Transactions of the ASME®

Transactions of the ASME, Journal of Engineering for Gas Turbines and Power (ISSN 0742-4795) is published quarterly (Jan., Apr., July, Oct.,) for $130.00 per year by The American Society of Mechanical Engineers, 345 East 47th Street, New York, NY 10017. Second class postage paid at New York, NY and additional mailing offices. POSTMASTER: Send address changes to Transactions of the ASME, Journal of Engineering for Gas Turbines and Power, c/o THE AMERICAN SOCIETY OF MECHANICAL ENGINEERS, 22 Law Drive, Box 2300, Fairfield, NJ 07007-2300.
CHANGES OF ADDRESS must be received at Society headquarters seven weeks before they are to be effective. Please send old label and new address.
PRICES: To members, $40.00, annually: to nonmembers, $136.00. Add $24.00 for postage to countries outside the United States and Canada.
STATEMENT from By-Laws. The Society shall not be responsible for statements or opinions advanced in papers or . . . printed in its publications (B 7.1, para. 3).

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

Journal of Engineering for Gas Turbines and Power

Published Quarterly by The American Society of Mechanical Engineers
VOLUME 115 • NUMBER 4 • OCTOBER 1993

TECHNICAL PAPERS

ANNOUNCEMENTS

Private Vehicles for Personal Transportation

C. A. Amann
KAB Engineering,
Bloomfield Hills, MI 48304

In the United States, private personal transportation has become dominated by the automobile, a platform supported on four wheels and propelled by an internal combustion engine (ICE). Some of the reasons why this combination has emerged as the preferred choice are reviewed. Since urban air quality has become an issue, the ICE has kept pace with progressively more stringent exhaust-emissions regulations. Future emissions standards will encourage the use of alternative fuels and battery-electric propulsion. Looking far into the future, the depletion of fossil-fuel resources and/or definitive evidence that greenhouse gases are actually changing the global climate would foster a shift toward nuclear and solar energy. The automobile platform is compatible with such a shift. The ICE and the electric motor remain as potential motive sources, although they would face some difficult challenges.

Introduction

Personal transportation is one of the cornerstones of modern life. It is used to earn a living and to take a vacation, to shop for the necessities of life and to enjoy entertainment and recreation, to access doctor and barber, school and church, friends and relatives, and for myriad other purposes.

Personal transportation may be classified as either private or public. The latter category includes buses, trains, and airplanes operated on schedules not under the direct control of the traveler. Private, rather than public, transportation accounts for the majority of personal trips made in the U.S. Vehicles for private transportation are the subject of this paper. The scope is further restricted to travel over land, thus excluding privately owned boats and aircraft.

The original mode of personal transportation was human locomotion. This imposed a speed-distance envelope on travel that depended on the physical condition of the traveler and the degree of exertion acceptable to him, the nature of the terrain, and the weather.

If a significant load was carried, the speed-distance tradeoff was affected adversely. There was also an upper limit to how much an individual could carry, independent of speed and distance. Domesticating beasts of burden therefore represented a giant forward step in personal transportation.

A large variety of animal species has been enlisted to assist in personal transportation, but the one most common in the U.S. has been the horse. Some insight into the value of the horse in personal transportation is provided by Fig. 1. On this plot of average speed versus distance, dashed loci of constant trip time are superimposed. The "running" curve indicates current world-record performance for human athletes, a sort of upper limit for human locomotion. The generally negative slope of this envelope illustrates the drop in speed that accompanies fatigue. Below this running envelope is a second labeled "walking." It reflects the world-record performance of race walkers. This is still well above the typical human speed on foot. For example, the two "soldier" points mark military expectations for the trained World War II Expert Infantryman carrying a full field pack. The filled square is for the average person who walks to work, covering 0.19 mi at 1.4 mi/h (0.3 km at 2.25 km/h) [1].

Well above the running-limit envelope is one for record race horses, extrapolated to a 50-mile point for horses traveling cross-country. As true for the human runner or walker, the average speed of the horse decreases with distance traveled. Even after discounting the record speeds on horseback for more typical travel by horse, that mode of transportation offers a substantial speed advantage over typical travel on foot. At

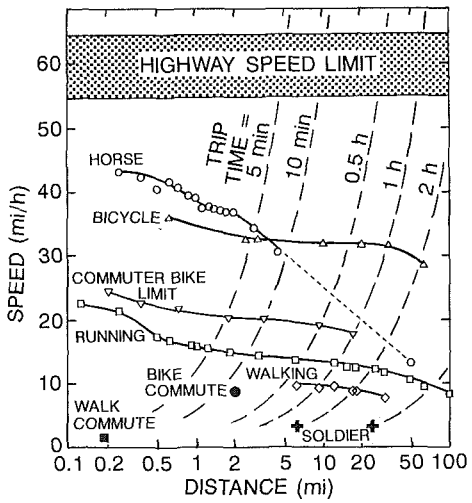


Fig. 1 Speed versus distance for various means of personal transportation

Contributed by the Internal Combustion Engine Division and presented at the 14th Annual Fall Technical Conference, Waterloo, Iowa, October 4-7, 1992. Manuscript received by the Internal Combustion Engine Division July 1, 1992. Associate Technical Editor: J. A. Caton.

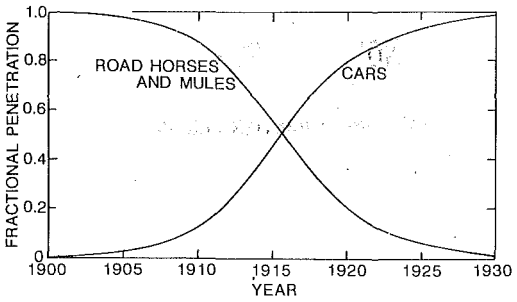


Fig. 2 Historical transition from equine mode to automobile for private transportation in the U.S.

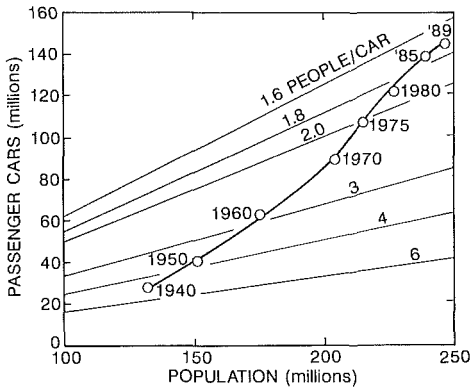


Fig. 3 Historical increase in number of cars with population in the U.S.

the same time, travel by horse spares the human the physical exertion of travel on foot.

Crossing over the horse envelope in Fig. 1 is the record-performance curve for humans racing on bicycles. It is noteworthy that beyond a distance of about 5 mi (8 km), or a trip time of about 10 min, a world-class athlete on a bicycle can outspeed a horse. It is also evident that a human can more than double his best running speed with the help of a bicycle. Again, though, the bicycle is not normally driven at record speeds.

The bike "commuter limit" curve in Fig. 1 is the calculated limit for a healthy man [2] riding a commuter bike [3] while seated upright. Even this is well above what a commuter would willingly achieve in normal travel. For example, the average person biking to work travels 2 mi (3.2 km) at a speed of 9 mi/h (14.5 km/h), as represented in Fig. 1. It is nevertheless remarkable what that 19th century invention, the bicycle, can do for human travel capability, given a proper roadbed.

The shaded band at the top of Fig. 1 marks the legal highway speed limit in the U.S. of 55 to 65 mi/h (88 to 105 km/h). The modern automobile is quite capable of sustaining such speeds over long distances, without the droop characterizing the speed–distance envelope for human and animal propulsion. Thus on a proper roadbed, the automobile can travel faster and farther than these other options while carrying cargo loads, when required, that would be intolerable to a riding horse, or to a human either walking or riding a bicycle.

In the history of this nation, the speed characteristics portrayed in Fig. 1 have contributed to a shift from reliance on horses and bicycles for personal transportation to dependence on automobiles. At the turn of the century there were many more horses than cars, and there were over 200 bicycles built in the U.S. for every new automobile. That situation has long since been reversed, however. Taking the sum of the number of road horses and mules (i.e., excluding those dedicated to farming) and the number of automobiles as a crude measure of transportation mode [4], the transition is illustrated in Fig. 2, where percentage penetration of each mode is plotted against years.

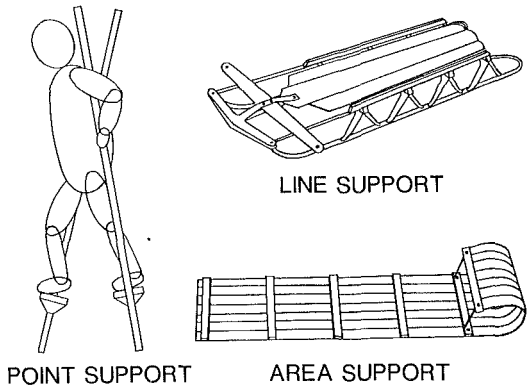


Fig. 4 Examples of point, line, and area support of a load

As the horse was retired from significant use in private transportation, the number of passenger cars grew faster than the national population, as shown in Fig. 3. The points of cars versus population are labeled by year, and lines of constant people per car are superimposed. While there were nearly 5 people per car in 1940, that ratio dropped below 2 in the late 1970s and appears to be leveling off at about 1.7. A possible explanation for such asymptotic behavior has been advanced by Lave [5]. By expanding the vehicle count to include vans and trucks used in personal transportation and contracting the population count to exclude non-drivers, it appears that the asymptote about coincides with one personal-transport vehicle per licensed driver.

Thus, private personal transportation in the U.S. has become dominated by a vehicle platform supported on four wheels and propelled by an internal combustion engine. In this paper the rationale favoring the four-wheeled platform is first examined. Then the justification for choosing an internal combustion engine is reviewed. Finally, possible future propulsion systems for the era beyond fossil fuels are discussed.

The Vehicle Platform

Important considerations in selecting the load-bearing platform of the vehicle include its support, stability, and directional control.

Platform Support. The contact footprint made on the ground may, in the limit, approach a point, a line, or an area. Examples are illustrated in Fig. 4. Point support is approached by stilts, line support by the runners on a sled, and area support by the undersurface of a toboggan.

In concept, the pneumatic tire of the automobile exemplifies point support. Neglecting the generally minor load-bearing contribution of the tire sidewalls, the footprint area of a pneumatic tire equals the product of its inflation pressure and the load it supports. Thus a 4000-lb (1816-kg) passenger car resting on tires inflated to 30 psi (207 kPa) creates on the ground a total footprint area of 0.96 sq ft (0.086 sq m). This is only about 1 percent of the plan area (length times width) of the vehicle platform.

The wheel is a unique point support because it allows forward motion of the vehicle with rolling rather than sliding friction. Thus the tractive effort required to transport a load is small compared to that with sliding-point support. Such support is exemplified by the travois of the American Indian, a load-supporting A-frame comprised of two poles, with the horse at the apex providing propulsion and the opposite ends of the poles dragging across the ground.

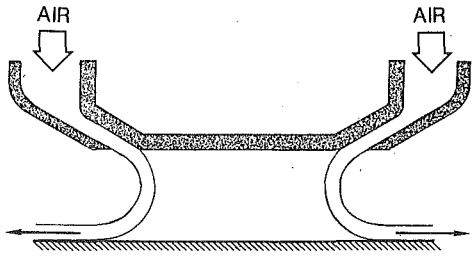


Fig. 5 Area support with an air cushion

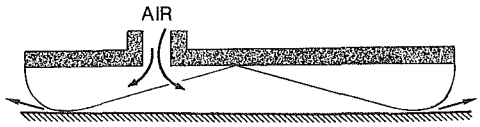


Fig. 6 Area support with an air film

The point-support example of Fig. 4, stilts, avoids the sliding of the travois but would require a complex walking linkage for mechanization. Such mechanisms have severe speed limitations compared to the rotating wheel.

Line contact of such elements as the sled runner and the ice-skate blade are special cases of rubbing friction, made practical only because the interface with the ground is lubricated by the water on which the supports glide. Such runners and blades are unsuited for use on dry land.

The tracked vehicle, which in the limit approaches line support, is adapted to overland operation because, at least in forward motion, the track avoids sliding over the ground. As the track is widened to decrease ground pressure for better operation over unprepared terrain, it moves from the line-support concept toward the area-support concept. Indeed, a snowmobile with a track width approaching that of the platform can hardly be considered an example of line support.

The area-supported toboggan, like the sled and the ice skate, owes its utility to lubrication by snow and is not suited for general-purpose use. Satisfactory area support can be provided by using air as the lubricant, however. Such a vehicle may ride on either a cushion of air having a thickness measured in inches (tenths of meters) or a film of air having a thickness measured in thousandths of an inch (hundredths of a millimeter).

The Hovercraft concept, represented by the cross section of Fig. 5, provides an air cushion. An annular jet-sheet of air around the periphery of the platform is angled inward but must then deflect outward to escape the vehicle undersurface. By applying a radial-equilibrium analysis to the curved jet-sheet [6], it can be shown that with atmospheric pressure acting on the outside of the curved sheet (bend inner streamline), the pressure underneath the platform (bend outer streamline) must be above atmospheric. This super-atmospheric pressure under the platform is what supports it above the ground.

The Hovair principle [7] for creating an air film under the platform is illustrated in cross section in Fig. 6. For a circular platform, a circular plastic sheet is secured to the platform periphery. The center of the sheet is fastened to the center of the platform. One or more holes in the sheet, made near the center of the assembly, connect two air volumes. One is the cavity between the sheet upper surface and the platform above it. The other is the space between the sheet lower surface and the ground that is bounded by a ring of minimum ground clearance. Pressurized air delivered into the upper cavity inflates the sheet to a sort of toroidal shape, then passes through the hole(s) in the sheet to the volume beneath it, from which the air escapes radially through the minimum-clearance ring to provide the film of lubricating air.

Platform Stability. The static stability of a point-supported load depends on the number of support points. The unicycle and the pogo stick, two examples of single-point support, are statically unstable in all directions. To maintain either of these devices upright at a relatively fixed position on the ground, the operator must make continuous adjustments. Obviously, a single-wheeled vehicle is a poor choice for general transportation.

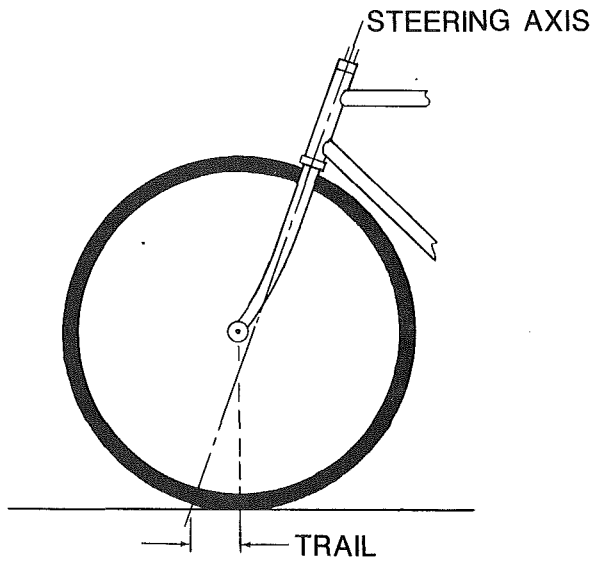


Fig. 7 Trail dimension on a bicycle

The stability of a two-wheeled vehicle depends on whether the wheels are located in tandem or in parallel planes. When the two wheels are in tandem, as exemplified by the bicycle, the vehicle is stable in pitch (longitudinal tipping). If properly designed, a bicycle ridden by an experienced operator is also stable in roll (lateral tipping) as long as the vehicle is moving forward at sufficient speed and the center of gravity is maintained in the plane of the wheels.

An important design factor in providing that stability is the trail, illustrated in Fig. 7. Trail is the horizontal dimension from the intersection of the steering axis with the ground back to the point of tangency of the wheel with the ground. In the modern bicycle, trail generally ranges from 45 to 70 mm, with the larger dimension increasing high-speed stability at the expense of responsiveness when maneuvering [8]. The feasibility of riding with hands removed from the handlebars demonstrates the extent of lateral stability.

Vehicles supported on a pair of parallel wheels include the Roman chariot, the ox cart, and the rickshaw. These vehicles are stable in roll. However, they are unstable in pitch, as evidenced by use of the propulsion medium in these examples—horse, ox, and human, respectively—to maintain balance.

Static stability in both pitch and roll is possible with three supports, whether the pair of parallel wheels is followed by a single trailing wheel or preceded by the single wheel. Before World War II most airplanes had the single wheel (or alternatively, a tailskid) trailing, but modern airplanes usually have the single wheel leading, in the manner of a child's tricycle.

One problem encountered with four-point support that is avoided with three-point support is that arising from operating on a non-planar surface. Anyone who has experienced use of a four-legged table with one short leg can appreciate this problem. In the modern road vehicle it is accommodated by the spring element in the suspension system. This allows four-point ground contact when one of the wheels is located in a modest depression in the ground plane.

Another aspect of stability is the tolerance of the platform for asymmetric loading. In the case of the vehicle with two wheels in tandem, that tolerance is poor. If the center of gravity

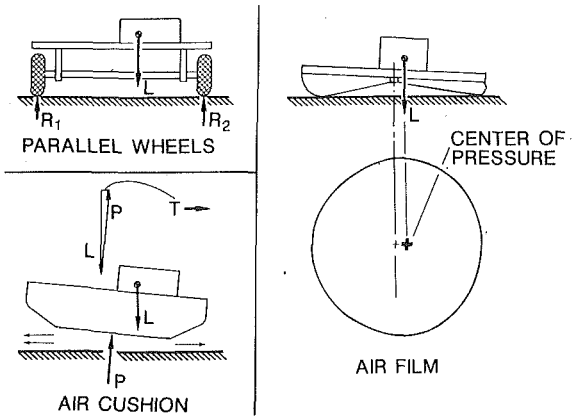


Fig. 8 Balancing an asymmetric load with parallel-wheel support, air-cushion support, and air-film support

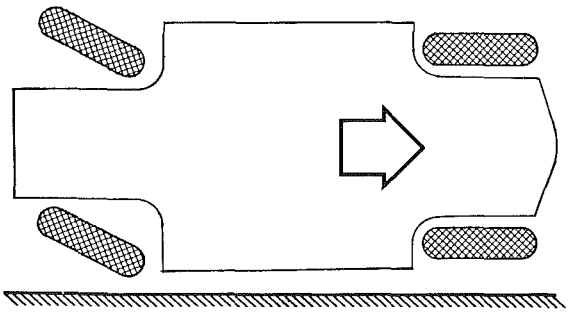


Fig. 9 Driving forward a vehicle with rear-wheel steering positioned parallel to a barrier

moves out of the plane of the wheels, the bicycle must be headed in the direction of the imbalance to avoid falling. The ability of the operator to steer a bicycle without using the handlebar merely by leaning in the direction of the turn demonstrates this coupling between imbalance and heading. That characteristic removes the tandem-wheel platform from general-purpose platform support.

With wheels in parallel planes that characteristic is avoided. As illustrated in Fig. 8, an off-center load is balanced by asymmetric loading of the wheels. A platform supported on parallel tracks enjoys the same advantage.

When the load is supported on an air cushion, as also represented in Fig. 8, a problem arises. The resultant of the pressure force on the underside of the platform (P) acts through the center of the platform. Vectorially summing this pressure force with the off-center load (L) yields a transverse force (T) that leads to tilting of the platform. That transverse force exists because the horizontal momentum of the greater airflow exiting the high side exceeds that of the lesser amount escaping the low side. In the absence of frictional resistance, as would be provided by tires on a wheeled vehicle, that transverse force causes the platform to move in the direction of the low side.

When an off-center load is applied to a platform supported by an air film under a flexible sheet, as illustrated in Fig. 8, the air-filled sheet assumes an asymmetric shape, moving the center of pressure of the footprint into alignment with the load. The degree of asymmetric loading tolerated by such a pad is severely restricted by the geometry of the flexible sheet. Consequently, the ability of three-point support to handle loading asymmetry can be utilized by supporting the platform on three pads, each one loaded centrally.

This solution is not free of problems, however. For any specified load, a given pad design functions satisfactorily only if the pressure of the air supplied to it is between two limiting values [7]. The lower bound is the flotation pressure necessary to lift the load, and the upper bound is a stability limit beyond which a vertical hopping motion is encountered. If two pads under a given platform receive significantly different loads, the air supply pressure required to float the heavier one may be above the stability limit for the lightly loaded one. For general-purpose use, therefore, a sophisticated automatic air-management system would be required.

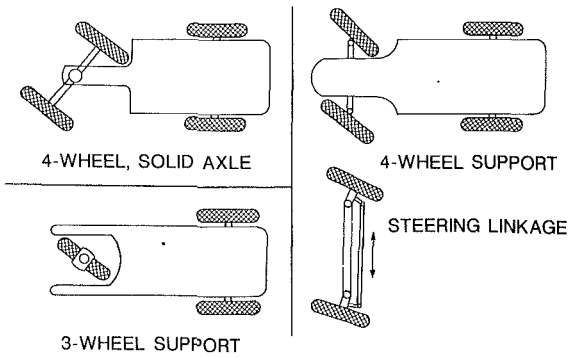


Fig. 10 Effect of steering on platform utility for (*a*) four-wheel support with a solid axle, (*b*) four-wheel support with a steering linkage, (*c*) three-wheel support

Directional Control. In a platform supported on a pair of tandem wheels, or on three wheels or four, the driver normally controls the direction of travel by steering the front wheel(s). Although steering with the rear wheels is possible, the reason behind a preference for front-wheel steering is illustrated in Fig. 9. In driving a parked vehicle away from a parallel barrier like a curb or wall, rear-wheel steering invites a collision with the barrier.

The coupling between roll stability and direction of travel in the bicycle has already been noted. For three- and four-wheeled vehicles, these functions are decoupled.

A low center of gravity is desirable in vehicles supported on more than two wheels to avoid rollover accidents. For this and other reasons, the platform of the vehicle is typically supported at a distance less than one wheel diameter above the ground.

Steering such a four-wheeled vehicle with a solid front axle, reminiscent of a horse-drawn wagon of yesteryear as illustrated in Fig. 10, seriously depreciates the utility of the platform by occupying space that would otherwise be a load-carrying area. This shortcoming has been eased by the modern steering linkage, whereby each front wheel pivots on its own steering knuckle, also shown in Fig. 10. For the three-wheeled vehicle of Fig. 10, platform-loading flexibility is again sacrificed because the front wheel encroaches on space otherwise available for cargo.

Although the air-cushion supported platform avoids the issue of interference between wheels and load-carrying space, it carries an equivalent penalty in loading utility because of its voluminous air-handling system. This handicap is more severe in small vehicles the size of an automobile than in huge vehicles such as those employing this principle that are used to cross the English Channel. The air-film supported platform escapes this problem because of its much lower airflow requirement.

The turning of a two-tracked vehicle is represented by Fig. 11. The platform can be spun around its own center by turning the two tracks in opposite directions, as indicated. However, when the linear velocity of the center of the track at radius R_1 is compared to that near the end of the track, R_2, it is evident that considerable rubbing must take place between the track and the ground. Comparing such a vehicle to a tire-supported automobile in normal driving, it is realized that substantially greater wear due to rubbing would occur in the former than in the latter.

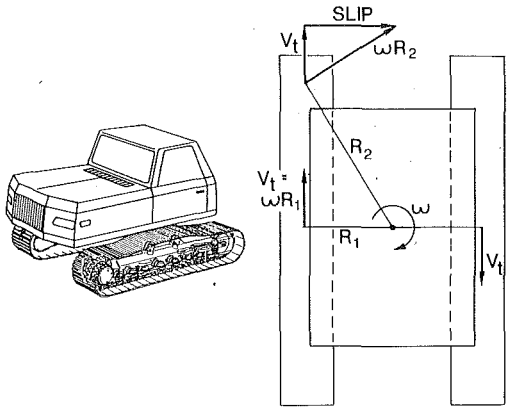


Fig. 11 Pivoting a two-track vehicle, illustrating rubbing between track and ground plane

Steering a platform supported entirely on air must be accomplished either by aerodynamic surfaces, such as with a rudder, or by effecting a directional change in the momentum of an air stream, as by directing a jet normal to the direction of travel. The turning moment produced by a conventional rudder varies with the square of vehicle velocity, making the rudder ineffective at low vehicle speeds.

The shortcoming of aerodynamic steering of these types becomes apparent when control precision is considered. Two cars can pass each other traveling in opposite directions at a legal relative speed of over 100 mi/h (161 km/h) with a separation distance of only a car width, thanks to friction between tire and road. When that friction is eliminated by employing air support, a side gust of wind can easily blow the vehicle off course, and such close-proximity operation is out of the question.

To avoid this hazard, it has often been proposed to run high-speed air-supported vehicles in a guidance trough or some variation of this scheme. That is the equivalent of operation on a limited-access track and removes the air-supported vehicle from the field of private transportation.

Overall Assessment. Given the foregoing background, it is possible to assess the leading options discussed with respect to various qualities desired in a personal transportation vehicle. The prime candidates are listed across the top of Table 1. The qualities considered, listed vertically, are (*a*) static stability in pitch and roll, (*b*) the flexibility of the support system in accepting off-center platform loading, (*c*) the extent to which a platform of specified length and width can accept cargo without interference from the support system, (*d*) the ability to provide precise directional control without adverse effects on the support system and/or ground plane, and (*e*) the quality requirements of the road bed. This fifth characteristic has not been discussed, but it is obvious that a point-support device like the wheel that is intended to roll over the ground benefits from a reasonably smooth, hard roadbed. This restriction is eased for air-cushion support, but not for air-film support.

Reviewing Table 1, the most favorable options for general use appear to be either four wheels or two tracks. The tracked vehicle operates better on unprepared surfaces than the wheel but suffers from excessive rubbing between track and roadbed during turns. Not brought out in this skeletal review is the noise that can arise with high-speed operation of a track.

The wheel is marked down for its roadbed requirement, but it has the advantage that a comprehensive network of prepared roads already exists in the U.S. One concludes from Table 1 that four-wheel support is the most sensible choice for a general-purpose personal private transportation vehicle.

Although the bicycle loses out in Table 1, its deficiency in roll stability is easily managed by an experienced operator. It remains a valid special-purpose option as long as the cargo is limited to the driver and a minimal additional load. The bicycle is not the vehicle of choice for hauling a steamer trunk, however. In addition, its comfort and safety, attributes not included in Table 1, are not competitive with those of a standard automobile.

Table 1 Comparison of platform-support options (OK = acceptable, X = deficient)

	Bicycle	Tricycle	Four-Wheeler	Two-Track	Air Cushion	Air Film
Static stability	X	OK	OK	OK	OK	OK
Loading flexibility	X	OK	OK	OK	X	X
Loading utility	X	X	OK	OK	X	OK
Directional control	OK	OK	OK	X	X	X
Roadbed requirement	X	X	X	OK	OK	X

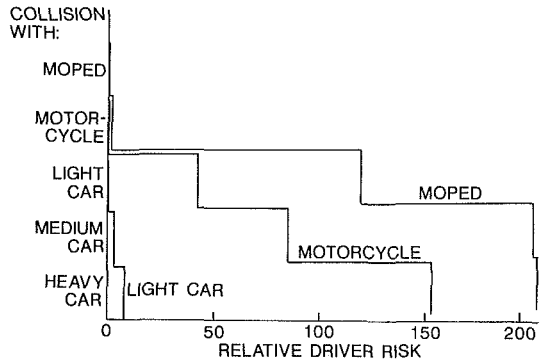


Fig. 12 Driver-fatality risk in two-vehicle collision of moped, motorcycle, and light car with heavier vehicles listed at the left

To ease the propulsion task of the human-powered bicycle, it has been motorized into the moped, motorscooter, and motorcycle. Although these engine-powered two-wheelers are safer than the standard bicycle when mixed with automobile traffic on public roads, the risk in riding on them instead of in an automobile is still considerable. For example, Evans and Frick have reviewed U.S. accident statistics from 1975–89 and expressed the relative risk of driver fatality for mopeds (including motorscooters), motorcycles and four-wheeled vehicles of various weights in collisions with heavier vehicles [9]. In these terms, a relative risk of 85 for a motorcycle colliding with a medium-weight car means that the motorcyclist is 85 times as likely to suffer fatal injury as the driver of the car. The relative risk to drivers of mopeds, motorcycles, and light cars of 1014 kg mean weight is shown in Fig. 12 for two-vehicle collisions with heavier vehicles. In this breakdown, medium and heavy cars have mean weights of 1428 and 1833 kg, respectively. The comparative safety problem for operators of two-wheeled vehicles in mixed traffic is evident.

Platform Propulsion

Once a wheeled platform has been chosen for operation on a road network, there are four ways to propel it: animal power, human power, jet momentum, and the driven wheel.

Animal Power. The historical rejection of the horse for widespread personal transportation has already been traced in Fig. 2. Given the expansion of personal transportation represented in Fig. 3, this switch in preferred transportation mode has been fortunate because it is unlikely that the environment could sustain a nation moved by horses today. If passenger cars were replaced one-for-one by horses, the resulting mass of solid waste requiring disposal would exceed a half billion

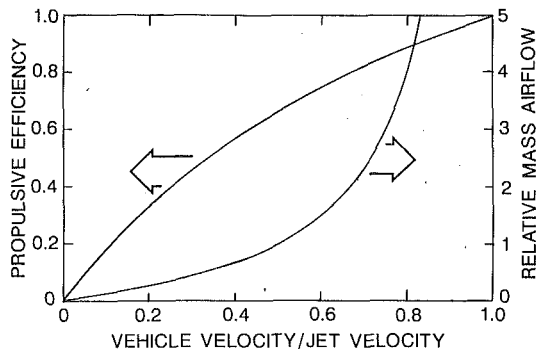


Fig. 13 Effect of vehicle/jet velocity ratio on propulsive efficiency and on mass airflow required per unit of thrust developed

tons per year, or over 350 g/mi. How fortunate that the automobile has rescued the U.S. from horse-based personal transportation.

Human Power. The bicycle is the most efficient way of applying human power to personal transportation. It is used for about half of the daily trips in Beijing, China, and in Delft, Netherlands, but only 8 percent of those in Manhattan, U.S.A. [10]. Nationwide, only about 1 percent of U.S. commuters travel to work on a bicycle or its motorized derivative, the motorcycle [1]. Bicycling remains popular as a sport in the U.S., but not for essential travel because of discomfort, the physical effort required, and the aforementioned concern about safety in mixed traffic on public roads.

However, the modern automobile has borrowed two of its important features from the bicycle technology that preceded it [11]. Dunlop, a Scottish veterinarian, popularized the pneumatic tire in Belfast in 1888. In England, Starley applied a differential to a two-person tricycle with side-by-side seating. Starley did this to even the torque distribution between the two parallel driving wheels when the greater strength of his son, who pedaled one of the wheels while he pedaled the other, made it difficult to maintain a straight course.

Jet Momentum. Propulsive thrust is produced by accelerating a mass of air opposite the direction of travel. The turbojet and the engine-driven propeller are two common examples of thrust generation.

Propulsive efficiency is a measure of how effectively the kinetic energy generated in the jet stream is transferred to the vehicle. The inefficiency appears in the form of residual kinetic energy in the jet. As shown in Fig. 13, when the vehicle is stationary, the propulsive efficiency is zero because none of the jet energy is transferred to the vehicle. As the vehicle velocity approaches the jet velocity, the propulsive efficiency approaches 100 percent, but the mass airflow required to produce the thrust approaches infinity. This is illustrated by the second curve in Fig. 13, which shows the relative airflow required per unit of thrust force at a fixed vehicle speed. For high efficiency, therefore, one desires a small difference between jet and vehicle velocities and a large jet mass. For land-based vehicles this points to a large-diameter propeller rather than a turbojet.

The practical consequences of this can be visualized by imagining a queue of passenger cars awaiting a red traffic light to turn green. When that happens, the large volume of air to be moved and the associated noise would create an intolerable situation. Trying to back a wheeled vehicle with propeller propulsion into a tight curbside parking slot is equally taxing to the imagination. Automotive propulsion by jet momentum is feasible but hardly practical.

Driven Wheel. In the history of the passenger car, three means for powering its wheels have enjoyed significant market penetration: the external-combustion engine, the electric motor, and the internal-combustion engine. Interestingly, all three saw concurrent use in automobiles at the turn of the century.

External-Combustion Engine. Since the external-combustion steam engine preceded the internal-combustion engine (ICE) in history, that it outnumbered the ICE in production cars of 1900 is no surprise. The steam engine enjoyed an insensitivity to fuel quality not shared by the ICE, but in a cold engine it took 20 to 45 minutes to raise enough steam pressure to drive.

The early steam engines had no condenser, and required frequent stops for water. Later, the condenser was added. However, the closed cycle of the external-combustion steam engine demands that all of the cycle inefficiency be rejected to the coolant. (Note that because of the external combustion employed to deliver input heat to the closed thermodynamic cycle, additional heat is rejected in the burner exhaust stream, but that loss is not inherent to the closed thermodynamic cycle.) This contrasts to the open-cycle ICE, in which much of the cycle heat rejection exits with the exhaust gas. A result of this difference is that the steam condenser must be much larger than the radiator of an ICE of equal rated power. So demanding was the heat-rejection requirement of the steam engine that its condenser was typically inadequate for the job at high power output. Under these circumstances, some of the exhaust steam leaving the expander was released to the atmosphere to avoid a buildup of back pressure. Thus a tank of makeup water had to be carried on board and occasionally refilled along with the fuel tank.

In all the years the steam engine has existed, a suitable antifreeze for the water that is compatible with the temperatures encountered in the cycle has never been found. Hence, in winter climates the steam car could not be parked outdoors unattended for long periods without draining the water.

As the ICE improved, the steam engine became bulky and heavy by comparison, and its fuel economy was inferior. It was driven from the market by the ICE in the 1920s.

The late 1960s witnessed a renaissance of interest in the steam car, driven by the belief that its continuous combustion process promised lower exhaust emissions than possible with the intermittent combustion of the ICE. Two modern steam cars built in 1969 indicated that the promised emission advantage did not come automatically [12].

Upon starting these modern steam engines from room temperature, it took 1 to 2 min to raise steam pressure, but 2.5 to 6 min before the system was sufficiently warm to realize the performance potential of the engine. The fuel consumed during starting and warmup is a serious flaw present to some degree in all external-combustion engines. In the two steam cars cited, fuel spent during starting and warmup ranged from a quarter to a half gallon. To put this in context, a contemporary car with an urban fuel economy of 20 mi/gal consumes only 0.38 gal of fuel on the 7.5-mi Federal Urban Driving Schedule, including starting and warmup.

In addition to disappointing emissions and fuel economy, these steam engines showed poor power/weight and power/volume ratios, leading to substandard performance by current standards. Clearly, the ICE has passed by the steam engine, and the market-based decision of the 1920s favoring the ICE has not been revoked.

The Stirling engine is an external-combustion engine in which the preferred working fluid is high-pressure gaseous hydrogen rather than water/steam. The thermodynamic efficiency of the ideal Stirling cycle equals that of a Carnot cycle operating between the same temperature limits, so the Stirling engine promises greater efficiency than the steam-based Rankine cycle. The 1984 experimental evaluation of a car powered by a Stirling engine then under development revealed many of the same problems identified for the modern steam car, although gen-

erally to a lesser degree of severity [13]. Power/weight and power/volume ratios were low compared to the ICE, and urban fuel economy was not competitive, thanks largely to the fuel consumed during starting and warmup. Again the closed-cycle engine failed to catch the ICE, and there is no reason to expect that it will as a primary passenger-car powerplant.

Electric Motor. Second to the steam car in 1900 production was the battery-electric. It cost more to operate and had poorer performance than its competitors. However, it started promptly, unlike the steam engine, and it did not require hand cranking, as did the ICE of that day. It was quiet, in contrast to the ICE, and emitted to smelly fumes. In that age before the existence of a developed network of paved roads, its short range between battery chargings was less of an impediment than today. Similarly, the fact that it took all night to recharge batteries was not so serious a drawback.

Since 1900 the ICE acquired an electric starter, silenced its exhaust effectively with a muffler, and improved its combustion to make exhaust gases more tolerable. The road network improved in both mileage and quality, leading to longer trips. By the 1920s, the battery-electric car joined the steam car in extinction.

Internal-Combustion Engine. Reasons that the ICE displaced competing propulsion options during the first quarter of this century are clear from the above discussion. The early spark-ignition ICE was troubled by the phenomenon of combustion knock. As knowledge of combustion grew, however, that problem was made manageable, largely through the introduction of leaded gasoline, improvements to petroleum refining, and better combustion-chamber design.

In the early 1950s a new issue arose—the contribution of the engine to air quality. The decade of the 1970s was greeted with nationwide tailpipe emission standards. Successively tighter standards have led to positive crankcase ventilation, exhaust-gas recirculation, the catalytic converter, the air pump, the evaporative canister, electronic controls, and replacement of the traditional carburetor with fuel injection. Given current tailpipe standards, it takes 25 contemporary car engines to emit as much tailpipe hydrocarbon (HC) and carbon monoxide (CO) as did a single car engine from the precontrol era. At the same time, the emission of nitrogen oxides (NO_x) has been decreased 75 percent.

New federal standards will be phased in beginning in 1994. In recognition of the lower reactivity of methane in the chemistry of photochemical smog, that hydrocarbon will no longer be counted. The new non-methane hydrocarbon (NMHC) standard is only 61 percent of the present total HC standard, and the NO_x standard will be decreased by 60 percent.

In California, which has the poorest air quality in the nation, even stricter standards are coming. That state has developed a successively cleaner series of low-emission vehicle standards that culminate in the ZEV (Zero-Emission Vehicle). ZEV is currently a synonym for an electric vehicle (EV). With no tailpipe emissions and no hydrocarbon fuel stored on board, the EV is a ZEV locally. However, given the origins of U.S. electricity, stack emissions at central power stations dictate that the EV is not a ZEV on a regional basis. California requires that 2 percent of the sales of major manufacturers be ZEVs in 1998, that fraction escalating to 10 percent in 2003.

Federal law prohibits 50 different sets of emission standards for 50 different states; each state must accept either the federal regulation or the California standards. Many states having trouble with air quality, mostly in the northeast, have either opted for or are considering adopting California standards. If those states all accept the California standards, close to half of the new-car production could be affected. As the law now stands, that would force a considerable number of EVs onto the road.

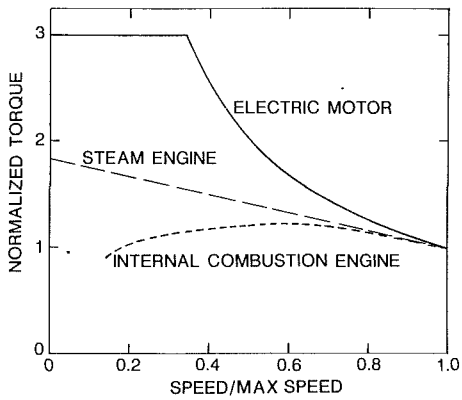


Fig. 14 Torque characteristics of a DC electric motor, an external-combustion engine, and an internal-combustion engine

In addition to lowering emission standards further and regulating gasoline composition in cities with the worst air quality, the federal Clean Air Act Amendment of 1990 calls for some use of alternative fuels by the 1996 model year. Included are methanol (usually blended with 15 percent gasoline to make M85), natural gas, reformulated gasoline, electricity, and hydrogen. With the exception of electricity, all of these are compatible with the ICE. The first four options are likely to receive the greatest attention in the ICE community.

Atmospheric ozone concentration is the normal index of urban smog. In recognition of the fact that not all HC species have the same ozone-forming potential (OFP), California has set out to assign an OFP value to each of the over 150 different species found in ICE exhaust. This will lead to a set of reactivity adjustment factors (RAF) for each alternative fuel when used in each of the California low-emission-vehicle categories. The measured tailpipe non-methane organic gases (NMOG = HC − methane + oxygenated HC species like aldehyde) will then be multiplied by the appropriate RAF to adjust the NMOG for regulatory purposes. This technique should help the alternative fuels to meet future California tailpipe standards.

The most stringent California emission standards short of the ZEV are those of the ULEV (Ultra-Low Emission Vehicle). Its requirements are 0.04/1.7/0.2/0.008 g/mi NMOG/CO/NO_x/formaldehyde. With the help of alternative fuels, including reformulated gasoline, the ULEV standards are not necessarily beyond the potential of the ICE, although they certainly involve a high degree of uncertainty. The ICE/EV hybrid would incur less risk of compliance. Given the anticipated slow penetration of the pure EV, the ICE will continue to play a strong role in personal transportation well into the 21st century, certainly as the vehicle prime mover and perhaps contributing a share of the propulsion duty through its application to hybrids.

Torque Characteristics. One important trait of the prime mover used to drive the wheels of a road vehicle is its full-load torque characteristic, for it is the difference between the torque delivered to the driving wheels and the road-load torque requirement of the vehicle that establishes its ability to climb hills, and to accelerate. In Fig. 14 typical torque curves, normalized to the torque at maximum power, are plotted against the fraction of maximum speed for a DC motor, a steam engine, and a spark-ignition engine.

In the upper part of its speed range, field current is controlled at fixed armature voltage and the electric motor is capable of producing the hyperbolic torque curve of constant power shown in Fig. 14. At low speeds its torque is normally restricted to an essentially constant value to avoid overheating the motor, with armature voltage being varied at full field strength. This is the general shape of an ideal torque curve for vehicle pro-

pulsion because if the power were held constant over the complete range of speeds, the torque would be infinite at zero speed. The tires cannot accept torques approaching infinity at low vehicle speeds without slipping.

The steam-engine torque curve is for a four-cylinder uniflow expander with a poppet intake valve in the head and exhaust ports in the cylinder walls [14]. It rises continuously toward a maximum at zero engine speed. Although this is a well-behaved torque curve for vehicle propulsion, the torque rise from rated to zero speed is sufficiently small that a transmission is required behind the engine to provide adequate performance. By stepping down the rotational speed of the driveshaft in its lowest gears, the transmission multiplies the torque delivered to the wheels at low vehicle speeds.

The ICE is seen to have the poorest torque curve. It extrapolates to zero at some low engine speed. This is because the torque developed depends on the mass of air trapped in the cylinder during each cycle, and with typical valve timing, the engine cylinder becomes an ineffective air pump at low speeds. Actually, rough running disallows operation in this regime. In contrast, the steam engine escapes this fate because it draws its working fluid from what amounts to a reservoir of high-pressure steam in the steam generator, which provides a supply not directly linked to expander speed. Given this inability to operate at engine speeds approaching zero, the ICE must be coupled to a transmission that includes some form of slipping element for accommodating the speed difference between an idling engine and a stationary driveshaft. That difference is normally handled with a clutch in manual transmissions or a torque converter in automatics.

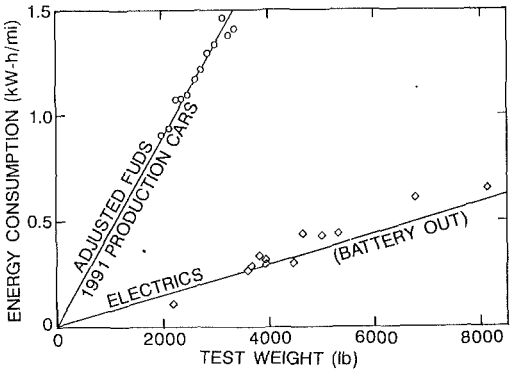


Fig. 15 Required energy in city driving for (*a*) battery-electric (energy out of battery) and (*b*) production cars with automatic transmissions (energy in gasoline)

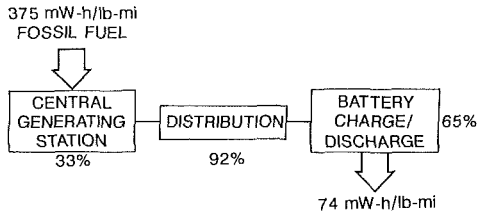


Fig. 16 Typical energy losses from fossil fuel input to battery output for battery-electric car

Efficiency Characteristics. A second characteristic of importance in the wheel-driving system is its energy efficiency. In a car it is difficult for an external-combustion engine to achieve fuel economy competitive with that of the ICE for several reasons.

First, the external-combustions engine normally operates on a closed cycle, which requires the heat generated in combustion to be transferred across the physical boundary of a heat exchanger. As previously noted, this first requires energy to be deposited in an initially cold heat exchanger before the working fluid becomes hot enough to perform significant work, and that stored heat is dissipated on shutdown.

Second, efficiency depends on both the maximum and minimum temperatures of the working fluid in the cycle. The maximum cycle temperature is restricted to a level lower than the temperature limit of the material in the heat exchanger that separates the combustion flame from the working fluid. Similarly, the minimum cycle temperature must be greater than ambient because of the temperature drop across the cycle cooler, which is the condenser in the steam engine. That elevation of minimum cycle temperature above ambient is minimized by enlarging the cooler, but there is little room for that in a passenger car.

Third, energy is discharged from the external-combustion path in the form of exhaust gas at above ambient temperature. This represents an energy loss that is not included in the normally calculated thermodynamic-cycle efficiency.

Fourth, the auxiliary load is normally greater than for an ICE engine, including, as it does, a combustion-air blower and cooling fan(s). The cooling fan(s) must be of larger capacity than for an ICE radiator because of the larger heat rejection attending a closed cycle, which lacks an exhaust heat sink.

Comparing the energy efficiencies of EVs to production cars with ICEs is always awkward because the vehicles are not equal. The production car is capable of cross-country driving; the near-term EV is not. Nevertheless, attempts at such comparisons can be instructive.

Wang and Deluchi have tabulated test results in city driving for a dozen different EVs [15]. The energy drawn from the battery per lb of vehicle test weight for each mile traveled is plotted against test weight in Fig. 15. Also shown are data points from the energy consumptions of the most efficient 1991 production cars with automatic transmissions in various EPA test-weight classes as tested on the FUDS (Federal Urban Driving Schedule). These points were calculated from the measured urban mi/gal after a 10 percent depreciation for real-world (as opposed to chassis-dynamometer) driving. The slopes of the two lines are 74 and 450 mW-h/lb-mi, respectively. But ratioing these coefficients does not imply that the EV is six times as efficient as the production car.

The EV encounters a number of energy losses between the battery terminals and the fossil-fuel energy entering the central power station, be it in the form of coal, natural gas, or oil. The energy chain is illustrated in Fig. 16. The current average efficiency for the conversion of fossil fuel energy to electricity is about 33 percent [15]. Typically, another 8 percent is lost in delivering the electricity to the wall plug. Next come the losses in the battery charger and in the battery during both charging and discharging. The average ratio of energy extracted from the battery to energy supplied from the wall plug has been measured at 0.67 for 15 test points in a car with a lead-acid battery [16] and separately reported as 0.624 for tests of five different EVs [17]. Combining these typical efficiencies from Fig. 16, about 5 kW-h must be supplied into the central station for every kW-h extracted from the battery.

Carrying the gasoline chain back to the corresponding point, i.e., crude oil entering the refinery, refining and distribution involves about a 10 percent energy loss.

Applying these factors, the energy required becomes 370 mW-h/lb-mi for the EV and 500 mW-h/lb-mi for the car with a gasoline ICE. Again, though, the ratio of these coefficients does not provide a proper comparison. In a sample of eight EVs with room-temperature batteries, the battery weight averaged nearly 30 percent of the vehicle test weight [17]. In an ICE car, the weight of the filled fuel tank is typically less than 5 percent of the vehicle test weight. Therefore, comparing the two types of cars at equal test weights is unrealistic.

Table 2 1989 emissions of a 4200-lb electric vehicle (g/mi in urban driving)

	VOC	CO	NOx	SO_2
1989 standards	0.41*	3.4	1.0	--
EV/coal	<0.01	0.06	1.84	4.81
EV/oil	0.03	0.10	1.05	2.64
EV/gas	0.01	0.10	1.21	<0.01

*0.41 g/mi total hydrocarbons

Wang and Deluchi matched eleven EVs using room-temperature batteries with their ICE-powered counterparts, finding that, primarily as a result of battery weight, the EV averaged 43 percent heavier [15]. Dividing the 500 mW-h/lb-mi for the ICE car by 1.4 to put it on the same car-size basis as the EV, the ICE coefficient becomes essentially equal to that of the EV. In other words, there is no great efficiency difference in city driving between the contemporary ICE car and the contemporary EV of comparable size. Of course, an EV weighing 40 percent more than an ICE car of equal size remains inferior in payload capacity. Moreover, the ICE car still offers better performance, faster replenishment of energy when the onboard energy supply has been expended, and an ability to make cross-country trips that is denied the EV.

Emission Characteristics. A third important trait of prime movers is their emission contribution. When the steam engine received its second wind in the 1970s, the reasoning was that its steady combustion should make emissions control easier than in the ICE. In 1975, however, the catalytic converter was added to the ICE. Because it can remove over 90 percent of the regulated emissions from the exhaust stream, that 1970s argument has lost much of its appeal.

The emissions associated with the EV depend on the source of electricity. If it is generated from water or nuclear power, then the EV is indeed clean. In the U.S., however, 54 percent of the electricity currently comes from coal combustion, 4 percent from oil, and 9 percent from gas [18]. Using central-station emission values from EPRI [18], it is possible to calculate the g/mi of VOC (Volatile Organic Compounds), CO, NO_x, and sulfur dioxide associated with a 4200-lb EV, which might be considered a replacement for a 3000-lb ICE car. In Table 2, the results are compared to the current federal standards.

It is apparent from Table 2 that the EV offers substantial benefits in VOC and CO in trade for a modest increase in NO_x. Sulfur dioxide emissions from gasoline-fueled ICEs are sufficiently low that there has been no need to regulate them, so clearly the EV would increase that contribution to acid rain substantially. Additionally, the particulate emissions associated with an EV operated on coal-based electricity would increase markedly compared to a gasoline-fueled ICE car.

EPRI projects that in post-1995 powerplants, NO_x will be reduced by a factor of 2 to 3.5 for all three fossil fuels listed in Table 2. That still leaves the EV above the Tier I standard of 0.4 g/mi that will be in effect federally for passenger cars in that time frame. In addition, a nearly fivefold reduction in sulfur dioxide is anticipated for the post-1995 coal plant [18].

Alternative ICEs. The earlier discussion of ICEs focused on the spark-ignition engine because it has so dominated automotive propulsion for the past 70 years. Two alternative ICEs that deserve mention are the compression-ignition diesel and the continuous-combustion regenerative gas turbine.

When it comes to distance covered on a gallon of fuel, the diesel engine is the uncontested winner. First, it gains about a 13 percent advantage over the gasoline ICE because of the greater energy content of a gallon of diesel fuel compared to gasoline. The diffusion combustion of its fuel spray avoids the knock problem accompanying the premixed combustion of the traditional spark-ignition ICE, so the diesel is free to use a higher compression ratio. Third, diffusion combustion allows load to be controlled by varying the overall fuel-air ratio in the cylinder, thus avoiding the part-load pumping loss of the traditional spark-ignition ICE during the intake stroke.

The diesel automobile first appeared on the market over 50 years ago. More popular in Europe than in the U.S., its market penetration in the U.S. peaked at 6 percent in 1981 and is now less than 1 percent. The comparatively low cost of automotive fuel in the U.S. and the unfavorable price differential between gasoline and diesel fuel have discouraged sale of the more expensive diesel engine in U.S. passenger cars.

Despite keen federal interest in higher automotive fuel economy, evolving emissions regulations have done little to encourage continued development of the passenger-car diesel by U.S. companies. The diesel has found a niche market in light- and medium-duty trucks and vans, however, some of which are used in personal transportation.

The diesel continues strong in the heavy-duty truck market. The most challenging emissions regulations for all classes of diesel have been those for NO_x and particulate matter. Technologies being pursued in the heavy-duty segment to satisfy more stringent standards include improved turbocharging and aftercooling, higher injection pressures, and increased flexibility in fuel injection, both its timing and its rate characteristic.

Particulate traps are being explored to satisfy the regulation of that emission, but cost and durability experience have been disappointing. Less costly routes to lower particulates include decreased oil consumption, reformulated fuel, and use of an oxidizing catalytic converter to decrease the soluble fraction of the exhaust particulate matter. Unfortunately, nearly all such measures increase the cost disadvantage of the diesel relative to the spark-ignition ICE.

A high-risk version of the conventional diesel that is being researched is the LHR (Low Heat Rejection) diesel, sometimes inappropriately called the "adiabatic" diesel [19]. In this concept, the traditional liquid-cooling system is eliminated, necessitating the use of ceramics for the walls of the cylinder and the combustion chamber to accommodate the increased temperature that results. The hotter walls lead to a reduction in volumetric efficiency, so turbocharging must be incorporated to compensate for the lost power.

Thermodynamics dictates that only a small fraction of the coolant heat rejection thus eliminated can be converted directly to crankshaft work. Most of the energy conserved by eliminating the coolant appears in the exhaust in the form of increased gas temperature. To capitalize on this energy, the LHR diesel is normally compounded, i.e., a second turbine geared to the crankshaft is added downstream of the turbocharger turbine. Such a compounding system is quite ineffective at light loads, so the LHR diesel appears poorly suited to the passenger car.

Other impediments to acceptance are compliance with stringent NO_x emission standards, identification of a lubricant that will withstand the hotter cylinder-wall temperature, and development of structural ceramics or ceramic coatings combining increased reliability with acceptable cost. The uncooled LHR diesel will not appear in personal-transportation vehicles until it has been commercialized in the heavy-duty automotive application, if ever.

Another high-risk ICE, being supported by the federal government, is the regenerative gas turbine. Both two-shaft and single-shaft versions are being explored. The single-shaft gas turbine is unsatisfactory for automotive use unless coupled with a continuously variable transmission, parallel development of which appears to be unsupported.

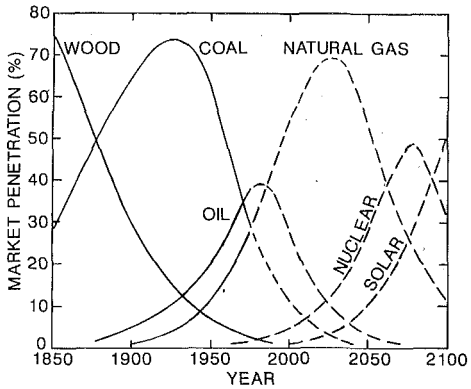


Fig. 17 Partitioning of world energy use, historical (solid curves) and projected (dashed curves)

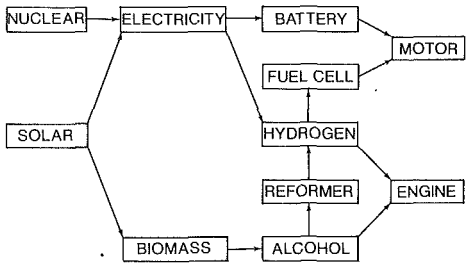


Fig. 18 Options for application of nuclear and solar energy to automobile propulsion

The two-shaft automotive gas turbine has 40 years of development history behind it. This is approximately twice the length of time it took the spark-ignition ICE to dislodge the steam car and the battery-electric from the marketplace. Still the first commercially viable automotive gas turbine has yet to appear.

The two-shaft gas turbine enjoys smooth, vibration-free operation. It has an admirable torque characteristic, although the rise in full-load torque with decreasing output speed is still insufficient to avoid the need for at least two transmission-gear ratios. Principal barriers for early passenger-car gas turbines were noncompetitive fuel economy, slow acceleration response, and excessive NO_x emissions.

The new thrust in automotive gas turbines has been toward use of structural ceramics to allow a higher turbine inlet temperature. The lower density of structural ceramics, compared to high-temperature metallic alloys, promises some help with acceleration response. Higher limiting temperature increases cycle efficiency, promotes increased NO_x emissions, and decreases engine size by boosting the power delivered per unit mass of airflow.

The increased turbine inlet temperature calls for ceramics with greater reliability and lower cost. Current and future NO_x standards for passenger cars seem to dictate abandonment of the diffusion-burning fuel spray in the gas-turbine combustor. A prevaporized premixed combustor has met a 0.4 g/mi NO_x standard in a car tested in a laboratory setting at lower turbine inlet temperatures than those sought with ceramics [20]. However, future higher burner-inlet temperatures may demand that an innovative new approach to combustion be developed.

The reduced airflow requirement accompanying a higher peak cycle temperature, other factors remaining constant, decreases the size of the compressor, the turbine, and the entire engine. Although that would help engine packaging, it is not yet clear that all other relevant factors would indeed remain constant. The decreased turbomachinery size accompanying reduced airflow threatens lower turbomachinery efficiencies. Like peak cycle temperature, turbomachinery efficiency is important to cycle efficiency and brake specific air consumption. It is conceivable that in a small engine of specified output, the fuel-economy benefit anticipated from higher peak cycle temperature could be nullified by losses in component efficiencies [21].

Although impressive progress has been made with the reliability of ceramic components, questions about fuel economy and emissions cannot be resolved satisfactorily until an automobile with a ceramic gas turbine can be subjected to transient testing. With present effort being focused on the development of suitable ceramics, that milestone is not imminent.

Beyond Fossil Fuels

The petroleum fueling almost all automobiles today is a nonrenewable resource. Twentieth-century history has been punctuated with predictions that depletion of that resource is just a few years away, but those termination dates have passed and oil continues to flow. Early prognosticators did not foresee such comparatively recent developments as the Alaskan pipeline or extraction of oil from beneath the North Sea. Nevertheless, the day must come when the recovery of remaining oil will be so costly that alternative energy sources will become economically attractive. The only question is when.

One projection of future world energy partitioning is offered in Fig. 17 [22]. The solid curves represent smoothed historical data for market penetration by wood, coal, oil, and natural gas. The dashed curves are predictions showing the eventual decay in penetration of those resources as they are replaced by nuclear and solar sources. Although the timing shown for these penetrations may again be incorrect, an eventual shift toward these emerging resources seems reasonable as fossil fuel is depleted.

Paths whereby nuclear and solar energy can power automobiles are shown in Fig. 18. Nuclear electricity can be used in a battery-electric car, or it can be used to electrolyze hydrogen from water. Hydrogen can then be used either in a fuel cell-electric car or an ICE car.

Solar energy is a poor match for the battery-electric because it is available only when sunshine exists, but EV operators will want to recharge batteries when the sun is not shining. However, solar electricity can be stored through electrolysis, in the form of hydrogen, for use in either a fuel cell or an ICE. Solar energy can also be used to grow alcohol fuel for an ICE. Alternatively, alcohol can be reformed on board a vehicle to release hydrogen for a fuel cell.

In addition to depletion of fossil-fuel resources, a second factor that could encourage movement in the directions suggested by Fig. 18 is the potential for global warming [23]. Our globe is surrounded by a tropospheric shield of greenhouse gases that reflect a fraction of the earth's infrared radiation back to keep our world at a habitable temperature. Combustion of fossil fuels adds carbon dioxide to the environment, and carbon dioxide is the primary greenhouse gas. More greenhouse gases in the shield means increased earthbound reflection of energy and a warmer world.

That the average atmospheric carbon dioxide concentration has been rising as fossil fuel has been consumed is an accepted fact. Whether that trend is responsible for a corresponding increase in world temperature is more controversial. Long before humans existed in sufficient number to influence climate, the average temperature of the earth is known to have fluctuated up and down, along with atmospheric carbon dioxide concentration. Which of these was cause and which was effect is less certain.

For the quarter century beginning in 1940, average temperature decreased. For the following quarter century, it increased. Yet average carbon dioxide concentration rose steadily

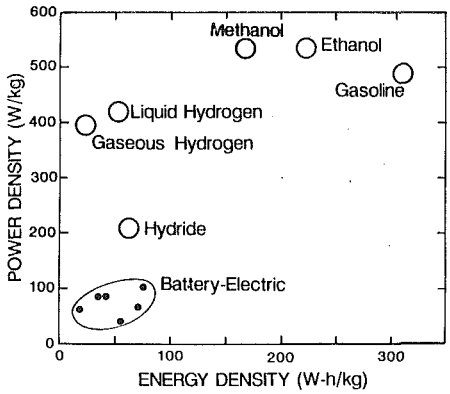


Fig. 19 Power density versus energy density for various automotive propulsion systems

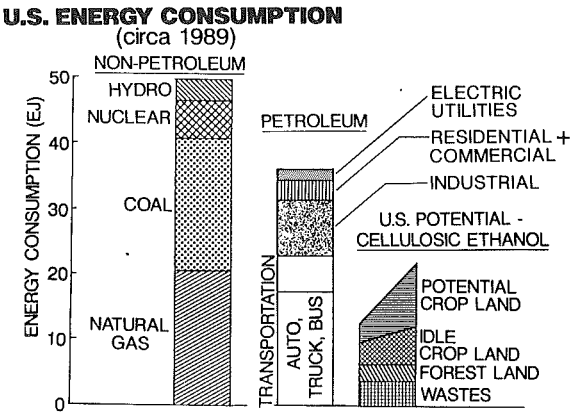


Fig. 20 U.S. energy consumption from nonpetroleum sources (left), petroleum sources (center), and estimated potential for cellulosic ethanol (right)

during both periods. Although the part that greater carbon dioxide concentration plays in the greenhouse effect is understood, the significance of such potentially ameliorating effects as cloud cover, and the ocean as a carbon dioxide sink, are less clear.

Should continued atmospheric studies confirm that fossil-fuel combustion is indeed a threat to global climate, nuclear and solar electricity and/or hydrogen avoid production of carbon dioxide. The carbon dioxide from combustion of biomass alcohol can be made part of a closed system in which replacement biomass, during its growth period, absorbs the carbon dioxide produced by combustion of biomass alcohol. Thus with no net increase in carbon dioxide, there is no greenhouse-gas contribution. Of course, burning biomass without replacing it, as occurs in the clearing of rain forests, does increase the greenhouse-gas inventory.

Biomass/ICE. Ethanol is made in the U.S. today from corn, but the process is too expensive to make ethanol a competitive primary fuel. The total energy available from U.S. corn-based ethanol is also small compared to the gasoline-energy consumption of the U.S. fleet. These two characteristics presently relegate corn-based ethanol to its role as an environmentally acceptable octane-enhancing additive in gasohol.

Methanol is the current alcohol of choice for the U.S. fleet. The present direction in the U.S. is toward M85, a blend containing 15 percent gasoline. Although methanol can be made from wood, current resources and economics point toward methanol from natural gas.

Whether the alcohol used is ethanol or methanol, shortened range is a disadvantage. In Fig. 19, power density is plotted against energy density for a variety of options. Power density is an indicator of performance, energy density of range. The gasoline bubble is a reference point for a contemporary automobile, based on the power and range *after* the gasoline energy has undergone conversion in the powertrain, and debiting the fuel with the mass of the powertrain needed to effect that conversion [23].

The ethanol and methanol bubbles are for the same-displacement engine and the same size of fuel tank. Power is up slightly, an advantage of alcohol fuels. However, range is down substantially, primarily because the energy content of a gallon of alcohol is approximately 2/3 and 1/2 that of a gallon of gasoline for ethanol and methanol, respectively. To minimize the range penalty, ethanol is obviously preferred over methanol.

The National Renewable Energy Laboratory is developing a process to convert cellulose into ethanol. If the process evolves satisfactorily, it has been estimated that enough ethanol might be produced from domestic sources at a reasonable cost to satisfy the current automotive fleet [24], as illustrated in Fig. 20, where recent national energy consumption for nonpetroleum and petroleum resources are represented by the two bars on the left. The unshaded portion of the center bar denotes transportation energy. Most of it goes to automobiles, trucks, and buses, as indicated. The balance is used in airplanes, railroads, ships, and other miscellaneous applications.

The bar at the right in Fig. 20 shows an estimate of the cellulose-based ethanol potential from (*a*) municipal, agricultural, and forest wastes, (*b*) wood from forestland, (*c*) crops grown on land now idled by federal policy, and (*d*) new cropland [24].

If the ethanol projections prove to be true, solar/biomass fuel could go a long way toward meeting the energy demand for transportation. Unfortunately, this route does not free the nation totally from current exhaust-emission concerns.

Hydrogen/ICE. If the ICE is fueled with hydrogen from electrolysis via either nuclear or solar electricity, the tailpipe emissions of HC and CO are virtually eliminated because the lubricating oil becomes the only source of carbon. NO_x emissions would require control, however. The ready ignitability of hydrogen makes the engine prone to backfiring, but ways exist to manage that phenomenon.

The biggest problem with hydrogen is that of on-board storage. The severity of this problem is indicated by the low energy density of the hydrogen bubble in Fig. 19, which represents a volume of gaseous hydrogen stored at 3000 psi (20 MPa) in the space normally occupied by a gasoline tank. Power is down relative to gasoline because of the space occupied in the intake manifold by the hydrogen gas.

Liquid hydrogen (LH_2) appears to be much better than gaseous hydrogen in Fig. 19, but it is still a poor match for gasoline. It must be kept at 20 K ($-423°F$) to avoid boiloff. The release of vapor that is bound to occur creates a safety concern in enclosed spaces like garages. Another drawback to liquid hydrogen is its unsuitability for pipeline transmission.

Hydrogen is storable in a metal hydride, where it is bound in the metal and released upon heating. Conceptually, that heat is available from the engine exhaust gas. Hydride storage involves a large weight penalty, however, which contributes to its poor showing in Fig. 19. The hydrides most effective in storing hydrogen require a temperature source hotter than engine exhaust gas under many driving conditions.

Battery/EV. In the first quarter of this century, the ICE car drove the battery-EV from the marketplace because the EV had (*a*) poor range, coupled with a long recharge time, (*b*) poor performance because of its low power/weight ratio, and (*c*) high cost associated with limited battery life. The prior discussion of contemporary EVs suggested that these problems

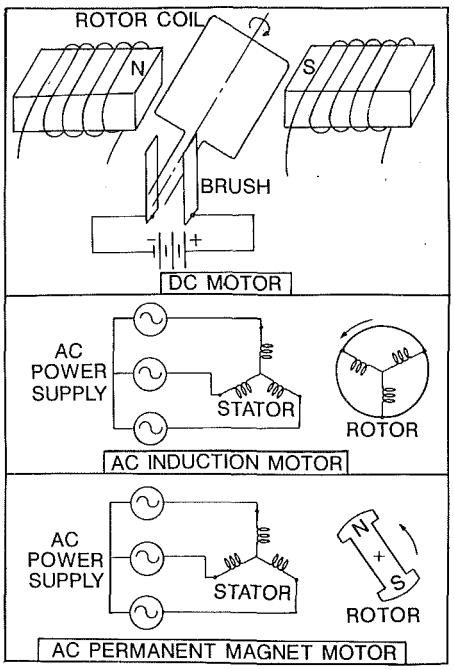


Fig. 21 Schematics of three types of electric motor suitable for tractive use

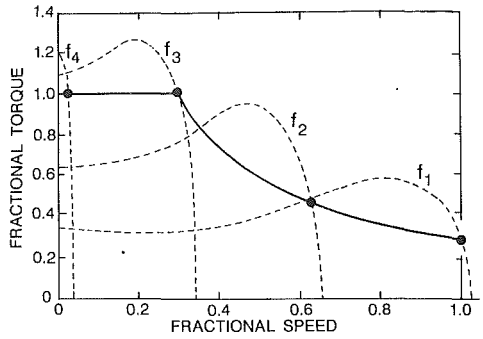


Fig. 22 Securing an ideal torque-speed characteristic from an AC induction motor via solid-state control of frequency and voltage

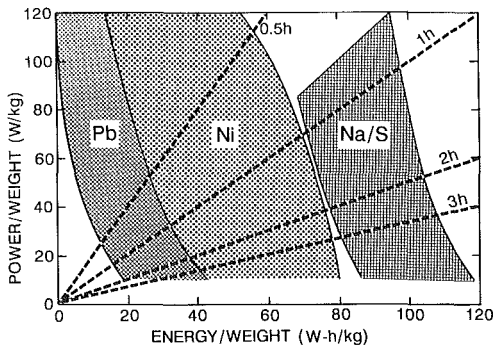


Fig. 23 General range of power density versus energy density for various battery classes

have not disappeared. Also, the EV showed no particular efficiency advantage. As for emissions, it claimed a definite advantage on a local basis, but on a regional basis it seemed merely to trade a reduction in some emissions for an increase in others.

If the electricity comes exclusively from a nuclear source, the central-station efficiency concern is diminished because no fossil fuel is consumed. Emissions, both local and regional, are eliminated in exchange for a radioactive-waste disposal problem. The other EV problems listed remain as technical challenges, primarily associated with the battery.

One of the challenges is system weight. To decrease the weight of the motor, its speed can be increased. The DC motor has the excellent torque characteristic illustrated in Fig. 14. However, as shown in the schematic of Fig. 21, it relies on brushes for commutation, and brushes wear. The speed of the DC motor is limited to keep brush maintenance at an acceptable level.

Solid-state electronics enable use of the AC induction motor, also diagramed in Fig. 21, for traction purposes. This option avoids brushes, facilitating a higher speed for reduced size and weight. As suggested in Fig. 22, when such a motor is operated at constant frequency f_1, its torque rises rapidly from zero at the synchronous (no slip) speed to a maximum with some modest degree of slip. The preferred operating point lies between the synchronous speed and the peak-torque speed. By employing a modulating inverter, voltage and frequency can be changed to provide curves f_2, f_3, and f_4. The electronic control system can manage frequency and voltage to provide the solid torque-speed curve of Fig. 22, which resembles the curve for the DC motor shown in Fig. 14.

The brushless permanent-magnet motor provides another high-speed option. Also shown schematically in Fig. 21, it, too, uses a variable-frequency inverter-type controller to match the needs of the vehicle.

By far the heaviest element in the battery-electric system is the battery pack, not the motor and controller. The gravity of the situation is illustrated by the battery-electric patch in Fig. 19, calculated for a variety of advanced batteries (at 50 percent depth of discharge, circa 1988) in combination with the corresponding motor and controller.

Great improvements have been made in the utility of the lead–acid battery that powered the EV of yesteryear. Sealed maintenance-free lead–acid batteries are now commonplace, and in some designs the liquid electrolyte has been replaced by a gel. Despite this progress, the lead–acid battery lags behind in energy density compared to some of the advanced batteries under development.

This is illustrated in Fig. 23, where approximate bands of power density versus energy density are plotted for three general classes of battery in their current state of development. Lines of constant discharge time are superimposed. Using the right boundary of the sodium-sulfur band as an example, it is seen that if energy is extracted from the battery at a rate of 37 W/kg, the battery charge will last for 3 h, supplying 109 W-h/kg. If energy is extracted at a higher rate of 97 W/kg, however, the battery is discharged in 1 h while supplying only 98 W-h/kg. If the energy were stored instead as a volume of liquid fuel, the total energy available would be independent of its rate of usage.

The Pb band in Fig. 23 is for various versions of the lead–acid battery. The Ni band includes Ni–Zn, Ni–Fe, and Ni–Cd pairs. These nickel batteries offer a better performance/range tradeoff than the lead–acid type, but are not without shortcomings. The nickel electrode is significantly more expensive than one made of lead. The Ni–Zn battery falls short of the lead–acid on cycle life, i.e., the number of discharges before requiring replacement. The Ni–Fe battery gives off hydrogen during recharging, which creates a safety concern. This battery also involves the maintenance requirement of adding water periodically. The Ni–Cd battery is frequently cited for its fast-charge capability, although the 15-min charge that has been claimed for it would require voltage and current levels far in excess of what is normally available. A 400-V supply providing a current of 160 A is typical for such fast recharging. The high toxicity of cadmium, considered much worse than lead, is also a concern, with means for recycling it not yet completely clear.

The high-temperature liquid sodium–sulfur battery has the

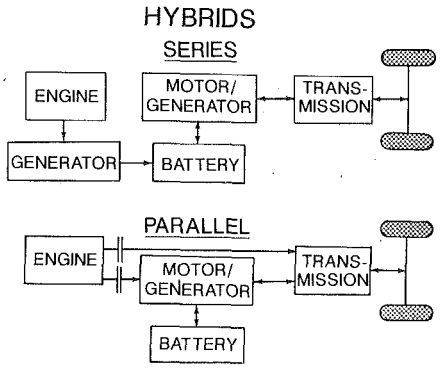


Fig. 24 Typical component arrangements for internal-combustion engine/electric hybrids

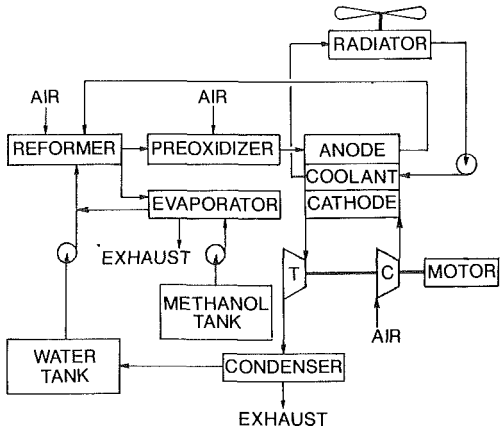


Fig. 25 Schematic of PEM (Proton Exchanger Membrane) fuel cell operated on reformed methanol

best energy density, as illustrated in Fig. 23. However, it must be kept at 350 °C (660 °F). This means that a car cannot simply be parked outdoors for a week in northern winter climates without risking a frozen battery. That may not be a problem in some fleet applications, but it imposes an impediment in general-purpose personal transportation. Concern also exists over safety in traffic accidents, should a collision ever release molten sodium and sulfur to the surroundings.

Although such advanced batteries offer significant gains in range over the lead–acid type, the technology-forcing timetable of the California ZEV will likely lead to ZEVs being dominated initially by the modern lead–acid battery. This battery has some characteristics, both good and bad, that set it apart from the gasoline ICE. Many of these traits are shared with other battery types.

In contrast to the typical contemporary ICE vehicle, the EV requires no energy consumption at idle—unless, of course, accessories are being operated. Supplying air conditioning or compartment heat from the battery severely depreciates range.

During unpowered decelerations and braking, the motor of the EV can be used as a generator to transfer vehicle kinetic energy back into the battery—like putting gasoline back into the tank. The effectiveness of this energy recovery is limited, however. There are losses in converting vehicle kinetic energy into stored electrical energy, and the energy that is accepted by a partially discharged battery may be limited by the high rate at which braking energy becomes available. Battery damage can result from too high a charging rate.

The energy available from a battery decreases as the rate at which it is withdrawn increases, as previously discussed in connection with Fig. 23. This is like having a gasoline tank that shrinks in size if the ICE is worked hard. Countering that, the energy available from the battery at a given rate of withdrawal increases if the battery is discharged intermittently, with intervening periods of rest, rather than continuously. Such intermittent operation is typical of much urban driving.

The available energy, or vehicle range, decreases at low ambient temperatures. At a given temperature, the available power—hence vehicle performance—deteriorates as the depth of discharge increases. This is like having a car in which the ability to merge into freeway traffic becomes poorer and poorer as the gasoline tank approaches empty.

If one or two cells in a series of cells comprising a battery deteriorate faster than the others, then those others must work harder to maintain a specified power-output level. This leads to a reduction in the energy capacity of the battery pack, a real-world phenomenon seldom apparent in carefully controlled laboratory tests.

The limited range and slow recharge time of the battery encourage fast recharging rates in automotive service. At least for some types of battery, this practice shortens useful battery life.

ICE/Electric Hybrid. The poor performance and range of the contemporary EV have created an interest in hybrids that combine an ICE with an electric power train within the same vehicle. The ICE could operate on any of the long-range alternative fuels. As the California regulations are originally formulated, such a hybrid does not satisfy the definition of a ZEV, although it could qualify as a ULEV. If California continues to disqualify the hybrid as a ZEV and the ICE vehicle succeeds in meeting ULEV standards, then the incentive for pursuing the hybrid would diminish.

Two of the many possible component arrangements for an ICE/electric hybrid are diagrammed in Fig. 24. In the series hybrid, all of the power delivered to the wheels passes through the electric motor, which draws its energy from the battery pack. Operating the motor as a generator during deceleration and braking allows partial recharging. The energy runs the generator to replenish battery energy and can be shut down when the state of charge is above a specified minimum. In the series hybrid, performance is limited by the electric powertrain, but range is limited by the ICE fuel-tank capacity.

In the parallel system shown in Fig. 24, both the motor and the ICE can deliver power to the driven wheels. The motor can also be operated as an engine-driven generator for battery recharging when the engine delivers more power than the vehicle requires. Performance is limited by the sum of ICE and electric-system power, and range again depends on fuel-tank capacity.

The ultracapacitor now undergoing development has interesting potential for the series hybrid [25]. One of the advantages of either type of hybrid is that the increased range facilitated by the ICE allows use of a smaller battery-energy capacity. With the performance of the series hybrid dependent on the electric powertrain, the downsized battery calls for a much greater power density to maintain performance.

Although present ultracapacitors have an energy density of less than 10 W-h/kg, a power density as high as 1000 W/kg is anticipated [25]. Ultracapacitors can also be charged and discharged in 30 s, the same order as the time required for typical acceleration in a light-duty vehicle. Thus in a series hybrid, the ultracapacitor might provide the normally encountered spurts of power, then be recharged from the battery when vehicle power demand is down.

Fuel Cell/EV. In a reversal of the electrolysis process, the fuel cell generates electricity by combining hydrogen and oxygen into water. The hydrogen might some day come from

nuclear or solar electrolysis of water. The oxygen is available from ambient air.

The efficiency of the fuel cell is not limited by the Carnot cycle in the manner of the ICE. Efficiency is about double that of the ICE, so the problem of on-board hydrogen storage is eased somewhat. HC, CO, and NO_x are not present in the exhaust stream if the required hydrogen comes from an on-board storage tank.

Of course, nuclear- and solar-based hydrogen are not available today, so in the U.S., vehicular fuel cells are being developed to use hydrogen extracted from methanol on board the vehicle. The methanol is first decomposed into CO and hydrogen over a catalyst at a temperature of at least 200°C (392°F). This endothermic reforming reaction is followed by an exothermic reaction in which the CO is reacted with water over a catalyst to produce carbon dioxide and additional hydrogen.

Several types of fuel cell exist [26], but the two being explored for vehicular use are the phosphoric acid fuel cell (PAFC) and the proton-exchange membrane (PEM) fuel cell.

The PAFC operates at a temperature of about 190°C (375°F). Starting from room temperature necessitates using a fuel burner to preheat the fuel cell and currently takes about an hour. This may be tolerable for a transit bus, but not in a general-purpose personal-transportation vehicle.

The PEM system, diagramed in Fig. 25, is less developed than the PAFC but is currently preferred because it can be started from room temperature in a matter of minutes, since its lower operating temperature of about 80°C (175°F) avoids the need for preheating. Unfortunately, its present construction is very expensive.

The fuel and oxidant streams must be humidified to avoid dehydration of the membrane above 50°C. This requires a tank of mineral-free water to be carried, raising the same concern about operation in subfreezing temperatures that plagued the steam engine. To minimize the capacity of the water tank, a condensing system is added for water recovery. A turbocompressor maintains the fuel cell at 3–5 atm to aid in the membrane humidification process. The low operating temperature of the PEM makes the catalyst sensitive to CO poisoning. This creates a concern when the hydrogen comes from processed methanol. To maintain the CO content at a couple of ppm or less, a preferential oxidizer is added after the reformer. The system complexity involved in reversing the simple electrolysis process is striking.

As long as the fuel cell derives its hydrogen from reformed methanol, the fuel-cell system has a response problem because the reformer cannot follow the load transients typical of automotive operation. Consequently, the fuel cell is run at a constant load and paired with a battery that accommodates the load peaks, then draws energy from the fuel cell for recharging when the vehicle requirement falls below the output of the fuel cell.

In its current state of development, an automotive propulsion system based on a PAFC or a PEM fuel cell would suffer from low power density and be very expensive. The vehicular fuel cell is presently quite immature and faces a long development path. Availability of a preproduction prototype may be a decade away.

Summary

Private personal transportation is embedded in the modern lifestyle. A vehicle platform supported on four wheels, at least two of which are driven by an internal-combustion engine, has emerged as the preferred conveyance. In the U.S. that vehicle, the automobile, is expected to dominate private personal transportation well into the next century. An inroad from the battery-electric vehicle can be expected, however, legislated in the interest of local air quality.

Looking farther into the next century, the supply of inexpensive fossil fuel will diminish, opening opportunities for nuclear and solar energy. Nuclear electricity could power battery-electric cars. Either nuclear or solar electricity could be used to produce hydrogen from water, with the hydrogen being consumed either in an internal-combustion engine or a fuel cell-electric system. Solar biomass is also convertible to alcohol for use in an engine. Alternatively, the alcohol can be reformed to provide hydrogen for either a fuel cell or an engine. None of these long-range options is attractive today, but all are subjects of current research.

References

1 Davis, S. C., and Hu, P. S., "Transportation Data Book: Edition 11," Oak Ridge National Laboratory ORNL-6649, Jan. 1991.

2 Drela, M., and Langford, J. S., "Human-Powered Flight," *Scientific American*, Vol. 253, No. 5, 1985, pp. 144–151.

3 Gross, A. C., Kyle, C. R., and Malewicki, D. J., "The Aerodynamics of Human-Powered Land Vehicles," *Scientific American*, Vol. 249, Nov. 6, 1983, pp. 142–145, 148–152.

4 Ausubel, J. H., "Regularities in Technological Development," *Technology and Environment*, J. H. Ausubel and H. E. Sladovich, eds., National Academy Press, Washington, DC, 1989.

5 Lave, C. A., "Future Growth of Auto Travel," *Energy and the Environment in the 21st Century*, J. W. Tester, D. O. Wood and N. A. Ferrari, eds., The MIT Press, Cambridge, MA, 1991.

6 Amann, C. A., and Scheel, J. W., "Development of the GMR Research Ground Effect Test Vehicle," Paper 370B, SAE Summer Meeting, 1961.

7 Amann, C. A., Mackie, H. A., and Scheel, J. W., "The Operating Principles of the Air Caster Ground Effect Device," *General Motors Engineering Journal*, Vol. 11, No. 1, 1964, pp. 2–7.

8 Kyle, C. R., and Olsen, J. N., "Bicycle Stability," *Bicycling*, Mar. 1990, pp. 134, 136–7.

9 Evans, L., and Frick, M. C., "Mass Ratio and Relative Driver Fatality Risk in Two-Vehicle Crashes," *Accident Analysis and Prevention*, Vol. 25, 1993, pp. 213–224.

10 Lowe, M. D., "The Bicycle: Vehicle for a Small Planet," Worldwatch Paper 90, Worldwatch Institute, Washington, DC, 1989.

11 Wilson, S. S., "Bicycle Technology," *Scientific American*, Vol. 228, Nov. 3, 1973, pp. 81–91.

12 Vickers, P. T., Mondt, J. R., Haverdink, W. H., and Wade, W. R., "General Motors' Steam Powered Passenger Cars—Emissions, Fuel Economy and Performance," *SAE Trans.*, Vol. 79, 1970, pp. 2099–2120.

13 Haverdink, W. H., Heffner, F. E., and Amann, C. A., "Assessment of an Experimental Stirling-Engine-Powered Automobile," *Proceedings of the 22nd Automotive Technology Development Contractors' Coordination Meeting*, Dept. of Energy, 1984, pp. 151–166.

14 Amann, C. A., Sheridan, D. C., Sagi, C. J., and Skellenger, G. D., "The Uniflow Steam Expander—Its Relation to Efficiency of the SE-101 Powerplant," *IECEC Proceedings*, 1972, pp. 960–970.

15 Wang, Q., and Deluchi, M. A., "Comparison of Primary Energy Consumption by Gasoline Internal Combustion Engine Vehicles and Electric Vehicles," Paper No. 910804, Transportation Research Board 70th Annual Meeting, Jan. 1991.

16 Conover, R. C., Hardy, K. S., and Sandberg, J. J., "Vehicle Testing of Near-Term Batteries," SAE Paper No. 800201, 1980.

17 Deluchi, M. A., Wang, Q., and Sperling, D., "Electric Vehicles: Performance, Life-Cycle Costs, Emissions, and Recharging Requirements," *Transportation Research A*, Vol. 23, No. 3, 1989, pp. 225–278.

18 Electric Power Research Institute, "Electric Van and Gasoline Van Emissions: A Comparison," EPRI TB.CU.177.10.89, 1989.

19 Amann, C. A., "Promises and Challenges of the Low-Heat-Rejection Diesel," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 110, July 1988, pp. 475–481.

20 Collman, J. S., Amann, C. A., Matthews, C. C., Stettler, R. J., and Verkamp, F. J., "The GT-225—An Engine for Passenger-Car Gas-Turbine Research," *SAE Trans.*, Vol. 84, 1975, pp. 690–712.

21 Murphy, T. E., "Power System Optimization for Passenger Cars," SAE Paper No. 850030, 1985.

22 Dietrich, W. R., "Criteria of Future Fuels for Stationary Gas and Diesel Engine Drives," *Engine and Environment: Which Fuel for the Future?*, AVL List GmbH, Graz, Austria, 1991, pp. 221–241.

23 Amann, C. A., "The Passenger Car and the Greenhouse Effect," *SAE Trans.*, Vol. 99, Sect. 6, 1990, pp. 1646–1665.

24 Lynd, L. R., Cushman, J. H., Nichols, R. J., and Wyman, C. E., "Fuel Ethanol From Cellulosic Biomass," *Science*, Vol. 251, No. 4999, 1991, pp. 1318–1323.

25 Burke, A. F., "Hybrid/Electric Vehicle Design Options and Evaluations," SAE Paper No. 920447, 1992.

26 Krumpelt, M., and Christianson, C. C., "An Assessment and Comparison of Fuel Cells for Transportation Applications," Argonne National Laboratory ANL-89/28, 1989.

T. Jaroszczyk

J. Wake

M. J. Connor

Nelson Industries, Inc.,
Stoughton, WI 53589

Factors Affecting the Performance of Engine Air Filters

Abrasive particles entering an engine because of inadequate air filtration can cause excessive wear, which may lead to premature engine failure. Despite the importance of filtration in engine systems, there is little understanding of the dynamics of the filtration process. Often, limited space is available for an engine air induction system. Therefore, filters are designed in smaller packages, resulting in higher aerosol velocities through the primary filter material. High aerosol velocities may cause dust re-entrainment and increase the amount of dust penetrating the filter. Our experiments with cellulose and synthetic-type filter media show examples of dust re-entrainment for fine and coarse dust. Conditions for dust particle re-entrainment are discussed.

I Introduction

A properly designed engine air cleaner reduces the dust concentration in the engine intake air to an acceptable level, has low restriction, attenuates engine noise, is not a source of flow noise, and has high operational reliability. It should be simple, small, inexpensive and easy to service. Besides performance characteristics, a filter design should comply with environmental concerns; materials used for filter construction should be recyclable. In order to minimize waste, high dust capacity filters with a life of more than 100,000–150,000 kilometers should be specified for general applications.

Better air filtration improves engine durability and performance by reducing engine wear (Khorshid and Nawwar, 1991; Fodor, 1982; Grigoryev and Ponomaryov, 1976). Engines equipped with proper filtration systems last longer and are more fuel efficient. In addition, they will consume less oil and, in turn, create less emissions.

Engine air filters operate at variable flow rates and in unpredictable environmental conditions. These factors have a significant influence on filter behavior. Unfortunately, many variables involved in the filtration process change randomly. Therefore, theoretical modeling of engine air filtration is very difficult, and research is predominantly experimental. Since experimentation is expensive, few published studies are available.

In this paper we concentrate on engine air filters operating at high flow rates in order to understand the limitations caused by dust re-entrainment. Because high-permeability filter media are typically required in modern engine air filtration, we selected multilayer nonwoven synthetic media and double-layer cellulose media for this study. Test aerosols were obtained by dispersing standard SAE fine and SAE coarse dusts and ultrafine 0–10 μm dust. We show that, despite popular belief, filter efficiency may dramatically decrease at higher dust loadings as a result of dust re-entrainment. We also show that the prediction of filter life based on test results using only one dust may be inaccurate.

II Engine Environment

Motor vehicle environments are extremely varied. Dust concentration, particle size distribution, chemical composition in a variety of geographic locations, season of the year, operating mode, location of the air intake, industrial activities, and filter efficiency influence filter performance.

The engine air intake system is the major source of contaminants entering the engine combustion chamber. Of the dust particles entering an engine, approximately 90 percent exit with the exhaust (Grigoryev and Ponomaryov, 1976). The type of road surface, roadside soil, traffic, and industrial activities all affect contaminant concentration and chemical composition.

The average dust concentration found in residential areas varies from 0.01 to 1 mg/m^3, in industrial areas from 0.1 to 10 mg/m^3, on highways 0.01 to 100 mg/m^3, agricultural areas 1–100 mg/m^3, on dirt roads 10 to 100 mg/m^3, and in dust storms from 100 to 10,000 mg/m^3 (Cockle et al., 1974; Pochtarov, 1957; Nietzold, 1979; Koffman, 1953; Dyakov, 1975; Mayev and Ponomaryov, 1971; Jaroszczyk, 1987; Caskey, 1982).

The actual dust concentration in the air entering a filter depends not only on environmental dust concentration but also on the air inlet location and its design. Location of the air inlet close to the engine exhaust system, facing an open space toward the road surface, close to the road surface, or in a space with high negative pressure can increase dust concentration and size of particles ingested.

The density of dust particles depends on their chemical composition. Densities vary from 1.8 g/cm^3 for carbon black to approximately 3 g/cm^3 for dusts containing significant amounts of iron oxide and alumina. Chemically, dust is a mixture of minerals, organics, and water. Minerals containing silica, alumina, and ferric oxide have a hardness greater than that of steel and are a major constituent of dust. Other chemicals include CaO, MgO, NaO, FeO, K_2O, MnO, etc. (Jaroszczyk,

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1992. Paper No. 93-ICE-1. Associate Technical Editor: J. A. Caton.

1987). In some industrial areas, or in areas with heavy diesel motor vehicle traffic, air contaminants may contain 10–85 percent soot (Baczewski and Jaroszczyk, 1981). Soot tends to form aggregates on filter media surfaces. If soot concentration is high, paper-type filters generally clog rapidly (Morton, 1970). Abraded particles from tires and vehicle brake systems may add substantially to the contaminants in areas with heavy vehicle traffic.

III Engine Wear Caused by Dust Particles

Engine wear is a continuous process leading to changes in the mass, chemical composition, and crystalline structure, as well as stress in the outer layers of rubbing surfaces. Changes in these parameters may be caused by the relative motion of the engine parts, the severity of the environment, and the efficiency of the filters. Inadequate air filtration is the major cause of engine failure due to increased wear by abrasive dust particles (Grigoryev and Ponomaryov, 1976).

When abrasive particles pass between two surfaces, they can break off metal chips or remain embedded in the outer layer of engine parts. Particles larger than the oil film thickness will impair the rubbing surfaces; smaller particles will be washed down by the oil, but may still weaken the oil film (Koffman, 1953). The thickness of the oil film varies from 1 to 50 μm under normal conditions (Needelman and Madhavan, 1988), but may approach 0 μm under dynamic load and high temperature. In this case, even a 1 μm particle can scratch the oil boundary layer formed on engine part surfaces.

Hard mineral particles such as silica, alumina, iron oxide, metal wear debris, and abrasive soot wear engine parts through microcutting, plowing, scratching, and erosion (Needelman and Madhavan, 1988). Organic oxidation products and other acidic substances accelerate corrosion. Oil droplets, mineral, and carbon black particles deposited on a mass flow sensor may affect its calibration. Large abrasive soot particles may break down the protective film formed on engine parts by oil additives. Soot may also absorb oil additives which can result in engine component wear (Needelman and Madhavan, 1988). Air contaminants and oxidized oil may form aggregates that can destroy the oil film.

The degree of engine wear caused by dust particles depends on their hardness, shape, roughness, size, and concentration. Because these particles are usually polyhedral or semi-spherical with sharp edges, they are ideal abrasives. Koffman (1953) described engine wear as $W = c_p^{0.675}$ for concentrations ranging from 30 to 1000 mg/m^3. Other results show that engine wear is directly related to air filter efficiencies for efficiencies over 99 percent (Grigoryev et al., 1975).

Wear increases exponentially with less efficient filters. The wear rate is approximately eight times greater for a 98 percent efficient filter and 18 times greater for a filter with an efficiency of 97.5 percent compared to a 99 percent efficient filter (Grigoryev et al., 1975). Every gram of mineral dust having a majority of particles below 30 μm removes 2.5–5 μm of an engine cylinder (Grigoryev and Ponomaryov, 1976).

Particles varying from 1–5 μm to 15–20 μm are observed to produce the greatest engine wear (Mayev and Ponomaryov, 1971; Pochtarov, 1957; Watson et al., 1955; Coates and Tull, 1963). However, components with larger clearances (e.g., bearings) were more affected by 20–40 μm particles, while wrist bearings were most sensitive to 20–80 μm particles (Midler et al., 1981). Koffman (1953) found that even 0–2.5 μm dust fed at rate of 7.5 mg/m^3 for 32 hours caused substantial wear of the engine cylinder (76 μm at the top of piston ring travel). The rate of piston ring wear for 5–10 μm dust particles was observed to be 4.5 times greater than for 0–5 μm particles, over 2 times greater than for 10–20 μm dust, and 2.3 times greater than for 0–80 μm SAE fine standard test dust (Thomas and Culbert, 1968).

IV Factors Affecting Air Filter Performance

Air filtration theory assumes very low dust concentration, small dust particles, and re-entrainment-free filtration at a constant flow rate. In contrast, air filters in engine air intake systems are subject to variable flow rates, high aerosol velocities, flow pulsations, mechanical vibrations, and various types of contaminant with a broad range of sizes and concentrations. Therefore, general filtration models describing real filter performance do not exist. However, there are numerous simplified filtration theories and experimental models, which provide a basis for dealing with some problems in filter design and development. Some of them will be discussed later.

Contaminant removal efficiency, pressure drop or flow restriction, and filter life or dust-holding capacity are the primary performance characteristics of an operational filter. They are dependent upon the properties of the filter media, air contaminants, and aerosol flow. A complete model of the filtration process must consider the aerosol flow field in the filter housing, the precleaner, and the filter element, the physical structure of the filter media, and dynamic changes in the filter during contaminant accumulation.

Pressure Drop. Pressure drop is usually the most important factor in the set of filter design assumptions. A low pressure drop is essential for optimal engine performance. Usually, the maximum acceptable engine power loss due to restriction of the engine air induction system is 2–3 percent. The maximum air intake system restriction is generally 3.8–5 kPa for gasoline engines, 5–7.6 kPa for turbocharged diesel engines, and 6.3–7.6 kPa for naturally aspirated diesel engines (Morton, 1970). The total pressure drop in the engine air induction system is the sum of pressure losses in its components.

