hz, associated with sensor 8 only. In contrast, the 250 gcm responses involved rotor/auxiliary bearing contact between 16 and 30 Hz on both run-up and rundown. Figure 5 shows examples of the measured orbits. The 50% misalignment condition corresponds with a translation of the nondriven end bush to the left by 0.45 mm relative to the origin. The orbits are forward whirl in the anti-clockwise sense and are of increased complexity compared to those of Fig. 3. At least two different types of contact mode are seen. At 22 and 29 Hz, misalignment induces more bounce on the left-hand orbit side with rub on the right-hand side.

Magnetic Bearing-Induced Contact

The previous results indicate how nonlinear rotor dynamic response behavior is sensitive to the unbalance condition. The unbalance tests were undertaken using an experimental system specifically designed to exhibit rotor flexural modes below 30 Hz. Typical industrial systems may operate well above this speed, which could present practical difficulties, such as frictional heating, in the undertaking of similar tests. As an alternative, it is proposed to carry out tests at an artificially low running speed, and then to superimpose circular forcing from a magnetic bearing at a frequency equivalent to a realistic operating speed. This is more controllable than the direct unbalance approach since the applied forces can be removed quickly in the event of severe rotor dynamic response. To validate the procedure, an artificially low running speed of 2 Hz was selected to guarantee positive slip between the rotor and auxiliary bushes/bearings for responses in the

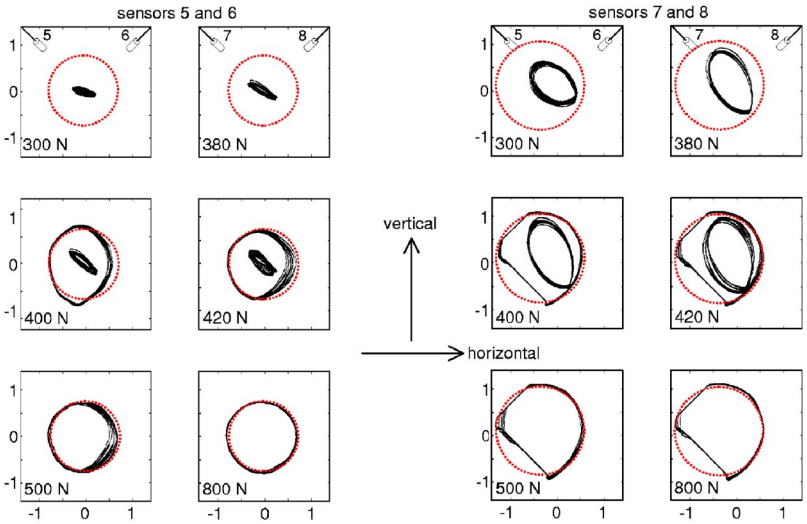


Fig. 8 Measured orbits with 50% misalignment condition under circular forcing from nondriven end magnetic bearing at 24 Hz with rotor speed 2 Hz. Orbit dimensions are in millimeters within nominal clearance circles (· · ·).

previous section. Positive forward circular forcing was then applied through the nondriven end magnetic bearing at 24 Hz because this coincided with a transition region between linear and nonlinear response for the unbalance tests. The circular force amplitude was varied over a range sufficient to induce contact.

Figure 6 shows the amplitudes and phases at 24 Hz extracted from measured signals at sensor 8. For 0% misalignment, the first audible contacts were experienced at 550 N forcing amplitude and persisted through to the chosen upper limit of 800 N. The forcing amplitude was then slowly reduced, and contacts were audible until 510 N. This can be concluded from the amplitude and phase data presented in Fig. 6. The plots show clear jumps in amplitude and phase, indicating deviation from linear behavior, at the force amplitudes stated. It was also found that contacts occurred at more that one auxiliary bush/bearing location. For the 50% misalignment condition, a 300 N amplitude was sufficient to induce contact with the nondriven end bush. The contact tended to occur at a localized point, which rotated clockwise until the force amplitude reached 420 N. Thereafter, a change in the mode of response occurred, which persisted to the upper limit of 800 N. As the force amplitude was slowly reduced, the second nonlinear mode was evident until 400 N, after which the response reverted to the localized point contact response. Selected orbits are shown in Figs. 7 and 8. For force amplitudes at which two types of response mode occurred, orbits for both cases are shown. It is now clear that the characteristics of the orbits induced by unbalance in Fig. 3 are also evident in the magnetic bearing-induced orbits of Fig. 7. Similarly, comparison of the orbits in Figs. 5 and 8 indicates that the nonlinear responses are equivalent. The reproduction of the complex detail in the orbits at the nondriven end bush is significant.

Conclusions

Magnetic bearings have been used to reproduce the characteristics of nonlinear orbits involving contact that are generated by appreciable unbalance. It has been shown that the rotor dynamic responses induced by contacts with auxiliary bearings and bushes could be correlated with those induced by circular forces applied through a magnetic bearing. Significantly, the responses generated by a magnetic bearing were undertaken at an artificially low running speed, thereby removing practical difficulties of obtaining direct unbalance responses at higher speeds. Comparable results were obtained for an end bush, both in alignment and misaligned by 50% of the radial clearance.

An alternative experimental approach for the study of rotor contact with auxiliary bearings or bushes has been demonstrated. Moreover, the complex orbit characteristics involving asymmetric bounce and rub were obtained without the need to perform detailed nonlinear dynamic analysis and simulation. The potential exists to identify appropriate control forces to restore a rotor from a condition of dynamic contact to one without contact. These will depend on the level of contact and degree of misalignment and will be the subject of future control strategy development.

Acknowledgment

The authors gratefully acknowledge the support of the Engineering and Physical Sciences Research Council of the UK in supporting this work though Grant No. GR/R45277/01.

Nomenclature

$\mathbf{B}_c, \mathbf{B}_u$ = force distribution matrices
$\mathbf{C}(\Omega)$ = rotor damping/gyroscopic matrix
$\mathbf{C}_m$ = transformation matrix
$\mathbf{f}_c$ = vector of contact force components, N
$\mathbf{f}_u$ = vector of unbalance force components, N
$\mathbf{f}_w$ = vector of static force (e.g., weight) components, N
k = signifies kth synchronous period
$\mathbf{K}$ = rotor stiffness matrix
$\mathbf{K}_\alpha$ = PID gain matrices ($\alpha = P, I, D$)
$\mathbf{M}$ = rotor mass matrix
r_k = amplitude of kth synchronous component, m
t = time variable, s
T = synchronous period ($2\pi/\Omega$), s
$\mathbf{u}$ = magnetic bearing force vector, N
$\mathbf{u}_r, \mathbf{U}_r$ = applied magnetic bearing forces, N
$\mathbf{x}$ = rotor displacement vector
$\mathbf{x}_m$ = vector of measured displacement states
y = general measurement signal
ϕ_k = phase of kth synchronous component, rad
ω, Ω = running speed/forcing frequency, rad/s

References

[1] Bartha, A. R., 1998, "Dry Friction Induced Backward Whirl: Theory and Experiment," *Proc. of 5th IFToMM Conf. Rotor Dynamics*, Darmstadt, Vieweg Publishers, Braunschweig, pp. 756–767.
[2] Newkirk, B. L., 1926, "Shaft Rubbing," Mech. Eng. (Am. Soc. Mech. Eng.), **48**, pp. 830–832.
[3] Johnson, D. C., 1962, "Synchronous Whirl of a Vertical Shaft Having Clearance in One Bearing," J. Mech. Eng. Sci., **4**(1), pp. 85–93.
[4] Black, H. F., 1968, "Interaction of a Whirling Rotor With a Vibrating Stator Across a Clearance Annulus," J. Mech. Eng. Sci., **10**(1), pp. 1–12.
[5] Childs, D. W., 1979, "Rub Induced Parametric Excitation in Rotors," ASME J. Mech. Des., **10**, pp. 640–644.
[6] Muszynska, A., 1984, "Partial Lateral Rotor to Stator Rubs," Paper C281/84, *Proc. of 3rd Int. Conf. Vibrations in Rotating Machinery*, University of York, UK, Mechanical Engineering Publications Limited, London, pp. 327–335.
[7] Ehrich, F. F., 1988, "High Order Subharmonic Response of High Speed Rotors in Bearing Clearance," ASME J. Vib., Acoust., Stress, Reliab. Des., **110**, pp. 9–16.
[8] Gonsalves, D. H., Neilson, R. D., and Barr, A. D. S., 1995, "A Study of the Response of a Discontinuously Nonlinear Rotor System," Nonlinear Dyn., **7**, pp. 451–470.
[9] Wang, X., and Noah, S. T., 1998, "Nonlinear Dynamics of a Magnetically Supported Rotor on Safety Auxiliary Bearings," ASME J. Vibr. Acoust., **120**, pp. 596–606.
[10] Wu, F., and Flowers, G. T., 1993, "An Experimental Study of the Influence of Disk Flexibility and Rubbing on Rotordynamics," *Proc. of ASME Conf. Vibrations in Rotating Systems*, ASME, New York, DE-Vol. 60, pp. 19–26.
[11] Lawen, J. L., and Flowers, G. T., 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.
[12] Lawen, J. L., and Flowers, G. T., 1999, "Interaction Dynamics Between a Flexible Rotor and an Auxiliary Clearance Bearing," ASME J. Vibr. Acoust., **121**, pp. 183–189.
[13] Xie, H., Flowers, G. T., Feng, L., and Lawrence, C., 1999, "Steady-State Dynamic Behavior of a Flexible Rotor With Auxiliary Support From a Clearance Bearing," ASME J. Vibr. Acoust., **121**, pp. 78–83.
[14] Fumagalli, M., Varadi, P., and Schweitzer, G., 1994, "Impact Dynamics of High Speed Rotors in Retainer Bearings and Measurement Concepts," *Proc. of 4th Int. Symp. Magnetic Bearings*, ETH Zurich, vdf Hochschulverlag AG an der ETH, Zurich, pp. 239–244.
[15] Fumagalli, M., and Schweitzer, G., 1996, "Measurements on a Rotor Contacting its Housing," *Proc. of 6th Int. Conf. Vibrations in Rotating Machinery*, University of Oxford, UK, Mechanical Engineering Publications Limited, London, Paper No. C500/085/96, pp. 779–788.
[16] Kirk, R. G., Swanson, E. E., Kavarana, F. H., and Wang, X., 1994, "Rotor Drop Test Stand for AMB Rotating Machinery, Part 1: Description of Test Stand and Initial Results," *Proc. of 4th Int. Symp. Magnetic Bearings*, ETH Zurich, vdf Hochschulverlag AG an der ETH, Zurich, pp. 207–212.
[17] Kirk, R. G., 1999, "Evaluation of AMB Turbomachinery Auxiliary Bearings," ASME J. Vibr. Acoust., **121**, pp. 156–161.
[18] Schmied, J., and Pradetto, J. C., 1992, "Behavior of a One Ton Rotor Being Dropped into Auxiliary Bearings," *Proc. of 3rd Int. Symp. Magnetic Bearings*, Alexandria, VA, Technomic Publishing, Lancaster, PA, pp. 145–156.
[19] Isaksson, J. L., 1996, "Unbalance Response Analysis of a Two-Mass Rotor System Including Rubbing Effects," *Proc. of 6th Int. Conf. Vibrations in Rotating Machinery*, University of Oxford, UK, Mechanical Engineering Publications Limited, London, Paper No. C500/021/96, pp. 395–403.
[20] Markert, R., and Wegener, G., 1998, "Transient Vibration of Elastic Rotors in Retainer Bearings," *Proc. of ISROMAC-7*, Hawaii, Bird Rock Publishing House, Honolulu, pp. 764–774.
[21] Wegener, G., Markert, R., and Pothmann, K., 1998, "Steady-State-Analysis of a Multi-Disk or Continuous Rotor with One Retainer Bearing," *Proc. of 5th IFToMM Conf. Rotor Dynamics*, Darmstadt, Vieweg Publishers, Braunschweig, pp. 816–828.
[22] Wegener, G., and Markert, R., 1998, "Influence of Contact and Impacts on the Dynamics of an Elastic Rotor With an Elastic Retainer Bearing," *Proc. of Euromech Colloquium on Dynamics of Vibro-Impact Systems*, Springer Pub-

lishing, Heidelberg, pp. 89–98.
[23] Ecker, H., 1997, "Steady State Orbits of an AMB-Supported Rigid Rotor Contacting the Backup Bearings," *Proc. of MAG'97*, Alexandria, VA, Technomic Publishing, Lancaster, PA, pp. 129–138.
[24] Ecker, H., 1998, "Nonlinear Stability Analysis of a Single Mass Rotor Contacting a Rigid Auxiliary Bearing," *Proc. of 5th IFToMM Conf. Rotor Dynamics*, Darmstadt, Vieweg Publishers, Braunschweig, pp. 790–801.
[25] Cuesta, E. N., Rastelli, V. R., Medina, L. U., Montbrun, N. I., and Diaz, S. E., 2002, "Non-Linear Behaviors in the Motion of a Magnetically Supported Rotor on the Catcher Bearing During Levitation Loss, an Experimental Description," ASME Turbo Expo, Amsterdam, ASME Paper No. GT-2002-30293.
[26] Cuesta, E. N., Medina, L. U., Rastelli, V. R., Montbrun, N. I., and Diaz, S. E., 2003, "A Simple Kinematic Model for the Behavior of a Magnetically Levitated Rotor Operating in Overload Regime," ASME Turbo Expo, Atlanta, ASME Paper No. GT-2003-38024.
[27] Keogh, P. S., Cole, M. O. T., Sakinkaya, M. N., and Burrows, C. R., 2002, "On the Control of Synchronous Vibration in Rotor/Magnetic Bearing Systems Involving Auxiliary Bearing Contact," ASME Turbo Expo, Amsterdam, ASME Paper No. GT-2002-30292.
[28] Keogh, P. S., and Cole, M. O. T., 2003, "Rotor Vibration with Auxiliary Bearing Contact in Magnetic Bearing Systems, Part I: Synchronous Dynamics," Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci., **217**, pp. 377–392.
[29] Cole, M. O. T., and Keogh, P. S., 2003, "Rotor Vibration With Auxiliary Bearing Contact in Magnetic Bearing Systems, Part II: Robust Synchronous Control for Rotor Position Recovery," Proc. Inst. Mech. Eng., Part C: J. Mech. Eng. Sci., **217**, pp. 393–409.
[30] Chen, H. M., Walton, J. F., II, and Heshmat, H., 2002, "Test of a Zero Clearance Auxiliary Bearing," *Proc. of MAG'97*, Alexandria, VA, Technomic Publishing, Lancaster, PA, pp. 111–119.
[31] Swanson, E. E., and Heshmat, H., 2002, "Oil-Free Foil Bearings as a Reliable, High Performance Backup Bearing for Active Magnetic Bearings," ASME Turbo Expo, Amsterdam, ASME Paper No. GT-2002-30291.

Stainless Steels With Improved Oxidation Resistance for Recuperators

Bruce A. Pint
Oak Ridge National Laboratory,
Metals and Ceramics Division,
Oak Ridge, TN 37831-6156
e-mail: pintba@ornl.gov

New materials are being evaluated to replace type 347 stainless steel in microturbine recuperators operating at higher temperatures in order to increase the efficiency of the microturbine. Commercial alloys 120 and 625 are being tested along with potentially lower cost substitutes, such as Fe-20Cr-25Ni and Fe-20Cr-20Ni. Long-term testing of these materials at 650–700 °C shows excellent corrosion resistance to a simulated exhaust gas environment. Testing at 800 °C has been used to further differentiate the performance of the various materials. The depletion of Cr from foils of these materials is being used to evaluate the rate of attack. Although those alloys with the highest Ni and Cr contents have longer lives in this environment, lower alloyed steels may have sufficient protection at a lower cost. [DOI: 10.1115/1.2056531]

Introduction

The August 2003 blackout in eastern North America exposed the critical problem of overloaded transmission lines. Utilities, faced with deregulation, have not sufficiently invested in the power transmission grid and any upgrades will require significant time and resources. An alternative solution to this critical problem is distributed generation. Producing electricity closer to the user reduces the load on the transmission grid. Among the possible generation technologies, small (25–300 kW) gas turbines or microturbines [1,2] can operate on multiple fuels while producing lower emissions than similar-sized reciprocating engines. Even before this recent crisis, microturbines were finding applications in manufacturing and service industries concerned with power quality, using sour gas from oilfields and with producers of biogas, including landfills, wastewater treatment facilities, and farms [3–5]. One of the added benefits of on-site generation is that waste heat from the turbine can be used for heating water, climate control, or a manufacturing process (such as drying).

Two current problems facing microturbines are cost and efficiency (currently $<30\%$). Because of their small size and simple, single-shaft design, there are few opportunities to improve the efficiency of microturbines except by increasing the efficiency of the recuperator (i.e., heat exchanger) and by increasing the turbine inlet temperature, which also would increase the temperature of the gas entering the recuperator [6]. Thus, the recuperator, which already is the most expensive part of the system, is one focus area of the U.S. Department of Energy's Distributed Energy and Electrical Reliability Program to improve the efficiency of microturbines [7].

There are a wide range of recuperator designs, but most employ some thin-walled components in order to maximize specific heat transfer [6,8]. Type 347 stainless-steel foil (75–125 μm, 3–5 mil thick) has been the material of choice for recuperators because of its combination of creep and corrosion resistance. However, as recuperator temperatures have moved above $\approx$600 °C (1100 °F), durability problems have been encountered. Thus, a materials testing and development program was begun at Oak Ridge National Laboratory (ORNL) to address materials issues for higher efficiency recuperators [8–20]. Although it has been possible to improve the creep properties of type 347 stainless steel by modifying the processing and chemistry [9,17,19], it has not been possible to significantly improve its corrosion resistance. Like other Fe-based alloys, type 347 stainless steel is subject to accelerated attack (AA) when exposed to steam or water vapor at $\geq$650 °C [10,11,13,15,16,19–27], which is present in turbine exhaust gas. Although this type of attack can be tolerated in, for example, thick-walled steam tubing, thin cross sections are quickly consumed because of their limited reservoir of Cr, which is needed to form a protective external scale. Thus, in order to improve efficiency and increase the operating temperature of the recuperator above 600 °C while maintaining durability of $\approx$40,000 hr, a new foil material will be required. Corrosion will be a primary issue in determining durability.

The performance of two groups of alloys will be presented in this paper. The first group are more highly alloyed commercial alloys, a Fe-based alloy, 120 and a Ni-based alloy, 625. These alloys are 3–5 times more expensive than type 347 stainless steel. The second group are stainless steels, not currently produced commercially as foil, that have slightly higher Cr and Ni contents than type 347 stainless steel. These alloys are hoped to result in a $<50\%$ higher material cost. As expected, the first group has better corrosion resistance than the second group. However, because cost is a major issue with the implementation of microturbines, the second group may be key to allowing recuperator operating temperatures to move up to 700 °C. Thus, this group is the main focus of the ORNL alloy development program which is now moving into mill scale foil production.

Experimental Procedure

The chemical compositions of selected alloys examined in this study are listed in Table 1. Some of the materials were obtained from commercial vendors in foil form, while others were obtained in thicker sections and then hot and cold rolled at ORNL to $\approx$100 μm thickness with average grain sizes given in Table 1. Model alloys were vacuum induction melted at ORNL and hot and cold rolled to 1.25 mm sheet. After the final cold-rolling step, the sheets were annealed in Ar for 2 min at 1000 °C. Selected alloys were then rolled to foil under similar conditions as used for the commercial alloys.

Foil specimens ($\approx$12 mm $\times$ 18 mm $\times$ 100 μm) were tested in the as-rolled conditions and similar-sized sheet specimens were polished to a 600 SiC grit finish. The specimens were cleaned in

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 13–17, 2004, Paper No. GT2004-53627.

Table 1 Alloy chemical compositions (weight percent) and average grain sizes (microns) of the 100 μm foil materials

	Cr	Ni	Mn	Si	Other	Grain Size (μm)
Type 347	17.8	9.9	1.6	0.5	0.5Nb	5
20/25/Nb	20.3	24.7	1.0	0.4	1.5Mo,0.2Nb	16
120	24.7	37.6	0.7	0.2	0.6Nb, 0.3Mo, 0.2Cu	23
625	23.1	63.8	0.04	0.2	8.9Mo, 4Nb, 3Fe	12
Fe-20/20+MS	19.8	19.8	1.7	0.3	0.01Ce	25
Fe-20/20+4Mn	20.9	20.8	3.8	0.2	0.3Nb,0.3Cu,0.3Mo	10
Fe-15/15+Al	15.1	15.8	4.8	0.2	3.8Al, 4Cu, 0.4Nb	30

acetone and methanol prior to oxidation and mass changes were measured using a Mettler-Toledo model AG245 balance. Exposures were 100 hr cycles at 650, 700, or 800 °C.

Oxidation exposures in humid air were conducted by flowing the gas at 850 cc/min through an alumina tube that was inside a resistively heated tube furnace. Distilled water was atomized into the flowing gas stream above its condensation temperature and heated to the reaction temperature within the alumina tube. Water was collected and measured after flowing through the tube to calculate its concentration and calibrate the amount of injected water. A water content of 10 vol. % was used for these experiments. For testing in air, the alumina furnace tubes were not sealed. Up to 40 specimens were positioned in alumina boats in the furnace hot zone so as to expose the specimen faces to the flowing gas. After oxidation, selected specimens were Cu-plated and sectioned for metallographic analysis and electron probe microanalysis (EPMA) to determine Cr depletion.

Results

Overview of Oxidation in Water Vapor. The effect of water vapor currently is being studied by several research groups but there is no widely accepted mechanism for its role in reducing the corrosion resistance of chromia-forming stainless steels. Figure 1 gives a schematic representation of the current understanding about the role of water vapor on the oxidation of an austenitic stainless steel like type 347. In dry laboratory air, a protective Cr-rich oxide scale forms on the surface of type 347 stainless steel and thickens with time following a parabolic rate law. Foil specimens (100 μm thick) have been exposed for over 40,000 hr at 650 °C and 25,000 h at 700 °C to confirm the long-term behavior. The addition of water vapor to the gas stream leads to effects such as faster oxide scale growth and increased evaporation of Cr_2O_3 from the scale as $CrO_2(OH)_2$ [25]. The net specimen mass change ($\Delta M_{specimen}$) measured after exposure in this type of environment can then be simply expressed as

$$\Delta M_{specimen} = \Delta M_{oxide\ growth} - \Delta M_{evaporation} - \Delta M_{spallation} \quad (1)$$

Thus, if the loss due to evaporation is high enough, then a net mass loss is measured without any oxide scale spallation. Both mechanisms result in an increased Cr consumption rate in the metal compared to oxidation in dry air. Because the diffusion rate of Cr in the metal is not fast enough relative to the consumption rate, a Cr depleted region forms in the metal near the surface [27–29]. After some incubation time [10,23,27], nodules of FeO_x begin to form. Whether the nodules form due to the Cr depletion in the adjacent metal substrate or some other mechanism has not been determined. However, it has been demonstrated that increasing the Cr and/or Ni content in the alloy increases the time to AA or prevents it from occurring [10,15,16,27]. After the nodules form, they tend to grow laterally (Fig. 1(*c*)) until a complete layer of Fe-rich oxide forms [30]. This results in a large specimen mass gain. An underlying layer of $(Fe,Cr)O_x$ also is observed. When these layers thicken sufficiently, they tend to spall during thermal cycling ($\Delta M_{spallation}$), which can lead to large specimen mass losses.

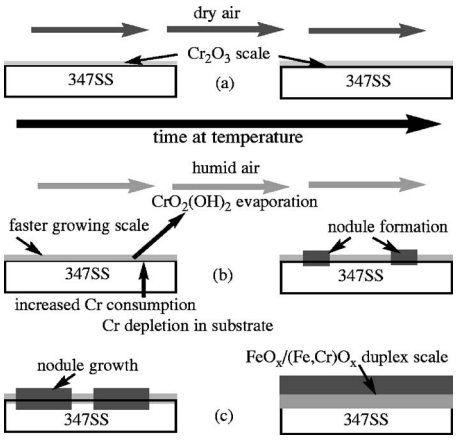


Fig. 1 Schematic of oxidation of type 347 stainless steel in dry air (*a*) and humid air (*b,c*). With minimal water vapor in the environment (*a*), a protective scale forms that thickens with time at temperature. With the addition of water vapor (*b*), there is an increase in the scale growth rate and increased evaporation of Cr_2O_3 as $CrO_2(OH)_2$. These result in increased Cr consumption and depletion of Cr in the substrate near the surface. With continued exposure, nodules of FeO_x begin to form and grow with time (*c*). Eventually, the nodules grow together forming a duplex scale with an outer FeO_x layer and an inner $(Fe,Cr)O_x$ layer.

Oxidation Results at 650 °C. Mass gain data for various alloy foils at 650 °C in humid air are shown in Fig. 2. Several versions of type 347 stainless steel all showed AA before 2000 hr following the mechanism shown in Fig. 1. Foils with higher Cr contents,

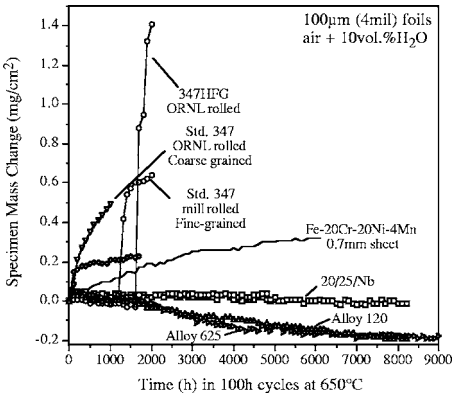


Fig. 2 Specimen mass gains for various foil (100 μm thick) materials during 100 hr cycles in humid air at 650 °C. Several versions of type 347 stainless steel showed AA after <2000 hr of exposure, whereas more highly alloyed materials have not shown AA after 8000 hr.

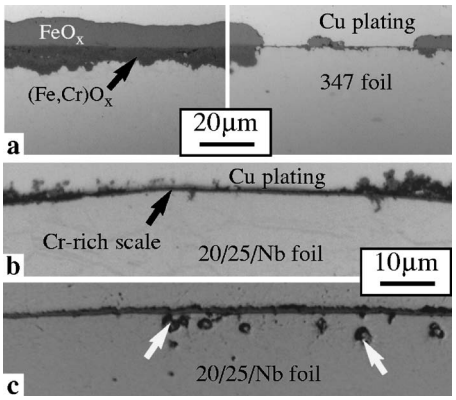


Fig. 3 Light microscopy of polished cross sections of 100 μm foils after exposure in humid air: (*a*) commercial 347 foil after 2000 hr at 650 °C, (*b*) 20/25/Nb foil after 5000 hr at 650 °C, and (*c*) 20/25/Nb foil after 5000 hr at 700 °C

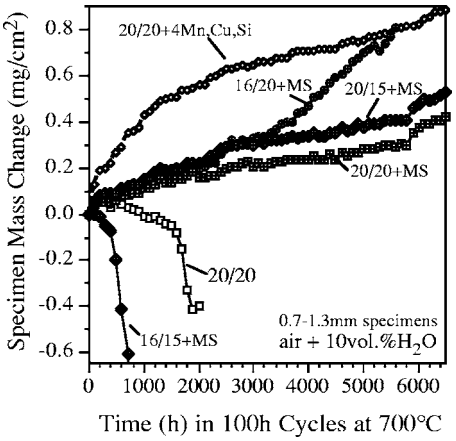


Fig. 5 Specimen mass changes for model Fe-Cr-Ni alloys (specified by their Cr/Ni contents) during 100 hr cycles at 700 °C in humid air. Designations with +MS indicate Mn and Si additions.

20/25/Nb, 120, 625 all showed low mass gains after more than 8000 hr of testing. The small mass losses for the latter two alloys are attributed to volatilization of $CrO_2(OH)_2$ rather than scale spallation. Sheet specimens (0.7 mm thick) of a developmental alloy, Fe-20Cr-20Ni-4Mn also are being evaluated under these conditions; a specimen has not shown AA after a 6000 hr exposure. Its higher mass gain compared to the higher alloyed foils is attributed to its high Mn content (Table 1).

An example of the scale morphology formed by AA is shown in cross-section in Fig. 3(*a*). In some regions of the specimen, a complete duplex layer has formed while in other areas nodules of various sizes are present. This was illustrated schematically in Fig. 1. In contrast, one specimen of 20/25/Nb foil was stopped after 5000 hr for characterization and showed a thin protective scale (Fig. 3(*b*)).

Oxidation Results at 700 °C. At higher test temperatures, differences in corrosion behavior are observed because of the multiple temperature-dependent mechanisms responsible for AA: diffusion in the alloy, growth of the external scale, and volatilization of $CrO_2(OH)_2$. The biggest difference between the 650 and 700 °C results is that AA was not observed on mill rolled 347 foil until after 7500 hr (Fig. 4). Type 347 foil rolled at ORNL with a coarser grain size began to show AA immediately [16]. As at 650°C, the foil specimens of higher alloyed materials all show low mass gains or slight mass losses. Again, a specimen mass loss is attributed to evaporation of the Cr-rich reaction product being greater than the mass loss due to oxide scale formation (Eq. (1)). Two versions of alloy 120 are being tested at this temperature. One foil was made commercially and the other other was rolled at ORNL. The ORNL-rolled 120 foil had a 13 μm average grain size, which was somewhat smaller than the commercially rolled foil (Table 1). However, unlike the 347 foils, the 120 foils are showing similar performance after >8000 hr of testing. Testing is continuing on the foils with low mass gains. However, one specimen of 20/25/Nb foil was removed from the test for characterization after 5000 hr. The thin surface oxide is shown in Fig. 3(*c*). Additional characterization is being performed to examine the features in the metal below the surface scale (arrows). These could be voids or some type of internal oxidation.

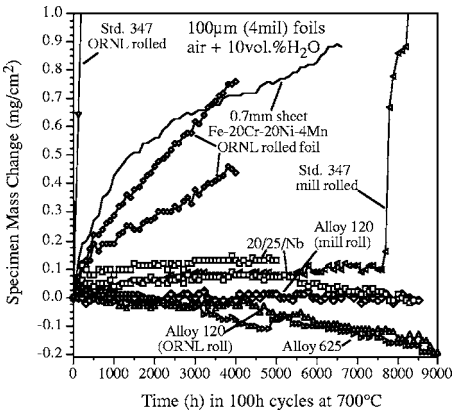


Fig. 4 Specimen mass gains for various foil (100 μm thick) materials during 100 hr cycles in humid air at 700 °C

Developmental alloys also are being evaluated at 700 °C. The mass gains for both foil and sheet specimens of a Fe-20Cr-20Ni-4Mn are shown in Fig. 4. As at 650 °C, the higher mass gains for this material are attributed to faster scale growth [20], primarily due to the higher Mn content. The mass gain also may be higher because a Mn-rich scale is believed to suppress volatilization [31], which is reducing the mass gains of the other alloys.

In order to better understand the role of alloy composition, especially minor alloy additions, testing also is continuing on sheet specimens (1.2 mm thick) of model alloys. Figure 5 shows results for several model alloys with and without minor alloy additions. A specimen of Fe-20Cr-20Ni showed AA after <2000 hr. Additions of Mn and Si have been found to be the most critical in preventing AA. However, for a base alloy of Fe-16Cr-15Ni, the Mn and Si additions did not prevent AA [16,20]. At longer times, the specimen of Fe-16Cr-20Ni+Mn/Si is beginning to show increased mass gain typical of AA. Specimens with 20Cr and Mn and Si have not shown AA after more than 5000 hr of testing. The mass gain for the specimen of Fe-20Cr-20Ni-1.7Mn-0.2Si is much lower than that for the developmental alloy with 3.8%Mn and other elements to increase the creep strength. Further work is needed to optimize the composition of this alloy.

Oxidation Results at 800 °C. Testing also is being conducted at higher temperatures in order to accelerate the corrosion testing and because higher temperature recuperator materials may be needed [6]. This temperature is beyond the capability of type 347 stainless steel from a creep strength standpoint and AA was observed for all 347 foils in less than 1000 hr (Fig. 6). Several 20/25/Nb foil specimens have been run with varying results. In one case, the mass gain remained relatively low, whereas in the

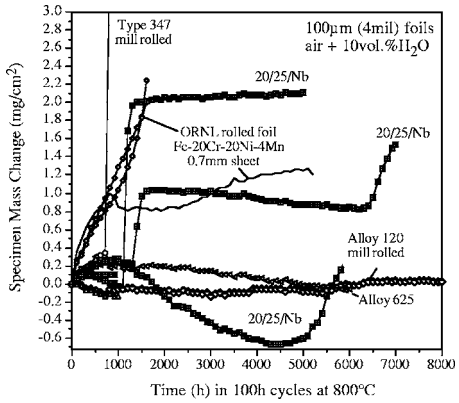


Fig. 6 Specimen mass gains for various foil (100 μm thick) materials during 100 hr cycles in humid air at 800 °C

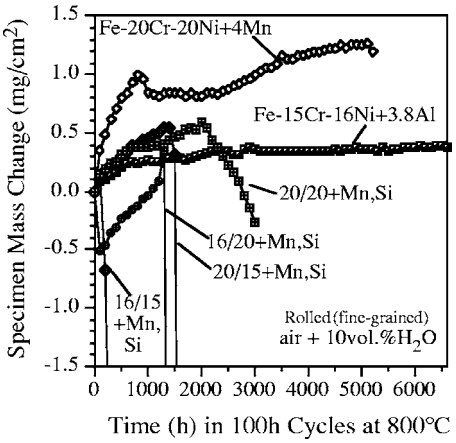


Fig. 8 Specimen mass gains for various model alloys (1.2 mm thick specimens) with different Cr/Ni contents and additions of Mn and Si during 100 hr cycles in humid air at 800 °C. Adding 3.8% Al significantly improved the resistance to AA.

other cases discrete jumps in the mass gain were observed followed by low mass gains or slight mass losses. These rapid increases appear to be localized nodule formation where the nodules did not continue to grow. One foil specimen was removed from the test after 5000 hr, while two others have been run to 6000 and 7000 hr, respectively, where they began to exhibit a continuous increase in mass gain suggesting the onset of AA. Figure 7(*a*) shows the 20/25/Nb foil specimen after 5000 hr. Some oxide nodules were observed, but most of the scale was uniform and 3–4 μm thick. Figures 7(*b*) and 7(*c*) show the 20/25/Nb specimens exposed for 6000 hr. In some areas, large oxide nodules were observed and, based on the mass gain increase over the last 1500 hr of exposure, these nodules were growing. In areas without nodules, the scale was thicker and more convoluted than after 5000 hr, Fig. 7(*c*). Despite the relatively low mass gain for this specimen, the formation of a relatively thick scale suggests that significant mass losses occurred due to evaporation (see Eq. (1)). The mass increase for a $\sim$6 μm thick scale should be 1 mg/cm^2.

Foil specimens of alloy 625 and 120 have not exhibited any rapid mass increases during testing at 800 °C. Instead, they have shown relatively low mass gains or slight mass losses typical of the combination of scale growth and evaporation. The foil specimen of alloy 120 has been tested past 7500 hr without showing any signs of AA as was observed for 20/25/Nb. The longer time before AA for this alloy was expected because of its higher Cr and Ni contents. Exposure of the alloy 625 specimen was stopped after 6000 hr for characterization. A cross section of this foil is shown in Fig. 7(*d*). Again, a substantial oxide scale has formed suggesting that significant loss of Cr occurred in order to result in a net mass loss after 6000 hr (Fig. 6).

Sheet specimens of the developmental Fe-20Cr-20Ni-4Mn alloy also are being tested (Fig. 6). The sheet specimen had a higher rate of mass gain during the first 1000 hr than the foil materials, but at longer times the rate of increase has decreased. Foil specimens of this material have only been tested for 2000 hr at this time and have shown a relatively high mass gain rate.

Testing sheet specimens of model alloys at 800 °C has helped differentiate their performance (Fig. 8). The specimen with Fe-20Cr-20Ni has shown a longer time to AA than any of the other materials with lower Cr or Ni contents and Mn and Si additions. Alumina scales are known to be more resistant to humid environments because Al_2O_3 is less susceptible to hydroxide formation. This has been shown for both alumina-forming alloys [11,13] and aluminide coatings [32]. One alloy has been made with a 3.8% Al addition and has shown better resistance to AA than the other alloys. Because the corrosion resistance is due to the Al, the Cr and Ni contents were lowered to 15–16 % (Table 1).

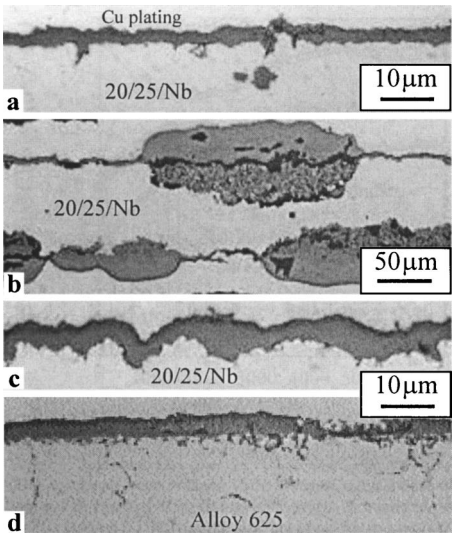


Fig. 7 Light microscopy of polished cross sections of ORNL-rolled 100 μm foils after exposure in humid air at 800 °C: (*a*) 20/25/Nb for 5000 hr (*b*,*c*) 20/25/Nb for 6000 hr, and (*d*) alloy 625 for 6000 hr

Characterization by EPMA. Because of the inconsistencies between the mass gain data in Fig. 6 and the cross sections in Fig. 7, additional characterization was performed on the foil specimens to determine the extent of Cr depletion. Figure 9 shows Cr EPMA profiles for several 20/25/Nb specimens. The specimen exposed for 6000 hr in humid air (squares) shows a significant Cr loss along with an overall loss in gage thickness to 80–85 μm from 100 μm. (The mass gain for this specimen is shown in Fig. 6 and the cross section in Figs. 7(*b*) and 7(*c*).) Overall, $\sim$45% of the initial Cr reservoir has been consumed based on the starting Cr content and specimen thickness. (Some of the data points showing higher Cr contents are attributed to Cr-rich σ phase or carbide particles.) In contrast, after a 5000 hr exposure in humid air (circles in Fig. 9), the Cr content is significantly higher with only $\sim$28% of the Cr consumed. One interesting feature of this profile is the strong Cr depletion near both surfaces. This may be the precursor to AA that has already occurred in the specimen exposed for an additional 1000 hr. In order to illustrate the role of

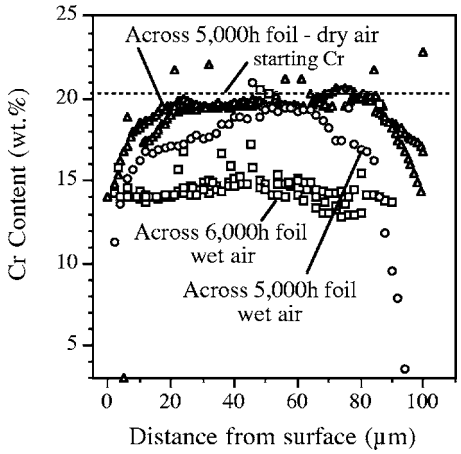


Fig. 9 Electron microprobe Cr profiles across 20/25/Nb foils after various exposures in humid and dry air at 800 °C. The addition of water vapor to the environment significantly increased the Cr consumption rate from the foils.

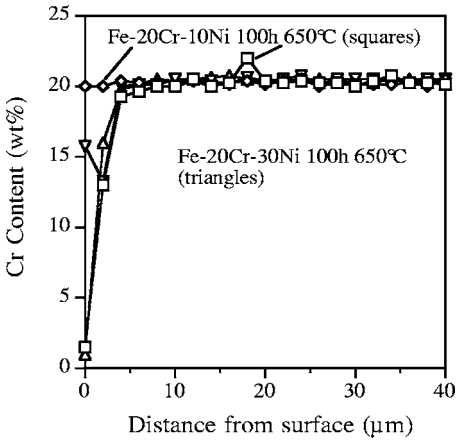


Fig. 11 Electron microprobe Cr profiles into model alloys after 100 hr exposures in humid air at 650 °C. The different Ni contents did not significantly change the profiles.

water vapor, a foil exposed for 5000 hr in laboratory air also was sectioned and profiled (triangles in Fig. 9). This specimen showed only a minimal Cr depletion, mainly near the surfaces. Only 9% of the Cr was consumed from the estimated starting reservoir, suggesting that the addition of 10% water vapor increased the Cr consumption rate by a factor of 3 at 800 °C. Clearly, this material could not be expected to meet any reasonably long recuperator durability goal at this temperature.

Because of the high Cr depletion observed in 20/25/Nb foils, profiles were made of 625 foils to see if similar depletion was occurring. Figure 10 shows profiles for two different 625 foils that were 120–130 μm thick. The profiles are more erratic in this material because of the formation of intermetallic phases rich in Mo or Nb which increases the variability of the Cr content (e.g., higher Cr content in the phases low in Mo and Nb). For the specimen exposed for 6000 hr in humid air at 800 °C, Cr depletion was found, especially near the surfaces. Based on the starting Cr reservoir, approximately 25% of the Cr was consumed. In contrast, a specimen exposed for 5000 hr in laboratory air showed much less depletion, especially near the metal-scale interface and only 12% of the Cr was consumed. A third specimen (not shown) exposed for only 1000 hr in humid air showed only 8% Cr consumption.

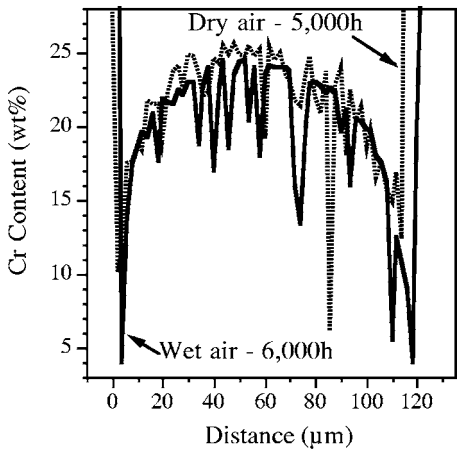


Fig. 10 Electron microprobe Cr profiles across 625 foils after various exposures at 800 °C. The addition of water vapor to the environment increased the Cr consumption rate from the foils.

Studying the Cr depletion appears to be a useful method for quantifying the rate of attack and studying the role of water vapor. One area of interest has been the role of Ni in improving resistance to AA. Initial work has compared the Cr depletion profiles in two model alloys Fe-20Cr-10Ni and Fe-20Cr-30Ni after 100 hr at 650 °C (Fig. 11). The profiles showed a large amount of variability after this short exposure, but no clear difference between the two Ni contents. Further work is being conducted to study this effect.

Discussion

If microturbines are to achieve higher electrical generation efficiencies, it is very likely that higher turbine inlet temperatures will be required. This increase will require the recuperator to operate at a higher temperature than type 347 can withstand, and a new material or materials will be required. The type of recuperator will determine the exact properties needed for this material but most use thin-walled components that need durability on the order of 40,000 hr. Thus, oxidation resistance is a primary requirement.

Laboratory testing of various foils has shown that there are materials available that can withstand much higher temperatures than type 347 stainless steel. Certainly a Ni-based alloy, like alloy 625, will meet most creep and corrosion requirements for a recuperator. Since AA is associated with the rapid oxidation of Fe, it is unclear if water-vapor-assisted corrosion is even a major concern for Ni-based alloys. A highly alloyed material, like alloy 120, also is very corrosion resistant and has not shown any AA in any of the tests conducted to date. However, the important secondary issue then becomes cost.

Ideally, the lowest cost solution would be a minor alloy addition or small change to type 347 stainless steel. Some improvement in 650 °C resistance to AA has been observed in modified composition 347 foils with improved creep strength [19]. However, the corrosion performance at 700 °C was not as promising.

Based on model alloy results, boosting the Cr and Ni contents into the 20% range provides the most likely opportunity to significantly improve the oxidation resistance at a modest cost increase. Therefore, most of the development work has focused on Fe-20Cr-20Ni and Fe-20Cr-25Ni. Alloys with these base compositions have shown excellent corrosion resistance at 650–700 °C in humid air for test times >5000 hr.

The primary means of assessing performance in these long-term tests has been specimen mass change. Because of the extremely

rapid mass gain increase associated with AA, this has been a good indicator of the difference between protective scale formation and AA. However, as the emphasis changes to more corrosion-resistant compositions, reliance on mass gain may not be sufficient to truly evaluate performance and especially to predict the useful lifetime of the material. (For a foil material, the component lifetime likely will be similar to the incubation time before the onset of AA because, once AA begins, the rate of metal consumption will be greatly increased, e.g., Fig. 9.) Lifetime predictions will require data extrapolation because the long-term goal is a recuperator lifetime of $\geqslant 40,000$ hr.

Comparing the specimen mass changes (Fig. 6) to the scales formed at 800 °C (Fig. 7) indicates an inability of mass change to truly assess performance. Measuring the residual Cr profiles in foil specimens appears to be a more reliable means of assessing performance and predicting lifetime and has been used in other studies [27]. These results for foils exposed at 800 °C show depletion to a degree, which brings into question whether any iron-based, chromia-forming alloy can be used in this application at 800 °C. Thus, the development of an Fe-based austenitic alumina-forming alloy may be an excellent long-term goal for this program.

Regarding the 800 °C depletion rates, it is important to emphasize that these foils were exposed under very specific conditions. The proposed mechanism for corrosion in this environment suggests that changing the gas composition and velocity will alter the rate of Cr consumption. Exhaust gas should have lower water vapor content, which would presumably decrease the evaporation rate, but a much higher gas velocity in the recuperator should increase the rate. These are issues that need to be explored in the ORNL microturbine test facility [14]. Ultimately, knowing the Cr depletion rate as a function of temperature, gas composition, and velocity will allow for the best lifetime predictions and alloy selection. Also, determining the Cr depletion rate will be the best way of optimizing the additions of Mn and Si. Current reliance on mass change may produce misleading conclusions. However, Cr depletion measurements are easier to quantify on foil specimens than on sheet specimens where the Cr reservoir is much larger.

Summary

The accelerated corrosion attack associated with the presence of water vapor in exhaust gas limits the temperature at which type 347 stainless steel can be used for a microturbine recuperator. Foil specimens of several candidate alloys are being studied in a long-term testing program. These alloys generally show better corrosion resistance in these environments at 650–800 °C. A representative Ni-base alloy, 625, shows excellent corrosion resistance in foil form as did alloy 120 (Fe-37Ni-25Cr). However, these materials may be too expensive for microturbines. Therefore, less expensive alternatives based on Fe-20Cr-20Ni and Fe-20Cr-25Ni are being explored in the ORNL alloy development program. Initial results show excellent corrosion resistance for these materials in humid air. However, the Cr depletion rates at 800 °C suggest that these materials will be limited to $\approx$700–725 °C if 40,000 hr durability is required. Alloys in this composition range are being selected for mill-scale fabrication and eventual engine testing.

Acknowledgment

The author would like to thank J. D. Vought, G. W. Garner, K. S. Reeves, H. Longmire, and L. Walker at ORNL for assistance with the experimental work and P. J. Maziasz, I. G. Wright, and P. F. Tortorelli for comments on the manuscript. This research was sponsored by the U.S. Department of Energy, Distributed Energy and Electrical Reliability Program under contract No. DE-AC05-00OR22725 with UT-Battelle, LLC.

References

[1] Watts, J. H., 1999, "Microturbines: A New Class of Gas Turbine Engine," Global Gas Turbine News, **39**(1), pp. 4–8.
[2] Hamilton, S. L., 1999, "Microturbines Poised to Go Commercial," Mod. Power Syst., **19**(9), pp. 21–22.
[3] Industry News, 1999, "PanCanadian Puts Four High-Tech Microturbines to Work in Alberta Oil Fields," J. Can. Pet. Technol., **38**(12), p. 64.
[4] Industry News, 2001, "Two California Companies Set an Example With Use of Microturbines," Pollut. Eng., **33**(8) p. 7.
[5] Field, K., 2002, "The Inner Workings of Microturbines," BioCycle, **43**(12), p. 13.
[6] McDonald, C. F., 2003, "Recuperator Considerations for Future Higher Efficiency Microturbines," Appl. Therm. Eng., **23**, pp. 1463–1487.
[7] Office of Energy Efficiency and Renewable Energy, Office of Power Technologies, March 2000, "Advanced Microturbine Systems Program, Plan for Fiscal Years 2000 Through 2006," U. S. Department of Energy, Washington, DC.
[8] Omatete, O., Maziasz, P. J., Pint, B. A., and Stinton, D. P., 2000, "Recuperators for Advanced Microturbines," Report No. ORNL/TM-2000/304, Oak Ridge National Laboratory, Oak Ridge, TN.
[9] Maziasz, P. J., Swindeman, R. W., Montague, J. P., Fitzpatrick, M., Browning, P. F., Grubb, J. F., and Klug, R. C., 1999, "Improved Creep-Resistance of Austenitic Stainless Steel for Compact Gas Turbine Recuperators," Mater. High. Temp., **16**, pp. 207–212.
[10] Pint, B. A., and Rakowski, J. M., 2000, "Effect of Water Vapor on the Oxidation Resistance of Stainless Steels," NACE Corrosion 2000, Orlando, NACE Paper No. 00-259.
[11] Pint, B. A., Swindeman, R. W., More, K. L., and Tortorelli, P. F., 2001, "Materials Selection for High Temperature (750–1000 °C) Metallic Recuperators for Improved Efficiency Microturbines," International Gas Turbine and Aeroengine Congress and Exhibition, New Orleans, ASME Paper No. 2001-GT-445.
[12] Harper, M. A., Maziasz, P. J., Smith, G. D., and Swindeman, R. W., 2001, "Materials Selection for High Temperature Metal Recuperators," International Gas Turbine and Aeroengine Congress and Exhibition, New Orleans, ASME Paper No. 2001-GT-540.
[13] Pint, B. A., More, K. L., and Tortorelli, P. F., 2002, "The Effect of Water Vapor on Oxidation Performance of Alloys Used in Recuperators," Int. Gas Turbine and Aeroengine Congress and Exhibition, Amsterdam, June, ASME Paper No. 2002-GT-30543.
[14] Lara-Curzio, E., Maziasz, P. J., Pint, B. A., Stewart, M., Hamrin D., Lipovich, N., and DeMore, D., 2002, "Test Facility for Screening and Evaluating Candidate Materials for Advanced Microturbine Recuperators," Int. Gas Turbine and Aeroengine Congress and Exhibition, Amsterdam, June, ASME Paper No. 2002-GT-30581.
[15] Pint, B. A., and Peraldi, R., 2003, "The Effect of Alloy Composition on the Performance of Stainless Steels in Exhaust Gas Environments," NACE Corrosion 2003, San Diego, March, NACE Paper No. 03-499.
[16] Pint, B. A., and Peraldi, R., 2003, "Factors Affecting Corrosion Resistance of Recuperator Alloys," Int. Gas Turbine and Aeroengine Congress and Exhibition, Atlanta, June, ASME Paper No. GT2003-38692.
[17] Maziasz, P. J., and Swindeman, R. W., 2003, "Selecting and Developing Advanced Alloys for Creep-Resistance for Microturbine Recuperator Applications," J. Eng. Gas Turbines Power, **125**, pp. 51–58.
[18] Maziasz, P. J., Pint, B. A., Swindeman, R. W., More, K. L. and Lara-Curzio, E., 2003, "Selection, Development and Testing of Stainless Steels and Alloys for High-Temperature Recuperator Applications," Int. Gas Turbine and Aeroengine Congress and Exhibition, Atlanta, June, ASME Paper No. GT2003-38762.
[19] Maziasz, P. J., Swindeman, R. W., Shingledecker, J. P., More, K. L., Pint, B. A., Lara-Curzio, E., and Evans, N. D., 2003, "Improving High-Temperature Performance of Austenitic Stainless Steels for Advanced Microturbine Recuperators," *Parsons 2003, Engineering Issues in Turbine Machinery, Power Plant and Renewables*, Maney, London, pp. 1057–73.
[20] Pint, B. A., 2004, "The Long-Term Performance of Model Austenitic Alloys in Humid Air," NACE Corrosion 2004, New Orleans, LA, March NACE Paper No. 04-530.
[21] Caplan, D., and Cohen, M., 1959, "High Temperature Oxidation of Chromium-Nickel Steels," Corrosion (Houston), **15**, pp. 57–62.
[22] Douglass, D. L., Kofstad, P., Rahmel, A., and Wood, G. C., 1996, "International Workshop on High Temperature Corrosion," Oxid. Met., **45**, pp.529–545.
[23] Shen, J., Zhou, L., and Li, T., 1997, "High-Temperature Oxidation of Fe-Cr Alloys in Wet Oxygen," Oxid. Met., **48**, pp. 347–356.
[24] Quadakkers, W. J., and Ennis, P. J., 1998, "The Oxidation Behavior of Ferritic and Austenitic Steels in Simulated Power Plant Service Environments," *Materials for Advanced Power Engineering 1998*, J. Leconte-Beckers, F. Schubert, and P. J. Ennis (Eds.), Forschungszentrum, Jülich, Germany, pp. 123–138.
[25] Asteman, H., Svensson, J.-E., Norell, M., and Johansson, L.-G., 2000, "Influence of Water Vapor and Flow Rate on the High-Temperature Oxidation of 304L: Effect of Chromium Oxide Hydroxide Evaporation," Oxid. Met., **54**, pp. 11–26.
[26] Henry, S., Galerie, A., and Antoni, L., 2001, "Abnormal Oxidation of

Stabilized Ferritic Stainless Steels in Water Vapor," Mater. Sci. Forum, **369-372** pp. 353–60.
[27] Rakowski, J. M., 2003, "The Oxidation of Metal Alloy Foils in the Presence of Water Vapor," Int. Gas Turbine and Aeroengine Congress and Exhibition, Atlanta, June, ASME Paper No. GT2003-38059.
[28] Gleeson, B., and Harper, M. A., 1998, "The Long-Term, Cyclic-Oxidation Behavior of Selected Chromia-Forming Alloys," Oxid. Met., **49**, pp. 373–99.
[29] Evans, H. E., Donaldson, A. T., and Gilmour, T. C., 1999, "Mechanisms of Breakaway Oxidation and Application to a Chromia-Forming Steel," Oxid. Met., **52**, pp. 379–402.
[30] Rakowski, J. M., and Pint, B. A., 2000, "Observations on the Effect of Water Vapor on the Elevated Temperature Oxidation of Austenitic Stainless Steel Foil," NACE Corrosion 2000, Orlando, March, NACE Paper No. 00-517.
[31] Quadakkers, W. J., Malkow, T., Piron-Abellan, J., Flesch, U., Shemet, V., and Singheiser, L., 2000, "Suitability of Ferritic Steels for Application as Construction Materials for SOFC Interconnects," *Proc. of 4th Eur. Solid Oxide Fuel Cell Forum*, Vol. 2, A. J. McEvoy, ed., Elsevier, Amsterdam, pp. 827–836.
[32] Pint, B. A., Zhang, Y., Tortorelli, P. F., Haynes, J. A., and Wright, I. G., 2001, "Evaluation of Iron-Aluminide CVD Coatings for High Temperature Corrosion Protection," Mater. High. Temp., **18**, pp.185–192.

Study of High Load Operation Limit Expansion for Gasoline Compression Ignition Engines

Koudai Yoshizawa
Atsushi Teraji
Hiroshi Miyakubo
Koichi Yamaguchi
Tomonori Urushihara

Nissan Research Center,
Nissan Motor Co., Ltd.,
1, Natsushima-cho,
Yokosuka,
Kanagawa 237-8523, Japan

In this research, combustion characteristics of gasoline compression ignition engines have been analyzed numerically and experimentally with the aim of expanding the high load operation limit. The mechanism limiting high load operation under homogeneous charge compression ignition (HCCI) combustion was clarified. It was confirmed that retarding the combustion timing from top dead center (TDC) is an effective way to prevent knocking. However, with retarded combustion, combustion timing is substantially influenced by cycle-to-cycle variation of in-cylinder conditions. Therefore, an ignition timing control method is required to achieve stable retarded combustion. Using numerical analysis, it was found that ignition timing control could be achieved by creating a fuel-rich zone at the center of the cylinder. The fuel-rich zone works as an ignition source to ignite the surrounding fuel-lean zone. In this way, combustion consists of two separate auto-ignitions and is thus called two-step combustion. In the simulation, the high load operation limit was expanded using two-step combustion. An engine system identical to a direct-injection gasoline (DIG) engine was then used to validate two-step combustion experimentally. An air-fuel distribution was created by splitting fuel injection into first and second injections. The spark plug was used to ignite the first combustion. This combustion process might better be called spark-ignited compression ignition combustion (SI-CI combustion). Using the spark plug, stable two-step combustion was achieved, thereby validating a means of expanding the operation limit of gasoline compression ignition engines toward a higher load range. [DOI: 10.1115/1.1805548]

Introduction

In recent years, homogeneous charge compression ignition (HCCI) has been the focus of much interest as a key technology for obtaining low emissions and high efficiency simultaneously [1,2]. Nitrogen oxide (NO_x) emissions can be greatly reduced with this combustion process [3,4]. Researchers have been examining auto-ignition combustion using a variety of fuels, such as diesel fuel [5–7], methanol [8], and natural gas [9]. There are some reports about using a dual fuel mixture for auto-ignition combustion [10]. With regard to compression auto-ignition combustion of gasoline fuel, Onishi et al. [11] and Ishibashi et al. [12] reported an improvement in combustion under low loads in a two-stroke engine, and Aoyama et al. [13] achieved compression auto-ignition combustion in a four-stroke engine using a high compression ratio of 17.4:1. With all of the fuels and combustion processes that have been tried, one major issue is that the engine speed and load range for obtaining stable combustion is narrow due to the occurrence of knocking and misfiring [14,15]. It is difficult to control the ignition timing and combustion speed, since these combustion characteristics are greatly affected by the engine operating parameters, including the engine speed, engine load, air-fuel ratio, intake temperature, and intake pressure [16,17].

To expand the high load operation limit, it is necessary to analyze the combustion characteristics of high load operation. In previous research, the authors developed a new auto-ignition combustion model for performing one-dimensional engine cycle simulations to investigate the characteristics of compression ignition combustion in gasoline engines [18]. In the present research, this engine cycle simulation method was used to analyze combustion characteristics under high load operation. The results suggested that knocking was caused by a rapid pressure increase in terms of $dP/dt_{\max}$. It was confirmed that slow or moderate combustion and retarded ignition timing are desirable for preventing knocking.

A statistical analysis was then conducted using various experimental data to clarify the phenomena determining the high load limit. It was found that with HCCI combustion, a stable combustion region was defined by three limits, namely, the engine stability limit, knocking limit, and misfire limit. The knocking limit was defined by rapid-pressure increase. The misfire limit was defined by the auto-ignition limit measured on the basis of combustion efficiency. Although slow or moderate combustion and retarded ignition timing are desirable for preventing knocking, the ignition timing, and combustion duration have a strong correlation with the air-fuel ratio. It was observed that high load operation was accompanied by a decrease in the air-fuel ratio. The richer air-fuel ratio resulted in earlier ignition and rapid combustion. These correlations are the cause of the high load operation limit with HCCI combustion.

Numerical and experimental analyses were conducted with the aim of expanding the high load limit. First, the possibility of expanding the high load limit by slow combustion was examined. It was found that the combustion duration has a strong correlation with engine stability. Even though a slower combustion speed is required to achieve higher load operation, it is difficult to reduce the combustion speed. Second, the possibility of expanding the high load limit by retarded combustion was examined. With retarded combustion, the variation of in-cylinder conditions has a much larger impact on ignition timing compared with combustion near TDC. The retarded combustion was not robust enough to accommodate the cycle-to-cycle variation of in-cylinder conditions. These phenomena are the main factors limiting high load operation under HCCI combustion. Therefore, it was confirmed that, to achieve stable retarded combustion, an ignition timing-control method is required.

Combustion control methods were then investigated by numeri-

Contributed by the IC Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received June 10, 2003; final manuscript received February 9, 2004. Assoc. Editor: D. Assanis.

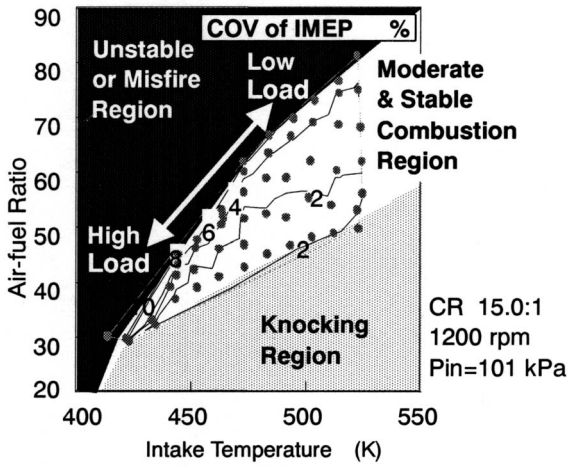


Fig. 1 Distribution of combustion regions [14]

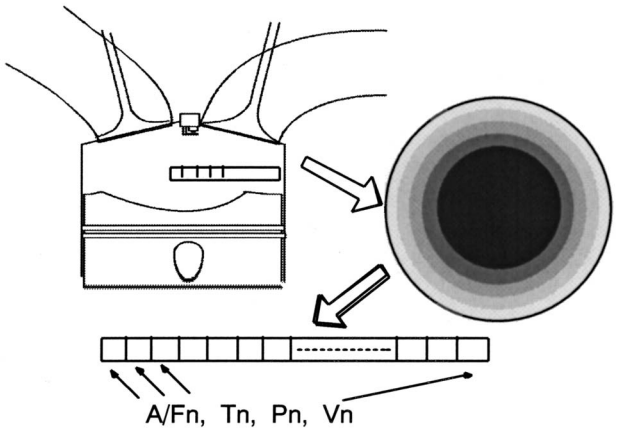


Fig. 2 Schematic diagram of computational grids [18]

cal analysis under a condition where an air-fuel stratification was induced in the cylinder. It was found that ignition timing control could be achieved by creating a fuel-rich zone at the center of the cylinder. The fuel-rich zone serves as an ignition source to ignite the surrounding fuel-lean zone. In this way, combustion consists of two separate auto-ignitions and is thus called two-step combustion. In the simulation, the high load operation limit was expanded using two-step combustion.

An engine system identical to a direct-injection gasoline (DIG) engine was used to demonstrate two-step combustion experimentally [19]. An air-fuel distribution was created by splitting fuel injection into first and second injections. The spark plug was then used to ignite the first combustion. This combustion process might better be called spark-ignited compression ignition combustion (SI-CI combustion). Using the spark plug, stable two-step combustion was demonstrated experimentally. Experimental results confirmed that two-step combustion achieved with the SI-CI combustion process is effective in expanding the higher load operation limit for gasoline compression ignition engines.

Background

One major issue in gasoline compression ignition engines is that the engine speed and load range for obtaining stable combustion is narrow due to the occurrence of knocking and misfiring. Figure 1 shows the distribution of combustion regions measured experimentally [14]. Under this load condition, the intake air temperature was increased to induce auto-ignition. The figure is divided into two major regions, one characterized by unstable combustion or misfiring at low intake air temperatures and the other characterized by stable combustion at higher intake air temperatures. The latter zone is further divided into a moderate combustion region positioned on the lean mixture concentration side and a knocking region positioned on the rich mixture concentration side. The coefficient of variation (COV) of the indicated mean effective pressure (IMEP) is the standard deviation of IMEP for 400 cycles and is used here as an index to define engine stability. In the moderate and stable combustion region, the COV of the IMEP contour is shown. The COV of IMEP is calculated by

$$\text{COV of IMEP} = \frac{\left[\frac{1}{400}\sum_{i=1}^{400}(\text{IMEP}_i - \text{IMEP}_{\text{ave}})^2\right]^{0.5}}{\text{IMEP}_{\text{ave}}} \quad (1)$$

where IMEP_i is IMEP of the ith cycle and IMEP_{ave} the average IMEP of 400 cycles. The moderate and stable region becomes narrower under high-load conditions due to the occurrence of knocking and misfiring.

To expand the high load operation limit, it is necessary to analyze the combustion characteristics of high load operation with respect to various engine operating conditions. In previous research, the authors developed a new engine cycle simulation procedure using an auto-ignition combustion model to investigate the characteristics of compression ignition combustion in gasoline engines [18]. A multizone model was selected to strike an acceptable balance between accuracy and CPU time for performing the calculations. Figure 2 shows the computational grids schematically. The cylinder volume was divided cylindrically into 500 cells from the bore center to the cylinder wall. Each cell had its own air-fuel ratio, temperature, and volume. This feature makes it possible to simulate the temperature and air-fuel distribution in the cylinder. A shell model was used to predict the auto-ignition timing [20,21]. However, calculating shell model reactions requires considerable CPU time to run a combustion simulation with the computational grids in Fig. 2. Therefore, different models were used for predicting the auto-ignition timing and performing the combustion calculations. The auto-ignition timing was predicted with the Livengood-Wu integral [22] and was calculated by

$$\int \frac{1}{\tau} dt = 1.0 \quad (2)$$

where τ is the ignition delay. In the model, it is assumed that the progress of combustion reactions is linear and accumulative. Auto-ignition occurs when the integral term on the left-hand side of the equation becomes 1.0. The accumulated $1/\tau$ represents how close the mixture is to auto-ignition. The shell model was used to estimate the ignition delay τ in a constant volume chamber under a wide range of pressure and temperature conditions. In the shell model, combustion consists of two-stage reactions, namely, low-temperature (cool flame) reactions and high-temperature (main) reactions. In this calculation, the ignition delay is defined by the occurrence of the main reactions. Therefore, the ignition delay up to the onset of the main reactions is taken into account. An ignition-delay database was generated for each air-fuel ratio, pressure, and temperature. Figure 3 shows one example of the ignition delay predicted with the shell model for an air-fuel ratio of 30:1. The auto-ignition timing was estimated by integrating $1/\tau$ along the pressure and temperature time history as shown in the figure. Once the ignited cells have been detected with the Livengood-Wu integral, the combustion calculations are then executed using thermodynamic equations [23]. In this simulation, predicted results for IMEP, hydrocarbon (HC) emissions, and nitrogen oxide (NO_x) emissions agreed well with the experimental results [18]. This multizone auto-ignition combustion model is called the General Engine Optimization Model (GENOM), which was used in this research to conduct numerical analyses.

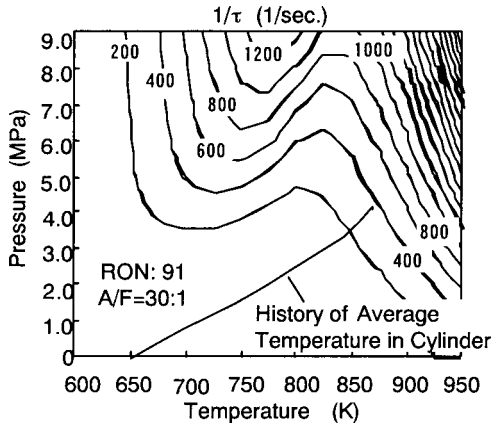


Fig. 3 Ignition delay predicted with shell model [18]

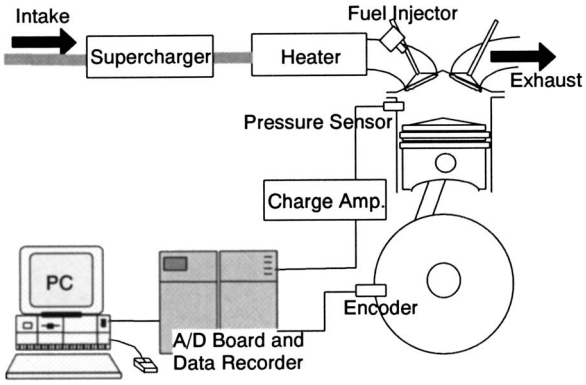


Fig. 5 Experimental apparatus

Mechanism of High Load Limit Under HCCI Combustion

Knocking Phenomena. The stable combustion region was estimated with the GENOM simulation model. Figure 4 compares the predicted and measured stable combustion regions. In the experiment, the knocking limit was defined on the basis of an analysis of pressure oscillations in the cylinder pressure wave form using an oscilloscope. In the simulation, the knocking limit was accurately predicted using the maximum rate of increase in pressure, $dP/dt_{\max}$. It is assumed that the in-cylinder pressure oscillations that led to knocking in the experiment were caused by a rapid pressure increase defined by $dP/dt_{\max}=4300$ MPa/s.

Statistical Analysis Using Empirical Equations. Using the experimental results, a statistical analysis was conducted to clarify the mechanism of the high load limit under HCCI combustion. The experimental apparatus used in this validation is outlined in Fig. 5. The fuel injector was installed in the intake port, where the fuel and air were mixed before being inducted into the single-cylinder engine in an attempt to achieve homogenous charge compression ignition. Intake air was naturally aspirated or supercharged at wide open throttle (WOT) and could be heated by means of a heater. The engine specifications are given in Table 1. The compression ratio was varied from 15:1 to 18.6:1 by changing the piston height. Regular gasoline (91 RON) was used as the test fuel. Table 2 shows the engine operating conditions that were examined. The cylinder pressure was measured with a pressure sensor and transferred via a charge amplifier and a data recorder to a PC terminal for analysis. The recorded cylinder pressure data were used to calculate the heat release rate. For each operating condition, data were recorded for 400 cycles to obtain average pressure profiles. As noted above, the COV of IMEP was also calculated based on the data for 400 cycles with Eq. (1). As shown in Tables 1 and 2, the engine operating parameters that were investigated included the compression ratio, intake air temperature, intake pressure, and air-fuel ratio.

The heat release rate pattern was investigated experimentally to define the heat release pattern in the statistical analysis. Figure 6 compares the heat release rates measured experimentally. With natural aspiration, the shape of the heat release rate was an isosceles triangle. On the other hand, in some tests conducted with supercharging, combustion continued slowly even after the main combustion period was finished. It is assumed that the slow com-

Table 1 Engine specifications

Engine type	4-stroke single cylinder
Bore×stroke	86×86 mm
Displacement	500 cm^3
Compression ratio	15.0:1–18.6:1
Throttle opening	WOT
Fuel	Regular Gasoline (91 RON)

Table 2 Engine operating conditions

Engine speed	1200–2400 rpm
Air-fuel ratio	30:1–120:1
Intake temperature	300–430 K
Intake pressure	101–152 kPa

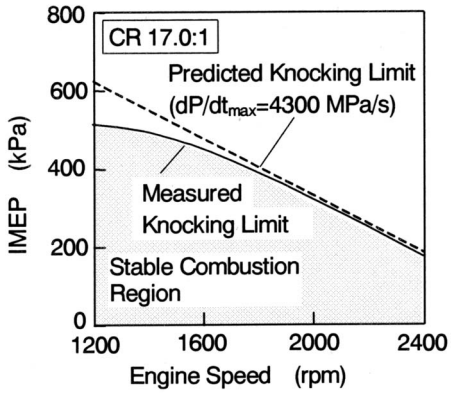


Fig. 4 Comparison of predicted and measured stable combustion regions

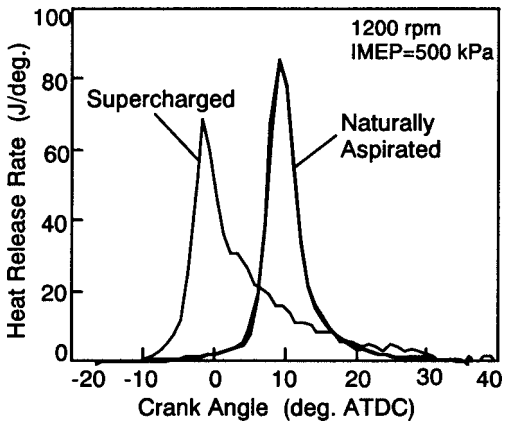


Fig. 6 Comparison of heat release rates

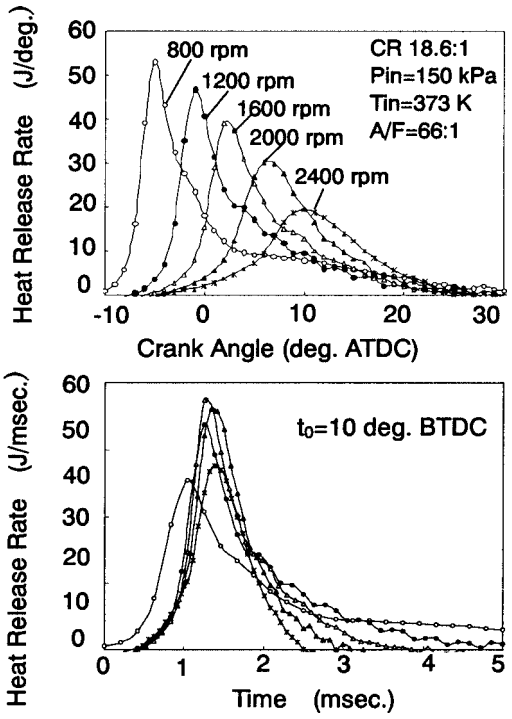


Fig. 7 Effect of engine speed on heat release rate profiles

bustion was due to oxidation of unburned fuel. Yet, even in the supercharged cases, the heat release rate of the main combustion period was simulated with an isosceles triangle.

Figure 7 shows the effect of the engine speed on the heat release rate profiles at an air-fuel ratio of 66:1. As the engine speed increased, the heat release duration became longer on a crank-angle basis. However, the heat release rate profiles coincided on a time basis. In other words, the combustion speed was independent of the engine speed. Heat release rates were analyzed for different air-fuel ratio conditions and similar results were obtained. The combustion speed was thus determined by the air-fuel ratio, pressure, and temperature. From these results, it is assumed that combustion was mainly controlled by the chemical reactions, though flow turbulence would also have some impact on combustion characteristics through the mixing effect.

A statistical analysis was then conducted to predict the stable combustion region. As shown in Fig. 1, the stable combustion region was defined by the knocking limit and the stability limit. First, the possibility of predicting the knocking limit was examined. Figure 8 shows a schematic diagram of the heat release rate pattern used in the analysis. The heat release rate pattern was simulated in the shape of an isosceles triangle, as indicated in Fig.

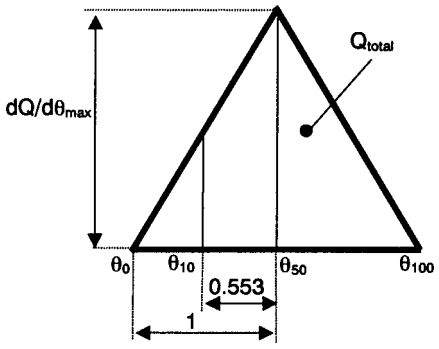


Fig. 8 Schematic diagram of heat release rate pattern

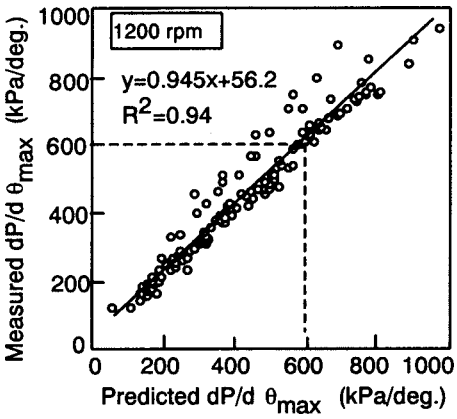


Fig. 9 Correlation between predicted and measured $dP/d\theta_{max}$

6. The crank angle interval from 10% to 50% mass burned, θ_{10-50}, was found to be 55.3% of that from 0% to 50% mass burned, θ_{0-50}. This mass burned ratio of 55.3% was estimated by correlation with the experimental data. It was assumed that the maximum heat release rate, $dQ/d\theta_{max}$, occurred at the 50% mass burned timing, θ_{50}. The maximum heat release rate, $dQ/d\theta_{max}$, was calculated by

$$\frac{dQ}{d\theta_{max}} = \frac{0.553 \cdot Q_{total}}{\theta_{10-50}} \tag{3}$$

As shown in Fig. 4, dP/dt_{max}, the rate of increase in maximum pressure per second, is an index that defines knocking intensity. In this analysis, experimental data obtained at an engine speed of 1200 rpm were used, and $dP/d\theta_{max}$, the rate of increase in maximum pressure per crank angle degree, was used as the knocking index. The following equation derived from the law of energy conservation and the gas phase equation was used to estimate $dP/d\theta$.

$$\frac{dP}{d\theta} = \frac{\kappa - 1}{V}\frac{dQ}{d\theta} - \frac{\kappa P}{V}\frac{dV}{d\theta} \tag{4}$$

where k is the specific heat ratio, V the in-cylinder volume, P the in-cylinder pressure, and $dV/d\theta$ the rate of increase in volume per crank angle degree. As shown in Fig. 8, it was assumed that $dQ/d\theta_{max}$ occurred at θ_{50}. Accordingly, $dP/d\theta_{max}$ was estimated by

$$\frac{dP}{d\theta_{max}} = \frac{\kappa_{50} - 1}{V_{50}} \cdot \frac{dQ}{d\theta_{50}} - \frac{\kappa_{50} P_{50}}{V_{50}} \cdot \frac{dV}{d\theta_{50}} \tag{5}$$

where κ_{50} is the specific heat ratio at θ_{50}, V_{50} the in-cylinder volume at θ_{50}, $dQ/d\theta_{50}$ the heat release rate at θ_{50}, P_{50} the in-cylinder pressure at θ_{50}, and $dV/d\theta_{50}$ the rate of increase in volume at θ_{50}. The value of P_{50} was calculated from the gas phase equation. Figure 9 shows the correlation between the predicted and measured $dP/d\theta_{max}$. The measured data were obtained for various conditions as shown in Table 2. The values of $dP/d\theta_{max}$ were well predicted as a result of using the measured mass burned θ_{10-50} and combustion timing θ_{50} in the calculations. Most of the divergent data were for supercharged cases. As indicated in Fig. 6, with supercharging, slow combustion occurred after the main combustion period. Therefore, $dQ/d\theta_{max}$ occurred earlier than at θ_{50} and $dP/d\theta_{max}$ was underestimated with Eq. (5). If $dP/d\theta_{max}$ were calculated at the crank angle of $dQ/d\theta_{max}$, the correlation coefficient R^2 would become 0.99. In this way, the knocking limit was predicted on the basis of θ_{10-50} and θ_{50}.

