PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A
MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Statistical Aspects of Design: Risk Assessment and Structural Integrity [and Discussion]

C. A. Rau, P. M. Besuner and Peter Hirsch

Phil. Trans. R. Soc. Lond. A 1981 **299**, 111-130
doi: 10.1098/rsta.1981.0012

Phil. Trans. R. Soc. Lond. A **299**, 111–130 (1981) [111]
Printed in Great Britain

Statistical aspects of design: risk assessment and structural integrity

By C. A. Rau, Jr and P. M. Besuner
Failure Analysis Associates, 750 *Welch Road*, *Suite* 116, *Palo Alto*, *California* 94034, *U.S.A.*

Statistical variations in input parameters that affect structural reliability have historically been incorporated approximately in engineering designs by application of safety factors. Increased concerns over the injury potential and costs of licensing, insurance, field repairs or recalls, and product liability claims now demand more quantitative evaluation of possible flaws or unusual usage conditions that might result from statistical variations or uncertainties. This paper describes the basic concepts of probabilistic fracture mechanics that are used to assess and control risk. Recent developments in combined analysis methods are presented that utilize field experience data with probabilistic analysis to improve the accuracy of the structural integrity predictions. Several specific examples are described that illustrate how these probabilistic methods are used to assess risk and to provide a quantitative basis for establishing design, operation or maintenance allowables. These procedures, which realistically model the actual statistical variations that exist, can eliminate unnecessarily conservative approximations and often achieve improved reliability at reduced cost.

Introduction

Our knowledge of the metallurgical and mechanical parameters that control cracking and fracture has expanded markedly in recent years. Today, engineers can make very detailed computer calculations of crack growth and instability conditions in complex structures. However, uncertainties in design analyses, material characteristics, non-destructive inspection techniques and service environments continue to complicate the assessment of structural reliability under actual service conditions.

Most engineers acknowledge the statistical variation of key design parameters; however, a significant number maintain that deterministic 'worst case' assumptions should be made whenever the parameter cannot be precisely defined. Such very conservative assumptions are both technically unrealistic and economically unacceptable. This paper will summarize several realistic approaches of incorporating statistical variations by utilizing probabilistic fracture mechanics (p.f.m.) and combined (engineering and statistical) analysis (c.a.) of field experience. The p.f.m. and c.a. approaches have been successfully applied to predict the structural integrity and make quantitative decisions regarding design, operating and maintenance allowables. Two examples will be described to illustrate the basic concepts of risk assessment and its implementation for structural integrity assurance.

Basic concepts

Risk

Risk can be defined as the product of failure probability (P_F) and failure severity (S) should it occur:

$$\text{risk} = P_F S. \tag{1}$$

High risk can result from a large number of moderately severe failures or a low number of very severe failures. In any case, risk is a measure of the total probable impact of failures. There are three basic ways to assess risk. The simplest approach is often simply to extrapolate from past failure data and experience and ignore explicit physical details of the system. This statistical analysis or 'data-base' approach has been taken by several investigations (Freudenthal *et al.* 1966; Rau *et al.* 1977; Mann *et al.* 1974; Cohen 1965, Whittaker & Besuner 1969) to evaluate reliability. The approach is straightforward given a large enough base of relevant data, since a direct relation with actual field experience does not require the use of analytical models or laboratory tests. However, the data base approach does not predict the effects of systematic changes in key parameters, since detailed data on the effect of such changes normally do not exist. The second approach is probabilistic engineering analysis, of which probabilistic fracture mechanics (p.f.m.) is a subset. P.f.m. relies on deterministic models (e.g. stress analysis and material) that relate the residual static strength or fatigue lifetime of a cracked structure to controllable engineering parameters. The statistical variations of each model parameter are quantified and combined to compute P_F and S for the entire system. The third approach, called combined analysis (c.a.), utilizes aspects of both the 'data-base' and p.f.m. approaches. Specifically, it utilizes the minimum amount of engineering modelling required to supplement the routine statistical analysis of service data on frequency and severity of cracking, failure and success.

Fracture mechanics

Cracks present or that may develop in a component are characterized by the stress intensity factor (K), which is proportional to applied stress (σ) and the square root of crack size (a) That is,

$$K \propto \sigma\sqrt{(\pi a)}. \tag{2a}$$

Unstable fracture can occur when K is equal to the fracture toughness of the material (K_c). In the simplest cases, residual fracture strength (σ_f) for a specified crack size or critical crack size (a_c) for specified stress may be determined from (2a):

$$\sigma_f \approx K_c/\sqrt{(\pi a \lambda)}, \tag{2b}$$

$$a_c = (1/\pi\lambda)(K_c/\sigma)^2, \tag{2c}$$

provided fracture occurs before net section yield. Toughness (K_c) is a function of material properties, section size and operating environment, while λ is a factor that depends on the shape of the crack and structure and the type of loading.

When stable crack growth can occur by fatigue, stress corrosion or other modes, the residual life or failure time (t_f) of a part containing an initial crack of size a can be obtained by integrating the crack growth equation for the mode involved. In general,

$$\mathrm{d}a/\mathrm{d}t = f(K); \tag{3a}$$

$$\mathrm{d}a/\mathrm{d}N = f(\Delta K, K_{\mathrm{mean}}). \tag{3b}$$

Often, an approximation by a simple power law is adequate,

$$\mathrm{d}a/\mathrm{d}t = AK^n, \tag{3c}$$

where constants A and n depend on loading mode, type of crack growth, material properties and environment. By rearranging (3c) and integrating, failure time can be calculated as follows:

$$t_f = \int_0^{t_f} \mathrm{d}t = \int_{a_i}^{a_c} \mathrm{d}a/AK^n \tag{4a}$$

$$= 1/A\{\sigma\surd(\pi\lambda)\}^n \int_{a_i}^{a_c} \mathrm{d}a/a^{\frac{1}{2}n} \tag{4b}$$

$$= 1/mA\{\sigma\surd(\pi\lambda)\}^n[(1/a_i)^m - (1/a_c)^m]; \quad m = \tfrac{1}{2}n - 1. \tag{4c}$$

In principle, (1) to (4) can be used to determine the future performance of a component or structure based on known or postulated cracks. They are used routinely to determine design allowable stresses, non-destructive testing limits and even to assess whether a crack detected during maintenance or overhaul should be repaired and, if so, when. Unfortunately, however, uncertainties in any of the equation parameters produce inaccuracies or errors in t_f or σ_f.

To ensure safety, fracture mechanics is often applied, assuming worst-case conditions. For instance, the initial crack size for design computations might be assumed as the maximum crack size that might be undetected by non-destructive inspection (n.d.i.), and fracture toughness might be assumed to be the lowest possible value in a particular environment, even though this condition may never or rarely occur.

Probabilistic fracture mechanics

Many satisfactory components are rejected by worst-case engineering models. P.f.m. accounts for the actual statistical variation in critical parameters based on test or field experience. Various probabilistic structural analysis (Freudenthal *et al.* 1966) and probabilistic fracture mechanics (Rau *et al.* 1977; Besuner & Tetelman 1977) representations are possible, which calculate the probability of failure (P(F[σ, t]) as a function of time or stress rather than a single conservative estimate of life or strength. One convenient representation (Johnson 1977b) is to write the component failure probability P(F) in terms of several key conditional failure probabilities. That is, P(F) = probability of component acceptance multiplied by probability of a non-rejected defect causing failure. Mathematically,

$$P(\mathrm{F}) = \exp(-\mathrm{PN_R})[1-\exp(-\mathrm{PN_F})], \tag{5a}$$

where $\mathrm{PN_R}$ is the probable number of rejection sites in a component,

$$\mathrm{PN_R} = \int_0^\infty \mathrm{pn}(a)P(\mathrm{R}\backslash a, S)\,\mathrm{d}a, \tag{5b}$$

and $\mathrm{PN_F}$ is the probable number of component failure sites in a component not rejected by inspection and causing failure,

$$\mathrm{PN_F} = \int_0^\infty \mathrm{pn}(a)[1-P(\mathrm{R}\backslash a, S)]P(\mathrm{F}\backslash a)\,\mathrm{d}a, \tag{5c}$$

where pn(a)da is the pre-inspection material quality or the probable number of defects of size $a \pm \frac{1}{2}\mathrm{d}a$, $P(\mathrm{R}\backslash a, S)$ is the *rejection probability* of a component given that an actual defect of size a exists and that inspection level is set at size S, and $P(\mathrm{F}\backslash a)$ is the probability of component failure given an existing defect of size a.