For a new air cleaner, the filter housing, including the air inlet and the flow sensor, is usually the main source of pressure loss. Our experiments have shown that at a nominal engine flow rate, the pressure drop of a filter housing is 1.75 kPa, and the pressure drop of the mass flow sensor 1.50 kPa, while the pressure drop of a clean panel-type filter element was only

Nomenclature

P = filter penetration = $1-E$ (E = filter efficiency)
P_o = initial filter penetration
V = air velocity
W = engine wear
c_p = dust concentration
d_f = mean fiber diameter, measured microscopically
d_{fe} = hydraulic effective fiber diameter
h = media thickness
k = clogging coefficient, which depends on media and dust properties
k_p = permeability coefficient
m = media basis weight
m_D = mass of dust collected
ΔP_{dl} = pressure drop of dust-loaded filter media
ΔP_m = media pressure drop
ΔP_{PL} = pressure drop of a panel element
α = pleat angle
β = filter solidity (or packing density) = volume of fibers/volume of filter
μ = air dynamic viscosity
ρ = air density
ρ_f = fiber density

0.47 kPa. The pressure drop of a conical-type filter element (U.S. Patent 5,106,397) with the same filter media surface area as the panel element was even lower at 0.22 kPa.

During engine operation, as a result of contaminant loading, the filter element becomes the major source of pressure loss in the air induction system. The terminal pressure drop of dust-loaded elements is generally specified in the range of 2.5–5.0 kPa.

Flow restriction of the filter inlet, housing and ducts depends on air velocity, type of design, shape, surface roughness, and air density.

Losses increase significantly due to the turbulent motion of the air at higher velocities, since inertial force is proportional to the velocity squared. This explains why the pressure drop of a filter housing without a filter element may be higher than with the filter element, which works like a flow straightener in the assembly.

Data necessary for calculating the pressure losses in engine air induction systems are provided by Fried and Idelchik (1989) in their recent publication and more fully by Idelchik (1986).

During field operations, the pressure drop of an engine air induction system is affected by contaminants collected by the filter. The total pressure drop of dust-loaded filter media is the sum of the media pressure drop and the dust layer pressure drop.

An ideal dust cake forms only on the surface of special nonwoven media used in self-cleaning type filters. In other engine filtration, the dust is collected partly inside the filter media but with the majority on the surface (Stinson et al., 1989; Davies, 1973; Wake and Jaroszczyk, 1991; Jaroszczyk and Wake, 1991).

Two approaches have been used to describe the pressure drop across fibrous filter media: the capillary model employing Darcy's equation, and the fiber drag theory, which considers the aerodynamic drag on individual fibers in a filter.

The modified Darcy equation includes a linear component to describe pressure drop for Reynolds numbers up to approximately 0.5 (air velocity less than 30–50 cm/s, depending on the *fiber diameter*) and the inertia component. The second part becomes greater than the first viscous part at Re > 2–5, or air velocities over 2–3 m/s. The equation is expressed as follows (Claes and de Bruyne, 1975):

$$\Delta P_m = \frac{1}{k_p} \cdot \mu \cdot V \cdot h + \frac{\rho}{k_p} \cdot V^2 \tag{1}$$

The permeability coefficient depends on media properties. The drag theory uses the drag force acting on a fiber. Because filtration theory deals with low-velocity filtration, theoretical models do not usually include the inertial part of the pressure drop equation. A more complete discussion of pressure drop models is given by Davies (1973), Pich (1987), and Flagan and Seinfeld (1988).

A model based on the drag theory was developed for a cellulose paper-type filter with filter media solidity in the range of 0.11–0.33 and average pore diameter of 12–84 μm (Myedvyedyev et al., 1984). The pressure drop is described as:

$$\Delta P_m = \frac{\mu \cdot V \cdot m \cdot h}{d_{fe}^2 \cdot \rho_f (-0.984 \ln \beta - 0.47)} \tag{2}$$

Davies (1952) developed his theoretical–experimental model for fibrous media with porosities of 70 and 98 percent:

$$\Delta P_m = \frac{\mu \cdot V \cdot h}{d_f^2/4} [16 \cdot \beta^{1.5}(1 + 56\beta^3)] \tag{3}$$

The most important factors affecting pressure drop across filter media are fiber diameter, media porosity, and velocity, in the case of high-velocity filtration in automotive air filter applications.

Current filter media for engine air filters are made of fibers with varying diameters. Their size distribution is described by multimodal distribution. There is no literature available on pressure drop for this type of media.

Two general models follow for pressure drop increase in dust-loaded filter media (Midler et al., 1981; Pich, 1987; Nowicki, 1971; Davies, 1973). For fine dusts containing particles less than 6 μm:

$$\Delta P_{dl} = \Delta P_m \cdot \exp(k_1 \cdot m_D) \tag{4}$$

For coarser dusts:

$$\Delta P_{dl} = \Delta P_m \cdot k_2 \cdot m_D \tag{5}$$

Constants k_1 and k_2 depend on fiber diameter, filter media porosity, dust particle size, and filter thickness.

The pressure drop of a pleated filter element depends on filter element shape, pleat configuration, corrugation, embossment, and type of flow, which is associated with the Reynolds number. Uneven flow distribution affects filter performance, including pressure drop. In the case of automotive panel filters, pleat spacing is equal to approximately 3–4 mm for cellulose-type media and 4–8 mm for synthetic media. Hartig (1973) (in Nietzold, 1979) showed that the pressure drop of a panel element, ΔP_{PL}, relates to pleat angle, α, as

$$\Delta P_{PL} = \frac{\Delta P_o}{\cos \alpha} \tag{6}$$

where ΔP_o is the pressure drop for $\alpha = 0$.

Filter Efficiency. Cellulose-type filter media, followed by dual-layer cellulose-synthetic and multilayer synthetic nonwovens predominate in air filtration for engines. All of these media are fibrous. Other media, such as high-porosity depth-type nonwoven media, are used only in prefilter applications. Their role in engine air filtration is now marginal. When higher dust capacity becomes more important, these types of media may find new applications in air filtration for engines, especially the multilayer and multicomponent synthetics with surface treatment.

Filter efficiency in fibrous filter media is thoroughly described by filtration theory (Chen, 1955; Pich, 1987; Flagan and Seinfeld, 1988). However, classical filtration theory discusses collision rather than retention efficiency. Therefore, reentrainment of collected dust particles, particle bouncing and seepage, which are characteristic of the dynamic filtration process in the field, are not described therein. Filter efficiency is a function of aerosol velocity, filter media physical properties, surface properties, physical and chemical properties of the air contaminants, flow pulsation, mechanical vibration, humidity, temperature, and filter system design.

Four principal filtration mechanisms are considered: diffusion, impaction, interception, and electrostatic mechanisms. The overall filter efficiency is dominated by these mechanisms, plus adhesion of dust to the fibers, and drag forces. Mechanical vibration, flow pulsation, and pressure forces also play important roles in overall efficiency at high dust loadings. The mathematical models reviewed are those of Pich (1987) and Löffler (1971).

Since the overall efficiency is a combination of the efficiencies of all filtration mechanisms, a minimum occurs at a particle size of 0.1 to 0.5 μm, depending on aerosol velocity (Lee and Liu, 1982). For polydisperse silica dust and high-porosity synthetic nonwovens, the minimum efficiency is at a velocity of 10–20 cm/s (Wirski, 1974).

The collision efficiency given by models describing filtration mechanisms is equal to the collision efficiency for some liquid particles (Ellenbecker et al., 1980). For solid particles, the efficiency deviates from the theoretical curve at Stokes numbers over approximately 1. In the inertial region of air filtration, where particles have significant momentum, the adhesion probability factor is less than 1. Particles can bounce off fibers,

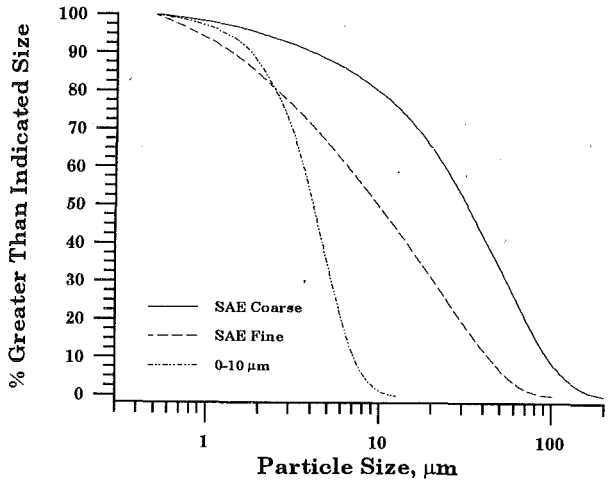


Fig. 1 Dust particle size distribution by mass

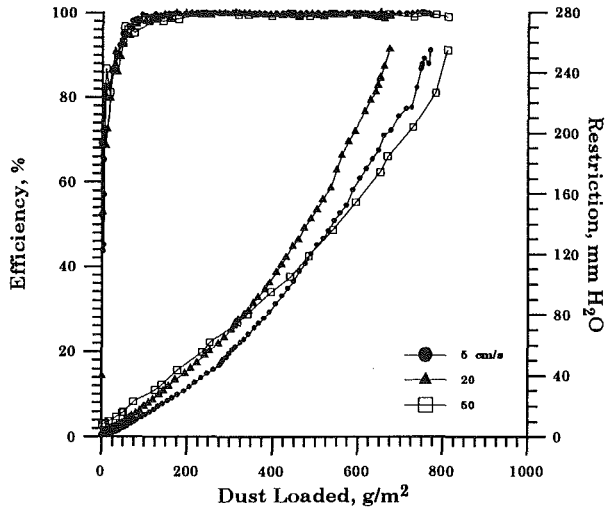


Fig. 2 Performance of filter media A for SAE fine dust

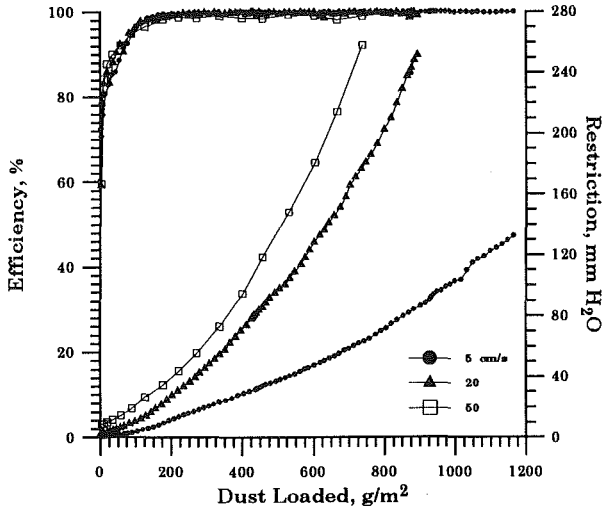


Fig. 3 Performance of filter media A for SAE coarse dust

decreasing filter efficiency (Sayfert, 1988). The rebounded particles have less energy than before collision; therefore, they are usually captured by downstream layers of fibers if the filter medium has sufficient thickness.

Filter efficiency changes with dust accumulation, which alters the filter structure. This typically results in increased filter efficiency at low aerosol velocities. A filter cake usually covers the entire media surface at low velocities and forms additional granular-type filter layers.

Table 1 Filter media properties

Property	Media A (Synthetic)	Media B (Cellulose-synthetic)
Basis Weight g/m^2	195.3	191.8
Fazier Permeability cfm/ft^2	145	124
Thickness . mm	1.78	2.25

The general equation describing filter efficiency for dust-loaded media is given as (Nowicki, 1971):

$$P = P_o \cdot \exp(-km_D) \tag{7}$$

At high velocities, the filtration process becomes more complicated due to excessive pressure drop across the filter, drag, and particle inertial forces, which may cause particle re-entrainment. There are no theoretical models describing dynamic filter performance at high aerosol velocities.

Re-entrainment. During use, a filter's packing density or solidity increases with time as a result of contaminant deposition. This typically increases filter efficiency. However, dust accumulation can result in decreased filter efficiency caused by dust particle re-entrainment.

Re-entrainment is an effect of limited adhesive forces, which are difficult to predict. Re-entrainment takes place when external forces overcome the attractive forces between the captured particle and fibers.

Since particle aggregates have a considerably larger surface area than single particles, they may be blown off as a result of increased drag forces, while single, fine particles remain attached to the fiber. Direct calculation or measurement of the forces responsible for particle detachment is nearly impossible for real filters. Media heterogeneity, flow pulsations, and mechanical vibrations play a significant role in this process.

Through the repetitive process of re-entrainment and re-deposition, particles can migrate through a filter. This migration is called seepage (Leith and Ellenbecker, 1980). Seepage is a principal mechanism of dust penetration in many applications of fibrous cellulose-type filter media. The theoretical background describing dust particle re-entrainment is complex, resulting in an inaccurate prediction of filter performance in the field. As a result, efforts to understand better the changes in dust removal as a function of time have been limited primarily to experimental work.

V Experimental

Standard gravimetric test methods provide information on the final and initial filter efficiencies. The final or overall efficiency is measured at the terminal pressure drop while the initial efficiency is measured at 20 g or more of dust loading, depending on the nominal filter flow rate. Gravimetric testing does not provide information concerning dynamic changes in filter performance characteristics caused by deposited dust. Since new engines have become more sophisticated, data on dust particles penetrating a filter are needed to understand better how the filter protects the engine.

Three types of test dusts were selected for testing: standard SAE coarse and SAE fine dusts specified by SAE and engine and vehicle manufacturers, and 0–10 μm dust. 0–10 μm dust is representative of conditions for a primary or secondary filter for a multistage air cleaner equipped with a cyclonic precleaner. All dusts have similar chemical compositions. The test dust particle size distributions are shown in Fig. 1.

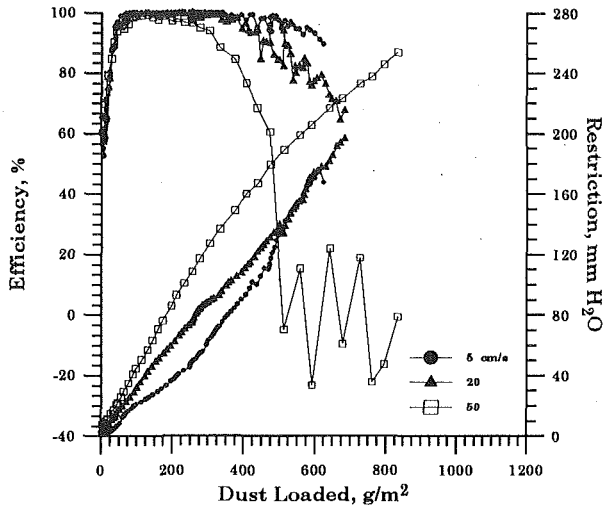


Fig. 4 Performance of filter media A for 0–10 μm test dust

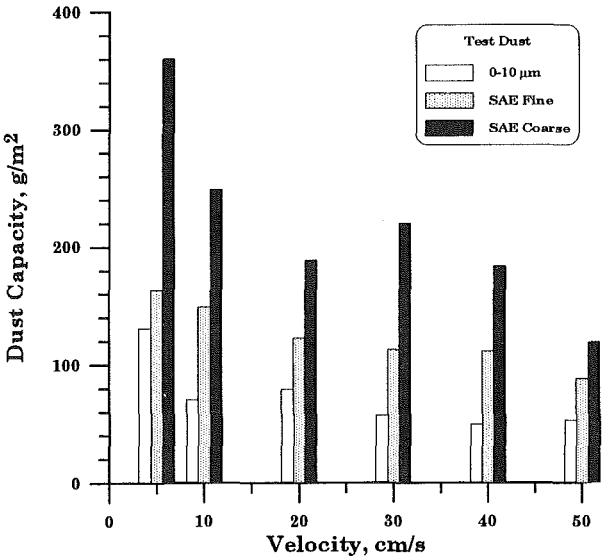


Fig. 5 Dust-holding capacity of media A at a restriction of 25 mm H_2O

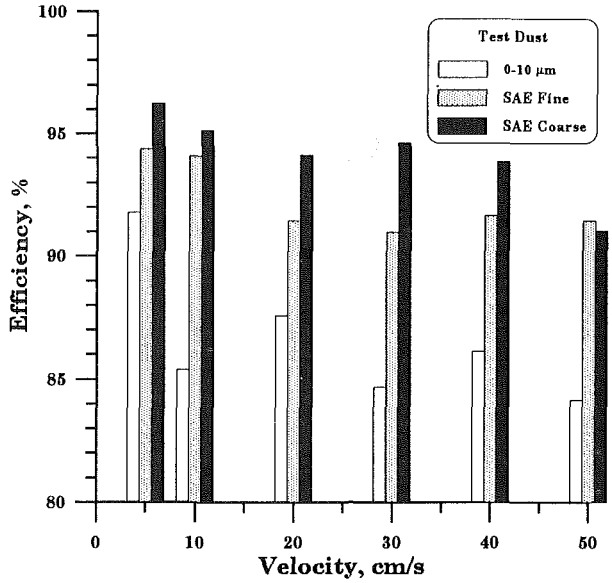


Fig. 6 Efficiency of media A at a restriction of 25 mm H_2O

Since new, synthetic or synthetic-cellulose type media are becoming more popular in modern engine air filter designs, we used samples of these media in our experimental work. Their properties are shown in Table 1.

The dust cake formed by SAE fine (Fig. 2) is relatively dense and mechanically strong. This dust contains a large number of fine particles, which form a dense cake. There are also larger particles up to approximately 80 μm, which are mainly responsible for the mechanical strength of the cake. The fine features of the cake results in good efficiency while the mechanical strength ensures that the cake will not collapse easily.

In the case of SAE coarse (Fig. 3), the dust cake is more permeable than for SAE fine; therefore, a substantial number of particles can penetrate this cake at high aerosol velocities, resulting in lower filter efficiency at 50 cm/s. The initial efficiency is slightly higher than for SAE fine because this dust contains a greater number of large particles, which are collected more easily by this type of medium.

Real filtration problems occur with 0–10 μm dust, especially at high aerosol velocities (Fig. 4). The dust cake collapses due to increased static pressure differences across the dust-loaded media, inertial force as a result of particle kinetic energy, and viscous drag acting on the dust cake. Because the external forces are greater at higher velocities, particle reentrainment due to cake collapse is significant at high velocities. Even negative efficiencies are sometimes measured at higher dust loadings. Dust reentrainment is less evident for the fine and coarse dusts because the cakes formed by large particles are mechanically stronger than in the case of 0–10 μm dust.

Figures 5 and 6 show how difficult it is to predict filter performance for different dusts. For a given filter, the amount of dust collected, aerosol velocity, and dust particle size distribution are major factors affecting filter performance. Depending on dust size distribution, different dust loadings are required in order to gain a filter restriction of 25 mm H_2O (250 Pa). Therefore, it is unrealistic to describe dust loading in the field with fine dust particles using SAE coarse dust, or even SAE fine dust in some cases.

Filter efficiency (Fig. 6) also depends on the type of dust and aerosol velocity. Due to increased re-entrainment at higher velocities, efficiency is usually lower. However, in the case of 0–10 μm dust, filter efficiency changes in a random pattern as a function of aerosol velocity. This can be explained by the weak, unstable dust cake and its different geometric structure at different velocities. Even relatively minor mechanical motion is sufficient to detach captured particles (Grant et al., 1989), not including other factors. Therefore, the variation in filter efficiency in the case of a weak 0–10 μm dust cake becomes significant.

Figure 7 shows filter efficiency versus particle size and dust loading for SAE fine dust at a velocity of 5 cm/s. It is assumed that the dust particles do not change their size passing through the filter. It is a very stable process at this relatively low velocity. This process becomes unstable at a velocity of 50 cm/s for 0–10 μm dust (Figs. 8 and 9). At loadings over 450 g/m^2, the efficiency is negative due to seepage (Fig. 4). Particles of all sizes penetrate the media.

Lessmann and Rodman (1987) found that the overall filter efficiency for SAE fine dust is a linear function of the m/d_{fA}^2 ratio, where m = the basis weight of the filter media and d_{fA} = the average diameter of the fiber in the filter. No general model is available to determine filter efficiency and penetration for other polydisperse dust.

Pinholes represent a particularly active area of dust re-entrainment. The random spatial arrangement of fibers in nonwoven synthetic filter media and cellulose papers results in a wide pore size distribution. Large pores, offering paths of low restriction to aerosol flow, act as the source of pinholes (Bauman, 1987; Dennis and Surprenant, 1978). As other regions

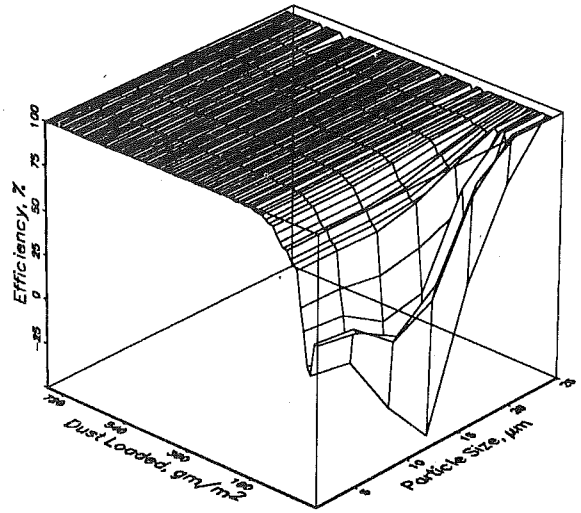


Fig. 7 Efficiency of media A at a velocity of 5 cm/s for SAE fine dust

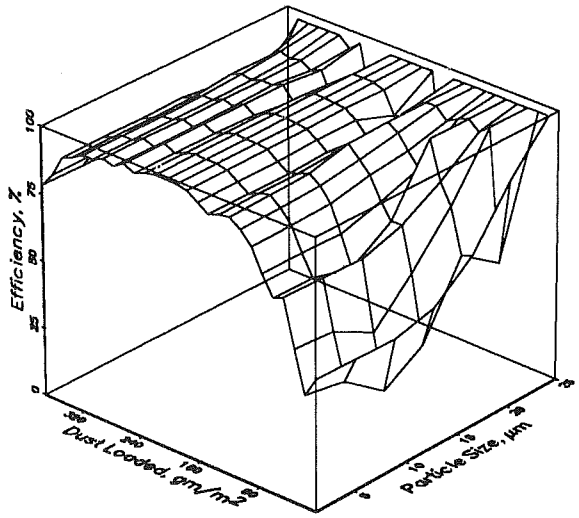


Fig. 8 Efficiency of media A at a velocity of 50 cm/s for 0–10 μm dust

of the filter become increasingly restricted by dust accumulation, the velocity increases through these pores, resulting in penetration of both the uncollected and re-entrained particles.

Pinholes have been observed for fine and coarse dusts in high-porosity surface-type and depth-type nonwoven synthetics and cellulose papers. The number of pinholes increases with increased aerosol velocity. At a velocity of 50 cm/s, the pinholes cover practically the entire surface of the filter. The weak dust cake collapses in larger pores. In some large pores, fibers are covered with dust deposits formed by very fine particles, leaving the space between fibers open to the aerosol flow. This is a result of the flow pattern around a fiber. The very fine particles remain in the boundary layer where flow velocity decreases to zero; therefore, the aerodynamic drag is weaker than the adhesive force and cannot drive the particles through the media.

There are fewer large pores during filtration of coarser dust. In this case, fine and coarse particles form compact dust aggregates, which fit tightly into the pores of the filter media. Only very large pores remain open. However, numerous, very small pinholes cover the entire surface area of the media in this case, also.

Cellulose-type media are used in automotive air filtration at moderate velocities up to approximately 25 cm/s. Due to space limitations in new car designs, higher velocities through the filter media are required. Our experimental investigations have

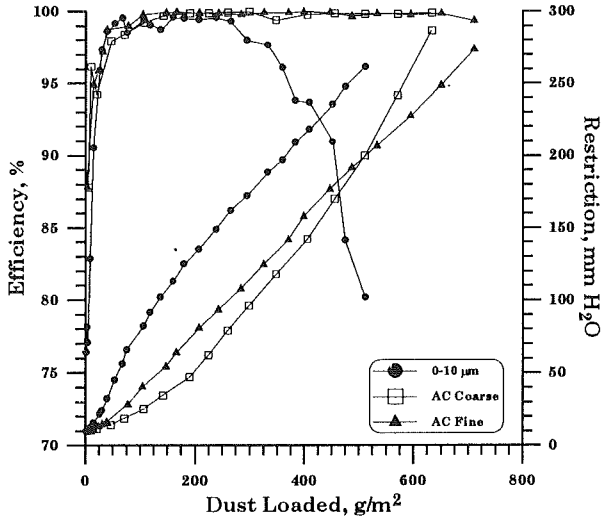


Fig. 9 Performance of media B at 50 cm/s

shown that the media behave like the nonwoven synthetics. Figure 9 shows the filter performance of cellulose-synthetic type media for all three types of test dust. The efficiency characteristic curves show small variations due to re-entrainment of the coarse and fine dusts. Significant dust disloading takes place in the case of 0–10 μm dust. In spite of the evident re-entrainment for dust loadings over approximately 250 g/m^2, the pressure drop is still higher than for more permeable dust cakes formed by the coarse and fine dusts. This is a result of numerous pinholes formed under these test conditions. At dust loadings over 500 g/m^2, the dust deposit of SAE coarse fills up the pores so that the filter becomes more restrictive than for SAE fine dust.

VI Conclusions

Engine wear caused by dust particles depends on dust concentration and particle size distribution. The wear rate can be minimized by employing finer filters. Design, selection, and evaluation of air filters for engines is very complex. Filter performance characteristics are affected by the type of dust and flow velocity in a rather random manner at high aerosol velocities. The filter itself can become a source of dust particles from dust re-entrainment. Due to limited understanding of the filtration process in high-permeability media during fine dust filtration at high velocities, there is no reliable method for simple filter selection nor ratings.

Acknowledgments

The authors wish to thank Nelson Industries, Inc. and our colleagues for their support. Special thanks are due to P. Steaffens, T. Sonsalla, K. Anderson, and B. Schwandt for their help in various aspects of this work.

References

Baczewski, K., and Jaroszczyk, T., 1981, "Laboratory and Field Testing of Air Filters," *Proceedings, Filtech Conference*, London, pp. 197–202.

Baumann, R. E., 1987, "The Evidence for Wormholes in Deep Bed Filters," *Proceedings, Filtech Conference*, Utrecht, Holland, Sept. 23–25, pp. 151–164.

Caskey, M. R., 1982, "Dust and Sand Protection for Marine Gas Turbines," *Transactions of the ASME*, Vol. 104 (April), pp. 260–267.

Chen, C. Y., 1955, "Filtration of Aerosols by Fibrous Media," *Chem. Reviews*, Vol. 55, pp. 595–623.

Claes, J., and de Bruyne, R., 1975, "Demisting With Metal Fibre Webs and Felts," *Proceedings of the Filtration Society Conference*, Sept. 16–19, London, pp. 9–13.

Coates, G. L. A., and Tull, E. V., 1963, "A Search for the Perfect Engine Air Cleaner," SAE Technical Paper No. 305B, pp. 50–57.

Cockle, G. R., Houser, F. N., and Koch, E. M., eds., 1974, *Car and Locomotive Cyclopedia of American Practices*, Centennial Edition, Simmons-Boardman Publishing Corp., New York, pp. S16-60.

Davies, C. N., 1952, "The Separation of Airborne Dust and Particles," *Proceedings, Institute of Mechanical Engineers*, Vol. 113(5), pp. 185-213.

Davies, C. N., 1973, *Air Filtration*, Academic Press, New York.

Dennis, R., and Surprenant, N. F., 1978, "Particulate Control Highlights: Research on Fabric Filtration Technology," EPA Report No. 600/8-78-005d.

Dyakov, R. A., 1975, *Air Cleaning in Diesels* [in Russian], Mashinostroyenye, Leningrad.

Ellenbecker, M. J., Leith, D., and Price, J. M., 1980, "Impaction and Particle Bounce at High Stokes Numbers," *J. Air Pollut. Control Assoc.*, Vol. 30(11), pp. 1224-1227.

Flagan, R. C., and Seinfeld, J. H., 1988, *Fundamentals of Air Pollution Engineering*, Prentice Hall, Inc., Englewood Cliffs, NJ.

Fodor, J., 1982, "Improving the Economy of I. C. Engines by Controlling the Contaminants Through Filtration," *Proceedings, World Filtration Congress III*, Uplands Press Ltd., Croydon, United Kingdom, pp. 707-711.

Fried, E., and Idelchik, I. E., 1989, *Flow Resistance: A Design Guide for Engineers*, Hemisphere Publishing Corp., New York.

Grant, D. C., Liu, B. Y. H., and Fisher, W. G., 1989, "Particle Capture Mechanisms in Gases and Liquids: An Analysis of Operative Mechanism of Membrane/Fibrous Filters," *J. Environ. Science*, pp. 43-51.

Grigoryev, M. A., Sokolov, V. V., Borisova, G. V., and Ruzayev, I. G., 1975, "Experimental-Theoretical Investigation of Oil and Air Filter Performance Effect on Engine Wear" [in Russian], *NAMI Journal*, Moscow, pp. 34-43.

Grigoryev, M. A., and Ponomaryov, N. N., 1976, *Wear and Engine Life* [in Russian], Mashinostroyenye.

Idelchik, I. E., 1986, *Handbook of Hydraulic Resistance*, 2nd ed., Hemisphere Publishing Corp., Washington, D.C., 1966 [translation from Russian].

Jaroszczyk, T., 1987, "Experimental Study of Dust Distribution Around Motor Vehicles in Motion," *Particulate and Multiphase Processes*, Vol. 2, T. Ariman and T. N. Veziroglu, eds., Hemisphere, New York, pp. 417-426.

Jaroszczyk, T., Hoops, R. H., and Kreikebaum, G., 1987, "Measurement of Engine Air Filter Efficiency Using a Continuous Aerosol Monitoring System," SAE Tech. Paper 872268.

Jaroszczyk, T., and Wake, J., 1991, "Critical Aerosol Velocity in Nonwoven Filtration," *Proceedings, Tappi 1991 Nonwoven Conference*, pp. 125-135.

Khorshid, E. A., and Nawwar, A. M., 1991, "A Review of the Effect of Sand Dust and Filtration on Automobile Engine Wear," *Wear*, Vol. 141, pp. 349-371.

Koffman, J. L., 1953, "The Cleaning of Engine Air," *Gas and Oil Power*, Vol. 48, pp. 60-63.

Lee, K. W., and Liu, B. Y. H., 1982, "Theoretical Study of Aerosol Filtration by Fibrous Filters," *Aerosol Science & Technology*, Vol. 1, pp. 147-161.

Leith, D., and Ellenbecker, M. J., 1980, "Theory for Penetration in a Pulse-Jet Cleaned Fabric Filter," *J. Air Pollut. Control Assoc.*, Vol. 30 (8), pp. 877-881.

Lessman, R. C., and Rodman, C. A., 1987, "Automotive Nonwoven Filter Media, Their Construction and Filter Mechanisms," presented at INDA-TECH 87, pp. 1-16.

Löffler, F., 1970, "Separation Efficiency and Pressure Loss of Filter Materials of Different Structure, at Different Conditions," *Staub-Reinhalt, Luft* (English edition), Vol. 30 (12), pp. 27-31.

Mayev, V. E., and Ponomaryov, N. N., 1971, *Air Cleaners for Automotive and Tractor Engines* [in Russian], Mashinostroyenye, Moscow.

Midler, F. L., Armini, A. J., and Jones, G. W., 1981, "The Use of Surface Layer Activation Wear Monitoring for Filter Design and Evaluation," SAE Technical Paper No. 810329.

Morton, D. W., 1970, "Match the Engine Intake System to the Application and Environment," *Engineering Know-How in Engine Design—Part 18 (SP-359)*, Society of Automotive Engineers, pp. 1-6.

Myedvyedyev, V. N., Nyevolin, V. F., and Kagan, M. R., 1984, "A Method of Pressure Drop Calculation in Cellulose-Type Filters for Dust Free Air and Under Dust Loading" [in Russian] *Engine Construction*, No. 5, pp. 25-27.

Needelman, W. M., and Madhaven, P. V., 1988, "Review of Lubricant Contamination and Diesel Engine Wear," SAE Technical Paper No. 881827.

Nietzold, I., 1979, *Luftfiltration* (Polish translation), VEB Verlag Technik, Berlin.

Nowicki, M., 1971, "The Process of Non-steady Aerosol Filtration in Fibrous Filters" [in Polish], *Proceedings, International Symposium of Polish Academy of Sciences*, Zabrze, Poland, Dec. 15-18, 1971.

Pich, J., 1987, "Gas Filtration Theory," *Filtration: Principles and Practices*, 2nd ed., M. J. Matteson and C. Orr, eds., Marcel Dekker, Inc., New York, 1987.

Pochtarov, N. F., 1957, *The Effect of Air Dustiness on Piston Wear* [in Russian], Military Publisher, Department of Defense, Moscow.

Sayfert, N., 1988, "Collection, Bouncing, and Adhesion on Crossed Fibers in Mono- and Multi-layer Design," *Aerosol Science & Technology*, Vol. 9, pp. 201-211.

Stenhouse, J. I. T., Broom, G. P., and Chard, N. T. J., 1978, "High Inertia Fibrous Filtration—Optimum Conditions," *Filtration & Separation*, pp. 138-146.

Stinson, J. A., Meyers, M. N., Jaroszczyk, T., and Verdegan, B. M., 1989, "Temporal Changes in Oil and Air Filter Performance Due to Dust Deposition," *Filtration & Separation*, pp. 368-371.

Thomas, G. E., and Culbert, R. M., 1968, "Ingested Dust, Filters, and Diesel Engine Ring Wear," SAE Technical Paper No. 680536, pp. 362-369.

Wake, J., and Jaroszczyk, T., 1991, "Experimental Study of Dust Filtration in Surface-Type Nonwovens," *Particulate Science & Technology*, Vol. 9, pp. 31-44.

Watson, C. E., Hanley, F. J., and Burchell, R. W., 1955, "Abrasive Wear of Piston Rings," *SAE Transactions*, Vol. 63, pp. 717-726.

Wirski, W., 1974, "The Filtration Velocity Influence on Filter Efficiency of Fibrous Filters" [in Polish], *Bull. Mil. Acad. Technol. Warsaw*, Vol. 23(9), pp. 105-111.

Progress in Vapor Phase Lubrication Technology

D. G. Placek
FMC Corporation,
Princeton, NJ 08543

T. Freiheit
Detroit Diesel Corporation,
Detroit, MI 48239

Improving the efficiency of engine performance will require the design of systems with higher operating temperatures and pressures. These conditions will stress traditional lubricants beyond their current performance capabilities, and require the development of improved methods for friction and wear reduction. The most revolutionary approach to high-temperature lubrication is the concept of vapor phase delivery. An ashless organic compound can be vaporized by the heat of the operating engine or a carrier gas, and introduced into the ring zone of the cylinder. The vapor condenses or chemically binds with the piston ring or cylinder wall, and provides boundary lubrication. A minute amount of lubricant is constantly introduced in order to maintain a lubricating film. Each stroke of the piston shears off a portion of the lubricant layer, but condensing vapor continually replaces the surface film. Lubricant contributions to exhaust emissions are expected to be lower than those currently resulting from liquid lubricants. Vapor phase lubrication is an emerging concept that may be the key to the development of a commercial low heat rejection engine with improved energy efficiency and reduced emissions. The Department of Energy continues to fund research at a variety of industrial and academic institutions. Basic concepts and recent developments in the field of vapor phase lubrication will be reviewed.

Introduction

Improving the energy efficiency of the internal combustion engine will require new designs that are smaller, lighter, and consume less fuel. Current engine technology allows approximately one third of the energy generated to escape through exhaust gases, and another third is intentionally absorbed by the cooling system. If engines are to become more efficient, new designs must utilize more of the heat of combustion that is rejected to the surrounding air.

Engines with improved efficiency will operate at higher temperatures and pressures. These increases can be achieved by eliminating the external cooling system, insulating the combustion chamber, and turbocharging with exhaust gases. All of these modifications can make positive contributions to energy efficiency, but they each pose considerable challenges to existing materials of construction and lubrication technology. The concept of an adiabatic (no heat loss) engine was pursued, but found to be impractical. When an engine is insulated to run hotter, it will heat up the air that is drawn into the combustion chamber, reducing its density. The additional work required to compress the hot air cancels out the efficiency gains obtained from insulation. A review of early engine designs and testing (Reddy et al., 1990) indicates that there is a practical limit on the amount of heat that can be transformed into power.

Efficiency gains can still be realized by insulating the combustion chamber with ceramic materials, and capturing some of the additional energy available by eliminating the external cooling system. The concept is now referred to as the Low Heat Rejection (LHR) Engine, and has made significant strides due to the support of the U.S. Department of Energy and a wide variety of academic, industrial, and government laboratories. The U.S. Department of Energy established the Heavy Duty Transport Technology Program in 1982, and continues to sponsor research into the development of an improved diesel engine.

The Heavy Duty Transport Technology Program has established several goals for a heavy duty diesel engine design that will be commercialized by the year 2000 (Fairbanks, 1991). Fuel consumption should be reduced to the target of 0.25 lb/bhp-hr, which is a 22 percent reduction from the 1991 average of 0.32 lb/bhp-hr, and a 30 percent reduction from the 1982 average of 0.35 lb/bhp-hr. The new technology should be immediately applicable to off-highway, in-land marine, railroad, and light-heavy duty engines. All new engine designs must meet the proposed EPA emission standards that have been set for 1994 and 1998. In addition, the potential use of alternate fuels (methanol and/or natural gas) must be considered. Table 1 contains a comparison of current and projected diesel engine performance parameters.

A significant amount of research and development is still required to tackle challenges in the areas of thermal barrier coating, ceramic materials development, ceramic component design and fabrication, and engine tribology. Reducing the friction and wear associated with ceramics has been studied in a variety of ways. In general, ceramic on ceramic rubbing contacts result in high wear (Braza et al., 1989; Kawamura, 1990; Ajayi et al., 1992), due to the hardness of the materials.

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1992. Paper No. 93-ICE-19. Associate Technical Editor: J. A. Caton.

Table 1 Heavy-duty diesel engine performance*

Engine performance	1991 Commercial production	2000 (Projected) LHR engines
BSFC (lb/bhp-hr) (brake specific fuel consumption)	0.32	0.25
Peak cylinder pressure (PSI)	2000	2800–3400
BMEP (PSI) (brake mean effective pressure)	310	435
Fuel injection pressure (PSI)	20,000	25,000
Turbocharger efficiency	58 percent	74 percent
Turbocompounding	No	Yes
Top ring reversal Temperature (°F)	400	730–850
EPA emission limits (gm/bhp-hr)		
Particulate	0.25	0.05
No_x	5.0	2.5
Hydrocarbons	1.3	1.3
Carbon monoxide	15.5	15.5

*Fairbanks (1991).

Research continues in the area of "self-lubricated" or low friction ceramics. Improvements can be achieved by modifying ceramic grain size or surface topography (polishing), altering ceramic composition (cermets), or the application of surface coatings. Solid, liquid, and vapor phase lubrication schemes have been evaluated to improve the life of ceramic parts in motion.

Surface coating and solid lubricants can reduce friction and wear, but they are slowly consumed as surface layers are sheared away. In addition, a coating must have an expansion coefficient that is similar to the substrate, or fatigue cracks will appear as the materials cycle through temperature extremes. This limits the number of coatings that can be used with any given ceramic substrate. As engine parts wear, nonvolatile debris will build up in contact points and must be removed. Scuffing and seizures occur as a result of accumulated wear debris or deposits that get into the wear track. Most commercial solid lubricants also have a limited operating range, as they will oxidize at temperatures above 400°C (Sliney, 1987).

Liquid lubricants represent the current state of the art for the lubrication of combustion engines. However, significantly higher demands will be placed upon a liquid lubricant if it is to successfully perform in an LHR engine. A single fluid will be used to lubricate moving parts and also to remove heat from critical areas. Some initial LHR engine designs utilized "precision" cooling, which eliminated the external block water cooling system providing oil cooling only to certain critical areas that require lower operating temperatures. Lubricating oil is circulated to cool hot spots in the cylinder head and liner, and heat is dissipated through an oil–air heat exchanger. Thermal insulation is provided by ceramic thermal barrier coatings on critical combustion chamber components. This type of design calls for a highly stable liquid lubricant that can withstand top ring reversal area temperatures of 400°C or higher, and sump temperatures of 250°C. Mineral oil lubricants are clearly unacceptable for this range of temperature operation, as are most lubricant additives currently used for wear reduction and oxidation stability.

Researchers developing high stability liquid lubricants are optimizing synthetic hydrocarbon technology, but all hydrocarbons have ultimate limitations. At temperatures above 400°C, carbon–carbon bonds are broken and liquid lubricants are subject to rapid decomposition. Researchers anticipate that liquid lubricants will degrade under LHR engine conditions, and so any functional product must degrade into volatile fractions that do not cause deposits, sludge, or varnish. The additives used in these liquid lubricants must be ashless, and have stability and volatility that are equal or superior to the base fluid. Additive development for high temperature lubricants is an additional area of concern that may ultimately limit the use of liquid lubricants (Perez et al., 1991).

The concept of vapor phase lubrication has been proven an effective way to reduce friction and wear at temperatures above 200°C. Organic vapors have been used to lubricate molten metal die casting operations at temperatures as high as 1600°C (Klaus and Lai, 1976), and the concept is predicted to be useful at temperatures as high as 1900°C (Klaus and Duda, 1991). Phosphate esters were successfully used as vapor phase lubricants in a commercial aluminum die casting operation (Lai, 1974; Klaus and Lai, 1983), which suggested that the concept might find use in other high-temperature lubrication applications. The DOE continues to support research into LHR engines lubricated with vaporized organic compounds, and several industrial and academic laboratories are advancing the technology with the intention of producing commercial engines within ten years.

Vapor Phase Lubrication Concept Development

An ashless organic compound can be vaporized by a carrier gas, or directly by engine heat, and introduced into the piston ring zone of the cylinder. The lubricant must have sufficient stability to be vaporized and transported into the contact zone without decomposing, and it must not form deposits once consumed as a boundary wear film. As the vapor enters the ring zone, it condenses or chemically binds with the piston ring and/or cylinder liner. Ideally, the rate of vapor flow is carefully controlled so that the rate of new film formation is equal to the rate of existing film consumption. Each stroke of the piston shears off a layer of the lubricant film, which is immediately replenished by condensing vapor. There is no excess lubricant available, as with a liquid, so lubricant contributions to the exhaust could actually be reduced.

The earliest work in this area was performed by Fein and Kreuz (1965), which indicated that cyclohexane vapor reduced wear better than cyclohexane liquid in four-ball wear tests. The VPL concept was expanded and proven effective in a variety of laboratory bench tests and prototype engines starting in 1984. Initial testing demonstrated that the friction and wear of steel bearings could be reduced by introducing tricresyl phosphate vapors (Klaus et al., 1989).

Efforts to optimize the lubrication process have focused on lubricant interaction with engine materials. Ceramic materials are obvious candidates for use in LHR engines due to their stability and inertness in extreme environments. Unfortunately, this makes the application of existing lubricating technology impossible. None of the ceramic materials tested to date allow for the formation of a chemically bound lubricant film. Organic vapors form lubricating films on ceramics merely by condensation. As a liquid film accumulates, it may degrade to form a solid or polymeric surface layer. These hydrodynamic lubricating films offer little wear protection in severe boundary wear conditions like those encountered at the temperatures and pressures found in an LHR engine. Organophosphates such as zinc dialkyldithiophosphates (ZDP's) and phosphate esters have been used for many years as anti-wear and extreme pressure additives in engine lubricants and industrial oils. The phosphorus portion of the molecule reacts with a steel surface to form an extremely effective lubricating film of iron phosphate. Ashless phosphorus compounds delivered as vapors rapidly form boundary lubricant films on surfaces that contain iron or copper (Klaus et al., 1989). Film formation is accelerated when the lubricant is catalytically attracted to the surface. Rates of film formation on various substrates have been studied from 200°C to 900°C (Gunsel et al., 1989; Makki and Graham, 1990; Klaus and Duda, 1991). Table 2 is a compilation of lubricating compounds and substrates that have been studied

Table 2 Vapor phase lubricants and substrates tested

Lubricants
Phosphate esters
TBP—Tributyl phosphate
TCP—Tricresyl phosphate
TBPP—t-butylphenyl phosphate
Organic esters
TMPTH—Trimethylolpropane-triheptanoate (Polyol ester)
Organic ethers
PPE—Polyphenyl ether
PFPE—Perfluoropolyethers
PGE—Polyglycol ether
Hydrocarbons
Mineral oil
Super-refined mineral oil
PAO—Polyalphaolefin
Ethylene
Substrates
Stainless steel
Carbon steel
Iron
Silicon nitride
Silicon carbide
Silicon dioxide (quartz)
Sialon (Si-Al-O-N)
Copper
Aluminum
Nickel
Nickel aluminide
Tungsten

and reported on in the literature. The studies indicate that hydrocarbons, esters, and ethers condense to form films on all substrates at a constant rate, which is dependent upon vapor density and surface temperature. Hydrocarbons will decompose to produce a carbonaceous or graphitic surface layer, which can significantly reduce the coefficient of friction and rate of wear in laboratory tests (Lauer and Dwyer, 1989, 1990).

Phosphate esters also exhibit a constant rate of film formation on all ceramics and certain metals, but have an accelerated rate on iron or copper containing surfaces. Iron phosphate boundary layers are well known for their ability to provide effective protection against wear. The phosphate esters vaporized onto active substrates form tenacious, polymeric films compared to hydrocarbons which form loose flaky films (Makki and Graham, 1991). All organic lubricants form flaky deposits on nonactive surfaces. Multiple research efforts have created a basic understanding of the mechanism of film formation on standard ceramic and metallic surfaces (Klaus, 1988; Klaus et al., 1989, 1990; Gunsel et al., 1989; Makki and Graham, 1990, 1991; Lauer and Dwyer, 1990).

Vaporized phosphate esters also offer safety advantages over hydrocarbons, as they will not sustain combustion and do not represent the same degree of fire or explosion hazard. Phosphate esters currently find use as fire-resistant hydraulic fluids and flame-retardant additives for plastics. The combination of stability, reactive film formation, and safety characteristics led to the development of a test engine with ceramic piston rings and a cast iron cylinder liner lubricated with phosphate ester vapor. This combination has been demonstrated to be a practical design for LHR engines that can operate with a relatively low level of wear.

The lubricating performance of any vapor can be altered by manipulating certain variables. Lubricant vapor concentration, carrier gas composition, vapor temperature, and surface temperature play important roles in optimizing performance. Insufficient vapor will obviously result in poor lubrication, but excess vapor can have the same effect. Close control of lubricant vapor deposition rate is critical, as low deposition rates have actually been found to cause higher rates of wear than no lubrication at all. The selection of an air or an inert gas environment will determine if oxidation or thermal degradation will occur. Tribological concerns should dictate which mode of decomposition is desired, and which mode should be prevented. Detailed studies of vapor deposition kinetics have been performed that consider many of these parameters (Klaus et al., 1989; Gunsel et al., 1989).

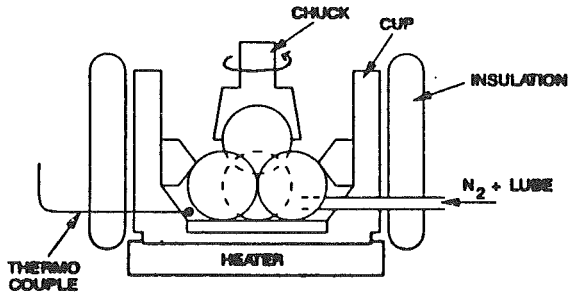


Fig. 1 Four-ball wear tester in vapor lubrication mode

Vapor Phase Lubrication in the Development of Low Heat Rejection Engines

Lubricants for commercial LHR engines are expected to experience top piston ring reversal area temperatures from 400°C to 550°C. The concept of vapor phase lubrication appears to be a promising technique that will allow standard organic lubricant compounds to be useful at these extreme conditions.

The first practical demonstration of high-temperature lubrication of both metals and ceramics was achieved with the four-ball wear tester. In this test, three balls are locked into a stationary position and maintained at a constant temperature. A fourth ball is loaded against the three and rotated at constant speed for a set period of time (Fig. 1). Flat disks can be substituted for the three lower balls if necessary. The average wear scar size is recorded for comparison. Data generated in four-ball wear tests have accurately predicted the actual performance of industrial hydraulic fluids and automotive or diesel engine oils (Klaus and Perez, 1983). Four-ball wear tests have recently been correlated as accurate predictors of diesel engine simulators and ceramic test engines (Klaus and Duda, 1991).

Four-ball wear tests conducted with steel-on-steel and steel-on-ceramic demonstrate that many conventional lubricants can be delivered in the vapor phase to reduce wear. In fact, 0.5 mole percent tricresyl phosphate (TCP) delivered as a vapor at 370°C provided better wear reduction than a commercial 10W-30 diesel engine oil, an industrial EP oil, or an R&O hydraulic fluid, used under full liquid emersion at 75°C (Klaus, 1988; Klaus et al., 1989). The phenomenon of low steady-state wear in the four-ball test is an excellent indication that the lubricant is forming an effective boundary lubricant film in the contact zone (Klaus and Perez, 1983). This type of performance is typically seen when phosphorus antiwear additives are present. Hydrocarbons do not chemically bind with the steel surface to create a boundary antiwear film, and do not exhibit a reduced wear rate after the break-in period.

Additional four-ball work has demonstrated that mixtures of TCP and mineral oils can be delivered as vapors to reduce wear better than either individual lubricant (Klaus and Duda, 1991). The synergy obtained from a mixture of phosphate ester and mineral oil has been previously identified in liquid lubricants, and is therefore not surprising. The "friction polymer" formed from a combination of iron phosphate and oxidized hydrocarbon is an excellent antiwear boundary lubricant.

These data suggest that an LHR engine should be designed to operate with a single liquid lubricant in the lower portion of the engine, which could be vaporized into the cylinder for upper engine lubrication. Sump temperatures should be kept low enough to allow the use of mineral oil or synthetic hydrocarbon fluids. An existing LHR engine design by Detroit

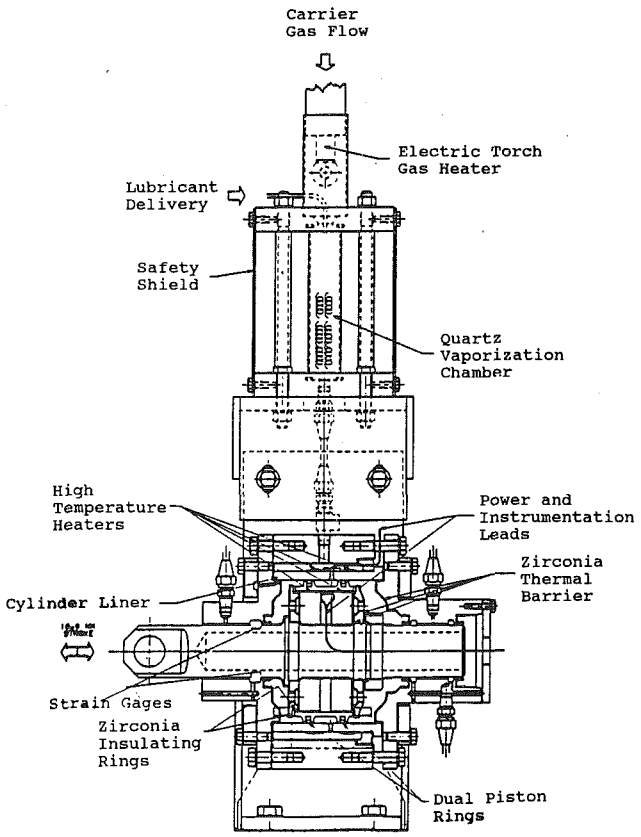


Fig. 2 Cylinder kit tribology test fixture and vapor generator diagram

Diesel eliminates the need for a camshaft by using electrohydraulic valve actuators. The piston ring contact zone would be the only hot spot in the upper portion of the engine that could not be served by liquid lubricants.

Four-ball wear testing has also been used to optimize lubricant vapor concentrations and flow rates required for maximum wear reduction (Klaus et al., 1983). Excessive vapor concentration can lead to higher rates of wear, which is believed to be related to deposit formation. Excess lubricant film not consumed in the wear track can build up, trap wear debris, and become abrasive (Klaus and Duda, 1991).

A minimum lubricant film is also desirable for exhaust emission considerations. In current engine design, excess liquid lubricant is scraped up the cylinder wall into the combustion zone. Lubricating oil that does not burn completely and is vaporized into the exhaust becomes a significant contributor to particulate and hydrocarbon emissions. The absence of excess lubricant will eliminate this source of undesirable emissions. In addition, a thick lubricant film can absorb a small portion of the fuel injected into the chamber, and allow it to escape combustion. When the chamber is exhausted, these solubilized hydrocarbons are volatilized from the lubricant film, contributing to hydrocarbon emissions.

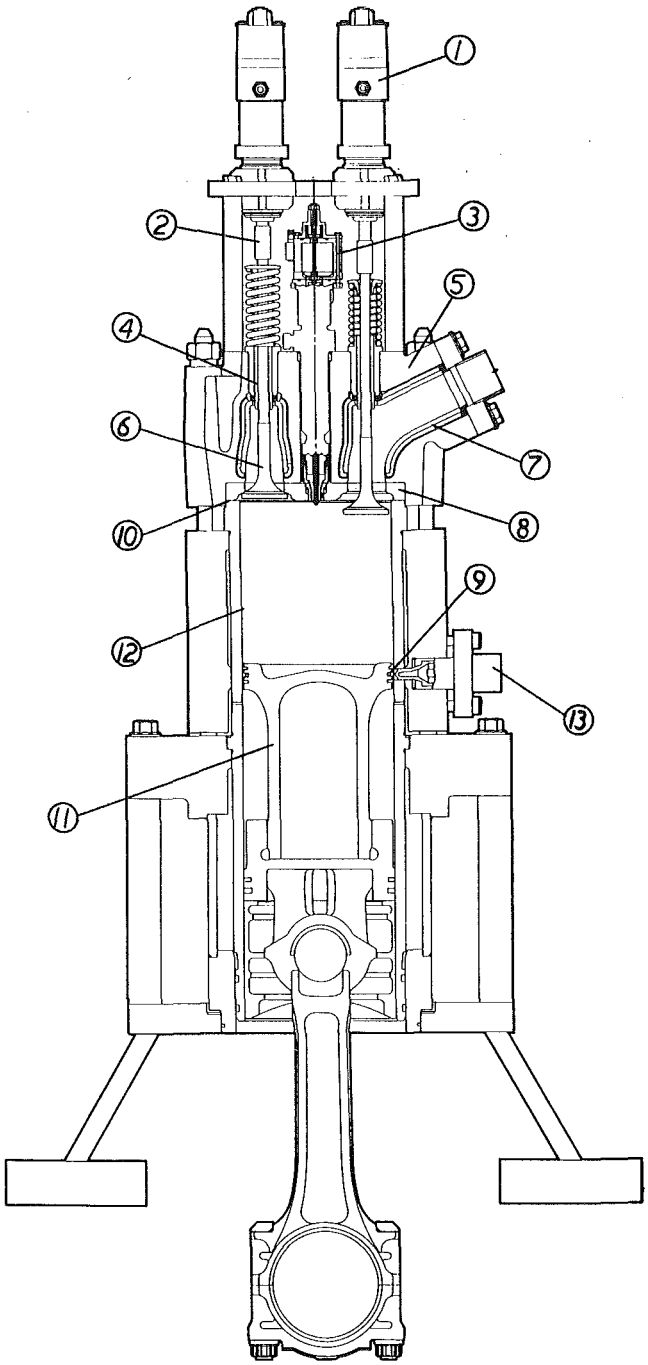
Other laboratory bench tests have been used to demonstrate the effective friction reducing capabilities of low-molecular-weight hydrocarbons like ethylene. A pin on disk tribometer indicates that ethylene vapors directed onto ceramic surfaces at temperatures above 450°C can quickly and dramatically reduce the coefficient of friction from 0.6 to 0.05. This reduction is significant, as the coefficient of friction in current engine technology is on the order of 0.05. Hydrocarbons like ethylene, benzene, and cyclohexane will decompose to form graphitic lubricating films that can significantly reduce friction and wear in ceramic contact points at high temperatures (Lauer and Dwyer, 1990).

The development of a practical high-temperature cylinder kit/lubricant bench test represents a significant technological hurdle. The initial screening of vapor phase lubricants requires a test apparatus capable of simulating the actual ring/liner interface in an LHR engine. Detroit Diesel designed a Cylinder Kit Tribology Test Fixture (CKTTF) that will allow the control of ring speed, ring normal force, and in-cylinder temperatures (Fig. 2). The fixture allows two piston rings to reciprocate against a section of cylinder liner at temperatures close to those observed in actual engine cylinder kits. The vaporized lubricant is introduced in between two piston rings, which reciprocate over a 10 mm wear track. The test rig has been helpful in evaluating the relative performance of different organic vapors, as well as differences in ring coatings and surface roughness of liners.

Several phosphate esters have been evaluated in the CKTTF with ceramic coated rings on cross-hatched and ground finish cast iron liners (Groeneweg et al., 1991). At a liner temperature of 350°C, tests were conducted comparing phosphate ester vapor lubrication with no intentional lubrication (NIL) conditions. In the NIL condition, nothing is intentionally introduced into the ring zone for the purpose of lubrication. Some compounds originating from uncombusted fuel, exhaust blowby, or lower engine lubricant volatiles may migrate into the piston ring wear track during normal operation. The presence of these compounds cannot be completely eliminated, but they are not believed to offer any friction or wear reduction. The NIL condition represents the frictional condition expected from direct metal on metal or ceramic on metal contact in an operating engine. The phosphate esters performed as the four-ball wear tests predicted, significantly reducing both wear and the coefficient of friction (Groeneweg et al., 1991). Optimum vapor delivery rates were determined that minimized wear, as measured by weight loss and surface topography. A lubricant concentration of approximately 0.1 mole percent tricresyl phosphate minimized the wear on ground finish liners, with only minor changes observed in the ring and cylinder profile traces. In addition, the surface profiles of the vapor phase lubricated rings resembled typical examples of conventionally lubricated piston rings. Tests carried out at both low and high ring loading indicate that wear coefficients associated with high ring load (1210 N) were lower than those measured at low ring load (220 N) (Weber, 1990; Groeneweg et al., 1991). The high ring loads in this test correlate to the actual BMEP of a high output diesel engine (1.48 MPa). The CKTTF test fixture allows various lubricant/ceramic/metal combinations to be screened with greater confidence before proceeding to actual engine testing.

Several methods of vapor delivery have been proposed for an actual engine. The vapor could be carried to the cylinder walls by the charge air, or as a component of the fuel. Both of these methods are undesirable as they would result in excessive lubricant consumption. Only a small portion of the lubricant would actually reach the cylinder wall due to the short residence time in the combustion chamber. The bulk of the lubricant would interfere with or be partially consumed in combustion, then contribute to exhaust emissions. In the initial demonstration of vapor phase lubrication in an operating engine, it was determined that the lubricant should be vaporized external to the engine and delivered directly to the ring zone in a nitrogen gas stream. This technique allows for precise control of lubricant concentration and feed rate. An inert carrier gas (nitrogen) was selected to avoid oxidation of the lubricant prior to its delivery.

Detroit Diesel designed a single-cylinder test bed (SCTB), which has been used to demonstrate vapor phase lubrication in an operating engine. This test bed currently represents the most advanced simulation of an LHR engine utilizing the vapor phase lubrication technique. As shown in Fig. 3, the SCTB includes a variety of experimental engine systems and components, including: electrohydraulic valve actuation, common



Key	Description
1	Electrohydraulic Valve Actuator
2	Floating Valve Bridge (not guided)
3	Common Rail Fuel Injector
4	Silicon Nitride Valve Guide
5	High Temperature Cylinder Head
6	Silicon Nitride Valve
7	Cast-in-place Silicon Nitride Port
8	Silicon Nitride Firedeck
9	Cermet Piston Ring
10	Gasketless Interface
11	Silicon Nitride Piston Cap
12	Nitralloy Liner
13	**Electronically Controlled In-cylinder Lubricant Injector**

Fig. 3 Single-cylinder test bed diagram

rail fuel injection, ceramic piston, valves, valve guides and fire deck, cermet compression piston ring, and cast-in-place ceramic port shields.

The piston/cylinder system was designed to demonstrate the use of vapor phase lubrication in the "hot" zone at the top of the liner, and liquid lubrication in the physically separate "cool" zone at the bottom of the liner. The height of the piston did not allow the "hot" compression rings to overlap the "cool" oil control rings at the bottom of the liner.

Comparative tests of vapor phase lubrication (VPL) versus no intentional lubrication (NIL) were conducted on the SCTB (Groeneweg et al., 1992). Utilizing tricresyl phosphate vapor at 0.1 mol percent, the VPL concept was successfully demonstrated in an uncooled engine at light loads with top ring reversal temperatures exceeding 400°C. Significant reductions in wear coefficients were exhibited for both the ring and liner compared to the NIL condition. The VPL ring running surface appeared clean and relatively smooth whereas the NIL ring showed significant amounts of adhered liner material and surface damage. Post-test inspection of the piston rings identified that the NIL ring had fractured into two pieces. Furthermore, the VPL ring surface appeared suitable for continued testing, while the NIL ring would have been unusable even if it had not fractured.

The testing of VPL at higher power conditions was arduous due to exhaust blowby problems. As power output was increased, the temperature and loading on the top compression ring increased, resulting in extreme temperature gradients in the ring and liner. Top ring reversal (TRR) temperatures of up to 470°C were measured. The thermal gradients produced ring and liner distortions, which caused excessive, erratic, and unstable levels of blowby (Weber, 1989). Modifications to the injection system improved combustion and reduced in-cylinder temperatures, with the TRR temperature dropping to 340°C. Further increases in power and intake boost pressures continued to produce problematic blowby, ultimately requiring redesign of the ring pack (Freiheit, 1991). Despite these problems, one build of the SCTB operated up to 16 percent of rated power for 17.1 hours with no evidence of liner scuffing or scoring. Analysis of the liner confirmed the presence of tricresyl phosphate and its reaction products, with similar amounts of iron and phosphorus compared to samples previously taken from the CKTTF and earlier engine builds.

In total, over 100 hours of engine operation at up to 50 percent of rated power were conducted with vapor phase lubrication and this excessive blowby condition (Weber, 1990). These positive vapor phase lubrication results led to a more practical design of the VPL delivery system. The nitrogen carrier gas stream was replaced with an electronically controlled liquid lubricant injector (Weber, 1990). The liquid phosphate ester is introduced into an inter-ring lubricant pocket and dispersed by the heat of combustion and motion of the piston. This design permits a direct delivery of the lubricant to the critical ring pack area. The versatility of this lubricant delivery system will also allow for the direct injection of liquid lubricant at start-up and lower temperature conditions. This new lubricant injector design may also be applicable to current diesel designs, where a controlled delivery of liquid lubricant may reduce oil consumption as well as particulate and hydrocarbon emissions.

Comparative tests of the electronically controlled lubricant injector and the original nitrogen carrier gas method on the SCTB have indicated that the injector concept is promising (Freiheit, 1992). A lubricant injector test run conducted at light loads with a higher tension ring pack designed for improved blowby control exhibited reduced ring wear and blowby compared to the carrier gas technique. However, increases in friction and exhaust emissions were observed. Blowby levels were stable, but not as low as the design target. Wear rates for the compression rings in both the externally vaporized and

lubricant injector runs were reduced due to vapor phase lubrication, but were much higher than wear rates encountered in current commercial diesel engines. The ceramic piston domes for both tests fractured in the ring belt area sometime during the test run. These conditions were a result of high blowby causing lubricant starvation in the ring pack, especially for the top compression ring. In the lubricant injector run, the average inter-ring lubricant concentration was estimated to be one fifth the target of 0.1 mol percent. CKTTF testing suggests that insufficient lubrication may result in higher wear rates than no intentional lubrication.

The performance of vapor phase lubrication technology in laboratory wear and single-cylinder engine tests has ensured that it will continue to be a promising technique for high-temperature lubrication in LHR engines. Future vapor phase lubrication work must address the characterization of high engine loads and cold, start-up conditions. Optimum lubricants and lubricant delivery rates must be determined for a wide range of engine operating conditions. Finally, since it is necessary to control finely the amount of lubricant delivered to the ring pack, a practical, closed-loop control system must be developed for all vapor phase lubrication delivery systems.

Conclusions

The concept of vapor phase lubrication continues to show promise as a practical method for reducing friction and wear in high-temperature contacts. This technology will continue to be explored in the development of high-efficiency, low-heat-rejection engines. A great deal of research has led to a basic understanding of lubricating vapor interaction with metals and ceramics. Engine design efforts have been supported by the development of high-temperature bench tests, which allow vapor deposition fundamentals and materials tribology to be investigated. These developments have allowed the concept of vapor phase lubrication to advance to actual engine testing. Vapor phase lubrication may be a key design parameter that will lead to the commercialization of a next generation LHR engine.

References

Ajayi, O. O., Erdemir, A., Hsieh, J. H., Erck, R. A., and Fenske, G. R., 1992, "Combined Solid and Liquid Lubrication of Silicon Nitride Under Boundary Conditions," *Lubrication Engineering*, Vol. 48, No. 7, pp. 584–591.

Braza, J. F., Cheng, H. S., and Fine, M. E., 1989, "Silicon Nitride Wear Mechanisms: Rolling and Sliding Contact," *Tribology Transactions*, Vol. 32, No. 4, pp. 439–446.

Fairbanks, J. W., 1991, "Overview of the Heavy Duty Transport Technology Program," *Proceedings of the Annual Automotive Technology Development Contractors Coordination Meeting*, Dearborn, MI, SAE Publication P-256.

Fein, R. S., and Kreuz, K. L., 1965, "Chemistry of Boundary Lubrication of Steel by Hydrocarbons," *ASLE Transactions*, Vol. 8, pp. 29–38.

Freiheit, T., 1991, "Advanced Component Development and Emissions Reduction for Heavy Duty Diesel Engines," *Proceedings of the Annual Automotive Technology Development Contractors Coordination Meeting*, Dearborn, MI, SAE Publication P-256.

Freiheit, T., 1992, "Advanced Component Development and Emissions Reduction for Heavy Duty Diesel Engines," *Proceedings of the Annual Automotive Technology Development Contractors Coordination Meeting*, Dearborn, MI.

Groeneweg, M., Hakim, N., Barber, G. C., and Klaus, E., 1991, "Vapor Delivered Lubrication of Diesel Engines—Cylinder Kit Rig Simulation," *Lubrication Engineering*, Vol. 47, No. 12, pp. 1035–1039.

Groeneweg, M., Barber, G. C., and Freiheit, T., 1992, "Vapor Delivered Lubrication of Diesel Engines—Uncooled Single Cylinder Engine Test," SAE Technical Paper No. 922351.

Gunsel, S., Klaus, E. E., and Bruce, R. W., 1989, "Friction Characteristics of Vapor Deposited Lubricant Films," SAE Technical Paper No. 890148.

Kawamura, H., 1990, "Study of Construction and Tribology in Heat Insulated Ceramic Engine," SAE Technical Paper No. 900624.

Klaus, E. E., and Lai, C. W., 1976, "Method of Die Casting Metals," U.S. Patent No. 3978908.

Klaus, E. E., and Lai, C. W., 1983, "Vapor Deposited Lubricants for Extreme Conditions," ASLE Special Publication 15, "Solid and Liquid Lubricants for Extreme Environments," 75–79.

Klaus, E. E., and Perez, J. M., 1983, "Comparative Evaluation of Several Hydraulic Fluids in Operational Equipment, a Full-Scale Pump Test Stand, and the Four-Ball Wear Tester," SAE Technical Paper No. 831680.

Klaus, E. E., 1988, "Vapor Delivery—A Tehnique Designed for High Temperature Lubrication," DOE/EC-88/3.

Klaus, E. E., Jeng, G. S., and Duda, J. L., 1989, "A Study of Tricresyl Phosphate as a Vapor Delivered Lubricant," *Lubrication Engineering*, Vol. 45, No. 11, pp. 717–723.

Klaus, E. E., Phillips, J., Lin, S. C., Wu, N. L., and Duda, J. L., 1990, "Structure of Films Formed During Deposition of Lubrication Molecules in Iron and Silicon Carbide," STLE *Tribology Transactions*, Vol. 33, No. 1, pp. 25–32.

Lai, C. W., 1974, "Lubrication Systems for Stainless Steel Die Casting," M. S. Thesis, The Pennsylvania State University.

Lauer, J. L., and Dwyer, S. R., 1989, "Continuous High Temperature Lubrication of Ceramic by Carbon Generated Catalytically," *Mat. Res., Soc. Symp. Proc.*, Vol. 140, pp. 363–368.

Lauer, J. L., and Dwyer, S. R., 1990, "High Temperature Lubrication by Carbon Continuously Replenished by Surface Reaction With Carbonaceous Gases—Comparison of Metallic (Ni) and Ceramic (Si_3N_4) Surfaces," SAE Technical Paper No. 900686.

Makki, J. F., and Graham, E. E., 1990, "Vapor Phase Deposition on High Temperature Surfaces," *Tribology Transactions*, Vol. 33, No. 4, pp. 595–603.

Makki, J., and Graham, E., 1991, "Formation of Solid Films From the Vapor Phase on High Temperature Surfaces," *Lubrication Engineering*, Vol. 47, No. 3, pp. 199–206.

Perez, J. M., Ku, C. S., and Hsu, S. M., 1991, "High Temperature Liquid Lubricant for Advanced Engines," SAE Technical Paper No. 910454.

Reddy, C. S., Domingo, N., and Graves, R. L., 1990, "Low Heat Rejection Engine Research Status: Where Do We Go From Here?" SAE Technical Paper No. 900620.

Sliney, H. E., 1987, "Coatings for Friction and Wear Control at High Temperatures," *Surface and Coatings Technology*, Vol. 33, pp. 243–244.

Weber, K., 1989, "Advanced Low Heat Rejection Diesel Technology Development," *Proceedings of the Annual Automotive Technology Development Contractors Coordination Meeting*, Dearborn, MI, SAE Publication P-230.

Weber, K., "Advanced Low Heat Rejection Diesel Technology Development," *Proceedings of the Annual Automotive Technology Development Contractors Coordination Meeting*, Dearborn, MI, SAE Publication P-243.

Electroplated Overlays for Crankshaft Bearings

D. R. Eastham
T&N Technology Ltd.,
Cawston, United Kingdom

Overlays of either lead–indium or lead–tin–copper are electroplated onto both lead–bronze and aluminum alloy crankshaft bearings to improve seizure resistance and conformability during the initial running-in period. In addition, both the corrosion resistance, particularly of lead–bronze, and the effective fatigue strength of the composite bearing are improved by this layer. The life of the overlay is largely dependent upon the diffusion rate of the low melting point species to the substrate. Thus, migration of either the indium or the tin will determine both the corrosion and wear rates of the overlay. Owing to the processing requirements, aluminum bearings require a nickel or copper interlayer prior to final overlaying with either of the lead alloys. For diffusion control reasons, when depositing lead–tin–copper onto lead–bronze it is usual to have a thin nickel dam to retard the formation of copper–tin intermetallics, which under given conditions may reduce the overall strength and adhesion; lead–indium does not require such a dam on lead–bronze. The principal differences between the two overlays lie in their respective fatigue and wear properties. Thus, lead–indium has a higher fatigue strength but lower wear resistance than lead–tin–copper. This paper compares these two major overlays and considers the selection criteria for the overlay employed.

Introduction

The tribological properties of engineering components may often be improved by the deposition of an overlay or a modification to the surface by a thermochemical treatment. Factors that may be improved by this means, and thereby increase the operational life of the component, include wear resistance, frictional characteristics, and corrosion resistance. The principal causes of wear are adhesion and abrasion, the former occurring from the welding and subsequent shearing of the surface asperities. The latter is covered by two situations known either as two-body abrasion or three-body abrasion; in both a soft surface is ploughed by a relatively harder surface. With two-body abrasion a rougher hard surface slides against a softer surface while with three-body rough hard particles trapped between two sliding surfaces cause either to be abraded. Typical of adhesive wear in the internal combustion engine is the microwelding of a piston ring to a piston or the scuffing of a connecting rod bearing on a crank pin.

In general the property of abrasion resistance is usually conferred upon a component by the use of a hard overlay such as chromium or a surface treatment, which increases the hardness. Resistance to adhesive wear is best achieved by improved lubrication, including low-friction coatings.

The use of a soft coating for engine shell bearings is a peculiar compromise, conferring as it does both low friction and some wear resistance. However, to achieve the latter it does require full hydrodynamic lubrication; problems invariably occur when this is not the situation.

Contributed by the International Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1992. Paper No. 93-ICE-26. Associate Technical Editor: J. A. Caton.

Overlay Design and Requirements

Bearing material selection for any given application is invariably a compromise of surface properties, mechanical strength, and corrosion resistance. The original metallic bearing materials such as tin-based white metal (Babbitt) had very good bearing and corrosion properties, but lacked the mechanical strength necessary for the modern, more highly loaded engine. It is difficult to achieve these requirements in a single material, although for more modestly loaded applications aluminum–tin–silicon alloys are proving very effective, as indicated by Massey et al. (1991).

The normal compromise is to use a high strength bearing lining, usually based on lead–bronze, onto which is electrodeposited a thin layer having good bearing properties; this layer is traditionally lead rich and contains alloying additions of materials such as tin, indium, and copper. These overlays have the functions of:

1 Preventing or reducing corrosive attack of the underlying lead–bronze bearing by oxidized or sulfur-containing lubricating oils. The lead can be leached out of the structure leaving a weaker, seizure- and fatigue-prone structure. This can be more of a problem with medium-speed diesel applications where high sulfur-containing residual fuels are often used.

2 Reducing the tendency toward solid phase welding of asperities by promoting sliding between surfaces without local surface melting. This resistance to what is known as scuffing, seizure, or galling depends upon three criteria:

(*i*) Compatibility or the inherent tendency of a material to resist solid-phase welding. This is of major concern if the oil film between the journal and the bearing is disrupted. It is also of concern with cold starting.

(*ii*) Conformability or the ability of the material to adjust

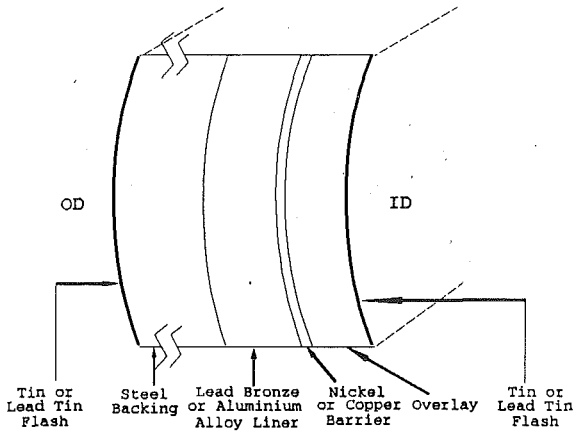


Fig. 1 Diagrammatic section of bearing

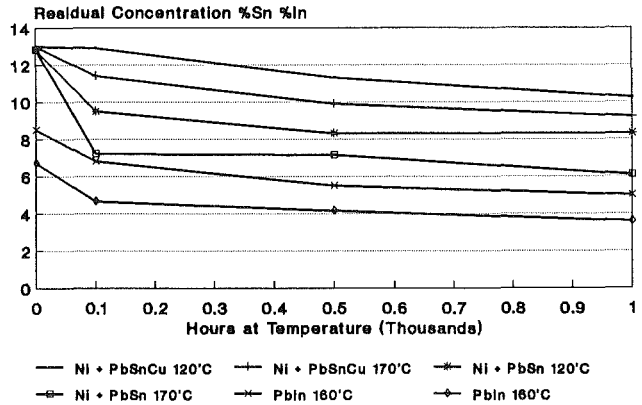


Fig. 2 Diffusion in lead-tin-copper and lead-indium overlays on lead bronze

to shaft deflection and minor misalignment by either wear or deformation. Some wear is inevitable, and although resistance to wear is important, it should not be so high as to prevent conforming.

(*iii*) Embeddability or the overlay's ability to cover small dirt particles with a thin soft film to prevent scoring damage.

Some damage to the overlay surface is certain to occur during the bearing life and a wear/score pattern will develop. It is generally intended that the overlay should last the life of the bearing, but there is a sacrificial element involved and the bearing will often continue to run adequately even though overlay damage has occurred.

The Manufacture of Soft Overlays

The manufacture of the modern bearing involves a complex system consisting of a strong steel backing lined with a lead-bronze or aluminum bearing layer overlaid with an electroplated deposit of a lead-based alloy. Either nickel or copper is often used as a bonding or diffusion barrier layer, and the whole product may be flashed with a thin coating of tin or lead-tin; the system is shown diagrammatically in Fig. 1.

In this paper we are concerned with the soft electroplated overlay. Historically, there are three compositions, PbSn10, PbSn10Cu2.5, and PbIn8, although there are variants on the lead-tin-copper theme with tin levels up to 18 percent and copper levels as high as six percent. Of these materials the standard lead-tin-copper and lead-indium are the most common.

These alloys are all produced by electrodeposition, but whereas the lead-tin and lead-tin-copper are produced by codeposition from a single solution containing all the relevant metallic species, the lead-indium is produced by separate deposition followed by a heat treatment to diffuse the elements together.

Most bearings normally require a layer of nickel of the order of 1–2 μm thickness. In the case of a lead-bronze bearing this nickel acts as a barrier to the diffusion of tin from the overlay and the subsequent formation of copper-tin compounds, which can occasionally result in the loss of adhesion. Diffusion still occurs but only nickel-tin intermetallics are formed. Perrin (1965) suggested that the manufacturing route for lead-bronze had a part to play in determining whether this adhesion loss occurred. A nickel interlayer is also used with aluminum alloy bearings, but for electrochemical manufacturing reasons rather than diffusion. Other interlayers such as copper and its alloys, e.g., brass and bronze, have also been used.

Overlay Diffusion

Bearing overlay systems currently designed for use on lead-bronze bearings, with or without a nickel interlayer, and having an overlay based on lead but containing tin or indium, are metallurgically unstable. Within an hour or so of running in an engine the overlay composition and the distribution of the low melting point phase will have changed and will continue to do so throughout the bearing life. This effect has been recognized for many years and has been reported by, among others, Semlitsch (1970), Eastham (1982), Pratt and Whitney (1983), Bierlein and De Hart (1983), and Subramanyan (1992).

Such diffusion has an effect upon the corrosion resistance, mechanical properties, and adhesion of the overlay. It is necessary, therefore, to design the overlay system in such a way that no long-term deleterious effects occur; this is particularly the situation with corrosion.

Diffusion Characteristics of Lead-Based Overlays. Typical engine bearing operating temperatures vary from 90°C for medium speed diesel to approaching 200°C for high-performance gasoline. The effect of extended heat treatments at temperatures between 120°C and 170°C for the lead-tin and lead-tin-copper systems has been shown in the previous referenced papers, but a comparison with lead-indium is shown in Fig. 2. Historically the lead-tin-copper system is always plated onto a nickel interlay, while lead-indium is not. Thus, the basic diffusion process produces a nickel-tin compound with the former metallurgy and a copper-indium compound with the latter.

It is clear that the diffusion rate of tin in lead-tin-copper is significantly less than in lead-tin when starting from similar tin levels of about 12.5–13 percent. Thus, at 170°C the tin level fell from 13 to 9.3 percent after 1000 hours in lead-tin-copper while in lead-tin it had decreased to 6.2 percent. Indium is usually deposited at levels between 7 and 10 percent. The results from Fig. 2 show reductions of indium from 9 to 5 percent and 7.5 to 3.5 percent after 1000 hours at 160°C.

The effect of the heat treatment on lead-tin-copper is demonstrated in Fig. 3 where an intermediate diffusion layer of nickel-copper-tin compound has been produced, which is believed to have the identity $(NiCu)_3Sn_4$ based on work by Kay (1976).

Indium diffuses into lead-bronze and forms the intermetallic Cu_4In_3. Figure 4 shows the effect of heat treatment for 500 hours at 160°C while Figs. 5 and 6 are the equivalent indium and lead energy dispersive x-ray analytical dot maps, respectively. The diffusion rate of indium without nickel barrier is very similar to that of tin from lead-tin-copper with nickel barrier (Sakamoto et al., 1989). The greatest loss is from the lead-tin overlay. The affinity of copper for tin slows down the movement of tin from a lead-tin-copper layer; on very long heat treatments examples of copper-tin intermetallics have

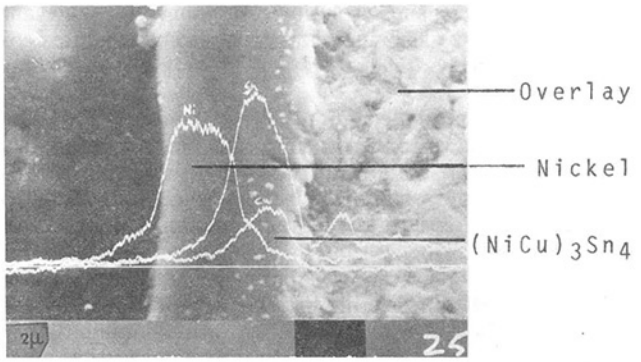


Fig. 3 Heat treatment of PbSn10Cu2, 1000 hours at 120°C

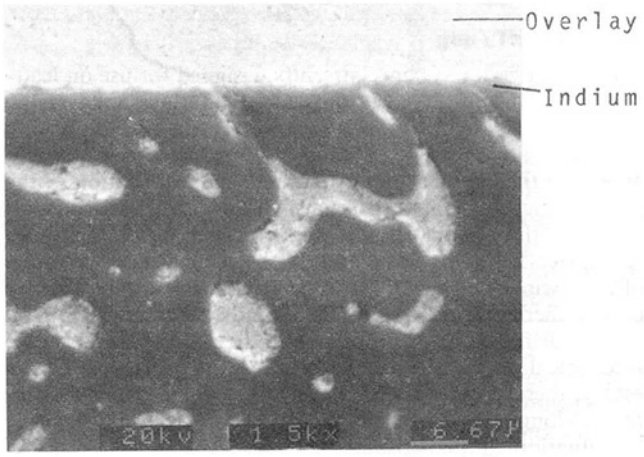


Fig. 4 Heat treatment of PbIn, 500 hours at 160°C

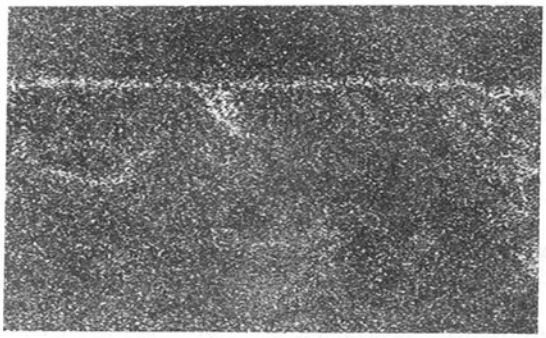


Fig. 5 Heat treatment of PbIn, indium analysis

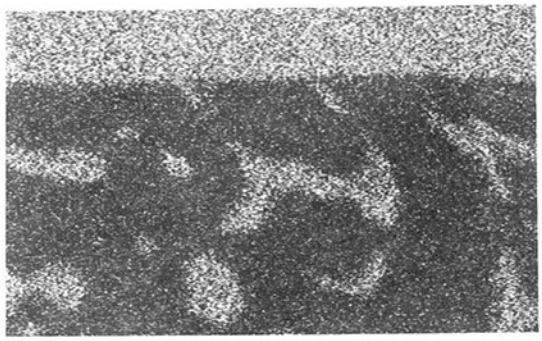


Fig. 6 Heat treatment of PbIn, lead analysis

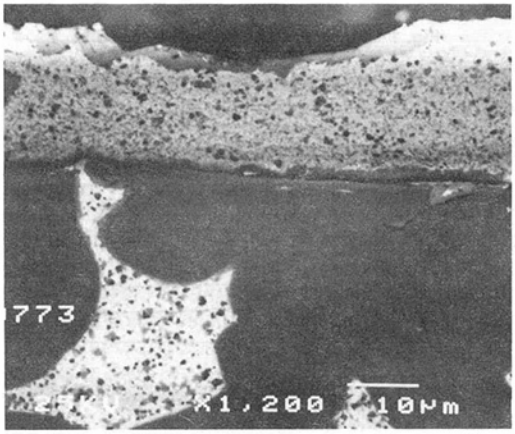


Fig. 7 Overlay bond deterioration on a non-nickel barriered lead-bronze bearing

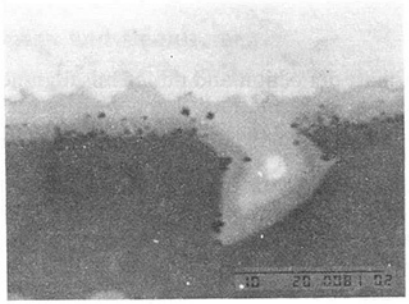


Fig. 8 Diffusion porosity, 4500 hours, 100°C

been found in the overlay, which may have an effect on both the overlay's wear resistance and corrosion performance.

The nickel interlayer is often referred to as a diffusion barrier, preventing loss of tin from the overlay. The results show that this is not the situation. Indeed, iron would provide a much superior barrier to diffusion as shown by Kay (1979), although the lack of a good-quality production electrolyte has prevented this from becoming a commercial reality.

Effect of Diffusion on Bond Adhesion. Without the nickel interlayer, copper–tin intermetallics will form at the interface between the overlay and the lead–bronze. Under certain circumstances this can result in bond deterioration over several thousand hours operation, more often in medium-speed diesel operations rather than automotive. When the tin diffuses from the overlay, it forms at least one intermetallic and often two, Cu_3Sn and Cu_6Sn_5. A black line has often been observed between the intermetallics and the lining, although its nature has never been conclusively determined, Fig. 7. The intermetallics form around the lead, causing it to extrude upward and exert pressure on the weakened interface, forcing the bond to give way.

Figure 8 shows the examination of a medium-speed diesel bearing, which had been in operation for 4500 hours at about 100°C. The two intermetallics can be clearly observed, together with porosity on the overlay side of the interface, i.e., on the side of the interface from which there has been a net loss of atoms; presumably the tin diffuses more quickly from the overlay than the copper does from the substrate. This is an example of the Kirkendall effect and confirms the vacancy mechanism of diffusion whereby a lattice defect may interchange places more frequently with one species of atom than the other.

A similar effect is demonstrated in Fig. 9, which is of a bearing heat treated at 160°C for 2000 hours. Here, the porosity is beginning to join together and resemble the so-called black line, which in the instance would result in a void and ultimately bond deterioration.

Fig. 9 Diffusion extended porosity, 2000 hours at 160°C

Although this may explain what happens, it does not explain why some materials are more prone to the problem than others. There is circumstantial evidence to suggest that it is connected to the size of lead particle in the lining, with the finer being more susceptible. However, as yet insufficient evidence has been gathered to confirm this unequivocally.

The overall effect is that a nickel barrier is an essential part of the structure for a cast lead-bronze bearing having a lead-tin-copper overlay. The nickel is effective in preventing copper-tin intermetallic formation. It is recognized that a lead-indium overlay does not need such an interlayer even though the intermetallic forms around the lead, and there are no known examples of such bond deterioration with this system.

Overlay Corrosion

Wilson and Shone (1970) showed that the corrosion resistance of a bearing overlay was heavily dependent upon the presence of the alloying element and that minimum levels of 3 percent tin and 5 percent indium were essential if corrosion was not to become a serious problem. Thus, the diffusion characteristics discussed previously play an important part in ensuring adequate corrosion resistance throughout the life of the engine.

A comparison of the corrosion resistance of a number of different bearing overlays was carried out in medicinal white oil (MWO) at 120°C with aeration being produced by stirring, Fig. 10. MWO was used as it is a pure material, having no inhibitors, and degrades rapidly, producing an aggressive medium.

The results show that the PbIn5 gave extremely good performance, being marginally better than PbSnCu with a nickel interlayer. However, PbIn2.5 gave poor results, being worse than PbSnCu without nickel. There is some indication that the lead-tin-copper with a nickel barrier is not so critically dependent upon tin level with one formulation containing 5.5 percent tin behaving very similarly to one having 10 percent tin. It is essential though that the lead-tin-copper based systems have a nickel interlayer if the best resistance is to be achieved.

Both indium and tin diffuse into the substrate through the lead pools on the lining surface when no nickel barrier is present. The indium in these lead areas can offer secondary corrosion protection if and when the overlay is worn through. The presence of the nickel barrier with lead-tin-copper prevents diffusion but the superior wear resistance of this system ensures that the secondary corrosion resistance is not often a requirement.

Wilson and Shone (1970) postulated that the corrosion protection was in part due to the presence of either tin or indium oxide at the alloy surface. Diffusion to the surface of either of these two elements is critical in ensuring a reduced corrosion rate. Alloys with less than about 3 percent tin or 5 percent indium appear to be unable to sustain a continuous oxide film on the surface.

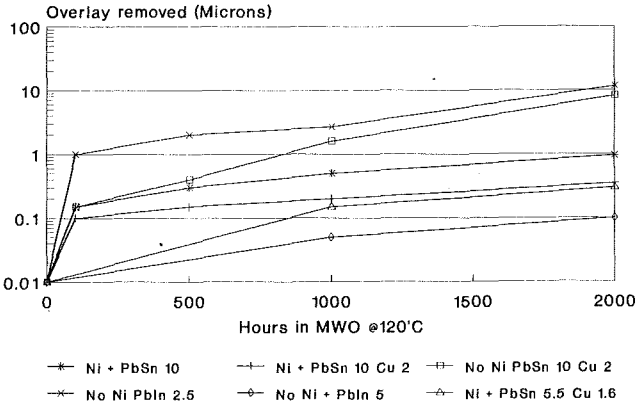


Fig. 10 Corrosion of lead-tin-copper and lead-indium

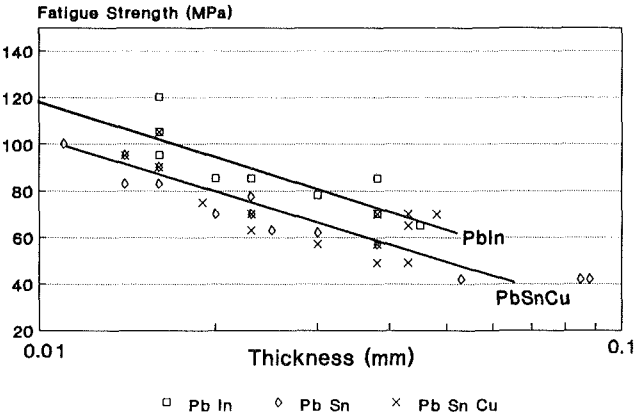


Fig. 11 Relationship between fatigue strength and overlay thickness

In general corrosion is not a major source of problem with automotive applications where distillate fuels are used. The major problems have been in medium-speed diesel operation where residual fuels and degraded oils with relatively high sulfur contents are used, and where expected engine life is considerably longer.

Overlay Hardness

The functions of embeddability, conformability, and wear resistance are dependent to some degree on hardness. The former two require a low hardness, while the latter is perhaps enhanced by a harder condition. The as-plated hardnesses of lead-tin-copper and lead-indium are 13-15 HV and 8-10 Hv, respectively (5 g load).

Operations at temperature with the subsequent diffusion of tin or indium affects the hardness. Thus after 500 hours at 160°C a PbSn10Cu2 overlay fell to about 9-11 Hv and a PbIn8 overlay to 6-7 Hv.

Mechanical Properties

Fatigue Strength. The fatigue strength of these overlays has been assessed using the Sapphire machine, which was described by Wilson (1967-68). In this test a bearing is housed in a connecting rod and run against an eccentric ground on the test shaft. The load is applied to the rod and hence to the bearing by an hydraulically loaded piston. After testing for 20 hours (3×10^6 cycles) at a given load the bearings are removed and examined for fatigue damage. If no fatigue has occurred the assembly is rebuilt, the load increased, and the test continued for a further 20 hours. The fatigue rating is the load at which damage is first observed.

Figure 11 shows the relationship between fatigue strength

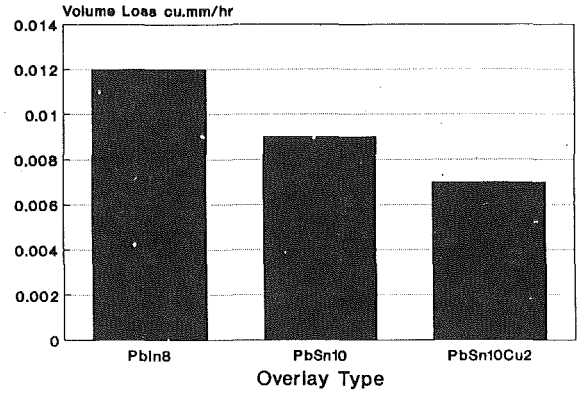


Fig. 12 Overlay wear

and the overlay thickness for both types of overlay. It can be seen that the lead–indium is about 15–20 percent stronger than the lead–tin–copper at the same thickness. Lead–tin and lead–tin–copper have similar fatigue strengths, although the ternary alloy has a longer life at a specific load (Yamada et al., 1973).

The superiority of lead–indium is believed to occur because of rapid recrystallization even at room temperature. In this way incipient fatigue defects are relieved or blunted and do not mature into structure-damaging cracks; the lead–tin–copper overlay does not behave in this way.

The damage mechanism found with lead–indium is often associated with a form of erosion and occurs over a long period of time. Indeed it is probably associated with internal oxidation of the grain boundaries (Rapp, 1965) leading to a loss in strength if the critical level of indium in the overlay is not maintained.

The higher tin-containing materials do not suffer from this erosion phenomenon (James, 1981).

Wear Resistance. Overlay wear resistance has been measured on the Sapphire machine at relatively low specific loads. The test consists of a running-in period of 20 hours at 48 MPa plus 20 hours at 55 MPa followed by five cycles each of 20 hours at 62 MPa. Average overlay wear is assessed by weight loss, which is converted to volume loss assuming the theoretical density of the material concerned. Comparative wear tests are shown in Fig. 12, which gives results after the running-in period for the lead–tin–copper and lead–indium overlays. The tin-based material is seen to have a significantly, superior wear resistance, 0.007 mm^3/h volume loss compared to 0.012 mm^3/h for lead–indium.

