JOURNAL OF AIRCRAFT
Vol. 43, No. 1, January–February 2006

Structural Safety Measures for Airplanes

Erdem Acar,* Amit Kale,* and Raphael T. Haftka†
University of Florida, Gainesville, Florida 32611-6250
and
W. Jefferson Stroud‡
NASA Langley Research Center, Hampton, Virginia 23681-2199

Passenger aircraft structural design is based on a safety factor of 1.5, and this safety factor alone is equivalent to a probability of failure of between 10^{-2} and 10^{-3}. Yet airliners are much safer, with crashes caused by structural failure being extremely rare based on accident records. The probability of structural failure of transport aircraft is of the order of 10^{-8} per flight segment. This paper looks at two additional contributions to safety—the use of conservative material properties and certification tests—using a simple model of structural failure. We find that the three safety measures together might be able to reduce the calculated probability of failure to about 10^{-7}. Additional measures, such as conservative load specifications, might be responsible for the higher safety encountered in practice, explaining why passenger aircraft are so structurally safe. In addition, the paper sheds light on the effectiveness of certification tests for improving safety. It is found that certification tests reduce the calculated failure probabilities by reducing the modeling error. We find that these tests are most effective when safety factors are low and when most of the uncertainty is caused by systemic errors rather than variability.

I. Introduction

IN the past few years, there has been growing interest in applying probability methods to aircraft structural design (e.g., Refs. 1–4). However, many engineers are skeptical of our ability to calculate the probability of failure of structural designs for the following reasons. First, data on statistical variability in material properties, geometry, and loading distributions are not always available in full (e.g., joint distributions), and it has been shown that insufficient information can lead to large errors in probability calculations (e.g., Refs. 5 and 6). Second, the magnitude of errors in calculating loads and predicting structural response is not known precisely, and there is no consensus on how to model these errors in a probabilistic setting. As a result of these concerns, it is possible that transition to probability-based design will be gradual. In such circumstances it is important to understand the impact of existing design practices on safety. This paper is a first attempt to explore the effects of various safety measures taken during aircraft structural design using the deterministic design approach based on Federal Aviation Administration (FAA) regulations.

The safety measures that we include here are 1) the use of safety factors, 2) the use of conservative material properties (A-basis), and 3) the use of final certification tests. These safety measures are representative rather than all inclusive. For example, the use of A-basis properties is a representative measure for the use of conservative material properties. We do not include in this discussion the additional safety caused by structural redundancy and caused by conservative design load specification. The use of A-basis property rather than B-basis is because we did not include redundancy. FAA suggests that (FAR 25.613) when there is a single failure path, A-basis properties should be employed, but in case of multiple failure paths B-basis properties are to be used. The effect of the three individual safety measures and their combined effect on the probability of structural failure of the aircraft are demonstrated. We use Monte Carlo simulations to calculate the effect of these safety measures on the probability of failure of a structural component.

We start with a structural design employing all considered safety measures. The effects of variability in geometry, loads, and material properties are readily incorporated by the appropriate random variables. However, there is also uncertainty because of various errors such as modeling errors in the analysis. These errors are fixed but unknown for a given airplane. To simulate these epistemic uncertainties, we transform the error into a random variable by considering the design of multiple aircraft models. As a consequence, for each model the structure is different. It is as if we pretend that there are hundreds of companies (Airbus, Boeing, etc.), each designing essentially the same airplane, but each having different errors in their structural analysis. This assumption is only a device to model lack of knowledge or errors in probabilistic setting. However, pretending that the distribution represents a large number of aircraft companies helps to motivate the probabilistic setting.

For each model we simulate certification testing. If the airplane passes the test, then an entire fleet of airplanes with the same design is assumed to be built with different members of the fleet having different geometry, loads, and material properties based on assumed models for variability in these properties. That is, the uncertainty caused by variability is simulated by considering multiple realizations of the same design, and the uncertainty caused by errors is simulated by designing different structures to carry the same loads.

We consider only stress failure caused by extreme loads, which can be simulated by an unstiffened panel designed under uniaxial loads. No testing of components prior to certification is analyzed for this simple example.

II. Structural Uncertainties

A good analysis of different sources of uncertainty is provided by Oberkampf et al.[7] Here we simplify the classification, with a view to the question of how to control uncertainty. We propose in Table 1 a classification that distinguishes between 1) uncertainties that apply equally to the entire fleet of an aircraft model and 2) uncertainties that vary for the individual aircraft. The distinction is important because safety measures usually target one or the other. Although type 2) are random uncertainties that can be readily modeled probabilistically, type 1 are fixed for a given aircraft model (e.g., Boeing 737-400), but they are largely unknown.

Presented as Paper 2004-1629 at the AIAA/ASME/ASCE/AHS/ASC 45th Structures, Structural Dynamics, and Materials Conference, Palm Springs, CA, 19–22 April 2004; received 3 November 2004; accepted for publication 16 May 2005.

*Graduate Research Assistant, Department of Mechanical and Aerospace Engineering. Student Member AIAA.

†Distinguished Professor, Department of Mechanical and Aerospace Engineering. Fellow AIAA.

‡Aerospace Engineer. Associate Fellow AIAA.

Table 1 Uncertainty classification

Type of uncertainty	Spread	Cause	Remedies
Systemic error (modeling errors)	Entire fleet of components designed using the model	Errors in predicting structural failure and differences between properties used in design and average fleet properties	Testing and simulation to improve math model and the solution
Variability	Individual component level	Variability in tooling, manufacturing process, and flying environments	Improve tooling and construction Quality control

That is, the uncertainty in the failure of a structural member can also be divided into two types: systemic errors and variability. Systemic errors reflect inaccurate modeling of physical phenomena, errors in structural analysis, errors in load calculations, or use of materials and tooling in construction that are different from those specified by the designer. Systemic errors affect all of the copies of the structural components made and are therefore fleet-level uncertainties. They can reflect differences in analysis, manufacturing, and operation of the aircraft from an ideal. The ideal aircraft is an aircraft designed assuming that it is possible to perfectly predict structural loads and structural failure for a given structure, that there are no biases in the average material properties and dimensions of the fleet with respect to design specifications, and that there exists an operating environment that on average agrees with the design specifications. The other type of uncertainty reflects variability in material properties, geometry, or loading between different copies of the same structure and is called here individual uncertainty.

III. Safety Measures

Aircraft structural design is still done, by and large, using code-based design rather than probabilistic approaches. Safety is improved through conservative design practices that include use of safety factors and conservative material properties. It is also improved by tests of components and certification tests that can reveal inadequacies in analysis or construction. In the following we detail some of these safety measures:

Safety margin: Traditionally all aircraft structures are designed with a safety factor to withstand 1.5 times the limit load without failure.

A-basis properties: To account for uncertainty in material properties, the FAA recommends the use of conservative material properties. This is determined by testing a specified number of coupons selected at random from a batch of material. The A-basis property is determined by calculating the value of a material property exceeded by 99% of the population with 95% confidence.