Note that failure probability depends strongly on three probability distributions. The probability of failure given a crack of size a, $P(\mathrm{F}\backslash a)$, is generally understood when discussing p.f.m. However, most design engineers do not fully comprehend the comparable impact of the inspection uncertainty distribution and the pre-inspection flaw size distribution, nor the specific procedures for evaluating them quantitatively.

Inspection uncertainty

One of the key concepts that until recently limited application of p.f.m. analysis was a quantitative understanding of inspection (n.d.i.) uncertainty (Besuner & Tetelman 1977; Johnson 1977*b*; Packman *et al.* 1969; Graham & Tetelman 1974; Tang 1973). Typically, an inspector sets a level of sensitivity (S) as shown in figure 1, which ideally should locate all imperfections of size greater than S and not indicate the presence of smaller defects. In (5*b*) and (5*c*), this is quantitatively stated as $P(\mathrm{R}\backslash a, S)$, the probability of rejection for actual flaws of size a at inspection level S. For an ideal inspection, $P(\mathrm{R}\backslash a, S) = 1$ for all cracks of size $a > S$ and $P(\mathrm{R}\backslash a, S) = 0$ for all cracks of size $a < S$. As indicated in figure 1, however, there is a finite probability of rejecting components with flaws smaller than S and a finite probability of not rejecting some with flaws larger than S for a measured eddy current inspection.

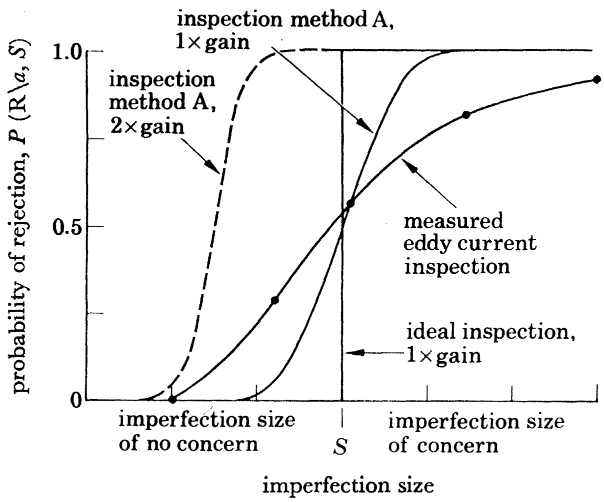

FIGURE 1. Comparison of the inspection uncertainty, characterized by the probability of rejection as a function of flaw size, for ideal and real inspections.

Two other inspection methods are also illustrated. Method A employs the same sensitivity (S) as the eddy current technique, i.e. it rejects imperfections of size S half of the time. However, it has less inspection uncertainty and more closely approaches the ideal, zero uncertainty inspection. The dotted line indicates the performance of method A when used at a higher sensitivity but with the same inspection uncertainty.

Pre-inspection material quality

The distribution of flaws that will get through the n.d.i. and into service components depends upon both the inspection uncertainties, $P(\mathrm{R}\backslash a, S)$, and the distribution of flaws in the material before inspection (Johnson 1977*a*), pn(a). Some approaches to fracture mechanics design very conservatively assume that the probability of flaws of size a occurring in service is equal to the probability of not detecting a flaw of size a_1, which is one minus the probability of rejection, $P(\mathrm{R}\backslash a_1, S)$.

More realistically, when the inspection procedures in figure 1 are applied to a component that initially contains a probable number pn(a) of imperfections of various size a, the distribution of flaw sizes after inspection should be modified as shown in figure 2. The ideal inspection would eliminate all imperfections of size greater than S. However, with real inspections some imperfections larger than S will go undetected while some defects of no concern, $a < S$, will cause rejection. It is this probability distribution of imperfections after inspection, which is the product of the pre-inspection flaw distribution (pn(a)) and inspection reliability ($1-P(\mathrm{R}\backslash a, S)$), that is the actual distribution of initial flaw sizes (a_1) for a p.f.m. calculation of failure probability.

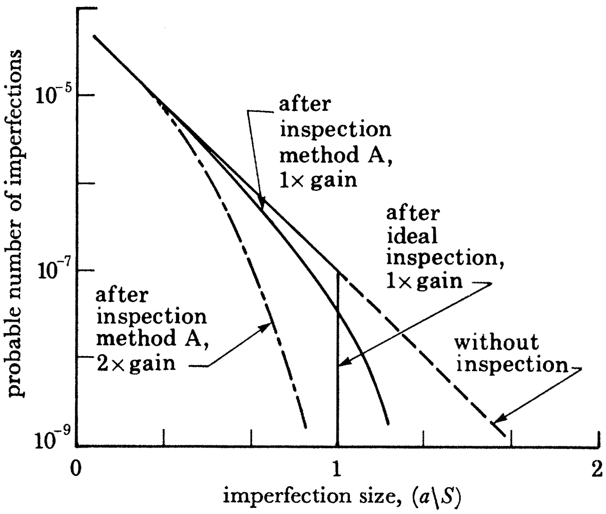

FIGURE 2. The effect of various inspections on the distribution of imperfection sizes going into service.

Retirement for cause

Many components are designed for finite life or load cycle limits, selected to ensure that an acceptably small fraction will fail within this life. Because of many statistical variations, the component lives vary markedly. A life extension strategy based on retirement for cause (r.f.c.) rather than removal from service after a specified time enables most components to remain in operation without failure beyond the initially specified design life for the population as a whole (figure 3). R.f.c. procedures have been developed (Besuner *et al.* 1977, 1978*a*, *b*; Rau 1978; Hill *et al.* 1980) by deterministic fracture mechanics models and p.f.m.

The p.f.m. approach to r.f.c. is more realistic (less conservative) in that it specifies (1) the probability of occurrence of various flaw sizes rather than a maximum flaw size, (2) a probability distribution of steady and cyclic stresses instead of a maximum steady and cyclic stress, and (3) a probability distribution of materials properties rather than minimum properties. The p.f.m. calculation yields a failure probability (reliability) as a function of time instead of one specific crack propagation curve. The use of p.f.m. enables selection of an accept–reject inspection size and inspection interval to assure sufficiently low failure probability rather than a safety factor based upon engineering judgement.

Implementation of r.f.c.–p.f.m. requires extensive collection of statistical data to develop appropriate probability distributions for inspection, service loads, local stress concentrations, and material crack propagation and fracture toughness properties. Large uncertainties in any

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

input parameter necessitate the use of conservative bounds of its probability distribution, thereby reducing the reliable life extension possible from the r.f.c. concept. However, the use of conservative bounds of a parameter's probability distribution, rather than of a parameter's value, often allow significantly more life extension than is possible with the use of deterministic models.