The wear resistance of the lead–indium has always been seen to be the weak spot of an otherwise high performance overlay. Recent work has shown that the addition of a third element such as tin can markedly increase the wear resistance while not reducing the fatigue strength (Massey et al., 1991).

Discussion

An overlay can offer the advantages of improved surface properties and improved strength characteristics. However, the quality improvements must be offset against the cost of application, and it can be seen from the foregoing that control of composition and thickness is an essential if the optimum compromise of fatigue, wear resistance, corrosion resistance, and anti-seizure characteristics is to be obtained.

The cost of an automotive bearing can be increased by anything from 10 to 30 percent, and this must be considered carefully when specifying the plate required. Thus, for performance application where a high fatigue strength is required, then the obvious choice should be for lead–indium, and indeed it is the preferred material for Formula 1 and Indianapolis racing. However, wear resistance is the advantage of the lead–tin–copper overlay and may be the preferred choice for applications where that property is of primary importance.

The major controlling factor for all these properties is diffusion of the low melting point species, tin or indium, and in order to achieve the correct level of corrosion resistance the minimum value of tin and indium must be present initially.

The lead–indium overlay benefits from not requiring a nickel interlayer. There is some evidence that nickel has relatively poor seizure characteristics, although the nickel–tin intermetallic formed by diffusion offers both good compatibility and corrosion resistance.

It should be remembered that the engine is the ultimate arbiter and the choice of overlay will depend upon the engine's characteristics and the particular compromise it requires. Overlays are an essential for the high-strength lead–bronze and aluminum–silicon based materials in order to offer improved conformability.

Electrodeposition of overlays is a recognized step in the manufacture of high strength plain bearings; advances in the future will be aimed at giving improvements in resistance to both corrosion and wear. Such changes may require the development of composite materials and the use of physical vapor-deposition processes.

References

Bierlein, J. C., and DeHart, A. O., 1983, SAE Technical Paper No. 830064.

Eastham, D. R., and Crooks, C. S., 1982, *Trans. Inst. Metal Finish*, Vol. 60, pp. 9–13.

James, R. D., 1981, PhD Thesis, Coventry (Lanchester) Polytechnic.

Kay, P. J., and MacKay, C. A., 1976, *Trans. Inst. Metal Finish*, Vol. 54, p. 68.

Kay, P. J., and MacKay, C. A., 1979, *Trans. Inst. Metal Finish*, Vol. 57, p. 169.

Massey, I. D., MacQuarrie, N. A., Coston, N. F., and Eastham, D. R., 1991, *Leeds-Lyon Symposium*, Tribology Series 18, Elsevier, Amsterdam.

Perrin, H., 1965, "Bearing Problems in Internal Combustion Four Stroke Rail Traction Engines," CIMAC 65.

Pratt, G. C., and Whitney, W. J., 1983, SAE Paper No. 830063.

Rapp, R. A., 1965, *Corrosion*, Vol. 21, Dec.

Sakamoto, M., Wada, M., and Ogita, Y., 1989, SAE Paper No. 890555, pp. 948–956.

Semlitsch, M., 1970, *Mikrochimica Acta* (*Wein*), Suppl. IV, pp. 157–169.

Subramanyan, P. K., 1992, ASME Paper No. 92-ICE-8.

Wilson, R. W., and Shone, E. B., 1970, *Anti-Corrosion*, Aug., pp. 9–14.

Wilson, W. H., 1967–68, *Proc. Inst. Mech. Eng.*, Vol. 182, No. 3G, p. 129.

Yamada, K., Tanaka, K., Mori, S., and Morita, S., 1973, presented at the 10th International Congress on Combustion Engines, Washington, DC.

Prediction Technique for Stress and Vibration of Nonlinear Supported, Rotating Crankshafts

H.-H. Priebsch

J. Affenzeller

AVL List GesmbH,
Graz, Austria

S. Gran
Graz, Austria

Design analyses of crankshafts, including bearings, are necessary for both the layout of new engines and the modification of existing engines (increased power output, etc.). To improve the existing calculation systems for crankshafts and bearings, AVL has developed a new method. This method enables the coupled vibrations in the torsional, bending, and axial directions, including gyroscopic effects, to be analyzed. For simulation of multibearing effects, the bearing models consider both the hydrodynamic oil film and the stiffness of the bearing structure. The calculation of forced vibrations is carried out using the gas and mass forces acting upon the rotating crankshaft. Comparisons of calculated to measured results demonstrate the accuracy of this calculation technique. The method can be used for passenger car, truck, and medium speed engines. In this paper examples of truck and passenger car engine applications confirm the additional possibilities for the estimation of crankshaft dynamics. Also the improvement of the results obtained from the new technique compared with those from classical calculation methods is described.

1 Introduction

The success of a new engine design is dependent on the durability, which is mainly influenced by the behavior of the crankshaft and the bearings. Therefore, great effort in simulation works is required for analyses of engine construction during the design stage. This is valid for both the layout of new engines and the improvement of existing engines. The high level of detailed refinement of engine design requires sophisticated tools for calculation and simulation systems. This is particularly true of a crankshaft and its bearings, for which a number of calculation methods exist. Nevertheless, for modern engine design, optimization problems occur due to some assumptions used in traditional calculation methods conflicting with current design demands. Without doubt, as improvement of engines using simulation and prediction techniques becomes more effective, the closer the calculation methods will be to the real physical events.

For analyses of bending stresses of crankshafts, the common method assumes statically determined supports [1]. Within this method gas and mass forces due to adjacent cranks only act on each bearing and crankshaft journal, respectively. The torsional vibrations are often calculated using a crankshaft model reduced to a plain shaft vibrating under moments due to gas and mass forces [2]. These methods include a decoupling of torsional and bending vibrations. The main disadvantage of this assumption is that a model created in this manner also neglects the coupling of bending and longitudinal vibrations. Most of these calculation methods use the influence of the engine structure known from experience, for the primary approximation.

Other common calculation methods for crankshafts use mean bearing stiffnesses [3–5]. This assumption is acceptable in the first approximation for torsional vibrations only, but not for bending vibrations. Also effects of the oil-film nonlinearity lead to stiffness increases by a factor of 1000, and are usually neglected for a multibearing crankshaft.

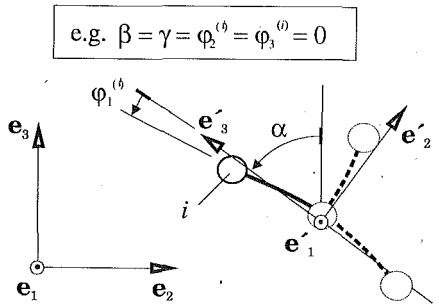
Bending modes and frequencies may be greatly influenced by the stiffnesses of the bearing structure and the main bearing oil film. The frequencies of the bending modes may be changed by 30 percent or more when the actual stiffness of the bearings are considered.

For this reason AVL has developed the calculation system "KWDYN" to be able to predict crankshaft dynamics considering a multibearing rotating crankshaft. For these calculations a matrix method is used modeling the different parts of the crankshaft, the bearings, and the influence of the stiffness and damping of the engine structure. Matrix modeling methods include both the global motion and deformation. This fact was found to be an advantage compared with Finite Element Methods (FEM) [6]. With the use of FEM calculations for the global motion and elastic deformation must be solved separately. However, matrix methods enable the equations to be solved simultaneously.

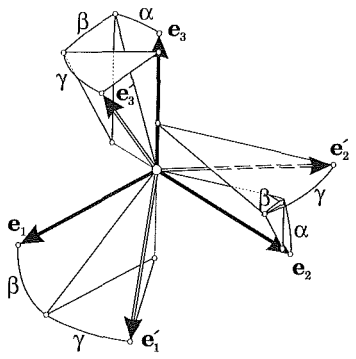
2 Calculation Procedure

To calculate the coupled bending and torsional and longitudinal vibrations of crankshafts with regard to their nonlinear bearings and rotation, an efficient procedure has been devel-

Contributed by the Internal Combustion Engine Division and presented at the 14th Annual Fall Technical Conference, Waterloo, Iowa, October 4–7, 1992. Manuscript received by the Internal Combustion Engine Division July 1, 1992. Associate Technical Editor: J. A. Caton.



a



b

Fig. 1 Definition of the angles: $\varphi_1^{(i)}$, $\varphi_2^{(i)}$, $\varphi_3^{(i)}$ and α, β, γ; transformation matrix R; definition of the structure's orientation

oped. This procedure allows determination of the motions of chainlike rotating structures built from massless beam and rigid single masses under nonlinear bearing conditions.

To achieve this, the classical transfer matrix method [16] was modified. This matrix method offers the advantage of calculating in one run the small vibrational motion of the elastic structure as well as the gross rigid body motion. Furthermore, we assume that while the chainlike structure may carry out large translations and rotations, its elastic distortion will be small and can be described by means of the linear theory of elasticity.

The modified transfer matrix method will need only two different elements: a beam element and a mass element.

2.1 Mass Element. In a general form, the equations of motion of a rigid body (element number i) freed from the chainlike structure will read as follows in the inertial system:

$$m^{(i)}\ddot{\mathbf{x}}^{(i)} = \mathbf{Q}^{(i)} - \mathbf{Q}^{(i-1)} + \mathbf{Q}_a^{(i)}$$
$$\mathbf{I}^{(i)} \bullet \dot{\boldsymbol{\omega}}^{(i)} + \dot{\mathbf{I}}^{(i)} \bullet \boldsymbol{\omega}^{(i)} = \mathbf{M}^{(i)} - \mathbf{M}^{(i-1)} + \mathbf{M}_a^{(i)} \quad (1)$$

We assumed that at the positive boundary (i) the internal forces and moments will be written positively in the positive direction of the coordinates and at a negative boundary ($i-1$) pointing to the negative direction of coordinates. The boundaries of the element are said to be (i) and ($i-1$).

The entire elastic structure's orientation is determined by the three angles, α, β, and γ (Fig. 1(b)). The actual position of the freed rigid body will then be determined by the small angles $\varphi_1^{(i)}$, $\varphi_2^{(i)}$, and $\varphi_3^{(i)}$ (Fig. 1(a)), which will be treated as infinitesimal angles. The structure's orientation needs to be determined by approximation only. There is only one restriction, namely, that Eq. (2) always must be fulfilled for a structure consisting of n elements. Up to this limit the angles $\varphi_j^{(i)}$ can be considered as infinitesimal angles with sufficient accuracy:

$$\varphi_j^{(i)} \leq \varphi_{\max} \approx 0.1 \hat{\approx} 6 \text{ deg} \quad j = 1, 2, 3 \quad i = 1, 2, \ldots, n \quad (2)$$

Thus, the freed body's angular velocity ω can be divided into the reference angular velocity Ω and in a small share $\dot{\varphi}$, which is generated by the distortion of this body with reference to the entire elastic structure's orientation defined by α, β, and γ:

$$\omega = \Omega + \dot{\varphi} \quad (3)$$

The reference angular velocity Ω, which is valid for the total structure, need not be determined by integration, but can be found by differentiation. The angles α, β, and γ refer to the inertial system and define a transformation matrix $\mathbf{R}$, which transforms the coordinates of any tensors from the initial position into the actual orientation in the inertial system. The reference angular velocity Ω may be given explicitly as the transformation matrix $\mathbf{R}$ will always be available in the formula.

$$\begin{bmatrix} 0 & -\Omega_3 & \Omega_2 \\ \Omega_3 & 0 & -\Omega_1 \\ -\Omega_2 & \Omega_1 & 0 \end{bmatrix} = \dot{\mathbf{R}} \bullet \mathbf{R}^T \quad (4a)$$

By specially selecting α, β, and γ (Fig. 1(b)) it may be shown that the simple relation (4b) will apply:

$$\Omega = \begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{bmatrix} = \begin{bmatrix} \dot{\alpha}\cos\beta\cos\gamma - \dot{\beta}\sin\gamma \\ \dot{\alpha}\cos\beta\sin\gamma + \dot{\beta}\cos\gamma \\ \dot{\gamma} - \dot{\alpha}\sin\beta \end{bmatrix} \quad (4b)$$

Then we define that the system of coordinates ($\mathbf{e}'_1$, $\mathbf{e}'_2$, $\mathbf{e}'_3$) belonging to the reference orientation of the entire structure will be generated by subsequent rotations of the angles α, β, and γ around the $\mathbf{e}_1$, $\mathbf{e}_2$ and $\mathbf{e}_3$ axis of the inertial system. The rotations are shown in Fig. 1.

The special transformation matrix $\mathbf{R}$ belonging to it will read as follows:

Nomenclature

$\frac{d()}{dt} \equiv \dot{()}$

$\frac{d^2()}{dt^2} \equiv \ddot{()}$

m = mass of an element
t = time
$\mathbf{e}_i$ = (1×3) unit vectors of the inertial coordinate system
$\mathbf{e}'_i$ = (1×3) unit vectors—definition of the structure's orientation
$\mathbf{x}$ = (1×3) displacement vector (transverse)
$\mathbf{z}$ = (1×13) state vector
$\mathbf{E}$ = (6×6) or (3×3) unit matrix
$\mathbf{M}$ = (1×3) vector of internal moments
$\mathbf{M}_a$ = (1×3) vector of external moments
$\mathbf{Q}$ = (1×3) vector of internal forces
$\mathbf{Q}_a$ = (1×3) vector of external forces
$\mathbf{R}$ = (3×3) transformation matrix
$\mathbf{U}$ = (13×13) transfer matrix
φ = (1×3) vector of angular displacement
ω = (1×3) vector of angular velocity of the mass element
Ω = (1×3) vector of reference angular velocity (structure)

$$\mathbf{R}=\begin{bmatrix} \cos\beta\cos\gamma & -\cos\alpha\sin\gamma+\sin\alpha\sin\beta\cos\gamma & \sin\alpha\sin\gamma+\cos\alpha\sin\beta\cos\gamma \\ \cos\beta\sin\gamma & \cos\alpha\cos\gamma+\sin\alpha\sin\beta\sin\gamma & -\sin\alpha\cos\gamma+\cos\alpha\sin\beta\sin\gamma \\ -\sin\beta & \sin\alpha\cos\beta & \cos\alpha\cos\beta \end{bmatrix} \quad (5)$$

The inertia tensor of the freed body in the entire elastic structure's orientation defined by $\mathbf{R}$ can be written as follows:

$$\underline{\mathbf{I}}^{(i)}=\mathbf{R}\bullet\mathbf{I}_0^{(i)}\bullet\mathbf{R}^T. \quad (6)$$

$\mathbf{I}_0^{(i)}$ designates the body's inertia tensor belonging to $\alpha=\beta=\gamma=0$. The actual inertia tensor measured in the inertial system can be determined with Eq. (7) in an extremely good, nonlinear approximation:

$$\mathbf{I}^{(i)}=\underline{\mathbf{I}}^{(i)}+\mathbf{A}_\varphi^{(i)}\bullet\underline{\mathbf{I}}^{(i)}-\underline{\mathbf{I}}^{(i)}\bullet\mathbf{A}_\varphi^{(i)};\quad \mathbf{A}_\varphi^{(i)}=\begin{bmatrix} 0 & -\varphi_3 & \varphi_2 \\ \varphi_3 & 0 & -\varphi_1 \\ -\varphi_2 & \varphi_1 & 0 \end{bmatrix}^{(i)} \quad (7)$$

With this, the first derivation in time of the inertia tensor appearing in the dynamic Eq. (1) can be laid down as follows:

$$\dot{\mathbf{I}}^{(i)}=\mathbf{A}_\omega^{(i)}\bullet\mathbf{I}^{(i)}-\mathbf{I}^{(i)}\bullet\mathbf{A}_\omega^{(i)};$$

$$\mathbf{A}_\omega^{(i)}=\begin{bmatrix} 0 & -\Omega_3-\dot{\varphi}_3 & \Omega_2+\dot{\varphi}_2 \\ \Omega_3+\dot{\varphi}_3 & 0 & -\Omega_1-\dot{\varphi}_1 \\ -\Omega_2-\dot{\varphi}_2 & \Omega_1+\dot{\varphi}_1 & 0 \end{bmatrix}^{(i)} \quad (8)$$

In a general form, an implicit direct integration statement will be brought in the dynamic Eq. (1) to eliminate partly the derivations in time. The basic idea of these integration methods is to express by approximation the derivations in time of a function $f(t)$ at a time t through the function to the time itself and through values of this function dated back in time $f(t-\Delta t), f(t-2\Delta t), \ldots$ or through the derivations in time of this function at these times. All direct implicit methods of integration can be traced back to the following formal form:

$$\dot{f}(t)=\mu(\Delta t)\bullet f(t)-\tau_f;\quad \ddot{f}(t)=\xi(\Delta t)\bullet f(t)-\eta_f \quad (9)$$

The parameters μ and ξ only depend on the time step size Δt and the functions τ_f and η_f at the actual time t are always known. Appendix I gives the integration parameters μ and ξ as well as the functions τ_f and η_f for the most common methods.

If a vector, the so-called state vector $\mathbf{z}$, is defined according to Eq. (10),

$$\mathbf{z}=[\mathbf{v}\quad \mathbf{p}\quad 1]^T;\quad \begin{cases}\mathbf{v}=[x_1\quad x_2\quad x_3\quad \varphi_1\quad \varphi_2\quad \varphi_3] \\ \mathbf{p}=[Q_1\quad Q_2\quad Q_3\quad M_1\quad M_2\quad M_3]\end{cases} \quad (10)$$

Equation (1) can be written in the form of Eq. (11) by means of the integration statement with regard to $\mathbf{v}^{(i-1)}\equiv\mathbf{v}^{(i)}$ and Eqs. (3)–(8):

$$\mathbf{z}^{(i)}=\mathbf{U}^{(i)}\bullet\mathbf{z}^{(i-1)} \quad (11)$$

The 13th coordinate of the state vector ($=1$) is used to take into account external forces and components due to the integration method (see sec. 2.3). In Eq. (11) $\mathbf{U}^{(i)}$ represents the modified transfer matrix of the general mass element in the form:

$$\mathbf{U}^{(i)}=\begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{U}_{pv} & \mathbf{E} & \mathbf{u}_p \\ \mathbf{0} & \mathbf{0} & \mathbf{1} \end{bmatrix}^{(i)} \quad (12a)$$

The partial matrix $\mathbf{U}_{pv}^{(i)}$ and the inhomogeneous part of the set of equations $\mathbf{u}_p^{(i)}$ will read as follows:

$$\mathbf{U}_{pv}^{(i)}=\begin{bmatrix} m\xi\mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mu\dot{\mathbf{I}}+\xi\mathbf{I} \end{bmatrix}^{(i)} \quad (12b)$$

$$\mathbf{u}_p^{(i)}=\begin{bmatrix} \mathbf{Q}_a+m\eta_x \\ \mathbf{M}_a+\dot{\mathbf{I}}\bullet(\tau_\varphi-\Omega)+\mathbf{I}\bullet(\eta_\varphi-\dot{\Omega}) \end{bmatrix}^{(i)} \quad (12c)$$

In the transfer matrix, the magnitudes $\mathbf{I}$ and $\dot{\mathbf{I}}$ depending on φ and $\dot{\varphi}$ will appear. However, this will not create any new problems, because the structure is already nonlinear. The nonlinear hydrodynamic reaction forces of the bearing are replaced by external forces. They will be corrected iteratively until they have been predetermined accurately. Simultaneously, $\mathbf{I}$ and $\dot{\mathbf{I}}$ are transformed. Due to the fact that $\mathbf{I}$ and $\dot{\mathbf{I}}$ appear as factors of φ in the transfer matrix, an implicit iteration algorithm is created, which safely converges. If the structure carries out large rotations around only one axis (e.g., crankshaft and connecting rod), the mass element transfer matrices can be brought in a form, where the elements of matrix $\mathbf{U}_{pv}^{(i)}$ are constant during the entire calculation and this iteration is no longer necessary. Appendix I gives an example of such a transfer matrix (flywheel element).

2.2 Beam Element. By integration of the differential equation of the elastic line, the transfer matrices of any beam elements arbitrarily orientated in space can be determined. If $\mathbf{U}$ represents the beam transfer matrix in position $\alpha=\beta=\gamma=0$, the required beam transfer matrix in the inertial system will read as follows (the transfer matrix for a straight beam is shown in Appendix I):

$$\mathbf{U}=\mathbf{R}\bullet\underline{\mathbf{U}}\bullet\mathbf{R}^T \quad (13)$$

2.3 Solution of the Equations of Motion. For a chainlike structure from n elements (beam and mass elements) all state vectors at the element's boundaries 1 to $n-1$ can be eliminated by simple matrix multiplications. First, all transfer matrices have to be generated according to the actual and given orientation (α, β, and γ), which is permissible according to Eq. (2). For this, the vectors η_x, τ_φ, and η_φ appearing in the 13th column of the transfer matrices can be expressed by a corresponding integration statement (see Appendix I) by means of already calculated displacements or velocities. At the start of a calculation, these vectors are defined by the initial conditions:

$$\mathbf{z}^{(n)}=\mathbf{U}^{(ges)}\bullet\mathbf{z}^{(0)}$$

$$\mathbf{U}^{(ges)}=\mathbf{U}^{(n)}\bullet\mathbf{U}^{(n-1)}\bullet\ldots\bullet\mathbf{U}^{(2)}\bullet\mathbf{U}^{(1)}=\prod_{i=n}^{1}\mathbf{U}^{(i)} \quad (14)$$

The (13×13) total transfer matrix $\mathbf{U}^{(ges)}$ represents the linear connection between the two ends of the structure. From the 24 unknowns, the unknown components of the state vectors $\mathbf{z}^{(0)}$ and $\mathbf{z}^{(n)}$, it is necessary to eliminate 12 magnitudes to solve the inhomogeneous (12×12) set of Eqs. (14) clearly. At a free end of the structure $\mathbf{p}\equiv\mathbf{0}$ will apply for the internal forces; at a fixed end $\mathbf{v}\equiv\mathbf{0}$ will apply. It is even possible to give a time curve of the boundary conditions. For further observation, we assume the two ends of the structure to be free which is the most frequently occurring case. With the partitioned total transfer matrix $\mathbf{U}^{(ges)}$

$$\mathbf{U}^{(ges)}=\begin{bmatrix} \mathbf{U}_{vv} & \mathbf{U}_{vp} & \mathbf{u}_v \\ \mathbf{U}_{pv} & \mathbf{U}_{pp} & \mathbf{u}_p \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}^{(ges)} \quad (15)$$

the inhomogeneous (6×6) set of equations for the determination of the state vector's $\mathbf{z}^{(0)}$ unknown components will read

$$\mathbf{U}_{pv}^{(ges)} \cdot \mathbf{v}^{(0)} = -\mathbf{u}_p^{(ges)} \tag{16}$$

Thus, the state vector is known and all displacements and internal forces within the structure can be determined through continuous multiplication by the transfer matrices. With the distortions of all elements, condition (2) is to be controlled and the reference position of the structure to be corrected, if necessary. This can be effected by averaging these distortions over all elements:

$$\bar{\varphi}_j = \frac{1}{n} \sum_{k=0}^{n} \varphi_j^{(k)} \quad j = 1, 2, 3 \tag{17}$$

The new orientation of the elastic structure then will be defined by the angles $\bar{\alpha}$, $\bar{\beta}$, and $\bar{\gamma}$ according to Eq. (18). In most cases, this adaptation will not be necessary within a time t as in practice the calculation is done with small computing time steps Δt and a sufficiently accurate orientation can be found for a new time $t + \Delta t$ by extrapolation.

$$\bar{\alpha} = \alpha + \bar{\varphi}_1, \quad \bar{\beta} = \beta + \bar{\varphi}_2, \quad \bar{\gamma} = \gamma + \bar{\varphi}_3 \tag{18}$$

2.4 Macro Elements. When applying the method of transfer matrices, two measures are to be taken to eliminate the known numerical problems with structures containing a high number of elements. On the one hand, the element transfer matrices are made dimensionless with a mean element length, a mean area second moment, and the elasticity module of the structure, and on the other hand so-called macro elements are introduced.

For this, the structure consisting of n elements will be divided up into q parts, the macro elements, with the permitted number of elements. From experience we know that 40 elements in dimensionless form can be processed without any numerical problems. The element boundaries of the macro elements have the numbers $k_j, j = 0, 1, \ldots, q$ with $k_0 = 0$ and $k_q = n$. The macro element transfer matrices $\mathbf{U}_M$ are determined by

$$\mathbf{U}_M^{(j)} = \prod_{i=k_j}^{k_{j-1}+1} \mathbf{U}^{(i)} \quad j = 1, 2, \ldots, q \tag{19}$$

These transfer matrices will be partitioned:

$$\mathbf{U}_M = \begin{bmatrix} \tilde{\mathbf{U}}_M & \tilde{\mathbf{u}}_M \\ \mathbf{0} & 1 \end{bmatrix}; \quad \mathbf{z}_M = \begin{bmatrix} \tilde{\mathbf{z}}_M \\ 1 \end{bmatrix} \tag{20}$$

With this, the following tapelike set of equations can be generated from which, after inclusion of the boundary conditions, the state vectors at the macro element boundaries can be determined with sufficient accuracy.

$$\begin{bmatrix} \tilde{\mathbf{U}}_M^{(1)} & -\mathbf{E} & & & & \\ & \tilde{\mathbf{U}}_M^{(2)} & -\mathbf{E} & & & \\ & & \cdot & \cdot & & \\ & & & \cdot & \cdot & \\ & & & & \tilde{\mathbf{U}}_M^{(q)} & -\mathbf{E} \end{bmatrix} \cdot \begin{bmatrix} \tilde{\mathbf{z}}_M^{(1)} \\ \tilde{\mathbf{z}}_M^{(2)} \\ \cdot \\ \cdot \\ \tilde{\mathbf{z}}_M^{(q)} \end{bmatrix} = \begin{bmatrix} \tilde{\mathbf{u}}_M^{(1)} \\ \tilde{\mathbf{u}}_M^{(2)} \\ \cdot \\ \cdot \\ \tilde{\mathbf{u}}_M^{(q)} \end{bmatrix} \tag{21}$$

This set of equations can be solved most efficiently due to its bandlike form. Within the macro elements the state vectors are determined by continuous multiplication by means of the found state vectors of the macro element boundaries.

2.5 Branching. The so far assumed chainlike structure can easily be extended to chainlike structures with branchings (e.g., counterweight at a crankshaft). For this, the branched part will be reduced to an individual transfer matrix of the main structure. For the (corresponding to Fig. 2) freed branched parts of the structure, the transfer matrix relation between the two ends of this structure will read as follows:

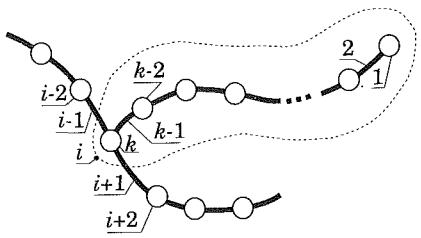


Fig. 2 Branched structure

$$\begin{bmatrix} \bar{\mathbf{v}}^{(k)} \\ \bar{\mathbf{p}}^{(k)} \\ 1 \end{bmatrix} = \begin{bmatrix} \bar{\mathbf{U}}_{vv} & \bar{\mathbf{U}}_{vp} & \bar{\mathbf{u}}_v \\ \bar{\mathbf{U}}_{pv} & \bar{\mathbf{U}}_{pp} & \bar{\mathbf{u}}_p \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}^{(ges)} \begin{bmatrix} \bar{\mathbf{v}}^{(0)} \\ \mathbf{0} \\ 1 \end{bmatrix}$$

$$\text{with } \bar{\mathbf{U}}^{(ges)} = \prod_{j=k}^{1} \bar{\mathbf{U}}^{(j)} \tag{22}$$

From this equation changes of the internal forces $\bar{\mathbf{p}}^{(k)}$ can be calculated due to the branched structural part

$$\bar{\mathbf{p}}^{(k)} = \bar{\mathbf{U}}_{pv}^{(ges)} \cdot \bar{\mathbf{U}}_{vv}^{(ges)^{-1}} \cdot \bar{\mathbf{v}}^{(k)} + \bar{\mathbf{u}}_p^{(ges)} - \bar{\mathbf{U}}_{pv}^{(ges)} \cdot \bar{\mathbf{U}}_{vv}^{(ges)^{-1}} \cdot \bar{\mathbf{u}}_v^{(ges)} \tag{23}$$

With this for the transfer matrix of the branched part reduced in the main structure with regard to $\mathbf{v}^{(i-1)} \equiv \mathbf{v}^{(i)}$ there will follow

$$\mathbf{U}^{(i)} = \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \bar{\mathbf{U}}_{pv}^{(ges)} \cdot \bar{\mathbf{U}}_{vv}^{(ges)^{-1}} & \mathbf{E} & \bar{\mathbf{u}}_p^{(ges)} - \bar{\mathbf{U}}_{pv}^{(ges)} \cdot \bar{\mathbf{U}}_{vv}^{(ges)^{-1}} \cdot \bar{\mathbf{u}}_v^{(ges)} \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix}^{(i)} \tag{24}$$

3 Description of Calculation Models

3.1 Crankshaft. The matrices are built up starting with a Three-dimensional-Finite Element model of the shaft, cranks, and webs (Fig. 5). Using this model to calculate the deformations under unit loads and moments, the elemental stiffness distribution along the crankshaft can be defined. Mass and inertia moments are reduced to the centers of gravity of selected nodal points of the crankshaft using CAD. A more detailed description of this step is given in [7, 8]. After the definition of this centerline model of the crankshaft, mass, damping, and stiffness values enable an elemental matrix model of the whole crankshaft to be generated. This model includes all dynamic properties of the crankshaft including gyroscopic effects, models for the flywheel, and the damping moments at the pulley.

To check the accuracy of this calculation model, a comparison of calculated and measured free-free modes was carried out. The calculated modes considered interaction between bending, longitudinal, and torsional vibrations. These comparisons also helped to refine the model. Typical mean absolute differences of measured and calculated natural frequencies of up to 2000 Hz were less than 4 percent.

Figure 4 shows selected modes of a crankshaft of a four-cylinder passenger car engine. The results are shown for first, second, and third-order bending modes of the crankshaft under free-free conditions. Differences in the natural frequencies are less than 4 percent (Fig. 4). Differences of the mode shapes found from the comparison of measured and calculated results were negligible.

To allow for calculation in the time domain with this crankshaft model, time integration terms for acceleration and velocity are automatically generated in the matrices. The time

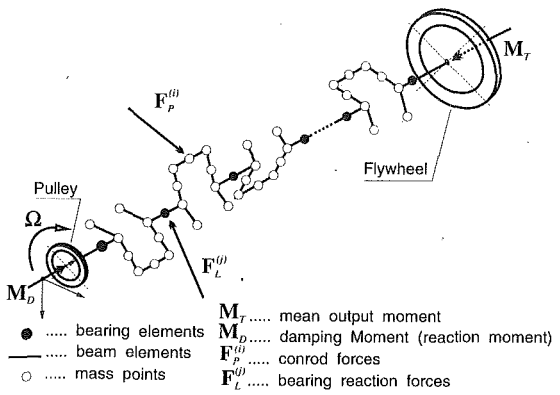
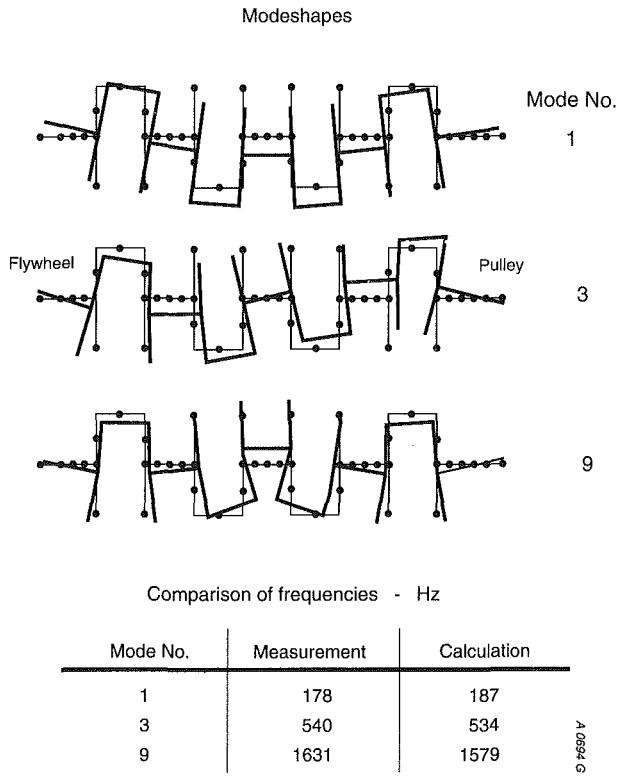


Fig. 3 Crankshaft model showing acting forces and moments for the transient multibearing calculation over 720 deg CA



Comparison of frequencies - Hz

Mode No.	Measurement	Calculation
1	178	187
3	540	534
9	1631	1579

A 0694 G

Fig. 4 Comparison of measured and calculated results of natural bending modes and frequencies for a four cylinder passenger car engine

history calculation is necessary to consider nonlinear effects in the bearing oilfilms. For the forced vibration calculation, all external forces and moments acting at the cranks, and the crankshaft front and rear ends, are applied (Fig. 3). At the clutch side of the flywheel mean effective torque is considered.

3.2 Bearings. For the calculations described in this paper, a hydrodynamic model was used for the main bearings. To check the accuracy of this bearing simulation technique, a comparison of two principle models for oil film and bearings was performed. The first model enables the elastohydrodynamic effects in the bearing oil film, taking into account the interaction of pressure distribution and bore deformation of the bearing, to be considered. The second model uses a hydrodynamic calculation method for the oil film, and for stiffness and damping properties of the bearing structure. The principle differences of these two stiffness models can be seen from Fig. 6.

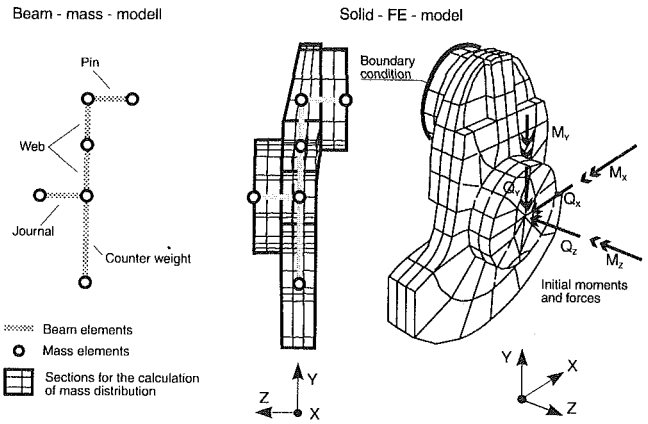


Fig. 5 Conversion of solid Finite Element model of one half crank throw to dynamic beam-mass model

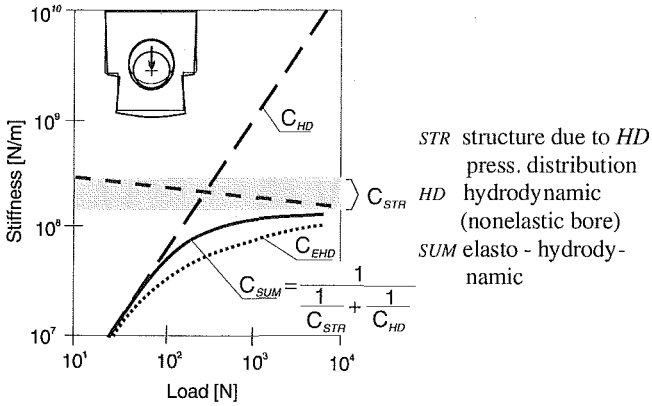


Fig. 6 Comparison of calculated stiffnesses of the main bearing structure and oil film, 2.3 l passenger car engine

In Fig. 6 the hydrodynamic stiffness C_{HD}, the stiffness of the bearing cap due to *HD*, pressure distribution C_{STR}, and stiffness resulting from interaction of the elasticity of the structure and hydrodynamic oil film C_{EHD} are plotted. These calculations were carried out on the main bearing of a passenger car engine. The details of the calculation of C_{EHD} are described in [8].

Furthermore, Fig. 6 shows that the resulting stiffness C_{SUM} of oil film and structure is close to the oil-film stiffness C_{HD} at low loads. At high loads the resulting stiffness approaches the stiffness of the bearing structure asymptotically.

At mean load forces and journal eccentricities respectively the actual stiffness of oil film and bearing C_{EHD} is less than C_{SUM}, because of the different pressure distribution of the oil film. For the calculation of C_{EHD} the interaction between the hydrodynamic oil film pressure and the deformation of the bearing construction was taken into account.

Both bearing models, the hydrodynamic and the elastohydrodynamic, simulate nonlinear properties of the oil film, solving the Reynolds equation and the extended Reynolds equation, respectively. For crankshaft stress and vibration analyses it could be found that hydrodynamic models are usually of sufficient accuracy. Therefore, in the following examples of the application of KWDYN, the hydrodynamic models are used together with structure stiffness of bearing construction.

Figure 7 explains the bearing model. This model consists of hydrodynamic behavior of the oil film due to journal motion and velocity in a rigid bore. Structure stiffness is applied using

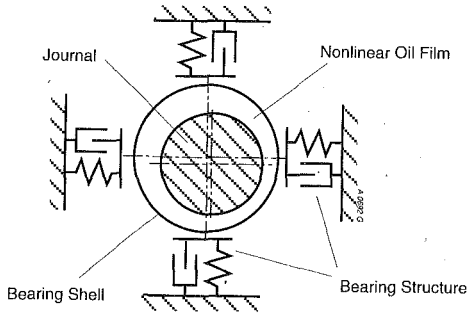


Fig. 7 Main bearing model for the calculation of transient multibearing crankshaft vibrations over 720 deg CA

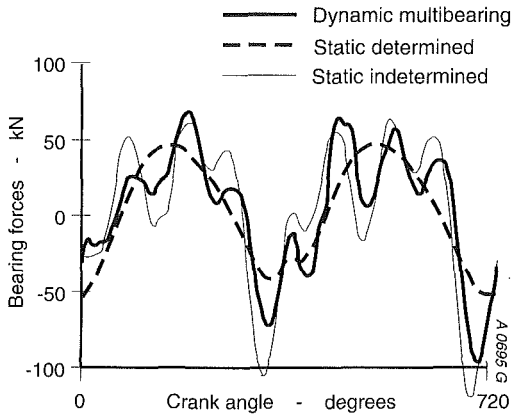


Fig. 8 Comparison of calculated main bearing forces using different calculation methods and models for a six cylinder truck diesel engine at 3000 rpm (overspeed)

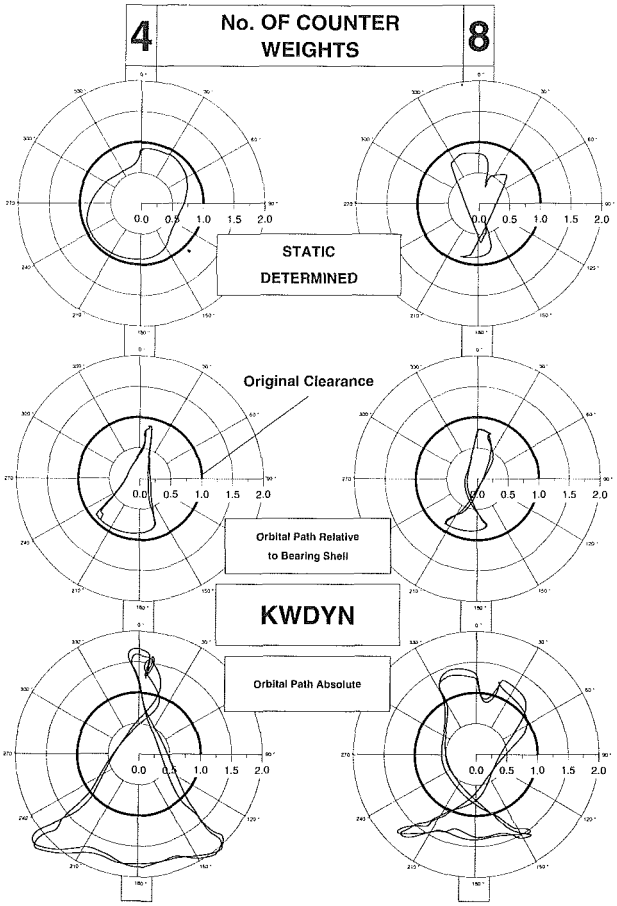


Fig. 9 Orbital path of crankshaft journal No. 3 calculated by different calculation methods for a four-cylinder 4.5 l inline DI diesel engine at 4000 rpm, overspeed

four pairs of springs and dampers. It is necessary to use different stiffnesses in different directions according to the structural behavior. In this model a torsional spring is also included to simulate the interaction of crankshaft bending vibration and vibration of the bearing cap in the crankshaft axis direction.

In connecting these bearing models to the crankshaft model described above, a system of nonlinear equations is created, representing the dynamic behavior of the vibration system and the interaction between rotating shaft and bearings (the basic analysis is found in [14]). This interaction between shaft vibration and oil film reaction force can only be solved iteratively. The vibration system has to be calculated in a time domain using a step-by-step method of implicit-explicit type and the iteration has to be done in each time step. In this iterative process bearing reaction forces due to displacement and velocities of the journals in the bearings are calculated. These reaction forces together with initial forces and moments lead to a new set of journal velocities and crankshaft deformations. Convergence is controlled by the comparison of reaction forces calculated in two sequential iteration steps.

4 Comparison of Calculation Methods

A comparative analysis was carried out to find what influence different calculation methods had on the vibration and stress results. Models of crankshaft, oil film, and bearings of a six cylinder truck diesel engine were used for these calculations as follows:

- (*a*) Statically determined model with hydrodynamic oil film.
- (*b*) Statically undetermined model with mean stiffness values for oilfilm and bearing structure.
- (*c*) Dynamic multibearing (AVL's program KWDYN) including hydrodynamic oil film and structure stiffnesses.

Figure 8 shows the calculation results for the bearing reaction forces of this engine's No. 7 main bearing under loads at 3000 rpm (overspeed). The statically determined method clearly leads to forces dominated by first-order mass forces. The maximum amplitude of the bearing forces is about 65 percent of those using dynamic multibearing method. The increase of maximum forces is dominated by dynamic effects, which are shown mainly in natural mode vibration of the crankshaft in this example.

The statically undetermined method also leads to these dynamic effects, but with increased force amplitudes compared with the results of KWDYN. Here one must recognize, that all calculation systems using the hydrodynamic models of the bearing oil film assume actual damping effects (version (*a*) and (*c*) in this comparison). Besides this effect a change of frequency can be seen for the statically undetermined system compared to results of KWDYN. This effect is caused by the different stiffness of the bearings, which is linear in the case of statically undetermined system and nonlinear in program KWDYN. This fact also leads to different natural modeshapes and frequencies in both calculation methods.

Figure 9 shows calculation results for the journal orbital path in the bearing oil film obtained by different mathematical models. The comparative analysis was carried out for a four-cylinder 4.5 1 inline DI diesel engine at 4000 rpm (overspeed) for four and eight counterweights. The hydrodynamic calculations were done for bearing No. 3 with both a statically determined and a dynamic multibearing system. Having the same balance ratio for the whole crankshaft, the orbital path

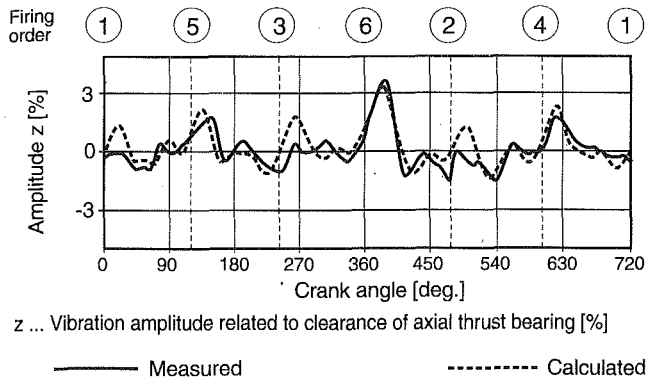


Fig. 10 Comparison of measured and calculated results of flywheel vibration in the crankshaft axis direction for a six cylinder truck diesel engine at 1300 rpm and full load

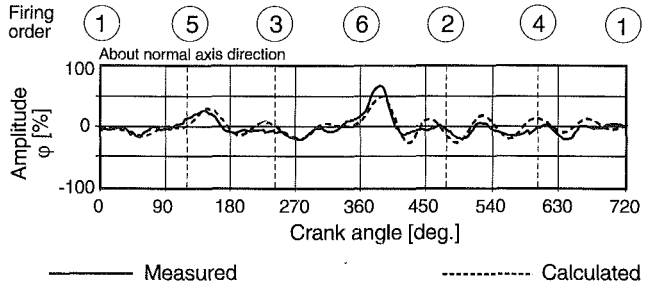


Fig. 11 Comparison of measured and calculated results of wobble motion of the flywheel for a six cylinder truck diesel engine at 1300 rpm and full load

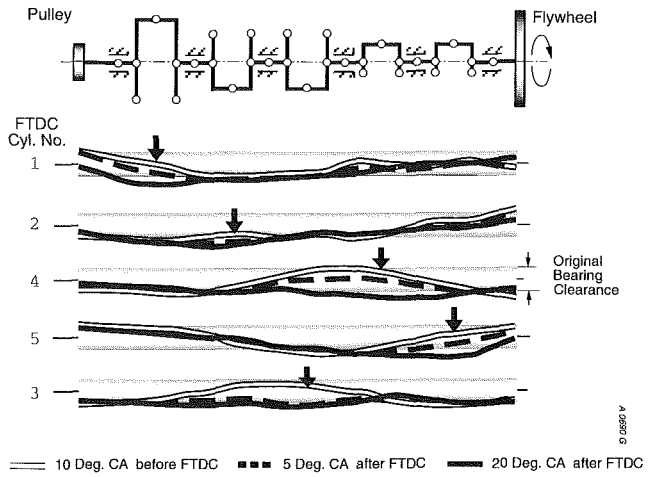


Fig. 12 Calculated bending vibration for the whole crankshaft over an engine cycle for 2.3 l passenger car engine at 5000 rpm and full load

of the center bearing shows clearly effects of locally unbalanced rotating mass forces in this bearing compared to the static determined model. The absolute motion of the journal is of course very different due to the elasticity of the bearing structure.

5 Measurement and Calculation Results at the Flywheel

An important result of this calculation technique is the analysis of the flywheel motion, often discussed in the literature [10–13, 15].

To determine the accuracy of the calculation results in this

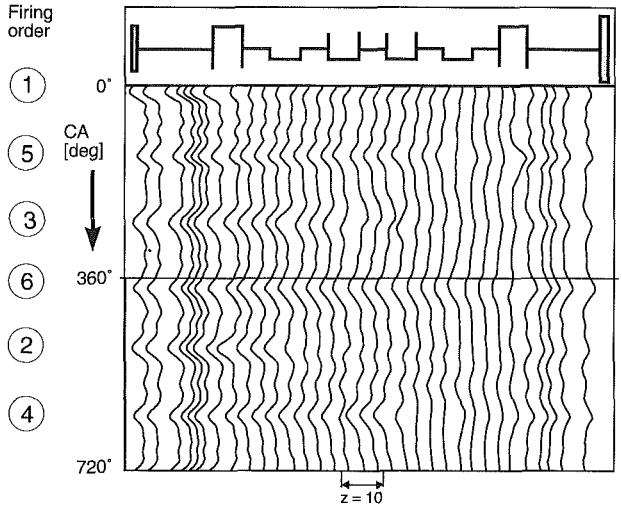


Fig. 13 Calculated axial vibration over crankshaft length for a six cylinder truck diesel engine at 1300 rpm and full load

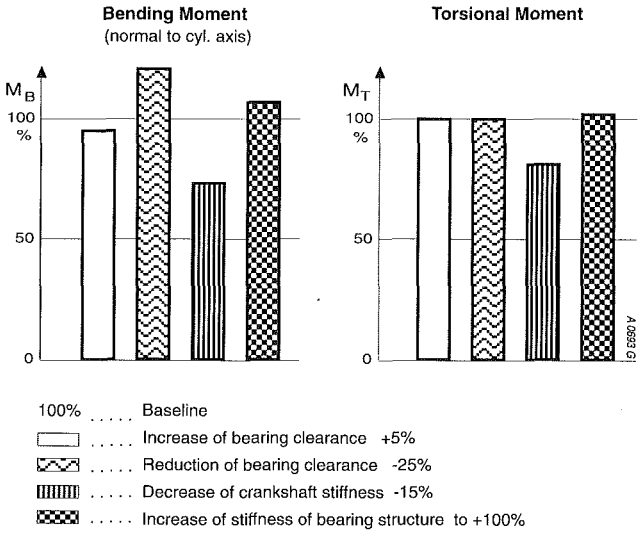


Fig. 14 Influence of principle design modifications on calculated stress moments of journal No. 7 for a six cylinder truck diesel engine at 3000 rpm (overspeed)

view, comparison with measured results obtained on a fired engine were made. Figures 10 and 11 show typical results of these comparisons for a six cylinder truck diesel engine at 1300 rpm and full load.

The measurement of the radial, axial, and wobble vibrations of the flywheel were obtained from inductive displacement pickups. These pickups were mounted on the flywheel housing.

The comparisons show good agreement especially for the main vibration events at FTDC of cylinders No. 4, 5, and 6 situated next to the flywheel. These vibrations are also influenced by the flywheel position, and the gear train width at the rear side of the engine, respectively.

Increased distances of the flywheel to the engine due to gear train width showed increased radial vibrations of the flywheel, particularly at overspeed. The wobbling motions were slightly reduced at the same engine condition.

Thus through simulating different flywheel positions with this method, their influence on durability and acoustical problems can be analyzed.

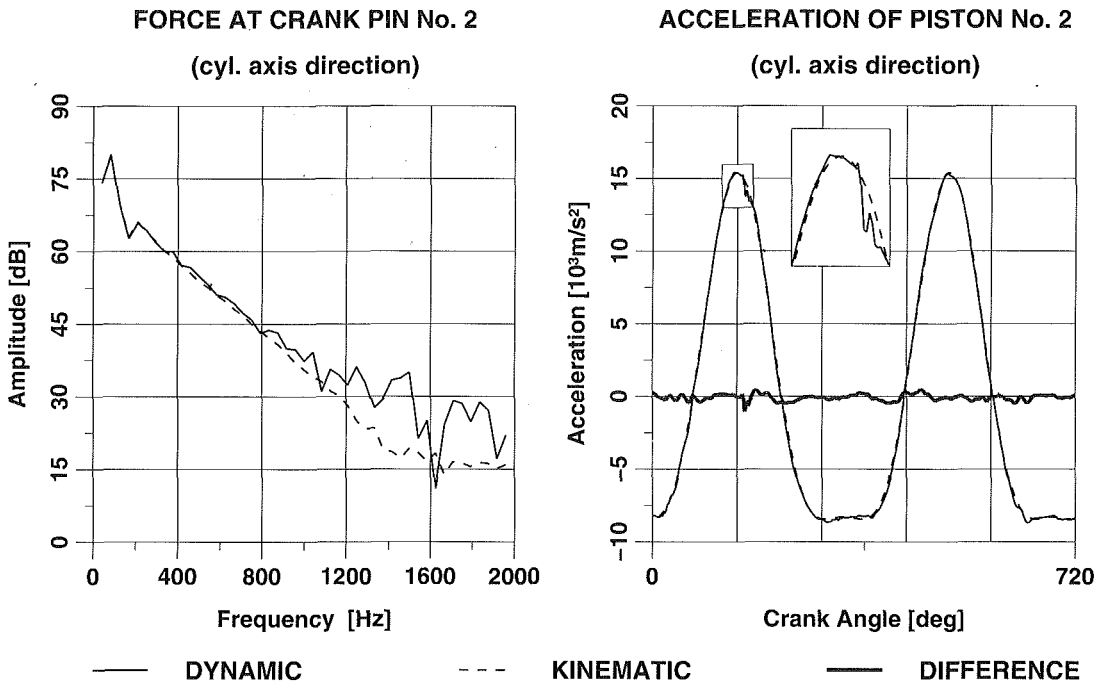


Fig. 15 Excitation forces calculated with (dynamically) and w/o (kinematically) interactions to conrod and piston, crank No. 2, 2.4 1 passenger car 5000 rpm, full load

6 Crankshaft Vibration

To illustrate the further advantages of this prediction technique, Fig. 12 shows calculated results of the crankshaft vibrations of a 2.3 1 passenger car engine at 5000 rpm and full load. In this figure the deformed crankshaft centerline 10 deg (crank angle) before, 5 deg and 20 deg after FTDC of each cylinder are plotted. Of course the ignition of a cylinder mainly causes deformation of the crankshaft close to that cylinder, but reactions due to vibrations of the crankshaft can be seen over the whole length.

Due to the elasticity of the main bearing structure, journal vibrations may cause the original bearing clearance to be exceeded. It can be seen from Fig. 12 that the maximum eccentricity of journals in the bearings is about twice the original clearance of the bearings. Thus, this method can be used to analyze interaction effects of the stiffnesses of the crankshaft and engine structure.

The thrust bearing of an engine should be placed in order to prevent acoustic problems. Finding the optimal position for the axial thrust bearing may be of importance in the design stage of engines.

Therefore postprocessing of KWDYN enables to plot axial vibrations over the whole crankshaft length. Figure 13 shows these axial vibrations over the whole engine cycle for a six cylinder truck diesel engine at 1300 rpm and full load. Longitudinal vibration waves, caused by the ignition of each cylinder, run between the crankshaft ends and result in maximum vibration amplitudes at the pulley. For this example the optimum position for the axial thrust bearing is at main bearing No. 6.

Finally moments and forces for the stress analysis can be plotted directly from the results of these calculations. The moments in torsional and bending directions as well as the shearing forces can be obtained for each beam of the crankshaft model.

Figure 14 shows an example of analyses of principle design modifications of crankshaft and bearings of a six cylinder truck diesel engine. In this engine the maximum stresses were found at overspeed. The results confirm a dominating influence of the crankshaft stiffness. Bending moments of the crankshaft are also influenced by the change of bearing clearance, which of course means a change of bearing stiffness. A change of structure stiffness influences both the bending moments and the torsional moments of the crankshaft.

7 Vibration of Cranktrain Including Conrod and Piston

In the running engine an interaction between the dynamics of piston and conrod on the one hand and the vibration of the rotating crankshaft on the other exists. The amount of the differences of the excitation forces on the crank between pure kinematic and dynamic crank loads is analyzed for a 2.4 liter passenger car engine. For this analysis the models for the cranktrain and the bearings were extended by conrod and piston models. Each conrod is represented by three masses and two beams. The piston node is guided in the cylinder axis direction at the cylinder centerline. The coupling of the structure parts in the conrod bearings is performed by stiffness and damping forces. The forces (in the oil film) are nonlinear depending on relative movements and velocities. The nonlinear bearing characteristics are approximated by a precalculation of the hydrodynamic bearings.

The results of these calculations at 5000 rpm and full load are shown in Fig. 15. The results indicate that at frequencies lower than 800 Hz the excitation forces upon the crank No. 2 are negligibly affected by the dynamics of conrod and piston. At frequencies higher than 800 Hz the calculated dynamic forces increase compared to the kinematic ones. The differences are 10 dB and more and thus showing considerable effects for noise excitation.

Resulting differences between dynamic and kinematic piston motion are known from experiments and range up to 0.35 mm for this engine.

8 Conclusions

For the design analyses of the stress and vibration of crankshafts a newly developed calculation system is introduced in this paper. To demonstrate the advantage of this method for the simulation of the forced vibrations of crankshafts, results

of applications for the analyses of passenger car and truck diesel engines were discussed. From these results and further investigations at AVL the following conclusions can be drawn:

- The described calculation system KWDYN enables the analysis of crankshaft vibrations in the view of:
  - bearing properties
  - flywheel position and gyroscopic effects
  - stiffness of engine construction
  - coupling of torsional, bending, and axial vibration
- The results show influences of engine structure on bending and torsional vibrations and stress moments. These influences depend on the ratio of both the crankshaft and the engine structure stiffnesses.
- The influence of a displaced flywheel position due to increased gear train width increases mainly the radial vibration, but not its wobbling motions.
- The calculation program enables us to consider also elements for conrods, pistons, and their bearings.

Acknowledgments

The authors wish to thank their colleagues at AVL for their contribution to the work described in this paper.

References

1 Cuppo, E., and Gaudio, R., 1975, "Crankshaft Bending Stresses: Experimental Investigations and Calculation Methods," CIMAC Barcelona.

2 Gross, W., and Hussmann, A. W., 1966, "Forces in the Main Bearings of Multicylinder Engines," SAE Paper No. 660756.

3 Donath, G., and Seidemann, H., 1987, "Auslegung von Dieselmotor-Kurbelwellen: Vergleich gemessener und gerechneter Spannungen—Teil 1 und 2 (Design of Diesel Engine Crankshafts: Comparison of Measured and Calculated Stresses—Parts 1 and 2)," *MTZ*, Vol. 48, pp. 6 and 11.

4 Parlevliet, Th., 1981, "Modell zur Berechnung der erzwungenen Biege- und Torsionsschwingungen von Kurbelwellen unter Berücksichtigung der Ölverdrängungsdämpfung und Steifigkeit in den Grundlagern (Model for the Calculation of Forced Bending and Torsional Vibrations Considering Squeeze Damping and Stiffness in the Main Bearing Oilfilms)," Thesis, Munich, Germany.

5 Svoboda, M., Ulrich, H., and Steinmüller, G., 1986, "Schwingungsberechnung bei der Entwicklung von Pkw-Motoren (Vibration Analysis Developing Passenger Car Engines)," VDI-Berichte No. 113.

6 Gran, S., 1989, "Berechnung der rotierender Kurbelwelle bei mehrfacher, nichtlinearer Gleitlagerung (Calculation of Rotating Crankshaft Considering Nonlinearity of their Slider Bearings)," *Dipl. Arbeit*, TU Graz, Austria.

7 Priebsch, H. H., Affenzeller, J., and Kuipers, G., 1990, "Prediction Techniques of Vibration Transfer in Engines," C420-023, IMechE, London.

8 Priebsch, H. H., Affenzeller, J., and Kuipers, G., 1990, "Structure Borne Noise Prediction Techniques," SAE Paper No. 900019.

9 Schweiger, W., Vollan, A., and Dirschmid, W., 1982, "Zur Strukturdynamik von Kurbelwellen (Structure Dynamics of Crankshafts)," VDI-Berichte No. 444.

10 Yamashita, T., Takahara, M., Tsujimura, A., 1980, "A Convenient Calculation Method on Bending Stress of Crankshaft," VDI-Berichte No. 370.

11 Okamura, H., Sogabe, K., 1988, "Dynamic Stiffness Matrix Method for the Three-Dimensional Analysis of Crankshaft Vibrations," C23/88 IMechE.

12 Yoshikawa, K., 1986, "Vibration of Crankshafts at High Engine Speeds," *FISITA*.

13 Hodgetts, D., 1976, "The Whirl Modes of Vibration of a Crankshaft," IMechE C216/76.

14 Priebsch, H. H., 1980, "Rotordynamik bei grosser Unwucht mit Berücksichtigung nichtlinearer Feder- und Dämpfungseigenschaften der Lagerung (Rotordynamics With High Eccentricity Ratios Considering Nonlinear Spring and Damping Characteristics in the Bearings)," Thesis, Graz, Austria.

15 Heath, A. R., and McNamara, P. M., 1990, "Crankshaft Stress Analysis—Combination of Finite Element and Classical Analysis Techniques," ASME JOURNAL OF ENGINEERING OF GAS TURBINES AND POWER, Vol. 112, No. 3.

16 Pestel, E. C., and Leckie, F. A., 1963, *Matrix Methods in Elastomechanics*, McGraw-Hill, New York.

17 Bathe, K.-J., 1990, *Finite Elemente Method*, Springer-Verlag.

APPENDIX

Three-Point Difference Method

$$\mu = 3/(2\Delta t) \quad \xi = 1/(\Delta t^2)$$

$$\tau_f = -\mu/3 \bullet [-4f(t-\Delta t) + f(t-2\Delta t)]$$

$$\eta_f = -\xi \bullet [-2f(t-\Delta t) + f(t-2\Delta t)] \qquad (25a)$$

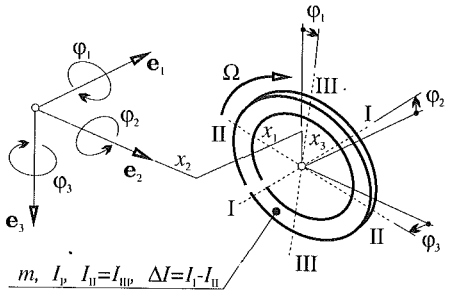


Fig. 16 Mass element—flywheel

Houbold's Method

$$\mu = 11/6\Delta t \quad \xi = 2/\Delta t^2$$

$$\tau_f = -\mu/11[-18f(t-\Delta t) + 9f(t-2\Delta t) - 2f(t-3\Delta t)]$$

$$\eta_f = -\xi/2[-5f(t-\Delta t) + 4f(t-2\Delta t) - f(t-3\Delta t)] \qquad (25b)$$

Newmark's Method [17] ($\delta \geq 1/2$, $\alpha \geq 1/2(1/4 + \delta)^2$)

$$\mu \equiv a_1 = \delta/(\alpha\Delta t) \quad \xi \equiv a_0 = 1/(\alpha\Delta t^2)$$

$$\tau_f = a_1 \bullet f(t-\Delta t) + a_4 \bullet \dot{f}(t-\Delta t) + a_5 \bullet \ddot{f}(t-\Delta t)$$

$$\eta_f = a_0 \bullet f(t-\Delta t) + a_2 \bullet \dot{f}(t-\Delta t) + a_3 \bullet \ddot{f}(t-\Delta t)$$

$$a_0 = 1/(\alpha\Delta t^2) \quad a_2 = 1/(\alpha\Delta t) \quad a_4 = \delta/\alpha - 1$$

$$a_1 = \delta/(\alpha\Delta t) \quad a_3 = 1/(2\alpha) - 1 \quad a_5 = \Delta t/2(\delta/\alpha - 2) \qquad (25c)$$

Example of a Mass Element Transfer Matrix—Flywheel

Figure 16 shows a flywheel mass element.

$$\mathbf{U} = \begin{bmatrix} \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{U}_{pv} & \mathbf{E} & \mathbf{u}_p \\ \mathbf{0} & \mathbf{0} & 1 \end{bmatrix} \qquad (26)$$

$$\mathbf{U}_{pv} = \begin{bmatrix} m\xi & 0 & 0 & 0 & 0 & 0 \\ 0 & m\xi & 0 & 0 & 0 & 0 \\ 0 & 0 & m\xi & 0 & 0 & 0 \\ 0 & 0 & 0 & I_1\xi & 0 & 0 \\ 0 & 0 & 0 & 0 & I_{II}\xi & \Delta I\mu \\ 0 & 0 & 0 & 0 & -\Delta I\mu & I_{II}\xi \end{bmatrix} \qquad (27a)$$

$$\mathbf{u}_p = \begin{bmatrix} m\eta_{x_1} + \mathbf{Q}_{1,a} \\ m\eta_{x_2} + Q_{2,a} \\ m\eta_{x_3} + Q_{3,a} \\ I_1\eta_{\varphi_1} + M_{1,a} \\ I_{II}\eta_{\varphi_2} + \Delta I\tau_{\varphi_3} + M_{2,a} \\ I_{II}\eta_{\varphi_3} - \Delta I\tau_{\varphi_2} + M_{3,a} \end{bmatrix} \qquad (27b)$$

Example of a Beam Transfer Matrix

The transfer matrix for a straight uniform beam (Fig. 17) with principal axis parallel oriented to the vectors of the $\mathbf{e}_i$ axis is given by

$$\underline{\mathbf{U}} = \begin{bmatrix} \mathbf{E} & \mathbf{B}_I & \mathbf{B}_{II} & -\mathbf{B}_{III}^T & \mathbf{b} \\ \mathbf{0} & \mathbf{E} & \mathbf{B}_{III} & \mathbf{B}_{IV} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{E} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{B}_I & \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{0} & 1 \end{bmatrix} \qquad (28)$$

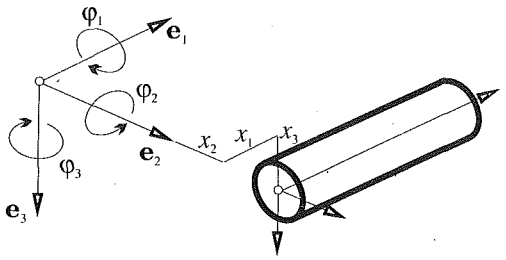


Fig. 17 Straight uniform beam element

The associated partial (3×3) matrices are:

$$\mathbf{B}_{\mathrm{I}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & L \\ 0 & -L & 0 \end{bmatrix} \tag{29a}$$

$$\mathbf{B}_{\mathrm{II}} = \begin{bmatrix} L/(EA) & 0 & 0 \\ 0 & -L^3/(6EI_3) + \kappa_2 L/GA) & 0 \\ 0 & -L & -L^3/(6EI_2) + \kappa_3 L/(GA) \end{bmatrix} \tag{29b}$$

$$\mathbf{B}_{\mathrm{III}} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & L^2/(2EI_3) \\ 0 & -L^2/(2EI_2) & 0 \end{bmatrix} \tag{29c}$$

$$\mathbf{B}_{\mathrm{IV}} = \begin{bmatrix} L/(GI_t) & 0 & 0 \\ 0 & L/(EI_2) & 0 \\ 0 & 0 & L/(EI_3) \end{bmatrix} \tag{29d}$$

$$\mathbf{b} = \begin{bmatrix} L \\ 0 \\ 0 \end{bmatrix} \tag{29e}$$

E = modulus of elasticity of the beam material
G = thrust modulus of the beam material
L = length of the beam element
A = section area of the beam
I_2, I_3 = second degree moment of area around the $\mathbf{e}_2$, $\mathbf{e}_3$ axis
I_t = torsional moment of area of second degree
κ_2, κ_3 = thrust adjustment figure in $\mathbf{e}_2$, $\mathbf{e}_3$ direction

Field Testing to Validate Models Used in Explaining a Piston Problem in a Large Diesel Engine

M. J. Graddage
Ricardo Consulting Engineers Ltd.,
Shoreham-by-Sea,
West Sussex, United Kingdom

F. J. Czysz
Pennsylvania Power & Light Co.,
Allentown, PA 18101

A. Killinger
MPR Associates Inc.,
Washington, DC

Two crankcase explosions occurred within one month in diesel engines that drive large emergency generator sets at a nuclear power plant in Eastern Pennsylvania. As a result, the electric utility conducted an extensive investigation to determine the root cause(s) of the problem. Initial inspections confirmed that the crankcase explosions were the result of pistons and liners becoming overheated. The technical challenge was to establish why the pistons and liners were overheating when other engines of the same type did not appear to have the problem in the same duty. Analytical models of piston motion, engine start, and run thermodynamics, and a finite element analysis of piston distortion during engine start and load transients were developed. Preliminary work with these models predicted a feature of the piston design that could adversely affect lubrication conditions during a rapid start and load transient. Final input data to refine the models were needed and these were obtained from tests carried out on a similar diesel generator operated by a municipality in Iowa. This paper describes the successful accomplishment of the field tests using state-of-the-art instrumentation and recording equipment. It also shows how the modeling and test work identified wear at certain locations on the piston skirt as the origin of distress leading to the crankcase explosions. Unfavorable engine starting and loading conditions as well as less than desirable piston skirt-to-liner lubrication conditions in the engines at the nuclear power plant have been identified as the root causes and corrective action has been initiated.

1 Introduction

In a one-month period, crankcase explosions occurred in two of the five emergency diesel generator set engines at the Susquehanna nuclear power plant in Eastern Pennsylvania operated by the Pennsylvania Power and Light Company (PP&L). The engines were being subjected to a routine monthly test at the time to demonstrate their continued availability, a procedure that involved starting the engines from stand-by conditions and bringing them up to full load within 90 seconds. The engines, 16 cylinder units rated at 4 MW, are a type licensed by the U.S. Nuclear Regulatory Commission for this duty and are installed at eight other nuclear power plants. At these nuclear power plants, as well as other stationary and marine applications, the engines have a generally good reliability record.

A preliminary investigation [1] quickly established that the crankcase explosions were a result of pistons and liners becoming overheated due to localized scuffing on the upper diameter of the piston skirt and liner on the antithrust side (the side opposite that carrying the major side forces due to conrod angularity on the power stroke). However, no defect in the pistons, liners, or other causal factors could be easily identified that would clearly explain why the scuffing should have occurred in this unusual position. Up until this time crankcase explosions had not been experienced by any of the other engines operating in the same duty. While introducing changes to operating procedures and interim modifications designed to improve the lubrication conditions of the piston and liner, PP&L therefore also initiated an in-depth investigation to establish the root cause of the problem.

This paper provides an outline of the work done to assist the electric utility to examine the operating conditions of the piston and liner inside the diesel engine. The paper describes how analytical modeling and field test work were combined successfully to identify the root cause of the crankcase explosions.

2 Method of Approach

Because of the limited availability of the engines for test work at the nuclear power plant, the primary method of investigating the piston problem was analytical. A Finite Element (FE) model of the piston and liner was used to predict piston distortions and piston/liner contact pressures. The rapid start and loading sequence regularly employed at the nuclear power plant was considered a potential causal factor. Therefore, to show how the piston/liner contact changes through the operating cycle, analyses were carried out for stand-by conditions,

Contributed by the Internal Combustion Engine Division and presented at the 14th Annual Fall Technical Conference, Waterloo, Iowa, October 4–7, 1992. Manuscript received by the Internal Combustion Engine Division July 1, 1992. Associate Technical Editor: J. A. Caton.

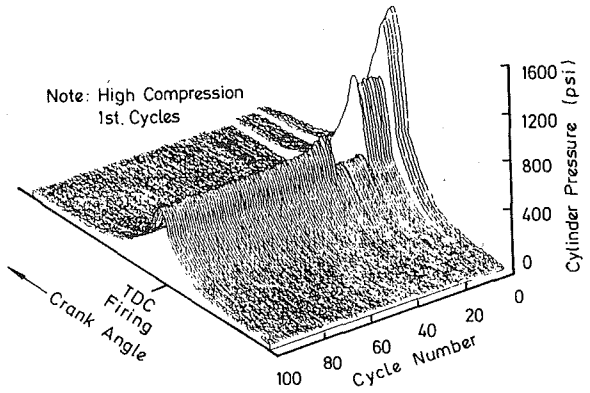


Fig. 1 Measured cylinder pressure of nuclear plant engines during start

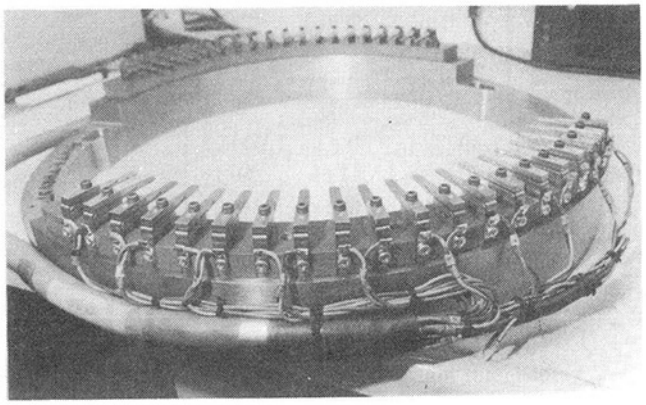


Fig. 2 Contact gear leaf springs on carrier ring

full-load steady state, and two intermediate points in the loading transient.

The starting point for the distortion and contact patch analysis was component temperatures and piston forces and attitudes. These were determined by the simulation of piston motion and component thermal conditions, both at steady state and throughout the start-up and loading transients. As this type of analysis had previously been carried out (with considerable success) only for steady load, steady temperature conditions, a program of test work involving temperature and motion measurements was undertaken to provide validation for the results of the simulation.

The validation test work was carried out on a diesel generator operated by a municipality in Sumner, IA. Although this engine was not identical to the nuclear power plant engines, it provided data for similar operating conditions. The actual conditions were then assessed using the analytical models with only a small amount of extrapolation.

3 Preliminary Measurements

Before undertaking the simulation work, a series of measurements were taken from the nuclear emergency diesel generator engines during typical start and running conditions. A comprehensive package of instrumentation and data logging equipment was employed to measure and record in-cylinder pressure, injector needle lift and other more general temperatures and pressures, etc. The results of these tests provided baseline data for the analysis work but also highlighted a characteristic of the air starting system of the engine, which caused compression and maximum cylinder pressures to be increased significantly during the first few cycles of a start; see Fig. 1. This finding was considered significant as the loadings on the anti-thrust side of the piston produced by the cylinder pressure would be increased to near those seen at full load, but at a lower speed, where lubrication and the beneficial effects of inertia forces (see Section 5.1) would not be so great.

Some imbalance in tuning of the nuclear power plant engine was also revealed by the measurements, but neither this nor any other unusual characteristics were identified that could be shown to be the sole root cause of the piston problem.

Once the baseline data had been recorded from the engine at the nuclear power plant, further measurements were taken from an engine operated by a municipality in Iowa. This engine had been identified as potentially suitable for use as a validation tool, as it was similar to those at the nuclear facility and employed the same pistons and liners. It was a 12-cylinder dual-fuel engine operating at 514 rpm whereas the nuclear power plant engines were 16 and 20-cylinder models running on diesel fuel at 600 rpm. The inertia of the generator was also higher than that of the nuclear power plant engines and the air starting system did not exhibit the same pressure "boosting" characteristics. However, the full-load bmeps of the engines were similar and the tests, coupled with supporting piston motion analysis work, established that the engine could be modified and operated in such a way as to provide the data necessary to validate the computer models.

Table 1 Main features of nuclear power plant and test* engines

Bore:	13.5 ins
Stroke:	16.5 ins
Rated Speed:	600 rev/min (514 rev/min)
Cylinders:	16 at 45° included vee angle (12 cylinders)
Continuous rating:	4000kW electrical (2748 kW electrical)
Rated BMEP:	13.6 bar (14.6 bar)
Piston Type:	One piece cast iron, tin plated skirt
Connecting Rod Configuration:	Master and Slave

* Figures in parenthesis refer to test engine.

4 Engine Tests to Validate the Modeling Work

4.1 Engine. Some details of the main features of the engines at the nuclear power plant are given in Table 1 compared with those of the test engine.

4.2 Instrumentation. A comprehensive package of instrumentation was installed on the test engine to monitor start air system operation, engine performance parameters, component temperatures, and piston motion. Details of some of the instrumentation and data logging systems employed are given in the following sections.

4.2.1 Component Temperatures. Piston and liner temperatures were measured by thermocouples. Twenty were installed in a piston and seven in a liner, both components being installed in the same position in the engine.

Signals from the piston thermocouples were communicated to the data logging system through intermittent contact gear designed and manufactured by Ricardo. The thermocouple leads were fixed to the inside of the piston and brought down to contact pins on the lower edge of the piston skirt. At the bottom dead center position these pins made contact with leaf springs attached to the bottom of the liner; see Fig. 2. During the brief period of contact the signal from each thermocouple charged up a capacitor. The time constant of the capacitor circuit was chosen such that a steady voltage was maintained between contacts and it was this voltage that was logged. The liner thermocouples were wired directly to the data logger.

4.2.2 Piston Motion. Piston motion was investigated using eight Micro-epsilon eddy current proximity probes mounted in the liner. The probes were tuned to give an output proportional to the gap when only the skirt of the tin-plated piston was opposite the probe. The passage of the rings and ring lands was not monitored. The main objective of the measurements

Table 2 Range of tests completed

Hours Since Last Run	Pre-Start Rack/Governor Setting	Idle Time	Rate of Loading	Remarks
18	0 mm 0.55	5 mins	0-2670 kW in 10 mins	'Normal' start
5 ½	45mm 0.65	40 secs	0-2600 kW in 60 secs	Rapid start from 'warm'
½	45 mm 0.65	1½ mins	0-2900 kW in 60 secs	Rapid start from 'hot'
16	45 mm 0.65	1 ½ mins	0-2900 kW in 60 secs	Rapid start from 'cold'
½	0 mm 0.55	5 mins	750 kW every15 mins	Slow start and load

Table 3 Variation in prestart temperatures

		Max	Min
Piston (crown edge)	°C	100	25
Piston (crown centre)	°C	65	25
Liner (mid-stroke)	°C	60	25
Oil	°C	65	25
Coolant	°C	50	25

was to monitor piston motion in the thrust/antithrust plane, and, therefore, six of the eight sensors were in this plane (four on one side, two on the other). These probes were distributed such that a minimum of three probes were adjacent to the piston skirt at any instant. This ensured that both angular and translational data were available at all times. Two probes were placed in the crankshaft axis direction to give an indication of motion in the orthogonal plane. Signals from the probes were input to a signal conditioning unit, which linearized the probe output and corrected for probe temperature.

4.2.3 Data Logging Systems. Cylinder pressure, needle lift, and promixity probe signals from the test cylinder were logged at each degree of engine crank angle with a Ricardo Configurable High Speed data Acquisition System for Engines (CHASE) [2]. The CHASE system comprised a front end multichannel acquisition unit, communicating with a PC, which provided data processing, display, archiving, and user interface functions. Ten channels of analog data were acquired synchronously on a trigger from a digital encoder coupled to the after end of the generator shaft.

For general purpose data acquisition of temperature, pressure, etc., a system comprising a number of Isolated Measurement Pods (IMPs) located adjacent to the engine, communicating with a PC via a serial data link, was employed. Each IMP was a self-contained multichannel data acquisition unit, providing multiplexing, analog-to-digital conversion (A/D), scaling, cold junction compensation, and communication functions, all housed inside a robust enclosure.

Logging data every degree for every cycle of a test run was not necessary and would have produced impractically large quantities of data. In order to avoid this and to have data throughout each test run, a selective sampling technique was employed: The CHASE high-speed logging system was configured to count engine cycles and was programmed with a map of which cycle numbers to record. A typical logging regime was the first 50 engine cycles followed by 5 cycles every 50 throughout the rest of the test. The general purpose logger was programmed to record data every second for the first part of the test and then every 3, 10, or 30 seconds thereafter as appropriate.

4.3 Test Procedure. The engine was set up prior to test to run in the diesel-only mode. The original dual-fuel pumps were retained but the camshafts were retarded in order to produce in-cylinder conditions closer to those of the nuclear power plant engines. Fuel injector nozzles were installed that were identical to those used in diesel engines at nuclear power plants.

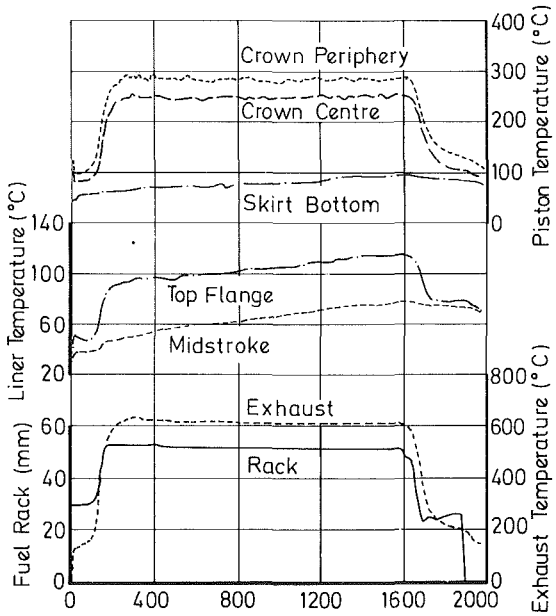


Fig. 3 Data recorded during start and load test

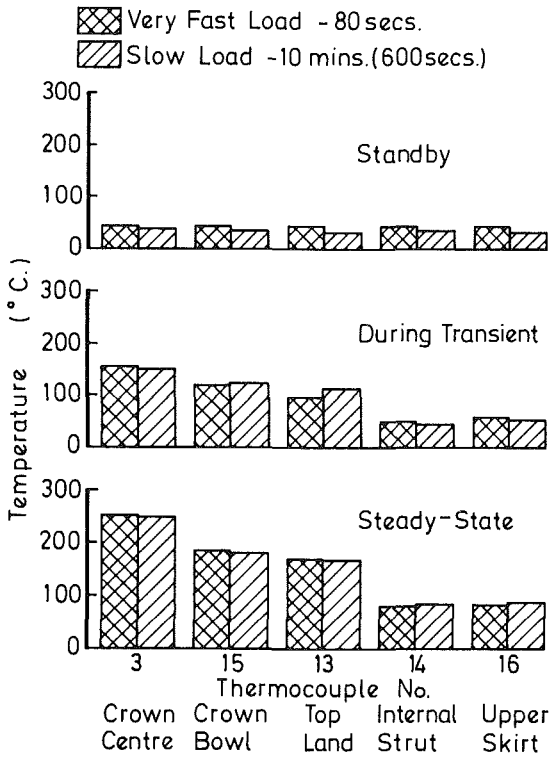


Fig. 4 Comparison of piston temperature distribution for fast and slow loadings

The engine was run-in and then a series of tests were undertaken covering different motoring and starting conditions and zero to full power loading times varying from one hour down to only 60 seconds. A summary of some of the tests carried out is given in Table 2 and the variation in prestart conditions achieved is summarized in Table 3.

For more of the tests a maximum power setting of 2900 kWe was employed (5.5 percent overload on the normal full power of the test engine in the diesel mode). The intention was to ensure that in-cylinder conditions of the test engine were at least equal to or slightly more severe than those of the nuclear power plant engines.

Table 4 Comparison of piston side thrust at two speeds

	Maximum Piston Side Thrust (kN) 600 rev/min	514 rev/min
Top of Skirt		
Major thrust side	34.2 @ 30	36.2 @ 30
Minor thrust side	13.0 @ 600	13.3 @ 700
Bottom of Skirt		
Major thrust side	22.0 @ 30	23.3 @ 30
Minor thrust side	5.2 @ 400	5.0 @ 700

Additionally, the difference in timing of the piston cross-overs is less than one crank degree.

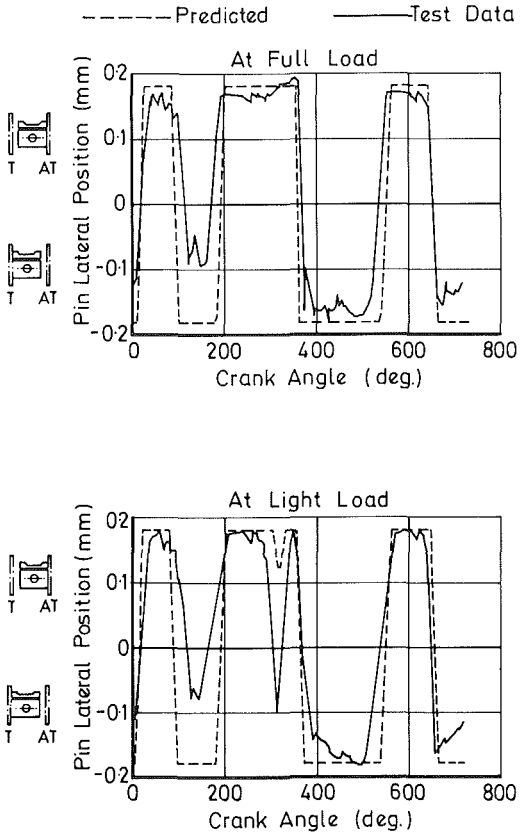


Fig. 5 Comparison of measured and predicted piston motion

4.4 Test Results. Typical plots of piston temperature versus time are shown in Fig. 3. The three piston thermocouples shown cover the full range of readings from all 20 thermocouples. The crown temperatures react quickly to the application of load and reach maximum temperature within two minutes of full load being applied. The response of the skirt temperatures, represented by the bottom temperature in Fig. 3, is much slower and is similar to that of the oil and coolant temperatures.

The temperature of the piston at five positions (from crown center to skirt bottom) measured before start, during transient, and at full power for two tests with widely different loading rates are compared in Fig. 4. This plot gives an indication of the effect of loading rate on the temperature distribution within the piston. The upper plot shows that the prestart temperatures of the two tests were a little different but confirms that the piston was at uniform temperature before each test. The lower plot shows that there was a similar temperature gradient throughout the piston at the end of each test. The center plot shows a point where the crown center thermocouple had reached a temperature approximately halfway between the stand-by and steady-state conditions during each test. The time after

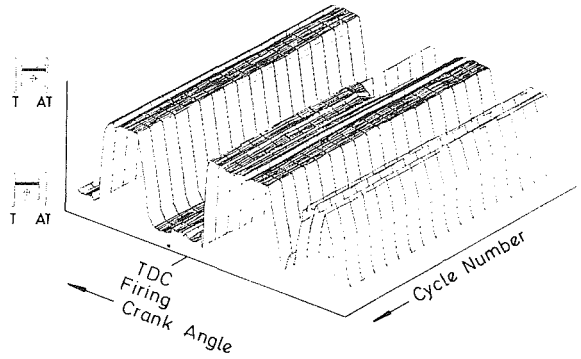


Fig. 6 Piston lateral movement every 100th cycle during a slow transient from 1500 kW to 2250 kW

start at which this temperature was reached was different for each test due to the difference in loading rates. However, it can be seen from the plot of the other thermocouple readings at this "midpoint" that the temperature distribution throughout the piston differed for each test but by only a small amount (approximately 20°C). This finding shows that the thermal gradient down the piston is controlled primarily by crown temperature and not loading rate.

Two of the seven liner thermocouple temperatures are shown in the middle plot of Fig. 3. Although the top flange temperature changed quickly, the temperatures of the liner in that part of the cylinder swept by the piston skirt responded to the increase in load much more slowly and in a similar way to those of the upper and lower piston skirt. This suggests that liner/piston skirt clearances do not change greatly as the engine warms up (coefficients of thermal expansion are similar).

5 Analysis Work

In the following sections, selected results of the piston motion, thermal transient, and distortion and contact patch modeling work carried out are shown and compared to the results of the test work covered in Section 4. The modeling methods employed are described in more detail in (3).

5.1 Piston Motion. Piston Secondary motion, tilt, and side forces for use in the finite element study of piston/liner contact during engine start-up were calculated using a piston motion program (PISMOT). Input data for the program included component dimensions, masses, coefficients of friction, and cylinder pressure. Calculations were carried out for a range of friction, and cylinder pressure. Calculations were carried out for a range of operating conditions and speeds and with a variety of cylinder pressure loadings. Piston motion and side forces were calculated for the whole of the 720 deg operating cycle at these conditions.

The results of the calculations confirmed that the difference in piston motion between the test engine at its rated speed of 514 rpm and that of the nuclear power plant engines at their 600 rpm rated speed was almost indiscernible and that on loadings quite small; see Table 4.

The predictions of typical full load and light load cycles of piston motion are compared in Fig. 5. It can be seen that the predictions are in generally good agreement with the measured motions, particularly when gas pressure, and hence piston side loading, is high.

It was noted that both prediction and measurement showed that the piston is held flat against the antithrust side for the whole of the compression stroke at full load due to the increased compression pressures produced by turbocharger boost at this load. At light load, the compression pressures are not high enough to keep the piston pressed hard against the antithrust side for the whole of the compression stroke, the piston being "lifted off" the cylinder wall in the final half of the

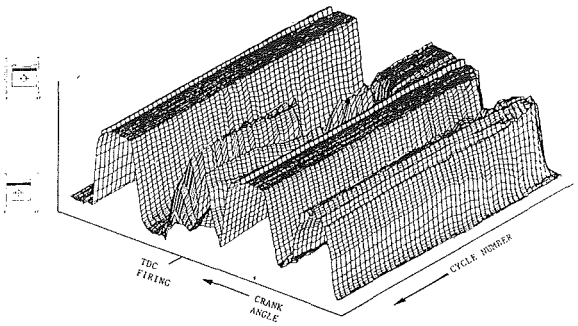


Fig. 7 Piston lateral movement of the 1st 50 cycles of a typical fast start

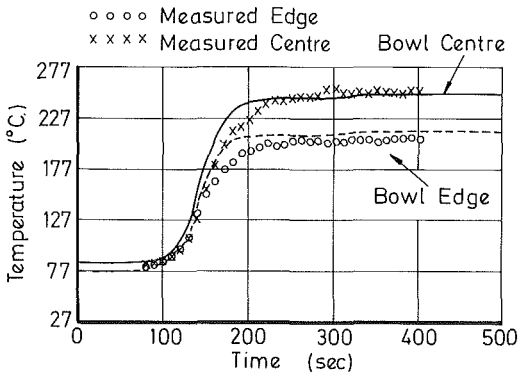


Fig. 8 Comparison of predicted (lines) and measured (symbols) temperature histories for the test engine piston during transient.

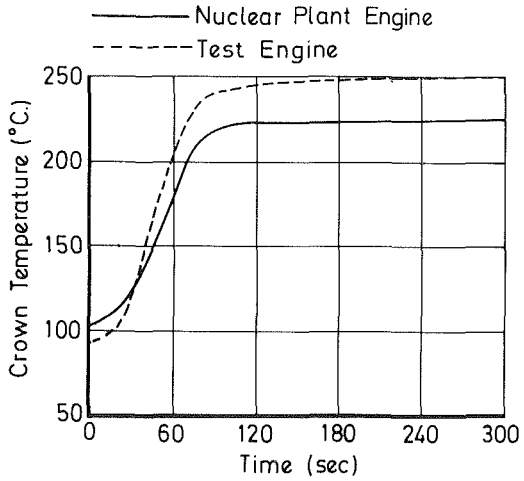


Fig. 9 Comparison of predicted piston transient temperatures for nuclear plant and test engines

stroke by inertia forces. This effect is also well illustrated by the three-dimensional plot of measured piston motion shown in Fig. 6. Here piston secondary motion is plotted every 100th cycle over a period in which the engine load was increased from 1500 to 2250 kW. It can be seen that piston motion changes during this load step due to the increased boost and resulting compression pressures. This plot also shows that piston secondary motion is very consistent from cycle to cycle at the same load.

Measured piston motion for the first 50 consecutive cycles recorded during a typical start of the test engine is shown in Fig. 7. For the first 18 or so cycles, as the engine accelerates from rest and inertia forces are relatively low, compression pressures, even without the aid of boost from the turbocharger, are sufficient to hold the piston against the antithrust side of the cylinder during most of the compression stroke. However, as the speed of the engine rises, the effects of inertia become more pronounced; the compression pressures can no longer hold the piston against the antithrust side for the whole of the compression stroke.

The effect of component temperatures on piston motion was negligible. The skirt warms up relatively slowly, at a similar rate to the liner and, therefore, the skirt-to-liner clearance does not change much as the engine warms up. Although the shape of the piston changes due to the temperature increase, this does not appear to affect piston motion significantly.

The results of the piston motion measurements successfully validated the motion modeling work and hence validated the use of the forces calculated by the simulation in the examination of piston distortion and piston/liner contact patch.

5.2 Transient Thermal Analysis. Before the distortion and contact patch analysis could be undertaken it was necessary to identify the "worst case" time steps in the start up and loading transient and provide boundary conditions for these steps in addition to the steady-state information more usually employed for the distortion analysis. Additionally, it was necessary to use modeling to compare the rate of thermal loadings of the test engine with those of the nuclear power plant engines. Ricardo carried out the performance/thermal simulations using their IRIS program [4].

The preliminary phase of the work indicated that temperature gradients and, therefore, distortions were greatest at full-power, steady-state conditions. Component temperatures rose from stand-by conditions to the full power state gradually without unusual excursions. This was confirmed by the test work, Fig. 3. Therefore, three thermal condition time-steps were identified for the preliminary distortion analyses; stand-by, full load, and a point after start, which was the earliest time that full fueling had been applied to the test engine.

Once the engine test work was completed, validation of the preliminary analysis work and adjustment of the model to match the measured conditions was undertaken and a further intermediate time step was also identified for the distortion analysis. The predicted heat fluxes were modified to match the calculated temperature map with the piston temperature measurements and the resulting comparison between predicted and measured values is shown in Fig. 8.

A comparison of the thermal response of test and the nuclear power plant engines using the models is shown in Fig. 9. It can be seen that the objective of producing a slightly more severe thermal transient on the test engine than that seen at the nuclear power plant was achieved. The maximum steady-state piston crown temperatures for the test engine condition were 20–25°C higher than normal operation at the nuclear power plant but the temperature distribution through the piston was very similar.

Validation of the transient thermal analysis confirmed the heat flux values used in the finite element model to determine distortions.

5.3 Distortion and Contact Patch Analysis. The objective of the distortion and contact patch analysis was to calculate piston/liner contact conditions and thereby assess the possibility of scuffing occurring during a start up and rapid loading transient.

The method employed was to use a three-dimensional FE model of the piston to assess the contact conditions with the piston flat against the bore near TDC, at the point of maximum side thrust as identified by the motion analysis work. The following cases were examined in detail:

(A) Start up at 100 rpm.

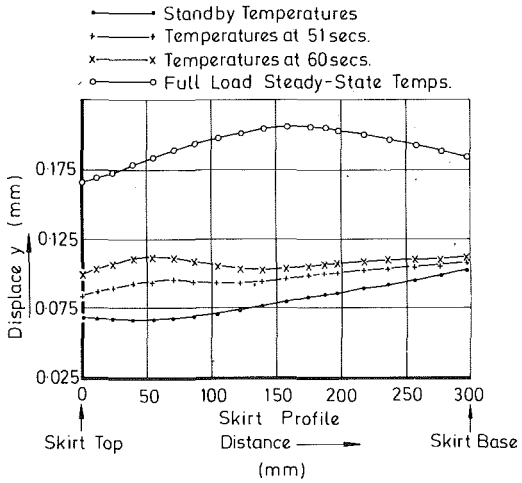


Fig. 10 Skirt radial distortion plot on thrust/nonthrust axis from stand-by to full-load steady-state conditions

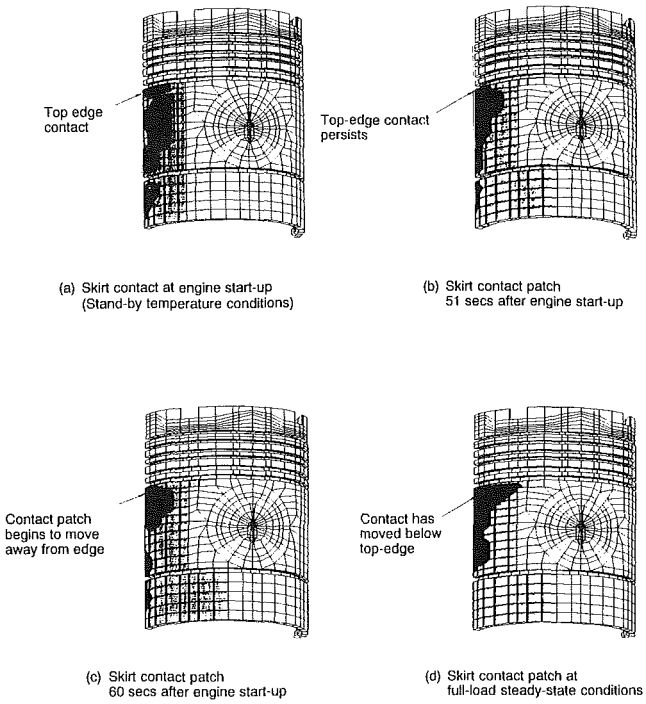


Fig. 11 Skirt contact patch comparison from stand-by to full-load steady-state conditions

(B) Two transient conditions at 600 rpm, full load (nominally 51 and 60 seconds after start).
(C) Steady-state condition at 600 rpm, full load.