The possibility of predicting the stability limit was then examined. In the experiment, engine stability was evaluated using the COV of IMEP, which was predicted on the basis of a multivari-

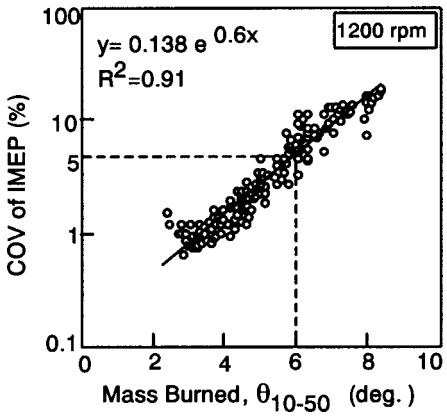


Fig. 10 Correlation between mass burned, θ_{10-50}, and COV of IMEP

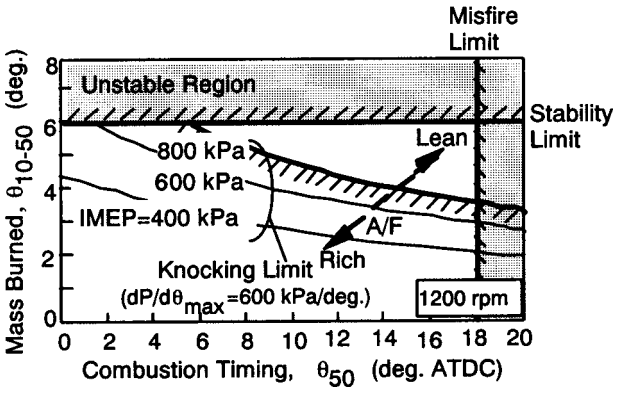


Fig. 12 Effect of combustion timing, θ_{50}, and mass burned, θ_{10-50}, on stable combustion region (predicted)

able analysis. It was found that the COV of IMEP had a strong correlation with mass burned θ_{10-50} and it was expressed by

$$\text{COV of IMEP} = 0.138e^{0.6\cdot\theta_{10-50}} \quad (6)$$

Figure 10 shows the correlation between mass burned θ_{10-50} and the COV of IMEP. The latter was well predicted with mass burned θ_{10-50}. A multivariable analysis was also conducted at an engine speed of 2400 rpm, and the strong correlation between the COV of IMEP and mass burned θ_{10-50} was again observed. In this way, the stability limit was predicted with mass burned θ_{10-50}. In the experiment, the stability limit was set to the COV of IMEP =5%. Therefore, the required mass burned θ_{10-50} was less than 6 crank angle deg, as shown in Fig. 10.

In addition to the aforementioned stability limit, another limit defined by combustion efficiency was observed. Figure 11 shows the effects of the combustion timing on engine performance as investigated with the GENOM simulation model. As seen in the figure, dP/dt_{max} decreased as the ignition timing was further retarded from TDC. On the other hand, HC emissions increased as the ignition timing was further retarded from TDC, presumably because there was less time for the fuel to advance the reactions.

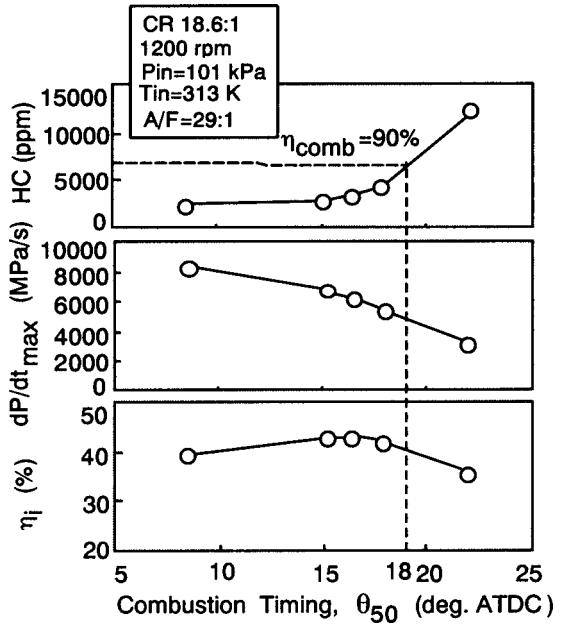


Fig. 11 Effect of ignition timing, θ_{10}, on HC, dP/dt_{max} and η_i (predicted)

In this way, the stable combustion region was also limited by combustion efficiency calculated in terms of the unburned hydrocarbon level. This combustion efficiency limit is referred to here as the misfire limit. In the analysis, the misfire limit was set at combustion efficiency=90%, which is equivalent to an HC level of about 7000 ppm under high load operation. That required that θ_{50} should be earlier than 18 deg ATDC. As mentioned earlier, the results of the statistical analysis suggested that the stable region of HCCI combustion was defined by three limits, namely, the knocking limit, stability limit, and misfire limit.

Figure 12 shows the stable combustion region. The knocking limit, namely, $dP/d\theta_{max}=600$ kPa/deg, was calculated by Eq. (5). The stability limit was defined by mass burned $\theta_{10-50}=6$ crank angle deg, as indicated in Fig. 10. The misfire limit was defined by $\theta_{50}=18$ deg ATDC, as indicated in Fig. 11. The stable combustion region was found to have a triangular shape when defined as a function of θ_{50} and θ_{10-50}. The figure shows the knocking limit calculated by Eq. (5). for each of the IMEP conditions. As shown in Fig. 12, a larger IMEP requires an increasingly longer combustion duration and retarded combustion timing. However, the ignition timing and combustion duration have a strong correlation with the air-fuel ratio. High load operation is accompanied by a decrease in the air-fuel ratio, which has the effect of advancing the ignition timing and promoting rapid combustion as shown in the figure. These correlations are one factor limiting high load operation with HCCI combustion.

The knocking limit was defined by $dP/dt_{max}=4300$ MPa/s. At an engine speed of 1200 rpm, the knocking limit was defined by $dP/d\theta_{max}=600$ kPa/deg, and at an engine speed of 2400 rpm, it was defined by $dP/d\theta_{max}=300$ kPa/deg. Therefore, higher engine speeds require a longer combustion duration and a more retarded combustion timing, the same as for stable operation under higher loads. On the other hand, the misfire limit was defined by the combustion timing θ_{50} that was determined by the in-cylinder pressure and temperature needed for the chemical reactions, as indicated in Fig. 11. These trends explain why the stable combustion region becomes narrower at higher engine speeds, as shown in Fig. 4. The foregoing results suggest that there is a potential difficulty in trying to expand the stable combustion region toward higher engine loads or higher engine speeds by HCCI combustion.

Combustion Concepts for Preventing Knocking. As seen in the foregoing section, the following two combustion concepts are effective in preventing knocking:

- Slow combustion with a longer combustion duration
- Retarded combustion with a retarded ignition timing

First, the possibility of expanding the high load limit with slow combustion was examined. As shown in Fig. 10, the combustion duration has a strong correlation with engine stability. A slower

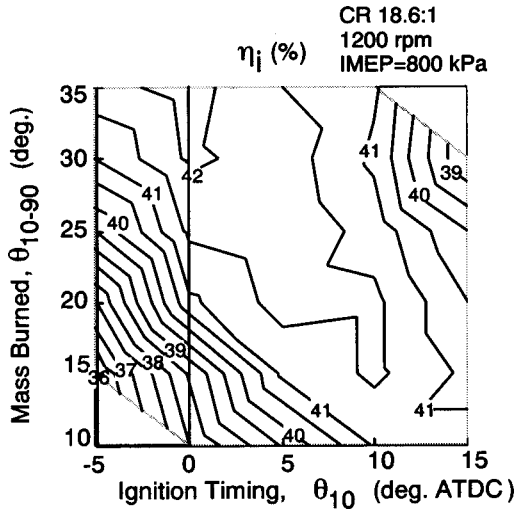


Fig. 13 Effect of ignition timing, θ_{10}, and mass burned, θ_{10-90}, on η_i (predicted)

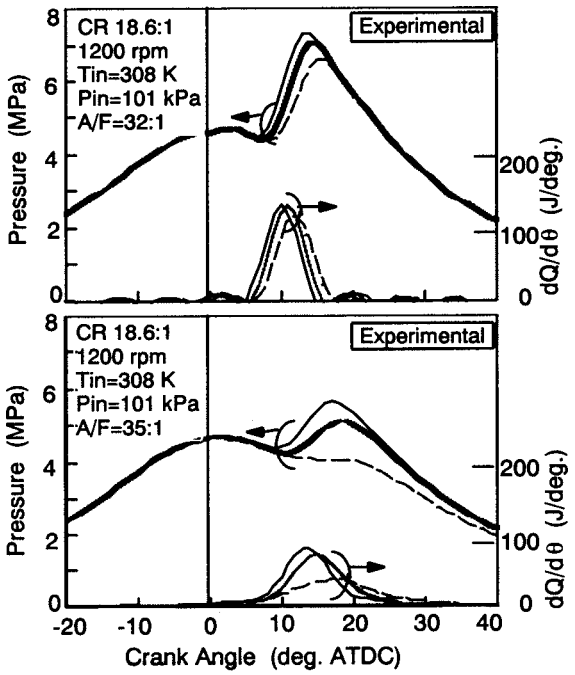


Fig. 14 Effect of combustion timing on pressure and heat release rate profile variation

combustion speed is required to achieve higher load operation. However, it would be difficult to reduce the combustion speed without causing engine stability to deteriorate.

Second, the possibility of expanding the high load limit was examined with retarded combustion. It is assumed that retarded combustion caused the indicated thermal efficiency to deteriorate because the pressure profile diverged from the ideal Otto cycle. In order to clarify this, the effect of retarded combustion on thermal efficiency was investigated. In the analysis, combustion was accomplished artificially with a predetermined ignition timing and combustion duration in order to investigate their effect on thermal efficiency. The heat release rate was simulated with an isosceles triangle shape as shown in Fig. 6. It was also assumed that complete combustion occurred, so the effect of unburned hydrocarbons was ignored. Figure 13 shows the indicated thermal efficiency η_i, under a wide range of ignition timings θ_{10} and combustion durations θ_{10-90}. The indicated thermal efficiency, η_i, was defined as the ratio of the actual work per cycle to the heating value of the fuel injected into the cylinder. The gray lines in the figure show the boundary of the region that was analyzed. The indicated thermal efficiency stayed at the same level until the ignition timing was further retarded to around 10 deg ATDC. It was confirmed that the cooling loss decreased with retarded combustion and that is why the indicated thermal efficiency stayed at the same level. These results indicated that retarded combustion is a promising way to prevent knocking without sacrificing thermal efficiency.

Figure 14 shows the cycle data measured experimentally for the pressure and heat release rate profiles. Three typical cycles are shown in the figure to clarify the variation in the pressure and heat release rate profiles. In the experiment, the air-fuel ratio was varied slightly to change the combustion timing. When combustion occurred just after TDC, the pressure and heat release rate profiles showed little variation and stable combustion was achieved. On the other hand, when combustion occurred at a more retarded timing, the variation in the pressure and heat release rate profiles became very large and combustion became unstable. The combustion timing shifted with each cycle. However, as indicated in Fig. 13, under a condition of complete combustion, the combustion timing shift shown in Fig. 14 did not affect the indicated thermal efficiency and IMEP. Therefore, it is assumed that the variation in IMEP under retarded combustion was caused by a difference in the unburned hydrocarbon quantity that was largely influenced by the combustion timing as shown in Fig. 11.

Numerical analyses were then conducted to clarify the cause of cycle-to-cycle variation. As indicated in Fig. 7, chemical reactions play a large role in HCCI combustion and the following three parameters represent the in-cylinder conditions that control the reaction rate:

- In-cylinder pressure
- Air-fuel ratio
- In-cylinder temperature

It is assumed that cycle-to-cycle variation of these in-cylinder conditions was the cause of unstable combustion. First, the possibility of in-cylinder pressure variation was investigated using a naturally aspirated engine that was operated under the conditions in Fig. 14 at WOT. The results suggested that cycle-to-cycle variation of the in-cylinder pressure was not likely to occur.

Second, the possibility of air-fuel ratio variation was investigated. Variation of the in-cylinder air-fuel ratio was measured using laser-induced fluorescence (LIF) and the in-cylinder HC measurement before combustion [24]. It was confirmed that the air-fuel ratio would vary in a range of -0.5 to $+0.5$ when it was around 32–35:1. It was assumed that the in-cylinder air-fuel ratio variation was caused by either fuel injection quantity variation or intake air flow variation. Cycle-to-cycle variation of the air-fuel ratio was then investigated by GENOM simulations. Since GENOM is designed for engine cycle simulations, it cannot simulate cycle-to-cycle variation directly. Therefore, engine-cycle simulations were run three times with different engine operating conditions. Figure 15 shows the effect of air-fuel ratio variation on the pressure and heat release rate profiles. The air-fuel ratio was changed artificially to simulate its cycle-to-cycle variation in the cylinder. When combustion occurred just after TDC, variation of the air-fuel ratio had little impact on the combustion timing. On the other hand, when combustion occurred at a more retarded timing, variation of the air-fuel ratio had a large impact on the combustion timing, shifting it in the same manner as in Fig. 14.

Third, the possibility of in-cylinder temperature variation was investigated. There was a possibility that the in-cylinder temperature at IVC varied from -1 K to $+1$ K due to variation of the internal EGR rate. Therefore, cycle-to-cycle variation of the in-cylinder temperature was investigated on the basis of GENOM simulations. Figure 16 shows the effect of inlet temperature varia-

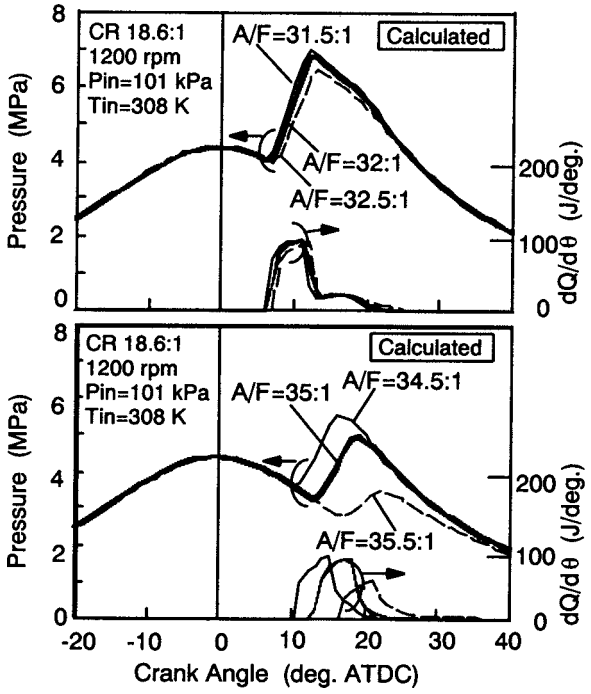


Fig. 15 Effect of air-fuel ratio variation on pressure and heat release rate profiles

tion on the pressure and heat release rate profiles. The inlet temperature was changed artificially to simulate cycle-to-cycle variation of the in-cylinder temperature. The same trend was observed. When combustion occurred just after TDC, variation of the inlet temperature had little impact on the combustion timing. On the other hand, when combustion occurred at a more retarded timing, inlet temperature variation had a large impact on the combustion timing, shifting it in the same manner as in Fig. 15.

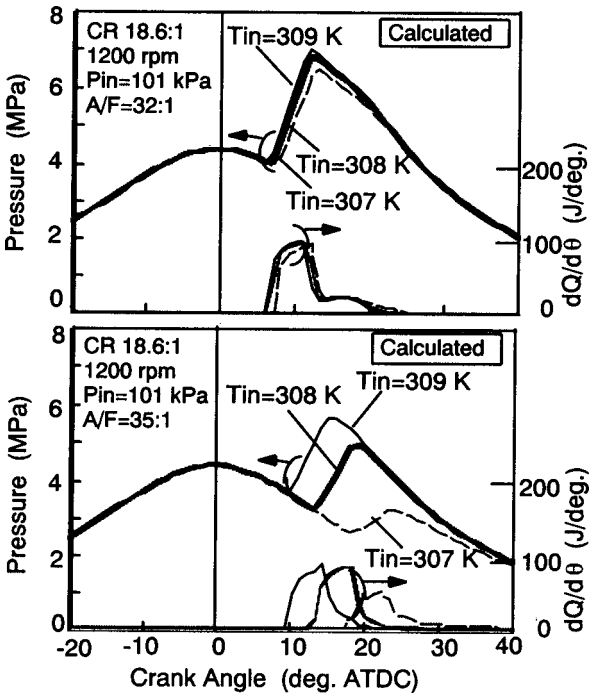


Fig. 16 Effect of inlet temperature variation on pressure and heat release rate profiles

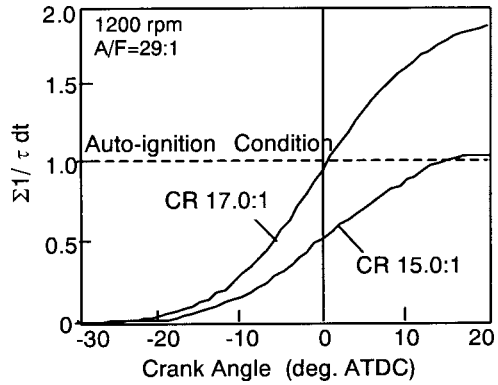


Fig. 17 Comparison of $\Sigma 1/\tau\ dt$ profiles (predicted)

These results suggested that combustion timing variation under a state of retarded combustion was caused by variation in the in-cylinder conditions. It is desirable to reduce cycle-to-cycle variation of the in-cylinder conditions through future development work. However, it is difficult to eliminate cycle-to-cycle variation of the in-cylinder conditions with current injectors or valve timing control devices. Therefore, it is necessary to develop a robust combustion process that can accommodate such cycle-to-cycle variations.

Figure 17 compares the $\Sigma 1/\tau dt$ profiles for a cell near the bore center. As seen in Eq. (2), $\Sigma 1/\tau dt$ represents how close the mixture is to auto-ignition. As shown in the figure, the chemical reactions progressed rapidly mainly around TDC when the in-cylinder pressure and temperature were high. In this analysis, the compression ratio was changed to vary the combustion timing. When auto-ignition occurred at TDC, the slope where $\Sigma 1/\tau dt$ crossed the auto-ignition condition was steep. Therefore, even though in-cylinder condition variations that affected the $\Sigma 1/\tau dt$ profile occurred, the auto-ignition timing did not change much. On the other hand, when ignition occurred at a retarded timing, the slope where $\Sigma 1/\tau dt$ crossed the auto-ignition condition was gentler. If in-cylinder condition variations that affected the $\Sigma 1/\tau dt$ profile occurred, the auto-ignition timing would change greatly. These results suggested that retarded combustion was not robust enough to accommodate the cycle-to-cycle variation of in-cylinder conditions. These phenomena are the main factors limiting high load operation under HCCI combustion. Therefore, it was assumed that some method of controlling the ignition timing was needed to achieve stable retarded combustion.

Optimum In-Cylinder Fuel-Distribution for Ignition Timing Control

Ignition Timing Control Method. Although it is easy to control the ignition timing in numerical simulations, it is a very challenging task under gasoline compression ignition combustion in an actual engine. The spark plug is an ignition timing-control method for spark ignition (SI) gasoline engines and the high-pressure injection system serves the same function in diesel engines. Methods of controlling the ignition timing in HCCI combustion have been reported in the literature, such as the use of an intake air heater, supercharger, variable valve timing control, variable compression ratio control, or a dual fuel mixture. However, these methods have severe disadvantages in transient operation or they require a more complicated engine system than conventional SI gasoline engines or diesel engines. Therefore, in this analysis, the effect of the in-cylinder fuel distribution on ignition timing control was investigated in detail for a simple engine system. The fuel distribution was controlled under the assumption that direct injection of gasoline fuel was applied [19].

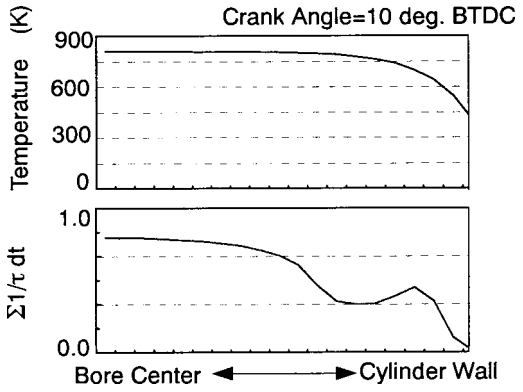


Fig. 18 Temperature and $\Sigma 1/\tau\ dt$ distributions [18]

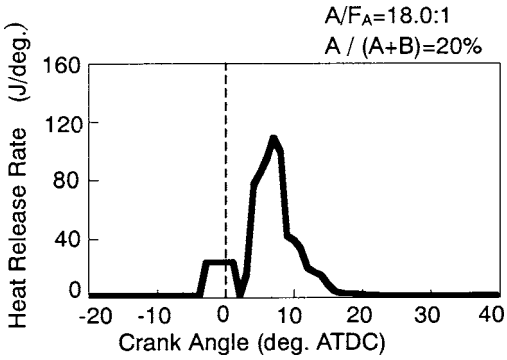


Fig. 20 Calculated heat release rate of two-step combustion (predicted)

Stratified Charge for Ignition Timing Control. To investigate the optimum fuel distribution for ignition timing control, first, a numerical analysis was made of high load operation under HCCI combustion. Figure 18 shows one example of the temperature and accumulated $1/\tau$ distributions in the cylinder at 10 deg BTDC under high load conditions. The data represent the conditions in the cylinder just before auto-ignition. It is seen that $\Sigma 1/\tau dt$ is high in the bore center, suggesting that auto-ignition is likely to occur there. There is another peak near the cylinder wall, which is attributed to the Negative Temperature Coefficient (NTC) region. This tendency agrees well with the results reported by Nishiwaki et al. [25]. The temperature gradient is rather small, and the temperature is uniform at the bore center. Based on the analytical data, it is assumed that creating a fuel distribution at the bore center to take advantage of the higher temperature region there would be more effective in controlling the ignition timing than creating a fuel-rich zone near the cylinder wall where the temperature is lower and rich fuel would be a source of large unburned HC emissions.

Figure 19 shows a schematic diagram of the air-fuel ratio distribution that was investigated. The in-cylinder volume, consisting of 500 cells, was divided into two zones, namely, a bore center zone denoted as A and an outer zone denoted as B. To ignite the bore center cells, the air-fuel ratio in zone A, A/F_A, was made relatively richer than that in the outer zone B, although it was still lean. The volume ratio of zone A was set at 20% of the total fuel amount. It was assumed that the heat release at the first auto-ignition in zone A would stimulate the chemical reactions in zone B through the pressure and temperature increase, making it possible to ignite the cells under a leaner air-fuel condition than that of HCCI combustion. It was assumed that condition would help to lower the NO_x emission level.

Figure 20 shows the calculated heat release rate. As intended, the fuel-rich zone in zone A worked as an ignition source to ignite the surrounding fuel-lean zone in zone B. In this way, combustion was accomplished by two separate auto-ignitions. The main combustion timing was retarded due to the time lag between the first and second auto-ignitions. In addition to that, rapid combustion that causes knocking was also prevented due to the separate heat release periods. Therefore, dP/dt_{max} was greatly reduced. This auto-ignition combustion process is called two-step combustion. In this way, creating a fuel-rich zone at the center of the cylinder for two-step combustion is effective for ignition timing control.

A numerical analysis was then conducted to examine the effect of ignition timing control on retarded combustion. Figure 21 shows the effect of ignition timing control on the pressure and the heat release rate profiles. In this analysis, the ignition timing was set at a predetermined crank angle under the assumption that the fuel at the bore center would be ignited forcibly by some means. The in-cylinder conditions were the same as in Fig. 15. One-tenth of the total fuel was ignited by combustion at 10 deg ATDC. As the results indicate, even when combustion occurred at a retarded timing, cycle-to-cycle variation of the air-fuel ratio had little impact on the combustion timing, which stayed at the same position. These results suggested that an ignition timing control with two-step combustion would be effective in achieving robust retarded combustion.

Experimental Validation of Two-Step Combustion

As the final step, two-step combustion was validated experimentally using the engine configuration shown in Fig. 22. For two-step combustion, the injector was placed in the combustion chamber for direct injection. In addition, a spark plug was installed for ignition timing control. This engine system is the same as a direct-injection gasoline (DIG) engine [19]. The two fuel distribution zones in Fig. 19 were created by splitting fuel injection into first and second injections. To ignite the bore center zone

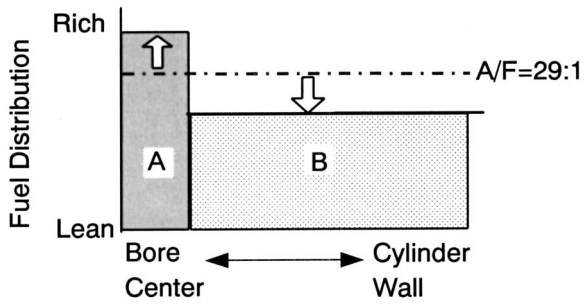


Fig. 19 Fuel distribution in the cylinder

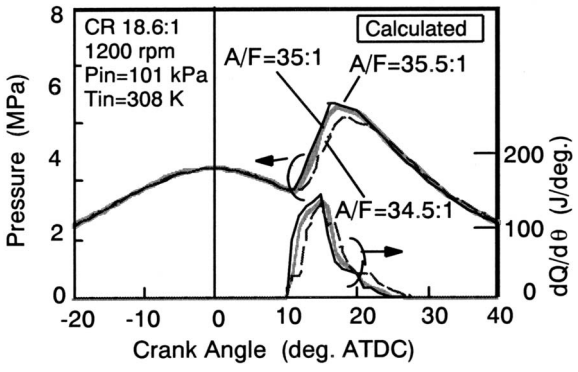


Fig. 21 Effect of ignition timing control with two-step combustion

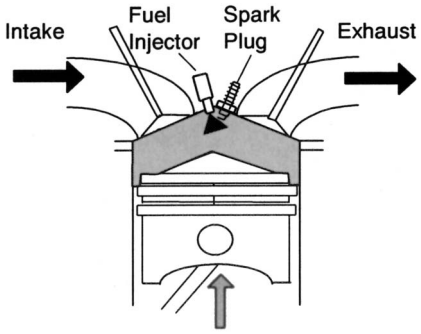


Fig. 22 Engine configuration for SI-CI combustion

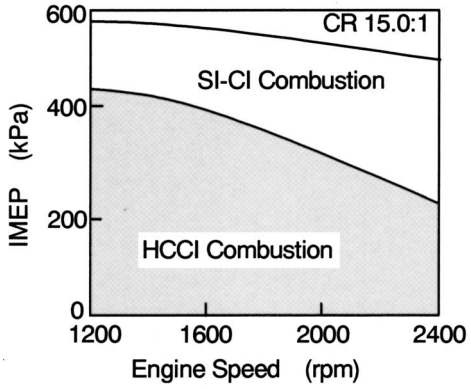


Fig. 24 Comparison of stable combustion region with SI-CI combustion and HCCI combustion

by spark ignition, the air-fuel ratio near the spark plug was made relatively richer than that of the periphery zone, though it was still lean. The first injection occurred during the intake stroke to form a lean air-fuel mixture throughout the combustion chamber. The second injection took place in the compression stroke to form a rich air-fuel mixture near the bore center. Spark ignition occurred right after the second injection. The spark timing determined the ignition timing. In this two-step combustion concept, the spark plug would be used only for igniting a small quantity of a fuel-rich mixture at the bore center and the remaining fuel would be ignited by auto-ignition.

Figure 23 shows the pressure and heat release rate profiles obtained experimentally. These experimental results suggest that two different types of combustion occurred. From −20 to 10 deg ATDC, rather slow combustion due to flame propagation was seen. Considering the combustion speed, it is assumed that spark-ignited combustion took place during this period. After 10 deg ATDC, rapid combustion with a heat release rate in the form of an isosceles triangle was observed. It is assumed that compression combustion took place during this period as indicated in Fig. 6. As

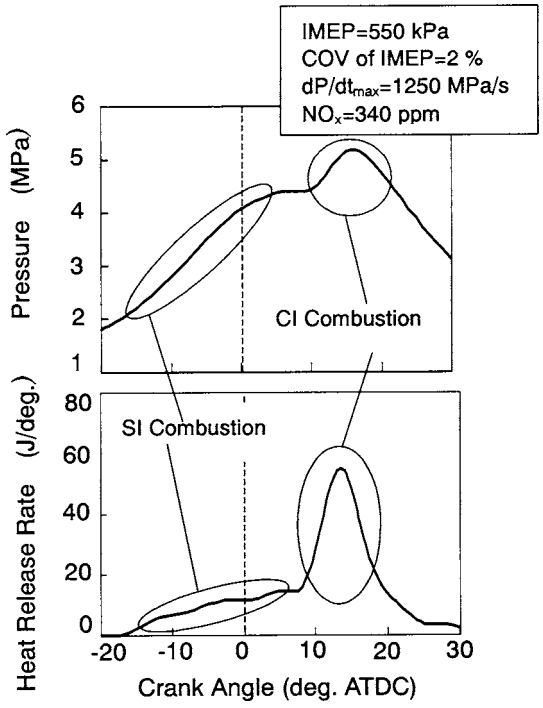


Fig. 23 Measured pressure and heat release rate profiles of SI-CI combustion

planned, spark-ignited combustion of the fuel-rich zone in the bore center worked as an ignition source to ignite the surrounding fuel-lean zone. This combustion process might better be called spark-ignited compression ignition combustion (SI-CI combustion) or spark-assisted CI combustion. The main CI combustion timing was retarded due to the time lag between SI combustion and CI combustion. The value of dP/dt_{max} was greatly reduced due to retarded combustion, although the NO_x emission level of 350 ppm was a little higher than expected. The overall air-fuel ratio was about 26:1. However, the charge composition near the spark-plug was made relatively richer than that of the periphery for SI combustion. Therefore, it is assumed that SI combustion was the main source of NO_x formation. Stable two-step combustion was achieved with the COV of IMEP of 2%. It is assumed that a stable first combustion process was the main reason for stable combustion.

Figure 24 compares the stable combustion region with SI-CI combustion and HCCI combustion. Both stable combustion regions were measured experimentally. With SI-CI combustion, the high load operation limit was expanded, resulting in a much larger stable combustion region than that seen for HCCI combustion. There is concern about an increase in NO_x due to SI combustion. Therefore, as shown in Fig. 19, it is necessary to reduce the quantity of rich fuel for SI combustion at the bore center to the minimum level needed to facilitate compression ignition of the surrounding fuel. The introduction of EGR gas would also be effective in suppressing NO_x formation. These are issues that will be investigated in future work.

Summary and Conclusions

Numerical and experimental analyses of gasoline compression ignition engines were conducted with the aim of expanding the high load operation limit. The following conclusions can be drawn from the results.

1. It was clarified that the stable region of HCCI combustion was defined by the knocking limit, stability limit, and misfire limit. The combustion duration and combustion timing have a strong correlation with the air-fuel ratio. A richer air-fuel ratio results in a shorter combustion duration and a more advanced combustion timing. This tendency explains the potential difficulty of achieving high load operation with HCCI combustion, especially at higher engine speeds.
2. Engine stability has a strong correlation with the combustion duration, and it is difficult to prevent knocking with slow combustion. Retarded combustion is an effective way to prevent knocking without sacrificing thermal efficiency.

3. Numerical analysis results suggested that the instability of retarded combustion was caused by variation of in-cylinder conditions, which caused the combustion timing to vary. Therefore, an ignition timing control method is required to achieve stable retarded combustion.
4. It was found that ignition timing control could be achieved by creating a fuel-rich zone at the center of the cylinder, which works as an ignition source to ignite the surrounding fuel-lean zone.
5. Two-step combustion achieved with ignition timing control combined with the above-mentioned in-cylinder air-fuel distribution is effective in preventing knocking and deterioration of the NO_x emission level.
6. This two-step combustion of gasoline fuel was achieved using the spark plug to ignite the first combustion. Experimental results confirmed that two-step combustion is effective in expanding the high load operation limit of gasoline compression ignition engines.

Acknowledgment

The authors would like to thank all of the people who helped to conduct the numerical and experimental analyses described here and who contributed to the writing of this paper.

Nomenclature

dP/dt_{max} = rate of increase in maximum pressure per second
$dP/d\theta_{max}$ = rate of increase in maximum pressure per crank angle deg
dQ/dt_{max} = maximum heat release rate per second
$dQ/d\theta$ = heat release rate per crank angle deg
$dQ/d\theta_{max}$ = maximum heat release rate per crank angle deg
$dQ/d\theta_{50}$ = heat release rate at θ_{50}
$dV/d\theta$ = rate of increase in volume per crank angle deg
$dV/d\theta_{50}$ = rate of increase in volume per crank angle deg at θ_{50}
$IMEP_{ave}$ = average IMEP of 400 cycles
$IMEP_i$ = IMEP of *i*th cycle
P = in-cylinder pressure
P_{in} = intake pressure
P_{50} = in-cylinder pressure at θ_{50}
Q_{total} = total heat release
R = correlation coefficient
T = in-cylinder temperature
T_{in} = intake temperature
V = in-cylinder volume
V_{50} = in-cylinder volume at θ_{50}
A/F_A = air-fuel ratio in region A
η_{comb} = combustion efficiency
η_i = indicated thermal efficiency
κ = specific heat ratio
κ_{50} = specific heat ratio at θ_{50}
λ = excess air ratio
θ_0 = combustion start timing
θ_{10} = 10% mass burned timing
θ_{50} = 50% mass burned timing
θ_{100} = combustion end timing
θ_{0-50} = crank angle interval from 0–50% mass burned
θ_{10-50} = crank angle interval from 10–50% mass burned
θ_{10-90} = crank angle interval from 10–90% mass burned
τ = ignition delay

Abbreviations and Acronyms

A/D = analog/digital
A/F = air-fuel ratio
ATDC = after TDC
BTDC = before TDC
CI = compression ignition
COV = coefficient of variation
CPU = central processing unit
CR = compression ratio
DIG = direct injection gasoline
EGR = exhaust gas recirculation
GENOM = general engine optimization model
HC = hydrocarbon
HCCI = homogenous charge compression ignition
IVC = intake valve closing
IMEP = indicated mean effective pressure
LIF = laser-induced fluorescence
NO_x = nitrogen oxides
NTC = negative temperature coefficient
PC = personal computer
RON = research octane number
SI = spark ignition
SI-CI = spark-ignited compression ignition
TDC = top dead center
WOT = wide open throttle

References

[1] Najt, P. M., and Foster, D. E., 1983, "Compression-Ignited Homogeneous Charge Combustion," SAE Paper 830264.
[2] Thring, R. H., 1989, "Homogeneous-Charge Compression-Ignition (HCCI) Engines," SAE Paper 892068.
[3] Kimura, S., Matsui, Y., and Koike, M., 1998, "New Combustion Concept for Simultaneous Reduction of NO_x and Particulate Emissions From Small DI Diesel Engines," *Proc. 1998 FISTA World Automotive Congress*, SIA, Paris, Paper No. F98T132.
[4] Masahiro, F., Ohta, Y., Kono, M., and Hasegawa, M., 1998, "An Ultra-Lean Premixed Compression-Ignition Engine Concept and Its Characteristics," *Proc. Fourth International Symposium COMODIA*, JSME, Kyoto, pp. 193–197.
[5] Yanagihara, H., Satou, Y., and Mizuta, J., 1996, "A Simultaneous Reduction of NO_x and Soot in Diesel Engines Under a New Combustion System (Uniform Bulky Combustion System-UNIBUS)," *17th Int. Vienna Motor Symposium*, OVK, Vienna, pp. 303–314.
[6] Takeda, Y., Nakagome, K., and Niimura, K., 1996, "Emission Characteristics of Premixed Lean Diesel Combustion With Extremely Early Stage Fuel Injection," SAE Paper 961163.
[7] Gray, A. W., and Ryan, T. W., 1997, "Homogeneous Charge Compression Ignition (HCCI) of Diesel Fuel," SAE Paper 971676.
[8] Iida, N., 1994, "Combustion Analysis of Methanol-Fueled Active Thermo-Atmosphere Combustion (ATAC) Engine Using a Spectroscopic Observation," SAE Paper 940684.
[9] Christensen, M., Johansson, B., and Einewall, P., 1997, "Homogeneous Charge Compression Ignition (HCCI) Using Isooctane, Ethanol and Natural Gas—A Comparison With Spark Ignition Operation," SAE Paper 972874.
[10] Daisho, Y., Yaeo, T., Koseki, T., Saito, T., Kihara, R., and Quiros, N. E., 1995, "Combustion and Exhaust Emissions in a Direct-Injection Diesel Engine Dual-Fueled With Natural Gas," SAE Paper 950465.
[11] Onishi, S., Jo, S. H., Shoda, K., Jo, P. D., and Kato, S., 1979, "Active Thermo-Atmosphere Combustion (ATAC)—A New Combustion Process for Internal Combustion Engines," SAE Paper 790501.
[12] Ishibashi, Y., and Asai, M., 1998, "A Low Pressure Pneumatic Direct Injection Two-Stroke Engine by Activated Radical Combustion Concept," SAE Paper 980757.
[13] Aoyama, T., Hattori, Y., Mizuta, J., and Sato, Y., 1996, "An Experimental Study on Premixed-Charge Compression Ignition Gasoline Engine," SAE Paper 960081.
[14] Hiraya, K., Hasegawa, K., Urushihara, T., Iiyama, A., and Itoh, T., 2002. "A Study on Gasoline Fueled Compression Ignition Engine—A Trial of Operation Region Expansion," SAE Paper 2002-01-0416.
[15] Mandokoro, Y., Nakano, M., and Kubo, S., 1998, "Combustion Characteristics of Homogeneous Charge Compression Ignition Engines (I)," *Proc. 36th Combustion Symposium* (in Japanese), Combustion Society of Japan, Sappore, pp. 130–132.
[16] Christensen, M., Hultqvist, A., and Johansson, B., 1999, "Demonstrating the Multi Fuel Capability of a Homogeneous Charge Compression Ignition Engine With Variable Compression Ratio," SAE Paper 1999-01-3679.
[17] Noda, T., and Foster, D., 2001, "A Numerical Study to Control Combustion Duration of Hydrogen-Fueled HCCI by Using Multi-Zone Chemical Kinetics Simulation," SAE Paper 20001-01-0250.
[18] Yoshizawa, K., Teraji, A., Aochi, E., Kubo, M., and Kimura, S., 2002, "Numerical Analysis of Combustion in Gasoline Compression Ignition Engines," SAE Paper 2002-01-2865.
[19] Takagi, Y., Itoh, T., Muranaka, S., Iiyama, A., Iwakiri, Y., Urushihara, T., and Naitoh, K., 1998, "Simultaneous Attainment of Low Fuel Consumption, High Output Power and Low Exhaust Emissions in Direct Injection SI Engines,"

SAE Paper 980149.
[20] Halsted, M. P., Kirsch, L. J., Prothero, A., and Quinn, C. P., 1975, ''A Mathematical Model for Hydrocarbon Autoignition at High Pressures,'' Proc. R. Soc. London, Ser. A, **346**, pp. 515–538.
[21] Halsted, M. P., Kirsch, L. J., and Quinn, C. P., 1977, ''The Autoignition of Hydrocarbon Fuels at High Temperatures and Pressures—Fitting of a Mathematical Model,'' Combust. Flame, **30**, pp. 45–60.
[22] Livengood, J. C., and Wu, P. C., 1955, ''Correlation of Autoignition Phenomenon in Internal Combustion Engines and Rapid Compression Machines,'' *Proc. 5th International Symposium on Combustion*, The Combustion Institute, Pittsburgh. pp. 347–356.
[23] Muranaka, S., Takagi, Y., and Ishida, T., 1987, ''Factors Limiting the Improvement in Thermal Efficiency of S.I. Engine at High Compression Ratio,'' SAE Transaction 870548.
[24] Kakuhou, A., Urishihara, T., Itoh, T., and Takagi, Y., 1999, ''Characteristics of Mixture Formation in a Direct Injection SI Engine With Optimized In-cylinder Swirl Air Motion,'' SAE Paper 1999-01-0505.
[25] Nishiwaki, K., 1998, ''Modeling Engine Heat Transfer and Flame-Wall Interaction,'' *Proc. Fourth International Symposium COMODIA*, JSME, Kyoto, pp. 35–44.

Quasidimensional Modeling of Direct Injection Diesel Engine Nitric Oxide, Soot, and Unburned Hydrocarbon Emissions

Dohoy Jung
e-mail: dohoy@umich.edu

Dennis N. Assanis
e-mail: assanis@umich.edu

Department of Mechanical Engineering,
The University of Michigan,
1231 Beal Avenue,
Ann Arbor, MI 48109-2133

In this study we report the development and validation of phenomenological models for predicting direct injection (DI) diesel engine emissions, including nitric oxide (NO), soot, and unburned hydrocarbons (HC), using a full engine cycle simulation. The cycle simulation developed earlier by the authors (D. Jung and D. N. Assanis, 2001, SAE Transactions: Journal of Engines, 2001-01-1246) features a quasidimensional, multizone, spray combustion model to account for transient spray evolution, fuel–air mixing, ignition and combustion. The Zeldovich mechanism is used for predicting NO emissions. Soot formation and oxidation is calculated with a semiempirical, two-rate equation model. Unburned HC emissions models account for three major HC sources in DI diesel engines: (1) leaned-out fuel during the ignition delay, (2) fuel yielded by the sac volume and nozzle hole, and (3) overpenetrated fuel. The emissions models have been validated against experimental data obtained from representative heavy-duty DI diesel engines. It is shown that the models can predict the emissions with reasonable accuracy. Following validation, the usefulness of the cycle simulation as a practical design tool is demonstrated with a case study of the effect of the discharge coefficient of the injector nozzle on pollutant emissions. [DOI: 10.1115/1.2056027]

Introduction

Diesel engines are known to be the most efficient energy converter of fossil fuels for vehicle propulsion, with a projected thermal efficiency potential of 55%. Recently the diesel engine has been evolving from being a rough, noisy and polluting engine to a high-technology, clean propulsion system. Assisted by technologies such as advanced control systems, high-pressure, multiple, fuel injections, new boosting systems, exhaust gas recirculation, and novel aftertreatment systems, the next generation, advanced diesel engines have the potential for being not only fuel economical but also environmentally friendly. As the available technical options for low emissions and high fuel economy increase, exploring the options through hardware experiments alone would not only be a very expensive and time-consuming proposition, but would also run the risk of not yielding a globally-optimized solution at the systems level. Instead, use of validated simulation tools would offer an efficient design methodology for optimizing inherent tradeoffs and hence achieving the desired performance, fuel economy, and emissions targets.

While zero-dimensional models of diesel combustion have been successfully used for predictions of engine performance and fuel economy [1–4], quasi-dimensional, multi-zone models [1,5–17] and multi-dimensional models [18–23] are better suited for predictions of diesel engine emissions. Among them, quasidimensional, multizone models can be effectively used as an engine development tool because they combine some of the advantages of zero-dimensional and multidimensional models. Unlike multidimensional models, only the fuel spray is divided into a number of zones, thus significantly reducing the computational load on calculating ambient gas conditions. In addition, quasidimensional models depend on a blend of fundamental theory, similarity considerations, and experimental data to describe spray penetration and evolution instead of solving the full momentum equation, which is one of the dominant reasons for the computational inefficiency of multidimensional models. Accordingly, quasidimensional models can provide sufficient spatial information to predict emissions while requiring significantly less computing resources compared to multidimensional models.

Most of the earlier multizone concept models (e.g., [5–8]) were focused only on predicting heat release during diesel combustion without including any emission submodels. Emission submodels were later added mainly for NO and soot emissions in a number of studies [9–13,15–17]. However, diesel engine cycle simulations with submodels for an HC emission prediction are rarely found in the literature. This is probably because of the fact that HC emissions from diesel engines are relatively smaller than those from spark-ignition (SI) engines. Hence, HC emissions from diesel engines had not received much attention until regulations became more stringent. Another reason stems from the fact that the quantification of unburned HC emissions generated from various sources is not straightforward.

Among the sources that cause HC emissions from diesel engines, overmixing of fuel beyond the lean flammability limit during the ignition delay period and undermixing of fuel dripping from the nozzle sac volume after the completion of injection are the most important factors under normal operating conditions [24]. In addition to those factors, the effect of wall quenching due to spray overpenetration cannot be ignored, especially in the case of small DI diesel engines [25,26]. Cyclic misfire and local undermixing influence HC emissions only under extreme operating conditions like cold start or extended idling. The contribution of fuel absorption/desorption in the lubricating oil is relatively small in diesel engines since the fuel is not premixed and only a limited portion of the fuel spray contacts the cylinder wall. In SI engines, premixed air–fuel mixture in the crevices is a major source of unburned HC. However, the crevice volume effect on HC emissions is negligible in diesel engines because the diesel spray remains in the piston bowl and rarely reaches crevice regions. Accordingly, the major sources of HC emissions can be identified as:

Contributed by the Internal Combustion Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received February 9, 2004; final manuscript received June 20, 2005. Assoc. Editor: D. Assanis.

(1) over-leaned fuel during the ignition delay, (2) undermixed fuel yielded by the sac volume and nozzle hole, and (3) overpenetrated fuel.

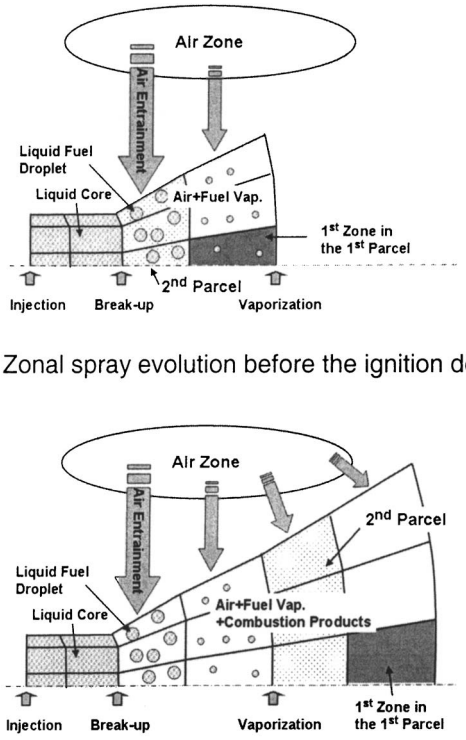
Kuo and his co-workers [26] estimated the contributions of overmixing and overpenetration on HC emissions from a diesel engine using a gas-jet combustion model. Lakshminarayanan, et al. [27] estimated HC emissions due to overmixing and sac volume effects based on their spray jet model. These two studies suggested practical semiempirical models for the estimation of HC emissions, however their works were limited in scope and not incorporated in an engine cycle simulation. On the other hand, Bazari [17] attempted to predict HC emissions using a multizone engine cycle model, based on the assumption that the difference injected and oxidized fuel is exhausted as HC emissions. However, this simplified approach produced predictions that could not follow the trend of experimental data. Carter et al. [28] developed submodels for overmixing and overpenetration effects and implemented them into an engine simulation program developed at the University of Bath. Nevertheless, their study did not account for the sac volume effect and several critical conditions, including equivalence ratio distribution, mixing rate coefficients and lean combustion limit, needed to be specified by the user as input to their model.

Our objective in the present study is to develop comprehensive models for predicting NO, soot, and unburned HC emissions from a direct-injected diesel engine using a quasidimensional, multizone engine cycle simulation [1]. Special emphasis is placed on considering all three major sources of HC emissions in DI diesel engines: (1) leaned-out fuel during the ignition delay, (2) fuel yielded by the sac volume and nozzle hole, and (3) overpenetrated fuel. The NO, soot, and HC emissions models are thoroughly validated with experimental data obtained from representative heavy-duty, DI diesel engines over a wide range of operation conditions. Subsequently, a case study of the effect of the injector nozzle discharge coefficient on pollutant emissions is conducted to demonstrate the usefulness of the cycle simulation as a practical design tool.

Quasidimensional Multizone Combustion Model

In this section, the concept of the quasidimensional, multizone, DI diesel combustion simulation developed by the authors [1] is briefly described for completeness of the paper. Figure 1 illustrates the evolution of fuel parcels and zones in the multizone model. Fuel injected into the combustion chamber according to the fuel injection schedule forms a parcel during each time step that moves in the spray axial direction. Each fuel parcel is further divided into small zones that are distributed in the radial direction. The total number of zones in the radial direction is fixed regardless of the amount of fuel injected or the time step used. However, the total number of zones in the spray direction equals the number of spray parcels, and is therefore determined by the injection duration and the computational time step size.

The fuel injected into the chamber is initially assumed to form a liquid core that travels at a speed equal to the fuel injection speed until the fuel breakup time elapses. After that, the injected fuel is distributed within a spray angle that is unique to each spray parcel and varies from one time step to another, depending on the conditions. The zone angle, i.e., the injection direction of each zone, is determined by dividing the spray angle with the number of radial zones. Each zone can be located relative to the injector hole by tracking the zone angle and penetration of each zone. The velocity of each zone is calculated by temporal differentiation of the correlation for spray tip penetration. Since quasidimensional models depend on empirical correlations to describe spray penetration instead of solving the full momentum equation for computational efficiency, it is very important to use proper spray penetration correlations. In this study, the correlations proposed by



(a) Zonal spray evolution before the ignition delay

(b) Zonal spray evolution after the ignition delay

Fig. 1 The evolution of fuel parcels and zones in the multizone model

Hiroyasu et al. [9] were modified to generalize their correlations so as to handle nozzles with arbitrary discharge coefficients. The new correlations are as follows:

(a) Before breakup, $0 < t < t_b$

$$S = C_d\left(\frac{2\,\Delta P}{\rho_f}\right)^{0.5} t \tag{1}$$

(b) after breakup, $t_b \leqslant t$

$$S = 2.95\left(\frac{\Delta P}{\rho_a}\right)^{0.25} (d_n t)^{0.5} \tag{2}$$

where breakup time, t_b, is

$$t_b = 4.351\frac{\rho_f d_n}{C_d^2(\rho_a\,\Delta P)^{0.5}} \tag{3}$$

To validate the spray penetration correlations, high-pressure chamber experimental data collected by Dan et al. [29] were compared with our predictions. The injector nozzle has a minisac volume design with 0.2 mm diameter and 5.5 l/d ratio for high injection pressure. The injection pressure was varied from 55 to 120 MPa. The ambient gas was nitrogen and the fuel was *n*-tridecane ($C_{13}H_{28}$). More detailed injector specification, experiment conditions, and fuel properties can be found in [29]. Figure 2 compares predictions with measurements of the spray tip penetration as a function of time from start of injection and injection pressure. It can be concluded that spray tip penetrations predicted by the multizone model exhibit good overall agreement with experimental data over a wide range of injection pressures. One notable discrepancy is that the multizone model tends to overpredict spray tip penetration in the very early stages of fuel injection.

Following breakup, it is assumed that the fuel spray atomizes to fine droplets, each with a diameter equal to the Sauter mean di-

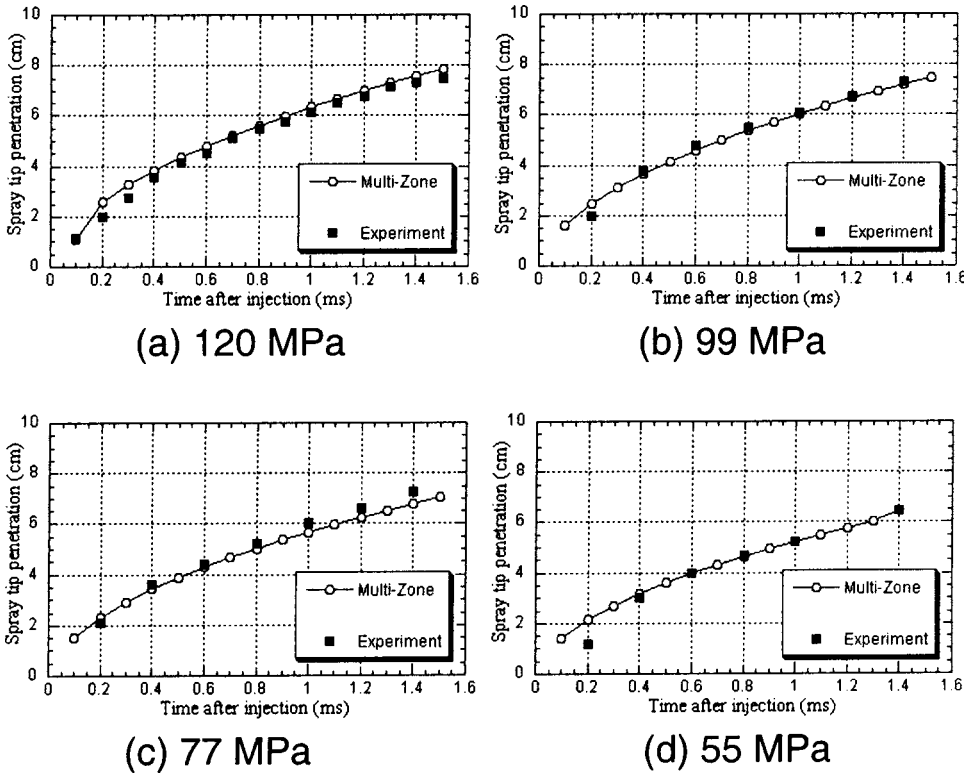


Fig. 2 A comparison of predicted and measured temporal spray tip penetration for a range of injection pressures

ameter (SMD). The effect of droplet size distribution in a spray parcel is neglected. However, droplet sizes in different parcels may vary according to the cylinder conditions at the moment of injection. All calculations related to droplet evaporation are based on SMD.

The air entrainment rate depends on the physical position of each zone, with centerline zones receiving less and edge zones receiving more air. The amount of entrained air is calculated based on conservation of momentum applied to each zone. It is assumed that the momentum of the zone at any instant is equal to the momentum imparted in the zone upon nozzle exit. Since the mass of fuel injected in each zone and its injection velocity are prescribed, and its subsequent velocity can be calculated, the amount of air entrained can be computed by the momentum conservation equation. The equation for air entrainment rate can be obtained by [1]:

$$\dot{m}_a = -\frac{m_f u_i}{(dS/dt)^2}\frac{d^2S}{dt^2} \tag{4}$$

It is assumed that fuel droplets begin to evaporate immediately after breakup occurs. Both heat and mass transfer for a single evaporating droplet are considered in order to compute the instantaneous droplet temperature, rate of evaporation, and droplet diameter.

Combustion is assumed to start individually in each zone after the lapse of the ignition delay period. The ignition delay is measured from the start of injection and is calculated based on zonal temperature and pressure. During the ignition delay period, some of the injected fuel is evaporated and mixed with air, forming a combustible mixture. In the early stage, combustion occurs under premixed conditions. Premixed combustion is assumed to occur until the amount of fuel evaporated at the end of the ignition delay period has been consumed. When the entire initial fuel vapor has been consumed, combustion is assumed to be controlled by diffusion of air into fuel zones.

Each zone is treated as an open system, with mass, energy, and species equations solved for each control volume. Instead of solving the full momentum equation, spray penetration is tracked by empirical correlations. Momentum conservation is used to calculate the amount of air entrained into each zone, as explained above. Critical submodels, such as for the heat release and emissions processes, contain empirical constants that need to be calibrated against experiments. The detail calibration process is described in [1] and the same constants are used in this study.

Consistent with the scope of this paper, the subsequent sections are devoted to the development, validation and use of the NO, soot, and HC emissions models.

Pollutant Emissions

Nitric Oxide Emissions. While nitric oxide (NO) and nitrogen dioxide are usually grouped together as NO_x emissions, NO is predominant in diesel engines [24]. Therefore, only NO formation is considered in the present study. The principal source of NO is the oxidation of atmospheric nitrogen. The formation and destruction processes are not part of the fuel combustion process. However, the reactions that produce NO take place in an environment created by the combustion reactions, so the two processes are still intimately linked. Nitric oxide forms throughout the high-temperature burned gases behind the flame through chemical reactions involving nitrogen and oxygen atoms and molecules, which do not attain chemical equilibrium. The principal reactions governing the formation of NO from molecular nitrogen and its destruction are [30]:

$$N + NO \overset{k_1}{\leftrightarrow} N_2 + O \tag{5}$$

$$O + NO \overset{k_2}{\leftrightarrow} O_2 + N \tag{6}$$

$$N + OH \overset{k_3}{\leftrightarrow} NO + H \tag{7}$$

where forward rate constants, k_1, k_2, and k_3 are in m^3/kmole/s:

$$k_1 = 1.6 \times 10^{10} \tag{8}$$

$$k_2 = 1.5 \times 10^6 T \exp\left(-\frac{19500}{T}\right) \tag{9}$$

$$k_3 = 4.1 \times 10^{10} \tag{10}$$

This is often called the extended Zeldovich's mechanism. Zeldovich was the first to suggest the importance of reactions (5) and (6). Lavoie et al. [30] added reaction (7) to the mechanism. The rate of change of NO concentration is expressed as [24]:

$$\frac{d[NO]}{dt} = \frac{2R_1\{1 - ([NO]/[NO]_e)^2\}}{1 + ([NO]/[NO]_e)R_1/(R_2 + R_3)} \tag{11}$$

where

$$R_1 = k_1[NO]_e[N]_e \tag{12}$$

$$R_2 = k_2[NO]_e[O]_e \tag{13}$$

$$R_3 = k_3[N]_e[OH]_e \tag{14}$$

[ ] in Eq. (11)–Eq. (14) denotes species concentration in kmole/m^3 and the subscript e denotes equilibrium concentration. The NO formation in each zone can be calculated with Eq. (11). Then total amount of NO is calculated by summing up NO formed in each zone.

Soot Formation and Oxidation. Soot forms in the rich unburned fuel-containing core of the sprays, within the flame region, where the fuel vapor is heated by mixing with hot burned gases. Soot then oxidizes in the flame zone when it contacts unburned oxygen. Therefore, the concentration of soot in the exhaust is governed by the formation and oxidation of soot during the engine cycle, i.e.,

$$\frac{dm_s}{dt} = \frac{dm_{sf}}{dt} - \frac{dm_{so}}{dt} \tag{15}$$

The general fact that the net soot formation rate is primarily affected by pressure, temperature, and equivalence ratio has been fairly well established. However, the details of the mechanism leading to soot formation are not known. Consequently, semi-empirical, two-rate equation models have often been used to describe soot dynamics. In particular, the soot formation model proposed by Hiroyasu et al. [9] is used in many multizone models. The formation rate is calculated by assuming a first-order reaction of vaporized fuel, m_{fg}, as:

$$\frac{dm_{sf}}{dt} = A_f m_{fg} P^{0.5} \exp\left(\frac{-E_{sf}}{\bar{R}T}\right) \tag{16}$$

The soot oxidation is predicted by assuming a second-order reaction between soot, m_s, and oxygen.

$$\frac{dm_{so}}{dt} = A_o m_s \frac{P_{ox}}{P} P^{1.8} \exp\left(\frac{-E_{so}}{\bar{R}T}\right) \tag{17}$$

where $E_{sf}=5.23\times10^4$ kJ/kmol, $E_{so}=5.86\times10^4$ kJ/kmol. A_f and A_0 are constants that are determined by matching the calculated soot with the measured data in the exhaust gas. The total amount of soot emitted is obtained by summing up soot mass calculated with Eq. (15) in each zone.

Hydrocarbon Emissions. In the following subsections, phenomenological models are introduced to describe the three major sources of HC emissions from diesel engines: (1) leaned-out fuel during the ignition delay, (2) fuel yielded by the sac volume and nozzle hole, and (3) overpenetrated fuel.

Effect of Overmixing. Some of the fuel injected into the combustion chamber during the ignition delay period is vaporized and mixed with air beyond a lean limit of equivalence ratio, ϕ. The value of ϕ_1 is determined by the lean limit for flammability or by the lean limit for self-ignition (whichever gives the lower value of ϕ_1) at the prevailing pressures (3–10 MPa) and temperature (800–1500 K) of the lean mixtures in the diesel combustion chamber. Under these conditions, the lean flammability limit is expected to be on the order of $\phi=0.1-0.3$ and to be relatively insensitive to the range of temperatures and pressures at present of interest. The data on lean limit for self-ignition suggest a similar insensitivity to the range of temperatures and pressures of interest, so that as a first approximation it can be assumed that ϕ_1 is constant [31]. The lean limit is assumed 0.2 in this study. Note that while the over-leaned fuel can escape the primary combustion process as unburned or only partially reacted, thermal oxidation of some of these hydrocarbons can occur during the expansion process.

The first step for the calculation of the HC emissions from the over-leaned mixture is to estimate the over-leaned fuel at the start of combustion. However, a challenge for multizone spray models is that they cannot calculate the amount of over-leaned fuel precisely unless the spatial resolution is extremely high. Since increasing resolution will increase computational load, a different approach is proposed in this study. The average fuel concentration of a parcel, which is a set of zones introduced to the combustion chamber at a given time step, can be calculated using the information from the multizone spray model as follows:

$$\bar{C} = \frac{\sum_{i=1}^{nz} m_{f,i}}{\sum_{i=1}^{nz} m_{f,i} + \sum_{i=1}^{nz} m_{a,i}} \tag{18}$$

where $\bar{C}=\int c(r)dr/r_m$. The local fuel concentration $C(r)$ within each parcel is assumed as proposed by Yu et al. [32].

$$\frac{C(r)}{C_c} = 1 - \left(\frac{r}{r_m}\right)^{1.5} \tag{19}$$

The radius of the spray which is r_m in Eq. (19) can be calculated with spray penetration and spray angle. Since the average fuel concentration within a parcel is known from Eq. (18), the fuel concentration at the centerline, C_c, and the fuel distribution can be calculated from Eq. (19). Accordingly, the equivalence ratio distribution can be calculated and the fuel mass beyond the lean flammability limit can be obtained for each parcel.

Some of the over-leaned fuel is then oxidized during the expansion process. In general, the reaction rate can be represented in the form of

$$\frac{d[HC]}{dt} = -A_1[HC]^a[O_2]^b \exp\left(\frac{-E}{\bar{R}T}\right) \tag{20}$$

For SI engines, HC oxidation rates have been determined in a number of different studies and several empirical correlations of the data in the form of overall reaction rate equations under exhaust gas conditions (near stoichiometric air-fuel ratio and atmospheric pressure) have been proposed [33]. However, correlations for diesel-like conditions are rare. Furthermore the correlations for SI engines cannot be applied to diesel engine cases since the fuel composition, conditions and processes of HC emissions formation in diesel engines are substantially different. For the over-leaned condition, the variation of the concentration of the oxygen will be very small [31], so a simplified expression of the following form is assumed:

$$\frac{d[\mathrm{HC}]}{dt} = -A_2[\mathrm{HC}]^a \exp\left(\frac{-E}{\bar{R}T}\right) \tag{21}$$

By using the following relation,

$$[i] = \frac{m_i}{M_i V} \tag{22}$$

Eq. (21) can be rewritten as

$$\frac{dm_{\mathrm{HC}}}{dt} = -A_3 \frac{m_{\mathrm{HC}}^a}{V} \exp\left(\frac{-E}{\bar{R}T}\right) \tag{23}$$

where A_3 is 1.161×10^{11}; a is 2; and E is 7.79×10^4 kJ/kmol.

The total mass of the overleaned fuel within each parcel at the start of combustion is used as the initial HC value for the time integration of Eq. (23). Hence, the final HC emissions from the overleaned fuel can be obtained by integrating Eq. (23) over the expansion stroke:

$$m_{\mathrm{HC,ol}} = \int_{\exp} \frac{dm_{\mathrm{HC}}}{dt} \tag{24}$$

Effect of Sac and Nozzle Hole Volume. At the end of the injection period, the injection pressure decreases rapidly. Consequently, right after the needle seats, the pressure in the small volume in the tip of the injector, known as the sac volume, could be lower than the cylinder pressure. Consequently, some fuel is retained in the sac volume and nozzle holes. However, the sac volume is not always completely full of fuel due to cavities induced by low pressures in the sac when injection terminates rapidly. As combustion and expansion proceed, this residual fuel issues from the nozzle sac as large liquid drops at a very low rate with relatively low momentum. This fuel evaporates and mixes very slowly, and hence the corresponding mixture element is fuel rich until too late in the engine cycle. Therefore, most of the fuel from the sac and hole volume is assumed to appear as HC emissions.

In order to account for the fact that the sac and nozzle hole volume is not fully filled with fuel due to the cavitation, a fraction factor is introduced as in Eq. (25). The value of the fraction factor, c_{sac}, is assumed to be 0.12, as suggested by Lakshminarayanan et al. [27].

$$m_{\mathrm{HC,sac}} = c_{\mathrm{sac}} \rho_f V_{\mathrm{sac}} \tag{25}$$

Effect of Overpenetration. In case of large-bore engines, the overleaned fuel during the ignition delay period and undermixed fuel from the sac and nozzle holes post injection are known to be the major sources of HC emissions [31,32]. However, in smaller bore engines, Kuo et al. [26] found that the contribution of the overleaned fuel to HC emissions could be only 10% to 30% of the total HC emissions. Another significant source of HC emissions from small diesel engines applied to passenger cars is identified as the HC emissions from the overpenetrated fuel. The overpenetrated fuel includes fuel vapor deflected by the walls, the liquid fuel rebounded from the walls, and the liquid fuel stuck on the walls. A part of the overpenetrated fuel will be consumed in the combustion process. However, the remainder of the impinging fuel, and especially the fraction stuck on the piston bowl, could potentially be emitted from the engine as HC emissions. Kuo et al. [26] found that this could account for 0% to 65% of the total HC emissions from a small bore engine. They also observed that approximately 2% of the fuel mass that overpenetrated before the start of combustion ended up as unburned HC emissions. Figure 3 illustrates an overpenetrated fuel spray zone.

Because of the difficulty of decomposing the overpenetrated fuel into individual components, all of the overpenetrated fuel was used to quantify the extent of wall impingement. To calculate the mass of overpenetrated fuel, the spray zones are initially allowed to penetrate without the restriction of wall boundaries; a subsequent geometric check, flags the zones that would pass the walls as overpenetrated zones. The total mass of overpenetrated fuel is the summation of fuel mass from all overpenetrated zones. Accordingly, HC emissions due to overpenetration are given by

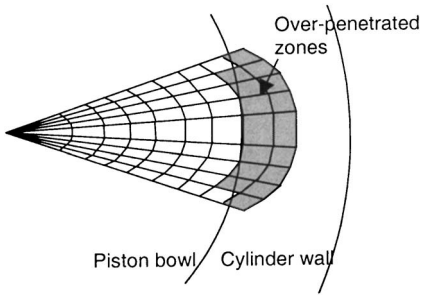


Fig. 3 Overpenetrated fuel spray

$$m_{\mathrm{HC},op} = c_{op} \sum_{i=op}^{\mathrm{zones}} m_{f,i} \tag{26}$$

where c_{op} is the fraction of the overpenetrated fuel that is emitted as HC emissions, which is assumed to be 0.02.

Model Validation

Experimental Data. NO and HC emissions predictions from the multizone model have been compared against corresponding measurements on an in-line, six cylinder, turbocharged, intercooled, water-cooled, direct injection diesel engine at the Automotive Research Center located in the University of Michigan. Additional validation of the multizone model's predictions was conducted using reported data from an independent engine setup at the Engine Research Center of the University of Wisconsin [34]. The latter work on a modern single-cylinder, four-diesel stroke engine was selected for the completeness in reporting soot and NO emissions over a range of injection timings. Primary specifications of the two engines are listed in Table 1. The multicylinder, DDC series 60 engine has a relatively quiescent low-swirl combustion chamber with a shallow "Mexican hat" bowl-in piston and produces very high injection pressures, delivered by unit injectors. The fuel injection timing and duration are electronically controlled. The single cylinder engine setup is based on a version of the Caterpillar 3406 production engine with simulated turbocharging.

Nitric Oxide Emissions. To verify the accuracy of the model over a wide range of operating conditions, predicted NO emissions were compared with experimental data acquired on the multicylinder engine for a range of loads varying from 10% to 100%, at a speed of 2100 rpm. As shown in Fig. 4, the model predicted very satisfactorily the trend of NO variation with changes of engine load, and also showed good quantitative agreement with experimental data. The fidelity of the model has been explored further by comparing its predictions with data over a range of speeds at a constant load. The engine speed has been varied from 900 to 2100 rpm in increments of 300 rpm. The load was set at

Table 1 Engine specifications and data availability

Engine	DDC 60	CAT 3406
Type	Multi cylinder	Single cylinder
Bore (mm)	130	137
Stroke (mm)	160	165
Comp. ratio (−)	15	15
Rated speed (rpm)	2100	2100
Rated power (kW)	350	54

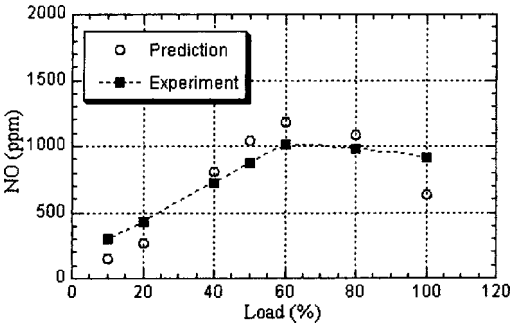


Fig. 4 Predicted and measured NO emissions over a range of loads at 2100 rpm (multicylinder engine)

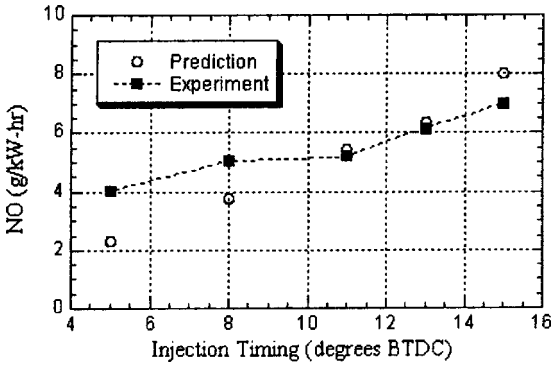


Fig. 6 Predicted and measured NO emissions over a range of injection timings (single-cylinder engine)

50%. As shown in Fig. 5, the model was able to qualitatively predict the trend correctly; however, it overpredicted NO at low engine speeds.

Following model validation with respect to the multicylinder engine, a comparison of predictions with measurements from the single cylinder engine was carried out over a range of injection timings. The operating conditions are listed in Table 2. The conditions were chosen to reflect common loads placed on the engine (1600 rpm and equivalence ratio of 0.45, or approximately 80% load), as well as the appropriate inlet temperature and pressure conditions. Predictions of NO emissions were compared with experimental data for injection timings of 5, 8, 11, 13, and 15 deg BTDC. As shown in Fig. 6, the model predicted the measured trend correctly, although it underpredicted NO at retarded injection timings.

Soot Formation and Oxidation. For the soot model, the preexponential constant, A_f, in Eq. (16) is set here at 150, which is the same value used by Patterson et al. [22]. The other preexponential constant, A_O, in Eq. (17) is calibrated (A_O=2250) to match the soot level with measurements under the baseline operating condition (11 deg BTDC). Comparisons of predicted soot emissions with measurements from the single cylinder engine were carried out over a range of injection timings (from 5 to 15 deg CA BTDC). As shown in Fig. 7, the predicted values are in excellent agreement with measurements.

Hydrocarbon Emissions. The HC emissions model was validated against experimental data acquired on the multicylinder engine for a range of engine loads and speeds. Figure 8 compares the predicted and measured HC emissions over a range of loads at 2100 rpm. As can be seen in Fig. 8, the model predicted the exact same trend for the HC variation with changes in engine load, as exhibited in the experiments; furthermore, predictions showed very good quantitative agreement with experimental data. Similarly, Fig. 9 indicates that the model is capable of predicting accurately HC emissions over a range of engine speeds at 50% load.

In Fig. 10, the contribution of the sources to HC emissions was compared. Clearly, overmixing during the ignition delay period is the most dominant source of HC emissions in this case. The

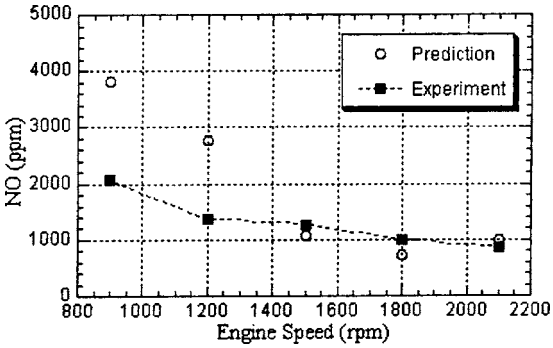


Fig. 5 Predicted and measured NO emissions over a range of speeds at 50% load (multicylinder engine)

Table 2 Operating conditions

Parameter	Value
Engine speed (rpm)	1600
Equivalence ratio	0.45
Injection timing (deg. BTDC)	5, 8, 11, 13, 15
Inlet air temperature (deg. C)	36
Inlet air pressure (kPa)	184
Exhaust back pressure (kPa)	159

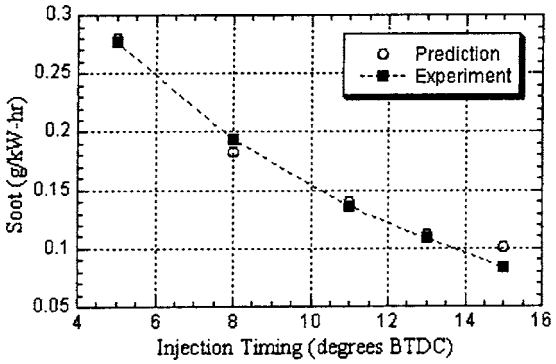


Fig. 7 Predicted and measured soot emissions over a range of injection timings (single-cylinder engine)

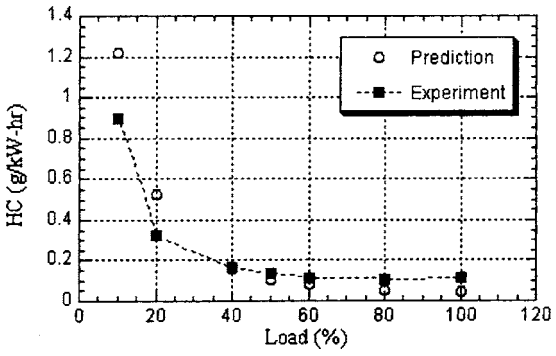


Fig. 8 Predicted and measured HC emissions over a range of loads at 2100 rpm (multicylinder engine)

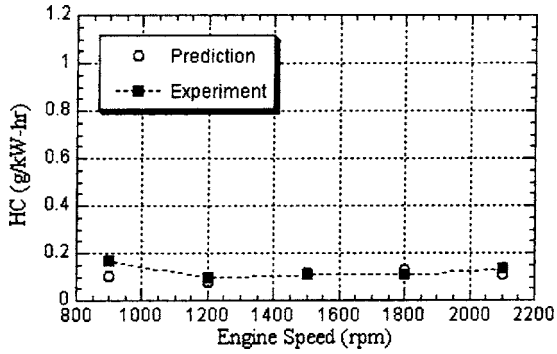


Fig. 9 Predicted and measured HC emissions over a range of speeds at 50% load (multicylinder engine)

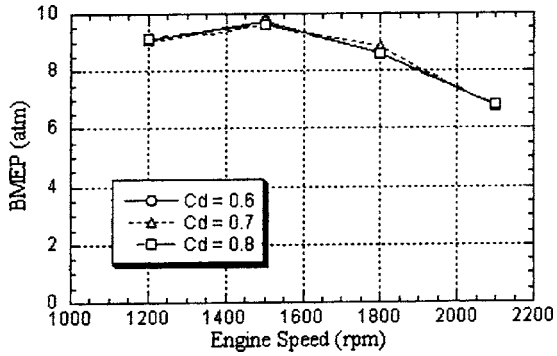


Fig. 11 The effect of injector nozzle discharge coefficient; Brake Mean Effective Pressure (BMEP)

nozzle of the engine tested is of the VCO (valve covered orifice) type, which has no sac volume. Accordingly, HC emissions due to sac and hole volumes are smaller than the HC caused by the overmixing effect. It should be noted that the results showed that no effect of overpenetration, with the exception of the 10% load condition. This is reasonable because the tested engine is a relatively large-bore, heavy-duty engine.