Component and certification tests: Component tests and certification tests of major structural components reduce stress and material uncertainties for given extreme loads caused by inadequate structural models. These tests are conducted in a building-block procedure. First, individual coupons are tested, and then a subassembly is tested followed by a full-scale test of the entire structure. Because these tests cannot apply every load condition to the structure, they leave uncertainties with respect to some loading conditions. It is possible to reduce the probability of failure by performing more tests to reduce uncertainty or by extra structural weight to reduce stresses. If certification tests were designed together with the structure, it is possible that additional tests would become cost effective because they would allow reduced structural weight.

We simulate the effect of these three safety measures by assuming the statistical distribution of the uncertainties and incorporating them in approximate probability calculations and Monte Carlo simulation. For variability the simulation is straightforward. However, although systemic errors are uncertain at the time of the design, they will not vary for a single structural component on a particular aircraft. Therefore, to simulate the uncertainty, we assume that we have a large number of nominally identical aircraft being designed (e.g., by Airbus, Boeing, Bombardier, etc.), with the errors being fixed for each aircraft. This creates a two-level Monte Carlo simulation, with different aircraft models being considered at the upper level and different instances of the same aircraft at the lower level.

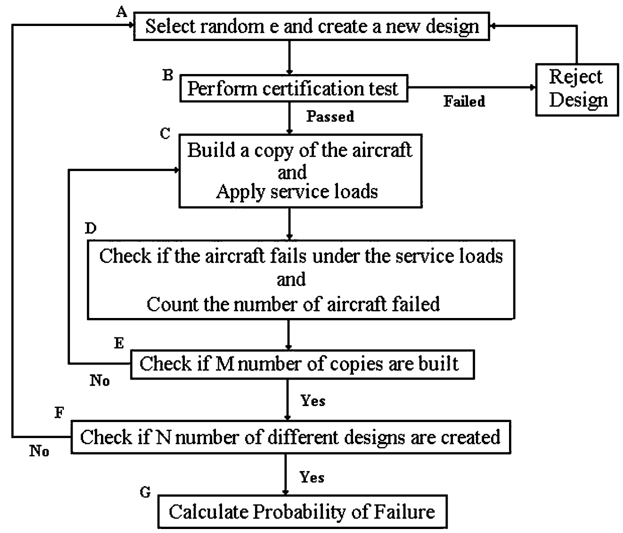

Fig. 1 Flowchart for Monte Carlo simulation of panel design and failure.

To illustrate the procedure, we consider a simple example of an unstiffened panel designed for strength under uniaxial tensile loads. This will still simulate reasonably well more complex configurations, such as stiffened panels subject to stress constraints. Aircraft structures have more complex failure modes, such as fatigue and fracture, that require substantially different treatment and the consideration of the effects of inspections (see Ref. 8). However, this simple example serves to further our understanding of the interaction between various safety measures. The procedure is summarized in Fig. 1, which is described in detail in the next section.

IV. Panel Example Definition

A. Design and Certification Testing

We assume that we have N different aircraft models, that is, we have N different companies producing a model with systemic errors. We consider a generic panel to represent the entire aircraft structure. The true stress σ_{true} is found from the equation

$$\sigma_{\text{true}} = P/wt \quad (1)$$

where P is the applied load on the panel of width w and thickness t. In a more general situation, Eq. (1) can apply to a small element in a more complex component.

When errors are included in the analysis, the true stress in the panel is different from the calculated stress. We include the errors by introducing an error factor e while computing the stress as

$$\sigma_{\text{calc}} = (1+e)\sigma_{\text{true}} \quad (2)$$

Positive values of e yield conservative estimates of the true stress, and negative values yield unconservative stress estimation. The other random variables account for variability. Combining Eqs. (1) and (2), the stress in the panel is calculated as

$$\sigma_{\text{calc}} = (1+e)(P/wt) \quad (3)$$

The design thickness is determined so that the calculated stress in the panel is equal to material allowable stress for a design load P_d multiplied by a safety factor S_F; hence, the design thickness of the panel is calculated from Eq. (3) as

$$t_{\text{design}} = (1+e)(S_F P_d / w_{\text{design}} \sigma_a) \quad (4)$$

where the design panel width w_{design} is taken here to be 1.0 m and σ_a is the material stress allowable obtained from testing a batch of coupons according to procedures that depend on design practices. Here, we assume that A-basis properties are used (Appendix A). During the design process, the only random quantities are σ_a and e. The thickness obtained from Eq. (4) (step A in Fig. 1) is the nominal thickness for a given aircraft model. The actual thickness will vary because of individual–level manufacturing uncertainties.

After the panel has been designed (that is, thickness determined) from Eq. (4), we simulate certification testing for the aircraft. Here we assume that the panel will not be built with complete fidelity to the design due to variability in geometry (width and thickness). The panel is then loaded with the design axial force of (S_F times P_d), and the stress in the panel is recorded. If this stress exceeds the failure stress (itself a random variable; see Table 2.) then the design is rejected; otherwise, it is certified for use. That is, the airplane is certified (step B in Fig. 1) if the following inequality is satisfied:

$$\sigma - \sigma_f = S_F P_d / wt - \sigma_f \leq 0 \tag{5}$$

and we can build multiple copies of the airplane. We subject the panel in each airplane to actual random maximum (over a lifetime) service loads (step D) and decide whether it fails using Eq. (6):

$$P \geq R = tw\sigma_f \tag{6}$$

Here, P is the applied load, and R is the resistance or load capacity of the structure in terms of the random width w, thickness t, and failure stress σ_f. A summary of the distributions for the random variables used in design and certification is listed in Table 2.

This procedure of design and testing is repeated (steps A and B) for N different aircraft models. For each new model, a different random error factor e is picked for the design, and different allowable properties are generated from coupon testing (Appendix A). Then in the testing, different thicknesses and widths, and different failure stresses are generated at random from their distributions.

Table 2 Distribution of random variables used for panel design and certification

Variables	Distribution	Mean	Scatter
Plate width w	Uniform	1.0	(1%) bounds
Plate thickness t	Uniform	t_{design}	(3%) bounds
Failure stress σ_f	Lognormal	150.0	10% coefficient of variation
Service load P	Lognormal	100.0	10% coefficient of variation
Error factor e	Uniform	0.0	10% to 50%

B. Effect of Certification Tests on Distribution of Error Factor *e*

One can argue that the way certification tests reduce the probability of failure is by changing the distribution of the error factor e. Without certification testing, we assume symmetric distribution of this error factor. However, designs based on unconservative models are more likely to fail certification, and so the distribution of e becomes conservative for structures that pass certification. To quantify this effect, we calculated the updated distribution of the error factor e. The updated distribution is calculated analytically by Bayesian updating by making some approximations, and Monte Carlo simulations are conducted to check the validity of those approximations.