The calculations show that the sidewall surface of the unloaded piston skirt is concave in the cylinder axial direction at stand-by temperature but rapidly changes to a better convex shape when the piston has reached final full load temperatures. Convex profiles assist the formation of hydrodynamic oil films in either direction of piston travel. The concave shape at stand-by conditions is a result of stresses caused by the pressing in of the wrist pin bush and differential expansion between the bush and piston materials.

The unloaded piston skirt profiles for the four conditions are compared in Fig. 10. The corresponding contact patches under worst case antithrust side loading are shown in Fig. 11. It can be seen how the concave profile at stand-by temperatures (*a*) will inhibit the formation of oil films on the compression stroke, because of the "sharp" edge on the leading edge of the skirt. Conditions on the downstroke are much more conducive to oil film formation because the gradient at the leading edge is less. In addition, the bottom of the skirt is more flexible than the top, it will therefore more readily conform to the liner bore.

The unloaded skirt profile at the first intermediate time step (51 seconds) has a small lead-in at the top of the skirt. However, predictions for the piston under side loading, Fig. 11 (*b*), show that the skirt has been flattened and that oil film generation would be poor. However, by the second intermediate time step (60 seconds), the skirt profile has improved and it becomes much better as the piston warms up to the full-load steady-state condition where the profile and contact patch are much more acceptable.

Calculations of peak skirt/liner contact pressure for the four cases were also carried out and these confirmed that contact pressures on the antithrust side were highest at the top of the skirt during the early stages of starting. The increased compression pressures seen at the nuclear power plant during starting (Section 3), coupled with a lack of the relieving effects of inertia at low speed, combined to produce this result. However, the peak contact pressure on the thrust side produced by firing loads is substantially higher than that seen on the antithrust side. Markings seen on pistons demonstrate that the contact pressures on the thrust side can be sustained, whereas the contact on the antithrust side leads to very heavy bedding and wear. The explanation lies in the difference in lubrication conditions. On the thrust side high loads occur when the piston is on the downstroke and conditions are favorable for the formation of a hydrodynamic oil film. On the antithrust side the load is applied during the upstroke, and is therefore applied via a boundary lubricated contact, which has a far lower load bearing capacity.

6 Discussion

The crankcase explosions at the nuclear power plant were undoubtedly caused when skirt rubbing loads exceeded the capacity of the skirt lubrication at some time prior to failure. A reduction in the quality of the lubrication conditions, an increase in skirt loadings, or a combination of both could have caused the problem. Equally, there are many possible reasons why skirt loadings and/or lubrication conditions could be adversely affected. The data presented in this paper cover only part of the investigation carried out, but they show how a feature of the piston design was identified that would adversely affect skirt lubrication during start and loading transients.

From the motion analysis it can be seen that the piston is held against the antithrust side throughout the compression stroke when the compression pressures are increased by the turbocharger as load is increased, or by the operation of the start air system. When compression pressures are lower the piston moves away from the liner toward the end of its compression stroke, thus relieving the loads on the antithrust side. The shape of the piston during the early stages of a rapid start and load is also not as favorable as it is at normal operation conditions. Therefore, it can clearly be seen that, due to the particular design of the piston, the operating conditions for the piston skirt are most arduous during the start and early part of a rapid loading. A slightly less rapid rate of loading allows the piston time to change to a more favorable shape before its motion becomes more arduous.

However, this finding could not be held solely responsible for the crankcase explosions, as this type of piston has been employed for a number of years in many different applications with a very low reported incidence of seizures and/or crankcase explosions. The problem at this nuclear power plant was that a combination of circumstances conspired to reduce the quality of the lubrication conditions:

- A lubricating oil was being used that, although normally

perfectly satisfactory for this type of engine and application, nevertheless had a slightly lower extreme pressure rating than oils used at other nuclear power plants.

- The monthly test procedure employed was severe, in terms of rate of loading after start, and was used more frequently than other power plants.
- The particular starting air system configuration employed, which enabled the engine to consistently meet the rapid starting requirement (start to synchronous speed in less than 10 seconds), caused an increase in compression pressures and, therefore, piston side loads in the first few cycles of start.

With the resulting reduction in quality of the lubrication conditions, the engines were put at greater risk of scuffing than most of the other users of this type of engine and this could explain why the problem appeared in diesel engines serving as rapid start, emergency generator sets.

The electric utilities have now amended the testing procedure of the engines and have also improved the quality of piston skirt lubrication by removing a lower oil control ring and piston pin end caps and changing to an oil with a higher extreme pressure rating. They are continuing to monitor the performance of the pistons and are also considering, for the long term, the adoption of a revised piston design.

7 Conclusion

The investigation involved the cooperative effort of at least five very diverse organizations, including the engine manufacturer, the electric utility, two engineering firms, and a municipal electric plant. An engine at a rather remote site in Iowa was modified, instrumented, and tested in under six months in order to support the continued operation of a nuclear power plant.

The work has demonstrated that a combination of advanced engine simulation and modeling techniques, validated by test work on a similar but not identical engine, can successfully identify the root cause of a piston problem that could not be investigated directly.

With the recent and continued advancement of modeling techniques coupled with the availability of ever more sophisticated but easy-to-use instrumentation, the designer will have the necessary tools to investigate potential problems with new component designs, particularly for those to be used in unusual or more arduous applications. The aim for the future must be to make full use of these tools to ensure development and testing time are reduced to a minimum.

Acknowledgments

In preparing this paper, the authors would like to acknowledge the contribution of their colleagues and associates at the companies involved in the work. In particular:

W. Rhoades of Pennsylvania Power and Light Company, J. J. Fowler, M. R. Stott, J. R. Thomas, A. D. S. Bedi, and R. Keribar of Ricardo International, D. G. Evans, C. L. Haller and L. E. Lehman of MPR Associates, Inc., and W. Bohle, J. Duhrkopf, T. Duhrkopf, and R. Jergens of the Sumner Municipal Light Plant, Sumner, Iowa.

The authors would also like to thank the management of the companies involved in the work described, for their permission to publish this paper.

References

1 Csysz, F., "Crankcase Overpressurization at Susquehanna," EPRI Seminar on Diesel Generator Operation, Maintenance and Testing, 1990.

2 Wibberley, D., "Ricardo CHASE Hardware Overview," Ricardo DP 88/1884.

3 Smith, A. V., and Clarke, D. P., "Engine Component Optimisation—The Development of a Predictive CAE Based Model of Piston/Liner Interaction," FISITA 92, CIMAC Seminar—Computation and Simulation.

4 Ricardo Publication, "Computer-Aided Engineering Software for Engines and Propulsion Systems," 1999.

Development of a High-Output Dual-Fuel Engine

P. R. Danyluk
Coltec Industries, Inc.,
Fairbanks Morse Engineering Division,
Beloit, WI 53511

This paper presents the results of a new dual-fuel engine development program. The engine is the largest commercially available in terms of power output (650 hp/cyl) and features very low emissions (1 g/hp-hr NO_x) and excellent fuel consumption (43 percent thermal efficiency). A two-cylinder turbocharged prototype was designed and built for the initial development. Results from testing on 18-cylinder production versions are also reported.

Introduction

In recent years engine development activities have been driven by uprating, fuel consumption improvement, and emissions reductions. In the late 1980's work was started on the development of a new dual-fuel engine, which would have high power output, low fuel consumption, and low emissions. Like other dual-fuel engines this new dual-fuel engine would borrow its structure from a successful diesel engine. In this case the Colt-Pielstick PC2.6 engine was chosen. The development steps that occurred in this program were:

1 Selection of the basic engine design.
2 Design and fabrication of a two-cylinder prototype.
3 Baseline testing.
4 Design refinement.
5 Commercial model testing of full-size engines.

The rest of the paper describes the design and development events that have led up to the latest product offering rated at 282 psi BMEP, 5900 Btu/hp hr fuel consumption, 1 g/hp hr NO_x utilizing natural gas pressures available from pipeline distribution and with 1 percent pilot oil. This latest version of the engine is called the Enviro Design™ PC2.5 engine by Fairbanks Morse.

Discussion

By the mid-1980s most commercially available large dual-fuel engines were fully developed and no new dual-fuel engine had been introduced since the early 1970s. Dual-fuel engine ratings had gone as far as the basic diesel engine upon which they were based would allow in terms of structural limitations. This included the PC2.3 dual-fuel engine, which had been in commercial service since 1971. Further uprating of the PC2.3 would require higher firing pressures if fuel consumption penalties were to be avoided. Fuel consumption penalties would guarantee commercial failure. For this reason the PC2.6 with its higher firing pressure capability was chosen to replace the PC2.3 as the basic building block of the uprated new dual-fuel engine. Table 1 shows the history of the PC2 series engine in both diesel and dual-fuel ratings.

The PC2.6 engine is an underslung crankshaft design with exceptional ruggedness providing exceptional conservatism based on actual firing pressures compared to maximum allowable structurally limited firing pressure. Note in Table 1 that the highest dual-fuel rating of 650 hp/cyl compares to the normal diesel rating of 737 hp/cyl and the sprint rating of 850 hp/cyl. At 650 hp/cyl firing pressures are 1800 psi compared to a structural limitation of around 2400 psi. This is a 25 percent safety factor providing one of the most conservatively rated dual-fuel engines ever developed.

Table 1 History of PC2 series engines

KEY EVENT	HP/CYL	DATE	BMEP
PC2 DIESEL	465	1964	209
PC2.2 DIESEL	500	1966	218
PC2.3 DIESEL	535	1969	234
PC2.3 DUAL FUEL	535	1970	234
PC2.5 DIESEL	650	1971	284
PC2.6 DIESEL	737	1980	322
PC2.6 BTC	850	1982	371
PC2.5 DUAL FUEL	600	1988	262
PC2.5 DUAL FUEL	650	1992	284

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1993. Paper No. 93-ICE-20. Associate Technical Editor: J. A. Caton.

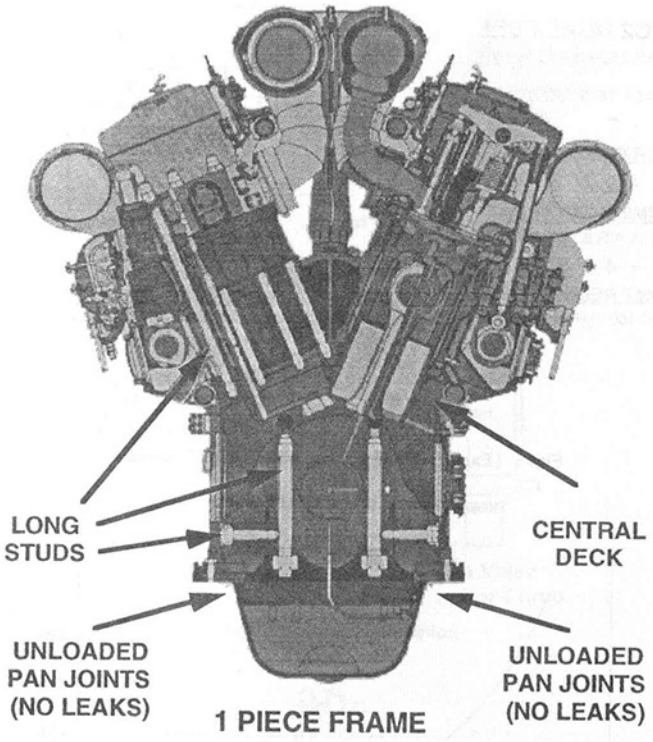


Fig. 1 PC2 engine cross section

Table 2 Firing pressures

ENGINE MODEL	STRUCTURAL LIMIT PSI	FULL LOAD FIRING PRESS. PSI	SAFETY MARGIN %	MAX OUTPUT KW
PC2.3	1800	1400	22	6900
PC2.5	2100	1500	28	8380
PC2.6	2400	1900	21	9500

Figure 1 shows the cross section of the PC2.6. In Fig. 1 note the one piece thick central deck, which runs the full length of the engine. Thick partitions connect to the central deck and carry the main bearings and crank and long studs thread into the central deck and extend up through the top of the cylinder head. This basic structure provides the extremely high firing pressure capability sought in the diesel engine chosen for uprating the 535 hp/cyl PC2.3 rating to the current PC2.6 dual-fuel ratings. Figure 2 shows the thick central deck at the beginning of the frame fabrication process.

With this as an introduction we will now summarize details of the five steps outlined in the Introduction.

1 Selection of Basic Engine Design. The PC2.6 engine was chosen from among its predecessors to provide the lowest cost structure that also incorporates the highest uprating potential. Uprating potential is firing pressure capability. Table 2 shows a comparison of the three engine models considered and the firing pressure capabilities of each.

In Table 2 the column entitled firing pressure is the diesel mode operating firing pressure that goes with that particular engine model rating and structural limit is the continuous firing pressure capability of the engine structure. The safety factor is the comparison of the operating versus structural limit columns.

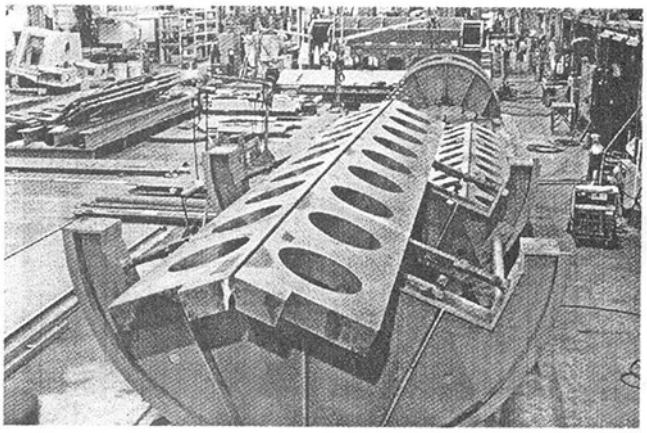


Fig. 2 Central deck

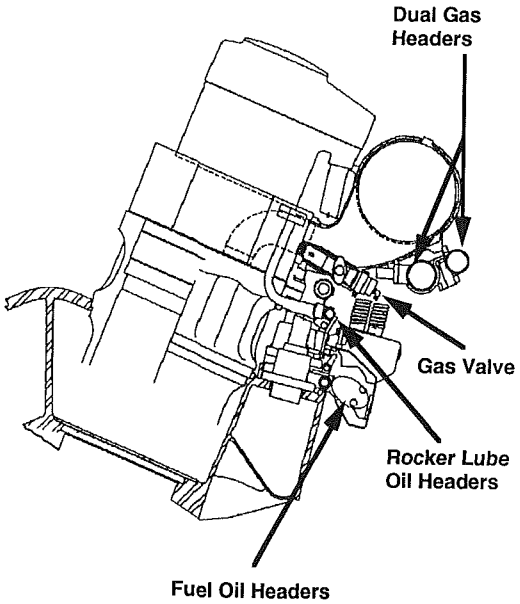


Fig. 3 Cylinder cross section

The basic dual-fuel engine layout is shown by Fig. 3. The figure shows the basic diesel cross section with the header locations and gas valve-in-intake manifold. The gas valve location was chosen for a number of reasons. Since extremely low NO_x is one of our goals, *extremely* homogenous lean mixtures are desirable. By placing the gas valve in the intake runner, mixing is promoted at the gas valve, across the intake valve, and by in-cylinder motion. Other reasons are the more favorable environment as compared to in-cylinder injection or placing it on the inlet valve. By avoiding putting the gas valve on the inlet valve we can also provide for independence from the inlet valve event, which can lead to reduced unburned hydrocarbon emissions. Initial gas valve actuation and timing design was selected to be hydraulic utilizing a small jerk pump both for the actuation and timing. The ultimate design is intended to be electrohydraulic to eliminate costly linkage and lay shafts and to simplify the governor.

High-pressure gas injection was also considered, as was carburation with low-pressure gas, but both were dropped from consideration. It was also felt that high-pressure gas injection directly into the cylinder would have many more disadvantages than advantages. One advantage is avoidance of detonation by diesel-like injection without allowing time for compression heating of an air fuel mixture. The advantages include the fact that lean combustion itself is going away from the detonation region, extremely lean homogeneous mixtures required for low NO_x (under 0.25 g/hp hr) are almost attainable with the high-

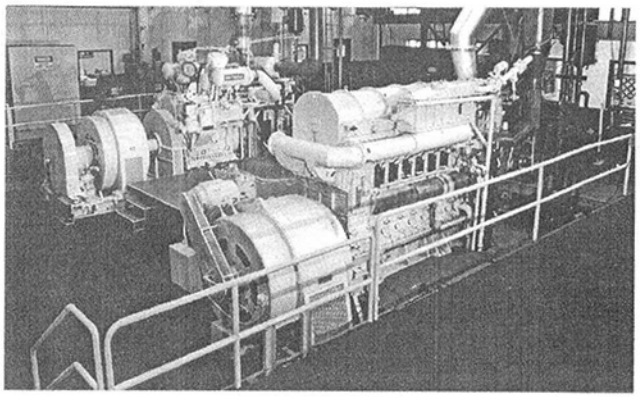


Fig. 4 Two-cylinder turbocharged engine

Table 3 PC2.6 dual-fuel engine ratings

HP/CYL	% PILOT	NO_X UNITS	BSFC UNITS	BMEP UNITS
600	5	5	5980	262
600	1	1	5920	262
650	1	1	5900	284

pressure gas injection because of diesel-like combustion, the compression horsepower needed for compressing fuel gas to 3500 psi, and the problems of multistage gas compression into the hypercritical region. Carburation also has significant drawbacks, not the least of which are having to reduce valve overlap greatly, which results in much higher combustion chamber, exhaust valve and exhaust system and preturbo temperatures, operation closer to detonation, and lack of load response or throttle response. Associated with either the high-pressure gas injection or carburation approach is the obvious safety objection as well.

2 Design and Fabrication of a Two-Cylinder Prototype. In 1987 it was decided to proceed with the new dual-fuel engine and also that a prototype was necessary for development testing. The two-cylinder engine configuration was chosen over a single cylinder because of the existence of a design for a nonoperational two-cylinder engine. An in-line six lacked the flexbility and ease of modification and the two cylinder has lower operating costs.

In 1987 the project commenced with modifying the existing nonoperational two-cylinder design developed and to provide turbocharging. The challenge of turbocharing a high-output, irregular firing two-cylinder engine was handled successfully. The turbo chosen was a MAN NA34 with a two-hole exhaust inlet. The turbo was fitted with the capability to assist the boost with shop air, an option not required in testing to date. The dual-fueled engine configuration is shown in Fig. 4 and was first operated on gas in December of 1988. The first few months of operations of the two-cylinder engine were used to "calibrate" its performance to full-size PC2.5 and PC2.6 engine performance in diesel mode and PC2.3 full-size dual-fuel engine performance. By mid-1989 performance parameters required for operation on gas above 600 hp/cyl at thermal efficiencies above 43 percent were being established. Since starting development testing in mid-1989, three new PC2.6DF engine ratings have been established as shown in Table 3.

3 Baseline Testing. By late 1989, development testing of the first PC2.6F engine model characterized by 600 hp/cyl with DEMA 10 percent overload to 660 hp/cyl (288 BMEP) has been completed. Fuel consumption on the two-cylinder engine was consistently less than 5900 Btu/hp hr at 5 g/hp-hr NO_x levels and under 6200 Btu/hp hr at the 3 g/hp-hr NO_x

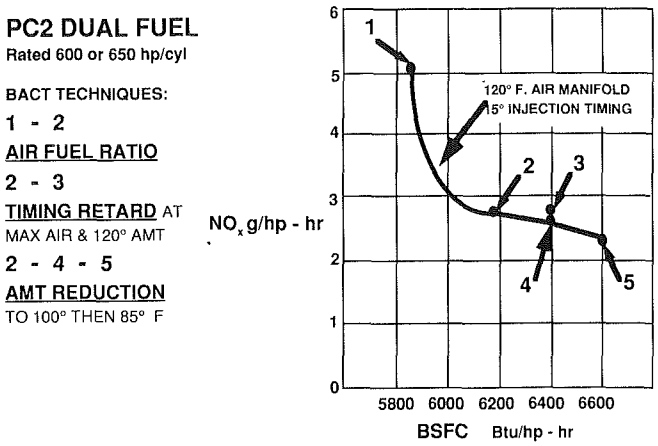


Fig. 5 Expected NO_x versus BSFC-BACT response

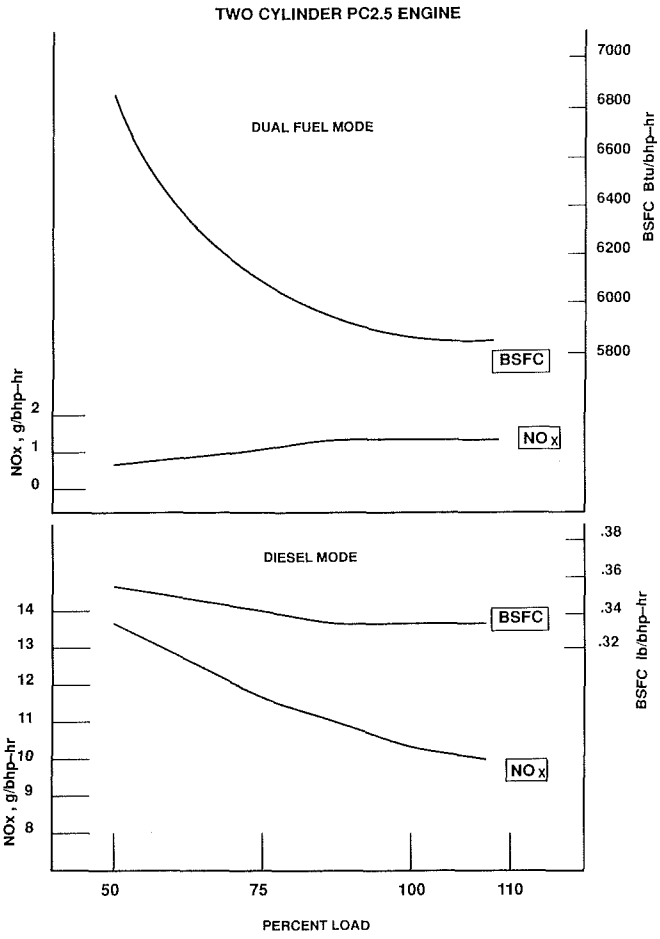


Fig. 6 Effect of derate

level. Response to the four Best Available Control Technique (BACT), techniques of lean air/fuel ratio, timing retard, air manifold temperature reduction and derate are shown by Figs. 5 and 6 with Fig. 5 showing the engine response to the first three BACT techniques and Fig. 6 showing the response to derate, all from the baseline 5 g/hp-hr NO_x level at 600 hp/cyl.

Note in Fig. 5 that going from point 1 to point 2 shows considerable effect of learning the mixture in gas mode on NO_x and BSFC. Once the advantage of the lean mixture is taken then little benefit is gained from retard shown by path 2 to 3 or by air manifold temperature reduction of 120° to 100° to 85°F shown by path 2 to 4 to 5.

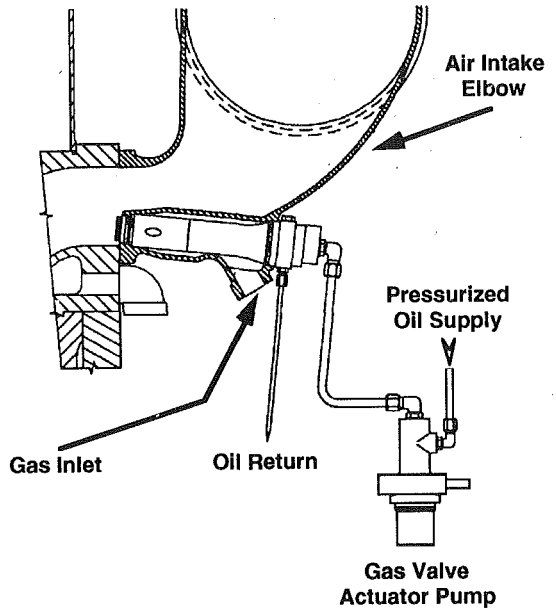


Fig. 7 Gas valve location

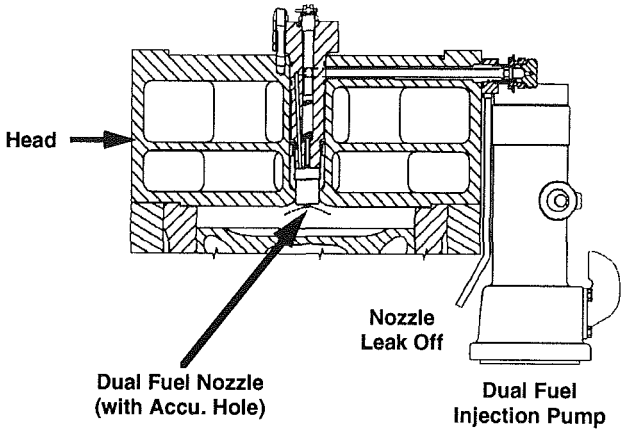


Fig. 8 Head cross section

Figure 7 shows the gas valve design in more detail and Fig. 8 shows the very conventional cylinder head layout for this version of the 600 hp/cyl engine. The cylinder head and fuel injection system are identical to the diesel version of the engine except for a small accumulator volume added to the injector body to control cavitation erosion of the needle and seat at pilot injection quantities (5 percent pilot oil levels).

Two 18-cylinder engines of this first configuration have been factory tested and delivered to Bay City, Michigan. Factory test guarantees of 5980 Btu/hp-hr at 5 g/hp-hr NO_x levels were easily met by these first two full size engines. The first engine at Bay City is expected to be on line by Feb. 1993 and the second by June 1993.

In 1989 we began performance testing on the lab two-cylinder engine to establish operating parameters and engine settings required for a rating of 650 hp/cyl continuous (284 BMEP) with DEMA 10 percent overload rating at 715 hp/cyl (312 BMEP). Several hundred hours of operation were successfully accomplished at this rating. In the configuration shown by Fig. 7 and 8 at 650 hp/cyl base rating, the NO_x is 5 g/hp-hr at 5900 Btu/hp-hr fuel guarantee on a full-size engine. Engine response to BACT is as shown by Figs. 5 and 6.

In 1988, work on a prechamber version of the 600 hp/cyl engine utilizing reduced pilot oil rate was begun. This reduced

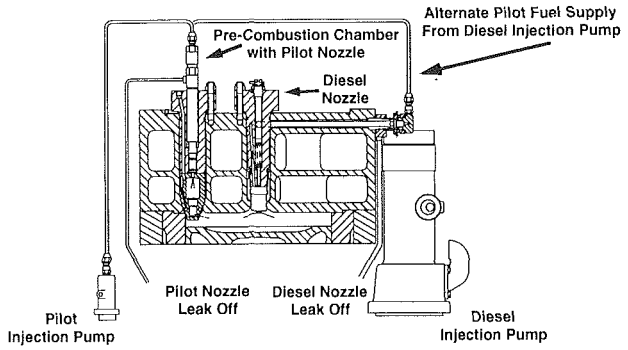


Fig. 9 Early design prechamber head

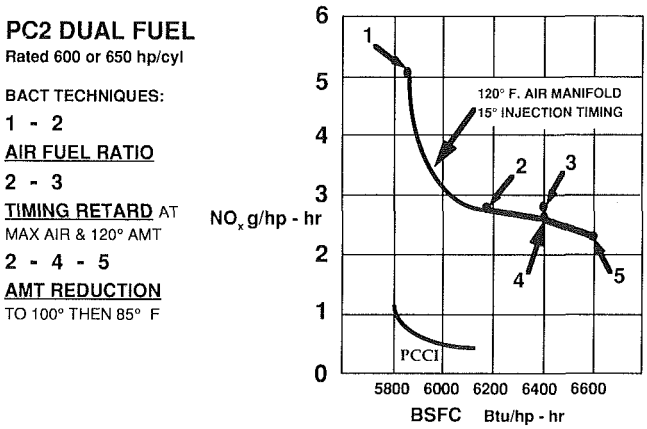


Fig. 10 Prechamber performance

pilot oil activity was undertaken because of the NO_x advantage of reduced pilot as suggested by Karim (1983) and Beshouri and Workman (1990). As an aside, the work of Blizzard et al. (1991) occurred in parallel with that discussed here. The first version of the prechamber with reduced pilot oil cylinder head is shown by Fig. 9. Note that two separate oil injection systems are shown. This is done because the conventional diesel injection system is limited to a turn down of about 5 percent pilot oil rate and we wanted to operate at 1 percent and less. The large fuel pump is a 30-mm-dia plunger and the small pump is 12.7-mm-dia by comparison. The large injector has 10 holes of 0.55 mm diameter and the small injector is an R size pintle. The prechamber volume is sized on the basis of stoichiometric combustion for a 1 percent pilot. In operation, the engine used only the large oil injection system in diesel mode and both the small and large oil system in gas mode. Since the large injector is water cooled there is no chance for fuel varnish buildup causing sticking of the needle while shut down for prolonged periods of gas operation. In the early stages of the design of the prechamber close attention was paid to achieving good cooling. A cooling circuit in parallel with the main diesel injector is used. Cooling is via two annular cavities and multiple drilled holes communicating between them. The cooling design is so effective that material of construction is low carbon steel and not the high temperature materials usually required in the design of prechambers.

The performance of the fully developed first version of the prechamber PC2.6DF is shown by Fig. 10 and labeled PCCI for "Precombustion Chamber Compression Ignition." This shows a comparison to the PC2.6DF performance from Fig. 7. Based upon these two-cylinder results we began commercial offering of 5920 Btu/hp-hr BSFC with 1 g/hp-hr NO_x, 600 hp/cyl continuous at 1 percent pilot oil.

4 Design Refinements. During the early running of the PCCI on the two-cylinder engine NO_x levels and fuel con-

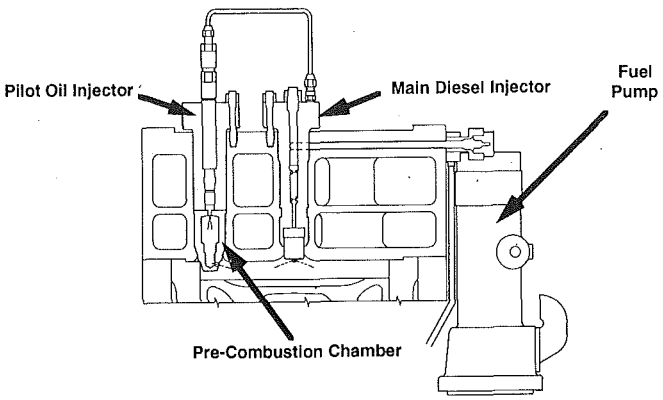


Fig. 11 Latest design PCCI

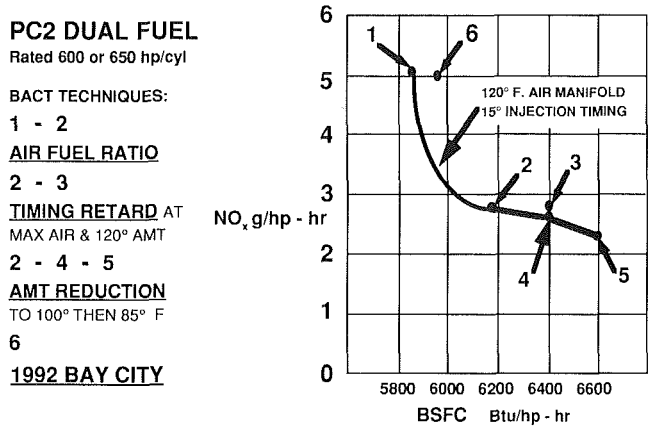


Fig. 12 Full-size engine performance

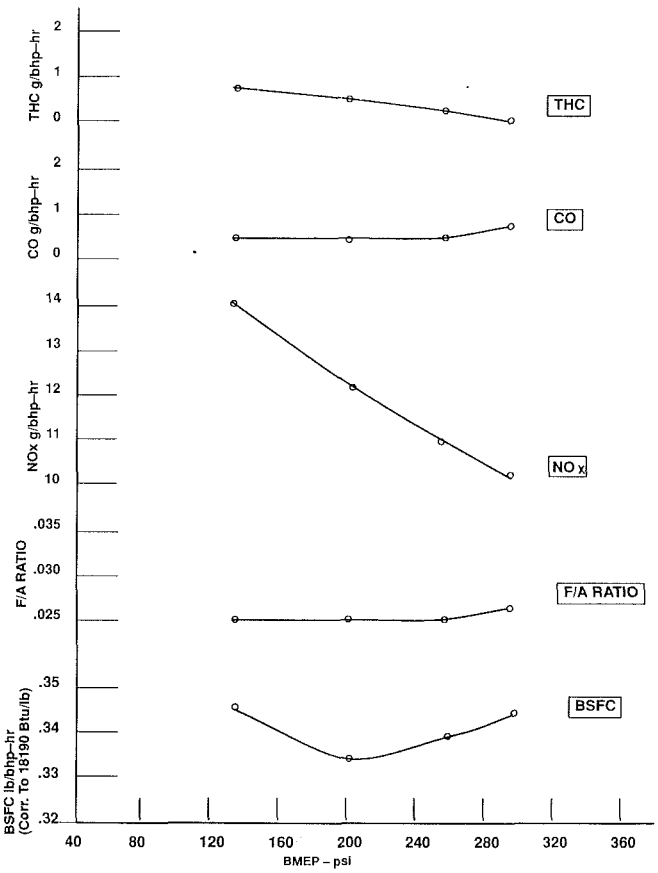


Fig. 13 Diesel performance 18 PC2.5DF; no prechamber, 5 percent pilot, diesel mode

sumption levels achieved were impressive. This encouraged us to take a close look at improving and simplifying the rather complicated design characterized by two fuel pumps, two fuel injectors, and an experimental electro hydraulic gas valve. The small cam-actuated jerk pump shown in Fig. 7 for actuating the gas injector was replaced with a small diesel fuel pump, which provided the reduced pilot oil to the injector in the PCCI. The small diesel pump referred to above is called the pilot pump in Fig. 9. For developmental testing, the two-cylinder setup as just described was quite acceptable but needed refinement for a production design.

The stage was set for two events to happen. The first was the development of a new diesel fuel system to eliminate the small pilot pump and have the main pump operate both the main diesel and pilot injector at each cylinder location. The second event was the design and development of a simple and reliable electrohydraulic gas valve, which would eliminate much of the mechanical control equipment on the engine, i.e., lay shafts, rod ends, bell cranks associated with the existing Fig. 7 hydraulic gas valve. The electronic gas valve in undergoing final engine and bench testing at the present time. It will be incorporated into future engines.

The diesel injection system development has been completed and will be provided on the first PCCI engine now under construction. Figure 11 shows the head cross section with the new one-pump two-injector system. Fuel pump operating pressure output and difference in opening pressure on the two nozzles determine whether either the pilot injector or both injectors are operating. A small shuttle valve in the pilot injector meters the pilot quantity independent of the main pump at pilot setting.

After successful operation of the two-cylinder engine utilizing the head configuration of Fig. 11 and gas valve configuration of Fig. 7 several hundred hours of endurance testing at 100 and 110 percent load were completed with no adverse events.

5 Commercial Model Testing of Full Size Engines. In Section 2 we described how we "calibrated" the turbocharged two-cylinder engine against full-sized engine performance. The only thing that the two-cylinder engine cannot do very well is duplicate parasitics, duplicate exhaust manifold pressure conditions and afford as broad a range of turbo performance as does a full size in-line or Vee engine. Figure 12 shows as point 6 how our first and second 18-cylinder engine compared to our 2-cylinder engine performance at full load. The results were very gratifying in terms of both fuel consumption and NO_x having been guaranteed at 5980 Btu/hp-hr BSFC and 5 g/hp-hr NO_x. These engines were 5 percent pilot oil no-prechamber engines built and shipped to Bay City Michigan for a commercial power plant application.

Performance at part load for the 18-cylinder engines set up to produce 5 g/hp-hr NO_x at full load on gas is as shown in Fig. 13 and 14 for diesel and gas mode, respectively. These data are representative of either of the two engines built and is test floor data with a "green" engine. Fuel rate would be expected to drop another 1 percent or so after a few hundred hours.

Summary

In summary, the PC2.5 dual-fuel engine has been built upon the success of its PC2.3 predecessor, borrowing many of its design features such as the jerk pump used with the hydraulic actuation system for the gas valves. An important development was the location of the gas valve in the intake manifold out of the cylinder head. This is a much less hostile environment for the gas valve and one that promotes extreme homogeneity of the mixture required for low-NO_x, lean combustion. The

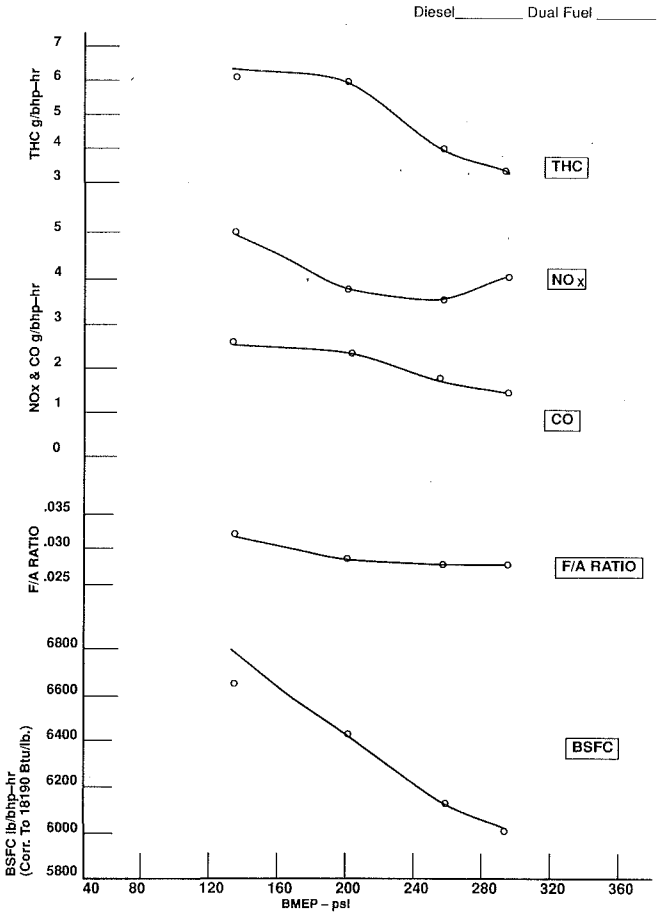


Fig. 14 Dual-fuel performance 18PC2.DF; no prechamber, 5 percent pilot, gas mode

structural limits of the PC2 series diesel engine, upon which the dual-fuel version is based, are well in excess of the loads imposed upon it by the gas combustion, resulting in a very conservative design.

Much attention has been given to cooling of the main injector and precombustion chamber to avoid not only the need for high-temperature materials of construction but also fouling. The fuel system is extremely simple and economic allowing one pump to operate the two injectors reliably with switchover built into the hydraulics. Retrofit of the system also has been considered in the design and development. Other than fitting of the extra cam, conversion consists of only minor adaptations of bolt-on components and tubing. The new PC2.5 engine promises a new era in dual-fuel engines characterized by high horsepower, excellent BSFC, and dramatically reduced emissions.

Acknowledgments

The author would like to thank Fairbanks Morse for support of this interesting development work and for permission to publish the results. I would also like to acknowledge the guidance and detail work of Gene Kasel and Frank Aboujaoude of the Fairbanks R & D lab and Hans Stauffer of Fairbanks for his assistance with engine layout, gas valve, and injection equipment design.

References

Beshouri, G. M., and Workman, J., 1990, "Single Cylinder Testing of a High Pressure Electronic Pilot Fuel Injector for Low NO_x Emission Dual Fuel Engines. Part II," ASME Paper No. 90-ICE-31.

Blizzard, D. T., Schaub, F. S., and Smith, J. G., 1991, "Development of the Cooper Bessemer Cleanburn Gas-Diesel (Dual Fuel) Engine," ASME ICE-Vol. 15, p. 89.

Karim, G. A., 1983, "The Dual Fuel Engine of the Compression Ignition Type—Prospects, Problems, and Solutions—a Review," SAE Paper No. 831073.

An Opposed-Piston Diesel Engine

J. K. Parker

S. R. Bell

D. M. Davis

Department of Mechanical Engineering,
University of Alabama,
Tuscaloosa, AL 35487

Typical conventional diesel engine designs are based on arrangements of single piston and cylinder sets placed sequentially either in-line or offset ("V") along the crankshaft. The development of other engines, such as the opposed piston type, has been motivated by potential advantages seen in such designs, which may not be viable in conventional in-line or V engine arrangements. Several alternatives to conventional engine design have been investigated in the past and some aspects of these designs have been utilized by engine manufacturers. The design and development of a proof-of-concept opposed piston diesel engine is summarized in this paper. An overview of opposed-piston engines is presented from early developments to current designs. The engine developed in this work is a two stroke and uses four pistons, which move in two parallel cylinders that straddle a single crankshaft. A prechamber equipped with a single fuel injector connects the two cylinders, forming a single combustion chamber. The methodology of the engine development process is discussed along with details of component design. Experimental evaluations of the assembled proof-of-concept engine were used for determining feasibility of the design concept. An electric dynamometer was used to motor the engine and for loading purposes. The dynamometer is instrumented for monitoring both speed and torque. Engine parameters measured include air flow rate, fuel consumption rate, inlet air and exhaust temperatures, and instantaneous cylinder gas pressure as a function of crank position. The results of several testing runs are presented and discussed.

Background

An opposed-piston engine is one in which two pistons oppose each other in each cylinder, each piston receiving the combustion forces and transmitting those forces to an output shaft; see Fig. 1. The method of coupling the forces to a common output shaft has been a major difference among the many designs attempted. There are several distinct benefits in the basic design of opposed-piston engines that have caused a great deal of interest in their development. They are typically very well balanced, mostly due to the symmetry of design, which establishes similar masses moving in opposite directions at all times. By incorporating piston-controlled intake and exhaust ports, the opposite-piston design uses fewer moving parts than conventional engines while providing an excellent means of uniflow scavenging. The total stroke of an opposed-piston engine is divided between two pistons; thus, piston speed and piston and cylinder wear are greatly reduced. Engine life is, therefore, increased while maintenance efforts are decreased. These and other advantages make the opposed-piston more efficient than its single-piston counterpart, both thermodynamically and mechanically, and thus opposed-piston engines are often smaller and lighter than conventional engines with comparable power ratings.

Although still considered a novel concept, the opposed-piston principle for diesel engines was introduced not long after the invention of the diesel engine itself. Opposed pistons were first used experimentally in early steam engines, and then in several gas engines beginning in 1874 [1]. Several inventors patented opposed-piston gas engines in England during the 1880s, but Wilhelm von Oechelhauser and Hugo Junkers are credited for the original development of these engines due to extensive work done at the Experimental Station for Gas Engines in Dessau. There, in 1892, they developed the first two-stroke opposed-piston engine. Oechelhauser and Junkers also worked independently on opposed-piston engines, resulting in the patent of the Oechelhauser gas engine in 1896 and of the Junkers gas engine in 1901 [1]. The concept was quickly applied to diesel engines by several researchers, including Junkers.

In 1907 Junkers developed his first opposed-piston oil engine using the design of the original Oechelhauser and Junkers gas engine patented in 1892. After several major changes, the standard Junkers design was established and was used extensively by Junkers and others in the years that followed [1]. Although others were unsuccessful with this design, Junkers continued to improve and produce opposed-piston engines, and in 1950 the Junkers "Jumo 205E," the powerplant of

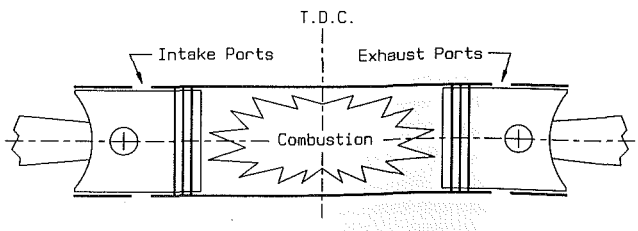


Fig. 1 Schematic of opposed piston engine

Contributed by the Internal Combustion Engine Division and presented at the 14th Annual Fall Technical Conference, Waterloo, Iowa, October 4–7, 1992. Manuscript received by the Internal Combustion Engine Division July 1, 1992. Associate Technical Editor: J. A. Caton.

many WWII German aircraft, had the highest thermal loading on the piston of any engine in production [2, 3].

Inspired by moderate success in the automotive and aircraft industries during the early 1900s by Gobron Brillié, Faccioli, Lucas, and Junkers, British marine engineers began to develop large diesel engines of this type for use in the shipping industry [1, 2]. The opposed-piston engine was found to be well suited for marine applications due to excellent low rpm operation and balancing, as well as its reduced size and weight. Although somewhat delayed by WWI, both Doxford and Fullager were successful in designing large opposed-piston engines [4]. Each of these engine manufacturers used very different concepts in coupling the pistons to establish a common output shaft. The Doxford engine, originally designed by Karl Otto Keller for Doxford and Sons, of England, quickly became a very popular powerplant for the shipping industry. Several Fullager engines were produced during the early 1920s by Cammell Laird and Co., but this manufacturer discontinued production due to design problems. The English Electric Co. also produced Fullager engines and in 1931 redesigned the engine and continued to make additional improvements, resulting in a very dependable engine, which was used for power generation in numerous electric power plants. The available literature indicates that The English Electric Co. was still producing Fullager Engines in 1946 [1].

Sulzer Brothers quickly became a leader in opposed-piston research and development with a new design, which incorporated rocker arms on each piston, which relayed the piston forces to a crankshaft located centrally beneath them. An advantage of this design is that the port timing is easily controlled by the offset of the crankshaft beneath the rocker arms [5]. This design feature was used by several other engine builders, but no others were as successful as Sulzer in finding a market for their new engines.

Sterling Engine Co. of Buffalo, NY, produced a Michell design crankless opposed-piston diesel engine in 1934 [6]. This engine used the angle disk plate principle, which reduces the internal moving parts of this unique engine. Due to the absence of piston side thrust from a crankshaft and connecting rod, cylinder wear was greatly reduced in these engines. Also, this design allowed the combustion forces on the pistons to be transmitted directly to the output shaft, negating the mechanical losses usually incurred with a crankshaft and connecting rod. These facts made this small engine very dependable with few maintenance requirements, and also extremely well balanced. It, therefore, found widespread use with stationary generators, compressors, etc.., as well as some marine applications [1], and was produced for several years.

Certainly the most successful opposed-piston engine producer in the U.S. is Fairbanks-Morse, who began producing opposed-piston diesel engines for U.S. Navy submarines during the late 1930s and continues to produce engines of similar design today. The advent of WWII caused an increased demand that prevented them from entering the commercial market until 1944, but by then the F-M38D8-1/8 was well proven and readily accepted [7]. It was used extensively in commercial shipping and also found application in several locomotives as well as stationary applications. The success of this large opposed-piston diesel caused a demand for a smaller, more portable engine of similar design, which Fairbanks-Morse met with the 38F5-1/4 in 1953 [8]. Although slightly different in design, these smaller opposed-piston engines proved just as dependable as their predecessors, and were extensively used as portable generators and compressors.

In 1958, Fairbanks-Morse found that installation of a turbocharger on the larger engines produced a 50 percent increase in output while decreasing specific fuel consumption. Also, use of a turbocharger nearly eliminated high-altitude power losses experienced by normally aspirated engines. These improvements were made with minimal increases in size and weight of the powerplant [9]. Turbocharging was thus designed into F-M's next generation of opposed-piston engines. The model 38A20, announced in 1965, used an upper (exhaust) piston that was much smaller than the lower piston. This engine was available in 6 and 9 cylinder in-line arrangements, and in a V-12 arrangement, all turbocharged. It had two crankshafts, geared together, the upper shaft being driven by a rocker arm and connecting rod from the exhaust piston. This design produced 1000 hp per cylinder, and found substantial application in stationary power generating plants [10].

Harland and Wolff, Ltd., of Belfast introduced a new 7500 bhp opposed-piston marine diesel engine in 1950. This heavily researched engine employed a single crankshaft with eccentrics attached to a crosshead through connecting rods. It also utilized two specially designed rotary blowers to supply its seven cylinders and had two opposing injectors in each cylinder to provide optimum combustion [11]. The Harland and Wolff engine was used in several ships and also in many stationary power plants during the next few years.

During the 1960s, Professor F. J. Wallace et al. performed several experiments on opposed-piston engines, the first of which was an assessment of high boost operation of the Commer TS-3. Excellent results were experienced at a boost pressure ratio of 3/1, with brake thermal efficiencies as high as 40 percent. Later, Wallace performed a comparison of two and four-stroke conventioned diesel engines with an opposed-piston two-stroke diesel by computer modeling. Conclusions were that the opposed-piston engine offered more power per size, weight, and cost than the conventional engines, while operating more efficiently and with lower thermal stresses [12].

The many previously stated advantages of opposed-piston engines certainly make it a legitimate avenue of research. The success of several manufacturers indicates that increased implementation of these engines is likely, while the failure of others indicates a need for continued development.

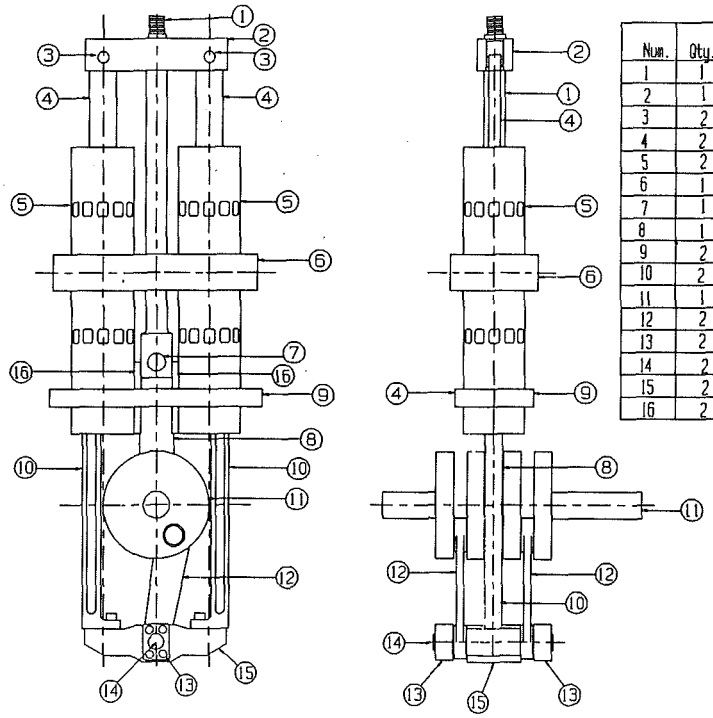
The Laitram Corporation currently has three U.S. [13–15] and several foreign patents for an opposed piston internal combustion engine. The engine is a two-stroke diesel, which uses four pistons, two cylinders, and a single crankshaft in its minimum configuration. Two opposed pistons work in each of the cylinders, which straddle a common crankshaft. Upper and lower crossarms connect the four pistons to the crankshaft. The two cylinders are connected by a common precombustion chamber, which further serves to balance the loads between them. Since the intake and exhaust timing are controlled by ports, no complicated valve arrangements are necessary. The current configuration of the "H-4" engine is shown in Fig. 2.

The University of Alabama has been developing a "proof-of-concept" H-4 engine since October 1988. A computer model of the kinematic, dynamic, and thermodynamic properties of the engine was initially developed to serve as a design tool [16]. After several of the major engine parameters had been selected, engine components were designed and manufactured. With the exception of the pistons and fuel injection system, all engine components were fabricated at The University of Alabama. Final assembly and preliminary testing of the "proof-of-concept" engine began in October 1989.

Project Description

The overall goal of the project has been to develop a proof-of-concept engine that demonstrated technical feasibility. Minimal effort in optimizing component design was made for this initial proof-of-concept engine. The proof-of-concept engine does not include optimized design nor performance, but serves as an evaluation of the H-4 engine concept. The development project was conducted in two phases. The first phase of this work had the following objectives:

1 Background development; thermodynamic analysis, dynamic analysis, establish major parameters;



Num.	Qty.	Part Description
1	1	Center Connecting Rod
2	1	Upper Piston Bridge
3	2	Upper Piston Rod Wrist Pin
4	2	Upper Piston Rod
5	2	Cylinder Liner
6	1	Center Cylinder Mounting Bracket
7	1	Center Connecting Rod Wrist Pin
8	1	Upper Connecting Rod
9	2	Lower Cylinder Mounting Bracket
10	2	Lower Piston Rods
11	1	Crankshaft
12	2	Lower Connecting Rod
13	2	Lower Bridge Bearing Block
14	2	Lower Connecting Rod Wrist Pin
15	2	Lower Piston Bridge
16	2	Lower Connecting Rod Bearing Pad

Fig. 2 Major components of H-4 engine

2 Conceptual design, component design/identification, final design;
3 Component manufacturing/acquisition, engine assembly;
4 Component testing and engine instrumentation.

The majority of tasks in Phase I work were completed on schedule, although component testing extended beyond the proposed schedule.

Phase II of the project was initiated during the spring of 1990. The overall objectives of this second phase were to:

1 Develop information through proof-of-concept engine testing for evaluating the technical feasibility of the H-4 engine,
2 Identify critical components for redesign and incorporate upgraded components into proof-of-concept engine, and
3 Identify overall design concepts for next generation proof-of-concept engine.

The goal of the second phase was to provide information for evaluating the H-4 engine feasibility. This information was developed primarily from engine testing; however, some extrapolation of these data was necessary in order to account for the nonoptimized condition of the engine.

Engine Design Features

Figure 2 shows a view of the current engine, which is a two-stroke design and uses four pistons, which move in two parallel cylinders that straddle a single crankshaft. An open frame was selected due to ease of fabrication and to allow modifications; see Fig. 3. No castings or forgings were used in the design; all components are machined using lathes and milling machines. Aluminum was used extensively throughout the engine, primarily because it was much easier to machine. Oil lubrication was used extensively for all bearing surfaces, but no attempt was made to collect and recycle the oil with a closed-loop system. A standard four-cylinder Lucas fuel injector pump was modified and used (with three of the outputs disabled) with a single pintle type fuel injector to feed the prechamber connecting the two cylinders. Standard roller bearing pillow

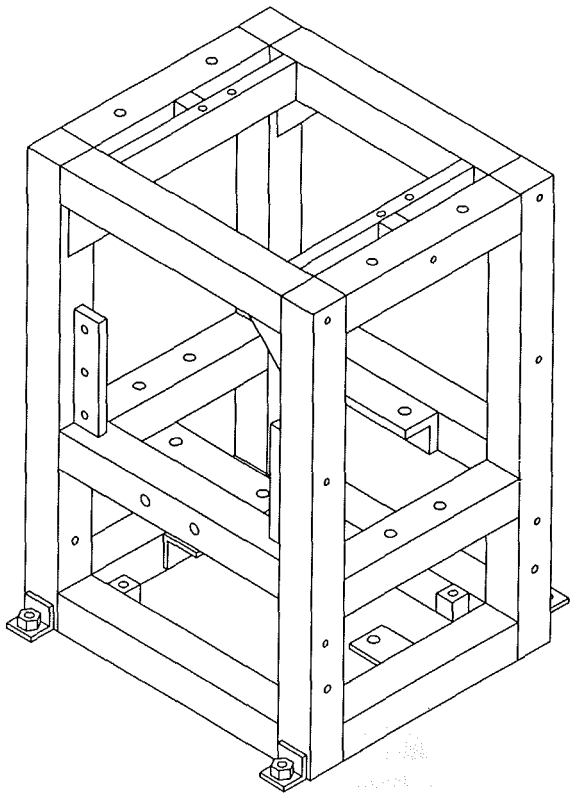


Fig. 3 Open engine frame

blocks were used as the main journals, due to the relatively low loading inherent in the design. Starting and motoring were provided by a General Electric dynamometer. It is important to note that the current design of the engine was intended to prove the concept, not to develop a production or even prototype opposed-piston engine.

The prechamber is removable and connects tangentially to

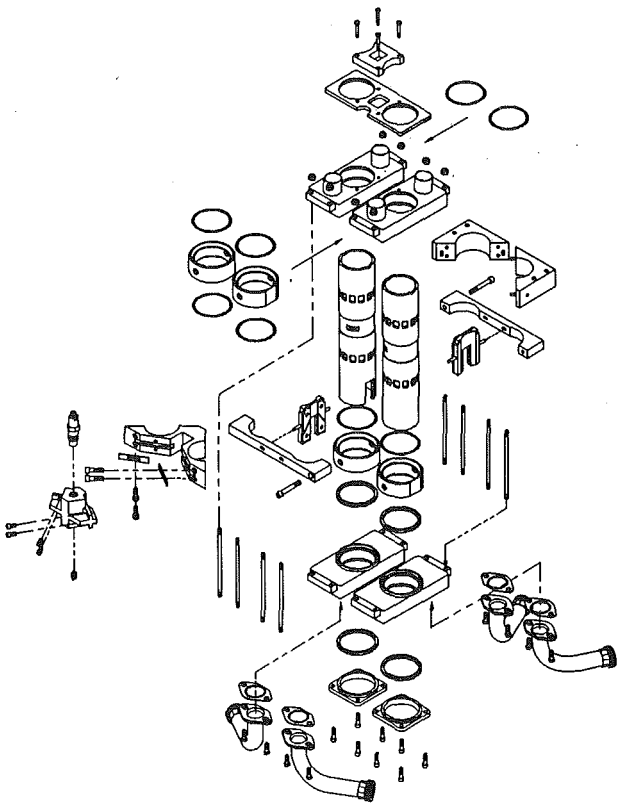


Fig. 4 Exploded view of H-4 engine

each of the liners through two short transfer ports. The upper two pistons connect to an upper bridge (item 2 in Fig. 2) and the bridge to the center connecting rod (item 1 in Fig. 2), which transfer the combustion force to the upper crankrod and crankshaft. Two sets of bearing supports are used to constrain the movement of the center connecting rod. The lower set of pistons are connected to a lower bridge, which is also constrained to vertical movement using a linear bearing arrangement. The lower bridge is connected to the crankshaft using two crankrods, which connect to the two outside crank throws straddling the center crankrod and throw. The cylinder liners were manufactured from steel tubing bored to match the pistons, which were obtained from a commercial vendor. Intake ports were machined at the upper midsection of the liners and exhaust ports at the lower. Intake and exhaust manifolds slip over the liners for routing the intake air and exhaust. The engine is water cooled using two slip-on water jackets, which cover the outside portion of the liner between the intake and exhaust manifolds. The prechamber fits along the center of liners. Figure 4 shows an exploded view of the major components connecting to the liners.

The engine design offers several potential attributes over conventional inline or V engine designs, which served as motivation for the work. First, the opposed piston design should lead to lower net loads on the crankshaft. The axial loads exerted by the upper piston set due to gas pressure are counteracted by the axial loads from the lower piston set. Inertial loads due to the rotating and translating masses of the upper and lower piston sets are also canceled to some degree. The crankshaft bearings can, therefore, be physically smaller than for a conventional engine with similarly sized pistons. This leads to lower weight and less material cost for the engine block. Smaller crankshaft loads could also lead to less bearing wear, which reduces maintenance requirements.

Vibration forces generated by the engine should also be minimized when compared to conventional engines of the same size. If the lag angle between primary crankshaft throws is 180 deg, then complete balancing of all pressure and inertial loads would be possible. The "proof-of-concept" engine currently uses a 150 deg lag angle to facilitate scavenging of the exhaust products from the cylinder. Therefore, the primary and secondary shaking forces are not completely canceled, but can be reduced over conventional engine characteristics.

The side or thrust loading on the individual pistons of the engine are reduced by several orders of magnitude. The piston skirts are not required to act as bearing surfaces as in conventional engines. A much lighter weight piston could likely be used with greatly reduced wear on the pistons and rings.

The engine design has eliminated the cylinder head and associated valve equipment for a reduction in moving mechanisms to accomplished scavenging. The elimination of valves could lead to lowered maintainence requirements and reduced cost of manufacturing and assembly. The lack of a cylinder head has a potential benefit with respect to thermal performance as the surface area bounding the working gas is reduced. Improved thermal efficiency can translate to lowered fuel consumption for a given engine.

The scavenging process for the engine is a uniflow design, which is also desirable. Gas enters the cylinder through the upper intake ports and exits through the lower exhaust ports. Although most other opposed engine designs also share some of these advantages, they commonly use two crankshafts or are coupled by complicated external gearing or pushrod arrangements.

The minimal configuration for the H-4 engine uses a single fuel injector to feed four pistons. This reduces the cost and complexity of the fuel injection system for the engine. The large piston-to-injector ratio would also be advantageous in potential special applications, such as dual-fuel engines.

Initial testing of the engine identified excessive bending (buckling) in the lower piston rod, which was used as a linear bearing as well. A new lower piston rod with approximately the same weight but 2–3 times the bending strength was designed and machined. Also, a new one-piece lower bridge was designed to reduce weight and accept the new lower piston rods; see Fig. 5. The horizontal bearing surfaces for the lower bridge were moved from the piston rods to the central pins of the bridge. Better lubrication of the bearing and more precise positioning at the point of force generation were distinct improvements. The lower bridge pins accepted two bronze bearing pads, which were positioned between hardened steel plates forming the linear bearing for the lower bridge. This lower bearing was force lubricated and was rigidly attached to the engine frame and crankshaft bearing supports.

Facility Description

The H-4 engine was set up for testing in the Internal Combustion Engines Laboratory at The University of Alabama. Figure 6 is a schematic of the experimental arrangement, illustrating the instrumentation pertinent to the engine testing presenting in this paper. As shown, the engine was attached to an electric dynamometer, which was equipped with a scale to measure the loading force of the dynamometer, thus providing a means of calculating the brake engine torque produced. Engine speed was also monitored using a speed transducer located on the opposite end of the dynamometer.

The intake system included a surge drum and a roots type blower to provide air to the inlet manifolds. An ASME long radius nozzle was used in the inlet of the surge drum to obtain air flow rates. The pressure drop across the nozzle was measured with a U-tube mamometer and a calibration curve was used to convert the pressure drop to air flow rate. The fuel flow rate measurement was obtained by recording the time for consumption of a given amount of fuel through a small buret graduated in milliliters.

The fuel injector was fitted with a pintle lift indicator in

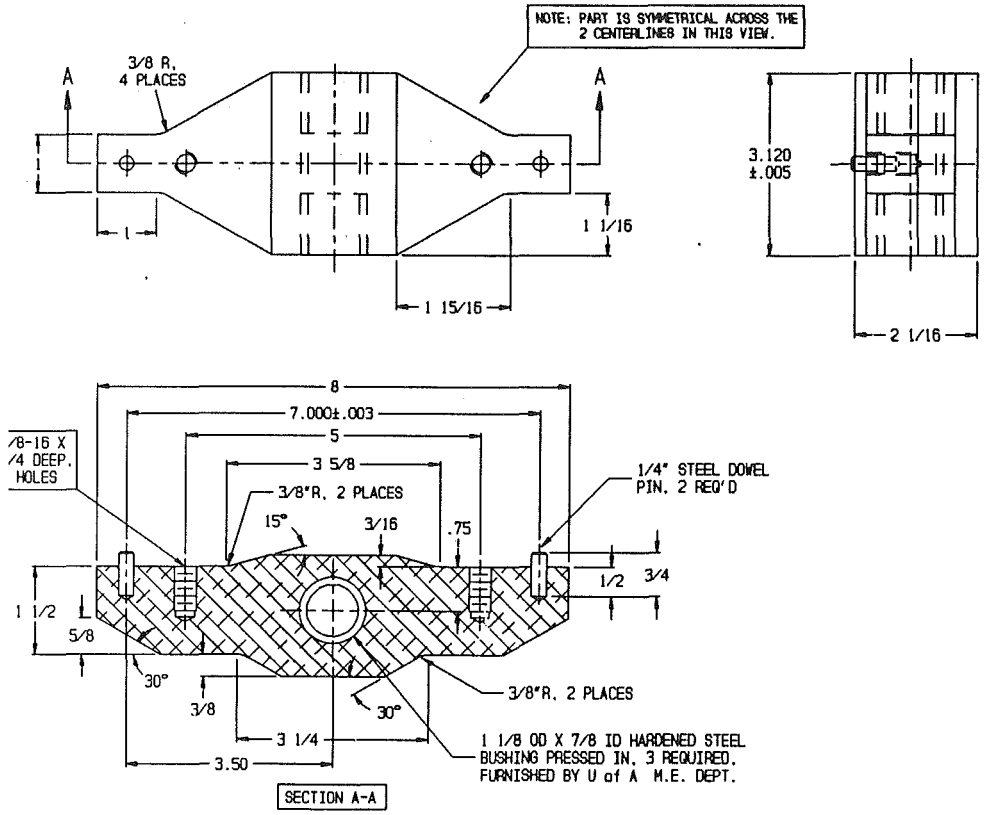


Fig. 5 H-4 lower bridge

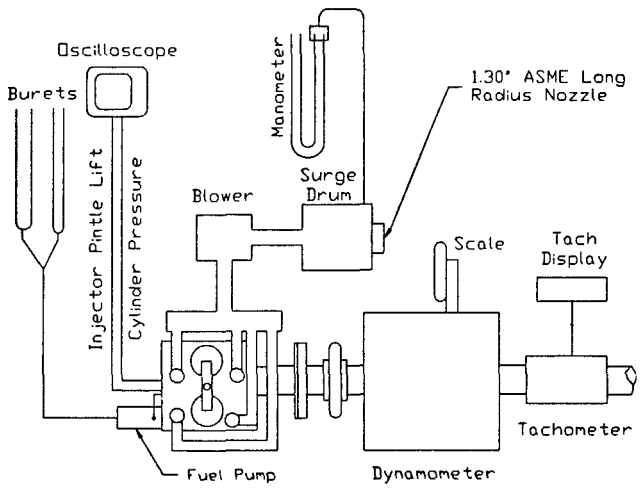


Fig. 6 Experimental test stand

order to monitor injector timing. A pressure transducer was installed in the prechamber through a port opposite the injector. The transducer allows measurement of the combustion chamber pressure, which allows more in-depth study of the combustion event. An encoder attached to the fuel pump shaft, which operates at engine speed, was used to record crankshaft position so that the pintle lift and cylinder pressure could be referenced to crankangle locations. Data from the pintle lift indicator, the pressure transducer, and the encoder were recorded using a digital storage oscilloscope. These data provided a means by which to explore the time history of various parameters within the combustion chamber, including pressure, heat release rates, injection timing, and combustion duration.

Testing Results

An initial firing test was conducted after all component upgrades and motoring tests were completed. The engine was motored to a speed of approximately 500 rpm by the dynamometer before diesel injection began. After the crack was engaged the engine began to run smoothly and accelerated to a speed of about 700–750 rpm before the fueling rate was decreased. The indicated power produced by the engine at this speed was in the range of 7–8 hp, which is approximately the same as the friction power.

Close inspection of the engine after the initial firing test uncovered two problems. The lower bridge used two cantilever pins to attach the connecting rods of the crankshaft. This design was not adequate and allowed the hardened steel bushings to separate from the aluminum block. Also, the addition of the hardened steel bushings left an inadequate cross section of aluminum to withstand the bending load applied as the engine fired. Consequently a small amount of bending was detected in the lower bridge. These two problems were detected before any damage to other engine components occurred. A minor redesign of the lower bridge to increase the aluminum cross section at the pins was completed and is shown in Fig. 5. A single pin was used to connect the lower bridge to the connecting rods of the crankshaft.

Figure 7 shows the brake torque and brake horsepower-output for six tests conducted after completing the required modifications discussed above. The results shown in the figure are selected data at which the equivalence ratio was approximately 0.36. Since the full-load equivalence ratio is expected to be higher, this figure represents a part-load condition. Engine speeds for these tests ranged from 550 to 660 rpm, and resulted in brake torque values of 25 to 35 ft-lbf with corresponding BHP values of 3 to 4 hp. As expected, the torque curve is essentially flat (due to the small range of test speeds) while the power curve increases with engine speed. The air

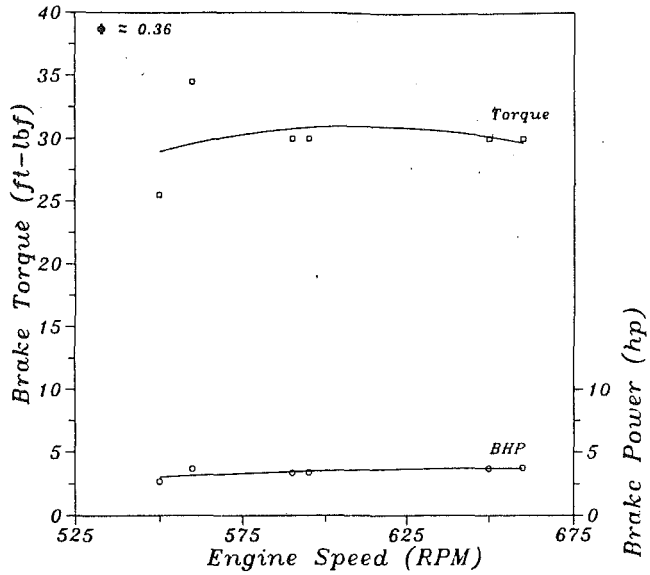


Fig. 7 Brake torque and brake power as a function of engine speed for equivalence ratio of 0.36

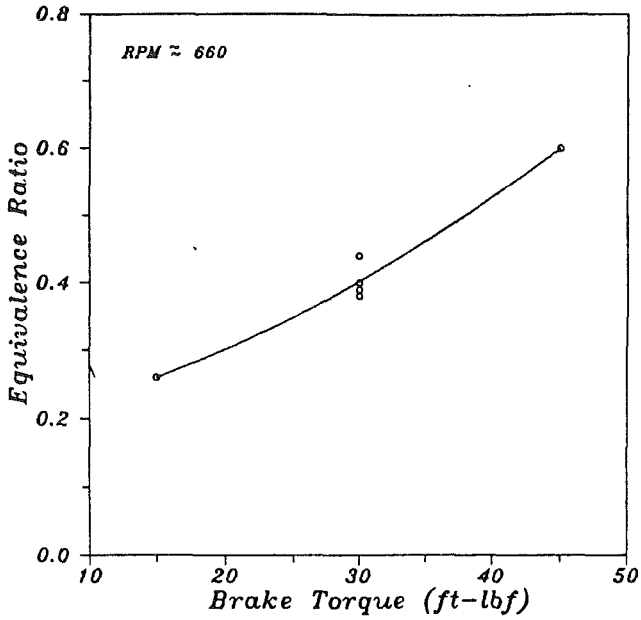


Fig. 9 Equivalence ratio as a function of brake torque at 660 rpm

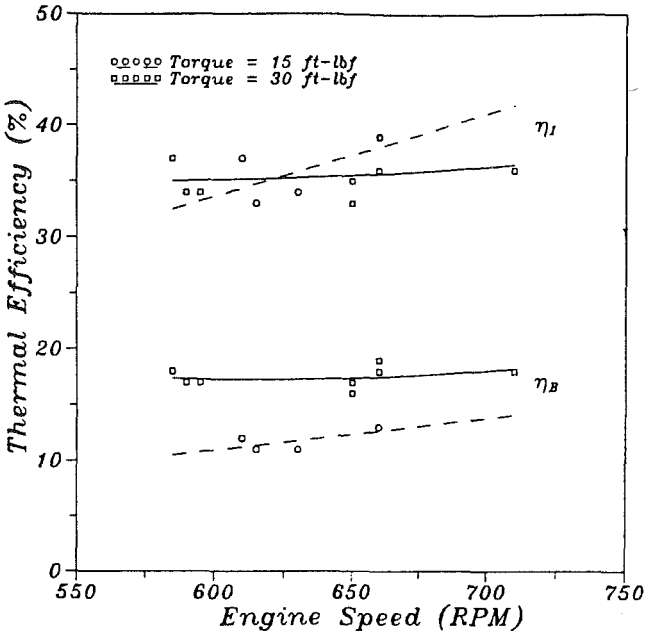


Fig. 8 Indicated and brake thermal efficiency for experimental and simulation results as a function of engine speed for 15 and 30 ft-lbf torque

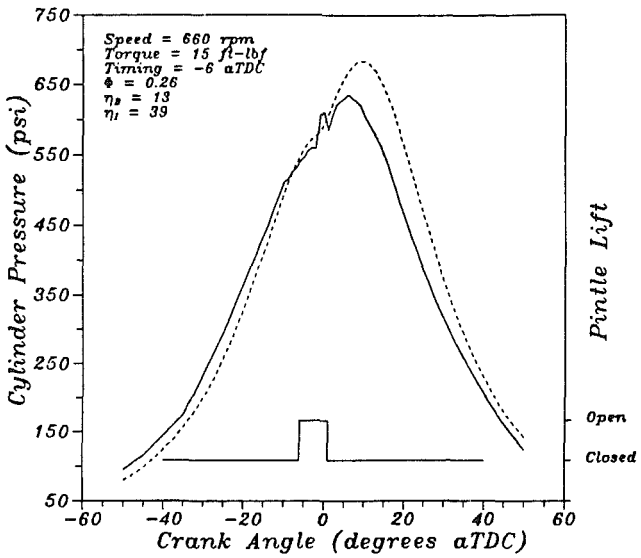


Fig. 10 Cylinder pressure and needle lift as a function of crank angle for experimental and simulation results for a torque of 15 ft-lbf

flow rate supplied by the external roots type blower was found to be strongly dependent on the operating speed of the engine and limited testing to a maximum speed of under 1000 rpm as a result of inadequate charging air at higher speeds.

In Fig. 8, brake and indicated values of thermal efficiency are plotted at torque values of 15 and 30 ft-lbf as a function of engine speed. The indicated fuel efficiencies are approximately equal for the two loads, while the brake values are considerably higher for the 30 ft-lbf curve. This is indicative of the fact that friction losses are relatively constant at any given speed, so increases in output beyond the friction load are reflected as increases in brake thermal efficiency. A single test at 45 ft-lbf of torque yielded a similar value of indicated thermal efficiency, while the brake value was only slightly higher than those given by the 30 ft-lbf curve. This would indicate that the best brake efficiency attainable with the current proof-of-concept engine is probably not much greater than the 30 ft-lbf efficiencies shown in the figure.

Equivalence ratio versus brake torque at 660 rpm is plotted in Fig. 9. The expected increase in torque with increased equivalence ratio is exhibited, as greater values of equivalence ratio correlate with a greater mass of fuel being injected, and thus more energy is available to convert to usable work (brake torque).

Figures 10 and 11 show cylinder pressure and needle lift as a function of crank angle position for the engine at 660 rpm for loads of 15 and 45 ft-lbf, respectively. Pressure data from a computer simulation of the engine are also shown on the figures. The pressures in the engine tests were measured with a transducer inserted in the prechamber. The simulation results shown for comparison are based on a single combustion chamber model or a direct injection model and, therefore, differences in the results are expected. A difference in comparison is particularly evident at ignition when the prechamber expe-

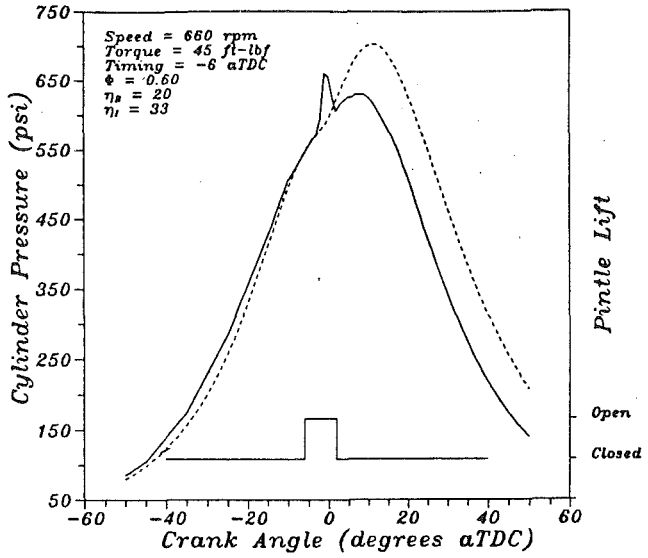


Fig. 11 Cylinder pressure and needle lift as a function of crank angle for experimental and simulation results for torque of 45 ft-lbf

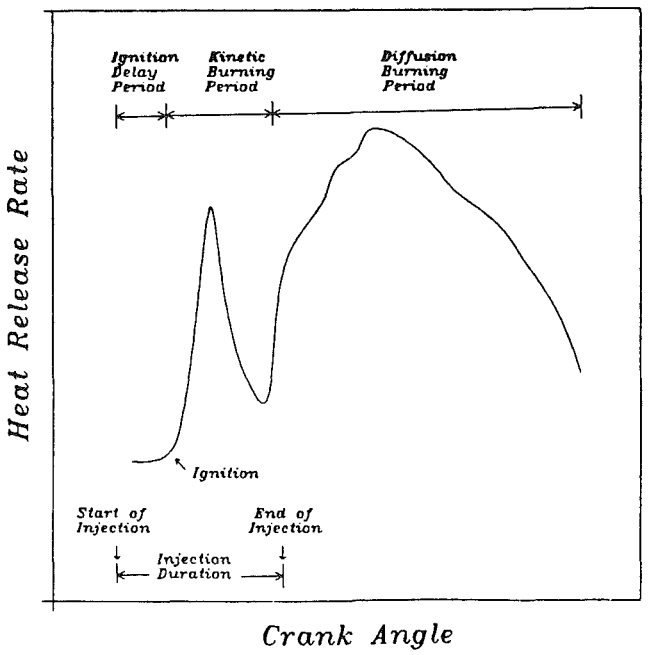


Fig. 12 Typical heat release rate schedule as a function of crank angle

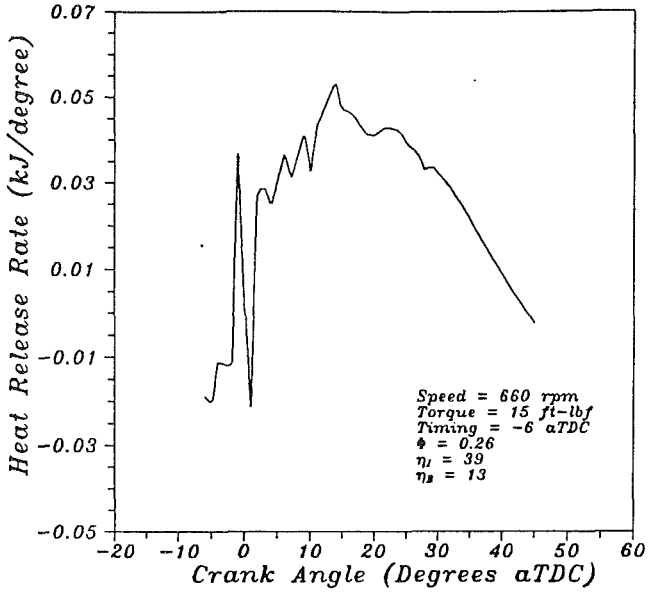


Fig. 13 Heat release rate as a function of crank angle for torque of 15 ft-lbf

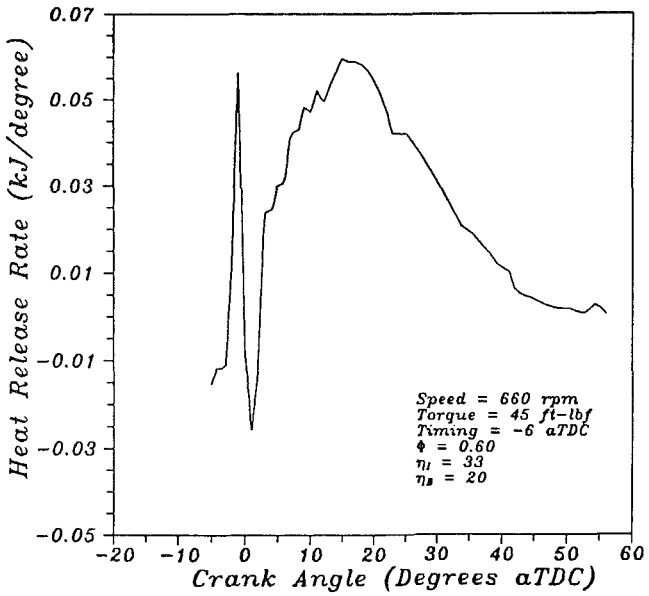


Fig. 14 Heat release rate as a function of crank angle for torque of 45 ft-lbf

riences a more abrupt pressure spike typical of prechamber engines.

For the results of Fig. 10, the injection start timing was −6 CA aTDC as measured by the position of the bottom pistons (the pistons have a lag angle and therefore top dead center is not reached at the same rotational angle for the top and bottom pistons) and the duration is indicated on the needle lift diagram. For this case, the equivalence ratio was 0.26, indicated fuel efficiency was 39 percent, and the brake fuel efficiency was 11 percent.

For the 45 ft-lbf case, the injection timing was −6 CA aTDC, equivalence ratio was 0.6, the indicated fuel efficiency was 33 percent, and the brake fuel efficiency was 20 percent. The injection duration was increased with the fueling increase for this case.

Further analysis of engine operation was made by estimating the instantaneous heat release rate of the burning mixture from the pressure data. To accomplish this, a simple one-zone computer program was written that uses the experimental pressure data as input and then using a first law of thermodynamics model of the engine and estimates what the necessary heat release was to produce the measured pressure. A typical heat release schedule is shown in Fig. 12 with three periods of combustion noted. An ignition delay period is shown, during which a portion of the injected fuel vaporizes and mixes with air prior to combustion, followed by a very rapid heat release rate. This rapid increase (called kinetic burning) is caused by the sudden ignition of the combustible mixture that has formed during the delay period, and is followed by a decrease in heat release rate as the amount of combustible mixture. Since injection continues through this period, combustion continues and a diffusion or mixing limited burning period follows. The heat release rates calculated from the measured cylinder pressure for the cases reported in Figs. 10 and 11 are shown in Figs. 13 and 14. As expected, slightly higher values of heat release rate for the 45 ft-lbf case compared to the 15 ft-lbf

case were found. This is expected, as the integration of these curves yields the total energy released from the mixture, which is proportional to the output of the engine. These figures show a strong pronounced kinetic burning peak, which resulted from the sharp rise in pressure seen in Figs. 10 and 11 followed by the diffusion or mixing-controlled period.

Summary

The design and development of a novel opposed-piston engine has been presented. The history of the opposed-piston engine, from its inception in the late 1800s to the present, has also been summarized. After several months of effort, the engine was successfully operated and shown to be capable of sustained operation. The results of initial testing have been presented and discussed. The proof-of-concept engine demonstrated reasonable power levels and thermal efficiencies (compared to expectations). Further efforts will be directed to more complete and thorough testing of the engine.

Acknowledgments

This work was supported by contract through the Laitram Corporation. The authors wish to express appreciation to J. M. Lapeyre, Jr., and his staff for their support, advice, and consideration during the difficult periods of this research.

References

1 Wilson, W. K., "The History of the Opposed-Piston Marine Oil Engine," *The Institute of Marine Engineers Transactions*, Vol. 58, Oct.-Nov., 1946.

2 Allen, O. F., *The Modern Diesel*, Prentice-Hall, NY, 1947.

3 Middleton, J. H. D., "The High-Speed Light-Weight Diesel Engine," Vol. 190, Nov. 17, 1950.

4 "The Opposed-Piston Principle for High-Power Diesel Engines," *Power*, Vol. 57, Jan. 9, 1923.

5 "Features of Diesel Engines With Opposed Pistons," *Power Plant Engineering*, Vol. 45, May 1941.

6 "Crankless, Opposed-Piston Diesel Engine," *Power*, Vol. 78, Feb., 1934.

7 Ruskin, P., "Novel Design Features of Fairbanks-Morse Opposed-Piston Engines," *Automotive and Aviation Industries*, Vol. 92, Jun. 15, 1945.

8 "F-M's Small OP Engine," *Diesel Power and Diesel Transportation*, Vol. 31, May, 1953.

9 "Turbocharging the O-P Engine," *Diesel Power*, Vol. 36, Mar., 1958.

10 "New Opposed-Piston Diesel," *Mechanical Engineering*, Vol. 87, Jun. 1965.

11 "Eccentric-Type Opposed-Piston Marine Diesel Engine," *Engineering*, Vol. 169, Jun. 9, 1950.

12 Smyth, R., and Wallace, F. J., "Comparative Performance Assessment of Various Compression-Ignition Engine Configurations in Combination With Compressors and Turbines," *Proc. Instn. Mechanical Engineers*, Vol. 181, 1966-67.

13 Lapeyre, J. M., "Internal Combustion Engines," United States Patent No. 4,543,917.

14 Lapeyre, J. M., "Internal Combustion Engine," United States Patent No. 4,566,408.

15 Lapeyre, J. M., "Internal Combustion Engine," United States Patent No. 4,732,115.

16 Bell, S. R., Parker, J. K., and Lapeyre, J. M., "Development of a Proof-of-Concept Opposed Piston Engine," ASME Paper No. 90-ICE-7.

Combined Air-Oil Cooling on a Supercharged TC & IC TAM Diesel Engine

F. Trenc
TAM Research and Development Institute,
Maribor, Slovenia

R. Pavletič
Department of Mechanical Engineering,
University of Ljubljana,
Ljubljana, Slovenia

In order to reduce the maximum cylinder wall temperatures of an air-cooled TC&IC diesel engine with large longitudinal and circumferential temperature gradients, a curved, squared cross-sectional channel supplied with engine lubrication oil was introduced into the upper part of the cylinder wall. Numerical analyses of the heat transfer within the baseline air-cooled cylinder and intensive experimental work helped to understand the temperature situation in the cylinder at diverse engine running conditions. The results of the combined cooling were greatly affected by the design, dimensions, position of the channel, and the distribution of the cooling oil flow, and are presented in the paper.

Description of the Problem

In order to establish the problem of inadequate temperature distribution in an air-cooled cylinder and consequently to solve it, analysis of the baseline cylinder temperature distribution must be completed first. The baseline cylinder was therefore equipped with more than 70 temperature sensors (thermocouples). Some temperatures served as temperature boundary conditions and some of them to check the accuracy of the numerical model of the heat transfer within the cylinder wall.

The position of the particular temperature sensors in the baseline cylinder wall is shown in Fig. 1; thermocouples defining the internal cylinder wall temperatures were distanced 1.2 mm from the cylinder working surface. Basic measurements were performed on a single and four-cylinder TC & IC air-cooled diesel engine at diverse engine running conditions, at well-stabilized loads, speeds, and temperatures. Results of measurements showed pronounced asymmetric circumferential temperature distribution, relatively high maximum cylinder temperatures and high longitudinal temperature gradients of the cylinder inner wall. Temperature asymmetry is mainly provoked by the influence of the intake and exhaust channel, by the shortages of the cooling surfaces design, and the physical properties of the cooling air. Figure 2 shows circumferential wall temperature distribution for diverse cylinder cross sections at full load and peak torque engine speed (according to the symbols in Fig. 1). Critical transverse section I, which corresponds to the top position of the first piston ring (when the piston reaches its TDC), determines most of the limiting temperature values of the cylinder wall. Circumferential asymmetry of temperature is defined as the biggest temperature difference in the particular cross section and reaches its critical peak value of nearly 35°C in section I. Wall temperature distributions in the longitudinal cross sections "3" (the hottest) and "8" (the coolest) are typical too; maximum average longitudinal temperature gradients in the upper part of the cylinder reach almost 1°C/mm. The influence of the entire cooling fin insulation of the lower half of the cylinder is interesting; wall temperatures in that region increase by approximately 6°C only, whereas the top cylinder area temperatures remain unchanged (Fig. 2).

On-field application of an engine is mostly characterized by sudden and frequent changes of speed and load. According to the results of the experiments performed, only a very thin layer of the cylinder wall reacts simultaneously, due to the thermal capacity of the cylinder material, to sudden changes of the engine thermal load. The "information" of the higher thermal load propagates slowly throughout the cylinder wall until it reaches the cooling fins. During the period of thermal stabilization, local temperature differences in the cylinder can reach even higher values compared to those at stabilized conditions and the highest loads. Local temperature differences can provoke additional thermal stresses and deformations, and can therefore influence the reliability and longevity of the engine.

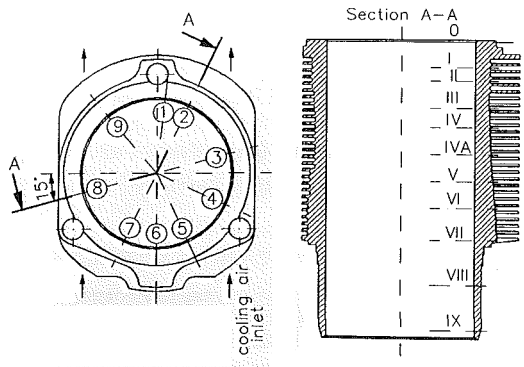


Fig. 1 Position of the thermocouples in the baseline cylinder wall

Contributed by the Internal Combustion Engine Division and presented at the 14th Annual Fall Technical Conference, Waterloo, Iowa, October 4–7, 1992. Manuscript received by the Internal Combustion Engine Division July 1, 1992. Associate Technical Editor: J. A. Caton.

Some typical engine load-time simulations have been performed on the laboratory test stand to determine their influence on the time–temperature distribution in the cylinder wall. Temperatures at diverse places were measured simultaneously throughout the stabilization period and approximated by a polynomial of the fourth degree for further calculations. It was shown by Trenc (1992) that the error introduced by this approximation was smaller than 1°C.

The main purpose of the physical model of the two-dimensional time-dependent heat transfer within the air-cooled cylinder was to prepare a good basis for the later three-dimensional model of the heat transfer in the combined air–oil cooled cylinder. Temperature boundary conditions at the inner cylinder wall and convective boundary conditions for the outer cooling surfaces were used for the calculations. A three-dimensional system was reduced to a two-dimensional system; time-dependent calculations were performed for individual longitudinal cylinder sections. The basic numerical model was described by Damjanić and Owen (1982) and by Owen and Damjanić (1983), suitable software "THENOL" was published by Damjanić (1991), while heat transfer problems and the Finite Element Analysis were from Šelih (1990). Temperature-dependent physical properties of the cylinder alloy were taken into account for the calculations. Results of the computations showed that thermal conductivity of the cylinder wall had a minor influence on the heat transfer results; the influence of the heat transfer coefficient was much more pronounced. The accuracy of the computed temperatures was checked by the measured temperatures within the cylinder wall. A comparison between time-dependent measured and calculated temperatures in one particular point of the cylinder (upper part of the cylinder and the vicinity of the outer cylinder wall) is presented in Fig. 4. Results in Fig. 3 showed that only the upper part of the cylinder was thermally overloaded and should therefore be analyzed carefully. Numerical results on Fig. 4 thus cover only the portion from the top to the transverse section IVA. Finally, the temperature distribution of particular longitudinal cylinder sections can be presented; Fig. 5 represents computed temperatures in the two most interesting longitudinal cross sections, "3" and "8," after 13 minutes of temperature stabilization. The shape of the particular isotherms was checked by the measured control temperatures and showed suitable accuracy of the applied numerical model.

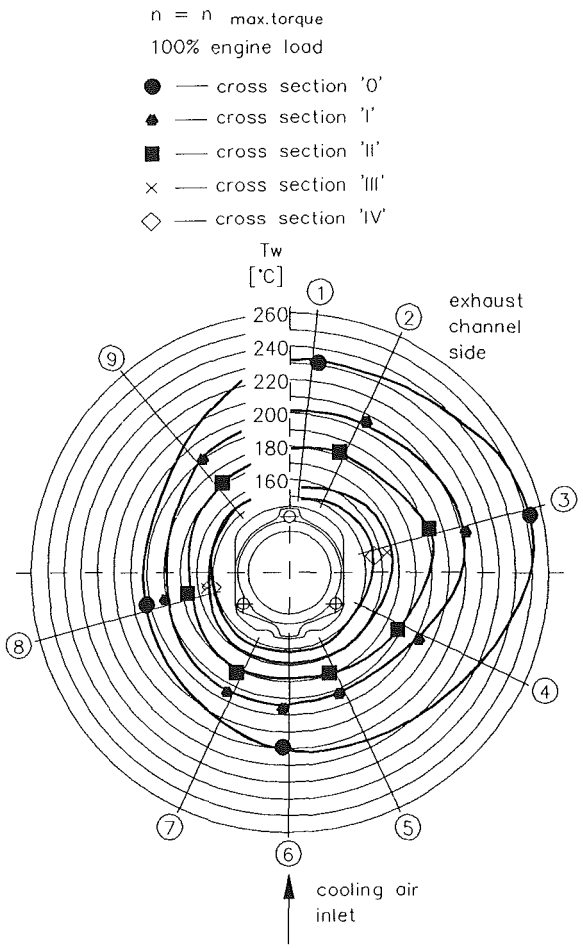


Fig. 2 Circular temperature distribution of the cylinder wall for diverse transverse sections

The analyses performed showed the following features of the baseline air-cooled cylinder:

- Longitudinal-axial temperature gradients of the inner cylinder wall showed their extreme values in sections "3" and "8" that were mainly influenced by the exhaust and by the intake channel. Asymmetry of the circumferential cylinder wall temperatures in the critical transverse section "I" exceeded 35°C.
- Wall temperatures decrease (in the axial direction from the top of the cylinder) intensively to the transverse section "IV," until they reach stable values.
- Maximum local wall temperatures exceed 210°C in the transverse section "I."
- Omitting the air cooling (thermal insulation of the cooling fins) in the lower half of the cylinder had a negligible effect on the temperature distribution in the upper part of the cylinder.

