



Design of validation experiments for life prediction models



Dan Ao, Zhen Hu, Sankaran Mahadevan*

Department of Civil and Environmental Engineering Vanderbilt University, Nashville, TN 37235, USA

ARTICLE INFO

Keywords:

Model validation
Life prediction
Experimental design
Accelerated life testing
Information theory

ABSTRACT

This paper proposes a novel validation experiment design optimization (VEDO) method for the assurance of life prediction model, which is one of the key steps in guaranteeing the reliable design of products in meeting the target service life. Life testing data collected from experiments are important for the validation of time-dependent models. However, directly collecting life data for model validation at the operating stress level is usually time-consuming and expensive. In order to overcome this challenge, the accelerated life testing (ALT) method is employed in the proposed method to collect data for model validation. The connection between ALT and model validation is established first; then a VEDO model is developed using the prior information obtained from the computer simulation model. In the VEDO model, the information gain for model validation is maximized within the testing budget and available testing chamber constraints. The obtained optimal number of tests and testing stress levels are designed to maximize the confidence in the validation results. Various sources of uncertainty such as prediction uncertainty, uncertainty of prior information, and observation errors are included within the optimization process in order to improve the robustness of validation experiment design. A composite helicopter rotor hub component is used to demonstrate the effectiveness of the proposed VEDO method.

1. Introduction

The development process of engineering products and systems often includes testing to assess their reliability and service life. In many cases, products and systems may be designed for long service lives and high reliability, and the tests might be expensive for complicated systems. Two types of strategies can be pursued to accelerate the development process and reduce the product development cost, namely accelerated testing and the use of computer simulation models. Accelerated testing subjects the system to elevated levels of demand, thus accelerating its failure; then empirical or assumed relationships are used to estimate reliability or service life under normal operating conditions. Computational models try to emulate the physics of the system and use simulation with assumed distributions of the inputs to predict the product life distribution. However, the life prediction model needs to be validated to make sure it can well represent the actual physical system, before it is applied to the product development. The subject of this paper is effective design of accelerated tests for the validation of life prediction models, thus speeding up the validation process and quantifying the confidence in the estimation of system reliability. Availability of such a quantitative accelerated approach will significantly support product/system development and certification.

Model validation is the process of quantifying the agreement

between model prediction and experimental data in order to guarantee that the prediction model can well represent the actual physical system [1]. Model validation can be performed qualitatively (i.e., graphical comparison) or quantitatively (using a validation metric). Many validation metrics have been proposed and investigated during the past decades, such as mean-based methods [2], hypothesis testing-based methods [3,4], area metric [5], and distance or reliability metric [6,7]. In addition to the above metrics, Kullback-Leibler (K-L) divergence is also widely used to measure the discrepancy between two distributions [8]. Based on the quantity of interest and the specific purpose, analysts may select different validation metrics for different problems.

In engineering settings, the validation problems can be either time-independent or time-dependent. Time-dependent problems refer to problems where the output is a function of time, such as fatigue crack growth [9], and stresses and deformation in structures under time-varying loads [10,11]. For time-dependent problems, the service life of products can be predicted from simulation models. For example, a crack growth model can be used to predict the fatigue life of a structure [12] and time-dependent reliability analysis can estimate the time to failure (TTF) of the structure under various sources of uncertainty [13,14]. Validation of time-dependent prediction models (e.g., life prediction model) is usually more challenging than that of time-

* Correspondence to: 272 Jacobs Hall, VU Mailbox: PMB 351831, Nashville, TN 37235, USA.
E-mail address: sankaran.mahadevan@vanderbilt.edu (S. Mahadevan).

independent models since collection of life data usually requires more effort [15]. Zhang and Mahadevan [15] proposed a Bayesian hypothesis testing-based method to describe the agreement of life interval predictions and observations [15].

In general, two issues have to be considered in the validation testing of life prediction models. The first issue is the financial aspect (testing cost) for validation of products designed for near-infinite service life. The second issue is that the testing conditions/environment be relevant to the operational conditions/environment of the product. The research of this paper is driven by the consideration of the first issue, namely, testing cost. There are two challenges in implementing the method presented in Ref. [15] for the validation of life prediction models belonging to the first issue. First, directly collecting life testing data for validation of time-dependent prediction models of products designed for near-infinite service life is time-consuming and expensive [16]. For problems with very low probability of failure, collecting life testing data under the operational condition will significantly delay the validation process. Second, due to budget constraints, only a limited number of tests may be possible; in that case, confidence in the validation result will depend on the number of tests. Therefore, the data will not only affect the lifecycle cost of product development but also the validation results [17]. In addition, budget and testing facility constraints affect the process of collecting data. In this situation, how to design the experiments to efficiently and effectively collect validation data is an important issue that needs to be addressed.

Experimental design is a process of determining the optimal experimental input settings to maximize the information gain from the experiments. Experimental design methods have been investigated for calibration, validation, or both in the past decades. For example, Huan and Marzouk developed experimental design methods for model calibration using information-theoretic measures and gradient-based stochastic optimization techniques [18,19]; Hu and Mahadevan proposed a calibration experiment design method for time-dependent reliability analysis [20]; Hu et al. investigated the effect of model uncertainty on the results of calibration experiment design [21]; Jiang and Mahadevan proposed a Bayesian cross entropy-based method for validation experiment design of computer simulation models [22]; Jiang and Mahadevan also developed a method to minimize the decision risk, by using likelihood ratio as a validation metric for assessment [17]; and Farrell and Oden applied the calibration and validation experiment design to semiconductor manufacturing [23]. The calibration and validation experiment design methods reviewed above, however, are difficult to apply in the validation of time-dependent prediction models due to the two challenges discussed earlier.

In order to overcome the first challenge (long life), this paper considers the use of accelerated testing, i.e., life testing data collected at stress levels higher than normal stress level, to validate the life prediction model. Here, the stress level refers to any factor that can be used in testing to reduce the time to failure, which can be temperature, force, voltage, humidity, etc [20]. The life testing data are collected at higher stress levels first. Based on that, the life distribution at the normal operational stress level is estimated according to a stress-life relationship. This process is called accelerated life testing (ALT) [20]. By comparing the life distribution obtained from ALT and the prediction model from computer simulation, model validity of life distribution model obtained from the computer simulation is assessed. Based on the proposed ALT-based validation method, the second challenge (variability of test data) is addressed through the formulation of a validation experiment design optimization (VEDO) model, which minimizes the uncertainty in the validation result under the budget and testing chamber constraints. In order to account for various sources of uncertainty present in VEDO, the objective function is formulated in the context of robust design optimization based on pre-posterior variance estimation. Finally, the optimization model is solved using the Fisher information matrix, surrogate modeling, and a

genetic algorithm. The obtained optimal testing stress levels and number of tests are expected to maximize the information gain for model validation.

The contributions of this paper are summarized as: (1) formulation of an ALT-based validation approach for time-dependent prediction models in order to reduce the time and cost of model validation; (2) formulation of a novel VEDO model to determine the optimal testing stress levels and number of tests for validation; and (3) integration of ALT statistical model, Bayesian approach, surrogate modeling, and experimental design optimization to solve the proposed VEDO model.

The reminder of paper is organized as follows. In Section 2, background concepts for validation of time-dependent prediction models are briefly introduced. In Section 3, we discuss the proposed VEDO method, which includes ALT-based validation method and VEDO model. A composite rotorcraft hub component example is studied in Section 4 to demonstrate the proposed method. Section 5 provides concluding remarks.

2. Background

2.1. Model validation

In the presence of uncertainty, several approaches have been investigated for quantitative model validation in the literature, and each method has advantages and disadvantages. Current model validation methods can be roughly classified into two groups: namely *hypothesis testing-based methods* and *non-hypothesis testing-based methods* [6].

In hypothesis testing, we decide the plausibility of two hypotheses - the null hypothesis (H_0) and the alternative hypothesis (H_1). The hypothesis testing can be based on classical or Bayesian statistics [6]. For given observation data y_D , in Bayesian hypothesis testing, the Bayes factor [24] given below is used as the validation metric:

$$B = \Pr(y_D|H_0)/\Pr(y_D|H_1) \quad (1)$$

The Bayes factor metric was originally developed to compare the data support for two physics models. It has been extended to compare two competing probability distribution models with uncertain parameters as follows [4]:

$$B = \Pr(y_D|H_0)/\Pr(y_D|H_1) = \int \Pr(y_D|\theta_i)\pi(\theta_i)d\theta_i / \int \Pr(y_D|\theta_j)\pi(\theta_j)d\theta_j \quad (2)$$

Bayesian hypothesis testing has also been investigated for equality hypotheses [3], interval hypotheses [3], and for validation data from fully or partially characterized experiments [25].

In non-hypothesis testing-based methods, the commonly studied validation metrics include the Mahalanobis distance [26], K-L divergence [27], area metric-based methods [5,28], and reliability-based metric [6,7]. Here, we briefly review the area metric and reliability-based metric since they have clear physical or probabilistic interpretations in terms of model validity, and both can be applied to validation of a model with multiple input variables using data from discrete test combinations [6].

2.1.1. Area metric

The area metric proposed by Ferson et al. [1,6] quantifies the mismatch between prediction and observation data using the area between the cumulative distribution (CDF) of model prediction and experimental data. This area can be expressed as

$$d(F_{ym}, S_{yn}) = \int_{-\infty}^{+\infty} |F_{ym}(y) - S_{yn}(y)| dy \quad (3)$$

where F_{ym} represents the CDF of model prediction and S_{yn} is the empirical CDF of experimental data.

The area metric can also be transformed from the physical space to probability space using the “u-pooling” procedure [29,30]. After the

transformation, the area metric is given by

$$d(F_u, S_u) = \int_0^1 |F_u - S_u| du \quad (4)$$

where F_u is the empirical CDF of all the u_i 's and S_u is the CDF of standard uniform random variable.