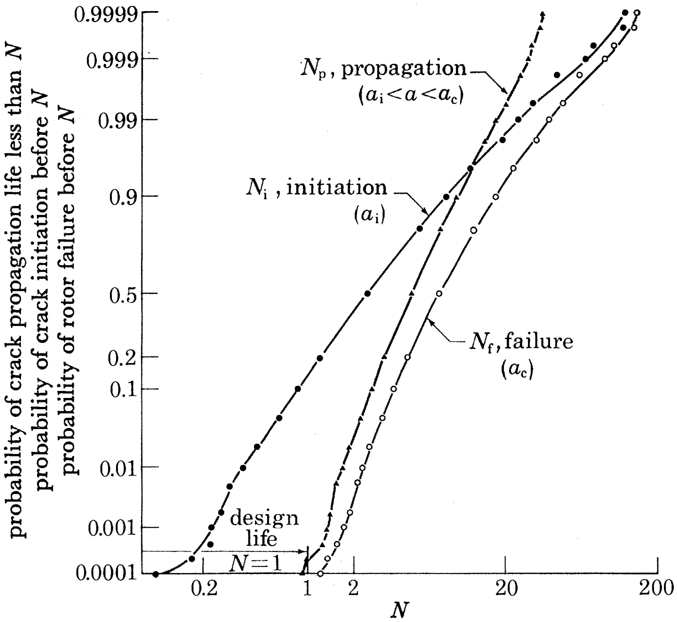

FIGURE 3. Comparison of the variability of initiation lives, propagation lives and total cycles to failure for a hypothetical but realistic population of turbine rotors.

Combined analysis

Errors in life prediction can be largely overcome with a modified procedure that makes more direct use of past operating experience as reflected in inspection information to establish the r.f.c. strategy. The approach is a logical outgrowth of a statistical engineering method that we call 'combined analysis' (c.a.) (Besuner *et al.* 1975; Sorensen & Besuner 1977; Tetelman & Besuner 1977). The c.a. method fits or calibrates an incomplete engineering model against available in-service data. The mathematics of c.a. is the same as that of mathematical regression analysis, but the formulation and use of c.a. in practical applications require sound engineering judgements. The formulation of the engineering model, specification of which parameters are known and unknown, and details of solving for the unknown parameters from the test or field data must be accomplished to minimize the impact of uncertainties upon c.a.-based decisions.

C.a. uses the minimum amount of engineering modelling required to supplement the routine statistical analysis of actual in-service data on the frequency and severity of crack occurrence, failures and successes. For example, with respect to r.f.c. procedures based on c.a., figure 3 shows simulated results for a hypothetical but realistic population of turbine rotors. Conventional design life is established to ensure a very low failure rate for the entire population. A significant number of rotors, say 10%, crack but do not fail, while the balance do not even develop cracks during the conventional design life.

An r.f.c. procedure based on p.f.m., would calculate the distribution of crack propagation lives, N_p, and use it to establish life extension allowables and n.d.i. requirements. R.f.c./p.f.m would rely upon a complete engineering model including, for example, an accurate stress analysis. However, as will be shown in the turbine disk example to follow, r.f.c.–c.a. does not necessarily require accurate stresses if adequate substitute information such as thorough field inspection data is available. C.a. would utilize all available in-service data on N_p and initiation life, N_i, along with laboratory data and engineering models which relate N_i, N_p and failure life, N_f, as shown in figure 3. C.a. procedures can also continuously account for new service or test data by refitting and, if needed, reformulating the incomplete engineering models with the larger data set. Thus, errors in the model or in materials data are not as influential in c.a.–r.f.c. life extension evaluations as they are in p.f.m.–r.f.c. approaches because c.a. provides continual calibration with actual experience. The basic approach of incorporating a 'fudge' factor in the design calculations to 'force' agreement with actual test and field experience is a common design approach. The c.a. approach simply provides a more formal and systematic basis for incorporating actual performance data into the life prediction.

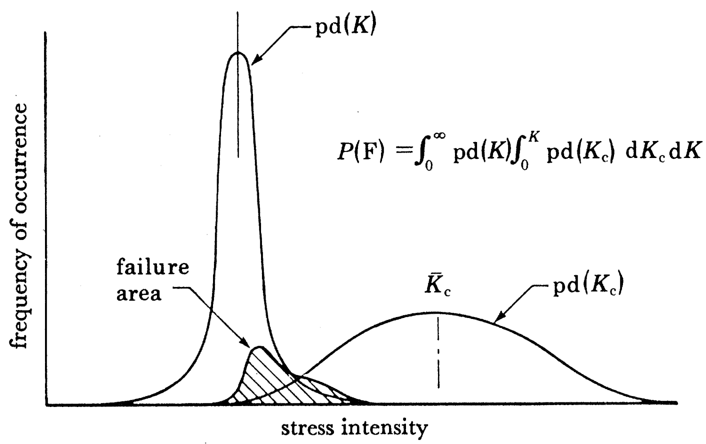

FIGURE 4. Failure probability without stable crack growth.

Decision making

Whether design or r.f.c. evaluations are based on f.m., p.f.m. or c.a., strategy selection should be based on optimizing cost and reliability, subject to reasonable constraints upon injury risk (see, for example, Besuner *et al.* 1975; Starr 1971; Thomas *et al.* 1975; Tetelman 1977). The effects of realistic inspection, analysis and service uncertainties must be quantified and incorporated. Experience shows that both brittle fracture and fatigue phenomena are probabilistic in nature. Even without time-dependent crack growth, components will exhibit a distribution of fracture strengths p.d.(σ_f) and be subjected to a distribution of service stresses p.d.(σ). Fracture occurs for $K > K_c$ and the probability of this occurring is given by the double integral of the product of K and K_c distributions, p.d.(K) and p.d.(K_c), where they overlap as shown in figure 4.

With time-dependent crack growth, p.f.m. is conceptually similar but more complicated to implement. Each of the stress (σ), materials (A, n, a_c) and crack (λ, a_f) parameters in (4*b*) have a probability distribution. The resulting probability distribution of time to failure, p.d.(t_f), is computed by combining the probability distributions of all parameters, some of which are not independent. Calculation of p.d.(t_f) is often complex enough to require powerful numerical procedures such as Monte Carlo simulation by computer (Besuner & Tetelman 1977).

To minimize cost and maximize efficiency, the safe life for the structure is set at a point just before risk reaches an unacceptable value, which of course depends upon both failure probability and the consequences of failure if it occurs. For critical components, the acceptable failure rate may be that associated with abusive loading only, in which case t is set by the time where failure rate just begins to rise owing to wear-out modes such as fatigue, creep and stress corrosion. For non-critical components, higher failure rates may be acceptable. In either case, for optimized design or r.f.c. conditions a range of crack sizes may exist in components with acceptable risk because the cracking rate usually exceeds the failure rate.

Total expected cost, C_t, associated with continuing in service a population of N similar components subject to one potential failure mode is

$$C_t(x_i, a_j) = \sum_{L=1} C_L(x_i, a_j)\mathrm{PN}_L, \tag{6}$$

where C_L represents the average event costs and PN_L the probable number of units undergoing event L. Each cost and probable number is a function of i system variables, x_i, which are controllable and j non-controllable system parameters. The mathematical problem is to find the combination of x_i that produce the lowest $C_t(x_i, a_j)$. Decisions can then be made on the basis of the options that result in minimum total cost, including indirect and intangible ones associated with failure. Specifically, the unit costs are for inspection (C_i), analysis (C_A), repair (C_R), replacement (C_B) and failure (C_F). C_i, C_R and C_B include downtime cost, which may be the dominant cost, especially in large systems, as well as the cost of hardware and personnel involved in performing each task. It is not difficult to obtain monetary estimates for most C_L. However, failure costs, C_F, are much more complex. This parameter includes the average cost of property damage associated with a failure plus the cost of personal injury and fatality. It is very difficult to predict accurately the average cost of property damage arising out of the failure on one component in a system, since the physical consequences of the failure could range over a spectrum from trivial to catastrophic. The cost of assessing personal injury is even more complex because, apart from dealing with a whole spectrum of possible consequences, there is a historical reluctance to assign monetary values to injuries and death. Viable suggestions for evaluating failure severity and comparing with acceptable risks without directly assigning monetary value have been developed (Besuner *et al.* 1975; Starr 1971; Tetelman 1977). Table 1 summarizes some typical risks of fatality. One suggested approach is to minimize C_t subject to the constraint that the risk be no larger than that demonstrated in service by the same or similar equipment.

