ELSEVIER

Fusion Engineering and Design 27 (1995) 191-197

Fusion Engineering and Design

Application of reliability analysis method to fusion component testing

Alice Y. Ying, Mohamed A. Abdou

Mechanical, Aerospace and Nuclear Engineering Department, University of California, Los Angeles, CA 90024-1597, USA

Abstract

The purpose of this paper is to discuss a basic reliability analysis method that is beneficial to fusion component test planning. Emphasis is placed on formulating the problem to be solved and on determining the required test time and the number of test articles needed to verify a performance or a failure criterion. While an increasing failure rate is likely to be the case for the fusion component resulting from irradiation effects, analysis is performed to show how the life characteristics impact the test requirements. Taking into account the available data from similar technology experiences, calculations based on the Bayesian approach indicate a possible saving on the test time requirement.

1. Introduction

The best way to measure system reliability is to test completed products or components, under conditions that simulate real life, until failure occurs. One simply cannot assess reliability without data, and of course, the more data available, the more confidence one will have in the estimated reliability level. Unfortunately, extensive testing is often considered undesirable because it results in expenditure of too much time and money. Thus, the need for "do-it-smarter" [1] consideration in developing the testing program becomes an important factor. The objectives of this paper are to utilize reliability concepts and analyses: (a) to determine meaningful guidelines for defining an effective and self-consistent test program for development of fusion components, (b) to serve as a tool for defining the testing requirements (on the major characteristics of the testing facility), and (c) to evaluate the need for fusion testing in large test volumes.

The testing process involves combining all sources of knowledge structurally, redefining and solving the problem as the test proceeds. Clearly, the goal is to produce the best possible fusion technology components, so that a high component reliability and subsequently a high plant availability can be achieved. In the development of a test plan, an understanding of the types of data and accuracy of the data to be obtained are required in order to achieve a reliability goal. Of course, this precision requirement determines how many test articles must be placed in the test and how long the test must be performed. If tests are directed toward the design risk areas, higher reliability can result from less test effort. Implementing the failure mode, effect and criticality analysis [2] allows the test efforts to be directed toward the highest risk areas. Furthermore, to maximize the benefits of the test efforts, emphasis should be placed on monitoring the component selection/design considerations, as opposed to running them and counting the number of failures. The real intent of the test is to

SSDI 0920-3796(94)00149-9

identify design weaknesses, flaws and failure modes and then fix them. This has led to a testing strategy for fusion nuclear components involving scoping, performance evaluation and reliability growth phases as presented in Ref. [3].

Major input to the preparation of the test plan can be gained from a review of available data from similar technology experience. Such data may help the designer to draw a reliability level at the design stage, to define a testing reliability goal, and may also pinpoint potential trouble areas. Moreover, this available information can be summarized in the prior information before the start of the experiment based on the Bayesian method [4]. The theory combines actual testing data with such prior information to get new estimates (believed to be more precise) or Bayesian confidence limits for the parameters with relatively little testing effort. Although a Bayesian analysis strongly depends on the validity of the model and prior distributions, analysis is performed here to show how the best-judged available information can be used to reduce the cumulative test time requirement.

2. Data requirements

2.1. Data needs for availability analysis

The minimum parameters of interest (such as in the case of exponential life distribution) for availability assessment are mean time between failures (MTBF) and mean down time (MDT). The accuracy of the availability analysis is a function of the quality of the data from the device being modeled; the precision is limited by the level of data available. These data obtained from non-fusion experience or estimated based on the judgement of experts familiar with the development of similar components are useful for rough estimates of plant availability (at the design stage) or, as is more commonly the case, for apportionment of the availability of each component to achieve the plant target availability. In most cases, the simulation models require some data elements in addition to MTBF and MDT, such as component failure characteristics and the associated probability density function, or a distribution model. (Often, one is forced to select a distribution model without having enough data actually to verify its appropriateness.) Building up these data banks is essential, as all the methods of the computation would be useless if we did not have at our disposal numerical values for the various parameters (failure rate, mean down time, etc.).