2.1.2. Reliability-based metric

When both model output and experimental data involve randomness, the difference between the responses for the same input is also random. For given distributions of model output and experimental data, the difference can also be described as a distribution. The reliability metric proposed by Rebba and Mahadevan [6,7] expresses the validation result as a simple probability. The reliability metric (r) is easy to interpret and compute. It indicates the probability that the difference (d) between observed data (Y_D) and model prediction (Y_M) is less than an accuracy requirement (ε), and is defined as [6,7].

$$r = P(-\varepsilon < d < \varepsilon), \quad d = Y_D - Y_M \quad (5)$$

A detailed review of various validation metrics is available in Ref. [6], and a detailed comparison of the area metric and reliability metric are given in Ref. [31].

Both the area metric and reliability-based metric can quantify the distribution mismatch over the entire validation domain. The area metric is simply the integration of the CDF difference between two distributions (predicted and observed) over the validation domain [5], while the reliability-based metric focuses on the disagreement in each small data interval and interprets the difference in terms of probability. The area metric is used in this paper for the purpose of illustration, but the proposed method can also be used with other validation metrics.

2.2. Validation of life prediction models

Defining the time-dependent prediction model as $g(\mathbf{S}, \mathbf{X}, \mathbf{Y}(t), t)$, where $\mathbf{S} = [S_1, S_2, \dots, S_c]$ is a vector of controllable variables (e.g., stress level) in experiments, $\mathbf{X} = [X_1, X_2, \dots, X_n]$ is a vector of other input random variables, $\mathbf{Y}(t) = [Y_1(t), Y_2(t), \dots, Y_m(t)]$ is a vector of input stochastic processes, and t is time, we have the output prediction function as $G(t) = g(\mathbf{S}, \mathbf{X}, \mathbf{Y}(t), t)$.

If we can directly measure the output $G(t)$ in experiments, the time-dependent prediction model $G(t) = g(\mathbf{S}, \mathbf{X}, \mathbf{Y}(t), t)$ can be validated directly using the validation metrics reviewed in Section 2.1. For some problems, however, it is difficult to directly measure $G(t)$ in experiments, such as the energy release rate in the numerical example presented in Section 4. Besides, $G(t)$ sometimes may not directly be the quantity of interest (QoI); instead, $G(t)$ may be the input for a life prediction model, and the QoI may be product life. In this situation, we have the life prediction model as

$$L = g_{life}(G(t), \boldsymbol{\theta}) = g_{life}(g(\mathbf{S}, \mathbf{X}, \mathbf{Y}(t), t), \boldsymbol{\theta}) \quad (6)$$

where L is life, $\boldsymbol{\theta}$ is a vector of parameters of the life prediction model, and $g_{life}(\cdot)$ is the life prediction model. For a given value of \mathbf{S} , L is a random variable due to the uncertainty in \mathbf{X} and $\mathbf{Y}(t)$,

The application of the validation metrics reviewed in Section 2.1 to the validation of the life prediction model of time-dependent problems is not straightforward if we do not have a lot of life testing data. In order to overcome this challenge, Zhang and Mahadevan [15] extended the Bayesian hypothesis testing to time-dependent problems by assessing the agreement of the life intervals between predictions and observations. From the model prediction, the probability of life falling in a given life interval is described as [15].

$$p_i^0 = F_L(l_i) - F_L(l_{i-1}) \quad (7)$$

where $F_L(\cdot)$ is the CDF of life distribution, l_{i-1} and l_i are the lower and upper bounds of the life interval, and p_i^0 is the probability value based on model prediction. Then the Bayes factor can be computed as [15].

$$B_i = (p_i^0)^{k_i} (1 - p_i^0)^{n-k_i} / \int_0^1 (p_i)^{k_i} (1 - p_i)^{n-k_i} f(p_i) dp_i \quad (8)$$

where B_i is the Bayes factor of the i -th interval, p_i is the true probability that life falls in the i -th interval, uniformly distributed between $[0,1]$, k_i is the number of observations counted in the i -th interval, n is the total number of observations. More details about the method can be found in [15].

Even though the extension of the validation metrics reviewed in Section 2.1 to life prediction models is possible, we observe two challenges that need to be addressed. (1) For products designed for near-infinite service life, it is not practical for us to collect life testing data directly because it is very time-consuming. Accelerated life testing (ALT) is often employed; however, in this case, how to validate the life prediction model? To answer this question, this paper will first investigate the connection between ALT and model validation through the fusion of prior information (i.e. prior distribution and stress-life function from the simulation model) and experimental data (i.e. from ALT). Based on that, a validation method based on ALT data is developed. (2) Given limited resources, the validation results may be affected by the experimental data. How should we perform the optimal experiments to collect the most useful data for validation? In this paper, a new validation experiment design optimization (VEDO) method is proposed to address the second challenge.

3. Proposed methodology

In this section, we first develop an approach to validate life prediction models based on ALT. Following that, we study how to design the validation experiments.

3.1. Validation of life prediction models using ALT

Here, the basic concept of ALT is introduced first. Following that, model validation using ALT data is discussed.

3.1.1. Accelerated life testing (ALT)

ALT is widely used in the industry to estimate the reliability of a product under nominal stress level within a reasonable testing time-frame. Here, the stress is the controllable experimental variable \mathbf{S} , which can be temperature, load, voltage, etc. [32]. Based on testing data obtained at stress levels that are higher than the nominal stress level, the life distribution at the nominal stress level is estimated. In ALT, the life time is often assumed to follow log-location-scale distributions (such as Weibull distribution, exponential distribution, and lognormal distribution) [33–35], piecewise exponential distribution [32], or Gompertz-Makeham distribution [32]. The log-location-scale distributions with CDF given as below are commonly used [36]:

$$F_L(t) = F_{\log}((\log(t) - \alpha)/\lambda) \quad (9)$$

where α and λ are the location and scale parameters of log time and $F_{\log}(\cdot)$ is a log scale distribution. When $F_{\log}(\cdot) = 1 - \exp(-\exp(\cdot))$, $F_L(t)$ is Weibull distribution; when $F_{\log}(\cdot)$ is the CDF of a standard normal variable, $F_L(t)$ is lognormal distribution [36]. Based on the distribution assumption, distribution parameters at different testing stress levels are estimated. The distribution parameters are then fitted as a function of stress level in the logarithmic coordinate as shown in Fig. 1. Based on the fitted function, the life distribution at the nominal stress level is estimated.

If the Weibull distribution is used to describe the life distribution, we have the probability density function (PDF) of life as $f_L(t) = \lambda \alpha^{-\lambda} t^{\lambda-1} \exp[-(t/\alpha)^\lambda]$. It is typically assumed that λ is a constant and $\log(\alpha)$ is a function of stress level \mathbf{S} . Commonly used functions are linear and quadratic functions. When a linear function is used, we have

$$\log(\alpha(\mathbf{S})) = \boldsymbol{\beta}^T \mathbf{S} = \beta_0 + \beta_1 S_1 + \dots + \beta_i S_i \quad (10)$$

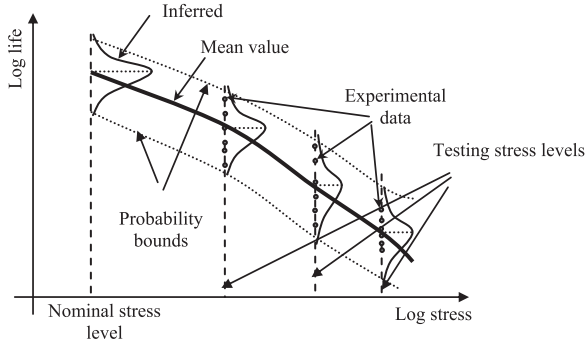


Fig. 1. Stress-life relationship in ALT.

in which β is a vector of coefficients and S_i is the i -th stress variable.

Based on the life testing data at the higher stress levels, the coefficients β and λ are estimated and thus the life distribution at the nominal stress level is approximated as indicated in Fig. 1. Note that, censored data (such as right censoring and interval censoring) are often encountered in ALT [36,37]. In this paper, we assume that all the specimens are tested until failure. Cases involving censored data will be investigated in our future work. Note that ALT has been extensively studied in the literature [38–42]; also, statistical software tools such as the ALT module of the JMP software and the ALTA module of ReliaSoft software offer the capability of accelerated testing design. The focus of this paper, however, is not about ALT. This paper concentrates on how to use ALT to validate the life distribution predicted by a physics-based computer simulation model (such as the combination of finite-element simulation and fatigue damage accumulation model to predict life in mechanical components), and how to optimize the ALT plan for this model validation. In the subsequent sections, we will discuss how to perform validation using ALT and how to optimize the testing plan.

3.1.2. Validation using ALT

In order to validate the life prediction model at the nominal stress level S_0 , we first get the predicted life distribution using $L = g_{life}(G(t), \theta) = g_{life}(g(S_0, \mathbf{X}, \mathbf{Y}(t), t), \theta)$ by propagating the uncertainty in \mathbf{X} and $\mathbf{Y}(t)$ through $g_{life}(\cdot)$. This uncertainty propagation can be achieved through Monte Carlo simulation (MCS) by considering the uncertainty sources in the simulation model such as input variable uncertainty (aleatory uncertainty) and model parameter uncertainty (epistemic uncertainty). We define the obtained life distribution L from the simulation model as $F_p(t, S_0)$. Note that for the purposes of $F_p(t, S_0)$ computation, the distributions corresponding to input random variables and model parameters are treated the same way. Next, we need to validate $F_p(t, S_0)$ using ALT.

As discussed in Section 3.1.1, the selection of life distribution type (i.e. Weibull, lognormal, or exponential) and the stress-parameter function (i.e. linear, quadratic) is important in ALT. The selection will affect the life distribution inferred at the nominal stress level. The Weibull distribution and the empirical linear relationship given in Eq. (10) are commonly used in ALT-based methods. Since the physics-informed simulation model can predict the life distribution $F_p(t, S)$, $\forall S \in \Omega_S$, where Ω_S is the domain of the stress levels, in this paper, we select the life distribution type and stress-life relationship using the information obtained from the simulation model. Note that the assumption of distribution type and stress-life relationship affects the validation result. In the absence of any other information, the simulation model can be used to guide the selection of distribution type and stress-life relationship. Validation experiment design accounting for uncertainty in both distribution type and stress-life relationship will be studied in our future work.

