54th AIAA/ASME/ASCE/AHS/ASC Structures, Structural Dynamics, and Materials Conference AIAA 2013-1819
April 8-11, 2013, Boston, Massachusetts

The effect of the number of coupon and element tests on conservativeness

Chanyoung Park[1], Nam H. Kim[2], and Raphael T. Haftka[3]
Department of Mechanical and Aerospace Engineering
University of Florida, Gainesville, FL 32611

Abstract

Structural elements, such as stiffened panels are designed based on material strength data obtained from coupon tests, with a failure theory that generalizes 1D stress failure to 3D stress field. Variability in material properties is captured by dozens of coupon tests, but there remains epistemic uncertainty due to error in a failure theory, which is reduced by element tests. However, the uncertainties are not completely removed by tests because the number of structural tests is finite. Therefore designs of structural elements require conservative estimate of the failure stress of structural elements to compensate the remaining uncertainties. Consequently, the design weight is a function of the number of coupon and element tests.

A key question, addressed here, is whether it is more important to increase the number of coupon tests or element tests if we want to design light structures by reducing the remaining uncertainty in failure prediction. A convolution approach that allows efficient estimation of the combined uncertainty from coupon tests and the failure theory is developed. Then the methodology is applied to typical values of the variability in material properties (COV 7%) and the error in the failure theory (±5% and ±1%). It is found that the uncertainty in failure theory dominates for the case of having ±5% error, so that increasing the number of element tests is more effective than increasing the number of coupon tests. However, for the case of having small ±1% errors, the number of element tests has similar influence to that the number of coupon tests.

Nomenclature

b_e	= Error bound for failure theory
b_σ	= Estimated bound for standard deviation of structural element
$\hat{e}_{k,Ptrue}$	= Possible true error in failure theory
$f^{init}(\mu_{e,Ptrue}, \sigma_{e,Ptrue})$	= Initial joint PDF for given mean and standard deviation of structural element
$f_{k,Ptrue}\left(e_{k,Ptrue}\right)$	= PDF for given possible true error in failure theory
$f_{\mu c,Ptrue}\left(\mu_{c,Ptrue}\right)$	= PDF for given possible true mean of material strength
$f_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\right)$	= PDF for given possible true mean of structural strength

[1] Graduate Research Assistant, Email) cy.park@ufl.edu
[2] Associate Professor, Corresponding author, Email) nkim@ufl.edu, Tel)1-352-575-0665, Fax)1-352-392-7303
[3] Distinguished Professor, Email) haftka@ufl.edu

$f_{\sigma c,Ptrue}\left(\sigma_{c,Ptrue}\right)$ = PDF for given possible true standard deviation of material strength

$f_{\sigma e,Ptrue}\left(\sigma_{e,Ptrue}\right)$ = PDF for given possible true standard deviation of structural strength

$f^{upd}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)$ = Updated joint PDF for given mean and standard deviation of structural element

$f^{upd}_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\right)$ = Updated marginal distribution for given mean of structural element

$f^{upd}_{\sigma e,Ptrue}\left(\sigma_{e,Ptrue}\right)$ = Updated marginal distribution for given standard deviation of structural element

k_{calc} = Calculated ratio of structural element strength to material strength

$\hat{k}_{Ptrue}$ = Possible true structural element strength to material strength

k_{true} = True ratio of structural element strength to material strength

$l^{i}_{test}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)$ = Likelihood function of i^{th} test for given mean and standard deviation of structural element

$\mu_{0.05}$ = Mean of 5th percentile of the mean element strength for given test results

$\hat{\mu}_{c,Ptrue}$ = Possible true mean of material strength

$\mu_{c,test}$ = Measured mean of material strength from coupon test

$\mu_{c,true}$ = True mean of material strength

$\hat{\mu}_{e,Ptrue}$ = Possible true mean of structural element strength

$\mu_{e,test}$ = Measured mean of structural element strength from coupon test

$\mu_{e,true}$ = True mean of structural element strength

n_c = The number of coupon tests

n_e = The number of element tests

PUD = Probability of unconservative design

$\hat{\sigma}_{c,Ptrue}$ = Possible true standard deviation of material strength

$\sigma_{c,test}$ = Measured standard deviation of material strength from coupon test

$\sigma_{c,true}$ = True standard deviation of material strength

$\hat{\sigma}_{e,Ptrue}$ = Possible true standard deviation of structural element strength

$\sigma_{e,test}$ = Measured standard deviation of structural element strength from coupon test

$\sigma_{e,true}$ = True standard deviation of structural element strength

$\tau_{0.05}$ = 5th percentile of the mean element strength for given test results

$\hat{\tau}_{c,Ptrue}$ = Possible true material strength

$\hat{\tau}_{c,true}$ = True material strength

$\hat{\tau}_{e,Ptrue}$ = Possible true structural element strength

$\hat{\tau}_{e,true}$ = True structural element strength

$w_{0.95}$ = 95th percentile of the weight penalty for given test results

Superscripts

init	= Initial distribution (prior distribution)
upd	= Updated distribution (posterior distribution)
Subscripts	
calc	= Calculated value using a theory
Ptrue	= Possible true estimate reflecting epistemic uncertainty of estimation process
test	= Measured value from a test
true	= True value

I. Introduction

Predicting variability in failure stress is important to ensure structural safety in engineering designs. Due to natural variability in material properties and manufacturing processes, identically manufactured structures have variability in failure stress, which brings unavoidable randomness.. Due to its incontrollable nature, the variability can be categorized as aleatory uncertainty. The randomness has to be compensated with conservativeness in structural design

The Element is a basic unit composing aircraft structure. Predicting randomness in failure stress of elements is important in aircraft structural design. The randomness in failure stress can be mathematically modeled with a statistical distribution. Since distributions are typically defined by their parameters, predicting the distribution of the failure stress is equivalent to predicting distribution parameters (e.g. predicting mean and standard deviation of failure stress that follows a normal distribution).

However, predicting the statistical distribution of element strength is hampered by two epistemic uncertainties. Since structural elements are under multi-axial stress state, the distribution of element failure stress has to be predicted by translating the distribution of material strength using a failure theory. First, the distribution of material strength has to be estimated based on a finite number of coupon tests; thus, incurring a sampling error. Second, element failure has to be predicted using a failure theory that causes some inevitable errors. In contrast to irreducible aleatory uncertainty, these uncertainties are reducible with various measures. They are categorized as epistemic uncertainty. The two epistemic uncertainties lead to uncertainty in parameters of the predicted distribution of element failure stress.

Epistemic uncertainty is often treated conservatively in the literature of probabilistic design. Noh *et al*. [1] compensated for epistemic uncertainty caused by the finite number of samples with a confidence level of 97.5%. Matsumura *et al.* [2] and Villanueva *et al*. [3] considered the effect of epistemic uncertainty in a computer model on estimating probability of failure of an integrated thermal protection system of a space vehicle and demanded 95% confidence for the epistemic uncertainty.

For ensuring integrity of structural elements, conservativeness in element design, compensates for the aleatory uncertainty and the two epistemic uncertainties. Here, a distinction is made between irreducible and reducible conservativeness in design. Since conservativeness due to the aleatory uncertainty is unavoidable, it is irreducible. On the contrary, since conservativeness due to the epistemic uncertainties is reducible, design weight increase due to the epistemic uncertainties is also reducible.

A Number of coupon and a number of element tests are carried out to reduce the epistemic uncertainties, and thus, the weight penalty. The epistemic uncertainty from predicting material strength distribution can be reduced by increasing the number of coupon tests. The epistemic uncertainty from a failure theory can be reduced by carrying out element tests. The reduced epistemic uncertainties after tests is reflected in reduced uncertainty in parameters of the element failure stress distribution, which in turn leads the reduced weight penalty. Typicall numbers of tests for coupon and element tests are 50 and 3, respectively.

In an overview of future structures technology for military aircraft, Joseph *et al*. [4] noted that a progressive uncertainty reduction model, which is seen in building-block tests, can be a feasible solution today, since complete replacement of traditional tests with computational models is not feasible yet. Lincoln *et al*. [5] pointed out that building-block tests play a key role in reducing errors in failure prediction of composite structures due to large uncertainty in computational models. They noted that the use of probabilistic methods can significantly lower the test cost by reducing the scope of the test program.