One difficult problem associated with the reliability measure from a practical standpoint is the selection of a distribution model. Unless one has considerable test data, it is difficult to determine whether the proper model is, for instance, Weibull, log normal, or gamma. A faulty assumption of failure rate characteristics can cause incorrect reliability predictions and interpretation of test results. A fusion component might have a decreasing failure rate during the beginning of its life (a typical behavior for commercial products) and an increasing failure rate as the time proceeds. Therefore, in addition to obtaining hard data to demonstrate component reliability, the test should also be designed to accommodate the component life characteristics.

2.2. Data needs for a probabilistic design methodology to design reliability

Special technical publications on the statistical aspects of material strength data have revealed, for example, that the probability distributions of the ultimate tensile, yield and endurance strengths of steel are normally distributed and that the strength properties of structural alloy materials often tend to follow a log normal distribution [5]. However, manufacturing processes such as heat treatment and surface finishing and temperature could alter this distribution function. For instance, the Weibull distribution is found for ferrous materials which have undergone heat treatment. Furthermore, for a given material with a certain surface finish the distribution function may be Weibull, but an alteration surface finish might cause the distribution to change to the largest extreme value distribution. On the other hand, the load (or stress) is by no means constant. Such variations might be due to the needs of different applications (such as load after operation under normal conditions), instantaneous transient force from plasma disruption, or short-term unexpected power exertion. These varied natures of strength (used to indicate any agency resisting failure) and stress (used to indicate any agency inducing failure) distributions disclose a new design methodology called probabilistic design [4,6].

For a given component that has a certain stress-resisting capacity, the concept of reliability is that, if the stress induced by the operating conditions exceeds this capacity (strength), failure results. In the probabilistic design approach, this reliability is determined by the strength and stress random variables. In the case of normal density functions, the reliability R is given as:

$$R = \frac{1}{\sqrt{2\pi}} \int_{-z_0}^{\infty} e^{-z^2/2} \mathrm{d}z \qquad (1)$$

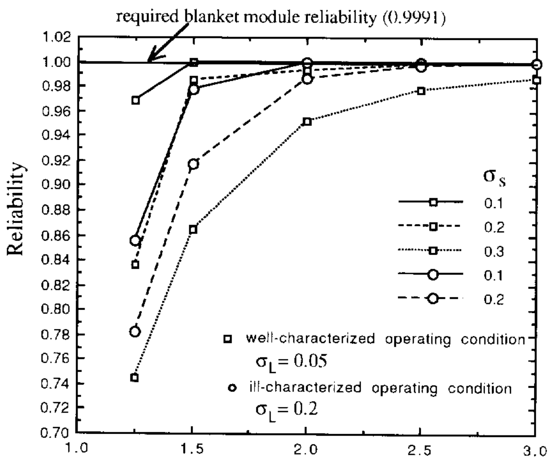

Fig. 1. Reliability as a function of the safety factor for different variations of strengths and loads.

where z_0 and z are defined as:

$$z_0 = \frac{\mu_y}{\sigma_y}; \quad z = \frac{y - \mu_y}{\sigma_y} \tag{2}$$

and

$$\mu_y = \mu_S - \mu_L \qquad \sigma_y = \sqrt{\sigma_S{}^2 + \sigma_L{}^2} \tag{3}$$

where μ_L is the mean value of the stress, μ_S is the mean value of the strength, σ_L is the standard deviation of the stress, and σ_S is the standard deviation of the strength. Fig. 1 shows that the variation in reliability is related to different magnitudes of variability in strength and stress random variables. Also shown in the figure is the required blanket module reliability for achieving an in-vessel blanket system reliability of 0.9 in the case where there are 120 modules in the blanket system [3].

For a well-characterized operating condition (such as $\sigma_L = 5\%$), the required standard deviation of the strength is about 10% for a safety factor of 1.5, and about 17.5% for a safety factor of 2.5 (where safety factor is defined as μ_S/μ_L). In contrast, for an ill-characterized operating condition (such as $\sigma_L = 20\%$), the accuracy requirement increases more significantly in the lower safety factor. Furthermore, the reliability can be limited if the safety factor is not adequate. The results show that, in order to achieve a blanket system reliability of 0.9, the minimum required safety factor is about 1.65 (100% accuracy requirement on the mean of the strength). It is evident that we need a high degree of accuracy, not only for the mean value but also for the standard deviation (as in the case of normal distribution) in order to arrive at a good estimate of the reliability. A more stringent requirement (a higher statistical significance) means that a larger sample size (a larger number of test articles) is needed in the test.