Suppose that the distribution type is selected as $F_{ALT}(t, \lambda, \alpha)$, where $F_{ALT}(t, \lambda, \alpha)$ is the CDF of the distribution, and λ and α are the

distribution parameters. The stress-parameter relationship is selected as $\alpha = g_\alpha(S, \beta)$, where β is the parameter of the function. Based on the life testing data at stress levels $S > S_0$, we can estimate the values of β and λ . Once we have β and λ , the CDF of life distribution at the nominal stress level S_0 is estimated from ALT as $F_{ALT}(t, \lambda, \alpha(S_0))$, where $\alpha(S_0) = g_\alpha(S_0, \beta)$.

The distribution $F_{ALT}(t, \lambda, \alpha(S_0))$ obtained from ALT statistical model of the experimental data is similar to the empirical CDF presented in Eq. (3). We can therefore validate $F_p(t, S_0)$ obtained from the computer simulations by computing the area between these two distributions as

$$d(F_p, F_{ALT})|\lambda, \beta = \int_{-\infty}^{+\infty} |F_p(t, S_0) - F_{ALT}(t, \lambda, \alpha(S_0))| dt \quad (11)$$

where $d(F_p, F_{ALT})|\lambda, \beta$ is the area metric conditioned on λ and β . Note that the validation metric is not limited to the area metric. Other validation metrics can also be used. The above validation is based on the comparison between life distribution obtained from the ALT statistical model (empirical distribution) and its counterpart obtained from the computer simulation model; the latter is the focus of validation. The empirical life distribution model is based on the experimental data at higher stress levels using ALT. Here we mainly focus on model validation at the nominal stress level. However, the proposed approach can also be extended to other stress levels.

Even though it is easier to collect life testing data using ALT, we still have limited testing data due to budget constraints. The limited data will lead to the uncertainty in λ and β . As shown in Fig. 2, for a given realization of λ and β , we can obtain the life distribution at the nominal stress level from ALT, which gives a particular value of the validation metric. To account for the uncertainty in λ and β , we use the expected area as the validation metric in the ALT-based validation method. The expected value is computed as

$$\bar{d}(F_p, F_{ALT}) = \int_{\Omega_\lambda} \int_{\Omega_\beta} [d(F_p, F_{ALT})|\lambda, \beta] f_\lambda(\lambda) f_\beta(\beta) d\lambda d\beta \quad (12)$$

in which Ω_λ is the domain of λ , Ω_β is the domain of β , $f_\lambda(\lambda)$ and $f_\beta(\beta)$ are the PDF of λ and β , respectively. The PDFs $f_\lambda(\lambda)$ and $f_\beta(\beta)$ are obtained from Bayesian calibration, which will be discussed in Section 3.2.

The uncertainty in the validation metric can be quantified using the standard deviation of $d(F_p, F_{ALT})$ as below:

$$\sigma_d^2(F_p, F_{ALT}) = \int_{\Omega_\lambda} \int_{\Omega_\beta} [d(F_p, F_{ALT})|\lambda, \beta - \bar{d}(F_p, F_{ALT})]^2 f_\lambda(\lambda) f_\beta(\beta) d\lambda d\beta \quad (13)$$

where the CDF values of F_p are obtained from the life distribution estimate obtained from the computer simulation model.

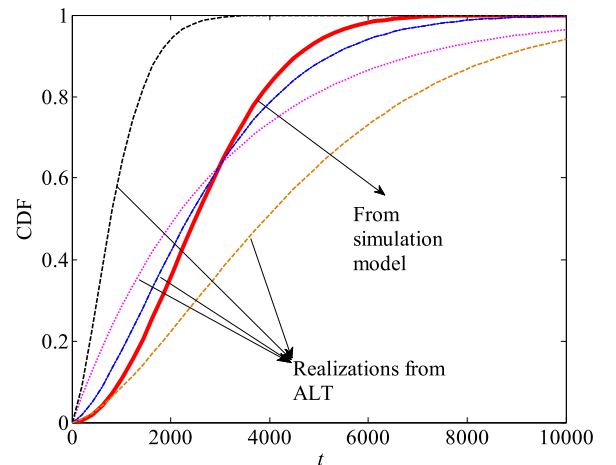


Fig. 2. Illustration of validation using ALT.

A smaller value of $\sigma_d^2(F_p, F_{ALT})$ implies a higher confidence in the model validation result. From the above discussion, it can be found that the uncertainty of λ and β is affected by the input values at which the testing data is collected, which will then affect $\sigma_d^2(F_p, F_{ALT})$.

In summary, there are four key steps in validation using ALT. First, we select the life distribution type and stress-life relationship using the information obtained from the simulation model. Second, we collect life testing data at elevated testing stress levels, and calibrate the parameter λ and stress-life function coefficients β . During this process, epistemic uncertainty due to limited testing data is also quantified. Third, we evaluate the ALT-based life distribution at the normal operating stress level ($F_{ALT}(t, \lambda, \alpha(S_0))$) according to the updated stress-life relationship. We can get a family of ALT-based life distributions due to uncertainty in the calibrated parameters (i.e. λ and β). Finally, validation result with uncertainty is obtained by comparing the life distribution from simulation model ($F_p(t, S_0)$) with that from ALT data ($F_{ALT}(t, \lambda, \alpha(S_0))$) at the normal operating stress level. Obviously, the validation result will be affected by the experimental data if testing resources are limited. In the next subsection, we will discuss how to design the optimal ALT experimental inputs so that $\sigma_d^2(F_p, F_{ALT})$ is minimized.

3.2. Validation experiment design optimization (VEDO)

3.2.1. VEDO optimization model formulation

In order to optimize the ALT inputs, the design variables include the testing stress levels ($S_i = S_i, i = 1, 2, \dots, n_S$, where n_S is the number of testing stress levels) and the number of tests at each stress level ($N_i = N_i, i = 1, 2, \dots, n_S$). Since testing is yet to be done, we need to use synthetic data for VEDO; this synthetic data can be generated based on an assumed ALT model. Note that the computational simulation model is not used to generate the synthetic data for VEDO because the simulation model is not validated yet. Suppose the true values of λ and β are λ^* and β^* ; for a given stress level S_i , we will have the CDF of life distribution as $F_{ALT}(t, \lambda^*, \alpha(S_i))$, where $\alpha(S_i) = g_a(S_i, \beta^*)$ and the PDF of life distribution as $f_{ALT}(t, \lambda^*, \alpha(S_i))$. If N_i tests are performed, we will have the life data at stress level S_i as $T_i = [T_i^{(1)}, T_i^{(2)}, \dots, T_i^{(N_i)}]$, where $T_i^{(j)}$ is the life testing data at stress level S_i obtained from the j -th experiment. The synthetic data $T_i^{(j)}$ is generated based on the following relationship:

$$T_i^{(j)} = \mu_{T_i} + \varepsilon_{obs} \quad (14)$$

where μ_{T_i} is a random variable with PDF given as $f_{ALT}(\mu_{T_i}, \lambda^*, \alpha(S_i))$ and ε_{obs} is observation error.

Each $T_i^{(j)}, j = 1, 2, \dots, N_i$, is a random variable with the PDF given as

$$f_{T_i^{(j)}}(t) = \int_0^\infty \phi((t - \mu)/\sigma_{obs}) f_{ALT}(\mu, \lambda^*, \alpha(S_i)) d\mu \quad (15)$$

where $\phi(\cdot)$ is the PDF of a standard normal variable, σ_{obs} is the standard deviation of the observation error.

Based on Eq. (15), the likelihood function of obtaining the life testing data $\mathbf{T}_{ALT} = [\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_{n_S}]$ from a testing plan $\mathbf{S}_t = [S_1, \dots, S_{n_S}]$ and $\mathbf{N}_t = [N_1, \dots, N_{n_S}]$ is given by

$$P(\mathbf{T}_{ALT}|\lambda, \beta) = \prod_{i=1}^{n_S} \prod_{j=1}^{N_i} f_{T_i^{(j)}}(T_i^{(j)}) \quad (16)$$

With the above likelihood function, the posterior distributions of λ and β are then obtained using Bayes' theorem as

$$f''(\lambda, \beta) = \frac{P(\mathbf{T}_{ALT}|\lambda, \beta) f'(\lambda, \beta)}{\int P(\mathbf{T}_{ALT}|\lambda, \beta) f'(\lambda, \beta) d\lambda d\beta} \quad (17)$$

in which $f'(\lambda, \beta)$ is the prior distribution of the ALT parameters and $f''(\lambda, \beta)$ is the posterior distribution.

Note that the above posterior distribution is conditioned on the

testing data $\mathbf{T}_{ALT} = [\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_{n_S}]$. Since synthetic data are generated using Eq. (14), there is uncertainty in \mathbf{T}_{ALT} . In order to account for the uncertainty in \mathbf{T}_{ALT} , the expected posterior distribution is used and is given by

$$\bar{f}''(\lambda, \beta) = \int [f''(\lambda, \beta) | \mathbf{T}_{ALT}] f(\mathbf{T}_{ALT}) d\mathbf{T}_{ALT} \quad (18)$$

where $f(\mathbf{T}_{ALT})$ is the joint PDF of \mathbf{T}_{ALT} computed using Eq. (15) and $f''(\lambda, \beta) | \mathbf{T}_{ALT}$ is the posterior distribution conditioned on \mathbf{T}_{ALT} . For given \mathbf{T}_{ALT} , $f''(\lambda, \beta) | \mathbf{T}_{ALT}$ is computed using Eqs. (16) and (17).

Once the expected posterior distribution $\bar{f}''(\lambda, \beta)$ is available, the standard deviation of the validation metric ($\sigma_d(F_p, F_{ALT})$) can be computed using Eq. (13). Then the VEDO model is formulated as

$$\begin{aligned} \min_{\mathbf{S}_t, \mathbf{N}_t} & \sigma_d(F_p, F_{ALT}) \\ \text{s. t.} & C_{test} \leq C_{budget} \\ & N_{min} \leq N_i \leq N_{max}, \quad \forall i = 1, 2, \dots, n_S \\ & S_{min} \leq S_i \leq S_{max}, \quad \forall i = 1, 2, \dots, n_S \\ & \sum_{i=1}^{n_S} N_i \leq N_{max} \end{aligned} \quad (19)$$

where C_{test} is the total cost of ALT, C_{budget} is the allowable budget, S_{min} and S_{max} are the minimum and maximum testing stress level, and N_{min} and N_{max} are the minimum and maximum number of tests. The constraint $C_{test} \leq C_{budget}$ is the budget constraint. $N_{min} \leq N_i \leq N_{max}$ and $\sum_{i=1}^{n_S} N_i \leq N_{max}$ are the number of testing chambers constraint, and $S_{min} \leq S_i \leq S_{max}$ is the testing stress level constraint. The objective function in Eq. (19) implies that we are trying to minimize the uncertainty in the validation result due to limited testing data.