Bayesian updating is a commonly used technique to obtain updated (or posterior) distribution of a random variable upon obtaining new information about the random variable. The new information here is that the panel has passed the certification test.

Using Bayes' theorem, the updated (posterior) distribution $f^U(\theta)$ of a random variable θ is obtained from the initial (prior) distribution $f^I(\theta)$ based on new information as

$$f^U(\theta) = \frac{P(\in|\theta) f^I(\theta)}{\int_{-\infty}^{\infty} P(\in|\theta) f^I(\theta)\,\mathrm{d}\theta} \tag{7}$$

where $P(\in|\theta)$ is the conditional probability of observing the experimental data $\in$ given that the value of the random variable is θ.

For our case, the posterior distribution $f^U(e)$ of the error factor e is given as

$$f^U(e) = \frac{P(C|e)\, f^I(e)}{\int_{-b}^{b} P(C|e) f^I(e)\,\mathrm{d}e} \tag{8}$$

where C is the event of passing certification and $P(C|e)$ is the probability of passing certification for a given e. Initially, e is assumed to be uniformly distributed. The procedure of calculation of $P(C|e)$ is described in Appendix B, where we approximate the distribution of the geometrical variables t and w as log-normal, taking advantage of the fact that their coefficient of variation is small compared to that of the failure stress (see Table 2).

We illustrate the effect of certification tests for the panels designed with A-basis material properties. An initial and updated distribution plot of error factor e with 50% bound is shown in Fig. 2. Monte Carlo simulation with 50,000 aircraft models is also shown. Figure 2 shows that the certification tests greatly reduce the probability of negative error, hence eliminating most unconservative designs. As seen from the figure, the approximate distribution calculated by the analytical approach matches well the distribution obtained from Monte Carlo simulations.

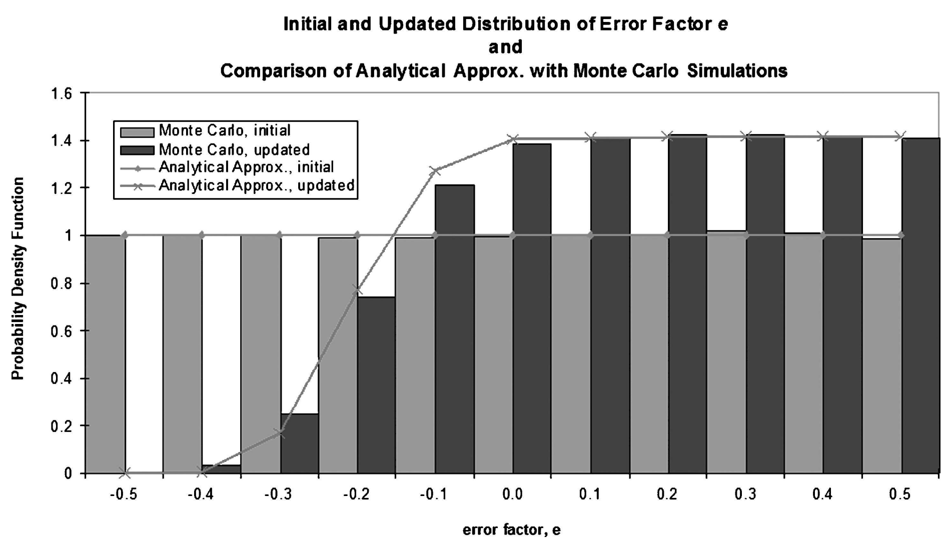

Fig. 2 Initial and updated probability distribution functions of error factor *e*. Error bound is 50% and Monte Carlo simulation done with a sample of 50,000.

C. Probability of Failure Calculation by Analytical Approximation

The stress analysis represented by Eq. (1) is trivial, so that the computational cost of Monte Carlo simulation of the probability of failure is not high. However, it is desirable to obtain also analytical probabilities that can be used for more complex stress analysis and to check the Monte Carlo simulations.

To take advantage of simplifying approximations of the distribution of the geometry parameters, it is convenient to perform the probability calculation in two stages, corresponding to the inner and outer loops of Fig. 1. That is, we first obtain expressions for the probability of failure of a single aircraft model (that is, given e and allowable stress). We then calculate the probability of failure over all aircraft models.

The mean value of the probability of failure over all aircraft models is calculated as

$$\hat{P}_f = \int P_f(t_{\text{design}}) f(t_{\text{design}})\, \mathrm{d}t_{\text{design}} \tag{9}$$

where t_{design} is the nondeterministic distribution parameter and $f(t_{\text{design}})$ is the probability density function of parameter t_{design}.

It is important to have a measure of variability in this probability from one aircraft model to another. The standard deviation of failure probability gives a measure of this variability. In addition, it provides information on how accurate is the probability of failure obtained from Monte Carlo simulations. The standard deviation can be calculated from

$$\sigma_{P_f} = \left\{ \int [P_f(t_{\text{design}}) - \hat{P}_f]^2 f(t_{\text{design}})\, \mathrm{d}t_{\text{design}} \right\}^{\frac{1}{2}} \tag{10}$$

D. Probability of Failure Calculation by Monte Carlo Simulations

The inner loop in Fig. 1 (steps C–E) represents the simulation of a population of M airplanes (hence panels) that all have the same design. However, each panel is different because of variability in geometry, failure stress, and loading (step D). We subject the panel in each airplane to actual random maximum (over a lifetime) service loads (step E) and calculate whether it fails using Eq. (6).

For airplane models that pass certification, we count the number of panels failed. The failure probability is calculated by dividing the number of failures by the number of airplane models that passed certification, times the number of copies of each model.

The analytical approximation for the probability of failure suffers because of the approximations used, whereas the Monte Carlo simulation is subject to sampling errors, especially for low probabilities of failure. Using large samples, though, can reduce the latter. Therefore, we compared the two methods for a relatively large sample of 1000 aircraft models with 100,000 instances of each model. In addition, the comparison is performed for the case where mean material properties (rather than A-basis properties) are used for the design, so that the probability of failure is high enough for the Monte Carlo simulation to capture it accurately. Table 3 shows the results for this case.

The last column of Table 3 shows the percent error of the analytical approximation compared to Monte Carlo simulations. It is seen that the analytical approximation is in good agreement with the values obtained through Monte Carlo simulations. Further investigations with larger Monte Carlo simulations (100,000 samples) showed that these differences are mostly due to limited Monte Carlo sampling. It is remarkable that the standard deviation of the probability of failure is almost twice the average value of the probability (the ratio, the coefficient of variation, is about 170%) before certification and about six times larger after certification. This indicates huge variability in the probability of failure for different aircraft models, and this is because of the large error bound $e = 50\%$. With 1000 different aircraft models N, the standard deviation in the Monte Carlo estimates is about 3%, and the differences between the Monte Carlo simulation and the analytical approximation are of that order.