The Effect of Injector Nozzle Discharge Coefficient on Performance and Emissions

Regardless of the application and size of a given engine and the type of fuel injection system used, the fuel injector design is of paramount importance in DI diesel engines. More specifically, injector nozzle design parameters, such as hole diameter, length/diameter ratio, hole inlet chamfer radius, and sac volume play an extremely important role in the mixture preparation and combustion processes. In addition, variation in the machining of injector holes can affect spray and mixing in two ways: the variation of main design parameters, e.g., the hole diameter, will result in different flow conditions, while microvariations of the hole surface or variations of the roughness of edges at the hole entrance will lead to change of flow resistance, and accordingly discharge coefficient. Yoda and Tsuda [35] simulated the effect of deburred inlet corner of the injector hole, which is also referred to as a nozzle inlet chamfer radius, on the flow discharge coefficient with a CFD code. According to their results, the discharge coefficient can be increased from 0.6 (which is the case with sharp corner) to as much as 0.84. This implies that deburring the inlet corner of the nozzle can improve the discharge coefficient significantly and eventually alter the spray evolution and combustion processes. Depending on the injector hole diameter, the discharge coefficient and injection pressure spray penetration and air entrainment are altered and eventually combustion rates are changed. Even though the effect of microscopic variation of the hole surface, or the variation of the roughness of edges at the hole entrance, cannot be explicitly simulated with the quasidimensional model, the effect of those microscopic variations can be represented by the discharge coefficient, as previously done by Yoda and Tsuda [35]. In the present study, the effect of the discharge coefficient of the injector nozzle on engine performance and emissions was investigated in order to demonstrate the application of the cycle simulation as a practical numerical tool for engine design and development.

The numerical experiments for the study of the effect of the discharge coefficient on engine performance and emissions were performed by modeling the DDC 60 series, six cylinder, DI diesel engine. Basic engine specifications were given in Table 1. All input parameters were kept the same, except for the discharge coefficient, which was varied from 0.6 to 0.8. For each value of the discharge coefficient, engine performance and emissions were predicted for engine speeds ranging from 1200 to 2100 rpm at 50% load.

As shown in Fig. 11, engine performance, which is represented by brake mean effective pressure (BMEP), does not show any significant changes as the discharge coefficient is varied over the range of engine speeds. This is primarily because of the fact that the amount of fuel injected was kept the same for all values of the discharge coefficient. Predicted HC emissions are relatively more sensitive to the change of the discharge coefficient compared to engine performance, as shown in Fig. 12. However, NO and soot emissions are much more sensitive to changes in the discharge coefficient compared to the sensitivity of HC emissions. NO-soot tradeoff curves for each discharge coefficient are plotted in Fig. 13. Clearly, both NO and soot emissions are reduced as the discharge coefficient is increased. This interesting result can be explained with the following reasoning. In Fig. 14, the accumulated mass of air entrained into the spray is plotted for the different

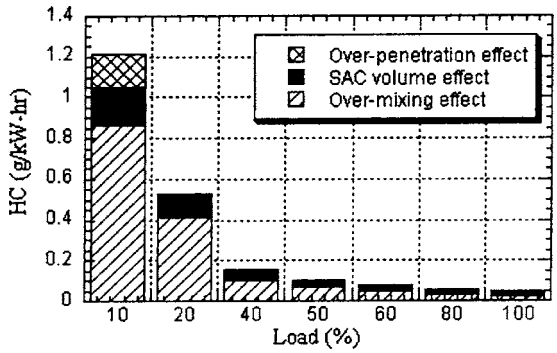


Fig. 10 Sources of HC emissions over a range of loads at 2100 rpm (multicylinder engine)

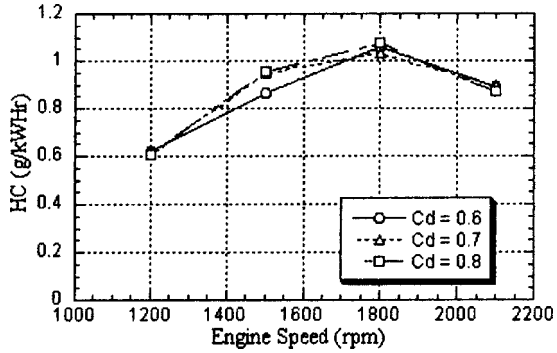


Fig. 12 The effect of injector nozzle discharge coefficient; HC emissions

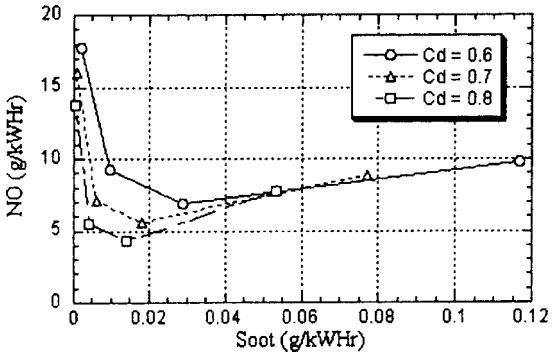


Fig. 13 The effect of injector nozzle discharge coefficient: NO-soot tradeoff curves

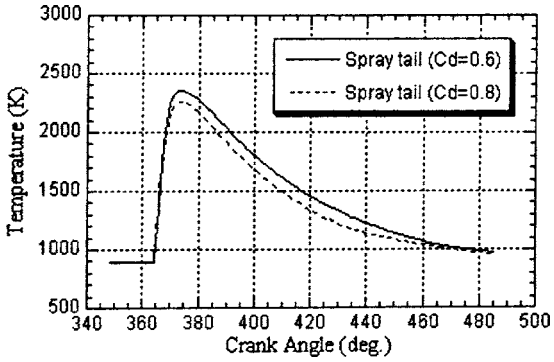


Fig. 16 The effect of injector nozzle discharge coefficient: Spray tail temperature

values of the discharge coefficient. The figure shows that the rate of air entrainment into the spray increases as the discharge coefficient increases. Accordingly, the temperature of the spray drops, which is evidenced by Figs. 15 and 16. Therefore, fuel and air mixing is improved by increasing the discharge coefficient. This increases the available air for soot oxidation and lowers the temperature of the spray, which, in turn, reduces NO formation. However, increased fuel and air mixing and lower temperatures generally cause increased HC emissions for higher discharge coefficients.

Conclusions

Phenomenological models for DI diesel engine emissions, including NO, soot, and unburned HC, have been implemented into a full engine cycle simulation based on a quasidimensional, multizonal spray combustion model. The following three major sources of HC emissions have been modeled in this study: (1) leaned-out fuel during the ignition delay, (2) fuel yielded by the sac volume, and (3) overpenetrated fuel. A comprehensive validation of the proposed models against experimental data obtained from representative heavy-duty DI diesel engines has been conducted over a range of engine speeds, loads, and injection timings. The application of the validated simulation has been demonstrated with a case study of the effect of the injector nozzle discharge coefficient on engine performance and emissions. The major conclusions are the following:

- The extended Zeldovich model predicted satisfactorily the observed trends of NO variation with changes in engine load, speed, and injection timing for both multi- and single cylinder engines. The comparison of predicted and measured data also demonstrated reasonable quantitative agreement between them.
- Predicted soot emissions based on semiempirical, two-rate equation models showed excellent agreement with measurements from a single cylinder engine over a range of injection timings.
- The model showed promising capability in predicting HC emissions accurately over a range of engine speeds and loads. The specific large bore engine examined in this study revealed that overmixing during the ignition delay period accounted for the bulk of the HC emissions, as the fuel spray did not overpenetrate in most cases.
- Additional effort is required to assess the fidelity of each model across a wider range of operating conditions and engine types, including high-speed, small-bore diesel engines.
- The case study showed that improving the discharge coefficient of the injector nozzle hole can reduce NO and soot emissions simultaneously, while maintaining the same engine performance. The NO and soot benefit comes at a modest penalty in HC emissions.
- The case study demonstrated that the quasidimensional, multizone engine cycle simulation model with the emissions submodels has the capability to be used as an effective design and development tool to predict engine performance and emissions.

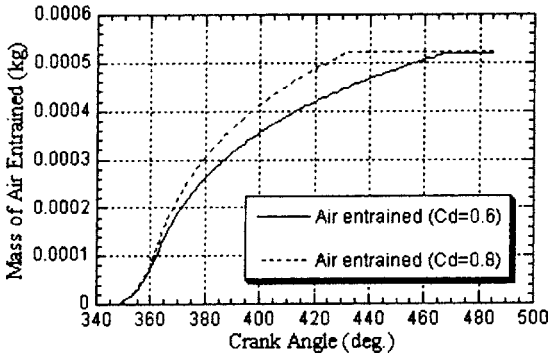


Fig. 14 The effect of injector nozzle discharge coefficient: Air entrainment

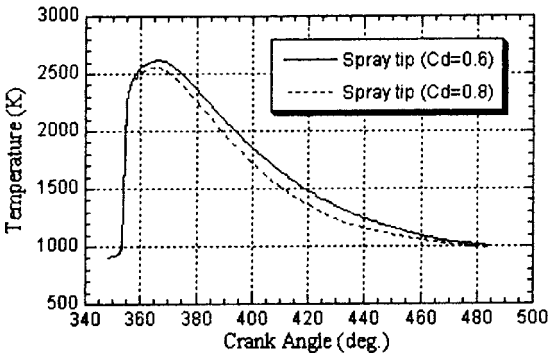


Fig. 15 The effect of injector nozzle discharge coefficient: Spray tip temperature

Acknowledgment

The authors would like to acknowledge the technical and financial support of the Automotive Research Center (ARC) by the National Automotive Center (NAC) located within the U.S. Army Tank-Automotive Research, Development and Engineering Center (TARDEC). The ARC is a U.S. Army Center of Excellence for Automotive Research at the University of Michigan, currently in partnership with seven other universities nationwide. Timothy J.

Jacobs and Dr. Zoran Filipi are gratefully acknowledged for providing data acquired on the ARC heavy-duty engine test cell for use in model validation.

Nomenclature

A = constant
$\bar{C}$ = average fuel mass concentration
C = mass concentration
C_d = discharge coefficient
c = fraction factor
d_n = nozzle hole diameter
E = activation energy
H = hydrogen
HC = hydrocarbon
k = forward rate constant
M = molecular weight
m = mass
N = nitrogen
NO = nitric oxide
nz = number of zones in a parcel
O = oxygen
OH = hydroxyl
P = pressure
R = gas constant
$\bar{R}$ = universal gas constant
r = radial distance from the center line
S = spray penetration
T = temperature
t = time
u_i = injection velocity
V = volume
ρ = density

Superscripts

a = exponent
b = exponent

Subscripts

a = air
c = centerline
e = equilibrium
exp = expansion stroke
i = index
f = fuel
m = maximum
ol = overleaned
op = overpenetration
ox = oxygen
s = soot emitted
sac = sac volume
sf = soot formed
so = soot oxidized

References

[1] Jung, D., and Assanis, D. N., 2001, "Multi-Zone DI Diesel Spray Combustion Model for Cycle Simulation Studies of Engine Performance and Emissions," *SAE Transactions: Journal of Engines*, SAE Paper No. 2001-01-1246.

[2] Assanis, D. N., and Heywood, J. B., 1986, "Development and Use of a Computer Simulation of the Turbocompounded Diesel System for Engine Performance and Component Heat Transfer Studies," SAE Paper No. 860329.

[3] Krieger, R. B., and Borman, G. L., 1966, "The Computation of Apparent Heat Release From Internal Combustion Engines," ASME Paper No. 66-WA/DGP-4.

[4] Foster, D. E., 1985, "An Overview of Zero-Dimensional Thermodynamic Models for I.C. Engine Data Analysis," SAE Paper No. 852070.

[5] Whitehouse, N. D., and Sareen, B. K., 1974, "Prediction of Heat Release in a Quiescent Chamber Diesel Engine Allowing for Fuel/Air Mixing," SAE Paper No. 740084.

[6] Chiu, W. S., Shahed, S. M., and Lyn, W. T., 1976, "A Transient Spray Mixing Model for Diesel Combustion," SAE Paper No. 760128.

[7] Merguerdichian, M., and Watson, N., 1978, "Prediction of Mixture Formation and Heat Release in Diesel Engines," SAE Paper No. 780225.

[8] Dent, J. C., and Mehta, P. S., 1981, "Phenomenological Combustion Model For a Quiescent Chamber Diesel Engine," SAE Paper No. 811235.

[9] Hiroyasu, H., Kadota, T., and Arai, M., 1983, "Development and Use of a Spray Combustion Modeling to Predict Diesel Engine Efficiency and Pollutant Emissions (Part 1 Combustion Modeling)," Bull. JSME, **26**, pp. 569–575.

[10] Hiroyasu, H., Kadota, T., and Arai, M., 1983b, "Development and Use of a Spray Combustion Modeling to Predict Diesel Engine Efficiency and Pollutant Emissions (Part 2 Computational Procedure and Parametric Study)," Bull. JSME, **26**, pp. 576–583.

[11] Bazari, Z., 1992, "A DI Diesel Combustion and Emission Predictive Capability for Use in Cycle Simulation," SAE Paper No. 920462.

[12] Yoshizaki, T., Nishida, K., and Hiroyasu, H., 1996, "Approach to Low NO_x and Smoke Emission Engines by Using Phenomenological Simulation," SAE Paper No. 930612.

[13] Morel, T., and Wahiduzzaman, S., 1996, "Modeling of Diesel Combustion and Emissions," *96 FISITA Proceedings*, 26th International Congress, Praha, Czech Republic, 17–21 June.

[14] Bhaskar, T., and Mehta, P. S., 1998, "A Multi-Zone Diesel Combustion Model Using Eddy Dissipation Concept," *Proceedings of the Fourth International Symposium on Diagnostics and Modeling of Combustion in IC Engines*, COMODIA 98, Kyoto, Japan.

[15] Rakopoulos, C. D., and Hountalas, D. T., 1998, "Development and Validation of a 3-D Multi-Zone Combustion Model for the Prediction of DI Diesel Engines Performance and Pollutants Emissions," SAE Paper No. 98102.

[16] Gao, Z., and Schreiber, W., 2000, "A Multizone Analysis of Soot and NO_x Emission in a D.I. Diesel Engine as a Function of Engine Load, Wall Temperature, and Intake Air O_2 Content," ASME Paper No. 2000-ICE-314.

[17] Tauzia, X., Hetet, J.-F., Chesse, P., Inozu, B., and Roy, P., 2000, "The Use of Phenomenological, Multizone Combustion Model to Investigate Emissions From Marine Diesel Engines," ASME Paper No. 2000-ICE-325.

[18] Oran, E. S., and Boris, J. P., 1981, "Detailed Modeling of Combustion Systems," Prog. Energy Combust. Sci., **7**, pp. 1–72.

[19] Bracco, F. V., 1985, "Modeling of Engine Sprays," SAE Paper No. 850394.

[20] Amsden, A. A., Ramshaw, J. D., O'Rourke, P. J., and Dukowicz, J. K., 1985, "KIVA: A Computer Program for Two- and Three-Dimensional Fluid Flow With Chemical Reactions and Fuel Sprays," Los Alamos National Laboratory Report LA-10245-MS.

[21] Amsden, A. A., Butler, T. D., and O'Rourke, P. J., 1987, "KIVA-II Computer Program for Transient Multidimensional Chemically Reactive Flows With Sprays," SAE Paper No. 872072.

[22] Patterson, M. A., Kong, S. C., Hampson, G. J., and Reitz, R. D., 1994, "Modeling the Effects of Fuel Injection Characteristics on Diesel Engine Soot and NO_x Emissions," SAE Paper No. 940523.

[23] Varnavas, C., and Assanis, D. N., 1996, "A High Temperature and High Pressure Evaporation Model for the KIVA-3 Code," SAE Paper No. 960629.

[24] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.

[25] Matsui, Y., and Sugihara, K., 1987, "Sources of Hydrocarbon Emissions From a Small Direct Injection Diesel Engine," SAE Paper No. 871613.

[26] Kuo, T.-W., Wu, K.-J., and Henningsen, S., 1988, "Effects of Fuel Overpenetration and Overmixing During Ignition Delay Period on Hydrocarbon Emissions From a Small Open-Chamber Diesel Engine," ASME J. Eng. Gas Turbines Power, **110**, pp. 453–461.

[27] Lakshminarayanan, P. A., Nayak, N., Dingare, S. V., and Dani, A. D., 2000, "Predicting HC Emissions From DI Diesel Engines," ASME Paper No. 2000-ICE-327.

[28] Carter, S. J., Hawley, G., Slowley, J., and Charlton, S. J., 1996, "Initial Development of a Predictive Hydrocarbon Emissions Model for a Diesel Engine," *Proc. Intersociety of Energy Conversion Engineering Conf.*, Vol. 3, IEEE, pp. 1962–1967.

[29] Dan, T., Takagishi, S., Senda, J., and Fujimoto, H., 1997, "Effect of Ambient Gas Properties for Characteristics of Non-Reacting Diesel Fuel Spray," SAE Paper No. 970352.

[30] Lavoie, G. A., Heywood, J. B., and Keck, J. C., 1970, "Experimental and Theoretical Investigation of Nitric Oxide Formation in Internal Combustion Engines," Combust. Sci. Technol., **1**, pp. 313–326.

[31] Greeves, G., Khan, I. M., Wang, C. H. T., and Fenne, I., 1977, "Origins of Hydrocarbon Emissions From Diesel Engines," SAE Paper No. 770259.

[32] Yu, R. C., Wong, V. W., and Shahed, S. M., 1980, "Sources of Hydrocarbon Emissions From Direct Injection Diesel Engines," SAE Paper No. 800048.

[33] Lavoie, G. A., 1978, "Correlations of Combustion Data for S.I. Engine Calculations-Laminar Flame Speed, Quench Distance and Global Reaction Rates," SAE Paper No. 780229.

[34] 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.

[35] Yoda, T., and Tsuda, T., 1997, "Influence of Injection Nozzle Improvement on DI Diesel Engine," SAE Paper No. 970356.

Post-Combustion In-Cylinder Vaporization During Cranking and Startup in a Port-Fuel–Injected Spark Ignition Engine

Jim S. Cowart
Mechanical Engineering Department,
U.S. Naval Academy,
Annapolis, MD 21402

During port-fuel–injected (PFI) spark-ignition (SI) engine startup and warm-up fuel accounting continues to be a challenge. Excess fuel must be injected for a near stoichiometric combustion charge. The "extra" fuel that does not contribute to the combustion process may stay in the intake port or as liquid films on the combustion chamber walls. Some of this combustion chamber wall liquid fuel is transported to the engine's oil sump and some of this liquid fuel escapes combustion and evolves during the expansion and exhaust strokes. Experiments were performed to investigate and quantify this emerging in-cylinder fuel vapor post-combustion cycle by cycle during engine startup. It is believed that this fuel vapor is evaporating from cylinder surfaces and emerging from cylinder crevices. A fast in-cylinder diagnostic, the fast flame ionization detector, was used to measure this behavior. Substantial post-combustion fuel vapor was measured during engine startup. The amount of post-combustion fuel vapor that develops relative to the in-cylinder precombustion fuel charge is on the order of one for cold starting (0 °C) and decreases to $\sim\frac{1}{3}$ *for hot starting engine cycles. Fuel accounting suggests that the intake port puddle forms quickly, over the first few engine cranking cycles. Analysis suggests that sufficient charge temperature and crevice oxygen exists to at least partially oxidize the majority of this post-combustion fuel vapor such that engine out hydrocarbons are not excessive.* [DOI: 10.1115/1.2061307]

Introduction and Background

As engine emission regulations continue to tighten, the need to focus on improved engine cranking and start-up has become evident. In the early years of fuel-injected engines, engine controller start-up calibrations always errored on the fuel-rich side of fuel delivery, as this is more safe and stable for smooth engine operation and vehicle driveability. However, this approach is no longer acceptable in meeting ultralow emission standards. Current and future engine emission standards require excellent fuel control as engine misfires from the first engine cycle cannot be tolerated. Thus, continued research into the area of spark-ignition (SI) port-fuel–injected (PFI) engine mixture preparation is necessary to further optimize injected fuel control.

One challenging area of PFI engine mixture preparation surrounds engine cranking and startup. In order to attain acceptable combustion, often, many times the required stoichiometric fuel (vapor) mass must be injected. Unfortunately, it is unclear as to where much of this fuel is being delayed or transported. Fuel accounting during startup can be extremely difficult as much of the injected fuel can remain as liquid films in both the intake port and on in-cylinder combustion chamber and piston surfaces. Additionally, some in-cylinder liquid fuel can migrate to the engine's sump and become entrained by the engine's oil. Therefore, engine developers and researchers are frequently uncertain as to the exact paths of injected fuel.

Within this context, this research work explores in-cylinder fuel behavior during engine startup after the main flame has passed (e.g., post-combustion). A fast in-cylinder diagnostic, the fast flame ionization detector (FFID) is used both before and after the main combustion event to discern not only the amount of precombustion fuel vapor, but also in-cylinder fuel that evolves as vapor post-combustion before the exhaust event.

The FFID has been used as an in-cylinder fuel concentration diagnostic over the past 15 years to understand the mixture preparation process. The precombustion FFID voltage signal, which is proportional to the unburned hydrocarbon (UHC) concentration, has been frequently connected to injected fuel behavior in order to more fully understand the formation of combustible mixtures. Numerous studies, principally dealing with the engine warm-up process have enlightened engine engineers and researchers as to the physical nature of mixture preparation [1–4]. More recently, engine cranking and startup processes have been observed [5,6].

The use of in-cylinder FFIDs to observe post-combustion processes has been very limited. An early study qualitatively observed crevice outgassing post-combustion [7]. Another study focused on cylinder wall-layer hydrocarbons using an in-cylinder FFID [8]. More recently the FFID was used for a post-combustion in-cylinder UHC measurement in a stabilized engine [9].

Thus, the objective of this work is to quantify in-cylinder fuel that is vaporized post-combustion to more fully understand in-cylinder liquid fuel behavior during engine start-up. Currently, very little is known about this post-combustion fuel behavior.

Experimental Setup and Instrumentation

The engine used in this study was a nominally stock production Ford four-cylinder (I4) 2.0 L Zetec DOHC (16-valve), which has a 9.6 compression ratio, 85 mm bore, 88 mm stroke, and was equipped with both production (modified for dyno operation) engine control unit (ECU) and a custom PFI system. Cylinder 4's

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received February 1, 2005; final manuscript received August 5, 2005. Assoc. Editor: D. Assanis.

behavior was observed during this study because of its ease of access. This engine was connected to a Froude Consine AG80 eddy-current dynamometer with a Digalog controller on a test bed. The entire engine-dyno test bed was contained within an insulated and thermally controlled space to provide colder than ambient starts when desired. Although the engine's accessories (power steering pump, AC compressor, and alternator) were not used the dyno did provide an additional load (dyno inertia $\sim\frac{1}{5}$ the inertia of the engine's flywheel). Consequently, it is believed that the engine's start-up behavior is representative of a production engine. The throttle was connected to a stepper motor and controlled with a PC, but was kept closed during these start-up experiments. Further engine and test-cell details can be found in Refs. [5,10,11]. The fuel used was California Phase II summer, and fuel control was calibrated to provide slightly fuel vapor rich in-cylinder fueling during the start-up.

A FFID (Cambustion HFR 400) was used to measure both in-cylinder and exhaust hydrocarbon mole fractions. A FID basically acts as a carbon counter. Both the in-cylinder and exhaust probes were located on cylinder 4 in the Zetec engine. Reference [10] describes, in detail, the experimental arrangement and calibration methods (including the treatment of residual gas effects on FFID output), and Ref. [12] discusses, in depth, the FFID unit alone.

The FFID has a time response of one milliseconds and an accuracy of ±1%. The most significant delay is from the sample transport from the combustion chamber to the FFID. This 0.002 mm i.d. tube is 5 cm long providing for a small sample delay that is adjusted for and described below. It should be noted also that the sample probe extends 2 mm into the combustion chamber and is a point measurement. Previous testing [5] with open-valve fuel injection during starting showed significant inhomogeneity of the measured in-cylinder charge, which was not apparent for closed-valve injection. Thus, in this study with closed-valve injection significant charge nonuniformity is not believed to exist.

For these experiments, the FFID constant pressure (CP) chamber in the FFID head was set at 1 atm absolute, so the signal is only accurate for compression and expansion, not the intake and exhaust strokes. Calibration of the FFID has been described in Refs. [5,10]; however, this referenced work investigated precombustion UHCs. Since the number of moles of charge changes after combustion, the precombustion calibration must be modified by the ratio of burned moles in the cylinder charge to precombustion unburned moles of charge as shown in Eqs. (1)–(3) where x is mole fraction and N_i is the moles of constituent i

$$\text{Voltage}_{\text{FFID}} \sim \text{x}_{\text{HC}} \tag{1}$$

$$x_{\text{HC}} = N_{\text{HCpre}}/N_{\text{total-pre}} = N_{\text{HCpost}}/N_{\text{total-post}} = N_{\text{HCpre}}(N_{\text{burned}}/N_{\text{unburned}})/N_{\text{total-post}} \tag{2}$$

$$N_{\text{HCpost}} = N_{\text{HCpre}}(N_{\text{burned}}/N_{\text{unburned}}) \tag{3}$$

Fuel vapor mass results from these expressions for fuel vapor moles. Since the FFID measures fuel concentration (or fuel mole fraction), a previously obtained volumetric efficiency calibration [5] was used to obtain absolute fuel mass.

Results

As mentioned in Ref. [7] the in-cylinder FFID signal can reemerge after the main flame has consumed the cylinder's fuel and air charge because of outgassing of UHCs from crevices. However, in this early study, the reemergent FFID signal was just a small fraction of the initial FFID signal magnitude (~10% or less) in a stabilized engine. Figure 1 shows the raw FFID signal data from an ambient engine start for the first 13 cycles of engine operation from cylinder 4. The preflame UHC FFID peak is shown. Also shown is the post-combustion FFID peak, which is occurring for a substantial portion of the exhaust stroke and is of a significant magnitude (although the post-combustion peak is steadily decreasing with passing engine cycles following startup). From this raw data it can be seen that the UHC concentration that develops post-combustion is on the order of the precombustion UHC concentration (near stoichiometric air-fuel ratio) decreasing to $\sim\frac{1}{2}$ the precombustion level after ~2 s of engine operation. Thus, the data indicate that significant fuel is escaping the main combustion event (as seen in Ref. [13] as liquid films) and is vaporizing during the expansion stroke for these start-up engine cycles. Possible crevice outgassing may also be a contributor to this post-combustion fuel vapor.

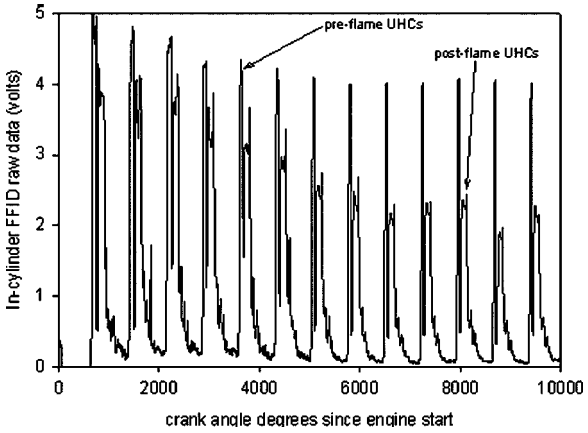


Fig. 1 Raw in-cylinder FFID voltage signal during an ambient engine start. Thirteen engine cycles are shown.

In Fig. 2, the first three engine cycles shown above are graphed in expanded scale. The corresponding simultaneous pressure trace is shown, as well with superimposed bottom dead-Center compression (BDCC) pulses. The precombustion FFID peak plateau is seen to exist just prior to peak in-cylinder pressures. After peak pressure the FFID voltage drops as a result of the main flame consuming the in-cylinder charge of UHCs. Next, a short period exists in which the in-cylinder FFID signal is low and then rises rapidly due to what is believed to be post-combustion in-cylinder liquid fuel film vaporization and crevice outgassing. The post-combustion signal falls off rapidly as the exhaust valve opens (EVO) due to a loss in the pressure gradient between the engine's cylinder and FFID CP chamber. Thus, the FFID data from EVO until shortly after the next intake valve closing (IVC) is not useful.

The data in Fig. 2 are those collected by a NI (National Instruments) Lab View data acquisition system. In reality, the FFID signal includes some transport delays between the engine's cylinder and FFID. The sample tube is very short (5 cm) with the FFID head placed directly up against the back of the engine's cylinder

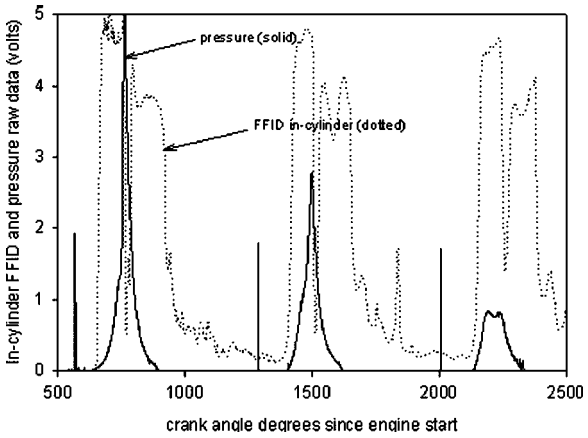


Fig. 2 First three cycles from Fig. 1 shown in expanded scale

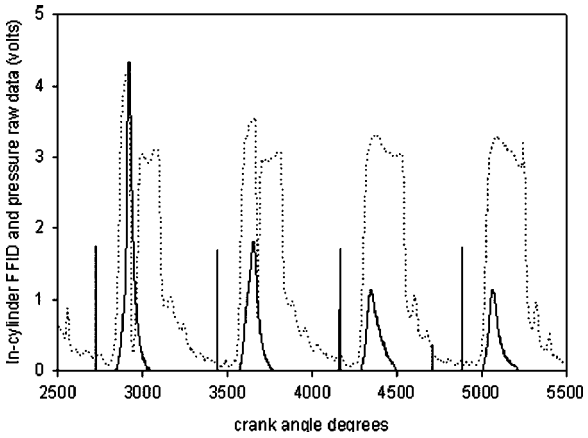


Fig. 3 Raw FFID data from ambient engine start, with two firing cycles and two misfiring cycles

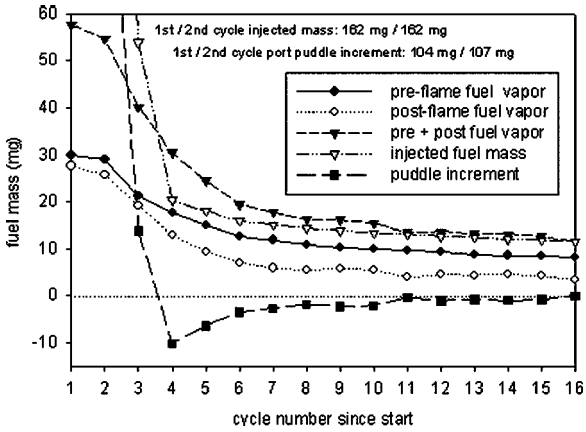


Fig. 5 Second characteristic ambient engine start

head. Thus, based on Refs. [14,15], the estimated sample delay for both the FFID and sample transfer tube is 30 crank-angle degrees at 1200 rpm. This delay is quite small when viewed on the abscissa scales of Figs. 1 and 2. Figure 3 corrects for this delay and shows both the FFID signal and pressure signals for a different ambient start, during which two positive IMEP (indicated mean effective pressure) cycles are followed by two negative IMEP cycles that are misfiring. It can be seen in the misfiring cycles that the rapidly falling FFID trace during combustion is not evident, and a slightly varying in-cylinder fuel concentration results over the compression and expansion strokes. It is interesting to note that the misfiring cycles do not continue to evolve further fuel vapor during the expansion stroke (unlike the firing cycles). This suggests that the observed post-combustion fuel vapor generation is a result of heat and turbulence characteristics from combustion.

Figure 4 shows the analyzed FFID in-cylinder data of an ambient engine start (20°C) for the first 16 cycles after startup (~3 s of real time since cranking commences). The solid circles show the measured fuel vapor mass in-cylinder prior to combustion. Since the manifold pressure is getting drawn down over the first six engine cycles (from 1 atm to ~0.4 bar) the fuel mass (by the calibrated engine controller) appropriately reduces fuel injection mass such that the in-cylinder air-fuel ratio (AFR) is just slightly fuel rich for each of the engine cycles. One atmosphere inlet pressure requires 28 mg of fuel for stoichiometric operation, and idle (0.4 bar) intake pressure requires 11 mg for stoichiometric combustion. The fuel injected mass is quite large for the first three engine cycles in order to obtain positive IMEP cycles and then reduces to ~40% excess fueling once the engine is above 1000 rpm. The post-combustion fuel mass is shown by the open circles, cycle by cycle. It is seen to be just slightly less in magnitude than the precombustion unburned fuel level, decreasing to $\sim\frac{1}{2}$ the precombustion fuel level after 3 s. This mass is a substantial amount of post-combustion fuel that is emerging into the engine's cylinder during the expansion stroke.

Also shown in Fig. 4 is the intake port puddle increment, assuming that no fuel is lost to the crankcase. This is likely an upper limit since some unburned fuel does make its way to the oil sump—crankcase [16]. It can be seen that the intake port is getting "wetted" over the first three cycles (heavy fueling cranking cycles). However, starting with cycle 4, fuel accounting suggests that the intake port puddle is reducing in magnitude modestly, cycle by cycle, and reaching equilibrium by cycle 9. Figure 5 shows a second ambient engine start with different fueling for the first three engine cranking cycles; however, the above described in-cylinder fuel behavior both pre- and post-combustion is very similar.

Cold engine starting at 0 °C was also performed. The results from a cold start are shown in both Figs. 6 and 7. Figure 6 shows the first five engine cycles, whereas Fig. 7 shows the first 3 s of startup (~16 engine cycles). Very heavy injected fueling (300 + mg) over ten times the stoichiometric in-cylinder fuel vapor mass is required for the first engine cycle in order to obtain a near stoichiometric mixture in the cylinder. Intake port puddle increments are also substantial during the early engine cycles during

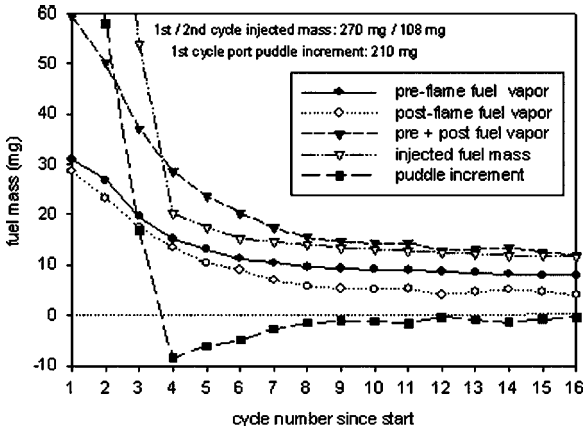


Fig. 4 Ambient engine start (20 °C) cycle-by-cycle analysis of pre- and post-combustion fuel vapor mass levels. Injected fuel mass and puddle increment mass are also shown.

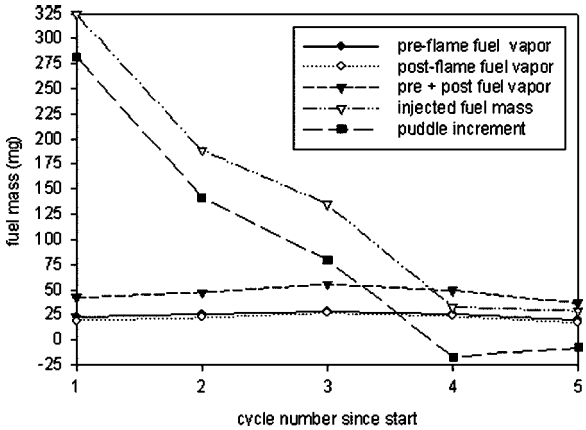


Fig. 6 Cold engine start (0 °C) cycle-by-cycle analysis of pre- and post-combustion fuel vapor mass levels. Injected fuel mass and puddle increment mass are also shown.

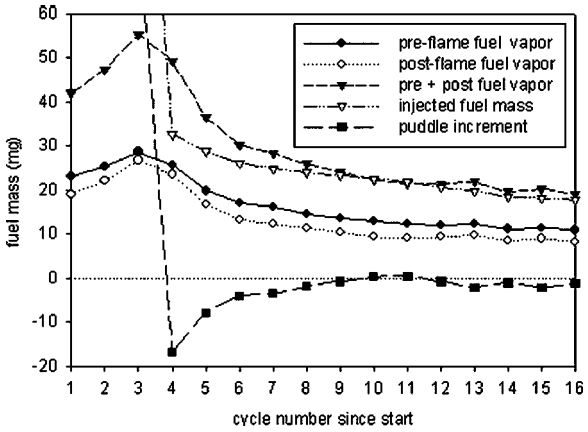


Fig. 7 First 3 s of engine running for the data shown in Fig. 6

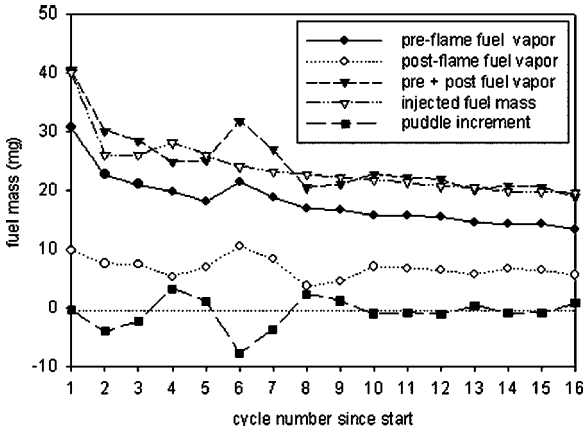


Fig. 9 Hot engine start (90 °C) cycle-by-cycle analysis of pre- and post-combustion fuel vapor mass levels. Injected fuel mass and puddle increment mass are also shown.

startup. A reduced ordinate scaling in Fig. 7 shows the in-cylinder fuel vapor behavior more clearly. The in-cylinder fuel vapor mass pre-combustion is slightly lean for the first two engine cycles, becomes near stoichiometric, and then appropriately decreases with decreasing intake manifold pressure. The post-combustion evolving fuel vapor is just slightly less than the precombustion fuel, thus more appears to be evolving for this cold start as compared to ambient starting. Figure 8 shows a second characteristic cold start with similar behavior to that just observed and described.

Figures 9 and 10 show two different hot starts (90 °C) that are characteristic of the many collected hot starts. As can be seen, the in-cylinder pre-combustion fuel vapor decreases appropriately from ~30 to 14 mg at 0.4 bar inlet pressure. The post-combustion evolving fuel vapor is much less than for the ambient and cold starts above. For hot starting, the post-combustion vapor is ~1/3 of the precombustion in-cylinder fuel vapor mass during these first few seconds of starting. This hot start behavior is expected as much less fuel is injected and with hot engine surfaces, less fuel is anticipated to remain as liquid films within the engine.

A comparison of the above three starting temperatures is shown in Fig. 11 (first presented figure for each of the above temperatures). On the ordinate is displayed postflame fuel vapor normalized to preflame fuel vapor as a function of engine cycles from start. Engine starting at 0°C shows 80–90 % post-combustion fuel vapor shortly after start, decreasing only modestly to 75% after 3 s of engine operation. The ambient (20 °C) engine start has similar post-combustion fuel vapor levels to that observed at 0 °C, however, shortly after startup, the post-combustion fuel vaporization drops to one-half of the precombustion (near stoichiometric) fuel levels. For hot starting the post-combustion fuel vapor levels are roughly steady at ~1/3 of the precombustion fuel vapor level.

Figure 12 shows a comparison of normalized post-combustion fuel level with respect to injected fuel over the three temperature starting cases. For all starting temperatures studied, the post-combustion fuel vapor levels are a small fraction of the injected fuel mass for the first few cycles of cranking. Thereafter, the colder engine cases see a greater fraction of injected fuel evolving

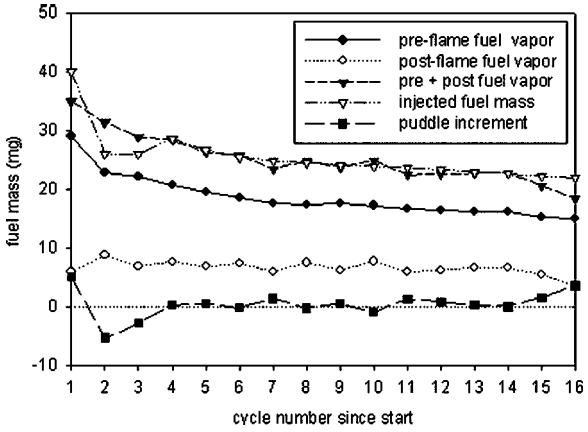


Fig. 10 Second characteristic hot engine start

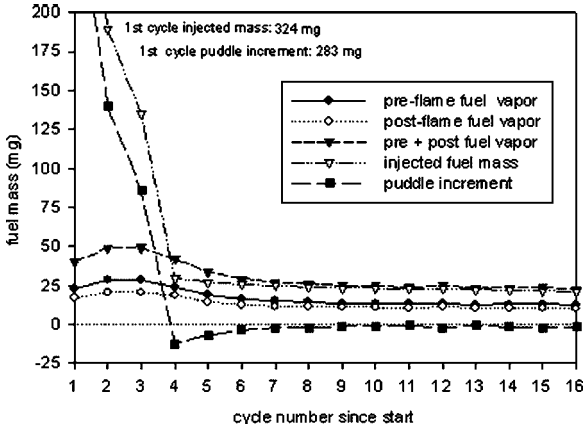


Fig. 8 Second characteristic cold engine start

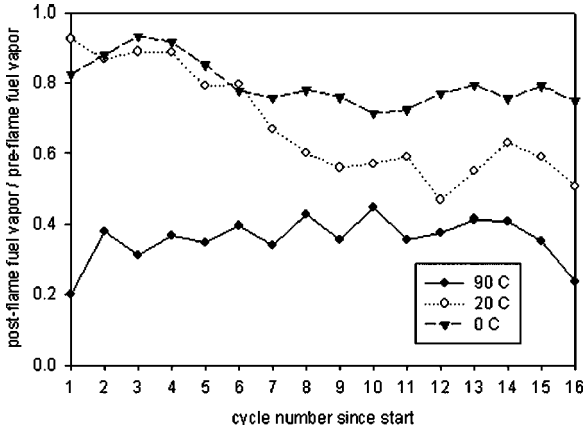


Fig. 11 Post-combustion fuel vapor mass normalized to its associated precombustion fuel vapor mass over the first 3 s of engine running

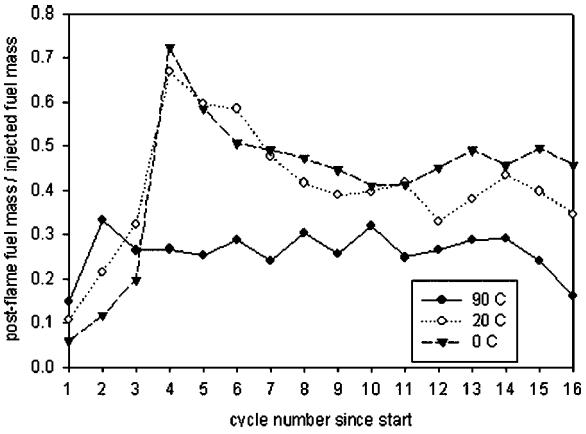


Fig. 12 Post-combustion fuel vapor mass normalized to its corresponding injected mass over the first 3 s of engine running

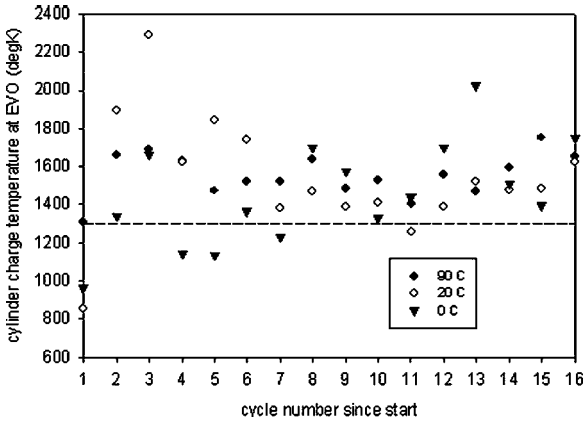


Fig. 13 Heat release model predicted cylinder charge temperature at EVO over the first 3 s of engine running

as post-combustion fuel vapor in subsequent cycles. This is expected due to the increased injected enrichment with the ambient and cold starting cases that is required to achieve near stoichiometric precombustion fuel vapor mass.

Discussion

From the above experiments it is evident that significant amounts of fuel vapor are emerging into the post-combustion cylinder charge after the main combustion event has passed. The amount of this post-combustion fuel vapor is from $\frac{1}{3}$ to nearly a full cylinder's charge amount of fuel. Since this level of post-combustion UHC fuel is on the order of tens of thousands to 10^5 ppmHC_1, two questions result since the engine out emissions (cycle by cycle from the first cranking cycle, as described in [5]) are only approximately twice stabilized engine levels during start-up. Since engine out exhaust port UHC levels are much lower than this measured post-combustion UHC concentration both sufficient oxygen and sufficient temperature must exist for substantial oxidation of most of this post-combustion fuel vapor.

The main source of oxygen-air that escapes combustion is due to the top-land piston crevice. Warmed up engines have a top land crevice volume that is characteristically 1% of the clearance volume [17]. From the author's measurements of cold engines, this crevice volume can approach 3% of the clearance volume, since the piston is cold and has a much larger thermal coefficient of expansion as compared to the cylinder bore. FIDs do not detect carbon monoxide (CO) [18], so enough oxygen must escape the main combustion event to at least form CO. The partial oxidation reaction for iso-octane gives

$$C_8H_{18} + 4(O_2 + 3.77N_2) - > 8\ CO + 9H_2 \quad (4)$$

Thus,

$$(4 \text{ moles air}/1 \text{ mole fuel})(29 \text{ gm/mole-air}/114 \text{ gm/mole-fuel}) = 1.0 \text{ gm air/gm fuel} \quad (5)$$

For this engine, the range of 3% crevice to clearance volume gives 50 mg of air escaping combustion in the top-land crevice at 1 atm inlet pressure, falling to ~20 mg at 0.4 bar inlet pressure. Since partial oxidation requires a 1:1 mass ratio for post-combustion fuel vapor to air, it is reasonable that sufficient oxygen exists to at least partially oxidize the post-combustion fuel vapor.

The second element necessary for post-combustion oxidation is sufficient temperature. Reference [19] shows the low-end "cutoff" temperature for HC oxidation to be 1300 K. A heat-release analysis was performed (with appropriate heat transfer) [20] in order to provide an indication of in-cylinder temperatures from these starting engine cycles. Shown in Fig. 13 are the heat release model predicted cylinder charge temperatures at EVO for the all three starting temperatures. It is seen that by the time of EVO, sufficient temperature has existed for partial oxidation to occur. Thus, it is believed that despite significant in-cylinder post-combustion fuel vapor levels, sufficient oxygen and temperature exist for most of the post-combustion fuel to be at least partially oxidized.

Conclusions

This experimental work has shown that in-cylinder fuel that escapes combustion by either residing on cool cylinder and head walls or in cylinder crevices evolves post-combustion in substantial levels from 1/3 to approximately a stoichiometric charge amount of fuel, increasing with decreasing starting engine temperature. Fuel accounting also shows that intake port puddles are likely to form during just the first few cranking cycles, with their associated heavy fuel injection levels. Intake port puddle equilibrium is achieved ~1 s after cranking begins. It is believed that sufficient oxygen escapes combustion and that significant expansion temperatures exist to at least partially oxidize the majority of this post-combustion emerging fuel vapor.

Acknowledgment

The author wishes to thank Dr. Eric Curtis of Ford Motor Company for the encouragement to pursue this analysis. The experimental work was performed by the author while at MIT's Laboratory for 21st Century Energy, and was funded by the Ford Motor Company and by an Industrial Consortium on Engine and Fuels Research, whose members are Daimler Chrysler Corp., Ford Motor Co., General Motors, Exxon-Mobil Corp., and Delphi Inc.. The analytical work presented within was supported by the United States Naval Academy. The author would also like to thank Lt. Armbrester for his grammatical critique of this manuscript.

References

[1] Sleightholme, G. R., 1990, "In Cylinder Measurements of Charge Inhomogenity in a Spark Ignition Engine," Society of Automotive Engineers (SAE) Paper No. 900484.
[2] Fox, J. W., Min, K. D., Cheng, W. K., Heywood, J. B. and Collings, N., 1992, "Mixture Preparation in a SI Engine With Port Fuel Injection During Starting and Warm Up," SAE Paper No. 922170.
[3] Shin, Y., Min, K. D. and Cheng, W. K., 1995, "Visualization of Mixture Preparation in a Port Fuel Engine During Engine Warm-Up," SAE Paper No. 952481.
[4] Meyer, R. C., and Thring, R. H., 1995, "Mixture Prepation Measurements," SAE Paper No. 950069.
[5] Castaing, B. M., Cowart, J. S. and Cheng, W. K., 2000, "Fuel Metering Effects on HC Emissions and Engine Stability During Cranking and Startup in a PFI SI Engine," SAE Paper No. 2000-01-2836.
[6] Cowart, J., 2004, "Air-Fuel Ratio Measurement Diagnostics During Cranking

and Startup in a Port-Fuel-Injected Spark-Ignition Engine," SAE Paper No. 2004-01-1915.
[7] Galliot, F., Cheng, W. K., Cheng, C., Sztenderowicz, M., and Heywood, J. B., 1990, "In-Cylinder Measurements of Residual Gas Concentration in a Spark Ignition Engine," SAE Paper No. 900485.
[8] Peckham, M., and Collings, N., 1992, "Study of Engine Wall Layer Hydrocarbons With a Fast Response FID," SAE Paper No. 922237.
[9] Liu, H., and Wallace, J. S., 1999, "Instantaneous In-Cylinder Hydrocarbon Concentration Measurement During the Post-Flame Period in an SI Engine," SAE Paper No. 1999-01-3577.
[10] Cowart, J. S., and Cheng W. K., [YEAR] "Throttle Movement Rate Effects on Transient Fuel Compensation in a Port Fuel Injected SI Engine," SAE Paper No. 2000-01-1937.
[11] Cowart, J., 2002, "A Comparison of Transient Air-Fuel Measurement Techniques," SAE Paper No. 2002-01-2753.
[12] Cheng, W. K, Summers, T., and Collings, N., 1998, "The Fast-Response Flame Ionization Detector," Prog. Energy Combust. Sci., **24**, pp. 89–124.
[13] Koenig, M. H., Stanglmaier, R. H., Hall, M. J., and Matthews, R. D., 1997, "Mixture Preparation During Cranking in a Port-Injected 4-Valve SI Engine," SAE Paper No. 972982.
[14] Zhao, H., and Ladommatos, N., 2001, *Engine Combustion Instrumentation and Diagnostics*, SAE, Warrendale, PA.
[15] Fox, J. W., Cheng, W. K., and Heywood, J. B., 1993, "A Model for Predicting Residual Gas Fraction in Spark-Ignition Engines," SAE Paper No. 931025.
[16] Shayler, P. J., Winborn, L. D., and Scarisbrick, A., 2000, "Fuel Transport to the Crankcase, Oil Dilution and HC Return With Breather Flow During the Cold Operation of a SI Engine," SAE Paper No. 2000-01-1235.
[17] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.
[18] McNair, H. M. and Miller, J. M., 1997, *Basic Gas Chromatography*, Wiley, New York.
[19] Yang, W., Zheng, J., Miller, D. L., and Cernansky, N. P., 2000, "Tracer Fuel Injection Studies on Exhaust Port Hydrocarbon Oxidation: Part II," SAE Paper No. 2000-01-1945.
[20] Gatowski, J. A., Balles, E. N., Chun, K. M., Nelson, F. E., Ekchian, J. A., and Heywood J. B., 1984, "Heat Release Analysis of Engine Pressure Data," SAE Paper No. 841359.

Experimental Validation of an Improved Squish Velocity Model for Bowl-in-Piston Combustion Chambers

P. Lappas[1]
e-mail: lappas@mech.ubc.ca

R. L. Evans

Department of Mechanical Engineering,
The University of British Columbia,
Vancouver, BC V6T 1Z4, Canada

A simple two-zone mass transfer model was used to predict the mean squish velocity history at the rim of a conventional bowl-in-piston combustion chamber. The chamber's geometry produces gas flow that converges radially inwards ("squish") as TDC (top dead center) is approached. The squish flow generates turbulence, which can be used to enhance the combustion rate. When compared with PIV (particle image velocimetry) measurements, the peak squish velocity at the bowl rim was 12% less than the value predicted by the simple mass transfer model. After a thorough examination, the assumption of uniform density in the simple model was strongly suspected to be the cause of this discrepancy. Improvements were made to the simple model to account for density variations that are caused by nonuniform heat transfer in the combustion chamber. The revised model yielded velocities that were in close agreement with PIV measurements.
[DOI: 10.1115/1.2130730]

1 Introduction

Combustion chambers exhibiting squish flow are advantageous for spark-ignited lean-burn engines because the turbulence generated by the squish near TDC (top dead center) enhances the burn rate. Previous attempts to develop models of the squish flow have not been successful in accurately representing the evolution of the squish flow in the chamber. For example, Fansler and French [1] found that their model overpredicted their squish velocity measurements by 20%–40%. The authors also noted similar discrepancies in previous studies [2–4]. In this study, independent experiments were performed that were designed to minimize flow measurement uncertainties. The experimental results were compared with a simple squish model and the discrepancies were used to create a new two-zone flow model which is representative of the actual flow.

2 The Ideal Squish Velocity Model

The motion of a piston (with a centered and symmetrical bowl) during the compression stroke is shown in Fig. 1.

As shown in Fig. 2, the combustion chamber can be divided into two separate control volumes. One can find the squish velocity by first applying a mass balance to zone 1 in Fig. 2:

$$V_1 \frac{d\rho_1}{dt} = v_p A_1 \rho_1 - v_s A_s \rho_2 \tag{1}$$

Similarly a mass balance applied to zone 2 yields

$$V_2 \frac{d\rho_2}{dt} = v_p A_2 \rho_2 + v_s A_s \rho_2 \tag{2}$$

where

V_1 and V_2 are the volumes in zone 1 and zone 2, respectively,
ρ_1 and ρ_2 are the densities in zone 1 and zone 2, respectively,
t denotes time,
v_p is the piston speed,
A_1 is the area of the squish surface on the piston crown,
A_2 is the area of the piston bowl opening,
v_s is the squish velocity (radially inwards), and
A_s is the area of intersection between zones 1 and 2.

Since the squish ratio $[A_1/(A_1+A_2)]$ is 75% in this study, A_2/A_1 equals 1/3.

If the assumption is made that the density in the cylinder is uniform, division of Eq. (2) by Eq. (1) yields

$$\frac{v_s}{v_p} = \frac{\frac{V_2}{V_1} - \frac{A_2}{A_1}}{\frac{A_s}{A_1}\left(1 + \frac{V_2}{V_1}\right)} = \frac{\frac{V_2}{V_1} - \frac{1}{3}}{\frac{A_s}{A_1}\left(1 + \frac{V_2}{V_1}\right)} \tag{3}$$

With a cross-sectional area that is around 1% of the bore area, the top-land crevice was not included in the ideal model because it makes a negligible contribution ($<1\%$) to the squish velocity. Since A_1 is a constant and V_2, V_1, v_p, and A_s are known functions of crank angle, the squish velocity can be determined for any given crank angle.

3 Experimental Validation

3.1 The Test Volume. The primary component of the experimental system is the UBCRICM (University of British Columbia Rapid Intake and Compression Machine) because it contains the test volume. This machine generates the physical phenomenon of interest: turbulent squish flow in a combustion chamber. Although motored engines can be used to generate similar flows, their initial flow conditions are not accurately known and they vary randomly from cycle to cycle. The UBCRICM eliminates these uncertainties because the initial flow field is always quiescent.

[1]Corresponding author.

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received February 24, 2004; final manuscript received July 16, 2005. Review conducted by D. Assanis.

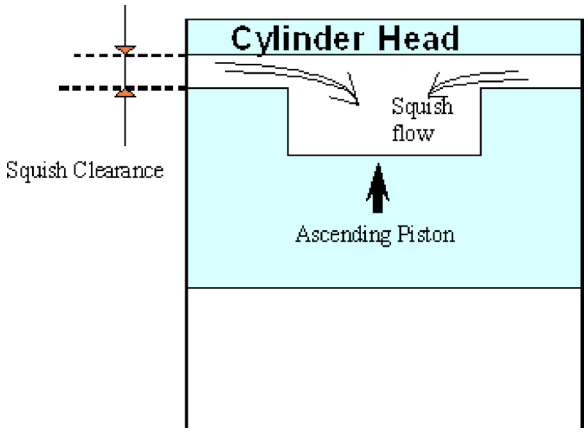


Fig. 1 Simple example of squish flow in bowl-in-piston combustion chamber

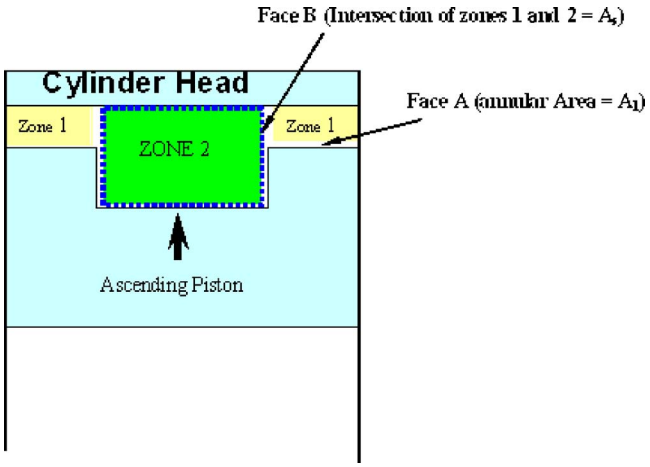


Fig. 2 The two control volumes (zones) assigned inside the combustion chamber. Mass conservation on these zones can yield the average squish velocity through area A_s

The UBCRICM comprises a single cylinder whose piston follows the motion of the wrist pin on a connecting rod and crank linkage. The crank is turned by the linear motion of a rack that meshes with a crankshaft pinion. The rack is driven from one end by a pneumatic cylinder while the other end is connected to a hydraulic ram. The ram pushes hydraulic fluid through an orifice at a constant flow rate, thus ensuring a constant crank speed. The UBCRICM crank normally rotates 360 deg when it is fired, with the piston starting at top dead center (TDC), so that both an intake and compression stroke can be executed. It was reconfigured so that the piston started at bottom dead center (BDC) to complete the compression stroke first. The pressure in the pneumatic cylinder was set so that the effective crank speed was 800 rev/min (2.66 m/s mean piston speed). A detailed description of the UBCRICM is found in the design report by Döhring [5] and modifications made to the machine for the experiments reported herein are documented by Lappas [6].

The cylinder, piston, and cylinder head were made of transparent acrylic to enable optical velocity measurements. Three piston rings and a head gasket, all of which were lubricated O-rings, formed a tight seal for the cylinder. No inlet or exhaust valves were present, but a small port drilled through the cylinder wall just above the piston crown at BDC allowed air to enter the cylinder. While the port ensured ambient cylinder pressure at the start of the compression stroke, it was also used to introduce microscopic particles (for PIV velocity measurements) before being plugged for the start of the compression stroke. The particle seeding system is shown in Fig. 3.

The combustion chamber used for this study was a "bowl-in-piston" type, as illustrated in Fig. 4.

The main geometric features of the combustion chamber are defined as

bowl area—The area of the piston bowl opening.
squish area—The cylinder bore area minus the bowl area.
squish ratio—The ratio of squish area to cylinder bore area.
piston clearance—The closest distance between the piston and the cylinder head at a given crank angle.
raised bowl volume—The volume of the bowl plus the product of the bowl area and the piston clearance.
squish volume—The volume of the combustion chamber minus the raised bowl volume.
squish flow area—The area that is formed by the intersection of the squish volume and the raised bowl volume. This is the area through which the squished gases flow when exiting the squish volume.

Unless described otherwise, the squish velocity for the bowl-in-piston chamber is the ensemble mean velocity of gas through the squish flow area. The combustion chamber had a squish clearance of 3 mm, a compression ratio of 8.5:1, and a squish ratio of 75%.

Velocity measurements were generally made in 5 deg intervals between 30 deg BTDC (before TDC) and TDC during the compression stroke. As described by Heywood [7], squish velocity becomes significant during this period and a maximum is typically reached near 10 deg BTDC. Turbulence production was also expected to peak during this period because of significant shear gradients caused by the squish flow.

Experimental data were obtained by PIV (particle image velocimetry). The technique enabled the in-cylinder flow to be visualized and measured in planes parallel to the cylinder head. Approximately ten tests were performed for each crank angle.

3.2 The PIV System. The PIV system was developed from separate components and it was specifically built for the UBCRICM. The system is shown in Fig. 5 and it worked by pulsing a thin laser sheet twice through the transparent combustion chamber. The sheet, which was approximately 0.5 mm thick, was aligned parallel to the cylinder head and the laser pulse separation time was typically below 200 ms. The double pulsing sequence was initiated by the crank angle trigger that was set on an optical encoder mounted on the crankshaft of the UBCRICM. For each crank angle of interest the sheet was placed midway between the piston crown and the head. Figure 4 shows the laser sheet position denoted by plane A. Microscopic particles that were seeded into the combustion chamber and that followed the local flow inside it were illuminated each time the laser pulsed. For each of the two laser pulses, a synchronized digital camera recorded an image of the illuminated particles. The resulting image pair was later analyzed using cross-correlation software, which tracks the displacements of gridded groups of particles. A sample image pair is shown in Fig. 6. In the cross-correlation process, the spatial pattern of the particle group found inside each grid cell (or "interrogation window") serves as the identifying feature with which the group's movement is tracked. The outcome is a velocity field map of the flow components parallel to and inside the laser sheet.

The twin pulse laser unit was a Gemini PIV 120-15 made by New Wave. Each laser head in the unit is capable of delivering 120 mJ per pulse and the nominal pulse duration is 5 ns. The wavelength of the light emitted is 532 nm. The signal generator was a Berkeley Nucleonics model 500D. Once triggered by the crank angle sensor, the signal generator provided synchronized pulses to the lasers and the camera's frame grabber (Matrox Meteor II-Digital). The camera was a Roper Scientific ES 1.0 CCD monochrome digital camera with 1008×1018 pixels per image

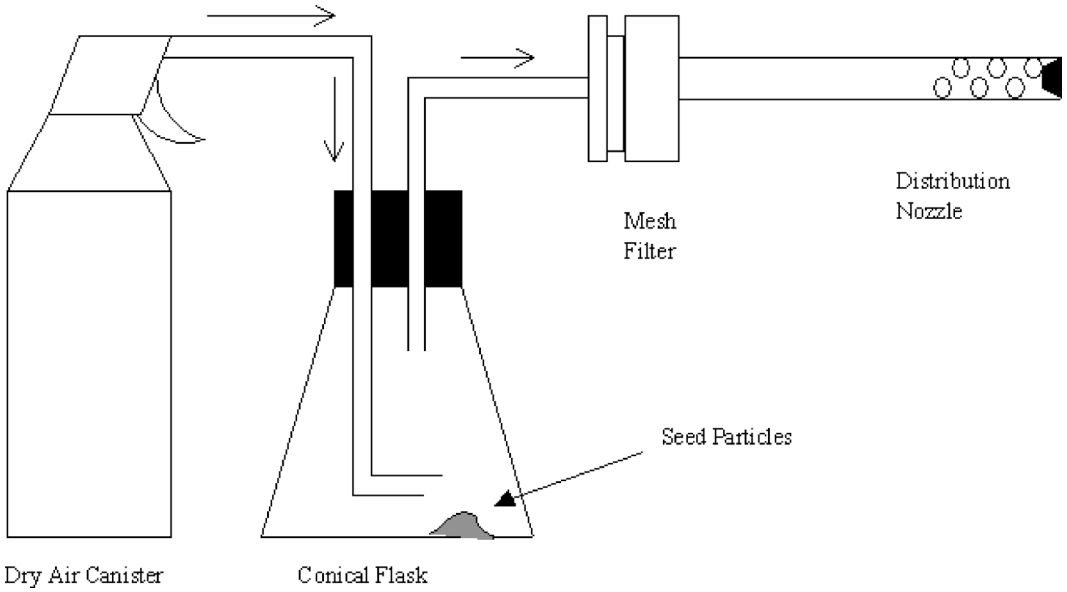


Fig. 3 The particle seeding system. The conical flask doubles as a cyclone separator

and a bit depth of 10 bits.

3.2.1 Beam Delivery. The purpose of the PIV optical assembly was to form a thin laser sheet, which could be easily translated to a desired plane in the combustion chamber. As shown in Fig. 4, this plane was always made to bisect the squish gap for the crank angle being studied. The translation of the laser sheet was achieved by mounting a 45 deg laser mirror on a traverse that moved parallel to the cylinder axis. The mirror was placed at the

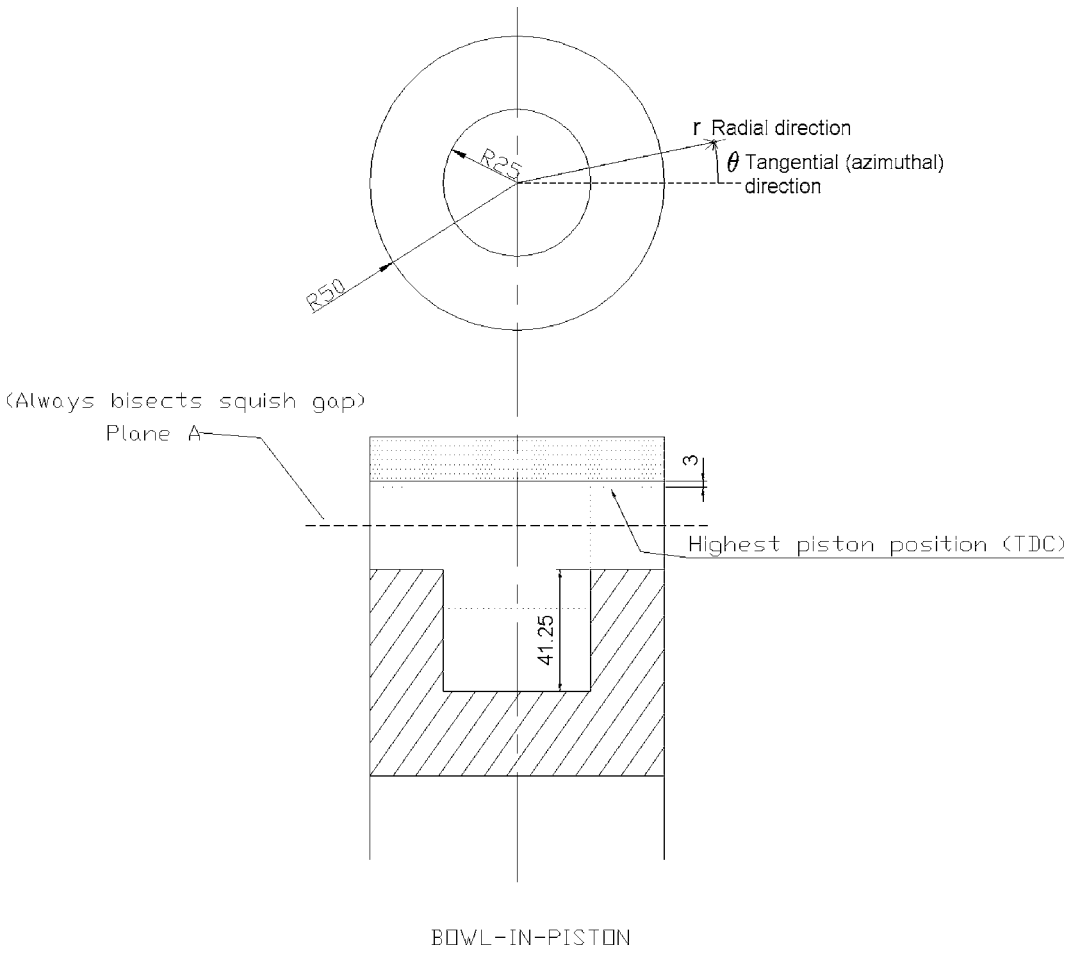


Fig. 4 PIV measurement plane (plane A) for the bowl-in-piston combustion chamber. Dimensions are in millimeters. The polar coordinate system in any plane parallel to the cylinder head is also described

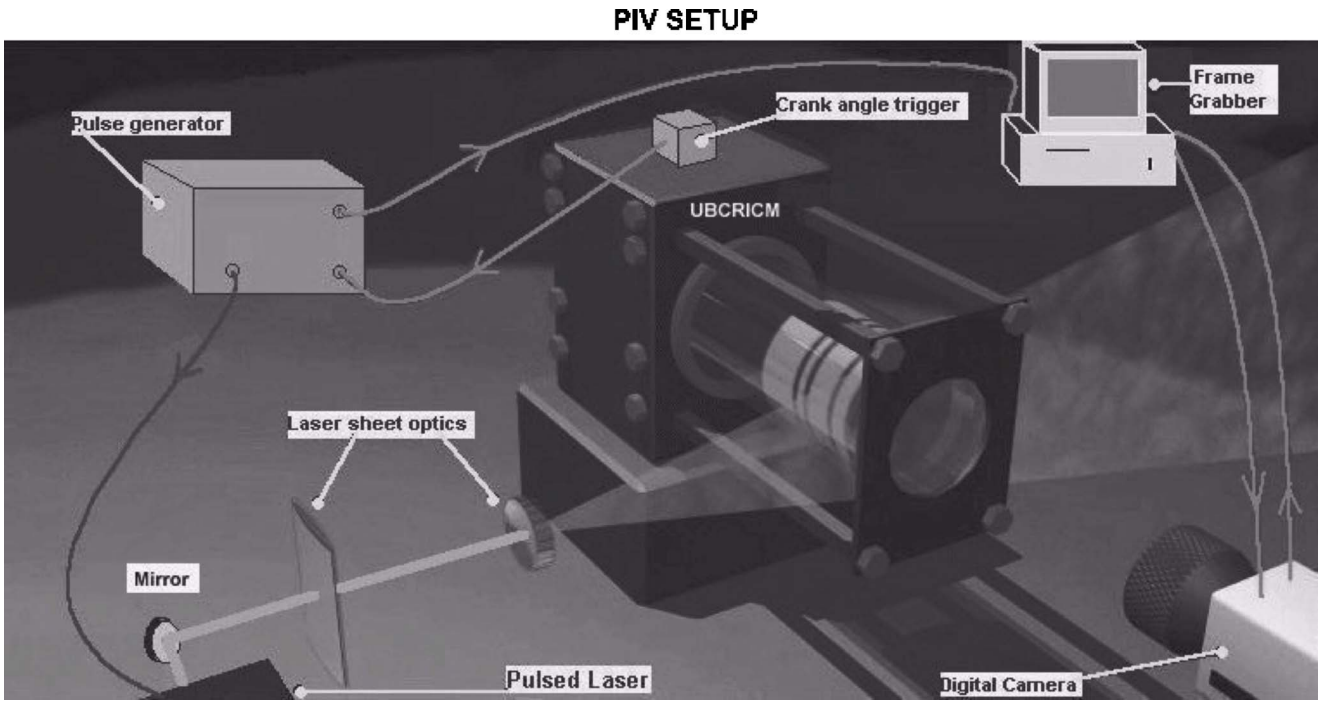


Fig. 5 Illustration of PIV system

same height as the cylinder axis and the distance between the mirror and the cylinder (approximately 1 m) was minimized to reduce the positional sensitivity of the laser sheet to any pointing instabilities in the laser. The traverse was mounted exclusively on a steel tower secured to the ground, to reduce mirror vibration that would cause unwanted movement of the light sheet.

The laser sheet was formed by two lenses (a positive cylindrical lens and a double concave lens) that were mounted on the traverse between the mirror and the test volume. Essentially, the lens pair contracts and then collimates the laser beam in the horizontal direction while it diverges the beam in the vertical direction. The cylinder axis was horizontal and normal to the laser sheet.

3.2.2 Particle Seeding Method. Commonly available constant flow-rate particle seeders are best suited to steady-flow systems because the ratio of particles delivered to the air admitted is kept at a fixed level. These types of seeders, however, are not suitable for the momentary event executed by the UBCRICM. Instead, a method was devised to "prime" the cylinder with a metered number of dispersed particles before "firing." The method entailed intermittently squirting dry air into a fixed number of solid seed particles resting inside a flask so that after a short period, the agitated particles would mix with the air. Dry air was used to avoid condensation on the cylinder walls and to prevent possible agglomeration of the seed particles. Using electrically grounded aluminium tubing, the agitated mixture would exit the flask and flow through a fine mesh. After the mesh the aerosol would flow into the cylinder through the small inlet port.

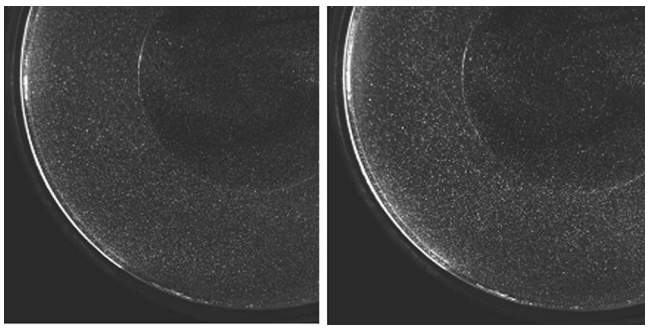


Fig. 6 Raw PIV image pair captured for bowl-in-piston chamber at 25 deg BTDC. The rightmost photograph was taken 170 μs after the left one. Displacements of particle groups in the left image are determined by spatial cross correlation analysis with the right image

The mesh filter was intended to break up any potentially clumped particles. The mesh was also used to block unbreakable particles larger than 3 μm. In order to ensure a homogeneous distribution of seed particles, a nozzle was used to inject the particles deep into the cylinder. The nozzle was made by perforating a 75 mm long tube and blocking one end. The OD (outer diameter) of the tube was smaller than the port diameter to provide an exit for the displaced cylinder air while the seed particles were being injected.

3.2.3 Seed Particle Selection. Solid polymer particles were selected for the current study because of their ability to withstand the extreme conditions at TDC and because of their narrow size distribution. The typically high degree of sphericity of the polymer particles is also an advantage because the light scattering characteristics are repeatable and the aerodynamic behaviour may be reliably predicted. The relatively low density of polymers also allows for large particle sizes that enhance light scatter while not compromising their flow following ability.