Note that the optimization model given in Eq. (19) is formulated based on the assumption that the true values of λ and β are λ^* and β^* . In reality, however, we do not know the true values of λ and β before testing. For given prior distributions of λ and β , the true values can be any values within the prior distributions. Instead of using arbitrarily assumed prior distributions, in this paper, the simulation model is used to provide the prior information. Since the optimization model given in Eq. (19) is conditioned on λ and β , the VEDO model is further modified as below to make the optimization robust to the uncertainty in λ and β .

$$\begin{aligned} \min_{\mathbf{S}_t, \mathbf{N}_t} & f_{obj}(\mathbf{S}_t, \mathbf{N}_t) = w_1 E_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})] + (1 - w_1) \sigma_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})] \\ \text{s. t.} & C_{test} \leq C_{budget} \\ & N_{min} \leq N_i \leq N_{max}, \quad \forall i = 1, 2, \dots, n_S \\ & S_{min} \leq S_i \leq S_{max}, \quad \forall i = 1, 2, \dots, n_S \\ & \sum_{i=1}^{n_S} N_i \leq N_{max} \end{aligned} \quad (20)$$

in which $w_1 \in [0, 1]$ is the weight, $E_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})]$ and $\sigma_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})]$ are the expectation and standard deviation of $\sigma_d(F_p, F_{ALT})$ by considering the uncertainty in the prior distribution of λ and β . The objective function now consists of three elements, namely $\sigma_d(F_p, F_{ALT})$, and mean and standard deviation of $\sigma_d(F_p, F_{ALT})$ by considering the uncertainty in the priors of λ and β since the true values of λ and β can be any value in their prior distributions. $\sigma_d(F_p, F_{ALT})$ accounts for the uncertainty in the validation result due to limited testing data. The mean and standard deviation of $\sigma_d(F_p, F_{ALT})$ are both affected by the prior distributions. The new objective function can therefore account for uncertainty due to both limited testing data and the prior distributions. The weights of $E_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})]$ and $\sigma_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})]$ are used to reflect decision makers' preference on the prior distribution, which is similar to a robust design optimization problem.

By solving Eq. (20), the expected uncertainty in our estimation of validation metric will be minimized and thus our confidence in the validation result is maximized. However, directly solving Eq. (20) is very challenging due to the involvement of Eqs. (17), (18) and C_{test} in the constraint. Next, we will discuss how to efficiently solve the VEDO model.

3.2.2. Solution of the VEDO model

We first summarize the challenges in solving the VEDO model given in Eq. (20). After that, we will discuss how to address those challenges. There are three main challenges in solving Eq. (20):

1. Suppose the true values of λ and β are λ^* and β^* . How to efficiently obtain the expected posterior distribution $\bar{f}^p(\lambda, \beta)$ by considering the uncertainty in the synthetic observation data?
2. How to estimate the expected testing cost C_{test} used in the constraint?
3. Given challenges 1 and 2, how to efficiently solve the optimization model?

(a) Estimation of $\bar{f}^p(\lambda, \beta)$ using Fisher information matrix

A key step in the VEDO method is to estimate the posterior distribution $\bar{f}^p(\lambda, \beta)$. Some analysts directly use a sampling-based Bayesian updating approach (e.g. particle filter, MCMC, slice sampling etc.) to get the posterior distribution [18], and then evaluate the quantity of interest. These sampling approaches, however, are time-consuming, and may significantly delay the design process. The Fisher information matrix, which infers the posterior distribution analytically, has been widely used in test planning in industrial engineering and reliability engineering [43].

For a q -dimensional calibration parameter vector $\alpha = [\alpha_1, \alpha_2, \dots, \alpha_q]$, when the sample size of observation data is large (> 30), it is assumed that the posterior distribution follows a multivariate normal distribution given by $N(\alpha^*, \Sigma)$, where α^* is the maximum likelihood estimate (MLE) of the calibration parameter and Σ is the covariance matrix of the posterior distribution. (The posterior distribution can be assumed to be multivariate normal when the number of observations is large compared to the number of parameters being updated [44]. The assumption of a multivariate normal posterior distribution for model parameters is also common in the field of experiment design [43]). The Fisher information matrix is a way to calculate Σ associated with the maximum likelihood estimate, and to measure the amount of information the data carries relative to the parameters being estimated [45], which is represented as smaller values of the elements in the covariance matrix Σ when more information is carried by the data.

The (i, j) -th element of the Fisher information matrix is the negative second-order derivative of the log-likelihood function given as [45]

$$I_{ij}(\alpha) = -\frac{\partial^2}{\partial \alpha_i \partial \alpha_j} \log(L(D|\alpha)), \quad i, j = 1, 2, \dots, q \quad (21)$$

where D is the observation data and $L(D|\alpha)$ is the likelihood function.

By considering the uncertainty in the observation data, the (i, j) -th element of the *expected Fisher information matrix* with n_{obs} observations is given by [45]

$$I_{ij}(\alpha) = -n_{obs} E_{\alpha}^D \left[\frac{\partial^2}{\partial \alpha_i \partial \alpha_j} \log(L(D|\alpha)) \right], \quad i, j = 1, 2, \dots, q \quad (22)$$

Once the Fisher information matrix is available, the covariance matrix of calibration parameters can be estimated. In this paper, the Fisher information matrix is extended to validation experiment design.

When λ^*, β^* are used to generate the synthetic data, then the

maximum likelihood estimates (MLE) of λ and β are λ^* and β^* . The covariance matrix Σ is estimated based on the expected Fisher information matrix as

$$\Sigma = [I(\lambda, \beta)]_{\lambda, \beta=\lambda^*, \beta^*}^{-1} \quad (23)$$

where $I(\lambda, \beta)$ is the expected Fisher information matrix.

Defining $\mathbf{v} = [\lambda, \beta]$, the (i, j) -th element of $I(\lambda, \beta) = I(\mathbf{v})$ is given by

$$I_{ij}(\lambda, \beta) = I_{ij}(\mathbf{v}) = \frac{\partial^2}{\partial v_i \partial v_j} \log(P(\mathbf{T}_{ALT}|\mathbf{v})) \quad (24)$$

Substituting Eq. (16) into Eq. (24), we have

$$I_{ij}(\lambda, \beta) = \sum_{k=1}^{n_S} \sum_{r=1}^{N_k} \frac{\partial^2}{\partial v_i \partial v_j} \log(P(T_k^{(r)}|\mathbf{v})) \quad (25)$$

Note that k and r in Eq. (25) correspond to i and j in Eq. (16).

Since $T_k^{(r)}, r = 1, 2, \dots, N_k$ are identical independent random variables (i.i.d), the above equation is rewritten as

$$I_{ij}(\lambda, \beta) = \sum_{k=1}^{n_S} N_k \frac{\partial^2}{\partial v_i \partial v_j} \log(P(T_k^{(1)}|\mathbf{v})) \quad (26)$$

where $P(T_k^{(1)}|\mathbf{v}) = P(T_k^{(1)}|\lambda, \beta)$ is computed using Eq. (16).

After considering the uncertainty in $T_k^{(1)}$, the (i, j) -th element of $I(\lambda, \beta) = I(\mathbf{v})$ is given by

$$I_{ij}(\lambda, \beta) = \sum_{k=1}^{n_S} N_k \int \left[\frac{\partial^2}{\partial v_i \partial v_j} \log(P(t_k|\mathbf{v})) \right] f_{T_k^{(1)}}(t_k) dt_k \quad (27)$$

where $f_{T_k^{(1)}}(t)$ is given in Eq. (15).

The above equation can be further transformed as

$$I_{ij}(\lambda, \beta) = \sum_{k=1}^{n_S} N_k \int_0^1 \left[\frac{\partial^2}{\partial v_i \partial v_j} \log(P(F_{T_k^{(1)}}^{-1}(P_k)|\mathbf{v})) \right] dP_k \quad (28)$$

where $F_{T_k^{(1)}}^{-1}(P_k)$ is the inverse CDF of $T_k^{(1)}$.

Suppose that n_G Gauss quadrature points are used to solve the one-dimensional integration given in Eq. (28). Then we have

$$I_{ij}(\lambda, \beta) = \sum_{k=1}^{n_S} N_k \left[\sum_{q=1}^{n_G} \omega_q \frac{\partial^2}{\partial v_i \partial v_j} \log(P(t_q|\mathbf{v})) \right] \quad (29)$$

in which $t_q = F_{T_k^{(1)}}^{-1}(\frac{g_q+1}{2})$ is the life corresponding to the Gauss quadrature point $g_q, q = 1, 2, \dots, n_G$, and ω_q is the weight of g_q [46].