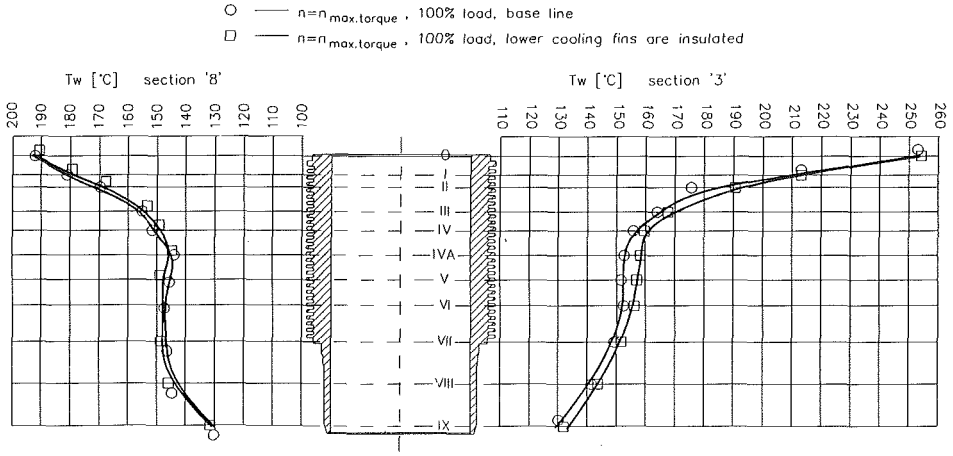


Fig. 3 Longitudinal cylinder wall temperature distribution; influence of the partial cooling surface insulation

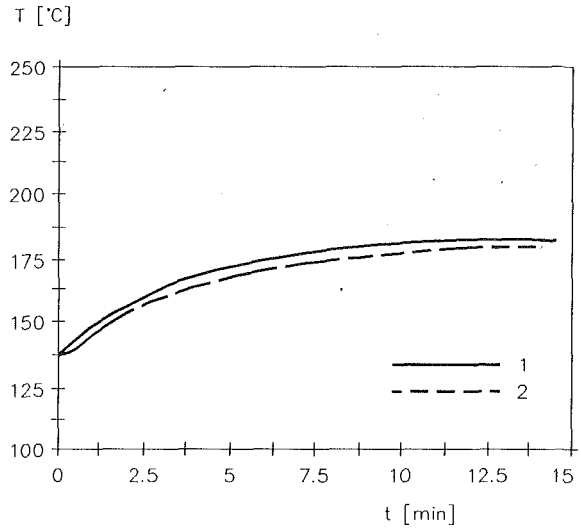


Fig. 4 Temperature time distribution of a point in the cylinder wall: (1) measured, (2) calculated

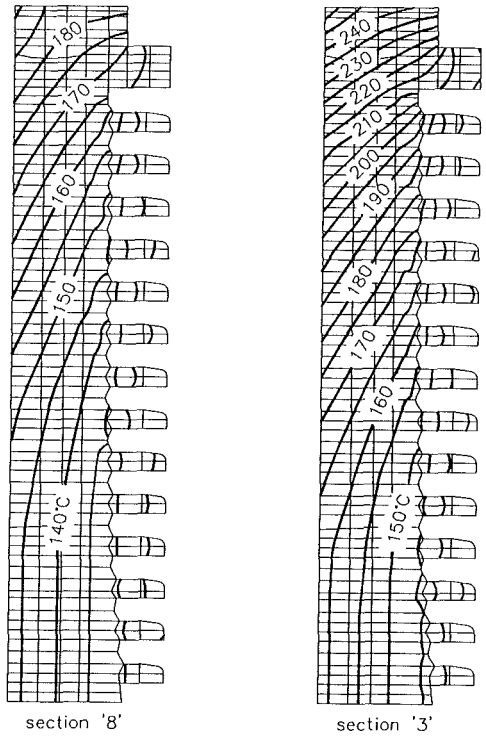


Fig. 5 Calculated temperature distribution in two longitudinal cylinder sections

Taking into account all the above-mentioned, one can conclude that thermal problems of an air-cooled cylinder involve only a small and limited portion of the cylinder wall surface: from the top (contact surface with the cylinder head) to the transverse section "I" (less than 25 percent of the cylinder height!).

Solution of the Problem and the Results

According to the results in chapter 2, a curved, square cross section and horizontal channel supplied with lubricating oil was introduced into the upper part of the cylinder wall. Dimensions and position of the channel consequently followed the results of the analyses of the baseline air-cooled cylinder and the analyses of the heat transfer problems in laminar flows

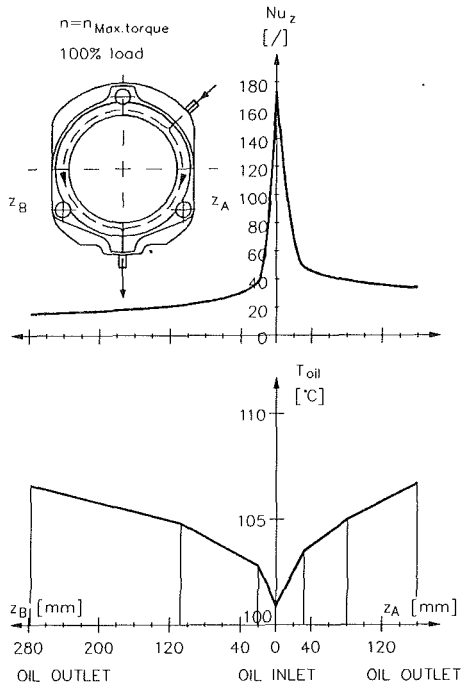


Fig. 6 Average local Nusselt number Nu_z and cooling oil temperature distribution for two branches of the oil channel

of viscous fluids within the short, curved channels reported by Trenc (1992).

Diverse forms of uniform oil flow and split flow combined cooling systems were tested on a single-cylinder engine. Detailed analysis of the heat transfer in the channel, based on statistical treatment of the experimental data, showed that local values of the heat transfer coefficient varied intensively along the branches of the channel (Trenc, 1992). Experiments were performed at diverse engine running conditions and diverse cooling oil flows. As the result of multiple regression, a nondimensional mathematical model for the heat transfer coefficient, depending on the Reynolds (Dean), Prandtl, and Grashof numbers was established (Trenc, 1992). Variation of the local heat transfer coefficient can be used to suit local demands for supplementary cooling and can therefore help to balance unequal circumferential temperature distribution in the cylinder wall. Figure 6 shows the distribution of the average (for the four channel walls) local Nusselt number (Nu_z).

Local oil temperatures for the two channel branches are presented as well; oil temperature increase of only 6°C was observed despite the high engine loads. On the other hand, no local overheating of the cooling oil was noticed. Results of the combined air-oil cooling are presented in Figs. 7 and 8. Figure 7 shows the influence of the combined cooling on the temperature distribution in the inner cylinder wall at the typical engine loads. Comparison was made for air and a combined cooled cylinder. A pronounced asymmetry of the polar temperature distribution in the air-cooled cylinder was substantially leveled in the combined cooled cylinder to only 10°C. Maximum wall temperatures in the transverse section "I" of the combined cooled cylinder never exceeded 180°C even at 10 percent overload (specific engine power corresponds to 24 kW/l); their level corresponded to the partial (67 percent) load temperature level of the baseline cylinder. Similar or better results were obtained at the engine rated speed. Axial cylinder wall temperatures for the baseline, combined, and water-cooled cylinders are presented in Fig. 8. Temperature gradients are reduced for the combined cooled cylinder (see the enclosed table) and the circumferential temperature asymmetry is even less pronounced when compared with the water-cooled version.

Temperatures in the lower half of the combined cooled cylinder are substantially higher compared to those of the water-cooled one, but not too high (corrosion problem due to the water vapor condensation can be better controlled).

Supplementary oil cooling is self-controlled; higher engine loads result in higher average cooling oil temperatures. As the result, higher heat transfer coefficient values are obtained and more intensive cooling tends to decrease cylinder wall temperatures. There is an optimum quantity of the oil required for the combined cooled cylinder: 6 to 10 kg/min are suitable for the rated specific power of 22 to 25 kW/l engine swept volume. The influence of the cooling oil flow on the circumferential wall temperature distribution is presented in Fig. 9. Besides, cooling oil inlet–outlet temperature increase and maximum cylinder wall temperature differences in the transverse section "I" are presented in the enclosed table. Although optimum oil flow was substantially increased, no adequate wall temperature reduction was observed.

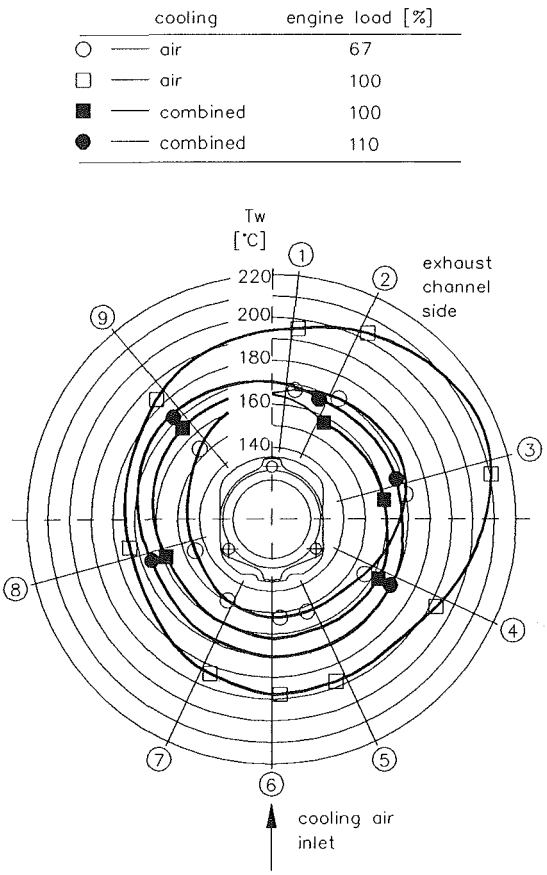


Fig. 7 Circumferential temperature distribution of the cylinder wall for air and combined cooled engine

Experimental Methods

The experiments performed were basically linked to the engine performance data at diverse running conditions. Measurements on the single and four-cylinder engines were performed on a fully electronically controlled test stand and supported with AVL-PUMA 4 control system. Simultaneous temperature measurements were performed with a digital temperature data logger and with a data acquisition system HP 3052A.

0.2-mm diameter NiCr–Ni thermocouples were used for temperature measurements; the whole measurement chain was calibrated by known temperature properties of the Pb–Sn binary alloy.

Discussion

High temperatures and temperature gradients within the upper part of the air-cooled cylinder do not tolerate any increase of the engine specific power. Axial and circumferential wall temperature gradients can provoke additional thermal stresses and deformations of the cylinder.

Analyses of the results performed on the baseline cylinder showed that the above-mentioned temperature irregularities occurred only in a limited upper part of the cylinder wall, and that maximum temperature values occurred in the vicinity of the exhaust channel. Solution of the problem should therefore include intensive local and controlled heat transfer. A horizontal curved channel fed with lubrication oil and introduced into the critical upper part of the air-cooled cylinder wall was used to save the problem. Dimensions, position, and intensive heat transfer in the laminar oil flow in the curved channel helped to reduce maximum cylinder temperatures.

Circumferential wall temperature asymmetry was reduced by unequal and controlled heat transfer within two channel branches. Longitudinal temperature gradients were simulta-

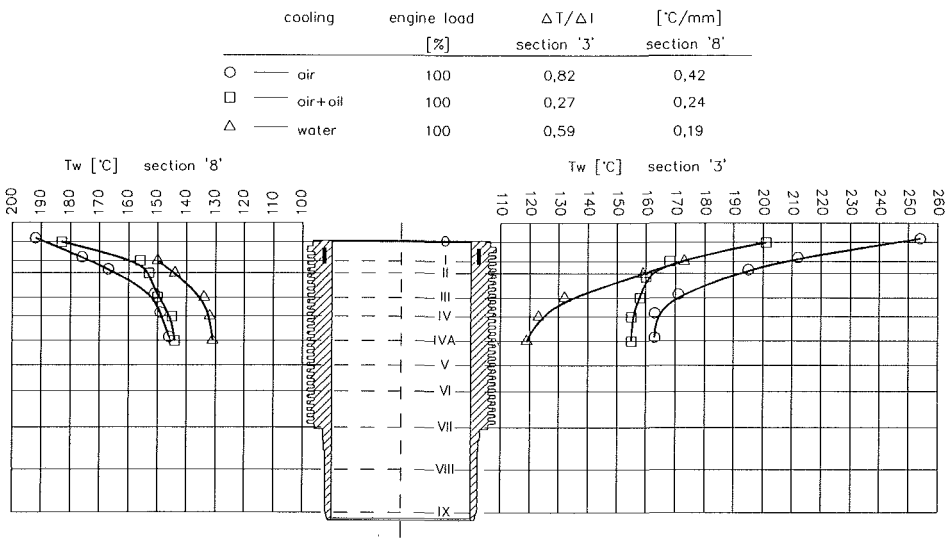


Fig. 8 Typical longitudinal temperature distribution for the air-cooled, combined, and water-cooled cylinder at the same engine running conditions

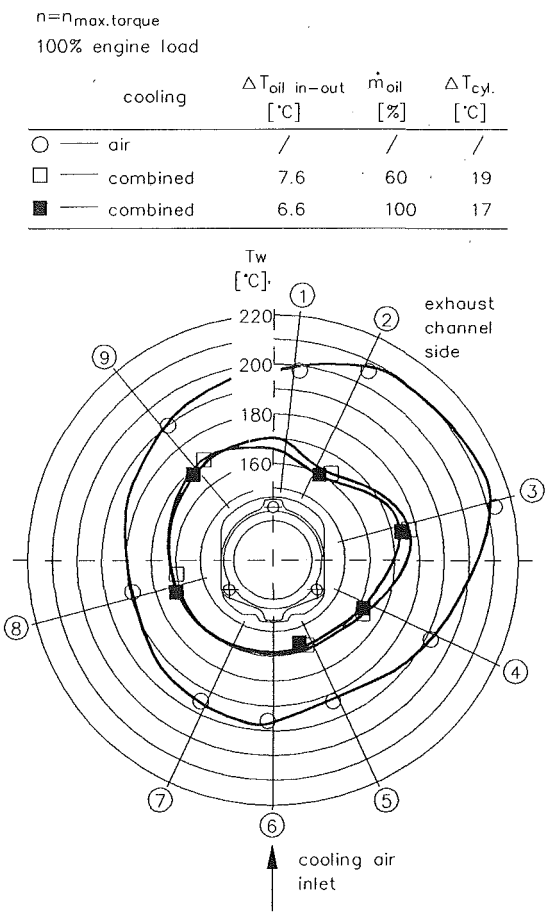


Fig. 9 Influence of the oil mass flow on the circumferential temperature distribution for air and combined cooled cylinders

neously reduced to one third of their initial value. The oil cooling component is self-controlled: Increase of the engine load is mainly compensated by the simultaneous increase of the heat transfer coefficient, and the wall temperatures remain more or less unchanged. Cooling oil flow required for optimum combined cooling presents no special problem; oil pump capacity is slightly increased. The heat released into the cooling oil requires a bigger capacity oil cooler. However, the resulting increase of the oil temperature in the oil pan does not exceed a few degrees centigrade.

Conclusions

A new combined air–oil cooling system applied on the prototype TAM TC & iC Diesel engine cylinders is presented in the paper and its main features are:

- The intensity of the additional oil cooling follows local requirements for cooling.
- Construction of the channel and the control of the cooling oil flow within the channel branches is very simple.
- Optimization of the combined cooling system results in the decrease of maximum rated and local cylinder wall temperatures by almost 40°C, critical longitudinal cylinder wall gradients are reduced by more than three times, and maximum circumferential wall temperature differences by nearly 40°C.
- New temperature distribution of the combined cooled cylinder makes possible an additional and safe increase of the engine specific power by more than 10 percent; the thermal level of the newly designed cylinder is at the same time identical to the level of the baseline air-cooled cylinder, but at the engine load, which is as low as 67 percent of the rated value.
- Temperature distribution in the upper part of the combined cooled cylinder wall is very similar to that of the typical water-cooled one.
- The quantity of cooling air is less important for the combined cooling; the capacity of the cooling air fan can therefore be reduced together with the required driving power.
- Reduction and redesign of the cylinder cooling fins makes possible simpler, more rigid construction, and an even larger cylinder bore.
- No special additional investments are required for the production of the combined cooled engines. Bigger and competitive-rated specific power and more reliable operation of the engine can be expected at the same time. Low specific fuel consumption, good ecological picture (complying with EURO1 requirements), and all other specific advantages of the air-cooled engines give favorable prospects for the air-cooled engines in the future.

Acknowledgments

The authors wish to record their appreciation of the TAM R&D Institute, which permitted the publishing of this paper. Moreover, the valuable financial support of the Ministry of Science and Technology of the Republic of Slovenia is gratefully acknowledged.

References

Damjanić, F. B., and Owen, D. R. J., 1982, "Practical Considerations for Thermal Transient Finite Element Analysis Using Isoparametric Elements," *Nuclear Engineering and Design*, Vol. 69, pp. 109–126.

Damjanić, F. B., 1991, "THENOL" Package, Version 2.01, Manual for Users, FAGG—University of Ljubljana, Slovenija.

Owen, D. R. J., and Damajanić, F. B., 1983, "Reduced Numerical-Integration in Thermal Transient Finite Element Analysis," *Computers and Structures*, Vol. 17, pp. 261–276.

Selih, J., 1990, "Numerical Methods for Engineers—Thermal Analysis of the Air Cooled Engine Cylinder," Magister Work No. 70, FAGG—University of Ljubljana, Slovenija.

Trenc, F., 1992, "Analysis of the Temperature Distribution in an Air Cooled Diesel Engine," Dr. Thesis, No. 137/D, Faculty of Mechanical Engineering, University of Ljubljana, Slovenija.

Carbonyl Compounds and PAH Emissions From CNG Heavy-Duty Engine

M. Gambino

R. Cericola

P. Corbo

S. Iannaccone

Istituto Motori C.N.R.,
Naples, Italy

Previous works carried out in Istituto Motori laboratories have shown that natural gas is a suitable fuel for general means of transportation. This is because of its favorable effects on engine performance and pollutant emissions. The natural gas fueled engine provided the same performance as the diesel engine, met R49 emission standards, and showed very low smoke levels. On the other hand, it is well known that internal combustion engines emit some components that are harmful for human health, such as carbonyl compounds and polycyclic aromatic hydrocarbons (PAH). This paper shows the results of carbonyl compounds and PAH emissions analysis for a heavy-duty Otto cycle engine fueled with natural gas. The engine was tested using the R49 cycle that is used to measure the regulated emissions. The test analysis has been compared with an analysis of a diesel engine, tested under the same conditions. Total PAH emissions from the CNG engine were about three orders of magnitude lower than from the diesel engine. Formaldehyde emission from the CNG engine was about ten times as much as from the diesel engine, while emissions of other carbonyl compounds were comparable.

Introduction

In metropolitan areas, vehicles are the prevailing source of atmospheric pollution. In many cities the use of private cars is restricted, and consequently, mobility is based on the public transportation system. Therefore, great attention must be paid to reducing emissions from mass transportation vehicles. In this context, the administration of the Tuscany Region of Italy took the initiative to verify the benefits of using CNG instead of Diesel fuel in urban bus engines. For this purpose Istituto Motori carried out research regarding the conversion of a Diesel engine to a natural gas engine. A comparison between performance and pollutant emissions before and after the conversion was conducted. In order to realize maximum benefit from the comparison of the two fuels, the exhaust emission analysis was performed not only on regulated substances (HC, CO, NO_x, smoke), but also on some unregulated compounds such as carbonyl compounds and PAH. These compounds, which are considered highly harmful to human health, were identified and quantified by means of a specific sampling and analysis procedure.

CNG Engine

The goal of developing a CNG spark ignition engine with low pollutant emissions can be achieved by one of two basic approaches: using three-way catalyst technology, which requires the engine to be operated at a stoichiometric air–fuel ratio, or using the lean burn technology. The first solution needs a very accurate control of the air–fuel ratio, which is particularly difficult to realize in a methane-fueled engine, where the window of catalyst efficiency is narrower than in a gasoline engine [1]. Moreover, the commercial catalytic converters showed low conversion effect on methane [2], which is the main component of total hydrocarbons present in the exhaust gas of a CNG engine. On the other hand, results of several experiments conducted at the Istituto Motori on different types of heavy-duty natural gas engines showed the possibility of meeting emission limits with the lean burn solution [3]. In addition, a lean burn engine generally showed higher thermal efficiency and lower exhaust temperature than a stoichiometrically adjusted engine. For these reasons the lean burn option appeared to be more viable and suitable, and was chosen in the present work. Engine technical characteristics are reported in Table 1.

A disk-shaped combustion chamber with two spark plugs has been adopted. The compression ratio was lowered to 9:1,

Table 1 CNG turbocharged engine data

Six cylinder turbocharged
Displacement 9.5l
Bore × stroke 120 × 140 mm
Compression ratio 9:1 (16:1)
Max power 154 kW@1900 rpm (154 kW@2050 rpm)
Max torque 776 Nm@1800 rpm (882 Nm@1100 rpm)
Max brake efficiency 0.359 (0.390)
Electronic ignition with two spark plugs per cylinder
CNG venturi carburetor DELTEC
Boost pressure 0.6 bar
Valve overlap angle 0 deg (49.8 deg)

Note: The values in brackets refer to the diesel engine.

Contributed by the Internal Combustion Engine Division and presented at the 14th Annual Fall Technical Conference, Waterloo, Iowa, October 4–7, 1992. Manuscript received by the Internal Combustion Engine Division August 1, 1993. Associate Technical Editor: J. A. Caton.

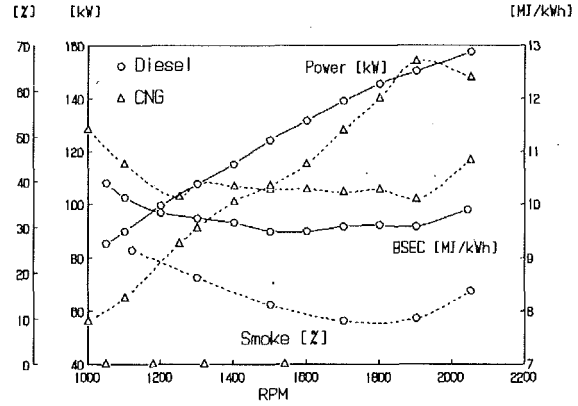


Fig. 1 Power, BSEC, and smoke for CNG and Diesel engine

Table 2 R49 test for CNG and Diesel engine

	HC	CO g/k Wh	NO_x
CNG	2.0	2.3	11.9
Diesel	1.0	1.4	15.3

while valve overlap was eliminated in order to prevent the fresh mixture from escaping through the exhaust valve. A liquid-cooled turbocharger was also adopted.

Engine Performance Tests

The two engines (CNG and Diesel) were tested using a Schenck U1-25H hydraulic dynamometer. Inlet air flow was measured with Ricardo Viscous Air Flow Meters, and natural gas flow was calculated with a Tartarini Instromet 782-2X volumetric flow meter.

In spite of the lean mixtures used, the power of the CNG engine was comparable with the Diesel version, as shown in Fig. 1. On the other hand brake specific energy consumption (BSEC) was slightly higher for the CNG version, due to the lower compression ratio. For exhaust gas analysis the following instrumentation was used:

- Beckman 404 HFID (flame ionization detector) for hydrocarbons;
- Beckman 880 NDIR (non dispersive infrared analyzer) for CO and CO_2;
- Beckman 955 CLA (chemiluminescent analyzer) for NO_x;
- Hartridge MK III smoke meter to measure exhaust gas opacity.

Regulated gas emissions were measured according to 88/77 CEE R49 procedure (13-mode cycle), while smoke opacity was evaluated following the R24 procedure.

Pollutants emitted during the R49 cycle were lower than the limits for both engines. In Table 2 the comparison between the gaseous emissions of the two engines is reported.

The high HC emission obtained from the CNG version can be attributed to the combustion chamber shape, which was not specifically optimized for natural gas. However, it is important to mention that the hydrocarbons emitted from the CNG engine were mainly constituted of methane, which is not toxic.

On the other hand, the lower NO_x emission from the natural gas engine was due to using lean mixtures and also to the different type of combustion. Figure 1 shows that exhaust opacity is negligible (2 percent Hartridge) for the CNG engine, while it varies from 10 to 30 percent for the Diesel engine.

Starting in July 1992, stricter emission limits will be imposed in Europe on heavy-duty engines. In order to meet the new limit standards, some improvements of the engine components are necessary, in particular, the combustion chamber shape. In fact, a microturbulence combustion chamber should allow extremely lean mixtures to be used, resulting in very low emissions and good thermal efficiency [4]. In addition, appreciable reduction of pollutant emissions should be achieved by the electronic control of the gas feed system.

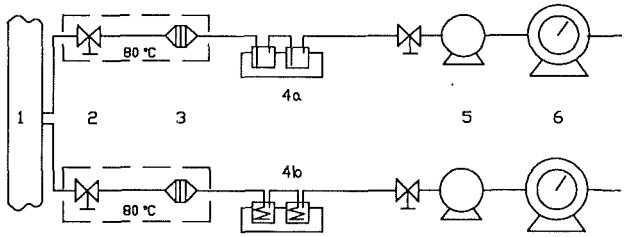


Fig. 2 Sampling train for carbonyl compounds and PAH collection: (1) exhaust pipe; (2) two-way valve; (3) filter; (*4a*) impingers; (*4b*) condensers; (5) vacuum pump; (6) dry gas meter

Unregulated Emissions

Sampling and Analysis Procedure. The sample components train is illustrated in Fig. 2. Unregulated emissions were collected by two stainless steel probes that were located in the exhaust pipe at 1.5 m from the turbocharger.

The sampling and quantitative analysis of carbonyl compounds were performed by using the 2,4-dinitrophenylhydrazine (DNPH) method [5]. The compounds were collected by bubbling the engine exhaust gas through an acetonitrile solution of DNPH reagent, cooled at 0°C. A filter placed upstream of the impingers stopped the solid particles. The face temperature of this filter was about 80°C. Samples of the solution containing the hydrazone derivatives were directly injected into the Hewlett-Packard high-pressure liquid chromatograph, which was equipped with an ultraviolet detector.

The analysis of the carbonyl compounds showed an appreciable quantity of several compounds, some of which are considered significant in the IARC classification [6].

PAH can be present in the exhaust of an engine in both the particulate matter and the gas phase. For this reason, the sampling line was comprised of a Pallflex filter type TX40H120-WW for soot collection and two condensers assembled in series (0°C and −20°C) to stop vapor phase PAH. The outside temperature of the filter was maintained at about 80°C. After sampling the line, the condensers were washed with methylene chloride, in order to recover any organic material lost along the line. The condensed water was separated by phase separator filters. Soxhlet extraction with methylene chloride was performed on the filter. The organic material contained in the extract was added to that trapped in the condensers, then all the solvent was roto-evaporated. The total organic material obtained by the described procedure was analyzed by a Hewlett-Packard GCMS (gas chromatography–mass spectrometry) using the SIM technique (selected ion monitoring). According to EPA method 610 and IARC classification, only 16 PAH were identified and quantified. Detecting this group of compounds is a significant indicator of the toxicity grade of engine emissions.

The unregulated emissions were evaluated using the same 13-mode cycle used for regulated emissions.

Test Results and Data Analysis. A summary of unregulated emission results is given in Tables 3 and 4. PAH emission data (Table 3) show a significant difference between CNG and Diesel engines.

PAH emission data of the Diesel engine are in a good agreement with the work conducted by Hemmerlein et al. on the same 13-mode cycle [7].

Table 3 shows that production of the PAH group from the Diesel engine was three orders of magnitude higher than from the CNG engine. In particular, differences from 1 to 3 orders of magnitude can be noticed for those PAH to which corre-

Table 3 Polycyclic aromatic hydrocarbons (PAH) emissions on 13-mode cycle (R49)

	CNG	Diesel
	μg/k Wh	
Naphthalene	$1.76 \cdot 10^{-1}$	$9.95 \cdot 10^{+1}$
Acenaphthylene	$1.63 \cdot 10^{-1}$	$1.62 \cdot 10^{+2}$
Acenaphthene	$5.61 \cdot 10^{-2}$	$8.28 \cdot 10^{+1}$
Fluorene	$1.80 \cdot 10^{-1}$	$2.52 \cdot 10^{+2}$
Phenanthrene	$8.65 \cdot 10^{-1}$	$1.03 \cdot 10^{+3}$
Anthracene	$1.74 \cdot 10^{-1}$	$7.63 \cdot 10^{+1}$
Fluoranthene	$2.70 \cdot 10^{-1}$	$6.70 \cdot 10^{+1}$
Pyrene	$2.90 \cdot 10^{-1}$	$1.19 \cdot 10^{+2}$
Chrysene	$9.44 \cdot 10^{-2}$	$6.75 \cdot 10^{+0}$
Benz(a)anthracene (2A)	$2.60 \cdot 10^{-1}$	$2.13 \cdot 10^{+1}$
Benzo(b)fluoranthene + Benzo(k)fluoranthene (2B)	$6.09 \cdot 10^{-2}$	$4.53 \cdot 10^{+0}$
Benzo(a)pyrene(2A)	$1.08 \cdot 10^{-1}$	$3.90 \cdot 10^{-1}$
Indeno(1,2,3-cd)pyrene (2B)	$1.09 \cdot 10^{-2}$	$5.61 \cdot 10^{-1}$
Dibenz(a,b)anthracene (2A)	$7.43 \cdot 10^{-3}$	$2.90 \cdot 10^{-1}$
Benzo(ghi)pyrene	$1.18 \cdot 10^{-1}$	$6.23 \cdot 10^{-1}$
Total Polycyclic Aromatic Hydrocarbons	$2.83 \cdot 10^{+0}$	$1.92 \cdot 10^{+3}$

Table 4 Carbonyl compounds emissions on 13-mode cycle (R49)

	CNG	Diesel
	mg/k Wh	
Formaldehyde (2A)	32.48	3.83
Acetaldehyde (2B)	1.29	1.82
Acrolein (3)	0.32	0.30
Acetone	1.76	0.58
Propionaldehyde	0.18	0.12
Crotonaldehyde	0.12	0.42
Methacrolein	0.11	0.07
Butyraldehyde + 2-Methyl ethyl ketone	0.20	0.15
Benzaldehyde	0.30	0.57
Total carbonyl compounds	36.76	7.86

spond a higher toxicological risk in IARC classification: benz(a)anthracene, benzo(a)pyrene, and dibenz(a,h)-anthracene, which are present in Group 2A of such a classification, and benzo(b)fluoranthene and indeno(1,2,3-cd)pyrene, included in Group 2B.

Concerning total carbonyl compounds, data of Table 4 show that emissions from the CNG engine are almost an order of magnitude higher than from the Diesel engine.

This difference is mainly due to emission of formaldehyde (Group 2A in IARC classification), which can be considered a typical emission of a methane-fueled engine operating in lean burn conditions [4]. The other compounds included in the IARC classification showed that a lower emission of 30 percent for acetaldehyde (Group 2B) was obtained from the CNG engine, while emission values of acrolein (Group 2B) measured for the two engines were comparable.

In order to evaluate the difference observed for the unregulated emission values of the two engines and its impact on health, the relative toxicity of the compounds included in Group 2A of IARC classification was estimated. A relative toxicity factor was calculated using the threshold limit values (TLV) established for concentration of harmful substances in the workplace. In particular, the TLV used for formaldehyde is that fixed by the American Conference of Governmental Industrial Hygienist (ACGIH) [9]. It is defined as the time-weighted average concentration (TWA) of harmful substance, to which most workers can be exposed without negative effects on health. However, as ACGIH does not establish any TLV for PAH, the value of maximum allowable concentration (MAC) fixed in Russian regulation and suggested in France for benzo(a)pyrene [10–12] was used for the three PAH included in Group 2A of the IARC classification. Assuming TLV and MAC values as a measure of the relative toxicity of the considered compounds, the relative toxicity factor for PAH is four orders of magnitude higher than for formaldehyde, as shown in Table 5.

Table 5 Relative toxicity of unregulated emissions

	HCHO	PAH*
TLV-TWA	1.5 mg/m^3	
MAC		0.15 μg/m^3
Relative toxicity factor	1	10^4
13-mode emission		
CNG	32.5 mg/k Wh	0.38 μg/k Wh
DIESEL	3.8 mg/k Wh	21.98 μg/k Wh
Toxicity index		
CNG	32.5	4
DIESEL	3.8	220
Total index (HCHO + PAH*)		
CNG	36	
DIESEL	224	

* Benz(a)anthracene + benzo(a)pyrene + dibenz(a,h)anthracene.

As a consequence, taking into account the emission values as calculated on the 13-mode cycle, it is possible to estimate a toxicity index from the emission value for the relative toxicity factor. Table 5 shows that the total toxicity index, calculated for unregulated compounds included in Group 2A of the IARC classification, is about one order of magnitude lower for the CNG engine than for the Diesel engine.

Conclusion

The CNG engine has proved to be a promising solution for reducing pollutant emissions from urban transit vehicles. In fact, the CNG engine tested in this work (a bus Diesel engine converted to work with CNG) showed very low smoke. CO, HC, and NO_x emissions were lower than R49 regulation limits. Total hydrocarbons was mainly composed of harmless methane.

Furthermore, a toxicity index, evaluated on the base of PAH and carbonyl compound emissions, is about one order of magnitude lower for the CNG engine than for the Diesel version.

For the lean burn CNG engine to meet the new European emission standards, new designs are needed for both the combustion chamber and the gas feed system.

References

1 Klimstra, J., "Catalytic Converters for Natural Gas Fueled Engines—A Measurement and Control Problem," SAE Paper No. 872165, 1987.

2 Gambino, M., Iannaccone, S., and Unich, A., "Performances and Emissions of Heavy Duty Spark Ignition Engines Fueled With Methane," *International Conference on Natural Gas and Liquefied Petroleum Gas as Fuel for Internal Combustion Engines*, Kiev, Sept. 1991.

3 Gambino, M., Iannaccone, S., and Unich, A., "Heavy Duty Spark Ignition Engines Fueled With Methane," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 113, No. 3, 1991, pp. 359–364.

4 Kingston Jones, M. G., and Heaton, D. M., "Nebula Combustion System for Lean-Burn Spark Ignited Gas Engines," SAE Paper 890211, 1989.

5 Lipari, F., and Swarin, S. J., "Determination of Formaldehyde and Other Aldehydes Automobile Exhaust With an Improved 2,4-Dinitrophenylhydrazine Method," *J. Chromatogr.*, Vol. 247, 1982, pp. 297–306.

6 IARC Monographs on the Evaluation of Carcinogenic Risks to Humans, Vol. 46, 1988.

7 Hemmerlein, N., Korte, V., and Richter, H., "Investigations Concerning Use of Rapeseed Oil as Alternative Fuel for Diesel Engine," Technical Report, Porsche R&D Centre, 1979.

8 Grechi, D., Poggi, A., and Bellini, V., "Da Gasolio a Metano: Effetti sulla Concentrazione Atmosferica di Inquinanti Tossici per l'Uomo e sulla Rumorosità (From Gasoil to Methane: Effect of Toxic Pollutants on Human Health and Noise)," Conference on Using of Natural Gas for Public Transport, Florence, Italy, May 1992.

9 Giornale degli Igienisti Industriali (Journal of Industrial Hygienists), Threshold Limit Values ACGIH 1986-1987, Vol. 12, 1, 1987.

10 Shabad, L. M., "On the So-Called MAC (Maximal Allowable Concentrations) for Carcinogenic Hydrocarbons," *Neoplasma*, Vol. 22, 1975, pp. 459–468.

11 "Occupational Exposure Limits for Airborne Toxic Substances," Occupational Safety and Health Series, No. 37, International Labour Office, Geneva, 1980.

12 "Valeurs Limités d'Exposition Professionelle aux Substances Dangereuses en France (Limit Values of Professional Exposure to Dangerous Substances in France)," INRS, *Cah. Notes Docum. No. 133, 1980, pp. 691–706.*

Dynamic and Static Flow Analysis of a Gasoline Fuel Injector

J. L. Chen

G. Chen

M. Wells

Electrical and Fuel Handling Division,
Ford Motor Company,
Ypsilanti, MI 48197

The effect of the electrical pulse width and the fuel temperature on dynamic flow and static flow rate of a gasoline fuel injector has been numerically and experimentally investigated. In numerical analysis, the physical domain covers the region from upstream of the valve seat to the injector exit. The three-dimensional unsteady Navier–Stokes equations in a curvilinear coordinate system are solved. Due to the needle movement in fuel injection, the physical domain is considered as a function of time. In the experimental study, the test stand consisting of a hydraulic system and an electrical system was designed to meet the requirements of Society of Automotive Engineers. The pulse width of 0.97–7.5 ms and the temperature of 20–100°C were used to study the pintle injector performance. Predicted dynamic flow and static flow rates show higher values at a temperature of 80°C, which are consistent with the test results.

Introduction

The gasoline fuel injection system injects fuel to the intake port of the engine. Various pulse widths can be sent to the fuel injector according to the actual engine operating conditions as sensed by an electronic control unit. As such, the fuel injection system can have more uniform spray distribution, more rapid engine response to changes in throttle positions, and more precise control of air/fuel ratio during cold start and engine warm-up (Heywood, 1988). To achieve these benefits, the fuel injector is required to provide a wide dynamic flow range and a minute variation in stroke-to-stroke fuel quantity. For a small engine, the low-end linearity between the dynamic flow and the pulse width are especially important for the fuel injector, because it has to meter a small amount of fuel accurately at a short pulse.

To deliver the liquid fuel, the solenoid coil of the injector, activated by a pulsed electronic signal, generates a magnetic field to lift the needle. This creates a passage between the needle and the valve seat. The fuel is driven by the regulated pressure gradient between the fuel rail and the injector exit. The fuel injector is required to deliver fuel with minimal stroke-to-stroke variation and a spray cone with good atomization. These are essential for smooth low idle, low raw emissions and good drivability for the cold engines.

To improve the fuel injector performance, several companies have been developing new concepts and modifying the designs. Greiner et al. (1987) at Robert Bosch GmbH developed a multihole injector to reduce buildup of deposit and a bottom-feed injector to handle under-hood high temperature. Andrighetti and Gallup (1987) at Lucas CAV employed a low-mass flat armature and an open-orifice valve seat to improve stability of air/fuel ratio. Okamoto et al. (1992) at Hitachi Ltd. developed a two-stream injector by implementing an adapter near the orifice to split the fuel flow. Although these developments provide interesting progress in improving the injector performance, some basic issues are still not resolved. No information is yet available regarding how the fuel temperature affects the flow field and pressure distribution inside the fuel injector and how much of the fuel is delivered during the opening cycle, the fully open stage, and the closing cycle. This is partly because it is very difficult to probe the flow field and pressure in the injector (e.g., a typical pintle diameter is 0.46 mm).

In view of this, a computational study was carried out to analyze the temperature effects on fuel delivery and to determine the mass flow at different stages of fuel injection. In the meantime, the test data of dynamic and static flow rate were employed to validate the numerical modeling. The fuel temperatures ranging from 20–100°C and the pulse widths of 0.97–7.5 ms are employed to investigate the pintle injector performance at two different operating conditions, i.e., the steady and transient states. At the steady-state operation, the valve of the fuel injector is held fully opened and the delivered fuel rate is called the static flow rate. In the transient state operation, the needle movement is controlled by the electronic pulse signal and the fuel mass delivered during a pulse is called the dynamic flow.

Test Facility

The test stand used to evaluate the gasoline fuel injector consists of a hydraulic system and an electrical system (see Fig. 1) that are designed to meet the requirements of the Society of Automotive Engineers (SAE, 1992). The hydraulic system supplies the test fluid to the injector inlet port at a stable flow rate and pressure, whereas the electrical system is to provide electrical pulses in controlling the valve movement of fuel injector(s).

Contributed by the Internal Combustion Engine Division and presented at the 14th Annual Fall Technical Conference, Waterloo, Iowa, October 4–7, 1992. Manuscript received by the Internal Combustion Engine Division July 1, 1992. Associate Technical Editor: J. A. Caton.

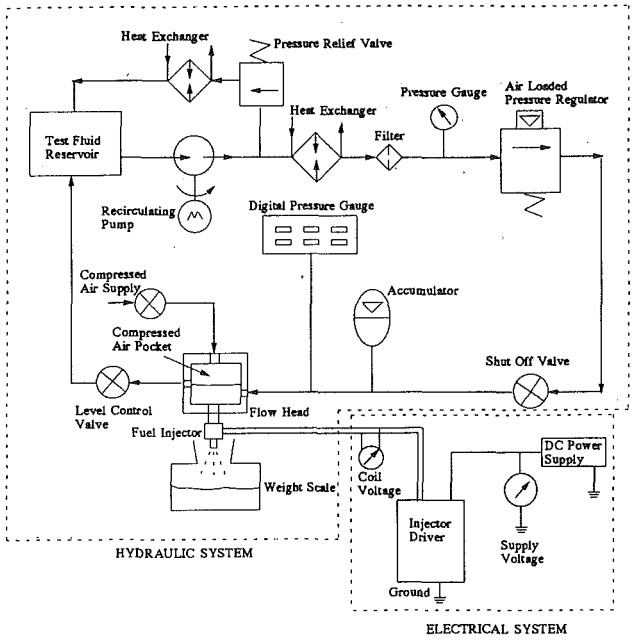


Fig. 1 Hydraulic and electrical schematics of test stand for gasoline fuel injectors

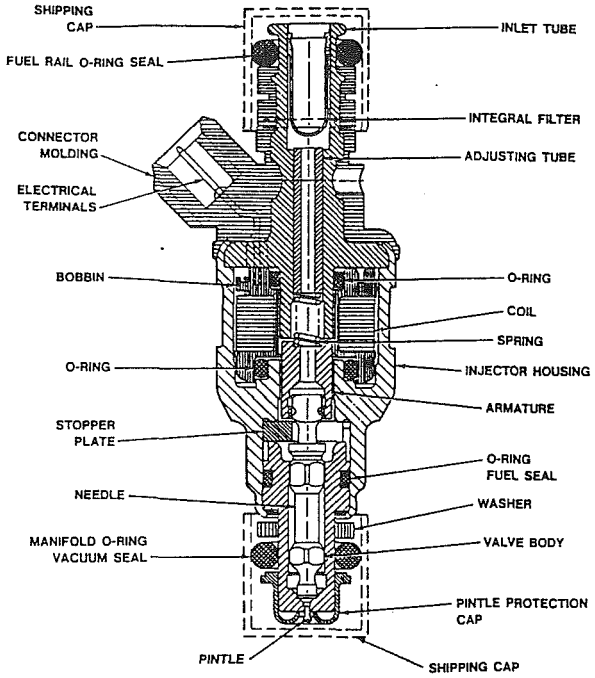


Fig. 2 Schematics of a typical pintle gasoline fuel injector

The hydraulic system is capable of handling test fluids as *n*-heptane, indolene, and mineral spirits. During the test, the test fluid in the fluid reservoir is pumped out to the fluid line by a recirculating pump. The fluid pump is capable of delivering 20 g/s of fluid at a stable pressure of 700 kPa. The excessive test fluid is sent back to the fluid reservoir through a bypass fluid line under the control of a pressure relief valve. To maintain the fluid temperature variation within 1°C, two heat exchangers are installed in the main fluid line and the bypass line, respectively. Downstream of the heat exchanger, there is a filter capable of removing solid particulates greater than 5 μm in order to avoid contamination in the fuel injector. The fluid pressure is reduced to the desired pressure by an air-loaded pressure regulator, which can achieve an accuracy of ± 0.1 kPa in the range of 40–500 kPa. The fuel pressure regulator is able to maintain the differential pressure across the injector at a value of 270 kPa and with a variation less than 0.5 percent throughout the test. A shut-off valve is located between the regulator and the accumulator for emergency shutdown. To eliminate any effect of fluid head pressure on injector inlet pressure, the pressure gage is installed in the same horizontal plane. In addition, an air-dampened fluid head is employed to reduce the pressure pulsation caused by the pump. This is achieved by precharging air on the top of the injector mount to form an air pocket with a minimum volume of 30 cc at fuel system pressure. The spray generated by the injector is collected by a test fixture beneath the injector exit. A weight scale is employed to measure the mass of the fuel that drains into the fixture. In order to provide repeatable readings, the test fixture has to be totally wet down by a minimum of 10,000 pulses.

The electrical system consists of an injector driver, a DC power supply unit, and voltage meters. The injector driver is an electronic circuit, which can supply voltage pulses to the fuel injector with an accuracy of ± 0.001 ms. Also, the time period elapsed from the beginning of the pulse to the next pulse is 10 ms. The DC power supply unit provides a voltage of 14 $\pm$ 0.05 VDC to the injector driver.

Mathematical and Numerical Models

The pintle injector (see Fig. 2) is chosen in this study because of its simplicity in design near the injector exit and thus fewer

Nomenclature

A_e = east surface area of a grid
C_D = discharge coefficient (i.e., ratio of actual flow rate to theoretical flow rate)
d_o = orifice diameter
F_u, G_u, H_u = local geometry factors in i, j, k directions
i, j, k = components in BFC (body-fitted coordinates)
i_x = location of the symmetric plane in BFC coordinates
j_y = location of the needle wall in BFC coordinates
k_v = location of the valve seat in BFC coordinates
k_z = location of the injector exit in BFC coordinates
l_o = orifice length
p = referenced pressure with respect to injector exit pressure
p_f = pressure at the fuel rail
r, θ, z = cylindrical coordinates
u, v, w = velocity components in Cartesian coordinates
u_i, v_j, w_k = velocity components in BFC coordinates
u_e, v_n, w_h = velocities at east, north, and high surfaces of a grid
V = vector form velocity
w_d = needle lift velocity in cylindrical coordinates
u_e^* = mass-flux velocity at east surface of a grid in BFC coordinates
α = inclination angle of needle surface in cylindrical coordinates
μ = dynamic viscosity
ρ_e = density at east surface of a grid

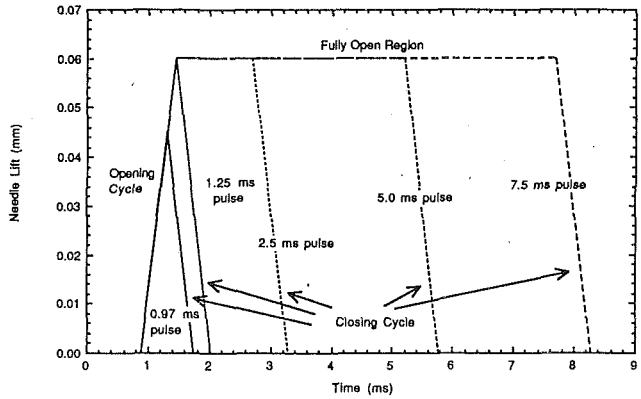


Fig. 3 Transient needle lift for various pulse widths

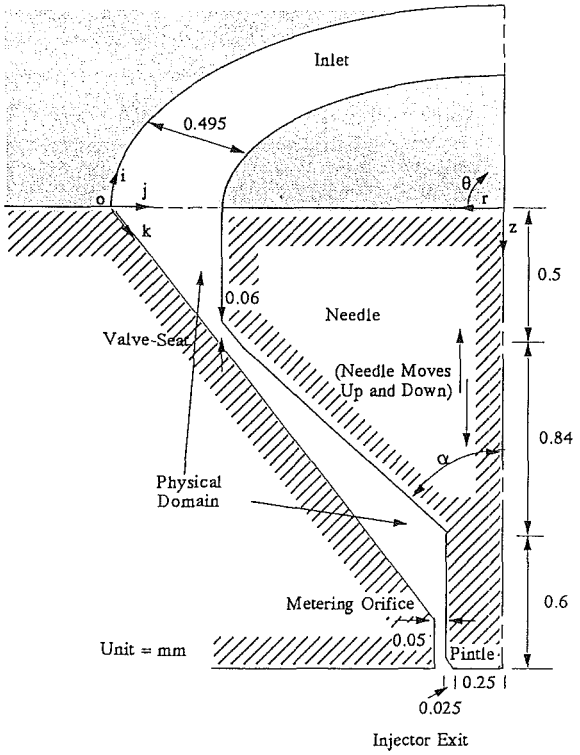


Fig. 4 Physical domain of the pintle injector at the needle lift of 0.06 mm (fully open)

parameters controlling the flow field, which is more suitable for fundamental study. Also, better understanding of the pintle injector can provide more basic design guidelines for the advanced injectors. Prior to the fuel injection, the needle is pressed against the valve seat by the helical spring to seal the flow passage. After the pulse is sent to the solenoid, it takes 0.9 ms of delay for the solenoid coil to generate enough force to overcome the spring force. The needle is therefore lifted to open the valve to start the opening cycle. The needle continues to lift and fully opens the valve at 1.45 ms, which sets on the fully open region. When the pulse is over, the needle still holds its position for another 0.2 ms because of the residual magnetic field in the coil. The valve then begins the closing cycle and takes an additional 0.58 ms to close the valve fully, which marks the end of fuel injection. The transient needle lifts at various pulse widths, ranging from 0.97 to 7.5 ms, are shown in Fig. 3.

The physical domain of the governing equations covers the region from upstream of the valve-seat to the injector exit (Fig. 4). This is because most of the pressure drop is expected to occur in this region and the design of this region controls the accuracy of fuel metering. The fluid dynamics in the injector is analyzed by solving the transient, three-dimensional Navier-Stokes equations. Since the flow temperature and density inside the injector maintain almost the same during the tests, the flow is considered as isothermal and incompressible in the computational analysis. The fuel density and kinematic viscosity at 20, 40, 60, 80, and 100°C are 7.34×10^{-4}, 7.24×10^{-4}, 7.1×10^{-4}, 6.97×10^{-4}, and 6.83×10^{-4} g/mm^3 and 0.6, 0.5, 0.4, 0.32, and 0.28 mm^2/s, respectively. Both viscosity and density decrease with temperature.

The Mean Flow Equations. The time-averaged equations for conservations of mass and momentum in a transient, three dimensional flow can be expressed in the tensor notation as

$$\frac{\partial U_i}{\partial X_i} = 0 \tag{1}$$

$$\frac{\partial(\rho U_i)}{\partial t} + \frac{\partial(\rho U_i U_j)}{\partial X_j} = -\frac{\partial P}{\partial X_i} + \frac{\partial}{\partial X_j}\left[\mu_l\left(\frac{\partial U_i}{\partial X_j} + \frac{\partial U_j}{\partial X_i}\right) - \rho\overline{u_i u_j}\right] \tag{2}$$

where the term $-\rho\overline{u_i u_j}$ is the Reynolds stress, which stands for *the contribution of turbulent motions to the mean stress.*

The Turbulence Model. In the present work, the Reynolds stress is calculated by adopting the Boussinesq turbulent eddy viscosity concept (Hinze, 1975) in which the Reynolds stress is expressed as

$$-\rho\overline{u_i u_j} = \mu_t\left(\frac{\partial U_i}{\partial X_j} + \frac{\partial U_j}{\partial X_i}\right) - \frac{2}{3}\delta_{ij}\rho k \tag{3}$$

where δ_{ij} is the Kronecker-Delta function and μ_t is the turbulent viscosity. The k-ϵ two-equation model (Launder and Spalding, 1972) is employed to calculate the turbulent viscosity, where the turbulent viscosity is expressed in terms of turbulent kinetic energy, k, and turbulent dissipation energy, ϵ, as follows:

$$\mu_t = C_\mu \rho \epsilon^2 / k \tag{4}$$

where $C_\mu = 0.09$. According to the two-equation turbulence model, the governing equations for k and ϵ are expressed as

$$\frac{\partial(\rho k)}{\partial t} + \frac{\partial(\rho k U_j)}{\partial X_j} = \frac{\partial}{\partial X_j}\left[(\mu_l + \mu_t/\sigma_k)\frac{\partial k}{\partial X_j}\right] - \rho\overline{u_i u_j}\frac{\partial U_i}{\partial X_j} - \rho\epsilon \tag{5}$$

$$\frac{\partial(\rho\epsilon)}{\partial t} + \frac{\partial(\rho\epsilon U_j)}{\partial X_j} = \frac{\partial}{\partial X_j}\left[(\mu_l + \mu_t/\sigma_\epsilon)\frac{\partial \epsilon}{\partial X_j}\right] - C_1\frac{\epsilon}{k}\rho\overline{u_i u_j}\frac{\partial U_i}{\partial X_j} - C_{2\rho}\frac{\epsilon^2}{k} \tag{6}$$

where $C_1 = 1.44$ and $C_2 = 1.92$ are further constants in this model. Also, $\delta_k = 1.0$ and $\delta_\epsilon = 1.217$ are turbulent Prandtl numbers for k and ϵ, respectively.

Grid Distribution. Due to the symmetry in the θ direction, only a quarter of the injector is considered as the physical domain in the numerical analysis. The whole physical domain is then divided to $15 \times 8 \times 53$ cells for the computational domain (see Fig. 5). In order to simulate the realistic physical domain, a curvilinear coordinate system, the Body-Fitted Coordinates (BFC), is employed to connect eight adjacent nodes to form a six-faced cell (Sugiura et al., 1990). The three coordinates in the BFC are i, j, and k and the origin is located at the intersection of the inlet region, the valve-seat plane, and the symmetric plane (see Fig. 4). The BFC grids can be viewed as the squeezed, stretched, bent, and twist orthogonal grids. In each grid, the conservations of mass and momentum are

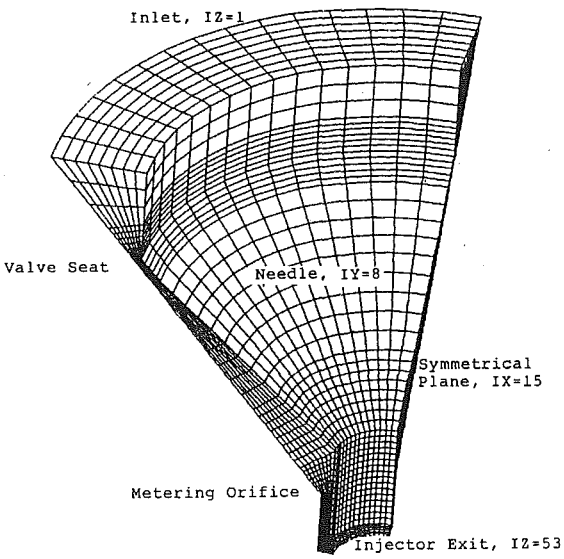


Fig. 5 Grid distribution in the computational domain of the fuel injection

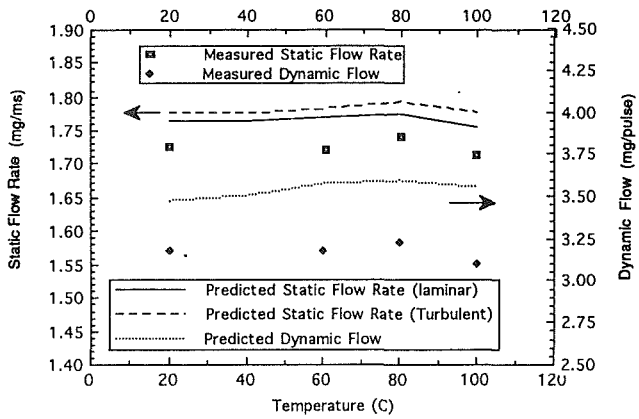


Fig. 6 Comparison of predicted and measured dynamic flow and static flow rate at various temperatures with pulse width of 2.5 ms

employed to form a set of algebraic equations. However, the direction of the velocity resolute in the BFC coordinates is not always normal to the corresponding surface of the grid. Thus, the velocity used in the mass-flux formula involves not only u, but also v and w. This new mass-flux velocity (say at the east surface of the grid), u_e^*, can be expressed in terms of the local grid geometry, i.e.,

$$u_e^* = F_u u_e + G_u v_n + H_u w_h \tag{7}$$

where F_u, G_u, and H_u are the functions of the local geometry of the grid in the i, j, and k directions, respectively, and u_e, v_n, and w_h are the velocities at the east, north, and high surfaces of the grid. A detailed expression of the local geometry functions, F_u, G_u, and H_u can be found from Rosten and Spalding (1986). The effective mass flux across the east surface is then expressed as

$$m_e = \rho_e A_e u_e^* \tag{8}$$

A similar treatment can be applied to the west, north, south, high, and low surfaces of the grid.

Initial and Boundary Conditions. In the event of fuel injection, the physical domain is a function of time in the opening and closing cycles. The grid sizes at these two stages are allowed to stretch and contract corresponding to the needle velocity, whereas the grid number is fixed. At the same time, the wall velocity is introduced along the needle body as a result of needle movement. The wall velocity in BFC is expressed in terms of the surface inclination angle and needle velocity in the cylindrical coordinate system. The initial and boundary conditions are expressed as follows.

Initial Conditions (at $t = 0$): At upstream of the valve (i.e., $i = 0$ to $i = i_x$, $j = 0$ to j_y, and $k = 0$ to k_v)

$$u_i = v_j = w_k = 0, \quad p = p_f, \tag{9}$$

at downstream of the valve (i.e., $i = 0$ to $i = i_x$, $j = 0$ to j_y, and $k = k_v$ to k_z)

$$u_i = v_j = w_k = 0, \quad p = 0, \tag{10}$$

where k_v is the location of the valve seat in the k coordinate and p_f is the pressure at the fuel rail, which is 270 kPa. The pressure value in present study is the referenced value with respect to the pressure at the injector exit.

Boundary Conditions (at $t > 0$): At the inlet (i.e., $i = 0$ to i_x, $j = 0$ to j_y, and $k = 0$),

$$u_i = v_j = 0, \quad p = p_f, \tag{11}$$

Note that the value of the inlet velocity w_j in this study is not given and is part of the solution.

At the outlet (i.e., $i = 0$ to i_x, $j = 0$ to j_y, and $k = k_z$),

$$u_i = v_j = 0, \quad p = 0, \tag{12}$$

At the symmetric planes (i.e., $i = 0$ and i_x, $j = 0$ to j_y, and $k = 0$ to k_z),

$$\partial u_i/\partial i = \partial v_j/\partial i = \partial w_k/\partial i, \tag{13}$$

At the valve-seat wall (i.e., $i = 0$ to i_x, $j = 0$, and $k = 0$ to k_z),

$$u_i = v_j = w_k = 0, \tag{14}$$

At the needle body wall (i.e., $i = 0$ to i_x, $j = j_y$, and $k = 0$ to k_z),

$$u_i = v_j = 0, \quad w_k = -w_d \cos(\alpha), \tag{15}$$

where w_d is the needle lift velocity and α is the inclination angle of the needle body surface in cylindrical coordinates.

The SIMPLEST algorithm and the staggered grid pattern are applied to integrate the finite-difference equations to obtain the flowfield and pressure distribution. The solvers of the PHOENICS computer code are applied to facilitate the numerical solution (Spalding, 1989).

Results and Discussion

Two different operating conditions, i.e., the valve held fully opened and the valve controlled by a pulse ranging from 0.97 to 7.5 ms, with temperatures of 20–100°C, are used to study the pulse width and temperature effects on fuel delivery at different stages.

The measured and predicted static flow rates and dynamic flows at different temperatures with a pulse of 2.5 ms are shown in Fig. 6. The M10 fluid was used as the test fluid. The predicted static flow rates in laminar flow and turbulent flow at 20°C are 1.764 g/s and 1.778 g/s, respectively, which are very close to the measured value of 1.74 g/s. At a temperature of 20–100°C, the static flow rates predicted by laminar flow and turbulent flow show similar trends and only differ by less than 2 percent. This is because the maximum Reynolds number ranges from 2500 to 3700 at the injector exit, which is not large enough to produce strong turbulence effects. Therefore, even though the flow is not totally laminar, the flow rate predicted by the laminar flow is more accurate than that by turbulent flow. In view of this, the laminar flow approach is employed for the following numerical analysis. As the temperature moves higher, the predicted static flow rate in laminar flow increases to the maximum value of 1.775 g/s at 80°C and drops to 1.754 g/s at 100°C. This could be the result of a decrease in both

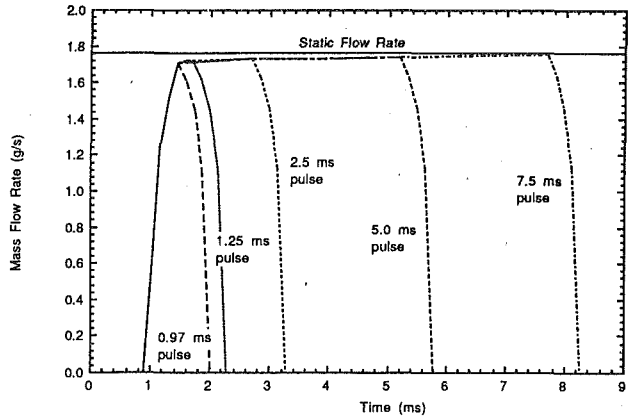


Fig. 7 Mass flow rate during fuel injection at various pulse widths with temperature of 20°C

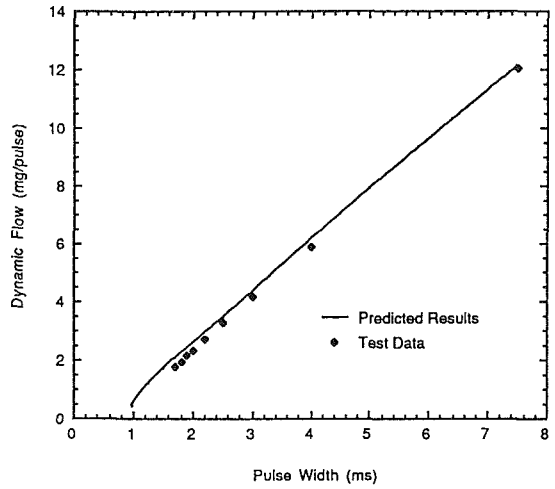


Fig. 8 Dynamic flow at various pulse widths and temperature of 20°C

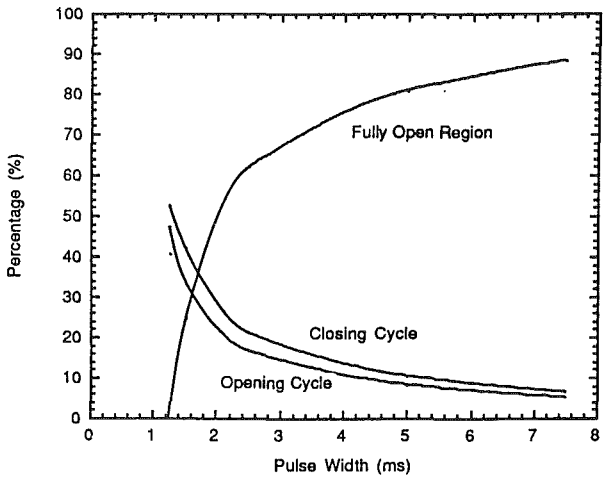


Fig. 9 Percentage of dynamic flow at various stages of fuel injection with temperature of 20°C

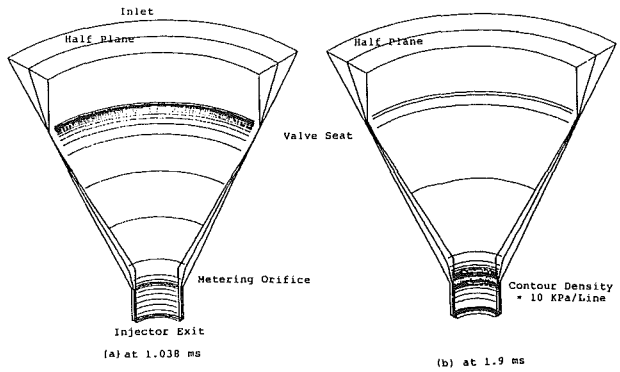


Fig. 10 Pressure distribution at the half plane of flow passage with temperature of 20°C: (*a*) at 1.038 ms with needle lift of 0.015 mm; (*b*) at 1.9 ms with needle lift of 0.06 mm

viscosity and density of the test fluid (M-10) at a high temperature. Lower viscosity tends to increase the flow velocity, whereas the lower density reduces the mass flow rate. The effects of the viscosity on mass flow rate are greater than that of the density before 80°C. Another possible mechanism accounting for the reducing flow rate at high temperatures is vaporization of test fluid. At high temperatures the vapor pressure of the test fluid is much smaller; therefore it is possible the static pressure falls below the vapor pressure. However, the present single-phase CFD analysis is not able to verify the vaporization effect. Although the predicted dynamic flow is 9–14.5 percent higher than the measured data, the overall trends of these two are very similar; they increase with temperature below 80°C and decrease above 80°C. Compared to the test data, this computational analysis predicted reasonably good results.

The predicted mass flow rate during fuel injection at various pulse widths and a temperature of 20°C is shown in Fig. 7. In the opening cycle, the flow rate goes up as the valve gradually opens. The flow rate at the fully open region increases slightly with the pulse width. At a long pulse width of 7.5 ms, the curve of transient flow rate at the fully open region finally merges to that of static flow rate. Note that the general shape of the mass flow rate curves is similar to that of needle lift, which directly controls the fuel delivery. The dynamic flow at different pulse widths with a temperature of 20°C is shown in Fig. 8. Linearity between the dynamic flow and pulse width can be observed at 1.5–7.5 ms with a slope of 0.433 mg/ms. The dynamic flow decreases quickly when the pulse width is smaller than 1.25 ms. This low-end nonlinearity is caused by a large amount of fuel delivered at the opening and closing stages. Compared with the test results, the numerical data predict good general trends and values. From Fig. 9, one can see that at short pulse widths (<2.5 ms), the percentage of dynamic flow delivered in the fully open region increases very quickly with the pulse width. About 50 percent of the fuel is injected at the fully open region when the pulse width is 2.0 ms at 20°C, but about 89 percent of the fuel at a longer pulse of 7.5 ms. Also, even when the valve opening time (Δt = 0.55 ms) and closing time (Δt = 0.58 ms) are almost the same, more fuel is delivered in the closing cycle. This is because at the beginning of the opening cycle, the pressure wave takes a certain time to move to the injector exit in order to drive the flow, but at the closing cycle, the flow field responded quickly to the needle movement.

Figures 10(*a*) and 10(*b*) show pressure distributions at the half plane of the flow passage (i.e., $j = 0.5 j_y$) at a temperature of 20°C in an early stage (1.038 ms) and in the fully open stage (1.0 ms), respectively. At 1.0375 ms with the needle lift of 0.015 mm, most of the pressure gradient occurs at the valve seat because of the narrowest flow passage. As the time goes to 1.9 ms and the needle lift is 0.06 mm, the larger pressure gradient region moves to the metering orifice because the smallest flow passage is now at the orifice and a large pressure head has been converted to kinetic energy.

Table 1 shows the temperature effect on the flow delivered at different stages of fuel injection at a pulse width of 2.5 ms. A significant 40 percent of the fuel is delivered in the opening

Table 1 Effect of temperature on dynamic flow at different stages with a pulse width of 2.5 ms

Temperature (C)	Dynamic Flow (mg/pulse)	Percentange (%) of Dynamic Flow at		
		Opening Cycle	Fully Open	Closing Cycle
20	3.467	16.82	62.0	21.18
40	3.503	17.09	61.77	21.14
60	3.568	17.23	61.57	21.20
80	3.585	17.28	61.50	21.22
100	3.549	17.48	61.27	21.25

cycle (i.e., 16.82–17.48 percent) and closing cycle (i.e., 21.14–21.25 percent). An increase in temperature tends to increase the percentages of flow delivered at the opening and closing cycles. This is because at lower fluid viscosity the fuel is easier to move through narrow valve passages at the opening and closing cycles. As the temperature increases from 20°C to 100°C, the dynamic flow at the opening and closing cycles increase by 0.66 and 0.07 percent, respectively.

Figure 11 shows the temperature effect on spatial distribution of the w_k velocity (velocity of the k component in the BFC) at the half plane of the flow passage at 1.9 ms. The injector inlet corresponds to the origin (i.e., 0) in the dimensionless unit of k/k_z and the injector exit to 1 in k/k_z. The w_k velocity increases to the first peak at the valve seat, and then drops a little in the vicinity of the metering orifice, and finally increases drastically to the maximum value of 29–32 m/s. Near the injector exit, the w_k velocity drops a little because of the sudden expanded flow passage created by the pintle. The flow velocity is maximal at a temperature of 100°C as a result of lowest fluid viscosity. When the needle is fully open at 1.9 ms, the flow velocities at the injector exit at temperature of 20, 40, 60, 80, and 100°C are 18.9, 19.3, 20.0, 20.5, and 20.6 mm/ms, which results in the maximum discharge coefficient (i.e., the ratio of the actual flow rate to theoretical flow rate) of 0.696, 0.707, 0.724, 0.735, and 0.732, respectively. These values are close to those predicted by the empirical equation for a plain orifice atomizer proposed by Nakayama (1961),

$$C_D = \frac{\mathrm{Re}^{5/6}}{17.11\ l_o/d_o + 1.65\ \mathrm{Re}^{0.8}} \tag{16}$$

where C_D is discharge coefficient at high Reynolds number, and l_o and d_o are the orifice length and diameter, respectively. From Eq.(16), one can calculate that for Re = 2500–3700 with $l_o/d_o = 4$, the discharge coefficient ranges from 0.729 to 0.754, which are slightly larger than predicted values. This is because the pintle injector has a divergent exit, which results in a larger exit hydraulic diameter and a lower ratio of l_o/d_o. However, the basic functions of the metering orifice in the pintle injector and the plain orifice in the atomizer are very similar according to the close values in discharge coefficient.

Conclusions

Based on the analysis, the following conclusions can be drawn:

1 At a temperature range of 20–100°C, both static flow rate and dynamic flow are maximal at the temperature of 80°C with a discharge coefficient of 0.735.

2 At a pulse width smaller than 2.5 ms, the percentage of the fuel delivered at the fully open region increases rapidly with the pulse width.

3 The temperature tends to increase the percentages of fuel delivered at the opening and closing cycles. More flow is delivered at the closing cycle than at the opening cycle.

4 Pressure and w_k velocity show small variation in the region between the valve seat and the metering orifice when the injector is fully open.

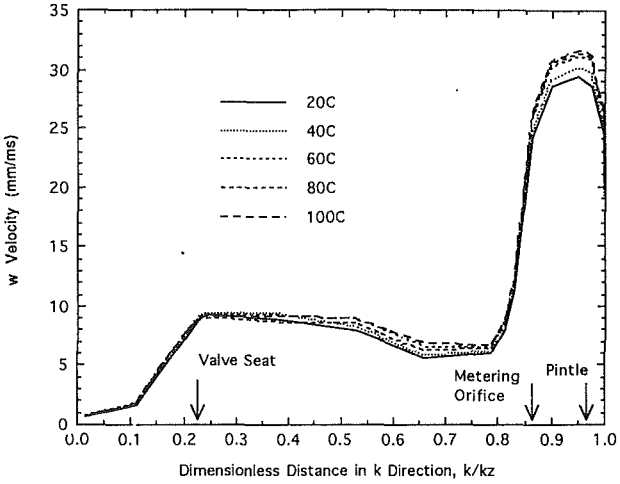


Fig. 11 Distribution of w velocity at 1.9 ms with various temperature and pulse width of 2.5 ms

References

Andrighetti, J., and Gallup, D., 1987, "Design-Development of the Lucas CAV Multipoint Gasoline Injector," SAE Paper No. 870127.

Greiner, M., Romann, P., and Steinbrenner, U., 1987, "BOSCH Fuel Injectors—New Developments," SAE Paper No. 870124.

Heywood, J. B., 1988, *Internal Combustion Fundamentals*, Chap. 7, McGraw-Hill, New York.

Hinze, J. O., 1975, *Turbulence*, McGraw-Hill, New York.

Launder, B. E., and Spalding, D. B., 1972, *Mathematical Models of Turbulence*, Academic Press, New York.

Nakayama, Y., 1961, "Action of the Fluid in the Air Micrometer: First Report, Characteristics of Small Diameter Nozzle and Orifice," *Bull. Japan Soc. Mech. Eng.*, Vol. 4, pp. 516–524.

Okamoto, Y., Arai, N., Nakagawa, K., Kosuge, T., and Atago, T., 1992, "Atomization Characteristics of Two-Stream Injector for 4-Valve Engines," SAE Paper No. 920705.

Rosten, H., and Spalding, D. B., 1986, *The PHOENICS Equations*, CHAM TR/99, CHAM Limited, London, UK, Oct.

SAE, 1992, "Gasoline Fuel Injector—SAE J1832 NOV89," SAE Handbook, Vol. 3, pp. 24.246 - 24.262, Society of Automotive Engineers, Warrendale, PA.

Spalding, D. B., 1989, *The PHOENICS Beginner's Guide*, CHAM TR/100, CHAM Limited, London, UK, Dec.

Sugiura, S., Yamada, T., Inoue, T., Morinishi, K., and Satofuka, N., 1990, "Numerical Analysis of Flow in the Induction System of an Internal Combustion Engine—Multi-dimensional Calculation Using a New Method of Lines," SAE Paper No. 900255.

Unsteady Flow Calculation of Sophisticated Exhaust Systems Using a Multibranch Junction Model

T. Bulaty

M. Widenhorn

ASEA Brown Boveri Ltd.,
Baden, Switzerland

The flow in a complex engine manifold is computed by generalizing the three-branch to the n*-branch junction model. Energy-related pressure losses are implemented in the one-dimensional continuity equation and the effects of choked flow and changing flow patterns are considered. The computational results are compared with measurements on an engine with multi-entry pulse converter and compact single pipe exhaust systems. The influence of several parameters on the fuel consumption is determined.*

Introduction

There are two basic solutions of the conservation laws (mass, energy, and momentum equations) for the one-dimensional unsteady flow: characteristics and differential methods. The computational advantages of the differential method are well known and it is becoming the state of the art. It utilizes Taylor's series, for the conservation laws, where differentials are replaced by differences. Usually two terms of the series are used, and therefore the method is classified as second-order accurate.

Many of those methods can be related to the family of the Lax–Wendroff method. These methods can inherently handle discontinuities but cause small under- or overshoots (explanation in Bulaty and Niessner, 1985). Pressure shocks are quite unusual in intake or exhaust systems. But to avoid the errors in the calculated turbine energy, caused by distortion of temperature lines for scavenged supercharged engines, the flux correction must be introduced (Niessner and Bulaty, 1981).

The characteristic values necessary for the solution of boundary conditions are usually obtained by linear interpolation between mesh points at the previous time step. Hence, the solution has only first-order accuracy at the boundaries.

Computer Model of a Multibranch Junction

The intake and exhaust system of any multicylinder IC engine always contains junctions. They connect pipe elements and mutually influence each other. They also modify the flow and therefore the pipe affects the cylinder. To calculate those phenomena it is necessary to simulate the junctions with high accuracy while calculating the one-dimensional unsteady flow. This is very difficult because of the very complex flow pattern in the junction, especially when the number of branches exceeds three. The junctions are probably the worst-modeled elements in present day programs.

Models. The simplest model neglects losses in the junction (Benson et al., 1963–4) and thus simplifies the calculation of the characteristics at a boundary. The extension of this model from three to n branches causes no problems (Woollatt, 1965). However, the model is suitable only for very small losses, i.e., at very low flow velocities and for the same entropy level in all branches. Volume models with small volume (Görg, 1982) or virtual volume with fictive pressure and temperature (Corberán, 1992) have only very limited applicability. Models with two momentum equations require one (Benson et al., 1963–64) or even two (Bingham and Blair, 1985) empirical coefficients. These are quite difficult to measure and the whole solution is numerically sensitive. The use of flow coefficients, for nozzle-type flow between two branches (Seifert, 1978), cannot describe ejector effects in pulse converters. When used in a more common form (Linnhoff et al., 1987) another empirical component of one momentum equation is needed. The generalization of these models to more branches is almost impossible.

Relative pressure losses are used as a function of different values: dynamic head (Watson and Janota, 1971), Mach number M (Dadone, 1973; Alexander, 1975) or M square (Chan, 1985). There are also versions using a hydraulic analogy (Morimune et al., 1981) or modified constant pressure (Takizawa et al., 1982). The Mach number or mass flow dependencies are in good agreement with measurement, which can be performed with reasonable effort. If necessary, they can be generalized, as we show later.

The last group of methods shown are quite similar, expressing pressure losses dependent on M or mass flow ratios. While the results are also very comparable, the methods should be evaluated according to the following criteria:

- flexibility for numerical application

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1992. Paper No. 93-ICE-14. Associate Technical Editor: J. A. Caton.

- difficulty in matching the results to experimental data
- complexity appropriate to sophisticated geometries
- computing resources required
- existing application and available experimental data

Our evaluation suggests that the method of pressure losses is the most suitable. It delivers good results with minimal effort.

Watson's model (1970) was selected, in the midseventies, and used at the Turbocharging department of BBC (now ABB Turbo Systems Ltd.). During further development, we applied Dadone's (1973) formulation and dependency on M. The database of pressure losses was composed from published data and enhanced with in-house measurements. The modern supercharging systems include not only pulse converters with ejector effects but also multi-entry pulse converters and compact single pipe pulse systems. Their successful simulation required development of an extended model (Bulaty, 1982, 1988; Bulaty and Niessner, 1985).

Three-Branch Model. Depending on the flow direction in n branches, there are many (2^n–2) different flow patterns (FP) possible. The three-branch junction has FP = 6, consisting of three joining and three dividing flows. To identify the FP, we classify the flow according to total pressures in the branches. However, the possibility of the flow pattern switching, during the necessary iterations, must be kept in mind.

The method of characteristics is almost always applied at the boundaries. Hence, it is convenient to formulate the continuity and energy equations with the same dimensionless variables as those used for characteristics. The two missing conditions are the pressure losses between two branches. Their choice is either related to the common flow branch or to the highest or lowest characteristic value, respectively.

Dadone (1973) extensively measured T-junctions (of same diameter and one straight pipe) and related the losses to M in the common pipe. Reformulating his formula we can see the physical meaning: The pressure loss of the branch L is expressed as the part of the dynamic head in the common pipe C (kappa being isentropic gas exponent). This is referred to as the energy related pressure loss (ERPL):

$$\mathrm{ERPL} = (p_L/p_C - 1)/\mathrm{M}_C^2 = -\mathrm{kappa}/2 * (p_C - p_L)/p_{\mathrm{dyn}C}$$

The measured loss dependencies are replaced by parabolic coefficients with the help of curve fitting. Figure 1 shows ERPLs for a symmetric pulse converter. Our program has built in T-junction loss coefficients, but permits any other geometry to be input.

There is a complex formula, if pressure losses are implemented into the continuity equation. It can be very effectively solved by using the modified regula-falsi method. The pressure losses of the previous step are used as the initial value. With the losses and characteristic values on the boundary, all the state values in pipes are calculated and reiterated. Convergence is obtained when the difference to the previous step is within the demanded accuracy, and no further flow pattern change occurs.

***n*-Branch Model.** It is possible to reduce the measurement effort significantly by utilizing the knowledge of the flow situations. In engine applications, joining flow prevails in the exhaust while dividing flow occurs in the intake systems. Analyzing the most frequent patterns and those with longest durations, the complex junction can always be reduced to the joining or dividing flow in three pipes. Hence, we look for pipes with the two highest and the lowest characteristic value in the exhaust or the highest and two lowest values in the intake system. For further branches, we focus on the two most important branches and systematically vary the remaining branch, always solving only three-branch junction equations. It is very logical to analyze pipes accordingly and, therefore, not surprising that this idea was later independently used by others (Bingham and Blair, 1985).

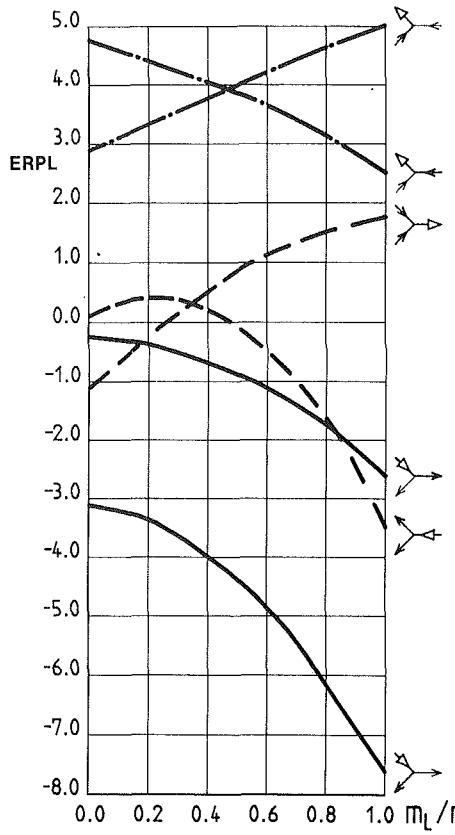


Fig. 1 Energy-related pressure losses (ERPL) of a symmetrical pulse converter

The highly supercharged engines began using multi-entry pulse systems at the end of the sixties. Because of the described complexity of FP and measurement prerequisites, the realistic calculation appeared later. The usual five-branch space symmetric multi-entry pulse converter has 30 FPs and requires solution of four to eight equations! After eliminating nonphysical FPs and utilizing the symmetry, eight FPs still remain (Fig. 2). Hence, we analyzed some engine applications on FPs, their duration and strength as well as their influence on the engine results. The goal of the pulse converter is to prevent wave reflection at the turbine entry and backflow into cylinders during the scavenging, but not, as often believed, to utilize ejector effects. Therefore, the pulses running to the turbine are much stronger than those running back. In fact there are only one or two pipes where backflow simultaneously prevails. Also there is a unique flow direction in the so-called mixing pipe leading to the turbine. We use a specific queuing algorithm, to classify the mixing pipe, the jet pipe (with the strongest flow) and side pipes (all others) and thus sequentially always solve only a three-branch junction (Fig. 3).

Special Problems. While measuring the first pulse converter models, we observed rising losses, without higher throughput, and detected the presence of the critical flow. This prompted the development of a numerical method to predict and compute the choked flow, as well as to determine the effective throat area. The program was thus enhanced to calculate the limit of the influence of the downward pressure. This problem often occurs when calculating compact single pipe pulse systems.

The smooth switching between different FPs is quite important for stable numerical solution. This is difficult to calculate for fitted curves of ERPLs and only iteratively possible (van Damme, 1986).

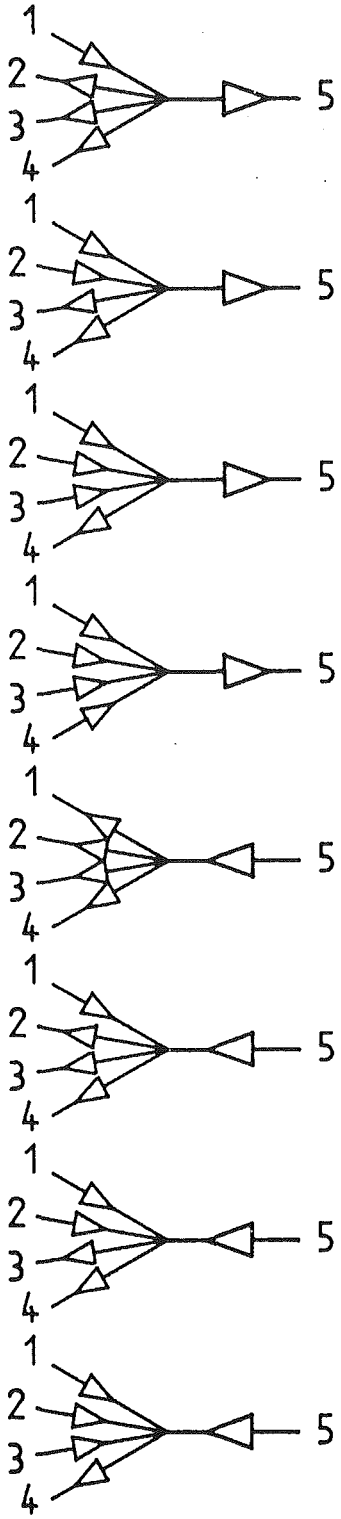


Fig. 2 Flow patterns in a five-branch junction

It is often necessary to calculate geometries that have not yet been measured. Then we interpolate within the matrix of systematically performed measurements. Using also relative pipe areas, it is frequently possible to recalculate losses for small deviations from the measured geometry. Larger deviations as well as extrapolation can be dangerous because of numerical instabilities.

The fitted curves can lead to nonphysical losses at very low flow velocities. We take this into account by calculating only above a certain M number in the pipe.

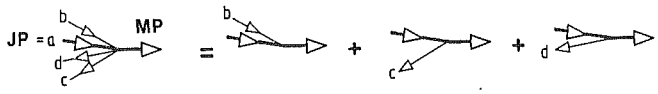


Fig. 3 Reduction of the *n*-branch to three-branch junction (JP = jet pipe, MP = mixing pipe)

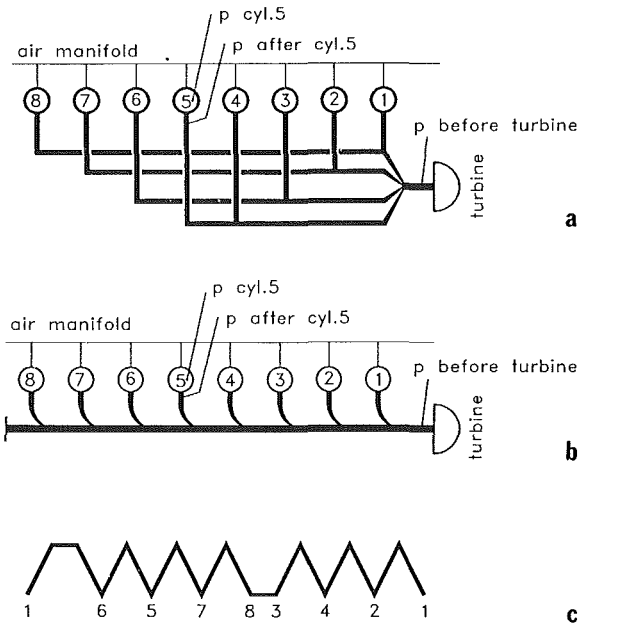


Fig. 4 Pipe systems of the multipulse (*a*) and the compact single pipe pulse system (*b*) as well as firing order (*c*)

The numerical damping necessary in each step, as well as the loss changes, are used to accelerate convergence. Further, an upper limit, for the maximum number of iterations, prevents large computing times when oscillations occur.

Comparison of the Calculations with Measurements

Measurements on an engine (V16, 1800 rpm, bore 170 mm, bmep = 18 bar), having two different exhaust systems (Fig. 4), are compared with the results obtained by our program system, using the junction model described above. Because the cylinder banks of the *V* engine are thermodynamically disconnected, we can simulate only bank (*B*) having an uneven firing order according to Fig. 4(*c*).

Multi-entry Pulse Converter Exhaust System. In this system, cylinder pairs are connected to common pipes via *T*-junctions, all of the same diameter. These relatively narrow pipes are interconnected in a so-called multi-entry pulse converter, just before turbine entry (Fig. 4(*a*)). This arrangement should prevent reflected waves from disturbing the scavenging. The turbine entry flange area is 1.7 times the cross section of the pipes, while in the converter the nozzle-to-pipe area ratio is only 0.7. This geometry was in our database of measured junctions. The pressure was measured in and after the cylinder No. 5, before turbine and in the air receiver (Fig. 5). The negligible pressure fluctuations in the receiver justify the assumption of constant pressure in the inlet. The agreement between measured and calculated pressure in the cylinder is good. The signal resulting from the valve impact on the seat at the exhaust end of course cannot be calculated. Further, the pressure after the cylinder can be well reproduced to recognize all important effects. The waves at the turbine entry are well described too, only the systematic deviation could not be explained.

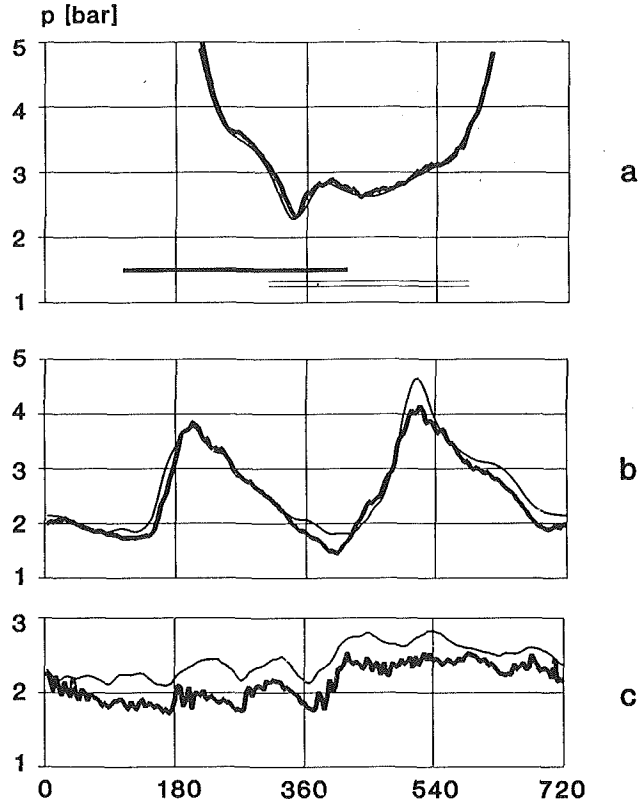


Fig. 5 Multipulse system measurements (thick line) compared to the calculation (thin line): (*a*) pressure in cylinder, (*b*) after cylinder, (*c*) before turbine entry

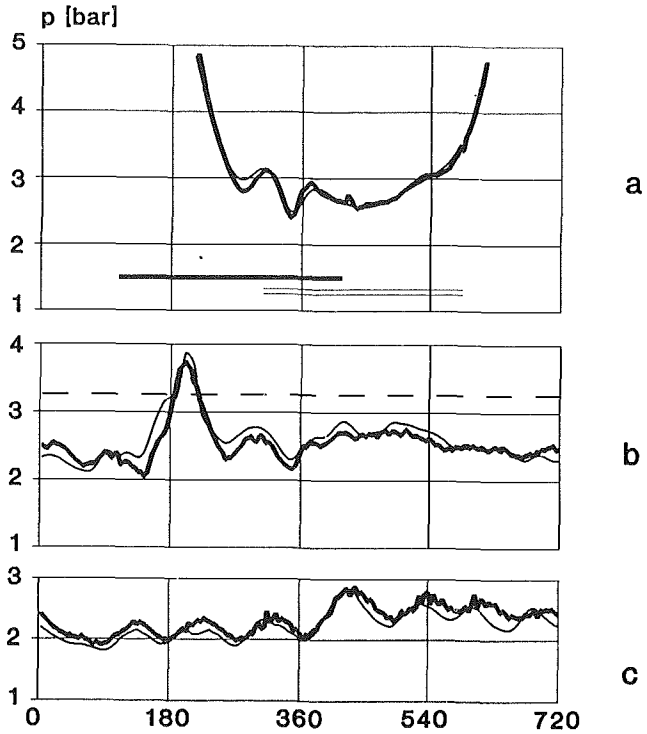


Fig. 6 Compact single-pipe pulse system measurements compared to the calculation (comments see Fig. 5)

Compact Single-Pipe Pulse System. Each cylinder discharges via port and elbow pipe into a elbow nozzle, before entering the larger single common manifold, Fig. 4(*b*). There is no area change until the turbine nozzle ring. The nozzles have 75 percent of the elbow cross-sectional area and enter the single pipe at approximately 30 deg. The area of the single pipe is 2.5 times that of the elbows. We had measured ERPLs for this geometry, too. Again the agreement of the cylinder pressure is very good and valve impacts of exhaust and intake closing can be seen (Fig. 6). There is an important effect, in the pressure after the cylinder, near bottom dead center (Fig. 6(*b*)). This pressure rise is caused by critical flow in the downward nozzle. Using the effective nozzle area from the blow test rig in the program, results in good reproduction by the calculation. In Fig. 6(*c*) we can observe very good coincidence in form and even level for the pressure at the turbine entry.

These comparisons of measurements and calculation prove the ability of our program to correctly calculate even quite sophisticated exhaust pipe systems. This is facilitated by special treatment of the unsteady flow calculation (Niessner and Bulaty, 1981, 1986) and by thorough specification of model data.

Single-Pipe System Geometry Variation

First we varied two of the most important parameters of the single pipe system, i.e., (1) the geometric nozzle to single pipe and (2) elbow to single pipe area ratios, while keeping the angle of elbow entry unchanged at 30 deg. The range of the variation was limited by available junction model measurements. This enabled us to use two single-pipe diameters of 100 and 120 mm and three nozzle ratios of 60, 75, and 90 percent. The measured ratio of the effective nozzle to elbow cross-sectional area as well as elbow area and engine data were kept constant during the whole analysis.

Operation points considered were the nominal and 40 percent power, corresponding to propeller law. For each geometry, the scavenging pressure was adjusted by change of the turbine nozzle area to obtain comparative maximum cylinder pressures. But, at part load, the turbine area of each corresponding nominal point was held constant.

When a single pipe system is chosen, the goal of the optimization is to reach uniform scavenging and thermal load of all cylinders with minimum exhaust manifold volume. This must be attained without penalty in fuel consumption, when compared with the constant pressure system. The first parameters are evaluated by total air throughput, air/fuel ratio, and temperatures after cylinder. The calculated differences of the former parameters were relatively small so only changes of the specific fuel consumption will be presented.

The estimation of the engine friction losses at partial load, and hence the fuel consumption, are slightly uncertain; but the relative changes are correct. In Fig. 7(*a*), it is related to the full-load consumption for the smaller diameter of the single pipe. The use of the larger single-pipe diameter shows an improvement over the whole range. The fuel consumption depends strongly on the nozzle to pipe area ratio at full load but shows less variation at part load. Fuel consumption is lower at larger nozzle areas but increasing the nozzle pipe area beyond approx. 0.75 results in only marginal improvement.

Therefore, we also calculated the system with the narrow single pipe but no constriction in the junctions. The promising results were extended by further variation of the entry angle of these junctions (Fig. 7(*b*)). The calculations show that there would be an improvement in the fuel consumption and other important values, if the unconstricted elbows entered at a smaller angle into the single pipe.