Examples of the risk assessment

As described previously, there are three ways to compute risk: data base, probabilistic fracture mechanics and combined analysis. Two examples will be described to illustrate these three approaches to risk assessment and optimized decisions regarding structural reliability. The first example utilizes the data base approach to evaluate alternative lorry tyre rim designs and regulatory suggestions. In the second example, a fleet of jet engine disks is examined at the end of the vendor-specified design life where no failures have occurred in any of the disks.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

However, cracks have been found in some of the rim slot blade attachments, which are inspectable. Both p.f.m. and the c.a. methods are considered to establish r.f.c. strategies and evaluate the cost savings possible by returning most of the disks for additional service, with full appreciation of the fact that owing to uncertainties in inspection, modelling, and material properties, some disks may fail during the extended life.

Lorry tyre rims

Often, the most difficult and controversial area of applied reliability and risk analysis is the setting of safety criteria and the answer to the question, 'How safe is safe enough?' The difficulty arises because of (1) the difference in risk acceptance among people, ranging from absolutists who demand zero risk to adventurers who enjoy risk taking as a benefit, (2) the question of who is at risk, and (3) the difference between rhetoric, which accompanies risk controversy in business decisions (including product litigations), and actual risk acceptance by individuals and society as a whole.

An interesting example that explores the safety question is a recent study in which we were asked to evaluate alternative methods to reduce the risk associated with explosive separation of tyres from multipiece wheels (rims) used on lorries, buses and other large vehicles. Two methods proposed are:

(1) Total recall of all multipiece wheels and replacement with single-piece wheels, as proposed by the National Highway Transportation Safety Administration (N.H.T.S.A.);

(2) Upgrading of tyre maintenance and changing procedures proposed by the Occupational Safety and Health Administration (O.S.H.A.).

The risk analysis in this case was based on the data base approach, and made use of statistical analysis of actual field experience. We studied 378 accidents associated with multipiece wheels, with the use of reports obtained from Firestone, from the Insurance Institute for Highway Safety, and from the National Highway Traffic Safety Administration. In each case we extracted the failure conditions and accident consequences. We defined the total usage of both single and multipiece rims and the corresponding usage and success data. Our analysis revealed that the bulk of the risk associated with use of the multipiece rim is centred in the maintenance activity: 80% of the mishaps occur during maintenance, and the fatality rate for those mishaps is in excess of three times the moving vehicle fatality rate. However, the risk associated with wheel maintenance is not high.

Multipiece wheels are involved in maintenance accidents approximately once every million maintenance events, and death is caused in approximately one-sixth of these incidents. Thus, changing a multipiece rim has a fatality risk of far less than one in a million. Recognizing that the maintenance tyre specialist is exposed to risk of wheel-related mishap for periods of approximately 20 or 30 minutes per wheel in maintenance, and only one-sixth of mishaps are fatal, his exposure to risk (per tyre change or per unit time) is less than that which he readily accepts when driving to work or travelling on commercial aircraft.

To illustrate how small this risk is, a series of normal activities are listed in table 1 that increase one's chance of death by one in a million. Our analyses also revealed that effective promulgation of the proposed O.S.H.A. standard would potentially prevent 68% of the accidents and 84% of the fatalities currently associated with multipiece wheels. The cost of implementing the O.S.H.A. standard appears insignificant to moderate and more cost effective than multipiece rim replacement with single-piece rims. We concluded that the high cost of replacement of the multipiece rim components is completely unjustified in the light of the lack of evidence that a change in rim type would reduce the already small rate of fatalities. Furthermore, the risk posed to society by the multipiece rim is so low that an increase by a factor of 3 in the number of known accidents would not significantly affect the conclusions.

Table 1. Activities that increase the chance of death by one in a million

smoking 1.4 cigarettes (cancer, heart disease)
spending 1 hour in a coal mine (black lung disease)
spending 3 hours in a coal mine (accident)
living for 2 days in New York or Boston (air pollution)
travelling 10 miles by bicycle (accident)
living 2 months in average stone or brick building (cancer from natural radioactivity)
living for 2 months with a cigarette smoker (cancer, heart disease)
eating 40 tablespoonsful of peanut butter (liver cancer from aflatoxin B)
drinking Miami drinking water for 1 year (cancer from chloroform)
eating 100 charcoal broiled steaks (cancer from benzopyrene)
travelling 30 miles by car (accident)
changing 6 multipiece wheel tyres (accident before O.S.H.A. standard)
changing 34 multipiece wheel tyres (accident after O.S.H.A. standard)

Certainly a single-piece rim does not suffer explosive separation of the side ring. But single-piece rims suffer from tyre bead explosions that virtually never occur on multipiece rims. A single-piece rim presents the unavoidable problem that a broken or weakened bead cannot be inspected after assembly, whereas most multipiece rims can be inspected both visually and tactilely after assembly. A second hazard that may arise owing to the exclusive use of tubeless tyres (which would be dictated by the single-piece rim) is the greater possibility of tyre burst due to overpressurization used to 'pop' and correctly seat the tyre bead.

These two risks associated with single-piece wheel–tubeless tyre servicing have actually been experienced by one major tyre manufacturer, who reports 36 known bead breaks and 72 known tyre explosions over approximately 10 years. Also, a major manufacturer of car tyre servicing machines indicated awareness of at least 50 mishaps related to tyre overpressurization to force the tyre beads over the head lugs.

A conservative estimate can be made of the cost effectiveness of risk reduction by substituting single-piece rims for multipiece rims. We assume that all 71 fatalities attributed to multipiece rims in the last 20 years would have been saved (round up to 80), at a cost of $\$10 \times 10^9$. The latter figure is a conservative fraction of the manufacturer's estimated cost to change rims and assumes the phase-out of multipiece rims to be gradual. From these variables, U.S.A. society would pay \$125000000 for each prevented fatality.

To put this in perspective, table 2 compares several highway safety measures that reduce fatalities, figures assembled and estimated by the Secretary of Transportation, to which we have added the two safety measures concerned with multipiece wheels. Replacing the multipiece with the single-piece rim would be more than ten times less cost effective than any listed contemplated use of public funds for highway risk reduction. A similar comparison can be made for injury accidents.

The argument cannot be made that private firms and not the public would bear the cost of a rim changeover. Ultimately, most or all of this cost would be passed on to the consumer in terms of higher freight and haulage rates.

All of the 302 mishaps related to multipiece wheel maintenance were examined to determine what provisions, if any, might reduce the risk of injury associated with explosive separation. It was determined that 172 of those accidents (36 fatalities, 130 injuries, and 6 with no injuries) could have been prevented by following five simple rim handling work rules.

Inflation: no multipiece wheel should be inflated without use of a fixed safety cage (and clip-on air chuck), or portable safety device to restrain any missile associated with separation. Portable pressure control device and inflation stick should be used to allow the maintainer to stand other than in line with the wheel. Hands should be kept out of the safety cage.

Hammer: no wheel, partly or fully inflated, should be struck with any object, nor should the side and or locking rings be pried.