3. Failure modes, effect and criticality analysis (FMECA) implication

Identification of the failure modes and consequences at the levels of component, submodule, or even element (described as a constituent) have a profound effect on the reliability and maintainability of the plant. A given component (such as a neutral beam injector) may have a poor failure rate; however, this effect can be mitigated or nullified by incorporating redundancy into the design (which will have the disadvantage of reducing the breeding blanket coverage). Certain failure modes can be tolerated if the reactor can be operated in a degraded mode (lower power output) until the next available maintenance period. On the other hand, some failures (e.g. loss of the impurity control) will cause a chain of events that has a detrimental effect on many components. Through these failure modes, effect and criticality analysis processes, identified as FMECA [2,6], those critical constituents which contribute most to the plant failure can be recognized. An appropriate analysis consists of construction of an overall plant reliability block diagram as well as a quantitative calculation based on the constituents' failure rates. A severity factor indicating the seriousness of the effect can be added to criticality analysis to differentiate failures which result in plant unavailability from those which are tolerable. The FMECA process gives the designer insight into the reliability structure of a complex system, and underlines the comparative strengths and weaknesses of the various subassemblies of the system.

At the beginning stage of fusion testing, a major emphasis of the test should be given to developing critical constituents at power generating plant operating conditions. In particular, development testing of those constituents which are elements of a large and expansive component has significant benefits for the economic aspect. For example, a past study shows that blanket designs involving a large number of welds have the largest failure rates [7]. While welds are unavoidable in blanket designs, a scoping test and a reliability growth test are essential to increase blanket reliability. The scoping test can be formulated as a selection device with the goal of identifying the best manufacturing technique in conjunction with the best quality assurance process for the welds in the fusion environment. The reliability growth test requires an iterative procedure of analyze/fix/test, aiming to improve the optimum reliability of the concept.

4. Test duration

Test duration, such as cumulative test time, has been used as a test requirement for measuring a component's reliability at a specified confidence level. For example, if the Poisson distribution is used to design tests in which a component's MTBF is greater than that specified, the MTBF can be demonstrated at a confidence level of 80%. This leads to a total test time of 1.609 times the specified MTBF, assuming no failures occur during the test, or 2.994 times the specified MTBF if one failure occurs during the test. In this case, the components are assumed to experience a constant failure rate and, consequently, this required total test time could be spread between several test components. However, for a component following the Weibull distribution, the constant failure rate assumption is often inappropriate; hence, the test time cannot be spread arbitrarily between the components under test. The required test time per test unit under a test is estimated according to the Weibull scale parameter b and the sample size n as [8]

$$t = k\left[\frac{B_{r,c}}{n}\right]^{1/b} \quad (4)$$

where B is the Poisson distribution confidence factor at number of failures r, c is the confidence interval and k is the characteristic life factor, which is estimated as:

$$k = \frac{1}{\Gamma(1 + 1/b)} \quad (5)$$

where Γ is the gamma function.

A calculated summary of the individual test time requirements to achieve an 80% confidence level with different component failures as a function of sample size is shown in Fig. 2 for different shape factors. A shape factor of $b > 1$ indicates an increasing failure rate, $b < 1$ a decreasing failure rate (early-life failures) and $b = 1$ a constant failure rate, as shown in Fig. 3. (This ability to describe increasing or decreasing failure rates contributed to making the Weibull distribution popular for life data analysis.) The individual test times required for simultaneously testing 5 and 10 test articles are respectively about 0.77 and 0.45 times the MTBF (3.85 and 2.25 years for MTBF = 5 years), for a shape factor $b = 2$. If one failure occurs during the test, these individual test times increase to 1 and 0.6 × MTBF, respectively. The test time per test unit drops significantly as the sample size increases for a shape factor less than 1. This implies that, if a large sample size was used to detect the early-life characteristics, it would result in a tremendous saving in the test time required. Test time savings resulting from one additional sample decrease as b increases. When assessing wear-out concerns, it is normally better to test a smaller number of units to the approximate mean life usage than to test a larger number to only a portion of this usage.