When we show the advantage of the 75 percent or unconstricted cases, it must be emphasized that all calculations only evaluate steady-state operation. There might be another optimum resulting from transient consideration. Also the results and their interpretation may vary when another firing order, number of cylinders or geometric engine dimensions were used. Nevertheless, the computer simulation is a great help when optimization by parameter variation should be made. The oc-

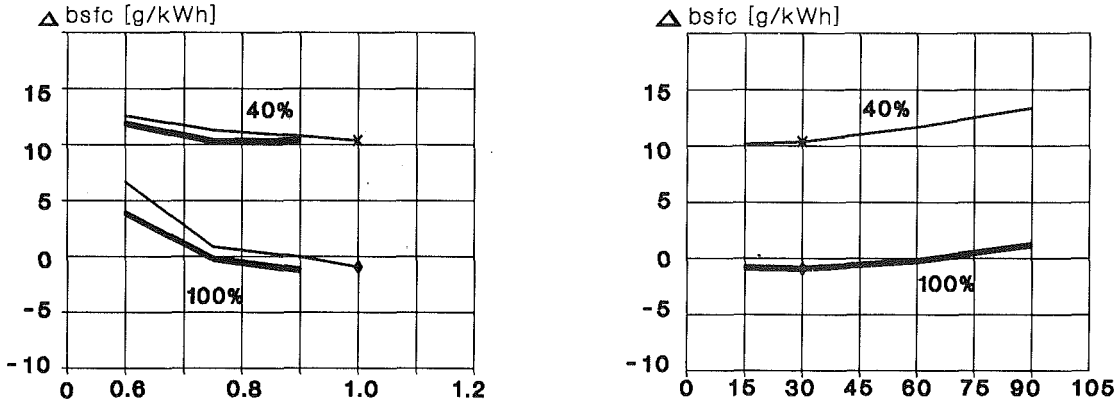


Fig. 7 Fuel consumption dependence on nozzle area constriction (*a*) and deflection angle (*b*) at full and part load for 120 mm (thick line) and 100 mm (thin line) manifold diameter

casional comparison between the calculation and measurements improves the usefulness of this powerful tool.

Conclusions

It is possible to generalize a three-branch to an n-branch junction model and successfully use it for the one-dimensional unsteady flow calculation in sophisticated exhaust systems. Accounting for chocked flow, excellent agreements between calculation and measurements were obtained.

References

Alexander, G. I., 1975, "An Investigation of Pulse Converters in a Turbocharged Diesel Engine," PhD Thesis, University of Manchester, United Kingdom.

Benson, R. S., Woollatt, D., and Woods, W. A., 1963-64, "Unsteady Flow in Simple Branch System," *IME*, Vol. 178, Pt. 3I, pp. 24-49.

Bingham, J. F., and Blair, G. P., 1985, "An Improved Branched Pipe Model for Multi-cylinder Automotive Engine Calculation," *Proc. IME*, Vol. 199D1, pp. 65-77.

Bulaty, T., 1982, "Ein Programmsystem zur Berechnung des Zusammenwirkens von Verbrennungsmotoren und Abgasturboladern," *MTZ*, Vol. 43, pp. 535-543.

Bulaty, T., and Niessner, H., 1985, "Calculation of 1-D Unsteady Flows in Pipe Systems of I.C. Engines," ASME *Journal of Fluids Engineering*, Vol. 107, pp. 407-412.

Bulaty, T., 1988, "The Use of Computer Simulation to Improve Turbocharging of Diesel Engines," ASME Book No. GOO465, pp. 327-33.

Chan, C. L., 1985, "An Investigation of the Performance of a Diesel Engine Fitted With Compact Manifold," PhD Thesis, University of Manchester, United Kingdom.

Corberán, J. M., 1992, "A New Constant Pressure Model for N-Branch Junctions," *Proc IME*, Vol. 206, Paper No. D02191.

Dadone, A., 1973, "Perdite di carico nelle giunzioni," *ATA*, Vol. 264, pp. 214-224.

Görg, K. A., 1982, "Berechnung instationärer Strömungsvorgänge in Rohrleitungen an Verbrennungsmotoren unter besonderer Berücksichtigung der Mehrfachverzweigung," Diss. Ruhr Univ. Bochum, Heft 82.8, Germany.

Linnhoff, H.-J., Peters, A., and Schindler, P., 1987, "Rechnerische Simulation der Injektorströmung am Turbinenbypass eines aufgeladenen Dieselmotors," *MTZ*, Vol. 48, pp. 395-402.

Morimune, T., Hirayama, N., and Maeda, T., 1981, "Study of Compressible High Speed Gas Flow in Piping System," *Bull. JSME*, Vol. 24, pp. 2082-2089.

Niessner, H., and Bulaty, T., 1981, "A Family of Flux-Correction Methods to Avoid Overshoot Occurring With Solutions of Unsteady Flow Problems," *Proc. of GAMM 4th Conference*, pp. 241-250.

Niessner, H., and Bulaty, T., 1986, "Improving Conservation of Mass, Momentum and Energy by Appropriate Flux Correction on Boundaries," *ZAMM*, Vol. 66, T. 250-251.

Seifert, H., 1978, "Erfahrungen mit einem mathematischen Modell zur Simulation von Arbeitsverfahren in Verbrennungsmotoren," *MTZ*, Vol. 39, pp. 321-326.

Takizawa, M., Uno, T., Oue, T., and Yura, T., 1982, "A Study of Gas Exchange Process Simulation of an Automotive Multi-cylinder I.C. Engine," SAE Paper No. 820410.

van Damme, G., 1986, "Numerische Simulatie van de Multipulse Converter," Diss. Rijksuniversiteit Gent, Belgium.

Watson, N., 1970, "Pulse Converters for Turbocharged Diesel Engines," PhD Thesis, London University, United Kingdom.

Watson, N., and Janota, M. S., 1971, "Non-steady Flow in an Exhaust System With Pulse Converter Junction," *Symposium on Internal Flows*, University of Salford, United Kingdom, Paper 27, pp. D17-28.

Woollatt, D., 1965, "An Approximate Theory for the Transmission and Reflection of Simple Waves at Area Change and Junctions in Pipes," *Int. J. Mech. Sci.*, Vol. 7, pp. 777-783.

D. Assanis

E. Karvounis

University of Illinois at Urbana-Champaign, Urbana, IL 61801

R. Sekar

W. Marr

Argonne National Laboratory, Argonne, IL 60439

Heat Release Analysis of Oxygen-Enriched Diesel Combustion

A heat release correlation for oxygen-enriched diesel combustion is being developed through heat release analysis of cylinder pressure data from a single-cylinder diesel engine operating under various levels of oxygen enrichment. Results show that standard combustion correlations available in the literature do not accurately describe oxygen-enriched diesel combustion. A novel functional form is therefore proposed, which is shown to reproduce measured heat release profiles closely, under different operating conditions and levels of oxygen enrichment. The mathematical complexity of the associated curve-fitting problem is maintained at the same level of difficulty as for standard correlations. When the novel correlation is incorporated into a computer simulation of diesel engine operation with oxygen enrichment, the latter predicts pressure traces in excellent agreement with measured pressure data. This demonstrates the potential of the proposed combustion simulation to guide the application of oxygen-enriched technology successfully to a variety of multicylinder diesel systems.

Introduction

Argonne National Laboratory (ANL) has been studying the application of oxygen-enrichment technology to diesel engines, especially for industrial cogeneration systems (Cole et al., 1990; Sekar et al., 1990a, 1990b). An initial assessment was based on a literature survey and a modified simulation code (Assanis et al., 1990). The study indicated that use of oxygen-enriched combustion air can lead to significant improvements in power density, as well as reductions in particulate emissions, but at the expense of higher NO_x emissions. Oxygen enrichment would also lead to lower ignition delays, and the potential for burning lower grade fuels. Strategies for introducing water into the combustion chamber, such as in the form of fuel emulsions, could be used to control NO_x emissions. Therefore, analytical and experimental studies were undertaken to establish the optimal combination of oxygen level and diesel fuel properties.

Experimental work has been conducted on both single-cylinder and multicylinder, heavy-duty diesel engines (Sekar et al., 1990a, 1990b). Apart from demonstrating the viability of oxygen-enriched diesels, experimental results are being used to develop a reliable computer simulation of the oxygen-enriched diesel engine. Such a tool would be very useful in predicting the performance and emissions of full-size cogeneration systems, using larger diesel engines, as opposed to conducting time-consuming and expensive experimentation. The success of such simulation codes would, to a large extent, depend on their capability to predict changes in the combustion process due to oxygen enrichment.

A convenient method for analyzing engine combustion is through heat release analysis of cylinder pressure data. While heat release models do not generally provide spatial resolution, the global view they present still provides valuable insight into the nature of the combustion process under different engine conditions. For the purposes of engine simulation, combustion can then be modeled as a heat addition process, the rate of which is given by an empirical formula. The functional form of heat release relations is chosen to match model predictions with experimental data. Coefficients in these formulas, which may vary with engine design details and operating conditions, are determined empirically.

Two heat-release correlations, which have commonly been used in simulations of conventional diesel engines, have been proposed by Watson et al. (1980) and Miyamoto et al. (1985). The objective of the present work is to explore whether one of these correlations, or an alternative functional form, with appropriately tuned constants, can successfully reproduce heat release profiles characteristic of oxygen-enriched diesel combustion. The study will be based on comparisons of correlation-generated heat release profiles against the ones obtained from measured pressure data. Additionally, alternative heat release correlations will be incorporated into a computer simulation of the oxygen-enriched diesel engine. Therefore, the potential of the simulation to predict measured cylinder pressure traces will be assessed in each case.

Review of Standard Heat Release Correlations

Watson's Model. In this representation, diesel combustion is described as a rapid premixed burning phase followed by a slower, mixing-controlled burning phase. The most commonly used correlation of this kind has been, up to date, that proposed by Watson et al. (1980). In this approach, one algebraic function is used to describe the premixed heat-release phase, and a second function describes the mixing-controlled heat release phase. These two functions are weighed with a phase propor-

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31-February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1992. Paper No. 93-ICE-8. Associate Technical Editor: J. A. Caton.

tionality factor β, which is largely a function of the ignition delay. Thus

$$\frac{m_{f,b}(\tau)}{m_{f,0}} = \beta f_1 + (1-\beta) f_2 \tag{1}$$

where $m_{f,b}(\tau)$ is the mass of fuel burned, $m_{f,0}$ is the total mass of fuel injected, and τ is time from ignition nondimensionalized by the total Δt_{comb} allowed for combustion

$$\tau = \frac{t - t_{\text{ign}}}{\Delta t_{\text{comb}}} \tag{2}$$

t being the time of interest and t_{ign} the ignition time. The premixed-burning function is

$$f_1 = 1 - (1 - \tau^{K_1})^{K_2} \tag{3}$$

while the mixing-controlled function is

$$f_2 = 1 - \exp(-K_3 \tau^{K_4}) \tag{4}$$

with K_1, K_2, K_3, and K_4 being empirical coefficients. The proportionality factor β is related to the ignition delay t_{id} through

$$\beta = 1 - \frac{a\phi^b}{t_{id}^c} \tag{5}$$

where ϕ is the overall equivalence ratio and a, b, and c are empirical constants.

The values of constants K_1, K_2, K_3, K_4 depend on engine combustion chamber design and operating conditions. The following ranges for a, b, and c have also been suggested:

$$0.8 < a < 0.95 \quad 0.25 < b < 0.45 \quad 0.25 < c < 0.5 \tag{6}$$

A relatively long nominal combustion duration of $\Delta t_{\text{comb}} = 125°$ was proposed by Watson et al. (1980). It should be noted here that Eq. (1) represents the cumulative heat release. The heat release rate is obtained by differentiating Eq. (1) as follows:

$$\frac{\dot{m}_{f,b}(\tau)}{m_{f,0}} = \beta K_1 K_2 (1-\tau^{K_1})^{K_2-1} \tau^{K_1-1} + (1-\beta) K_3 K_4 \tau^{K_4-1} \exp(-K_3 \tau^{K_4}) \tag{7}$$

This is the correlation that is finally fitted to the experimentally determined heat release profiles.

Double Wiebe Function Model. Another approach to the mathematical description of diesel engine combustion rate is that of the double Wiebe function correlation (Miyamoto et al., 1985). Again the heat release period is divided into two subsequent phases. However, both phases are now fitted with Wiebe functions:

$$\frac{\dot{m}_{f,b}(\theta)}{m_{f,0}} = 6.9 \frac{Q_p}{\theta_p} (M_p+1) \left(\frac{\theta}{\theta_p}\right)^{M_p} \times \exp\left[-6.9\left(\frac{\theta}{\theta_p}\right)^{M_p+1}\right] + 6.9 \frac{Q_d}{\theta_d} (M_d+1) \left(\frac{\theta}{\theta_d}\right)^{M_d} \times \exp\left[-6.9\left(\frac{\theta}{\theta_d}\right)^{M_d+1}\right] \tag{8}$$

where θ is the crank angle after ignition, and six parameters—Q_p, θ_p, M_p, Q_d, θ_d, and M_d—are available for fitting. The increased number of adjustable parameters provides for greater flexibility and, thus, the potential for higher fitting success.

It should be noted here that, like Eq. (7), Eq. (8) resulted from the differentiation of a corresponding cumulative heat release formula:

$$\frac{m_{f,b}(\theta)}{m_{f,0}} = Q_p \left\{1 - \exp\left[-6.9\left(\frac{\theta}{\theta_p}\right)^{M_p+1}\right]\right\} + Q_d \left\{1 - \exp\left[-6.9\left(\frac{\theta}{\theta_d}\right)^{M_d+1}\right]\right\} \tag{9}$$

Table 1 Test engine specifications

Bore	137 mm
Stroke	165 mm
Displacement	2.44 L
Compression ratio	14.5
Engine Speed	1800 rpm
Injection Timing	33° btdc

It can be easily shown that the cumulative heat release profile described by Eq. (9) is explicitly normalized to unity if

$$Q_p + Q_d = 1 \tag{10}$$

Miyamoto et al. (1985) stated that the sum of Q_p and Q_d should be equal to the total heat released (if not normalized). However, such a consistency is not apparent in their plots, and their fitting methodology was not clarified. Therefore, in our work, an additional degree of freedom is introduced in the fitting procedure by computing Q_p and Q_d independently. If the fitting is successful, the resulting profile should be normalized. On the other hand, failure to recover the total heat release value would indicate poor fitting. Note that if Q_p and Q_d were instead computed subject to the constraint of Eq. (10), normalization would have been guaranteed, but the fitting would be even less accurate.

Curve-Fitting Heat Release Profiles Using Standard Correlations

A single-cylinder, four-stroke, direct-injection diesel engine was run with different oxygen levels (from 21 to 35 percent by volume) in the intake air. The major specifications of the engine, which is commonly employed in heavy-duty truck and off-highway applications, are given in Table 1. For each oxygen level, engine performance and emissions were recorded at a speed of 1800 rpm. Cylinder pressure diagrams were obtained with an AVL pressure transducer, an optical encoder, and a PC-based data-acquisition system developed at the University of Illinois. Cylinder pressure diagrams obtained from the experiments were used to generate heat release rate diagrams as a function of oxygen content and operating conditions. The heat release code that was used treats the combustion chamber contents as uniformly mixed and accounts for the effects of heat loss, crevice volume, and in-cylinder motion. More details on the heat release code are given by Bonne (1989).

As a starting point, it was explored whether the two standard heat release correlations could fit experimental heat release profiles. The Levenberg–Marquardt method has been employed as the basis for every fitting process. A detailed description of this method can be found from Press et al. (1989). This technique, however, cannot be utilized as a stand-alone algorithm to curve-fit heat release profiles. A typical diesel engine burning profile is characterized by the existence of two subsequent maxima, preceded and followed by exponentially rising and falling curves. Especially in the case of the double Wiebe function, where the two functionals are identical, any attempt to attack the fitting problem directly leads inherently to singularities. To overcome these difficulties, a fitting algorithm was devised to distinguish between the two maxima, and correspondingly fit the two functionals to each one of them.

Examples of attempts to describe experimental heat release profiles using Watson's correlation are given in Figs. 1(*a*)–1(*d*). Heat release rates have been normalized with respect to the cumulative heat released by the end of combustion. The operating conditions pertaining to the engines tests referenced in the above figures are summarized in Table 2.

Table 3 summarizes the optimum values of the parameters in the Watson correlation for each of the above tests. It should be noted that the values of these parameters are different from the ones recommended by Watson et al. (1980) for conventional

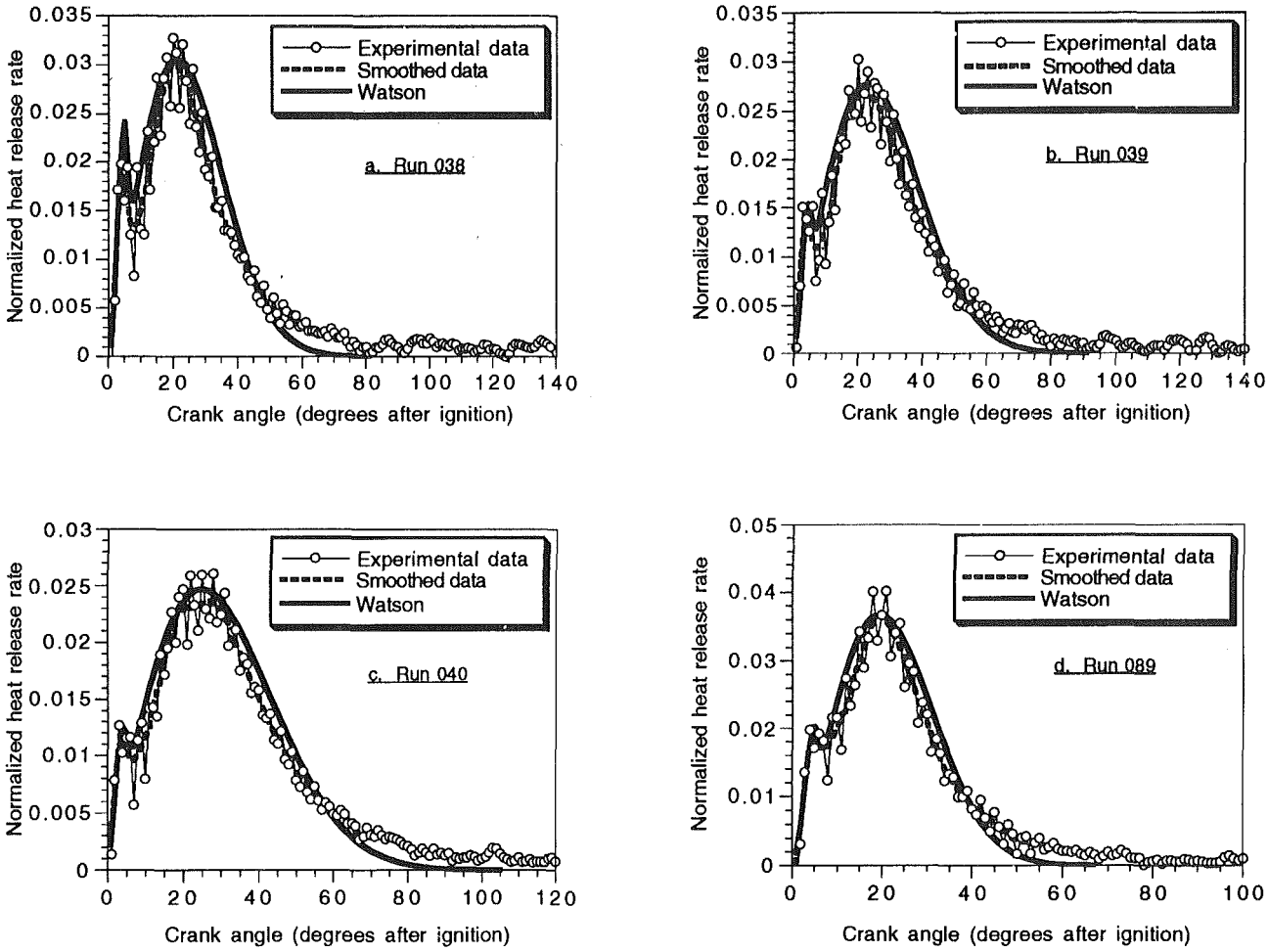


Fig. 1 Comparison of experimental heat release rates against curve-fitted profiles generated using Watson's correlation. Operating conditions are defined in Table 2.

Table 2 Engine running conditions for the fitting examples

	Run 038	Run 039	Run 040	Run 089
O_2(%)	25	25	25	30
O_2/Fuel	8.07	6.52	5.51	10.3
Equiv. ratio	0.41	0.51	0.61	0.32
Power (BHP)	50.1	60.6	70.7	50.0

Table 3 Parameters of Watson's correlation

	Run 038	Run 039	Run 040	Run 089
β	0.05485	0.02828	0.02320	0.03632
K_1	3.689	3.532	3.091	3.603
K_2	2.216×10^5	1.674×10^5	4.051×10^4	6.644×10^4
K_3	32.42	24.76	15.20	29.38
K_4	2.180	2.122	2.054	2.309
Δt_{comb}	123°	123°	130°	107°

diesel engines. Moreover, additional flexibility was introduced in the original correlation by replacing the typical combustion duration of $\Delta t_{comb} = 125°$ with a value more closely related to the experimentally determined heat release profiles. Starting from the time after the start of injection when the heat release rate becomes positive, the latter is integrated to give the cumulative heat release. Δt_{comb} is then the time elapsed until the computed integral does not change in magnitude more than a prescribed small amount between two subsequent time steps.

Clearly, the second functional always fails to fit the majority of the diffusion burning phase. It is concluded that the functional form of Watson's correlation is not flexible enough so as to fit oxygen-enriched combustion profiles. However, the explicit normalization of the model functions f_1 and f_2 (when integrated over the entire combustion period, each yields a unit area) is advantageous. When implemented into a zero-dimensional engine cycle simulation, the fitted curve will produce a cumulative heat release value equal to the actual one, regardless of the success of the fitting process.

Figures 2(*a*)–2(*d*) illustrate how double Wiebe functions fit the same set of measured heat release profiles, using the parameter values given in Table 4. While the double Wiebe function model describes the smoothed heat release curves better than Watson's model, it also fails to fit the combustion tail. Moreover, because of the independent weighing of the two functionals (through Q_p and Q_d) the combined function is not

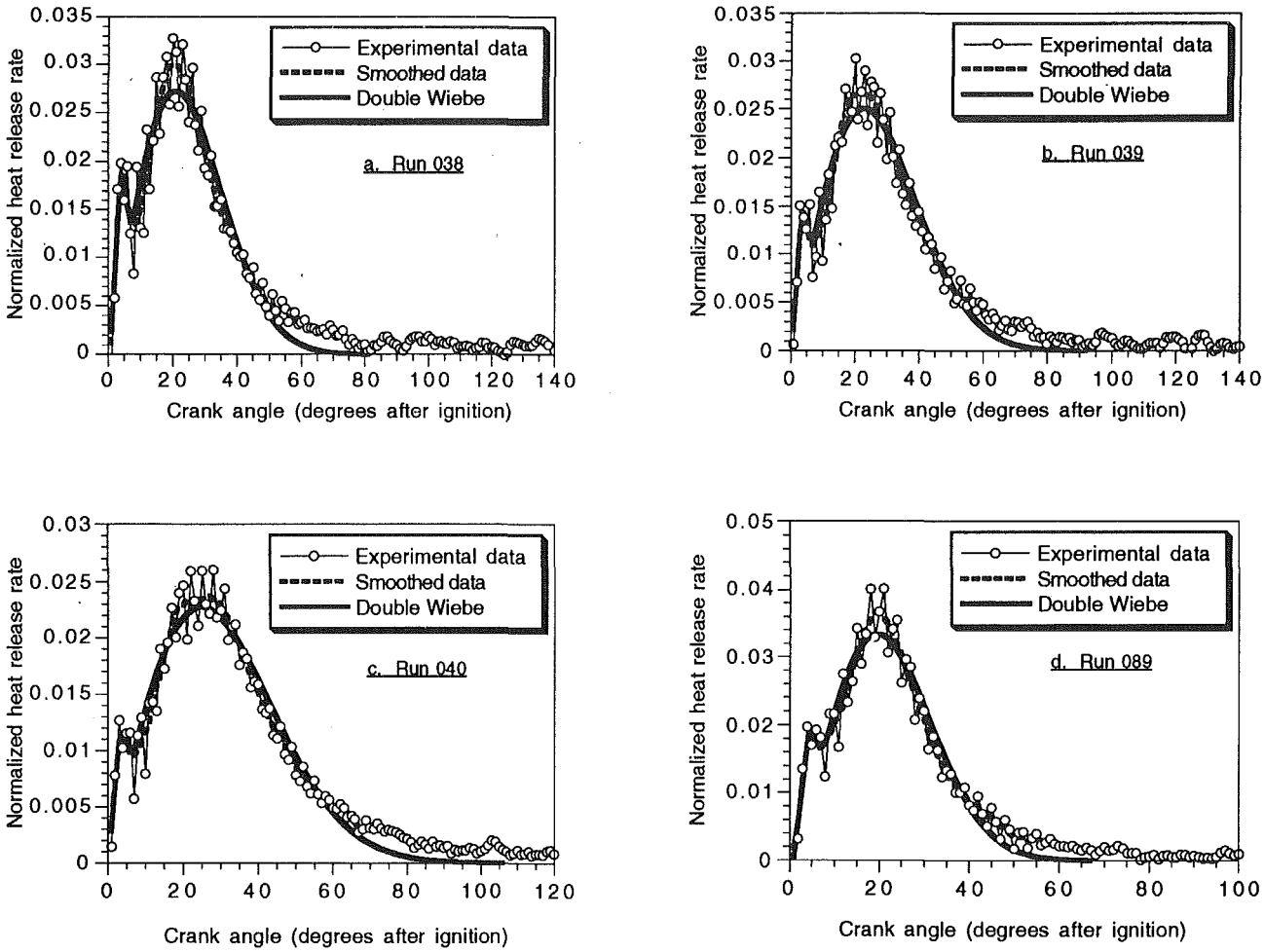


Fig. 2 Comparison of experimental heat release rates against curve-fitted profiles generated using double Wiebe correlation. Operating conditions are defined in Table 2.

Table 4 Parameters of the double Wiebe correlation

	Run 038	Run 039	Run 040	Run 089
Q_p	0.04678	0.03229	0.02528	0.03959
M_p	2.301	2.040	1.765	2.502
θ_p	8.241	7.920	7.912	8.538
Q_d	0.9546	0.9664	0.9763	0.9614
M_d	1.151	1.095	1.037	1.302
θ_d	69.23	78.77	89.64	58.27

explicitly normalized. Had the fitting been accurate, this should not present any difficulties in potential modeling applications since the experimental heat release profile is already normalized. However, if the fitting is poor, the fitted function will most likely not have a total unit area. This will result in a predicted cumulative heat release, which is different from the experimental one.

Formulation of Novel Heat Release Correlation

In seeking an improved curve-fitting, an additional "degree of freedom" was introduced in each of the original Wiebe functions, in the form of an additional free exponent. This way very good reproduction of the experimental curves was made possible. The modified functional form was

$$\frac{\dot{m}_{f,b}(\theta)}{m_{f,0}} = \theta^{B_p} \cdot \exp(A_p - C_p \cdot \theta^{D_p}) + \theta^{B_d} \cdot \exp(A_d - C_d \cdot \theta^{D_d}) \quad (11)$$

where A, B, C, and D are adjustable parameters, and subscripts p and d correspond to the premixed and diffusion phases, respectively.

Despite fitting success, the additional degrees of freedom make the mathematical problem of best-fitting the functional form of Eq. (11) much more difficult to solve. Therefore, considering computer time and effort, the last correlation appears to be less attractive. It provides, though, for a good starting point to reduce the problem to one having the same number of parameters with the original Wiebe functions, while still producing improved curve fits. Indeed, it was found that if a constant value of $D_p = D_d = 0.5$ was employed, satisfactory fittings, as good as those shown in Figs. 3(*a*)-3(*d*), can be achieved with considerably less computer effort. The values of the constants for the illustrated examples are given in Table 5. The following expression is therefore proposed to describe oxygen enriched diesel combustion:

$$\frac{\dot{m}_{f,b}(\theta)}{m_{f,0}} = \theta^{B_p} \cdot \exp(A_p - C_p \cdot \sqrt{\theta}) + \theta^{B_d} \cdot \exp(A_d - C_d \cdot \sqrt{\theta}) \quad (12)$$

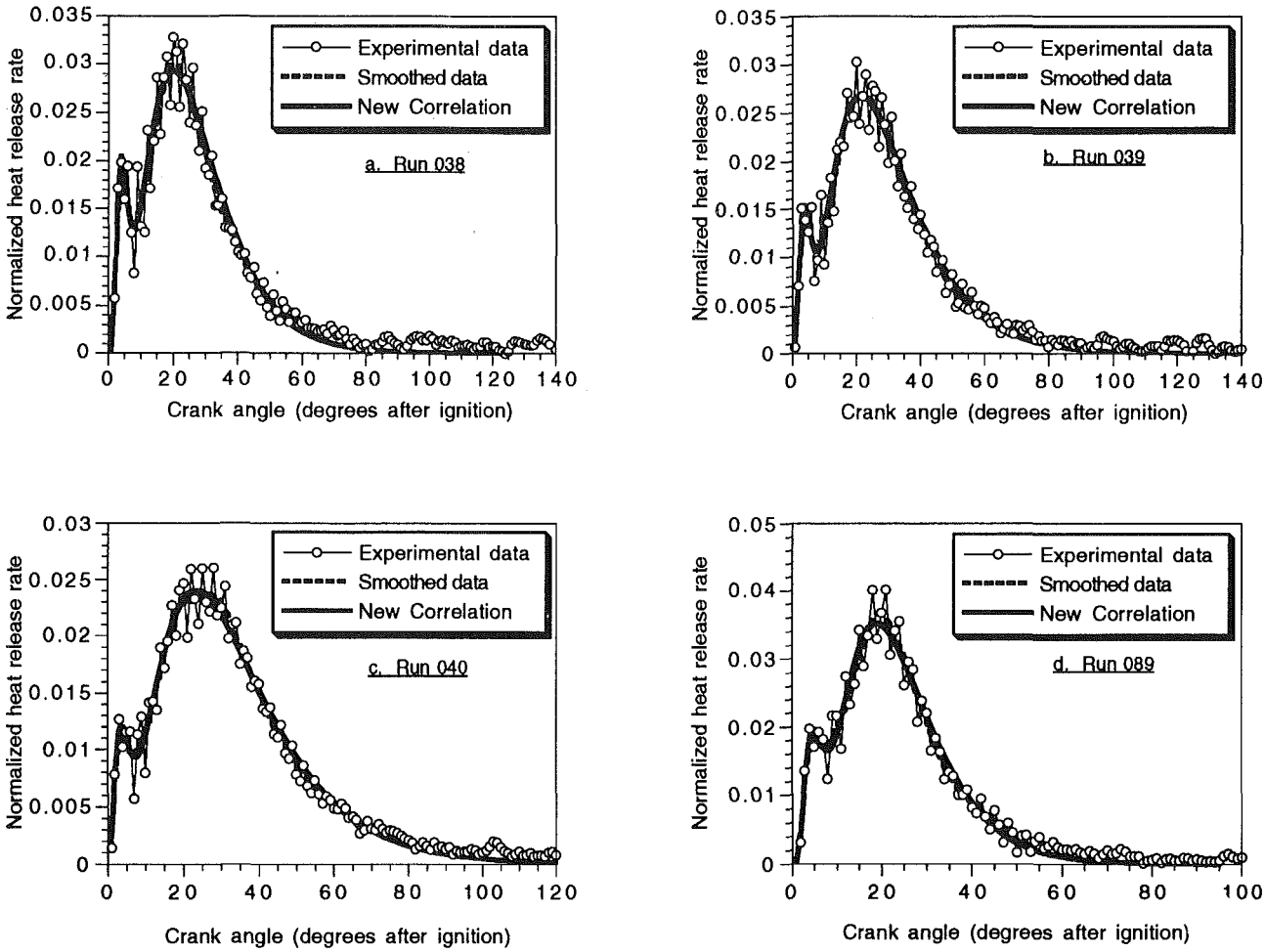


Fig. 3 Comparison of experimental heat release rates against curve-fitted profiles generated using the novel correlation of Eq. (12). Operating conditions are defined in Table 2.

Table 5 Parameters of the new correlation

	Run 038	Run 039	Run 040	Run 089
A_p	2.689	1.610	-0.9203	0.623
B_p	10.65	8.255	5.818	9.555
C_p	10.69	8.643	6.274	8.949
A_d	-10.48	-10.36	-10.37	-11.68
B_d	6.962	6.284	5.760	8.796
C_d	3.108	2.704	2.3818	4.030

Evaluation of Alternative Correlations Using an Engine Simulator

In order to verify the applicability of the proposed heat release correlation to simulation studies of oxygen-enriched combustion, it has been implemented into a comprehensive computer simulation of the turbocharged diesel engine, developed by Assanis. The parent code was validated against test results from a Cummins engine (Assanis and Heywood, 1986), and modified in order to allow for various levels of oxygen enrichment in the intake air and for operation with water-emulsified fuels (Assanis et al., 1990). This section briefly summarizes the main assumptions of the simulation. Additional details on the parent code can be found from Assanis and Heywood (1986).

The diesel four-stroke cycle is treated as a sequence of continuous processes: intake, compression, combustion (including expansion), and exhaust. Quasi-steady, adiabatic, one-dimensional flow equations are used to predict mass flows past the intake and exhaust valves. The compression process is defined to include the ignition delay period (i.e., the time interval between the start of the injection process and the ignition point). The total length of the ignition delay can either be specified or predicted using an Arrhenius expression based on the mean cylinder gas temperature and pressure during the delay period. Combustion is modeled as a uniformly distributed heat release process. The rate of heat release is specified using either one of the standard burning rate correlations (Watson's or double Wiebe), or the novel correlation proposed in this paper, as desired by the user. Heat transfer is included in all the engine processes. Convective heat transfer is modeled using available engine correlations based on turbulent flow in pipes. The characteristic velocity and length scales required to evaluate these correlations are obtained from a mean and turbulent kinetic energy model. Radiative heat transfer, based on the predicted flame temperature, is added during combustion.

Figures 4(*a*) and 4(*b*) compare simulated pressure traces against measured ones when the engine operates under 25 and 30 percent oxygen level (runs 40 and 89, respectively). Each of the three available heat release models has been used in turn to generate predictions. Adjustable model parameters were set

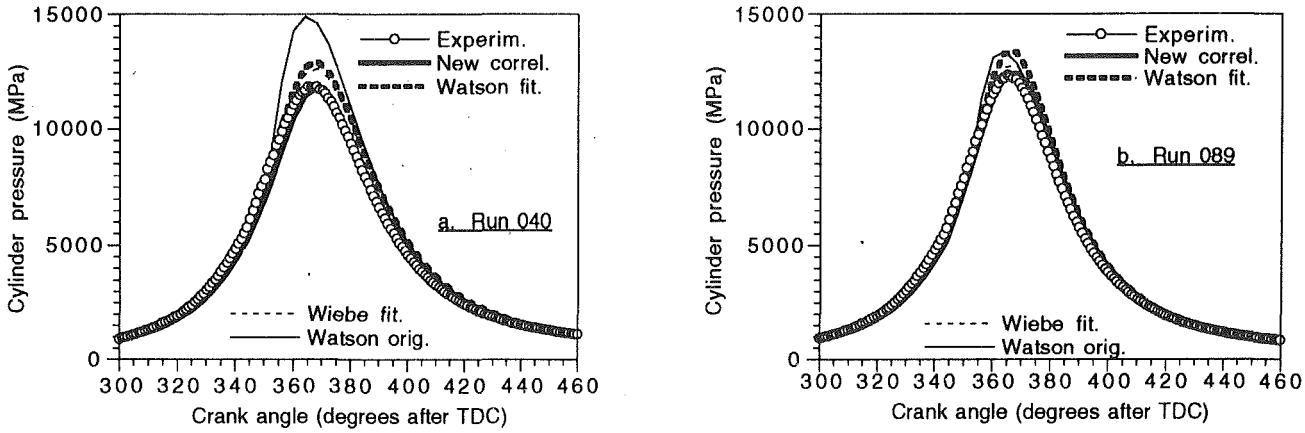


Fig. 4 Comparison of measured cylinder pressure traces against predictions from a computer simulation. Predictions are generated using alternative heat release models. Operating conditions are defined in Table 2. Parameter values are defined in Tables 3–5.

to their optimum values (see Tables 2, 3, 4) for each case. No fine tuning of the heat loss model was attempted in any case. Clearly, using the proposed heat release correlation produces pressure levels and rates of pressure rise that are in excellent agreement with measured data. Watson's correlation substantially overpredicts peak pressure levels, while the double Wiebe correlation produces results somewhere between Watson's model and the present model. Note also that results generated using Watson's model in conjunction with its recommended values for the adjustable parameters fail to match both rates of pressure rise and pressure levels. Overall, results confirm that, using the new heat release correlation, the simulation can provide sufficiently accurate predictions of trends and magnitudes to be useful in parametric studies assessing the effects of oxygen enrichment on engine performance.

Correlations of Combustion Model Parameters With Oxygen Content

Having established the merits of the novel functional form for describing oxygen-enriched heat release, it is worthwhile investigating the potential existence of some universal correlation between adjustable model parameters and engine operating conditions. In such a correlation, parameter values for a specific engine design and operating conditions could be accounted for by reference values. These values should be modified to capture the effects of changes in operating conditions on the combustion process. This way, extrapolation of engine behavior to different conditions would be made feasible through a computer model.

Lack of a sufficient number of data points for each factor that should be accounted for (load, speed, injection timing, oxygen level, water content, etc.), did not allow derivation of such a universal correlation from the present study. However, an attempt has been made to isolate the effect of oxygen content on combustion model parameters by studying heat release profiles from runs with similar fuel injection patterns. As a best approximation to the latter condition, runs at constant power were used, even though the amount of fuel injected per cycle somewhat varied due to the changes in thermal efficiency from run to run. The dependence of the values of the combustion correlation parameters on oxygen content (no water case) for two power levels corresponding to 100 and 50 percent load is illustrated in Fig. 5. The operating conditions pertaining to the engine tests referenced in the above figures are summarized in Tables 6(a) and 6(b).

Straight lines that best fit the full-load and part-load data have been superposed in Fig. 5. In our opinion, any attempt for higher order fitting of the data would be questionable for such a limited number of points. However, full-load and part-load lines seem to have similar slopes. This finding motivated an attempt to replace each set of lines with a straight line starting at some reference value (which was chosen to be that corresponding to 21 percent O_2) and having a slope equal to the average of the slopes for the 50 and 100 percent load lines. The appearing trends are summarized below:

$$B_p = B_{p,\text{ref}} - 0.460565 \times (\%\ O_2 - \%\ O_{2\text{ref}}) \quad (13b)$$

$$C_p = C_{p,\text{ref}} - 0.470640 \times (\%\ O_2 - \%\ O_{2\text{ref}}) \quad (13c)$$

$$A_d = A_{d,\text{ref}} - 0.217195 \times (\%\ O_2 - \%\ O_{2\text{ref}}) \quad (13d)$$

$$B_d = B_{d,\text{ref}} + 0.345115 \times (\%\ O_2 - \%\ O_{2\text{ref}}) \quad (13e)$$

$$C_d = C_{d,\text{ref}} + 0.176770 \times (\%\ O_2 - \%\ O_{2\text{ref}}) \quad (13f)$$

Overall, the parameters conformed to the above trends, which suggests that correlations (13a)–(13f) can be generalized in the following form:

$$\text{coeff} = \text{coeff}_{\text{ref}} + \text{const} \times (\%\ O_2 - \%\ O_{2\text{ref}}) \quad (14)$$

where $\text{coeff}_{\text{ref}}$ denotes the reference value of any of the parameters A, B, C at a corresponding reference level of oxygen enrichment and at given engine operating conditions.

Conclusions and Recommendations

In this paper, cylinder pressure data acquired on a single-cylinder engine are used to generate heat release rates for operation under various oxygen contents. These derived heat release rates are in turn used to explore the existence of a mathematical correlation that would accurately reproduce measured burning rate profiles. Two classical formulations, Watson's combustion model and the double Wiebe function approach, are tested. It has been found that both correlations failed to reproduce the diffusion burning phase closely. In particular, the correlations could not fit the gradual drop of the heat release rate observed in the combustion tail. This can result in significant errors, if these heat release models are to be implemented in an engine simulator. Moreover, in the case of the double Wiebe function, which is not explicitly normalized, poor fitting can result in errors in the computation of the total amount of fuel burned.

Introducing two additional degrees of freedom in the double Wiebe correlation can provide flexibility so as to fit the combustion tail accurately. Furthermore, by setting two of the parameters of the extended correlation to specific constants, a new heat release correlation is proposed. While the novel correlation has the same number of parameters as the double Wiebe correlation, it fits well all phases of the oxygen-enriched diesel combustion with acceptable computational effort. As expected, when the novel correlation is implemented into a

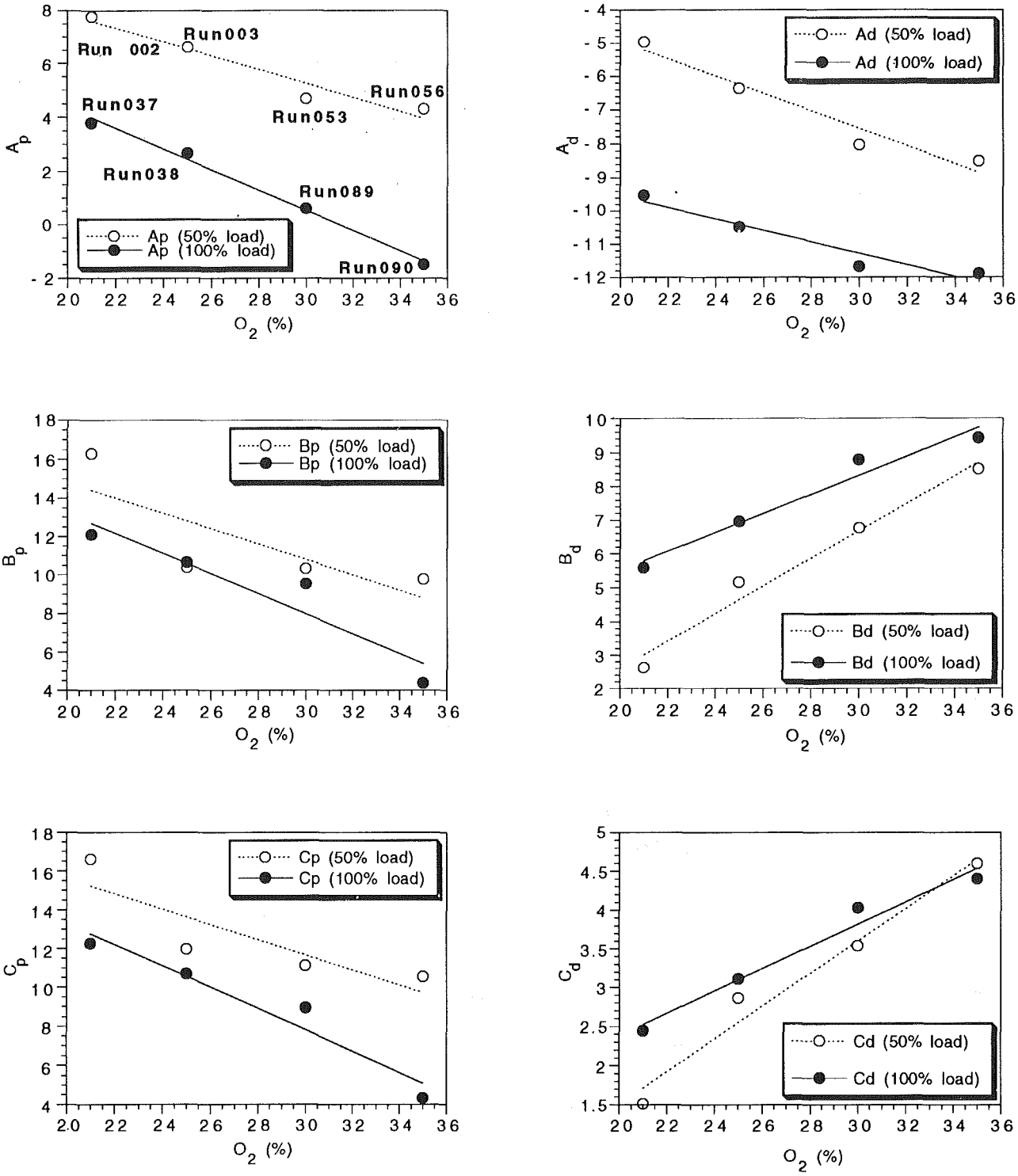


Fig. 5 Variation of heat release correlation parameters with oxygen level and engine load

Table 6(A) Engine running conditions for the correlation of parameters (100 percent load)

	Run 037	Run 038	Run 089	Run 090
O_2(%)	20.52	25.01	30.00	34.99
O_2/Fuel	6.684	8.074	10.33	11.74
Equiv. ratio	0.499	0.413	0.323	0.284
Power (BHP)	49.90	50.10	50.00	50.00

Table 6(B) Engine running conditions for the correlation of parameters (50 percent load)

	Run 002	Run 003	Run 053	Run 056
O_2(%)	20.59	25.13	29.53	35.07
O_2/Fuel	9.173	11.68	11.88	13.30
Equiv. ratio	0.363	0.285	0.281	0.251
Power (BHP)	25.02	25.00	25.10	24.90

computer simulation of the oxygen-enriched diesel engine, it can produce pressure traces in much better agreement with measurements than any of the standard correlations.

An attempt has also been made to correlate the parameters of the new heat release profile with oxygen content. Some simple correlations were identified. However, due to the limited number of data points available in this investigation, these correlations should be treated as preliminary. The methodology adopted here in correlating experimental data with simulation predictions should be extended to analyze results from tests with water-emulsified fuels. The measured heat release data for various oxygen and water content levels should be used to develop a more generalized heat release correlation to account for operation with oxygen enrichment and water-emulsified fuels over a range of conditions.

Acknowledgments

This work is supported by the U.S. Department of Energy, Assistant Secretary for Conservation and Renewable Energy, under contract W-31-109-Eng-38. Accordingly, the U.S. Government retains a nonexclusive, royalty-free license to publish or reproduce the published form of this contribution, or allow others to do so, for U.S. Government purposes.

References

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 Transactions*, Paper No. 860329.

Assanis, D. N., Sekar, R. R., Baker, D., Siambekos, C. T., Cole, R. L., and Marciniak, T. J., 1990, "Simulation Studies of Diesel Engine Performance With Oxygen Enriched Air and Water Emulsified Fuels," ASME Paper 90-ICE-17.

Bonne, M. A., 1989, "Development of Instrumentation and Analysis Techniques for Diesel Engine Combustion and Heat Transfer Studies," MS Thesis, University of Illinois, Urbana.

Cole, R. L., Sekar, R. R., Stodolsky, F., and Marciniak, T. J., 1990, "Technical and Economic Evaluation of Diesel Engine With Oxygen Enrichment and Water Injection," ASME Paper 90-ICE-1.

Miyamoto, N., Chikahisa, T., Murayama, T., and Sawyer, R., 1985, "Description and Analysis of Diesel Engine Rate of Combustion and Performance Using Wiebe's Functions," SAE Paper No. 850107.

Press, W. H., Flannery, B. P., Teukolsky, S. A., and Vetterling, W. T., 1989, *Numerical Recipes*, Cambridge University Press, New York.

Sekar, R. R., Marr, W. W., Schaus, J. E., Cole, R. L., and Marciniak, T. J., 1990a, "Diesel Engine Experiments With Oxygen Enrichment, Water Addition and Lower-Grade Fuel," presented at the Intersociety Energy Conversion Engineering Conference, Reno, NV.

Sekar, R. R., Marr, W. W., Schaus, J. E., Cole, R. L., Marciniak, T. J., and Eustis, J. N., 1990b, "Cylinder Pressure Analysis of a Diesel Engine Using Oxygen-Enriched Air and Emulsified Fuels," presented at the SAE Off-Highway and Powerplant Meeting, Milwaukee, WI.

Watson, N., Pilley, A. D., and Marzouk, M., 1980, "A Combustion Correlation for Diesel Engine Simulation," SAE Paper No. 800029.

T. W. Ryan III

T. J. Callahan

S. R. King

Southwest Research Institute,
6220 Culebra Road,
San Antonio, TX 78228-0510

Engine Knock Rating of Natural Gases—Methane Number

A procedure has been developed and documented for determining the methane number of gaseous fuels. The methane number provides an indication of the knock tendency of the fuel. An experimental test matrix was designed for quantifying the effects of ethane, propane, butane, and CO_2. A unique gas mixing and control system was developed to supply test gases to the engine and to control the equivalence ratio and engine operation. The results of the experiments agreed well with the limited data published in the literature. Predictive equations were developed for the methane number (MN) of gaseous fuels using the gas composition. The forms of these equations are suitable for incorporation in a computer program or a spreadsheet.

Introduction

Current interest in the use of natural gas in highway transportation vehicles stems from recent legislation in the United States that encourages the use of natural gas as one of the designated clean alternative fuels. The Alternative Motor Fuels Act (AMFA) specifically designates natural gas as a fuel endorsed by the Federal government as a clean alternative. The AMFA directs that the Federal government will perform several large-scale demonstrations of the use of natural gas in the Federal fleet of vehicles. In addition, recent amendments to the Clean Air Act also mention clean alternative fuels and specify a credit system designed to stimulate the use of these fuels in new equipment manufactured by the OEM's.

As the conversion work progresses and the OEM's prepare to manufacture new vehicles, it has become clear that the current specifications for quality control of natural gas have been based on its use as a burner fuel and are, therefore, not adequate to accommodate the specific new requirements imposed by its use as an engine fuel. In addition, it is obvious from the available literature that gas composition and, therefore, gas properties can vary significantly from one location to another. The variations in gas composition across the United States have been examined by AGA and the Gas Research Institute (GRI) [1]. Gas composition has been of historical concern to the gas engine manufacturers and users, in regards to the effects of composition on the knock resistance of the fuel mixture. Much of the background for this concern is embodied in two references, which present two different techniques for rating the knock resistance of natural gas. One procedure [2] was based on the use of mixtures of normal butane and methane as the reference fuels for knock rating in a CFR engine and in larger bore medium-speed engines. The other procedure [3] involves the use of hydrogen and methane as reference fuels, also in a CFR engine, as well as in two different larger bore engines.

Table 1 is a comparison of the results using the various knock rating methods for the C_1 to C_4 normal paraffins. Since these paraffins make up more than 98 percent of most natural gases, they represent the major constituents and the major contributors to the combustion characteristics of the natural gas.

The motor octane numbers (ON) were determined using the standard ASTM procedure, which uses iso-octane and *n*-heptane as the reference fuels, where 100 percent iso-octane equals 100 ON and 100 percent *n*-heptane equals 0 ON. The methane number (MN) scale is based on the molar percentages of methane and hydrogen, with neat methane equal to 100 MN. The butane number (BN) method is based on molar blends of methane and butane with neat methane equal to 0 BN. On this basis, it appears that the butane number technique provides the most sensitive scale for rating natural gases. It should be noted, however, the relationship between carbon number and butane number is much more nonlinear than either the octane or the methane number scales. This is clearly shown in Fig. 1, where the various scales are plotted versus the corresponding carbon numbers. The significance of this observation rests in the fact that the knock quality of natural gas is directly related to the composition, displaying increasing sensitivity with both concentration and carbon number. The higher the concentrations of the higher carbon number constituents, the more sensitive the fuel. In this light, the butane number scale appears to place strong emphasis on the butane concentration (going from 10 to 100 for the C_3 and C_4 paraffins, respectively). The

Table 1 Knock ratings for C_1 to C_4 normal paraffins

Gas	Motor octane No. [2]	Methane No. [3]	Butane No. [2]
Methane	122	100	0
Ethane	101	44	7.5
Propane	97	34	10.0
Butane	89	10	100.0

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1992. Paper No. 93-ICE-18. Associate Technical Editor: J. A. Caton.

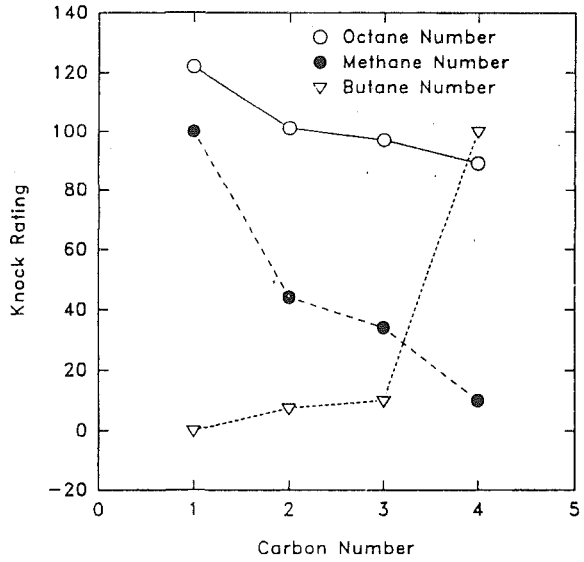


Fig. 1 Knock rating versus carbon number for normal paraffins

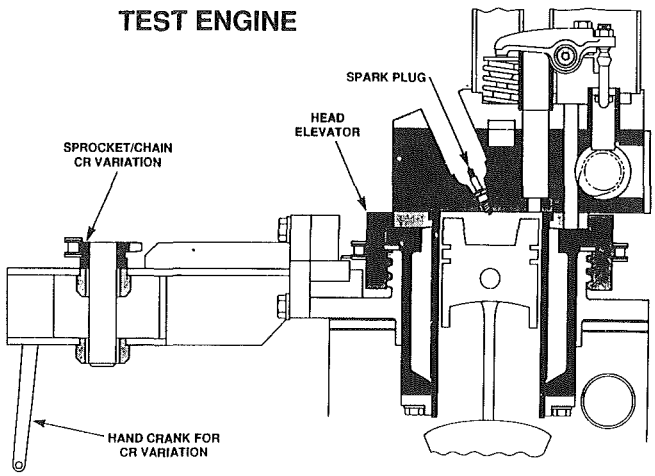


Fig. 2 Schematic of test engine

octane scale is more linear and would be a very convenient scale in terms of familiarity within the SI engine community. The problems with the octane rating method are that it was not designed to handle gaseous fuels and fuels that are less sensitive to knock than iso-octane. It should be noted that there appears to be a linear relationship between octane number and methane number. This relationship can be represented for these data by Eq. (1) with a coefficient of determination (R-square) of 0.9996. However, considering butane, based upon the octane rating of 89, it would seem that butane would make a good SI engine fuel. But from experience, butane is known to have undesirable knock qualities as indicated by a methane number of 10. The methane number procedure, because it was designed to handle gaseous fuels, appears to be the "best choice" for rating the knock sensitivity of natural gases. Close examination of the reported development work [3] for the methane number scale reveals, however, that the methane number procedure and the knock meter were not adequately described.

$$\text{MON} = 84.9 + 0.37\ (\text{MN}) \tag{1}$$

As indicated previously, most of the work reported to date has generally involved the use of only one gas composition, or has not even considered variations in composition. Dubel et al. [4] examined the effects of composition on methane number and engine performance optimization. They found that methane number provides a good indicator of the fuel performance in an engine in terms of the ability to achieve acceptable efficiency within the knock limited spark advance. Klimstera and Wolfing [5] examined various relationships between composition and mixture properties. Very little new information was reported in this work, but it did represent a good summary of the effect of composition on stoichiometric air–fuel ratio, heating value, methane number, knock sensitivity, and Wobbe number considerations. Results reported by the authors [6] indicate that the effects of composition on the performance and emissions are dominated by the effects of composition on the equivalence ratio at a fixed air–fuel ratio and the knock rating.

It appears that methane number is a very important parameter that will ultimately be incorporated in future fuel specifications for natural gas. It was felt, therefore, that a systematic study was needed to define, document, and streamline the methane number procedure. The goals of this work were to refine and document the methane number procedure and to develop a mathematical model for predicting the knock tendency of natural gas mixtures.

Experimental Apparatus and Procedures

Engine. The test engine, shown schematically in Fig. 2, was designed and constructed at SwRI specifically for use in the evaluation of fuel effects on engine combustion. Special features include variable compression ratio, variable valve timing, variable spark timing, and several access holes for installation of spark plugs and pressure transducers. The test cell was equipped with a high-speed data acquisition system and an exhaust gas analysis system.

The configuration of the combustion chamber was representative of current lean-burn heavy-duty natural gas engines, with a combustion bowl in the piston. The clearance volume was selected to provide a continuously variable compression ratio from 8.0:1 to 16.3:1. A spark plug was installed in the head, with the gap located slightly below the fire deck and at the center of the combustion chamber. A 10-mm NGK CR8E spark plug was used with a 0.045-in. spark gap. The engine was equipped with a Mallory High Fire Four ignition system and a spark control, which was manually adjusted to obtain different spark timings.

Air–Fuel Control System. One of the unique features of the experimental equipment was the use of a gas flow metering system to prepare fuel gas mixtures prior to the intake to the engine. The gas flow system consisted of seven mass flow controllers. Each controller regulated the mass flow of a pure compressed gas to obtain a desired concentration of that gas in a fuel blend.

The gas mixture composition was monitored and controlled using a PC-based data acquisition system to adjust the flow rates of the seven components to match the set point. The set point was defined by the concentrations of the fuel constituents and the equivalence ratio in the intake system to the engine. Calculations were performed to establish the required flow ranges of the seven control valves (one for each gas/vapor) based upon the desired concentrations and the estimated air flow of the engine.

The air–fuel control system consisted of the air supply and control side, and the fuel control and mixing side. The air system included a pressurized air supply, a pressure regulator, an air flow control valve, a surge tank, and the fuel/air mixing section. The fuel system consisted of the bottled gases, two stage regulators, mass flow controllers, a mixing tank, and the mixture supply line to the fuel/air mixing section.

The design strategy was based on the use of the dynamometer controller for engine speed control. The engine load was defined by the mass flow rate of air to the engine (effectively defined by the intake manifold pressure). The air flow was controlled using the air mass flow control valve based on de-

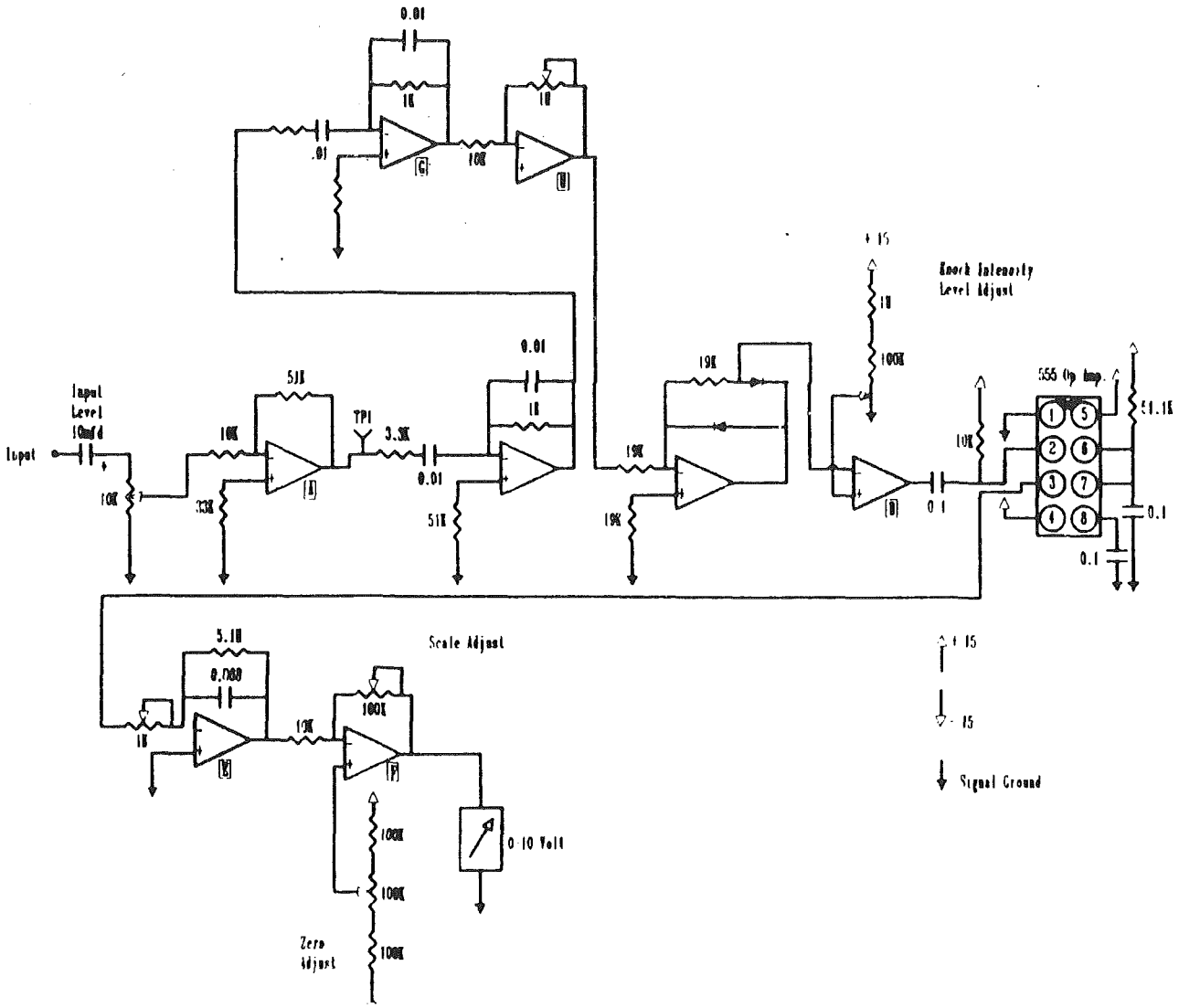


Fig. 3 Schematic of knock meter

fined set points of manifold air pressure (MAP). The set points for the individual gas flow control valves were computed and transmitted to the valves in order to maintain predefined mixture ratios. In this manner, the test gas composition and the overall air–fuel ratio were maintained independent of variation in the intake air mass flow rate.

The emission bench consisted of analyzers for oxygen, carbon dioxide, carbon monoxide, unburned hydrocarbons, and oxides of nitrogen. The output of each emission instrument was connected to the computer and the emissions were monitored on-line and used to compute the equivalence ratio. This added feature enabled the computer that was controlling the inlet gas composition and flow rate, to monitor the exhaust gas composition and to adjust the inlet conditions to achieve the desired stoichiometry based upon exhaust gas composition.

Knock Meter. Using the guidelines presented by Leiker et al. [3], a knock meter was designed and built. A schematic of the meter is presented in Fig. 3. The knock meter was designed to provide an indication of knock by monitoring the cylinder pressure signal. An AVL-12QP pressure transducer was flush mounted in the head. Briefly, the meter monitored the cylinder pressure signal, filtered out the pressure rise due to compression, and examined the remaining signal for pressure oscillations due to knock. A knocking signal was detected when the filtered pressure signal exceeded a threshold value. The threshold was set using a knock intensity level adjustment and a reference fuel.

Table 2 Engine conditions for methane number rating

Engine Parameter	Set Point
Equivalence ratio	0.6 and 1.0
Speed	900 rpm
Intake manifold pressure	0.099 MPa (14.5 psia), Φ = 1.0 0.148 MPa (21.4 psia), Φ = 0.7
Spark timing	15 deg BTDC, Φ = 1.0 25 deg BTDC, Φ = 0.6
Intake temperature	21°C (70°F)

Test Procedure. The test conditions for the methane number rating procedures are listed in Table 2. The knock meter was sensitive to engine speed, therefore all data were collected at one speed, 900 rpm. The calibration procedure consisted of three steps:

1 With no signal input, the zero potentiometer was adjusted to provide a meter reading of zero.
2 The full-scale level was adjusted by running the engine at the desired speed and connecting the cylinder pressure signal to the input of the meter. Then with the knock intensity level adjustment at its most sensitive setting, the scale adjustment potentiometer was adjusted to obtain a full-scale reading.

Table 3 Summary data for natural gas composition based on tabulation of Griffiths et al. [7]

	Methane	Ethane	Propane	Butane	Pentane	Hexane	Nitrogen	Carbon dioxide
Mean	93	3	1	0.5	0.1	0.5	1.5	0.5
Standard dev.	5.5	2.6	1.4	1.0	0.3	0.1	2.9	0.5
Min.	73	0	0	0	0	0	0	0
Max.	99	13	8	7	3	1	17	2
Range for test matrix design								
Min.	85	0	0	0	0	—	—	0
Max.	100	10	4	3	2	—	—	3

3 The knock intensity level was then set using a reference fuel and setting the meter to give a predefined meter reading at a given compression ratio.

A methane number (MN) calibration curve was developed for the engine by testing blends of methane (MN = 100) and hydrogen (MN = 0). The knock meter was used to determine the knock level for each methane–hydrogen blend while the engine compression ratio was varied. The knock meter was adjusted to give a reading of 50-percent full scale on methane at a compression ratio of 14.4. For the methane–hydrogen blends, the compression ratio was varied to achieve a 50-percent knock reading. The methane number of the reference fuel blends was plotted against the 50-percent knock compression ratio to obtain a calibration curve.

Test Fuel Matrix. Topical Report for GRI Contract 5011-345-0100, entitled "Effect of Fuel Gas Composition on Appliance Performance" [7] was used as the basis for selection of the test gas mixtures. The report contained tabulations of the analyses of 120 different natural gas samples collected throughout the United States. These composition data were entered into a computer data base for statistical analysis in order to establish the most representative ranges in the concentrations of the various test gases. A summary of this statistical analysis is presented in Table 3. Also shown in the table are the ranges selected for use in the design of the test matrix. It appears that the selected ranges are appropriate in terms of being representative of natural gas, as well as being sufficiently large to demonstrate the effects of these various constituents on the methane number. An Experimental Design Optimization Program, entitled "ECHIP" [8], was used in the design of the test matrix once the appropriate ranges of concentrations were established. ECHIP was designed specifically to handle mixture experimental designs and analyses that are not easily handled using more conventional statistical approaches. In this particular design, six independent variables were used: methane, ethane, propane, butane, pentane, and carbon dioxide. The test matrix [9] consisted of 31 individual test points with five repeats. The goal of the test matrix design was to optimize the number of test points for development of regressions of the various dependent variables in terms of the gas composition.

Experimental Results

The 31-gas test matrix has been used to develop expressions for the methane number in the diesel conversion configuration using both a stoichiometric and a lean mixture ratio. The results of these experiments are discussed separately in the following sections.

Stoichiometric Methane Number. The methane number determinations were performed by first calibrating the engine using mixtures of methane and hydrogen, following as closely as possible the procedures outlined by Leiker et al. [3]. As a review, the first problem that was encountered in these experiments was the fact that knock was not adequately defined and the knock meter was not described in [3]. Some development work was, therefore, required to develop a meter that provided an acceptable range of readings from no knock, through trace audible, to severe knock. The scale was adjusted to provide a reading of 50, on a scale of 0 to 100, at trace audible knock. This definition was used throughout the methane number experiments.

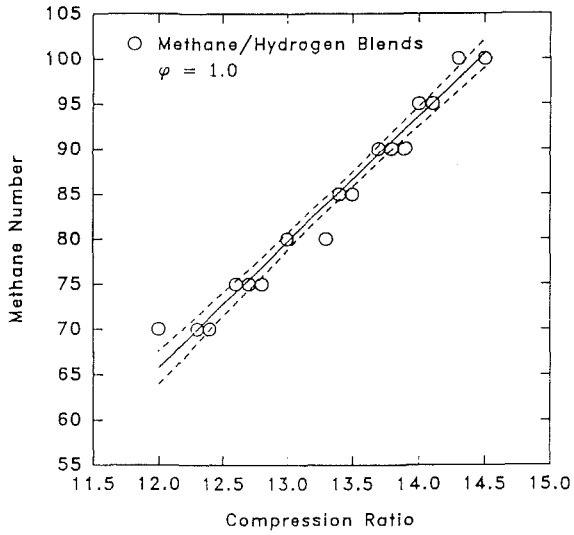


Fig. 4 Methane number calibration at Φ = 1.0

The conditions outlined by Leiker et al. [3], naturally aspirated, 900 rpm, spark timing of 15 deg BTDC, and an equivalence ratio, Φ, of unity, were adopted for this work. Φ is defined as the stoichiometric air–fuel ratio divided by the actual air–fuel ratio. The engine was first calibrated using mixtures of methane, with a defined methane number of 100, and hydrogen, with a defined methane number of 0. Methane number is defined in terms of these two gases as the volume percentage of methane in the blend with hydrogen. The calibration was repeated several times to examine the repeatability of the procedure. The calibration curve for the stoichiometric condition is presented in Fig. 4, where the solid line represents the least-squares fit to the data and the dashed lines show a 95-percent confidence interval about the regression line.

Each gas in the 31-gas test matrix was then evaluated using the same conditions and procedures as used for the calibration. The resulting data consisted of the compression ratio required for each gas to produce the same knock level (50 on the knock meter scale) as used to develop the calibration data. The resulting compression ratios were then used in the calibration equation to compute the corresponding methane numbers.

The observed, or measured, methane numbers were used in a regression analysis to develop a full quadratic response surface relating methane number to fuel composition. A quadratic response surface was chosen to give the best predictive equation based upon the experimental test matrix design. The form of the equation and the coefficients are presented in Table 4. The

Table 4 Prediction equation for methane number, Φ = 1.0

i	B_i	X_i
0	83.987	Intercept
1	10.496	Methane = (percent methane/100)
2	−171.740	Ethane = (percent ethane/100)
3	−261.121	Propane = (percent propane/100)
4	250.521	Butane = (percent butane/100)
5	158.922	Pentane = (percent pentane/100)
6	−77.889	CO_2 = (percent CO_2/100)
7	3.940	$(\text{Methane})^2$
8	754.632	$(\text{Ethane})^2$
9	−0.178	$(\text{Propane})^2$
10	−12012.767	$(\text{Butane})^2$
11	−13175.333	$(\text{Pentane})^2$
12	9503.235	$(CO_2)^2$
13	−55.778	Methane * ethane
14	1.774	Methane * propane
15	121.472	Methane * butane
16	−85.543	Methane * pentane
17	−117.579	Methane * CO_2
18	110.682	Ethane * propane
19	1110.558	Ethane * butane
20	187.904	Ethane * pentane
21	−348.815	Ethane * CO_2
22	−649.417	Propane * butane
23	3206.545	Propane * pentane
24	−2222.563	Propane * CO_2
25	3704.816	Butane * pentane
26	−2367.126	Butane * CO_2
27	6640.567	Pentane * CO_2

Equation valid in these ranges: methane—85–100 percent, ethane—0–10 percent, propane—0–9 percent, butane—0–3 percent, pentane—0–2 percent

equation: $\text{MN} = \sum_{i=0}^{27} B_i X_i$

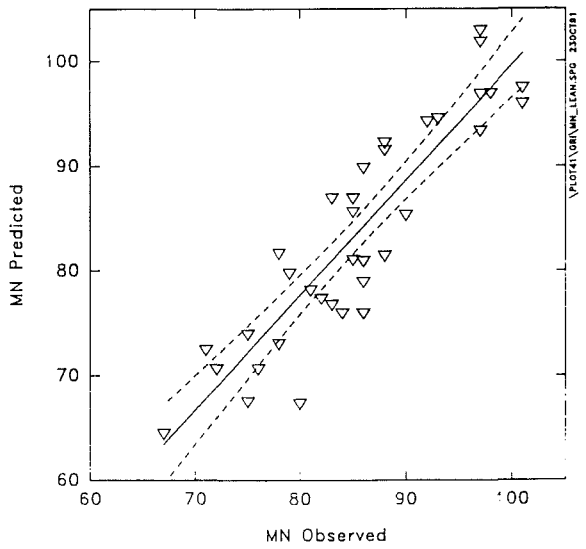


Fig. 5 Predicted methane numbers versus the measured values at Φ = 1.0

regression equation was used to calculate the methane number of the 31 gases. The predicted values are plotted versus the measured values in Fig. 5, where the solid line represents the least-squares fit to the data and the dashed lines show a 95 percent confidence interval about the regression line. The regression analysis indicated an R-square of 0.94 and a standard error of ±3.0 MN, in good agreement with the observed value. The variability in the standard error includes the variability of the engine and the ability to control the compositions of the calibration and the test gas mixtures accurately.

As was indicated previously, the range of concentrations used in the 31-gas matrix was based on the range of compositions of natural gas observed in the U.S. Use of the methane number equation presented in Table 4 represents interpolation of the data within the range of the test gases and extrapolation outside this range. The equation should, therefore, be used with caution if used for gases outside the ranges presented previously. A graphic presentation of the response surface is presented in Fig. 6. This figure was generated by assuming that the graph was composed only of methane, ethane, and propane, normally the three major constituents. This figure can be used to estimate methane number graphically if one assumes the major constituents of the gas correspond to those represented on the plot.

Table 5 Prediction equation for methane number of high-propane concentrations, Φ = 1.0

i	B_i	X_i
0	43.28	Intercept
1	28.53	Methane = (percent methane/100)
2	−29.91	Ethane = (percent ethane/100)
3	−121.32	Propane = (percent propane/100)
4	28.15	Methane^2
5	147.7	Ethane^2
6	272.9	Propane^2
7	−79.8	Methane * ethane
8	−91.5	Methane * propane
9	303.7	Ethane * propane

Equation valid in these ranges: methane—60–100 percent, ethane—0–2 percent, propane—0–40 percent,

$$\text{MN} = \sum_{i=0}^{9} B_i X_i$$

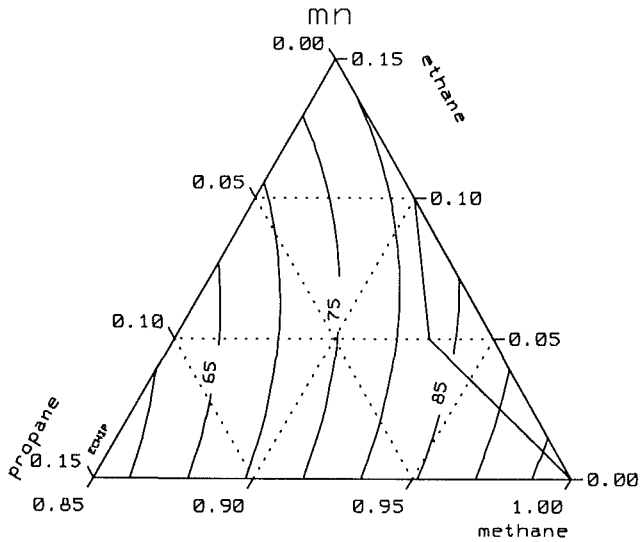


Fig. 6 Methane numbers for blends of methane, ethane, and propane, Φ = 1.0

Stoichiometric Propane Mixtures. In the U.S., gases that fall very far outside the mean in composition generally contain high concentrations of propane, resulting from the practice of some companies of using propane-air peak shaving. A special test matrix and regression equation was developed for use with high propane content gas mixtures. The fuel gas matrix was limited to three constituents: methane, ethane, and propane. The concentrations ranged from 60 to 100 percent methane, 0 to 20 percent ethane, and 0 to 40 percent propane. An eight-gas test matrix was designed for these experiments. The test conditions were again naturally aspirated, 900 rpm, 15 deg BTDC ST, and Φ = 1.0. The experimental data were statistically analyzed, and a regression equation was developed. The regression equation relating the methane number to the composition, presented in Table 5, has an R-Square of 0.99 and a standard error of ±2.3 MN. Figure 7 is a graphic represen-

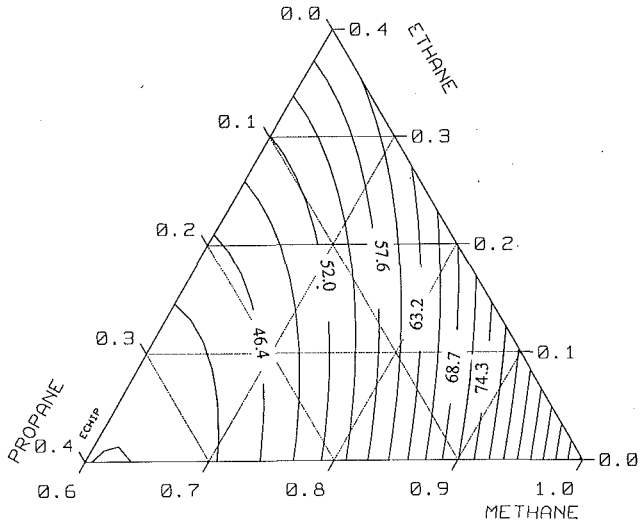


Fig. 7 Methane numbers for mixtures of methane and high levels of ethane and propane, $\Phi = 1.0$

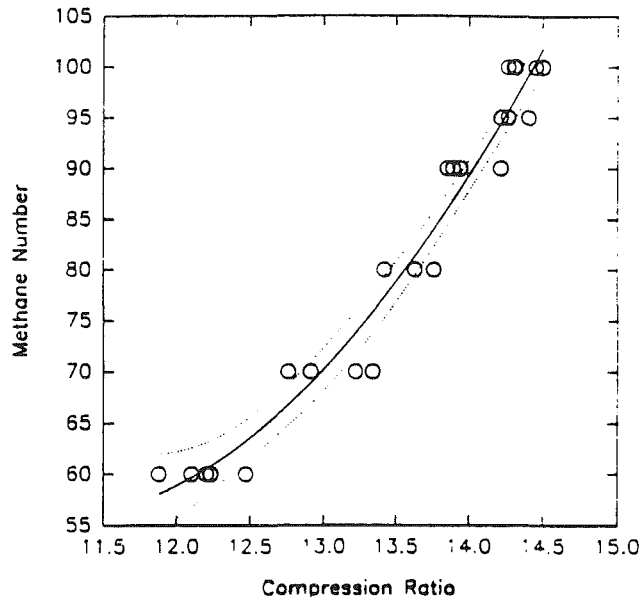


Fig. 8 Methane number calibration at $\Phi = 0.6$

tation of the results, including an indication of the range covered by the experimental data.

Lean Methane Numbers. A methane number technique was also developed for applications where fuel lean conversions are used. Lean operation and turbocharging are generally involved in these applications. The initial work involved development of the test condition that is representative of the application while at the same time producing the defining knock level over the same range of gas compositions used in the stoichiometric technique. Several different spark timings, boost pressure, and air-fuel ratios were examined in the initial experiments. The equivalence ratio was varied from 0.6 to 0.7, spark timing from 15 to 42 deg BTDC, and manifold pressure (MAP) from 0.099 to 0.160 MPa.

The initial experiments verified the findings of Dubel et al. [4] regarding the reduction of methane number requirement with leaner mixtures. It was also found that under some combinations of conditions the engine would not produce the required level of knock on methane, i.e., lean, retarded timing, and low boost pressures. At the other extreme, rich, advanced timing, and high boost, the engine would knock with all mixtures containing hydrogen. It was found that an acceptable calibration could be developed if the engine was operated at $\Phi = 0.6$, 43 deg BTDC spark timing, and a MAP of 0.148 MPa. The methane number versus compression ratio calibration is presented in Fig. 8.

The 31 test gases in the test matrix were rated at the lean condition. The resulting data were also used to develop an expression for the methane number as a function of the composition for the lean condition. The R-Square for the regression equation is 0.92, and the standard error is ± 4.9 MN. It is felt that the statistics for these experiments were not as good as those previously described because of larger variations in the engine operation resulting from an increase in the cycle-by-cycle variations in combustion due to operation in the proximity of the lean misfire limit. It is also felt, however, that this condition represents a realistic situation for lean engines. The regression equation is presented in Table 6.

Table 6 Prediction equation for methane number, $\Phi = 0.6$

i	B_i	X_i
0	67.545	Intercept
1	16.762	Methane = (percent/100)
2	−112.693	Ethane = (percent/100)
3	−353.541	Propane = (percent/100)
4	9.677	Butane = (percent/100)
5	−478.617	Pentane = (percent/100)
6	139.322	CO_2 = (percent/100)
7	19.006	$(\text{Methane})^2$
8	−110.852	$(\text{Ethane})^2$
9	6900.598	$(\text{Propane})^2$
10	−8012.389	$(\text{Butane})^2$
11	26496.674	$(\text{Pentane})^2$
12	3391.418	$(CO_2)^2$
13	−69.005	Methane * ethane
14	−354.447	Methane * propane
15	246.166	Methane * butane
16	−632.976	Methane * pentane
17	−250.103	Methane * CO_2
18	2096.323	Ethane * propane
19	1334.573	Ethane * butane
20	1652.291	Ethane * pentane
21	−446.539	Ethane * CO_2
22	−6758.749	Propane * butane
23	8837.239	Propane * pentane
24	5536.706	Propane * CO_2
25	−5961.901	Pentane * butane
26	278.997	Butane * CO_2
27	8718.1001	Pentane * CO_2

Equation valid in these ranges: methane—85–100 percent, ethane—0–10 percent, propane—0–9 percent, butane—0–3 percent, pentane—0–2 percent

equation: $\text{MN} = \sum_{i=0}^{27} B_i X_i$

Discussion

Three separate equations have been developed relating methane number of a gas blend to the gas composition. This in itself is a rather unique accomplishment, since similar equations do not exist relating fuel composition to any of the knock rating parameters (Octane Number or Butane Number). However, there are several questions that arise regarding the utility of the developed equations. Are the equations significantly different in a practical sense? How should these equations be applied, and what are the restrictions?

In the strictest sense, the equations should only be used for gases that have components that are in the equation. For example, it might not make sense to use the equation for high levels of ethane and propane to predict methane numbers for gases containing butane, CO_2, or N_2. On the other hand, how often does one encounter a natural gas with only methane, ethane, and propane? Do butane, CO_2, and N_2 have a major impact on the MN in the concentrations normally found in natural gas?

Table 7 Gas blend composition and methane numbers

Gas	Blend 1	Blend 2	Blend 3	Blend 4	Blend 5	Blend 6
Methane	99.7 percent	95 percent	90.1 percent	88.3 percent	85 percent	83.3 percent
Ethane	0.28 percent	3.0 percent	6.0 percent	7.8 percent	6.5 percent	3.0 percent
Propane	—	0.5 percent	0.7 percent	1.2 percent	3.0 percent	11.9 percent
Butane	—	0.3 percent	0.5 percent	0.3 percent	0.6 percent	0.6 percent
Pentane	—	0.1 percent	0.2 percent	—	0.2 percent	0.2 percent
CO_2	—	0.2 percent	0.7 percent	1.8 percent	1.0 percent	0.3 percent
Nitrogen	0.02 percent	0.8 percent	1.7 percent	0.6 percent	3.5 percent	0.7 percent
Hexane	—	0.1 percent	0.1 percent	—	0.2 percent	—
MN graphic[1]	100	85	77.5	—	71.0	—
MN Measured	94	85	79	78	72	57
Predicted MN[2]	98	91	85	81	78	(60)[6]
Predicted MN[3]	98	90	87	82	(88)[6]	(58)[6]
Predicted MN[4]	99	93	87	83	80	66
Predicted MN[5]	99	92	85	83	78	65

[1]Leiker et al. [4].
[2]MN (Table 4) general equation. Ignore N_2.
[3]MN (Table 4) general equation. (N_2 = CO_2).
[4]MN (Table 5) high C_2H_6, C_3H_8 equation. Use only CH_4, C_2H_6, C_3H_8.
[5]MN (Table 5) high C_2H_6, C_3H_8 equation. (CH_4 + CO_2), C_2H_6, (C_3H_8 + C_4H_{10} + C_5H_{12}).
[6]These points are extrapolated.

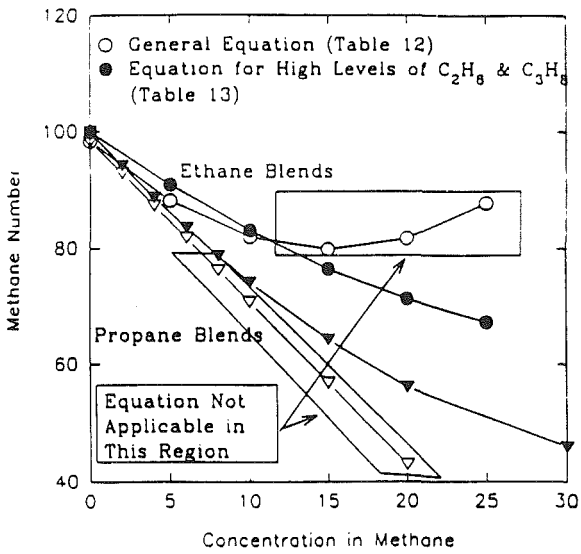


Fig. 9 Methane number predictions for high propane and ethane concentration using both the general and HEP equations

In order to compare the equations, first consider gases that have only those components that are in the equation. A comparison of the general equation for stoichiometric mixtures and the equation for high levels of ethane and propane (HEP) is shown in Fig. 9. This figure depicts the effects of increasing concentrations of ethane and propane in binary mixtures with methane. As expected, adding ethane or propane to methane resulted in a decrease in MN. Adding propane resulted in a larger decrease in MN than adding a corresponding amount of ethane.

Both the general equation and the HEP equation gave similar results for low levels of ethane (< 10 percent) and propane (< 6 percent). The HEP equation by design has a wider range of applicability and can be used for gases with up to 20 percent ethane and 40 percent propane. The general equation deviates significantly from the expected results at these higher levels. This deviation is a result of extrapolation beyond the range of experimental data, which should be avoided when using these equations.

A Practical Example. Six bottled test gas mixtures were also rated at the stoichiometric test condition using the methane number technique. These gases were chosen by GRI to be representative of the range of gases found throughout the United States. The compositions of the gases are shown in Table 7. Also shown in the table are the measured methane numbers and the MN calculated from the procedure outlined by Leiker et al. [3]. There appears to be good agreement between the measured MN and those calculated using Leiker et al. [3]. However, the procedure outlined by Leiker et al. [3] is time consuming to implement and not practical for evaluating large numbers of gases. The equations developed in this study lend themselves very well to implementation in a standard computer program or spreadsheet.

In order to use the equations presented in this paper, some decisions must be made regarding gas constituents that are not in the equations. Should they be ignored? Or perhaps grouped with other components? In the case of the general equation, the only major component that is not accounted for is nitrogen. Any hydrocarbons heavier than pentane that are present are typically very low in concentration and can be grouped with pentane. Therefore, the only question is what to do with the nitrogen present in the gas. Considering the six test blends, in general these gases have very low concentrations of nitrogen with the possible exception of Blend 5, which had 3.5 percent N_2.

There are several approaches to this problem. One approach would be to ignore N_2 and to re-distribute the concentration of the other components to sum to 100 percent. In this approach methane gets the majority of the concentration attributed to N_2 since it is the largest component. The other components receive only a nominal increase on an absolute basis. A second approach would be to group the N_2 with the CO_2, reasoning that both components act as diluents. The results, using these two approaches, are shown in Table 7. In general, since N_2 was low in concentration, both approaches yield similar results. The exception to this was Blend 5. For this blend, adding the N_2 to the CO_2 resulted in a CO_2 level that was outside of the experimental data and resulted in an extrapolated value of 89 compared to 78 by the first approach and a measured MN of 72. The measured results fall within a 95 percent confidence interval about the predicted value as shown in Fig. 10. In the figure, the MNs are plotted versus the methane content for each of the six blends. As indicated, the equations appear to overpredict the MN of the gas blends, but still provide the correct trend.

It should be noted that Blend 6 has a much higher propane content than was considered in the 31-gas test matrix. Fortuitously, the predicted value came very close to the measured value. The use of the general equation to predict this MN, however, should be considered a very serious mistake due to the extent of extrapolation required to arrive at this prediction. A more acceptable approach would be to use the equation for

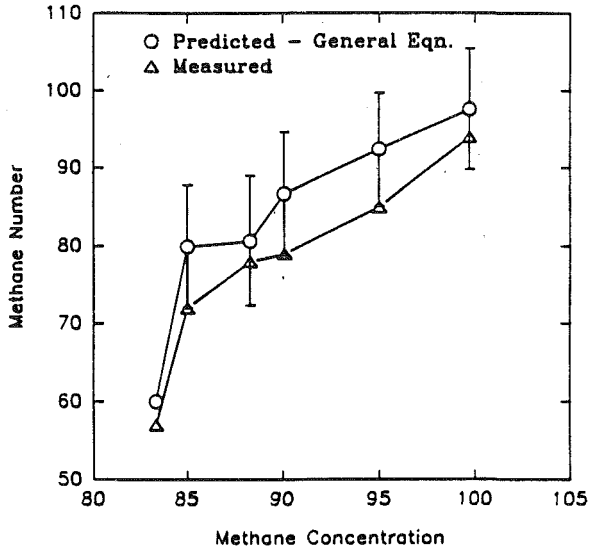


Fig. 10 Predicted versus measured methane numbers

gases with high ethane and propane levels (Table 5). The problem with this equation, obviously, is that it only includes methane, ethane, and propane, ignoring butane, pentane, CO_2, and N_2. Once again a decision must be made regarding the constituents not included in the equation. One approach would be to ignore all the other components and redistribute methane, ethane, and propane to sum to 100 percent. A second approach would be to group CO_2 and N_2 with methane and group the heavier hydrocarbons with propane. Predicted MNs for the six blends are included in Table 7. As indicated in the table, each of the above approaches gave similar results, provided that the range of the gas constituents did not exceed the bounds of experimental data.

Summary and Conclusions

A scaled single-cylinder version of a modern lean-burn natural gas conversion of a diesel engine was developed for use in this study. A computer-controlled gas mixing system was also designed and developed in the project to supply whatever gas mixture was desired for study in the engine. The system not only provided the correct proportions of six component gases, but it also controlled the air–fuel ratio to the engine. Since the equivalence ratio has a significant effect on the engine performance, the control system not only checked the mass flows of the various gases and air but also monitored the emissions and computed the equivalence ratio using the emissions data. A statistically designed test matrix of 31 different gas mixtures was used to develop relationships for the methane number at conditions that represent typical conditions in both lean burn and stoichiometric engines. The engine knock studies produced several results that are important for the future advancement of gas engines. They include:

1 A knock meter was developed that can be used to quantify knock intensity in gas fueled engines.
2 A procedure, developed previously, was redefined and developed for use in rating the knock sensitivity of natural gases in terms of methane number.
3 Mathematical expressions were developed for predicting the methane number of natural gases based on their composition. The expressions were developed for three different representative test conditions.
4 The experimental results demonstrated that the knock sensitivity of a fuel is primarily dependent on the composition, but is also affected by the engine operating condition, as indicated by the fact that the ratings were trendwise similar but systematically different for the different operating conditions.

Acknowledgments

The authors acknowledge the financial support of the Gas Research Institute. The authors would also like to thank Mr. Milan Maymar for his efforts in performing the experiments.

References

1 Liss, W. E., and Thrasher, W. H., "Natural Gas as a Stationary Engine and Vehicular Fuel," SAE Paper No. 912364, Oct. 1991.

2 Schaub, F. S., and Hubbard, R. L., "A Procedure for Calculating Fuel Gas Blend Knock Rating for Large-Bore Gas Engines and Predicting Engine Operation," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 107, 1985, pp. 922–930.

3 Leiker, M., Cartellieri, W., Christoph, K., Pfeifer, U., and Rankl, M., "Evaluation of Antiknocking Property of Gaseous Fuels by Means of Methane Number and Its Practical Application to Gas Engines," ASME Paper No. 72-DGP-4, April 1972.

4 Dubel, M., Schmillen, K., and Wackertapp, H., "Influence of Gas Composition on the Knocking Behavior of Spark-Ignited Gas Engines," presented at the 1983 International Gas Research Conference, 1983.

5 Klimstra, J., and Wolting, H. R., "Quality Aspects of Natural Gas as an Engine Fuel," SAE Paper No. 861578, 1986.

6 Ryan, T. W., III, and Callahan, T. J., "The Effects of Natural Gas Composition on Engine Combustion, Performance, and Emission," FISITA 92, Paper No. 925012, London, United Kingdom, June 1992.

7 Griffiths, J. C., Connelly, S. M., and DeRemer, R. B., "Effect of Fuel Gas Composition on Appliance Performance," GRI Contract 5011-345-0100, GRI Report No. 82/00337, Dec. 1982.

8 "ECHIP-Experimental Design in a Chip," ECHIP Inc., Copyright 1990, ECHIP Inc., 7460 Lancaster Pike, Suite 6, Hockesson, DE, 19707.

9 Ryan, T. W., III, and Callahan, T. J., "Effects of Gas Composition on Engine Performance and Emissions," GRI Topical Report GETA 91-13, Dec. 1991.

The Effect of Ambient Gas Temperature and Density on the Development and Wall Impingement of High-Injection-Pressure Diesel Fuel Sprays

Gong Yunyi

Liang Xuanming

National Engine Combustion Laboratory,
Tianjin University,
Tianjin, People's Republic of China

An investigation of the effect of ambient gas temperature and density on diesel fuel spray penetration, spray angle, and wall impingement at an injection pressure of 75–134 MPa was conducted in a constant-volume bomb with a reconstructed Cummins PT fuel system by using a high-speed photographic technique. The results show that penetration does not increase monotonically with injection pressure, and ambient temperature has more effect on a high-pressure spray than on those with conventional pressures. With the high temperature, the penetration of a high injection pressure spray is reduced a bit, while the spray angle increases obviously. When the high-pressure spray impinges on a wall at ordinary temperature, the rebounding droplets can hardly be seen, but at higher wall temperature, a cloud of dense spray will be observed near the wall, and sometimes a vapor layer will be formed between the spray and the wall. Based on experimental results, an empirical formula considering the effects of both the ambient temperature and injection pressure is presented.

Introduction

It is well known that a big increase in fuel injection pressure is one of the most effective approaches for improving the combustion process and performance of diesel engines [1, 2]. The conventional injection pressure of common diesel engine is about 40–50 MPa, and the raised injection pressure can reach above 100 MPa. At conventional injection pressure, a significant amount of research has been conducted on the law that spray penetration and spray angle, which are closely related to the combustion process, vary with ambient pressure and temperature [3]. While at high injection pressure, studies are still scarce, some results have been reported [4], though few people have researched the influence of ambient temperature. In this paper, we study the influence of ambient pressure and temperature upon spray penetration and spray angle at an injection pressure of 75–134 MPa by using high-speed Schlieren photography, and discovered that high-pressure sprays possess many characteristics, which are quite different from those of conventional pressure sprays.

The phenomenon of spray impingement on a wall has been paid wide attention in recent years. This paper deals with the process of high-pressure spray wall impingement at different wall temperature and exposes some characteristics when the spray touches the wall. Those characteristics may have an important influence upon the mixture formation.