There are also several studies investigating the effect of tests on safety and reducing uncertainty in computational models. Jiao and Moan [6] investigated the effect of proof tests on structural safety using Bayesian inference. They showed that proof tests reduce uncertainty in the strength of a structure, and thus provide a substantial reduction in the probability of failure. An *et al.* [7] investigated the effect of structural element tests on reducing uncertainty in element strength using Bayesian inference. Acar *et al.* [8] modeled a simplified building-block process with safety factors and knockdown factors. Bayesian inference is used to model the effect of structural element tests. They showed the effect of the number of tests on the design weight for the same probability of failure, and vice versa. Jiang and Mahadevan [9] studied the effect of tests in validating a computational model by obtaining an expected risk in terms of the decision cost. Urbina and Mahadevan [10] assessed the effects of system level tests for assessing reliability of complex systems. They built computational models of a system and predicted the performance of the system. Tests are then incorporated into the models to estimate the confidence in the performance of the systems. Park et al. [11] estimated uncertainty in computational models and developed a methodology to evaluate likelihood using both test data and a computational model. McFarland and Bichon [12] estimated probability of failure by incorporating test data for a bistable MEMS device.

The objective of this paper is to introduce a probabilistic approach to estimate parameters, mean and standard deviation, and their uncertainty of the distribution of element failure stress. Since approaches to estimate the mean and standard deviation are similar, this paper focuses on estimating the mean and its uncertainty after tests. Using the estimated uncertainty in the mean, a mean with 95% conservativeness is estimated. The mean with 95% conservativeness is a conservative estimate that is expected to be less conservative than the true mean with 95% probability. The weight penalty, a consequence of the conservativeness, is also calculated. The effect of the number of coupon and element tests on reducing the weight penalty is discussed. It is assumed that with an infinite number of coupons and elements, the sampling uncertainty and the uncertainty of a failure theory can be eliminated; therefore, the weight penalty becomes zero. The effect of the number of tests on the incurring weight penalty is compared to the case with an infinite number of tests.

The methodology is examined with typical values of material variability (7% COV) and element test variability (3% COV). The effect of the magnitude of error in failure theory is shown with two different magnitudes of errors in a failure theory, ±1% and ±5%. The effect of the number of tests on the weight penalty with different errors in a failure theory is also discussed. It is observed that the element test is influential to reduce the weight penalty for a failure theory with ±5% error. For a failure theory with ±1% error, coupon test becomes influential but still the effect of element test is not ignorable.

The paper is organized as follows: Section II introduces the building-block test process, which is composed of coupon and element test stages, used in this paper and sources of uncertainty. Section III provides uncertainty modeling of the building-block test process to estimate the element strength and its uncertainty. This section has three subsections: coupon tests, element design and element tests. Section IV introduces different measures that are used to evaluate the efficiency of different tests. Section V presents numerical results, followed by conclusions in Section VI.

II. Structural Uncertainties

For aircraft structures, the building-block test process (Fig. 1) is used to find design errors and to reduce uncertainties in design and manufacturing. At each level, analytical/numerical models are calibrated to account for discrepancies between model prediction and test results. Since the errors are unknown at the modeling stage, they may be modeled as uncertainty (epistemic), and test results may be used to reduce the uncertainty. Starting from simple coupon tests at the bottom level, structural complexity gradually increases further up the building-block pyramid. The number of tests gradually reduces from bottom to top; for example, 50 coupons, 3 elements, and 1 component. In higher-level tests, it is difficult to understand deviations from analytical predictions, tests are more expensive, and any design modification can be expensive. The building-block test process is designed to detect modeling errors at the possible lowest level.

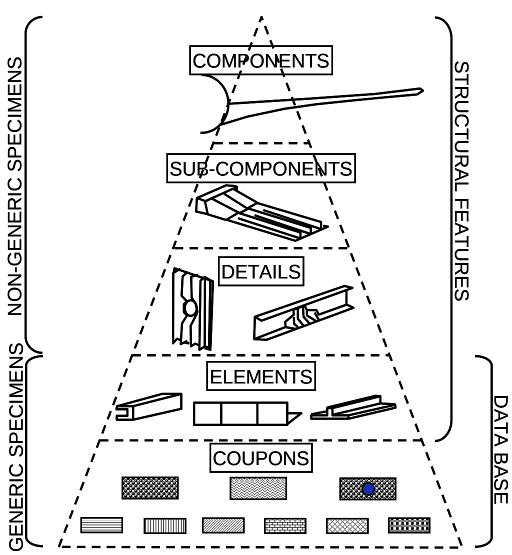


Figure 1: Building-block test process for aircraft structural components

Although building-block tests are designed to reduce uncertainty, it is difficult to quantify how much each level can contribute to uncertainty reduction, which is the main objective of this paper.

Once the contribution of each level to uncertainty reduction is understood, a design engineer can decide how to allocate resources to different levels in order to achieve the target reliability at minimum cost.

Although the actual building-block test process has many levels, this paper only considers coupon and element tests to demonstrate the effect of these tests on uncertainty reduction. Table 1 shows the objectives of these two tests and the sources of uncertainty.

Table 1. Sources of uncertainty in the building-block test process for estimating element strength

Test stage	Objectives	Uncertainty sources
Coupon test	Estimate nominal value and variability of material strength	Variability in material strength and sampling error due to a finite number of coupons
Element design	Estimate multi-axial strength based on a failure theory	Incomplete knowledge of failure mechanism: error in failure theory
Element test	Reduce uncertainty in the multi-axial strength	Sampling error due to a finite number of elements

In this paper, the failure stress of a structural element is simulated with randomly generated test results. True distributions are used only for generating test samples and assessing the estimations of the failure stress.

III. Modeling Uncertainty in the Building-Block Test Process

In order to model the two-level building-block test process, it is assumed that the strength of coupons and elements follows a normal distribution due to material variability. This assumption can easily be removed when actual test results are available, and the type of distribution can be identified using various statistical methods, such as the one in MIL-HDBK [20]. In the following subsections, uncertainties at each stage are modeled.

A. Coupon tests: Modeling uncertainty in estimating statistical properties

Due to inherent variability, the material strength shows a statistical distribution. Coupon tests are conducted to estimate the distribution and to determine regulatory (e.g., FAA) strength allowables (e.g., A-basis or B-basis) that compensate for the uncertainty. It is assumed that the true material strength, $\hat{\tau}_{c,true}$, follows a normal distribution as,

$$\hat{\tau}_{c,true} \sim N\left(\mu_{c,true}, \sigma_{c,true}\right) \tag{1}$$

where $\mu_{c,true}$ and $\sigma_{c,true}$ are, respectively, the mean and standard deviation of $\hat{\tau}_{c,true}$. The circumflex symbol represents a random variable. The subscript "*c*" is used to denote coupons. In this paper, Eq. (1) is only used for the purpose of simulating coupon tests; the true distribution is unknown to the designer.

Since the true distribution parameters are estimated with a finite number of coupons, the estimated parameters have sampling uncertainty (or error). Thus, it is natural to consider these parameters as distributions rather than deterministic values. In this paper, this estimated distribution is called the possible true distribution (PTD) of the parameter, and $\hat{\mu}_{c,Ptrue}$ and $\hat{\sigma}_{c,Ptrue}$ are random variables of the PTD of the mean and standard deviation, respectively. Note that $\hat{\mu}_{c,Ptrue}$ and $\hat{\sigma}_{c,Ptrue}$ depend on the number of coupons. With n_c coupons, $\hat{\mu}_{c,Ptrue}$ is nothing but the distribution of sample mean and can be estimated as

$$\hat{\mu}_{c,Ptrue} \sim N\left(\mu_{c,test}, \frac{\sigma_{c,test}}{\sqrt{n_c}}\right) \tag{2}$$

where $\mu_{c,test}$ and $\sigma_{c,test}$ are, respectively, the mean and standard deviation of coupons. With an infinite number of coupons, $\hat{\mu}_{c,Ptrue}$ will become a deterministic value; i.e., no sampling error.

It is also well known that the standard deviation $\hat{\sigma}_{c,Ptrue}$ follows a chi-distribution of order $n_c - 1$. In a way similar to the mean, $\hat{\sigma}_{c,Ptrue}$ can be estimated as

$$\hat{\sigma}_{c,Ptrue} \sim \frac{\sigma_{c,test}}{\sqrt{n_c - 1}} \chi\left(n_c - 1\right) \tag{3}$$

where $\chi\left(n_c - 1\right)$ is the chi-distribution of order $n_c - 1$.