V. Effect of Three Safety Measures on Probability of Failure

We next investigate the effect of other safety measures on failure probability of the panels using Monte Carlo simulations. We performed the simulation for a range of variability in error factor e for 500 airplane models (N samples in outer loop) and 20,000 copies of each airplane model (M samples in inner loop). Here, we compare the probability of failure of a structure designed with three safety measures (safety factor, conservative material property, and certification testing) to that of a structure designed without safety measures.

Table 4 presents the results when all safety measures are used for different bounds on the error. The second column shows the mean and standard deviation of design thicknesses generated for panels that passed certification. These panels correspond to the outer loop of Fig. 1. The variability in design thickness is caused by the

Table 3 Comparison of probability of failures P_f for panels designed using safety factor of 1.5, mean value for allowable stress, and error bound of 50%

Value	Analytical approximation	Monte Carlo simulation[a]	% error
Average value of P_f without certification P_{nt}	1.741×10^{-1}	1.787×10^{-1}	2.6
Standard deviation of P_{nt}	3.006×10^{-1}	3.035×10^{-1}	1.0
Average value of P_f with certification P_t	1.010×10^{-3}	1.094×10^{-3}	7.6
Standard deviation of P_t	6.16×10^{-3}	5.622×10^{-3}	9.6
Average value of initial error factor e^i	0.0000	−0.0058	——
Standard deviation of e^i	0.2887	0.2876	0.4
Average value of updated error factor e^{up}	0.2444	0.2416	1.2
Standard deviation of e^{up}	0.1578	0.1546	2.1

[a] $N = 1000$ and $M = 100{,}000$ is used in the Monte Carlo simulations.

Table 4 Probability of failure for different bounds on error e for panels designed using safety factor of 1.5 and A-basis property for allowable stress

Error bound e, %	Average design thickness after certification[a,b]	Certification failure rate, %	Probability of failure after certification $P_t \times 10^{-4}$[b]	Probability of failure without certification $P_{\text{nt}} \times 10^{-4}$[b]	Probability ratio P_t/P_{nt}	Probability difference $P_{\text{nt}} - P_t$
50	1.453 (0.18)	30.3	6.04 (8.30)	447 (2.63)	1.35×10^{-2}	4.41×10^{-2}
40	1.388 (0.17)	21.4	5.74 (7.01)	96.3 (2.78)	5.96×10^{-2}	9.06×10^{-3}
30	1.327 (0.15)	16.2	3.76 (2.91)	12.7 (2.65)	2.96×10^{-1}	8.94×10^{-4}
20	1.286 (0.11)	8.4	0.972 (2.55)	1.14 (2.24)	8.53×10^{-1}	1.68×10^{-5}
10	1.273 (0.06)	1.5	0.101 (1.53)	0.117 (1.50)	8.66×10^{-1}	1.57×10^{-6}

[a] Average over $N = 500$ models. Average design thickness without certification is 1.271.
[b] Numbers in parenthesis denote the coefficient of variation of the quantity, as obtained from the analytical approximation.

randomness in the error e and in the stress allowable. The average thickness before certification was 1.271, so that the column shows the conservative effect of certification testing. When the error bound is 10%, 98.5% of the panels pass certification (third column in Table 4), and the average thickness is increased by only 0.16% because of the certification process. On the other hand, when the error bound is 50%, 30% of the panels do not pass certification, and this raises the average thickness to 1.453. Thus, the increase in error bound has two opposite effects. Without certification testing, increasing the error bound greatly increases the probability of failure. For example, when the error bound changes from 30 to 50%, the probability of failure without certification changes from 0.00127 to 0.0447, or by a factor of 35. On the other hand, with the increased average thickness, after certification the probability increases only from 3.76×10^{-4} to 6.04×10^{-4}.

The effectiveness of the certification tests can be expressed by two measures of probability improvement. The first measure is the ratio of the probability of failure with the test P_t to the probability of failure without tests P_{nt}. The second measure is the difference of these probabilities. The ratio is a more useful indicator for low probabilities of failure, whereas the difference is more meaningful for high probabilities of failure. However, when P_t is high, the ratio can mislead. That is, an improvement from a probability of failure of 0.5 to 0.1 is more substantial than an improvement in probability of failure of 0.1 to 0.01 because it "saves" more airplanes. However, the ratio is more useful when the probabilities are small, and the difference is not very informative.

Table 4 shows that certification testing is more important for large error bounds e. For these higher values the number of panels that did not pass certification is higher, thereby reducing the failure probability for those that passed certification. Although the effect of component tests (building block tests) is not simulated, their main effect is to reduce the error magnitude e. This is primarily because of the usefulness of component tests in improving analytical models and revealing unmodeled failure modes. With that in mind, we note that the failure probability for the 50% error range is 6.0×10^{-4}, and it reduces to 1.0×10^{-5} for the 10% error range, that is, by a factor of 60.

The actual failure probability of aircraft panels is expected to be of the order of 10^{-8} per flight, much lower than the best number in the fourth column of Table 4. However, the number in Table 4 is for a lifetime for a single structural component. Assuming about 10,000 flights in the life of a panel and 100 independent structural components, this 10^{-5} failure probability for a panel will translate to a per flight probability of failure of 10^{-7} per airplane. This factor of 10 discrepancy is exacerbated by other failure modes like fatigue that have not been considered. However, other safety measures, such as conservative load specifications, can account for this discrepancy.

Table 5 shows results when average rather than conservative material properties are used. It can be seen from Table 5 that the average thickness determined using the mean value of allowable stress is lower than that determined using the A-basis value of allowable stress (Table 4). This is equivalent to adding an additional safety factor over an already existing safety factor of 1.5. For the distribution considered in this paper, a typical value of the safety factor caused by A-basis property is around 1.27.

Without the A-basis properties, the stress in the certification test is approximately equal to the average ultimate service stress, so that about 50% of the panels fail certification. When the errors are large, this raises substantially the average thickness of the panels that pass certification, so that for an error bound of 50% the certification test is equivalent to a safety factor of 1.244. Large errors produce some super-strong and some super-weak panels (see Fig. 3b). The super-weak panels are mostly caught by the certification tests, leaving the super-strong panels to reduce the probability of failure. Another way of looking at this effect is to note that when there are no errors, there is no point to the tests. Indeed, it can be seen that the probability of failure without certification tests improves with reduced error bound e, but that the reduced effect of the certification tests reverses the trend. Thus for this case we obtain the counterintuitive results that larger errors produce safer designs.

Comparing the first row of Table 5 to Table 3, we see the effect of the smaller sample for the Monte Carlo simulations. Table 3 was obtained with 1000 models and 100,000 copies per model, whereas Table 5 was obtained with 500 models, and 20,000 copies per model. The difference in the probability of failure after certification between the two tables is about 14%. However, the two values straddle the analytical approximation.