Experiment

Experiments were conducted in a constant-volume bomb using a reconstructed Cummins PT fuel system for generating diesel fuel pressure in excess of 130 MPa. The bomb was pressurized with nitrogen gas. An electric heater of 2.0 kW enabled the gas temperature to be increased to 800 K. The nozzle hole diameter is 0.18 mm. The fuel used in experiments is light diesel fuel with density of 0.84. The section scheme of the bomb is shown in Fig. 1. A detailed description of the bomb and the measurement of fuel injection rate has been given in [5].

Figure 2 illustrates the optical arrangement for the high-speed photography schematically. A He–Ne laser with an output power of 25 mW was used as the light source. An aperture was placed at the focus of the lens. The aperture size should be carefully selected. In this study, the diameter of the aperture is 3 mm.

The contour of developing spray was determined by high-speed photography with an adjusted Schlieren effect. Spray tip penetration and spray angles were measured from photographs. In this paper, we assume that the Lewis number $Le = \alpha/D = 1$. This means that the spray contour defined by density gradient and the contour defined by temperature gradient are identical.

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division September 21, 1992. Paper No. 93-ICE-21. Associate Technical Editor: J. A. Caton.

Results and Discussion

Effect of Ambient Density on Spray Penetration. For high-pressure spray, the relative value of the variation of ambient pressure is small, and in fact, the corresponding variation of the ambient density plays a chief role. Therefore, in this research density is used as a controlled parameter, and the variations of ambient pressure are reflected by the variation of density and temperature. The results show that with higher density, the spray tip penetration is shorter. The conclusion is basically in line with the previous work reported by other researchers.

Effect of Injection Pressure on Spray Penetration. Many researchers, according to the study of conventional pressure spray, proposed the following formula presenting the relationship between penetration S and injection pressure difference Δp [3, 6, 7]:

$$S \propto (\Delta p)^n$$

where n is a constant above zero. It is obvious from the formula that spray tip penetration at a given moment will increase monotonically with injection pressure. But Kamimoto has reported that at injection pressures above p_j = 80 MPa the increase in tip penetration seems to saturate [8]. In this paper, we found that the tip penetration at injection pressures of 134 MPa is smaller than that at injection pressure of 75 MPa. We believe that at an approximate injection pressure of 75–80 MPa the dependence of spray penetration will change from positive to negative. But because Kamimoto's experiment was carried out on a rapid compression machine and his data processing method is different from ours, one can hardly compare his results with ours directly. The present experimental results are shown in Fig. 3. For comparison, Varde's data obtained at an injection pressure of 140 MPa are also illustrated. They are relatively close to our data, which were obtained at an injection pressure of 134 MPa.

The above-mentioned phenomena can be explained by the fact that at high injection pressure, the atomization quality is improved, the fuel droplet's size and its inertia become small, so the droplet speed declines quickly; though the initial velocity of droplets is faster than that at ordinary injection pressure. When the former tendency prevails over the latter, the penetration will tend to decrease.

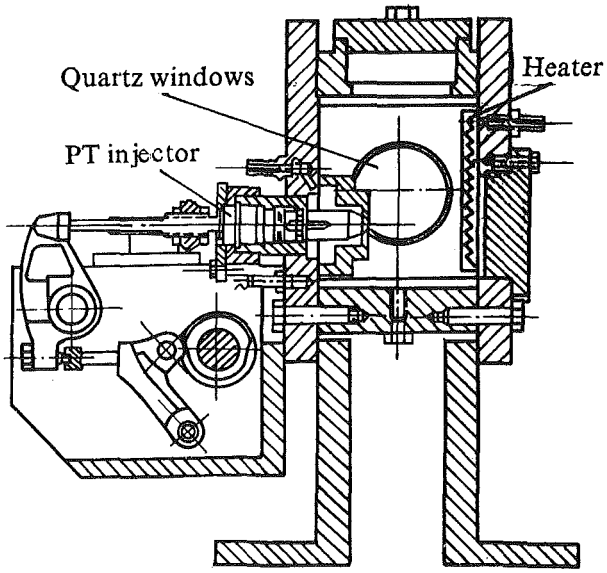


Fig. 1 Constant-volume bomb

Effect of Ambient Temperature on Spray Penetration. In various reported formulae calculating penetration, except Dent, the influence of ambient temperature is scarcely taken into consideration [3], but our research shows that the temperature influences high-pressure spray penetration rather distinctly, which is reduced with higher temperature. The explanation may be as follows: the diameter of high-pressure spray drops is much smaller than that at conventional pressure; hence, the spray evaporates more easily and loses its mass and momentum, which will reduce the penetration.

Empirical Formula. An empirical formula considering both the influences of ambient temperature and the injection pressure on penetration has been deduced:

$$S = A\left[\int \left(\frac{\Delta p}{\rho_a}\right)^{0.5} dt\right]^{0.5} \left(\frac{293}{T}\right)^{0.3}$$

$$A = 23.7/[(p_j/100)^3 + 7.91]$$

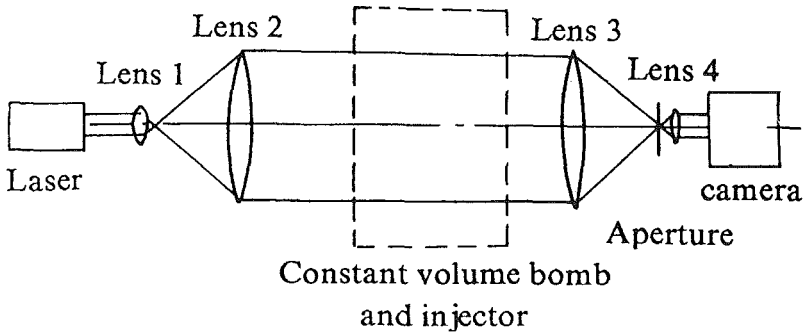


Fig. 2 Optical arrangement

Nomenclature

d = diameter of nozzle, m
D = diffusion coefficient
Le = Lewis number
n = constant
p_j = injection pressure, MPa
p_a = ambient pressure, MPa
$\Delta p = p_j - p_a$ = pressure difference, MPa
S = spray tip penetration, m
t = time delay related to the start of injection, ms
T = ambient temperature, K
α = thermal diffusivity
ρ_a = ambient density, kg/m^3
Ω = spray angle, deg

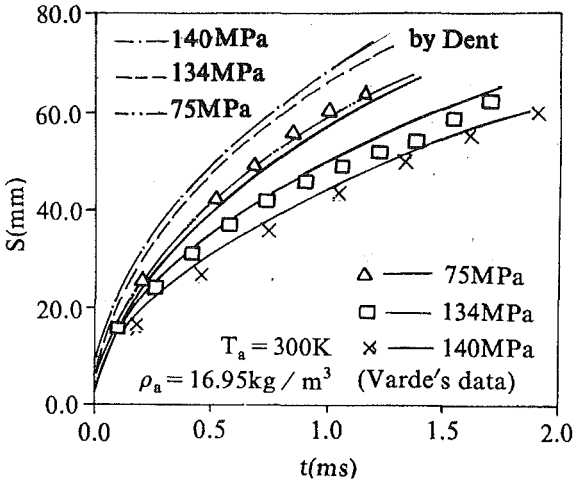


Fig. 3 Effect of injection pressure on penetration

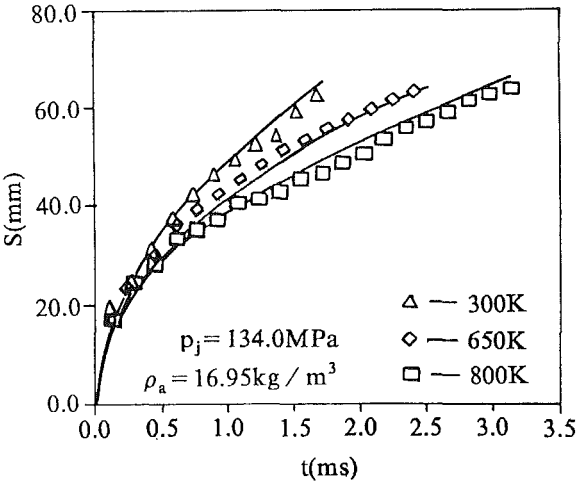


Fig. 4 Effect of ambient temperature on penetration

The factor A varies a little at injection pressures p_j ranging from 10 to 70 MPa; when $p_j = 10$ MPa, $A = 3$, which is close to the coefficient in the formulas of Dent [6] and Arai [7]. Predicted values of penetration obtained from the above formulae are illustrated by solid lines in Figs. 3 and 4, where one can see that predicted values conform fairly well with experimental data.

The comparison between present experimental results and calculated values from the relations given in [6] is shown in Fig. 3.

Effect of Injection Pressure and Ambient Condition on Spray Angle. During the spray developing period, no matter what the injection pressure and ambient condition are, the spray angle has a tendency to enlarge until reaching a relatively steady value. At the same ambient condition, an increasing of injection pressure will cause the spray angle to enlarge a bit; but at the same injection pressure, an increase of either medium density or temperature can make the high-pressure spray angle large. Figure 5 shows the visible contour development of high-pressure spray at different conditions, and the varying tendencies of the spray angle are given in Figs. 6–8.

Within the range of injection pressures p_j = 75–134 MPa, the law that spray angle varies with injection pressure and medium density is basically similar to that at ordinary injection pressure. But at the high injection pressure of 134 MPa, spray angle increases remarkably with higher temperature. Especially when both the ambient temperature and density are very high

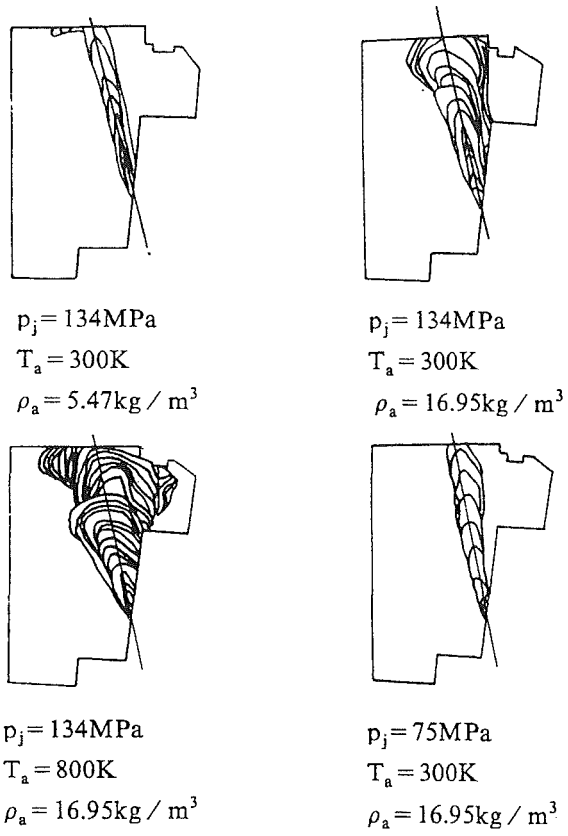


Fig. 5 Growth of spray envelope

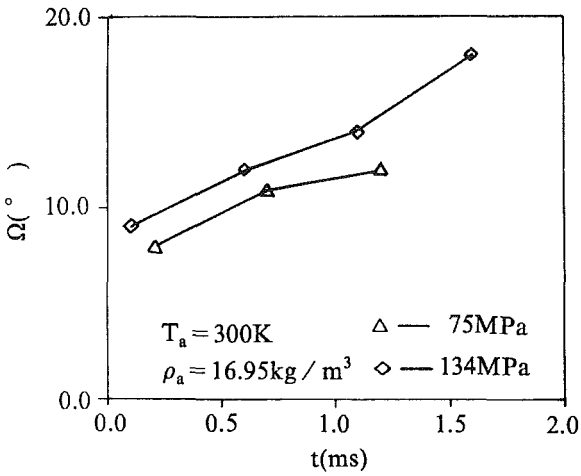


Fig. 6 Spray angle variation with time at different injection pressure

(for example, $T = 800$ K, $\rho_a = 16.95$ kg/m^3), the spray angle can expand to more than 30 deg. The phenomenon that spray angle expands sharply at high temperature can be regarded as an essential feature of high-pressure sprays. The reason for such a phenomenon lies mainly in the fact that the droplets of high-pressure spray can turn rapidly from liquid state into gaseous state in a high-temperature environment because of their small size; thus their volume will expand extremely. This is obviously quite beneficial for improving air utilization. Besides improving atomization and air entrainment, the expansion of spray volume may be the third and the most important favorable factor of the combustion process for high-pressure sprays.

Wall Impingement of High-Pressure Spray. The process when the spray impinges on a metal wall surface at injection

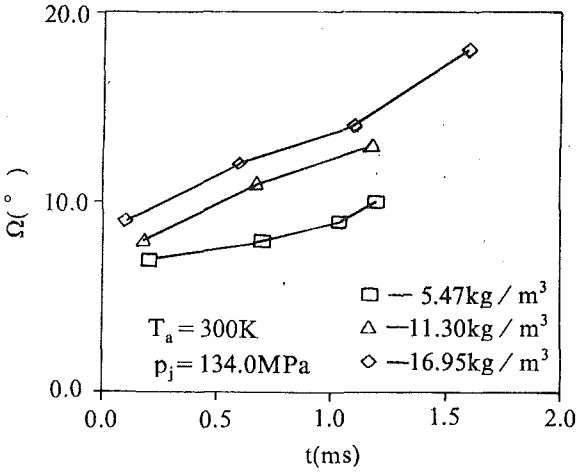


Fig. 7 Spray angle variation with time under different medium densities

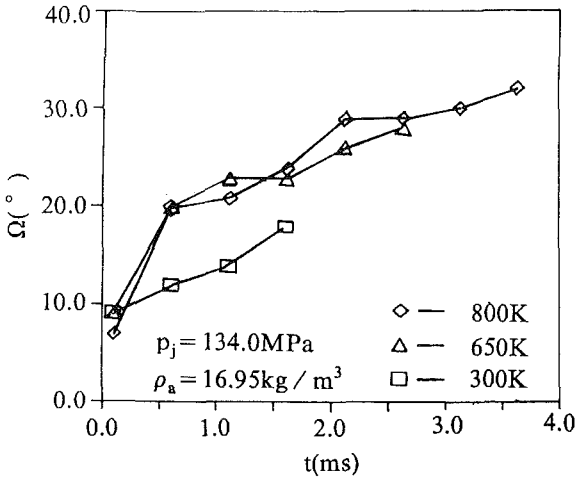


Fig. 8 Spray angle variation with time under different ambient temperatures

pressures from 75 to 134 MPa and wall temperatures from 300 to 800 K has been observed by high-speed photography. Figures 9(*a*) and 9(*b*) represent two typical patterns of high-pressure spray impinging on the wall at room temperature and high temperature, respectively. A new finding is that the phenomena at high injection pressure of 134 MPa have some different characteristics compared with those reported in [9–11] at ordinary injection pressures. For the high-pressure spray, at room temperature, one can hardly see the sign of spray developing along the wall surface by photography (see Fig. 9 (*a*)), though laser diffraction particle analyzer can detect the existence of liquid drops (see [5]). Only at a high temperature can one see results similar to what have been reported in the abovementioned references. The higher the temperature is, the thicker the fuel spray near the wall surface is, while the length developing along the surface reduces and a vortex appears in the front of the spray developing along the surface, i.e., it is similar to a conventional pressure spray, or gas wall impinging jet as reported in [12, 13]. With the wall surface at high temperature, a thin, transparent layer appears closely between the surface and spray. This layer seems to be mainly occupied by gas, whose appearance may influence both the spray wall impingement and heat transfer processes. The phenomena mentioned above can be seen from the photographs in Fig. 9[*b*].

Conclusions

1 Spray penetration does not increase monotonically with injection pressure. The penetration is smaller at an injection pressure of 134 MPa than that at 75 MPa.

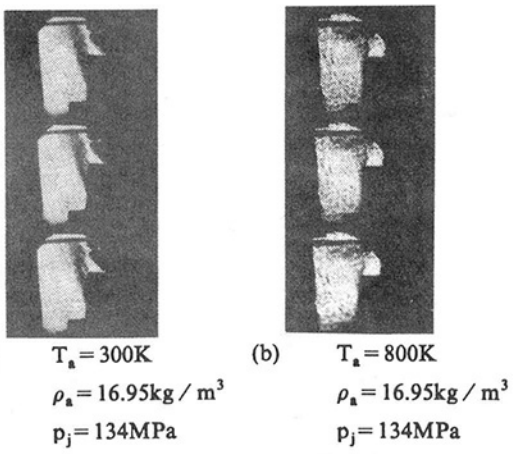


Fig. 9 Wall impingement at different temperatures

2 High-pressure spray penetration is influenced by both the ambient pressure and the temperature. When the pressure or temperature increases, the penetration decreases.

3 High-pressure spray angles expand sharply with higher temperatures, which differs evidently from conventional pressure spray and gas jet. This characteristic benefits the improvement in air utilization.

4 When high-pressure spray arrives at the wall, if the wall surface has a lower temperature, the rebound is not obvious, and the phenomenon of spray developing along the surface is invisible. Only at a higher surface temperature can a spray layer developing along the surface be seen. When the wall temperature reaches a certain degree, a relatively thin, transparent gas layer, which will have an influence upon the rebound of spray and heat transfer, appears between the spray and the wall.

Acknowledgments

This project is supported by National Science Foundation of China.

References

1 Beck, N. J., "High Pressure Fuel Injection—A Rational Approach to Diesel Engine Efficiency, Emissions and Economics," SAE Paper No. 830860, 1983.

2 Kato, T., et al., "Spray Characteristics and Combustion Improvement of D.I. Diesel Engine With High Pressure Fuel Injection," SAE Paper No. 890265, 1989.

3 Hiroyasu, H., "Diesel Engine Combustion and Its Modeling," presented at the International Symposium on Diagnostics and Modeling of Combustion in Reciprocating Engines, Tokyo, Japan, Sept. 4–6, 1985.

4 Varde, K., et al., "Diesel Fuel Spray Penetration at High Injection Pressures," SAE Paper No. 830448, 1983.

5 Yun-yi, Gong, et al., "An Investigation on Droplet Size Distribution and Evaporation of Diesel Fuel Sprays at High Injection Pressure by Using Laser Diagnostic Technique," SAE Paper No. 920090, 1992.

6 Dent, J. C., et al., "A Basis for the Comparison of Various Experimental Methods for Studying Penetration," SAE Paper No. 70-571, 1971.

7 Arai, M., et al., "Disintegrating Process and Spray Characterization of Fuel Jet Injected by a Diesel Nozzle," SAE Paper No. 840275, 1984.

8 Kamimoto, T., et al., "Effect of High Pressure Injection on Soot Formation Processes in a Rapid Compression Machine to Simulate Diesel Flames," SAE Paper No. 871610, 1987.

9 Fujimoto, H., et al., "Characteristics of a Diesel Spray Impinging on a Flat Wall," Proceedings of the 4th ICLASS—'88, Aug. 22–24, 1988, Sendai, Japan, pp. 439–446.

10 Kataura, N., et al., "Characteristics of a Diesel Spray Impinging on a Flat Wall," SAE Paper No. 890264.

11 Hiroyasu, H., et al., "Computed Tomographic Study on Internal Structure of a Diesel Spray Impinging on a Flat Wall," *International Symp. COMODIA*, Vol. 90, 1990, pp. 205–210.

12 Fujimoto, H., Senda, J., et al., "Characteristics of a Diesel Spray Impinging on a Flat Wall," *International Symp. COMODIA*, Vol. 90, 1990, pp. 193–198.

13 Tanabe, H., et al., "Experimental Study on Unsteady Wall Impinging Jet," SAE Paper No. 900605, 1990.

Multidimensional Modeling of Diesel Ignition and Combustion Using a Multistep Kinetics Model

S.-C. Kong
Graduate Student.

R. D. Reitz
Professor.

Engine Research Center,
Department of Mechanical Engineering,
University of Wisconsin—Madison,
Madison, WI 53706

Ignition and combustion mechanisms in diesel engines were studied using the KIVA code, with modifications to the combustion, heat transfer, crevice flow, and spray models. A laminar-and-turbulent characteristic-time combustion model that has been used successfully for spark-ignited engine studies was extended to allow predictions of ignition and combustion in diesel engines. A more accurate prediction of ignition delay was achieved by using a multistep chemical kinetics model. The Shell knock model was implemented for this purpose and was found to be capable of predicting successfully the autoignition of homogeneous mixtures in a rapid compression machine and diesel spray ignition under engine conditions. The physical significance of the model parameters is discussed and the sensitivity of results to the model constants is assessed. The ignition kinetics model was also applied to simulate the ignition process in a Cummins diesel engine. The post-ignition combustion was simulated using both a single-step Arrhenius kinetics model and also the characteristic-time model to account for the energy release during the mixing-controlled combustion phase. The present model differs from that used in earlier multidimensional computations of diesel ignition in that it also includes state-of-the-art turbulence and spray atomization models. In addition, in this study the model predictions are compared to engine data. It is found that good levels of agreement with the experimental data are obtained using the multistep chemical kinetics model for diesel ignition modeling. However, further study is needed of the effects of turbulent mixing on post-ignition combustion.

Introduction

Ignition delay is considered to be an important parameter in the operation of a diesel engine since it influences hydrocarbon and NO_x emissions. During the delay period, the injected fuel undergoes complex physical and chemical processes such as atomization, evaporation, mixing, and preliminary chemical reaction. Ignition takes place after the preparation and reaction of the fuel–air mixture leads to fast exothermic reaction. Several approaches have been made to model the autoignition phenomena in multidimensional modeling of diesel engines.

Single-step irreversible Arrhenius kinetics models have often been employed since it is easy to use them. Usually those models use an empirical reaction equation to account for the energy release. Since the physics of the mixing process is already taken into account by the fluid dynamic model equations, the single-step kinetics model has been found to work reasonably well with some adjustments to model constants for different engine operating conditions.

However, in order to make the ignition model more predictive, multistep chemical kinetics models must be considered. In the available models such as the Shell [1] and IFP [2] models, several necessary generic reactions are formulated for simulating the controlling elementary reactions, i.e., initiation, propagation, branching, and termination. The corresponding model constants are also fitted and adjusted to reproduce measured data in the development of these models. Such detailed kinetics models are believed to be able to capture more of the physics and chemistry of the autoignition event. Successful applications of these models can be found for modeling the diesel ignition process [3, 4].

Other more detailed chemistry models have also been developed for modeling the autoignition of hydrocarbons in homogeneous mixtures [5, 6]. Applications of these models to diesel spray ignition modeling are still very limited due to the large computer resources that are required.

For the purpose of engineering application, reduced reaction kinetics models are more useful for the multidimensional modeling of ignition chemistry. Among these models, the Shell model has been chosen to simulate the diesel spray ignition process for this study. This kinetics model was incorporated in the KIVA code [7] and was applied to reproduce ignition delay data of homogeneous mixtures in a rapid compression machine [1]. Next, the physical significance of the model parameters is discussed and the sensitivity of results to the model constants is assessed. The model is then applied to predict

Contributed by the Internal Combustion Engine Division and presented at the Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1992. Paper No. 93-ICE-22. Associate Technical Editor: J. A. Caton.

dodecane spray ignition delay measurements in a combustion bomb [8]. The above studies make use of two-dimensional calculations. Finally, three-dimensional calculations are presented using the Shell kinetics model coupled with a post-ignition combustion model to predict combustion in a Cummins diesel engine, and the results are compared with experimental data [9].

Model Formulation

The computations were performed using a modified version of the KIVA-II code, which solves the three-dimensional equations of transient chemically reactive fluid flows with sprays [7]. Turbulence was modeled using the standard k–ϵ turbulence model. However, the turbulence boundary conditions were modified to account for compressibility effects by using a coordinate transformation in the wall region [10]. Flow through the piston-cylinder-ring crevice was accounted for in the present study using a phenomenological crevice flow model [11]. Wall heat and momentum boundary conditions were modeled using the modified wall function model [10], which is based on the solutions of the one-dimensional unsteady energy equation. The spray model considered the breakup of the injected liquid parcels and resulting drops by using results from a stability analysis of liquid jets [12].

Implementations of the entire combustion model are emphasized in this study, including various ignition kinetics models, which will be discussed shortly. The post-ignition combustion was modeled using both a one-step Arrhenius kinetics model and a characteristic-time model [13]. Essentially the one-step Arrhenius kinetics model refers to the original KIVA combustion model [7].

With the characteristic-time model, the time rate of change of the partial density of species i, due to conversion from one chemical species to another, is given by

$$\frac{dY_i}{dt} = -\frac{Y_i - Y_i^*}{\tau_c} \tag{1}$$

where Y_i is the mass fraction of species i, Y_i^* is the local and instantaneous thermodynamic equilibrium value of the mass fraction, and τ_c is the characteristic time for the achievement of such equilibrium. The characteristic time τ_c is assumed to be the same for the seven species considered necessary to predict thermodynamic equilibrium temperatures accurately: fuel, O_2, N_2, CO_2, CO, H_2, H_2O. Among these seven species, six reactive species (i.e., all except N_2) are accounted for in order to solve the local and instantaneous thermodynamic equilibrium values Y_i^*. This characteristic-time model has demonstrated good performance in modeling combustion processes in spark-ignited engines [10, 11, 13], compression-ignited homogeneous-charge engines [14], and diesel engines [15, 16].

Ignition Kinetics

Unlike the case of conventional spark-ignited engines where combustion is initiated by a spark, the ignition model should be able to predict the autoignition process in a compression-ignited engine. The single-step Arrhenius kinetics model has been widely used to model ignition in diesel engine studies. In our previous studies, the characteristic-time combustion model has been used successfully to predict both the ignition and combustion processes in compression-ignited engines, including homogeneous-charge and diesel engines [14–16]. In these studies, the characteristic time scale for ignition was derived either from elementary initiation reactions or from suspended single droplet ignition delay measurements. Since the complex autoignition processes are represented by only one reaction with this approach, the pre-exponential constant in the reaction rate had to be modified for a best fit of engine data. The adjusted constant thus also accounts for fluid dynamics effects and the different pathways of the controlling reactions, plus the time needed to build up radical pools. In order to take better account of the complexity of these ignition processes, and to make the ignition model more predictive, more detailed ignition kinetics models for diesel ignition modeling are considered in this study.

In particular, the Shell model, which was developed to predict the knock in a gasoline engine, is used to predict the autoignition of hydrocarbon fuels [1]. This multistep reaction scheme and the corresponding model parameters were developed by comparison with rapid compression machine experiments. Homogeneous mixtures, with the octane quality and sensitivity of fuels similar to gasoline, were compressed and autoignited in the rapid compression machine. The temperature histories were calculated from the charge conditions and the recorded pressure traces. Hot ignition was defined when the temperature exceeded 1100 K or had a rise rate greater than 10^7 K/s. Extensive ignition delay measurements were performed for a wide range of end-of-compression conditions. A mathematical model was then developed to describe the autoignition of hydrocarbon fuels [1].

Eight generic reactions based on the degenerate branching characteristics of hydrocarbon autoignition were formulated with five generic species. Basically the rate constant of each reaction was of the Arrhenius form. The mathematical model for this generalized reaction set contains 26 kinetic parameters [1]. The generic species and reactions involved in the model are as follows:

$$\mathrm{RH} + \mathrm{O_2} \rightarrow 2\mathrm{R}^* \qquad K_q \tag{2}$$

$$\mathrm{R}^* \rightarrow \mathrm{R}^* + \mathrm{P} + \mathrm{Heat} \qquad K_p \tag{3}$$

$$\mathrm{R}^* \rightarrow \mathrm{R}^* + \mathrm{B} \qquad f_1 K_p \tag{4}$$

$$\mathrm{R}^* \rightarrow \mathrm{R}^* + \mathrm{Q} \qquad f_4 K_p \tag{5}$$

$$\mathrm{R}^* + \mathrm{Q} \rightarrow \mathrm{R}^* + \mathrm{B} \qquad f_2 K_p \tag{6}$$

$$\mathrm{B} \rightarrow 2\mathrm{R}^* \qquad K_b \tag{7}$$

$$\mathrm{R}^* \rightarrow \text{termination} \qquad f_3 K_p \tag{8}$$

$$2\mathrm{R}^* \rightarrow \text{termination} \qquad K_t \tag{9}$$

where RH is the hydrocarbon fuel (C_nH_{2m}), R^* is radical formed from fuel, B is branching agent, Q is labile intermediate species, and P is oxidized products, consisting of CO, CO_2, and H_2O in specified proportions. The expressions for K_q, K_p, K_b, K_t, f_1, f_2, f_3, f_4, etc., are given by Halstead et al. [1]. In addition, the local concentrations of O_2 and N_2 are needed to compute the reaction rates.

There have been several successful implementations for modeling engine knock using the Shell model [17, 18]. Natarajan and Bracco [17] tested the Shell model's ability to simulate cool flame phenomena in a stirred reactor. The model was also implemented into the REC code and was used to predict engine knock in two-dimensional calculations. It was noticed that some inhibition criteria had to be employed when the low-temperature kinetics model was coupled with a high-temperature flame chemistry model.

A similar situation was encountered when the Shell model was incorporated in the REC-P3 code in two-dimensional calculations by Schaperton and Lee [18]. The generic species, R^*, B, and Q, were assigned molecular weights for the consideration of mass balances. Another modification was to freeze the reaction rate above 950 K to prevent unreasonable oscillations of the branching agent, B, concentration because the branching reaction was found to be much faster than the production of B. This worked well since at such high temperatures the reaction tended to be in quasi-equilibrium.

A plausible implementation of the Shell model to predict

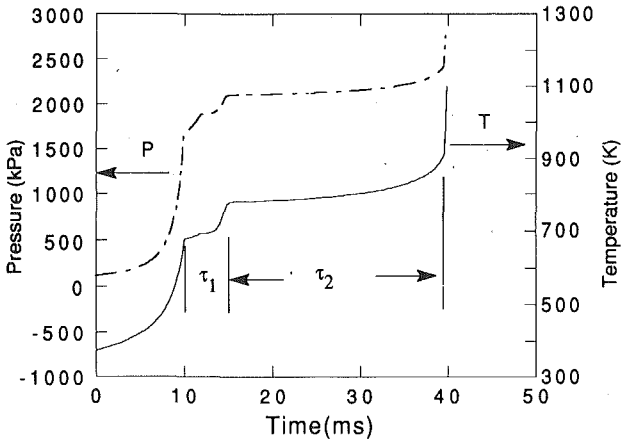


Fig. 1 Temperature and pressure history of the RCM for 100 RON (iso-octane) with pre-compression temperature 373 K and pressure 15 psia

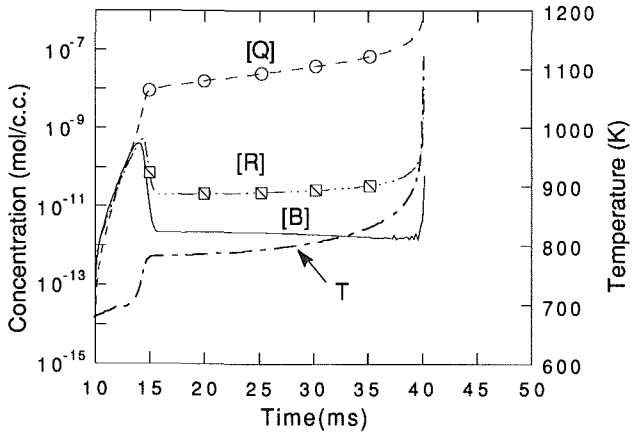


Fig. 2 Temperature and radical concentration history of RCM for 100 RON

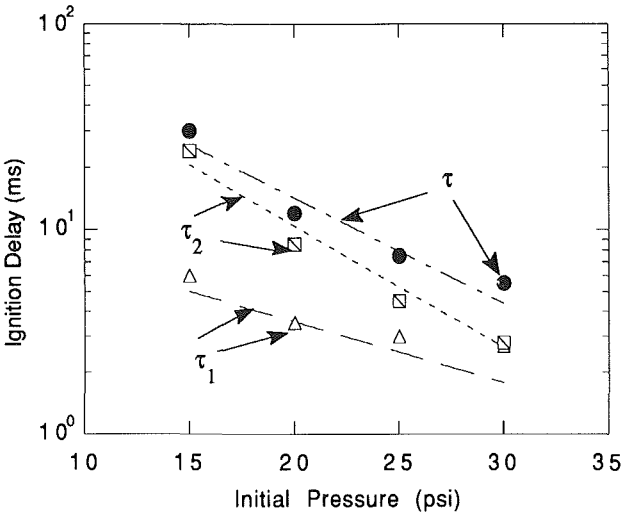


Fig. 3 Comparison between measured (symbols) and reproduced (lines) ignition delay data as a function of pre-compression pressures

spray ignition delay was made by Theobald and Cheng [3]. A method of adjusting the numerical timestep dynamically during ignition period was applied to avoid numerical problems. The model constants had to be modified to predict accurately the ignition delay of a dodecane spray in a combustion bomb in their two-dimensional axisymmetric computations using the KIVA-I code. With these modifications the model was applied to simulate spray combustion experiments in a rapid compression machine.

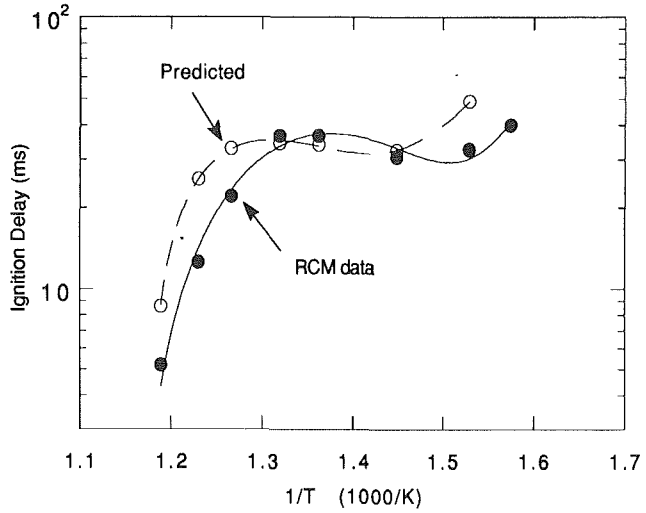


Fig. 4 Preliminary results of predicted ignition delay for the case P_i = 15 psi

In the present study, the Shell model is applied with some modifications for multidimensional calculations of diesel ignition coupled with KIVA-II solution. However, this study differs from the previous study of Theobald and Cheng [3] in two ways. One is the different features of KIVA-I and KIVA-II code that have been documented [10], including the use of the k–ϵ turbulence model and the improvements in numerical schemes, solution accuracy, and other physical submodels. The other is the modifications to the original KIVA-II code, which have been described in the previous section, including the use of a different spray model [12].

Results and Discussion

Rapid Compression Machine. The ignition delay measurements of iso-octane performed in the Thornton rapid compression machine [1] with different pre-compression charge pressures (15, 20, 25, and 30 psi) were chosen to test the model prediction. The first approach was to solve the six differential equations [1] derived from the reaction set to obtain the concentration of species RH, O_2, R^*, B, Q, and the temperature history. Figure 1 shows the typical pressure and temperature histories of the rapid compression machine starting from the beginning of compression. Two-stage ignition phenomena can be observed easily as indicated by the delay times, τ_1 and τ_2. The species concentration and temperature histories are shown in Fig. 2. The first-stage ignition delay is defined from the start of deceleration of pistons (10 ms) to the peak value of radical R^*. The second-stage ignition delay then follows up to the point where the temperature exceeds 1100 K.

Stiffness problems were indeed observed in some of these calculations where the concentration of the branching agent B oscillates somewhat just before the onset of hot ignition. Fortunately this does not affect the ignition prediction since the timestep was adjusted dynamically when the branching reaction becomes rapid. The reproduced ignition delay data are shown in Fig. 3 and the total delay $\tau = \tau_1 + \tau_2$.

Next the Shell model was incorporated into the KIVA code. A two-dimensional calculation was performed for ignition delay prediction using the experimental conditions and geometry of the Thornton RCM. In this case it is noted that some discrepancies are expected due to the different approximations inherent in zero- and multidimensional calculations. The kinetics scheme itself is expected to be qualitatively correct (in terms of its two-stage ignition and negative temperature coefficient features) but the nonhomogeneity of the cylinder charge temperature causes different heat transfer rates in the multi-

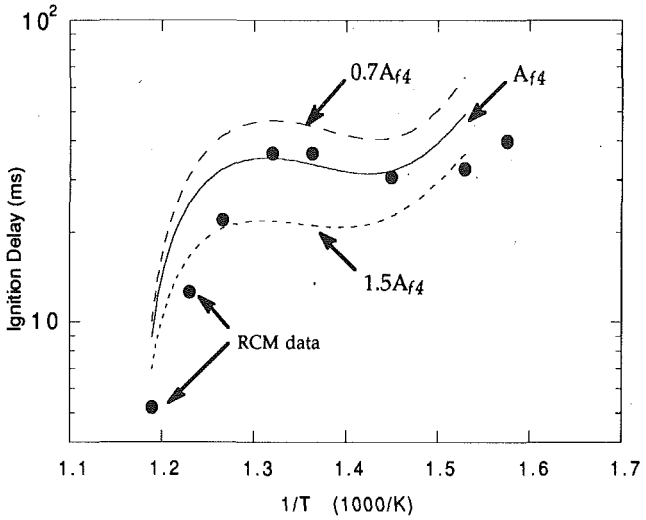


Fig. 5 Effects of model constant A_{f4} on the ignition delay

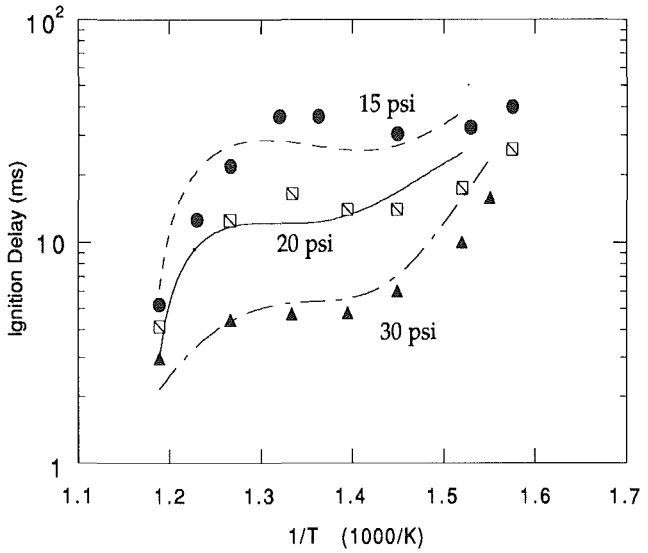


Fig. 6 Comparisons between predicted (lines) and measured (symbols) ignition delay for different pre-compression pressures using $A_{f4} = 1.9 \times 10^4$

dimensional calculations and this effect influences the prediction of ignition delay.

Ignition delay predictions are obtained as a function of the induction (end-of-compression) temperature. The autoignition event is signaled by observing the rapid temperature rise of the core gas. Generally speaking, the core gas temperature was found to be 50 K higher than the average temperature at the time of ignition. Figure 4 shows results from the implemented Shell model coupled with the two-dimensional KIVA solution using the original Shell model constants. For the case shown in Fig. 4, the diluent is nitrogen and the wall temperature is 373 K, which is chosen to be the baseline case for further sensitivity study.

Some discrepancies are observed in Fig. 4. It should be noted that in the Shell model formulation, the temperature was calculated from the measured pressure data, and the energy equation was formulated using an average temperature. However, in reality the core gas is responsible for the autoignition. If a higher gas temperature is used to start the calculation, the predicted curve would move to the right, in better agreement with the experiments. The same argument was made by Hu and Keck [5] in which the core gas temperature was assumed to be 80 K higher than the average value in their model.

Indeed there are other uncertainties associated with the modeling. Sensitivity studies, not discussed here, were performed to investigate the effects of the diluent composition and wall temperature, which could also influence the heat transfer characteristics. The results show that the effects of the diluent are much more significant during the compression stroke than in the induction period. The wall temperature also affects the predicted results. However, the most important parameters were found to be the model constants, as discussed next.

Sensitivity and Interpretation of Model Parameters. The philosophy of the Shell model is based on the fact that degenerate branching plays an important role in the autoignition of hydrocarbon fuels, and the observation of both cool flame and negative temperature coefficient phenomena in autoignition [19]. A chain-propagation cycle is formulated to describe the history of the branching agents. In addition, one initiation and both linear and quadratic termination appear in the model.

The initiation reaction is RH + $O_2 \rightarrow 2R^*$, and then the main propagation cycle follows, viz., $R^* \rightarrow R^*$, which produces the exothermicity. Since the degenerate branching plays a major role in the autoignition event, some routes appear to account for the formation and reaction of the branching agents, including the reaction, $R^* \rightarrow R^* + B$. This reaction is responsible for the first-stage ignition. However, the dominance of this reaction decreases after a certain amount of rise of temperature and cool flame phenomena are observed. Therefore, a termination reaction is formed to take the radicals out of the propagation route, $R^* \rightarrow$ out, whose reaction rate is also considered as a propagation rate ($f_3 K_p$). This reaction dominates in the moderate temperature range and slows down the overall reaction during the second induction period.

Another mechanism is used to form the branching agents that lead to hot ignition. The formation of the branching agents is related to the intermediate reactive species that are formed during and/or after the cool flame appearance. Thus, two reactions are added to account for the formation of the intermediate species and the secondary build-up of branching agents. They are $R^* \rightarrow R^* + Q$ and $R^* + Q \rightarrow R^* + B$. These two reactions are also parts of the propagation cycle. After the formation of the branching agents, the branching reaction comes into play through the reaction mechanism, $B \rightarrow 2R^*$. Finally, a quadratic termination reaction serves as the terminator of radicals, $2R^* \rightarrow$ out. The above reduced kinetics scheme describes the main features of hydrocarbon autoignition induced by degenerate branching.

Since all the Shell model constants were fitted under the specified experimental conditions and the corresponding approximations in the governing equations, certain model constants should be adjustable if a different submodel, i.e., heat transfer model, is used, as in this study.

When difficulties were encountered in extending the Shell model to multidimensional applications, some model parameters needed to be modified in order to predict existing ignition delay data correctly. Special attention was paid to the build-up of the branching agent B. In the model, the reaction rates of forming B in the first stage, $R^* \rightarrow R^* + B$, and the second stage, $R^* + Q \rightarrow R^* + B$, are very fast. However, the intention of the Shell model is to put the formation of Q, i.e., $R^* \rightarrow R^* + Q$, as the rate-limiting reaction for the following chain reactions, $R^* + Q \rightarrow R^* + B$ and then $B \rightarrow 2R^*$. In other words, the formation of the intermediate species Q is more responsible for the hot ignition than any other species. Figure 5 shows the effects of the pre-exponential constant A_{f4} on the overall ignition delay. It can be seen that the constant A_{f4} has a significant effect on the prediction of ignition delay. This sensitivity was also found in modeling diesel ignition [3].

After the above sensitivity study, it was decided to modify the constant A_{f4} slightly from 1.7×10^4 to 1.9×10^4. Three cases with different precompression pressures, i.e., 15, 20, and

Table 1 Combustion bomb experimental conditions

P_{amb}	T_{amb} range	Atmosphere
11 *atm*	670 - 900 K	Air
21 *atm*	670 - 900 K	Air
31 *atm*	670 - 900 K	Air
11 *atm*	295 K	N_2
31 *atm*	295 K	N_2

30 psi, respectively, were chosen to be studied. The results are shown in Fig. 6 where the negative temperature coefficient feature can be observed.

By reviewing the chemistry model of Benson [20] and earlier research by Salooja [21] for cool flame, some interpretation of the nature of the generic species and their roles in the Shell model is proposed here. As branching agents, peroxides are more important during the autoignition event of hydrocarbon fuels than aldehydes. Conclusively, the branching agent B is related to hydroperoxides (RO_2H) at low temperatures, and to hydrogen peroxide (H_2O_2) at high temperatures. The intermediate species Q is related to the aldehydes (RCHO) during first induction period and to the alkylperoxy radical (HO_2) and its isomerization products, which are responsible for the formation of H_2O_2, during the second induction period. The combined roles of HO_2 and H_2O_2 were also found to be crucial during high-pressure autoignition by Pitz and Westbrook [22] in their engine knock study.

Spray Ignition. The Shell model was further implemented to predict diesel spray ignition under enginelike conditions. This ignition chemistry model [1] was incorporated in KIVA [7] together with a spray breakup model [12] to compare with spray ignition delay measurements of Igura et al. [8]. The experimental conditions of Igura et al. are listed in the first three lines of Table 1. The combustion chamber was a cylindrical constant-volume bomb and the fuel was simulated using dodecane ($C_{12}H_{26}$) and the injection velocity was 100 m/s.

Assessment of the performance of the spray model was made by comparing the predicted and measured penetration history of a noncombusting spray as also shown in Table 1. The computation used a two-dimensional nonuniform axisymmetric mesh with 20 × 26 grid points in the radial and axial directions, respectively. The spray penetration as a function of time is shown in Fig. 7 for the noncombusting spray cases. Good levels of agreement are obtained using the current spray model.

The autoignition spray experiments [8] were done with the fuel injected along the axis of a cylindrical combustion bomb of 180 mm radius and 540 mm length. The cylinder walls were heated electrically and the initial air temperature and pressure were adjustable over the range of 670 ~ 900 K and 1 ~ 31 atm absolute, respectively. Ignition was detected by a photodetector such that the illumination delay time was measured. A single-orifice nozzle with diameter 0.5 mm and an injection pressure of 96.8 atm absolute were used.

The initial Sauter mean radius of the injected droplets was assumed to be 2.5 mm to allow comparisons of the results of Theobald and Cheng [3], who also used this assumption. The chemical sub-timestep was dynamically adjusted within the fluid timestep according to the production rate of branching agent B. Only a specified percentage change in branching agent concentration was allowed per chemical sub-timestep. In this way there could be up to 150 chemical subcycles within a fluid timestep during ignition period in some cases. Fortunately this large number of subcycles does not increase the computation time significantly since the ignition event is comparably short, and other spray phenomena such as the vaporization and collision processes consume most of the computer time.

The overall combustion model was formulated by interfacing

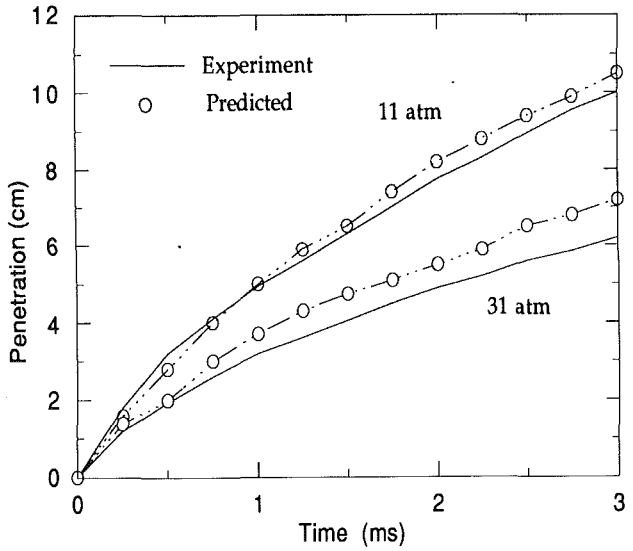


Fig. 7 Comparison between measured and predicted spray penetration distance as a function of time

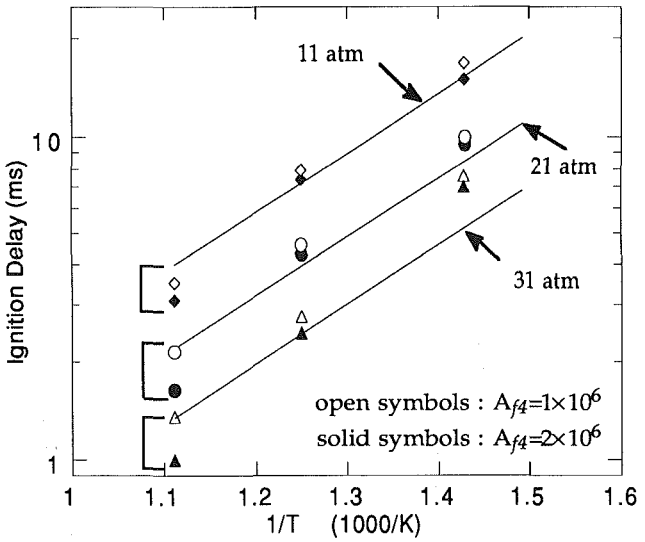


Fig. 8 Effects of the parameter A_{f4} on the ignition delay prediction of the dodecane spray (lines: measured; symbols: predicted)

the ignition chemistry model with a single-step high-temperature reaction model of the Arrhenius form [3]

$$\text{Rate} = 5 \times 10^{12}[C_{12}H_{26}][O_2]\exp(-15780/T) \qquad (10)$$

This single-step kinetics model is inhibited when the temperature is lower than 900 K to ensure that the ignition kinetics dominate at low temperatures. The completion of ignition is chosen to be the time when cell temperature is greater than 1100 K. Then the autoignition model is turned off, partially because the ignition model constants have not been validated at such high temperatures. The single-step kinetic combustion model, Eq. (10), is then used to account for the main stage heat release. In the temperature range between 900 and 1100 K, both the autoignition kinetics and single-step combustion model are used to compute the fuel consumption rate. According to the formulation, every cell undergoes its own ignition event. However, once some hot spots are formed, energy is diffused or convected by fluid transport and the cells in neighborhood of hot spots have shorter ignition events. The mixture burns due to either autoignition or flame propagation, or both.

Similar difficulties to the modeling of the RCM ignition

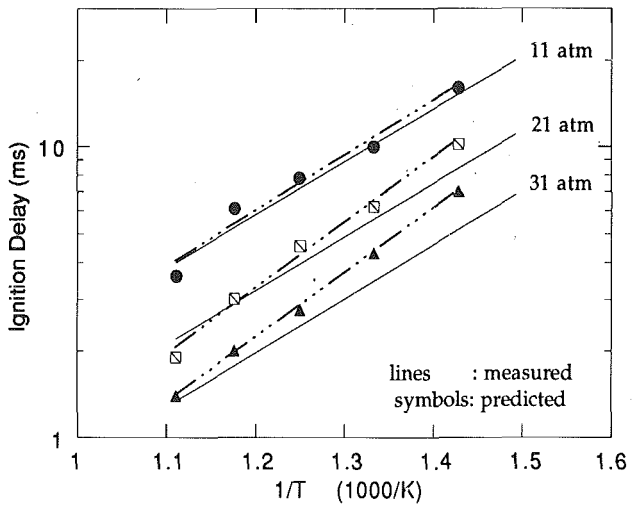


Fig. 9 Comparisons between measured (solid lines) and predicted (symbols and dashed lines) ignition delay

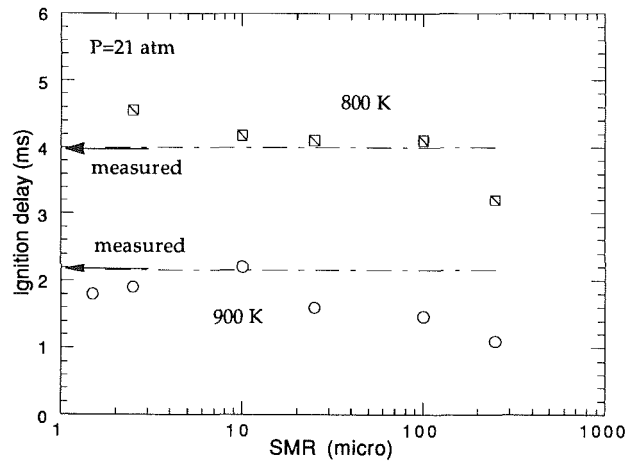


Fig. 10 Effects of initial drop sizes on ignition delay prediction

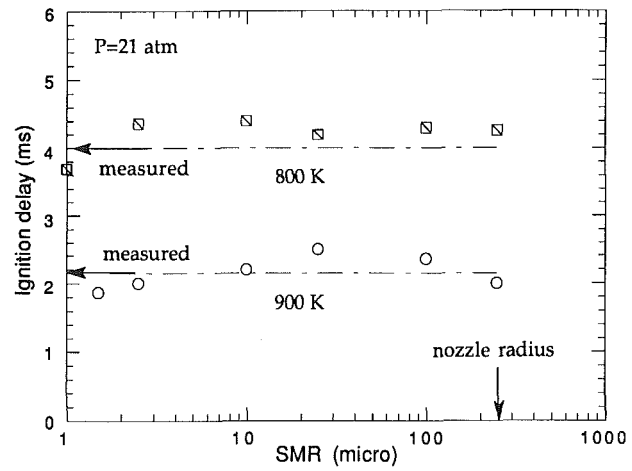


Fig. 11 Effects of initial drop sizes on ignition delay prediction using a revised model constant for the breakup time in atomization model

delay were observed. The model constant A_{f4} had to be adjusted to be able to reproduce the experimental results. Sensitivity of the predictions to this model constant was investigated for three cases with different charge bomb pressures over a wide range of initial temperatures. The results of sensitivity study are shown in Fig. 8 using 1×10^6 and 2×10^6 for A_{f4}.

Results from Fig. 8 show the predicted ignition delay is less sensitive to the model constant than in the case of homogeneous mixtures. This is attributed to the strong effects of the physical processes of evaporation and mixing. However, it is somewhat more sensitive to the value of the model parameter under high-pressure and/or high-temperature conditions. As a matter of fact, under high-temperature and pressure conditions, the optimum value of the model constant was more difficult to determine since the high rate of reaction caused some numerical difficulties.

For the dodecane spray, a suitable value of A_{f4} is found to be 1.3×10^6 in order to reproduce the measured ignition delay data. This value was chosen since it produces more accurate predictions in the high-pressure and temperature range, which is similar to diesel engine conditions. The predicted results are shown in Fig. 9 and general trends of increasing delay with decreasing pressure and/or temperature are well represented.

In addition, in order to investigate the effects of the injected drop size, a sensitivity study was also performed using different initial drop sizes for the computation. The results are shown in Fig. 10 for the 21 atm case. A general trend can be seen but the sensitivity is insignificant. In the spray breakup model, large injected drops soon break up into small drops [12]. This explains the lack of sensitivity to the injected drop size.

Other results showed that the ignition delay can also be predicted accurately using the nozzle radius as initial drop size (which is 250 μm in this study), and the results can be improved further by changing one spray model constant, which describes the drop breakup time [12]. The results shown in Fig. 11 use the revised breakup time model constant in the atomization model. It is worth noting that the new value of this constant, $B_2 = 1.73$ (the previous value was $B_2 = 10$ [12]), is determined from another study of the air-assisted atomization mechanism [23]. In principle, the current spray model allows the value for initial drop size to be predicted for different operating conditions.

Table2 Cummins NH engine specifications and operating parameters

Cylinder bore x stroke (mm)	139.7x152.4
Compression ratio	13.23
Displacement (L)	2.33
Number of spray nozzle orifices	8
Nozzle hole diameter (mm)	0.2
Spray angle (from head)	18 degrees
Combustion chamber	Quiescent
Piston crown	Mexican hat
Engine speed (rpm)	1500
Overall equivalence ratio	0.6
Fuel	Tetradecane
Injection duration	-18 ~11 atdc

Diesel Modeling. Validation of the ignition and combustion models makes use of results from a Cummins NH engine for which extensive experimental data are available with measurements of injection characteristics, cylinder pressure, and flame temperature from Yan and Borman [9]. The engine specifications and operating parameters are listed in Table 2. The initial conditions are calculated using a zero-dimensional cycle analysis program [24]. The injection velocity data were determined based on the injector link load versus crank angle data, assuming a constant nozzle hole discharge coefficient equal to 0.7. The cylinder gas pressure data from the experiments will be used for comparison in the present study.

The computations used a three-dimensional 45-deg sector mesh with fine numerical grid resolution. The mesh spacing was 23 × 25 × 18 cells in the radial, azimuthal, and axial directions, respectively, at the start of the computations, i.e., 150 deg BTDC. A perspective view of the computational mesh at top-dead-center together with the computed spray droplet locations is shown in Fig. 12.

Table 3 Descriptions of different cases

Case	comments
Case 1	original KIVA combustion model with one-step Arrhenius kinetics, i.e., Eq. (11)
Case 2	interfacing Shell and Arrhenius kinetics in between 900 - 1100 K, i.e., Shell plus Eq. (11)
Case 3	switch between Shell and characteristic-time model, which includes turbulence-mixing effects, at 1000 K, i.e., Shell plus Eq. (1).

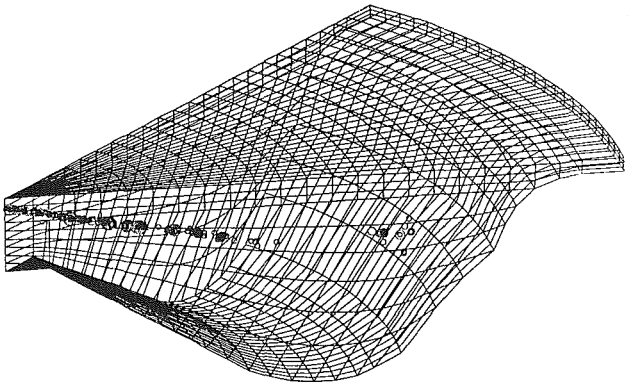


Fig. 12 Perspective view of the computation mesh at top-dead-center with the liquid fuel droplets

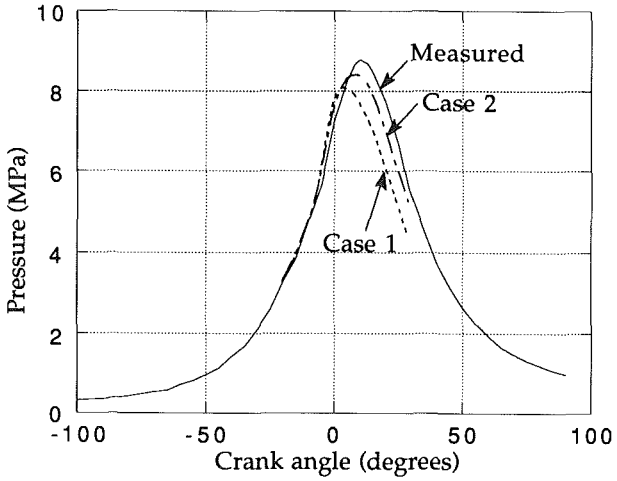


Fig. 13 Predicted pressure histories of Cases 1 and 2

In formulating the combustion model, use was made of the ignition delay measurements of a suspended droplet in high-temperature environments [25], from which an equivalent one-step overall reaction was formulated, i.e.,

$$\text{Rate} = A[C_{14}H_{30}]^{0.25}[O_2]^{1.5}\exp(-E/RT) \qquad (11)$$

where $A = 3.68 \times 10^{10}$ mol/cm^3-s and $E = 77.3$ kJ/mol. Preliminary results using the one-step Arrhenius kinetics model for modeling both the ignition and combustion processes (Case 1 in Table 3) are shown in Fig. 13. It should be noted that the model constant A had to be modified to best fit the engine experimental data, i.e., $A = 7.68 \times 10^8$ mol/cm^3-s [15].

For comparison, the Shell model was then implemented to serve as the ignition model with the same model constant A_{f4} as in the case of the combustion bomb spray ignition study. The same interfacing criteria were employed as discussed before that switch between the multistep ignition kinetics and high-temperature chemistry between 900 and 1100 K. The results of using this combined combustion model (Case 2 in Table 3) are also shown in Fig. 13. In this case, the high-temperature reaction was also formulated to use the original KIVA combustion model algorithm, with Eq. (11) for the energy release rate. From the comparisons in Fig. 13, superior performance of the implemented ignition and combustion models (Case 2) to that of the simple one-step reaction model (Case 1) can be observed.

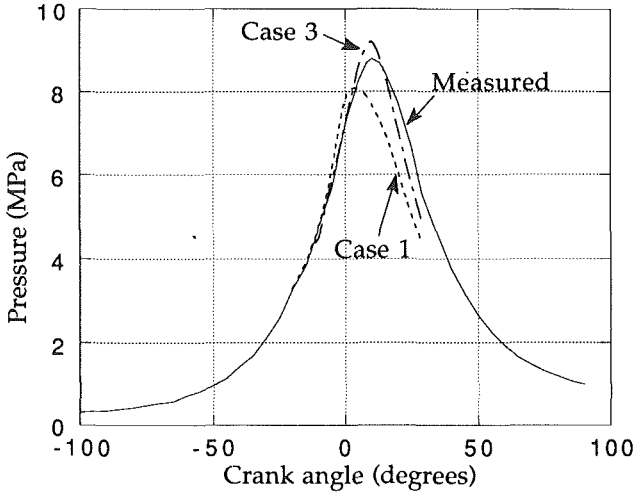


Fig. 14 Predicted pressure histories of Cases 1 and 3

Results from a sensitivity study of the effect of using different temperatures for the high-temperature limit of the autoignition model showed that the predicted pressure is insensitive to this temperature. When a higher temperature is used to switch off the autoignition model, i.e., replacing 1100 K by 1500 K, the peak pressure only slightly increases by 1.5 percent. A possible reason is that the ignition event for a high-temperature cell is extremely short. The time needed to reach 1100 K is almost the same as that to reach 1500 K. The fact that both the Shell model and the high-temperature chemistry model are responsible for the energy release during the period when each model is controlling the combustion rate suggests that the empirical energy release rate used in the Shell model propagation cycle (Eq. (3)) is quite close to that predicted by Arrhenius kinetics model (Eq. (11)).

However, the goal of the present study is to combine the Shell model and the characteristic-time model for the overall combustion modeling. Preliminary results using the characteristic-time combustion model for modeling both the ignition and combustion processes in our previous studies [15, 16] have also shown good levels of agreement. The characteristic time in Eq. (1) was derived from Eq. (11) and becomes

$$\tau_c = A^{-1}[C_{14}H_{30}]^{0.75}[O_2]^{-1.5}\exp(E/RT) \qquad (12)$$

Essentially this is also a single-step kinetics scheme for ignition modeling, which also accounts for the energy release during subsequent mixing-controlled combustion.

By using the single-step Arrhenius-type reaction (Cases 1 and 2), the turbulent mixing effects on post-ignition combustion were only taken account of by the diffusion terms in the fluid dynamic and spray models. Similarly, there is no turbulent time scale in Eq. (12) for the characteristic-time model [16]. However, for Case 3, the characteristic-time model that includes the effects of turbulence [27] was used with the Shell model to describe the high-temperature chemistry. In this case, the ignition model and the characteristic-time combustion model were switched at 1000 K. The results are shown in Fig. 14.

As can be seen in the figure, the predicted pressure is very close to the measured value during the early stages of combustion, i.e., before its peak value. Since the autoignition kinetics are forced to dominate up to 1000 K, it is seen by

comparing all the cases that the low-temperature chemistry model produces more accurate predictions during the ignition processes. It would be argued that interfacing low- and high-temperature chemistry models between 900 and 1100 K is more physically based as discussed before. However, in terms of ease-of-use in multidimensional modeling and the accuracy of the predicted pressures shown in Fig. 14, switching between the low- and high-temperature chemistry models at a temperature of 1000 K is preferred. In addition, the predicted cylinder gas pressure was found to be quite insensitive to the value of the threshold temperature. Only a 1 percent increase in peak pressure was observed when 1100 K was used for switching instead of 1000 K.

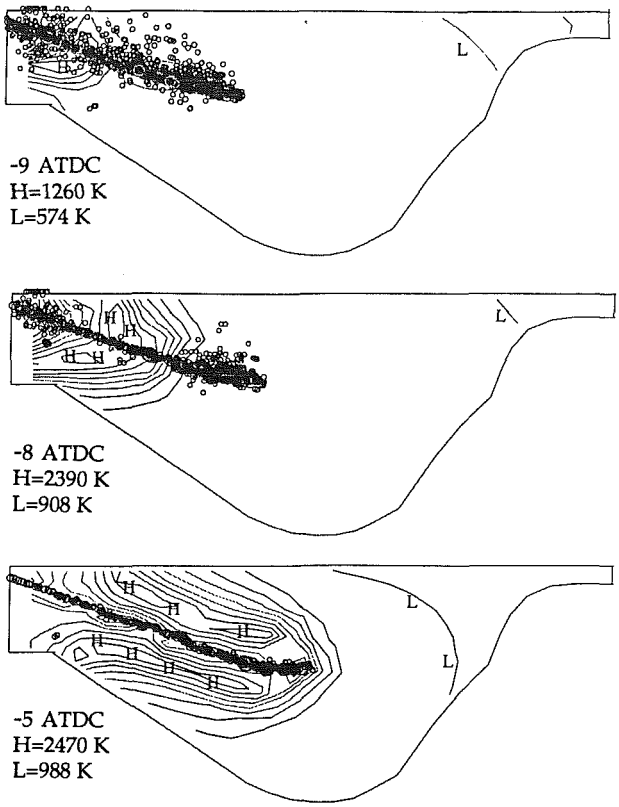


Fig. 15 Temperature contours and spray droplets distribution after ignition for Case 3 in Table 3

From the comparison of Cases 2 and 3, it indicates that the post-ignition results are sensitive to the details of the high-temperature chemistry model, although both cases use the same ignition model. However, further study is needed of the effects of turbulence and model parameters on post-ignition combustion.

From the temperature contours shown in Fig. 15 for Case 3 in Table 3, the spray ignition site is located one-third of the way between the nozzle exit and the spray tip. As also can be seen in Fig. 15, the spray autoignites at 9 deg BTDC with T_{high} = 1260 K, i.e., 9 crank angle deg after injection corresponding to an ignition delay time of 1 ms. The high-temperature regions, i.e., flames, quickly spread out downstream and have a value of T_{high} = 2390 K at 8 deg BTDC. Then the flames spread out further along the spray and the vaporized fuel/air mixture either autoignites or burns as the flame approaches. Finally the spray is surrounded by the flame. This sequence of events is in qualitative agreement with experimental studies of diesel spray ignition [26].

From the wide range of high-temperature contours surrounding the spray at 5 deg BTDC as in Fig. 15, the regions where the autoignition model is being used are very limited. As long as the energy release rate is accurately computed during the ignition period, a one-step reaction model would also produce reasonable predictions of the ignition process. This is probably the reason that a one-step reaction model can simulate the entire combustion processes reasonably well (Case 1) as shown in Fig. 13. However, from the reasonably long ignition delay (1 ms), the ignition chemistry should also be considered during this period.

More detailed information about the ignition events corresponding to Case 3 in Table 3 can be observed in Fig. 16, which shows the evolution of the cell temperature and radical concentrations, including R*, B, and Q, at selected locations in the combustion chamber. As can be seen, the temperature is already high at the time of injection, i.e., 18 deg BTDC. The high temperature promotes the quick build-up of radicals

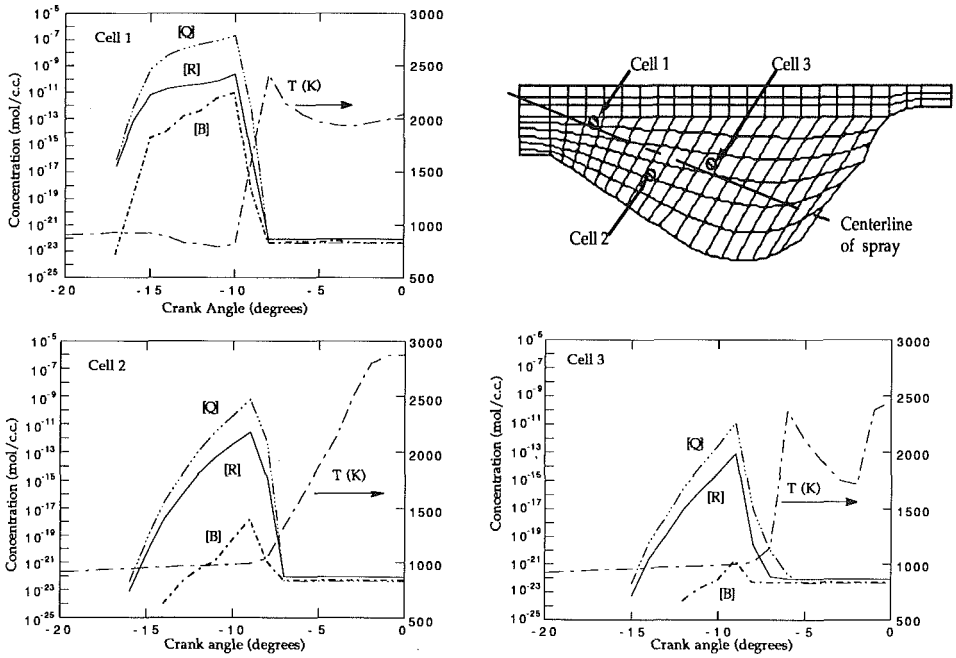


Fig. 16 Temperature and species concentration histories in specified cells obtained for Case 3 in Table 3

and then the reactions accelerate. A further slight increase of temperature produces a rapid decrease in the radical concentrations (when the ignition is taking place) followed by a subsequent fast temperature rise, i.e., hot ignition. Shortly after 1000 K is reached the autoignition model is deactivated and the radical concentrations are given an arbitrarily small value (10^{-22}) for numerical reasons.

The rise of the radical concentration is not as significant as it was in the RCM study because the high temperatures make the ignition occur more readily. A sequence of ignition events can be seen in Fig. 16. Combustion starts in Cells 1, 2, and 3 at 9, 8, and 8 deg BTDC, respectively. The temperature traces in Cells 1 and 3 oscillate somewhat because they are close to the liquid zone where considerable energy is required to vaporize the spray. The decreasing concentration of radicals from Cell 1 to Cell 3 suggests that the ignition chemistry becomes less significant downstream in the spray. The radical concentrations are so low that the combustion in Cell 3 is more likely to be due to the high-temperature flame chemistry. Although the high temperatures, which are apparent in the global temperature distributions presented in Fig. 15, might suggest that the ignition chemistry plays a minor role, the detailed evolution of the cell temperature and radical concentrations seen in Fig. 16 confirms that the low-temperature ignition chemistry still must be considered, especially near the liquid zone, in order to make the combustion modeling more accurate.

There are discrepancies between the measured and calculated pressure histories in the later stages of combustion. The predicted cylinder gas pressures indicate that the total energy available in the fuel has not been released. Some of this discrepancy can be attributed to inadequate numerical resolution [15]. In addition, the combustion model constant has not been optimized. However, further improvement of post-ignition combustion model is in progress, which takes account of turbulent mixing effects by making use of the turbulence kinetic energy and its dissipation rate. This modification in the combustion model leads to very much improved predictions of post-ignition combustion [27].

Summary and Conclusions

The Shell autoignition model was implemented in multidimensional calculations to predict the ignition of homogeneous mixtures in a rapid compression machine. The model was also used to simulate the ignition process of a dodecane spray in a combustion bomb under enginelike conditions. The results show that the Shell model is capable of predicting the autoignition of hydrocarbon fuels accurately.

The Shell model was also combined with different post-ignition combustion models. Experimental pressure data from a Cummins NH engine were compared to assess the model performance. Good levels of agreements with the experiments were obtained using the combined model to simulate the combustion process.

It was concluded that, for diesel ignition and combustion modeling, although a one-step kinetics model can be made to work quite well with empirical modifications on a case-by-case basis, the multistep kinetics model, which is derived from out-of-engine, low-temperature chemistry data, has a better performance for predicting the autoignition process.

In this study, the low- and high-temperature chemistry models were switched at 1000 K. The combined models properly account for the energy release during both the autoignition event and the early part of the subsequent combustion process. However, turbulent mixing effects during the later stages of combustion are needed in order to make the modeling of the entire combustion more accurate.

In order to investigate fuel effects on ignition, it may be of interest to consider even more detailed chemical kinetics models for future studies of diesel ignition since the present multistep chemistry model was found not to increase the computation time significantly.

Acknowledgments

The authors thank Dr. Z. Lian for his help in spray modeling. The authors also thank Professor Gary Borman for helpful comments. This work was supported under NASA-Lewis grant NAG 3-1087, W.T. Wintucky grant monitor. Support for computations from the Army Research Office, Cray Research, and the Caterpillar Engine Company are also appreciated.

References

1 Halstead, M., Kirsh, L., and Quinn, C., "The Autoignition of Hydrocarbon Fuels at High Temperatures and Pressures—Fitting of a Mathematical Model," *Combust. Flame*, Vol. 30, 1977, pp. 45–60.

2 Zellat, M., and Zeller, H., "Modelisation multidimensionnelle de l'autoinflammation dans un moteur Diesel—Premieres validations experimentales," Rapport IFP 55-551.

3 Theobald, M. A., and Cheng, W. K., "A Numerical Study of Diesel Ignition," ASME Paper No. 87-FE-2, 1987.

4 Zellat, M., Rolland, T., and Poplow, F., "Three Dimensional Modeling of Combustion and Soot Formation in an Indirect Injection Diesel Engine," SAE Paper No. 900254, 1990.

5 Hu, H., and Keck, J., "Autoignition of Adiabatically Compressed Combustible Gas Mixtures," SAE Paper No. 872110, 1987.

6 Blin-Simiand, N., Rigny, R., Viossat, V., Circan, S., and Sahetchian, K., "Autoignition of Hydrocarbon/Air Mixtures in a CFR Engine," *Twenty-fourth Symposium (International) on Combustion*, The Combustion Institute, 1992.

7 Amsden, A. A., O'Rourke, P. J., and Butler, T. D., "KIVA-II—A Computer Program for Chemically Reactive Flows With Sprays," Los Alamos National Labs., LA-11560-MS, 1989.

8 Igura, S., Kadota, T. and Hiroyasu, H., "Spontaneous Ignition Delay of Fuel Sprays in High Pressure Environments," *Trans. Japan Soc. Mech. Engr.*, Vol. 41, 1975, pp. 1559–1566.

9 Yan, J., and Borman, G. L., "Analysis and In-Cylinder Measurement of Particulate Radiant Emissions and Temperature in a Direct-Injection Diesel Engine," SAE Paper No. 881315, 1988.

10 Reitz, R. D., "Assessment of Wall Heat Transfer Models for Premixed-Charge Engine Combustion Computations," SAE Paper No. 910267, 1991.

11 Reitz, R. D., and Kuo, T. W., "Modeling of HC Emissions Due to Crevice Flows in Premixed-Charge Engines," SAE Paper No. 892085, 1989.

12 Reitz, R. D., "Modeling Atomization Processes in High-Pressure Vaporizing Sprays," *Atomisation and Spray Technology*, Vol. 3, 1987, pp. 309–337.

13 Abraham, J., Bracco, F. V., and Reitz, R. D., "Comparisons of Computed and Measured Premixed Charge Engine Combustion," *Combust. Flame*, Vol. 60, 1985, pp. 309–322.

14 Kong, S. C., Ayoub, N., and Reitz, R. D., "Modeling Combustion in Compression Ignition Homogeneous Charge Engines," SAE Paper No. 920512, 1992.

15 Gonzalez, M. A., Lian, Z. W., and Reitz, R. D., "Modeling Diesel Engine Spray Vaporization and Combustion," SAE Paper No. 920579, 1992.

16 Kong, S. C., "Modeling Ignition and Combustion Processes in Compression Ignited Engines," MS Thesis, Mechanical Engineering Department, University of Wisconsin—Madison, 1992.

17 Natarajan, B., and Bracco, F. V., "On Multidimensional Modeling of Auto-ignition in Spark-Ignition Engines," *Combust. Flame*, Vol. 57, 1984, pp. 179–197.

18 Schaperton, H., and Lee, W., "Multidimensional Modeling of Knocking Combustion in SI Engines," SAE Paper No. 85023, 1985.

19 Glassman, I., *Combustion*, Academic Press, 1987, pp. 73–79.

20 Benson, S. W., "The Kinetics and Thermochemistry of Chemical Oxidation With Application to Combustion and Flames," *Prog. Energy Combust. Sci.*, Vol. 7, 1981, p. 125.

21 Salooja, K. C., "The Degenerate Chain Branching Intermediate in Hydrocarbon Combustion: Some Evidence From Studies on the Isomeric Hexanes," *Combust. Flame*, Vol. 9, 1965.

22 Pitz, W. J., and Westbrook, C. K., "Chemical Kinetics of the High Pressure Oxidation of *n*-Butane and Its Relation to Engine Knock," *Combust. Flame*, Vol. 63, 1986, pp. 113–133.

23 Liu, A. B., and Reitz, R. D., "Mechanism of Air-Assisted Liquid Atomization," *Atomization and Sprays*, Vol. 3, 1993, pp. 55–75.

24 Lei, N., "A Cycle Simulation Program for the Dynamic Operation of a Single Cylinder Direct Injection Diesel Engine," MS Thesis, Mechanical Engineering Department, University of Wisconsin—Madison, 1988.

25 Bergeron, C. A., and Hallett, W. L. H., "Ignition Characteristics of Liquid Hydrocarbon Fuels as Single Droplets," *Canadian J. of Chem Engng*, Vol. 67, 1989, pp. 142–149.

26 Edwards, C. F., Siebers, D. L., and Hoskin, D. H., "A Study of the Autoignition Process of a Diesel Spray via High Speed Visualization," SAE Paper No. 920108, 1992.

27 Kong, S. C., and Reitz, R. D., "Spray Combustion Processes in Internal Combustion Engines," to be published in *Recent Advances in Spray Combustion*, K. K. Kuo, ed., AIAA series, 1994.

Progress on the Investigation of Coal-Water Slurry Fuel Combustion in a Medium-Speed Diesel Engine: Part 6—In-Cylinder Combustion Photography Studies

B. D. Hsu

D. P. Branyon

General Electric Company,
Transportation Systems,
Erie, PA 16531

In the GE 7FDL single cylinder research diesel engine, in-cylinder high-speed photographic studies were conducted on coal-water slurry (CWS) fuel combustion. Distinct flames of pilot and CWS combustion were noticed. It was proven that the coal fuel burns after piston impingement and secondary atomization. Agglomerated particles will develop when combustion conditions are not favorable. Cylinder pressure data were simultaneously recorded for each film frame. Heat release data can thus be produced for each photo study. Most of the findings of earlier combustion studies on engine performances were confirmed.

Introduction

Under the sponsorship of the U.S. Department of Energy, Morgantown Energy Technology Center, GE Transportation Systems has been conducting a proof of concept program to use coal-water slurry (CWS) fuel to power a diesel engine locomotive since 1988. As part of the program, extensive research and development effort has been devoted to the combustion studies of CWS fuels in the diesel engine. These efforts include the early combustion bomb study (Leonard and Fiske, 1986), and a series of single-cylinder engine studies on ignition (Hsu, 1988a), fuels (Hsu and Confer, 1991), and combustion (Hsu, 1988b; Hsu et al., 1989, 1992). A computer engine combustion model was developed (Wahiduzzaman et al., 1990, 1991) and CWS fuel injection spray development was studied (Seshadri et al., 1992). It was also planned to complete this series of studies with high-speed in-cylinder combustion photography to better understand and improve the engine performance.

Description of Test

Experimental Hardware, Data Reduction, and CWS Fuel Used. The basic test engine and data reduction processes have been described in detail before (Hsu and Confer, 1991). Briefly, the research single-cylinder engine has a 229 mm bore, a 267 mm stroke, and a rated speed of 1050 rpm. The combustion chamber arrangement has a side-mounted pilot diesel fuel injector and a centrally placed CWS injector as shown in Fig. 1. The pilot fuel injection system and the main CWS FIE are both electronically controlled. Their injection timings and quantities can be arbitrarily varied by the test engineer during engine operation.

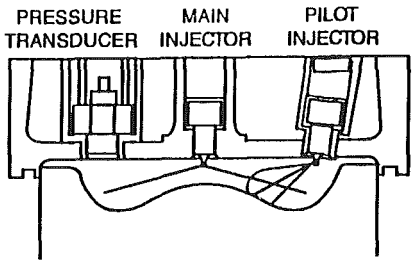


Fig. 1 Test engine combustion chamber layout

The AVL High Speed Combustion Photography System, Type 510 (Werlberger and Cartelliere, 1987) is used together with the general instrumentation at GE as described earlier by Hsu (1988a). Color photos are taken at 6000 frames per second (3000 fps half frame) by a Red Lake HYCAM Camera through a system of optical linkage and a wide angle endoscope probe developed by AVL. The AVL system supplies the beginning and subsequent crank angle markings to both the film and the high-speed data acquisition instruments. Consequently, the four to five cycles of combustion processes captured on the film can precisely correspond to the cylinder pressure and fuel injection information gathered on the high-speed data acquisition system. One cycle is selected from each run to compare the photographs and the heat release information calculated from the corresponding cylinder pressure data.

Two fuels (5 and 8 μm mean particle size) were studied. Both fuels were prepared by Otisca Industries, Inc. of New York using Kentucky Blue Gem coal. They were cleaned to

Contributed by the Internal Combustion Engine Division and presented at the 16th Annual Energy-Sources Technology Conference and Exhibition, Houston, Texas, January 31–February 4, 1993. Manuscript received by the Internal Combustion Engine Division August 1, 1992. Associate Technical Editor: J. A. Caton.

Table 1 Coal nominal lot analysis

Proximate Analysis		Ultimate Analysis	
% Ash	0.80	% Carbon	82.59
% Volatile	39.40	% Hydrogen	5.34
% Fixed Carbon	59.80	% Nitrogen	2.08
		% Chlorine	0.18
Particle Size		% Sulfur	1.01
Mass Mean Diameter (microns)	5 & 8	% Oxygen (diff.)	8.00
Heating Value	High Heating Value (kJ/kg)		34630

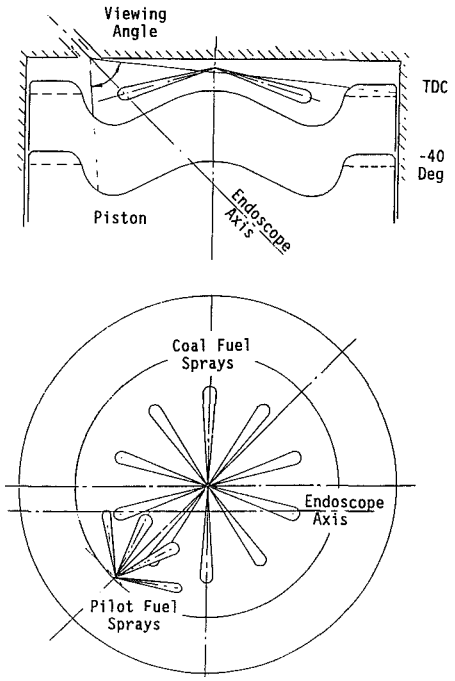


Fig. 2. Schematic of endoscope mounting in the combustion chamber

0.8 percent ash as shown in Table 1. The solid loadings of the slurry used were maintained at about 49 percent by weight.

Photo View Orientation. All the photos make use of the AVL 0 deg endoscope with 80 deg included angle lens. Figure 2 shows the schematic of endoscope mounting in the combustion chamber. Careful consistency of the optical linkage arrangement assures that a fixed view orientation of the sprays and the combustion chamber is obtained for all the photos. Figure 3 is the resultant view. Two of the pilot sprays would come into view from the lower left corner toward the center. Coal fuel sprays come in from the top. Depending on the position of the piston, the top part of the view might be one side of the piston bowl or one side of the cylinder liner (see Fig. 2). The bottom of the view always corresponds to one side of the piston bowl. This is confirmed both by statically using the fiber optic light source verification before taking pictures and by taking a flash photo of pilot fuel injection sprays before combustion with the engine running.

Cases Studied. Two cases using the 8 μm fuel and four cases using the 5 μm fuel were photographed. They are all run at the full engine speed of 1050 rpm and a fixed CWS high-pressure replenishing pump rack of 26 mm. In most cases, this corresponds to 2400–2600 mm^3/stroke of CWS fuel delivered at about 83 MPa start of injection pressure. When operating close to optimized efficiency, the engine will develop the normal full load of 2.0 MPa IMEP with this fuel rate. In the case

Table 2 Matrix of test cases

Ref Run #	Coal Particle Size Micron	CWS Injector Hole mm	Pilot Fuel Injection Start BTDC	Pilot Fuel Injection Duration CAD	Coal Fuel Injection Start BTDC	Coal Fuel Injection Duration CAD	Coal Fuel Injection Press MPa
4	-	-	12	7	-	-	-
5	-	-	26	7	-	-	-
6	8	0.46	26	6	28	37	89
8	8	0.41	26	7	29	55	83
9	5	0.46	10	7	27	36	83
10	5	0.46	18	7	25	36	81
11	5	0.46	9	6	41	35	81
12	5	0.46	9	7	26	51	71

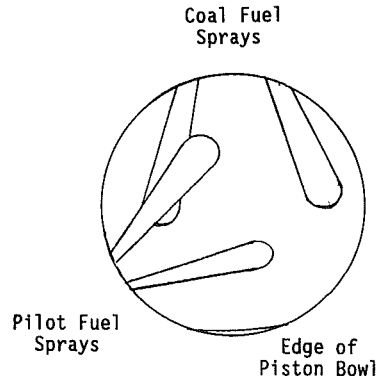


Fig. 3. Orientation of the endoscope view

of low injection pressure (70 MPa) the fuel rate increased to 2978 mm^3/stroke.

Pilot fuel is maintained at about 90 mm^3/stroke, which is 6.5–7.0 percent of total energy. Basically two pilot injection timings are used. They are a late timing of 9–12 crank angle deg (CAD) BTDC and an early timing of 25–26 CAD BTDC. In order to identify the pilot combustion in the photos, two runs were made with only pilot fuel injection using the two timings.

All runs made use of 10 hole CWS fuel injector nozzles. For the 8 μm fuel, two nozzle hole diameters (0.41 and 0.46 mm) were used. The test setup matrix is listed in Table 2.

The #6 run corresponds to the optimized injection timings for the 8 μm fuel using the ten-hole, 0.46-mm-dia hole size (10 $\times$ 0.46 mm) injector nozzle. It yielded an indicated specific fuel consumption (ISFC) of 8617 kJ/kWh. The combustion efficiency (CE) or the percent of fuel burned cannot be measured during a photography test due to the short running time. The previous test (Hsu et al., 1992) generally showed a 99.0 percent CE in this case. The #8 run was made to compare the combustion using different injector hole sizes. Due to the prolonged coal fuel injection duration of the smaller 0.41 mm holes, the cycle efficiency becomes lower showing an ISFC of 10,000 kJ/kWh. The CE in this case is about 98.0 percent. Optimized injection timing performance with 5 μm fuel is demonstrated in run #9. This case displays the desired "Delayed Ignition" in which the CWS fuel enjoys a long in-cylinder residence time before ignition (Hsu et al., 1992). The CE for this type of combustion can reach 99.5 percent and the ISFC for this case is the lowest of all cases at 7912 kJ/kWh. Run #10 shows the combustion characteristics of a case of shorter CWS in-cylinder residence time relative to run #9. This is done by advancing the pilot and therefore the ignition time of the CWS. It yields an ISFC of 8873 kJ/kWh and the CE is around 99 percent. The #11 run investigates the detrimental effect of

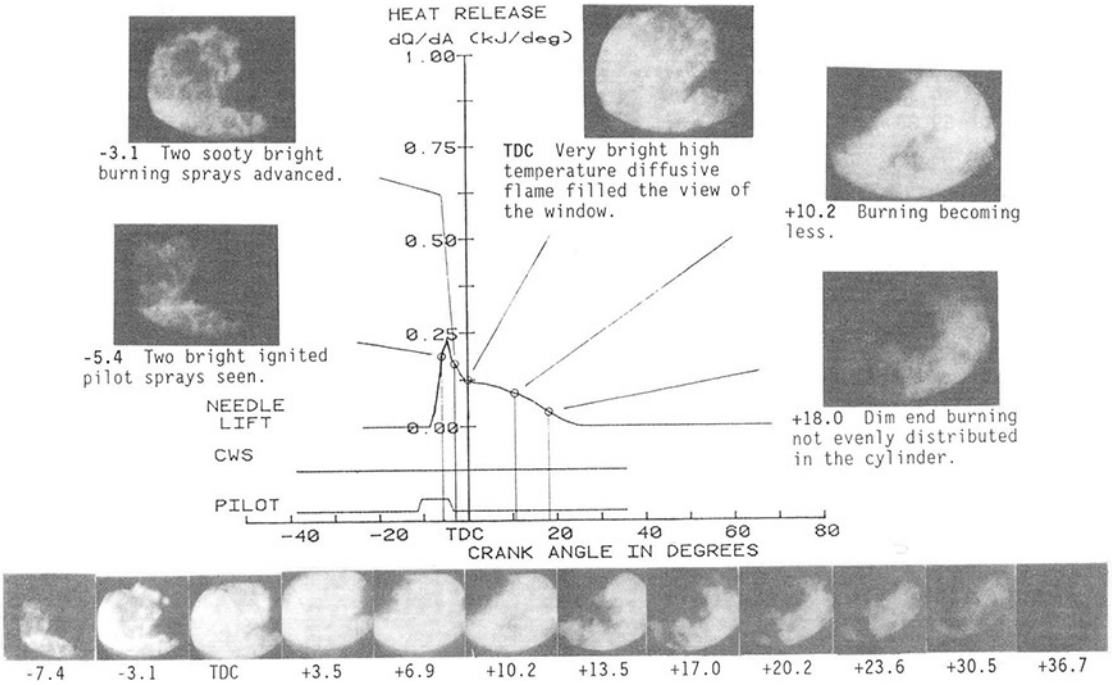


Fig. 4 Run #4, pilot fuel only; late injection timing (−12.0)

CWS fuel impinging on the piston liner (Hsu et al., 1992). Impingement is realized by using excessive coal fuel injection timing advance of 41 CAD BTDC. In this case the ISFC is 9123 kJ/kWh and CE drops to approximately 96 percent. The last run (#12) displays the low CWS injection pressure case. The injection duration is lengthened to obtain a low injection pressure of approximate 70 MPa in contrast to the over 80 MPa injection pressure in all the other cases. An ISFC of 9762 kJ/kWh is observed for this run and the reference CE is about 97 percent.

Presentation of Results

Pilot Fuel Combustion Only. Figure 4 is the photo and combustion heat release result of the late pilot timing case of run #4. In the lower portion of the figure, a selected sequence of the high-speed photos from one engine cycle is shown to describe the in-cylinder combustion process. Using the cylinder pressure information collected on the high-speed data acquisition system for this particular cycle, the heat release results were reduced and are shown on the top part of the figure. Several enlarged pictures are placed around the heat release curve. This type of display will be used for all subsequent cases. The sequence of photos shows a typical diesel fuel combustion scene. Information on the high-speed data acquisition system shows the injection needle starting to lift at −12 CAD. Distinct pressure rise is noticeable at about −7 CAD, indicating the start of apparent heat release. Weak illumination is detected on the photo that corresponds to −7.4 CAD. Very sooty combustion flames are seen to be advancing in the spray envelopes during this early period. The heat release information indicates the main kinetic burning period with high rate of heat release ends at about TDC. Due to most of the fuel being burned in the premixed state during this period, the flame is not the brightest. The flames become very bright as the combustion enters into the diffusion stage after TDC until about +10 to +12 CAD. This is an indication of very high local temperature. The rest of the sequence is the end burning period with low luminosity. Since there is no coal fuel injection in this case, the pictures also confirm the pilot spray jet orientation.

Figure 5 is the early pilot fuel injection case (26 CAD BTDC) of run #5. Due to the lower cylinder gas temperature at this early timing, the ignition delay becomes much longer. Instead of the 4–5 CAD delay of the previous case, it is now 7–8 CAD. Apparent heat release is not observed until piston reaches −18.5 CAD. A low level of luminosity is detected at about the same time. A heat release rate much higher than the previous case immediately follows as a consequence of more injected diesel fuel being evaporated and mixed with combustion air during the longer ignition delay period. Again, because of the mostly premixed burning at this point, luminosity was not very high. The combustion spray envelope boundary is not as clear and the flame does not appear sooty. After about −10.7 CAD, some high-temperature diffusive flames finally complete the combustion process. In this case, none of the photos is as bright as the brightest picture in the previous case.