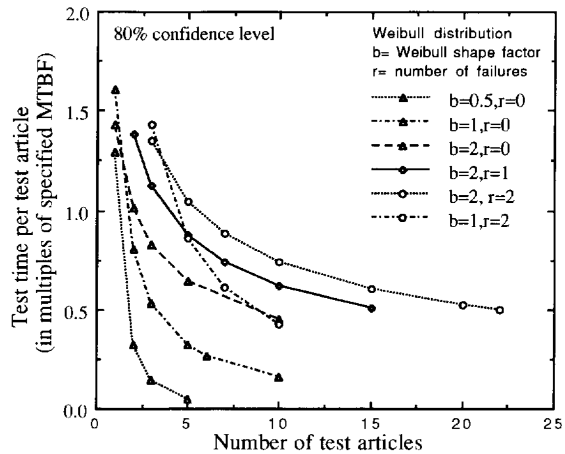

Fig. 2. Test time per test article as a function of number of test articles for different shape factors.

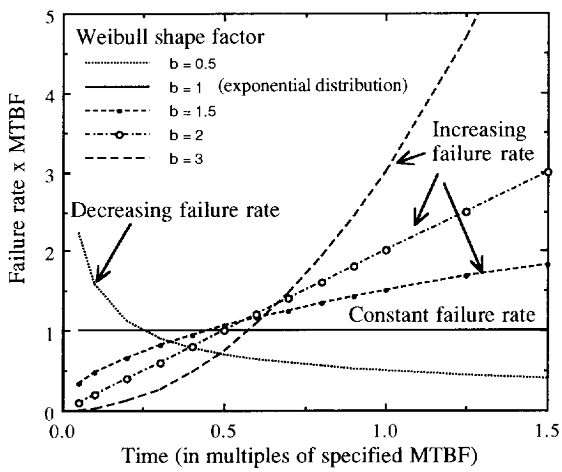

Fig. 3. Weibull life distribution failure rate characteristics for different shape factors.

5. Requirements on number of test articles (sample size)

In planning the experiments, test program planners are invariably confronted with the question: "What sample size do I need to verify a mean or a failure criterion?" Frequently, the sample size is determined by non-statistical considerations, such as limited budget or time, and by the number of units which are in service or are available for a test. When sample sizes are small, it is difficult to resolve whether the observed important differences are real (convincing). Furthermore, a small sample size makes the statistics too dependent on the precise value of a few individual observations or a low precision in estimating both sample mean and variance. In practice, one needs enough data (therefore, a large sample size) to produce both statistically convincing and practically significant observations.

There are several broad categories of problems for which systematic procedures have been developed to exercise control over sampling errors. Here, we focus on estimation and selection problems. Both are the basic test elements in fusion technology testing.

5.1. Estimation problems

This refers to problems in which we wish to deduce that the true but unknown value of a specified population parameter is contained within a bounded interval of given width. An example parameter of interest could be the mean of the primary stress intensity of a weld under irradiation, or the mean of the tritium breeding ratio of one particular blanket concept. In the estimation of parameters, sample sizes are selected to ensure satisfactory precision.

The sample size for estimating the mean θ within a factor of f with $100\gamma\%$ probability, assuming an exponential distribution, is approximated as [9]

$$n \approx [K\gamma/\ln(f)]^2 \qquad (6)$$

where $K\gamma$ is the $[100(1+\gamma)/2]$th standard normal percentile. Fig. 4 shows the required sample size for a range of f at different confidence levels. This simple approximation shows that a sample size of 50 is needed to obtain a percentage deviation of an unknown mean contained in the $\pm 20\%$ range at a confidence level of 90%, and of 11 for a confidence level of 75%. The required sample size is prohibitively large when a high precision is demanded ($n = 423$ for a mean to be estimated within 10%, with 95% probability). A more refined method given by Mace [10] shows that a slightly larger sample size is needed to achieve the same specified expected length of confidence intervals. As shown in the same figure, a sample size of 67 is needed, as compared to 50, for the mean to be estimated within 20% at a 90% confidence level.