Since a known distribution type is usually assumed for the life distribution, we have the analytical expression for $P(t_q|\mathbf{v})$, $\forall q = 1, 2, \dots, n_G, k = 1, 2, \dots, n_S$. When the commonly used Weibull distribution is employed, we have

$$\begin{aligned} P(t_q|\mathbf{v}) &= P(t_q|\lambda, \beta) \\ &= \int \phi\left(\frac{t_q - \mu}{\sigma_{obs}}\right) \lambda \alpha^{-\lambda} (S_k) \mu^{\lambda-1} \exp\left(-\left(\frac{\mu}{\alpha(S_k)}\right)^{\lambda}\right) d\mu \\ &= \lambda \alpha^{-\lambda} (S_k) \int \phi\left(\frac{t_q - \mu}{\sigma_{obs}}\right) \mu^{\lambda-1} \exp\left(-\left(\frac{\mu}{\alpha(S_k)}\right)^{\lambda}\right) d\mu \end{aligned} \quad (30)$$

where $\alpha(S_k) = g_{\alpha}(S_k, \beta)$ and μ is the mean of the synthetic experimental data as defined in Eq. (14). The logarithmic form of $P(t_q|\mathbf{v})$ is

$$\begin{aligned} \log(P(t_i|\mathbf{v})) &= \log(\lambda) - \lambda \log(\alpha(\mathbf{S}_k)) \\ &+ \log \left[\int \phi \left(\frac{t_i - \mu}{\sigma_{obs}} \right) \mu^{\lambda-1} \exp \left(- \left(\frac{\mu}{\alpha(\mathbf{S}_k)} \right)^\lambda \right) d\mu \right] \end{aligned} \quad (31)$$

The second-order derivatives in Eq. (26) can then be computed using the chain rule as

$$\frac{\partial^2 \log(P(t_i|\mathbf{v}))}{\partial \lambda \partial \beta_i} = \frac{\partial^2 \log(P(t_i|\mathbf{v}))}{\partial \lambda \partial \alpha(\mathbf{S}_k)} \frac{\partial g_\alpha(\mathbf{S}_k, \boldsymbol{\beta})}{\partial \beta_i}, \forall i = 1, 2, \dots, n_\beta \quad (32)$$

$$\frac{\partial^2 \log(P(t_i|\mathbf{v}))}{\partial^2 \lambda} = -\frac{1}{\lambda^2} - \frac{1}{F_l^2} \left(\frac{\partial F_l}{\partial \lambda} \right)^2 + \frac{1}{F_l} \frac{\partial^2 F_l}{\partial^2 \lambda} \quad (33)$$

and

$$\begin{aligned} \frac{\partial^2 \log(P(t_i|\mathbf{v}))}{\partial \beta_i \partial \beta_j} &= \frac{\partial^2 \log(P(t_i|\mathbf{v}))}{\partial^2 \alpha(\mathbf{S}_k)} \frac{\partial g_\alpha(\mathbf{S}_k, \boldsymbol{\beta})}{\partial \beta_i} \frac{\partial g_\alpha(\mathbf{S}_k, \boldsymbol{\beta})}{\partial \beta_j} \\ &+ \frac{\partial \log(P(t_i|\mathbf{v}))}{\partial \alpha(\mathbf{S}_k)} \frac{\partial^2 g_\alpha(\mathbf{S}_k, \boldsymbol{\beta})}{\partial \beta_i \partial \beta_j}, \forall i, j = 1, 2, \dots, n_\beta \end{aligned} \quad (34)$$

where $F_l = \int \phi \left(\frac{t_i - \mu}{\sigma_{obs}} \right) \mu^{\lambda-1} \exp \left(- \left(\frac{\mu}{\alpha(\mathbf{S}_k)} \right)^\lambda \right) d\mu$, n_β is the number of coefficients in the model $\alpha(\mathbf{S}_i) = g_\alpha(\mathbf{S}_i, \boldsymbol{\beta})$, and $\frac{\partial F_l}{\partial \lambda}$ and $\frac{\partial F_l}{\partial \alpha(\mathbf{S}_k)}$ are given by

$$\begin{aligned} \frac{\partial F_l}{\partial \lambda} &= \int \phi \left(\frac{t_i - \mu}{\sigma_{obs}} \right) \left[\log(\mu) - \log \left(\frac{\mu}{\alpha(\mathbf{S}_k)} \right) \left(\frac{\mu}{\alpha(\mathbf{S}_k)} \right)^\lambda \right] \mu^{\lambda-1} \\ &\times \exp \left(- \left(\frac{\mu}{\alpha(\mathbf{S}_k)} \right)^\lambda \right) d\mu \end{aligned} \quad (35)$$

$$\frac{\partial F_l}{\partial \alpha(\mathbf{S}_k)} = \int \phi \left(\frac{t_i - \mu}{\sigma_{obs}} \right) \mu^{\lambda-1} \exp \left(- \left(\frac{\mu}{\alpha(\mathbf{S}_k)} \right)^\lambda \right) \left(\frac{\mu}{\alpha(\mathbf{S}_k)} \right)^{\lambda-1} \frac{\lambda \mu}{\alpha^2(\mathbf{S}_k)} d\mu \quad (36)$$

The other derivatives required in Eqs. (32)–(34) can be derived similarly as in Eqs. (35) and (36). One reason that the above equations look complicated is due to the inclusion of observation error which introduces a one-dimensional integration in Eq. (15). When the observation error is not considered, Eqs. (30) through (36) can be simplified. Once the elements of the Fisher information matrix are available, the posterior distributions of λ and $\boldsymbol{\beta}$ can be approximated using $N(\mathbf{v}, \boldsymbol{\Sigma})$ and Eq. (23). Based on the posterior distribution, $\sigma_d(F_{simu}, F_{ALT})$ conditioned on given values of λ and $\boldsymbol{\beta}$ is estimated using Eq. (13). It should be noted that the domain of λ and $\boldsymbol{\beta}$ in multivariate normal distribution $N(\mathbf{v}, \boldsymbol{\Sigma})$ is $[-\infty, \infty]$. However, in Bayesian calibration, we assume prior distributions for the calibration parameters λ and $\boldsymbol{\beta}$. When non-informative uniform distributions are used as priors for λ and $\boldsymbol{\beta}$, the multivariate normal distribution needs to be truncated according to the lower and upper bounds of the priors and thus $\sigma_d(F_p, F_{ALT})|\lambda^*, \boldsymbol{\beta}^*$ is computed as

$$\sigma_d(F_p, F_{ALT})|\lambda^*, \boldsymbol{\beta}^* = \int_{\lambda_L}^{\lambda_U} \int_{\boldsymbol{\beta}_L}^{\boldsymbol{\beta}_U} [d(F_p, F_{ALT})|\lambda, \boldsymbol{\beta} - \bar{d}(F_p, F_{ALT})]^2 \mathcal{F}^{\pi}(\lambda, \boldsymbol{\beta}) d\lambda d\boldsymbol{\beta} \quad (37)$$

where λ_L and λ_U , $\boldsymbol{\beta}_L$ and $\boldsymbol{\beta}_U$ are the lower and upper bounds of λ and $\boldsymbol{\beta}$ priors, and $\mathcal{F}^{\pi}(\lambda, \boldsymbol{\beta})$ is computed using the multivariate normal distribution $N(\mathbf{v}, \boldsymbol{\Sigma})$.

Based on the above equation, the objective function in Eq. (20) is estimated as

$$E_{\lambda, \boldsymbol{\beta}}[\sigma_d(F_p, F_{ALT})] = \int_{\lambda_L}^{\lambda_U} \int_{\boldsymbol{\beta}_L}^{\boldsymbol{\beta}_U} [\sigma_d(F_p, F_{ALT})|\lambda^*, \boldsymbol{\beta}^*] f'(\lambda^*, \boldsymbol{\beta}^*) d\lambda^* d\boldsymbol{\beta}^* \quad (38)$$

$$\begin{aligned} \sigma_{\lambda, \boldsymbol{\beta}}^2[\sigma_d(F_p, F_{ALT})] &= \int_{\lambda_L}^{\lambda_U} \int_{\boldsymbol{\beta}_L}^{\boldsymbol{\beta}_U} (\sigma_d(F_p, F_{ALT})|\lambda^*, \boldsymbol{\beta}^* f'(\lambda^*, \boldsymbol{\beta}^*) \\ &- E_{\lambda, \boldsymbol{\beta}}[\sigma_d(F_p, F_{ALT})])^2 \\ &\boldsymbol{\beta}^*) d\lambda^* d\boldsymbol{\beta}^* \end{aligned} \quad (39)$$

The above equations can be solved using either numerical integration methods (when the dimension of calibration parameters is not high) or Monte Carlo simulation (MCS). Next, we will discuss how to estimate the constraints for the optimization in Eq. (20).

(b) Computing the expected testing cost C_{test}

The ALT cost consists of two parts: (1) manufacturing cost to make the test specimens; and (2) labor cost to conduct the tests. The former cost is only related to the number of tests, while the latter is determined by both the number of tests and the test durations. Since before any actual tests, we do not know the test durations, expected test durations will be used here. Thus the ALT cost is defined as

$$C_{test} = C_s \sum_{i=1}^{n_S} N_i + C_l \sum_{i=1}^{n_S} N_i t_{test}(\mathbf{S}_i) \quad (40)$$

where $t_{test}(\mathbf{S}_i)$ is the expected test duration at stress level \mathbf{S}_i , C_s is the manufacturing cost per specimen, C_l is the labor cost per unit test duration, and i is the number of testing stress levels.

During VEDO, we do not know $t_{test}(\mathbf{S}_i)$. If we have prior information on the ALT parameters λ and $\boldsymbol{\beta}$ (e.g., when the tests are designed adaptively), $t_{test}(\mathbf{S}_i)$ can be estimated based on the prior information. When we do not have any prior information about the ALT model, the simulation model $L = g_{life}(G(t), \boldsymbol{\theta}) = g_{life}(g(\mathbf{S}, \mathbf{X}, \mathbf{Y}(t), t), \boldsymbol{\theta})$ mentioned in Section 2.2 is used to estimate $t_{test}(\mathbf{S}_i)$. Based on the estimation of $t_{test}(\mathbf{S}_i)$, C_{test} can be estimated for a given testing plan.

(c) Solving the VEDO model based on (a) and (b)

Based on the above discussed methods, both the objective function and constraint function can be estimated for a given testing plan. Even if the objective and constraint can be estimated efficiently, we cannot directly use them for VEDO since the design variables include both discrete and continuous variables. In this situation, the VEDO model cannot be solved using traditional gradient-based optimization algorithms. The design optimization problem with both discrete and continuous variables can be solved using genetic algorithm (GA) or simulated annealing method. Since the optimization algorithm may need to evaluate the objective and constraints hundreds of times, it is still computationally expensive. In order to overcome this challenge, we build a surrogate model for the objective function. Considering that the objective function is a function of both continuous and integer variables, the Kriging method is used in this paper to construct the surrogate model. Swiler et al. [47] have shown that a Kriging model with special correlation functions generally performs better than other types of surrogate models for problems with both continuous and integer variables. In this paper, we relax the discrete design variables as continuous. A brief review of the Kriging surrogate modeling method is given as below.