Deflation: every wheel should be deflated before removal from the axle (both wheels when dual mount) and before assembly.

Mismatch: parts should be inspected to insure that they are compatible. Use of the rim chart is recommended.

Miscellaneous: severe corrosion or deformation of parts must result in substitution of new components before assembly.

Table 2. Cost of countermeasures to reduce risk

measure	cost per fatality forestalled/$
mandatory safety belt use	506
regulatory and warning signs	34000
guard rail	34100
median barriers	228000
improve tyre change practices by using O.S.H.A. proposal	1000000
periodic motor vehicle inspection	2120000
road markings and delineators	2700000
paved or stabilized shoulders	5800000
ban all multipiece wheels	125000000

All of the proposed work rules require no exercise of judgement. They demand only that a segmented procedure be followed in the servicing of multipiece wheels. Had they been observed in the 302 events examined, approximately 68% of the accidents and 84% of the fatal injuries would have been avoided.

The proposed O.S.H.A. standard is even more comprehensive than the five procedural accident prevention activities that we derived by examining the available accident data. Consequently, we endorsed the O.S.H.A. standard and we testified before an O.S.H.A. hearing panel that any safety-intended replacement of multipiece wheels with single-piece wheels could not possibly produce a decrease in risk sufficient to offset the enormous cost and disruption predicted.

Turbine disks

Life extension of turbine disks has been analysed by p.f.m. and c.a. These disks and similar rotating components usually have two or three potentially life-limiting regions: the bore (high nominal stresses), the bolt hole region (high concentrated stresses) and the rim slot (high thermal–mechanical and multiaxial local stresses).

In this example, which is hypothetical but based upon realistic data and analyses performed by the authors, disk life was limited by the initiation and growth of fatigue cracks in the rim. It was known that none of the disks had failed during their design-specified fatigue life, and each rotor had been inspected at the end of its design life. Those containing rim cracks were recorded and maximum crack size noted. It was also demonstrated that fracture mechanics analysis based on influence function techniques can describe the relation between crack size a, nominal stress and stress intensity factor K, and that the crack growth rate model and environmental parameters in (3) are well documented.

We have been developing analytical procedures (Besuner *et al.* 1975, 1977, 1978*a*, *b*; Rau 1978; Sorenson & Besuner 1977; Tetelman & Besuner 1977) to help to evaluate the effectiveness of any proposed equational life extension method based on r.f.c., considering service load, analysis and inspection uncertainties. In this evaluation, a probabilistic simulation model was applied to a population of 1000 gas turbine disks to determine the impact of (1) stress variations from disk to disk, (2) unknown precise age or past usage of each disk, (3) repeated as well as single inspections and (4) inspection intervals. In addition, load conditions were established by using both c.a. of inspection results and design calculations.

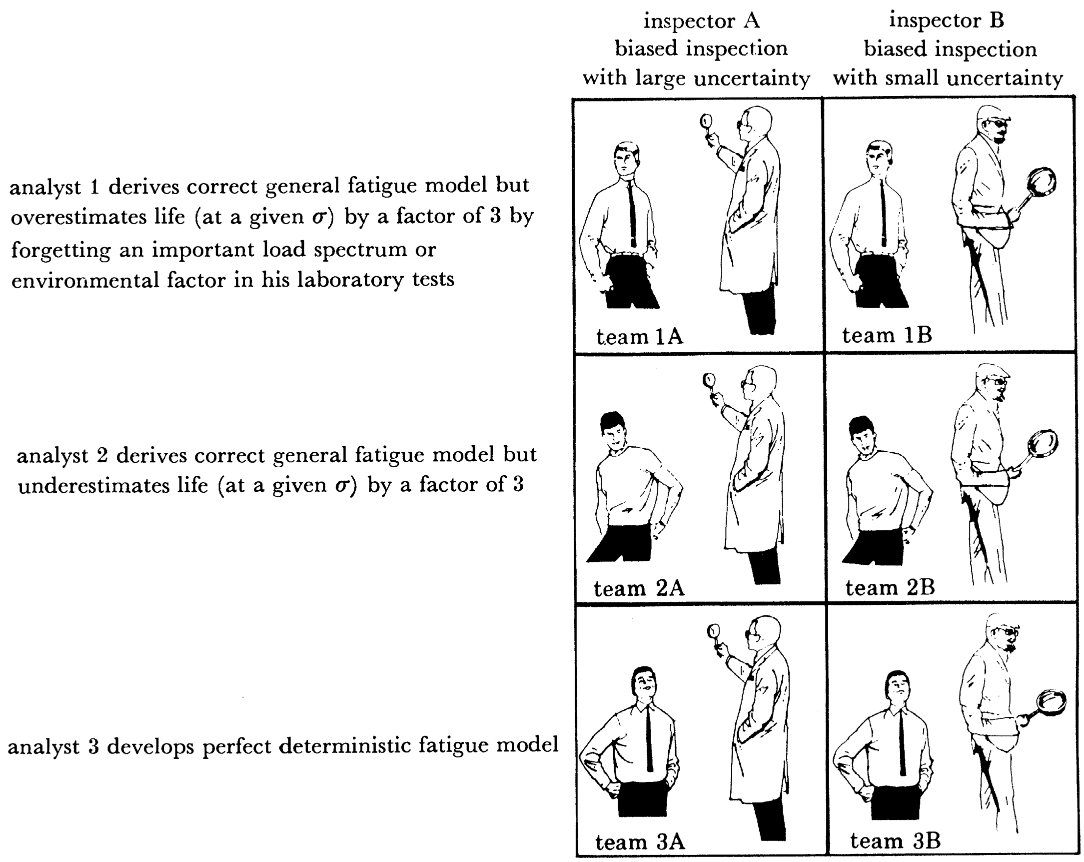

FIGURE 5. Hypothetical teams assigned to set a life limit for 1000 inspected turbine disks, based on retirement-for-cause procedures.

The effect of analysis and inspection uncertainties was evaluated by considering six hypothetical teams (figure 5) of inspectors and analysts with varying but well defined capabilities. The total expected cost savings possible by implementing r.f.c. procedures with each specific team were computed for 1000 simulated disks and normalized to obtain the average cost savings per rotor for each procedure and team.

Details of the equations used to simulate fatigue crack initiation, subcritical growth and brittle fracture are provided by Besuner *et al.* (1978*a*) and Sorenson & Besuner (1977).

Basically, the simulation was formulated to include significant fatigue life scatter at a given stress level, significant variation in the nominal stress level σ, and realistic relations among the times required for crack initiation and subcritical crack growth to various crack depths. Figure 3 was obtained during this simulation and does, in fact, show that significant life variations occur from disk to disk and that the first few failures occur between the design life, $N = 1$ and $N = 1.5$. This is intended to simulate a good initial life prediction or design analysis resulting in neither excessive failures nor severe overdesign.

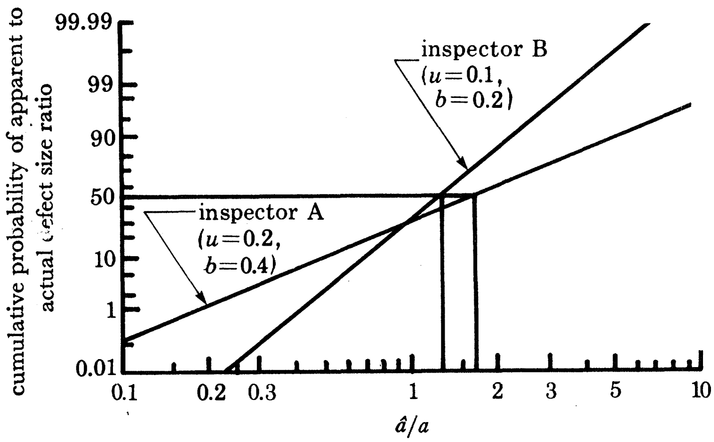

FIGURE 6. Inspection bias (u) and uncertainty (b) for inspectors A and B defined by the probability distribution of apparent to actual crack size; lg $(\hat{a}/a)$ = Gau (u, b).