The effects of building-block types of tests that are conducted before certification are not included in this study. These tests reduce the errors in analytical models. For instance, if there is 50% error in the analytical model the building-block type of tests can reduce this error to lower values. Hence, the difference between the rows of Table 4 can be viewed as indicating the benefits of reducing the error by building-block tests.

Table 6 shows the effect of not using a safety factor. Although certification tests improve the reliability, again in a general trend of high improvement with high error, the lack of safety factor of 1.5 limits the improvement. Comparing Tables 4 and 6, it can be seen that the safety factor reduces the probability of failure by two to three orders of magnitudes. It is interesting to note that the effect of the error bound on the probability of failure after certification is not monotonic. Certification is still done at the same loads, but the test airplane is designed for higher loads because of the conservative material properties. For high errors, many airplanes will

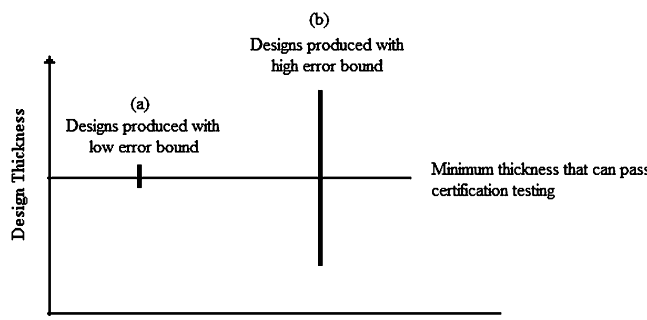

Fig. 3 Design thickness variation with low and high error bounds. Note that after certification testing only the designs above the minimum thickness are built and flown. Those on the right have a much higher average design thickness than those on the left.

Table 5 Probability of failure for different bounds on error *e* for panels designed using safety factor of 1.5 and mean value for allowable stress

Error bound e, %	Average design thickness after certification[a,b]	Certification failure rate, %[c]	Probability of failure after certification $P_t \times 10^{-4}$[b]	Probability of failure without certification $P_{\mathrm{nt}} \times 10^{-4}$[b]	Probability ratio P_t/P_{nt}	Probability difference $P_{\mathrm{nt}} - P_t$
50	1.244 (0.13)	50.1	9.44 (6.10)	1780 (1.73)	5.32×10^{-3}	1.77×10^{-1}
40	1.192 (0.11)	51.5	10.9 (5.43)	1060 (1.86)	1.02×10^{-2}	1.05×10^{-1}
30	1.137 (0.09)	52.1	15.1 (4.59)	451 (1.86)	3.36×10^{-2}	4.36×10^{-2}
20	1.080 (0.08)	52.9	22.9 (3.05)	142 (1.64)	1.61×10^{-1}	1.19×10^{-2}
10	1.025 (0.05)	47.0	27.1 (1.34)	41.8 (1.08)	6.48×10^{-1}	1.47×10^{-3}

[a]Average over $N = 500$ models.
[b]Numbers in parenthesis denote the coefficient of variation of the quantity, as obtained from the analytical approximation.
[c]With only 500 models, the standard deviation in the certification failure rate is about 2.2%. Thus, all of the numbers in this column are about 50%, as can be expected when mean material properties are used. Average design thickness without certification is 1.000.

Table 6 Probability of failure for different bounds on error *e* for safety factor of 1.0 and A-basis property for allowable stress

Error bound e, %	Average design thickness after certification[a,b,c]	Certification failure rate, %	Failure probability after certification $P_t \times 10^{-2}$ [b]	Failure probability with no certification $P_{nt} \times 10^{-2}$ [b]	Probability ratio P_t/P_{nt}	Probability difference $P_{nt} - P_t$
50	0.968 (0.18)	52.3	6.76 (2.00)	25.8 (1.29)	2.62×10^{-1}	1.91×10^{-1}
40	0.925 (0.17)	24.0	10.0 (1.81)	23.7 (1.33)	4.22×10^{-1}	1.37×10^{-1}
30	0.885 (0.15)	16.0	10.8 (1.58)	18.6 (1.33)	5.82×10^{-1}	7.50×10^{-2}
20	0.857 (0.11)	7.5	10.1 (1.25)	11.7 (1.18)	8.60×10^{-1}	7.00×10^{-3}
10	0.849 (0.06)	2.0	6.006 (0.76)	6.096 (0.76)	9.85×10^{-1}	9.01×10^{-4}

[a]Average over $N = 500$ models.
[b]Numbers in parenthesis denote the coefficient of variation of the quantity, as obtained from the analytical approximation.
[c]Average design thickness without certification is 0.847.

Table 7 Probability of failure for different error bounds for panels designed using safety factor of 1.0 and mean value for allowable stress

Error bound e, %	Average design thickness after certification[a,b]	Certification failure rate, %	Probability of failure after certification P_t	Probability of failure without certification P_{nt}	Probability ratio P_t/P_{nt}	Probability difference $P_{nt} - P_t$
50	0.830 (0.13)	50.7	0.151 (1.32)	0.598 (1.19)	2.53×10^{-1}	4.47×10^{-1}
40	0.794 (0.11)	51.6	0.185 (1.10)	0.519 (0.76)	3.56×10^{-1}	3.35×10^{-1}
30	0.758 (0.09)	53.2	0.237 (0.86)	0.514 (0.67)	4.61×10^{-1}	2.77×10^{-1}
20	0.720 (0.08)	49.8	0.333 (0.60)	0.510 (0.53)	6.53×10^{-1}	1.77×10^{-1}
10	0.683 (0.05)	50.4	0.429 (0.32)	0.510 (0.31)	8.41×10^{-1}	8.08×10^{-2}

[a]Average over $N = 500$ models. [b]Average design thickness without certification is 0.667.

Table 8 Probability of failure for uncertainty in failure stress for panels designed using safety factor of 1.5, 50% error bounds *e*, and A-basis property for allowable stress

Coefficient of variation of σ_f, %	Average design thickness without certification[a,b]	Average design thickness after certification[a,b]	Certification failure rate, %	Probability of failure after certification $P_t \times 10^{-4}$ [b]	Probability of failure without certification $P_{nt} \times 10^{-4}$ [b]	Probability ratio P_t/P_{nt}	Probability difference $P_{nt} - P_t$
0	0.993 (0.29)	1.250 (0.12)	52.6	0.011 (4.33)	1690 (1.90)	6.26×10^{-6}	1.69×10^{-1}
5	1.131 (0.29)	1.324 (0.15)	37.5	1.52 (14.3)	1060 (3.01)	1.44×10^{-3}	1.06×10^{-1}
10	1.274 (0.29)	1.454 (0.18)	35.5	6.05 (8.30)	540 (2.63)	1.12×10^{-2}	5.34×10^{-2}
15	1.376 (0.29)	1.578 (0.22)	24.0	21.9 (5.50)	222 (2.66)	9.88×10^{-2}	2.00×10^{-2}
20	1.608 (0.29)	1.729 (0.24)	16.5	47.5 (3.79)	133 (2.55)	3.58×10^{-1}	8.53×10^{-3}

[a]Average over $N = 500$ models. [b]Numbers in parenthesis denote the coefficient of variation of the quantity, as obtained from the analytical approximation.