The Kriging model of a black box function $g(\mathbf{d})$ is given by [48]

$$\hat{g}(\mathbf{d}) = \mathbf{h}(\mathbf{d})^T \boldsymbol{\gamma} + \varepsilon(\mathbf{d}) \quad (41)$$

where \mathbf{d} is the vector of input variables, $\boldsymbol{\gamma} = [\gamma_1, \gamma_2, \dots, \gamma_p]^T$ is a vector

of unknown coefficients, $\mathbf{h}(\mathbf{d}) = [h_1(\mathbf{d}), h_2(\mathbf{d}), \dots, h_p(\mathbf{d})]^T$ is a vector of regression functions, $\mathbf{h}(\mathbf{d})^T \boldsymbol{\gamma}$ is the trend of prediction, and $\varepsilon(\mathbf{d})$ is usually assumed to be a Gaussian process with zero mean and covariance $\text{Cov}[\varepsilon(\mathbf{d}_i), \varepsilon(\mathbf{d}_j)]$ given by

$$\text{Cov}[\varepsilon(\mathbf{d}_i), \varepsilon(\mathbf{d}_j)] = \sigma_\varepsilon^2 R(\mathbf{d}_i, \mathbf{d}_j) \quad (42)$$

in which σ_ε^2 is the process variance and $R(\cdot, \cdot)$ is the correlation function.

The coefficients $\boldsymbol{\gamma}$ can be estimated by using least squares, maximum likelihood, or Bayesian updating [48,49]. For a new point \mathbf{d} , the mean of the prediction is given by

$$\hat{g}(\mathbf{d}) = \mathbf{h}(\mathbf{d})^T \boldsymbol{\gamma} + \mathbf{r}(\mathbf{d})^T \mathbf{R}^{-1}(\mathbf{g} - \mathbf{H}\boldsymbol{\gamma}) \quad (43)$$

where \mathbf{g} and \mathbf{H} are $g(\mathbf{d})$ and $\mathbf{h}(\mathbf{d})^T$ evaluated at the training points, \mathbf{R} is the correlation matrix of training points with element $R(\mathbf{d}_i, \mathbf{d}_j)$, and

$$\mathbf{r}(\mathbf{d}) = [R(\mathbf{d}, \mathbf{d}_1), R(\mathbf{d}, \mathbf{d}_2), \dots, R(\mathbf{d}, \mathbf{d}_{n_t})] \quad (44)$$

in which $\mathbf{r}(\mathbf{d})$ is a vector containing the covariance between \mathbf{d} and each of the training points, and n_t is the number of training points.

The mean square error (MSE) of the prediction is given by [50]

$$\text{MSE}(\mathbf{d}) = \sigma_\varepsilon^2 \{1 - \mathbf{r}(\mathbf{d})^T \mathbf{R}^{-1} \mathbf{r}(\mathbf{d}) + [\mathbf{H}^T \mathbf{R}^{-1} \mathbf{r}(\mathbf{d}) - \mathbf{h}(\mathbf{d})]^T (\mathbf{H}^T \mathbf{R}^{-1} \mathbf{H})^{-1} [\mathbf{H}^T \mathbf{R}^{-1} \mathbf{r}(\mathbf{d}) - \mathbf{h}(\mathbf{d})]\} \quad (45)$$

where σ_ε^2 is also a parameter estimated using the least square method. More details about Kriging can be found in [48,49].

In the VEDO problem, the design variables are $\mathbf{d} = [\mathbf{S}_t, \mathbf{N}_t]$ and the black box function is $f_{obj}(\mathbf{S}_t, \mathbf{N}_t)$. In the surrogate modeling, the zero order trend function and the Gaussian correlation function are used. Once the surrogate model of $f_{obj}(\mathbf{S}_t, \mathbf{N}_t)$ is available, the VEDO model given in Eq. (20) can be solved using either the genetic algorithm (GA) [51,52] or an MCS-based method. Note that the above reviewed Kriging surrogate modeling method may not only be used for the modeling of the objective function, but also be applied to the time-dependent prediction model if the prediction model is computationally expensive.

In the next subsection, we will summarize the overall implementation procedure of the proposed VEDO method.

3.3. Implementation procedure

Fig. 3 shows the flowchart of the proposed VEDO method. The detailed implementation procedure is explained as follows. There are basically three modules.

Module 1: Time-dependent prediction model

- Obtain the prediction of life distribution at the nominal stress level $S_0 (F_p)$.
- Obtain the life distributions at different stress levels and determine the distribution type of the life distribution model and the stress-parameter function type (Sec. 3.1.1).
- Estimate the values of λ and β based on the predicted life distributions and thus provide prior distributions for λ and β .

Module 2: Estimation of the objective function

- Generate samples of λ and β based on the prior distributions.
- For each sample of λ and β

- Compute the Fisher information matrix $[\mathbf{I}(\lambda, \beta)]_{\lambda=\lambda^*, \beta=\beta^*}^{-1}$.
- Obtain the approximated posterior distribution of λ and β .
- Compute the standard deviation of the validation metric.

- Compute the objective function for a given testing plan.

Module 3: Optimization

- Build a surrogate model for the objective function using Kriging.
- Solve the optimization problem using the Genetic algorithm (GA).

The proposed VEDO method introduces an ALT-based validation method to reduce the testing time and cost in the validation of time-dependent prediction model; based on that, a novel VEDO model is developed to obtain the optimal testing inputs to reduce the uncertainty in the validation result; finally, a couple of techniques (Kriging and GA) are utilized to efficiently solve the VEDO model. The proposed validation methodology makes the validation of time-dependent pre-

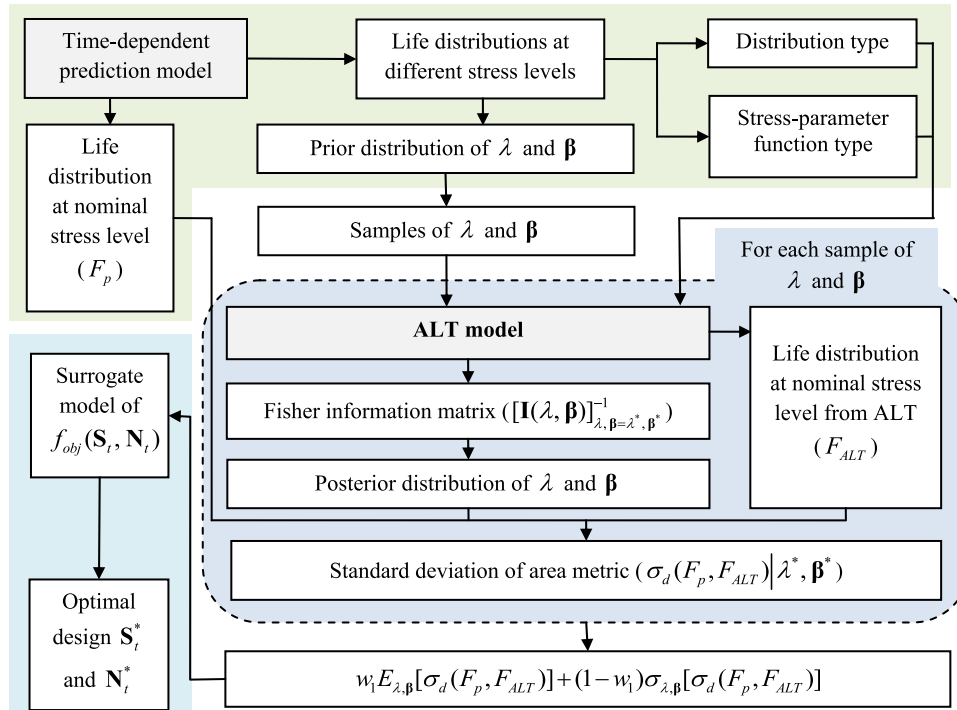


Fig. 3. Flowchart of the proposed VEDO method.

diction models affordable and thus accelerates the product development process.

4. Numerical example

4.1. Problem statement

A composite helicopter rotor hub yoke is used in this section to demonstrate the proposed methodology [53]. The component is subjected to fatigue due to centrifugal and bending loads during flights, leading to delamination between individual plies at the thick-to-taper transition resin pocket. The thick-taper-thin geometry is tailored to give the proper flapping flexure as shown in Fig. 4. The quantity of interest is the delamination fatigue life of the component. Fatigue life distribution at the nominal stress level is predicted using a simulation model, which needs to be validated.

The virtual crack closure technique (VCCT) is used to assess fatigue delamination. We first construct a finite element analysis (FEA) model for a half symmetric section of the hub geometry (as shown in Fig. 5(a)). In the thick-to-taper transition region where delamination occurs, the finite element mesh was refined (as shown in Fig. 5(b)) around crack tip and modeled with 8-node quadrilateral elements. The minimum element size in this region is 0.003. Such small size ensures that computation of the strain energy release rate function (G) does not suffer from numerical inconsistencies. We then assess the strain energy release rate associated with the fatigue delamination of the component under given geometry, material property, and loads using the VCCT theory. Fig. 5(c) shows the stress contour of the analysis result. Fig. 6 gives the results of strain energy release rate.

Based on the energy release rate, the fatigue life can be obtained based on the following function constructed from experimental data [54]:

$$G_{crit} = 448.56 - 58.571 \log(N) \quad (46)$$

where G_{crit} is the critical energy release rate and N is the number of load cycles.

In this problem, the time-dependent prediction model is the energy release rate prediction model from ABAQUS plus the life prediction model given in Eq. (46). The uncertain parameters in the life prediction model are given in Table 1. Due to the uncertainty in these parameters, the service life of the rotorcraft hub component is also uncertain. Note that model parameter uncertainty regarding to Eq. (46) is not considered in this example. The predicted life distribution at the nominal load level (20 kips) needs to be validated.

4.2. Validation experiment design optimization

The following cost information is assumed in the validation testing design optimization. The manufacturing cost of each test specimen is assumed to be \$150, the labor fee per test cycle is $\$1.5 \times 10^{-2}$, and the total budget is $\$1.5 \times 10^4$. The total number of test specimens is 70. The range of testing load levels is [22.5, 27.5] kips. The observation error of test cycles is $\varepsilon_{obs} \sim N(0, 10^2)$ cycles.