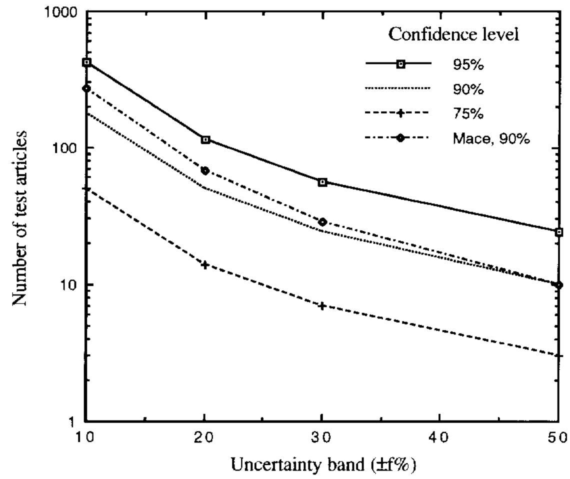

Fig. 4. Number of test articles required as a function of uncertainty band for different confidence levels.

5.2. Selection problems

Previous calculations indicate that 24 sample observations from each of three concepts should be made so that a best concept can be selected, for a ratio of 1.75 between the largest and the second largest means and a probability of correct selection of at least 95% [3,11]. According to the method proposed, the required sample size increases as the power of discrimination (the ratio of the largest mean to the second largest mean) decreases and becomes practically uneconomical (66 samples if the discrimination power is set at 1.4).

6. Application of Bayesian approach to testing

The Bayesian approach [4] takes account of additional information of a subjective nature, such as confidence in the manufacture or confidence in well-known older equipment from which the equipment under test has been derived. This subjective information is integrated with the objective data (testing data) by assigning a prior distribution to the parameters to be evaluated. This alleviates the necessity for a large amount of hard data, obtained by testing, in order to demonstrate a reliability level with a high degree of confidence. However, in general our subjective knowledge is somewhat vague, and it is impossible to specify a prior distribution very precisely. Furthermore, we have a large latitude in the selection of a priori distribution. The simplest case is using natural conjugates as a priori distribution, where the introduction of additional information does not modify the type of the distribution but simply alters the parameters of that distribution. In this case, the natural conjugate of an exponential distribution is a gamma distribution.

Recommended leakage failure rates based on non-fusion operating experience failure rates subjected to K-factor modification were published for various material and coolant combinations [12]. For example, a leakage (defined as coolant flow into the vacuum vessel sufficient to trigger a plasma disruption) failure rate value of 7.7×10^{-9} h^{-1} per meter tube with an error factor of three (defined as the square root of the 95% confidence failure rate divided by the 50% confidence failure rate) is suggested as a reference for use of preliminary ITER design and safety studies until further failure testing produces more accurate failure data. Here, the Bayesian approach is used to estimate the cumulative test time (in terms of tube length in meters times hours) required in a truncated sequential test to verify that the probability of the true failure rate falling within the range $7.7 \times 10^{-9}/3$ to $3 \times 7.7 \times 10^{-9}$ is equal to 95%. Notice that the test results can also be used to modify the estimated failure rates.

Fig. 5 shows the calculated required cumulative test time as a function of the number of failures for different shape parameters in the gamma distributions. The results show that the required cumulative test time (actual test time + scale parameter) ranges from 1.597×10^8 to 3.127×10^8 m h for zero to two failures, by assuming a shape factor equal to one (which gives a scale parameter of 9×10^7), and 2.41×10^8 to 3.8×10^8 m h for the same range of number of failures by assuming a shape factor of two (which gives a scale parameter of 2.2×10^8). It also shows that the test should be stopped, because of the inaccurate prior information, if the test results cross the termination point as shown in Fig. 5. The Bayesian estimate calculated from the test result at the test termination point is 1.8×10^{-8} h^{-1} per meter of tube length. Compared to the cases in which no prior information is considered, the required cumulative test times for demonstrating that the probability of the true failure rate being less than the Bayesian estimate equals 95% can be reduced by 15–180%. Furthermore, the analysis shows that the failure rate range as provided in Marshall and Cadwallader (1993) is more appropriately described by a gamma distribution with a shape factor equal to two.