A prerequisite to particle size selection is the estimation of the flow field, particularly in regions of extreme flow acceleration. The ideal two-zone mass transfer analysis of the bowl-in-piston chamber revealed that for peak squish velocity, the mean acceleration of a parcel of air traversing the squish zone is at most 2100 m/s^2. For a more realistic representation of the parcel's acceleration, a term for the turbulent velocity fluctuations was also added. This term was based on combustion chamber observations reported by Tabaczynski [8]. The author reported an upper bound on turbulence frequency of 10 kHz. The amplitude of the fluctuations was estimated from Anetor's [9] observation that tangential and radial turbulent fluctuations (v_t^{rms} and v_r^{rms}) were as high as

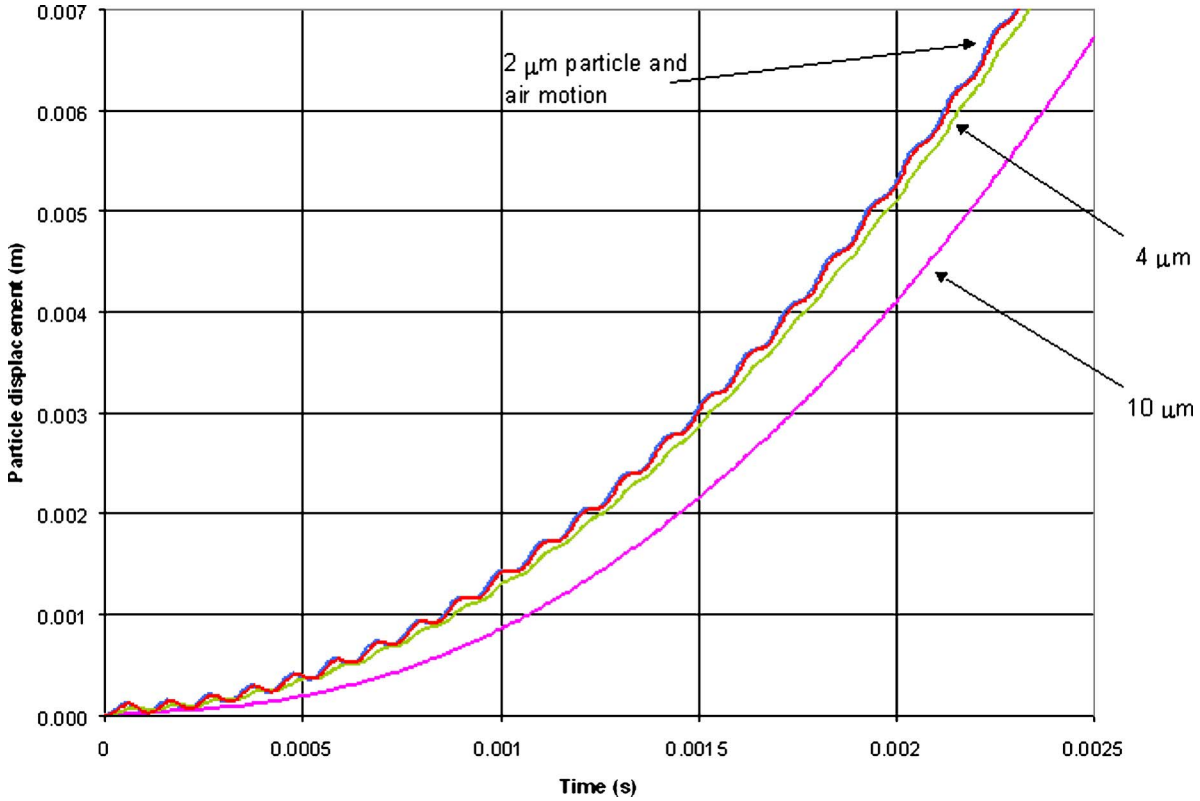


Fig. 7 The simulated displacement history for three differently sized particles following fluctuating squish flow inside the bowl-in-piston chamber. The motion of the air that the particles are responding to is also shown. The particles have a density of 1000 kg/m^3. The motion of the 2 micron particle coincides with that of the air.

the mean piston speed.

The acceleration of the gas parcel described above was twice integrated with respect to time to yield its displacement history. This history is shown in the graph of Fig. 7. The responses of three different particles to the accelerating air parcel are also shown in this graph. The particle responses were simulated by integration of the aerodynamic forces (Stokes' drag). The force from pressure gradients was neglected because of the small air to particle density ratio. Gravitational forces were also neglected as the settling velocities are less than 1 mm/s. While the particles were assumed to be spherical with a specific gravity of unity, they differed in diameter. The particle diameters were 2, 4, and 10 μm.

It is clear from Fig. 7 that the simulated 2 μm particle not only follows the mean acceleration of the fluid, but it also follows the high-frequency fluctuations. The suitability of this particle for 10 kHz flows is also documented by other researchers, including Melling [10]. Nevertheless, the 2 μm "high fidelity" particle is redundant for PIV in this study because the technique has limited spatial resolution and is unable to resolve the smaller turbulent scales [6]. For example, even though the camera image just covers a cylinder quadrant, a PIV interrogation window of 16×16 pixels corresponds to a 1 mm wide square in the object plane. This dimension is considerably larger than the 0.3 mm microscale cited by Tabaczynski [8]. In the current study this limitation makes PIV better suited to determining mean velocity rather than rms velocity fluctuations. Consequently, the 4 μm particle was considered superior for the purpose of PIV because it is a large particle (to maximize light scatter) that is able to follow the mean accelerating flow component and the low frequency components ($<$3 kHz). Accordingly, nylon microspheres from TSI Inc. with a nominal diameter of 4 μm were used for the PIV measurements in this study.

3.2.4 PIV Processing. Computer code made by Dantec Dynamics A/S called "Flowmanager" was used to extract velocity data from each PIV image pair. The computer code produced a two-dimensional array of vectors that described the velocity at the centroids of the interrogation windows covering the observation plane. This amounted to a vector array of 83 rows (x positions) by 84 columns (y positions): a total of 6972 evenly spaced velocity measurement points per PIV test.

FlowManager incorporates an iterative algorithm called "Adaptive Correlation" which has the advantage of using the displacement information determined from initially large interrogation windows to generate velocities from progressively smaller windows. As a result, the final vector density is high so that flows with large velocity gradients are spatially resolved. The initial interrogation window size in this study was 64×64 pixels while the final window size was 16×16 pixels. Since the largest particle displacement was around 15 pixels, there is compliance to the standard rule of thumb that the fastest particles do not move more than 25% of the (initial) interrogation window size during the pulse separation period.

The symmetry of the bowl-in-piston chamber was exploited to determine the mean velocity. The cylinder plane was divided into a number of 2 mm thick imaginary concentric rings that contained velocity vectors. The arithmetic mean of the vectors in each ring was used to represent the mean velocity at the ring's mid-circumference. Other useful quantities such as the standard deviation of the vectors per ring were also determined. The computer software chosen to plot the resultant vector fields was "Tecplot."

3.2.5 PIV Errors. The velocity vectors that FlowManager yielded were validated by a number of means. Perhaps the most important technique used was "peak validation" in which vectors that have a clearly dominant correlation peak were identified. Only PIV images that had passed validation were accepted. Sub-

sequent LDV (laser Doppler velocimetry) measurements taken near the bowl rim of a similar combustion chamber also matched their corresponding PIV result [6].

Assuming that the particles follow the flow with adequate fidelity and that the determined particle group displacements are valid, there are possible sources of PIV error for ensemble-mean velocities.

One source is related to the displacement resolution of particle groups. A commonly accepted value for this spatial resolution is 0.1 pixels [11]. With a typical laser pulse separation time of 200 μm and an image conversion factor of 15 pixels/mm, the resolution yields a velocity uncertainty of 33 mm/s. This represents a relative uncertainty of 3.3% for 1 m/s flows.

Another source of error is statistical in nature and it represents the difference between the measured mean and the true mean (determined from an infinite number of samples). The range of velocities in which the probability of finding the true ensemble mean is 90% is known as the 90% confidence interval for this mean. The interval can be described by the following probability statement, which is similar to the one found in Bhattacharyya and Johnson [12]:

$$\Pr\left(V_{\text{mean}} - t_{0.05}\frac{s}{\sqrt{N}} < \mu_v < V_{\text{mean}} + t_{0.05}\frac{s}{\sqrt{N}}\right) = 90\% \qquad (4)$$

where V_{mean} is the measured mean velocity,

N is the number of samples used to find V_{mean},
μ_v is the true mean velocity,
$t_{0.05}$ is the t value above which the area under the t distribution is 0.05,
s is the standard deviation in the sample set of velocities.

It follows from Eq. (4) that the 90% confidence interval for the ensemble mean is

$$V_{\text{mean}} \pm t_{0.05}\frac{s}{\sqrt{N}} \qquad (5)$$

As shown in Eq. (5), the statistical uncertainty depends on the standard deviation, which varies with crank angle and location. As a rough guide, however, the ensemble mean velocity in both the radial and tangential directions had a typical uncertainty of around 0.3 m/s.

4 Experimental Results and Discussion

Ensemble mean velocity maps determined from the PIV images are shown in Fig. 8. Each velocity map is split in two—the right half shows a vector field, while the left half mirrors the right side with contour lines of flow speed in the plane. Each PIV map does not cover the entire cylinder cross section because the raw PIV images were focused on a cylinder quadrant to maximize image resolution. Out-of-plane velocities are not shown. The broken circle surrounding the center of each velocity map represents the piston bowl rim. It is evident from Fig. 8 that the predominant velocity in the bowl-in-piston chamber is radially inward (squish) for all crank angles tested. As shown by the concentric contour rings, the velocity distribution is symmetrical about the longitudinal axis and the peak squish velocity tends to penetrate the bowl volume as TDC is approached. The squish velocity is zero at the center of the chamber because of the flow symmetry. It is worth noting that although the squish velocity near the bowl rim subsides near TDC, residual inward flow is observed above the bowl region. By continuity arguments, this indicates a persisting recirculating flow inside the bowl.

Figure 9 shows a summary of the mean radial velocities observed in plane A of the combustion chamber. In this figure, the dependence of mean radial velocity on crank angle and radial position is shown. Figure 9 shows that the peak mean radial velocity inside the PIV plane reaches a maximum at around 10 deg BTDC. This maximum occurs at the bowl lip (r=25 mm). The ideal two-zone model that was created to estimate the squish velocity also predicted a maximum at the same crank angle; however, the experimental maximum was 12% lower than the ideal value. This discrepancy is seen in Fig. 10, which shows a comparison in the squish velocity history (at r=25 mm) between the ideal model and PIV results. The difference exceeds the maximum measurement uncertainty of 0.3 m/s in the mean radial velocity. Figure 11 shows the radial distribution of the statistical uncertainty in the mean radial velocity. The distribution is shown for the two different crank angles: 20 deg BTDC and TDC.

It is worth noting from Fig. 11 that near TDC the statistical uncertainty is considerably larger above the bowl than it is above the squish area. A similar spatial distribution was found for tangential velocity near TDC. The distribution is caused by larger variations in velocity detected above the bowl due to a suspected higher level of turbulence there. Poorer PIV image quality above the bowl cannot be blamed for the larger variations because the reverse was observed. In fact, the particles appeared less distinct above the squish area because the light sheet was very close to the head and piston at TDC and stray reflection from these surfaces was more apparent. More turbulence is expected in the piston bowl because of the shear gradient resulting from squish flow at the bowl lip. The high variability in measured velocity is also evident in the PIV velocity map near TDC shown in Fig. 8. In this figure, the velocity contour bands are broken and less defined than for other crank angles.

4.1 Causes of Discrepancy in Mean Velocity between Ideal Model and PIV. In order to systematically try to isolate the cause(s) of the discrepancy between predicted and measured squish velocities, the sources of error were grouped into the following three categories: (a) poor experimental generation of specified flow, (b) PIV measurement errors, and (c) numerical modeling errors. The categories are shown in the tree diagram of Fig. 12. Parameters affecting category (a) can be identified by doing a mass balance in the squish volume. Using the notation from the ideal model, the mass balance can be expressed as follows:

$$\rho_2 v_s A_s = \rho_1 A_1 v_p - V_1\frac{d\rho_1}{dt} - \dot{m}_l \qquad (6)$$

where the new term $\dot{m}_l$ denotes the mass loss rate due to leakage and crevice filling. The mass transfer by molecular diffusion is ignored because it is insignificant relative to the squish flow.

Acknowledging that the squish flow is turbulent, the appropriate velocities and densities used in Eq. (6) can be split into their mean and fluctuating components. If the result is then Reynolds averaged, the following relation is produced:

$$\bar{v}_s = \frac{\bar{\rho}_1 A_1 v_p - V_1\dfrac{d\bar{\rho}_1}{dt} - \dot{m}_l - \overline{\rho_2' v'}_s A_s}{\bar{\rho}_2 A_s} \qquad (7)$$

where overbars represent means and the primes denote fluctuations from the means.

It is worth noting that the last term in the numerator of Eq. (7), which arose from the averaging operation, represents the turbulent transport of mass from the squish volume. It follows from Eq. (7) that the mean squish velocity $\bar{v}_s$ at a given crank angle depends on initial and boundary conditions along with time-dependent geometric parameters (V_1, A_s, v_p) and mass losses ($\dot{m}_l$). The density and turbulence terms are not included as separate variables because they too depend on the above parameters (through the energy equation and the ideal gas law). The geometric parameters depend on variables such as squish gap, squish area, bowl diameter, stroke, crank speed, and crank angle. Careful examination of these independent variables shows that squish gap and crank speed are more prone to error. These two variables along with the

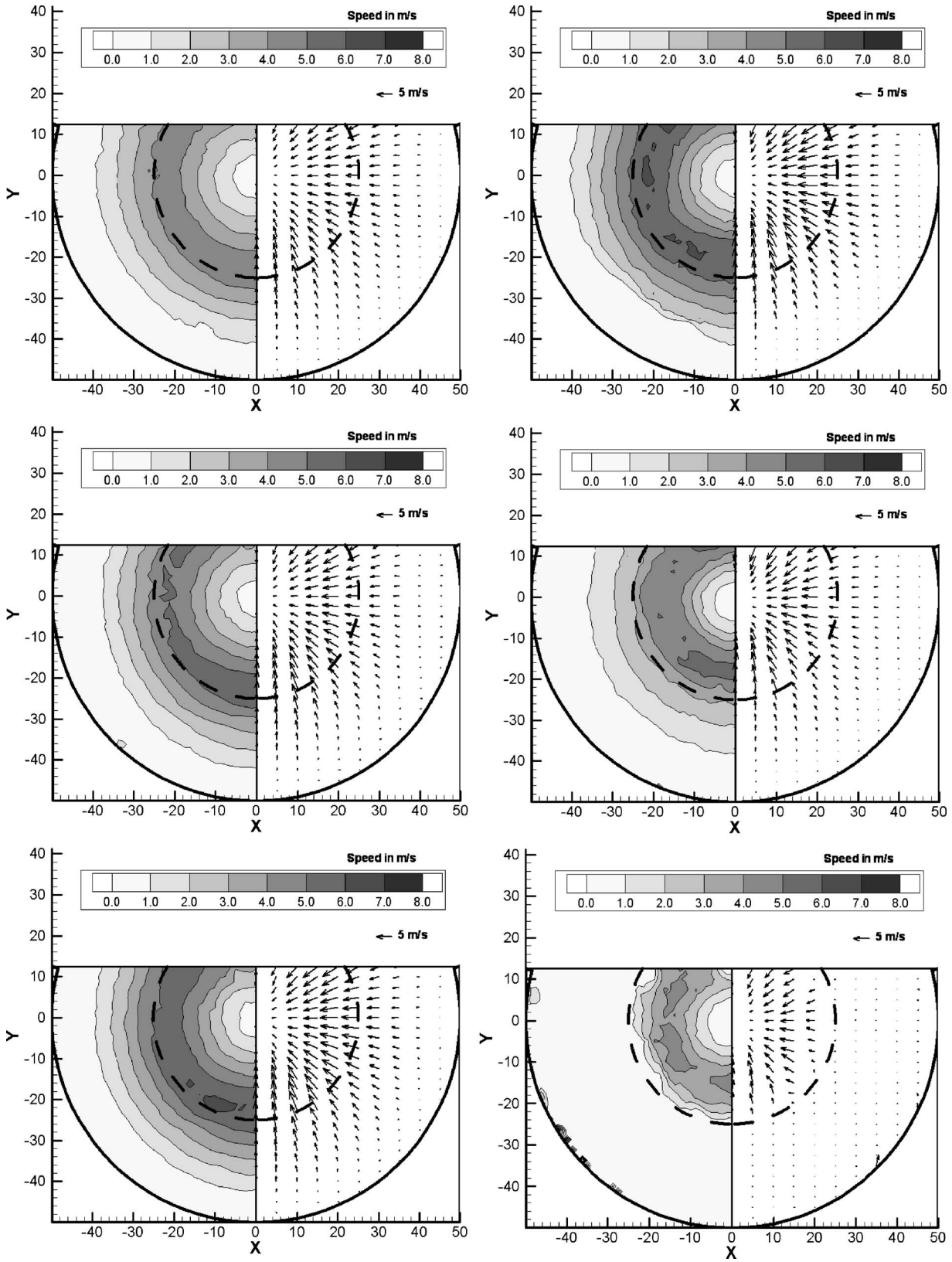


Fig. 8 Mean PIV velocity maps in combustion chamber. Units for *X* and *Y* axes are mm.

mass loss are shown in Fig. 12 under category A. The variables must have uncertainties that are small enough in order to generate a well-defined squish velocity history.

There are two relevant sources of PIV measurement uncertainty and, as shown in category B of Fig. 12, they are seed particle errors and PIV processing. The seed errors are related to distribution, visibility, and flow-following ability of the particles.

Examination of the ideal model developed earlier places suspi-

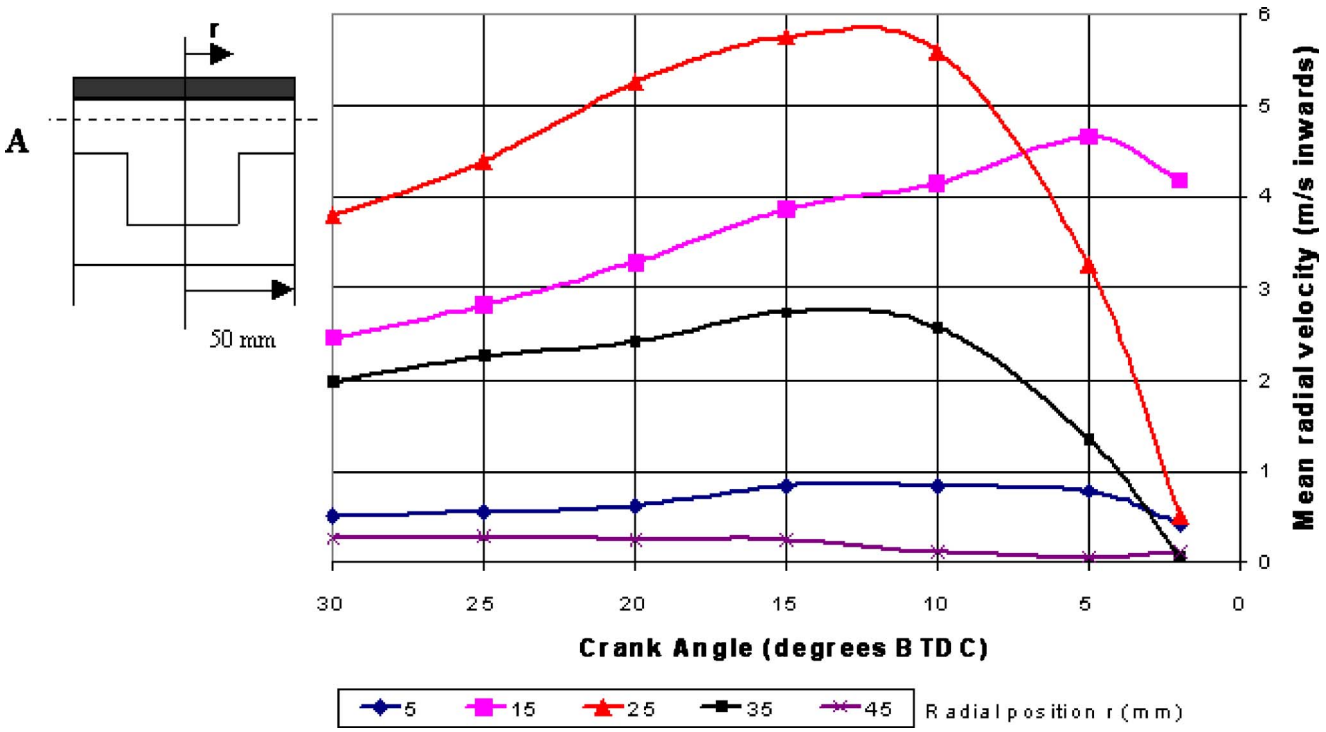


Fig. 9 Summary of PIV measurements in plane A for the bowl-in-piston chamber. The evolution of the mean radial squish velocity is shown for a set of radial positions.

cion on the key assumption that the gas density in the combustion chamber is uniform at any one time. Although pressure differentials in the nonreacting combustion chamber can be considered to have a negligible effect on density because of the subsonic velocities, temperatures may vary in the chamber due to nonuniform heat transfer. The nonuniform heat transfer could originate from local regions of relatively high turbulence (such as near the bowl rim) and, as a result, errors in the model may arise that fall into category C.

In the current study, a number of causes (of the disparity between numerical and experimental squish results) identified at the bottom level of Fig. 12 could be eliminated because of their rela-

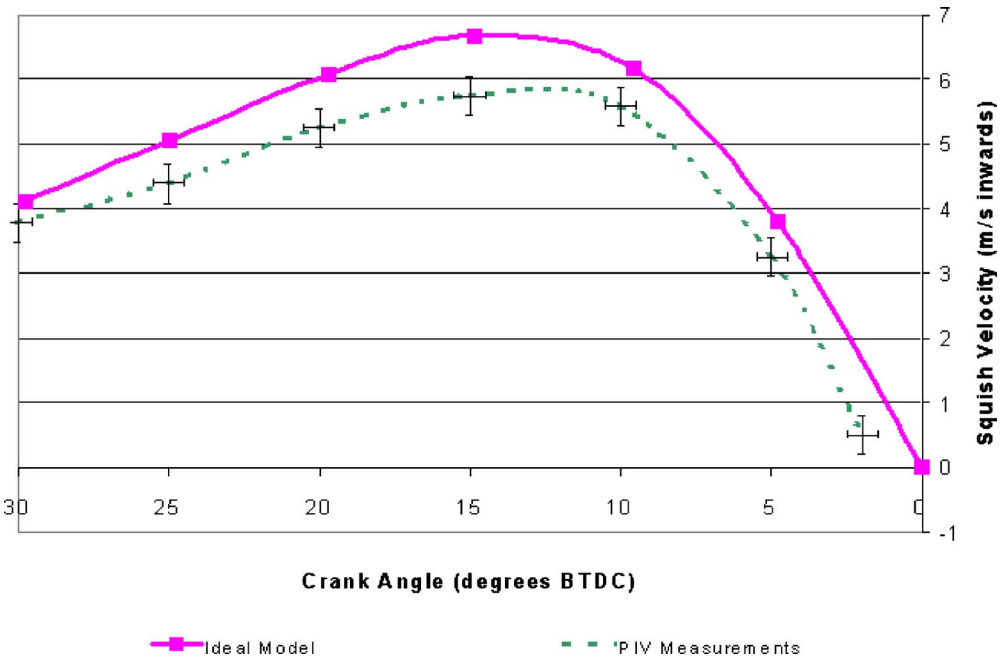


Fig. 10 Temporal evolution of the mean squish velocity for the bowl-in-piston chamber. Results from the ideal two-zone mass transfer model and from PIV results are shown for comparison. Crank speed is 800 rev/min.

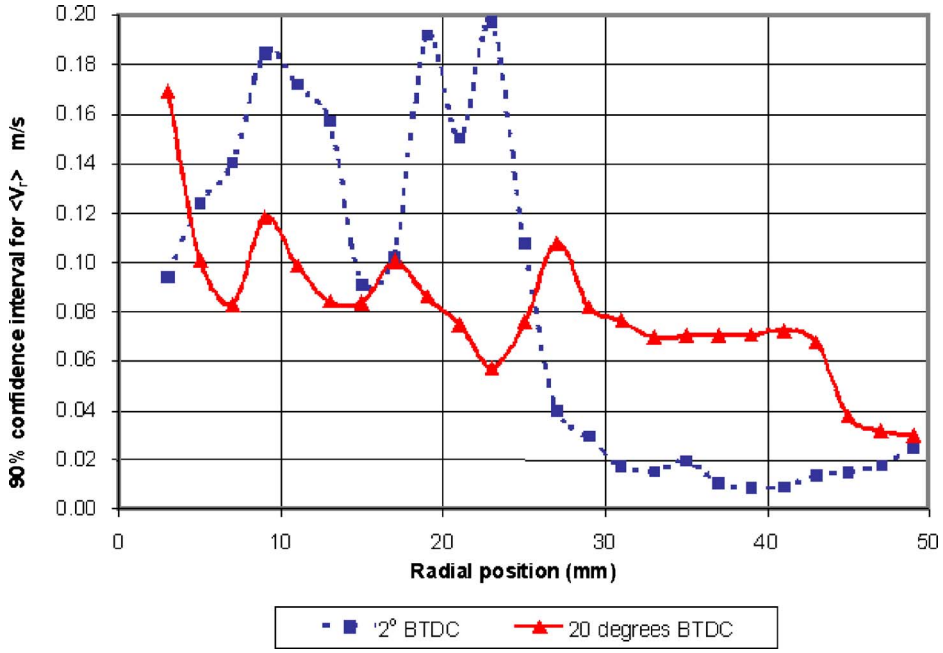


Fig. 11 The radial distribution (in plane A) of the 90% confidence interval for the mean radial velocity ($\langle V_r \rangle$) in the bowl-in-piston chamber. The distribution is shown for two different crank angles.

tive insignificance. For example, the uncertainty in crank speed was neglected because it was only 3%. The uncertainty was determined by monitoring the output of the crank angle sensor.

Rigorous static helium tests that were performed in the combustion chamber also indicated that leakage was negligible. The strongest evidence that leakage and crevice-flow were insignificant was the absence of PIV velocity vectors pointing toward and being near the cylinder wall. Any such vectors would have revealed flow into crevices by simple continuity arguments. Furthermore, short exposure photographs (of a side of the piston) during UBCRICM operation revealed a squish gap that matched model specifications.

The possibility of particle flow infidelity (to the bulk motion) did exist if the nylon PIV particles had a tendency to agglomerate. Nevertheless, this cause was also dismissed because a finer oil mist that was tried yielded similar velocity results to the nylon particles. The smaller size of the olive oil particles in the mist was readily confirmed by PIV image inspection. The oil mist was not regularly used for the PIV measurements because of the difficulty in repeating particle concentration and uniformity. PIV processing errors were also considered negligible causes for the discrepancy because individual particle displacements closely matched PIV results. The individual displacements were measured (in pixels) between the two successive PIV images on occasions when individual particles were discernible.

The remaining cause for the discrepancy that could not be eliminated without further investigation was incorrect representation of physical processes (in particular the process of nonuniform heat transfer).

5 An Improved Model for Predicting Squish Velocity

The acknowledgment that heat transfer may differ between zones 1 and 2 (the squish volume and the raised bowl volume, respectively) means that each zone may be modeled with distinctly different temperatures and densities. In order to solve for

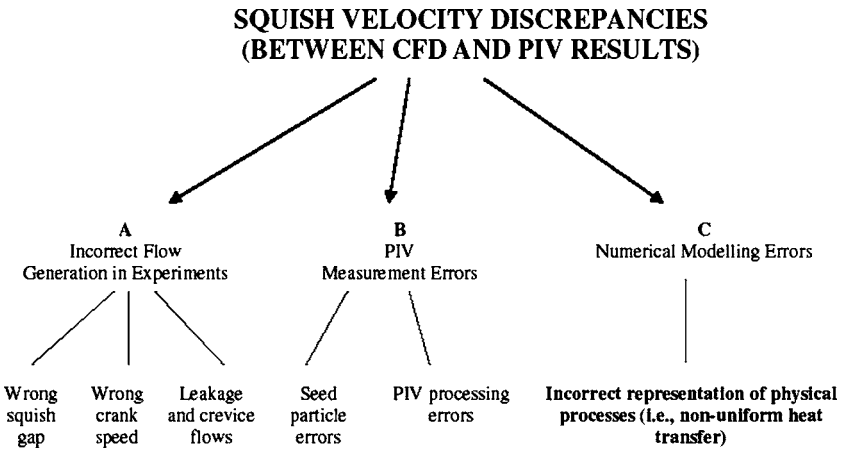


Fig. 12 Classification of possible causes for discrepancies in squish velocity between PIV results and mathematical model

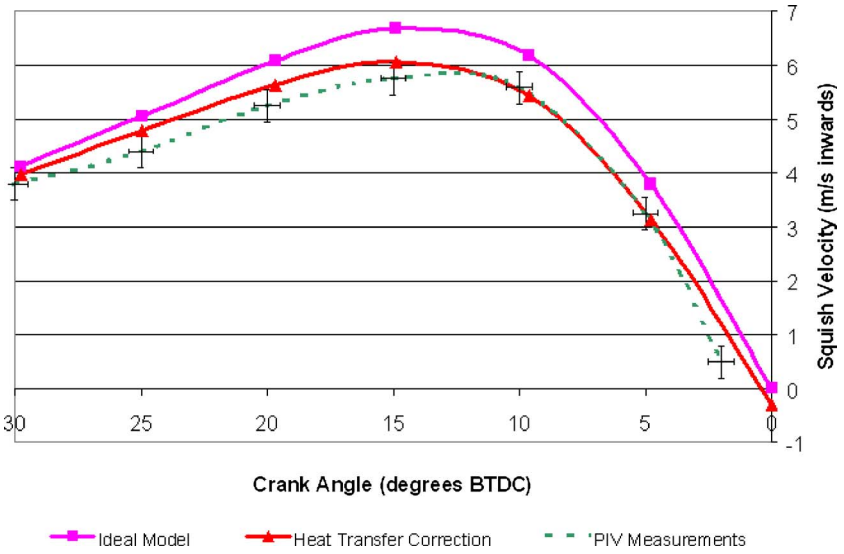


Fig. 13 Effect of heat transfer correction on predicted squish velocity history

these properties and squish velocity, mass conservation and energy conservation is applied to each zone. The ideal gas equation and the assumption of uniform pressure are then used to achieve closure. Using the notation from the ideal model, the following equations describe the solution procedure.

The mass balance for zone 1 is expressed as

$$V_1 d\rho_1 = v_p A_1 \rho_1 dt - dm_s \tag{8}$$

Similarly a mass balance applied to zone 2 yields

$$V_2 d\rho_2 = v_p A_2 \rho_2 dt + dm_s \tag{9}$$

where

dt is the time step,
dm_s is the mass exiting the squish volume during dt.

The energy balance for zone 1 is

$$-|Q_1| = \rho_1 V_1 C_V dT_1 + \frac{P dm_s}{\rho_1} - P A_1 v_p dt \tag{10}$$

Similarly, the energy balance for zone 2 is

$$-|Q_2| = \rho_2 V_2 C_V dT_2 - \frac{P dm_s}{\rho_1} - P A_2 v_p dt \tag{11}$$

where

Q_1 and Q_2 are the heat losses in zones 1 and 2, respectively,
dT_1 and dT_2 are the temperature increments in zones 1 and 2, respectively.

The heat loss terms are modeled by the function

$$Q_i = h_i A_{wi} (T_i - T_{wi}) dt \tag{12}$$

where i is the zone index

h is the convective heat transfer coefficient (found from the Annand correlation [13]),
T_{wi} is the inner wall temperature of the combustion chamber zone i,
A_{wi} is the wall surface area in zone i.

T_{wi} is calculated by numerically solving for unsteady heat conduction in the combustion chamber walls.

For closure, the ideal gas law is used for both zones with the assumption that the pressure (P) is uniform:

$$\frac{dP}{P} = \frac{d\rho_1}{\rho_1} + \frac{dT_1}{T_1} \tag{13}$$

$$\frac{dP}{P} = \frac{d\rho_2}{\rho_2} + \frac{dT_2}{T_2} \tag{14}$$

The above analysis contains six equations with six unknowns, namely $d\rho_1$, $d\rho_2$, dT_1, dT_2, dP, and dm_s. Therefore, all unknowns can be solved for each time step. The value of dm_s is used to determine the squish velocity, v_s, by the relation

$$dm_s = v_s \rho_2 A_s dt \tag{15}$$

The equations were numerically integrated and the improvement of the revised model is evident in Fig. 13.

The value of h_1 increased fivefold during the compression simulation mainly because of the influence of charge density in the Annand correlation. At 20 deg BTDC, h_1 was 873 W/m^2K. This value cannot be directly confirmed because of the absence of independent measurements on similar rapid compression machines. Nevertheless, the value was within 10% of the heat transfer coefficient (953 W/m^2K at 20 deg BTDC) estimated from motored data taken by Alkidas [14]. Alkidas used a conventional spark ignition engine operating at similar conditions to the UBCRICM. His engine had a disk-shaped combustion chamber and it operated with little swirl. The ratio for h_2/h_1 that was used in the revised model was 5 and it reflects the heat transfer increase in zone 2 due to turbulence. This ratio is very important because it influences the density ratio (and hence the squish velocity) between the squish volume and the bowl volume. The ratio was assigned after observing with PIV that, for crank angles between 30 deg BTDC and TDC, the turbulent fluctuations just inside the bowl lip were on average five times as high as in the squish zone. Based on the findings of Maciejewski and Moffat [15], the convective heat transfer coefficient was approximated to be proportional to the turbulent fluctuations. It should also be noted that when h_2 and h_1 were set to zero, the squish velocity history matched the ideal model's results, as expected.

6 Conclusions

Measurements of the squish velocity inside the bowl-in-piston combustion chamber and comparison of the results with output from numerical simulations yielded the following major conclusions:

1. Besides factors such as a particle's flow following ability and visibility, PIV image resolution is very important when determining the size of seed particles used. In many cases, when this resolution is coarser than the smaller scales of turbulence, the biggest (and therefore the most conspicuous) possible particles should be used that can follow turbulence scales corresponding to the resolution. Smaller particles would be redundant because, although they may be able to follow the smaller turbulence scales in the flow, these scales could not be resolved by the PIV system. Since the smaller scales are not usually resolved in PIV, rms velocities tend to be more uncertain than point measurements, making PIV more suitable for global mean velocity measurements. Ideally, PIV system resolution should be fine enough to discern a wide range of turbulent scales without compromising the system's ability to capture a flow's global detail.
2. Despite the simplicity of the ideal two-zone mass transfer model for simulating squish velocities inside the bowl-in-piston combustion chamber, the measured peak squish velocity was 12% less than the prediction. The simple model's assumption of uniform density throughout the cylinder was a strongly suspected cause of the discrepancy. An improved model for simulating the squish velocity was developed by accounting for the density changes that are caused by heat transfer variations from nonuniform turbulence in the combustion chamber. This improved model provided a more accurate prediction of the measured squish velocity in a bowl-in-piston-type combustion chamber.

Acknowledgment

The authors wish to acknowledge the financial support of the University of British Columbia and Westcoast Energy Inc. through scholarships (UGF Scholarship and the Jack Davis Scholarship, respectively) to P. Lappas.

References

[1] Fansler, T. D., and French, D. T., 1987, "Swirl, Squish and Turbulence in Stratified-Charge Engines: Laser Velocimetry Measurements and Implications for Combustion," SAE Paper No. 870371.

[2] Monaghan, M. L., and Pettifer, H. F., 1981, "Air Motion and Its Effects on Diesel Performance and Emissions," SAE Paper No. 810255.

[3] Bacon, D. M., Renshaw, J., and Walker, K. L., 1981, "The Use of In-Cylinder Modelling and LDA Techniques in Diesel Engine Design," *Proceedings of the International Conference on New Energy Conservation Technologies and their Commercialization*, J. P. Milhone and E. H. Willis, eds., Springer-Verlag, Berlin, pp. 2431–2440.

[4] Johns, R. J. R., 1985, "The Effect of Piston Bowl Offset on the Compression-Induced Air Motion in Direct Injection Diesel Engine Combustion Chambers," *Proceedings of the International Symposium on Diagnostics and Modeling of Combustion in Reciprocating Engines*, Tokyo, pp. 489–502.

[5] Döhring, K., 1986, "The Relative Effects of Intake and Compression Generated Turbulence on I.C. Engine Combustion Duration," M.A.Sc. thesis, University of British Columbia.

[6] Lappas, P., 2003, "An Experimental and Computational Study of Flow in the Squish Jet Combustion Chamber," Ph.D. thesis, University of British Columbia.

[7] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York, pp. 353–357.

[8] Tabaczynski, R. J., 1976, "Turbulence and Turbulent Combustion in Spark-Ignition Engines," Prog. Energy Combust. Sci., **2**, pp. 143–165.

[9] Anetor, L., 1994, "Experimental and Numerical Simulation of Charge Motion in Internal Combustion Engines," Ph.D. thesis, University of British Columbia.

[10] Melling, A., 1997, "Tracer Particles and Seeding for Particle Image Velocimetry," Meas. Sci. Technol., **8**(12), pp. 1406–1416.

[11] Keane, R. D., and Adrian, R. J., 1992, "Theory of Cross-Correlation Analysis of PIV Images," Appl. Sci. Res., **49**, pp. 191–215.

[12] Bhattacharyya, G. K., and Johnson, R. A., 1977, *Statistical Concepts and Methods*, John Wiley and Sons, New York, Chap. 8

[13] Annand, W. J. D., 1963, "Heat Transfer in the Cylinders of Reciprocating Internal Combustion Engines," Proc. Inst. Mech. Eng., **177**(36), pp. 973–990.

[14] Alkidas, A. C., 1980, "Heat Transfer Characteristics of a Spark-Ignition Engine," ASME J. Heat Transfer, **102**, pp. 189–193.

[15] Maciejewski, P. K., and Moffat, R. J., 1992, "Heat Transfer with Very High Free-Stream Turbulence: Part II—Analysis of Results," ASME J. Heat Transfer, **114**(4), pp. 834–839.

Experimental Study on Homogeneous Charge Compression Ignition Combustion With Fuel of Dimethyl Ether and Natural Gas

Mingfa Yao
e-mail: y_mingfa@tju.edu.cn

Zunqing Zheng

Jin Qin

State Key Laboratory of Engine Combustion,
Tianjin University,
Tianjin 300072, China

The homogeneous charge compression ignition (HCCI) combustion fueled by dimethyl ether (DME) and compressed natural gas (CNG) was investigated. The experimental work was carried out on a single-cylinder diesel engine. The results show that adjusting the proportions of DME and CNG is an effective technique for controlling HCCI combustion and extending the HCCI operating range. The combustion process of HCCI with dual fuel is characterized by a distinctive two-stage heat release process. As CNG flow rate increases, the magnitude of peak cylinder pressure and the peak heat release rate in the second stage goes up. As DME flow rate increases, the peak cylinder pressure, heat release rate, and NO_x emissions increase while THC and CO emissions decrease.
[DOI: 10.1115/1.2130731]

1 Introduction

Homogeneous charge compression ignition (HCCI) combustion has the potential to realize high efficiency and low emissions [1–7]. HCCI engines can have efficiencies as high as compression-ignition, direct-injection (CIDI) engines, while producing ultra-low oxides of nitrogen (NO_x) and particulate matter (PM) emissions. In some regards, the HCCI engine incorporates the best features of both spark ignition (SI) gasoline and CIDI engines. Like an SI engine, the charge is well mixed so that the particulate emissions are minimized and, similar to a CIDI engine, it is compression ignited and has no throttling losses, which leads to high efficiency. The HCCI engine is expected as a next generation engine to take over in place of the conventional internal combustion engine [8]. However, the HCCI smooth operation occurs only in a limited region. The operation range is limited by misfire and knocking.

Studies have shown that the ignition and combustion of HCCI are dominated by the kinetics of chemical reaction of the fuel with air. Controlling ignition timing is one of the challenging tasks in the operation of HCCI engines because HCCI ignition is determined by the charge mixture composition, its time-temperature history, and to a lesser extent pressure. Therefore, an indirect approach has to be applied for controlling ignition timing. General solutions to this problem include exhaust gas recirculation (EGR), variable compression ratio (VCR), and variable valve timing (VVT), etc. [9,10]. In addition, when the engine operates at high loads, the nature of HCCI combustion can lead to knock due to its high burning rates. On the other hand, operation on fuels with high octane number at light loads may cause incomplete combustion or misfire resulting from lower flame temperature caused by much smaller burning rates. Therefore, extending the operating range of HCCI engines is another challenge in the attempt to take advantage of HCCI technique. One approach to obtain a wider operation range is stratified combustion with multiple pulsed injections of fuel and water to optimize the mixing of charge [11,12]. Another important approach is to use two fuels that have different properties. An adjustment of the proportion between high cetane number fuel and high octane number fuel has been reported as a method for controlling the ignition timing in HCCI combustion [13,14].

Shudo et al. [13] have proposed a new HCCI combustion engine system fueled with DME and methanol-reformed gas (MRG). This system has demonstrated high thermal efficiency over a wide range of equivalence ratio. MRG effectively controls the heat release timing of the second stage by the high-temperature reactions in HCCI of DME to expand operable range of equivalence ratio and engine load. Chen et al. [14] have investigated the effects of HCCI combustion using the mixture of natural gas and DME in various proportions. The results show that the mixture could run the engine quietly and smoothly over a wide load range. The thermal efficiency was higher than that of fuel direct injected diesel when the DME proportion was optimized. For a fixed airflow only a small amount of DME was required to obtain auto-ignition, while further DME addition had little effect on advancing the start of ignition.

This paper reports the investigation of the HCCI combustion strategy on a dual fueled test engine using DME and CNG. The effects of DME and CNG concentrations on the combustion, performance and exhaust emissions were investigated in the case of same brake mean effective pressure (BMEP). The aim of this research was to analyze the influence of DME and CNG on HCCI combustion, engine performance, and pollutant emissions.

2 Experimental Setup and Apparatus

The experiment was conducted on a single-cylinder, water-cooled, direct-injection diesel engine, with bore and stroke both 115 mm, ω shape combustion chamber, 17 compression ratio, rated power of 14.7 kW at 2200 rev/min. An electronically controlled fuel injection system was used for fueling DME and CNG. The electromagnetic valve was mounted close to the intake port. Electronically controlled fuel injection allows the adjustment of the injected amount of each fuel according to the engine operating conditions. A homogenous mixture of DME and CNG with air was formed during the compression stroke. Ignition timing and combustion rate can be controlled by adjusting the relative con-

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received February 2, 2004; final manuscript received September 22, 2005. Review conducted by D. Assanis.

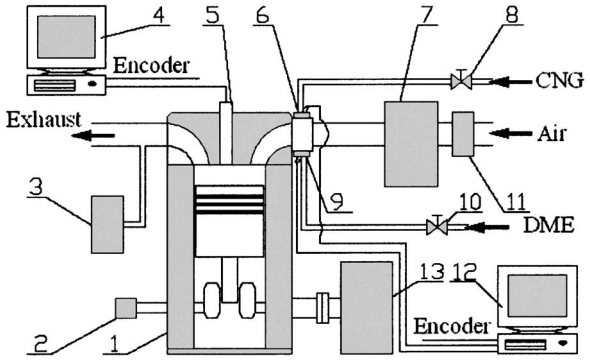


Fig. 1 Schematic of test setup

centration of DME and CNG, and both the lower and higher loads in HCCI operating range can be extended through this system. In order to reduce the effect of intake air cycling variation on the measurement of intake air flow rate, a large tank was added as a pressure stabilizer in the intake system. The flow meter was mounted at the inlet of the tank. Figure 1 illustrates the experimental setup. The cylinder pressure was recorded by a pressure transducer that was mounted at the cylinder head and connected to a Labview data acquisition system. The rates of heat release were calculated using a zero-dimensional model. A dynamometer was used to estimate the torque and, depending on that, brake mean effective pressure (BMEP) was calculated. An exhaust gas analyzer was adopted to measure the NO_x, CO, and THC emissions.

Test conditions were set as follows:

Engine speed: 1400 rev/min
Coolant temperature: 80°C
Oil temperature: 70°C
Inlet charge temperature: 16–18°C

3 Results and Discussions

In order to present the excess air ratio of blends of DME and CNG and the relative proportions of them, the total excess air ratio of two fuels and the proportions of DME and CNG are defined. The total excess air ratio is expressed as

$$\lambda_{\text{total}} = \dot{m}_{\text{air}}/(\dot{m}_{\text{CNG}} \cdot f_{\text{CNG\_}th} + \dot{m}_{\text{DME}} \cdot f_{\text{DME\_}th}) \quad (1)$$

The proportion of DME is defined by

$$R_{\text{DME}} = \frac{\dot{m}_{\text{DME}} \cdot Q_{\text{HV}_{\text{DME}}}}{\dot{m}_{\text{CNG}} \cdot Q_{\text{HV}_{\text{CNG}}} + \dot{m}_{\text{DME}} \cdot Q_{\text{HV}_{\text{DME}}}} \times 100\% \quad (2)$$

so the proportion of CNG can be defined by

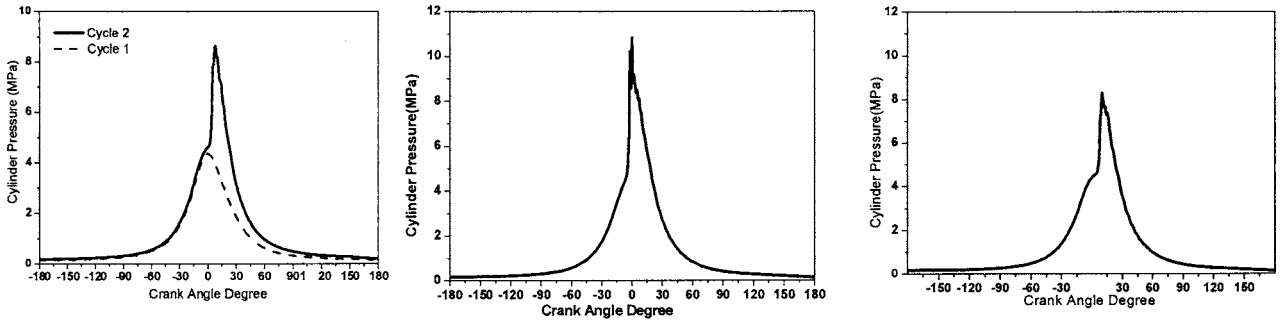
$$R_{\text{CNG}} = 1 - R_{\text{DME}} \quad (3)$$

where $\dot{m}_{\text{air}}$ is the air flow rate, $\dot{m}_{\text{CNG}}$ and $\dot{m}_{\text{DME}}$ represent the flow rates of CNG and DME, and $f_{\text{CNG\_}th}$ and $f_{\text{DME\_}th}$ are the theoretical air fuel ratios of CNG and DME. $Q_{\text{HV}_{\text{CNG}}}$ and $Q_{\text{HV}_{\text{DME}}}$ are the lower heating values of CNG and DME.

3.1 Ignition Limit and Knock Limit of HCCI for DME/CNG Dual Fuel. With a compression ratio of 17, the diesel engine used in this study could not achieve compression ignition of CNG due to its high temperature point for auto-ignition. In the HCCI process of DME/CNG dual fuel, ignition of CNG is achieved through heat release from the burning of DME, which first ignites through compression. Consequently, if the proportion of DME is too low, its heat release is not sufficient to ignite CNG, causing "misfire" or incomplete combustion as shown in Fig. 2(*a*). This situation results in large cycle-to-cycle variation, reduction in power, and dramatic increase in HC emission. As illustrated in Fig. 2(*b*), two situations could cause knocking combustion: (1) when DME quantity is very high, resulting in earlier ignition and more rapid burning; and (2) when CNG quantity is very high and DME quantity is high enough to ignite CNG, causing rapid combustion of CNG. Therefore, the HCCI engine fueled by DME/CNG is limited by knocking combustion as well. At a particular load, by controlling the injected quantity of each fuel to ensure normal combustion (see Fig. 2(*c*)), the engine can run smoothly and achieve a better thermal efficiency and a very low NO_x emissions.

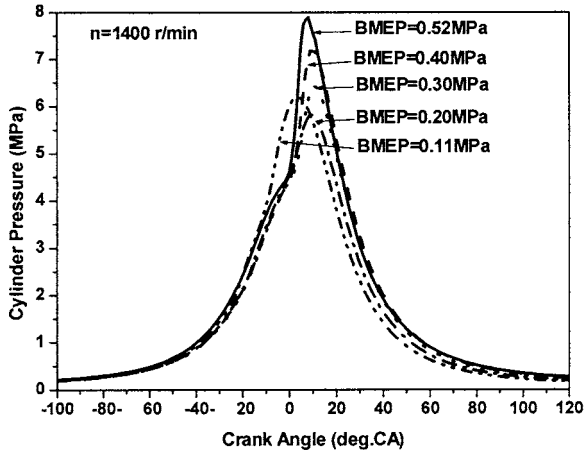
3.2 Combustion Characteristics

3.2.1 Effects of CNG Combustion Process. Chen et al. [14] reported that a certain amount of DME does not affect the lean burning limit. This suggests that the engine running at that load could be adjusted by changing CNG flow rate. A higher load requires a higher CNG flow rate. The cylinder pressure diagram and heat release rates of the engine running at various loads are shown in Fig. 3. Figure 3(*a*) gives the indicator diagrams for different BMEPs, and Fig. 3(*b*) shows the rates of heat release and lists the total excess air ratios λ_{total} corresponding to different BMEPs at 1400 rev/min. The results show that the HCCI process of DME/CNG dual fuel is characterized by a two-stage heat release process, and the second heat release begins near TDC. Due to the limitation introduced by knocking combustion, a lower $\dot{m}_{\text{DME}}$ was supplied during the experiment at BMEP of 0.52 MPa, while a higher $\dot{m}_{\text{DME}}$ was used at BMEP of 0.11 MPa. The $\dot{m}_{\text{DME}}$ was kept almost unchanged for other three situations. Table 1 lists the proportions and flow rates of DME and CNG (R_{DME}, R_{CNG}, $\dot{m}_{\text{DME}}$, $\dot{m}_{\text{CNG}}$) for different test points shown in Figs. 3 and 4. The results also show that the first peak of heat release rate and that of ignition timing are related to DME flow rate. A higher DME flow rate leads to earlier heat release and hence a higher peak in the

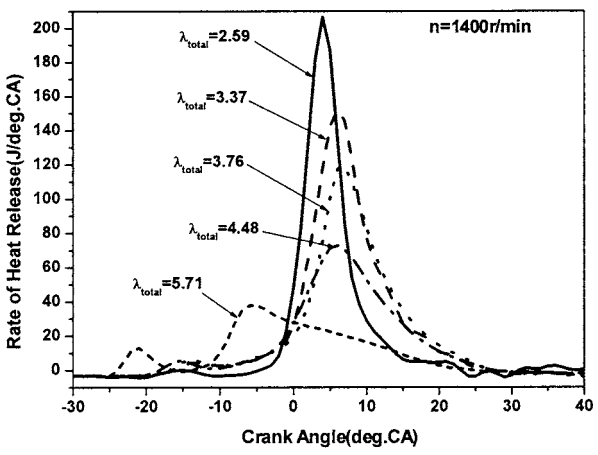


(a) *Partial combustion* (b) *Knocking combustion* (c) *Normal combustion*

Fig. 2 Abnormal and normal HCCI process for DME/CNG dual fueled engine



(a) *Cylinder pressure*



(b) *Rate of heat release*

Fig. 3 Effects of CNG on combustion process

first stage. A higher CNG flow rate results in peak pressure increased and peak value of heat release rate higher. The results also show that when DME flow rate is kept constant, almost no difference can be found in the time interval between the two peaks of heat release rate as CNG flow rate is adjusted.

Although values (such as 5% and 10%) of mass fraction burned are usually defined ignition, it is believed that the combustion efficiency is lower at leaner charge in HCCI combustion. In this investigation, ignition timing is defined by the crank angle at which 5% of the total heat has released by the burnt fuel, and the primary burning duration is defined as the period in which 5%–90% of the total heat has released. Figure 4 shows that the ignition timing (5%) varies a bit (within 1 degree of crank angle) and the primary burning duration (5%–90%) does not vary very much either when CNG flow rate is adjusted. It can be concluded that CNG flow rate has little effect on the primary burning duration. However, when the load is high (BMEP=0.52 MPa), since the mixture is relatively rich, the ignition timing still varies slightly. Nevertheless, heat release rate reaches its second peak earlier when the load is high, leading to a sharp reduction in primary burning duration. At higher loads, because of high concentration, the rapid of burning of mixture will cause knocking combustion, so it is desirable to find approaches that can help slow down the burning rate at higher load. The introduction of cold EGR may be an effective technique to control the burning rate.

On the other hand, at BMEP of 0.11 MPa, due to a higher DME flow rate, the ignition timing comes earlier and the second peak in heat release rate occurs prior to TDC, resulting in an even higher peak pressure than that at BMEP of 0.2 MPa. Additionally, the mixture being too lean prolongs the primary burning duration. It is evident that all these factors do not bring any benefits in terms of increasing the engine thermal efficiency.

The experimental results show that CNG affects both the first stage heat release (heat released by the low-temperature reaction of DME) and the second stage heat release (heat released by high-temperature reaction of DME and CNG). They are well in agreement with those from the chemical kinetic modeling work [15].

Table 1 The λ_{total}, R_{DME}, R_{CNG}, m_{DME}, m_{CNG} for Figs. 3 and 4

BMEP (MPa)	λ_{total}	R_{DME} (%)	R_{CNG} (%)	$\dot{m}_{DME}$ (kg/h)	$\dot{m}_{CNG}$ (kg/h)
0.52	2.59	22.1	77.9	0.523	1.007
0.40	3.37	33.5	66.5	0.612	0.665
0.30	3.76	37.4	62.6	0.615	0.563
0.20	4.48	44.5	55.5	0.616	0.420
0.11	5.71	60.4	39.6	0.663	0.237

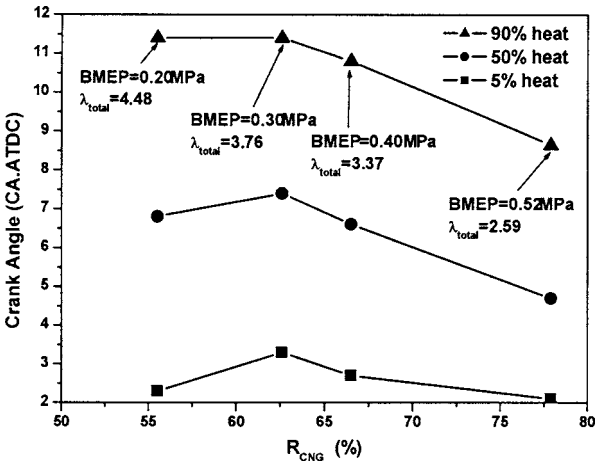
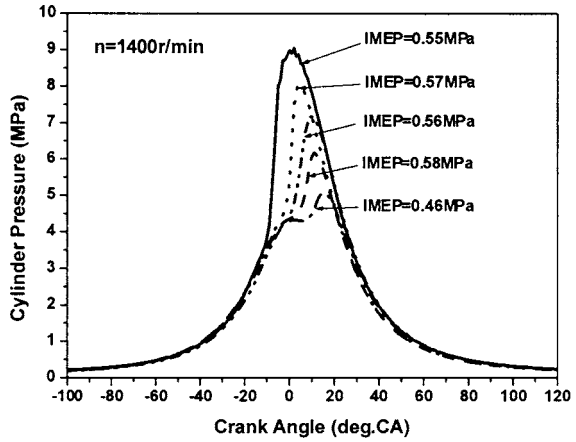


Fig. 4 Effects of CNG on burning rates

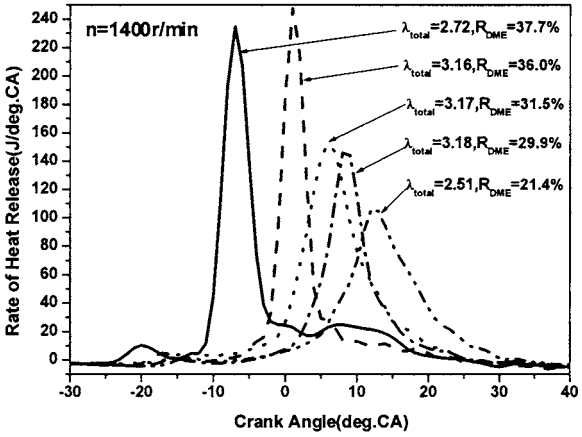
3.2.2 Effects of DME on Combustion Process. Figure 5 summarizes the results from a series of tests designed to investigate the effect of DME on the HCCI combustion process of DME/CNG dual fueled engine running at 1400 rev/min with BMEP about 0.4 MPa. Figure 6 illustrates the effect of DME on the mean temperature in cylinder. Table 2 lists the proportions and flow rates of DME and CNG. It can be seen that the maximum cylinder pressure and temperature increase as DME flow rate goes up; knock occurs when the DME flow rate is larger than 0.775 kg/h ($R_{DME}>37.7\%$) in the case of BMEP about 0.4 MPa. An increased DME flow rate advances the ignition timing forward, enhances heat release rate, and shortens both burning duration and the time interval between two peaks of heat release rate. Although

Table 2 The λ_{total}, R_{DME}, R_{CNG}, m_{DME}, m_{CNG} for Figs. 5 and 6

IMEP (MPa)	λ_{total}	R_{DME} (%)	R_{CNG} (%)	$\dot{m}_{DME}$ (kg/h)	$\dot{m}_{CNG}$ (kg/h)
0.46	2.51	21.4	78.6	0.518	1.050
0.58	3.18	29.9	70.1	0.573	0.741
0.56	3.17	31.5	68.5	0.606	0.726
0.57	3.16	36.0	64.0	0.697	0.684
0.55	2.72	37.7	62.3	0.775	0.706



(*a*) *Cylinder pressure*



(*b*) *Rate of heat release*

Fig. 5 Effects of DME on combustion process

$\lambda_{total}=2.51$ is the lowest excess air ratio (rich mixture) among the five test points, incomplete combustion occurs because of the lowest DME flow rate ($m_{DME}=0.518$ kg/h), and the indicated mean effective pressure (IMEP) decreases. Therefore, the principal function of DME in the dual-fuel system is to initiate the combustion of two fuels and to regulate the burning rate of the mixture.

Figure 7 shows that DME flow rate greatly affects the ignition timing and the burning rate in the first stage. As DME flow rate increases, ignition timing advances and the first peak in heat release rate goes up. As a result of the effect on the first stage, DME flow rate will also affect the second stage: the second peak appears earlier and becomes larger when a higher DME flow rate is used. As a summary, the effect of DME flow rate on the combustion process is due mainly to the fact that higher DME flow rate enhances the cool-flame heat release, resulting in elevated in-cylinder temperature (as shown in Fig. 6) and increased burning rate (as shown in Fig. 5(*b*)). Also, the enhancement in cool-flame chemistry yields more active radicals that will speed up the burning rate of the second stage and thus lead the heat release rate to its peak earlier [15].

3.3 Brake Thermal Efficiency (BTE)

3.3.1 Effects of DME on BTE. Figure 8(*a*) illustrates the effect of DME on BTE of the DME/CNG dual fuel HCCI engine. The results indicate that high BTE appears in normal combustion, and, between incomplete combustion and knocking combustion, DME proportion has no remarkable influence on BTE. However, when DME proportion is high enough, the engine will knock. Because knock breaks the thermal boundary layer, heat loss is increased [16], and consequently BTE is reduced. Tadashi et al. [17] also suggested that it is possible to reduce heat loss resulting from knocking combustion by controlling the ignition timing at an appropriate timing after TDC. Therefore, it is necessary to decrease DME flow rate at high engine load. On the other hand, when DME proportion is close to lean burning limit, incomplete combustion will happen. This situation will lead to large cycle-to-cycle variations and high HC emissions, and BTE will decrease greatly.

Figure 9 shows the knock limit and lean burning limit according to Fig. 8(*a*), including the relationship between λ_{total} and R_{DME} (Fig. 9(*a*)), and the relationship between λ_{total} and m_{DME} (Fig. 9(*b*)). Apparently, as engine load increases (BMEP increases), the DME proportions of knock limit and lean burn limit decrease, but DME flow rate of partial combustion limit increases slightly (Fig. 9(*b*)). That is to say, an increased CNG flow rate demands DME flow rate increasing to initiate CNG ignition. This result is different from Chen's result [14]. Furthermore, the DME flow rate of knock limit decreases as engine load increases. Therefore, as engine load increases, between the knock limit and lean burn limit, the acceptable variable ranges of DME flow rate and DME proportion decrease.

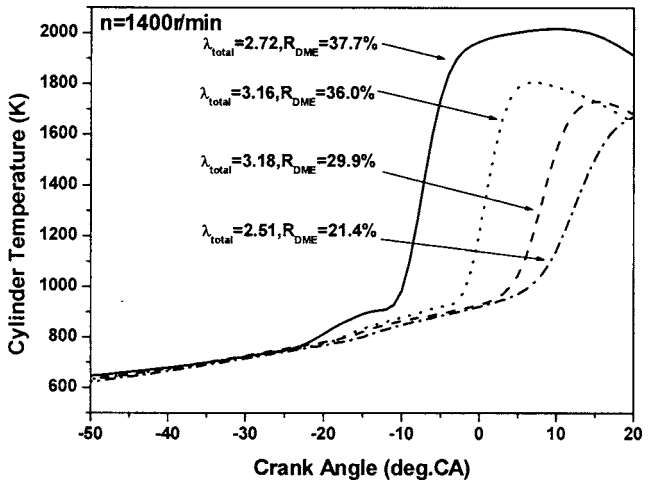


Fig. 6 Effects of DME on the temperature in cylinder

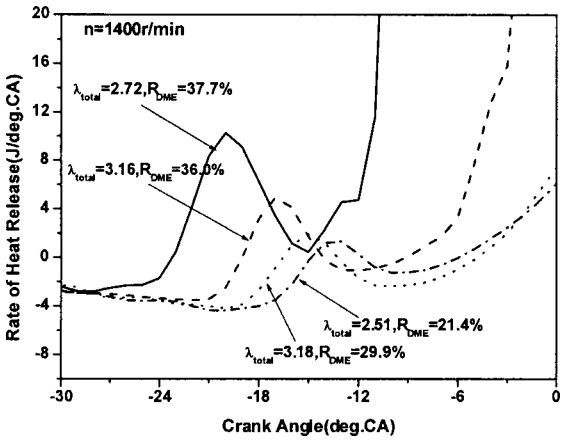
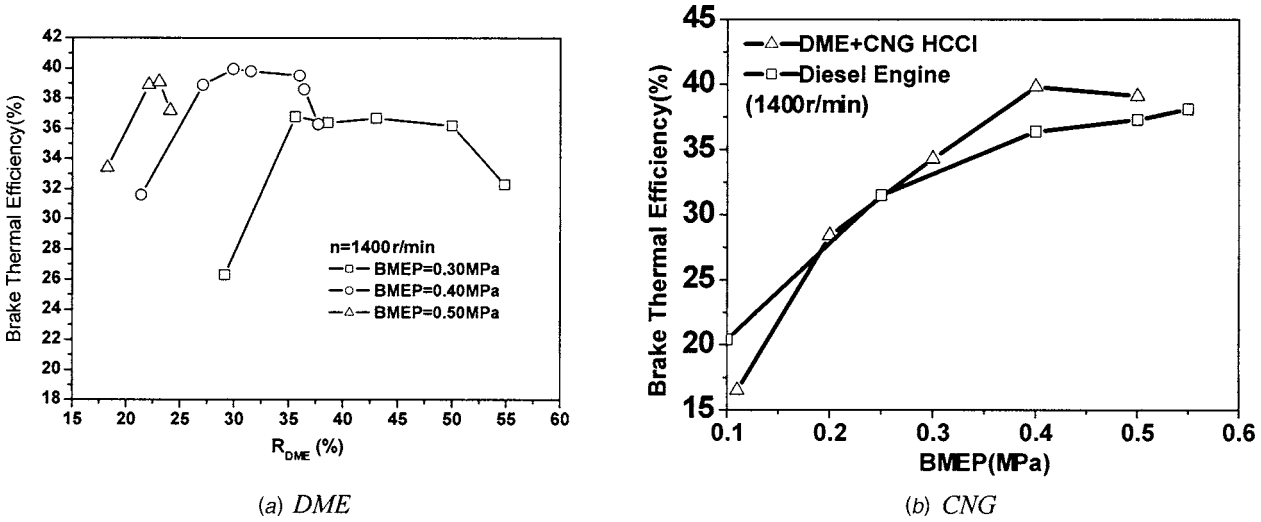
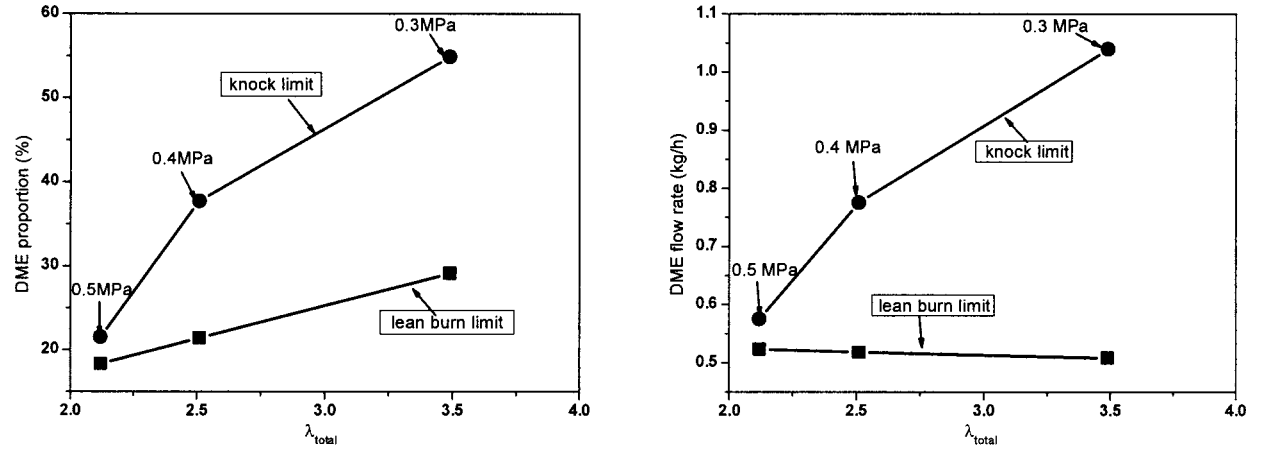


Fig. 7 Effects of DME on the first stage heat release



(*a*) *DME* (*b*) *CNG*

Fig. 8 Effects of DME and CNG proportions on BTE



(*a*) DME proportion (*b*) DME flow rate

Fig. 9 Knock limit and lean burning limit

3.3.2 Effects of CNG on BTE. Figure 8(*b*) shows the effect of CNG flow rate on BTE of the HCCI engine, and Table 3 lists the total excess air ratio, DME proportion and flow rate, etc. As mentioned above, if the DME flow rate is kept constant, the higher the amount of CNG is supplied, the greater BMEP is achieved. In this test, $\dot{m}_{DME}$ was set at about 0.610 kg/h. Results in Fig. 8(*b*) suggest that high BTE can be achieved by increasing the CNG flow rate. However, as the CNG flow rate gets closer to the knock limit, the BTE decreases slightly (i.e., BMEP=0.5 MPa). Figure 8(*b*) also indicates that the BTE of the HCCI engine is higher than that of original diesel engine at middle load. At light load, the HCCI combustion system must be optimized in order to improve the BTE of the engine. One effective technique is to use EGR. The effects of EGR on BTE are being investigated, and the results will be reported elsewhere.

Table 3 The λ_{total}, R_{DME}, R_{CNG}, $\dot{m}_{DME}$, and $\dot{m}_{CNG}$ for Fig. 8(*b*)

BMEP (MPa)	λ_{total}	R_{DME} (%)	R_{CNG} (%)	$\dot{m}_{DME}$ (kg/h)	$\dot{m}_{CNG}$ (kg/h)
0.11	5.71	60.4	39.6	0.663	0.237
0.20	4.48	44.5	55.5	0.616	0.420
0.30	3.89	38.6	61.4	0.609	0.534
0.40	3.17	31.5	68.5	0.606	0.726
0.50	2.48	23.1	76.9	0.564	1.038

3.4 Emissions

3.4.1 THC and CO Emissions. HCCI engines have inherently low emissions of NO_x and PM (particulate matter), but relatively high emissions of hydrocarbons (HC) and carbon monoxide (CO). In this study, HCCI combustion fueled with DME and CNG exhausted very high levels of THC and CO emissions compared to those from the original diesel (not shown in this paper). HC emissions from HCCI combustion at moderate and high loads are shown to originate primarily from crevices [18]. The amount of HC leaving the cylinder is therefore, as a first-order approximation, controlled by the amount of fuel trapped in the crevices at the time when bulk reactions freeze. On the other hand, the formation of CO is much more complex. CO is believed to be formed close to the walls where the temperature is high enough for the oxidation of HC starting, but the cooling from the wall prevents complete oxidation to CO_2 [19]. For this mechanism, both wall temperature and bulk temperature history have strong impact on CO emissions. Figure 10 illustrates how CO and THC emissions are dependent upon DME proportion. When DME pro-

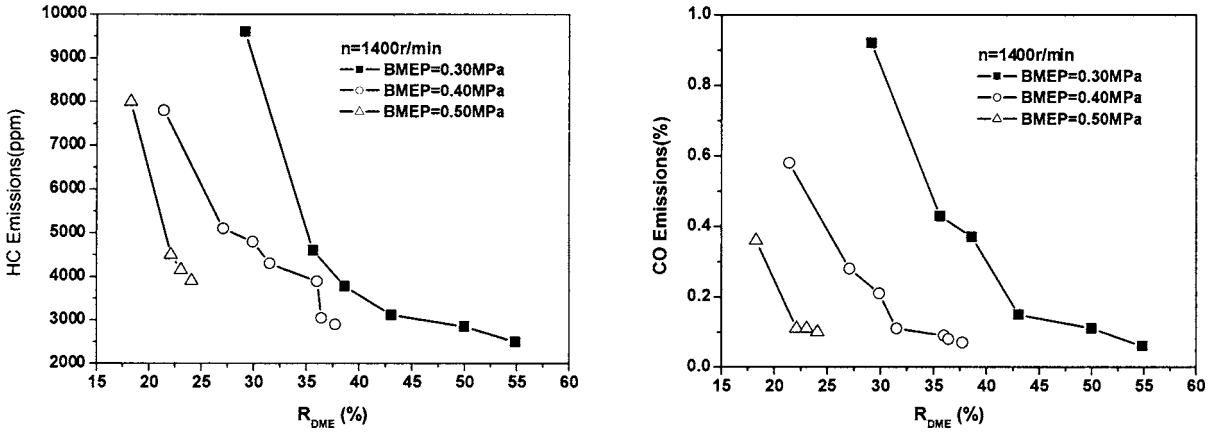


Fig. 10 Effect of DME proportion on THC and CO emissions

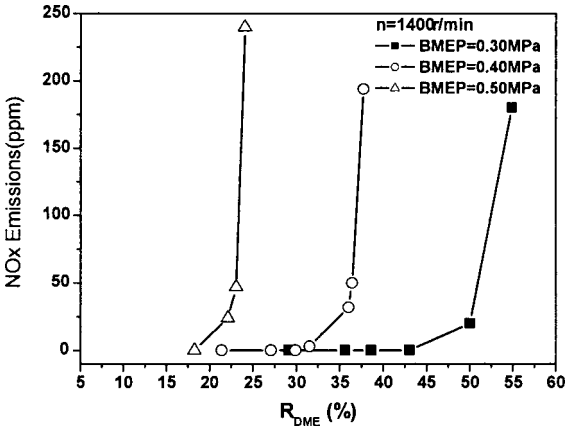


Fig. 11 Effects of DME proportion on NOx emissions

portion is adjusted to decrease at constant load mode (while CNG proportion is adjusted to increase accordingly), both the THC and CO emissions increase (Fig. 10). As the DME proportion gets closing to the ignition limit, THC and CO emissions increase rapidly. Therefore, high CO and THC emissions are the characters of partial combustion or misfire in HCCI.

3.4.2 NO_x *Emissions*. NO_x formation is mainly dependent on combustion temperature. The important attribute of HCCI is to allow combustion to occur at a very low temperature. NO_x emissions from the engine can be dramatically reduced. Figure 11 illustrates the effects of DME proportion on NO_x emissions. As mentioned above, the cylinder pressure increases as DME proportion goes up, leading to increased combustion temperature. Therefore, NO_x emissions are directly related to DME proportion, meaning NO_x emissions increase as DME proportion goes up. In addition, when the DME proportion reaches the knock limit, NO_x emissions increase rapidly. Therefore, high NO_x emission is another character of knocking combustion in HCCI.

4 Conclusions

This article reports experimental results related to the combustion characteristics, performances, and emissions characteristics of the HCCI engine with fuel of DME and CNG. Experimental results are summarized below.

1. Adjustment of the proportion of DME and CNG is an effective technique to control HCCI combustion and extend the HCCI operating range. The maximum BMEP of HCCI operation with fuel of DME and CNG dual fuel can reach 0.52 MPa at 1400 rpm.
2. The combustion process of HCCI with dual fuel DME/CNG is characterized by a distinctive two-stage heat release process. CNG affects heat release both in the first stage and in the second stage. As CNG flow rate increases, the magnitude of the second peak goes up, but the burning duration almost remains unchanged for a certain range of CNG flow rate.
3. The effect of DME on the ignition and combustion of HCCI process focuses on the first stage. As DME flow rate increases, auto-ignition occurs earlier, the two peak heat release rates increase, and the interval and burning duration between the two peak heat release rates become shorter.
4. When DME flow rate is very high, the engine will knock, causing BTE to be reduced; on the other hand, if DME flow rate is very small, incomplete burning will cause significant reduction in BTE. In normal combustion conditions, DME flow rate has little effect on BTE.
5. As DME flow rate increases, NO_x emissions increase and THC and CO emissions decrease. When the DME flow rate gets close to the knock limit, NO_x emissions increase rapidly. In contrast, if it gets close to the ignition limit, THC and CO emissions increase rapidly.

Acknowledgment

It is the state key project of fundamental research plan (973) of "new generation of engine combustion principle and approach to application of alternative fuels." Grant No. 2001CB209201. It is also founded by the National Natural Science Found of China (NSFC). The ratified number is 50106007.