Table 9 Probability of failure for uncertainty in failure stress for panels designed using safety factor of 1.5, 30% error bound *e*, and A-basis properties

Coefficient of variation of σ_f, %	Average design thickness without certification[a]	Average design thickness after certification[a]	Certification failure rate, %	Probability of failure after certification $P_t \times 10^{-4}$	Probability of failure without certification $P_{nt} \times 10^{-4}$	Probability ratio P_t/P_{nt}	Probability difference $P_{nt} - P_t$
0	0.997 (0.17)	1.149 (0.08)	49.9	0.022 (3.71)	245 (2.44)	8.98×10^{-5}	2.45×10^{-2}
5	1.139 (0.17)	1.231 (0.11)	31.0	0.145 (6.93)	45.5 (2.78)	3.19×10^{-3}	4.54×10^{-3}
10	1.271 (0.17)	1.325 (0.15)	16.0	3.93 (2.91)	12.1 (2.65)	3.26×10^{-1}	8.12×10^{-4}
15	1.439 (0.17)	1.464 (0.16)	7.4	6.79 (1.80)	10.1 (1.42)	6.70×10^{-1}	3.34×10^{-4}
20	1.624 (0.17)	1.641 (0.17)	4.5	7.37 (1.56)	8.96 (1.41)	8.23×10^{-1}	1.58×10^{-4}

[a]Average over $N = 500$ models.

still fail certification, but small errors will be masked by the conservative properties. Thus in Table 6, the certification failure rate varies from 52.3% for the largest errors to 2.0% for the smallest errors. At the highest error bound (50%), the certification process increases the average thickness from 0.847 to 0.968, and this drops to 0.885 for 30% error bound. This substantial drop in average certified model thicknesses increases the probability of failure. Below an error bound of 30%, the change in thickness is small, and then reducing errors reduces the probability of failure. This is because small negative errors are not caught by certification, but they still reduce the effective safety factor.

Table 7 shows results when the only safety measure is certification testing. Certification tests can reduce the probability of failure of panels by 45%; the highest improvement corresponds to the highest error. As can be expected, without certification tests and safety measures, the probability of failure is near 50%.

Tables 4–7 illustrate the probability of failure for a fixed 10% coefficient of variation in failure stress. The general conclusion that can be drawn from these results is that the error bound e is one of the main parameters affecting the efficacy of certification tests to improve reliability of panels. Next, we will explore how another parameter, variability, influences the efficacy of tests. This is accomplished by changing the coefficient of variation of failure stress σ_f between 0–20% and keeping the error bound constant.

The increase in the variability in failure stress has a large effect on the allowable stress because A-basis properties specify an allowable that is below 99% of the sample. Increased variability reduces the allowable stress and therefore increases the design thickness. It is seen from Tables 8–10 that when the variability increases from 0 to 20% the design thickness increases by more than 60%. In spite of this, the probability of failure still deteriorates. That is, the use

Table 10 Probability of failure for uncertainty in failure stress for panels designed using safety factor of 1.5, 20% error bounds e, and A-basis properties

Coefficient of variation of σ_f, %	Average design thickness without certification[a]	Average design thickness after certification[a]	Certification failure rate, %	Probability of failure after certification $P_t \times 10^{-4}$	Probability of failure without certification $P_{nt} \times 10^{-4}$	Probability ratio P_t/P_{nt}	Probability difference $P_{nt} - P_t$
0	1.007 (0.12)	1.099 (0.05)	47.9	0.04 (2.73)	25.2 (2.33)	1.60×10^{-3}	2.51×10^{-3}
5	1.127 (0.12)	1.174 (0.09)	21.0	0.253 (4.67)	2.33 (2.54)	1.09×10^{-1}	2.08×10^{-4}
10	1.279 (0.12)	1.285 (0.10)	7.0	0.538 (2.55)	1.55 (2.24)	3.47×10^{-1}	1.01×10^{-4}
15	1.436 (0.12)	1.442 (0.11)	3.5	1.17 (2.01)	1.70 (1.93)	6.91×10^{-1}	5.26×10^{-5}
20	1.623 (0.12)	1.629 (0.11)	3.1	1.91 (1.68)	2.40 (1.64)	7.93×10^{-1}	4.96×10^{-5}

[a]Average over $N = 500$ models.

of A-basis properties fails to fully compensate for the variability in material properties.

The variability in failure stress greatly changes the effect of certification tests. Although the average design thicknesses of the panels increase with the increase in variability, we see that when the variability is large the value of the tests is reduced because the tested aircraft can be greatly different from the airplanes in actual service. We indeed see from the Tables 8–10 that the effect of certification tests is reduced as the variability in the failure stress increases. Recall that the effect of certification tests is also reduced when the error e decreases. Indeed, Table 8 shows a much smaller effect of the tests than Table 10. Comparing the second and third columns of Tables 8–10, we see that as the bound of error decreases the change in the average value of design thicknesses of the panels becomes less, which is an indication of loss in the efficacy of certification tests.

P_{nt} and P_t results of Tables 8–10 corresponding to a 10% coefficient of variation of s_f are slightly different from the results presented in Table 3 with 50, 30, and 20% error bounds. This is an indication of the accuracy of Monte Carlo simulations. More accurate results can be obtained by increasing the sample size.

As mentioned earlier, in this study we did not include the effects of building-block types of tests that are conducted before certification. These tests not only reduce the errors in analytical models but also reduce the variability in material properties. For instance, if there is 20% coefficient of variation in the failure stress the building-block type of tests can reduce this error to lower values. Hence, this is another way of looking at Tables 8–10.

Up to now, both the probability difference $P_{nt} - P_t$ and the probability ratio P_t/P_{nt} seem to be good indicators of efficacy of tests. To allow easy visualization, we combined the errors and the variability in a single ratio (bounds on $e)/V_R(s/s_f)$ ratio (ratio of error bound e to the coefficient of variation of the stress ratio). The denominator accounts for the major contributors to the variability. The value in the denominator is a function of four variables: service load P, width w, thickness t, and failure stress s_f. Here, P and s_f have log-normal distributions, but w and t are uniformly distributed. Because the coefficient of variations of w and t is very small, they can also be treated as log-normally distributed to make calculation of the denominator easy while plotting the graphs. Because the standard deviations of the variables are small, the denominator is now the square root of the sum of the squares of coefficient of variations of the four variables just mentioned, that is,

$$V_R(\sigma/\sigma_f) \cong \sqrt{V_R^2(P) + V_R^2(w) + V_R^2(t) + V_R^2(\sigma_f)} \qquad (11)$$

The effective safety factor is the ratio of the design thickness of the component when safety measures (such as usage of A-basis values for material properties and safety factor) are applied to the thickness of the component when no safety measures are taken.