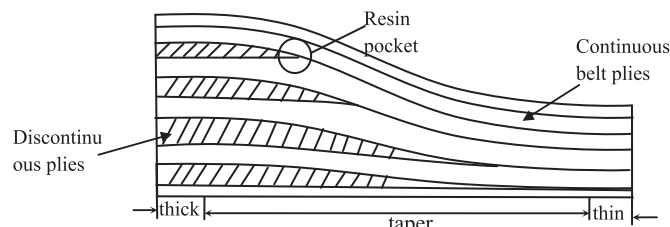


Fig. 4. Half of the symmetric section of tapered composite in a helicopter rotor hub.

4.2.1. Extract information from the simulation model

Before applying the proposed method, we need to gather some information from the simulation model. (1) The entire method is based on the distribution assumption of the fatigue life at a particular stress level, such as a Lognormal or Weibull distribution. This assumption is first verified using a probability paper analysis. In the rotorcraft hub component, Fig. 7 shows that the Weibull distribution can well describe the fatigue life predicted by the simulation model. (2) The simulation model also provides the load-life function, and prior information about the parameters. Fig. 8 shows the regression of the load-life function using least squares. Based on the regression, we describe the load-life function as a quadratic function:

$$\log(\alpha) = \beta_0 + \beta_1 S + \beta_2 S^2 \quad (47)$$

where S is the load level, α is the scale parameter at load level S , β_0 , β_1 and β_2 are estimated using least squares, based on data generated by the simulation model.

From regression, we get $\beta_0 = -2.4525$, $\beta_1 = 1.2008$, $\beta_2 = -0.0296$. The average shape parameter over different stress levels is obtained as $\lambda = 1.5$. Since we do not have any information about the ALT model, we provide prior distributions for the ALT model parameters based on the above information. The prior distributions are given by: $\lambda \sim U(0.8, 2.2)$, $\beta_0 \sim U(-3.5, -1.5)$, $\beta_1 \sim U(0.5, 2.0)$, $\beta_2 \sim U(-0.05, -0.01)$, where $U(\cdot, \cdot)$ stands for uniform distribution.

Before solving the optimization model, we also need to construct a Kriging surrogate model of the expected testing time (number of cycles to failure) using 50 training points, based on estimates of the expected testing time obtained from the simulation model. (If we have some information about the ALT model, we can also estimate the expected life using the ALT model.) Fig. 9 shows the relationship between load level and expected testing time.

4.2.2. VEDO

Based on the information provided above, the VEDO model is defined as

$$\begin{aligned} \min_{S_i, N_i} \quad & f_{obj}(S_i, N_i) = w_1 E_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})] + (1 - w_1) \sigma_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})] \\ \text{s. t.} \quad & C_s \sum_{i=1}^3 N_i + C_l \sum_{i=1}^3 N_i t_{test}(S_i) \leq \$1.5 \times 10^4 \\ & 12 \leq N_i \leq 70, \quad \forall i = 1, 2, 3 \\ & 22.5 \leq S_i \leq 27.5, \quad \forall i = 1, 2, 3 \\ & \sum_{i=1}^3 N_i \leq 70 \end{aligned} \quad (48)$$

Here $w_1 \in [0, 1]$ is the weight, $E_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})]$ and $\sigma_{\lambda, \beta}[\sigma_d(F_p, F_{ALT})]$ are the expectation and standard deviation of $\sigma_d(F_p, F_{ALT})$ by considering the uncertainty in the prior distribution of λ and β . The total cost of the tests (including manufacturing cost of test specimens and labor fee) should be less than the budget $\$1.5 \times 10^4$. The design of tests is also constrained by the total number of tests and least number of tests at each test level. The testing load levels S_i is within the range [22.5, 27.5] kips.

We perform validation experiment design assuming that tests will be conducted at three stress levels ($n_s = 3$), and set $w_1 = 0.5$, $C_s = 150$, $C_l = 1.5 \times 10^{-2}$, and $C_{budget} = 1.5 \times 10^4$. To solve the optimization model, we construct a surrogate model using the method presented in Section 3.2 using 300 training points. We then solve the optimization model using the method discussed in Section 3.3. Table 2 gives the optimal design obtained from VEDO. It shows that the optimal test load levels are: 25.03 kips, 22.53 kips, and 27.49 kips; and the optimal number of tests are 13, 22, and 13 respectively; the total cost corresponding to the optimal test plan is $\$1.44 \times 10^4$.

In order to verify the effectiveness of the results obtained from VEDO, assume that the true ALT model is as follows: the distribution type is Weibull, the shape parameter is $\lambda = 1.5$, and the scale parameter is given by

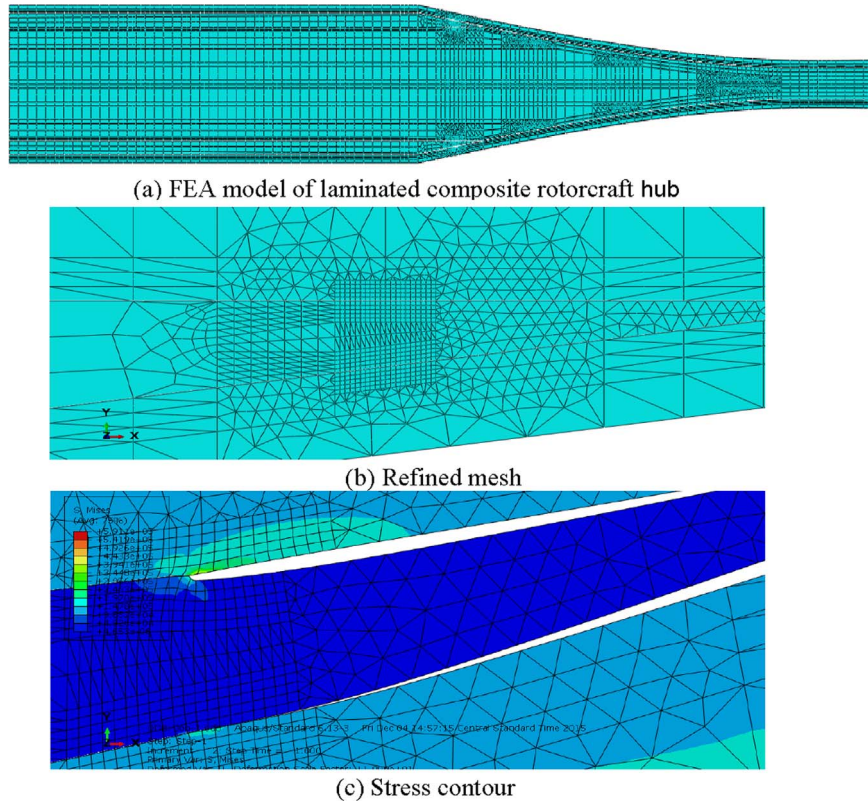


Fig. 5. Simulation of a laminated composite rotorcraft hub component.

$$\log(\alpha) = -2.45 + 1.20S + 0.03S^2 \quad (49)$$

Since the true model is assumed to be close to the simulation model, a good ALT plan is expected to have a smaller mean value of area metric between the empirical distribution and predicted distribution. We generate observation data from the assumed true model using the optimal ALT plan obtained from the proposed VEDO model and compute the area metric. The distribution of the area metric is shown in Fig. 10. At the same time, we generate observation data using three arbitrarily chosen ALT plans and compute the corresponding area metrics. In Case 1 and Case 2, we perform tests at a single stress level (i.e. all three stress levels are the same) at lower and upper bounds of the load range respectively. The number of tests are 40 for 22.5 kips, and 71 for 27.5 kips. In Case 3, the tests are conducted at 22.5 kips, 23 kips and 24 kips, with number of tests 22, 11, and 8. Fig. 10 and Table 3 give the validation result comparison between the optimal design and three arbitrarily chosen designs. The results show that the optimal design obtained from the proposed method not only obtains a better validation result in terms of mean, but also significantly reduces

Table 1

Uncertainty parameters of the rotorcraft hub example.

| Input parameter | Mean value | Standard deviation | Distribution type |
|-------------------------------|------------|--------------------|-------------------|
| Modulus (E_{11} , Msi) | 6.7 | 0.1 | Normal |
| Applied load (P , kips) | S | 0.5 | Normal |
| Applied angle (θ , °) | 10 | 0.5 | Normal |

Note: S is stress level. Nominal stress level is 20 kips, and the range of ALT stress level is [22.5–27.5] kips.

the uncertainty in the validation results, when compared with the other three cases. We can also see that the cost for optimal design is $\$1.44 \times 10^4$, while the costs for other three designs are $\$1.45 \times 10^4$, $\$1.45 \times 10^4$, and $\$1.44 \times 10^4$. The cost of the optimal design is slightly less than the other three cases.

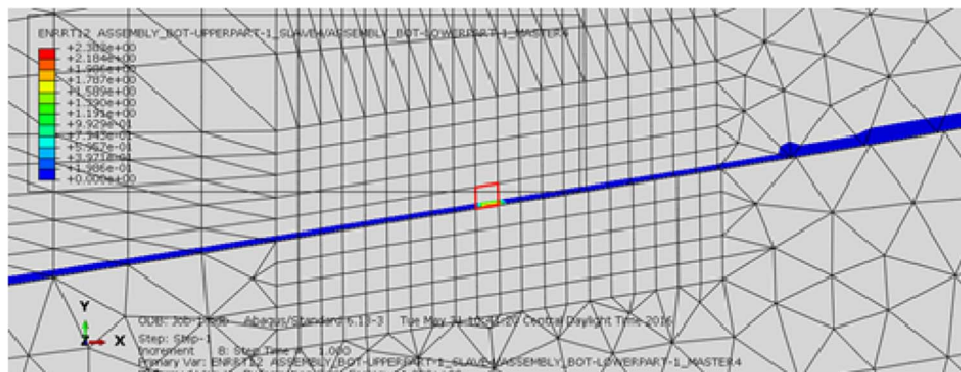


Fig. 6. Result of energy release rate obtained from ABAQUS simulation.