References

[1] Najt, P. M., and Foster, D. E., 1983, "Compression-Ignited Homogeneous Charge Combustion," SAE Paper No. 830264.
[2] Gray, A. W., III, and Ryan, T. W., III, 1997, "Homogeneous Charge Compression Ignition (HCCI) of Diesel Fuel," SAE Paper No. 971676.
[3] Aoyama, T., Hattori, Y., Mizuta, J., and Sato, Y., 1996, "An Experimental Study on Premixed-Charge Compression Ignition Gasoline Engine," SAE Paper No. 961160.
[4] Ryan, T. W., and Callahan, T. J., 1996, "Homogeneous Charge Compression Ignition of Diesel Fuel," SAE Paper No. 961160.
[5] Suzuki, H., Koike, N., and Odaka, M., 1998, "Combustion Control Method of Homogeneous Charge Diesel Engines," SAE Paper No. 980509.
[6] Oakley, A., Zhao, H., and Ladommatos, N., 2001, "Experimental Studies on Controlled Auto-Ignition (CAI) Combustion of Gasoline in a 4-Stroke Engine," SAE Paper No. 2001-01-1030, SP-1623.
[7] Kimura, S., Aoki, O., Kitahara, Y., and Aiyoshizawa, E., 2001, "Ultra-Clean

Combustion Technology Combining a Low-Temperature and Premixed Combustion Concept for Meeting Future Emission Standards," SAE Paper No. 2001-01-0200, SP-1623.
[8] U.S. Department of Energy Efficiency and Renewable Energy Office of Transportation Technologies, 2001, "Homogeneous Charge Compression Ignition (HCCI) Technology-A Report to the U.S. Congress."
[9] Christensen, M., Hultqvist, A., et al., 1999, "Demonstrating the Multi Fuel Capability of a Homogeneous Charge Compression Ignition Engine with Variable Compression Ratio," SAE Paper No. 1999-01-3679.
[10] Li, J., Zhao, H., and Ladommatos, N., 2001, "Research and Development of Controlled Auto-Ignition (CAI) Combustion in a 4-Stroke Multi-Cylinder Gasoline Engine," SAE Paper No. 2001-01-3608.
[11] Su, W., Lin, T., and Pei, Y., 2003, "A Compound Technology for HCCI Combustion in a DI Diesel Engine Based on the Multi-Pluse Injection and the BUMP Combustion Chamber," SAE Paper No. 2003-01-0741.
[12] Christensen, M., and Johansson, B., 1999, "Homogeneous Charge Compression Ignition with Water Injection," SAE Paper No. 1999-01-0182.
[13] Shudo, T., Ono, Y., and Takahashi, T., 2002, "Influence of Hydrogen and Carbon Monoxide on HCCI Combustion of Dimethyl Ether," SAE Paper No. 2002-01-2828.
[14] Chen, Z., Konno, M., Oguma, M., and Yanai, T., 2000, "Experimental Study of CI Natural-Gas/DME Homogeneous Charge Engine," SAE Paper No. 2000-01-0329.
[15] Qin, J., and Yao, M. F., 2004, "Numerical Study on the Dimethyl Ether/Compressed Natural Gas HCCI Combustion Using Chemical Kinetics Model," Combust. Sci. Technol., **10**(1), pp. 225–231.
[16] Tsurushima, T., Harada, A., Iwashiro, Y., Enomoto, Y., Asaumi, Y., and Aoyagi, Y., 2001, "Thermodynamic Characteristics of Premixed Compression Ignition Combustions," SAE Paper No. 2001-01-1891.
[17] Tsurushima, T., Kunishima, E., Asaumi, Y., and Aoyagi, Y., 2002, "The Effect of Knock on Heat Loss in Homogeneous Charge Compression Ignition Engines," SAE Paper No. 2002-01-0108, SP-2688.
[18] Christensen, M., Hultqvist, A., and Johansson, B., 2001, "The Effect of Toplland Geometry on Emissions of Unburned Hydrocarbons From a Homogeneous Charge Compression Ignition (HCCI) Engine," SAE Paper No. 2001-01-1893.
[19] Aceves, S., Martinez-Frias, J., Flowers, D., Smith, J., Dibble, R., Wright, J., and Hessel, R., 2001, "A Decoupled Model of Detailed Fluid Mechanics Followed by Detailed Chemical Kinetics for Prediction of Iso-Octane HCCI Combustion," SAE Paper No. 2001-01-3612.

Three-Dimensional Catalytic Regeneration Modeling of SiC Diesel Particulate Filters

George Pontikakis

Anastassios Stamatelos[1]
e-mail: stam@uth.gr

Mechanical Engineering Department,
University of Thessaly,
383 34 Volos, Greece

Increasingly stringent diesel particulate emissions standards have reestablished international interest in diesel filters, whose first series application dates back to 1985. Modern diesel engine technology, with computerized engine management systems and advanced, common rail injection systems, needs to be fully exploited to support efficient and durable diesel filter systems with catalytic aids, as standard equipment in passenger cars. Efficient system and components' optimization requires the use of mathematical models of diesel filter performance. The three-dimensional model for the regeneration of the diesel particulate filter presented in this paper has been developed as an engineering tool for the detailed design optimization of SiC diesel filters of modular structure. The 3-D modeling is achieved by interfacing an existing 1-D model to commercial finite element method software for the computation of the 3-D temperature field within the whole filter assembly, including the adhesive of the filter blocks, the insulation mat, and the metal canning. The 3-D model is applied to real-world component optimization studies of diesel filter systems. [DOI: 10.1115/1.2130732]

1 Introduction

Cellular ceramic diesel particulate filters with catalytic assistance have demonstrated the capacity to attain the extremely stringent particulate matter emissions of EPA Tier II and EURO 4–EURO 5 standards, both for passenger cars and for heavy duty vehicles. Diesel filter applications date back to the 1980s. Nevertheless, their series application was not realized until recently, because of problems regarding filter reliability and durability, which were associated with the regeneration of the filter. During regeneration, the accumulated particulate matter (also referred to as soot) is oxidized to CO_2 and CO. The soot oxidation reactions are followed by significant heat release, which may endanger the integrity of the filter under certain operating conditions. A typical scenario of filter regeneration that may lead to filter damage is illustrated in Fig. 1. A sudden deceleration of the vehicle occurring shortly after the onset of regeneration in a heavily loaded SiC diesel filter results in high oxygen concentration and low mass flow rates in the filter inlet gas stream. This leads to uncontrolled filter regeneration and excessive heat release that causes temperatures at the center of the filter to rise to about 900 °C (even higher, locally). In addition to the advantageous higher thermal conductivity, compared to Cordierite, SiC has a higher melting point, which adds to its durability. On the other hand, SiC has a higher thermal expansion coefficient, thus it is more prone to thermal stresses due to temperature gradients within the filter. SiC filters are usually made up of smaller filter blocks, which are glued together with adhesive cement. The adhesive normally has a lower thermal conductivity than SiC, a fact that inhibits conduction in the radial direction. SiC filters are increasingly employed in passenger cars, coupled with catalytic regeneration techniques (catalytic fuel additives or washcoated filters) [1]. The capabilities of modern diesel engine management and injection systems (post injection) are employed to increase exhaust gas temperature to achieve the onset of regeneration. Thus, the diesel filter forms the core device of an integrated system. The need for engineering modeling tools has thus emerged [2,3] to assist diesel filter and overall exhaust system design optimization.

Out of a variety of modeling approaches available, the final selection depends on the nature of the design problem posed to the engineer. Most prominent factors, though, should be the following:

- level of optimization detail
- data availability
- computational cost and cost of use
- model consistency and reliability

The level of optimization detail that the engineer pursues greatly influences the complexity of the model itself. For relatively simple or preliminary design and optimization, simpler models are preferred to more complicated ones. Difficulty in model development and validation grows rapidly with increase of model complexity, a fact that might create uncertainty about the reliability of highly complex models results. Furthermore, sophisticated models with many degrees of freedom also require detailed input data such as exhaust gas velocity profiles or complex reaction kinetics. More often than not, such data are not routinely available. More sophisticated models also require higher expertise from the user, higher computational cost, and are more labor intensive in their results post-processing and conclusions drawing.

Having the above in mind, we present in this paper a 3-D model for the regeneration of the diesel particulate filter, to be employed as a tool for the detailed optimization of the structure and geometry of SiC diesel filters of modular structure. The model is incorporated in an integrated methodology for the computer-aided engineering of the total exhaust treatment system, which also comprises pre- and post-processing procedures, code validation with design and quality assurance of supporting experiments and tests, model tuning procedure, and application in the design and optimization of target exhaust treatment systems.

The 3-D model described in this work is based on the CATWALL 1-D model, which has been extensively validated in the past [4–6]. It may be used either via the MATLAB/Simulink environment, or as a stand-alone tool, supported by a set of Excel macros for pre- and post-processing of data. The extension of the model to three dimensions is obtained by linking it to commercial

[1]To whom correspondence should be addressed.

Contributed by the Internal Combustion Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received February 27, 2004; final manuscript received August 15, 2005. Review conducted by D. Assanis.

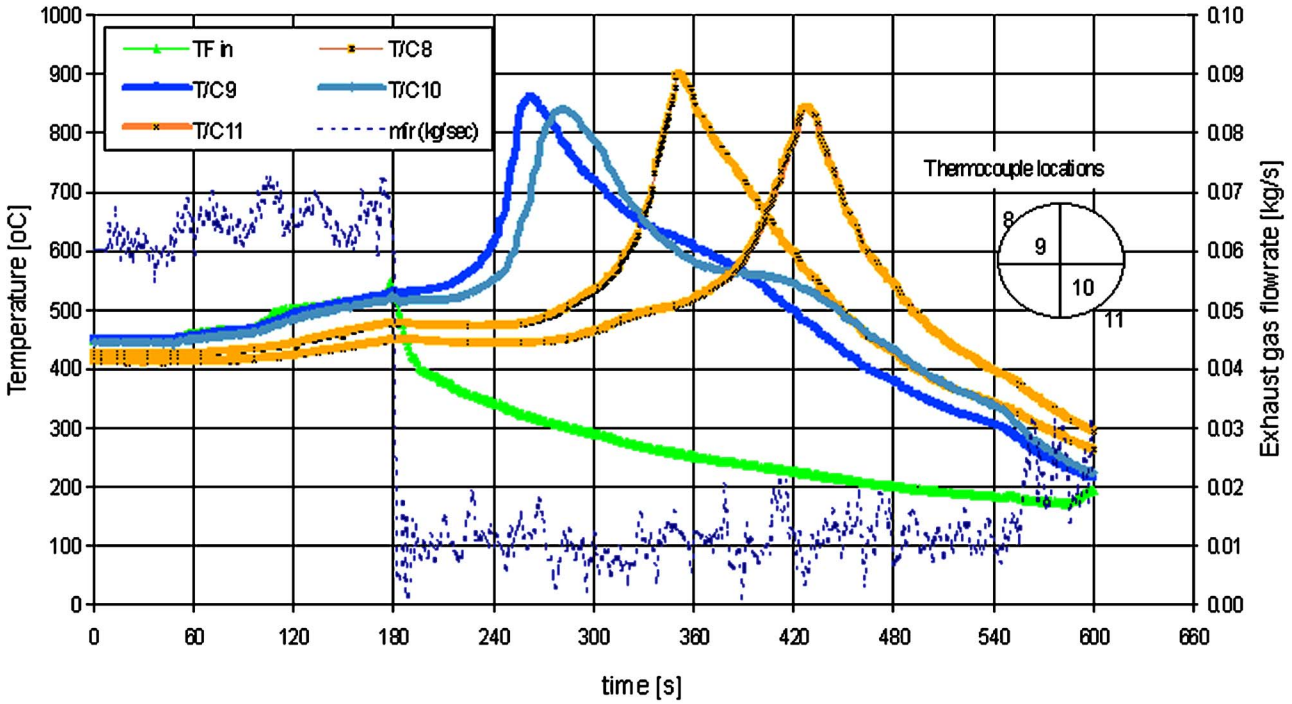


Fig. 1 Deceleration test with fuel additive. Initial SiC filter soot loading: (est.) 29 g. Initial engine operation at 2500 rpm—engine load 80 Nm, filter inlet temperature 500°C. Step decrease at t=180 s to 800 rpm—load 20 Nm.

finite element method (FEM) software for the computation of the 3-D temperature field within the whole filter assembly, including the adhesive of the filter blocks, the insulation mat, and the metal canning. The feasibility of this approach had been investigated previously [7]. The model incorporates a module for the computation of nonuniform flow distribution caused by nonuniform soot deposit of the different channels.

2 Modeling Catalytic Regeneration

2.1 General. Chemical phenomena in diesel filter regeneration include (a) the combustion of the soot layer because of thermal and catalytic reactions and (b) the adsorption, desorption, and combustion of hydrocarbons contained in the soot particles (volatile organic fraction). The primary physical phenomena of the filter operation are (a) the heat transfer in the filter and (b) the distribution of the flow field, which is affected by the distribution of the soot layer within the filter.

The first attempt to provide a mathematical description of the regeneration process of the soot layer was the pioneering work of Bissett [8]. Their first model was a zero-dimensional regeneration model. It was concerned with thermal soot combustion in one point of a filter channel, and flow and soot distribution within the channels of the filter were averaged. Despite its apparent simplicity, this work has been remarkable in that it essentially defined a robust mathematical framework for the description of the physical and chemical phenomena in the soot layer and filter wall, and some significant simplifying assumptions were established. The initial model was later extended to 1-D, also accounting for the axial heat conduction and the soot layer and flow distribution along the channel [9].

This work has been adopted and extended in [10], allowing the study and experimental validation of high space velocity thermal regeneration events, with a systematic methodology of filter loading assessment by energy balances. It was further extended to cover catalytic regeneration, in zero dimensions [11], and also in 1-D [12], while further refinements and improvements of the regeneration model were validated by extensive experimental data [4,5,13].

As regards the basic assumptions, and if we take out the chemical reactions submodel, the majority of the models in use today could be considered as extensions of [8,9]. Minor improvements have been presented, as the raising of the uniform soot and wall temperature across the wall [14]. A number of studies have been done focusing on the extension to 2-D [15–17]. However, the reaction scheme remained very simplified in all cases (thermal carbon oxidation, in spite of the fact that all diesel filter systems in use today are based on the use of catalysts) and extensive validation of a 2-D model was only presented in [16].

Since the regeneration process in the filter is essentially three-dimensional, especially at low flow rates, and significant flow and soot maldistribution are reported [18], which may also result in partial burning of the soot [19], a further extension of the above model has been attempted for the 3-D modeling of the heat conduction within the filter by linking the channel model [12] with commercial finite element software. In this work, we apply this concept to build a 3-D model for the regeneration process in the filter by interfacing the 1-D model with the ANSYS [20] finite element software package. The 1-D model features an inclusive reaction scheme for the thermal and catalytic regeneration of the soot. Analytical solutions of the pertinent balance equations are obtained, including an updated catalytic regeneration reactions scheme and a more detailed view of the soot layer geometry. Finally, the model is enhanced with a module to account for flow maldistribution at the filter inlet because of nonuniform soot loading of the filter, which is usually a result of incomplete filter regeneration.

It should be noted that the reaction rates that have been used in all filter regeneration models are essentially of the Arrhenius-type, i.e., the rates are analogous to the factor $k=A\exp(-E/R_gT)$. This implies that the activation energy E and the preexponential factor A must be determined for each reaction involved in the reaction scheme.

2.2 Geometry of the Channels and the Soot Layer. The filter consists of a ceramic grid of parallel channels. Adjacent

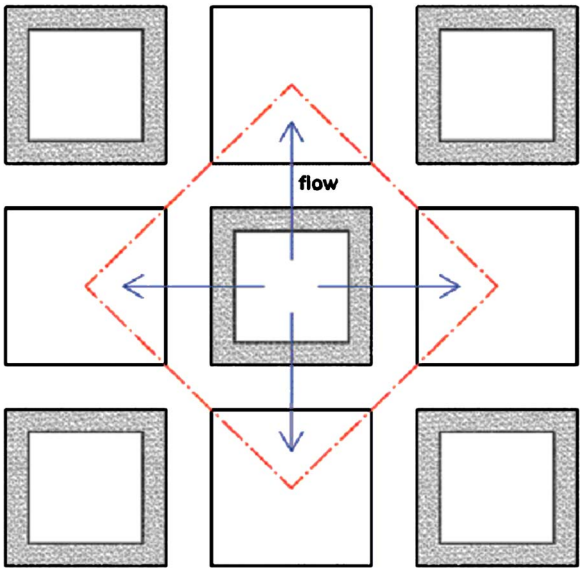


Fig. 2 Computational cell for a wall flow diesel particulate filter, consisting of one inlet channel and four adjacent outlet channels

channels are alternatively plugged at each end in order to force the diesel aerosol through the porous walls. Thus, diesel particulate matter is deposited on the four sides of each inlet channel. Because of the repeating geometry of the filter, one may restrict the mathematical treatment to a unit computational cell as seen in Fig. 2. This fundamental volume consists of one inlet channel and four quarters of the four adjacent outlet channels. The boundaries of the volume are assumed adiabatic and no flow occurs through them to the rest of the outlet channel. From the standpoint of modeling, the whole filter is constructed by repeating this fundamental volume.

To formulate the balance equations pertinent for regeneration modeling, we consider a small part of this fundamental volume, of length Δz. At this part of the channel, a layer of soot of mass m and thickness w has already accumulated. The wall's thickness is w_s and the length of the channel edge is ℓ. Exhaust gas flow is perpendicular to the soot layer and wall. Exhaust gas species are consumed or produced as gas flows through the soot layer. In this work, we shall use the subscript p to denote the physical properties of the soot layer and s to denote the corresponding properties of the wall.

It is evident from Fig. 3 that the soot layer consists of four

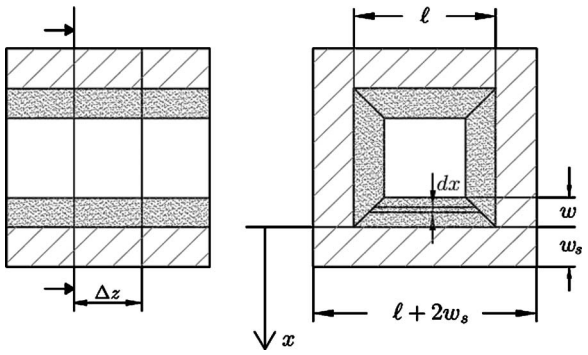


Fig. 3 Schematic representation of the cross section of a loaded wall flow channel assuming that the soot layer consists of four trapezoids

Table 1 Reactions and rate expressions of the regeneration model

	Reaction	Rate expression
1	$C+O_2 \rightarrow CO_2$	$r_1=k_1 y$
2	$C+0.5O_2 \rightarrow CO$	$r_2=k_2 y$
3	$C+4CeO_2 \rightarrow 2Ce_2O_3+CO_2$	$r_3=k_3 \psi$
4	$C+2CeO_2 \rightarrow Ce_2O_3+CO$	$r_4=k_4 \psi$
5	$Ce_2O_3+0.5O_2 \rightarrow 2CeO_2$	$r_5=k_5(1-\psi)y$
	where $k_i=A_i e^{-E/R_g T}$, $i=1\ldots5$	

trapezoids. Assuming that some soot mass m_p of density ρ_p is deposited on a square filter channel with inner side length ℓ, the deposit thickness can be calculated as

$$w_p = \frac{1}{2}\left(\ell - \sqrt{\ell^2 - \frac{m_p}{\rho_p \Delta z}}\right) \tag{1}$$

This relation is employed in the computation of the evolution of deposit thickness in the different channels and different nodes of each channel. As a consequence of the trapezoid distribution of the soot layer, the area perpendicular to the flow $A(x)$ (i.e., the area that the flow "faces" at each x) is also changing along x:

$$A(x) = \begin{cases} 4(2x+\ell)\Delta z, & -w_p \leqslant x \leqslant 0 \\ 4\ell\Delta z, & 0 \leqslant x \leqslant w_s \end{cases} \tag{2}$$

The volume of the soot layer V_p may be readily calculated from the above relationship

$$V_p = 4\Delta z(\ell - w_p)w_p \tag{3a}$$

while the volume of the substrate is

$$V_s = 4\Delta z(\ell + 0.5w_s)w_s \tag{3b}$$

Notice that it is common in the literature to consider only one of the four sides of the channel instead of the computational cell of Fig. 2. Such a simplification would also yield slightly simpler expressions for w_p, $A(x)$, and V_p, since the trapezoid shape of the soot layer would not be taken into account. In fact, this approach was adopted in earlier works, e.g. [8,9]. However, the total mass of the ceramic substrate and also the thickness of the soot layer w as a function of soot load would be incorrectly calculated, introducing some error in the thermal inertia and pressure drop calculation, respectively. Admittedly, these errors would not change the behavior of the model dramatically. Nevertheless, the extension described in Fig. 2 and Eqs. (1)–(3) was adopted because it does not complicate the model significantly.

2.3 Reaction Kinetics. The soot that is deposited in the filter walls consists mainly of carbon and adsorbed hydrocarbons (VOF). When a fuel additive is used to assist regeneration, it takes part in the combustion process in the engine's cylinder and the formation of particulate matter. Therefore, soot particles deposited in the trap contain bonded metal oxide particles, which react with the soot carbon and VOF at lower temperatures compared to thermal regeneration.

The kinetic model employed herein accounts for thermal soot oxidation by oxygen and for catalytic soot oxidation by a Ce-based fuel additive, but it does not yet include the effect for the VOF. The reactions along with their respective rate expressions are given in Table 1. The reaction rates for all reactions are assumed to depend on temperature according to a first-order Arrhenius-type expression.

For thermal oxidation modeling, two reactions are employed, accounting for the oxidation of carbon by exhaust gas oxygen to CO_2 and CO, respectively. Both reactions are assumed to be proportional to the oxygen concentration y in the exhaust gas. The latter has a significant distribution along the soot layer thickness,

since it is gradually depleted therein. For the catalytic oxidation of carbon, a model was first formulated in [11]. The implemented mechanism is only valid when the catalytic oxide has two oxidation states, which is true for ceria. The first step of the mechanism is oxidation of Ce_2O_3 by O_2 to produce CeO_2. The second step involves reaction of CeO_2 with soot carbon. Ceria returns from its four- to its three-valent state, while carbon monoxide is produced. The presence of the additive does not prevent direct oxidation of carbon by oxygen, if the temperature is sufficient for this reaction.

For the oxidation of Ce_2O_3 by the exhaust gas oxygen, the reaction rate is assumed proportional to the oxygen concentration y in the exhaust gas, similarly to the thermal oxidation of C. The reaction rates of C oxidation by CeO_2 are assumed proportional to the quantity ψ which represents the percentage of Ce being in its higher oxidation state (CeO_2):

$$\psi = \frac{\text{mol } CeO_2}{\text{mol Ce}} = \frac{\text{mol } CeO_2}{2 \times \text{mol } Ce_2O_3 + \text{mol } CeO_2} \tag{4}$$

An additional term ξ will also be needed in the next section, though it does not appear in the reaction rates. It is defined as the concentration of catalyst in the soot layer and is a function of metal additive concentration in the fuel. It may be expressed as

$$\xi = \frac{\text{mol Ce}}{\text{mol soot}} = \frac{2 \times \text{mol } Ce_2O_3 + \text{mol } CeO_2}{\text{mol C}} \tag{5}$$

As it is evident from Table 1, the quantity ψ plays a similar role for catalytic soot oxidation with the O_2 concentration y for thermal soot oxidation. It should be noted, though, that both ψ and ξ are viewed as averaged quantities along the soot layer thickness, contrary to y which is assumed to have a distribution. Averaging ξ is expected because Ce is uniformly distributed in the soot layer, since it is bonded with the soot particles during the fuel combustion process, contrary to O_2 which flows through the soot layer where it is gradually depleted. The rationale behind the averaging of ψ stems from the need to suppress the model's dimensions, in order to keep the feasibility of analytical integration of the balance equations along the x dimension.

2.4 Mass and Heat Balances in the Soot Layer and Ceramic Wall (Regeneration Submodel). As gas flows through the soot layer, the oxygen is gradually depleted because of reaction with carbon and ceria. An oxygen concentration profile is thus established through the soot layer, which is expressed by an oxygen mass balance:

$$\frac{\dot{m}_g}{A(x)}\frac{dy(x)}{dx} = -M_g(k_1 + 0.5k_2 + 0.5k_5(1-\psi))y(x) \tag{6}$$

The above mass balance may be solved to get the oxygen concentration profile $y(x)$, for the initial condition $y(x=-w_p)=y_{in}$, provided that the temperature of the soot layer is considered constant in a direction perpendicular to the soot layer (which is justified in the discussion about temperature boundary conditions later in this section). This is a fundamental assumption that is supported by the original work of Bissett [8] (see also the boundary conditions below). The solution yields the total consumption of oxygen $\Delta y = y(x=-w_p) - y(x=0)$ through the soot layer:

$$\Delta y = y_{in}\left[1 - \exp\left(-\frac{M_g(k_1 + 0.5k_2 + 0.5k_5(1-\psi))}{\dot{m}_g}V_p\right)\right] \tag{7}$$

The total consumption of oxygen may be used in order to calculate the rate of soot consumption because of both thermal oxidation with oxygen and catalytic oxidation with ceria. Specifically, the mass balance for the soot mass consumption is

$$\frac{\rho_p}{M_p}\frac{dV_p}{dt} = -\int_{-w_p}^{0}[(k_1 + 0.5k_2)y(x) + (k_3 + k_4)\psi]A(x)dx \tag{8}$$

or, splitting the integral,

$$\frac{\rho_p}{M_p}\frac{dV_p}{dt} = -(k_1 + 0.5k_2)\int_{-w_p}^{0} y(x)A(x)dx - (k_3 + k_4)\psi\int_{-w_p}^{0} A(x)dx \tag{9}$$

The first integral contained on the right-hand side of the above equation may be calculated from the oxygen mass balance (6):

$$\int_{-w_p}^{0} y(x)A(x)dx = \frac{\dot{m}_g}{M_g(k_1 + 0.5k_2 + 0.5k_5(1-\psi))}\Delta y \tag{10}$$

The second integral of the soot mass balance is the volume of the soot layer V_p. Substituting the above, we get the following rate of soot mass consumption:

$$\rho_p\frac{dV_p}{dt} = -\frac{M_p}{M_g}\frac{k_1 + 0.5k_2}{k_1 + 0.5k_2 + 0.5k_5(1-\psi)}\dot{m}_g\Delta y - M_p(k_3 + k_4)\psi V_p \tag{11}$$

A mass balance must also be formulated for ψ, which expresses the continuous transition of the catalytic additive between CeO_2 and Ce_2O_3 states. According to the CeO_2 mass balance, the rate of change depends on the additive oxidation and reduction reactions, i.e.,

$$\left(\int_{-w_p}^{0}\frac{\rho_p A(x)dx}{M_p}\right)\xi\frac{d\psi}{dt} = \int_{-w_p}^{0}[-(4k_3 + 2k_4)\psi + 2k_5 y(1-\psi)]A(x)dx \tag{12}$$

The above equation may be simplified and, with the substitution of (10), yields the total rate of change for ψ:

$$\xi\frac{d\psi}{dt} = -\frac{M_p}{\rho_p}(4k_3 + 2k_4)\psi + \frac{2k_5(1-\psi)}{k_1 + 0.5k_2 + k_5(1-\psi)}\frac{\dot{m}_g}{\rho_p V_p}\frac{M_p}{M_g}\Delta y \tag{13}$$

The two mass balance equations for soot and ceria should be solved for the following initial conditions: For $t=t_0$ we have $V_p = V_{p,0}$ and $\psi=\psi_0$.

To complete the model, a heat balance equation must be formulated. This expresses the accumulation of heat in the soot layer and the ceramic wall because of (a) heat generation by the chemical reactions, (b) heat conduction perpendicular to the soot layer and wall, and (c) heat exchange between the gaseous and the solid phase. Denoting the temperature of the solid phase (particle layer or substrate) as T, and the temperature of the exhaust gas as T_g, we have the following balance equations for the solid phase:

$$\rho_j c_{p,j} A(x)\frac{\partial T}{\partial t} = -\dot{q}(x)A(x) + \frac{\partial}{\partial x}\left(A(x)\lambda_j\frac{\partial T}{\partial x}\right) - \dot{m}_g c_{p,g}\frac{\partial T_g}{\partial x}, \quad j = p,s \tag{14}$$

where the heat flux $\dot{q}(x)$ [J/(m^{3}s)] is defined as the sum of the products of the reaction rates r_n [mol/(m^{3}s)] as defined in Table 1 and the reaction enthalpies ΔH_n (J/mol):

$$\dot{q}(x) = \begin{cases} \sum_{n=1}^{5} r_n \Delta H_n & \text{if } -w_p \leq x \leq 0 \\ 0 & \text{if } 0 \leq x \leq w_s \end{cases}$$

(No reactions occur as the gas flows through the ceramic wall, thus $\dot{q}(x)$ is zero in that region.)

Except for the solid phase temperature T, the above heat balance also contains the gas temperature T_g, which, in principle, has a separate profile $T_g(x)$ along the solid phase. Instead of solving for the two coupled temperature profiles of $T(x)$ and $T_g(x)$, Bissett and Shadmann proved that the gas temperature T_g adjusts to the solid temperature in a length scale several orders of magnitude smaller than the soot layer thickness. That is, $T_g = T$ everywhere

except at $x=-w_p$.

At $x=-w_p$, and at $x=+w_s$, according to Bissett and Shadmann, we have the following boundary conditions:

$$\text{At } x=-w_p \quad T_g=T_{g,\text{in}}$$

$$A(x=-w_p)\lambda_p\frac{\partial T}{\partial x}=\dot{m}_g c_{p,g}(T-T_g) \tag{15}$$

$$\text{At } x=+w_s \quad \frac{\partial T}{\partial x}=0$$

Using the above boundary condition at $x=-w_p$, we may integrate the second term of the right-hand side of the heat balance to get

$$\int_{-w_p}^{+w_s}\frac{\partial}{\partial x}\left(A(x)\lambda_j\frac{\partial T}{\partial x}\right)=\dot{m}_g c_{p,g}(T-T_{g,\text{in}})$$

Thus, the boundary conditions at $x=-w_p$ imply that all the heat content of the gas is immediately exchanged with the solid phase.

Furthermore, in the same work, Bissett and Shadmann also show mathematically that the temperature of the solid phase may be considered uniform along x. This is a critical result, and we have already used it in the integrations of the mass balances, since we have silently considered the reaction rate coefficients k (which depend strongly on temperature) as constant along x. This result also further simplifies the heat balance, since we readily find that the third term of the right-hand side is zero, since $dT/dx=0$.

Under these conditions, the heat balance (14) is integrated to yield the rate dQ/dt that heat is released (a) by chemical reactions and (b) by convection because of flow normal to the soot layer and wall.

$$\begin{aligned}\frac{dQ}{dt}&=(\rho_p c_{p,p}V_p+\rho_s c_{p,s}V_s)\\&=-\frac{[\Delta H_1 k_1+\Delta H_2 k_2+0.5\Delta H_5 k_5(1-\psi)]}{k_1+0.5k_2+0.5k_5(1-\psi)}\frac{\dot{m}_g}{M_g}\Delta y\\&\quad-(\Delta H_3 k_3+\Delta H_4 k_4)\psi V_p+\dot{m}_g c_{p,g}(T-T_g^{\text{in}})\end{aligned} \tag{16}$$

2.5 Pressure Drop Submodel. The total pressure drop across the particulate layer and the ceramic wall can be expressed as the sum of the porous ceramic substrate pressure drop and the soot layer pressure drop:

$$\Delta p=\Delta p_s+\Delta p_p \tag{17}$$

It is generally accepted [21] that pressure drop across the porous walls of the ceramic filters can be approximated by Darcy's law. According to this law, the total pressure drop due to flow through the ceramic wall and the soot layer can be approximated by the following simplified relation:

$$\Delta p=p_{\text{in}}-p_{\text{out}}=\frac{\mu u_w w_s}{k_s}+\frac{\mu u_w w_p}{k_p} \tag{18}$$

By using the concept of effective particulate layer thickness [22], the latter formula can be rewritten as follows:

$$\Delta p=p_{\text{in}}-p_{\text{out}}=\frac{\mu u_w w_s}{k_s}+\frac{\mu u_w m_p}{A_f(\rho k)_p} \tag{19}$$

where A_f is the mean filtration area. This formula allows, in principle, the backward approximate calculation of collected soot mass as function of measured filter backpressure at a certain engine and filter loading operation point, once an approximate value for the product $(\rho k)_p$, that is, (soot layer density) times (soot layer permeability) is known for the specific engine–filter–operation point combination.

2.6 1-D Modeling of the Flow Distribution and Pressure Losses in the Channel. The above pressure drop and regeneration model is a zero-dimensional approach to filter modeling, in the sense that it is valid for each point z along the filter axis, but does not provide any information regarding axial profiles of temperature, pressure drop, and mass distribution. This is accomplished below, employing the following balance equations: (i) the conservation of the exhaust gas mass flow, (ii) the conservation of the axial component of momentum of exhaust gas, and (iii) the conservation of the energy of the exhaust gas. Below, we shall use the subscripts "in" for the flow parallel to the inlet channel, "out" for the flow parallel to the outlet channel, and w for flow through the soot layer and the wall.

The exhaust gas mass balances state that, at each axial point of the inlet channel, the change of the exhaust gas flow is equal to the flow leaving the inlet channel via the porous wall. An exactly analogous balance is defined for the outlet channel.

$$\frac{\partial}{\partial z}(\rho_{\text{in}}u_{\text{in}})=-(4/\ell)\rho_w u_w, \quad \frac{\partial}{\partial z}(\rho_{\text{out}}u_{\text{out}})=+(4/\ell)\rho_w u_w \tag{20}$$

The balance equation for the axial component of the momentum of the gas states that the axial pressure variation is due to convective transport of the z component of momentum and the viscous drag forces exerted in the gas flow. The balance equation is the same for both channels:

$$\frac{\partial p}{\partial z}+\frac{\partial}{\partial z}(\rho_i u_i^2)=-\frac{a\mu u_i}{d_h^2}, \quad i=\text{in, out} \tag{21}$$

where d_h is the hydraulic diameter of the channel, which, for square channels, is equal to the channel's edge: $d_h=\ell$. The viscous drag is caused because there is a gradient of the axial velocity of the gas, as axial velocity reduces to zero at the walls. According to Bissett [9], these drag forces should be near those observed for an impermeable wall, because only a small fraction of the gas penetrates the wall at each axial point.

The above balances, together with the Darcy correlation (19), are used to calculate the distribution of the gas velocities in the inlet and outlet channels and the velocities normal to the soot layer and wall.

Once the velocities' distribution has been calculated, the model computes the energy exchanged between the gaseous and solid phase porous wall. The energy balances for the inlet and outlet channels are

$$\begin{aligned}c_{p,g}\frac{\partial}{\partial z}(\rho_{\text{in}}u_{\text{in}}T_{\text{in}})&=-(4/\ell)c_{p,g}\rho_w u_w T_{\text{in}}+h_{\text{in}}(4/\ell)(T-T_{\text{in}})\\c_{p,g}\frac{\partial}{\partial z}(\rho_{\text{out}}u_{\text{out}}T_{\text{out}})&=-(4/\ell)c_{p,g}\rho_w u_w T_w+h_{\text{out}}(4/\ell)(T-T_{\text{out}})\end{aligned} \tag{22}$$

The balance for the outlet channel differs in the second term that concerns energy transport via the flow through the porous wall. Apart from the sign of the term, the temperature is different, because the gas leaving the porous wall has the temperature of the wall.

Finally, the axial temperature distribution along the wall is calculated. The heat balance implements 1D heat conduction and source terms are because of gas-solid heat convection and reaction exothermy:

$$\begin{aligned}(\rho_p c_{p,p}+\rho_s c_{p,s})\frac{\partial T}{\partial t}&=\lambda\frac{\partial^2 T}{\partial z^2}+h_{\text{in}}(4/\ell)(T-T_{\text{in}})+h_{\text{out}}(4/\ell)(T-T_{\text{out}})\\&\quad+\frac{1}{V_p+V_s}\frac{dQ}{dt}\end{aligned} \tag{23}$$

In the above equation, the heat source term dQ/dt expresses heat exchange due to perpendicular gas flow and reaction exothermy; it is given by Eq. (16). Since this term refers to an axial element of length Δz, as seen in Fig. 3, it is divided by the sum of the

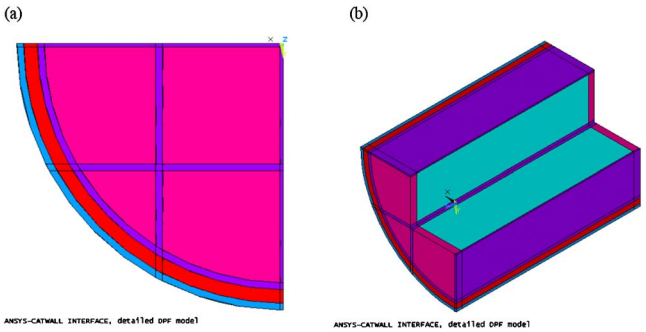


Fig. 4 Solid model of the filter. Different materials are designated by different shading.

elementary volumes of the substrate and the soot layer V_p+V_s, defined in Eqs. (3*a*) and (3*b*), respectively.

2.7 3-D Modeling of Heat Conduction and Inlet Flow Distribution. One-dimensional channel models assume that all channels of the filter operate under identical conditions. The three-dimensional modeling raises this assumption by accounting for three main phenomena that are neglected by 1-D modeling, namely, heat conduction between channels, heat losses to the ambient air, and nonuniformity of mass flow rate at the channels inlet. Also, it takes into account the nonsymmetric structure of the specific filter, which is made up of smaller SiC filter blocks, arranged in a Cartesian grid and glued together using a special adhesive cement layer (Fig. 4). The cement layer has different thermophysical properties than the SiC (see Table 2).

For the calculation of the 3D temperature field in the filter, the heat diffusion equation in three dimensions should be solved, which is [23]

$$\rho_s c_{p,s} \frac{\partial T}{\partial t} = \nabla \cdot (\lambda \nabla T) + \dot{q}_{\text{total}} \tag{24}$$

The source term $\dot{q}_{\text{total}}$ is the sum of all heat sources in the model. It is given by the sum of heat exchanged due to chemical reactions and convection owing to the flow *through* the soot layer and the wall (dQ/dt of Eq. (16)), and the heat convected between the solid phase and the gas flow in the inlet and outlet channels that is *parallel* to the wall:

$$\dot{q}_{\text{total}} = h_{\text{in}}(4/\ell)(T - T_{\text{in}}) + h_{\text{out}}(4/\ell)(T - T_{\text{out}}) + \frac{1}{V_p + V_s}\frac{dQ}{dt} \tag{25}$$

The solution of the three-dimensional heat balance (24) is accomplished by the finite element code, which calculates the three-dimensional temperature field in the filter relying on the 1-D model to provide the heat source terms involved in the balance. This requires the interfacing of the FEM code with the 1-D core model. The interfacing is possible using the user programmable feature (UPF) functionality of the FEM software suite that allows the user to write custom functions for properties or source terms of the elements. Such user subroutines are customarily written in FORTRAN, which is also the language of the CATWALL code implementation [5]. The calculation of the nonuniform temperature field in the filter allows the model to account for the so-called partial regenerations, which may leave significant portions of the deposited soot unburned, because of unfavorable conditions for regeneration propagation. Partial regenerations affect the distribution of the flow field at the inlet of the filter, since the flow is then channeled through the low flow resistance areas of the filter. A submodel has been included to take into account this effect. According to this submodel, the filter channels are viewed as a set of flow resistances in parallel. The objective of the submodel is to distribute the total mass flow rate at the inlet of the filter so that the pressure drop induced from each channel is the same. In fact, not every channel of the filter is treated separately, because of the excessive computational power required. Rather, the model separates sectors of channels that are treated identically. These sectors correspond to the meshing of the filter that is used from the FEM code for the temperature field calculation.

For the determination of the pressure drop of each channel sector, Eqs. (19) and (21) are used. Because of the quadratic term of Eq. (21) and the dependence of physical properties on the pressure and temperature of the gas at each point, the pressure drop Δp_i for each sector of channels is a nonlinear function of the mass flow rate $\dot{m}_i$ at the inlet of the sector: $\Delta p_i = f(\dot{m}_i)$. Moreover, there is not a closed expression for the function f, and, given $\dot{m}_i$, the pressure drop Δp_i may only be calculated by a trial-and-error procedure, following the usual procedure for pipe networks [24]. Nevertheless, the function f is *nearly* linear, because the pressure losses are mainly induced by the gas flow through the soot layer and the wall (Eq. (19)). Therefore, the total pressure losses for each channel may be approximately computed by a relation of the form $\Delta p_i = R_i \dot{m}_i$, where R_i is the flow resistance of the sector.

This fact has been exploited in order to develop an iterative procedure to compute the flow distribution at the filter inlet. The procedure features the following steps [6]:

Step 1. Provide an initial guess for the flow rate $\dot{m}_i$ of each sector. Assuming that the cross-sectional (frontal) area of each sector is A_i, we assume that $\dot{m}_i = \dot{m} A_i / A$ (uniform flow distribution).

Step 2. Calculate the induced pressure drop for each sector $\Delta p_i = f(\dot{m}_i)$.

Step 3. Calculate the approximate linear resistance R_i for each sector: $R_i = \Delta p_i / \dot{m}_i$.

Table 2 Thermophysical properties of SiC, adhesive, insulation mat and canning. Units: λ(Wm^{-1} K^{-1}), c_p(J kg^{-1} K^{-1}), ρ(kg m^{-3}). Void fractions for the calculation of bulk values: ε=0.375, ε_{plug}=0.688.

	SiC				Adhesive		Insulation		Canning	
Temperature (°C)	Real	λ Bulk (main)	Bulk (plug)	c_p	λ	c_p	λ	c_p	λ	c_p
25	57.1	21.4	39.3	686	0.187	740	0.04	1190	15.2	466.5
250	35.4	13.3	24.4	1005	0.222	1012	0.05	1190	19.4	538.2
500	24.6	9.2	16.9	1123	0.241	1141	0.08	1190	22.4	588.1
7500	19.8	7.4	13.6	1182	0.242	1188	0.12	1190	24.5	623.8
1000	15.0	5.6	10.3	1219	0.302	1200	0.17	1190	26.2	651.7
1250	13.0	4.9	8.9	1230	0.350	1205	0.23	1190	27.5	670.0
	1800	ρ 675	1238		ρ 137		ρ 300		ρ 7900	

Step 4. Calculate the approximate linear resistance of the whole filter: $(1/\bar{R})=\Sigma_{i=1}^{N}(1/R_i)$.

Step 5. Calculate an approximate value for the pressure drop of the whole filter: $\Delta p=\bar{R}\dot{m}$.

Step 6. Calculate new values for the flow rates of each sector: $\dot{m}_i'=\Delta p/R_i$.

Steps 2–6 are repeated until a set of flow rates for each sector is obtained so that the calculated induced pressure drop $\Delta p_i=f(\dot{m}_i)$ for every sector is the same. The above iterative procedure converges quickly (usually after three to ten iterations) to the desired set of mass flow rates.

3 3-D Heat Transfer Calculation

The approach selected to solve the problem of the calculation of the 3-D temperature field of the filter was to interface CATWALL with the commercial FEM software, by developing (a) a model of the diesel particulate filter in the ANSYS environment and (b) a set of interface subroutines that handle the task of data interchange between the two codes [6,25].

3.1 ANSYS Solid and Finite Element Model of the Filter. The solid model is the description of the geometric shape of the filter. The volumes comprising the solid model are then assigned material properties. The solid model is then meshed to generate the finite element model that is actually used by the FEM solver. The front of the solid model of the filter is given is Fig. 4. Usually, a quarter of the cylindrical filter needs to be modeled only, because of the symmetry of its structure. The solid model comprises different volumes for the main part of the SiC filter blocks, the inlet and outlet regions of the blocks (where the plugs are), the adhesive of the filter blocks, the insulation mat, and the canning. Depending on the volume, different material properties are assigned (that is, heat capacity, thermal conductivity, and density).

Evidently, for simplicity and speed of the computation, the solid model of the filter is simplified, in the sense that the details of the filter structure, with its grid of channels and plugs, is replaced by continuous volumes. This approach necessitates the use of bulk quantities for the density and the heat conductivity of the SiC filter blocks, which are computed using the void fraction of the filter structure:

$$\rho_{\text{bulk}}=\varepsilon\rho_{\text{SiC}} \quad \lambda_{\text{bulk}}=\varepsilon\lambda_{\text{SiC}} \tag{26}$$

Because of the presence of the plugs, the void fraction $\varepsilon_{\text{plug}}$ of the inlet and outlet regions of the filter blocks is different from the void fraction of their main part. Therefore, separate volumes are defined for the plug regions of the solid model. These volumes are assigned the bulk density and heat conductivity of the plug regions that were computed using $\varepsilon_{\text{plug}}$ instead of ε. The (real and bulk) temperature-dependent material properties of the SiC, adhesive, insulation mat, and canning that were used in this work are given in Table 2. In the second view of Fig. 4, one brick of the SiC filter as well as its surrounding adhesive have been removed, to reveal the different volumes of the solid model that correspond to the main and the plug region of the filter.

ANSYS meshes the solid model to construct the finite element model of the filter, which consists of elements and nodes. The element type that was chosen for meshing was SOLID70, which is a thermal solid element with eight nodes, with a single degree of freedom, temperature, at each node. The meshed finite element model of the specific filter that has been used in this work is illustrated in Fig. 5. In conjunction with grid generation, two fundamental points have to be noted: First, there is no restriction—that is imposed from the 1-D regeneration model—regarding the cross-sectional shape of the elements. Each row of elements along the axial direction that corresponds to SiC (see Fig. 5) is referred to as a sector. The obvious implication is that the 1-D code should be invoked as many times as the number of sectors, namely the number of elements at a cross-section of the filter. Thus, the finer the grid, the more times the 1-D model will be invoked, with direct consequences to the computational cost of the model. Second, the grid designed for the FEM analysis must be orthogonal along the axial direction, with fixed element length, which is an integer multiple or submultiple of the discretization length Δz of the 1-D code. Essentially, the finite element model of the face is extruded along the axial direction in order to produce the total finite element model. The requirement for axial orthogonality is imposed by the structure of the filter itself, while the fixed element length is imposed by a limitation of the 1-D code, which operates on a fixed axial discretization interval Δz. The above two requirements must be met, so that a mapping of the 1-D code and the ANSYS grids is possible, so that heat source and temperature information may be exchanged between the two grids. The latter point brings into the foreground the issue of interface implementation, which is discussed below.

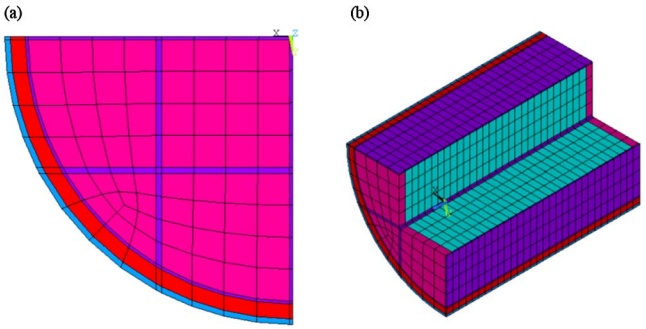


Fig. 5 Finite element model of the filter, produced after meshing the solid model. (*a*) Front view. (*b*) Side view, with the central filter block and adhesive removed.

3.2 Interface Implementation. Through the interface between the two codes, the element number, temperature, heat source terms, and time are communicated. Essentially, the FEM software gives the current time, element number, and element temperature as input to the main interface subroutine and expects to receive the element heat source term. The heat source terms are employed by the FEM software in order to compute the new temperature field in the filter. The task for the interface subroutines is to convert ANSYS input so that it is sensible for the 1-D code, subsequently call the 1-D code for the calculations, and finally reconvert 1-D code output and feed it back to the FEM software.

The concept of the interfacing is given schematically in Fig. 6. It is evident that the implementation of two basic modules was required, because of the noncompatible operation and grid design of the two codes. The restrictions of the grid design and the functionality of the two basic modules are discussed in brief below.

3.2.1 Grid Mapping Module. There are two coexisting (though incompatible) grids for the filter: The FEM grid consists of elements (and nodes for each element) while the 1-D code grid consists of sectors and nodes for each sector. Therefore, a mapping procedure must be implemented between the two grids to achieve exchange of information. Specifically, the procedure uses the internal numbering of elements of the FEM software. If the length for axial discretization is the same for both the FEM grid and the 1-D code grid, then each element number also corresponds to a single pair of CATWALL sector and node numbers. In this case, there is a one-to-one correspondence between the two grids, which also applies to the heat source and temperature data.

If the 1-D code grid is finer than the FEM grid (the usual case), then one element number corresponds to a set of 1-D code nodes. In this case, the sum of heat sources from the 1-D code is given to the FEM code as the element heat source term, and temperature data passed to the 1-D code is assumed equal for all nodes corre-

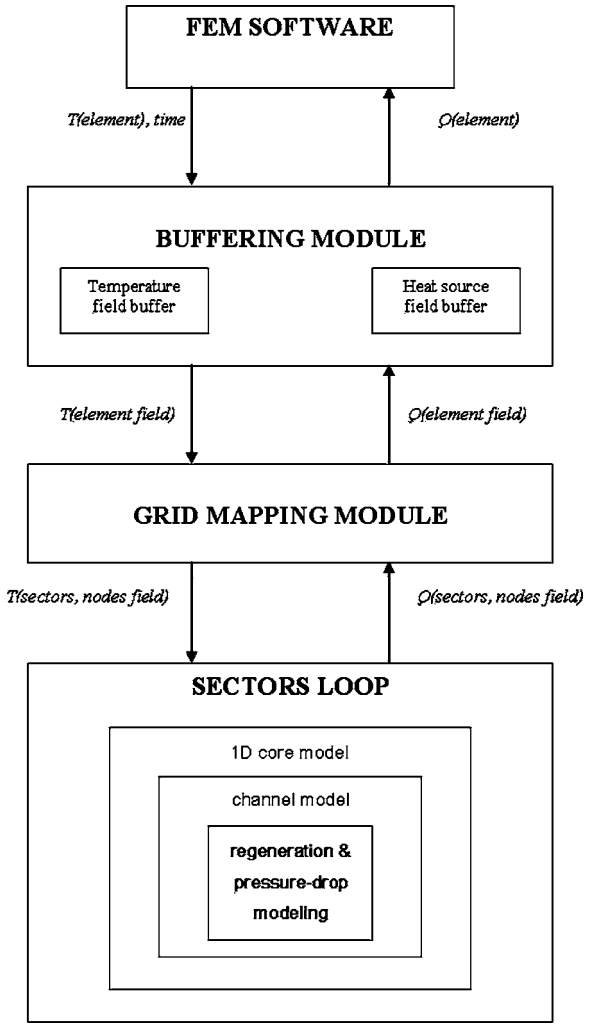


Fig. 6 ANSYS-CATWALL interfacing concept (flow diagram)

sponding to the specific element. The situation is inversed if the 1-D code grid is coarser than the FEM grid. Then the heat source calculated for a single 1-D code node is divided to the corresponding the FEM code elements, and the 1-D code node temperature is computed as the average of the corresponding element temperatures. These considerations should make clear why the discretization length of the FEM grid along the axial direction must be an integer (sub) multiple of the 1-D code discretization length.

3.2.2 Field Buffering Module. The FEM software may call the user subroutine more than once for each element. Each time that the user subroutine is called, the heat source term is requested and temperature is supplied back as a result for the next time step. Nevertheless, it is not efficient in terms of computational power to call the 1-D code for the sector to which the element belongs each time the FEM software requires it. The reason is that the 1-D code operates on the whole sector and not on a per-node basis. Therefore, a special module was implemented, which buffers the temperature and heat source fields. Whenever the FEM software takes a new time step, the 1-D code is called as many times as the numbers of the sectors, and the resulting heat source field is stored in a buffer. Subsequently, when a heat source term is requested, it is extracted from the buffer and passed to the FEM software. The temperature value that returns from the FEM software is stored to the temperature field buffer. At the next time step of the FEM software, the temperature field is passed from the buffer to the 1-D code and the next heat source field is computed.

Table 3 Engine technical data and diesel filter specifications

Engine type	HDI turbocharged engine
Cylinders	4, in-line
Displacement	1997 cm^3
Rated power/rpm	80 kW/4000 rpm
Rated torque/rpm	250 Nm/2000 rpm
Filter type	SiC 14/200 cpsi (cells per square inch)
Filter diameter × length	143.8×150 mm^2
Cell pitch	1.89 mm
Filter wall thickness	0.4 mm

3.3 Solution Procedure. The heat losses of the filter to the surrounding air are taken into account by imposing surface convective boundary conditions on the external nodes of the filter model. The value of the heat transfer coefficient between the filter's canning to the ambient air depends on the specific filter layout in exhaust system and vehicle speed, usually being in the range 20–100 $W/(m^2\,K)$. The highest values of the range typically include the effect of radiation exchange to the surroundings which becomes important when the can surface temperature exceeds 400°C.

An indicative running time of 2 s per 1 s of real filter operation is observed for meshing used in this work, which features $144 \times 50 = 7200$ elements in total (144 elements on each cross section) for the whole filter, which includes the 5.66×6 $in.^2$ SiC/adhesive filter structure, the insulation mat, and the canning.

Finally, a significant advantage of interfacing the 1-D filter regeneration code with FEM software to calculate the three-dimensional temperature field is the ability to carry out thermomechanical analysis. Calculation of thermal stresses in the filter is a very important feature of the model, since crack formation due to thermal stresses is a common reason of filter failure. Details of the stress analysis procedure are presented in [7]. Similar to the thermal analysis, the cell structure of the filter is modeled by a "homogenized" isotropic linear elastic material with Young's modulus $E=19$ GPa, Poisson's ratio $\nu=0.15$, thermal expansion coefficient $\alpha=4\times10^{-6}/°C$ and mass density $\rho=0.72$ g/cm^3. The adhesive is modeled also by an isotropic linear elastic material with $E=1.5$ GPa, $\nu=0.2$, $\alpha=2.7\times10^{-6}/°C$, and $\rho=1.26$ g/cm^3.

4 Exploitation of the 3-D Model Results in Detailed Filter Design

The assessment of the model's capabilities was based on the results of engine bench-loading–regeneration experiments performed on one of the laboratory engine test benches. The filter employed in these experiments was a 200 cpsi SiC filter fitted in the exhaust pipe of a 2-l displacement HDI turbocharged passenger car engine, about 600 mm after turbine exit. Engine and diesel filter specifications are summarized in Table 3. Exhaust temperatures were measured simultaneously, at the inlet and the exit of the filter, and inside the filter, along a filter diameter 15 mm deep from filter exit. Thermocouples are placed in the center of each filter block, along the diagonal of the filter. The exhaust emissions of CO, CO_2, NO_x, and HC were measured simultaneously before and after the filter. The O_2 concentration was calculated by the A/F ratio obtained from the UEGO sensor installed before the filter inlet. The experimental layout of the filter loading-regeneration experiments is presented in Fig. 7.

The kinetic parameters of the rate expressions that were used in this work were obtained by a validation study of the 1-D model [4]. In its turn, the latter study was supported by an extensive set of TGA analysis experiments of soot samples taken directly from the filter [26]. Values of the preexponential factor and the activation energy of each reaction are given in Table 4. In brief, the values of activation energies for catalytic oxidation correspond to dry soot oxidation for samples where the VOF content varies in

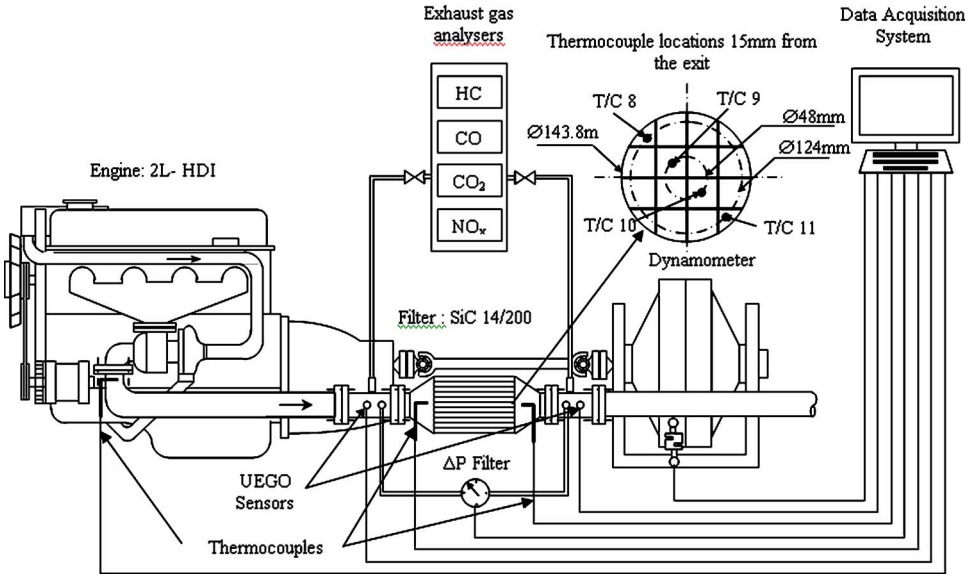


Fig. 7 Experimental layout. Engine and digitally controlled dynamometer installation are shown along with exhaust gas analyzers, main filter measurement lines, and data acquisition system.

the range between 2.5% and 8%. The preexponential factors were obtained by allowing a certain amount of tuning. Furthermore, the wall permeability, soot density, and soot permeability times density that were inserted in the 1-D model to simulate the pressure drop behavior of the validation experiments had the following values: wall permeability $K=2\times10^{-13}$ m^2, soot density $\rho_p=80$ kg/m^3, and soot permeability times density $(\rho K_p)=2.2\times10^{-11}$ kg/m.

In order to demonstrate the role of the 3-D filter regeneration model in the process of detailed filter design, a simulation of the case of Fig. 8 has been carried out. Obviously, compared to the limited information that may be obtained by the four thermocouples of the experiment, the code yields an excessive amount of information. Three-dimensional temperature fields, temperature-gradient fields for any direction, and inlet flow profiles are exited by the code and may be viewed for any filter section.

As a first step, eight snapshots of temperature fields for a section of the filter are presented in Fig. 9. The elements of the canning, the insulation mat, the surrounding adhesive, and one SiC block of the filter have been removed, to comparatively view regeneration characteristics in the center and the periphery of the filter. The onset of regeneration is observed near the filter exit, where the maximum temperatures are also observed. It becomes apparent that the evolution of regeneration in each one of the small periphery blocks is significantly different from that of a central block. Regeneration in the periphery of the filter is delayed, presumably due to the increased heat losses to the ambient air combined with the insulating effect of the adhesive surrounding the filter block, which inhibits the radial propagation of heat released in the central blocks. Thus, the specific design with periphery blocks is a bad design resulting in sustained incomplete regeneration.

Nevertheless, as illustrated in Fig. 8, the computed delay in the regeneration of the outlet blocks of the filter is significantly less than that measured delay of about 5 min. It is beyond the capacity of the model in its current state to accurately predict this delay. Possible causes for the discrepancy could be the following: First, the regeneration model lacks a mechanism to include the effect of the soot's VOF, which may be responsible for this type of regeneration, because it seems insensitive to the regeneration occurring in the inner filter block and seems more connected to the gradual elevation of the inlet gas temperature. Second, the initial filter temperature distribution and accumulated soot distribution may be significantly inhomogeneous, while the model assumes that the above distributions are uniform. Third, uncertainties regarding the thermophysical properties of the filter materials and especially the soot layer may negatively influence the model's predictive ability.

A second step in the detailed study of the filter regeneration behavior is the study of temperature gradients. In Fig. 10, the magnitude of the thermal gradient is plotted for three characteris-

Table 4 Kinetic parameters of the model's reaction rates

	Reaction	Pre-exponential factor (A) (mol/m^3, s)	Activation energy (E) (kJ/mol)
1	$C+O_2\rightarrow CO_2$	1E13	190
2	$C+0.5O_2\rightarrow CO$	5.5E10	150
3	$C+4CeO_2\rightarrow 2Ce_2O_3+CO_2$	4.5E11	120
4	$C+2CeO_2\rightarrow Ce_2O_3+CO$	4E8	80
5	$Ce_2O_3+0.5O_2\rightarrow 2CeO_2$	1E12	80

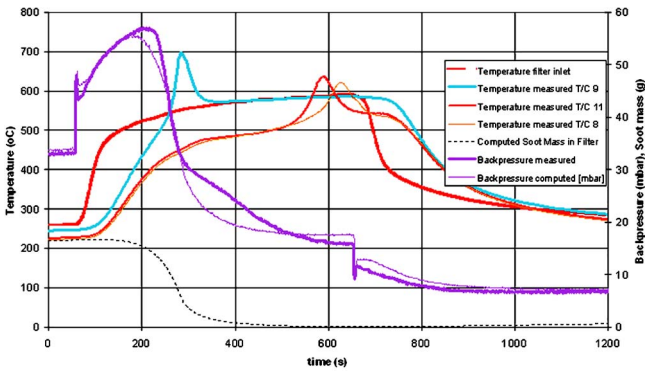


Fig. 8 Low mass flow rate regeneration case: Evolution of exhaust temperature at filter inlet, filter exit wall temperature at three characteristic points (thermocouples 8, 9, 11, see Fig. 7), computed evolution of soot mass in the filter, measured and calculated filter backpressure

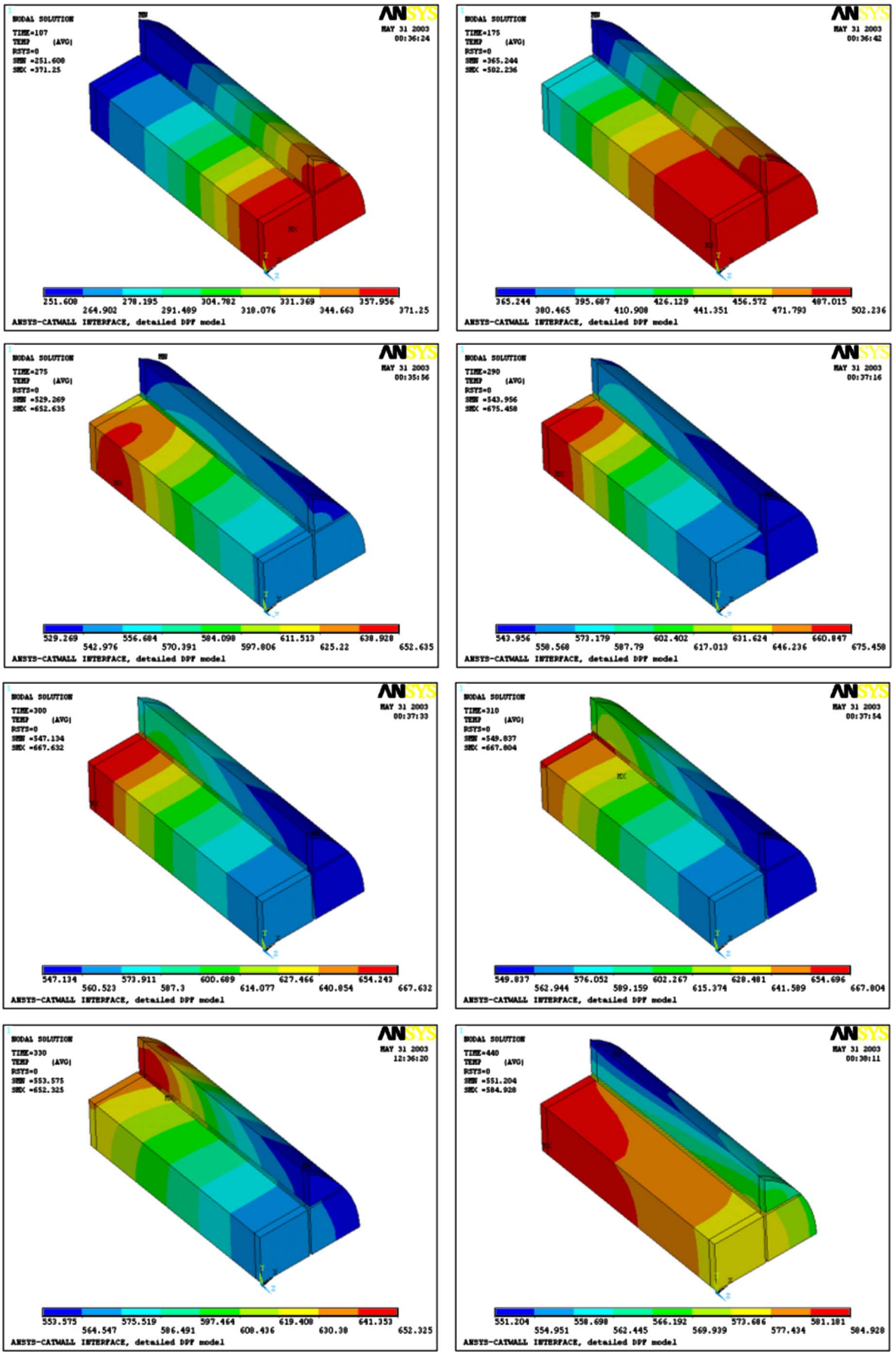


Fig. 9 Sequence of snapshot views of the predicted 3D temperature field of the SiC, in the interior of the filter, at the SiC—adhesive cement boundary (for the regeneration experiment of Fig. 8)

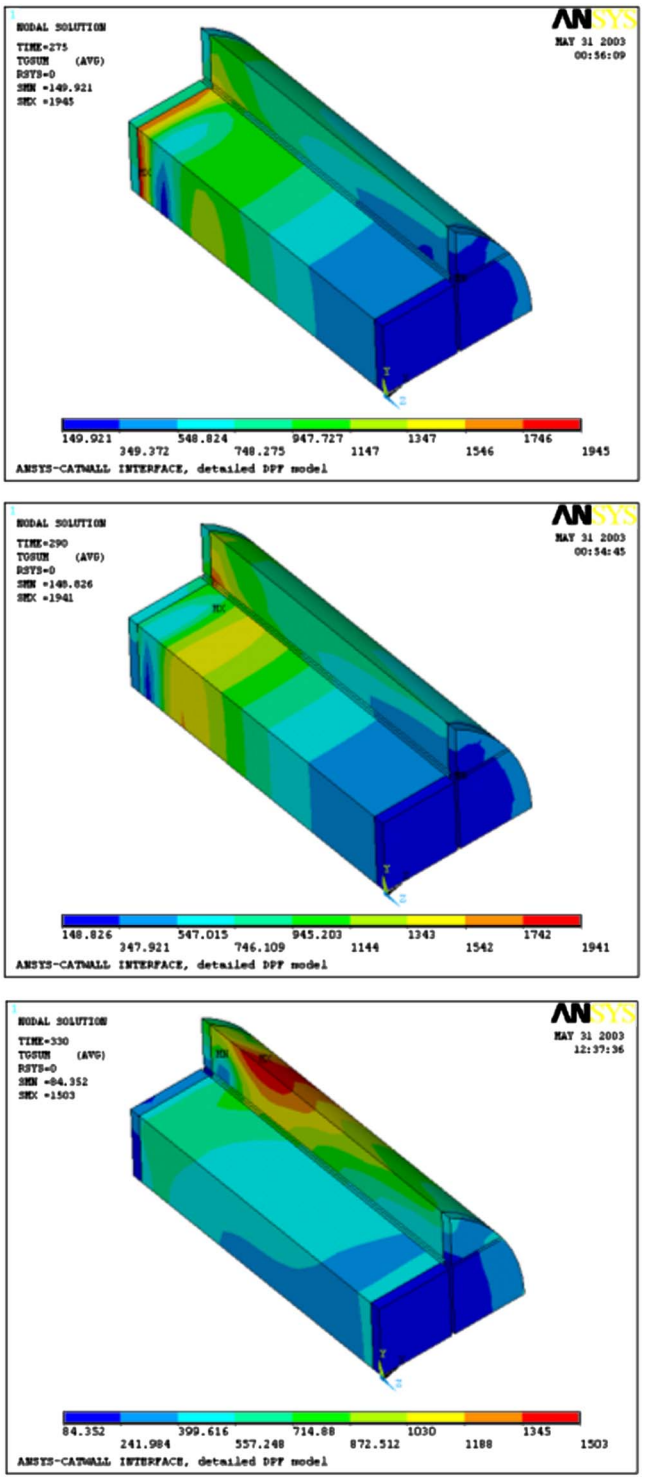


Fig. 10 Computed 3-D temperature gradient field in three characteristic time points (t=275, 290, and 330 s) for the regeneration test case of Fig. 8

tic time points. The maximum temperature gradients of the inner filter block are initially observed near the outlet of the filter, at the boundary where the channels' plugs start. This is a consequence of the higher heat capacity of the filter in the plug region. As the regeneration moves to the periphery of the filter, the same happens with the thermal gradients. Maximum thermal gradients are predicted at the small periphery block at the interior cement-SiC

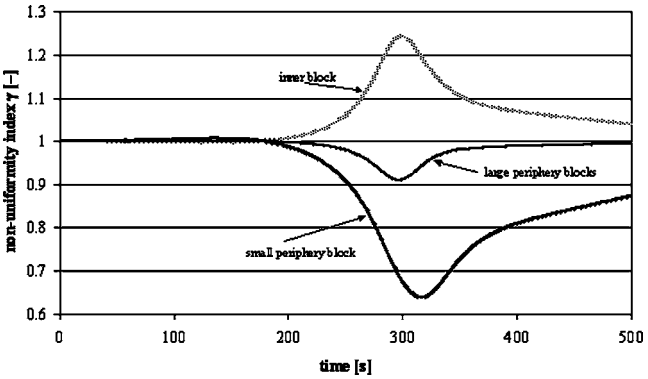


Fig. 11 Computed mass flow rate nonuniformity index during the first 500 s of the simulation of the experiment of Fig. 8, for the four bricks of the filter

interface. Later on, thermal gradients gradually diminish, although their maximum values still remain at the small periphery block of the filter.

The temperature gradient fields are employed in the computation of the thermal stress field of the filter. This process was discussed in more detail in [7]. Maximum thermal stresses of the order of 20 MPa have been computed in severe filter failure scenarios like that of Fig. 1. Such levels of thermal stresses lead to the failure of the filter material by cracking, as confirmed by the full-scale experiments. Thus, the identification of high thermal stress concentrations areas like those of Fig. 10 is very important for improvements in the detailed design of the filter.

The third step taken in this study is the examination of the flow profiles at the inlet of the filter, as affected by filter loading maldistribution. Specifically, incomplete regenerations lead to prominent flow maldistribution, which depends on the nonuniformity of the soot layer distribution and the corresponding flow resistances. To quantify this effect, we define an index of the mass flow rate nonuniformity for each brick, as follows:

$$\gamma = \frac{\text{flow entering the brick}}{\text{flow that would enter the brick if the flow profile were uniform}} = \frac{\dot{m}_i}{\dot{m}A_i/A}$$

The above index is unity for uniform flow. Values greater than unity imply higher mass flow rates than those expected from a uniform inlet flow profile, and vice versa.

In Fig. 11, this index is plotted through the first 500 s of the simulation for the four bricks that comprise the quarter of the filter. In the beginning of the simulation, all channels are assumed uniformly loaded with particulate, therefore the nonuniformity index equals unity for all bricks. However, during the regeneration, the model predicts a significant nonuniformity of the flow. Specifically, the flow is channeled mainly through the central brick of the filter, since regeneration occurs there first. Mass flow rates through the center brick are approximately double compared to those at the filter periphery. The above results of the inlet flow distribution are aligned with the experimental findings presented in [13,18].

A more detailed view of the nonuniformity of the inlet flow is also given in Fig. 12, where two snapshots of the percentage of the flow entering each sector of the model are presented. It is noticeable that the model also predicts a mass flow rate profile across each filter block, which declines towards the outlet of the filter. This effect is the result of the temperature field within the filter. Temperatures decline at the filter periphery due to the losses to ambient air, and thus less soot mass is consumed and the flow resistance is higher. The flow nonuniformity becomes milder after regeneration propagates to the periphery of the filter. It is impor-

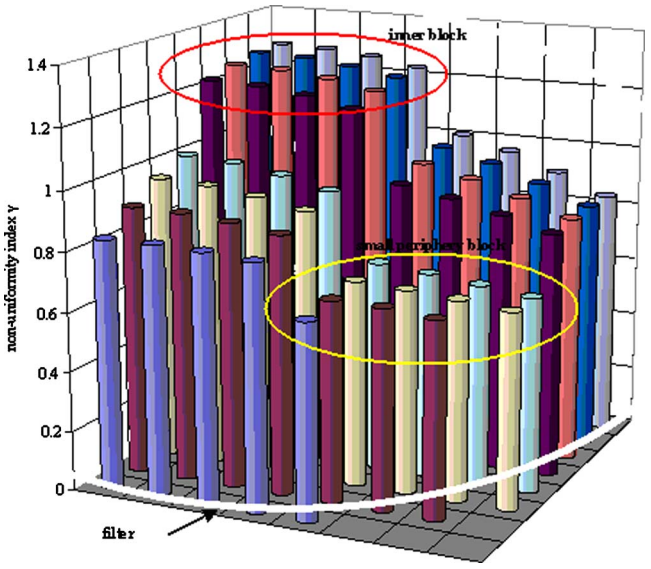


Fig. 12 Computed percentage of the flow entering each sector of the 3-D model at t=300 s for the regeneration experiment of Fig. 8

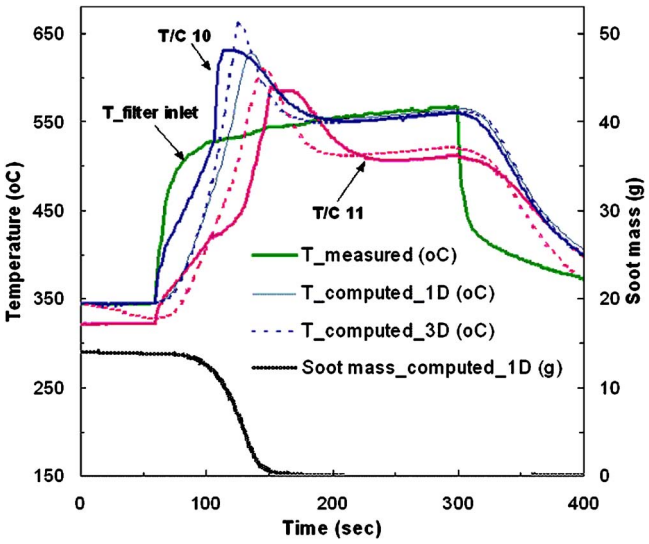


Fig. 14 Computational simulation of the regeneration of Fig. 13 by the 1-D and the 3-D codes. Predictions of measured temperatures by thermocouples T/C10 (filter central area) and T/C11 (near filter periphery) are presented for comparison. Predicted soot mass is also shown.

tant to note, though, that, according to Fig. 11, the flow profile remains nonuniform after the end of the regeneration. This may be due to incomplete regeneration in the periphery, but also to residual temperature differences remaining in the monolith that cause differences in exhaust gas density and thus mass flowrate among the channels. This flow maldistribution has an impact to the total pressure drop induced by the filter and is connected to the problem of estimating the soot load level of the filter employing pressure drop measurements. As a further example, an attempt is made to compute the evolution of the medium flowrate regeneration of Fig. 13. In this case the 3-D model successfully predicts the evolution of regeneration and the phase shift between center and periphery (Fig. 14) with the same kinetics parameters already presented in Table 4. The 1-D model predictions are also inserted in Fig. 14 for comparison.

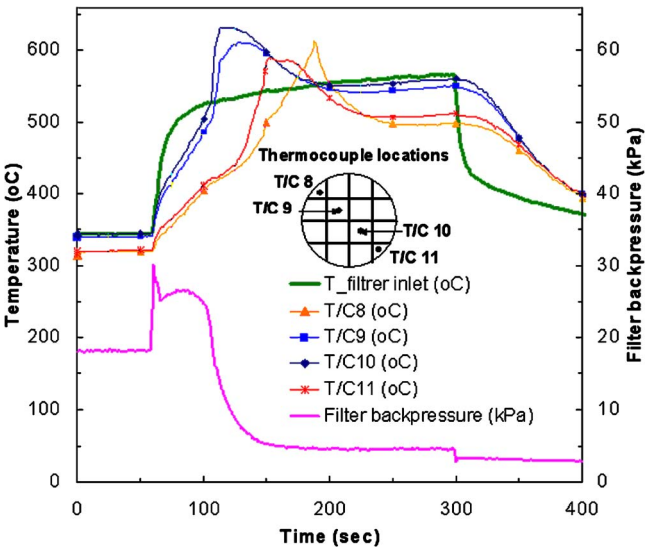


Fig. 13 Test results of a medium flow rate regeneration scenario (52 g/s). Measured temperatures along a filter diameter near the channels exit together with filter backpressure are presented during regeneration at 4000 rpm, 30 Nm. Moderate filter loading (6 g/l) has been accumulated with the engine running at 3000 rpm, 40 Nm with 25 ppm Ce-doped fuel.

Of course, the development of a 3-D regeneration model is not intended to substitute for the 1-D model, since the scope of each model is different. Rather, the two models should be viewed as complementary tools. As mentioned above, the application of the 3-D model is based on the original tuning of the reaction kinetics performed by the 1-D model, to save computation and manpower cost. Furthermore, the 1-D model is useful in the initial stage of the exhaust system design, where it should provide a rough view of the overall performance characteristics of the filter. The need for 3-D modeling comes to the foreground in detailed filter design and optimization studies, when local phenomena and especially thermal stress calculations are of interest. In its current state of development, and despite the acknowledged limitations, the 3-D filter regeneration model provides insight to the evolution of regeneration in the modular SiC filter and may be used for the determination of stress concentrations and the effect of critical material properties. This could be a valuable aid towards the identification of design flaws and decisive directions for research.

5 Conclusions and Future Work

A 3-D modeling approach for the catalytic regeneration of diesel filters was developed by interfacing a one-dimensional catalytic regeneration filter channel model with commercial FEM software. The 1-D base features a reaction scheme for thermal and catalytic regeneration. It has already been tested and validated extensively against characteristic cases, including failure scenarios.

The 3-D model is enhanced with a module for the determination of the flow distribution among the different channels of the filter. It extends the potential of the model towards the prediction of the full temperature field within the filter, the nonuniformity of the flow, and the effect of incomplete filter regeneration to the pressure drop of the filter. Also, stress analysis and determination of stress concentrations in the detailed design of diesel filters of modular design, with highly anisotropic structure, can be carried out based on the above. Characteristic validation cases are presented in this paper to demonstrate its capabilities and weak points. The limitations of the model are illustrated by purposely

selecting a difficult test scenario to demonstrate the effects of the volatile organic fraction content of the soot at low flow rates.