One characteristic of this disk life is that at a given age beyond the point of crack initiation, a 7–8% change in σ will change crack depth by a factor of approximately 3 and crack area by a factor of 9. Thus, for each in-service disk, an inspection error of a factor of 9 in the estimate of crack area is equivalent to a 7–8% analysis error in the effective nominal alternating stress. For this reason we evaluated the cost effectiveness of (1) a combined analysis r.f.c.–c.a. procedure in which the inspection results are used to infer a nominal stress level as well as (2) r.f.c.–p.f.m. in which the stress level is estimated from design calculations and usage variation.

Simulation of inspection uncertainty

Inspection uncertainty is simulated by calculating apparent crack depth ($\hat{a}$, where the hat denotes an estimated value) as a maximum of 0.001 in, or $\hat{a}_{\text{insp}}$. That is

$$\hat{a} = \text{maximum}\ (0.001\ \text{in}, \hat{a}_{\text{insp}}), \tag{11}$$

where $\hat{a}_{\text{insp}}$ is derived solely from the calibrated inspection signal given by

$$\lg\,(\hat{a}_{\text{insp}}/a_{\text{t}}) = \text{Gau}\,(b, u), \tag{12}$$

where Gau refers to a normal distribution with parameters b, the mean, and u, the standard deviation. The value of b, taken to be either 0.1 or 0.2, reflects positive bias in the inspection. For example, $b = 0.1$ implies that, typically, crack size will be overestimated by a factor of $10^{0.1} = 1.26$. This bias could represent conservative procedures or the fact that multiple rim slots create greater chance for a high, rather than low, estimate of the rotor's maximum crack size, a_{t}. The inspection uncertainty is reflected by u, taken to be either 0.2 or 0.4, and which is the logarithmic standard deviation of $\hat{a}_{\text{insp}}/a_t$. Equation (12) is plotted in figure 6 for the two sets of parameters b and u corresponding to inspectors A and B. Inspector A has a poorer inspection technique with larger bias and uncertainty than inspector B.

Simulation of analysis uncertainty

Various hypothetical analysts were considered, three of which are described below. The first analyst modelled the fatigue process by an expression (Besuner *et al.* 1978*a*; Sorenson & Besuner 1977) that overpredicts life by a factor of 3, on average. Such an error could be due, for example, to the use of inappropriate temperatures in laboratory fatigue tests. The second analyst modelled the fatigue process by an expression that underpredicts life by a factor of 3. The third analyst developed a nearly perfect deterministic model of the fatigue process that corresponds closely to median life.

It is assumed that each analyst has included all relevant failure modes. For example, a vibratory stress greater than expected could cause the effective critical crack depth a_c to be limited by the high frequency fatigue threshold level rather than the material's fracture toughness. The nominal cyclic σ was assumed to be normally distributed about a mean, $\overline{\sigma}$, with a coefficient of variation (v) of either 5 or 12.5%. That is,

$$\sigma = \text{Gau}(\overline{\sigma}, v\overline{\sigma}), \tag{13}$$

where $\overline{\sigma} = 34.2,\ 40 \text{ or } 47.1 \times 10^3 \text{ lbf in}^{-2}\dagger$

and $v = 0.05 \text{ or } 0.125.$

These r.f.c. procedures were programmed into a Monte Carlo simulation of 1000 rotors for each procedure and team. For each rotor, the program generates 'in-service' fatigue data, performs a specified r.f.c. procedure and makes random errors using the probability distributions for inspection, analysis and materials variations and checks for rotor failure. Costs are assigned to the various outcomes of the r.f.c. procedure for each jth rotor. Each time the rotor is inspected, a negative cost savings of $G_{j,\text{i}} = -\$2000$ is assigned. Each time rotor life is extended, a cost saving of $G_{j,\text{e}} = \$20\,000\ \hat{N}_\text{e}$ is assigned, since \$20 000 is the cost of the disk designed for one life unit and $\hat{N}_\text{e}$ is the perceived amount of life extension until either the next inspection, retirement or failure, whichever occurs.

Should failure occur before the rotor is retired, a negative cost savings of $G_{j,\text{f}} = -\$1\,500\,000$ is recorded. Although the estimate of the precise expected cost of failure, G_f is a complex and controversial subject, G_f is finite and failure probability is greater than zero. To insist otherwise is unrealistic and impractical.

The total r.f.c. cost saving for each rotor is obtained by the summation

$$G_j = f_\text{i} G_{j,\text{i}} + f_\text{e} G_{j,\text{e}} + f_\text{f} G_{j,\text{f}}, \tag{14}$$

where f_i, f_e and f_f represent the number of incidents for each type of cost gain for the jth rotor.

The expected average dollar gain (cost saving), G, for each r.f.c. procedure is then estimated by averaging all 1000 simulations for each procedure:

$$\overline{G} = \sum_{j=1}^{1000} \tfrac{1}{1000} G_j.$$

† 1 lbf in^{-2} ≈ 6.9 kPa.

The r.m.s. error of the $\bar{G}$ estimate, that is the sampling tolerance of standard deviation of $\bar{G}$, due to the use of a finite number of rotor simulations (1000) is estimated to be \$2000 near the optimum safety factor (or inspection interval) where simulated failure probability is on the order of 0.001. Thus, the results described below are accurate to about ± \$2000.

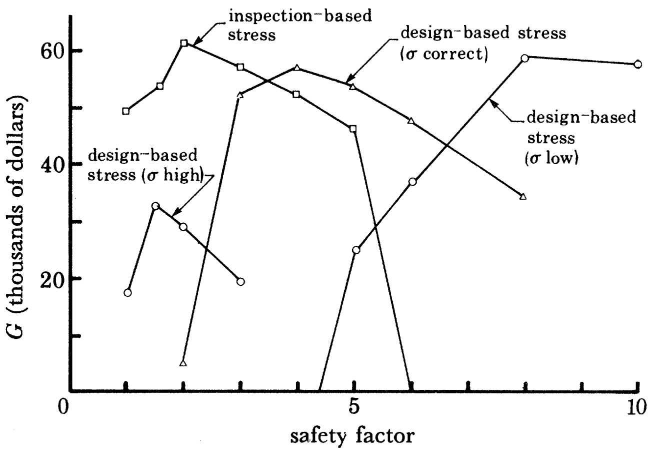

FIGURE 7. Comparison of the expected cost savings per rotor (G) with r.f.c. procedures with the use of various stress inputs and a 5% coefficient of variation on stress. Note that all curves are for team 3B.

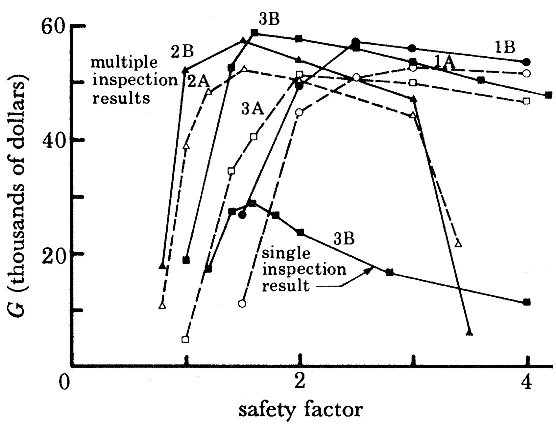

FIGURE 8. Expected cost savings per rotor (G) for single and multiple inspection r.f.c.–c.a. procedures with various analysis and inspection uncertainties and a coefficient of variation on stress of 12.5%. Note that analyst 3 and inspector B are the most accurate performers.