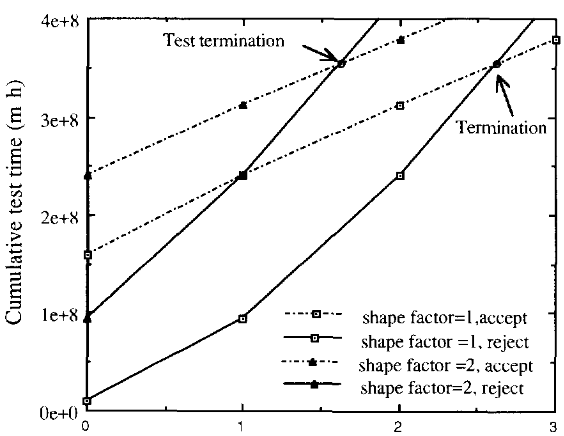

Fig. 5. Cumulative test time requirement for demonstrating the 95% probability of $7.7 \times 10^{-9}/3 <$ failure rate $< 3 \times 7.7 \times 10^{-9}$ in a Bayesian truncating sequential test as a function of number of failures for different gamma shape factors.

7. Summary

From the design reliability concept point of view, the magnitude of the variation plays an important role in determining the system reliability, particularly where the safety margin is not ample. Consequently, the test must be formulated not only to measure the mean value but to quantify the variation. A best component (or element) is chosen according to the largest mean load-resisting capacity combined with the smallest variation associated with this mean capacity. This implies that a large number of test articles in the test is required in order to assure a high accuracy.

Furthermore, the test must be constructed not only to count the failure rate but to detect component failure characteristics. This leads to a requirement for a longer test time per test article, when the failure rate increases as the operating time proceeds. An example calculation showed that each test article should be tested to about $0.45 \times$ MTBF for a Weibull shape factor of 2 to accept a specified MTBF at 80% confidence level, assuming a zero failure occurrence and a parallel test of 10 test articles. If one failure occurs during the test, individual test time increases to $0.6 \times$ MTBF.

The number of test articles to be placed in the test depends on the precision level of the parameter to be estimated, the confidence level requirement, testing objectives, and the component life characteristics. If the test is to ensure with a confidence level of 90% that the magnitude of the uncertainty in the mean (such as the tritium breeding ratio of one particular blanket concept) is to be contained within ±20%, the required number of test articles is about 50 (for Poisson distribution). On the other hand, it requires a test involving 24 test articles per concept in order to select the best concept within three design options with 95% confidence and a discrimination power of 1.75. A lower discrimination power (i.e. a small difference between different concepts) requires a larger sample size.

If it were possible to collect data from currently available fusion (or fission) technology experience, it would indeed be very useful in saving test time based on the Bayesian approach. An example calculation showed that a maximum saving at a factor of about two is possible to demonstrate an available stated water leakage failure rate.

Acknowledgement

This work was performed under US Department of Energy Contract DE-F603-86ER52123.

References

[1] F.W. Breyfogle III, Statistical Methods for Testing, Development, and Manufacturing, Wiley, 1992, Chapter 2.
[2] R. Bünde, Reliability and availability issues in NET, Fusion Eng. Des., 11 (1989) 139–150.
[3] M. Abdou, M. Peng, A. Ying and M. Tillack, Requirements and design envelope for volumetric neutron source fusion facilities for fusion nuclear technology development, IAEA Meeting, University of California, Los Angeles, 1993.
[4] K.C. Kapur and L.R. Lamberson, Reliability in Engineering Design, Wiley, New York, 1977, Chapter 13.
[5] American Society for Metals, Metals Handbook, Properties and Selection, Vol. 1, 8th edn., 1969.
[6] A.D.S. Carter, Mechanical Reliability, A Halsted Press Book, Wiley, New York, 2nd edn., 1986.
[7] R. Bünde, S. Fabritsiev and V. Rybin, Reliability of welds and brazed joints in blankets and its influence on availability, Fusion Eng. Des., 16 (1991) 59–72.
[8] Ref. [1], Chapter 8.
[9] W. Nelson, Applied Life Data Analysis, Wiley, 1982, Chapter 6.
[10] Mace, Sample Size Determination, Reinhold, 1964, Chapter 5.
[11] Ref. [10], Chapter 7.
[12] T.D. Marshall and L.C. Cadwallader, Recommended in-vessel tubing failure rates for the International Thermonuclear Experimental Reactor, 1993.