8 μm Coal Fuel Combustion. Figure 6 shows the case (run #6) of optimum operation for the 8 μm fuel under full-load conditions. The early type of pilot fuel ignition (Run #5) is seen at −20.0 CAD when the heat release becomes apparent. Shortly thereafter at −17.5 CAD on the fringes of the sooty pilot fuel combustion, dim coal fuel combustion can be detected. After about −10.0 CAD, full-blown coal fuel combustion starts to fill the viewing window. However, the luminosity is not as high as the previous two pilot fuel only cases, which indicates a lower local flame temperature. This is probably due to the water content in CWS fuel. The observed flame is now neither contained in the pilot nor in the coal fuel spray envelopes. Upon viewing the moving pictures of this period, burning fine particles are seen moving from the bottom of the view upward. According to the view orientation of Fig. 3, the flame is now moving from the rim of the piston bowl toward the center. Using data from a separate pressure vessel study (Caton et al., 1993) of CWS spray development, under the present conditions, the CWS fuel should start to impinge onto the piston bowl at −22.0 CAD, which is before the pilot fuel ignition. Therefore, the earlier conclusion of CWS fuel burning after secondary atomization from bouncing off the combustion chamber walls (Hsu et al., 1992) is actually observed here. This is further confirmed by the previously men-

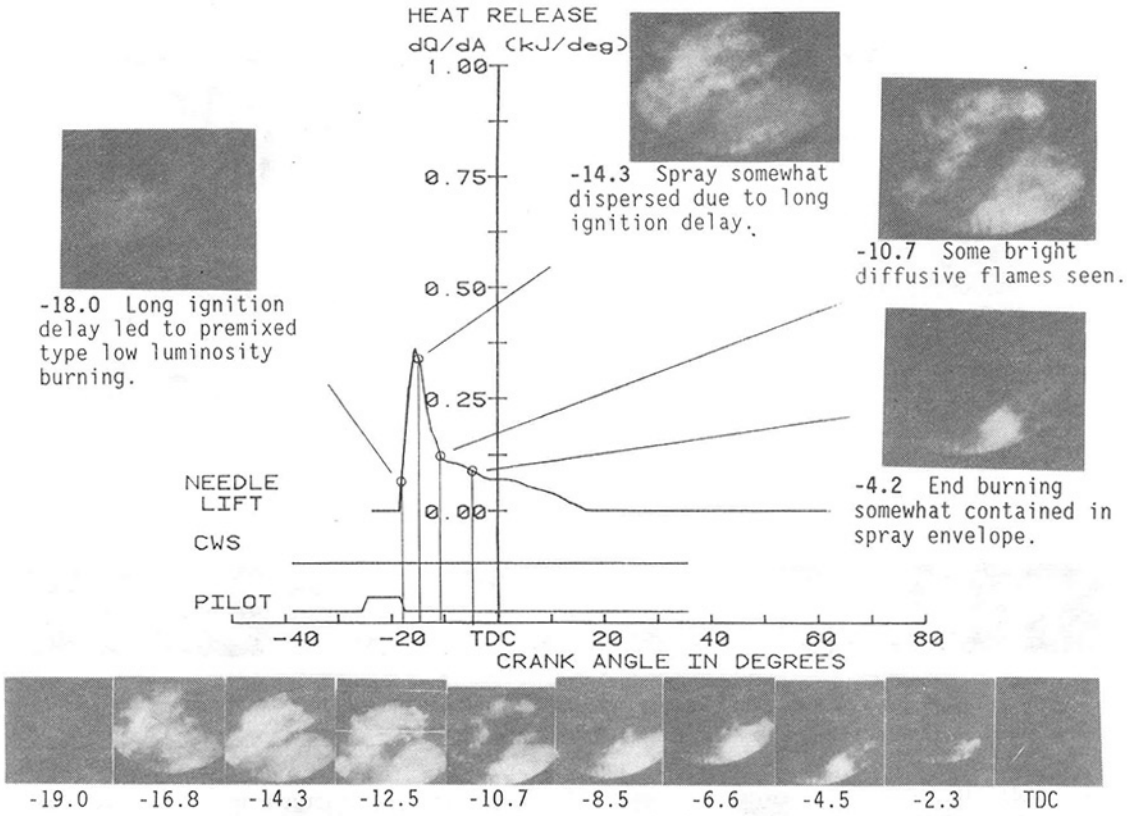


Fig. 5 Run #5, pilot fuel only; early injection timing (−26.0)

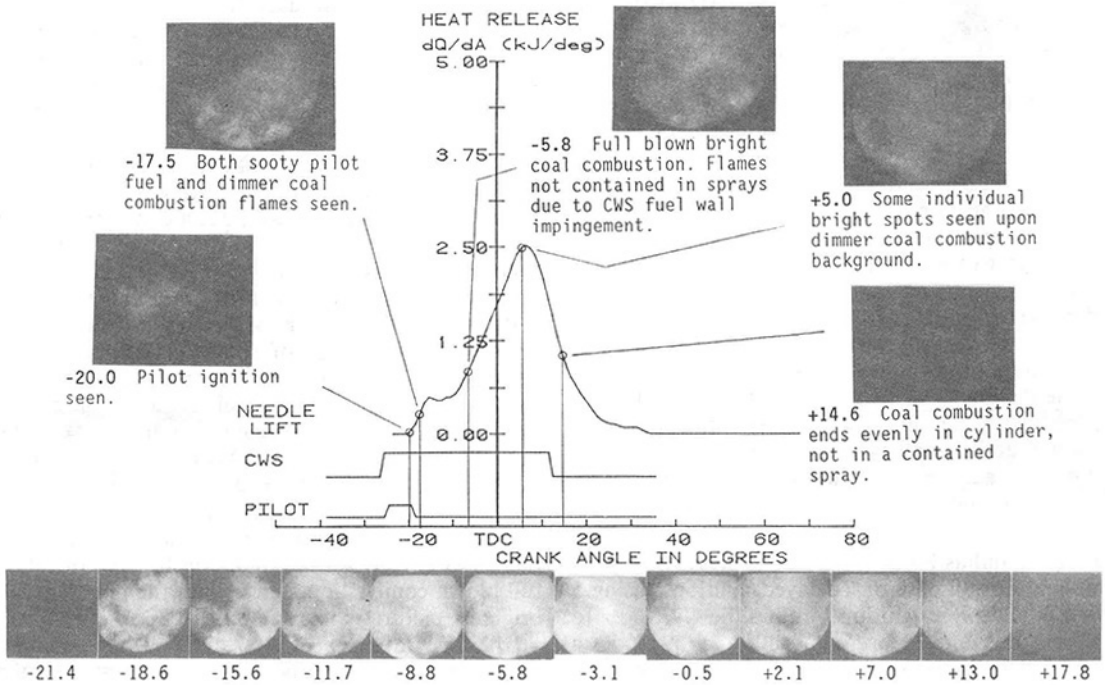


Fig. 6 Run #6, 8 μm fuel, optimum injection timings (−26 pilot, −28 CWS), 10 * 0.46 mm injector

tioned upward motion of the flame and burning particles. As the combustion ends in the cylinder, the brightness of the view dims rather evenly, which is different from the previous diesel fuel cases. This can again be attributed to the CWS fuel being more evenly distributed in the cylinder after impingement and not bounded in the original spray envelopes.

Figure 7 is the case (run #8) using the same 8 μm particle size fuel, but the CWS fuel injection rate is reduced by using a smaller diameter hole size nozzle (0.41 mm instead of the standard 0.46 mm). Due to less CWS fuel present in the cylinder (therefore less water quenching effect) at the time of pilot ignition, the pilot fuel combustion seemed to be brighter as shown in the −13.7 CAD picture on the upper part of the figure. Again, at this time, dimmer coal combustion is occurring outside of the pilot flame. Full-blown coal combustion is well developed after −5.2 CAD. The same bottom-to-top

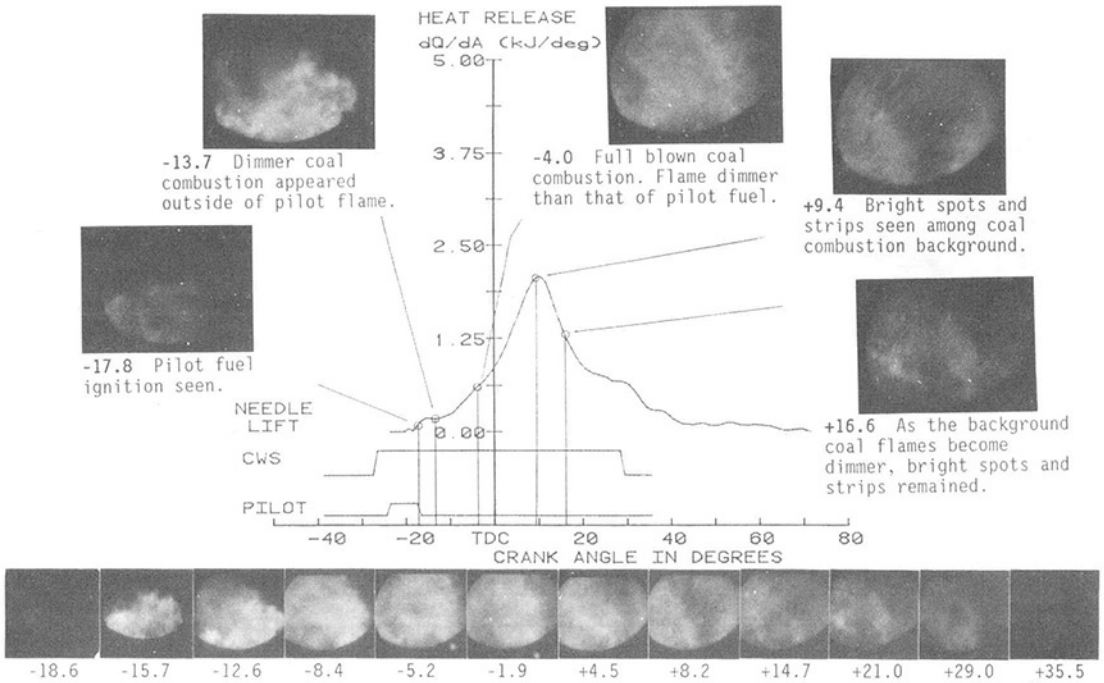


Fig. 7 Run #8, 8 μm fuel, −25 pilot and −29 CWS injection timings, 10 * 0.41 mm injector

movement of flame for this period is again evident when viewed with a movie scanner. After +6.0 CAD, in almost all the pictures, bright spots and strips are seen among the coal combustion background as shown in the +9.4 CAD picture on the upper part of the figure. This is believed to be the surface oxidation of char particles. Since individual particle are much too small to be identified, they must be the agglomerated particle after CWS fuel dehydration. The strips would be such particles moving across the lens in between camera shutter closings. These spots and strips can also be seen in some of the pictures of the previous run. However, they are not as evident. The particle agglomeration probably could be attributed to the lower maximum cylinder average temperature of 1646 K in the present case. For the previous run it was 1915 K. From +4.5 CAD on, as the coal combustion becomes dimmer, a dark streak appears from the top center extending to the bottom right of the view window. The CWS fuel needle lift trace on the diagram shows fuel injection lasted well into +30 CAD. The diagonal dark streak may be the cool CWS spray being freshly injected into the cylinder and heading for the combustion chamber walls. The long injection duration and the coal particle agglomeration are the main contributing factors of lower cycle and combustion efficiencies in this case.

5 μm Coal Fuel Combustion. Figure 8 displays results from the optimum combustion case of "delayed ignition" using 5 μm fuel (run #9). "Delayed ignition" allows the CWS fuel to have a long residence time before ignition, which provides a large amount of dehydrated coal available for burning once ignition starts. In this case, CWS is injected at −27 CAD and dehydrated by the cylinder air temperature developed by piston compression. Pilot fuel injection is delayed until very late in the cycle (−10 CAD). Combustion in the engine cylinder is actually started by coal fuel self-ignition as seen in the −7.6 CAD picture in the upper part of Fig. 8. Ignition of coal begins with the combustion of the volatiles in coal, which burn in the premixed state with low luminosity. The pilot fuel is ignited with short ignition delay and shows the initial sooty type flame of run #4 (late pilot injection only). Within a few CAD, the coal generates into full blown combustion as in the +3.0 CAD picture. Due to the large amount of dehydrated coal participating in combustion at this time, the luminosity of the picture is the highest among all coal runs. It is interesting to note that in all coal combustion cases, the brightest pictures always precedes the highest heat release rate. No evidence of coal particle agglomeration was seen in this run (cylinder temperature 1954 K). As the combustion becomes dimmer, CWS fuel spray is again visible in the pictures (+7.6 and on). In this case fuel impingement on the piston crown began at −21.0 CAD, which is well before combustion start. Nevertheless suitable combustion was still attained. The upward movement of the flames is detected in this as well as the rest of the coal fuel cases. For this run, combustion duration is the shortest and the most concentrated. As a result, the highest cycle and combustion efficiencies among all coal runs are obtained.

Run #10 shown in Fig. 9 was made to investigate the difference in combustion of shorter CWS fuel residence time before ignition. To achieve this, pilot fuel is injected earlier (−18.0 CAD) to ignite the coal fuel available at that time. In this case, the first sign of combustion is of the pilot burning type as shown in the −10.0 CAD picture. Due to less residence time for the CWS fuel to be dehydrated, coal is not ignited at the same time as the pilot. Coal combustion is detected later at about −6.7 CAD, at which time the brightest pilot type flame has already started to diminish. Coal fuel develops into full blown combustion slower than the previous case and the brightest picture is less luminous. Upon careful observation of the later (after +6 CAD) dimmer view, some barely detectable bright spots of burning agglomerated particles can be seen. However, this is much less evident than with the 8 μm fuel. At the latter stages of combustion (from +6 to +8 CAD on), burning of the wall-attached fuel is seen off the piston bowl. This became the only sign of burning at the end of the combustion period. In all the combustion movies taken, from time to time this kind of significant wall-attached fuel burning appeared. But it is not consistent, and does not contribute to any peak cylinder firing pressure change.

Figure 10 (run #11) shows the case of trying to increase CWS fuel residence time by excessive early CWS injection timing. Comparing to the case in Fig. 8 (run #9), the CWS fuel injection

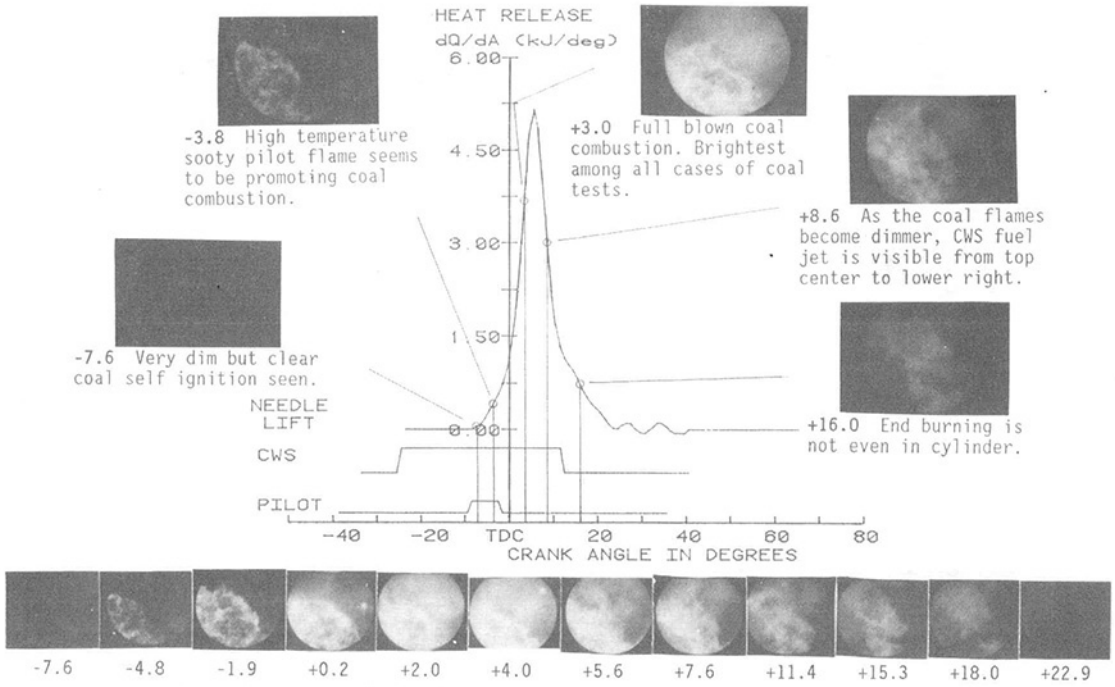


Fig. 8 Run #9, 5 μm fuel, optimum "delayed ignition" timings (−10 pilot, −27 CWS)

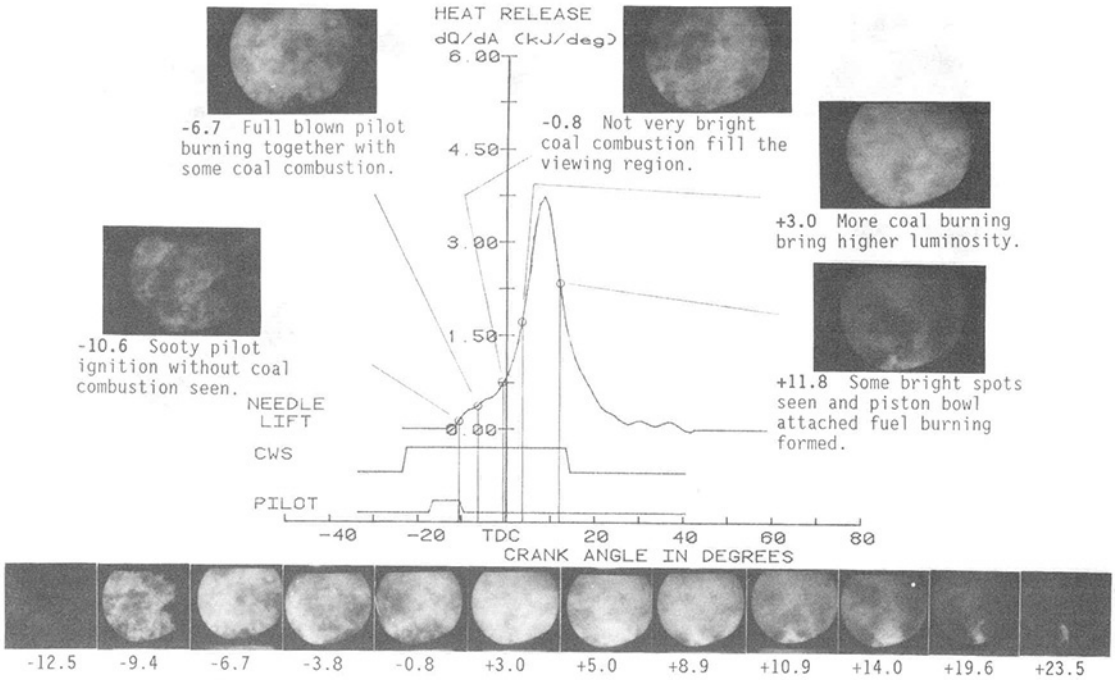


Fig. 9 Run #10, 5 μm fuel, early pilot ignition timing (−18 pilot, −25 CWS)

timing is advanced from −27 to −41 CAD. The pilot timing is about the same at −9 to −10 CAD. Data from the spray tip penetration study (Caton et al., 1993) show that, at this CWS injection timing, a portion of CWS fuel will arrive at the cylinder liner at −35 CAD before the piston can move up closer to the injector. Due to the cooler cylinder temperature caused by more CWS injected at the early timing, coal fuel self-ignition is not detected until much later at −3.5 CAD (compared to the −7.6 CAD in Fig. 8). Pilot fuel combustion is not seen until +1.2 CAD. This corresponds to approximately 10 deg CAD or 1.6 ms pilot fuel ignition delay time. In the previous Fig. 8 case, pilot fuel ignition was detected at −5.8 CAD, which corresponds to 4.2 CAD or 0.7 ms ignition delay. The difference in cylinder temperature during the ignition delay period is the main reason for the change in pilot fuel ignition delay. It is 830 K for the present case, and 950 K for the previous case. Following the ignition of pilot fuel, the available coal fuel quickly develops into full blown combustion as seen around +6.1 CAD. But the overall luminosity is not high due to the limited amount of dehydrated fuel present. Signs of particle agglomeration exist but are again not very evident. The end burning of this run does not show the general overall dimming in the entire view window. The bright patterns are results of the burning fuel that is attached to the combustion chamber walls (both the piston crown and cylinder liner). Both the cycle and combustion efficiencies are compromised by the fuel that was attached to the cooler cylinder liner walls.

The purpose of run #12 in Fig. 11 is to investigate the effect

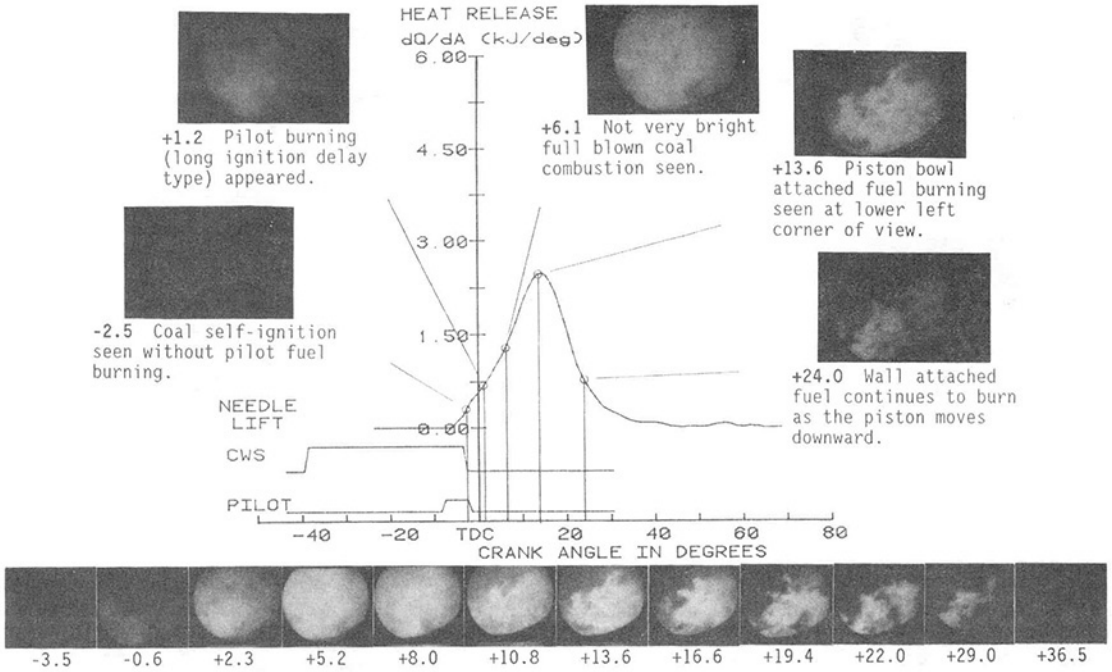


Fig. 10 Run #11, 5 μm fuel, CWS cylinder liner impingement timings (−9 pilot, −41 CWS)

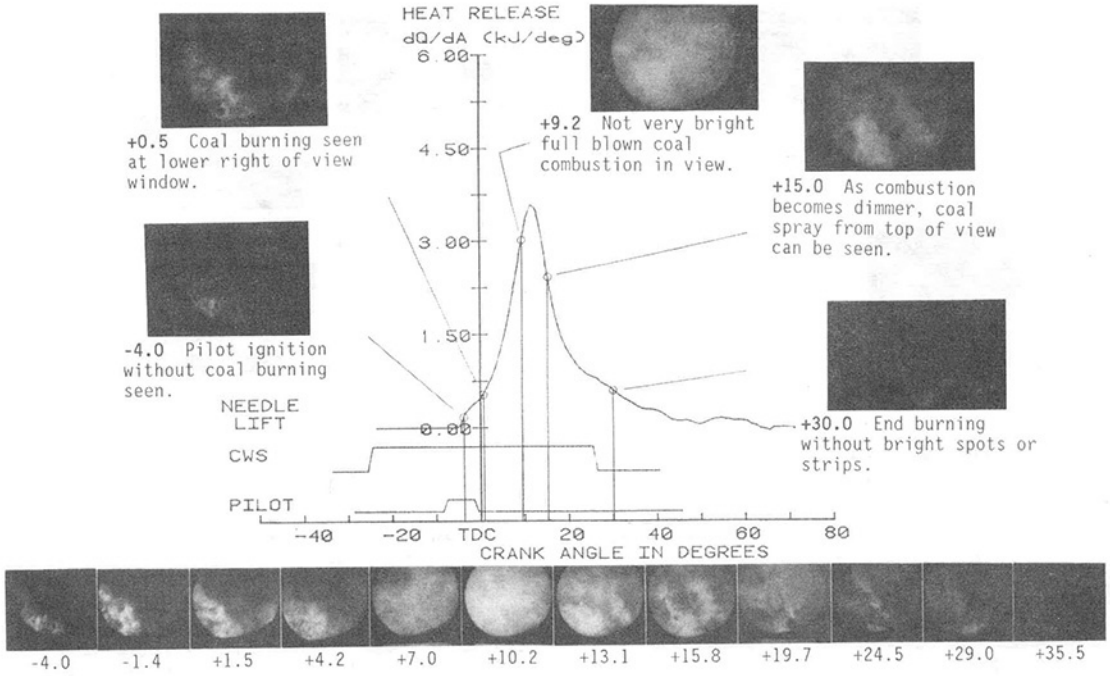


Fig. 11 Run #12, 5 μm fuel, low CWS injection pressure (70 MPa), −9 pilot and −26 CWS injection timings

of CWS fuel injection pressure. The injection starting timings of both the pilot and CWS fuel are essentially the same as the #9 run (Fig. 8). However, the CWS fuel injection duration is lengthened to reduce the injection pressure from 83 to 70 MPa. Unlike the higher injection pressure case, the first sign of burning is due to the pilot fuel at about −4.0 CAD with no clear burning of the coal fuel. Lower injection pressure and lower piston crown temperature are evaporating less water from the injected CWS fuel, thus making less dehydrated coal available for ignition. However, very soon after pilot ignition, clear signs of coal combustion co-exist together with the sooty pilot fuel flames. This is seen in the +0.5 CAD picture. As the coal fuel develops into full blown combustion as in the +9.2 CAD picture, it is not as bright as in the run #9 case. As the combustion luminosity becomes lower, coal fuel sprays can be seen (+15.0 CAD picture). The end of combustion in this case is again characterized by burning the wall attached coal fuel. Agglomerated particle burning is not very evident in this case.

Conclusions

High-speed photography of CWS fuel combustion in an actual diesel engine has been completed for the first time. Some of the results can be summarized as follows.

- The main findings of the earlier CWS fueled diesel engine combustion study (Hsu et al., 1992) have been confirmed. They include:

- Coal fuel is burned in the engine after wall impingement and secondary atomization of the coal fuel spray hitting the combustion chamber walls.
- Impingement of the CWS fuel onto the piston crown does not prohibit good combustion characteristics. Impingement onto the cylinder liner prolongs the combustion process.
- Fully utilizing the residence time of CWS fuel in the cylinder to evaporate the water content by "delayed ignition" brings intense and concentrated combustion. In this case, combustion can actually be initiated by the self ignition of CWS fuel after dehydration.

- Coal fuel burning in large agglomerated particle forms is usually associated with lower quality combustion. An optimum condition reduces the tendency toward this phenomenon.
- Large agglomerated particle burning is very pronounced for the 8 μm fuel case when the average cylinder temperature is low. It is not as evident with the 5 μm fuel. Future study to investigate the spray droplet diameter's dependence on coal particle size is recommended.
- Many times, the combustion of coal fuel ends with burning the wall-attached fuel off the piston bowl. However, no detrimental effect was seen.
- Within the range tested (under 89 MPa), higher injection pressure provides more dehydrated coal fuel at ignition time.

Acknowledgments

The support and permission to publish the results of this study from GE Transportation Systems and DOE Morgantown Energy Technology Center are gratefully acknowledged. The work on which this paper is based was conducted under DOE Contract No. DE-AC21-88MC23174. The support of the program from the other sponsors, Norfolk Southern Corporation, Pennsylvania State Energy Development Authority, and New York State Energy Research and Development Authority is also gratefully acknowledged. The authors would like to thank their many colleagues at GE Transportation Systems for their valuable contributions. The authors would also like to thank Mr. W. Cary Smith, Contracting Officer Technical Representative of DOE/METC, for his comments and encouragement in the course of this study.

References

Caton, J. A., Payne, S. E., Teraccina, D., and Kihm, K. D., 1993, "Coal Water Slurry Spray Characteristics of an Electronically Controlled Accumulator Fuel Injection System," presented at the Internal Combustion Engine Symposium of ASME ETCE, Houston, TX, Jan., 1993.

Hsu, B. D., 1988a, "Progress on the Investigation of Coal-Water Slurry Fuel in a Medium Speed Diesel Engine: Part 1—Ignition Studies," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 110, pp. 415–422.

Hsu, B. D., 1988b, "Progress on the Investigation of Coal-Water Slurry Fuel in a Medium Speed Diesel Engine: Part 2—Preliminary Full Load Test," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 110, pp. 423–430.

Hsu, B. D., Leonard, G. L., and Johnson, R. N., 1989, "Progress on the Investigation of Coal-Water Slurry Fuel in a Medium Speed Diesel Engine: Part 3—Accumulator Injector Performance," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 111, pp. 516–520.

Hsu, B. D., and Confer, G. L., 1991, "Progress on the Investigation of Coal-Water Slurry Fuel Combustion in a Medium-Speed Diesel Engine: Part 4—Fuels Effect," *Coal Fueled Diesel Engines*, ASME ICE-Vol. 14.

Hsu, B. D., Confer, G. L., and Shen, Z. J., 1992, "Progress on the Investigation of Coal-Water Slurry Fuel in a Medium Speed Diesel Engine: Part 5—Combustion Studies," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 114, pp. 515–521.

Leonard, G. L., and Fiske, G. H., 1986, "Combustion Characteristics of Coal/Water Mixtures in a Simulated Medium-Speed Diesel Engine Environment," ASME Paper No. 86-ICE-15.

Seshadri, A. K., Caton, J. A., and Kihm, K. D., 1992, "Coal Water Slurry Atomization Characteristics of a Positive Displacement Fuel Injection System," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 114, pp. 528–533.

Wahiduzzaman, S., Blumberg, P. N., Keribar, R. and Rackmil, C. I., 1990, "A Comprehensive Model for Pilot Ignited, Coal Water Mixture Combustion in a Direct Injection Diesel Engine," ASME JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER, Vol. 112, pp. 384–390.

Wahiduzzaman, S., Blumberg, P. N. and Hsu, B. D., 1991, "Simulation of Significant Design and Operating Characteristics of a Coal Fueled Locomotive Diesel Engine," *Coal Fueled Diesel Engines*, ASME ICE-Vol. 14.

Werlberger, P., and Cartelliere, W. P., 1987, "Fuel Injection and Combustion Phenomena in a High Speed DI Diesel Engine Observed by Means of Endoscopic High Speed Photography," S.A.E. Paper No. 870097.

Managing ISO 9000 Implementation

W. J. Lupear
Terrecorp, Inc.,
Clinton, IN 47842

Many companies worldwide are implementing and applying for registration for the ISO 9000 series of quality standards. This implementation has the same effect on an organization as any other process that is changing the methods, procedures, or the basic culture of the company. The management style necessary to make these changes successful is the same as that which is necessary for any other transformational process. If the correct management style is present, ISO 9000 implementation, and any other change being attempted, will be successful. Change can be accomplished without transformational management but the overall effect and the cost benefit will not be maximized. This transformational management is the most important element. The technical aspects (the ISO 9000 hows and whys) are just a matter of someone within the organization learning them and developing a plan and system for management. The correct management style to achieve change is the most difficult. Strategies need to be clearly defined and methods developed in order to control the projects. Roadblocks need to be clearly identified and action planned to compensate for any shortcomings.

Introduction

In 1979 the ISO Technical Committee 176 was charted with the task of coordinating the international activities of quality systems. The objective was to insure good control practices and procedures for the mutual benefit of the supplier and customer. The result was the publishing of the ISO 9000 series of quality standards in 1987. These quality standards were first adopted by the European Community but have gained widespread acceptance worldwide (51 countries). As these standards gain acceptability, more and more US companies are finding themselves in the position of changing and/or improving their quality system to comply with the ISO standards, especially if they deal in world markets. More customers are requiring that their supplies have ISO 9000 series certification.

The actual activities associated with developing and implementing an ISO 9000 system have been well documented in numerous books and magazines. The scope of this paper is aimed at the management methods, strategies, and implementation procedures that I've experienced in the ISO chase.

Management Method

Management must use a method or style to insure change is possible. The type of transformational management proposed by Kim Cameron in his book *Developing Management Skills* (1991) or by J. Juran in his book *Managerial Breakthrough* (1964) is the type of management style to which I am referring. This change process must be planned, developed, and then implemented. Most important of all, this change must have the commitment of all involved, including top management. No matter how skilled and knowledgeable an organization is on the subject of the ISO 9000 series of standards, it won't accept and achieve the results expected unless the proper management style and top management support exist.

Three criteria must exist before any implementation is started; Top management must be committed and be involved in all phases of the process, a system to allow for change must exist within the organization, and there must be a total involvement of all the people in the process of change and their needs addressed.

Management commitment will show all that the change is important and necessary for the success of the company, and warrants the resources applied to the change process. Management must not only charter the ISO implementation, but must also be involved in the day-to-day activities.

The system that will allow change needs to be tailored to each organizational culture. There are general guidelines for this change process; however, the specific recipe will have to be developed by someone with the expertise to clearly identify the culture and what methods are needed. Many times this individual will be someone from outside the company.

The employees must be a part of this process and need to be involved in all aspects of this change process. The people must be treated as associates, not just employees. Their input should be sought when developing the planning and linkage to everyday activities. They must have a thorough understanding of what the requirements of the ISO 9000 standards are, and how their specific responsibilities will be affected. It is management's responsibility to supply whatever is necessary to achieve this understanding. In addition to normal supervision, this could take the form of employee participation groups, formal or informal training, published procedures, published policies, corporate mission statements, and oper-

Contributed by the Internal Combustion Engine Division and presented at the 14th Annual Fall Technical Conference, Waterloo, Iowa, October 4–7, 1992. Manuscript received by the Internal Combustion Engine Division July 1, 1992. Associate Technical Editor: J. A. Caton.

ating principles. It is establishing the "state of self control" for all employees. This concept explained in *Managerial Breakthrough* (1964) and in the *Quality Control Handbook* (1974) by J. M. Juran is: know what the objectives are, know whether they are being met, and have the ability and authority to regulate performance in order to achieve the objectives. The ability and authority to regulate is a very important part of this concept. Many times the authority is not truly given by management. This concept is basic but it applies to ISO implementation, establishing good working practices, Statistic Process Control, or any change or new addition to an organization.

Strategies

A company involved in implementing a change such as ISO 9000 must develop a strategy for structuring this change, and monitoring and controlling the activities. The seven-step process for formulating fundamental transformation to an ISO 9000 system should be followed. These steps are: recognize the need for change, initiating an internal study to assess organizational needs, decide to pursue the change, formulate a holistic quality policy adapting the ISO requirements, commit to support the policy by all personnel, plan and schedule how to implement the system, and initiate implementation. How these individual steps are accomplished must be determined by each organization; however, I would like to suggest that existing project planning techniques can be used to accomplish these tasks.

Task Gantt, Resource Gantt, and PERT charts are just a few of the possible planning and control devices that can be used. A color-coded project status board might be considered for an organization that requires constant monitoring and updating. If this status board is prominently displayed, it has the added advantage of letting top management see whether all of the tasks are on schedule. Special reporting will not be necessary if top management intervention is necessary.

As complex and mutifaceted as ISO 9000 implementation is, some type of monitoring system must be used to insure all aspects of the standards have been addressed. Many companies have based their supplier audits on the ISO standards. These audit forms are ideal to use as a guide in doing a self-audit in preparation for registration or a customer audit.

If an in-house expert on ISO 9000 does not exist, the company can hire a consultant to help guide and steer the company. This is especially important through the planning phase. In addition, the consultant can bear the brunt of any adverse reaction and take the heat if the implementation is not successful.

Another alternative is to develop an internal expert through seminars and ISO specific training. This is time consuming and will take a project manager type integrator to be successful. If this path is chosen, several very hard questions need to be answered, as the candidates for this type of responsibility are hard to find. If the company is progressive and is involved with many R&D projects, this individual can be very valuable. In the present day competitive markets the need for highly specialized functions and tight integration between disciplines is critical. This integrator can fill that need. This person would also be ideal in the facilitation of employee involvement groups that would be used from the planning stages through implementation of the ISO 9000 standards.

Roadblocks

After developing the proper controls and systems, the possible roadblocks to success should be addressed. There are many tools that can be used to help identify problems and plan countermeasures. I use a method similar to Failure Mode, Effects, and Criticality Analysis (FMECA) as outlined in MIL-STD-1629A. This method is primarily for product design from concept through development but with modification it can be used for failures within a process/system implementation. The objective of my analysis is to identify possible failures with the implementation plan, identification of these failures as they occur, and countermeasures to eliminate or minimize them. This analysis should be performed throughout the planning and preliminary implementation. I have found that this is an ideal tool to use when involving the personnel affected by the ISO implementation. Their input is sought during the planning stage and brainstorming sessions can be used to suggest possible failure items on the FMECA list. This also has the added benefit of the personnel "buying-in" to the change.

Another tool to use is the fishbone diagram, also called cause and effect or Ishikawa diagrams. This is a very common tool used by Quality Control personnel but it can also be used to help dissect a problem and find the true cause of the problem. It is a very useful tool to be used by groups assigned the task of identifying and planning for corrective action to problems.

The choice of which methods are used to accomplish the task of controlling problems encountered in the ISO 9000 implementation process is a matter of personal preference and which methods best match the company culture. It's very important, however, that this task be performed and not taken for granted. Problems will occur and they can have a negative impact on the success of the project.

Implementation

After the proper planning has been accomplished, it is a matter of following the task list with midcourse corrections. The task list should have all the tasks identified, prioritized, and a time element established for each task. Predecessor and successor tasks should be clearly identified and the sequence understood.

One of the first tasks should be the formulation of the Quality Manual. This manual must fulfill the requirements of the ISO standard; however, it should truly reflect the philosophy and policies of the company. All personnel, especially top management, should have input into the development of this manual. After all, everyone in the organization will have to live with this document. Too many times someone will write a manual that will say all the right things, but it will only be lip service if everyone hasn't bought in to the new system. At this point in the implementation process, the company should search its corporate soul to identify what it wants to accomplish and does it truly want and need the ISO certification and the system that goes with it.

The second task will be to document the operating procedures. These procedures will cover the who, what, when, and where of the daily operations throughout the organization. At this point I would suggest that the individual departments put in writing, if it does not exist, the procedures that are presently being followed. These procedures can then be compared to the ISO standards to see if any modifications are needed. These modifications can then be documented and implemented into the day-to-day activities. If everyone was involved in the early phases of this process, they will fully understand why these change are necessary. There will be very little resistance to the change.

The third task is to write the specific work instructions. This would include specific instructions for machines, gages, instruments, audit and inspection, etc. If a company was to write these instructions for every work task, it could be a never-ending process. Only those instructions directly related to the ISO standard need to be documented.

Summary

If the right steps are taken early in the planning and developing phases, the actual implementation and certification will

be comparatively smooth. Management needs to be aware and act on the necessary activities that will insure a successful implementation. If the company is rigid and reluctant to change, the ISO implementation will be a fight and the results will be wanting. The company has to not only know what the destination is, but how they are going to get there.

The ISO 9000 Standards have their strengths and weaknesses. The standard is to bring companies together under one umbrella, avoiding confusion, and building one common communication tool. Registration is becoming a necessity in some markets, and will surely strengthen a company's position in any market. As a criticism, I do not feel that the ISO 9000 Standards are enough of a quality effort for a company dealing in world markets and pursuing constant improvement. A company needs to stay competitive and constantly improve their systems, including quality. ISO is just a start toward total quality management and improvement. Many companies have excellent quality systems already in place and getting ISO certified is a fairly simple process for them. Others have weak or nonexistent quality systems and it's a major struggle for them. Unfortunately, the companies with the poor quality systems are usually the ones with the cultures that have difficulty changing. If they understood the need for good quality systems, they would understand the purpose of the ISO 9000 Standards.

Understanding the concept of the ISO 9000 Standards and quality systems cannot be overemphasized. A company and all its people must understand the necessity of world class quality in order to survive in today's markets. If they do not understand and implement, they will have a difficult time surviving.

The question of cost is always a consideration in any project or proposal. The cost of installing an ISO 9000 system will vary depending on what level of quality a company has achieved before ISO. For some companies it will be a major expenditure, for others, only the cost of the activities of the assessor performing the ISO accreditation. My question is, how can any company afford not to pursue a better quality system? ISO 9000 is a good start toward a world class system. A company can save the cost of the certification process if it's not necessary for their market; however, every company should be striving toward improvement if they want to stay competitive.

A company's quality objectives should be more ambitious than the ISO standards. The Malcolm Baldrige National Quality Award is a standard of business excellence in the United States. The evaluation criteria for this award are an excellent guide for business to follow in development of their quality systems. They address such issues as strategic quality planning, human resource development and management, information and analysis, and the very important issues of customer focus and satisfaction. I would highly recommend that any company interested in improving their system use a copy of the Baldrige Award Criteria to plan and steer their activities.

References

Cameron, K., 1991, *Developing Management Skills*, Harper-Collins, New York.

Department of Commerce, NIST, 1992, "Malcolm Baldrige National Quality Award Criteria," American Society for Quality Control, Milwaukee, WI.

Department of Defense, 1980, "Procedures for Performing a Failure Mode, Effects and Criticality Analysis," MIL-STD-1629A, Navy Publishing and Printing Service, Philadephia, PA.

Juran, J. M., 1964, *Managerial Breakthrough*, McGraw-Hill, New York.

Juran, J. M., 1974, *Quality Control Handbook*, McGraw-Hill, New York.

Lamprecht, J. L., 1991, "ISO 9000 Implementation Strategies," *Quality*, Vol. 30, No. 11, Hitchcock Publishing Co., Carol Stream, IL.

Meckstroth, D. J., 1992, "The European Community's New Approach to Regulation of Product Standards and Quality Assurance (ISO 9000): What It Means for U.S. Manufacturers," MAPI Economic Report RE-218, Manufacturers' Alliance for Productivity and Innovation, Washington, DC.

Warchol, M. H., 1992, "Understanding ISO 9000: Part 1," *Modern Casting*, Vol. 82, No. 5, American Foundrymen's Society, Des Plaines, IL.

Wheeler, S., 1992, "Guidelines for Managing Organization Change," *Manufacturing Systems*, Vol. 10, No. 6, Hitchcock Publishing Co., Carol Stream, IL.

Modeling and Identification of the Combustion Pressure Process in Internal Combustion Engines: II—Experimental Results

F. T. Connolly
Department of Mechanical Engineering,
The Ohio State University,
Columbus, OH 43210-1107

A. E. Yagle
Department of Electrical Engineering
and Computer Science,
The University of Michigan,
Ann Arbor, MI 48109-2122

In Connolly and Yagle (1992, 1993) we presented a new model relating cylinder combustion pressure to crankshaft angular velocity in an internal combustion engine, primarily the fluctuations in velocity near the cylinder firing frequency. There are three aspects to this model. First, by changing the independent variable from time to crankshaft angle, a nonlinear differential equation becomes a linear first-order differential equation. Second, a new stochastic model for combustion pressure uses the sum of a deterministic waveform and a raised-cosine window amplitude-modulated by a Bernoulli–Gaussian random sequence, parametrizing the pressure by the sample modulating sequence. This results in a state equation for the square of angular velocity sampled every combustion, with the modulating sequence as input. Third, the inverse problem of reconstructing pressure from noisy angular velocity measurements was formulated as a state-space deconvolution problem, and solved using a Kalman-filter-based deconvolution algorithm. Simulation results in Connolly and Yagle (1992, 1993) show that the parametrized pressure can be deconvolved at low to moderate noise levels, and combustion misfires detected, all in real time. This paper presents and discusses experimental results that confirm this model, at least at the relatively low-speed, low-to-moderate load operating conditions analyzed. They show that cyclic combustion pressure variation is fairly well modeled and may be directly estimated from angular velocity measurements. They also show that the deconvolution algorithm is able to detect misfires and possibly classify their severity. Since the experimental data are taken from an actual V-6 automobile engine, and the algorithms are simple enough to be implemented in real time, these results are directly applicable to real-world combustion pressure identification.

1 Introduction

Legislatively and environmentally mandated requirements concerning reduction of exhaust gas emissions in internal combustion (IC) spark-ignited (SI) engines have become increasingly stringent, e.g., the Environmental Protection Agency (EPA) emissions standards and CARB (1989, 1991). This motivates more sophisticated analysis and modeling of IC-SI engine operation and dynamics. One direct measure of engine operation and performance is combustion pressure. In particular, *cyclic variability* in combustion pressure is a fundamental limitation of engine operation as the air-to-fuel (A/F) ratio becomes more efficient, i.e., is increased from the currently normal stoichiometric ratio (Al-Alousi and Karim, 1984; Belmont et al., 1986; Janula, 1983; Matekunas, 1983; Nohira et al., 1980; Patterson, 1966; Yamamoto and Misumi, 1987). Successful measurement of cyclic variability in combustion pressure would permit tighter and better control of such variability, which would lead to improved engine performance.

In Connolly and Yagle (1992, 1993) we presented a model relating combustion pressure to angular velocity in an IC-SI engine. This model incorporates cyclic variability and has three major features: (1) reformulation of a nonlinear differential equation relating the combustion pressure (P) to engine crankshaft angular velocity (Ω) with time t as the independent variable into a linear, first-order differential equation relating P to the *square* of angular velocity Ω^2 with crank angle θ as the independent variable; (2) a new stochastic model for combustion pressure P, which includes the effect of cyclic variability; and (3) application of signal processing system identification techniques to solve the inverse problem of determining the stochastic part of the P model (cyclic variability) from noisy observations of Ω. The works of Connolly and Yagle (1992, 1993) were based on Connolly's doctoral thesis (1992) and presented computer simulation results verifying the model. Connolly and Yagle (1992) also presented a few experimental results.

Contributed by the Internal Combustion Engine Division for publication in the JOURNAL OF ENGINEERING FOR GAS TURBINES AND POWER. Manuscript received by the Internal Combustion Engine Division December 1, 1992. Associate Technical Editor: J. A. Caton.

This paper presents further experimental results from Connolly (1992). The data were collected from a working production 3.8 liter V-16 IC engine mounted in an automobile. During the experimental tests the automobile was run on a chassis dynamometer. The results confirm the model presented in Connolly and Yagle (1992, 1993). They show that the model can indeed be used to reconstruct combustion pressure from actual noisy measurements of angular velocity. They also suggest the possibility of using the algorithm to detect misfires and classify their severity. This paper should be viewed as a companion paper to Connolly and Yagle (1992, 1993).

This paper is organized as follows. Section 2 briefly reviews the engine model and pressure identification method presented in Connolly and Yagle (1992, 1993). Sections 3 and 4 discuss how the data were collected, and present experimental results for the inverse and forward problems of reconstructing combustion pressure from angular velocity, and vice versa. Section 4 also presents results on reconstruction of continuous pressure waveforms from a deconvolved cyclic pressure variability sequence, detection of misfires, and classification of misfire severity. Section 5 concludes the paper with a summary.

2 Review of Model of Connolly and Yagle (1992, 1993)

This section briefly reviews the new engine and stochastic combustion pressure models developed in Connolly (1992) and presented in Connolly and Yagle (1992, 1993). The issue of interest in these models is the relationship between variations in cylinder combustion pressure from combustion to combustion, and the resulting variations in fluctuations of crankshaft angular velocity. In the absence of cyclic pressure variability, angular velocity fluctuations are assumed to be periodic, but cyclic variability in combustion pressure will produce variations in the periodicity of the angular velocity fluctuations. These models are developed for steady-state engine behavior. The goals of these models are to: (1) model the relationship between cyclic pressure variations and angular velocity variations—the *forward problem*; and (2) solve the *inverse problem* of identifying cyclic pressure variations from noisy observations of angular velocity.

2.1 Basic Equations and Assumptions. The engine model is based on a familiar *torque balance equation* relating the major torques produced by the engine to the angular acceleration of the crankshaft. This equation is (Lichty, 1967; Obert, 1968; Heywood, 1988)

$$J\ddot{\theta} = T_i(\theta) + T_r(\theta) + T_L(\theta) \tag{1}$$

where crankshaft angle θ is a function of time, i.e., $\theta = \theta(t)$, $\ddot{\theta} = d^2\theta/dt^2$, $T_i(\theta)$ is the indicated or combustion pressure torque, $T_r(\theta)$ is the torque due to reciprocating masses, $T_L(\theta)$ is the load torque, and J is the effective moment of inertia of the engine. Equation (1) is valid given the following assumptions:

1 The crankshaft is stiff.
2 The dynamics of the engine are modeled by the single lumped moment of inertia J.
3 $T_L(\theta)$ incorporates the mean value of all torques acting on the crankshaft not shown in Eq. (1).
4 Fluctuations in other torques not shown in Eq. (1) are negligible at the frequencies of interest (viz., near the firing frequency).
5 The engine is operating in steady-state (i.e., constant "average" engine speed over several engine cycles, constant $T_L(\theta)$, fully warmed, and fixed spark timing and fuel mixture).

We now present the model of Connolly and Yagle (1992, 1993) for a four-stroke, even-firing, six-cylinder engine, although in general any number of cylinders may be used. Detailed models of the torques (Lichty, 1967; Obert, 1968; Heywood, 1988) are substituted into Eq. (1), which becomes

$$J\ddot{\theta} = \sum_{i=1}^{6} [P_i(\theta) f_1(\theta - \phi_i)] + \sum_{i=1}^{6} M_{eq}[\dot{\theta}^2 f_2(\theta - \phi_i) + \ddot{\theta} f_3(\theta - \phi_i)] + T_L(\theta) \tag{2}$$

where $\phi_i = 2\pi/3(i-1)$, $P_i(\theta)$ is pressure in the ith cylinder, and M_{eq} is the mass of the reciprocating parts associated with each cylinder (assumed to be the same for all cylinders). The functions $f_1(\theta), f_2(\theta)$, and $f_3(\theta)$ are derived from the geometry of a single cylinder (e.g., see Lichty, 1967; Obert, 1968; Heywood, 1988). Equation (2) is the nonlinear engine model relating P to Ω in the time (t) domain.

2.2 Transformation From Nonlinear to Linear Differential Equation. Transforming the independent variable from time t to crankshaft angle θ, and the dependent variable from $\Omega(t)$ to $\Omega^2(\theta)$, Eq. (2) becomes

$$\frac{d}{d\theta}[x(\theta)] - A(\theta)x(\theta) = B(\theta), \tag{3}$$

where $x(\theta) = \Omega^2(\theta)$ and

Nomenclature

$A(\theta), B(\theta)$ = functions used in forward continuous model
$\bar{b}$ = parameter of forward discrete model
f_1, f_2, f_3 = engine geometry functions of θ
J = effective engine rotational moment of inertia
M_{eq} = equivalent mass of reciprocating parts
n = combustion number index
N = number of cylinders
$P, P(\theta)$ = combustion pressure waveform
$P_n^i(\theta)$ = pressure, ith cylinder, nth engine cycle
$\Delta P_n^i(\theta)$ = pressure deviation waveform
$\bar{P}^i$ = average pressure, ith cylinder
$\bar{P}$ = mean or average combustion pressure waveform
s_n = sign function for pressure parametrization
$T_i(\theta)$ = indicated torque
$T_r(\theta)$ = reciprocating mass torque
$T_L(\theta)$ = load torque
u_1, u_2, u_3 = functions of θ used in continuous model
$v(n)$ = discrete observation noise
x = Ω^2, square of angular velocity
$y(n)$ = noisy observation of $x(n)$
γ_n = starting angle of combustion
γ_{comb} = angular duration of combustion
θ = engine angular position
μ_n^i = "unsigned" cyclic variability sequence, ith cylinder
$\xi_c(\theta)$ = continuous stochastic part of combustion pressure
$\xi_n, \xi(n)$ = cyclic variability or Bernoulli–Gaussian sequence
ξ_n^i = cyclic variability sequence, ith cylinder
ϕ_i = phase angle of firing for ith cylinder
Ω = angular velocity waveform

$$A(\theta)=\frac{u_3(\theta)}{u_1(\theta)} \quad B(\theta)=\frac{u_2(\theta)}{u_1(\theta)}$$

$$u_1(\theta)=\left[\frac{J}{2}-M_{eq}\sum_{i=1}^{3} f_3(\theta-\phi_i)\right]$$

$$u_2(\theta)=\sum_{i=1}^{6}[P_i(\theta)f_1(\theta-\phi_i)]+T_L(\theta)$$

$$u_3(\theta)=2\dot{M}_{eq}\sum_{i=1}^{3} f_2(\theta-\phi_i). \tag{4}$$

This is a linear first-order θ-varying differential equation with a known function of combustion pressure as the forcing function (see $B(\theta)$ and $u_2(\theta)$ in Eq. (4)). This Eq. (3) has a closed-form solution, which may be cascaded with a square root operator to give $\Omega(\theta)$.

2.3 Stochastic Pressure Model. The new model of combustion pressure $P(\theta)$ is

$$P(\theta)=\bar{P}(\theta)+\xi_c(\theta) \tag{5}$$

where $\bar{P}(\theta)$ is the known mean pressure waveform. The stochastic part of combustion pressure $\xi_c(\theta)$ is defined as

$$\xi_c(\theta)=\begin{cases}\xi_n\dfrac{1}{2}(1-\cos(6\theta)), & \theta\in[\gamma_n,\ \gamma_n+\gamma_{\text{comb}})\\ 0 & \theta\in[\gamma_n+\gamma_{\text{comb}},\ \gamma_{n+1})\end{cases} \tag{6}$$

where ξ_n (or $\xi(n)$) is a Bernoulli-Gaussian random sequence with Bernoulli probability of success $p = 1$ and Gaussian amplitude $\mathfrak{N}(0, \sigma^2)$, n is the combustion-to-combustion index, $\gamma_n = n4\pi/N$ is the starting crankshaft angle of the nth combustion, $\gamma_{\text{comb}} = 2\pi/N$ is the angular duration of the combustion event, and N is the number of cylinders. This model parametrizes cyclic pressure variability due to all cylinders into a compact, efficient form using the sample modulating sequence ξ_n. It is based on the work of Rizzoni (1986, 1989).

2.4 State Equation and Stochastic Deconvolution. The closed-form solution of Eq. (3) was combined with the stochastic pressure model, Eqs. (5) and (6), to form a discrete recursion relating $\xi(n)$ to samples $x(n)$ of the square of angular velocity in every combustion. This state equation was then combined with an equation for noisy observations $y(n)$ of $x(\theta)$ at $\theta = n4\pi/N$ to give a state equation model of noisy observations

$$x(n+1)=x(n)+\bar{b}\xi(n)$$
$$y(n)=x(n)+v(n), \tag{7}$$

where $\bar{b}$ is a parameter that depends on the continuous model equations and the transformation from a continuous to discrete model and $v(n)$ is white observation noise uncorrelated with $\xi(n)$. See Connolly and Yagle (1992 or 1993) for further details on the calculation of $\bar{b}$. In Connolly (1992) and Connolly and Yagle (1992, 1993) a Kalman-filter-based stochastic deconvolution algorithm (Mendel, 1977) was applied to Eq. (7). The goal is to estimate $\xi(n)$ from noisy observations $y(n)$. Computer simulation results were presented in Connolly (1992) and Connolly and Yagle (1992, 1993), and a few experimental results were presented in Connolly and Yagle (1992). In the sequel, many experimental results from Connolly (1992) will be presented.

3 Experimental Results—Forward Problem

This section presents experimental results, in which real-world data are used in the model reviewed in Section 2 to solve the forward problem of computing angular velocity from measured combustion pressure. In the next section (Section 4) experimental results for the inverse problem of identifying the cyclic variability parametrizing sequence ξ_n from a single noisy observation of angular velocity in each combustion are presented. For the forward problem, the angular velocity computed from measured combustion pressure is compared to measured angular velocity. For the inverse problem, deconvolved ξ_n sequences computed from actual measurements of angular velocity using the stochastic deconvolution algorithm are compared to similar parametrizing sequences computed directly from measured pressure. The results indicate that, using the model of Section 2, both problems can be solved fairly well for the operating conditions analyzed.

3.1 Data Collection. The pressure and angular velocity data were collected from a 3.8 liter V-6 engine with an automatic transmission mounted in an automobile. The tests were performed with the vehicle on a chassis dynamometer. The combustion pressures were measured using Kistler 601B pressure transducers mounted in spark plug adapters conditioned through Kistler charge amplifiers (models 5004 and 504E1). The angular velocity was measured from the starter ring gear (142 teeth) using a magnetic pickup sensor and an analog frequency-to-voltage conversion instrument designed and built by the Vehicular Electronics Laboratory at the University of Michigan. The data were sampled and collected using a commercial general purpose 12-bit analog-to-digital (A/D) conversion high-speed data acquisition board mounted in an 80386-based portable computer at rates of up to 568 points per engine cycle. Sampling was synchronous with engine position (A/D trigger pulses were derived from the starter ring gear teeth sensor signal). Other than the sensors installed to measure engine position and combustion pressure, and cutoff switches in the fuel injector electrical connection circuits, no other modifications were made to the engine, engine control system, or automobile. The cutoff switches enable disconnection of an injector from the engine control module in order to induce a total fuel misfire. Further details are available in Connolly (1992).

3.2 Forward Problem Results. This section presents results of the forward problem solution, computation of angular velocity using measured pressure as input. The computed angular velocity is compared to the corresponding simultaneously measured angular velocity. The continuous engine model, Eq. (3), was used to derive a numerical solution to the forward problem: given measured samples of combustion pressure waveforms, compute the angular velocity waveform. The results given may be compared to the forward simulations using simulated pressure waveforms presented in Connolly (1992) and Connolly and Yagle (1992, 1993).

Figure 1 shows angular velocity computed from measured pressure compared to corresponding measured angular velocity, for five engine cycles. The upper plot shows the first three cycles, and the lower part the last two cycles. This data is for 1400 rpm with the transmission in drive and the parking brake and foot brake firmly depressed, or zero mph road wheel speed, and 130 N-m average indicated torque (computed from measured pressure). The measured velocity is indicated by the solid line, and the computed velocity (computed from measured pressure) by the dashed line. Figure 2 also shows measured and calculated angular velocity for five engine cycles from the same operating condition as Fig. 1. Measured velocity is indicated by the solid line, and computed velocity by the dashed line. Other operating conditions exhibited similar behavior, and are not shown.

The shapes of the simulated and measured angular velocity waveforms bear a very strong visual resemblance—note that peaks and dips in the periodic fluctuations of one waveform correspond to those in the other waveform. This indicates that

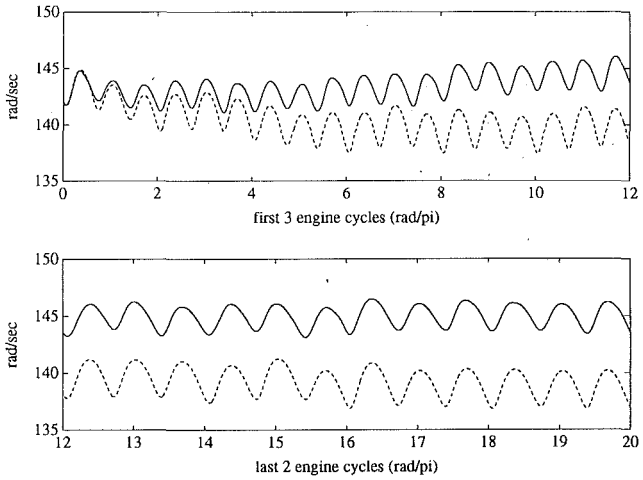


Fig. 1 Numerical simulation of angular velocity computed using measured pressure (- - - line) compared to corresponding measured angular velocity (——— line) at 1400 rpm, 0 mph (in drive with parking and foot brake set), 130 N·m average indicated torque. Note the "drift" in the upper plot.

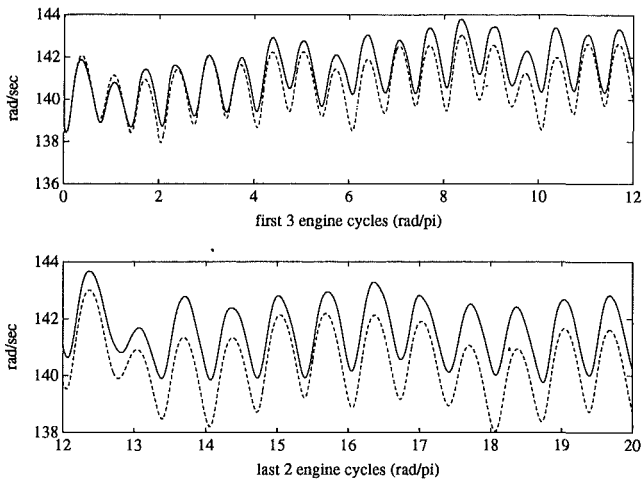


Fig. 2 Numerical simulation of angular velocity computed using measured pressure (- - - line) compared to corresponding measured angular velocity (——— line) under same operating condition as in Fig. 1

the forward model works fairly well for the conditions analyzed (low engine speed and low to moderate load), especially for cyclic variability information present in the angular velocity waveform. This reinforces the supposition that fluctuations in the periodicity of angular velocity are in large part due to cyclic pressure variability.

There is one major difference: a mismatch in low frequency "drift," which is especially apparent in the upper plot of Fig. 1. This drift mismatch between simulated and measured angular velocity is attributed to error in modeling short-term average angular velocity behavior (i.e., over several engine cycles). It was assumed that the load torque $T_L(\theta)$ and the mean $\overline{T}_i(\theta)$ of the indicated torque are constant and opposing in steady-state operation. In a real engine, this is not entirely true. There is always some fluctuation in $T_L(\theta)$ and $\overline{T}_i(\theta)$, where $\overline{T}_i(\theta)$ is defined as the average over several engine cycles. Longer term averages (over *many* engine cycles) of these two quantities must still balance in order to maintain steady-state operation. This drift mismatch can be removed by appropriate filtering of the waveforms (see Connolly, 1992).

4 Experimental Results—Inverse Problem

This section presents experimental results that deconvolve cyclic combustion variability from measured angular velocity. The results were compared to parametrizations computed directly from corresponding simultaneously measured combustion pressure data.

4.1 Pressure Parametrizations. To present the deconvolution and misfire detection results in Sections 4.2 and 4.3, the following mathematical definitions are made. Let $P_n^i(\theta)$, $n4\pi \le \theta < (n+1)4\pi$ be the absolute pressure of the ith cylinder in the nth engine cycle. Then the mean pressure waveform of M engine cycles for cylinder i is

$$\overline{P}^i(\theta) = \frac{1}{M}\sum_{n=0}^{M-1} P_n^i(\theta - n4\pi), \quad 0 \le \theta < 4\pi \tag{8}$$

and the total mean pressure waveform for all cylinders in an N cylinder engine is

$$\overline{P}(\theta) = \frac{1}{N}\sum_{i=1}^{N} \overline{P}^i(\theta), \quad 0 \le \theta < 4\pi. \tag{9}$$

We now define several methods for *parametrizing* cyclic combustion pressure variation at a given engine operating point (i.e., assuming that a given total mean pressure waveform and a set of individual mean pressure waveforms are valid for a given, unchanging operating condition). These parametrization methods are similar to the raised-cosine model for cyclic combustion variability in Eq. (6). In fact, the raised-cosine model itself is another method for parametrizing cyclic variability, based on a hypothetical model for pressure deviation. Here we determine from measured pressure some suitable parametrizations of cyclic pressure variability, based on the deviations of measured instantaneous pressure from a corresponding mean pressure. The purpose of these parametrizations is to provide comparisons between them and the ξ_n sequence computed from corresponding measured speed data using the stochastic deconvolution noted in Section 2. Direct waveform comparisons are discussed in Section 4.4.

The first two parametrizations are based on the deviation of measured pressure from the total mean pressure waveform. One method normalizes the pressure deviations to *unity amplitude*; the second normalizes them to the *energy* of the unity amplitude raised-cosine window (over one combustion interval).

Let the pressure deviations from total mean pressure be defined as

$$\Delta P_n^i(\theta) = \tilde{\overline{P}}(\theta) - P_n^i(\theta) \tag{10}$$

where $P_n^i(\theta)$ is measured absolute pressure (defined just previous to Eq. (8) above) and $\tilde{\overline{P}}(\theta)$ is defined as the periodic extension of $\overline{P}(\theta)$ (defined in (9) above). Define the following sign function:

$$s_n = \begin{cases} 1 & \text{if } |\min(\Delta P_n^i(\theta))| < |\max(\Delta P_n^i(\theta))| \\ -1 & \text{if } |\min(\Delta P_n^i(\theta))| > |\max(\Delta P_n^i(\theta))| \end{cases} \tag{11}$$

for $n4\pi \le \theta < (n+1)4\pi$. This defines whether $\Delta P_n^i(\theta)$ is "positive" or "negative" for a particular combustion. This distinction is made because $\Delta P_n^i(\theta)$ may lie above or below zero. Note that other definitions of the "sign" of $\Delta P_n^i(\theta)$ are possible.

Now it is necessary to normalize the $\Delta P_n^i(\theta)$, thereby defining the cyclic combustion variation parametrization. The first normalization is to unity *amplitude*. Then the parametrizing sequence ξ_n^i for pressure in the ith cylinder, indexed by engine cycle n, is the scalar sequence ξ_n^i where $\xi_n^i = s_n \mu_n^i$ and μ_n^i is defined by

$$\max\left(\frac{1}{\mu_n^i} s_n \Delta P_n^i(\theta)\right) = 1 \tag{12}$$

i.e., the normalized waveform in the argument of max(•) in

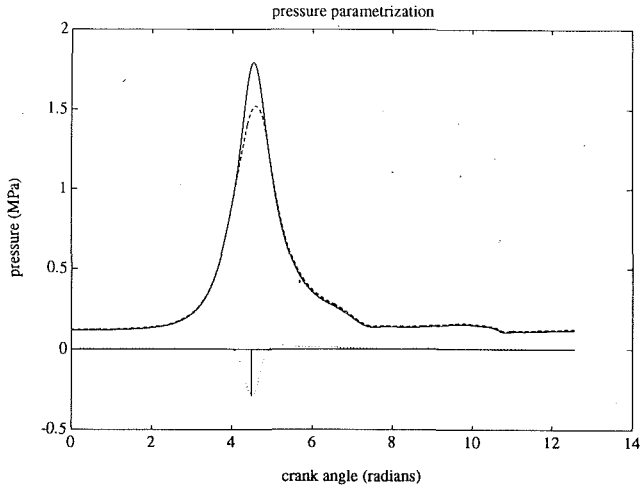


Fig. 3 Parametrization of combustion pressure for a single cylinder: mean pressure $\bar{P}(\theta)$ (——— line); pressure waveform for current combustion $P_n^i(\theta)$ (- - - line); "signed" pressure deviation $\Delta P_n^i(\theta)$ (••• line); corresponding parametrizing sequence value ξ_n^i (——— vertical line)

Eq. (12) has unity maximum amplitude. That is, for combustion n in cylinder i: $\Delta P_n^i(\theta)$ is the *signed* deviation waveform with "sign" s_n; $s_n \Delta P_n^i(\theta)$ is the *unsigned* deviation waveform; μ_n^i is the *unsigned* (or positive) number that normalizes $s_n \Delta P_n^i(\theta)$ to unity amplitude, i.e.,

$$\mu_n^i = \max(s_n \Delta P_n^i(\theta)) \tag{13}$$

(μ_n^i may be thought of as the "unsigned" parametrizing sequence); and finally $\xi_n^i = s_n \mu_n^i$ is the "signed" sequence parametrizing the pressure deviation (or cyclic variability) with "sign" s_n.

Figure 3 illustrates the process of normalizing a pressure deviation waveform to determine the parametrizing sequence value for a single combustion in one cylinder. The illustration is for measured pressure. Shown are the mean pressure (—— line), the current combustion pressure (- - - line), the pressure deviation waveform (••• line), and the pressure parametrizing sequence value corresponding to this combustion (——— vertical line). In this case the "sign" of the deviation waveform is negative. Figure 4 shows an example of parametrization of measured pressure for all six cylinders in one engine cycle, based on the total mean pressure waveform.

The second normalization (and hence parametrization) is the sequence of numbers ξ_n^i where $\xi_n^i = s_n \mu_n^i$ and μ_n^i is now defined by

$$\text{energy}\left(\frac{1}{\mu_n^i} s_n \Delta P_n^i(\theta)\right) = \text{energy}(r\cos(\theta));$$

$$\text{energy}(f(\theta)) \triangleq \int_0^{2\pi/3} f(\theta)^2 d\theta \tag{14}$$

i.e., the "unsigned" pressure deviation $s_n \Delta P_n^i(\theta)$ is normalized to have the same *energy* (over one combustion) as the unity amplitude raised-cosine window in Eq. (6), which is represented in Eq. (14) above by $r\cos(\theta)$. This is a slightly different parametrization than the unity amplitude normalization above (see Eq. (12)).

A third parametrization simply uses the *peak* pressures for each cylinder. The peak pressure occurs during the combustion interval, and its value differs between cylinders and in a given cylinder, due to cyclic combustion variation. This third parametrization is

$$\xi_n^i = \max\{\Delta P_n^i(\theta)\} - \bar{P}_{pk}, \quad n4\pi \le \theta < (n+1)4\pi \tag{15}$$

where $\bar{P}_{pk}$ is the average value of the peak pressures among all cylinders at the given operating condition.

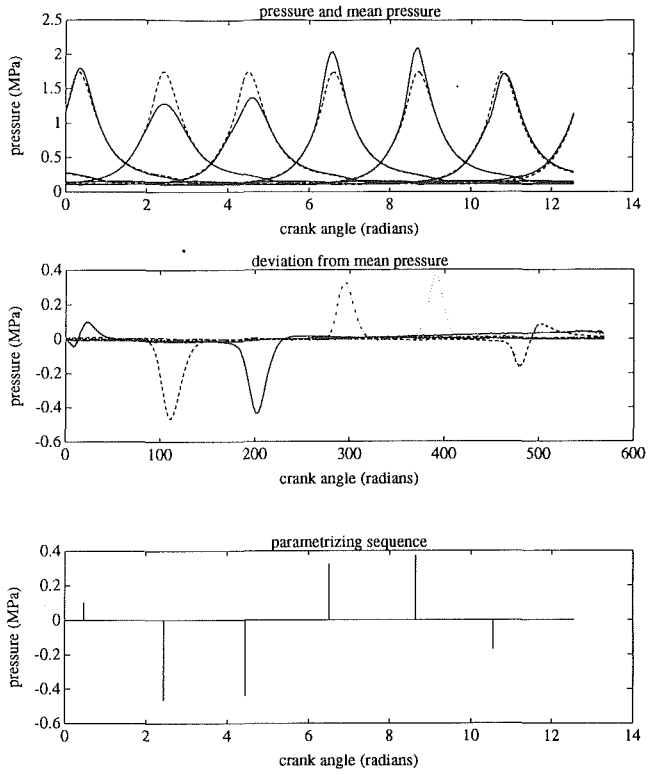


Fig. 4 Parametrization of measured combustion pressure for six cylinders. Upper plot: mean pressure $\bar{P}(\theta)$ (- - - lines); individual cylinder pressure waveforms for current engine cycle $P_n^i(\theta)$ (——— lines). Middle plot: "signed" pressure deviation waveforms $\Delta P_n^i(\theta)$. Lower plot: corresponding pressure parametrizing sequence ξ_n^i (——— vertical lines).

The total cyclic variability parametrization (representing all cylinders) is the "multiplexing," or interspersing, of the ξ_n^i for all cylinders in the correct firing order, resulting in a sequence ξ_n indexed by combustion number n. This idea is straightforward conceptually, but tedious to represent mathematically. These parametrizations given above are three of many possible different parametrizations of cyclic combustion pressure variability. The purpose of computing these parametrizing sequences is to allow evaluation of measured combustion data, and to provide comparison for the hypothetical parametrization (based on the raised-cosine window) computed by the deconvolution algorithm, i.e., the inverse problem solution.

4.2 Deconvolution Results. This section presents experimental results for the inverse problem of deconvolving the parametrizing sequence ξ_n from noisy observations of angular velocity Ω. The first operating point is approximately 1400 rpm with the transmission in drive and the parking brake and foot brake firmly depressed (zero mph road wheel speed). In other words, this is the load the torque converter presents to the engine with the transmission output at zero angular speed. The average value of indicated torque for this operating point (computed from measured pressure) is 130 N-m.

Figure 5 shows the three parametrizing sequences computed directly from the measured combustion pressure for 13 engine cycles (78 combustions). The parametrizations are roughly the same, with significant differences at certain combustions. Also, note combustion 39 in this data set. Analysis of pressure data showed that this combustion had low Indicated Mean Effective Pressure (IMEP), indicative of a "poor" combustion with a corresponding "large" negative pressure deviation. The parametrizations computed from measured pressure at other operating conditions have very similar behavior and are not shown.

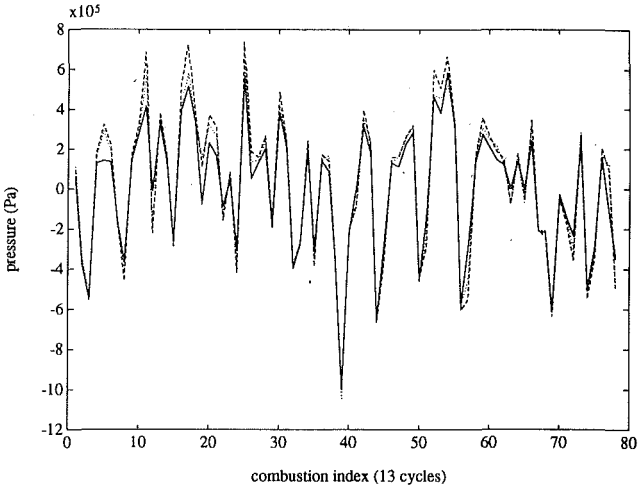


Fig. 5 Parametrizations of measured absolute pressure for 13 cycles, 1400 rpm, 0 mph (i.e., in drive with parking and foot brake set), and 130 N-m average indicated torque: ——— peak pressure; - - - amplitude *normalized*; ••• *energy normalized*

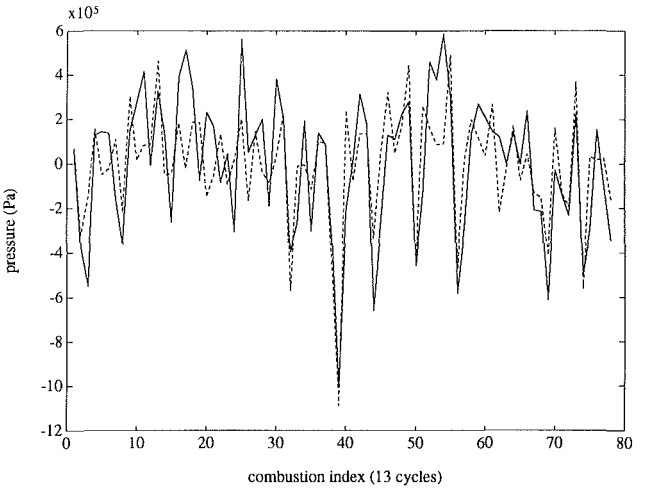


Fig. 6 Deconvolved parametrization (- - - line) plotted along with peak pressure parametrization (——— line) from Fig. 5 at operating point 1400 rpm, 0 mph (i.e., in drive with parking and foot brake set), and 130 N-m average indicated torque

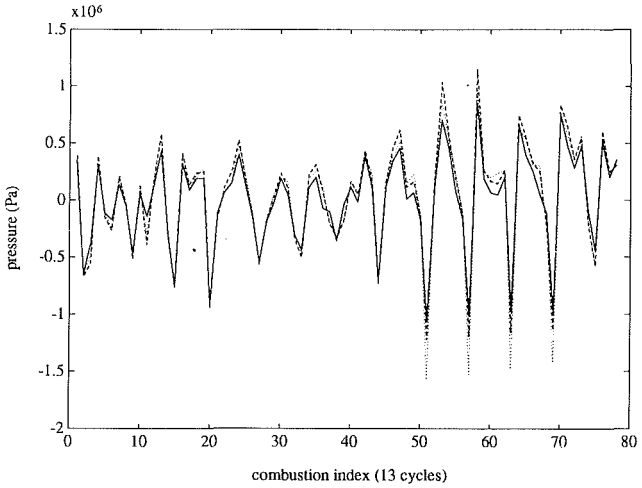


Fig. 7 Parametrizations of pressure at the operating point of Fig. 5 with an induced intermittent total misfire in cylinder 5: ——— peak pressure; - - amplitude normalized, ••• energy normalized

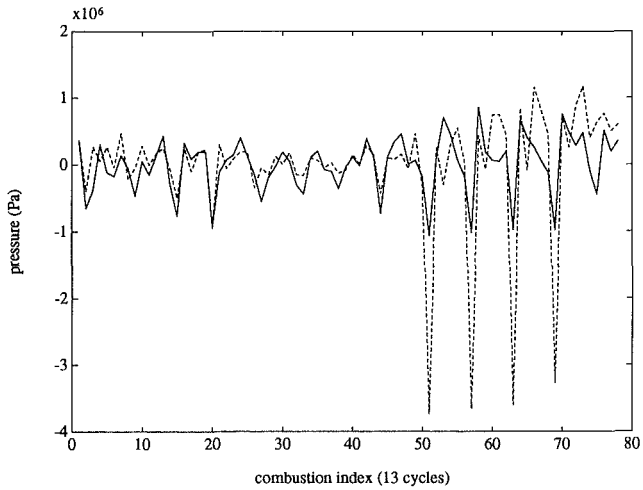


Fig. 8 Deconvolved parametrization (- - - line) plotted along with peak pressure parametrization (——— line) of Fig. 7

Figure 6 shows the output of the deconvolution algorithm plotted along with the corresponding measured *peak* pressure parametrization shown in Fig. 5. The input to the deconvolution algorithm is the corresponding measured instantaneous angular velocity sampled *once per combustion*. The deconvolution algorithm computes a parametrizing sequence that should be *very similar* to the parametrizing sequences computed directly from the measured pressure; note that we are not comparing quantities that are directly equivalent.

With this in mind, we see that the estimated sequence compares well to the sequence computed from measured pressure. Also, the deconvolved sequence is clearly able to distinguish combustion 39, because of its relatively large deviation from mean pressure. This indicates that the stochastic deconvolution method holds promise for distinguishing different modes of combustion (in real time) such as described in Martin et al. (1988), and for detection of misfiring cylinders.

4.3 Misfire Detection. Figure 7 shows parametrizations of measured pressure for the same operating point as Fig. 5, except that now there is an intermittent total misfire in cylinder 5, which was induced by toggling the respective fuel injector connection. Notice the significant negative values due to the misfires. Figure 8 shows a deconvolved sequence, computed from measured angular velocity, for the data set corresponding to the parametrizations of Fig. 7. The behavior of the deconvolved sequence is similar to that of Fig. 6, except that the induced misfires are severely overestimated. However, the deconvolution algorithm correctly estimates "poor" combustions with (negative) pressure deviation magnitudes similar to those of the misfiring cylinders, e.g., combustion 20.

This indicates that the deconvolution algorithm is sensitive to large magnitude outliers, and raises an important issue: How does the proposed stochastic pressure model handle "total" misfires, or a complete lack of combustion? Here misfires are not considered to be part of normal cyclic combustion variability. Therefore, a misfire is interpreted as an outlier datum, a case where the stochastic pressure model and the deconvolution algorithm are not designed to operate well. The results indicate that the deconvolution algorithm is sensitive to large-magnitude negative outliers (as would be produced by a "severe" or "total" misfire among normal cyclic variability), suggesting that the deconvolution algorithm may eventually also be useful as a misfire detector and misfire severity classifier. Further work and evaluation is required.

Results at other operating conditions are now shown. The first is an engine speed of 1900 rpm at a road speed of 30 mph

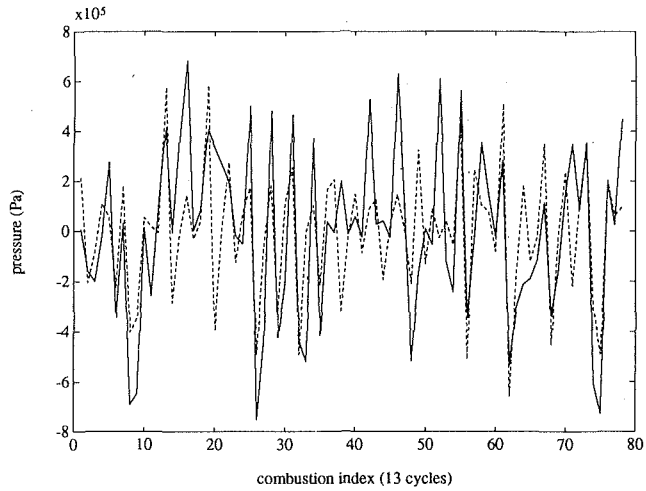


Fig. 9 Deconvolved parametrization (- - - line) plotted along with peak pressure parametrization (——— line) at operating point 1900 rpm, 30 mph, and 152 N-m average indicated torque

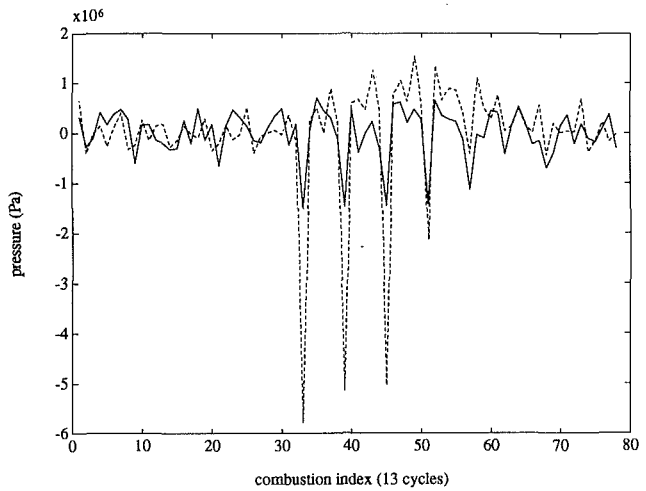


Fig. 10 Deconvolved parametrization (- - - line) plotted along with peak pressure parametrization (——— line) at the operating point of Fig. 9 with an induced intermittent total misfire

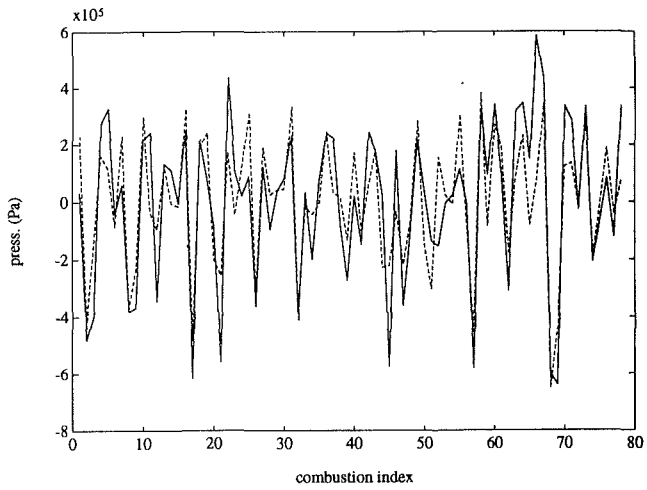


Fig. 11 Deconvolved parametrization (- - - line) plotted along with peak pressure parametrization (——— line) at the operating point 1575 rpm, 30 mph, and 108 N-m average indicated torque

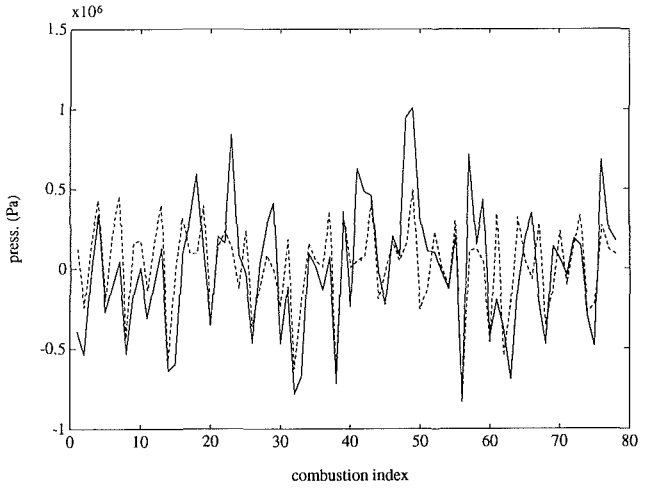


Fig. 12 Deconvolved parametrization (- - - line) plotted along with peak pressure parametrization (——— line) at the operating point 2150 rpm, 30 mph, and 190 N-m average indicated torque

with a computed average indicated torque of 152 N-m. Figure 9 shows the deconvolved sequence plotted along with the corresponding peak pressure parametrization for no induced misfires. Figure 10 shows the deconvolved sequence for the same operating point with an intermittent total misfire. Figures 11 and 12, respectively, show deconvolved parametrization sequences for the following operating conditions: (1) 1575 rpm, 30 mph, 108 N-m average indicated torque; and (2) 2150 rpm, 30 mph, 190 N-m average indicated torque. These data exhibit behavior similar to that discussed above for the other operating points.

4.4 Reconstructed Pressure Waveforms. The deconvolved sequence parametrizes combustion pressure at a given operating condition. Under our stochastic pressure model, this parametrization, along with a mean pressure waveform (or waveforms for individual-cylinder mean pressures) for that operating condition, and a template raised-cosine window representing the deviation from mean pressure, completely describes the combustion pressures at that operating condition. Thus, the deconvolved parametrizing sequence may be used to reconstruct continuous pressure waveforms for *all* cylinders from *one* observation of angular velocity in each combustion. The reconstruction accuracy, of course, depends on how well the model reflects the true pressures. In particular, the raised-cosine window used here to represent pressure deviation may not accurately represent the instantaneous behavior of each pressure deviation exactly *during* the combustion, but it may be able to represent the gross behavior integrated *over* the combustion interval fairly well.

Figure 13 shows a set of individual-cylinder pressure waveforms reconstructed from the deconvolved sequence shown in Fig. 6. The reconstructed pressures (••• lines) are obtained by adding the sequence of raised-cosine windows amplitude-modulated by the corresponding deconvolved pressure parametrizing sequence to the *total* mean pressure waveform $\overline{P}(\theta)$, with phase corresponding to the particular cylinder of the respective combustion. Over each combustion interval, the amplitude-modulated raised-cosine windows raise or lower the pressure (from the mean pressure) around the peak region (where the combustion occurs), thereby simulating combustion variability in the pressure waveforms. It should be emphasized again that these waveforms were estimated from *one* sample of angular velocity per combustion, taken at approximately the beginning of combustion. Pressure reconstructions at other operating conditions exhibit similar characteristics.

The decomposition of pressure into a cyclic variability pa-

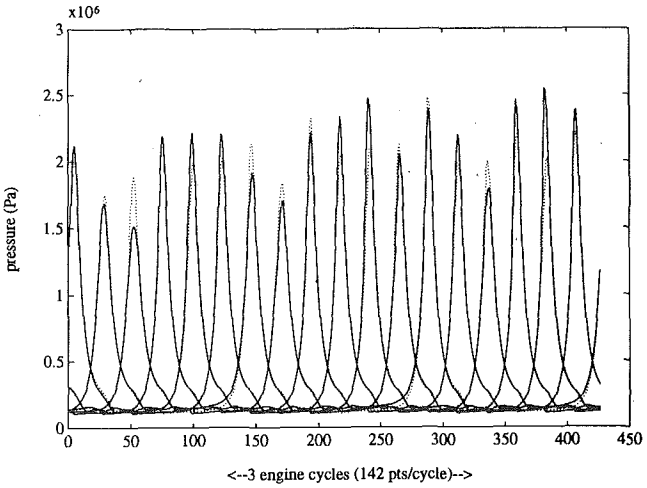


Fig. 13 Measured pressure (——— lines) and reconstructed pressure (••• lines) waveforms using the raised-cosine window amplitude-modulated by the deconvolved sequence from Fig. 6. The operating point is 1400 rpm, 0 mph (i.e., in drive with parking and foot brake set), and 130 N-m average indicated torque.

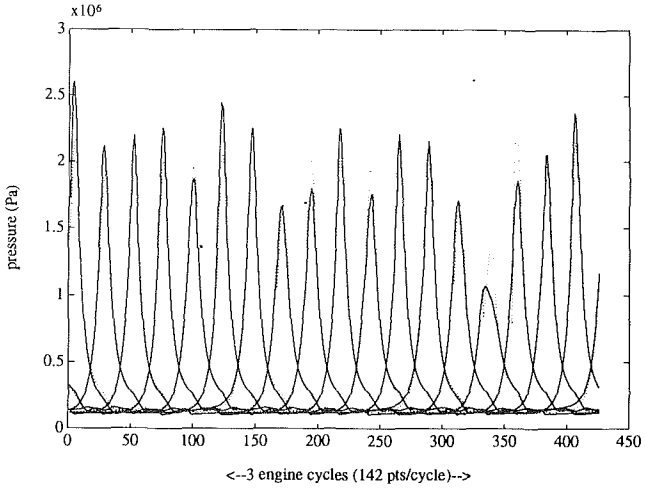


Fig. 14 Measured pressure (——— lines) and reconstructed pressure (••• lines) waveforms using the raised-cosine window amplitude-modulated by the deconvolved sequence from Fig. 6. The operating point is 1400 rpm, 0 mph (i.e., in drive with parking and foot brake set), and 130 N-m average indicated torque. Note the poor reconstruction of the fourth-from-last combustion shown, which corresponds to combustion 39 in Fig. 6—this resulted in two peaks in the reconstructed pressure during this combustion.

rametrizing sequence was illustrated in Fig. 4, which gave examples of pressure parametrizations for measured pressure. Pressure reconstruction is the reverse of this process, using the raised-cosine window as the pressure deviation template waveform. Note that the true deviations of measured pressure from mean pressure are not exactly raised-cosines, but the raised-cosine model strongly resembles the true deviations. Indeed, it was observed that the average of many measured pressure deviations visually resembles a raised-cosine waveform very strongly. It was for this reason, and analytic and computational simplicity, that the raised-cosine window was chosen to *model* pressure deviation.

It is important to note that the above continuous waveform reconstruction technique breaks down when large-magnitude negative pressure variations are estimated. The result is that "too large" a raised-cosine window is subtracted from the peak region of the mean pressure waveform for the particular combustion, resulting in a distorted waveform with two peaks for that combustion. This behavior is shown in Fig. 14, which shows the effect of over-estimating a large negative pressure deviation—the poor combustion number 39. This combustion is the fourth combustion from the last shown on Fig. 14.

5 Conclusion

Experimental results that confirm a combustion pressure to angular velocity model of an IC-SI engine have been presented. This model was presented in Connolly and Yagle (1992, 1993), and developed and tested (via computer simulation and experimental measurements) in Connolly (1992). In particular, the model relates a random sequence parametrizing cyclic combustion pressure variability to fluctuations in angular velocity. The model takes the form of a linear, first-order differential equation relating continuous combustion pressure to continuous angular velocity waveforms, as well as a discrete recursion, indexed by combustion number, relating a cyclic combustion pressure variability sequence to samples of angular velocity at each combustion.

The experimental results shown here verify the theoretical models and computer simulation results for the engine operating conditions analyzed (low engine speed and low to moderate engine load). Results for the *forward problem* show that continuous angular velocity waveforms numerically computed from measured pressure waveforms, using the continuous engine model, match the corresponding measured angular velocity waveforms fairly well. Specifically, variations in the periodicity of angular velocity (cyclic variability information) appear to be well modeled, although there is some low-frequency (below the firing frequency) "drift."

Results for the *inverse problem* show that the pressure variability sequence $\hat{\xi}_n$, deconvolved from measured angular velocity using the model, compares favorably to parametrizations computed directly from corresponding measured pressure. $\hat{\xi}_n$ is a hypothetical parametrization of cyclic pressure variability based on the raised-cosine model for pressure deviation waveforms. Parametrizations computed directly from measured pressure were developed, in order to facilitate comparison with the deconvolved sequences. It was observed that the deconvolved sequence models a "poor" combustion among normal combustions fairly well, but induced total misfires, while identified, are consistently overestimated. This indicates that the deconvolution algorithm may eventually be useful for the detection of misfires and the classification of combustion quality (i.e., from the standpoint of IMEP).

Continuous pressure waveforms reconstructed from the raised-cosine pressure deviation model and a deconvolved sequence were also presented. Considering that only *one* noisy observation of angular velocity in each combustion is used to estimate six continuous waveforms, we conclude that these waveform estimates do a reasonable job of modeling gross behavior *over* each combustion, but do not accurately model the instantaneous shape of the actual waveforms *during* the combustions.

Acknowledgments

The authors would like to thank Professor Giorgio Rizzoni for many useful conversations and for his assistance as co-chair of the first author's doctoral committee. The work of the first author was supported by assistantships from the University of Michigan Department of Electrical Engineering and Computer Science, Systems Division. The work of the second author was supported in part by the National Science Foundation under grant #MIP-8858082.

References

Al-Alousi, Y., and Karim, G., 1984, "Some Considerations of Cyclic Variations in Spark Ignited Engines Fueled With Gaseous Fuels," SAE Technical Report No. 840232.

Belmont, M., Hancock, M., and Buckingham, D., 1986, "Statistical Aspects of Cyclic Variability," SAE Technical Report No. 860324.

CARB, 1989, State of California Air Resources Board Mail-out No. 89-19,

"Malfunction and Diagnostic System Requirements Applicable to New California Passenger Cars, Light Duty Trucks, and Medium Duty Vehicles With Feedback Fuel Systems," May 5.

CARB, 1991, State of California Air Resources Board Mail-out No. 91-18, "Revised On-Board Diagnostics II (OBDII) Requirements," April 8.

Connolly, F., 1992, "Modeling and Identification of the Combustion Pressure Process in Internal Combustion Engines," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI.

Connolly, F., and Yagle, A., 1992, "Modeling and Identification of the Combustion Pressure Process in Internal Combustion Engines Using Engine Speed Fluctuations," *Proceedings, 1992 ASME Symposium on Advanced Transportation Technologies*, Anaheim, CA, Nov. 9–13.

Connolly, F., and Yagle, A., 1993, "Modeling and Identification of the Combustion Pressure Process in Internal Combustion Engines," *Mechanical Systems and Signal Processing*, Vol. 7, No. 5, Sept.

Heywood, J., 1988, *Internal Combustion Engine Fundamentals*, McGraw-Hill, New York.

Janula, J., 1983, "The Effect of Cyclic Dispersion of Combustion Pressure on Working Parameters of a Spark Ignited Engine," *International Journal of Vehicle Design*, Vol. 4, pp. 587–604.

Lichty, L., 1967, *Combustion Engine Processes*, McGraw Hill, New York.

Martin, J. K., Plee, S. L., and Remboski, D. J., 1988, "Burn Modes and Prior-Cycle Effects on Cyclic Variations in Lean-Burn Spark-Ignition Engine Combustion," SAE Technical Report No. 880201.

Matekunas, F., 1983, "Modes and Measures of Cyclic Combustion Variability," SAE Technical Report No. 830337.

Mendel, J. M., 1977, "White Noise Estimators for Seismic Data Processing in Oil Exploration," *IEEE Transactions on Automatic Control*, Vol. AC-22, pp. 694–706.

Nohira, H., Konomi, T., and Ishiyama, S., 1980, "Analysis of Cycle-by-Cycle Variation of Combustion From Pressure Indicator Data," *International Journal of Vehicle Design*, Vol. 1, pp. 131–142.

Obert, E., 1968, *Internal Combustion Engines*, International Textbook Company, Scranton, PA.

Patterson, D. J., 1966, "Cylinder Pressure Variations, a Fundamental Problem," SAE Technical Report No. 660129.

Rizzoni, G., 1986, "A Dynamic Model for the Internal Combustion Engine," Ph.D. Thesis, Department of Electrical Engineering and Computer Science, University of Michigan, Ann Arbor, MI.

Rizzoni, G., 1989, "A Stochastic Pressure Model for the Indicated Pressure Process and the Dynamics of the Internal Combustion Engine," *IEEE Transactions on Vehicular Technology*, Vol. 38, No. 3, Aug.

Yamamoto, H., and Misumi, M., 1987, "Analysis of Cyclic Combustion Variability in a Lean Operation Spark Ignited Engine," SAE Technical Report No. 87054.