It should become clear by the overall presentation that the development and validation of workable engineering models for diesel filter regeneration is significantly more demanding than the development of their computational cores that are usually published in the literature with simple validation test cases. Nevertheless, engineering models are already essential in engineering design of diesel filter systems and their accuracy and application range are constantly improving to address challenging design optimization tasks.

Acknowledgment

The authors would like to thank their collaborators, George Stratakis and George Konstantas, for supplying experimental data and running a number of simulation cases, as well as Sungtae Hong of IBIDEN Co. Ltd. for his useful discussions and comments, and assistance with ANSYS simulations.

Nomenclature

A = 1. Preexponential factor of reaction rate expression, mol/m^3s; 2. Cross-sectional area of the soot layer, m^2
c_p = specific heat capacity, J/kg·K
d_h = hydraulic diameter of channel, m
E = activation energy of reaction rate expression, J
ΔH = molar heat of reaction, J/mol
h = heat transfer coefficient, W/m^2 K
k = Arrhenius-type rate factor, mol/m^3 s
ℓ = length of inner channel edge, m
$\dot{m}$ = exhaust gas mass flow rate, kg/s
M = molecular mass, kg/mol
p = pressure, Pa
q = heat flux, W/m^3
Q = heat, J
r = rate of reaction, mol/m^3 s
R_g = universal gas constant, 8.314 J/mol·K
t = time, s
T = temperature, K
u_z = exhaust gas velocity, m/s
V = volume, m^3
w = thickness, m
y = oxygen molar fraction, −
z = distance from the filter inlet, m

Greek Letters

ε = void fraction
λ = thermal conductivity, W/m·K
μ = viscosity
ρ = density, kg/m^3
ψ = fractional extent of the catalytic additive oxidation, −
ξ = total concentration of catalytic additive in the soot layer, mol/m^3

Subscripts

g = gas
in = inlet channel
out = outlet channel
p = particle (soot) layer
s = substrate
w = wall

Abbreviations

DI = direct injection
DPF = diesel particulate filter
EGR = exhaust gas recirculation
HC = hydrocarbons
HDI = high pressure direct injection
LTTE = laboratory of thermodynamics and thermal engines
mfr = exhaust gas mass flow rate
TF = temperature of filter
T/C = thermocouple fitted on experimental layout
TGA = thermo-gravimetric analysis
VOF = volatile organic fraction

References

[1] Salvat, O., Marez, P., and Belot, G., 2000, "Passenger Car Serial Application of a Particulate Filter System on a Common Rail Direct Injection Diesel Engine," SAE Paper No. 2000-01-0473.
[2] Koltsakis, G. C., and Stamatelos, A. M., 1997, "Catalytic Automotive Exhaust Aftertreatment," Prog. Energy Combust. Sci., **23**, pp. 1–39.
[3] Depcik, C. D., 2003, "Modeling Reacting Gases and Aftertreatment Devices for Internal Combustion Engines," Ph.D. Thesis, University of Michigan.
[4] Stratakis, G. A., Pontikakis, G. N., and Stamatelos, A. M., 2004, "Experimental Validation of a Fuel Additive Assisted Regeneration Model in SiC Diesel Filters," Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.), **218**, pp. 203–216.
[5] University of Thessaly/LTTE, 2003, "CATWALL Diesel Particulate Filter Modeling Software," User's Manual, Version v4r3 Volos, September.
[6] Pontikakis, G., 2003, "Modeling, Reaction Schemes and Parameter Estimation in Catalytic Converters and Diesel Filters," Ph.D. thesis, Mechanical Engineering Dept., University of Thessaly, Volos (http://www.mie.uth.gr/labs/ltte/pubs/PhD_G_Pont.pdf).
[7] Pontikakis, G., Stamatelos, A., Bakasis, K., and Aravas, N., 2002, "3-D Catalytic Regeneration and Stress Modeling of Diesel Particulate Filters by ABAQUS FEM Software," SAE Paper No. 2002-01-1017.
[8] Bissett, E., and Shadman, F., 1985, "Thermal Regeneration of Diesel-Particulate Monolithic Filters," AIChE J., **31**(5), pp. 753–765.
[9] Bissett, E., 1984, "Mathematical Model of the Thermal Regeneration of a Wall-Flow Monolith Diesel Particulate Filter," Chem. Eng. Sci., **39**, p. 1232.
[10] Koltsakis, G. C., and Stamatelos, A. M., 1996, "Modeling Thermal Regeneration of Diesel Particulate Traps," AIChE J., **42**(6), pp. 1662–1672.
[11] Koltsakis, G. C., and Stamatelos, A. M., 1996, "Modeling Catalytic Regeneration of Diesel Particulate Traps," Ind. Eng. Chem. Res., **35**, pp. 2–13.
[12] Koltsakis, G. C., and Stamatelos, A. M., 1997, "Modes of Catalytic Regeneration in Diesel Particulate Filters," Ind. Eng. Chem. Res., **36**, pp. 4155–4165.
[13] Stratakis, G., 2004, "Experimental Investigation of Catalytic Soot Oxidation & Pressure Drop Characteristics in Wall-Flow Diesel Particulate Filters," Ph.D. thesis, Mechanical Engineering Dept., University of Thessaly (http://www.mie.uth.gr/labs/ltte/pubs/PhD_Stratakis.pdf).
[14] Haralampous, O., and Koltsakis, G., 2002, "Intra-layer Temperature Gradients During Regeneration of Diesel Particulate Filters," Chem. Eng. Sci., **57**, pp. 2345–2355.
[15] Aoki, H., Asano, A., Kurazono, K., Kobashi, K., and Sami., H., 1993, "Numerical Simulation Model for the Regeneration Process of a Wall-Flow Monolith Diesel Particulate Filter," SAE Paper No. 930364.
[16] Miyairi, Y., Miwa, S., Abe, F., Xu, Z., and Nakasuji, Y., 2001, "Numerical Study on Forced Regeneration of Wall-Flow Diesel Particulate Filter," SAE Paper No. 2001-01-0912.
[17] Opris, C. N., and Johnson, J. H., 1998, "A Two-Dimensional Computational Model Describing the Flow and Filtration Characteristics of a Ceramic Diesel Particulate Trap," SAE Paper No. 980545.
[18] Stratakis, G. A., and Stamatelos, A. M., 2004, "Flow Distribution Effects in the Loading and Catalytic Regeneration of Wall-Flow Diesel Particulate Filters," Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.), **218**, pp. 203–216.
[19] Gaiser, G., 2004, "Prediction of Pressure Drop in Diesel Particulate Filters Considering Ash Deposit and Partial Regenerations," SAE Paper No. 2004-01-0158.
[20] "Guide to ANSYS User Programmable Features. Guide to Interfacing With ANSYS," ANSYS 5.7 User's Manual, ANSYS Inc., Canonsburg, PA, May 2001.
[21] Sorenson, S. C., Hoj, J. W., and Stobbe, P., 1994, "Flow Characteristics of SiC Diesel Particulate Filter Materials," SAE Paper No. 940236.
[22] Ebener, S., and Florchinger, P., 2000, "Drukverlustmodel Fuer Keramische Dieselpartikelfilter," Motortech. Z., **61**(6), pp. 414–422.
[23] Incropera, F. P., and DeWitt, D. P., 1996, *Fundamentals of Heat and Mass Transfer*, 4th ed., John Wiley and Sons, New York.
[24] Stoecker, W. F., 1989, *Design of Thermal Systems*, 3rd ed., McGraw-Hill, New York.
[25] University of Thessaly/LTTE, 2003, "ANSYS 5.7-CATWALL v4r3 Interfacing: 3D Modeling of Diesel Particulate Filters," User's Manual. Version v4r3 Volos, December.
[26] Stratakis, G. A., and Stamatelos, A. M., 2003, "Thermogravimetric Analysis of Soot Emitted by a Modern Engine Run on Catalyst-Doped Fuel," Combust. Flame, **132**, pp. 157–169.

A Comprehensive Thermodynamic Approach to Acoustic Cavitation Simulation in High-Pressure Injection Systems by a Conservative Homogeneous Two-Phase Barotropic Flow Model

Andrea E. Catania
Alessandro Ferrari
Michele Manno
Ezio Spessa

Energetics Department
& IC Engines Advanced Laboratory,
Politecnico di Torino,
Torino, Italy

A general conservative numerical model for the simulation of transmission-line unsteady fluid dynamics has been developed and applied to high-pressure injection systems. A comprehensive thermodynamic approach for modeling acoustic cavitation, i.e., cavitation induced by wave propagation, was proposed on the basis of a conservative homogeneous two-phase barotropic flow model of a pure liquid, its vapor, and a gas, both dissolved and undissolved. A physically consistent sound speed equation was set in a closed analytical form of wide application. For the pure-liquid flow simulation outside the cavitation regions, or in the absence of these, temperature variations due to compressibility effects were taken into account, for the first time in injection system simulation, through a thermodynamic relation derived from the energy equation. Nevertheless, in the cavitating regions, an isothermal flow was retained consistently with negligible macroscopic thermal effects due to vaporization or condensation, because of the tiny amounts of liquid involved. A novel implicit, conservative, one-step, symmetrical, and trapezoidal scheme of second-order accuracy was employed to solve the partial differential equations governing the pipe flow. It can also be enhanced at a high-resolution level. The numerical model was applied to wave propagation and cavitation simulation in a high-pressure injection system of the pump-line-nozzle type for light and medium duty vehicles. The system was relevant to model assessment because, at part loads, it presented cavitating flow conditions that can be considered as severe, at least for a diesel injection system. The predicted time histories of pressure at two pipe locations and of injector needle lift were compared to experimental results, substantiating the validity and robustness of the developed conservative model in simulating acoustic cavitation inception and desinence with great accuracy degree. Cavitation transients and the flow discontinuities induced by them were numerically predicted and analyzed. [DOI: 10.1115/1.2056007]

Introduction

The numerical simulation of cavitating flows in engine fuel injection systems is useful for better understanding the complex time-dependent flow phenomena at the early development stage of such systems and to prevent erosion damages.

Excessive liquid pressure decrease, up to the vapor tension, which can be induced by the wave propagation motion, leads to oil evaporation that is called acoustic cavitation. This reduces the mass flow rate and affects the valve working characteristics. The collapse of the vapor bubbles can lead to noise in addition to the component detriment. Acoustic cavitation can be described by several models with different accuracy degree. Many papers deal with this topic ([1–10]).

An initial approach to the injection-system modeling is to take cavitation effects into account by drastic simplifications that neglect the mutual influence of the cavitating regions on the surrounding liquid. More specifically, the pressure is set equal to the vapor pressure value at those points where it drops below such a value ([11]). This can lead to inconsistent solutions close to the evaporation regions with effects that amplify in time.

Works on cavitating flow simulation, such as [2,6], which are based on one-dimensional models of pseudocavitating flows ([12,13]) split the continuity equations for the gaseous and liquid phases coupling them through a source term that drives the evaporation and condensation processes and that must be modeled. However, a model for this source term is not easy to be derived ([3,10]). An elementary thermodynamic equilibrium model was applied to simulate the source term in [10], while the energy equation was neglected for the homogeneous two-phase mixture considered there. Such a model led to the largest mass fraction of steam in comparison to the nonequilibrium cavitation models derived in [14].

Discrete-bubble cavitation models, which take oscillations of single bubbles or bubble clouds into account (e.g., [7]), are very difficult to be numerically handled, particularly for flows with high pressure gradients ([10]). Besides, these models imply many parameter assumptions, which can hardly be assessed by experiments, making the study of basic phenomena uncertain. Even

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 28, 2003; final manuscript received June 1, 2005. Assoc. Editor: D. Assanis.

though it is stated in [7] that bubbles are subjected to nonequilibrium processes, thermodynamic equilibrium laws are used such as the *Henry's law* for the dissolved gas. Furthermore, the microscale bubble-dynamic model in [7] needs a homogeneous-mixture macroscopic model where the mixture density is derived from the liquid density and from the rather arbitrary volumetric density and uniform radius of the undeformable bubbles, neglecting the density of the vapor and gas in the bubble.

Multidimensional modeling has been applied to flow field optimization in specific fuel-injector regions, such as the nozzle tip ([15]). However, this is cumbersome and unnecessary for the whole system simulation and acoustic cavitation numerical analysis.

To solve the nonlinear partial differential equations for one-dimensional fluid flow in pressure pipes, many numerical codes use the characteristics method, including the effects of convective and friction terms. However, this method is known to present difficulties when dealing with cavities in transient flows ([16]). The simple procedures applied to isolate the cavities with this method ([1,17]) are generally inappropriate and too cumbersome for injection-system cavitation analysis ([18]).

A survey of numerical algorithms used for solving pipe-flow equations is reported in [3,19].

Nonconservative implicit numerical models of the first- and of the second-order accuracy were employed in [3] and in [5,8], respectively, to simulate unsteady flow and cavitation in high-pressure injection systems. Implicit differentiation formulas that are suitable for problems of the stiff type were applied for the ordinary differential equations modeling the pump and injector. Therefore, implicit schemes were selected for discretizing the partial differential equations governing the pipe flow, so that a single algorithm could be built to solve all the equations of the system components at the same time with a fast direct resolution method. The cavitation simulation was based on an isothermal homogeneous mixture model of liquid and vapor, introducing the void fraction as an unknown variable. Two different equation sets were solved, one for the pressure and velocity variables, in the absence of cavitation, and the other for the void fraction and velocity in the presence of cavitation. The cavitating region was located by the point that first reached the vapor pressure and was determined by an iterative procedure. Though a good agreement between computational and experimental results was generally obtained ([3,5,8]), the model was difficult to be numerically handled.

Nonconservative models can assure satisfactory accuracy in wave propagation phenomena simulation, but they introduce fictitious source terms in the discretized equations that can play a sensible role when discontinuities in the flow-property space and time derivatives occur, as in the presence of cavitation ([20]).

The main objective of the present work was to develop an enhanced conservative model for cavitating and noncavitating one-dimensional flows of general application. The implicit, conservative, second-order accurate scheme detailed in [21,22] was applied to discretize the pipe-flow partial differential equations in divergence form. It was shown to provide very good resolution at the cavitation induced flow discontinuities, with reduced oscillation problems. Besides, since the unknown variables are density and momentum the same set of equations was applied both in the absence and in the presence of cavitation ([21,22]).

The model was tested and validated through its application to the transient flow simulation in the pressure pipe of a pump-line-nozzle diesel injection-system built by Bosch for light- and medium-duty commercial vehicles. The system was of particular relevance to the model assessment, because it was subjected to cavitating flow conditions that can be considered as severe at part loads, at least for a diesel injection system, according to Bosch. Furthermore, it was a challenging benchmark for the numerical simulation of a high-pressure fuel injection system, due to its complex fluid dynamics. The predicted injector needle lift and pressure time histories at two pipe locations, one close to the pump and the other close to the injector, were compared to experimental results, substantiating the validity and robustness of the conservative model in simulating the acoustic cavitation with a high degree of accuracy. Numerical results on cavitation inception, evolution, and desinence are presented and discussed.

Cavitation Simulation

In a slender pipe, the one-dimensional barotropic flow of a compressible pure liquid or of a homogeneous mixture of liquid and aeriform constituents requires only the mass and momentum conservation equations for its mathematical modeling ([3]). The energy equation is reduced to the thermodynamic state equation for the specific process the fluid is subjected to.

According to a conservative formulation, the flow model equations are written in the divergence form with density (ρ) and momentum (ρu) as variables, i.e., using a control volume approach, with respect to a pipe element, the mass-conservation and momentum-balance equations of the fluid are expressed in the following matrix form:

$$\frac{\partial \mathbf{w}}{\partial t} + \frac{\partial \mathbf{F}}{\partial x} = \mathbf{H} \tag{1}$$

where

$$\mathbf{w} = \begin{bmatrix} \rho \\ \rho u \end{bmatrix} \quad \mathbf{F}(\mathbf{w}) = \begin{bmatrix} \rho u \\ p + \rho u^2 \end{bmatrix} \quad \mathbf{H} = \begin{bmatrix} 0 \\ -\dfrac{4\tau_w}{d} \end{bmatrix} \tag{2}$$

ρ, u, and p are the respective average cross-sectional density, velocity, and pressure of either the liquid or the homogeneous mixture; τ_w is the wall shear stress and d is the pipe diameter; t is the time variable and x is the axial variable along the pipe.

The state and transformation equations of the fluid have to be added for a barotropic flow model.

The cavitating flow at a macroscopic level ([3,7,10]) can be taken as a homogenous mixture of a pure liquid, its vapor and a quantity of gas, both dissolved and undissolved. The mixture of liquid and gaseous constituents can be treated as a macroscopically homogeneous and isotropic system, with no slip between liquid and aeriform constituents, if the amount of these latter is very small and their bubbles are finely distributed. In this case, the mass and void fractions of the aeriform constituents can be introduced as variables:

$$\mu_x = \frac{m_x}{m}, \quad \alpha_x = \frac{V_x}{V}, \tag{3}$$

where x refers to either the vapor (v) or the undissolved gas (g), m and V designate the mass and volume of the mixture. With specific reference to the dissolved gas, its mass fraction will be indicated by $\mu_{dg}=m_{dg}/m$ and its void fraction is virtually nil. The effective properties of the liquid are experimentally determined and are not significantly influenced by the tiny amount of the dissolved gas, whenever this is present. It is worth pointing out that, due to the generally small concentration of the dissolved gas, the *Henry's law* can be assumed to govern the gas dissolution process in the liquid and thus can be used to determine μ_{dg} for given thermodynamic conditions of the liquid ([7]).

In general, the homogeneous mixture of a pure liquid, its vapor and a gas both dissolved and undissolved, that are finely distributed in it, can be treated as a pure phase at a macroscopic level, taking average local intensive properties of the mixture ([3,7,10]). More specifically, a mechanical and thermodynamic equilibrium model results ([2,3,10]) by neglecting the surface tension and viscosity in the static bubble-wall equilibrium and by considering a mass-averaged internal energy of the mixture leading to an average local temperature virtually equal to the liquid temperature, due to the tiny amount of gaseous constituents involved. Therefore, if the intensive properties of such a mixture are related

through a thermodynamic statement, describing the process it is subjected to (barotropic flow model), two intensive properties (namely, ρ and μ_g) are required to describe its internal state, i.e., the system has two degrees of freedom, according to the *Gibbs phase rule* ([23]). Besides, if a definite value (even nil) of the mass fraction of gas, both dissolved and undissolved, is assigned, the mixture reduces to a single degree-of-freedom system, that is, only one intensive property (namely, ρ) is required to describe its thermodynamic state.

Applying the primitive definition of specific volume to the mixture, as well as to each constituent, and introducing the mass fractions of Eqs. (3), the homogeneous mixture density can be expressed as a function of the pressure, the temperature, the vapor mass fraction and the undissolved gas mass fraction, by (Appendix A):

$$\rho = \frac{\rho_v \rho_g \rho_l}{\mu_v \rho_g \rho_l + \mu_g \rho_v \rho_l + (1 - \mu_v - \mu_g)\rho_v \rho_g} \tag{4}$$

Since for a barotropic flow, and generally in vaporization regions, the temperature T is a function of pressure, $T=T(p)$, in Eq. (4) the liquid density is a function of p, i.e., $\rho_l=\rho_l(p)$:

$$\rho_l = \frac{m_l}{V_l} \tag{5}$$

V_l is the volume of the liquid, including the dissolved gas.

The vapor and gas densities ρ_v and ρ_g, respectively, are given by the *Amagat's law of the additivity of volumes* ([23]), the vapor being treated as a perfect gas:

$$\rho_x = \frac{m_x}{V_x} = \frac{p}{R_x T} \tag{6}$$

where x designates either the vapor (v) or the undissolved gas (g). Thus, $\rho_v=\rho_v(p)$ and $\rho_g=\rho_g(p)$ because $T=T(p)$.

Let us consider a constant total mass fraction of dissolved and undissolved gas. The actual amount of dissolved gas that saturates the liquid can be expressed by the *Henry's law* ([24]):

$$\mu_{\mathrm{dg}} = K(p) p_g = K(p)(1 + \mu_v \rho_g / \mu_g \rho_v)^{-1} p \tag{7}$$

p_g is the gas partial-pressure, according to the *Dalton's Law* and $K(p)$ is related to the Henry's coefficient.

If μ_g is replaced in Eq. (4) by the complement to μ_{dg} with respect to the total gas mass-fraction, through Eq. (7), and the mixture density ρ is taken as the property that describes the so-obtained monovariant system, then the vapor mass fraction μ_v must be a function of ρ. This means that Eq. (4) connects the pressure p to the density ρ of the mixture, as it should be for a barotropic fluid flow. Therefore, we can calculate the pressure gradient as follows:

$$\frac{\partial p}{\partial x} = \frac{dp}{d\rho}\frac{\partial \rho}{\partial x} \tag{8}$$

In order to derive a general expression for the sound speed of the mixture, Eq. (1) is rewritten in the following form:

$$\frac{\partial \mathbf{w}}{\partial t} + \mathbf{A}\frac{\partial \mathbf{w}}{\partial x} = \mathbf{H} \tag{9}$$

where

$$\mathbf{A} = \begin{bmatrix} 0 & 1 \\ \dfrac{dp}{d\rho} - u^2 & 2u \end{bmatrix} \tag{10}$$

is the Jacobian matrix of the partial differential equation system. As can be easily verified, the eigenvalues λ of the matrix $\mathbf{A}$ are:

$$\lambda_{1,2} = u \pm \sqrt{\frac{dp}{d\rho}} \tag{11}$$

and therefore one can deduce that the sound speed of the homogeneous barotropic mixture is given by:

$$a = \sqrt{\frac{dp}{d\rho}} \tag{12}$$

in agreement with the thermodynamic definition of sound speed in a simple system ([23]).

The sound speed for the homogeneous barotropic mixture, with a fixed total mass fraction of gas (dissolved and undissolved), can thus be drawn from Eq. (12) by taking the derivative of the mixture density ρ (Eq. (4)) with respect to the pressure p, so that:

$$\frac{1}{\rho a^2} = \frac{\alpha_v}{\rho_v a_v^2} + \frac{\alpha_g}{\rho_g a_g^2} + \frac{1 - \alpha_v - \alpha_g}{\rho_l a_l^2} - \rho\left(\frac{1}{\rho_v} - \frac{1}{\rho_l}\right)\frac{d\mu_v}{dp} - \rho\left(\frac{1}{\rho_g} - \frac{1}{\rho_l}\right)\frac{d\mu_g}{dp} \tag{13}$$

where a_v, a_g, a_l are the sound speeds of the constituents for the specified thermodynamic evolution. Such an analytical form of the mixture sound speed a is a new one, because usually the pseudocavitation form ([2,6,10]) is considered. The first three terms on the right-hand side of Eq. (13) are related to the elasticity of each component of the mixture, whereas the last two terms refer to the vaporization or condensation and to the gas entrainment or dissolution processes. This can be easily verified by introducing an effective bulk elasticity modulus E_{eff} of the mixture for the detailed thermodynamic process, as is done in the Appendix B:

$$\frac{1}{E_{\mathrm{eff}}} = -\frac{1}{V}\frac{dV}{dp} \tag{14}$$

The knowledge of the derivatives $\mathrm{d}\mu_v/\mathrm{d}p$, $\mathrm{d}\mu_g/\mathrm{d}p$ is required to obtain the wave propagation speed from Eqs. (11)–(13).

The derivative of μ_g with respect to p can be obtained by differentiation of Eq. (7), because it is:

$$\frac{d\mu_g}{dp} = -\frac{d\mu_{\mathrm{dg}}}{dp} \tag{15}$$

As far as the term $\mathrm{d}\mu_v/\mathrm{d}p$ is concerned, a procedure for its evaluation is indicated in the Energy Considerations Section.

The following expressions are directly derived from Eq. (13) to show its general applicability for relevant cases.

Pure Vaporous Cavitation. Let us consider the case of a pure liquid under cavitation conditions in the absence of any quantity of undissolved gas ($\mu_g=0, \alpha_g=0$). In this case, Eq. (13) reduces to:

$$\frac{1}{\rho a^2} = \frac{\alpha_v}{\rho_v a_v^2} + \frac{1 - \alpha_v}{\rho_l a_l^2} - \rho\left(\frac{1}{\rho_v} - \frac{1}{\rho_l}\right)\frac{d\mu_v}{dp} \tag{16}$$

It is worth pointing out that the assumption of isothermal flow in the cavitation region is consistent with negligible thermal effects of the fluid phase change induced by wave propagation and is generally acceptable, owing to the very small amount of liquid involved. With the assumption of isothermal flow in the cavitation region, the pressure is constant and equal to the vapor pressure at the given temperature, so that ρ_v and ρ_l are constant for the above macroscale mixture model. Thus, Eq. (4), yielding

$$\rho = \frac{\rho_v \rho_l}{\mu_v \rho_l + (1 - \mu_v)\rho_v} \tag{17}$$

directly relates the vapor mass fraction μ_v to the mixture density ρ. In such a case, the following expression of the sound speed is obtained from Eqs. (12) and (17):

$$\frac{1}{\rho a^2} = -\rho\left(\frac{1}{\rho_v} - \frac{1}{\rho_l}\right)\frac{d\mu_v}{dp} \tag{18}$$

Since μ_v assumes variable non-negative values, its derivative with respect to p tends to infinity. Thus, Eq. (18) leads to:

$$a = 0 \tag{19}$$

as can be also directly deduced from Eq. (12).

Pure Gaseous Cavitation. Assume that pressure reductions take place above the liquid vapor tension, so that the vaporous cavitation does not occur and the cavities in the liquid contain virtually only undissolved gas. In this case $\mu_v \approx 0$, $\alpha_v \approx 0$.

From Eqs. (13) and (15) one obtains:

$$\frac{1}{\rho a^2} = \frac{\alpha_g}{\rho_g a_g^2} + \frac{1-\alpha_g}{\rho_l a_l^2} + \rho\left(\frac{1}{\rho_g} - \frac{1}{\rho_l}\right)\frac{d\mu_{\mathrm{dg}}}{dp} \tag{20}$$

The differential quotient in the last term can be evaluated by the derivative of Eq. (7) with respect to the pressure.

Pseudocavitation. This case relates to the evolution of a gas–liquid mixture with cavities containing a fixed amount of an undissoluble and uncondensable gas, under pressure variations such that the pressure always takes values that are greater than the vapor pressure. Therefore, the vaporous cavitation does not occur ($\mu_v \approx 0$ and $\alpha_v \approx 0$). The substantial difference between pseudocavitation and gaseous cavitation resides in the fact that pseudocavitation bubbles can grow or decrease in response to changes of the liquid pressure without any mass transfer across the bubble surface.

In this case, Eq. (13) yields:

$$\frac{1}{\rho a^2} = \frac{\alpha_g}{\rho_g a_g^2} + \frac{1-\alpha_g}{\rho_l a_l^2} \tag{21}$$

This expression is commonly applied by different authors [2,6,10,12,13] to derive the sound speed of mixtures of aeriform-liquid constituents under cavitating conditions, even though it is not consistent with cavitation phenomena description.

Pure Liquid Flow. Finally, for the case of pure liquid flow without any gaseous constituent ($\mu_g = 0, \mu_v = 0$) expression (13) yields:

$$a = a_l \tag{22}$$

that is, the sound speed of the liquid under the specific process.

Energy Considerations

For the one-dimensional flow of a compressible pure liquid or of a homogeneous mixture of liquid and aeriform constituents the thermal energy conservation equation can be written as:

$$\frac{dh}{dt} - \frac{1}{\rho}\frac{dp}{dt} = \dot{q} + \dot{l}_w \tag{23}$$

where h is the enthalpy per unit mass of the system, $\dot{q}$ is the heat power per unit mass transferred from the boundaries to system, and $\dot{l}_w$ is the viscous power dissipation per unit mass within the system.

Introducing the entropy, Eq. (23) gives rise to the following basic thermodynamic relation among the fluid state variables during any elemental process:

$$T\,ds = dh - \frac{dp}{\rho} \tag{24}$$

where T is the local temperature and s is the entropy per unit mass of the system.

Pure Liquid Flow. For a simple system, with specific reference to a pure liquid, the following general state equations can be easily derived:

$$dh = c_p dT + (1-\beta T)\frac{dp}{\rho} \tag{25}$$

$$dp = a_T^2\, d\rho + \beta E_T\, dT \tag{26}$$

where $a_T = \sqrt{E_T/\rho}$ is the sound speed at constant temperature,

$$\beta = \frac{1}{v}\left(\frac{\partial v}{\partial T}\right)_p = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial T}\right)_p \tag{27}$$

is the *volumetric coefficient of thermal expansion* or the *thermal expansivity*, and:

$$E_T = -v\left(\frac{\partial p}{\partial v}\right)_T = \rho\left(\frac{\partial p}{\partial \rho}\right)_T \tag{28}$$

is the *isothermal bulk modulus of elasticity* of the fluid.

Equation (26) relates the pressure variation to the density and temperature variations in a liquid for any process. Usually the term containing dT is neglected in the literature, so that the speed of sound in a liquid is generally evaluated with reference to an isothermal process. From Eqs. (24)–(26) it is easy to derive the general expression of the sound speed with reference to an isentropic process ($ds = 0$):

$$a_s = \sqrt{\left(\frac{dp}{d\rho}\right)_s} = \sqrt{\frac{c_p}{c_v}}\,a_T \tag{29}$$

The difference between the specific heat at constant pressure c_p and the specific heat at constant volume c_v can be higher than 15% for an oil of interest in diesel fuel simulation, such as, for example, the ISO 4113 oil. Thus, the relative difference between the squares of a_s and a_T can be of the same order of magnitude. Such a difference has an influence on the pressure prediction for isentropic and isothermal processes, as can be verified from Eq. (26).

It is worth linking separately the variations of pressure and density to the temperature variations in an isentropic process, through Eqs. (24)–(26), as follows:

$$\frac{dp}{E_T} = \frac{c_p}{\beta a_T^2}\frac{dT}{T}, \quad \frac{d\rho}{\rho} = \frac{c_v}{\beta a_T^2}\frac{dT}{T} \tag{30}$$

In the present approach, these relations were applied to evaluate the temperature variations of the liquid due to compressibility effects in high-pressure injection system simulation, for the first time in the literature.

Cavitating Flow. For the detailed homogeneous two-phase mixture of vapor–gas–liquid under cavitating conditions at a macroscopic level, the following thermal energy relation can be easily derived:

$$T\,ds = c_p\,dT - \left[\frac{1}{\rho} - \frac{\mu_l}{\rho_l}(1-\beta_l T)\right]dp + r\,d\mu_v + [r + T(s_g - s_v)]d\mu_g \tag{31}$$

where c_p is the average heat capacity of the mixture, defined as:

$$c_p = \mu_v c_{p_v} + \mu_g c_{p_g} + \mu_l c_{p_l} \tag{32}$$

ρ is the mixture density expressed by Eq. (4); μ_l, ρ_l, and β_l are the mass fraction, density, and expansivity, respectively, of the liquid; $r = r(T)$ is the condensation heat of the vapor.

Equation (31) and the thermodynamic process specification allow modeling the vapor source term, that is, the derivative of μ_v with respect to the pressure. In particular, for pure vaporous cavitation ($d\mu_g \approx 0$), under an isentropic process, Eq. (31) yields:

$$r\frac{d\mu_v}{dp} = \left[\frac{1}{\rho} - \frac{\mu_l}{\rho_l}(1-\beta_l T)\right] - c_p\frac{dT}{dp} \tag{33}$$

whereas, under an isenthalpic process ($dh = 0$), one has:

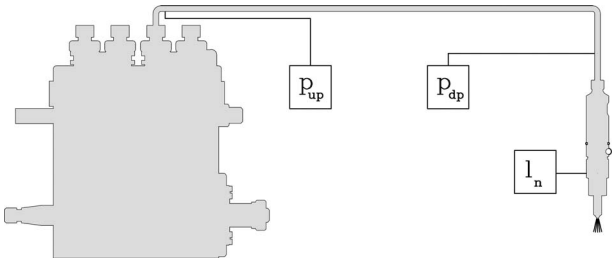


Fig. 1 Injection-system layout and measured quantities

$$r\frac{d\mu_v}{dp} = -\frac{\mu_l}{\rho_l}(1-\beta_l T) - c_p\frac{dT}{dp} \tag{34}$$

In the present work, an isothermal process was chosen to simulate the transients of vaporous acoustic cavitation in high-pressure injection systems, because of its physical consistency, simplicity, and accuracy in capturing the macroscopical effects.

Fuel Injection System

The high-pressure injection system, built by Bosch for light and medium duty vehicles, was made up of an in-line pump, a pressure pipe, and a single-spring injector. The injector nozzle was of the orifice type with five holes, presenting a reduced sac volume. In order to avoid the injector-nozzle reopening, the pump was equipped with a delivery valve of the relief-volume type and with a snubber valve downstream to it, in the forward flow direction, which was aimed at avoiding or attenuating cavitation occurrence inside the pipe ([25,26]). Cavitation can take place in the pipe subsequently to the spill port opening, because of the depression wave protraction and intensification due to the delivery valve configuration ([8]). In spite of the snubber-valve presence, the system underwent anomalous operating conditions of intense and widespread cavitation, mainly at part loads.

Figure 1 shows a schematic of the injection system layout with measurement quantities and locations. The experimental data used for comparison with computational results consisted in time histories of the injector-needle lift and of the pressure at two pipe locations: one close to the pump delivery outlet (p_{up}) and the other close to the injector inlet (p_{dp}). At these locations the average temperatures were also measured, showing almost the same value. The feeding test-oil temperature was kept under control.

The pump-plunger stroke was experimentally determined for two operating conditions: one at full load and the other at partial load, both at the pump speed of 1300 rpm. The cam-angle distributions of plunger lift, inlet- and spill-port cross-sectional areas were also measured. The measurements were carried out on a test bench of the type described in [5].

Mathematical Model

Pipe Flow. The fluid flowing through the pressure pipe was modeled as a pure liquid in the absence of cavitation or outside the regions under cavitation. In the pure liquid flow simulation, temperature variations due to compressibility effects were included, for the first time in injection system modeling and numerical analysis. To that end, the thermal energy conservation equation, Eq. (23), was applied together with an isentropic evolution, according to the hypothesis of an adiabatic flow and negligible friction-loss thermal effects, or rather, to the assumption that the heat yielded by the friction forces was entirely transferred outside through the pipe wall. The liquid temperatures were evaluated by Eqs. (30), linking pressure or density variations to temperature variations during the wave propagation through the pipe.

Whenever the pressure decreased up to the vapor tension at the local temperature, the fluid was modeled as a homogeneous mixture of liquid and vapor without undissolved gas ($\mu_g \approx 0$) (pure vaporous cavitation). The vapor was taken to be equally distributed within the cavitating region and, more specifically, within each of the computational cells.

As previously mentioned, an isothermal process was chosen to simulate the cavitation evolution for the following reasons: it is physically consistent with the negligible thermal effects of the tiny amount of vaporizing liquid; it is very simple to handle and accurate in capturing the macroscopical effects. Therefore, whenever at any location in the pipe the pressure reaches the value of the local vapor tension, the temperature at this point, that is equal to its saturation value, keeps constant until a vapor portion exists, and so the pressure does, consistently with the thermodynamic functional relation between temperature and pressure in a vaporizing pure-liquid region.

It is worth pointing out that, based on the results of some numerical tests, the simple approach of considering only a pure vaporous cavitation was chosen, neglecting the presence of any arbitrary amount of undissolved gas, as was done in previously developed nonconservative numerical models ([3,5,8]). In fact, besides its simplicity, this procedure was shown to yield physically consistent results, without any significant difference in the macroscopic effects with respect to the tested case in which the contemporary presence of a tiny fixed gas fraction is considered ([25]). This is supported by the good agreement between the predicted and measured pressure time histories at the above-mentioned pipe locations. In the present work, such macroscopic effects are of main interest, together with the robustness shown by the new numerical scheme.

Pump and Injector. The pressure-wave propagation along the injector-drilled passages is simulated with the foregoing pipe-flow model and a lumped mass model is applied to write the continuity equations for the pump and injector isobaric chambers. Further details on the rebuilt injection-system mathematical model are given in ([21,25]).

With reference to a generic chamber j, one can write:

$$Q_{\mathrm{in},j} - Q_{\mathrm{out},j} = \frac{1}{\rho_j}\frac{d\rho_j}{dt}V_j + \frac{dV_j}{dt} \tag{35}$$

where $Q_{\mathrm{in},j}$ and $Q_{\mathrm{out},j}$ are the volumetric flow rates coming into and out of the chamber j, respectively, and are modeled according to *Torricelli's law*.

A second-order linear system model was used to simulate the valves (the injector needle is also indicated as nozzle valve):

$$m_m\frac{d^2 l_m}{dt^2} + \bar{\beta}_m\frac{dl_m}{dt} + \bar{k}_m l_m + \bar{F}_0 = \sum_k p_k S_{m_k} \tag{36}$$

where

$$l_m < 0, \quad \bar{k}_m = k_m + k_{\mathrm{bs}}, \quad \bar{\beta}_m = \beta_m + \beta_{\mathrm{bs}} \quad \bar{F}_0 = F_0,$$

$$0 \leqslant l_m \leqslant l_M, \quad \bar{k}_m = k_m, \quad \bar{\beta}_m = \beta_m, \quad \bar{F}_0 = F_0,$$

$$l_m \geqslant l_M, \quad \bar{k}_m = k_m + k_{\mathrm{bs}}, \quad \bar{\beta}_m = \beta_m + \beta_{\mathrm{bs}}, \quad \bar{F}_0 = F_0 - k_{\mathrm{bs}} \cdot l_M,$$

with

$$\beta_m = 2\zeta_m\sqrt{k_m m_m}$$

Fluid Properties

The physical properties (elasticity modulus, density, sound speed, thermal expansivity, and kinematic viscosity) of the test-oil ISO 4113, that simulates the diesel fuel, are reported in Figs. 2–5 and 7. These compare the experimental data ([27]), reported by symbols, with the authors' model analytical relations (solid lines) used in the program. The properties are given as functions of the pressure p, taking the temperature T as a parameter.

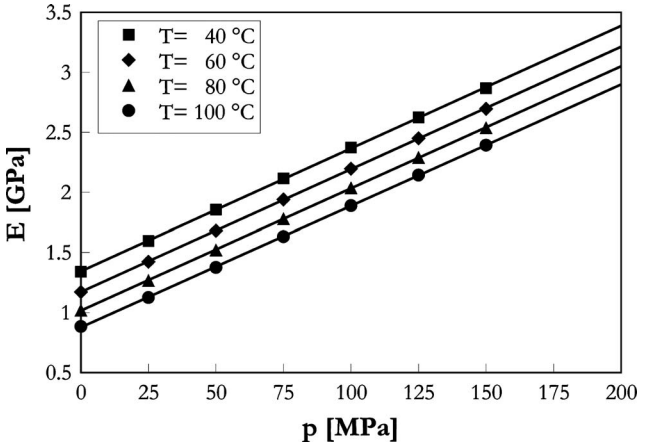


Fig. 2 Bulk modulus of elasticity: model Eq. (37) (solid lines) versus experimental data (symbols)

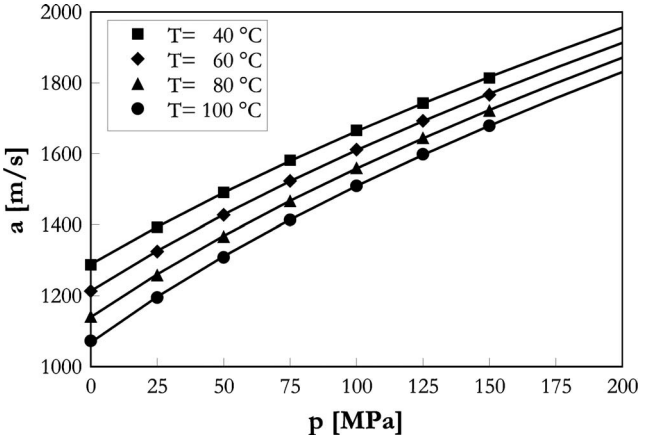


Fig. 5 Sound speed: Eq. (41) (solid lines) versus experimental data (symbols)

The bulk modulus of elasticity is very well described as a function of pressure and temperature by the following linear state equation:

$$E(p,T) = E_T(p) = E_0(T) + \chi(T)(p - p_0) \tag{37}$$

where p_0 is a reference pressure value and E_0 is the value of the elasticity modulus of the liquid at $p=p_0$. A plot of Eq. (37) and of experimental data at four temperatures in the range of interest for the present investigation is given by Fig. 2.

In order to obtain the relation between the oil density and pressure, we start from the standard definition of the elasticity bulk modulus of Eq. (28), which gives:

$$\left(\frac{\partial p}{\partial \rho}\right)_T = \frac{E_T(p)}{\rho} \tag{38}$$

A direct integration of Eq. (38) yields:

$$\rho(p,T) = \rho_T(p) = \rho_0(T)\left[\frac{E_T(p)}{E_0(T)}\right]^{1/\chi(T)} \tag{39}$$

where ρ_0 is the value of the density of the liquid for $p=p_0$, at the temperature T. The density distributions versus pressure are reported in Fig. 3 for the considered temperatures.

The volumetric coefficient of the liquid thermal expansion was determined by differentiation of the density (Eq. (39)) with respect to the temperature:

$$\beta(p,T) = -\frac{1}{\rho}\left(\frac{\partial \rho}{\partial T}\right)_p \tag{40}$$

A slight dependence of β on temperature can be observed in Fig. 4, plotting β versus pressure at 40°C and 100°C.

The sound speed as a function of pressure and temperature is given by (Fig. 5):

$$a(p,T) = a_T(p) = \sqrt{\frac{E_T(p)}{\rho}} \tag{41}$$

When cavitation occurs, without any undissolved gas, the pressure keeps constant at the vapor tension level, whereas the value of the mixture density is still variable and is calculated by the numerical algorithm solving the system given by Eq. (1). This value can be used to work out the void fraction:

$$\alpha = \frac{\rho_l - \rho}{\rho_l - \rho_v} \tag{42}$$

Thus, during cavitation the mixture compressibility is related to the phase change rather than to the elastic properties of the fluid.

Figure 6 shows the relation between pressure and density close to the saturation point, for a given temperature T.

It should to be pointed out that there is also a discontinuity in the sound speed at the saturation point. This is a consequence of the adopted model for the cavitating fluid, which is considered as a homogeneous mixture of a pure liquid and its vapor. More specifically, the sound speed drops to zero inside the cavitation regions (Eqs. (18) and (19)). Therefore, the perturbation waves

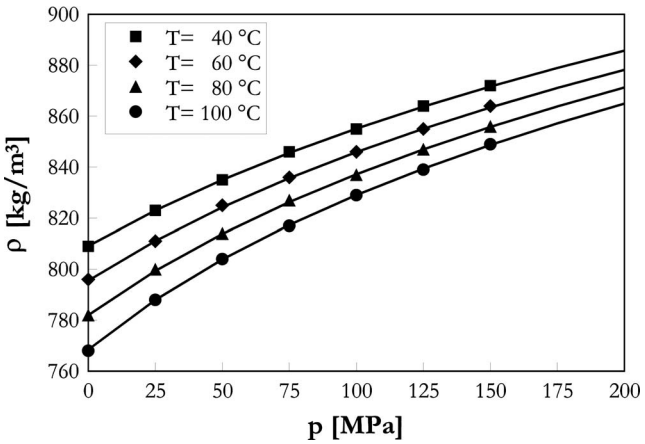


Fig. 3 Density: model Eq. (39) (solid lines) versus experimental data (symbols)

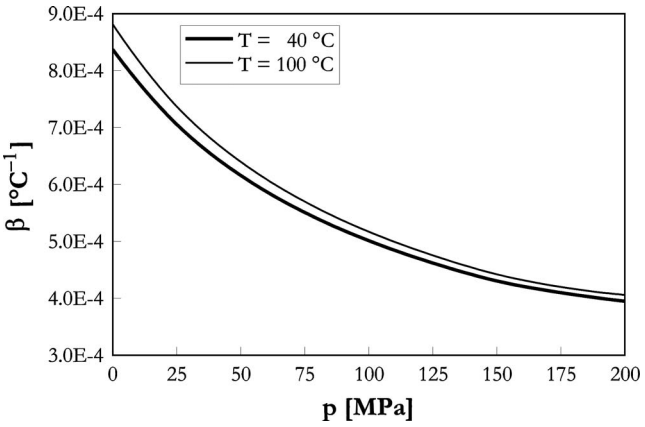


Fig. 4 Thermal expansivity

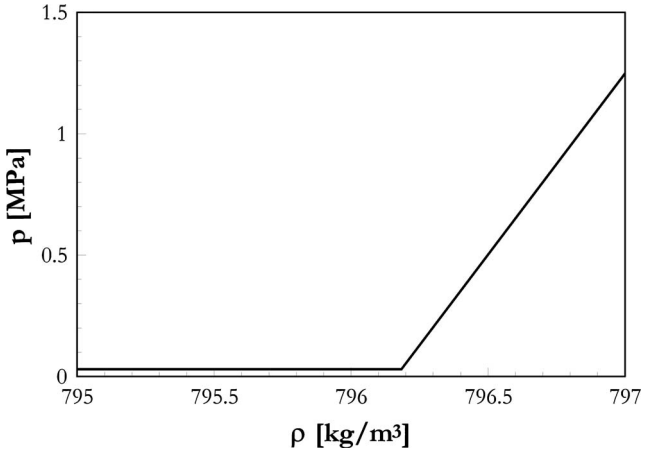


Fig. 6 Pressure as a function of density

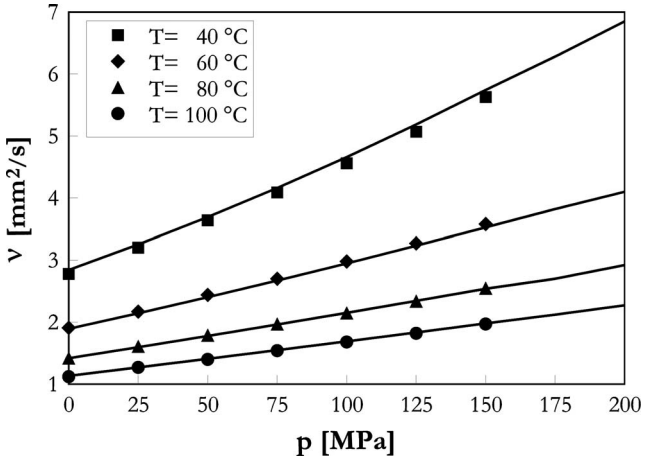


Fig. 7 Kinematic viscosity: model Eq. (43) (solid lines) versus experimental data (symbols)

propagate at the fluid velocity through the cavitation regions as long as the vapor component exists, according to the eigenvalues given by Eq. (11).

Finally, a good interpolation of the experimental data for the kinematic viscosity is given by the following quadratic polynomial (Fig. 7):

$$\nu(p,T) = k_0(T) + k_1(T)p + k_2(T)p^2 \tag{43}$$

Inside the cavitation region, the kinematic viscosity is held constant and equal to its value at the vapor pressure ([4]).

The analytical expressions of $\chi(T)$ and $E_0(T)$ in Eqs. (37) and (39), of $\rho_0(T)$ in Eq. (39) and of $k_0(T), k_1(T), k_2(T)$ in Eq. (43) are reported in [21].

Numerical Algorithm

A new finite-volume scheme was proposed for discretizing the hyperbolic partial differential system of Eq. (1). It is an implicit, conservative, one-step, symmetrical, and trapezoidal scheme of the second-order accuracy, which will also be referred to with the acronym ICOST (Implicit Conservative One-step Symmetric Trapezoidal) ([21,22]).

According to a finite-volume approach, the ICOST scheme can be written in the symmetrical conservative form given by the following expression:

$$\mathbf{w}_j^{n+1} = \mathbf{w}_j^n - \frac{\Delta t}{\Delta x}(\mathbf{\Phi}_{j+1/2}^{n+1/2} - \mathbf{\Phi}_{j-1/2}^{n+1/2}) + \Delta t\, \mathbf{H}_j^{n+1/2} \tag{44}$$

where the subscript j refers to the spatial grid locations, the superscript n designates the time level, $\mathbf{\Phi}$ is the numerical flux function defined as:

$$\mathbf{\Phi}_{j+1/2}^{n+1/2} = \frac{1}{2}(\mathbf{F}_{j+1}^n + \mathbf{F}_j^n) - \frac{1}{2}\left(\frac{\Delta t}{\Delta x}\right)[\mathbf{A}_{j+1/2}^{n+1}(\mathbf{F}_{j+1}^{n+1} - \mathbf{F}_j^{n+1})]$$

$$\mathbf{\Phi}_{j-1/2}^{n+1/2} = \frac{1}{2}(\mathbf{F}_j^n + \mathbf{F}_{j-1}^n) - \frac{1}{2}\left(\frac{\Delta t}{\Delta x}\right)[\mathbf{A}_{j-1/2}^{n+1}(\mathbf{F}_j^{n+1} - \mathbf{F}_{j-1}^{n+1})]$$

$\mathbf{A}$ is the Jacobian of the flux vector $\mathbf{F}=\mathbf{F}(\mathbf{w})$:

$$\mathbf{A} = \frac{d\mathbf{F}}{d\mathbf{w}} = \begin{bmatrix} 0 & 1 \\ a^2 - u^2 & 2u \end{bmatrix}$$

From the von Neumann stability analysis of the ICOST difference approximation to the wave equation, it follows that the proposed scheme can be considered to be unconditionally stable and therefore it is possible to select the mesh sizes Δt and Δx paying attention only to accuracy requirements. In order to prevent the generation of oscillations around discontinuities, second-order schemes usually introduce flux limiters ([28,29]), but the high-resolution schemes so obtained are more cumbersome to apply than are those without limiters (Figs. 8–15).

In the present work, the implementation of the high-resolution version of the ICOST method ([21]) does not involve differences in the simulation of the macroscopic effects with respect to the case without the limiters. In particular, as will be evidenced by the results (Figs. 11–15), in the presence of the flow discontinuities induced by cavitation inception and collapse, no significant oscillation problems are given by the ICOST scheme. On the other hand, available nonoscillatory schemes are either of the explicit type or too onerous to apply to injection systems, or are nonconservative ([8,30,31]). The necessity of a conservative numerical scheme is outlined in the Introduction and is examined in [20,29]. Besides, the choice of an implicit scheme is motivated by the following considerations.

For the discretization of the ordinary differential equations that model the pump and injector dynamics, BDF (Backward Differential Formulas) of the second-order accuracy were used, being suitable for problems of the stiff type ([3,5,8]). A second-order BDF approximation to the first-order time derivative of any time-function y, at the instant $n+1$, is given by:

$$\left.\frac{dy}{dt}\right|^{n+1} = \frac{1}{\Delta t}\left(\frac{1+2\xi}{1+\xi}y^{n+1} - (1+\xi)y^n + \frac{\xi^2}{1+\xi}y^{n-1}\right) \tag{45}$$

with $\xi=\Delta t/\Delta t'$ being $\Delta t=t^{n+1}-t^n$ and $\Delta t'=t^n-t^{n-1}$.

The difference analogues to the partial differential equations of the pipe flow and to the ordinary differential equations of the pump and injector led to a nonlinear algebraic equation system, which was linearized by the method of Newton–Raphson and solved by an efficient direct algorithm. At each step of time the primitive variables ρ, u, p were derived from the conservative variables $\rho, \rho u$, taken for discretization in both cases of the presence and the absence of cavitation. A robust, efficient, stable, and accurate numerical algorithm was thus attained, capable of essentially matching the nonstationary pipe fluid dynamics to the dynamics of the mechanical elements at pipe boundaries.

Results

Figures 8–10 illustrate the injection system performance at a pump speed of 1300 rpm at partial load. They report needle lift and pressure time-histories at two pipe locations, one close to the pump and the other close to the injector (Fig. 1).

The predicted distributions (solid line) are compared to the experimental results (circle symbol). The camshaft angle $(\theta-\theta_0)$, is reported as abscissa, θ_0 being a reference angle.

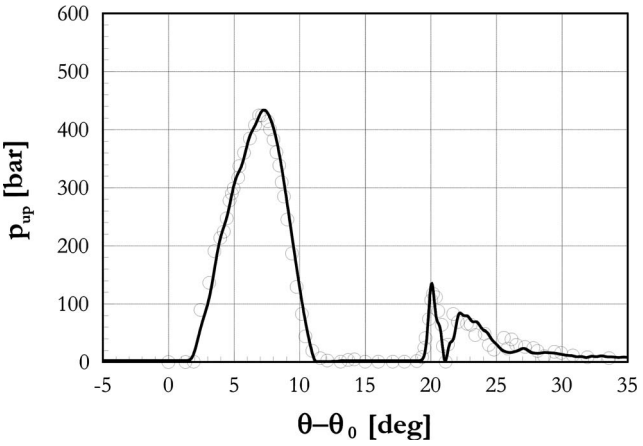


Fig. 8 Pipe pressure at the delivery outlet

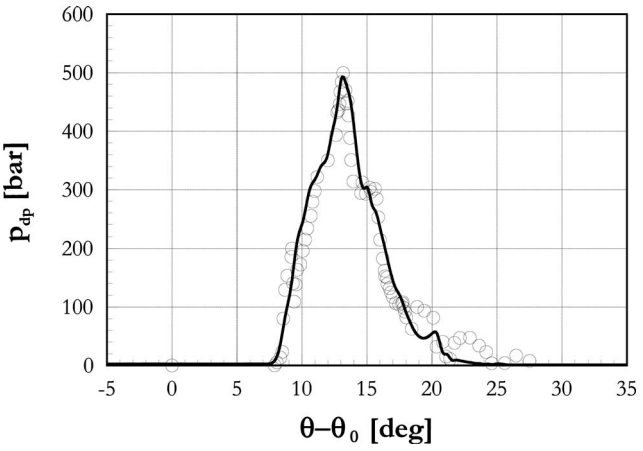


Fig. 9 Pipe pressure at the injector inlet

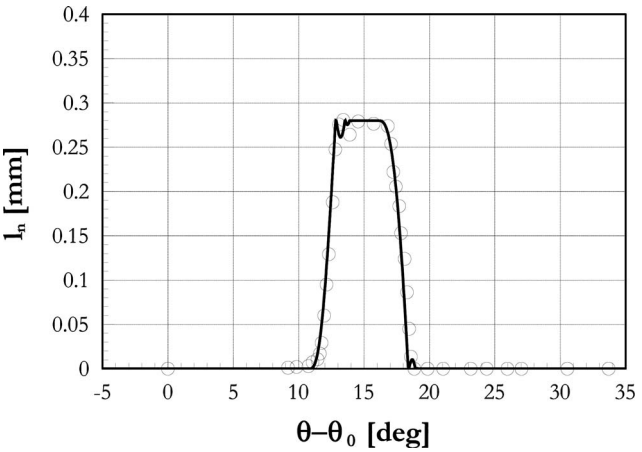


Fig. 10 Needle lift

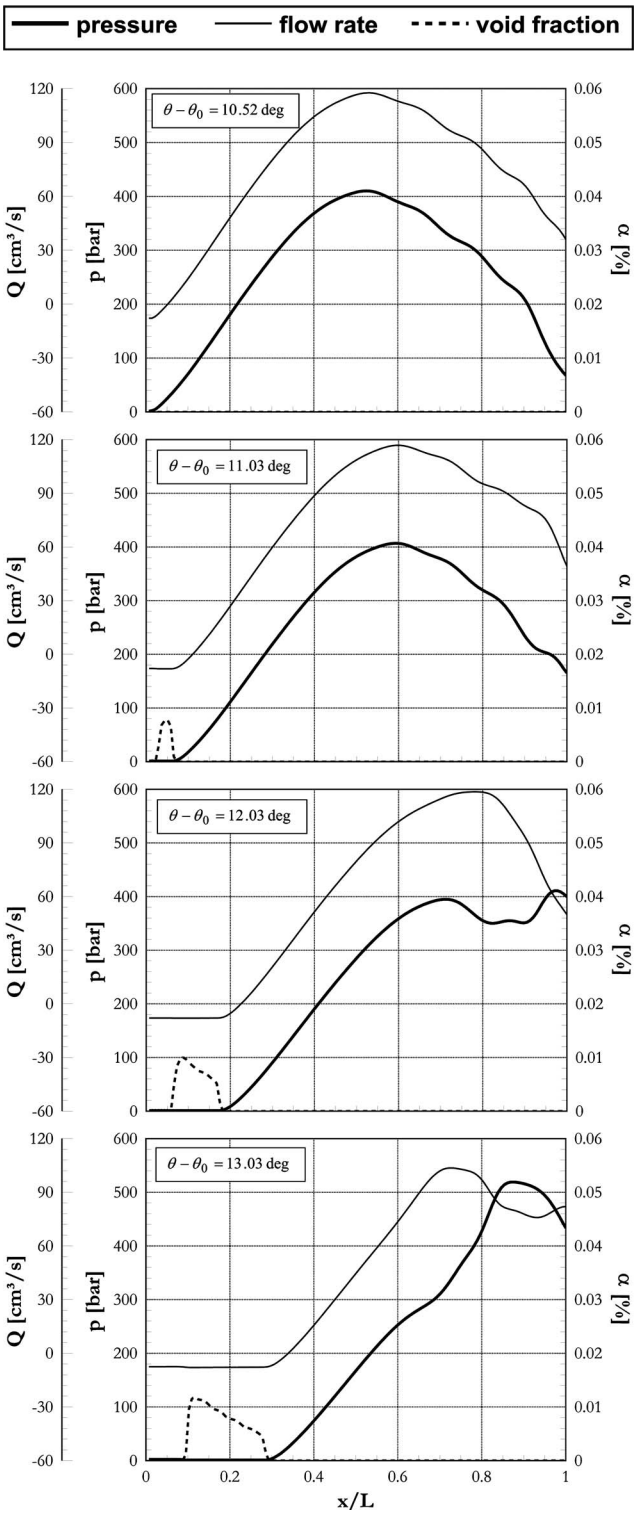


Fig. 11 Cavitation inception

The good agreement between computed and measured quantities shown by these results indicates a quite satisfactory degree of accuracy of the injection-system mathematical model.

Figure 8 indicates the occurrence of cavitation at the pump outlet. The fuel pressure reaches the value of the vapor tension at the cam angle $\theta-\theta_0 \approx 11$ deg and cavitation persists until the pressure wave coming back from the injector reaches the cavitating region, increasing the pressure there, so that the fluid leaves cavitation conditions at $\theta-\theta_0 \approx 19$ deg. As can be inferred from the figure, the model is capable of predicting the cavitation temporal duration and the intensity of the subsequent pressure peaks with high precision, substantiating the validity of the conservative model for simulating wave propagation phenomena in the presence of cavitation.

The pressure peak in Fig. 9 has a higher value than that of the previous figure because it is affected by an accumulation effect

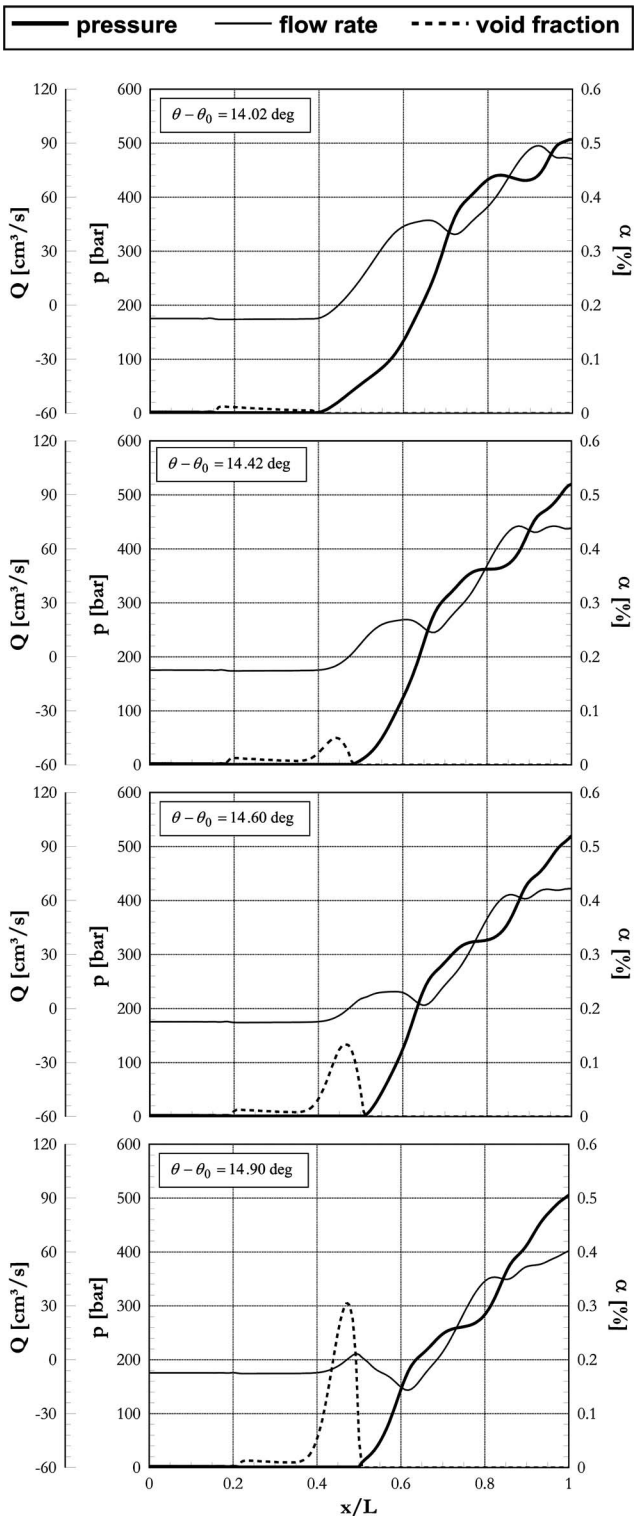


Fig. 12 Cavitation growth

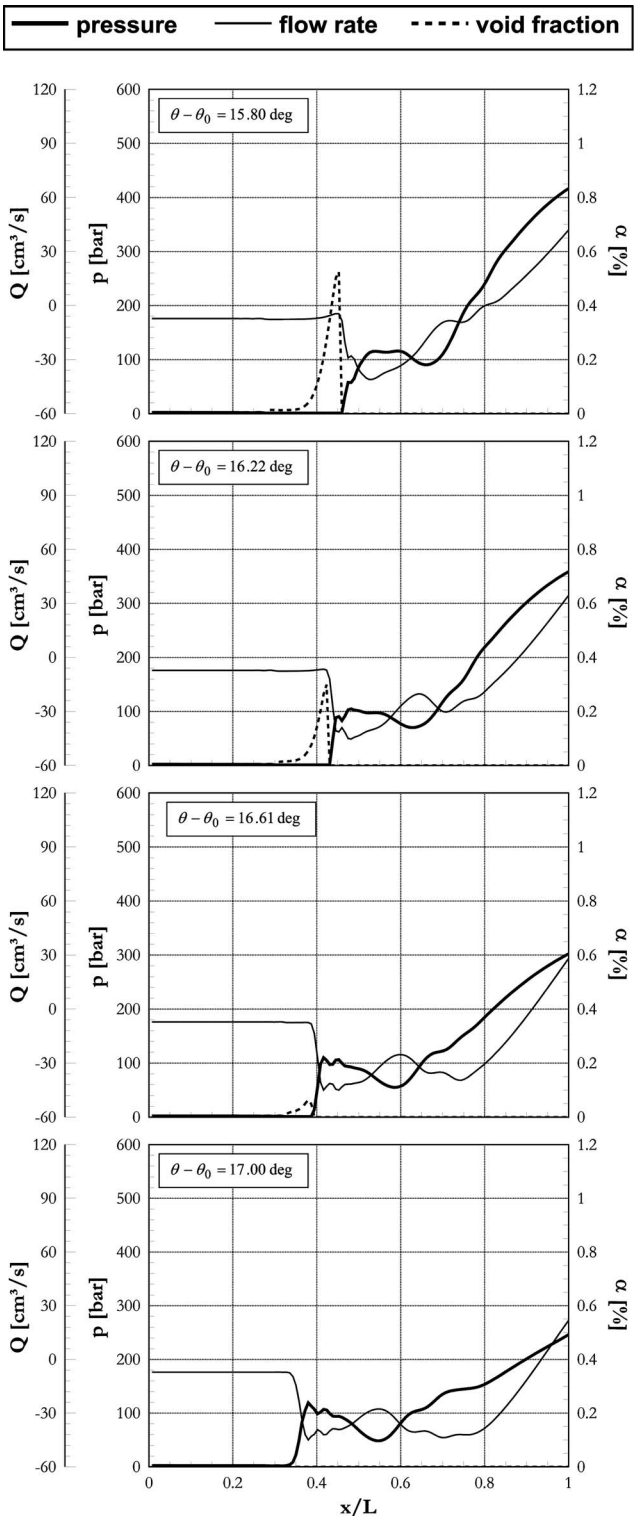


Fig. 13 Cavitation desinence

due to the reduced flow area that is present at the injector nozzle. Furthermore, owing to the wave propagation with the compressibility driven sound speed, the pressure peak shows a delay in time with respect to that of Fig. 8.

The result of Fig. 10 further validates the injector model, with the needle valve treated as a second-order linear system. The small oscillations that arise when the needle reaches its maximum lift and when the needle closes are due to the impact of the mobile element against the basement. Then, no reopening of the needle valve is observed after the injection end.

Figures 11–13, report significant sequences of the flow property distributions along the pipe illustrating the cavitation inception, growth and desinence stages at the abovementioned operating conditions. The abscissa reports the distance from the pump

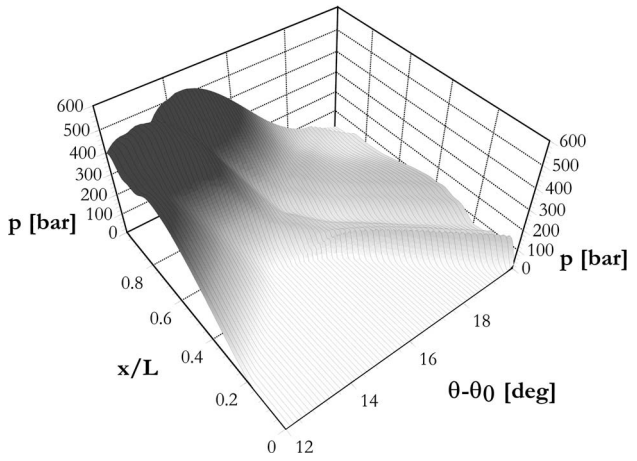


Fig. 14 Pressure along the pipe at different camshaft angles

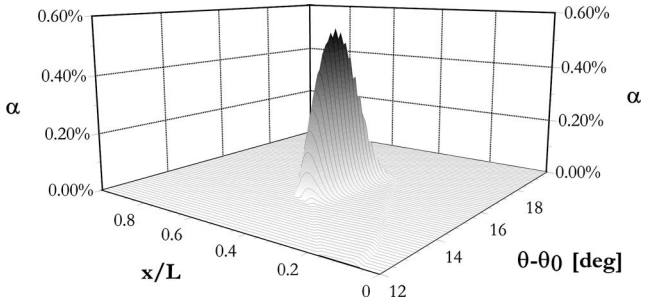


Fig. 15 Void fraction along the pipe at different camshaft angles

outlet, normalized to the length L of the connecting pipe and of the injector-drilled passage ($L \approx 1.1$ m). These figures show the pressure, void fraction, and flow rate patterns along the pipe at different instants of time corresponding to the indicated cam-angle coordinates $\theta - \theta_0$. The grid mesh sizes $\Delta x = 0.01$ m and $\Delta t = 10$ μs were employed for these predictions, requiring a computational time CPU of a few minutes on a Pentium 4 CPU, 2.4 Ghz, for a whole camshaft cycle.

The picture at the top in Fig. 11, for $\theta - \theta_0 = 10.52$ deg, represents the situation before the onset of cavitation, whereas the subsequent plots show the evolution of this phenomenon and the discontinuity induced by the cavitation in the flow properties. The rarefaction wave, which is responsible for the cavitation occurrence, is caused by the spill-port opening. The cavitation region travels toward the injector and enlarges with time, leaving behind a low-pressure noncavitating zone. The pressure values in this latter zone cannot be resolved because of the p-axis scale.

Figure 12 shows the subsequent evolution of the cavitation region, with a considerable growth of void fraction values while the above-mentioned rarefaction wave continues to travel toward the injector.

In Fig. 13, it can be observed how this rarefaction wave has been reversed in a pressure wave, sweeping away the formerly developed cavitation region. The figures clearly indicate a discontinuity surface, which separates cavitating from noncavitating regions. Such a discontinuity is recognizable as a shock, which travels with a sensibly reduced speed with respect to the sound speed of the liquid because the fluid properties in the cavitation zone are affected by the vapor–liquid phase change. In fact, the sound speed drops from the value a_l to zero across the flow discontinuity and therefore the supersonic flow existing inside the cavitating region becomes subsonic in the pure liquid zone. Thus a shock arises. On the other hand, in Fig. 11 there is no evidence of any shock occurrence because the subsonic flow at one pipe location becomes supersonic when the same location is reached by cavitation. Therefore, a rarefaction wave can be observed instead of a shock.

In order to allow a more complete overall view of wave propagation as well as of cavitation inception and evolution, Figs. 14 and 15 give a three-dimensional representation of the computed distributions of the pressure and void fraction along the pipe, in a significant range of cam angle, for the same operating conditions of Figs. 8–10. The x axis reports the non-dimensional distance along the pipe, the y axis the camshaft angles, and the z axis the pressure (Fig. 14) and the void-fraction distributions (Fig. 15). The lines drawn at constant values of the cam angle $\theta - \theta_0$ represent the pressure and void fraction distributions along the pipe at fixed instants of time. As is shown by these figures, the traveling cavitation region, occurring downstream to the delivery valve after its closure, is particularly diffuse and relatively high void-fraction values are reached along the pipe. Even for such cavitating conditions, that are considered to be severe for high-pressure injection systems, the numerical algorithm shows a stable behavior and does not give rise to significant oscillation problems.

Conclusion

A new conservative numerical model of general application for hydraulic-system unsteady fluid-dynamics simulation, was developed and assessed with specific reference to high-pressure injection systems. The thermodynamic approach to acoustic cavitation modeling that is based on a homogeneous two-phase barotropic flow model of a pure liquid, its vapor and a gas, both dissolved and undissolved, led to a consistent comprehensive sound-speed analytical form.

For the liquid flow simulation, the temperature variations due to compressibility effects were included, for the first time in injection system modeling. The thermal-energy conservation equation was applied in connection with an isentropic evolution consistently either with the hypothesis of an adiabatic flow and negligible friction-loss thermal effects or with the assumption that the heat produced by friction is fully transferred outside through the pipe wall.

To simulate the cavitation evolution, an isothermal process was chosen because it is in accordance with negligible temperature variations, due to the tiny vaporizing-liquid amount and also because it is very simple to handle. Furthermore, the approach of considering pure vaporous cavitation was shown to yield physically consistent macroscopic results by the very good agreement between the predicted and experimental pressure time histories at the considered pipe measurement locations.

The new second-order accurate conservative numerical scheme, which was applied for discretizing the partial differential equations governing the pipe flow, was shown to be robust in accurately predicting the wave propagation phenomena, and more specifically the flow discontinuities induced by the cavitation inception and desinence, with minor oscillation problems.

The mathematical model uses the same set of conservation equations both in the presence and in the absence of cavitation.

The numerical analysis of the acoustic cavitation transients in the pressure pipe indicated that the cavitation inception and growth were induced by rarefaction waves, which gave rise to a transition from subsonic to supersonic flows, whereas the cavitation decrease and desinence were caused by compression waves leading to a transition from supersonic to subsonic flows with shock occurrence.

Acknowledgment

Financial support to this research was provided by Fiat Research Center, Fiat-GM Powertrain and by MIUR (Ministero dell'Istruzione, dell'Università e della Ricerca) under the Cofin 2000 Program "Thermo-Fluid Dynamic Processes and Techniques for pollutant Emissions Reduction in IC Engines."

Nomenclature

a = sound speed
$\mathbf{A}$ = Jacobian matrix
c = specific heat
d = pipe diameter
E = bulk modulus of elasticity
F = spring force
$\mathbf{F}$ = flux vector
h = enthalpy per unit mass
$\mathbf{H}$ = source vector
k = spring rate
l = lift
$\dot{l}_w$ = viscous power dissipation per unit mass
L = length of pipe and injector-drilled passage
m = mass; mixture mass
p = pressure; average cross-sectional pressure
$\dot{q}$ = heat transfer rate per unit mass
Q = volumetric flow rate
s = entropy per unit mass
S = surface area perpendicular to the lift axis
t = time
T = temperature
u = average cross-sectional velocity
v = specific volume; mixture specific volume
V = volume; mixture volume
$\mathbf{w}$ = conservative variable vector
x = axial distance along pipe
α = void fraction; average cross-sectional void fraction
β = thermal expansivity; damping
Δt = time increment
Δx = distance increment
ζ = damping factor
θ = cam angle
μ = mass fraction; average cross-sectional mass fraction
ρ = density; average cross-sectional density; mixture density
τ_w = wall shear stress
$\mathbf{\Phi}$ = numerical flux vector

Subscripts

0 = reference value
bs = seats
dg = dissolved gas
dp = downstream-pipe location
g = gas; undissolved gas
l = liquid
m = mobile element
n = needle valve
M = maximum value
$p/T/s$ = at constant pressure/temperature/entropy per unit mass
up = upstream-pipe location
v = vapor; at constant volume
x = gas (g) or vapor (v)

Appendix A

Mixture Density. From the application of the primitive definition of specific volume v to the mixture, made up in general of the liquid fuel, its vapor and a quantity of gas, both dissolved and undissolved, one has:

$$v = \frac{1}{\rho} = \frac{V}{m} = \frac{V_v + V_g + V_l}{m} \tag{A1}$$

m is the mass of the mixture occupying the volume V, V_l is the volume of the liquid portion held in V, including the dissolved gas, and V_v, V_g are, respectively, the vapor and undissolved gas volumes.

Introducing the specific volumes of the liquid, vapor, and undissolved gas:

$$v_l = \frac{V_l}{m_l}, \quad v_v = \frac{V_v}{m_v}, \quad v_g = \frac{V_g}{m_g} \tag{A2}$$

where m_l, m_v, m_g refer to the liquid, vapor and undissolved gas masses held in the volume V (the sum of m_l, m_v, m_g is equal m).

Equations (A1) and (A2), taking the definition of vapor and undissolved-gas mass fractions into account (Eq. (3)), yield:

$$v = \mu_v v_v + \mu_g v_g + (1 - \mu_v - \mu_g) v_l \tag{A3}$$

Since the density is, by definition, the inverse of the specific volume, from Eq. (A3) one has:

$$\frac{1}{\rho} = \frac{\mu_v}{\rho_v} + \frac{\mu_g}{\rho_g} + \frac{(1 - \mu_v - \mu_g)}{\rho_l} \tag{A4}$$

from which the expression of the mixture density (Eq. (4)) can be easily derived.

Appendix B

Effective Bulk Elasticity Modulus of the Mixture. If a fixed mass of vapor-gas-liquid mixture is considered, its volume change in an infinitesimal interval of time can be evaluated according to the following contributions taking elasticity, vaporization or condensation and gas entrainment or dissolution into account:

$$dV = dV_v + dV_g + dV_l + dV_{\mathrm{pc}} + dV_{\mathrm{ed}} \tag{B1}$$

The first three terms (dV_v, dV_g, dV_l) on the right-hand side of Eq. (B1) are related to the elasticity properties of the mixture constituents, that is, vapor, gas and liquid, respectively:

$$dV_v = -V_v \frac{dp}{E_v} = -V\alpha_v \frac{dp}{E_v} \tag{B2}$$

$$dV_g = -V_g \frac{dp}{E_g} = -V\alpha_g \frac{dp}{E_g} \tag{B3}$$

$$dV_l = -V_l \frac{dp}{E_l} = -V(1 - \alpha_v - \alpha_g) \frac{dp}{E_l} \tag{B4}$$

The term dV_{pc} (the subscript stands for phase change) has to be ascribed to the vaporization or condensation of a portion of the liquid or vapor, respectively:

$$dV_{\mathrm{pc}} = \left(\frac{1}{\rho_v} - \frac{1}{\rho_l}\right) dm_v = \rho V \left(\frac{1}{\rho_v} - \frac{1}{\rho_l}\right) d\mu_v \tag{B5}$$

in which dm_v is the mass that either evaporates or condensates.