B. Element design: combining uncertainties

To design a structural element, the material strength from coupon tests must be generalized to multi-axial stress states using a failure theory. Since the failure theory is not perfect, additional error (i.e., epistemic uncertainty) is introduced, which needs to be combined with the sampling error in the coupon test. Since the uncertainty in element strength can be represented using the distributions of mean and standard deviation, the uncertainties of these two random variables are modeled separately [15].

A failure theory provides a relation between uni-axial strength and multi-axial strength. In this paper, this relation is represented using a prediction factor $k_{3d,true}$ as

$$\tau_{e,true} = k_{3d,true} \tau_{c,true} \tag{5}$$

where $\tau_{c,true}$ is a true uni-axial material strength, and $\tau_{e,true}$ is a true multi-axial equivalent strength. Subscript "*e*" is used to denote that the variable is for an element. For example, when von Mises criterion is used, $k_{3d,true} = 1$. The relation between the two mean values can be obtained from Eq. (5) as

$$\mu_{e,true} = k_{3d,true} \mu_{c,true} \tag{6}$$

Again, $k_{3d,true}$ is unknown to designers; only its estimate $k_{3d,calc}$ is given from the failure theory. Therefore, the epistemic uncertainty in the failure theory can represented using the PTD of the prediction factor as

$$\hat{k}_{3d,Ptrue} = \left(1 - \hat{e}_{k,Ptrue}\right) k_{3d,calc} \tag{7}$$

In the above equation, the error $\hat{e}_{k,Ptrue}$ is assumed to follow a normal distribution that has symmetric 95% confidence interval defined with $\pm b_e$ that is obtained error bounds of the failure prediction from experts. Then, the designer's estimated relationship corresponding to Eq. (6) can be written as

$$\hat{\mu}_{e,Ptrue} = \hat{k}_{3d,Ptrue}\hat{\mu}_{c,Ptrue} \tag{8}$$

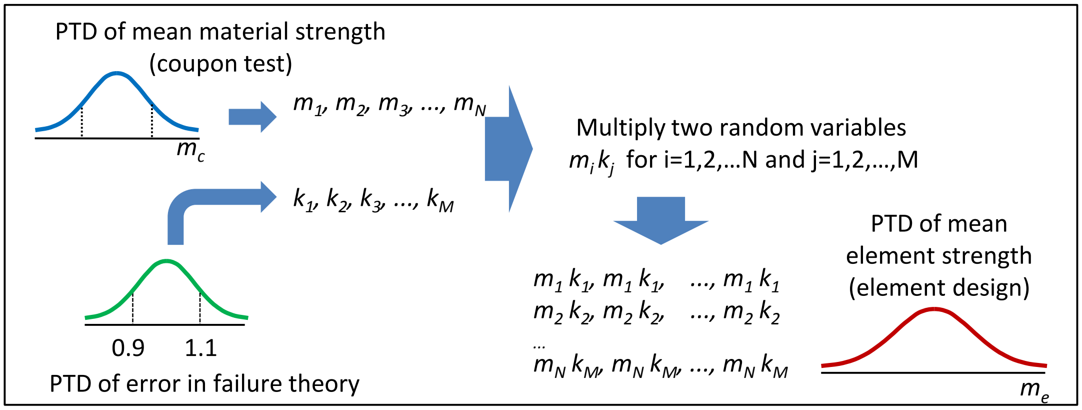


Figure 2: Process of estimating element mean strength

Figure 2 shows a process of obtaining $\hat{\mu}_{e,Ptrue}$ through MCS. First, N samples from $\hat{\mu}_{c,Ptrue}$ and M samples from $\hat{k}_{3d,Ptrue}$ are generated. Then, $\hat{\mu}_{e,Ptrue}$ is estimated from $N \times M$ samples that are obtained by taking every possible combination of the two sets of samples.

In this paper, a convolution integral is used to calculate the PDF of $\hat{\mu}_{e,Ptrue}$. The convolution integral provides an accurate PDF using numerical integration, whereas MCS brings in additional uncertainty. A comparison between MCS and the convolution integral is given in the illustrative example section. In the case of a normally distributed mean and normaly distributed error, the PDF of $\hat{\mu}_{e,Ptrue}$ can be written as (see Appendix A for detailed derivations)

$$f_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\right) = \int_{-\infty}^{\infty} \varphi\left(\mu_{e,Ptrue} \mid \mu_{c,Ptrue}, \frac{b_e}{\mu_{c,Ptrue}\Phi(0.975)}\right)\varphi\left(\mu_{c,Ptrue} \mid \mu_{c,test}, \frac{\sigma_{c,test}}{\sqrt{n_c}}\right) d\mu_{c,Ptrue} \tag{9}$$

where b_e is the error bound of $\hat{e}_{k,Ptrue}$ and $k_{3d,calc} = 1.0$ is assumed. $\Phi(0.975)$ is the standard normal distribution.

Unlike the mean, there is only a weak relation between the standard deviation of coupon strength and that of element strength. Usually test conditions are well controlled to minimize uncertainty; the standard deviation in the test is substantially smaller than that of material properties. The distribution of $\hat{\sigma}_{e,Ptrue}$ is defined as a uniform distribution with lower and upper bounds as

$$f_{\sigma e,Ptrue}\left(\sigma_{e,Ptrue}\right) = \frac{1}{\left(\sigma_e^{upper} - \sigma_e^{lower}\right)} I\left(\sigma_{e,Ptrue} \in \left[\sigma_e^{upper}, \sigma_e^{lower}\right]\right) \tag{10}$$

where $I(\bullet)$ is the indicator function, and σ_e^{upper} and σ_e^{lower} are upper and lower bounds of the standard deviation of element strength, respectively. These bounds are estimated to cover a true standard deviation of element test.

C. Element tests: Bayesian inference to reduce errors

The PTDs in Eqs. (8) and (9) are combined uncertainty from (a) material variability, (b) sampling errors in coupon tests and (c) error in the failure theory. Although material variability will always exist, the other two epistemic uncertainties can be reduced using element tests. In this section, the effect of element tests on reducing uncertainty is modeled using Bayesian inference.

For the purpose of Bayesian inference, Eqs. (8) and (9) are used as marginal prior distributions. Since no correlation information is available, these distributions are assumed to be independent. Therefore, the prior joint PDF is given as

$$f^{init}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)=f_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\right)\cdot f_{\sigma e,Ptrue}\left(\sigma_{e,Ptrue}\right) \tag{11}$$

In Bayesian inference, the updated joint PDF with n_e number of element tests is expressed as

$$f^{upd}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)=\frac{1}{A}\prod_{i=1}^{n_e}\ell_{test}^{i}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)f^{init}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right) \tag{12}$$

where A is a normalizing constant, and $\ell_{test}^{i}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)$ is the ith likelihood function for given $\mu_{e,Ptrue},\sigma_{e,Ptrue}$. From the assumption that the true element strength $\hat{\tau}_{e,true}$ follows a normal distribution and by ignoring errors associated with the test, the likelihood function can be defined as a probability of obtaining test result $\tau_{e,test}^{i}$ for given $\mu_{e,Ptrue}$ and $\sigma_{e,Ptrue}$ as

$$\ell_{test}^{i}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)=\varphi\left(\tau_{e,test}^{i}\mid\mu_{e,Ptrue},\sigma_{e,Ptrue}\right) \tag{13}$$

Note that the likelihood function is not a probability distribution, but a conditional probability. The numerical scheme to evaluate the updated joint PSF is explained in Appendix B.

Using the updated joint PDF, the marginal PDFs of $\mu_{e,Ptrue}$ and $\sigma_{e,Ptrue}$ can be obtained as

$$f_{\mu e,Ptrue}^{upd}\left(\mu_{e,Ptrue}\right)=\int_{0}^{\infty}f^{upd}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)d\sigma_{e,Ptrue} \tag{14}$$

$$f_{\sigma e,Ptrue}^{upd}\left(\sigma_{e,Ptrue}\right)=\int_{-\infty}^{\infty}f^{upd}\left(\mu_{e,Ptrue},\sigma_{e,Ptrue}\right)d\mu_{e,Ptrue} \tag{15}$$

The above distributions represent the uncertainty in estimating mean and standard deviation of the element strength. If a conservative prediction is wanted, the lower 5th percentile of $f_{\mu e,Ptrue}^{upd}$ can be used for the 95% confidence level. The mean values of distributions in Eqs. (14) and (15) are, respectively, the estimate of the mean strength and its standard deviation. The standard deviations of distributions in Eqs. (14) and (15) are measures of remaining uncertainty after the element tests.