Figures 4 and 5 present the P_t/P_{nt} ratio in visual formats. It can be seen that as expected the ratio decreases as the (bounds on $e)/V_R(s/s_f)$ ratio increases. However, these two figures do not give a clear indication of how certification tests are influenced by the effective safety factor.

Figures 6 and 7 show the probability difference $P_{nt} - P_t$. In these cases, the dependence on the effective safety factor is monotonic. As expected, it is seen that as the effective safety factor increases, the improvement in the safety of component decreases, meaning that

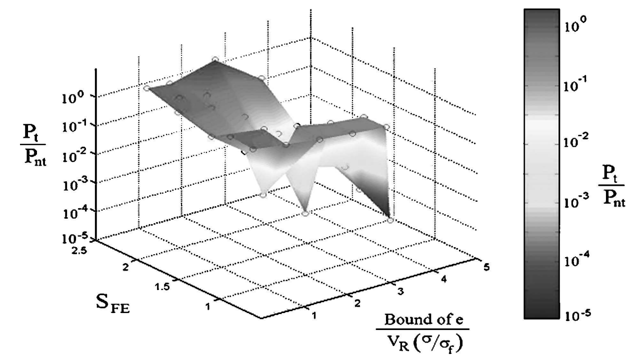

Fig. 4 Influence of effective safety factor, error, and variability on the probability ratio (three-dimensional view).

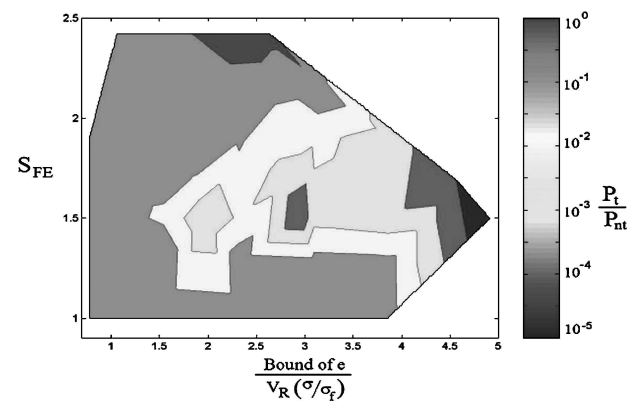

Fig. 5 Influence of effective safety factor, error and variability on the probability ratio (two-dimensional contour plot).

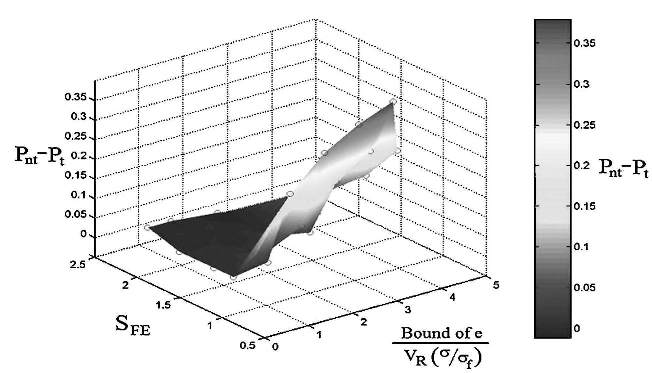

Fig. 6 Influence of effective safety factor, error and variability on the probability difference (three-dimensional view).

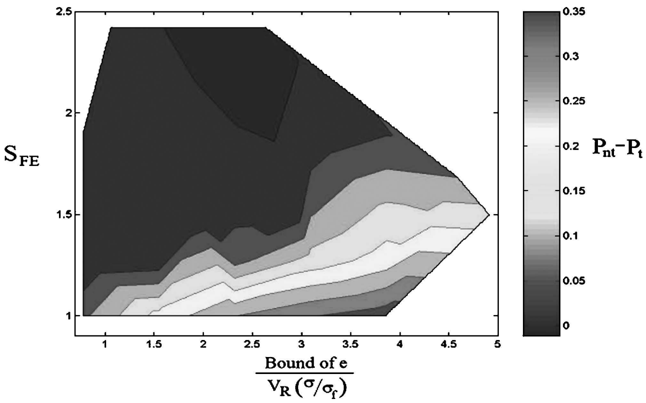

Fig. 7 Influence of effective safety factor, error and variability on the probability difference (two-dimensional contour plot).

the certification tests become less useful. The probability difference is more descriptive as it is proportional to the number of aircraft failures prevented by certification testing. The probability ratio lacks such clear physical interpretation, even though it is a more attractive measure when the probability of failure is very small.

Considering the results presented by Figs. 4–7, the probability difference $P_{\rm nt} - P_t$ is the more appropriate choice for expressing the effectiveness of tests.

VI. Conclusions

We have used a simple example of point stress design for yield to illustrate the effects of several safety measures taken in aircraft design: safety factors, conservative material properties, and certification tests. Analytical calculations and Monte Carlo simulation were performed to account for both fleet-level uncertainties (such as errors in analytical models) and individual uncertainties (such as variability in material properties).

It was seen that an increase of the systemic errors in the analysis causes an increase in the probability of failure. We found that the systemic errors can be reduced by the use of certification tests, thereby reducing the probability of failure. Also we found that design thicknesses of the panels increased as the bounds of systemic errors increased.

We found that the effect of certification tests is most important when errors in analytical models are high and when the variability between airplanes is low. This leads to the surprising result that in some situations larger error variability in analytical models reduces the probability of failure if certification tests are conducted. For the simple example analyzed here, the use of conservative (A-basis) properties was equivalent to a safety factor of up to 1.6, depending on the scatter in failure stresses.

The effectiveness of the certification tests is expressed by two measures of probability improvement. The ratio of the probability of failure with the test P_t to the probability of failure without tests $P_{\rm nt}$ is useful when P_t is small. The difference is more meaningful when the probability is high. Using these measures, we have shown that the effectiveness of certification tests increases when the ratio of error to variability is large and when the effective safety factor is small.

The effect of building-block type tests that are conducted before certification was not assessed here. However, these tests reduce the errors in the analytical models, and on that basis we determined that they can reduce the probability of failure by one or two orders of magnitude.

The calculated probabilities of failure with all of the considered safety margins explain why passenger aircraft are so safe structurally. They were still somewhat high—about 10^{-7}—compared to the probability of failure of actual aircraft structural components—about 10^{-8}. This might be caused by additional safety measures, such as conservative design loads or to the effect of design against additional failure modes.