The last term in Eq. (B1), that is dV_{ed}, refers to the gas entrainment or dissolution (as the subscript suggests):

$$dV_{\mathrm{ed}} = \left(\frac{1}{\rho_g} - \frac{1}{\rho_l}\right) dm_g = \rho V \left(\frac{1}{\rho_g} - \frac{1}{\rho_l}\right) d\mu_g, \tag{B6}$$

where dm_g indicates the mass of gas that either comes out of solution to form gas bubbles (entrainment), or dissolves in the liquid.

Therefore, the relative volume change is given by:

$$\frac{dV}{V} = -\frac{\alpha_v}{E_v} dp - \frac{\alpha_g}{E_g} dp - \frac{(1 - \alpha_v - \alpha_g)}{E_l} dp + \rho \left(\frac{1}{\rho_v} - \frac{1}{\rho_l}\right) d\mu_v + \rho \left(\frac{1}{\rho_g} - \frac{1}{\rho_l}\right) d\mu_g, \tag{B7}$$

so that an effective bulk modulus of the homogeneous mixture can be defined by the follows inverse expression (Eq. (14)):

$$\frac{1}{E_{\text{eff}}} = \frac{\alpha_v}{E_v} + \frac{\alpha_g}{E_g} + \frac{(1-\alpha_v-\alpha_g)}{E_l} + \rho\left(\frac{1}{\rho_v} - \frac{1}{\rho_l}\right)\frac{d\mu_v}{dp} + \rho\left(\frac{1}{\rho_g} - \frac{1}{\rho_l}\right)\frac{d\mu_g}{dp} \tag{B8}$$

From this relation and Eq. (13), it follows that:

$$a_v = \sqrt{\frac{E_v}{\rho_v}}, \quad a_g = \sqrt{\frac{E_g}{\rho_g}}, \quad a_l = \sqrt{\frac{E_l}{\rho_l}}, \quad \text{and } a = \sqrt{\frac{E_{\text{eff}}}{\rho}}$$

References

[1] Baltzer, R. A., 1967, "Column Separation Accompanying Liquid Transients in Pipes," ASME J. Basic Eng., **89**, pp. 837–846.

[2] Chaudhry, M. H., Bhallamudi, S. M., Martin, C. S., and Naghash, M., 1990, "Analysis of Transient Pressures in Bubbly, Homogeneous, Gas-Liquid Mixtures," ASME J. Fluids Eng., **112**, pp. 225–231.

[3] Catania, A. E., Dongiovanni, C., and Mittica, A., 1992, "Implicit Numerical Model of a High-Pressure Injection System," ASME J. Eng. Gas Turbines Power, **114**, pp. 534–543.

[4] Shu, J.-J., Edge, K. A., Burrows, C. R., and Xiao, S., 1993, "Transmission Line with Vaporous Cavitation," ASME Paper No. 93-WA/FPST-2.

[5] Catania, A. E., Dongiovanni, C., Mittica, A., Spessa, E., and Lovisolo, F., 1994, "Study of Unsteady Flow Phenomena in an Automotive Diesel Injection System," *Proceedings of the XXV FISITA Congress*, International Academic Publishers, Beijing, **1**, pp. 124–137.

[6] Hapke, I., and Iben, H., 1998, "Modellierung und Berechnung der Kavitation in Instationären Leitungsströmungen," Motortech. Z., **1**, pp. 60–64.

[7] Nguyen-Schaefer, H., and Sprafke, H., 1998, "Numerical Study on Interaction Effects of Bubbles Induced by Air-Release and Cavitation in Hydraulic Systems," *10th Bath International Fluid Power Workshop*, Research Studies Press, Wiley, New York.

[8] Catania, A. E., Dongiovanni, C., and Spessa, E., 2000, "Delivery-Valve Effects on the Performance of an Automotive Diesel Fuel-Injection System," SAE Trans., **108**, pp. 1399–1415.

[9] Beck, M., Iben, U., Mittwollen, N., Iben, H.-K., and Munz, C.-D., 2001, "On Solution of Conservation Equations in Cavitated Hydraulic Pipelines," *3rd International Symposium on Computational Technologies for Fluid/Thermal/Chemical Systems with Industrial Applications*, July 22–26, Atlanta, GA.

[10] Iben, U., Wrona, F., Munz, C. D., and Beck, M., 2002, "Cavitation in Hydraulic Tools Based on Thermodynamic Properties of Liquid and Gas," ASME J. Fluids Eng., **124**, pp. 1011–1017.

[11] Matsuoka, S., Yokota, K., Kamimoto, T., and Igoshi, M., 1976, "A Study of Fuel Injection Systems in Diesel Engines," SAE Paper No. 760551.

[12] Wallis, W., 1975, "*One Dimensional Two-Phase Flow*," McGraw Hill, New York.

[13] Kolev, N. I., 1986, "*Transiente Zweiphasen Strömung*," Springer-Verlag, Berlin.

[14] Iben, U., 2000, "Cavitation Modeling," *Analysis, Modeling and Simulation*, Gordon and Breach, London.

[15] Date, K., Manabe, M., Kano, H., Kato, M., and Oya, T., 1992, "Contribution of Fuel Flow Improvement in Nozzle to Spray Formation," SAE Paper No. 920622.

[16] Kalkwijk, J. P. T., and Kranenburg, C., 1971, "Cavitation in Horizontal Pipelines Due to Water Hammer," ASCE Journal of Hydraulics Division, **97**, pp. 1585–1605.

[17] Streeter, V. L., 1969, "Water Hammer Analysis," ASCE Journal of Hydraulics Division, **95**, pp. 1959–1972.

[18] Kumar, K., Gajendra Babu, M. K., Gaur, R. R., and Garg, R. D., 1983, "A Finite Difference Scheme for the Simulation of a Fuel Injection System," SAE Paper No. 831337.

[19] Anderson, D., Tannehill, J. C., and Pletcher, R. H., 1984, "*Computational Fluid Mechanics and Heat Transfer*," McGraw-Hill, New York.

[20] Ferrari, A., Manno, M., and Mittica, A., 2004, "Comparison Between Conservative and Nonconservative Models in the Simulation of High-Pressure Injection System Under Cavitating Conditions," CD ROM Proceedings of *7th Biennal ASME Conference Engineering System Design and Analysis*, July 19–22, Manchester, UK.

[21] Ferrari, A., 2004, "Development of a Model for Thermo-Fluid Dynamic Transient Simulation in High-Pressure Injection Systems. Equipment of a High-Performance Test Bench for Diesel Fuel Injection Systems: First Experimental Results on Common Rail System Dynamics," (in Italian), Ph.D. thesis, Politecnico di Torino.

[22] Manno, M., 2004, "Equipment of a High-Performance Test Bench for Diesel Fuel Injection-Systems: First Experimental Results on Common Rail System Dynamics. Development of Numerical Methods for High-Pressure Injection System Flow Transients," (in Italian), Ph.D. thesis, Politecnico di Torino.

[23] Bejan, A., 1997, "*Advanced Engineering Thermodynamics*," Wiley, New York.

[24] Atkins, P. W., "*Physical Chemistry*," 1986, Oxford University Press, Oxford.

[25] Catania, A. E., Ferrari, A., Manno, M., and Spessa, E., 2004, "Thermal Effect Simulation in High-Pressure Injection System Transient Flows," *SAE SP-1824 "Diesel Fuel Injection and Sprays*," pp. 44–59, SAE Paper No. 2004-01-0532, *SAE 2004 World Congress*, Detroit, MI, March 8–12.

[26] Bosch, R., GmbH, 1999, "*Diesel-Engine Management*," Bosch Technical Books, SAE international.

[27] Bosch, R., GmbH, private communication.

[28] Hirsch, C., 1988, *Numerical Computation of Internal and External Flows, Vol. 2—Computational Methods for Inviscid and Viscous Flows*, Wiley, New York.

[29] LeVeque, R. J., 1990, *Numerical Methods for Conservation Laws*, Birkhäuser-Verlag, Berlin.

[30] Chaudhry, M. H., and Hussaini, M. Y., 1985, "Second-Order Accurate Explicit Finite-Difference Schemes for Waterhammer Analysis," ASME J. Fluids Eng., **107**, pp. 523–529.

[31] Harten, A., Engquist, B., Osher, S., and Chakravarthy, S. R., 1987, "Uniformly High Order Accurate Essentially Non-Oscillatory Schemes, III," J. Comput. Phys., **71**, pp. 231–303.

Effects of Heat Release Mode on Emissions and Efficiencies of a Compound Diesel Homogeneous Charge Compression Ignition Combustion Engine

Wanhua Su
e-mail: whsu@tiu.edu.cn

Xiaoyu Zhang

Tiejian Lin

Yiqiang Pei

Hua Zhao

State Key Laboratory of Engines,
Tianjin University,
Tianjin 300072, P.R.C.

A compound diesel homogeneous charge compression ignition (HCCI) combustion system has been developed based on the combined combustion strategies of multiple injection strategy and a mixing enhanced combustion chamber design. In this work, a STAR-CD based, multidimensional modeling is conducted to understand and optimize the multiple injection processes. The parameters explored included injection timing, dwell time, and pulse width. Insight generated from this study provides guidelines on designing the multipulse injection rate pattern for optimization of fuel-air mixing. Various heat release modes created by different injection strategies are investigated by experimental comparison of combustion efficiency, heat loss, and thermal efficiency. It is demonstrated that the process of fuel evaporation and mixing are strongly influenced by pulse injection parameters. Through control of the parameters, the stratification and autoignition of the premixed mixture, and the heat release mode can be controlled. The dispersed mode of heat release created only by the compound diesel HCCI combustion is a flexible mode in combustion control. The thermal efficiency with this mode can reach approximately to as high as that of conventional diesel combustion, while the NO_x and smoke emissions can be reduced simultaneously and remarkably. [DOI: 10.1115/1.2032447]

Introduction

In order to meet increasingly stringent emission legislation and demand of energy saving, homogeneous charge compression ignition (HCCI) combustion has become an attractive technology of IC engines because of the potential to significantly reduce NO_x and soot emissions, while achieving high thermal efficiency [1]. For diesel-fueled HCCI engines, high compression ratio can be used to enhance autoignition at light load, and direct in-cylinder fuel injection can be used to control combustion timing. However, the low-temperature chemistry of diesel fuel tends to cause a rapid autoignition once compression temperature exceeds a threshold, thereby leading to an overly advanced combustion phasing. In addition, the low volatility of diesel fuel together with its propensity of autoignition attributed to high fuel cetane number makes it extremely difficult to form a homogeneous mixture required to explore the full benefit in simultaneous reduction of NO_x and soot emissions. Most already proposed diesel HCCI engines only require one fuel handling system, however, its high fuel injection pressure causes inevitable wall wetting, which is the source for increased unburned hydrocarbons and possible soot emissions. One fuel injection technology that can help homogeneous mixture formation without fuel wall impingement is in active demand.

Over the last few years, a lot of research on diesel HCCI combustion has been reported. They may be separated into three main categories, depending on fuelling techniques: (1) premixed HCCI, which involves port fuel injection [2–7]; (2) early direct-injection HCCI, in which fuel is injected sufficiently early during the compression stroke so that mixing can be completed prior to ignition [8–15]; (3) later injection HCCI, in which fuel is injected later than conventional operation, combined with high swirl levels for fast mixing and exhaust gas recirculation (EGR) for extended ignition delay [16,17].

In addition, HCCI combustion is normally limited to the low and medium load operations by the excessive rate of the pressure rise due to simultaneous heat release of the whole charge in the cylinder. Conventional diesel combustion is therefore still required at higher load conditions. The engine fuel handling system, therefore, should be well suitable in both conditions of HCCI combustion and conventional diesel combustion. Our previous work has presented a new HCCI combustion technology [15,18–20], which operates the engine with HCCI combustion at low loads, combined HCCI combustion and lean diffusion combustion at medium and high loads. Because of the engine characteristics of the multiinjection and the bump combustion chamber design, it is named as MULINBUMP system. In this study, numerical simulations are performed in order to understand and assist the optimization of the multiple injection process. Then, the experimental results are presented to discuss the combustion control and its effect on emissions and thermal efficiency.

MULINBUMP Combustion System

The test engine, as shown in Fig. 1, is based on a heavy-duty inline six-cylinder truck engine. In this study, the engine is modified so that the cylinder 1 is used as the test cylinder and the other five cylinders are used as a motor to drive the test cylinder. The test cylinder is equipped with the flexible injection rate common rail injector (FIRCRI) common rail fuel injection system [21], a separate air supply system and a separate exhaust system. The air supply system comprises an air compressor, a surge tank, an intake air filter, and a pressure control valve. A gate valve in the exhaust pipe of cylinder 1 is used to control the exhaust backpressure, which is normally caused by a turbocharger. Table 1 shows

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received March 3, 2004; final manuscript received October 14, 2005. Review conducted by D. Assanis.

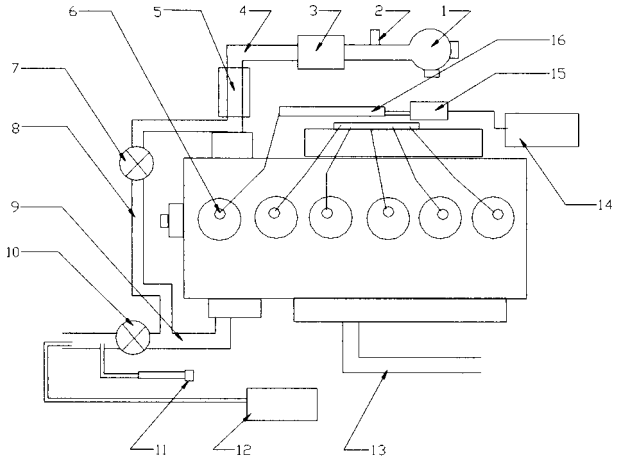


Fig. 1 Schematic of experimental setup

the engine specifications and Table 2 gives the specifications of the fuel injector. The typical experimental conditions are summarized in Table 3.

For the MULINBUMP Compound HCCI combustion, the novel injection strategy developed bases on the FIRCRI common rail fuel injection system, by which the injection is split up to ten pulse injections followed by a main injection at a predetermined timing around the engine top dead center (TDC). The pulse width, pulse number, dwell time, and the start timing are flexibly controlled for the formation of homogeneous mixture, combustion phasing, and rate of heat release. An underlying advantage of the multi-injection strategy is capable of preventing fuel spray impingement on the cylinder liner owing to short pulse width and disturbance by the consecutive injections.

Table 1 Specifications of the test engine

Engine type	Steyr WD 615, DI diesel
Bore	126 mm
Stroke	130 mm
Connecting road length	219 mm
Combustion chamber	BUMP; w/o BUMP
Piston bowl volume	85.3 cm^3
Squish height	1 mm
Compression ratio	15
Swirl ratio	1.8

Table 2 Specifications of the fuel injector

Type	FIRCRI common rail
Common rail pressure	80 MPa; 110 MPa
Number of injector holes	7
Nozzle hole diameter	0.23 mm
Length/diameter of the holes (l/d)	4.35
Injector's cone angle	150 deg
Injection modes	Multipulse injection Main injection

Results and Discussions

CFD Simulation of Mixing Process of the Multiple Injections. For better understanding of the roles of multiple injection parameters, the multidimensional CFD code STAR-CD is employed to understand and optimize the mixing process during the multiple injections. The numerical model incorporates the conservation of mass, momentum and energy. The small-scale turbulence effects are modeled indirectly within the Reynolds averaging framework, using renormalization group (RNG) κ-ε turbulence model with associated wall functions [22,23]. The major submodels relating to the spray process are the Huh atomization model [24], the Reitz-Diwakar breakup model [25,26], and the Bai spray impingement model [27]. All other associated models are left defaults. Details of these models can be found in the literature [28].

Due to the different sizes of the intake and exhaust valve, the seven-hole nozzle and the piston bowl in the real engine have an eccentricity from the cylinder axis. Since the intake flow and the valves are not simulated, the eccentricity is not taken into account for computational efficiency. That allows simulating only a sector of 51 deg according to one-seventh of the geometry. The numerical mesh at TDC is shown in Fig. 2. The numbers of grid cells at TDC are 34 in radial, 17 in azimuthal, and 13 in the axial direction. Three cell layers are left in the squish region at TDC. Because of the inherent restriction of droplet discrete model (DDM) spray model, the simulations are sensitive to mesh refinement. The applied mesh resolution has therefore been chosen such that it is close to the spray model optimized mesh resolution suggested by Gonzalez et al. [29] and Abraham [30].

The computations are started at the time of intake valve closure (IVC=154 deg BTDC). The average in-cylinder pressure and temperature at the beginning of the calculation are determined by running a cycle analysis program. The surface temperatures of the piston, cylinder head and liner are estimated as 525, 500, and 475 K, respectively. The computations used *n*-dodecane as the fuel because its properties are similar to those of the used fuel. The injection rate histories are given in Fig. 3. Based on the actual profile of the common rail system, injection profile 1, two designed injection profiles are presented, which are named injection profile 2 and injection profile 3. The dwell time is decreased from 8.4 deg CA to 5 deg CA for injection profile 2 and the pulse width is further decreased from 5.38 deg CA to 4.39 deg CA for injection profile 3 when the engine operates at 1400 rpm. Note that in this CFD study, the amount of total fuel injected is kept fixed at 31.4 mg.

As the fuel is injected into the cylinder, it mixes with the in-cylinder air. In order to assess the effectiveness of mixing, cell fuel-air equivalence ratio (ϕ) is computed at every instant according to the amount of fuel and air present in a given cell. Based on the computations, the evaporated fuel is divided into three parts: rich, flammable and lean by the intervals, $\phi<0.3$, $0.3<\phi<1.2$, and $\phi>1.2$. Note that these values are arbitrary, but as long as they are consistently employed for all cases, they would indicate useful trends. Furthermore, it is important to recognize that the reported mixing results do not account for the combustion process. Consequently, they should be treated as indicative of the effects of injection parameters on mixing rather than as exact figures for the cases under investigation.

Effect of Injection Timing. The injection timing is one of the most important control variables in controllable premixed charge compression ignition (CPCCI) combustion mode. In order to examine the effect of injection timing on the premixed charge formation, the present parametric study examines three different injection timing scenarios, one being the baseline case (90 deg BTDC) and the others being injecting earlier (110 deg BTDC) and later (70 deg BTDC). Figure 4 shows the results of the simulations for injection profile 1. As shown in the figure, four parameters are taken as indicating of evaporation, mixing, and stratification. One

Table 3 Engine experiments and operating conditions

Case no.	IMEP (kPa)	Intake pressure (kPa) P_{in}	Intake temperature (K) T_{in}	Multiple injection timing (deg. BTDC) θ_p	Number of injections N	Percentage of multiple injections	Pulse width (ms) Δt_1	Dwell between pulses (ms) Δt_2	Main injection timing (deg. BTDC) θ_m	Common rail pressure (MPa) P_{cr}	Combustion chamber
A1	320	120	303	70	4	100%	0.14	1		80	BUMP-V
A2	320	120	303	90	4	100%	0.14	1		80	BUMP-V
B3	670	120	303	60	3	40.5%	0.14	1	−3	80	BUMP-V
B4	670	120	303	70	3	40.5%	0.14	1	−3	80	BUMP-V
B5	670	120	303	90	3	40.5%	0.14	1	−3	80	BUMP-V
B6	670	120	303	110	3	40.5%	0.14	1	−3	80	BUMP-V
C7	930	130	303	110	4	53%	0.19	1.5	−3	80	BUMP-V
C8	930	130	303	110	4	53%	0.19	1.5	−8	80	BUMP-V
C9	930	130	303	110	4	47%	0.19	1.5	−10	80	BUMP-V
D10	930	130	303	110	4	53%	0.19	1.5	−8	80	BUMP-V
D11	930	190	303	220	6	55%	0.14	1.15	−8	80	BUMP-V
D12	930	150	303			0%			−8	110	BUMP-V
D13	930	150	303			0%			8	110	BUMP-V
D14	930	150	303			0%			−8	110	w/o BUMP
D15	930	150	303			0%			8	110	w/o BUMP

is the evaporation fraction, which is defined as the ratio of total evaporated mass to the total injected mass at the current time step. The other is a mass fraction, which is the ratio of the fuel mass in lean, flammable, or rich mixture part to the total evaporated mass, respectively.

Under the selected operating conditions, the results indicate two quite distinct features from the general knowledge of conventional engine operation. First, in the conventional one, as the fuel is injected, a rich fuel core zone is formed and continues to build up throughout the injection process. At the same time, as the jet plume penetrates into the cylinder and entrains air, its periphery is progressively being leaned out giving birth to flammable regions.

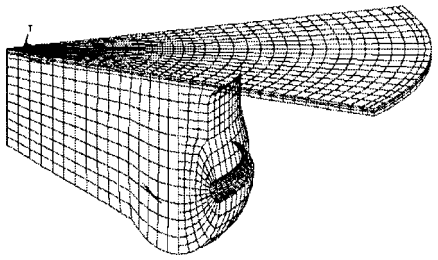


Fig. 2 Computational grid at TDC used for engine calculations

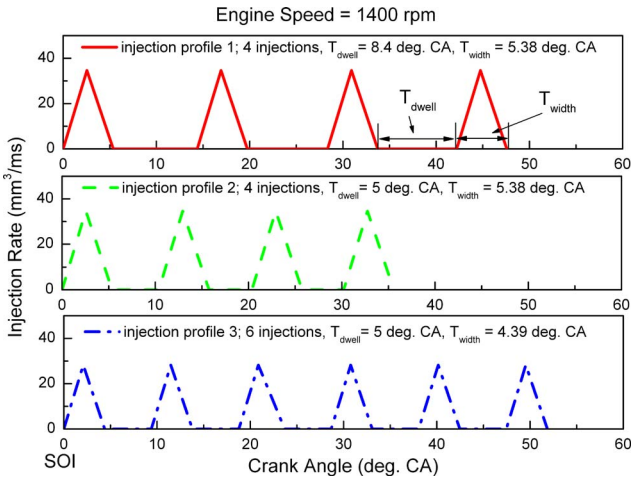


Fig. 3 Multiple injection profiles

However, it is not true in the early pulse injection case, as shown in Fig. 4. In fact, it turns out that evaporation would be substantially more efficient when properly retarded injection timing is used. The flammable regime is formed around 70 deg BTDC for the baseline and 60 deg BTDC for later injection case. For all cases, the rich fuel core zone is not formed until 40 deg BTDC, instead of at the beginning of injections. Second, one would expect that earlier injection results in longer evaporation time, as well as less stratification of the mixture, which can lead in more rapid combustion. However, inspection of Fig. 4 reveals the exact opposite; the evaporation is poor for the earlier injection case. The later injection scenario results in 36% lean mixture of total fuel injected (the product of the evaporation fraction and the fuel mass fraction of the lean mixture) by 20 deg BTDC and 54% lean mixture by TDC, fractions which are slightly higher than 35% and 29% for the baseline, and substantially higher than 23% and 21% for the earlier injection case.

In order to explore the reasons of the exact opposite trends, simulation results of distributions of equivalence ratio and temperature at 40, 20, and 0 deg BTDC are shown in Figs. 5 and 6, respectively. For the baseline, in-cylinder pressure and temperature at the instant of injection are very low. It is hard to find evaporated fuel mixture around the poor atomized jet. Most droplets formed by the first two injections penetrate and impinge on the cylinder liner quickly. Some droplets rebound from the wall and undergo second breakup process, and therefore the evaporation is improved. The rest of fuel droplets stick on the liner until impact of the droplets injected by the consecutive injections. All pulse jets undergo similar experiences, except evaporation for later pulse jets is improved and less fuel droplets impinge on the liner due to increased pressure and temperature. The fuel-air mixture concentrates in the cylinder periphery with liquid fuel droplets rather than distributes uniformly in the cylinder, because of poor atomization, lower evaporation and over penetration of the pulse jets. The low temperature in the cylinder periphery hinders the evaporation of fuel droplets. Just before ignition (20 deg BTDC), as observed in Figs. 5 and 6, a large portion of the mixture still distributes in the narrow space between the cylinder head and the piston. The temperature there is lower (T=650–775 K), in comparison with around 800–825 K in the central part of the cylinder. Obviously, the feature of concentration and temperature distribution caused by relatively earlier injection timing restrains the occurrence of autoignition of the mixture.

For the later injection timing case, the atomization and evaporation are improved, because of the increased in-cylinder pressure and temperature. After jet breakup, the spray angle increased rap-

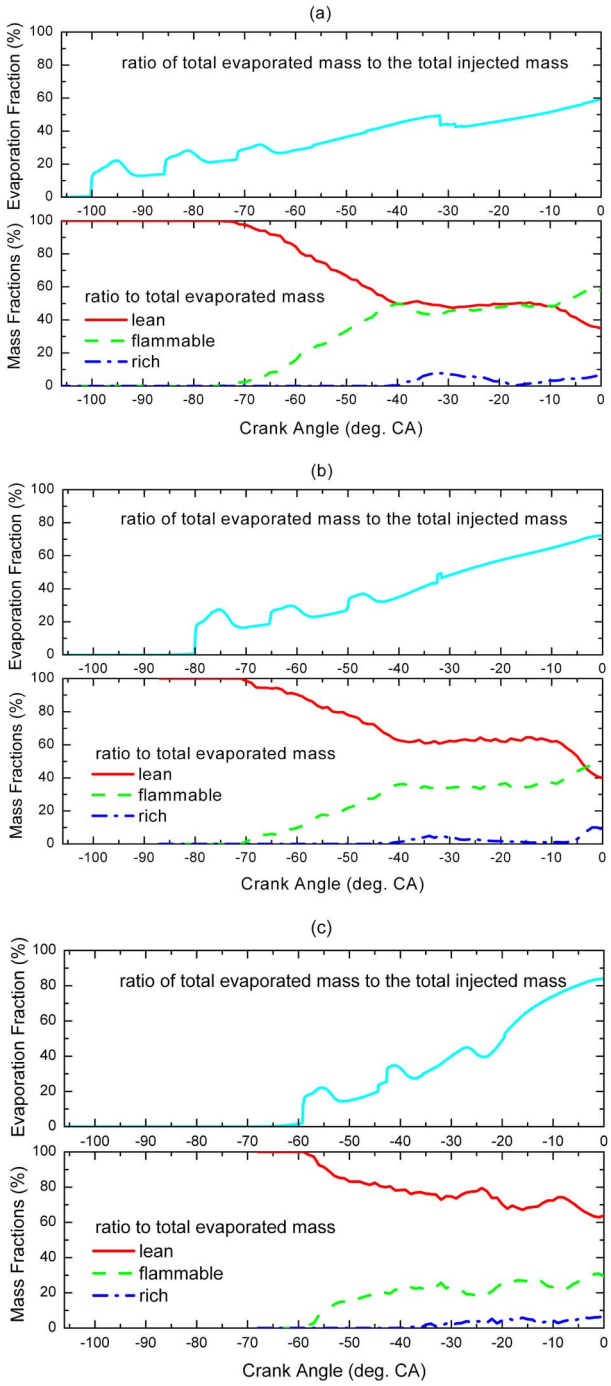


Fig. 4 Effect of fuel injection timing on fuel-air mixture formation: Injection profile 1. (*a*) Earlier injection, (*b*) baseline, (*c*) later injection.

idly and only a few of fuel droplets impinge on the liner. The following jets penetrate through the mixture parcel formed by the former pulse injections and medium scale turbulence motions are induced in the parcel. Therefore, the diffusion of the mixture parcel is dramatically enhanced. The resulting lean mixture distributes mainly in the piston bowl with concentration at $\phi = 0.06-0.21$ and temperature at $T = 775-825$ K at 20 deg BTDC. An earlier autoignition and a fast combustion rate can be reasonably concluded for the later injection case. These findings are also supported by Fig. 13 and agreed very well with previous experimental investigations [15,18,19].

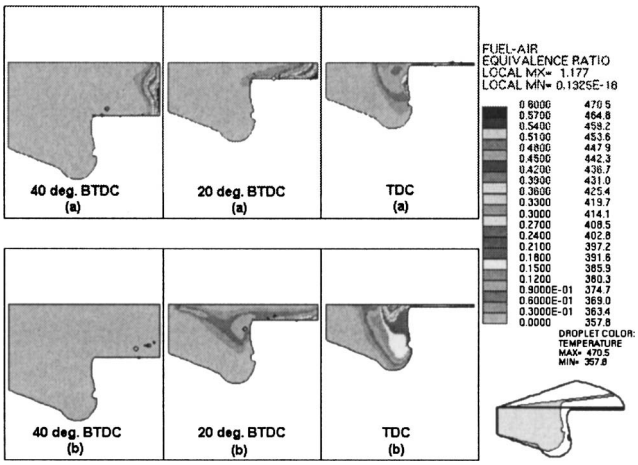


Fig. 5 Equivalence ratio distributions at 40, 20, and 0 deg BTDC. Perspective view. (*a*) Baseline, (*b*) later injection.

Effect of Pulse Dwell Time. This parametric study examines the effect of the pulse dwell time on the mixture stratification and fuel mixing. The dwell time is reduced from 8.4 deg CA for injection profile 1 to 5 deg CA for injection profile 2. Again, three referred injection timings are examined. Results of these simulations are shown in Figs. 7–9.

These computations clearly indicate that the pulse dwell time have a trifle impact on the fuel evaporation for all cases, but differently effect the mixture formation and mixture stratification for different injection timing cases. For the earlier injection timing, i.e., 110 deg BTDC, the fuel droplets penetrate to the cylinder periphery with scarcely evaporation. Therefore, the expected mixing enhancement disturbed by following injections is lacked. When retarding the multi-injection timing, higher in-cylinder pressure and temperature lead to better evaporation, and hence the disturbance by each other becomes more important. As shown in Figs. 4 and 7–9, when the short pulse dwell time is used, the lean fraction increases, from 35% to 36% by 20 deg BTDC for baseline case, and from 36% to 43% for later injection case. In particular, the flammable and lean portions of the mixture reach a combined 73% of the mixture at TDC, leaving the rich part with only 7% and 20% fuel droplets for later injection case. This clearly indicates that short pulse dwell time has a strong effect on mixing processes and the fuel stratification, especially for the later injection cases.

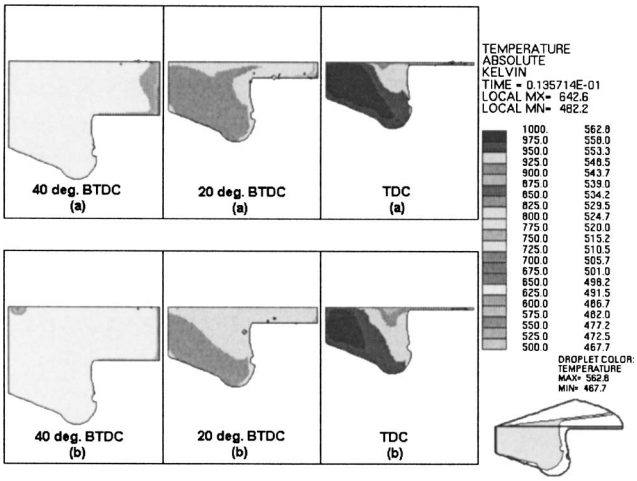


Fig. 6 Temperature distributions at 40, 20, and 0 deg BTDC. Perspective view. (*a*) Baseline, (*b*) later injection.

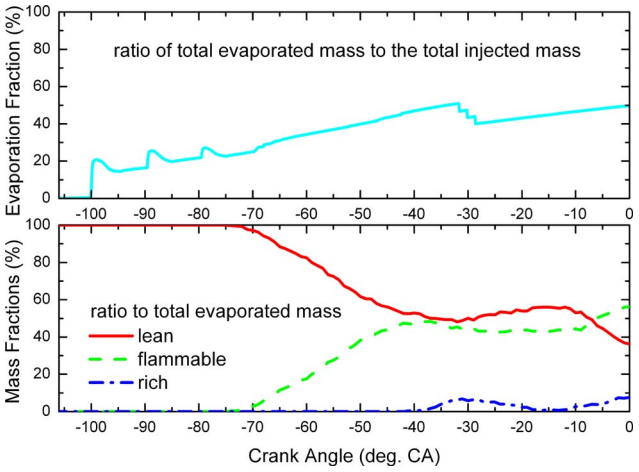


Fig. 7 Effect of pulse dwell time on fuel-air mixture formation, earlier injection, injection profile 2

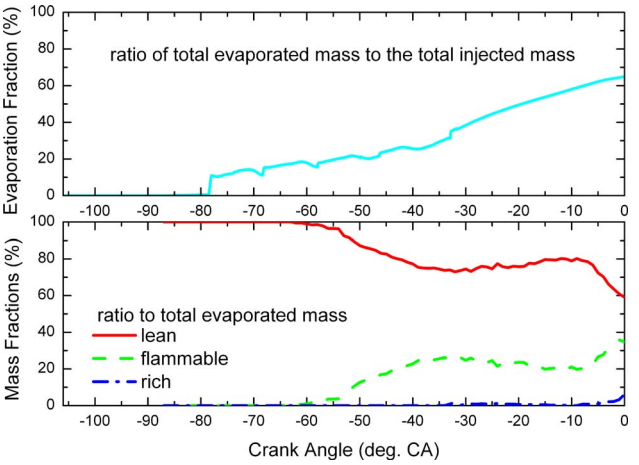


Fig. 10 Effect of pulse dwell time on fuel-air mixture formation, earlier injection, injection profile 3

Effect of the Pulse Width. The pulse width is also an important parameter that can be used to control mixing rate. In this part, the injection profile 2 and injection profile 3 are examined, i.e., the pulse width is reduced from 5.38 deg CA to 4.39 deg CA, while the pulse dwell time is kept fixed at 5 deg CA. Figures 10–12 show a dramatic difference in mixing when shorter pulse width is used, as expected.

Inspection of Figs. 7–12 reveals the effects of shorter pulse width on the lean mixture fraction for different injection timing

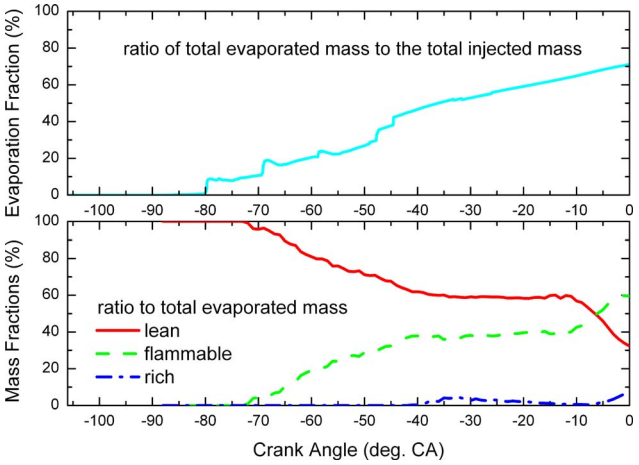


Fig. 8 Effect of pulse dwell time on fuel-air mixture formation, baseline, injection profile 2

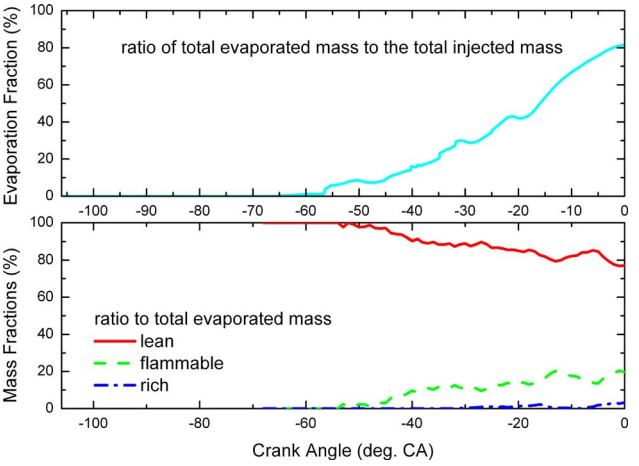


Fig. 11 Effect of pulse dwell time on fuel-air mixture formation, baseline, injection profile 3

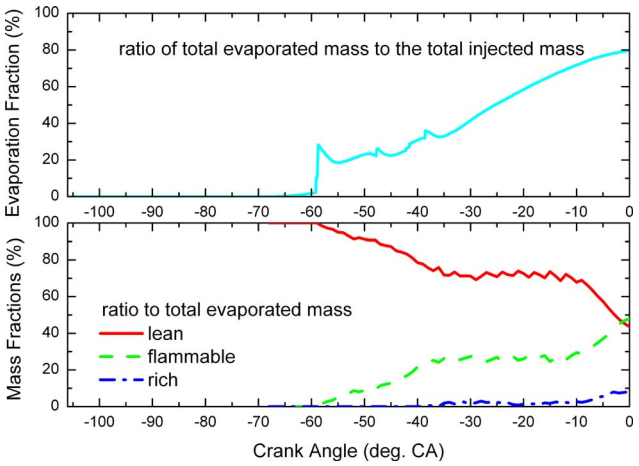


Fig. 9 Effect of pulse dwell time on fuel-air mixture formation, later injection, injection profile 2

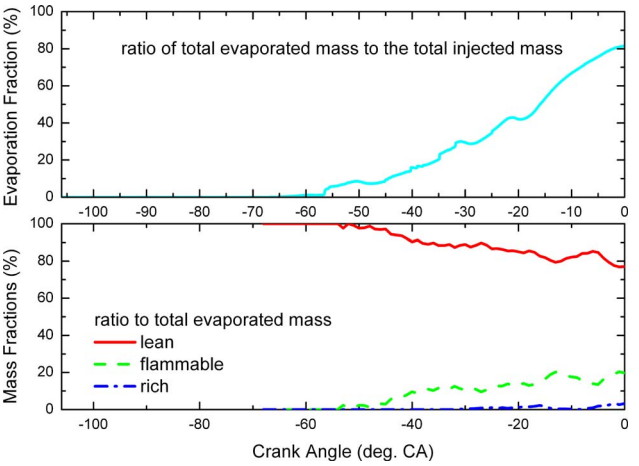


Fig. 12 Effect of pulse dwell time on fuel-air mixture formation, later injection, injection profile 3

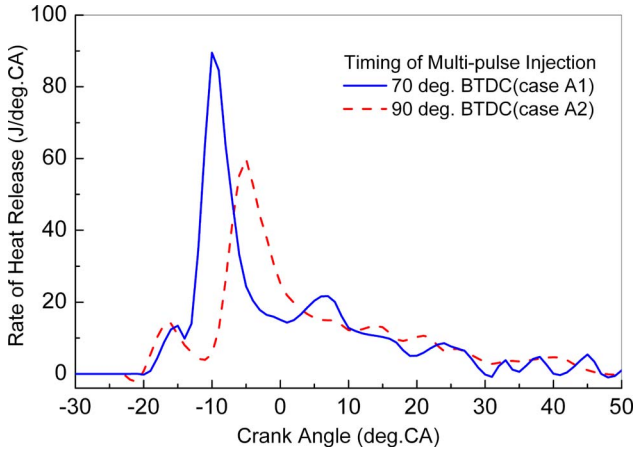


Fig. 13 ROHR at low load PCCI operation conditions by means of multiple injections (IMEP=0.32 MPa)

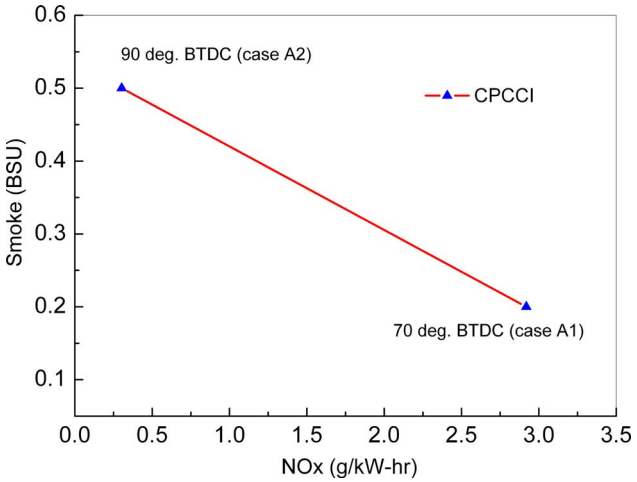


Fig. 14 NO_x and smoke emissions at low load HCCI operation conditions by means of multiple injections (IMEP=0.32 MPa)

cases. As the pulse width decreases the evaporation rate increases and then increases the lean mixture fraction from 23% to 31% at 20 deg BTDC for the earlier injection case. It is concluded that short pulse width will benefit to fuel evaporation and mixing of earlier injection case.

In summary, for the present design and operating conditions, the multiple injection timing is of paramount interests since it determines in-cylinder conditions of pressure and temperature. For earlier injection case, the pulse width shows a strong impact on evaporation and mixing rate. Hence, the short pulse width should be employed. The dwell time is an important factor and short dwell time will result in less stratification of the fuel for later injection case. Naturally, it is important to match and optimize the above three injection parameters, and then the direct-injection engine can operate in a well-premixed mode.

Effects of Rate of Heat Release on Emissions and Thermal Efficiency

Shape Analysis of the Heat Release Rate and its Effects on Emissions. The next portion of this work attempts to character the heat release mode, emissions and thermal efficiency of the MULINBUMP compound HCCI combustion from engine experiment results. As shown in Fig. 13, the engine operates in pure HCCI combustion modes for cases A1 and A2 of Table 2. The shape of the heat release rate is featured as a single peak curve. The rate of heat release is delayed and blunt for the earlier start timing of pulse injections (90 deg BTDC) and, on the contrary, advanced and concentrated for the later injection timing (70 deg BTDC). In correspondence with the feature of the rate of heat release, a very low NO_x emissions of only 0.3 g/kW hr is obtained at the case A2, while a higher NO_x emissions of 2.9 g/kW hr is observed at case A1, as shown in Fig. 14.

The compound combustion modes are obtained by cases B3–B6. About one-half of the fuel is injected by multipulse injection and a rather homogeneous fuel-air mixture is formed before ignition. Another half of the fuel is injected by the main injection at −3 deg BTDC, and then burns in a lean diffusion mode by means of the BUMP combustion chamber. Therefore, the combustion is divided into the HCCI combustion and the lean diffusion combustion. The curve of rate of heat release is correspondingly divided into two parts, as seen in Fig. 15. For a typical condition of a compound combustion, the two part of rate of heat release are quite equivalent, which is defined as a dispersed (DS) mode of rate of heat release. Whereas for the pure HCCI combustion or conventional diesel combustion, the curve of rate of heat release is usually with one rather concentrated peak. Therefore, it is called as a concentrated mode of rate of heat release. It is found repeatedly from the engine experiments that the cases with a DS mode of rate of heat release get lower emissions of NO_x and smoke, shown in Fig. 16, as an example. The identical character can be also observed in Figs. 17 and 18 even though the engine operates at higher loads.

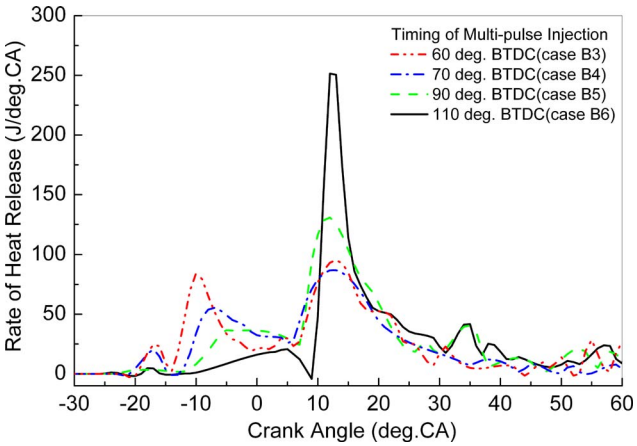


Fig. 15 Heat release results of cases B3–B6 (IMEP =0.67 MPa)

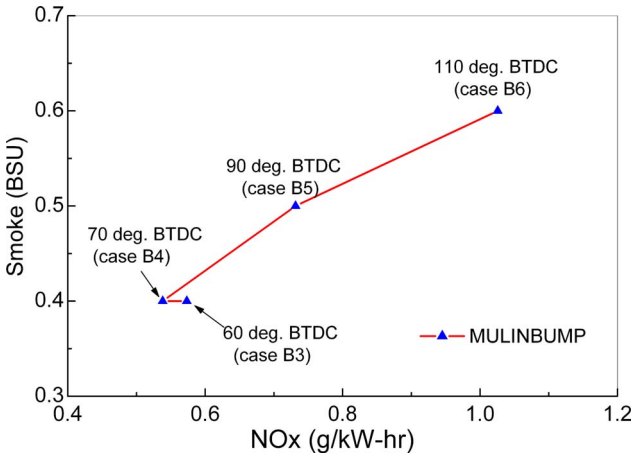


Fig. 16 NO_x and smoke emission results of cases B3–B6 (IMEP=0.67 MPa)

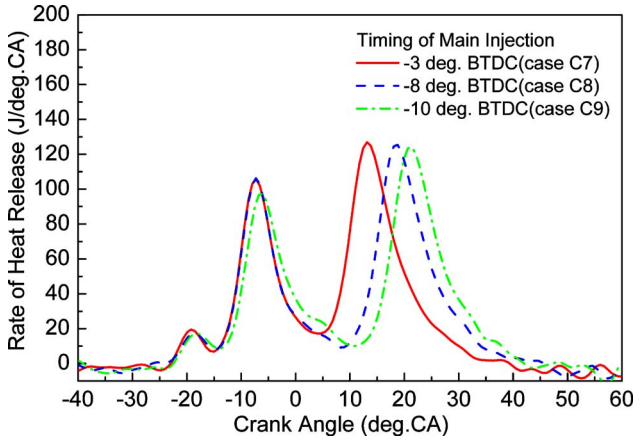


Fig. 17 Heat release results of cases C7–C9 (IMEP =0.93 MPa)

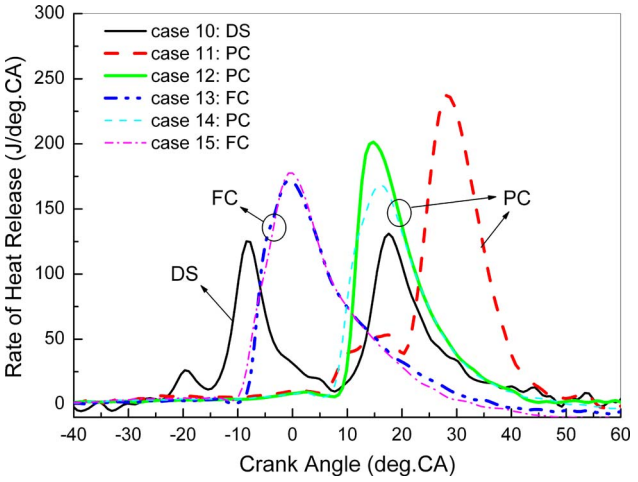


Fig. 19 Heat release results of cases D10–D15 (IMEP =0.93 MPa)

Efficiencies Discussions. Even though MULINBUMP-HCCI combustion has the advantage in simultaneous reduction of NO_x and smoke emissions, the significant increase in uHC and CO emissions will decrease the combustion efficiency in comparison with conventional diesel combustion. Therefore, it is concerned whether the thermal efficiency of the MULINBUMP diesel-HCCI combustion is high and it can be potential in industrial application. In this part, an analysis of combustion efficiency, heat loss and thermal efficiency is conducted based on engine experiments.

The combustion efficiency (η_c) is defined as

$$\eta_c = 1 - \frac{Q_{uf}}{Q_f}, \tag{1}$$

where Q_f is the heat value of the cycle fuel and Q_{uf} is the heat value lost due to uncompleted combustion of the cycle fuel, which equals the sum of the reaction heat of uHC and CO produced in an engine cycle. The specific reaction heat of uHC can be calculated by the low heat value of diesel fuel. The lost heat due to production of CO is calculated by calculation of the reaction heat deference between CO_2 and CO. The quantity of uHC and CO is measured by exhaust gas analyzer. The heat loss (η_{HL}) is defined as

$$\eta_{HL} = \frac{Q_w}{Q_f}, \tag{2}$$

where Q_w is the heat loss brought out by cooling water, which can be calculated by Sitkei formulas [31]. The indicated thermal efficiency (η_t) is defined as

$$\eta_t = \frac{W_t}{Q_f}, \tag{3}$$

where W_t is the piston work during all valve are closed, which is calculated through measured in-cylinder pressure data.

A combustion process in engines with higher combustion efficiency will release much heat of the fuel and the in-cylinder temperature will be higher. According to the analysis of an air standard cycle [32], an increase in temperature of the hot source will raise the thermal efficiency. However, it is also so important to locate the in-cylinder high temperature phase near the engine TDC, so that in-cylinder high-temperature charge will expand thoroughly before opening of the exhaust valve, resulting in higher thermal efficiency. Cases of group D in Table 2 are conducted at load of IMEP=0.93 MPa with variation of control parameters, including injection parameters, injection pressure, intake pressure, and so on. The cases of only single injection (cases D12–15) and the cases of a combustion chamber without the BUMP ring are also included. The performances of all these cases are optimized for the best thermal efficiency and emissions, respectively.

As shown in Fig. 19, the mode of rate of heat release may be classified as three modes. Case D10 is of a typical DS mode with two equivalent peaks of heat release rate. The first part of heat release rate represents the heat release of CPCCI combustion by the multipulse injection and the second part of heat release rate represents the heat release of the lean diffusion combustion by the main injection. Cases D13 and D15 are of a concentrated mode with a single peak of rate of heat release. Because the peak phases in cases D13 and D15 locate before the TDC, it is named as a FC mode, which is formed by earlier single injection timing (8 deg BTDC). Cases D12 and D14 are of PC modes with a single peak locating after the TDC. It is caused either by a retarded single fuel injection as cases D12 and D14 (–8 deg BTDC) or by a compound combustion with an over advanced multipulse injection and a retarded main injection as case D11. For over advanced multipulse injection, the pre-mixed mixture is too lean and cool to produce fast combustion. Therefore, most of the pre-mixed mixture formed by over advanced multipulse injection burns together

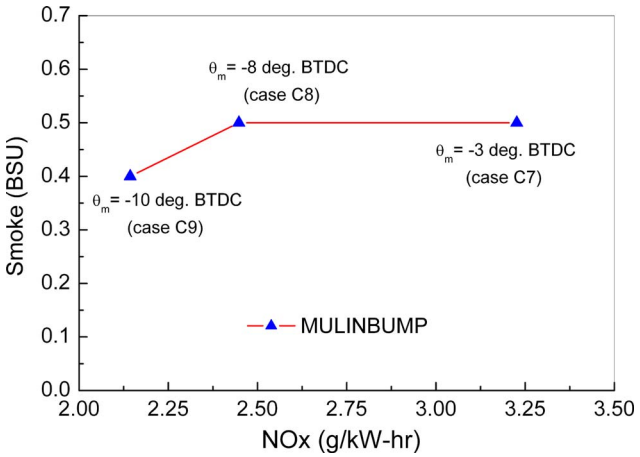


Fig. 18 NO_x and smoke emission results of cases C7–C9 (IMEP=0.93 MPa)

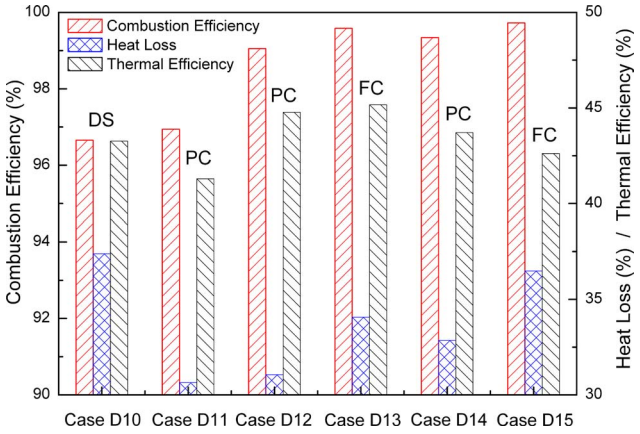


Fig. 20 Combustion efficiency, heat loss and thermal efficiency of cases D10–D15 (IMEP=0.93 MPa)

with the fuel injected by main injection.

As shown in Figs. 19–22, it is clear that the cycle in FC heat release mode gets the highest thermal efficiency, comparing to PC and DS modes, because of the highest in-cylinder temperature, the earliest phase of high-temperature and high combustion efficiency, although the heat loss is also rather high. The main problem for FC mode is high NO_x emissions. As a solution of NO_x reduction in DI diesel engines, a widely used technique is retarding the fuel injection timing, and burning the fuel in PC mode. Cases D12 and D14 with PC heat release mode reduce the NO_x emissions effectively, but get the lower thermal efficiency, because of a retarded phase of high temperature. Case D10 is a typical MULINBUMP combustion with DS heat release mode. The emissions of NO_x and smoke are significantly reduced at the same time. The MULINBUMP combustion with a DS heat release mode reveals higher heat loss because of a relatively longer duration of heat release and lower combustion efficiency due to higher emissions of uHC and CO, which worsen the thermal efficiency of the MULINBUMP combustion. However, as shown in Fig. 20, it is approximately as high as that in FC mode. Therefore, it is still a promising combustion mode from the comprehensive consideration of efficiency and emissions. A serious challenge to further development of the MULINBUMP compound combustion is to improve combustion efficiency and heat loss, while expanding the load range of CPCCI combustion.

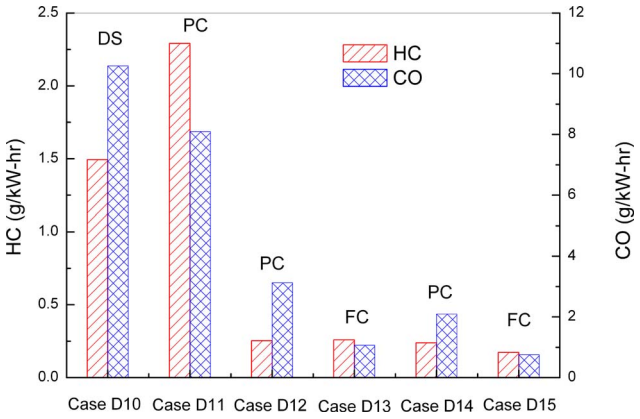


Fig. 21 uHC and CO emission results of cases D10–D15 (IMEP=0.93 MPa)

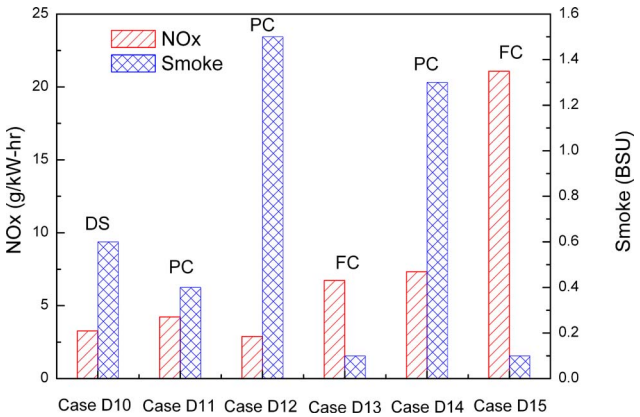


Fig. 22 NO_x and smoke emission results of cases D10–D15 (IMEP=0.93 MPa)

Summary

The MULINBUMP-HCCI compound combustion technology is developed by combination of controlled premixed charge compression ignition (CPCCI) and lean diffusion combustion (LDC). The results of engine experiments and CFD studies show that premixed mixture preparation may be well controlled by setting parameters of multipulse injection according to certain engine operation conditions. The MULINBUMP working cycle with dispersed heat release mode reveals that its thermal efficiency is approximately as high as a conventional diesel cycle with an advanced injection timing of 8 deg BTDC, while emissions of NO_x and smoke are reduced simultaneously. NO_x emissions are only 0.3 g/kW hr at lower loads and less than 3.25 g/kW hr at higher loads. Smoke emission is never more than 0.5 BSU. The combustion characteristics of the MULINBUMP-HCCI compound combustion technology can be summarized as follows:

1. Through control of the multipulse injection parameters, the distributions of concentration and temperature of the preinjected fuel can be flexibly controlled. The multiple injection timing is of paramount interests since it determines in-cylinder conditions of pressure and temperature, and consequently represents the relative effectiveness of fuel stratification for controlling CPCCI combustion phasing. For earlier injection case, the pulse width shows a strong impact on evaporation and mixing rate. Hence, the short pulse width should be employed. The dwell time is an important factor and short dwell time will result in more uniform stratification of the fuel for later injection case. Matching and optimizing of those injection parameters are necessary at a certain operation condition in order to reach the objective of higher efficiency and lower emissions.
2. By comparison of heat release mode of different diesel combustion, the heat release mode can be summarized as the forward concentrated (FC) mode, the post concentrated (PC) mode and the dispersed (DS) mode. The FC mode gives the highest thermal efficiency but with highest NO_x emissions. The PC mode may gets lower NO_x emissions, but with the drawback of lower thermal efficiency and higher smoke number. The DS mode is a flexible mode in combustion control. The thermal efficiency in DS mode can reach approximately to as high as that in FC mode, while the NO_x and smoke emissions can be reduced simultaneously and remarkably.

Acknowledgment

The research is supported by the National Natural Science Foundation of China through its key project funding and Overseas Outstanding Scholarship Award to H. Zhao, and by the Ministry of Science and Technology through its 973 National Key Project on HCCI Combustion Engines and Fuels.

References

[1] Zhao, Fuquan (Frank), et al., 2002, *Homogenous Charge Compression Ignition (HCCI) Engine: Key Research and Development Issues*, Society of Automotive Engineers, Inc., Warrendale, PA.

[2] Ryan III, Thomas W., and Callahan, Timothy J., 1996, "Homogeneous Charge Compression Ignition of Diesel Fuel," SAE Paper No. 961160.

[3] Gray III, Allen W., and Ryan III, Thomas W., 1997, "Homogeneous Charge Compress Auto-ignition (HCCI) of Diesel/Fuel," SAE Paper No. 971676.

[4] Suzuki, Hisakazu, Koike, Noriyuki, and Odaka, Matsuo, 1998, "Combustion Control Method of Homogeneous Charge Diesel Engines," SAE Paper No. 980509.

[5] Odaka, Matsuo, Suzuki, Hisakazu, Koike, Noriyuki, and Ishii, Hajime, 1999, "Search for Optimizing Control Method of Homogeneous Charge Diesel Combustion," SAE Paper No. 1999-01-0184.

[6] Kaneko, Naoya, Ando, Hirokazu, Ogawa, Hideyuki, and Miyamoto, Noboru, 2002, "Expansion of the Operating Range with in-Cylinder Water Injection in a Premixed Charge Compression Ignition Engine," SAE Paper No. 2002-01-1743.

[7] Peng, Zhijun, Zhao, Hua, and Ladommatos, Nicos, 2003, "Effects of Air/Fuel Ratios and EGR Rates on HCCI Combustion of N-heptane, a Diesel Type Fuel," SAE Paper No. 2003-01-0747.

[8] Akagawa, Hisashi, Miyamoto, Takeshi, Harada, Akira, Sasaki, Satoru, Shimazaki, Naoki, Hashizume, Takeshi, and Tsujimura, Kinji, 1999, "Approaches to Solve Problems of the Premixed Lean Diesel Combustion," SAE Paper No. 1999-01-0183.

[9] Nishijima, Yoshiaki, Asaumi, Yasuo, and Aoyagi, Yuzo, 2001, "Premixed Lean Diesel Combustion (PREDIC) Using Impingement Spray System," SAE Paper No. 2001-01-1892.

[10] Yokota, Haruyuki, Kudo, Yugo, Nakajima, Hiroshi, Kakegawa, Toshiaki, and Suzuki, Takashi, 1997, "A New Concept for Low Emission Diesel Combustion," SAE Paper No. 970891.

[11] Suzuki, Takashi, Kakegawa, Toshiaki, Hikino, Kiyoharu, and Obata, Atsuomi, 1997, "Development of Diesel Combustion for Commercial Vehicles," SAE Paper No. 972685.

[12] Iwabuchi, Yoshinori, Kawai, Kenji, Shoji, Takeshi, and Takeda, Yoshinaka, 1999, "Trial of New Concept Diesel Combustion System - Premixed Compression-ignition Combustion," SAE Paper No. 1999-01-0185.

[13] Yanagihara, H., 2001, "Ignition Timing Control at Toyota "UNIBUS" Combustion System," *Proceedings of the IFP International Congress on a New Generation of Engine Combustion Processes for the Future*, pp. 34–42.

[14] Walter, Bruno, and Gatellier, Bertrand, 2002, "Development of the High Power NADI Concept Using Dual Mode Diesel Combustion to Achieve Zero NO_x and Particulate Emissions," SAE Paper No. 2002-01-1744.

[15] Su, Wanhua, Lin, Tiejian, and Pei, Yiqiang, 2003, "A Compound Technology for HCCI Combustion in a DI Diesel Engine Based on the Multi-Pulse Injection and the BUMP Combustion Chamber," SAE Paper No. 2003-01-0741.

[16] Kimura, Shuji, Aoki, Osamu, Ogawa, Hiroshi, Muranaka, Shigeo, and Enomoto, Yoshiteru, 1999, "New Combustion Concept for Ultra-clean and High-efficiency Small DI Diesel Engines," SAE Paper No. 1999-01-3681.

[17] Kimura, Shuji, Aoki, Osamu, Kitahara, Yasuhisa, and Aiyoshizawa, Eiji, 2001, "Ultra-clean Combustion Technology Combining a Low-Temperature and Premixed Combustion Concept for Meeting Future Emission Standards," SAE Paper No. 2001-01-0200.

[18] Pei, Yiqiang, Su, Wanhua, and Lin, Tiejian, 2002, "The BUMP Combustion Chamber Presented Based on the Concept of Lean Diffusion Combustion in a D.I. Diesel Engine With Common Rail Fuel Injector," Transactions of CSICE **20**(5), pp. 381–386 (in Chinese).

[19] Lin, Tiejian, Su, Wanhua, and Pei, Yiqiang, 2002, "A New HCCI Combustion Technology in a DI Diesel Engine Based on Compound Control of Multi-Pulse Fuel Injection and the BUMP Combustion Chamber," Transactions of CSICE **20**(6), pp. 475–480 (in Chinese).

[20] Su, Wanhua, Lin, Rongwen, Xie, Hui, and Shi, Shaoxi, 1997, "Enhancement of Near Wall Mixing of an Impinging Jet by Means Bump on the Wall," SAE Paper No. 971616.

[21] Su, Wanhua, Wang, Yang, Lin, Tiejian, Xie, Hui, Pei, Yiqiang, Li, Shaoan, Cheng, Gang, and Wang, Xiaobo, 2001, "A Study of Effects of Design Parameters on Transient Response and Injection Rate Shaping for a Common Rail Injector System," SAE Paper No. 2001-01-3506.

[22] Han, Z. and Reitz, R. D., 1995, "Turbulence Modeling of Internal Combustion Engines Using RNG k-ε Models," Combust. Sci. Technol., **106**, pp. 267–295.

[23] Kong, S. C., Han, Z., and Reitz, R. D, 1995, "The Development and Application of Diesel Ignition and Combustion Model for Multi-Dimensional Engine Simulation," SAE Paper No. 950278.

[24] Huh, K. Y., and Gosman, A. D., 1991, "A Phenomenological Model of Diesel Spray Atomization," *Proceedings of the International Conference on Multiuphase Flows*, Sept. 24–27, Tsukuba, Japan.

[25] Reitz, R. D., and Diwakar, R., 1986, "Effect of Drop Breakup on Fuel Sprays," SAE Paper No. 860469.

[26] Reitz, R. D., and Diwakar, R., 1987, "Structure of High-Pressure Fuel Sprays," SAE Paper No. 870598.

[27] Bai, C., and Gosman, A. D., 1995, "Development of Methodology for Spray Impingement Simulation," SAE Paper No. 950283.

[28] "STAR-CD Methodology," 2001, CD Adapco Group.

[29] Gonzalez, M. A., Lian, Z. W., and Reitz, R. D., 1992, "Modeling Diesel Engine Spray Vaporization and Combustion," SAE Paper No. 920579.

[30] Abraham, J., 1997, "What is Adequate Resolution in the Numerical Computation of Transient Jets?" SAE Paper No. 970051.

[31] György, Sitkei, 1974, "Heat Transfer and Thermal Loading in Internal Combustion Engines," Akademiai Kaido Budapest.

[32] Benson, Rowland S., and Whitehouse, N. D. 1979, *Internal Combustion Engines*, Pergamon Press, Oxford.

Effect of Developing Turbulence and Markstein Number on the Propagation of Flames in Methane-Air Premixture

M. Z. Haq
Department of Mechanical Engineering,
Bangladesh University of Engineering and Technology,
Dhaka-1000, Bangladesh
e-mail: zahurul@me.buet.ac.bd

In spark ignition engines, initial flame kernel is wrinkled by a progressively increasing bandwidth of turbulence length scales until eventually the size of the flame kernel is sufficient for it to experience the entire turbulence spectrum. In the present study, an effective rms turbulence velocity as a function of time, estimated by integrating the nondimensional power spectrum density (psd) function for isotropic turbulence, is utilized to analyze the statistical distribution of flame front curvatures and turbulent burning velocities of flames propagating in methane-air premixtures. The distributions of flame front curvatures show these to become more dispersed as the effective turbulence velocity increases, and result in increased burning of premixtures. A decrease in the Markstein number also results in a further increase in curvature dispersion and enhanced burning, in line with the flame stability analysis. [DOI: 10.1115/1.2056537]

Introduction

In a turbulent premixture inside the engine cylinder, spark initiates reaction and a propagating flame front capable of overcoming the high geometric stretch. As the flame kernel grows and the influence of the geometric stretch is superseded by stretch due to aerodynamic strain, the flame is, at first, wrinkled by only the smallest scales of turbulence [1,2]. When the flame ball is small relative to the average eddy size, the high frequency smaller eddies are most effective in shearing and wrinkling the flame front, and the eddies larger than the flame kernel mostly propel the flame around [3]. The size of the eddy interacting with the initial flame kernel plays an important role in the early flame development. Although most of the mass in the combustion chamber burns in the later stage of the combustion, a large portion of the total combustion time is occupied by the early flame growth [4]. It is now widely accepted that the early flame development is an important factor affecting the cycle-to-cycle variations in the engine [5,6].

As the flame kernel continues to grow it becomes progressively wrinkled by the larger length scales with an associated increase in the turbulent burning velocity as a result of the contribution of an increasingly larger portion of the turbulence spectrum [1]. During this period, turbulent velocity effective in wrinkling the flame is u_k' ($u_k' \leqslant u'$), where u' is the rms turbulent velocity. After sufficient time, eventually the size of the kernel is sufficient for it to experience the entire turbulence spectrum. The effect of turbulence (embodied in u_k') on flame wrinkling is then fully developed and is equal to u' [2]. However, a reverse effect, "de-developing turbulence," might occur as the flame approaches the wall such that progressively smaller scales of turbulence are available to affect the flame [7].

It is now widely recognized that turbulent flames in spark-ignited engines can be treated as an array of laminar flamelets with no turbulence structure residing within them, and the flame front is considered as an interface separating cold reactants and the burned combustion products [8]. Flamelet combustion corresponds to chemical reaction occurring at fast time scales and short length scales relative to turbulence. In this situation, the flame is confined to a relatively thin layer and the dominant effect of the turbulence is to wrinkle the flame front and to increase the flame surface area and to strain the local flame [9]. Consequently, the effect of strain has been used to correlate the turbulent burning velocity [2]. However, flame front curvature strongly influences the distribution of fast diffusing radicals such as H, the concentration of which increases in regions of negative curvature because of flame focusing [10]. This results in a further increase in local displacement speed of the flame and plays a significant role in turbulent flame propagation. Therefore, recent analyses of turbulent flame in the flamelet regimes have involved the various spatial statistical properties of the flame surface [11]. Hence, the flame surface area and curvature statistics provide a complete geometrical description of the turbulent flame propagation in laminar flamelet regime [12]. In this regime, a turbulent flame then can be treated as an ensemble of laminar flamelets whose burning velocities are determined by the local flame stretch. Integration of the local burning rate over the flame surface gives the total burning rates of the turbulent flame [13].

Using two-dimensional Laser sheet imaging techniques, it is possible to visualize the turbulent flame front, derive the statistics of curvatures of flame fronts, and profile the distribution of burned and unburned gases at all stages of the flame propagation. The statistics thus obtained are required to model three-dimensional flame propagation and its interaction with the turbulent flow field. The present paper reports the effects of developing turbulence on the propagation of flames in the methane-air premixture at 300 K using images obtained with planar Mie scattering (PMS) technique. The sequences of two-dimensional images are analyzed to derive the statistics of flame front curvatures and turbulent burning velocities for rms turbulence intensities of 0.6 and 1.2 m/s, at two initial pressures of 0.1 and 0.5 MPa. Experimental conditions of the present study, corresponding mixture properties and turbulence parameters are summarized in Table 1.

Experimental and Analytical Techniques

Methane-air premixtures were ignited in a fan-stirred combustion vessel shown in Fig. 1. It is a 380 mm diameter spherical stainless steel vessel capable of withstanding the temperatures and pressures generated from explosions at initial pressures of up to 1.5 MPa and initial temperatures of up to 600 K. It has extensive

Contributed by the Internal Combustion Engine Division of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received January 24, 2004; final manuscript received February 18, 2005. Assoc. Editor: D. Assanis.

Table 1 Initial conditions and mixture properties of methane-air premixtures at 300 K

Initial conditions		Turbulence parameters			Mixture properties, from Ref. [14]			
P (MPa)	ϕ	u' (m/s)	η (mm)	λ (mm)	ρ_u (Kg/m^3)	ρ_b (Kg/m^3)	u_l (m/s)	Ma_b
0.1	0.8	0.6	0.142	2.93	1.12	0.167	0.265	11.8
0.1	1.0	0.6	0.142	2.93	1.11	0.148	0.36	21.7
0.1	1.0	1.2	0.085	2.07	1.11	0.148	0.36	21.7
0.5	1.0	0.6	0.043	1.31	5.55	0.732	0.195	−19.3
0.5	1.0	1.2	0.025	0.93	5.55	0.732	0.195	−19.3

optical access through three pairs of orthogonal windows of 150 mm diameter and 100 mm thickness, made with Schlieren quality glass. Turbulence was generated using four identical, eight-bladed, independently controlled fans, and these fans were symmetrically disposed in a regular tetrahedron configuration. The turbulence velocity field in the optically accessed central region was calibrated for different fan speeds using laser Doppler velocimetry (LDV) and found to be nearly isotropic with no significant mean flow (further details can be found in Ref. [15]). The rms turbulent velocity, u', was found to vary linearly with fan speed and the turbulence integral length scale of 20 mm was estimated from the spatial correlation.

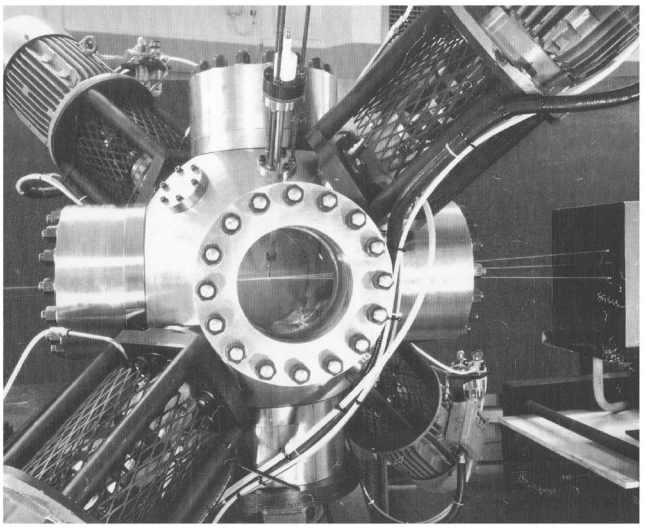


Fig. 1 The fan-stirred combustion vessel

The arrangement used for flame visualization using the planar Mie scattering (PMS) technique used in the present study is shown in Fig. 2. A copper vapor laser, model CU15-A, made by Oxford Lasers, operating simultaneously at wavelengths of 510.6 nm (green) and 578.2 nm (dark yellow), was used as a light source. The pulse rate was 4.5 kHz, the pulse duration was 15 ns, and the pulse energy was approximately 6 mJ. The light sheet was obtained using a 500 mm spherical bi-convex lens which focused the laser beam, followed by a 1000 mm plano-convex lens to expand it in one plane to form a sheet with an estimated thickness of 0.5 mm. The laser sheet was passed along a vertical plane just in front of the spark plug located in the center of the vessel and was Mie scattered from tobacco smoke particles. The particles evaporated, sublimed, or burned as they crossed the flame front and consequently did not scatter lights, and so provided a planar section of the complex three-dimensional flame front geometry. The scattered lights were imaged perpendicular to the incident light sheet and the images were recorded and stored using a Kodak Ektapro HS Motion Analyzer (Model 4540) at a rate of 4500 images per second. The light sensitivity of the camera was ISO 3000 at the high gain setting. It had a 256×256 pixel sensor, with an 8 bit level response. A 510.6 nm interference filter was attached to the camera lens to prevent flame-generated light from obscuring the images.

The electronic triggering features of the motion analyzer made it possible to store images prior to, after, and at both sides of the

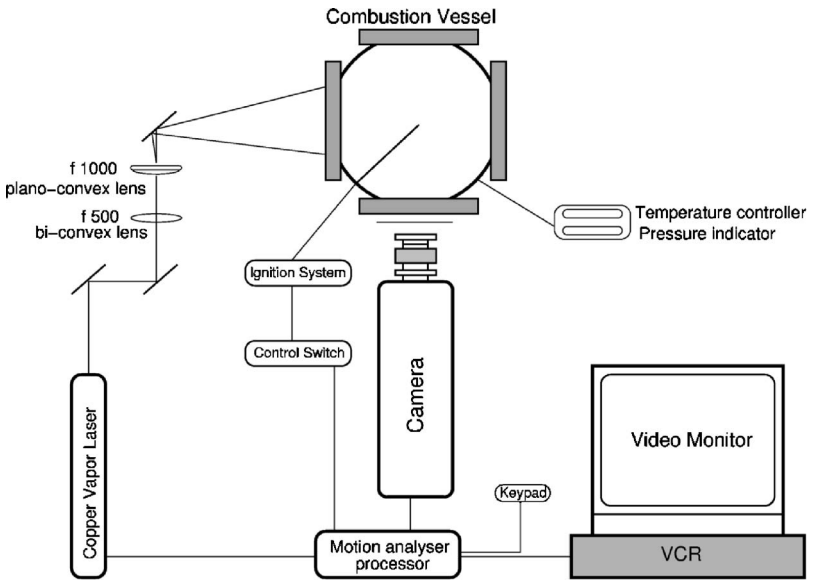


Fig. 2 Experimental setup for PMS visualization of turbulent flames

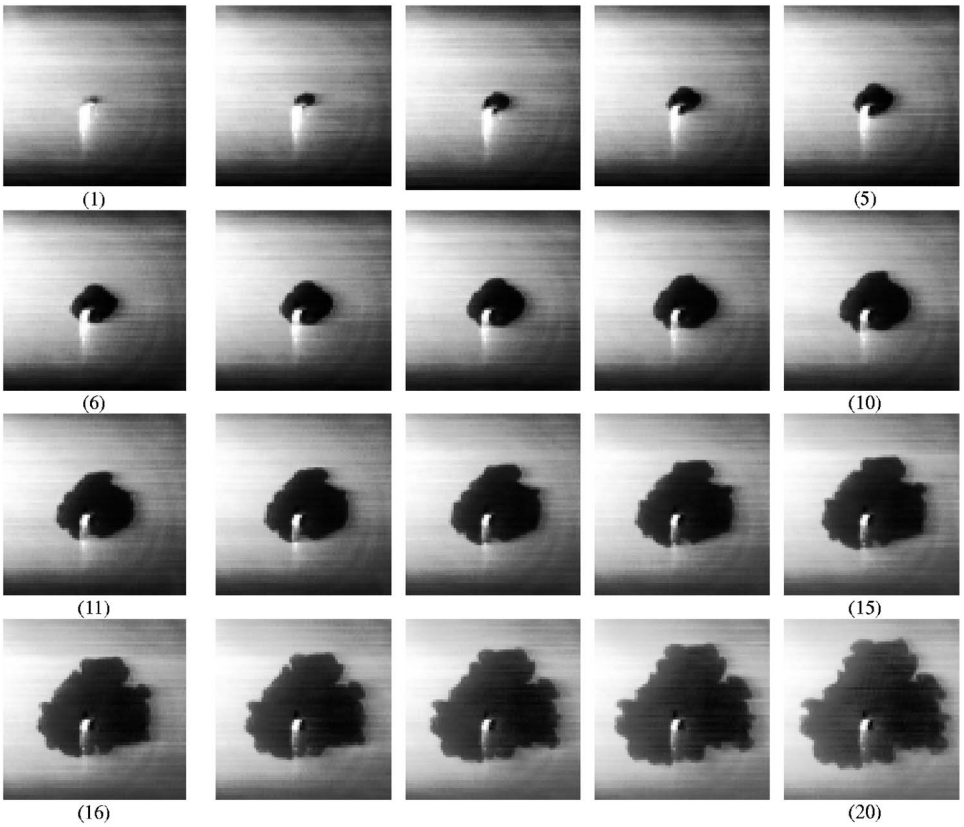


Fig. 3 Flame images for the propagation in methane-air premixture with an isotropic rms turbulent velocity of 0.6 m/s. Time interval between the images is 0.67 ms and real size of the full frame image is 128 mm.

trigger pulse. In the present study, "center recording mode" was used in which images were recorded until the trigger signal was received and then 1536 additional frames were recorded after the trigger. Prior to triggering the spark, the copper-vapor laser was synchronized to the camera-framing rate at 4.5 kHz. Upon receiving the spark signal through a control switch, the motion analyzer stored digital images in dynamic random access memory (DRAM) and these were available for immediate playback or for subsequent image processing. After a sequence of image processing steps, a clear cross-sectional view of the flame, identified as the frontier between the bright, unburned, and the dark, burned regions in the illuminated zone, were obtained. A front tracking algorithm was used to obtain the front coordinates from the images with a spatial resolution of 0.32 mm/pixel. Flame front coordinates were analyzed to measure turbulent burning parameters. All other aspects of equipment and the experimental techniques are reported in Ref. [15].