The average cost savings, $\bar{G}$, computed for each r.f.c. procedure and analyst–inspector team are summarized in figures 7 and 8 for various safety factors, which is closely related to inspection interval. Figure 7 shows the cost savings for three r.f.c.–p.f.m. procedures where the mean calculated design stresses are 47.1, 40 and 34.2×10^3 lbf in^{-2}, respectively, and the coefficient of variation is $v = 0.05$. The p.f.m.–r.f.c. approach does reasonably well provided the analyst calculates the correct stress value. However, if too high or too low a stress is chosen, much lower cost savings result, even with the best analyst–inspection team.

Three conclusions are evident from the average cost savings for single and multiple inspection-based stress r.f.c. procedures (figure 8): (1) a different optimum safety factor exists for each team using each procedure; (2) teams with the better inspectors do consistently better than their less able counterparts, but the latter still can achieve substantial cost savings; and (3) provided the safety factor is between two and three, multiple inspection r.f.c. procedures are much more effective than single inspections for all teams.

The maxima in the $\bar{G}$ curves are due to the trade-off between premature failures and premature retirement. The optimum safety factor (s.f.) or inspection interval represents the best balance between these competing effects and usually corresponds to failure rates on the order of one per 1000 rotors. The sharp drop in cost savings at the low s.f. end corresponds to too many failures. The more gradual decrease in cost savings at high s.f. reflects the increased cost of unnecessary inspections and more premature rotor retirements.

Figure 8 shows that the optimum s.f. is, unfortunately, a strong function of the accuracy of the analytical model. Since there is little or no knowledge of analysis error before implementing the r.f.c. procedure, there seems to be little chance of accurately choosing the optimum s.f. However, because the maxima are quite broad, a safety factor between two and three will provide substantial and near-optimum cost savings for all the teams evaluated so long as an r.f.c.–c.a., multiple inspection procedure is used.

The Monte Carlo simulation was also used to evaluate specific concerns. The analyst may not know exactly the past usage or effective age of the disk N_t at inspection time. Detailed simulations with varying degrees of unknown usage within the population showed that even when $\lg(\hat{N}_t/N_t) = \mathrm{Gau}\,(0, 0.3)$, that is an error by a factor of 2 for 68 % of the time, there was relatively small effect (12 % reduction) on the optimum cost savings by using the r.f.c.–c.a. procedure. If the analyst had neglected fatigue crack initiation but modelled crack growth correctly, in a deterministic sense, results are unsatisfactory (see Besuner *et al.* 1977). This analyst (model) is not bad for predicting the shortest life parts of the distribution (note in figure 3 the close proximity of the N_f and N_p curves at the early life end), but overall this analyst did poorly. The basic problem is that, within the context of an r.f.c.–c.a. application, this analyst severely underestimates the crack growth rate for most cases in which cracks are observed. For example, if a 0.002 in (0.005 cm) crack depth is observed at $N_t = 1.00$ life unit, a propagation-only model would imply a crack of size 0.001 in (0.0025 cm) at $N = 0$ and therefore an average growth rate of 10^{-3} in per life unit. However, if the 0.001 in crack actually initiated at $N = 0.9$, the true average growth rate would be ten times larger than calculated. Thus, calculated failure lives based on unrealistically slow growth rates would too often overestimate actual lives.

This example demonstrates that even with large inspection and analysis uncertainties, cost effective life extension at extremely low failure probabilities can be achieved by using r.f.c.–c.a. procedures that make full use of in-service structural fatigue data. Computer simulation of the fatigue crack initiation and growth process provides an inexpensive means of quantitatively evaluating the effect of errors and uncertainties, thereby enabling selection of an optimum r.f.c. strategy. R.f.c.–c.a. procedures with the use of stress calculated from inspection results, rather than conventional design analysis, can be less sensitive to uncertainties and will produce larger cost reductions in most applications.

MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES
THE ROYAL SOCIETY A
PHILOSOPHICAL TRANSACTIONS OF

Recent developments

The simulation analyses described in the turbine disk example have defined some requirements for non-destructive inspection and fracture mechanics development programmes to provide the specific input data to implement r.f.c. Specifically, the state-of-the-art capability of eddy current n.d.i. to detect and measure fatigue cracks in actual turbine disks is being quantified. Some 50 engine disks, containing 500 bolt holes, were removed from service after at least one design life time and inspected with conventional and advanced eddy current systems. The inspection results were recorded on magnetic tape and analysed. All 500 holes were electropolished and replicated, and the replicas examined on the optical microscope to locate and measure the actual 847 surface cracks. Seven hundred metallographic sections were prepared on 56 cracks to fully characterize crack shapes and features.

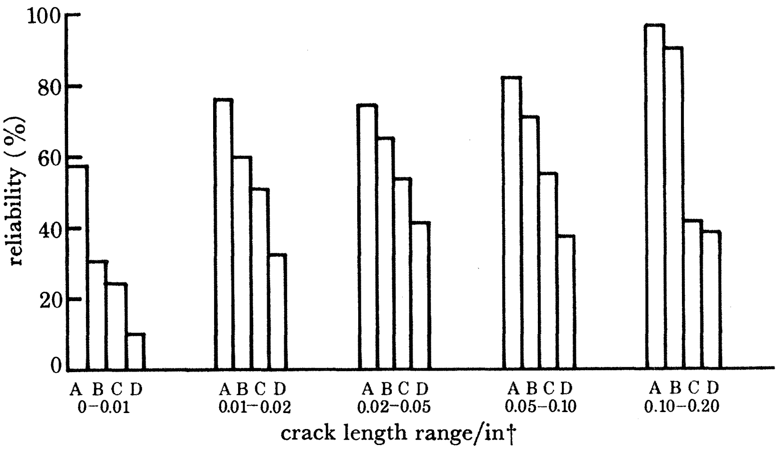

Figure 9. Inspection reliability as a function of actual crack size, and accuracy of the prediction required. Agreement criteria: A, indication; B, location (RI or RO); C, length (0.5–2 ×); D, position ($\pm 0.106 = \frac{1}{8}t$). † 1 in = 2.54 cm.

Figure 9 summarizes the inspection reliability results as a function of actual crack size range. The inspection reliability decreases significantly with decreasing crack size. The probability of accurately locating, measuring and positioning the defect is considerably lower than detection alone. Perhaps the most surprising observation is the relatively low inspection reliability, especially at small sizes. For example, the probability of correctly measuring a 0.015 in (0.037 cm) crack by current eddy current inspection systems after field exposure is only 30 %. However, an equally surprising conclusion is drawn from the r.f.c./c.a. simulations, which indicate that major cost savings are possible with this existing uncertainty.

Figure 10 summarizes the pre-inspection distribution of cracks in the same disk bolt hole population. The number of small cracks observed is very large compared with the number of larger cracks. These actual data reconfirm the importance of obtaining the actual pn(a) for input to equation (5) rather than assuming that the cracks are there if they might be missed by inspection.

Research is continuing to develop more efficient ways of estimating inspection reliability and pre-inspection material quality as well as of minimizing the impact of uncertainties through more extensive feedback and combined analysis of field data.