Appendix A: A-Basis Property Calculation

A-basis value is the value exceeded by 99% of the population with 95% confidence. This is given by

$$\text{A-basis} = \mu - s \times k_1 \tag{A1}$$

where μ is the mean, s is the standard deviation, and k_1 is the tolerance coefficient for normal distribution given by Eq. (A2):

$$k_1 = \left(z_{1-p} + \sqrt{z_{1-p}^2 - ab}\right)\Big/ a$$

$$a = 1 - z_{1-\gamma}^2\Big/2(N-1), \qquad b = z_{1-p}^2 - z_{1-\gamma}^2\Big/N \tag{A2}$$

where N is the sample size and z_{1-p} is the critical value of normal distribution that is exceeded with a probability of $1 - p$. The tolerance coefficient k_1 for a log-normal distribution is obtained by first transforming the log-normally distributed variable to a normally distributed variable. Equations (A1) and (A2) can be used to obtain an intermediate value. This value is then converted back to the log-normally distributed variable using inverse transformation.

To obtain the A-basis values, usually 40 or 50 coupons are randomly selected from a batch. Then, the mean and standard deviation of the failure stresses of these coupons are calculated and used in determining the A-basis allowable stress. The calculated A-basis value itself is also a random variable. For instance, when the failure stress is lognormal with 10% coefficient of variation and 40 tests are performed, the coefficient of variation of A-basis value is about 4%. To suppress the uncertainty in A-basis stress, we used 10,000 batches in this study.

Appendix B: Probability Calculations

Calculations of $P(C|e)$, the Probability of Passing Certification Test

$$P(C|e) = P(\sigma_f > \sigma) = P(\sigma_f > S_F P_d/wt)$$

$$= P(\sigma_f wt > S_F P_d) = P(R > S) \tag{B1}$$

where

$$R = \sigma_f t w \qquad \text{and} \qquad S = S_F P_d \tag{B2}$$

S is a deterministic value, and because the coefficient of variations of t and w is small compared to the coefficient of variation of σ_f, R can be treated as log normally distributed with parameters λ_R and ζ_R. Then,

$$\lambda_R(e) = \lambda_{\sigma_f} + \lambda_t(e) + \lambda_w \qquad \text{and} \qquad \zeta_R^2 = \zeta_{\sigma_f}^2 + \zeta_t^2 + \zeta_w^2 \tag{B3}$$

where

$$\lambda_t(e) = \ell\text{n}[t_{\rm design}(e)] - 0.5\zeta_t^2 \tag{B4}$$

where $t_{\rm design}$ is given in Eq. (4).

Then, $P(C|e)$ can be calculated as

$$P(C|e) = P(R > S) = \Phi\left[\frac{\lambda_R(e) - S}{\zeta_R}\right] = \Phi(\beta(e))$$

$$= \int_{-\infty}^{\beta(e)} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{x^2}{2}\right) \mathrm{d}x \tag{B5}$$

Calculations of Mean Value and Standard Deviation of Probability of Failure

Failure is predicted to occur when the resistance of the structure R of the problem is less than the load P; see Eq. (6). Then, the probability of failure is given as

$$P_f = Pr(R < P) \tag{B6}$$

The load P is log-normally distributed, and, as explained in Appendix B, the distribution of R can also be approximated by a log-normal distribution, which allows us to immediately obtain the probability of failure of a single aircraft model.

To calculate the probability of failure over all aircraft models, we take into account the fact that that t_{design} is a random variable. Then, the expected value of probability of failure is given as

$$\hat{P}_f = \int P_f(t_{\text{design}}) f(t_{\text{design}})\, \mathrm{d}t_{\text{design}} \tag{B7}$$

where t_{design} is the nondeterministic distribution parameter and $f(t_{\text{design}})$ is the probability density function of parameter t_{design}.

The standard deviation of failure probability can be calculated from

$$\sigma_{P_f} = \left[\int (P_f - \hat{P}_f)^2 f(P_f)\, \mathrm{d}P_f\right]^{\frac{1}{2}} \tag{B8}$$

where

$$P_f = P_f(t_{\text{design}}), \qquad f(P_f) = f(t_{\text{design}})\left|\frac{\mathrm{d}t_{\text{design}}}{\mathrm{d}P_f}\right|$$

$$\mathrm{d}P_f = \frac{1}{\mathrm{d}t_{\text{design}}/\mathrm{d}P_f}\, \mathrm{d}t_{\text{design}} \tag{B9}$$

Hence, Eq. (B8) can be rewritten as

$$\sigma_{P_f} = \left\{\int [P_f(t_{\text{design}}) - \hat{P}_f]^2 f(t_{\text{design}})\, \mathrm{d}t_{\text{design}}\right\}^{\frac{1}{2}} \tag{B10}$$

As seen from Eqs. (B7) and (B10), the mean and standard deviation of the probability of failure can be expressed in terms of the probability density function (PDF) f of the design thickness t_{design}. Therefore, we can perform the failure probability estimations to after calculating the PDF of t_{design}. The random variables contributing to t_{design} are [see Eq. (4)] e, w, and σ_a. Because the variations of w and σ_a are small compared to e, we neglect their contribution and calculate the PDF of t_{design} from the PDF of error factor e from

$$f(t_{\text{design}}) = f_e(e)\frac{\mathrm{d}e}{\mathrm{d}t_{\text{design}}} \tag{B11}$$

where $f_e(e)$ is the updated PDF of e.

Acknowledgment

This work was supported in part by NASA Cooperative Agreement NCC3-994, NASA University Research, Engineering and Technology Institute and NASA Langley Research Center Grant Number NAG1-03070.

References

[1]Lincoln, J. W., "Method for Computation of Structural Failure Probability for an Aircraft," ASD-TR-80-5035, Wright–Patterson AFB, OH, July 1980.

[2]Wirsching, P. H., "Literature Review on Mechanical Reliability and Probabilistic Design," *Probabilistic Structural Analysis Methods for Select Space Propulsion System Components (PSAM)*, NASA CR 189159, Vol. III, 1992.

[3]"Integration of Probabilistic Methods into the Design Process," Society of Automative Engineers, Aerospace Information Rept. 5080, 1997.

[4]Long, M. W., and Narciso, J. D., "Probabilistic Design Methodology for Composite Aircraft Structures." DOD/FAA/AR-99/2, Final Rept., June 1999.

[5]Ben-Haim, Y., and Elishakoff, I., *Convex Models of Uncertainty in Applied Mechanics*, Elsevier, Amsterdam, 1990.

[6]Neal, D. M., Matthews, W. T., and Vangel, M. G., 1992, "Uncertainties in Obtaining High Reliability from Stress-Strength Models," *Proceedings of the 9th DOD-/NASA/FAA Conference on Fibrous Composites in Structural Design*, Vol. 1, Dept. of Defence, Lake Tahoe, NV, 1991, pp. 503–521.

[7]Oberkampf, W. L., Deland, S. M., Rutherford, B. M., Diegert, K. V., and Alvin, K. F., "Estimation of Total Uncertainty in Modeling and Simulation," Sandia National Lab. Rept. SAND2000-0824, Albuquerque, NM, April 2000.

[8]Kale, A., Haftka, R. T., Papila, M., and Sankar, B. V., "Tradeoffs of Weight and Inspection Cost in Safe-Life Design," AIAA Paper 2003-1501, April 2003.