IV. Assessing the Merits of a Combination of Number of Element Tests and Number of Coupon Tests

The objective of this section is to assess the effect of coupon and element tests on reducing uncertainty, estimating conservative allowables, and weight penalty. For that purpose, a single set of test results is generated to calculate the conservative estimate of the mean strength and also to compute the weight penalty due to conservativeness. These results are compared with the weight obtained with an infinite number of coupon and element tests. It is also checked that if the resulting design is actually conservative as expected. Since the results with a single set of test measurements are likely to be biased due to sampling error, the above process is repeated (100,000 times) to deduce the average weight penalty and the probability of unconservative design (PUD).

The 5th percentile of the marginal PDF for the mean of element strength is a conservative estimate that is expected to be less than the true mean of element strength with the 95% confidence level, and its mean is a conservative estimator of the true mean of element strength $(\mu_{e,true} = 0.95)$. The 5th percentile of the mean element strength, $\tau_{0.05}$, can be calculated using Eq. (15) as

$$0.05 = \int_{-\infty}^{\tau_{0.05}} f_{\mu e,Ptrue}^{upd}\left(x\right) dx \qquad (16)$$

With an infinite number of tests, prediction should be the same with the true element mean, $\mu_{e,true}$, regardless of variability. If a truss member is designed with an axial load and the mean of element strength, the weight penalty due to the conservativeness in the 5th percentile is calculated as

$$w_i = \left(\frac{A_{0.05}^i}{A_\infty} - 1 \right) \times 100 = \left(\frac{F / \tau_{0.05}^i}{F / \mu_{e,true}} - 1 \right) \times 100 = \left(\frac{\mu_{e,true}}{\tau_{0.05}^i} - 1 \right) \times 100 \ (\%) \qquad (17)$$

where the index i represents i^{th} set of tests results.

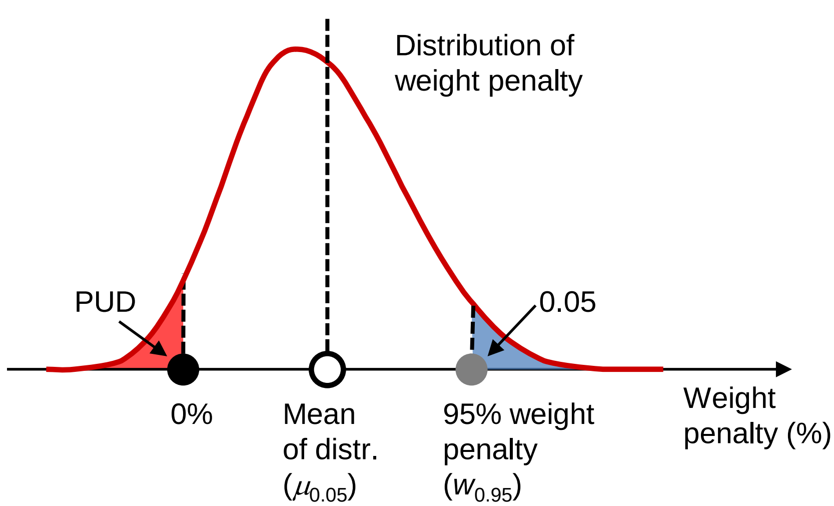


Figure 3. Distribution of weight penalty due to the variability in tests

When w_i is 3% for (10/5), it means that a design with 10 coupon tests and 5 element tests is 3% heavier than a design with an infinite number of tests. Negative weight penalty indicates that the design is unsafe because it implies that the 5th percentile mean strength is less than the true mean. The area of the weight penalty distribution on the left side to zero is the probability of unconservative design (PUD) (see Fig. 3).

Figure 3 illustrates the weight penalty distribution, mean weight penalty, 95% weight penalty and probability of PUD. The 0% weight penalty (the black filled circle) represents design weight with an infinite number of tests. The mean of the weight penalty (the hollow circle) represents expected conservativeness in the design. The 95% weight penalty (the grey filled circle) can be interpreted as a possible very conservative design weight with 5% probability. PUD (the area left to 0% weight penalty) is the true probability of unsafe design. When we design the truss with 5th percentile mean strength, we expect that the design will be unsafe with only 5% probability. PUD shows the actual probability of unsafe design. Those measures are calculated from N sets of test results, or N repetitions (N = 100,000 here) as follows:

$$\mu_{0.05} = \frac{1}{N}\sum_{i=1}^{N} w_i \tag{18}$$

$$0.95 = \frac{1}{N}\sum_{i=1}^{N} I\left(w_i < w_{0.95}\right) \tag{19}$$

$$\text{PUD} = \frac{1}{N}\sum_{i=1}^{N} I\left(\tau_{0.05}^{i} < \mu_{e,true}\right) \tag{20}$$

This procedure needs to be performed for different realizations of the epistemic uncertainty. Here, for illustration, we repeat it only for four values, 1% and 5% unconservative errors and 1% and 5% conservative errors. These appear to be sufficient to illustrate the effect of different values of the epistemic uncertainty.

V. Illustrative Example

In this section, the effect of the number of tests is investigated in two steps. First, the conservative mean of the element strength is predicted using a single set of tests, and then, average prediction is estimated with multiple sets of tests.

A. The effect of the number of tests with a single set of tests

In this section, estimation of mean element strength is illustrated with a single set of coupon and element tests. The test results were randomly generated from the true distributions defined in Table 2. The difference between the element mean and the coupon mean represents error in the failure theory as assumed in Eq. (7). Since $k_{3d,calc} = 1.0$ is assumed in this paper and

$k_{3d,true} = \mu_{e,true} / \mu_{c,true}$ is 0.95, the failure theory overestimates the element strength; that is, the error in the failure theory is unconservative. Randomly generated test results are given in Table 3. For example, for 10 coupons and 3 elements (10/3), the mean and standard deviation of coupons were 0.972 and 0.091, respectively, and the three element test results were first three data (i.e., 0.945, 0.955 and 0.987). The true distribution is only used for the purpose of simulating tests.

Table 2. True distributions of coupon and element tests

Test	Distribution	Parameters
Coupon test	Normal	$\mu_{c,true}$ = 1.0, COV 7%
Element test	Normal	$\mu_{e,true}$ = 0.95, COV 3%

Table 3. Statistics for coupon and element tests

No. of coupon tests	Coupon test	Element tests (order by sequence)
10	$\mu_{c,test} = 0.972$, $\sigma_{c,test} = 0.091$	0.945, 0.955, 0.987, 0.953, 0.935 $\mu_{e,test} = 0.955$, $\sigma_{e,test} = 0.0193$
50	$\mu_{c,test} = 1.004$, $\sigma_{c,test} = 0.073$	0.896, 0.981, 0.939, 0.998, 0.957 $\mu_{e,test} = 0.954$, $\sigma_{e,test} = 0.039$
90	$\mu_{c,test} = 1.001$, $\sigma_{c,test} = 0.070$	0.917, 0.989, 0.954, 0.939, 0.948 $\mu_{e,test} = 0.949$, $\sigma_{e,test} = 0.026$

To estimate the mean of element strength, the prior is constructed based on the coupon test results and error bounds as shown in Eqs. (10) and (11). Table 4 gives the standard deviation of normal error distribution for mean and bounds for standard deviation $\left[\sigma_e^{lower}, \sigma_e^{upper}\right]$. Recall that the error bounds represent the current estimate of the maximum error in the failure theory and the error bounds are used to establish the normal error distribution that the error bounds and symmetric 95% confidence interval have the same domain. Detailed procedure of numerical calculation is given in Appendix B.

Table 4. Error distributions of element tests

Error	Distribution	Std / Bounds
b_e	Normal	$0.1/\Phi(0.975)$
$\left[\sigma_e^{lower}, \sigma_e^{upper}\right]$	Uniform	$[0, 0.04]$

Table 5 summarizes the 5th percentile value ($\tau_{0.05}$) and the weight penalty after Bayesian update. It is observed that the effect of element tests is more significant than that of coupon tests. As the number of element tests increases between n_e = 1 and n_e = 5, weight penalty decreases from 4-6% to 1.4-2.3%, and 5th percentile strength converges to 0.95 monotonically. However the effect of the number of coupon tests is ambiguous and no clear trend is observed. This is because the error in the failure theory (Table 4) is much larger than that in measuring uncertainty of coupons. For the 50 and 90 coupons cases, n_e = 1 estimates more conservativeness than n_e = 0 because the particular element test results happen to be very conservative, as shown in Table 3 (0.896 and 0.917 from a normal distribution with the mean of 0.95 and the standard deviation of 0.0285).

Table 5. Estimates of the conservative element strength and the resulting weight penalty (compared to infinite number of tests) from a single set of test results ($\mu_{e,true}$ = 0.95: Unconservative +5% error in failure theory)

	n_c \ n_e	0	1	3	5
5th percentile	10	0.879	0.916	0.938	0.938
Weight penalty		8.1%	3.7%	1.2%	1.3%
5th percentile	50	0.917	0.882	0.917	0.934
Weight penalty		3.5%	7.7%	3.6%	1.7%
5th percentile	90	0.916	0.900	0.930	0.932
Weight penalty		3.7%	5.6%	2.1%	1.9%

B. The effect of the number of tests averaged over multiple sets of tests

The results from the previous subsection depend on the particular samples of coupons and elements. In order to measure the expected effect of tests, the same process is repeated 100,000 times. Different test results are randomly generated and used for each time. Weight penalties are generated with the 100,000 test sets. The effects of the number of tests on the weight penalty are analyzed with three measurements, mean of the weight penalty, 95th percentile of the weight penalty, and probability of unsafe design (PUD), described in Section IV. Two scenarios associated with epistemic uncertainty in the failure theory are considered. The first scenario addresses the effect of relatively large epistemic uncertainty in the failure theory ($b_e = \pm 10\%$) compared to that in coupon samples. With 7% COV in material strength, the uncertainty in the mean coupon strength is small even with 10 coupons. The second scenario examines the effect of relatively small epistemic uncertainty in the failure theory ($b_e = \pm 2\%$). Note that the bounds are used to establish normal error distributions. Each scenario is further divided into two cases: unconservative and conservative failure theory. The true mean of element tests and its error bounds are set to reflect each scenario as shown in Table 6, the other settings are the same with the previous single set example.

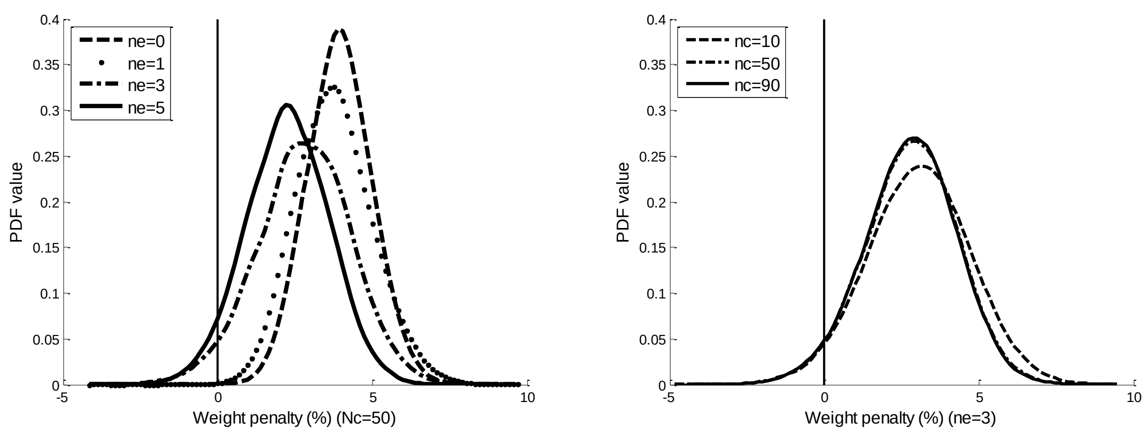
Table 6. Four scenarios associated with epistemic uncertainty in failure theory and corresponding example settings (COV of 7% in material strength is assumed)

Magnitude of error in failure theory	Failure theory	True mean of element test	Error distribution
Large epistemic uncertainty in failure theory	Unconservative	$\mu_{e,true}$ = 0.95	$N(0, 0.1/\Phi(0.975))$ (standard deviation of 5.1% error)
	Conservative	$\mu_{e,true}$ = 1.05	

Small epistemic uncertainty in failure theory	Unconservative	$\mu_{e,true} = 0.99$	$N(0, 0.02/\Phi(0.975))$ (standard deviation of 1.0% error)
	Conservative	$\mu_{e,true} = 1.01$	

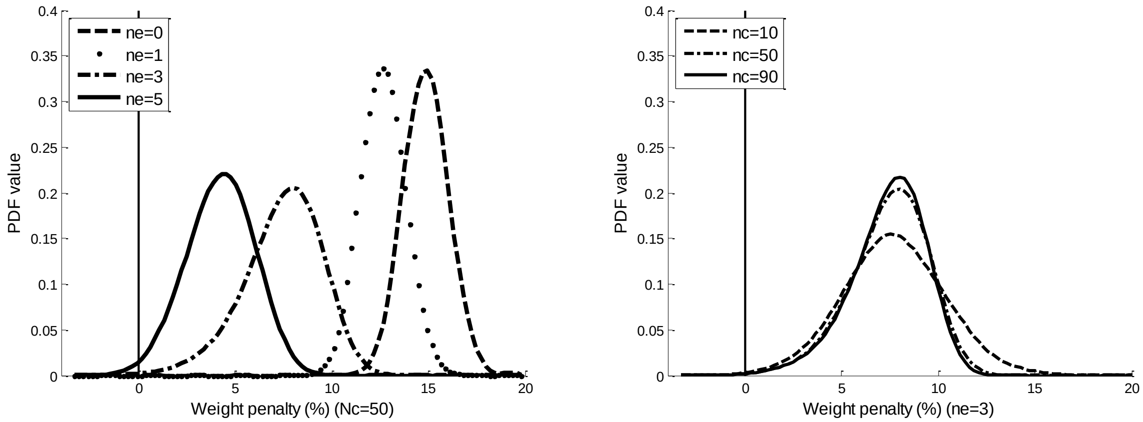
When the failure theory has relatively large epistemic uncertainty, the distributions of the weight penalties as functions of the number of tests are shown in Fig. 4 for both conservative and unconservative failure theories. n_c=50 and n_e=3 are assumed as the nominal numbers of tests. The effects of the number of element tests and the number of coupon tests are shown around the nominal numbers. Figure 4 shows that n_e is far more influential than n_c for shifting the distribution to less conservative region and narrowing it.

With no element tests, the distribution is narrow, since it represents only the sampling uncertainty in 50 coupon tests. As the number of element tests increases, the distribution is first widened for a single element test, because a single test is quite variable, and then, gradually narrowed. The updated distribution is also shifted closer to 0% weight penalty. For the unconservative case, Figure 4(a), the shift is small, because the conservativeness in the design with the unconservative failure theory is small. However, for the conservative case, Figure 4(c), the shift is large since the conservative failure theory provides very conservative design.



(a) n_e = 0,1,3,5 with n_c = 50 ($\mu_{e,true}$ = 0.95 and $b_e = \pm 10\%$)

(b) n_e = 3 with n_c = 10,50,90 ($\mu_{e,true}$ = 0.95 and $b_e = \pm 10\%$)



(c) n_e = 0,1,3,5 with n_c = 50 ($\mu_{e,true}$ = 1.05 and $b_e = \pm 10\%$)

(d) n_e = 3 with n_c = 10,50,90 ($\mu_{e,true}$ = 1.05 and $b_e = \pm 10\%$)

Figure 4. Distributions of weight penalties for comparison between the number of coupon tests and the number of element tests

Tables 7 and 8 summarize the distributions with three statistics, mean weight penalty, 95% weight penalty and PUD in terms of the number of tests.

Table 7. Mean, 95th percentile of weight penalties and probability of unsafe design (PUD) ($\mu_{e,true}$ = 0.95: Unconservative +5% error in failure theory)

	0	1	3	5
	Mean			
10	4.6%	4.3%	3.0%	2.3%
50	3.9%	3.8%	2.8%	2.3%
90	3.8%	3.7%	2.8%	2.2%
	95th perc. of weight penalty factor (extreme design weight)			
10	8.6%	7.5%	5.8%	4.7%
50	5.6%	5.9%	5.2%	4.4%
90	5.1%	5.7%	5.2%	4.4%
	Probability of unsafe design (PUD)			
10	2.2%	1.0%	3.6%	4.5%
50	0%	0.3%	3.7%	4.6%
90	0%	0%	3.7%	4.6%

Table 8. Mean, 95th percentile of weight penalties and probability of unsafe design (PUD) ($\mu_{e,true}$ = 1.05: Conservative -5% error in failure theory)

	0	1	3	5
	Mean			
10	15.6%	13.2%	7.7%	4.3%

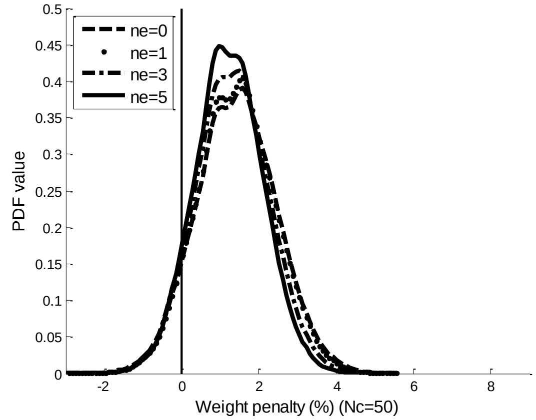
50	14.9%	12.6%	7.4%	4.2%
90	14.8%	12.6%	7.4%	4.2%
	95th perc. of weight penalty factor (extreme design weight)			
10	20.0%	17.5%	12.1%	7.5%
50	16.8%	14.5%	10.4%	7.0%
90	16.2%	14.0%	10.2%	6.9%
	Probability of unsafe design (PUD)			
10	0%	0%	0.2%	1.0%
50	0%	0%	0.2%	0.9%
90	0%	0%	0.2%	0.9%

We first consider the case of minimal testing with only 10 coupon tests and no element tests. For the case of unconservative failure theory, Table 7, minimal testing will cost us 4.6% weight penalty on average, and 2.2% chance that we will end up with unconservative design. For the case of conservative failure theory, Table 8, the weight penalty shoots up to 15.6% and we do not run the chance of unconservative design. The weight penalties with the 95th percentiles (corresponding to tests that happen to be on the conservative side) are about 11% higher.

Both the weight penalties and the PUD continue to drop substantially with more element tests. On the other hand, the effect of adding coupon tests is much smaller, and going from 50 to 90 coupon tests hardly make any difference.

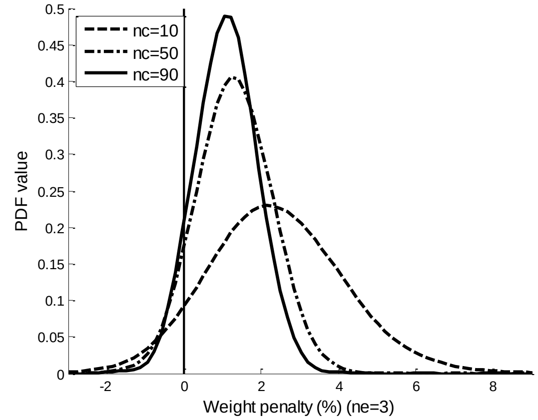
The fact that for this example, element tests are more important than coupon tests can be understood by observing the magnitude of two epistemic uncertainties. The variability in the strength is 7% (see Table 2), so even with 10 coupon tests, the standard deviation of the mean coupon strength is only 2.2% which is epistemic uncertainty in sampling. On the other hand, with ±10% error bounds, the standard deviation of the epistemic uncertainty in the failure theory is 5.1%. This is why element tests were more significant in reducing uncertainty. If, on the other hand the failure theory was much more accurate, then it is expected element tests to be less significant. For example, ±2% error bounds, the magnitude of the epistemic uncertainty in failure theory is merely 1.0%. With such an accurate failure theory, it turned out that the number of coupon tests becomes more influential than the number of element tests.

It turned out that increasing the number of element tests is more important than increasing the number of coupon tests when we have the large epistemic uncertainty (±10%) in the failure theory. However, when the epistemic uncertainty is small (±2%), the number of coupon tests becomes more influential than the number of element tests. In parallel to Fig. 4, Fig. 5 shows a comparison between the effect of n_c and the effect of n_e on the weight penalty when the error in the failure theory is small. It is clearly seen that the effect of the number of coupon tests is more influential than the number of element tests for decreasing chance of having very conservative designs and reducing the variation of design.



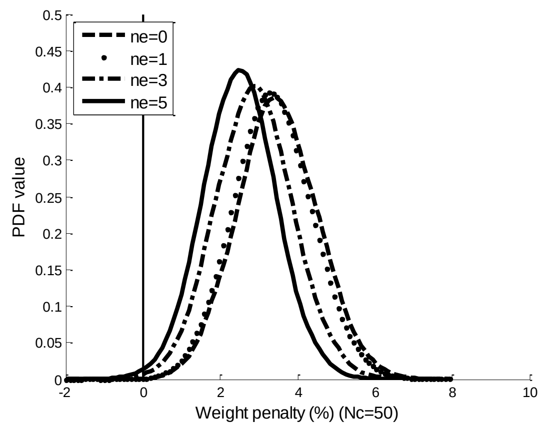
(a) n_e = 0,1,3,5 with n_c = 50

($\mu_{e,true}$ = 0.99 and $b_e = \pm 2\%$)



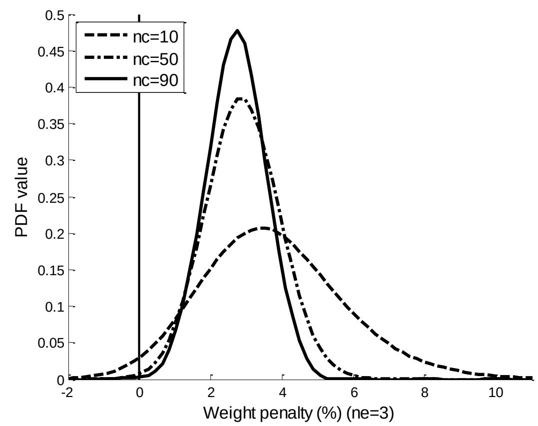
(b) n_e = 3 with n_c = 10,50,90

($\mu_{e,true}$ = 0.99 and $b_e = \pm 2\%$)



(a) n_e = 0,1,3,5 with n_c = 50

($\mu_{e,true}$ = 1.01 and $b_e = \pm 2\%$)



(b) n_e = 3 with n_c = 10,50,90

($\mu_{e,true}$ = 1.01 and $b_e = \pm 2\%$)

Figure 5. Distributions of weight penalties for comparison between the number of coupon tests and the number of element tests

Compared to Tables 7 and 8, the increased accuracy of the failure theory reduces substantially the penalty associated with no element tests. For 10 coupon tests, the weight penalty for no element tests is reduced from 4.6% to 3.2% for unconservative error (Tables 7 and 9) and from 15.6% to 5.2% for conservative errors (Tables 8 and 10). Also, because the epistemic uncertainties associated with the failure theory are no comparable to the epistemic uncertainties in the mean of the coupon tests, the effect of the number of coupon tests and element tests becomes comparable. Going from one element tests to five element tests for 10 coupon tests reduces the weight penalty from 3.2% to 2% (Table 9) and from 5.2% to 3.0% (Table 10). In comparison, increasing the number of coupon tests from 10 to 90 for one element test, reduces the 3.2% and 5.2% to 1.2% and 3.2%, respectively.

Table 9. Mean, 95th percentile of weight penalties and probability of unsafe design (PUD)
($\mu_{e,true}$ = 0.99: Unconservative +1% error in failure theory)

	0	1	3	5
	Mean			
10	3.2%	3.0%	2.4%	2.0%
50	1.4%	1.4%	1.3%	1.2%
90	1.2%	1.1%	1.1%	1.1%
	95th perc. of weight penalty factor (extreme design weight)			
10	7.4%	6.8%	5.4%	4.3%
50	3.1%	3.1%	2.9%	2.7%
90	2.4%	2.4%	2.4%	2.3%
	Probability of unsafe design (PUD)			
10	10.0%	8.9%	8.0%	7.0%
50	8.0%	7.6%	8.3%	7.9%
90	6.1%	6.2%	7.6%	7.5%

Table 10. Mean, 95th percentile of weight penalties and probability of unsafe design (PUD)
($\mu_{e,true}$ = 1.01 Conservative -1% error in failure theory)

	0	1	3	5
	Mean			
10	5.2%	4.9%	3.8%	3.0%
50	3.5%	3.3%	2.9%	2.5%
90	3.2%	3.1%	2.7%	2.4%
	95th perc. of weight penalty factor (extreme design weight)			
10	9.6%	8.9%	7.3%	5.8%
50	5.2%	5.1%	4.5%	4.0%
90	4.5%	4.4%	4.0%	3.6%
	Probability of unsafe design (PUD)			
10	1.5%	1.4%	1.9%	2.1%
50	0%	0%	0.2%	0.5%
90	0%	0%	0%	0.2%

C. The accuracy of convolution integral on calculating a conditional distribution

It has been shown that a double-loop MCS can be used to calculate the distribution in Eq. (9). However, MCS has a computational challenge in the tail region (low-probability region) as well as sampling error. For example, 10^{-4} level of probability can be hardly estimated with 10,000 samples. Different from MCS, convolution integral can calculate a nearly exact distribution without having sampling errors. In this section, the accuracy of convolution integral is compared with that of MCS.

In order to illustrate the advantage of convolution integral, the probability of the product of two random variables, $\hat{Z} = \hat{X} \times \hat{Y}$, are used. It is assumed that the two independent random variables are defined as $\hat{X} \sim N(1.1, 0.0096)$ and $\hat{Y} \sim U(0.9, 1.1)$. For MCS, one million samples are used to evaluate the probability at Z value of 0.955 and 0.975. Since MCS has sampling error, this process is repeated 1,000 times, and then, the mean and standard deviation are listed in Table 11. For convolution integral, the entire range is divided by 50 segments, and three-point Gauss quadrature is used in integrating Eq. (9) with b_e=0.1, $\mu_{c,test=}1.1$, and $\sigma_{c,test} / \sqrt{n_c} = 0.0096$. The results only differ by 0.2% when 400 segments are used. Different from MCS, there is no need for repetition because convolution integration does not have sampling error.

Table 11. Probability of Z at two different values

Z value		0.955	0.975
MCS	Mean	2.34×10^{-7}	6.77×10^{-4}
	COV	210.7%	3.9%
Convolution integral		2.40×10^{-7}	6.78×10^{-4}

When the probability is of the order of 10^{-4}, MCS has about 3.9% coefficient of variance (COV), while the convolution integral shows a very little calculation error. When the probability is of the order of 10^{-7}, the MCS with 1 million samples is not meaningful as reflected in the COV value of 210%. However, the convolution integral is still accurate, and the value can be obtained by a one time calculation. Note that the estimated error in the mean PF with 1000 repetition can be calculated as $4.93\times10^{-7}/1000^{0.5} = 1.56\times10^{-8}$.

VI. Conclusions

In this paper, the effect of the number of coupon and element tests on reducing conservativeness and weight penalties due to the uncertainty in structural element strength was studied. Two sources of epistemic uncertainties were considered: (a) the uncertainty in sampling in measuring material variability and (b) the uncertainty in the failure theory. A large number of coupon test reduce the uncertainty in measuring material variability, while element tests reduce the uncertainty in the failure theory. These uncertainties were combined using convolution integral, which is more accurate and robust than MCS. Then, Bayesian inference was used to update this uncertainty with element test results. Because test results can vary, a large number of simulations were used to obtain mean performance and distributions.

For a typical case of ±5% error in the failure theory, 7% and 3% coefficients of variation in material strength and element strength, element tests were found to be very important in reducing weight penalties from about 16% with no tests, to about 4% with five element tests. The effect of the number of coupon tests was much smaller because sampling uncertainty was much smaller than the uncertainty in the failure theory. For ±1% error in the failure theory and the same coefficient of variations in material strength and element strength, the effect of the number of coupons became comparable to that of element tests. The methodology developed would thus allow designers to estimate the weight benefits of tests and improvements in failure predictions.

Acknowledgement

Authors would like to thank the National Science Foundation for supporting this work under the grant CMMI-0856431.

References

[1] Noh, Y., Choi, K. K., Lee, I., Gorsich, D., and Lamb, D., "Reliability-based design optimization with confidence level under model uncertainty due to limited test data" *Structural and Multidisciplinary Optimization*, 43(4), pp. 443–458.

[2] Matsumura, T., Haftka, R.T. and Sankar, B.V. " Reliability Estimation Including Redesign Following Future Test for an Integrated Thermal Protection System" *9th World Congress on Structural and Multidisciplinary Optimization, Shizuoka, Japan, June 14-17, 2011*.

[3] Villanueva, D., Haftka, R.T. , Sankar, B.V. "Accounting for Future Redesign in the Optimization of an Integrated Thermal Protection System", *AIAA-2012-1933, 14th AIAA Non-Deterministic Approaches Conference, Honolulu, HI, 2012*.

[4] Joseph M. Manter and Donald B. Paul, Airframe Structures Technology for Future Systems, ICAS 2000.

[5] Lincoln, J.W., "USAF Experience in the Qualification of Composite Structures", Composite Structures; Theory and Practice, ASTM STP 1383, 2000, pp. 3-11.

[6] Jiao, G., and Moan, T., "Methods of reliability model updating through additional events," Structural Safety, 9(2), 1990, pp. 139-153.

[7] An, A., Acar, A., Haftka, R.T., Kim, N.H., Ifju, P.G, and Johnson, T.F., (2008) "Being Conservative with a Limited Number of Test Results," *Journal of Aircraft*, 45(6), pp. 1969-1975.

[8] Acar, E., Haftka, R.T., Kim, N.H.(2010) "Effects of Structural Tests on Aircraft Safety " AIAA Journal, 48(10), pp. 2235–2248.

[9] X. Jiang, S. Mahadevan, (2007) "Bayesian risk-based decision method for model validation under uncertainty", Reliability Engineering & System Safety, 92: pp. 707-718.

[10] A. Urbina, S. Mahadevan and T. L. Paez, (2011) "Quantification of margins and uncertainties of complex systems in the presence of aleatoric and epistemic uncertainty", Reliability Engineering & System Safety, 96 (9): pp. 1114-1125.

[11] I. Park, H. K. Amarchinta, R. V. Grandhi, (2010) "A Bayesian approach for quantification of model uncertainty", Reliability Engineering & System Safety, 95, pp. 777-785.

[12] J. M. McFarland, B. J. Bichon, Bayesian model averaging for reliability analysis with probability distribution model from uncertainty. 50th AIAA/ASME/ASCE/AHS/ASC structures, structural dynamics and material conference, Palm Springs, CA, 2009.

[13] M. Evans, N. Hastings, and B. Peacock, *Statistical distributions*, Wiley, New York, 1993.

[14] Kale, A.A., Haftka, R.T., and Sankar, B.V., (2008) "Efficient Reliability Based Design and Inspection of Stiffened Panels against Fatigue," *Journal of Aircraft*, 45(1), pp. 86-97

[15] Park, C., Matsumura, T., Haftka, R. T., Kim, N. H. and Acar.,E., "Modeling the effect of structural tests on uncertainty in estimated failure stress" 13th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, Fort Worth, Texas, Sept. 13-15, 2010.

[16] M. E. Johnson, G. L. Tietjen, and R. J. Beckman (1980) A New Family of Probability Distributions With Applications to Monte Carlo Studies, Journal of the American Statistical Association, 75(370), pp. 276- 279

[17] M. McDonald and S. Mahadevan, Uncertainty Quantification and Propagation in Multidisciplinary Analysis and Optimization, AIAA-2008-6038, 12th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference, 2008

[18] E.J. Dudewicz and Karian, Z.A, The Extended Generalized Lambda Distribution (EGLD) for Fitting Distribution with Moments, (1996) American Journal of Mathematical and Management Science 16, pp. 271-332.

[19] E. L. Lehmann and G. Casella, *Theory of point prediction*, Springer-Verlag, New York, 1998.

[20] Composite Materials Handbook MIL-HDBK-17-1F, "Composite Materials Handbook," ASTM Publications 2002

[21] R. E. Neapolitan, *Learning Bayesian Network*, Pearson Education, New Jersey, 2004.

[22] Oberkampf, W., J. Helton, C. Joslyn, S. Wojtkiewicz, and S. Ferson. Challenge problems: uncertainty in system response given uncertain parameters. *Reliability Engineering and System Safety*, 85(1-3) (2004)

[23] "Guidelines for Property Testing of Composites," Composite Materials Handbook MIL-HDBK-17, U.S. Dept. of Defense, Washington, D.C., 2002.

[24] "Bayesian Statistics," Adrian Raftery and Jeff Gill. One-day course for the American Sociological.

APPENDIX A: Statistical formulation of possible true distribution of mean and standard deviation of element strength

The PTD of element mean failure strength can be expressed as

$$f_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\right)=\int_{-\infty}^{\infty} f_{\mu e,Ptrue}\left(\mu_{e,Ptrue} \mid \mu_{c,Ptrue}\right) f_{\mu c,Ptrue}\left(\mu_{c,Ptrue}\right)\mathrm{d}\mu_{c,Ptrue} \tag{A1}$$

which is in the form of the convolution integral. The conditional PDF $f_{\mu e,Ptrue}\left(\mu_{e,Ptrue} \mid \mu_{c,Ptrue}\right)$ corresponds to the distribution of $\hat{k}_{3d,Ptrue}$. In the following, the two PDFs in the integrand will be explained.

In this paper, $k_{3d,calc}=1$ is used for simplicity, and it is assumed that $e_{k,Ptrue}$ follows a normal distribution that has symmetric 95% confidence interval defined with $\pm b_e$ bounds as

$$f_{k,Ptrue}\left(e_{k,Ptrue}\right)=\varphi\left(e_{k,Ptrue} \mid 0,\frac{b_e}{\Phi\left(0.975\right)}\right) \tag{A2}$$

where the notation $\varphi\left(x \mid a,b\right)$ denotes the value of normal PDF with mean *a* and standard deviation *b* at *x*.

By using Eq. (A2), $f_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\right)$ can be obtained from all possible combinations of random variables generated from $f_{k,Ptrue}\left(e_{k,Ptrue}\right)$ and $f_{\mu c,Ptrue}\left(\mu_{c,Ptrue}\right)$. For a given sample of $\mu_{c,Ptrue}$, the PTD of element failure strength can be regarded as a conditional PDF $f_{\mu_{e,Ptrue}}\left(\mu_{e,Ptrue} \mid \mu_{c,Ptrue}\right)$, which is derived from Eq. (A2).

$$f_{\mu e,Ptrue}\left(\mu_{e,Ptrue} \mid \mu_{c,Ptrue}\right)=\varphi\left(\mu_{e,Ptrue} \mid \mu_{c,Ptrue},\frac{\mu_{c,Ptrue}b_e}{\Phi\left(0.975\right)}\right) \tag{A3}$$

The PDF in Eq. (A3) represents the epistemic uncertainty in failure theory. The PTD $f_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\right)$ can be calculated by considering all possible values of $\mu_{c,Ptrue}$ with Eq. (A3).

PDF of the PTD of $\mu_{c,Ptrue}$ is calculated from coupon test results as

$$f_{\mu c,Ptrue}\left(\mu_{c,Ptrue}\right)=\varphi\left(\mu_{c,Ptrue}\mid\mu_{c,test},\frac{\sigma_{c,test}}{\sqrt{n_c}}\right) \qquad \text{(A4)}$$

Samples of $\mu_{c,Ptrue}$ are generated from Eq. (A4), which is then used in Eq. (A3) to generate samples of $\mu_{e,Ptrue}$. Figure A1 illustrates the conditional PDF of $\mu_{e,Ptrue}$ for a given sample of $\mu_{c,Ptrue}$, which is drawn from $f_{\mu c,Ptrue}\left(\mu_{c,Ptrue}\right)$ based on $\mu_{c,test}$. Note that $\mu_{e,true}$ is given as a unique value, and is expressed by the PTD $f_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\mid\mu_{c,Ptrue}\right)$.

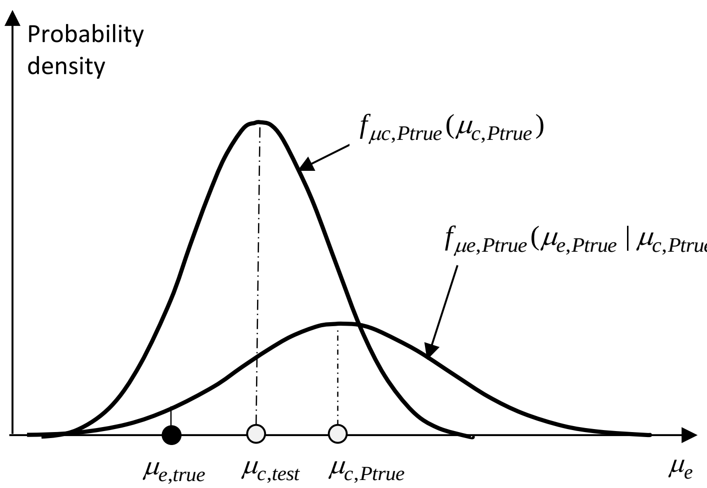


Figure. A1: The possible true distribution of mean failures strength of specimens and the conditional distribution of the element mean failure strength.

With Eq. (A3) and Eq. (A4), the convolution integral in Eq. (A2) can be directly integrated as

$$f_{\mu e,Ptrue}\left(\mu_{e,Ptrue}\right)=\int_{-\infty}^{\infty}\varphi\left(\mu_{e,Ptrue}\mid\mu_{c,Ptrue},\frac{\mu_{c,Ptrue}b_e}{\Phi\left(0.975\right)}\right)\varphi\left(\mu_{c,Ptrue}\mid\mu_{c,test},\frac{\sigma_{c,test}}{\sqrt{n_c}}\right)d\mu_{c,Ptrue} \qquad \text{(A5)}$$

The PDF in Eq. (A5) is a prior distribution of mean failure strength of elements, which includes the effect of uncertainty from failure theory as well as that of a finite number of samples.

APPENDIX B: Numerical scheme to obtain the presented results

For the mean element strength, a range of [0.78, 1.22] was found to be large enough for capturing the updated joint probability distribution. Since the initial distribution for the mean element strength has very little influence on posterior distribution on both tails. The standard deviation is bounded in [0, 0.04] as noted in Table 4. In order to calculate the updated distribution from Bayesian inference, each range is discretized into 200 equal intervals, and this discretization generates a 200 by 200 grid. The updated joint PDF is calculated at each grid point using Eq. (12). Then, the prior is updated using likelihood function with different numbers of element tests; i.e. n_e= 1, 3, and 5.

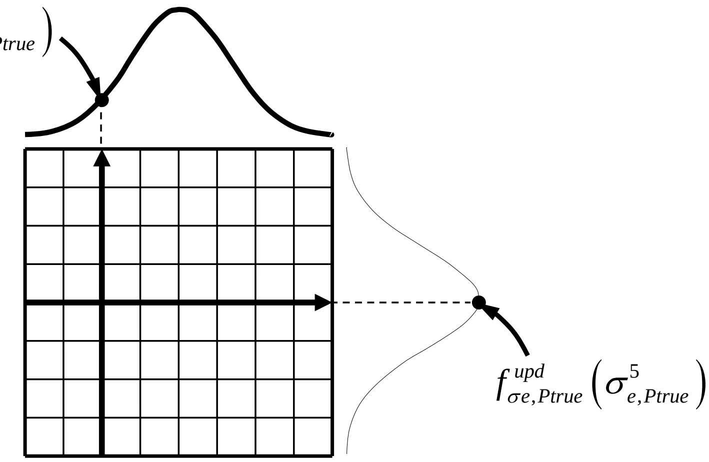


Figure. B1: A 8 by 8 grid for obtaining a joint PDF and its marginal PDFs

The marginal updated distributions are obtained using the updated joint distribution as expressed in Eqs. (14) and (15). For the updated marginal element mean distribution, conditional PDFs for given 201 mean element strength are integrated over 201 points using Gaussian quadrature with 2 points. Figure B1 shows an equivalent example that has 8 by 8 grid. The marginal distribution of the updated mean element strength is formed by calculating PDF values on 9 given mean values. $f^{upd}_{\mu e,Ptrue}\left(\mu^3_{e,Ptrue}\right)$ is equal to a value obtained by integrating a conditional PDF of the standard deviation for $\mu_{e,Ptrue} = \mu^3_{e,Ptrue}$ over the vertical arrow.