Shown in Fig. 3 are the images obtained from a flame propagating in stoichiometric methane-air mixture at an initial pressure of 0.1 MPa and initial temperature of 300 K. The flame front wrinkling caused by turbulence and the persistence of history are clearly visible. Shown in Fig. 4 are flame images obtained for different initial conditions where the effects of effective turbulence and Markstein number are evident. Hence, a propagating flame front is subjected to effects due to aerodynamic strain and curvature which together constitute flame stretch and change the flame surface area. For a laminar flame propagation, a burned gas Markstein length, L_b, is defined to account for the sensitivity of the laminar flame speed, S_n, to stretch, α, such that [14]

$$S_n = S_s - L_b\alpha \tag{1}$$

where S_s is the unstretched laminar flame speed and is obtained as the intercept value of S_n at $\alpha=0$ in the plot of S_n against α [14]. Markstein length is normalized by laminar flame thickness, δ_l, to produce Markstein number, Ma_b. The influence of kinetics, thermodynamics, thermodiffusion, and activation energy on the flame propagation are embodied in the Markstein number [15] and the

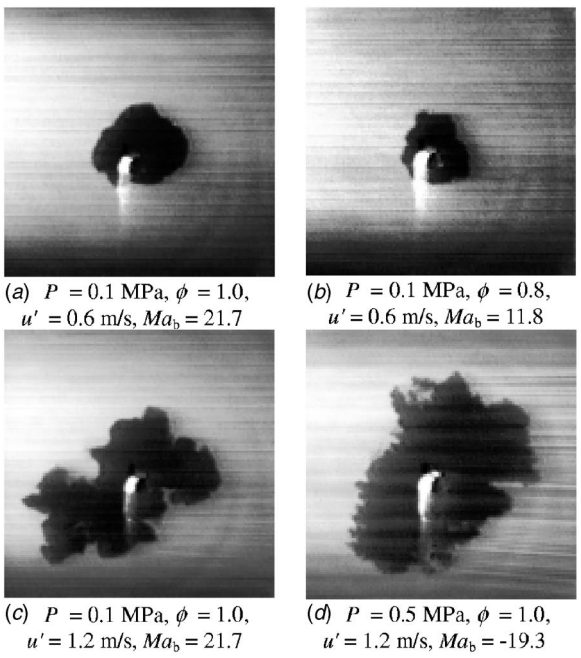


Fig. 4 Flame images at different initial conditions

Markstein number is now being widely used to analyze turbulent flame propagation [16] and the susceptibility to flame instability [17]. A high value of Markstein number is reported to stabilize and counter the perturbation in the flame fronts and lower values are also related to the early onset of instability at small flame radius in flame propagation [17].

Estimation of Flame Front Curvature. Analytically the flame front curvature, H, can be calculated using flame front coordinates (x,y), from [12,18]

$$H = \frac{d^2y/dx^2}{[1+(dy/dx)^2]^{3/2}} \tag{2}$$

However, the use of an independent variable, s, has been reported to provide improvement over this procedure [12], because the flame curvature then may be analyzed regardless of the slope and complexities of the flame surface. Here s is the distance along the flame front from a fixed origin located on the flame front. The curvature, H, can then be obtained from [12,18]

$$H = \left[\left(\frac{d^2x}{ds^2}\right)^2 + \left(\frac{d^2y}{ds^2}\right)^2\right]^{1/2} \tag{3}$$

Here the flame curvature is defined being positive if the flame element is convex to the reactant gases. The flame coordinates obtained from digital images required some smoothing of the flame front edge. In the present study, the Savitzky-Golay algorithm [19] is used for data smoothing and numerical differentiation.

Estimation of Turbulent Burning Velocity. For a spherical flame propagation, a mass rate of burning can be expressed in the form

$$\frac{dm}{dt} = -4\pi R_i^2 \rho_u u_t \tag{4}$$

where R_i is a mean general radius associated with the turbulent burning velocity, u_t, ρ_u is the unburned gas density, and m is the mass of the gas changing from unburned to burned. However, problems arise in defining *burned* and *unburned*, and also in selecting the most appropriate value of R_i. In the present study, a thin planar sheet cutting the mean spherical flame kernel is considered. This is shown in Fig. 5 and the general radius, R_i, lies between a root radius, R_r, and a tip radius, R_t. The circumference of radius R_r embraces the burned gas entirely, while the circumference of radius R_t has nothing but unburned gas outside it. It is also necessary to define the masses of unburned and burned gases associated with the different zones. Let m_{ui} be the mass of unburned gas within the general perimeter of radius R_i, m_{uo} be the mass of unburned gas outside perimeter R_i but within that of R_t, m_{bi} be the mass of burned gas inside the perimeter of R_i, m_{bo} be the mass of burned gas outside the perimeter of R_i but within R_t, and m_u be the mass remaining outside R_t. The definitions of the masses are also shown in Fig. 5.

In the flame images, the proportions of circumference occupied by unburned and burned gases were measured around the circumference for all the full range of radii at different instants during the flame propagation. From these flame images, values of the average volume fractions occupied by unburned gas at the different spherical radii r, $\sigma(r)$, were estimated and the assumption of isotropy enables these values to be expressed solely as a function of r. Hence, the mass of unburned gas within any general radius R_i, m_{ui}, is given by

$$m_{ui}(R_i) = 4\pi\rho_u \int_o^{R_i} \sigma(r) r^2 dr \tag{5}$$

The mass of unburned gas outside this general radius R_i, m_{uo}, is

$$m_{uo}(R_i) = 4\pi\rho_u \int_{R_i}^{Rt} \sigma(r) r^2 dr \tag{6}$$

Similarly, the mass of burned gas within any general radius R_i, m_{bi}, is

$$m_{bi}(R_i) = 4\pi\rho_b \int_o^{R_i} [1-\sigma(r)] r^2 dr \tag{7}$$

The mass of burned gas outside the general radius R_i, m_{bo}, is

$$m_{bo}(R_i) = 4\pi\rho_b \int_{R_i}^{Rt} [1-\sigma(r)] r^2 dr \tag{8}$$

Mass conservation implies that

$$\frac{dm_{ui}}{dt} + \frac{dm_{uo}}{dt} + \frac{dm_{bi}}{dt} + \frac{dm_{bo}}{dt} + \frac{dm_u}{dt} = 0 \tag{9}$$

In the present study, the value of R_i is defined such that the volume of unburned gas inside the circumference of radius, R_v, is equal to that is burned gas outside that circumference. Hence, the value of R_v is estimated by equating $m_{ui}(R_i)/\rho_u$ from Eq. (5) and $m_{bo}(R_i)/\rho_b$ from Eq. (8), computing its value by Bisection method [15].

A turbulent burning velocity, u_t, is defined at R_v such that it measures the mass rate of production of burned gas, which is related to the pressure development in engine cylinders, and is given by

$$\frac{d}{dt}(m_{bi} + m_{bo}) = 4\pi R_v^2 \rho_u u_t \tag{10}$$

Estimation of Effective rms Turbulent Velocity. At any instant, the effective rms turbulence velocity, u_k', was estimated by the integral of the dimensional power spectral density function,

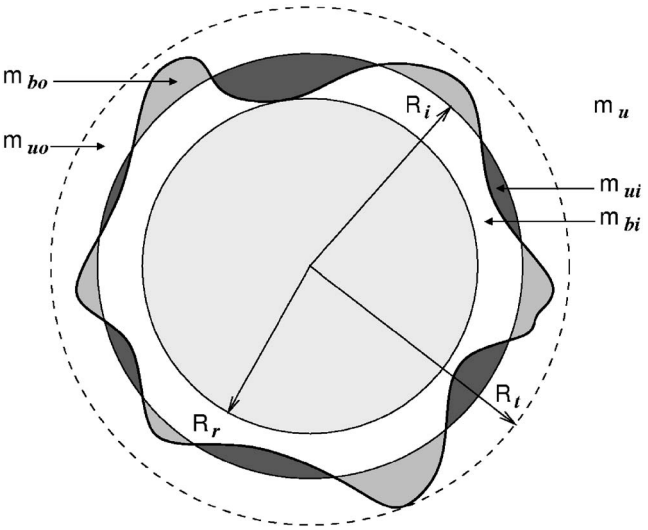


Fig. 5 Definition of reference radii and associated masses for a flame image

from the highest frequency to a threshold frequency given by the reciprocal of the elapsed time from ignition [2]. Scott [20] improved the nondimensional power spectrum by replacing dimensional frequency with a dimensionless wave number, $\bar{k}_\eta$, defined as

$$\bar{k}_\eta = k\eta \tag{11}$$

where k is the wave number and η is the Kolmogorov length scale. The wave number provides the link to the time scale. At the elapsed time, t_k, measured from ignition, the corresponding value of k is given by

$$k = \frac{2\pi}{u' t_k} \tag{12}$$

and Kolmogorov length scale, η, is defined as

$$\eta = (\nu^3/\varepsilon)^{1/4} \tag{13}$$

where ε is the specific rate of dissipation of turbulent kinetic energy and ν is the kinematic viscosity. Kolmogorov length scale is related to the Taylor macro-scale, λ, as

$$\frac{\eta}{\lambda} = 15^{-1/4}\,\mathrm{Re}_\lambda^{-1/2} \tag{14}$$

where Re_λ is the Reynolds number based on Taylor scale and is related to integral length scale, L, as

$$\frac{\lambda}{L} = \frac{A}{\mathrm{Re}_\lambda} \tag{15}$$

where A is a numeric constant. Scott [20] suggested the value of $A = 16 \pm 1.5$, and it gives a good fit to the measurements, over a wide range of the data [21].

The nondimensional turbulence power spectrum, $\bar{S}(\bar{k}_\eta)$, measured over a wide range of physical situations, reveals that the spectra at the higher wave numbers collapse to a universal form of $k^{-5/3}$ and, as the Reynolds number decreases, the spectra show shorter ranges of the universal behavior [21]. Scott [20] produced the following correlation for $\bar{S}(\bar{k}_\eta)$ as a function of $\bar{k}_\eta$ and Re_λ:

$$\bar{S}(\bar{k}_\eta) = \frac{0.01668\,\mathrm{Re}_\lambda^{2.5} + 3.74\,\mathrm{Re}_\lambda^{0.9} - 70\,\mathrm{Re}_\lambda^{-0.1}}{1 + (0.127\,\mathrm{Re}_\lambda^{1.5}\bar{k}_\eta)^{5/3} + (1.15\,\mathrm{Re}_\lambda^{0.622}\bar{k}_\eta)^4 + (1.27\,\mathrm{Re}_\lambda^{0.357}\bar{k}_\eta)^7} \tag{16}$$

The good agreement between Eq. (16) and the data of other research is seen in Fig. 6, where the symbols are the measured values reported in Ref. [21] and the lines correspond to the values obtained from the correlation.

At an elapsed time of t_k, the rms turbulent velocity effective in wrinkling the flame front, u'_k, is given by

$$\frac{u'_k}{u'} = \left(\frac{15^{0.5}}{\mathrm{Re}_\lambda} \int_{\bar{k}_\eta}^{\infty} \bar{S}(\bar{k}_\eta) d\bar{k}_\eta \right)^{1/2} \tag{17}$$

Values of u'_k/u' were obtained at different values of $\bar{k}_\eta$ by numerical integration of Eq. (17) using Eq. (16). Shown in Fig. 7 are such variations of u'_k/u' plotted against $\bar{k}_\eta^{-1}$ for a range of values of Re_λ. Hence, the values of u'_k/u' for different t_k are computed and are plotted in Fig. 8 for the experimental conditions of the present study. It is evident in Fig. 8 that the flames are not fully developed turbulence even after 15 ms elapsed time from ignition and it emphasizes the role of developing turbulence effects in the modeling of turbulent flame propagation in engines.

Experimental Results and Discussions

Shown in Fig. 9 are the probability density functions (pdfs) of curvatures of a stoichiometric methane-air flame fronts at three

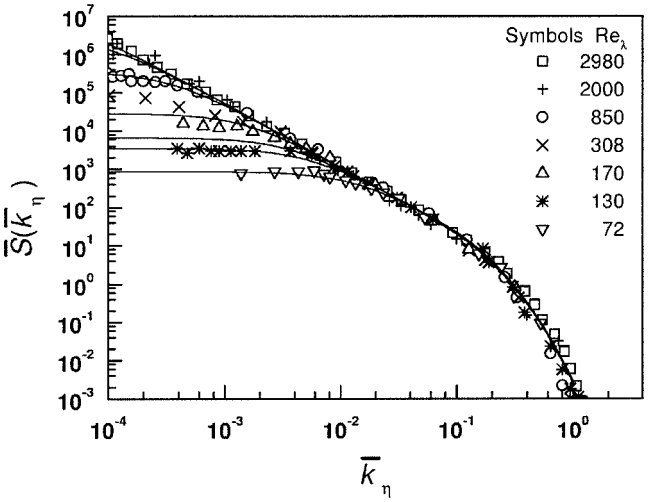


Fig. 6 Generalized psd function showing the spectrum of turbulent energy

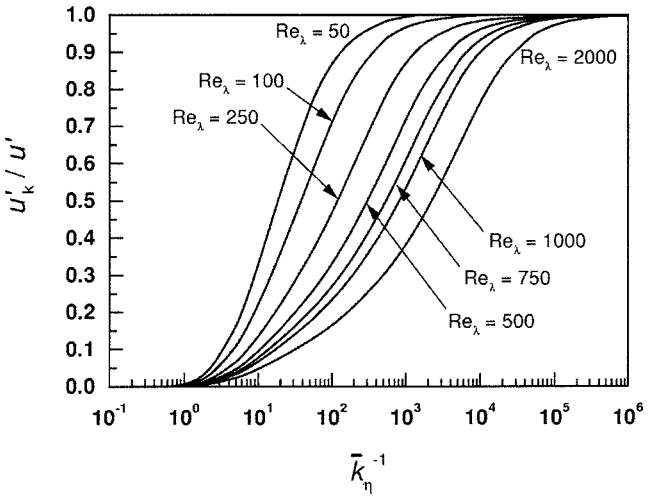


Fig. 7 Development of effective rms turbulent velocity

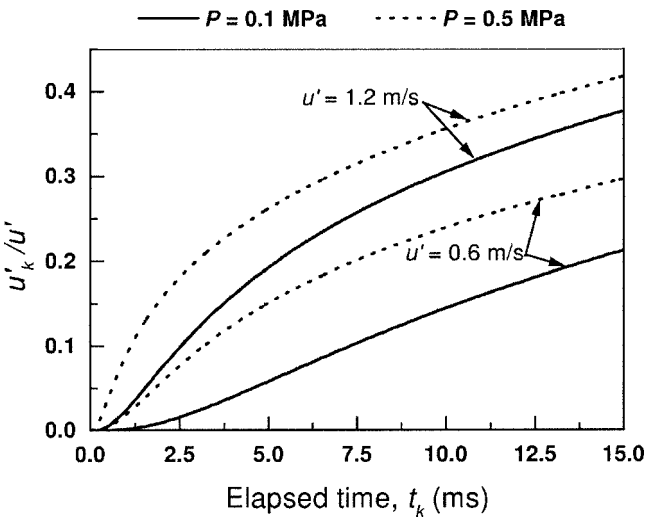


Fig. 8 Temporal development of the effective rms turbulent velocity

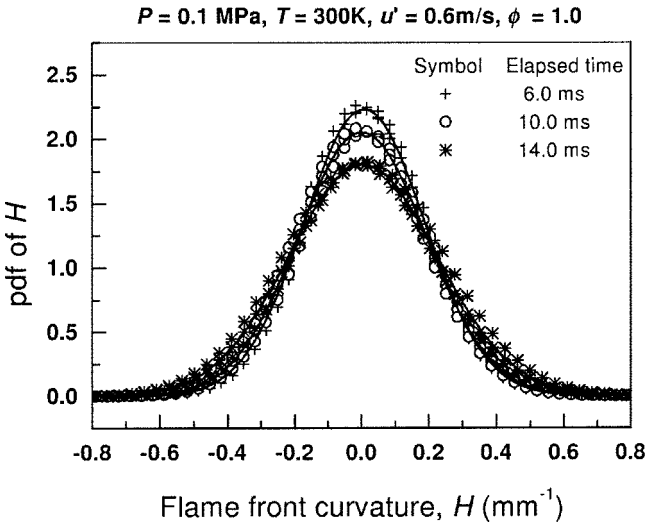


Fig. 9 Curvature pdfs for a developing methane-air flame front at three different elapsed times

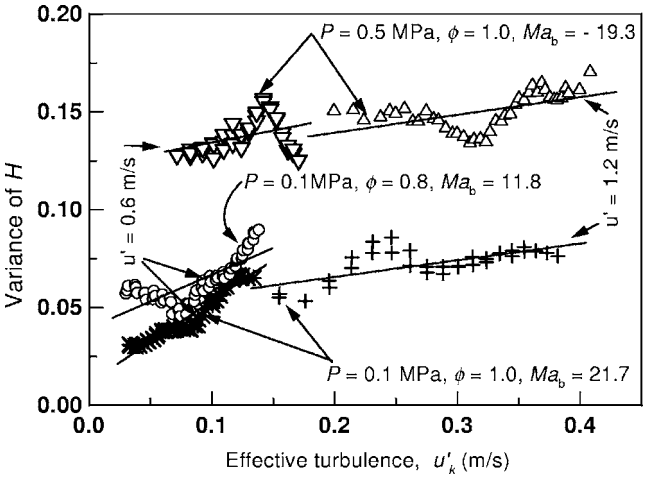


Fig. 11 Variance of curvature pdfs of methane-air flame fronts at different initial conditions of equivalence ratio, pressure, and rms turbulent velocity

different elapsed times. The pdfs are symmetric with respect to a very small negative mean value. The widths (variances) are increased with time, commensurate with increased flame wrinkling.

From such flame curvature pdfs, reported in Fig. 9, it is possible to compute the positive, negative, and overall mean flame curvatures H^+, H^-, and H^o, respectively. Shown in Fig. 10(*a*) are the values H^+, H^o, and H^- plotted against elapsed time for three different explosions at $u' = 0.6$ m/s. The positive and negative values of curvature clearly increase with time, while the mean remains close to zero. Shown in Fig. 10(*b*) are the values of variance of curvature, obtained from these flame propagation.

The influence of the effective rms turbulent velocity, u'_k, and Markstein number, Ma_b, on turbulent flame propagation are clearly revealed in Fig. 11, which shows an increasing variance of H as u'_k is increased. At a given value of u'_k and similar Markstein numbers, the magnitude of variance of H are similar, however lower values of Markstein numbers are associated with the higher variance of the curvature pdfs. This observation is in line with the flame stability analysis which reports that flames with lower value of Markstein numbers to show greater propensity to cellular flame to occur at smaller flame radii [15,17].

Shown in Fig. 12 are the values of normalized turbulent burning velocities defined at R_v, u_t/u_l plotted as a function of normalized effective rms turbulent velocity, u'_k/u_l for different initial conditions. Hence, laminar burning velocity, u_l, is an intrinsic property of combustible fuel-air premixture and is defined as "the velocity, relative to and normal to the flame front, with which the unburned gas moves into the front and is transformed into products under laminar flow conditions" [22]. All these dimensionless data, shown in Fig. 12, suggest that, for same values of u'_k/u_l, values of u_t/u_l are increased with a decrease in Markstein number, and for the same values of Markstein numbers the values of $u_t(R_v)/u_l$ are found to increase with an increase in the values of u'_k/u_l. These all signify the importance of both the effective rms velocity and Markstein number in the turbulent burning of methane-air premixtures.

Measured values of turbulent burning velocities and turbulent flame speeds can be related analytically when both of these values are defined at a reference radius, R_v. Hence, mass of the burned gas at any time can be written as $(4\pi/3)R_v^3\rho_b$. Hence the left-hand side of Eq. (10) can be written as

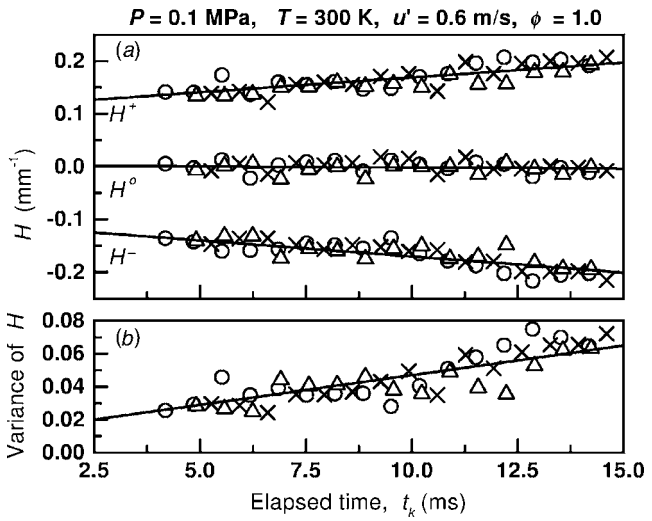


Fig. 10 (*a*) Positive, negative, and mean flame curvatures. (*b*) Variance of pdf of curvatures, plotted as a function of elapsed time from ignition.

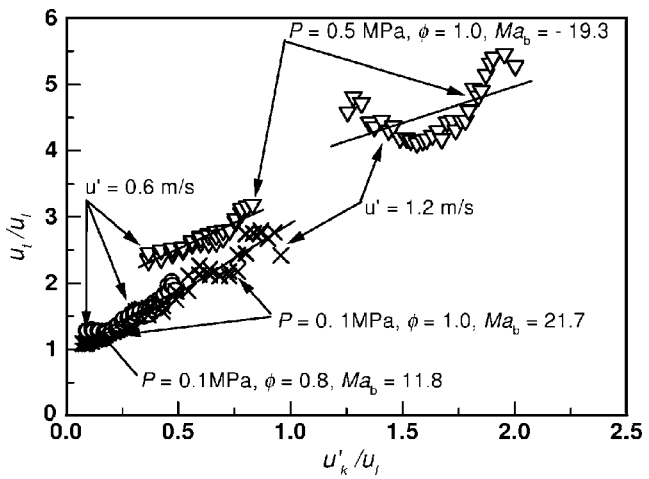


Fig. 12 Variation of $u_t(R_v)/u_l$ with u'_k/u_l for methane-air premixtures at different initial conditions

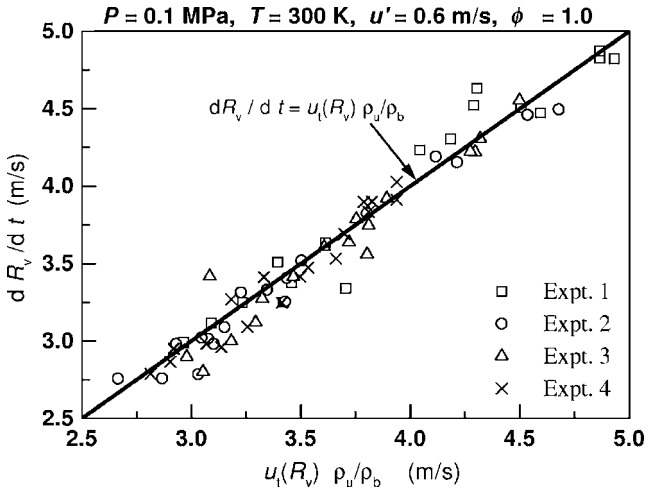


Fig. 13 Relationship between turbulent flame speeds and turbulent burning velocities defined at R_v

$$\frac{d}{dt}(m_{bi}+m_{bo})=\frac{d}{dt}\left(\frac{4\pi}{3}R_v^3\rho_b\right)=4\pi R_v^2\rho_b\frac{dR_v}{dt} \tag{18}$$

Equating Eqs. (10) and (18), turbulent flame speed, dR_v/dt, can be related to turbulent burning velocity, $u_t(R_v)$, by

$$\frac{dR_v}{dt}=\frac{\rho_u}{\rho_b}u_t \tag{19}$$

Equation (19) also has been verified using experimental data and an example is shown in Fig. 13, where the values of dR_v/dt are plotted as a function of $u_t(R_v)\rho_u/\rho_b$, using data obtained from four different experiments of stoichiometric methane-air premixtures at $u'=0.6$ m/s. The experimental results satisfy the relationship quite closely; the observed scatter probably arises from the numerical methods employed to estimate these values.

Conclusions

The paper reports the effects of developing turbulence and Markstein number on the propagating flame kernel in isotropic methane-air premixture. Two-dimensional PMS images are analyzed to derive the statistics of flame front curvatures and turbulent burning velocities. The major conclusions of the study are

- Flame front curvature distributions become more dispersed with increase in elapsed time from ignition.
- Consideration of developing turbulence is important in determining the pdf of curvatures for flame propagation in turbulent premixture.
- For same effective rms turbulent velocity, flame front wrinklings are reasonably similar as long as the corresponding Markstein numbers are similar.
- A decrease in Markstein number increases the flame front curvature dispersion, in line with the flame front stability studies. A flame with a negative Markstein number is more wrinkled than a flame with a positive Markstein number.
- Turbulent burning velocities increase with increase in effective rms turbulent velocity and with decrease in Markstein number.

Acknowledgment

The author acknowledges the friendship and co-operations of the members of the Combustion Group of the School of Mechanical Engineering, The University of Leeds, Leeds, UK.

Nomenclature

$\underline{H}$ = mean flame curvature
k_η = dimensionless wave number
L_b = burned gas Markstein length
Ma_b = burned gas Markstein number, $\mathrm{Ma}_b=L_b/\delta_l$
m_{bi} = mass of burned gas within any sphere of general radius, R_i
m_{bo} = mass of burned gas outside any sphere of general radius, R_i
m_u = mass of unburned gas outside radius, R_t
m_{ui} = mass of unburned gas within any sphere of general radius, R_i
m_{uo} = mass of unburned gas outside any sphere of general radius, R_i
P = pressure
R_i = any general radius between R_r and R_t
R_r = root radius, within which all gas is burned
R_t = tip radius, beyond which all the gas is unburned
u' = rms turbulent velocity
u_l = unstretched laminar burning velocity
u_k' = effective rms turbulent velocity acting on the flame
α = flame stretch
δ_l = laminar flame thickness
η = Kolmogorov length scale of turbulence
λ = Taylor length scale of turbulence
ρ_u = unburned gas mixture density
ρ_b = burned gas mixture density
ϕ = equivalence ratio

References

[1] Ting, D. S., Checkel, M. D., and Johnson, B., 1995, "The Importance of High-Frequency, Small-Eddy Turbulence in Spark Ignited, Premixed Engine Combustion," SAE Paper No. 952409.
[2] Abdel-Gayed, R. G., Bradley, D., and Lawes, M., 1987, "Turbulent Burning Velocities: A General Correlation in terms of Straining Rates," Proc. R. Soc. London, Ser. A, **414**, pp. 389–413.
[3] Abdel-Gayed, R. G., Al-Khishali, K, J., and Bradley, D., 1984, "Turbulent Burning Velocities and Flame Straining in Explosions," Proc. R. Soc. London, Ser. A, **391**, pp. 393–414.
[4] Checkel, M. D., and Ting, D. S., 1993, "Turbulence Effects on Developing Turbulent Flames in a Constant Volume Chamber," SAE Paper No. 930867.
[5] Shen, H., Hinze, P. C., and Ting, J. B., 1996, "A Study of Cycle-to-Cycle Variations in SI Engines Using a Modified Quasi-Dimensional Model," SAE Paper No. 961187.
[6] Ting, D. S., and Checkel, M. D., 1997, "The Importance of Turbulence Intensity, Eddy Size and Flame Size in Spark Ignited, Premixed Flame Growth," Proc. Inst. Mech. Eng., Part D (J. Automob. Eng.), **221**, pp. 83–86.
[7] Merdjani, S., and Sheppard, C. G. W., 1993, "Gasoline Engine Cycle Simulation Using Leeds Turbulent Burning Velocity Correlations," SAE Paper No. 932640.
[8] Bradley, D., 1992, "How Fast Can We Burn?," Proc. Combust. Inst., **24**, 247–262.
[9] Williams, F. A., 2000, "Progress in Knowledge of Flamelet Structure and Extinction," Prog. Energy Combust. Sci., **26**, pp. 657–682.
[10] Echekii, T., and Chen, J. H., 1996, "Brief Communication: Unsteady Strain Rate and Curvature Effects in Turbulent Premixed Methane-Air Flames," Combust. Flame, **106**, pp. 184–202.
[11] Haq, M. Z., Sheppard, C. G. W., Woolley, R., Greenhalgh, D., and Locket, R. D., 2002, "Wrinkling and Curvature of Laminar and Turbulent Premixed Flames," Combust. Flame, **131**, pp. 1–15.
[12] Lee, T. W., North, G. L., and Santavicca, D. A., 1993, "Surface Properties of Turbulent Premixed Propane/Air Flames at Various Lewis Numbers," Combust. Flame, **93**, pp. 275–288.
[13] Shepherd, I. G., and Cheng, R. K., 2001, "The Burning Rate of Premixed Flames in Moderate and Intense Turbulence," Combust. Flame, **127**, pp. 2066–2075.
[14] Gu, X. J., Haq, M. Z., Lawes, M., and Woolley, R., 2000, "Laminar Burning Velocity and Markstein Lengths of Methane-Air Mixtures," Combust. Flame, **121**, pp. 41–58.
[15] Haq, M. Z., 1998, "Fundamental Studies of Premixed Combustion," Ph.D. thesis, The University of Leeds, UK.

[16] Bradley, D., Lau, A. K. C., and Lawes, M., 1992, "Flame Stretch Rate as a Determinant of Turbulent Burning Velocity," Philos. Trans. R. Soc. London, Ser. A, **338**, pp. 359–387.
[17] Haq, M. Z., 2005, "Correlations of the Onset of Instabilities of Spherical Laminar Premixed Flames," ASME J. Heat Transfer, **127**, pp. 1410–1415.
[18] Haworth, D. C., and Poinsot, T. J., 1992, "Numerical Simulations of Lewis Number Effects of Turbulent Premixed Flames," J. Fluid Mech., **224**, pp. 405–436.
[19] Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, *Numerical Recipes in Fortran 77: The art of scientific computing*, Cambridge Univ. Press, Cambridge, UK.
[20] Scott, M. J., 1992, "Distributions of Strain Rate and Temperature in Turbulent Combustion," Ph.D. thesis, The University of Leeds, UK.
[21] McComb, W. D., 1990, *The Physics of Fluid Turbulence*, Oxford University Press, Oxford, UK.
[22] Heywood, J. B., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.

Journal of Engineering for Gas Turbines and Power

Technical Briefs

Tuning Your Fuel-Gas Delivery System

T. D. Newbound
e-mail: timothy.newbound@aramco.com

K. S. Al-Showiman

Saudi Aramco, R&D Center, Box 62, Dhahran 31311, Saudi Arabia

Saudi Aramco has focused more attention in recent years on fuel-gas conditioning for land-based industrial and aeroderivative combustion gas turbines (CGTs). Hydrocarbon dew points and entrained solids are two important fuel quality issues that frequently trouble CGT operators, partly because they cannot be guaranteed by the fuel suppliers and they are rarely monitored by the operators. This paper addresses these issues and offers some practical advice to optimize the design and operation of fuel gas delivery systems. Saudi Aramco has been testing an automated on-line dew point monitor capable of detecting both hydrocarbon and aqueous dew points in natural gas. Dew point monitoring, conducted at three locations on the fuel gas grid, revealed wide variations in the hydrocarbon and aqueous dew points due to a variety of conditions. Gas production and pipeline operations accounted for the most dramatic variations in dew points, but exposure of pipelines to the weather can also be important. Measurement of pipeline solids for the purpose of sizing and placement of particle filters have also been explored. Pipeline scraping, gas velocities, length of pipeline span, pipeline junctions, and control valves are all considerations for solid control strategies. The optimized design and operation of a CGT fuel system is highly dependent on dew point control and efficient removal of entrained pipeline solids. Practical experience in monitoring hydrocarbon and aqueous dew points, pipeline solids control, and optimizing fuel conditioning equipment are considered.
[DOI: 10.1115/1.2031267]

Introduction

Saudi Arabia's main source of electrical power is provided by combustion gas turbine generators (CGTGs). Most of the CGTGs operating in the Eastern and Central Regions of Saudi Arabia are currently fueled by natural gas. This has been made possible by a rapidly growing sweet gas industry comprised of five major gas processing plants and thousands of kilometers of gas pipelines. As new gas sources, processing facilities and pipelines have been added to the gas grid, the system as a whole has become increasingly more complex and diagnosis of irregularities in gas composition more challenging. While Saudi Aramco's sales gas specifications mirror sales gas standards worldwide, these standards do not guarantee adequate fuel condition for operating combustion gas turbines (CGTs).

One gas-fuel specification generally required by CGT manufacturers, but not part of sales gas specifications, is the 50°F superheat requirement. That is, inlet fuel-gas temperatures must be at least 50°F (28°C) above the dew point calculated from the fuel-gas composition. The safety margin imposed by the superheat margin mainly allows for Joules Thomson cooling as fuel expands through its path to the combustion chamber. There is some room for error in the measurement of fuel-gas compositions, but accuracy is important, especially in the C6+ range where small errors in the concentrations of heavy hydrocarbons can translate to large dew point shifts. Perhaps more important is the potential for gas compositions to vary within any given 24 h period due to gas process changes, pipeline operations, and even changes in the weather. Of course, the potential for these changes will depend on the sources of the fuel gas, complexity of the fuel gas system, and the environment.

Another requirement for CGT gas fuel is that it must be delivered free of solids. Some sales gas contracts have general statements addressing solids, but none to our knowledge have defined a quantitative limit for solids. Perhaps this is because there is no recognized standard for measuring solids entrained in natural gas pipelines.

Previous publications have addressed methods of measuring fuel-gas dew points and fuel treatment to eliminate liquids [1]. We have recently published findings from the first two phases of our dew point monitoring campaign [2]. These and other publications describe the distress to CGT hot gas path parts that can result from liquid carryover [3]. There are plenty of products on the market to treat most any fuel-gas condition, but careful diagnosis of the fuel-gas condition over an adequate time period is where many operators fail to zero in on their fuel treatment needs. Realization of this gap led Saudi Aramco to begin a dew point monitoring campaign on its fuel-gas grid.

Discussion

Dew Point Determinations. Detection of entrained liquids in natural gas-fuel supplies is difficult and not practical for continuous on-line fuel-gas analysis. However, determination of aqueous and hydrocarbon dew points can predict conditions that are favorable for liquid formation. Although it has been common practice for the design basis of CGT fuel systems to be based on single (spot) sample analyses, on-line monitoring has become much more practical, especially where multiple CGTs are operated on the same fuel source. Improvements and cost reductions in microprocessors over the past 10 years have made automated fuel monitoring more reliable and affordable.

The *gas chromatograph (GC)* is the most widely used on-line analytical tool for fuel-gas analysis, although it is usually used for

Contributed by the International Gas Turbine Institute of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received June 6, 2004; final manuscript received September 27, 2004. Review conducted by L. S. Langston.

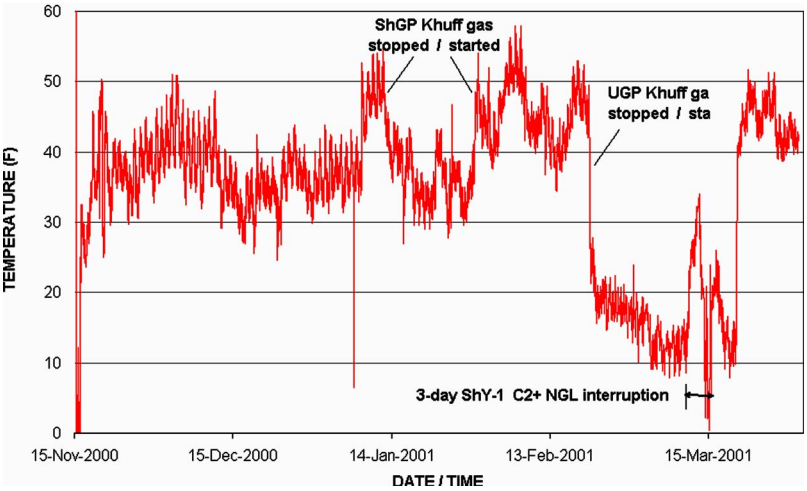


Fig. 1 Hydrocarbon dew points recorded over a three month period reveal dramatic changes that coincide with major operational events in the gas plants

heating value determinations (determines hydrocarbons groups up to C6+). There are commercial GCs equipped for extended gas analysis that are capable of resolving individual hydrocarbon groups out to C8 and C9+. These instruments now have a relatively short analytical cycle (5–10 min) and are capable of high reproducibility within 5°C of the actual hydrocarbon dew point. Furthermore, they tolerate corrosive gases such as H_2S and CO_2. However, without the extended capability to analyze C9+, they are of little value for dew point monitoring. GCs configured to monitor hydrocarbons cannot measure moisture in gas.

Direct dew point measurements using an automated Bureau-of-Mines dew point monitor can afford highly accurate dew point measurements for hydrocarbon dew points from −40°C and +35°C. Saudi Aramco has been operating an Ametek Model 241 dew point analyzer for five consecutive years at three sites on the fuel gas grid in Eastern and Central Provinces of Saudi Arabia. During the first two phases of monitoring activity, hydrocarbon dew points in sales gas were observed to vary as much as 40°F in one 24 h period primarily due to gas plant operations (Fig. 1). It was later shown that hydrocarbon dew points combined with CGTG control data, provided direct indications of liquid hydrocarbon carryover past the CGT fuel control valve (Fig. 2).

In the latest phase of monitoring on Saudi Arabia's East/West Pipeline, water dew points were observed for the first time above the hydrocarbon dew points. Previously, water dew points were not observed because they were always below the hydrocarbon dew points, and the monitor is programmed to terminate the cooling cycle when it finds the hydrocarbon dew point. Otherwise, the monitor would not readily recognize the water condensation point on surfaces that are already coated with liquid hydrocarbon.

Gas transported though the East/West Pipeline is primarily derived from three gas processing plants: UGP, HGP, and HDGP. Average daily production rates are compared in Fig. 3. Gas from these sources is combined in stages, eventually joining the pipeline that feeds the dew point monitoring site. Thus, tracing the source of moisture was not straightforward.

The first appearance of moisture (January 20, 2003, Fig. 4) coincides with a significant drop in the production rate of a gas processing module from UGP. This module processes nonassociated gas derived from a reservoir known as the "Khuff." The Khuff gas is processed at higher-pressure with higher-temperature liquid recovery and less stringent dehydration. Consequently, the

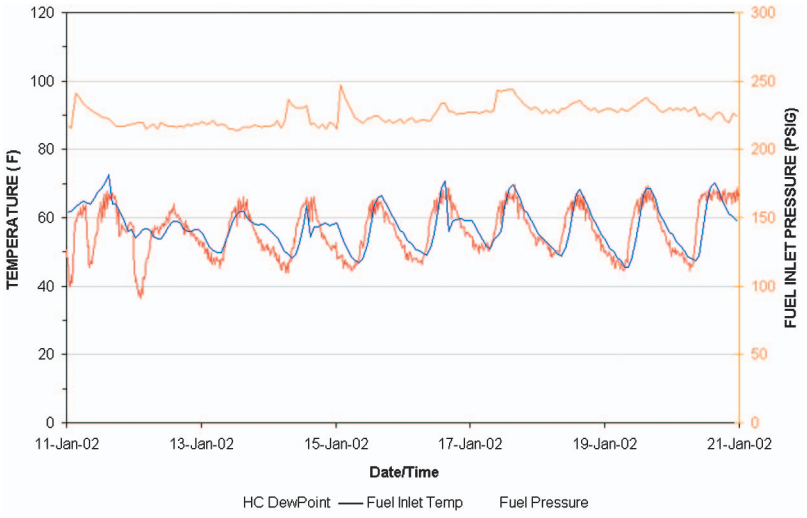


Fig. 2 (Color) Hydrocarbon dew points plotted with archived fuel inlet temperature and pressure for one CGTG unit at a power plant feed by the local fuel-gas system near the central Ghawar region gas plant UGP. Note that saturated fuel gas conditions coincide with elevated fuel pressure.

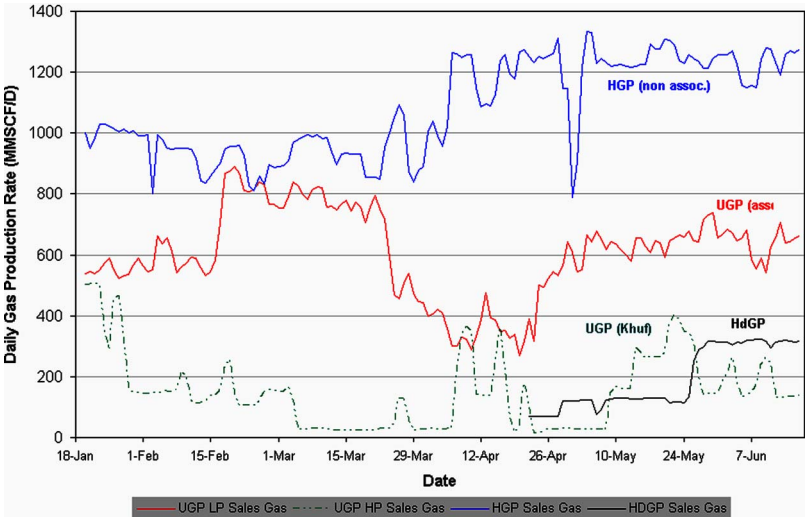


Fig. 3 Average daily production rates from South Ghawar Gas Plants

Khuff gas has more C6+ and has more moisture relative to gas processed from the low pressure associated gas processing modules. The reduction of Khuff gas produced from the UGP high-pressure module appears to coincide with a drop in the hydrocarbon dew point, although less dramatic than its effect in 2001 when HGP was not producing gas (cf. Fig. 1). In this case, HGP gas becomes a larger fraction of the overall sales gas, but this does not explain the high level of moisture.

Further inquiries led to information about a new pipeline installation from HDGP to UGP, which had just recently undergone a hydrotest. The hydrotest drying protocol were conducted during February and March before the pipeline was finally commissioned on April 28. The commissioning of the new pipeline coincided with the first production from the new gas plant HDGP. Following the opening of the new gas plant and pipeline to the sales gas system, the water dew points were recorded in a regular pattern exceeding the 7 lbs/MMSCF (20°F for 550 psig) sales gas specification for moisture (Fig. 5). Peak water dew points were recorded at 32°F, which at 550 psig equates to 11 lbs/MMSCF. The high moisture contents persisted for about two months. Fortunately, this event occurred during a relatively warm season when gas temperatures were above 85°F. Thus, no incidents from use of sales gas were reported as a consequence of high dew points.

By August, indications of high moisture had all but completely disappeared. Water dew points dropped below the hydrocarbon dew points, and hydrocarbon dew points appeared to stabilize in a familiar pattern, although the average dew points were significantly higher that those observed earlier that year (Fig. 6). This is expected as more Khuff gas is produced during the summer to meet the higher demand for energy.

Pipeline Solids. The Saudi Sales gas grid customers have experienced intermittent filter skid plugging, control valve erosion, and plant upsets due to entrained solids. Indeed, worldwide, corrosion solids appear to be the most prominent contamination problem for natural pipelines and compression equipment [4].

Entrainment of solids in natural gas streams involves a complex array of chemical and physical conditions. The composition of pipeline solids, aptly described as "black powder" is generally a blend of iron oxides and iron sulfides derived from pipeline mill scale and chemical sulfidation (reaction of hydrogen sulfide with iron). Moisture can create even more serious problems by activating otherwise dormant microbial corrosion from sulfate reducing

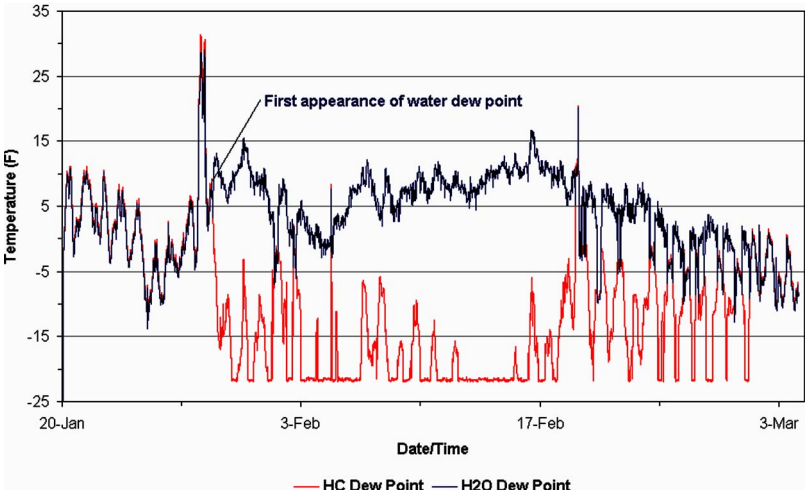


Fig. 4 On-line hydrocarbon and water dew points recorded with the Ametek 241 CE dew point monitor at Pump Station 3 (SRG 1 sales gas)

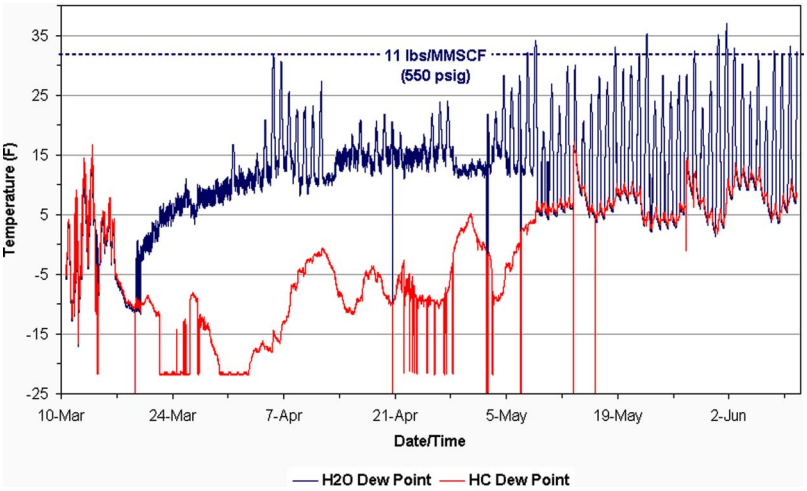


Fig. 5 On-line hydrocarbon and water dew points recorded with the Ametek 241 CE dew point monitor at E/W Pipeline Pump Station 3 from March 10–June 10, 2003

bacteria and acid producing bacteria. Solids from pipeline hydrotests have also contributed to the solids problem with a variety of iron hydroxides.

Since most CGT operators usually have little if any control over the pipeline operations upstream from their own fuel delivery piping, their only option may be to remove solids at the entrance of their fuel system by some form of filtration or centrifugal separation. Determining which filtration options will be most effective for a particular fuel gas system is often not a trivial task. Where there are existing filters in equivalent locations, or for filter upgrades, it is often possible to collect representative samples of the solids from spent filter cartridges out of existing filters. From these solids, it is possible to determine composition particle size distribution, grain density, and bulk density. Perhaps it would also be possible to estimate of the quantity of solids over a period of time. The difficulty in measuring quantities of solids at a particular location is that, like dew points, solids entrainment in gas streams may vary from one moment to the next. It can also vary widely from location to location.

One campaign to measure solids entrained at various points in the northern end of Saudi Aramco's Sales gas system was undertaken during summer of 2001. Samples were acquired from six different locations over a nine-hour period by flowing a steady state-stream of the gas through a solids trap at equivalent flow rates (~2 SCF/min), then measuring the mass of solids collected in the trap. The amounts varied from 0.006 to 94 g/MMSCF, or more than four orders of magnitude. Although the sampling method used in this study is not a recognized standard method, it illustrates the potential for large variations in the amount of entrained solids from one position to another on a fuel-gas grid.

Ostensibly, the entrainment of solids in a pipeline will be dependent on such conditions as flow velocity of gas, geometry of the piping upstream of the sampling point, and the size distribution and particle density of the solids. However, other conditions, such as the presence or absence of residual compressor oils or glycols, high hydrocarbon, or water dew points, can have significant effects on solids entrainment. Furthermore, pipeline operations, such as opening and closing pipeline bypasses and pipeline scraping, can release and entrain solids in sudden and severe slugs.

Filtering. The uncertainty that remains after a careful analysis of pipeline solids generally relates to sizing of filtration apparatus. For example, if large volumes of pipeline solids are expected, it

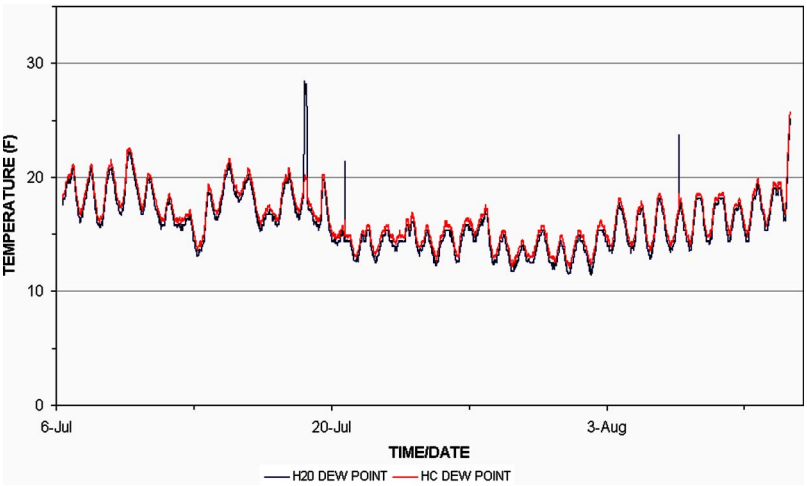


Fig. 6 On-line hydrocarbon and water dew points recorded with the Ametek 241 CE dew point monitor at E/W Pipeline Pump Station 3 from July 7–August 12, 2003

may be necessary to include a cyclone or magnetic trap upstream of the particle filter to remove large volumes of material and avoid having to change filter cartridges on a frequent basis. After particle filters have been put into service, it is important to closely monitor changes in pressure differential across the filter. Filter cartridges should be changed on a regular schedule or sooner if high-pressure differential indicates that the filter loading is high. A sudden drop in pressure differential to zero indicates that the filter has been breached and the element should be changed immediately.

Coalescing filters should be located between the particulate filters (upstream) and the gas superheater (downstream). Coalescing filters designed for natural gas-fuel systems are generally not designed for and should not be used as solid particulates filters. Coalescing filters with absolute ratings to 0.3 microns are common and available from several manufacturers. With adequate protection from solid particulates, more expensive coalescing filter elements should not require frequent replacement.

Superheating. The final stage of fuel conditioning should be the fuel superheater. To assure that the superheater is truly delivering superheated fuel, the feed to the superheater should be practically free of liquid and solid particulates above 0.3 microns. Regardless of what type of fuel superheater is being used, good prefiltration is important. This assures that heating actually provides superheat, and not just partial fuel vaporization.

Overheating fuel beyond the minimum superheat will have little effect on the health of the turbine, but is wasteful of energy. It can also decrease the overall power output of the turbine when the fuel supply pressure is limited. Depending upon the circumstances at a particular site, different strategies for applying superheat may be used.

In Phase 1 of our dew point monitoring program, fuel preheating was provided by refinery steam. Energy loss from the use of refinery steam was not as much of a concern as the health and reliability of the refinery's two utility CGTGs. Thus, the refinery Utilities chose to operate their fuel superheaters at 135°F year round.

Dew point monitoring during Phase 2 revealed that the fuel gas supply for a major power plant in the South Ghawar area was saturated in heavy hydrocarbons during cool weather (about 5 out of 12 months/year). Without fuel preheating, liquid hydrocarbons were reaching the CGT hot gas path. Gas-fired fuel preheaters that previously had been out of service because they were troublesome to operate, were put back into service, but only during the cool season. Power plant operators were able to set the superheater outlet temperature to approximate minimum superheat conditions and establish the time period that fuel superheating would be necessary.

In Phase 3, the dew points of Sales gas that feeds the entire central region of Saudi Arabia was monitored at a single site. The fuel systems in this region are new and most are equipped with electrical fuel superheaters. By continuing to monitor dew points at this single site, it will be possible to synchronize the entire fleet of operating CGT fuel superheaters in order to provide the minimum energy input to achieve minimum superheat. Although the power required for operating one superheater a few degrees higher than necessary may not amount to much, the savings in energy could be substantial with tight fuel temperature control of an entire fleet of a dozen or more CGTs.

Conclusions

This work emphasizes the benefits of on-line monitoring in gas-fuel systems as a means of preventing damage to CGT hot section components, while also saving energy and maximizing the power output potential of the CGT. In cases where it is suspected that damage has occurred to CGTs as a result of liquid carryover, it is often difficult to prove without recorded fuel monitoring.

Sales gas contracts do not provide adequate assurance that fuel properties will meet manufacturer's specifications for CGTs. Although specific limits for moisture, H_2S, and total sulfur are typically included in sales gas specifications, quantitative limits for entrained solids and hydrocarbon liquids cannot be guaranteed. Documented upset conditions have shown that even moisture cannot be controlled at times.

CGT operators are responsible for providing fuel to their CGTs that meets the manufacturer's specifications. Especially on complex fuel-gas grids, only continuous monitoring, not single sample analysis, provides an accurate representation of the range of fuel conditions. The savings from the reduction in CGT repair costs and energy savings in Saudi Arabia as a result of dew point monitoring has been estimated at over $3 million per year. The cost of dew point monitoring has been less that $120 thousand over the five-year study including the capital cost of the dew point monitor.

Entrained solids have been a growing problem as the Saudi sales gas grid has been expanding. Saudi Aramco has been actively pursuing remedies designed to improve the cleanliness of the sales gas system. Advanced designed pipeline scraping and chemical cleaning methods have been used in an ongoing effort to reduce pipeline solids. Consideration has been given to installing particle separators at strategic pipeline junctions. But, ultimately, the customers will be required to provide fuel system filtering to meet their specific fuel quality specifications.

Acknowledgment

The authors wish to express appreciation to Saudi Aramco, and the Saudi Arabian Ministry of Petroleum and Mineral Resources for granting permission to publish this material. Credit is due to many individuals at the three field locations who assisted in the installation and commissioning of the dew point monitoring instrumentation in their facilities. In this regard, we appreciate the cooperation of the plants' managements and their engineering, operations, and maintenance staffs.

References

[1] Wilkes, C., and Dean, A J., 1997, "Gas Fuel Conditioning Design Considerations for Utility Gas Turbines," ASME Paper No. 97-GT-227.
[2] Newbound, T. D., and Wagiealla, W. S., 2003, "On-Line Hydrocarbon Dew Point Monitoring in Fuel Gas," ASME Paper No. GT2003-38868.
[3] (a) Newbound, T. D., and Norek, R. S., 1997, "Contaminants in a Gas Turbine Fuel System and Thermal Shock," ASME Paper No. 97-GT-307. (b) Norek, R. S., 1992, "Operating Experience with Nozzles and Blades in MS-5002 Gas Turbines in a Desert Environment," *Proc. Steam and Combustion Turbine Blading Conference*, Electric Power Research Institute (Edison Electric Institute).
[4] Baldwin, R. M., 1998, "Black Powder in the Gas Industry," SRI Mechanical and Fluids Engineering Division Technical Assessment, Gas Machinery Research Council.

Gas Turbine and Generator Test Experience in Support of the Advanced Hull Form Inshore Demonstrator (AHFID)

Roger J. Carr, Jr.

Denis Lison

Fred T. Willett

DRS Power Technology Inc.,
Fitchburg, MA 01420

This paper describes a test program designed to assess the suitability of a TF40-based gas turbine generator set (genset) for use in the Advanced Hull Form Inshore Demonstrator (AHFID). An overall description of the test program and its objectives is given. Test results are presented. The relevance of the test to the greater program goal of demonstrating the viability of rim-driven propulsion technology is discussed. The genset demonstration test was successfully completed in 2003. The package performed as designed and met all test objectives. [DOI: 10.1115/1.2061327]

Introduction

Podded propulsion is believed by many experts to be the propulsion technology of the twenty-first century. Proposed applications for podded propulsion run the gamut of marine duty, from cruise ships to icebreakers. Pod drives are ship propulsors with two distinguishing characteristics: (i) an electric motor is located inside a pod outside of the ship hull, and (ii) the total unit pivots in a horizontal plane for steering and maneuvering the ship. To coin a phrase, podded propulsion is "thinking outside the hull."

Podded propulsion is enabled by the all-electric ship concept because electric power generated onboard the ship can be directed to the pod. Podded propulsion creates advantages in maneuverability, space requirements, and cost. Another advantage is increased manufacturing flexibility, not only because of the design, but also because the pod can be installed late in the production process. With podded propulsion, the ship's construction is no longer driven by the propulsion system. Likewise, the separation of propulsor from ship opens up the design envelope for propulsion components. Because the propulsor shaft has been eliminated, the wake distribution is more uniform. Another advantage is low propeller noise. The use of electric drive also means that a wide range of propeller speeds are possible, as are very low speeds. A podded propulsion concept is shown in Fig. 1.

Within the field of podded propulsion, a new concept has emerged [1,2]. Traditional podded propulsion has relied on hub-drive pods, where (as the name implies) the electric motor is located at the center of the pod, in the hub. The concept of rim-driven propulsor pods consists of a ducted, multiple blade-row propulsor with a permanent magnet (PM), radial field motor rotor mounted on the tips of the propulsor rotor blades, and the motor stator mounted within the propulsor duct. Compared to hub-driven pods, the rim-driven propulsor pod concept offers several benefits, including improved efficiency. The benefits translate to even greater flexibility in ship design.

The goal of the Advanced Hull Form Inshore Demonstrator (AHFID) test program is to provide an at-sea demonstration of rim-driven propulsion technology that will create a foundation for further developments in electric propulsion technology. One step in achieving that goal is the development and test of a gas turbine generator unit that will supply the necessary power to the drive motor.

The AHFID is a gas turbine generator set (genset) designed specifically suited to the marine environment. Its function is to provide electric power to a permanent magnet motor. Gas turbine output power is fed to a six-pole generator via a single-reduction, epicyclic gearbox system and flexible disk-pack-type coupling. The system is designed to use diesel fuel.

Configuration

The prime mover for the genset is a Vericor TF40 gas turbine capable of 2983 kW (4000 shp) continuous power and 3430 kW (4,600 shp) boost power at standard ISO conditions. The output shaft speed at maximum continuous output is 15,400 rpm. The TF40 is both ABS and US Navy certified. The TF40 was selected for its experience and because the risk of not meeting performance or operability goals was low.

An existing General Electric 700 V DC generator was modified into a 530 V, 350 Hz, four winding (three phases each) AC generator for the test. The original generator was rated at 3000 kW three-phase delta connected with a 700 V DC regulated output voltage. It is a six-pole wound field synchronous machine with a design speed of 7000 rpm. The generator consists of four independent three-phase delta connected windings. Each three-phase winding produces 530 V AC (rms), 350 Hz that is coupled through a power rectifier and interphase transformers to provide the 700 V DC regulated output voltage. The generator has two integral air coolers mounted above the stator, and the exciter is brushless. The dual automatic voltage regulator, which presently senses and regulates DC output voltage, is mounted integral to the generator housing. The generator bearings are the tilt-pad journal type. The generator lubrication oil system is common with that of the reduction gear. The generator operated in an isochronous mode only (no droop compensation required).

The generator was modified from the existing 700 V DC machine into one that produces four independent three-phase 530 V, 350 Hz loads. The modification to an AC machine included elimination of the existing power rectifier diodes and interphase transformers. Modification to the generator-sensing circuitry was required to maintain an analogous and isolated signal to the voltage regulator to provide regulated voltage under varying load conditions. Figure 2 shows the generator after refurbishment.

The reduction gearbox was an Artec-Maag MPF-13, 15,400 rpm input, 7000 rpm output single-helical unit. It was also chosen for its low risk and similarity to proven designs. Risk was further reduced through a factory test prior to installation. The reduction gearbox and lube-oil system were individually tested for ABS certification and were fully compliant.

A custom baseplate design was required for equipment mounting. The finished baseplate is shown in Fig. 3. Figure 4 shows the TF40 gas turbine mounted on the baseplate.

A pair of water-to-air coolers in the totally enclosed water-to-air-cooled (TEWAC) generator was used to cool the generator air. The coolers extract heat from the air at the rate of 8909 W (304,000 Btu/hr), using total water flow of 7.28 kg/s (16 lb m/s). The generator coolers are designed to operate with seawater at temperatures up to 38 °C (100 °F).

The gas turbine, gearbox, and generator were selected for their individual proven performance. Unknown and to be determined by testing were the acceptability of the generator conversion and the integrity and functionality of the control system.

Contributed by the International Gas Turbine Institute (IGTI) of ASME for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received October 1, 2003; final manuscript received March 1, 2004. IGTI Review Chair: A. J. Strazisar. Paper presented at the International Gas Turbine and Aeroengine Congress and Exhibition, Vienna, Austria, June 17, 2004, Paper No. GT2004-54006.

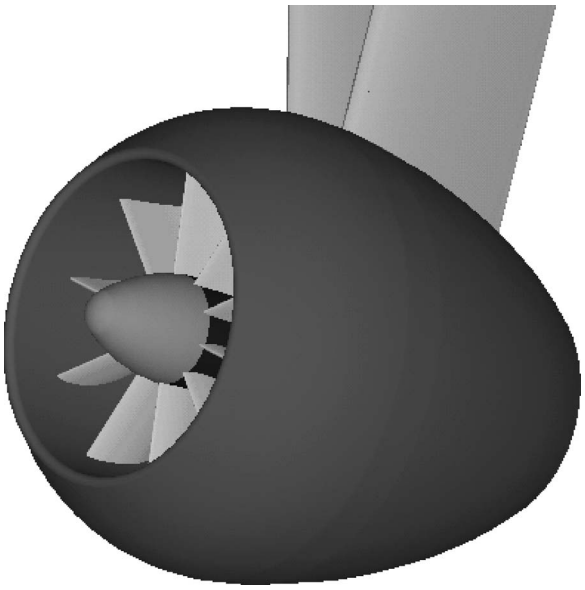


Fig. 1 Podded propulsion concept

Test Objectives

The greater rim-driven propulsor program objectives were to (i) demonstrate the seaworthiness of the rim-driven propulsor (RDP) concept, (ii) demonstrate hydrodynamic efficiency of the RDP concept, and (iii) investigate electromagnetic signature. These objectives pertain to the permanent magnet motor; however, the success of any large program depends on the success of individual parts. Since the turbine-generator set is the power source for the permanent magnet motor (PMM), ensuring its suitability was critical to the overall program.

The test was originally envisioned as a precursor to full testing of the genset and permanent magnet motor. Changes to the overall program scope resulted in the genset and motor being tested separately, but not together. The primary objective was to verify that the generator was capable of carrying full load. More specifically, the test was to verify that the generator insulation was adequate for the intended use of the genset. Since there were no load banks at the test site, the short circuit test at rated current verified that the generator was capable of full load at least with respect to the insulation system and temperatures.

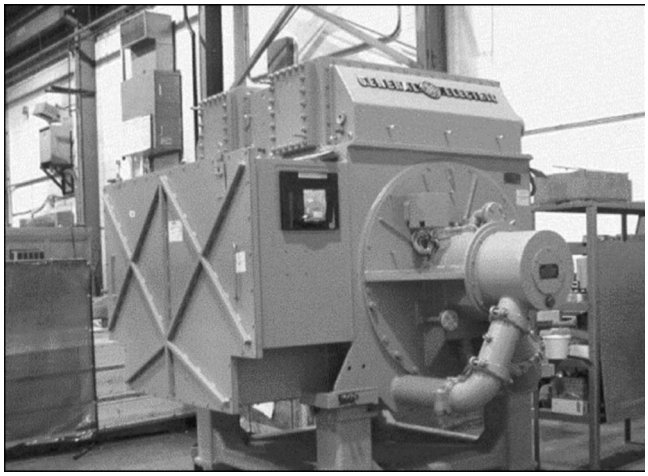


Fig. 2 Generator after conversion and refurbishment

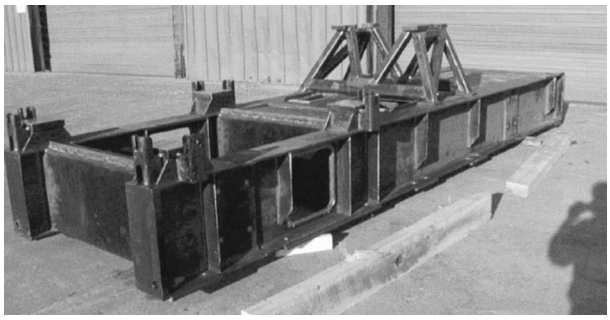


Fig. 3 Custom baseplate for equipment mounting

The short circuit test was useful in that it satisfied the test objectives. However, there were some shortcomings. The short circuit test was inadequate for evaluation of transient response, i.e., sudden load application and rejection. Transient response of the gas turbine and control system also could not be evaluated.

Other acceptance test objectives were verification of gas turbine genset safety, control, and monitoring criteria. In addition the test plan also included checks of no-load voltage saturation, open-circuit heat run, no-load voltage balance, synchronous impedance, and short-circuit at rated current.

Testing

The AHFID gas turbine-driven genset was tested in 2003 to verify the acceptability of controls, Human Machine Interface, electrical switchboard, gear, turbine, generator, and overall system operation. The test was conducted at General Tool Company in Cincinnati, OH. A simplified block diagram of the gas-turbine genset architecture is shown in Fig. 5. Figure 6 shows a (3D) solid model of the genset. The actual genset, as tested, is shown in Fig. 7.

The test sequence consisted of equipment power-up, static tests for remote operating panel (ROP) and local operating panel (LOP), safety shutdown tests, followed by functional tests from both the LOP and ROP, and finally a spin test of the system.

At equipment power-up, both local and remote operating panels were observed to operate satisfactorily. All displays were functional and reported expected values.

Static tests are functional tests of individual pieces of equipment to verify operation before attempting to run a complete functional system test. The starter wires to the gas turbine engine were removed and insulated to ensure the engine does not start during the static testing.

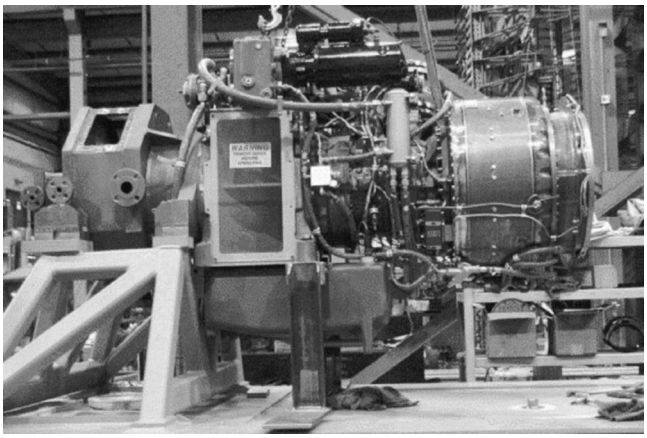


Fig. 4 TF40 gas turbine mounted on the baseplate

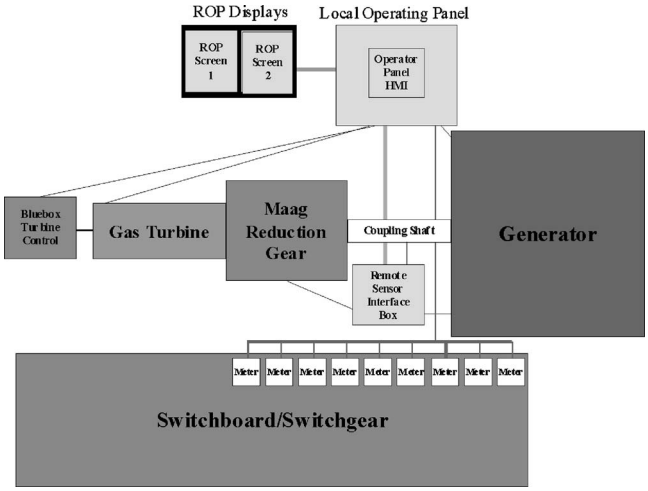


Fig. 5 Gas turbine genset architecture

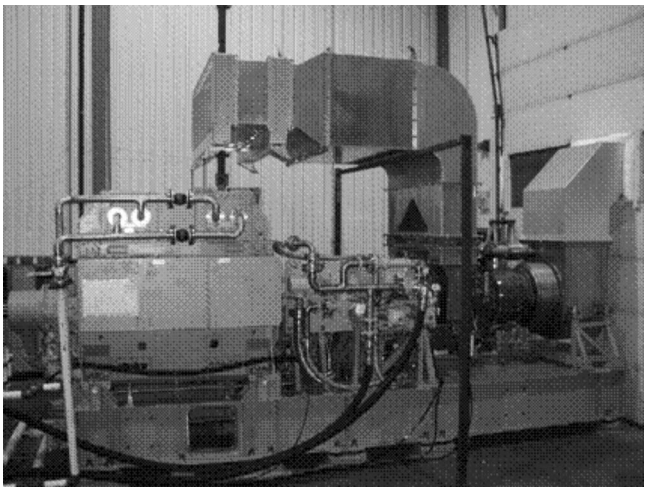


Fig. 7 Complete TF40 genset at test site

An important part of the test program was validation of the integrity of the refurbished generator. This was done in three parts: (i) a high pot test to verify suitability of the insulation, (ii) a Megger test to check for residual moisture, and (iii) a saturation test to characterize the generator. Results of the saturation test are shown in Fig. 8 and demonstrate satisfactory performance of the generator over the entire load range.

A screen traversal test—on each graphical page of the ROP—verified that each selection touch area was active and provided the functionality for which the touch area is labeled. This test was successfully completed for every graphical page on the ROP and also verified that each data point was receiving data from the programmable logic controller. Temperature, pressure, voltage, vibration, and status levels were verified.

Safety shutdown tests were used to verify that the gas turbine engine could be stopped in the event of an emergency or unexpected system operation. Emergency stop, normal stop, and system logic stops were tested by getting the gas turbine into a "ready-to-start" condition, resulting in a "ready-to-start" indication on the turbine status page on both the LOP and ROP status-page display.

The LOP has several switches and lights that monitor and control the system. The functionality of each was verified. A similar battery of tests was conducted for the ROP.

The system spin test followed. Gas turbine operation at idle and synchronous speeds was verified, as well as all alarm conditions and emergency stops. The no-load saturation test, open-circuit heat run, no-load voltage balance, short circuit saturation, and short circuit at rated current test procedures were successfully executed.

To evaluate the generator insulation, 24 dual-element resistance temperature devices (RTDs) were used. The RTDs were located at three axial locations within the generator stator and were circumferentially positioned to ensure that areas of concern were ad-

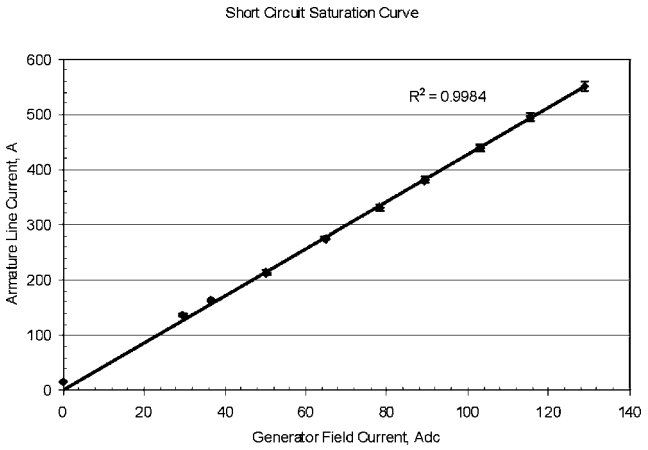


Fig. 8 Short-circuit saturation curve for the refurbished generator

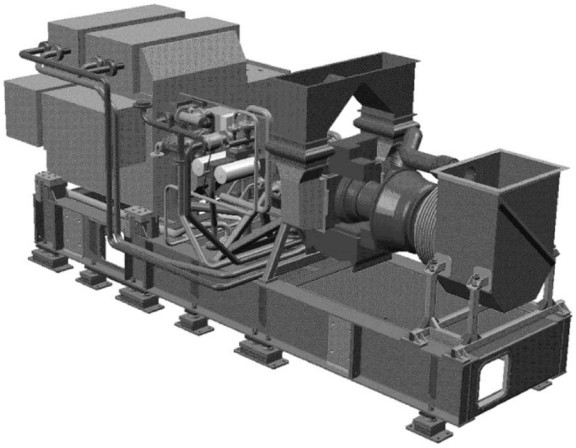


Fig. 6 Solid model depiction of TF40-based gas turbine genset

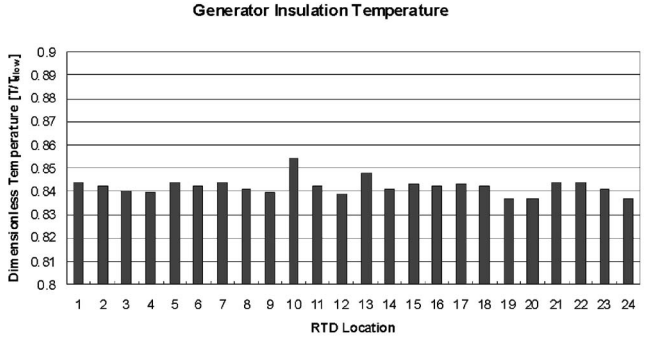


Fig. 9 Normalized steady-state full-load insulation temperatures

equately covered. Recorded steady-state full-load temperatures were compared against the maximum allowable insulation temperature. Figure 9 describes the steady-state results.

The entire system (gas turbine, reduction gearbox, generator, lube oil system, control system voltage regulator) completed a land-based qualification test in accordance with schedule and with fully satisfactory performance.

Summary

A TF40-based gas turbine genset was configured to provide electric power to a rim-driven propulsion pod. The design included a refurbished and reconfigured generator. Verification testing was conducted at a commercial land-based test site. All aspects of the test (gas turbine operation and electric power generation) were successfully completed.

Temperature measurements in the generator stator demonstrated the suitability of the refurbished generator for the application. Acceptable temperature margin was demonstrated at steady-state, full-load conditions. This proven test platform is available for future marine gas turbine and podded propulsion research and development.

Acknowledgment

The authors appreciate the encouragement and support of William Catlow.

References

[1] Van Blarcom, B., Hanhinen, J., and Mewis, F., 2002, "The Commercial Rim-Driven Permanent Magnet Motor Propulsor Pod," SNAME Annual Meeting, Boston.

[2] Waaler, C. M., Quadrini, M. A., and Peltzer, T. J., 2002, "Design and Manufacture of a 2100 Horsepower Electric Podded Propulsion System," Publication of Navatek, Ltd., Honolulu.