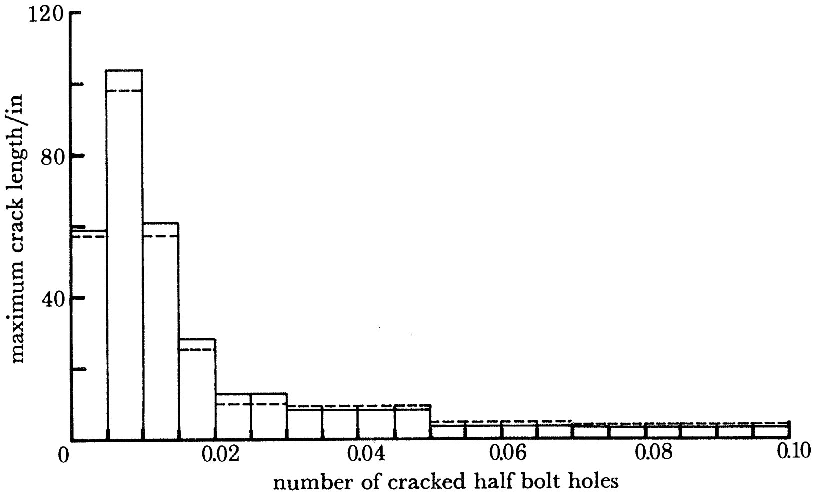

FIGURE 10. Pre-inspection material quality for disk bolt holes removed from service after more than one design lifetime. —, Raw data; – – –, 0.04 in linked.

SUMMARY AND CONCLUSIONS

Fracture mechanics methods that accurately evaluate the effect of actual or potential defects are now utilized in conjunction with traditional design analyses. The accuracy and applicable range of the fracture mechanics calculations continue to increase, and more extensive materials data become available each day. Nevertheless, uncertainties and errors remain in our input data and analyses, and the need to utilize probabilistic analysis techniques to assess realistically structural integrity is now accepted by many. Deterministic, 'worst case' assumptions are often both technically unrealistic and economically unacceptable. The application of probabilistic fracture mechanics (p.f.m.) does not necessarily make the component any more or less reliable, but it does quantify the impact of the uncertainties that actually exist and provides a quantitative basis for engineering decisions that must be made.

This paper describes the three approaches to assess accurately structural reliability. Extrapolation from the analysis of past data, if sufficient relevant experience exists, is a very accurate approach. When no field data exist, probabilistic engineering analysis (for example p.f.m.) can be used to quantify the structural reliability. The p.f.m. approach can be employed with any quantity of field data, but if insufficient materials and inspection performance data exist or inaccurate analyses are utilized, variability in the key parameters may be so large that the overall reliability results are unacceptable or inconclusive. The third approach, combined analysis, formally combines only the field data and engineering analyses most credible to the analyst to quantify structural integrity when neither the data base nor the engineering (p.f.m.) approaches alone can do so.

Two examples have been discussed, which illustrate these basic approaches. Combined analysis, particularly when utilized for life extension and retirement-for-cause, has been shown to identify inspection, repair and replacement strategies that can markedly lower total costs or improve reliability. Monte Carlo simulation techniques now exist that enable quantitative probabilistic analysis of various options so that optimum decisions can be made. The most surprising result with the combined analysis approach is that tremendous improvements are possible, even with very limited data and inaccurate analyses. No longer should engineers automatically claim 'insufficient data' as a reason for not employing a probabilistic analysis. It is precisely when sufficient uncertainties exist and data are limited that the probabilistic, combined analysis approach is most needed and very effective.

The authors acknowledge the technical leadership and creativity of the late Dr Alan S. Tetelman and coworkers, particularly K. Sorenson, D. Johnson, R. McCarthy and J. Finnegan, who performed some of the work described, and the Electric Power Research Institute, the Advanced Research Project Agency, and the National Wheel and Rim Association for their financial support.

References (Rau & Besuner)

Besuner, P. M. & Tetelman, A. S. 1977 *Nucl. Engng Design* **43**, 99.
Besuner, P. M., Sorenson, K. G. & Johnson, D. P. 1977 In *Fatigue life technology*, p. 53. A.S.M.E.
Besuner, P. M., Sorenson, K. G. & Johnson, D. P. 1978*a* Topical Report EPRI NP-855, part 1.
Besuner, P. M. *et al.* 1975 In *Proceedings of Risk–Benefit Methodology and Application Conference*, Asilomar, California (UCLA-Eng-7598).
Besuner, P. M. *et al.* 1978*b* Tropical Report EPRI NP-855, part 2.
Cohen, A. C. 1965 *Technometrics* **7**, 579–588.
Freudenthal, A. M., Garrelts, J. M. & Shinozuka, M. 1966 *ASCE struct. Div. J.* **92** (ST1).
Graham, T. W. & Tetelman, A. S. 1974 In *Proceedings, AIAA Conference*, Las Vegas.
Hill, R. J. *et al.* 1980 Air Force Material Laboratory Document. (Submitted.)
Johnson, D. P. 1977*a* *Nucl. Engng Design* **43**, 207.
Johnson, D. P. 1977*b* *Nucl. Engng Design* **43**, 219.
Mann, N. R., Schafer, R. E. & Singpurwalla, N. D. 1974 In *Methods for statistical analysis of reliability and life data*, pp. 419–465.
Packman, P. F. *et al.* 1969 *Metals Engng Q.*, pp. 52–62.
Rau Jr, C. A. 1978 Presented at ARPA/AFML Review of Progress in Quantitative NDE, San Diego, July.
Rau Jr, C. A. *et al.* 1977 *Nucl. Engng Design* **43**, 1.
Sorenson, K. G. & Besuner, P. M. 1977 In *Failure data and failure analysis: in power and processing industries* (ASME vol. PVP-PB-023), p. 77.
Starr, C. 1971 Presented at Colloquium on Benefit–Risk Relationship for Decision Making, Washington, D.C., April.
Tang, W. H. 1973 *J. Test. Eval.* **1**, 459.
Tetelman, A. S. 1977 In *Social consequences of engineering*, p. 122.
Tetelman, A. S. & Besuner, P. M. 1977 Presented at the Fourth International Conference on Fracture, Waterloo, Canada, June.
Thomas, J. M. *et al.* 1975 In *IEEE Reliability and Maintainability Symposium Proceedings*, pp. 255–262.
Whittaker, I. C. & Besuner, P. M. 1969 AFML TR-69-65. The Boeing Company.

Discussion

Sir Peter Hirsch, F.R.S. (*Department of Metallurgy and Science of Materials, Parks Road, Oxford, U.K.*). Probabilistic fracture mechanics seems to me to represent an important advance, in that risk can be quantified in terms of probability of failure, and in principle this could eventually replace safety factors. In practice, however, the application of the theory depends on availability of good input data; such data are scarce, and there is an urgent need to obtain reliable data on probability distributions of defects as a function of defect size, etc. For pressure vessels, the most severe types of failure may well be controlled by defects in the tails of the distribution curves, where the degree of uncertainty is particularly great. Thus, in practice, the use of safety factors will still be necessary.

C. A. RAU, JR AND P. M. BESUNER. We agree that there is an urgent need to obtain better probability distribution data for key variables, but we think that the advent of 'combined analysis' and related techniques, as described in the paper, justify our optimism in the applicability of probabilistic fracture mechanics with limited data. For example, the current absence of any 'infant mortality' failures due to large flaws in reactor pressure vessels can be used to bound the tail of the defect probability distribution. Combined analysis can take advantage of in-service experience to minimize the adverse impact of data limitations. For example, the model would clearly show the oft-neglected fact that for a static failure mode, the operation of 100 nominally identical pressure vessels, each for 1 year, provides far more information than the operation of one vessel for 100 years, We do not conceive of the elimination of the use of safety factors; we only propose that choice of the specific safety factor be established more quantitatively.

PHILOSOPHICAL TRANSACTIONS OF THE ROYAL SOCIETY A MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES