AIAA JOURNAL
Vol. 55, No. 10, October 2017

Sampling by Exploration and Replication for Estimating Experimental Strength of Composite Structures

Yiming Zhang*
University of Florida, Gainesville, Florida 32611-6250
Jaco F. Schutte†
Materials Sciences Corporation, Horsham, Pennsylvania 19044
Waruna P. Seneviratne‡
National Institute for Aviation Research, Wichita, Kansas 67260
and
Nam H. Kim§ and Raphael T. Haftka¶
University of Florida, Gainesville, Florida 32611-6250

DOI: 10.2514/1.J055862

The industry has to resort to experiments for practical design of composite laminates when physical models/simulations are inadequate for desirable accuracy or require excessive computational resources. Surrogates are used to predict strength of composite laminates in the design space by conducting an array of tests. It has been reported that an exploration strategy to test as many different configurations as possible is more effective than replication of fewer points for reducing test noise. The observation was based on analytical test functions and synthetic Gaussian noise. This paper first studies real experiments to check whether the previous observation stands. Test results of open-hole tension for composite plates that included 18 replicates per configuration were examined. A resampling procedure was developed that compared exploration and replication, and it was found that exploration proved to be more accurate for prediction than replication. Second, the major source of uncertainty for surrogate prediction was examined, which is variability of strength. The distribution of experimental open-hole tension strength was found to be not unambiguously independent and was identically distributed normal distribution as commonly assumed. The variation of specimen strength is correlated rather than independent at different configurations due to the between-batch variability. Consequently, the influence of distribution type was then investigated on an analytical test function with three synthetic distributions. The exploration strategy proved to be better than the replication strategy for all three distributions. It was found that the exploration strategy allows for higher-order polynomial surrogate to be used, which is a key point to improve characterization of a function with complex dependence on design parameters.

I. Introduction

COMPOSITE materials have been routinely used in aerospace applications due to their outstanding capability to be tailored to specific load paths and conditions, resulting in weight-efficient designs. To achieve weight savings, effective and accurate characterization of structural strength is essential. Significant progress has been achieved on composite mechanics in past decades. Failure criteria have demonstrated reasonable accuracy for predicting strength of benchmark structures. For example, Whitney and Nuismer [1] proposed a failure criterion to predict strength of composite laminates with a hole. Tsai and Wu [2] proposed a phenomenological material failure theory that is widely used for anisotropic composite materials. The World Wide Failure exercises [3,4] summarized the effort to provide experimental data and benchmark different modeling strategies for failure criteria of composites.

Various commercial software is available to simulate the response of composite structures and used for routine analysis. However, each new material system, structural configuration, and fabrication process requires a large, costly, and time-consuming program to obtain simulation with reasonable accuracy [5,6]. Innovative approaches are being developed to simulate specific structures with certain configurations such as open-hole tension test [7]. Large challenges still remain for simulating strength of composite structures with complicated failure mechanisms and variability in failure response due to progressive damage.

Industry has resorted to empirical approaches where testing has been the focus for the characterization of structural failure when simulations are inadequate for desirable accuracy or require excessive computational resources. The quantity of interest for experimental analysis may be as simple as an averaged pass/fail criteria based on a single load or may be extended to include mixed mode loadings and more involved statistical analysis such as strength allowables [8–10]. Handler et al. [11] discussed the experimental procedures to develop a test database for composite structures. Carlsson et al. [12] and the Composite Materials Handbook-17 provide in-depth guideline for systematic experimental analysis.

Experimental results are usually obtained for an array of configurations for different combinations of important design parameters. Then surrogate models (such as polynomial response surfaces) are often fit to the data to estimate the structural response over design space. Forrester et al. [13] discussed about approximating noisy data with Kriging. Glaz et al. [14] adopted multiple surrogate models for design optimization to reduce the vibration of rotor-blade problem. Chaudhuri and Haftka [15] proposed an adaptive sampling for reliability-based design optimization. In this context, a surrogate model serves three purposes. First, a surrogate allows for the statistical averaging or weighed averaging of multiple test results or replicates at the same configuration (i.e., multibatch testing). Second, it allows interpolation or extrapolation of the failure response at locations in the design space where test data are not available. Last, it yields a functional

Received 21 November 2016; revision received 11 April 2017; accepted for publication 31 May 2017; published online 31 July 2017.

*Doctoral Candidate, Mechanical and Aerospace Engineering, 333 MAE-A; yimingzhang521@ufl.edu. Student Member AIAA.

†Senior Engineer, Materials Sciences Corporation, 135 Rock Road; schutte@materials-sciences.com.

‡Research Scientist & Technical Director, Composites & Advanced Materials Laboratory, 1845 Fairmount Street; waruna@niar.wichita.edu.

§Professor, Mechanical and Aerospace Engineering, MAE-A; nkim@ufl.edu. Associate Fellow AIAA (Corresponding Author).

¶Distinguished Professor, Mechanical and Aerospace Engineering, MAE-A; haftka@ufl.edu. Fellow AIAA.

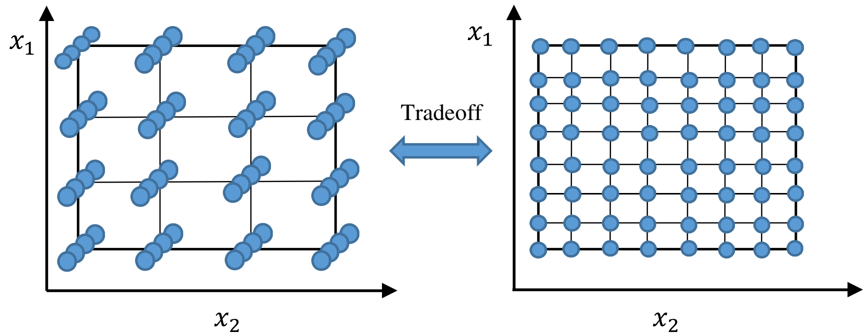


Fig. 1 Replication strategy vs exploration strategy using two design variables and 64 samples.

representation of a failure criterion, which can easily be incorporated in the design process.

Surrogate prediction consists of two parts: design of experiments and model development. Design of experiments, which is also termed sampling strategy, refers to the approach of selecting test configurations. Prediction accuracy of a surrogate can be significantly affected by sampling strategy. Wang and Shan [16] and Mackman et al. [17] reviewed popular approaches to generate design of experiments, which had significant effect on optimization results. For design of experiments, we may explore different test configurations in the design space or replicate the same test configuration multiple times. We would like to examine the tradeoff between exploration and replication when limited testing resources are available, as shown in Fig. 1. The replication strategy is represented by 4×4 test matrix with four replicates at each point, whereas the exploration strategy is represented by 8×8 test matrix with no replicates.

Exploration interrogates the design space to find potential unforeseen failure modes and unexpected responses, whereas replication at the same configuration quantifies experimental scatter and improves accuracy at test points. In the replication strategy, test variability is considered by averaging replicated tests, which may allow the use of an interpolating surrogate that passes through the averaged data points. Without replication, filtering out experimental noise is accomplished by a regression (i.e., least-squares fit) surrogate model in the exploration strategy. In the literature [18,19], it was found that exploration may be more efficient than replication in testing structural elements for a given budget. Exploration is more likely to reduce surprises from unrecognized failure modes, and it yields surrogate models that are more accurate when applied to approximation at untested designs. However, the results obtained by [18,19] were with manufactured data to the assumption that experimental strength was a random variable with normal distribution, independent and identically distributed.

This paper first presents a study of how this conclusion stands using real experimental results produced by the National Institute for Aviation Research on open-hole tension (OHT) tests [20,21] according to American Society for Testing and Materials (ASTM) standard [22]. These OHT experiments were intended to investigate the impact of the coupon width-to-hole diameter ratio on the failure of a composite panel. The failure mode and strength were evaluated at four structural configurations with replicates. We examined the effect of exploration versus replication on surrogate prediction based on subsets of experimental results. Different designs of experiments were generated with the same total number of tests.

The experimental data exhibited deviations from a normal distribution. Thus, we were prompted to consider other distributions that could be alternatives for characterization of composite strength, such as the Weibull distribution [23] and irregular distribution without analytical expression caused by batch-to-batch variability [24]. Influence of strength distributions on the sampling methods was investigated using one analytical function, composite laminate with highly nonlinear failure response. Polynomial response surface (PRS) was selected for surrogate modeling due to its robustness and excellence with approximating noisy data. The surrogate toolbox provided by [25] was adopted to develop PRS prediction.

The paper is arranged as follows. Section II introduces the OHT test configuration and surrogate prediction of OHT tests. Surrogate models are developed and compared based on exploration strategy and replication strategy. To compensate for experimental variability, we resampled subsets of experimental results, emphasizing either exploration or replication and repeated the modeling procedure. In Sec. III, we investigate the effect of distribution on sampling plans using one analytical function and three synthetic distributions. Exploration and replication strategies are compared. Section IV summarizes the results and details future work.

II. Exploration Versus Replication Sampling Schemes for Open-Hole-Tension Tests

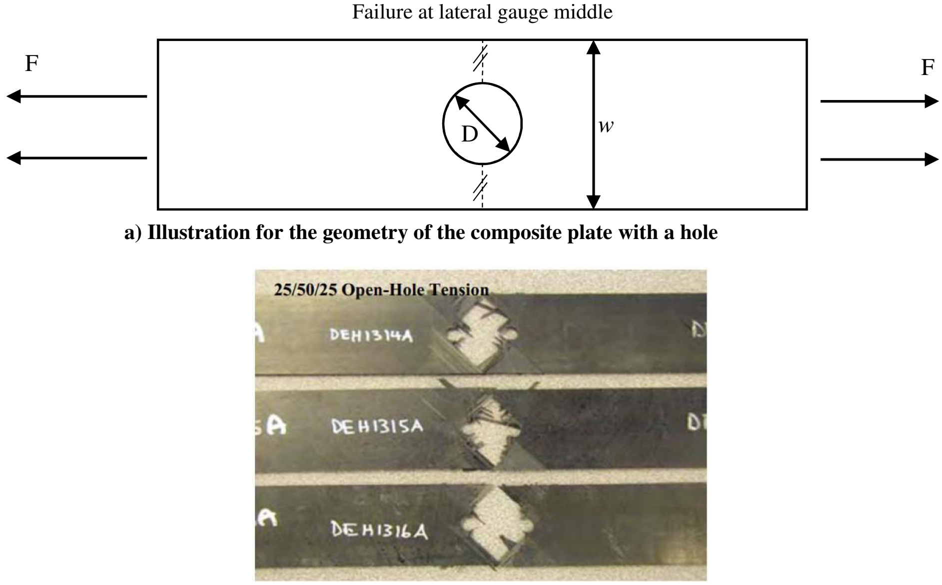
A. Open-Hole Tension Tests

The open-hole tension (OHT) test was designed to investigate the effect of hole size on the tensile strength of composite laminates. OHT tests were conducted according to ASTM D5766 [22]. The standard test specimen geometry ($w/D = 6$) is shown in Fig. 2. The width of the plate and diameter of the hole are denoted by w and D, respectively. The plates with a hole were made of three distinct material batches of Toray T700SC-12K-50C/#2510 plain weave fabric. The specimen laminate ply orientations were $[0/90/0/90/45/-45/90/0/90/0]$ with a nominal thickness of 0.172 in. Tests were conducted in ambient laboratory conditions with an as-fabricated moisture content. The tensile strength of the plate with a hole was calculated based on tensile load and the nominal cross-sectional area (disregarding the hole). Table 1 details the test matrix at four configurations, with $w/D = 3$, 4, 6, 8. Eighteen replicates were manufactured at each configuration to quantify strength variability. Three preimpregnated (“prepreg”) batches were adopted to quantify batch-to-batch variability. All prepreg batches were manufactured with the same processing specification but at different dates. The loading rate of the tension tests was 0.05 in./min. In all the tests, the dimension of D is the same, but the specimen width w is different. The smaller lateral gauge may result in higher possibility of manufacturing defects. Therefore, panels were examined using through-transmission ultrasonic C-scan before machining specimens to guarantee the manufacturing quality.

For the OHT tests, only a single failure mode was observed, lateral gauge middle, and the strength response is likely to be smooth without jump by visual check, as seen in Fig. 3. The strength of OHT tests was assumed to be a smooth curve perturbed by random noise. We focused on the prediction accuracy of mean strength with varying w/D ratio in this paper.

B. Experimental Strength Results and Polynomial Fit

Physical models [12,26] have been proposed to predict the strength of composite plate with a hole. We limited ourselves to data-driven approaches only as a valuable complement to design optimization when the physical models/simulations are inadequate for desirable accuracy or require excessive computational resources. A quadratic polynomial response surface (PRS) was selected to approximate the mean value of failure strength after trying other fits. The PRS was developed using all 72 specimen strength values at four w/D ratios. The OHT strength of each laminate is shown in Fig. 3, along with a



a) Illustration for the geometry of the composite plate with a hole

b) Typical tension failure mode of composite laminate with a hole

Fig. 2 OHT test specimen configuration and observed failure mode.

quadratic PRS fitted to the mean value of specimen strengths and 95% confidence interval of the prediction variance. The OHT strength increased gradually with w/D. Statistical information on the replicate OHT strengths were summarized in Table 2, where $\boldsymbol{S}_i$ denote tested strength of 18 replicates at $w/D = 3$, 4, 6, and 8. Detailed experimental results are provided in Supplement A.

In Table 3, we can see that the maximum discrepancy between surrogate prediction and mean strength was 0.8% or less. Different from the interpolation surrogate models (i.e., kriging) that pass all the samples, the PRS filters noise and preserves monotonicity in the design space. Therefore, a quadratic PRS seemed reasonably accurate to approximate OHT tests and was considered as the true function value in the following analysis of prediction accuracy. The differences between this PRS model and any other surrogate models based on subsets of OHT test results were assumed to be from sampling strategies.

C. Resampling Experimental Strength for Comparing Exploration and Replication

To compare sampling strategies for exploration and replication, we simulated situations of paucity of data, by resampling partial experimental strength out of the 72 available experimental results without replacement. For the exploration strategy, we sampled from all values of w/D, whereas for the replication strategy, we left out one value of w/D but had more replicates of the other three. This resampling approach was motivated by bootstrapping [27], which is a nonparametric approach for statistical inference. The strength of a composite material is often associated with large variability (as seen in Supplement A); the evaluation of sampling strategy should be interpreted in the context of stochastic effect. A resampling procedure makes the most use of limited experimental strength and is likely to enable insightful understanding of sampling strategy excluding stochastic effect. All the resampling plans were repeated 1000 times using Monte Carlo simulations to account for randomness. We compared surrogate prediction for the cases of a total of 12 samples and of 24 samples.

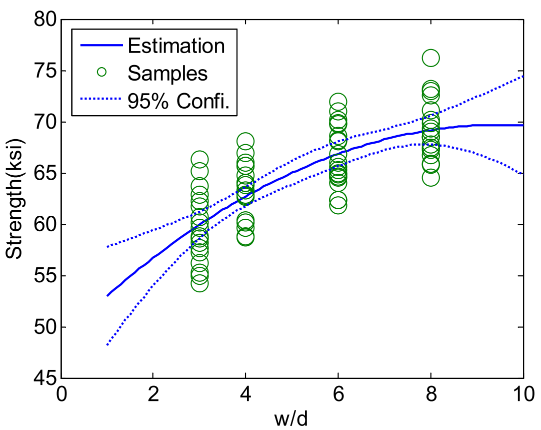


Fig. 3 Experimental data and quadratic polynomial response surface fitted to the data.

Table 1 w/D ratio test matrix (18 replicates are selected from three prepreg batches)

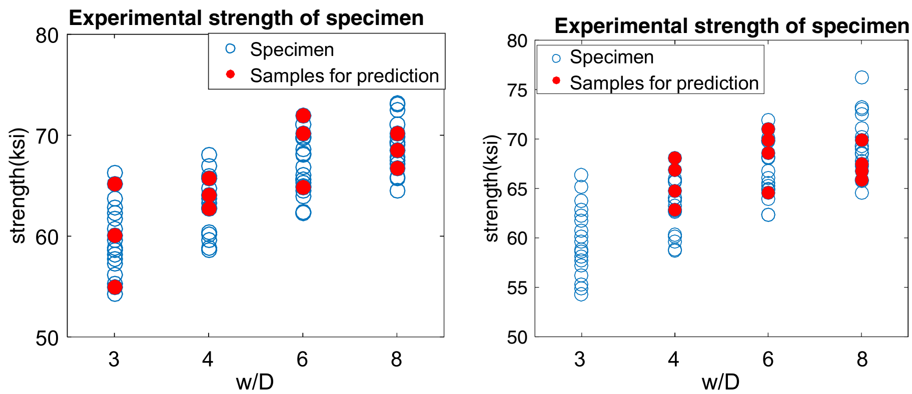
w, in.	D, in.	w/D	Replicates	Number of specimens
0.75	0.250	3	3×6	18
1.00	0.250	4	3×6	18
1.50	0.250	6	3×6	18
2.00	0.250	8	3×6	18
		Total		72

Table 2 Statistical properties of OHT strengths at given w/D with 18 replicates

Data set	Mean, ksi	Standard deviation, ksi	Coefficient of variation, %	Range of strength, ksi
$\boldsymbol{S}_3$	59.64	3.52	5.90	[54.27, 66.33]
$\boldsymbol{S}_4$	63.22	2.79	4.41	[58.66, 68.07]
$\boldsymbol{S}_6$	66.53	3.01	4.52	[61.83, 71.89]
$\boldsymbol{S}_8$	69.27	3.04	4.39	[64.51, 76.24]

Table 3 Relative difference between surrogate prediction and mean strength using quadratic PRS

w/d	Relative difference, %
3	0.4
4	0.8
6	0.5
8	0.2



a) Illustration of exploration strategy

b) Illustration of replication strategy while omitting w/D=3

Fig. 4 Illustration for one set of resampled experimental strength representing exploration or replication strategy using 12 points.

For the exploration strategy, with 12 samples, three samples were randomly selected from each of the $\boldsymbol{S}_3, \boldsymbol{S}_4, \boldsymbol{S}_6, \boldsymbol{S}_8$ sets. For the replication strategy, four samples were selected from each of the $\boldsymbol{S}_i, \boldsymbol{S}_j, \boldsymbol{S}_k$ sets, where $i \neq j \neq k$, as shown in Fig. 4. We skipped one of w/D and used four replicates for each of the remaining three. The replication strategy was repeated four times by omitting one w/D in turn. With 24 samples, the same scheme was used but with twice as many samples for each w/D.

A quadratic PRS was then constructed with the selected samples. The relative error $\mathrm{err}(x)$ in Eq. (1) was computed at the four test configurations to quantify the prediction accuracy:

$$\mathrm{err}(w/D_i) = \frac{\bar{f}(w/D_i) - f(w/D_i)}{f(w/D_i)} \tag{1}$$

where $\bar{f}(w/D)$ denoted predicted failure strength, and $f(w/D)$ denoted the true strength obtained from the previous PRS with all 72 samples. The root-mean-square error (RMSE) given in Eq. (2) was used to quantify the overall performance of surrogate prediction:

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{4} \mathrm{err}(w/D_i)^2}{4}} \tag{2}$$

Table 4 shows the accuracy of surrogate models using exploration and replication strategies with 12 and 24 total samples. The RMSE of relative error for 1000 sets of samples was computed at four configurations, $w/D = 3, 4, 6, 8$. It is obvious that, for this problem, the exploration strategy was more accurate than the replication strategy. When modeling surrogates using samples from ($w/D = 4$, 6, 8) and ($w/D = 3$, 4, 6), the estimation errors at $w/D = 3$ and $w/D = 8$ were significantly large due to the necessity for extrapolation. Also, the largest RMSE among the different sample selections occurred when extrapolation was required.

The influence of number of samples on prediction accuracy was also observed. When twice as many samples are used to average the noise, the RMSE is expected to be reduced by the square root of 2. Indeed, the errors were reduced by a factor ranging from 1.38 to 1.45. The error in surrogate prediction strongly depends on variability of experimental strength, and the results in Table 4 show that exploration filters out the noise in individual measurement as effectively as or more effectively than replication. We further discuss the effect of strength variability on surrogate models in the following section to make the study more general.

Table 4 Mean values[a] of relative RMSE for surrogate from exploration and replication

Sampling strategy	Exploration	Replication omitting specimen from specific w/D: $w/D = 3$	$w/D = 4$	$w/D = 6$	$w/D = 8$
12 samples	2.27	3.07	2.33	2.76	6.75
24 samples	1.57	2.22	1.63	1.89	4.70

[a]Mean values reported are based on 1000 sets of samples.

D. Identifying the Distribution Type of Open-Hole Tension Strength

In their previous work, Matsumura et al. [18] assumed that the distribution of strengths at the same configuration was normal, independent, and identically distributed (IID). Independent and identical normal distributions are frequently used for theoretical studies of stochastic effects. Here, we examined to what extent this assumption holds for OHT test data. We assume that all $\boldsymbol{S}_i$ follow the same type of distribution but possibly with different distribution parameters for each w/D. Three candidate distributions were selected to characterize variability of strength: a normal distribution, two-parameter Weibull distribution (which is commonly used for strength characterization), and a uniform distribution for comparison purposes.

The $\boldsymbol{S}_i$ distribution was analyzed using the Kolmogorov–Smirnov (K-S) test, which is a nonparametric test to quantify the goodness-of-fit between a given probability distribution and empirical distribution of samples [28]. The p-value of the K-S test indicates the probability that samples do not reject the hypothetical distribution. A high p-value denotes high probability that samples are from the hypothetical distribution. Effectiveness of K-S test strongly depend on the number of samples. Therefore, we pooled $\boldsymbol{S}_i$ together (i.e., from different w/Ds) for better estimation before applying the K-S test. $\boldsymbol{S}_i$ distributions had similar coefficients of variation (CVs) as seen in Table 2, which implied that it was possible to combine the four sets of $\boldsymbol{S}_i$ together to query the underlying distribution. Samples were transformed into the standard form, namely with location parameter equal to zero and scale parameter equal to one. Equation (3) was used to standardize $\boldsymbol{S}_i$ while assuming normal distribution or uniform distributions. Equation (4) was used to standardize $\boldsymbol{S}_i$ while assuming Weibull distribution:

$$\boldsymbol{N}_i = \frac{\boldsymbol{S}_i - \mathrm{mean}(\boldsymbol{S}_i)}{\mathrm{std}(\boldsymbol{S}_i)}, \qquad \text{where } i = 3, 4, 6, 8 \tag{3}$$

$$\boldsymbol{W}_i = \frac{\boldsymbol{S}_i}{\mathrm{std}(\boldsymbol{S}_i)}, \qquad \text{where } i = 3, 4, 6, 8 \tag{4}$$

The p-values for pooled strengths were computed for the three candidate distributions as shown in Fig. 5. A 5% significance level was adopted to exclude distributions that were less likely to occur (the red horizontal line in Fig. 5). Candidate distributions with a p-value lower than the significance level were rejected. The figure showed that there was a high chance that OHT strength follows a normal distribution, a lower chance that it came from the Weibull distribution, and the uniform distribution was rejected. Probability plots for OHT strengths while assuming normal distribution and

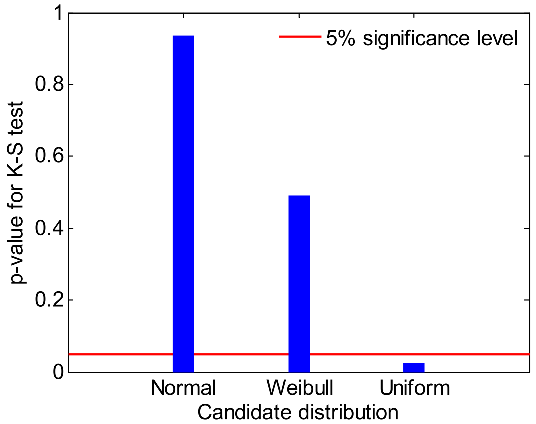


Fig. 5 K-S test statistic for OHT strength.

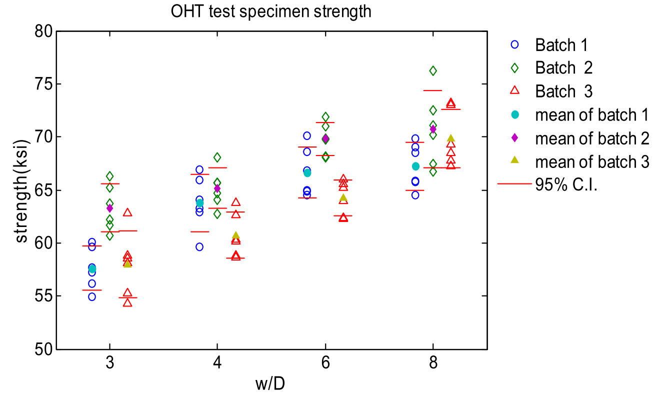


Fig. 7 Statistical properties of specimen strength.

Weibull distribution are provided in Fig. 6 for further comparison. The OHT strengths matched both distributions in general, except at the tails. The inconsistency at tails may be due to outliers in the experiments and the inability of statistical distributions to approximate physical phenomenon as well as the possibility that the real distribution may be different. We also repeated the analysis of fitting of the distributions considering samples from each batch independently by removing the bias associated with batches. The P values using normal distribution were around 0.9 for all the three batches, which were larger than or comparable with the P values using Weibull distribution. Therefore, normal distribution seemed desirable to approximate the variation of the specimen strength.

E. Effect of Between-Batch Variability

We examined the distribution of OHT test results and concluded that the assumption that structural strength followed IID normal distribution was not perfectly satisfied as assumed in Matsumura et al. [18,19]. The strength of composite laminates could be associated with significant variability, as seen in Sec. II.C. In practical designs, specimens may be made from different prepreg batches. Batch-to-batch variability could be a major source of strength variation as reported by ASTM standard. Notable batch-to-batch variability would invalidate the assumption of using IID normal random variability to approximate structural strength. Figure 7 compares $\boldsymbol{S}_i$ obtained from different batches. Let $\boldsymbol{b}_i^{w/D}$ be the strength of specimen tested from batch i at w/D, where $i = 1, 2, 3$, and $w/D = 3, 4, 6, 8$. At given w/D, maximum values of min $(\boldsymbol{b}_i^{w/D})$, max $(\boldsymbol{b}_i^{w/D})$, and mean $(\boldsymbol{b}_i^{w/D})$ were mostly from batch 2, indicating that specimens made from batch 2 tended to be stronger.

We assumed that $\boldsymbol{b}_i^{w/D}$ followed a normal distribution based on the results of the previous nonparametric analysis. $\boldsymbol{b}_i^{w/D}$ was likely to have different mean value with same w/D ratio. Welch's t-test was adopted to compare central values of two Gaussian populations [29]. This test assumes that the two populations have normal distributions and unequal variances. The p-value was used to indicate probability of the null hypothesis that the two population means are equal (using a two-tailed test). A high p-value represents high possibility that the two distributions under consideration are equal. Table 5 documents the p-value for Welch's t-test between specimens made from batch 1 and batch 2. The p-values at $w/D = 3, 6, 8$ are smaller than the 5% significance level (i.e., the assumption that the two samples were from the same distribution was rejected). The p-value at $w/D = 4$ was 0.15, which is slightly larger than significance level 0.05. Based on Fig. 7 and Table 5, strengths of specimens from same batch obviously shared systematic bias. The between-batch variability resulted in a violation of normality assumption and IID condition even though $\boldsymbol{S}_i$ was close to normal distribution in overall.

It would be interesting to check the effect of between-batch variability on prediction accuracy. This is an extended study on tradeoff between exploration of more batches and replication ignoring the between-batch variability. We assumed that each prepreg batch was equally important and selected at least one sample from each batch to conserve batch-to-batch variability. The sampling plan conserving batch-to-batch variability is termed three-batch sampling and can be viewed as emphasizing exploration, whereas random sampling may omit a batch and have more replicates in the other batches. We compared surrogate models based on random sampling and three-batch sampling. Table 6 shows the accuracy of surrogate models using exploration and replication strategies with 12 specimens. The relative RMSE for 1000 sets of samples were computed at four configurations, $w/D = 3, 4, 6, 8$. It is seen that we benefit from both types of exploration, using full set of w/Ds, and all three batches.

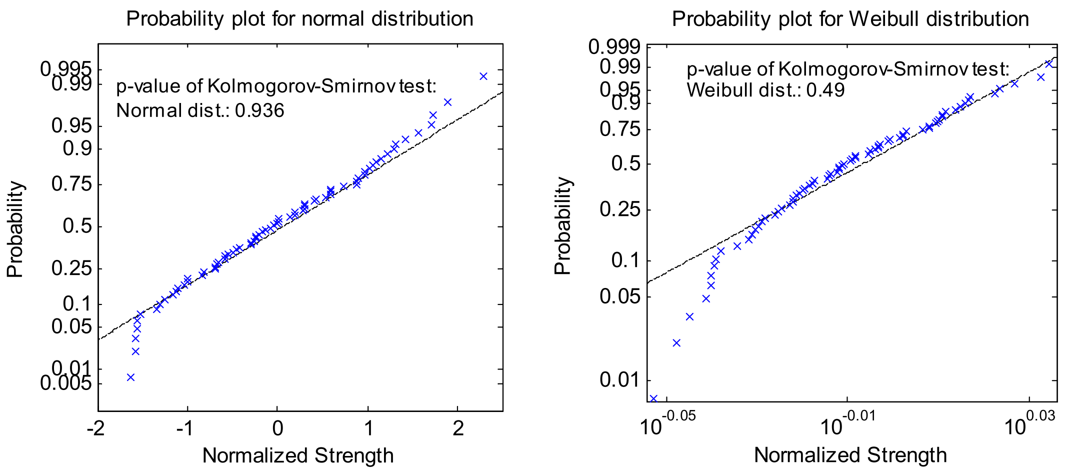


Fig. 6 Probability plot for a normal distribution and a Weibull distribution.

Table 5 Statistics of Welch's t-test between specimens from batch 1 and batch 2

Statistics	$w/D = 3$ $\boldsymbol{b}_1^{w/D}$	$w/D = 3$ $\boldsymbol{b}_2^{w/D}$	$w/D = 4$ $\boldsymbol{b}_1^{w/D}$	$w/D = 4$ $\boldsymbol{b}_2^{w/D}$	$w/D = 6$ $\boldsymbol{b}_1^{w/D}$	$w/D = 6$ $\boldsymbol{b}_2^{w/D}$	$w/D = 8$ $\boldsymbol{b}_1^{w/D}$	$w/D = 8$ $\boldsymbol{b}_2^{w/D}$
Mean value	57.6 ksi	63.3 ksi	63.8 ksi	65.2 ksi	66.7 ksi	69.8 ksi	67.3 ksi	70.7 ksi
p-value	0.0004		0.15		0.01		0.04	

Table 6 Mean values[a] of relative RMSE for surrogate with/without between-batch variability

Sampling strategy	Exploration	Replication omitting specimen from $w/D = i$: $w/D = 3$	$w/D = 4$	$w/D = 6$	$w/D = 8$
Random sampling	2.27	3.07	2.33	2.76	6.75
Three-batch sampling	1.62	2.45	1.69	2.14	4.9

[a]Mean values reported are based on 1000 sets of samples. Three replicates from four configurations are for exploration, four replicates from three configurations are for replication.

Table 7 Parameter settings of composite laminate

Parameter	Value
E_1	150 GPa
E_2	9 GPa
ν_{12}	0.34
G_{12}	4.6 GPa
Ply thickness	125 μm
$\varepsilon_{1\text{allow}}$	± 0.01
$\varepsilon_{2\text{allow}}$	± 0.01
$\gamma_{12\text{allow}}$	± 0.015

III. Distribution Type Influence on Sampling Strategies: Analytical Model of a Composite Plate

The experimental results indicated that the strength of composite laminates was not a perfectly independent and identically distributed Gaussian random variable. Therefore, we examined the influence of distribution type on sampling strategies using an analytical test function. Although the results aligned with the previous numerical studies showing that exploration was better than replication, the OHT case was relatively benign compared to the numerical examples of Matsumura et al. [18]. First, there is only one design parameter in the OHT test results; second, the failure stress varies smoothly, allowing accurate approximation with low-order polynomials; and third, the variation of the distribution is relatively small. To obtain further insight, we investigate sampling strategies and distribution types using one analytical function of the failure load of an unnotched composite laminate with highly nonlinear failure response. The geometry of the plate and the dependence of the failure load on the two design parameters are shown in Fig. 8.

The composite laminate structure had three ply angles $[0/-\theta/+\theta$ deg], which was intended to have a more complex failure response as shown in Fig. 8. The laminate is subject to forces N_x and N_y along the x and y axes defined by parameter α, where $N_x = (1-\alpha)F$ and $Ny = \alpha F$. There are two design parameters for this structure: $\theta \in [0, 90$ deg] and $\alpha \in [0, 0.5]$. Table 7 details the material properties and strain allowables. The strains are predicted by the classical lamination theory. Critical failure mode was due to ply axial strain. The failure loads were calculated analytically as in Matsumara et al. [18].

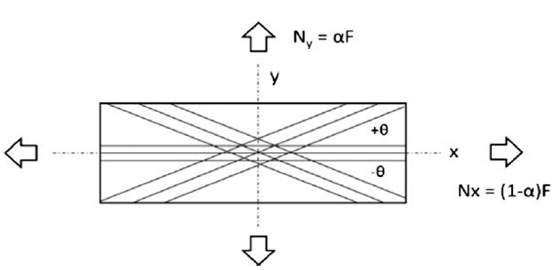
The analytical failure loads are perturbed with synthetic noise to imitate experimental variability. Three types of distribution were considered in this work: normal distribution, Weibull distribution with heavy tail, and multisource distribution to imitate batch-to-batch variability. The numerical study focused on average performance of surrogates based on the synthetic samples from Monte Carlo simulation. Variability of the performance of surrogates is attached in Supplement B for reference.

A. Synthetic Noise Using Three Types of Distribution

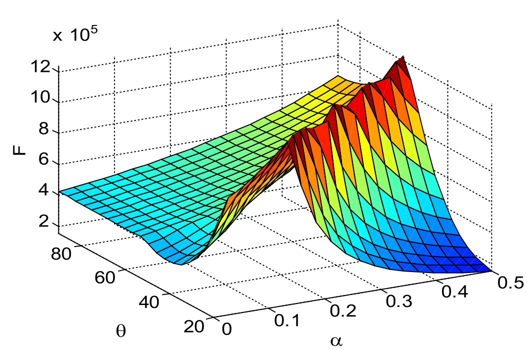
We considered three types of distributions to simulate strength variation. Normal and Weibull distributions were considered as two popular distributions to approximate the structural strength. The normal distribution was set to be $N(\mu_N(\theta,\alpha), \sigma_N(\theta,\alpha))$, where $\mu_N(\theta,\alpha)$ was the mean value of structural strength from analytical function at configuration (θ,α), and $\sigma_N(\theta,\alpha)$ was obtained assuming constant CV to be 0.05.

The Weibull distribution was first shaped to follow $W(a,b)$ with a heavy tail to amplify the difference with the normal distribution, where the scale parameter $a = 2$, the shape parameter $b = 1.5$. Then, $W(a,b)$ was scaled through a linear transformation to imitate the experimental variability of the analytical test function. Details of the Weibull distribution are given in the Appendix.

The multisource distribution was proposed to imitate between-batch variability based on findings from the OHT test results. We first

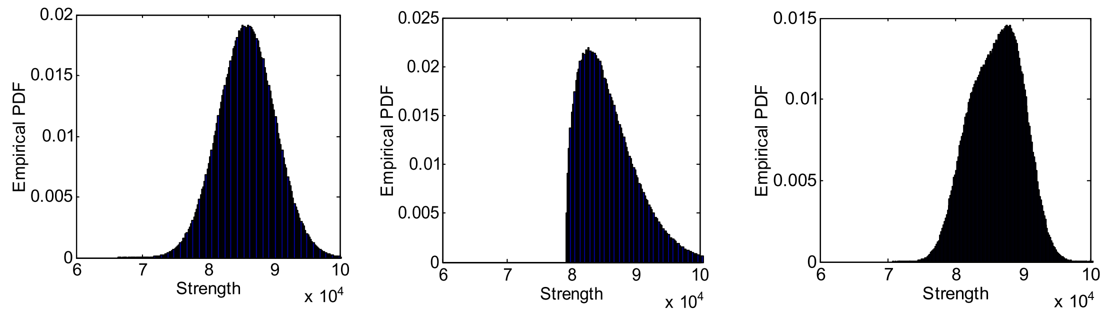


a) Illustration of composite laminate



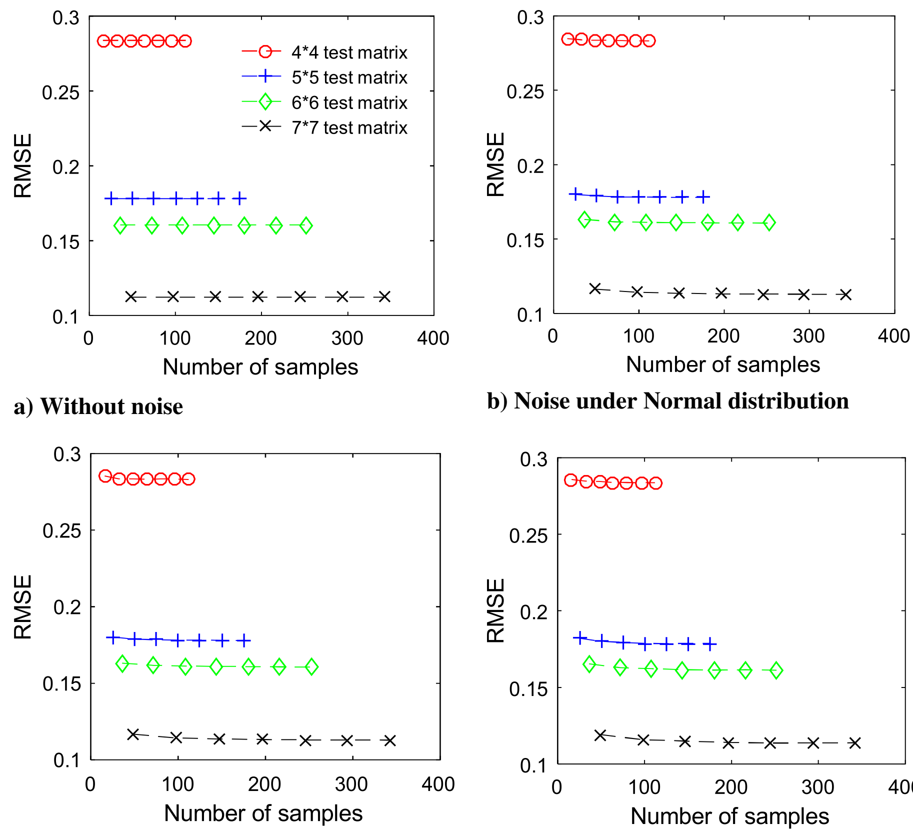
b) Response for critical failure load (N) in the design space

Fig. 8 Configuration and failure response of composite laminate. α is a constant to balance between horizontal and vertical loadings. θ is the orientation angle of lamina.



a) Normal distribution b) Weibull distribution c) Multisource distribution

Fig. 9 Normal distribution, Weibull distribution, and multisource distributions with mean value to be 8.5 ksi. The mean values of the five synthetic batches are 8.1, 8.5, 8.8, 8.4, and 9.0 ksi.



a) Without noise b) Noise under Normal distribution

c) Noise under Weibull distribution d) Noise under multisource distribution

Fig. 10 Mean value of relative RMSE based on 1000 sets of samples (number of samples indicates the total number of tests, including 1–7 replicates for each test point) with different distribution types.

generated five random values M_i as the mean strengths of five batches from $N(0, 0.035^2)$ and fixed to imitate the between-batch bias for the following analysis. The mean values were then scaled up through linear transformation at different configurations (θ, α) as shown in Eq. (5):

$$m_i(\theta, \alpha) = f(\theta, \alpha) \times M_i + f(\theta, \alpha) \tag{5}$$

Synthetic samples were then generated from $N(m_i(\theta, \alpha), (m_i(\theta, \alpha) \times 0.035)^2)$ assuming constant CV to be 0.035. Typical probability density functions of three distributions are illustrated in Fig. 9 for 8.5 ksi mean value of strength. The synthetic specimen strength in the design space is assumed to come from one of the five batches. The multisource distribution should be skewed due to between-batch for five specific distributions as seen in Fig. 9c.

B. Fitting Strategy

A polynomial response surface (PRS) surrogate was selected because of its robustness and good performance for predictions based on noisy data. Leave-one-out cross validation was adopted to select an appropriate polynomial order for each individual set of samples. PRS assumed constant variance instead of constant coefficient of variation. Therefore, function values of sampling points were preprocessed using logarithmic transformation with base 10 to obtain close-to-constant variance.

As in Matsumura et al. [18], our test matrices ranged from 4×4 to 7×7 with evenly spaced test points to investigate the effect of the density of matrix on the accuracy of approximation. For each test matrix, we replicated the same test configuration up to seven times. Each test matrix was generated and fitted 1000 times to compensate for randomness of samples and obtain the average accuracy. Accuracy of surrogate prediction was computed using a 20×20 matrix of test points (400 points in total). Relative RMSE from Eq. (4) was adopted to measure overall accuracy of the surrogate model.

C. Influence of Distribution Type on Sampling Strategies

We first developed surrogate models to approximate the analytical function without noise for reference. Figure 10a summarizes prediction accuracy of surrogate prediction at all test matrix without noise.

The typical orders for polynomial response surface selected using cross validation were quadratic, cubic, quartic, and quantic for the four sets of test matrix. The composite laminate has a highly nonlinear response (see Fig. 8) that greatly benefited from the higher-order polynomials made possible by using more test points. This example corresponds to a case that is more challenging to fit than the OHT case. The RMSE of relative error reduced from 0.2 to 0.11 with increasing points for exploration. While modeling the composite laminate strength, a typical PRS model based on 4×4 matrix was quadratic, and a typical PRS model based on 7×7 matrix was quintic. Note that, because there is no noise, RMSE remains constant through replication.

Noise was then added to the data, and the influence of noise distribution on sampling strategies was compared. Figure 10 documents the accuracy of surrogate models for the composite laminate. From Figs. 10b–10d, we can observe a clear advantage of exploration over replication with different distributions. The prediction accuracy did not change much with increasing replicates for the same configuration set. This was due to the highly nonlinear response that implies that surrogate model error is the major concern. Increasing replicates could mitigate mostly the effect of randomness (noise). Allocating samples for exploration enables high-order PRS for closer approximation to the response. That is, model error was more significant than sampling error corresponding to experimental variability for this example. We also examined the influence of magnitude of noise on prediction accuracy in Appendix A. Noise with a coefficient of variation varying from 0.05 to 0.2 only had a limited effect on predicative capability of surrogates comparing with model form.

IV. Conclusions

To predict the strength of many composite structural elements, it is necessary to perform structural strength tests for a matrix of configurations as function of loading, geometry, and material properties. Then, a surrogate fit allows interpolation or extrapolation to be performed to obtain strengths of untested configurations. For a given number of tests, a balance between exploration, meaning testing many structural configurations, and replication of each tested configuration must be struck to overcome the effects of noise. Previous work has shown that exploration is more effective than replication, especially when the functional form of response is complicated, but it was based on simulated test data assuming independent and identically distributed normal noise.

In this paper, the conclusion favoring exploration using actual test data of strength of open hole composite plates is first tested as function of the ratio of plate width to hole diameter, w/D. A resampling procedure is proposed based on subsets of experimental strength to compare exploration and replication. The experimental data consist of 72 samples, indicating failure strengths for three batches of laminates with six replicates per batch and four values of w/D. The two strategies were compared by using 12, 24, or all 72 samples. For each case, testing all four w/D configurations with fewer replicates was compared with testing only three of the four with more replicates. Using 12 or 24 samples randomly drawn from 72 samples, the accuracy of fitted surrogates was checked 1000 times in a Monte Carlo sampling strategy. The results showed that testing all four configurations with fewer replicates led to more accurate predictions than testing only three configurations with more replicates, thus confirming the previous simulated results.

The major source of uncertainty for prediction was then investigated, which is the variability of strength. It was found that the distribution of OHT strength is not perfectly independent and identically distributed normal distribution as commonly assumed. This was likely due to systematic bias in strength for different prepreg batches. To generalize the present observation on sampling strategies, the influence of distributions based on synthetic data was investigated using nonnormal distributions and correlations. Synthetic samples were generated based on the strength response for an unnotched composite laminate using normal distribution, Weibull distribution, and multisource distribution. Prediction of polynomial response surface (PRS) with fixed order/terms had similar accuracy while increasing replicates for different distributions. Introducing more replicates did not benefit prediction accuracy noticeably. An exploration strategy enables a higher-order PRS to be used, which is key to increase prediction accuracy for complex structural behavior.

These results are limited to the accuracy of the mean value of the failure strength. Future work will focus on estimating the tolerance limits, such as the B-basis using surrogate models for designing structures with a limited number of available experimental results. Additional composite test types to investigate the application of surrogate are also being sought.

Appendix: Weibull Distribution with a Heavy Tail

The Weibull distribution was first shaped to follow $W(a,b)$ with a heavy tail to amplify the difference with the normal distribution, where the scale parameter $a = 2$, and the shape parameter $b = 1.5$ as seen in Eq. (A1). Then, $W(a,b)$ was scaled through a linear transformation to imitate the experimental variability of the analytical test function according to Eqs. (A2–A4), where the strength of the analytical test function, $F_W^{-1}(p|a,b)$, is the inverse cumulative distribution at percentile p, and $\mu_W(a,b)$ is the mean of $W(a,b)$. $al(\theta,\alpha)$ from Eq. (A2) was adopted to scale up the variability of the synthetic strength, and $bl(\theta,\alpha)$ in Eq. (A3) was used to guarantee the mean value of the synthetic strength is $f(\theta,\alpha)$:

$$k \sim W(a,b) = \frac{b}{a}\left(\frac{k}{a}\right)^{b-1} e^{-(x/a)^b} \tag{A1}$$

$$al(\theta,\alpha) = \frac{0.2 f(\theta,\alpha)}{F_W^{-1}(97.5|a,b) - F_W^{-1}(2.5|a,b)} \tag{A2}$$

$$bl(\theta,\alpha) = f(\theta,\alpha) - al(\theta,\alpha) \times \mu_W(a,b) \tag{A3}$$

$$s(\theta,\alpha) = k \times al(\theta,\alpha) + bl(\theta,\alpha) \tag{A4}$$

Acknowledgments

This research was supported by DARPA contract W911QX-13-C-0137. Open-hole tension test data were furnished by the National Institute for Aviation Research at Wichita State University.

References

[1] Whitney, J. M., and Nuismer, R., “Stress Fracture Criteria for Laminated Composites Containing Stress Concentrations,” *Journal of Composite Materials*, Vol. 8, No. 3, 1974, pp. 253–265. doi:10.1177/002199837400800303

[2] Tsai, S. W., and Wu, E. M., “A General Theory of Strength for Anisotropic Materials,” *Journal of Composite Materials*, Vol. 5, No. 1, 1971, pp. 58–80. doi:10.1177/002199837100500106

[3] Hinton, M. J., Kaddour, A. S., and Soden, P. D., *Failure Criteria in Fibre Reinforced Polymer Composites: The World-Wide Failure Exercise*, Elsevier, Amersterdam, London, 2004, pp. 702–1045.

[4] Hinton, M., and Kaddour, A., “Triaxial Test Results for Fibre-Reinforced Composites: The Second World-Wide Failure Exercise Benchmark Data,” *Journal of Composite Materials*, Vol. 47, Nos. 6–7, 2013, pp. 653–678. doi:10.1177/0021998312459782

[5] Hallett, S. R., Green, B. G., Jiang, W.-G., Cheung, K. H., and Wisnom, M. R., “The Open Hole Tensile Test: A Challenge for Virtual Testing of Composites,” *International Journal of Fracture*, Vol. 158, No. 2, 2009, pp. 169–181. doi:10.1007/s10704-009-9333-8

[6] LLorca, J., González, C., Molina-Aldareguía, J. M., Segurado, J., Seltzer, R., Sket, F., Rodríguez, M., Sádaba, S., Muñoz, R., and Canal, L. P., “Multiscale Modeling of Composite Materials: A Roadmap Towards Virtual Testing,” *Advanced Materials*, Vol. 23, No. 44, 2011, pp. 5130–5147. doi:10.1002/adma.v23.44

[7] Hallett, S. R., Jiang, W.-G., and Wisnom, M. R., "Effect of Stacking Sequence on Open-Hole Tensile Strength of Composite Laminates," *AIAA Journal*, Vol. 47, No. 7, 2009, pp. 1692–1699. doi:10.2514/1.41244

[8] Rice, R. C., Goode, R., Bakuckas, J., and Thompson, S., "Development of MMPDS Handbook Aircraft Design Allowables," *Proceedings of the 7th Joint DOD/FAA/NASA Conference on Aging Aircraft*, New Orleans, LA, Sept. 2003.

[9] Tomblin, J., Sherraden, J., Seneviratne, W., and Raju, K., "A-Basis and B-Basis Design Allowables for Epoxy-Based Prepreg Toray T700SC-12K-50C/# 2510 Plain Weave Fabric," Advanced General Aviation Transport Experiments, National Institute for Aviation Research, Wichita State Univ., Wichita, KS, 2002.

[10] Mehrez, L., Moens, D., and Vandepitte, D., "Stochastic Identification of Composite Material Properties from Limited Experimental Databases, Part 1: Experimental Database Construction," *Mechanical Systems and Signal Processing*, Vol. 27, Feb. 2012, pp. 471–483. doi:10.1016/j.ymssp.2011.09.004

[11] Handler, J. J., Dustin, J. S., Liguore, S., Parrish, A., and Wanthal, S., "Development of a High Fidelity Test Database for Failure Prediction Methods Validation," *Proceedings of the American Society for Composites 2014—29th Technical Conference on Composite Materials*, Univ. of California, San Diego, La Jolla CA, 2014.

[12] Carlsson, L. A., Adams, D. F., and Pipes, R. B., *Experimental Characterization of Advanced Composite Materials*, CRC Press, Boca Raton, FL, 2014, pp. 231–248.

[13] Forrester, A. I. J., Keane, A. J., and Bressloff, N. W., "Design and Analysis of 'Noisy' Computer Experiments," *AIAA Journal*, Vol. 44, No. 10, 2006, pp. 2331–2339. doi:10.2514/1.20068

[14] Glaz, B., Goel, T., Liu, L., Friedmann, P. P., and Haftka, R. T., "Multiple-Surrogate Approach to Helicopter Rotor Blade Vibration Reduction," *AIAA Journal*, Vol. 47, No. 1, 2009, pp. 271–282. doi:10.2514/1.40291

[15] Chaudhuri, A., and Haftka, R. T., "Efficient Global Optimization with Adaptive Target Setting," *AIAA Journal*, Vol. 52, No. 7, 2014, pp. 1573–1578. doi:10.2514/1.J052930

[16] Wang, G. G., and Shan, S., "Review of Metamodeling Techniques in Support of Engineering Design Optimization," *Journal of Mechanical Design*, Vol. 129, No. 4, 2007, pp. 370–380. doi:10.1115/1.2429697

[17] Mackman, T., Allen, C., Ghoreyshi, M., and Badcock, K., "Comparison of Adaptive Sampling Methods for Generation of Surrogate Aerodynamic Models," *AIAA Journal*, Vol. 51, No. 4, 2013, pp. 797–808. doi:10.2514/1.J051607

[18] Matsumura, T., Haftka, R. T., and Kim, N. H., "Accurate Predictions from Noisy Data: Replication Versus Exploration with Applications to Structural Failure," *Structural and Multidisciplinary Optimization*, Vol. 51, No. 1, 2015, pp. 23–40.

[19] Matsumura, T., "Effective Safety Measures with Tests Followed by Design Correction for Aerospace Structures," Ph.D. Dissertation, Dept. of Mechanical and Aerospace Engineering, Univ. of Florida, Gainesville, FL, 2013.

[20] Zhang, Y., Haftka, R. T., Kim, N. H., Seneviartne, W., and Schutte, J. F., "Allocation of Samples Between Exploration and Replication for Open-Hole-Tension Test," *Proceedings of the American Society of Composites 30th Annual Technical Conference*, Michigan State Univ., East Lansing, MI, 2015.

[21] Tomblin, J., and Seneviratne, W., "Laminate Statistical Allowable Generation for Fiber-Reinforced Composite Materials: Lamina Variability Method," US Department of Transportation, Federal Aviation Administration (Office of Aviation Research and Development) Rept. DOT/FAA/AR-06/53, Washington, D.C., Jan. 2009.

[22] "Standard Test Method for Open-Hole Tensile Strength of Polymer Matrix Composite Laminates," American Society for Testing and Materials, Rept. D5766/D5766M-11, West Conshohocken, PA, 2011.

[23] Dirikolu, M. H., Aktaş, A., and Birgören, B., "Statistical Analysis of Fracture Strength of Composite Materials Using Weibull Distribution," *Turkish Journal of Engineering and Environmental Sciences*, Vol. 26, No. 1, 2001, pp. 45–48.

[24] Deo, R. B., and Saff, C. R., *Composite Materials: Testing and Design*, American Society for Testing and Materials, West Conshohocken, PA, 1996, pp. 358–370.

[25] FAC Viana, *SURROGATES Toolbox User's Guide*, Gainesville, FL, Ver. 3.0, 2011, https://sites.google.com/site/srgtstoolbox/.

[26] Wisnom, M., Khan, B., and Hallett, S., "Size Effects in Unnotched Tensile Strength of Unidirectional and Quasi-Isotropic Carbon/Epoxy Composites," *Composite Structures*, Vol. 84, No. 1, 2008, pp. 21–28. doi:10.1016/j.compstruct.2007.06.002

[27] Mooney, C. Z., Duval, R. D., and Duval, R., *Bootstrapping: A Nonparametric Approach to Statistical Inference*, Newbury Park, London, 1993, pp. 30–41.

[28] Chakravarti, I. M., and Laha, R. G., "Handbook of Methods of Applied Statistics," *Handbook of Methods of Applied Statistics*, Wiley, New York, 1967, pp. 392–394.

[29] Welch, B. L., "The Generalization of 'Student's' Problem When Several Different Population Variances Are Involved," *Biometrika*, Vol. 34, Nos. 1–2, 1947, pp. 28–35. doi:10.2307/2332510

C. Bisagni
Associate Editor