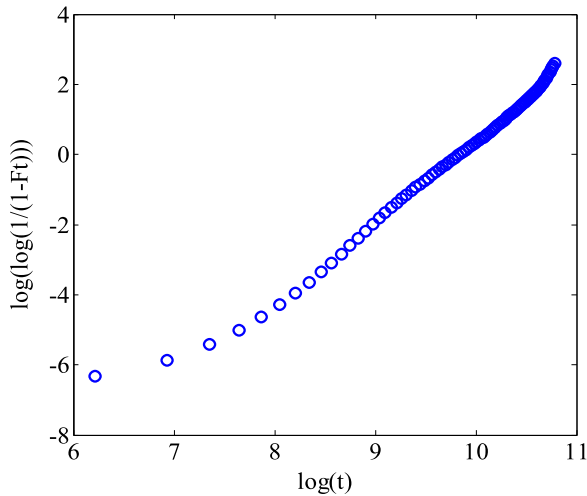


Fig. 7. Weibull probability plot.

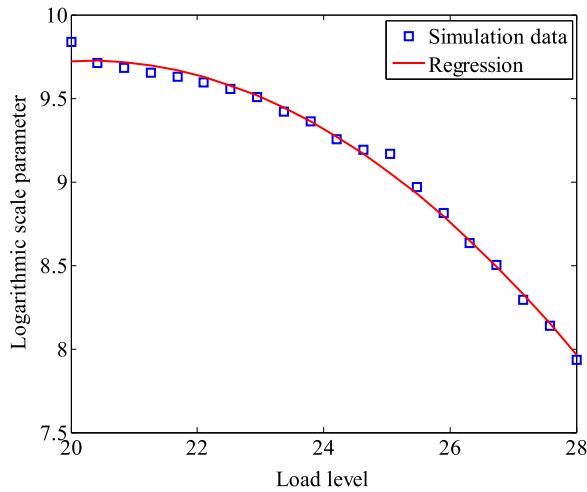


Fig. 8. Load-life relationship.

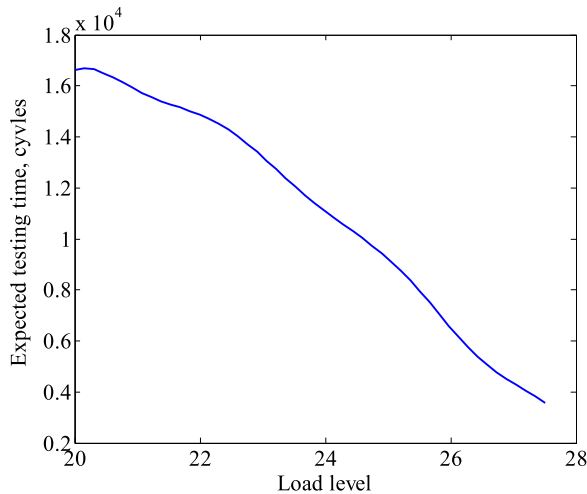


Fig. 9. Expected testing time at different stress levels.

Table 2
Results of VEDO.

| Variable | N_1 | N_2 | N_3 | S_1 | S_2 | S_3 | $C_{\cos t}$ |
|---------------|-------|-------|-------|-------|-------|-------|----------------------|
| Optimal value | 13 | 22 | 13 | 25.03 | 22.53 | 27.49 | $\$1.44 \times 10^4$ |

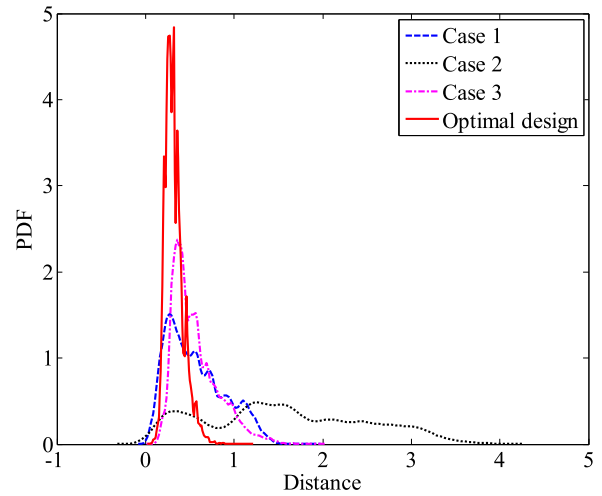


Fig. 10. Comparison of validation results from different designs.

Table 3
Comparison of VEDO result with arbitrary designs.

| Variable | N_1 | N_2 | N_3 | S_1 | S_2 | S_3 | E_d | σ_d | $C_{\cos t}$ |
|---------------|-------|-------|-------|-------|-------|-------|-------|------------|----------------------|
| Case 1 | 40 | 0 | 0 | 22.5 | 0 | 0 | 0.57 | 0.25 | $\$1.45 \times 10^4$ |
| Case 2 | 71 | 0 | 0 | 27.5 | 0 | 0 | 1.56 | 0.90 | $\$1.45 \times 10^4$ |
| Case 3 | 22 | 11 | 8 | 22.5 | 23 | 24 | 0.56 | 0.26 | $\$1.44 \times 10^4$ |
| Optimal value | 13 | 22 | 13 | 25.03 | 22.53 | 27.49 | 0.32 | 0.11 | $\$1.44 \times 10^4$ |

where “ E_d ” is the mean of the validation metric and “ σ_d ” is the standard deviation of validation metric.

5. Conclusion

Life testing data are required in the validation of time-dependent prediction models. However, products are usually designed to perform satisfactorily for a very long time. In this situation, collecting life testing data for validation under nominal stress levels will be very time-consuming and costly. In order to overcome this challenge, this paper proposes an accelerated life testing (ALT) based validation method which reduces the time and cost for the validation of time-dependent prediction models. Based on the ALT-based validation method, a validation experiment design optimization (VEDO) model is developed. By solving the VEDO model, the optimal testing stress levels and the number of tests at each stress level can be obtained. A composite rotorcraft hub example demonstrates that the proposed VEDO model can increase the confidence in the validation result by determining the optimal testing plan.

In this paper, the Weibull distribution and the stress-parameter relationship obtained from the simulation model were used as prior information for the experimental design. These assumptions affect the VEDO results. Relaxing these assumptions will be investigated in our future work. Extending the proposed VEDO method to systems with complicated configurations and multiple components is also a future need. In addition, considering censored data within VEDO requires investigation, since some tests may be terminated before failure due to time constraints.

Acknowledgement

This study was supported by funds from the National Science Foundation (Grant No. 1404823, CDSE Program). The support is gratefully acknowledged.

References

- [1] Ferson S, Oberkampf WL. Validation of imprecise probability models. *Int J Reliab*

- Saf 2009;3(1–3):3–22.
- [2] Kleijnen JP. Verification and validation of simulation models. *Eur J Oper Res* 1995;82(1):145–62.
 - [3] Rebba R, Mahadevan S, Huang S. Validation and error estimation of computational models. *Reliab Eng Syst Saf* 2006;91(10):1390–7.
 - [4] Rebba R, Mahadevan S. Model predictive capability assessment under uncertainty. *AIAA J* 2006;44(10):2376–84.
 - [5] Ferson S, Oberkampf WL, Ginzburg L. Model validation and predictive capability for the thermal challenge problem. *Comput Methods Appl Mech Eng* 2008;197(29):2408–30.
 - [6] Ling Y, Mahadevan S. Quantitative model validation techniques: new insights. *Reliab Eng Syst Saf* 2013;111:217–31.
 - [7] Rebba R, Mahadevan S. Computational methods for model reliability assessment. *Reliab Eng Syst Saf* 2008;93(8):1197–207.
 - [8] Moreno PJ, Ho PP, Vasconcelos N. "A Kullback-Leibler divergence based kernel for SVM classification in multimedia applications," In: *Proceedings of the Advances in neural information processing systems*, p. None.
 - [9] Zhang W, Liu Y. Investigation of incremental fatigue crack growth mechanisms using in situ SEM testing. *Int J Fatigue* 2012;42:14–23.
 - [10] Hu Z, Du X. Reliability analysis for hydrokinetic turbine blades. *Renew Energy* 2012;48:251–62.
 - [11] Singh A, Mourelatos Z, Nikolaidis E. time-dependent reliability of random Dynamic systems using time-series modeling and importance sampling. *SAE Int J Mater Manuf* 2011;4(1):929–46.
 - [12] Liu Y, Mahadevan S. Threshold stress intensity factor and crack growth rate prediction under mixed-mode loading. *Eng Fract Mech* 2007;74(3):332–45.
 - [13] Hu Z, Du X. Time-dependent reliability analysis with joint upcrossing rates. *Struct Multidiscip Optim* 2013;48(5):893–907.
 - [14] Hu Z, Du X. A sampling approach to extreme value distribution for time-dependent reliability analysis. *J Mech Des* 2013;135(7):071003.
 - [15] Zhang R, Mahadevan S. Bayesian methodology for reliability model acceptance. *Reliab Eng Syst Saf* 2003;80(1):95–103.
 - [16] Meeker WQ. Bayesian methods for planning Accelerated life tests. *Technometrics: A J Stat Phys, Chem Eng Sci* 2006;1:49–60.
 - [17] Jiang X, Mahadevan S. Bayesian risk-based decision method for model validation under uncertainty. *Reliab Eng Syst Saf* 2007;92(6):707–18.
 - [18] Huan X, Marzouk YM. Simulation-based optimal Bayesian experimental design for nonlinear systems. *J Comput Phys* 2013;232(1):288–317.
 - [19] Huan X, Marzouk Y. Gradient-based stochastic optimization methods in Bayesian experimental design. *Int J Uncertain Quantif* 2014;4:6.
 - [20] Hu Z, Mahadevan S. Accelerated life testing (ALT) design based on Computational reliability analysis. *Qual Reliab Eng Int* 2015.
 - [21] Hu Z, Ao D, Mahadevan S. Calibration experimental design considering field response and model uncertainty. *Comput Methods Appl Mech Eng* 2017.
 - [22] Jiang X, Mahadevan S. Bayesian cross-entropy methodology for optimal design of validation experiments. *Meas Sci Technol* 2006;17(7):1895.
 - [23] Oden JT, Prudencio EE, Bauman PT. Virtual model validation of complex multi-scale systems: applications to nonlinear elastostatics. *Comput Methods Appl Mech Eng* 2013;266:162–84.
 - [24] Mahadevan S, Rebba R. Validation of reliability computational models using Bayes networks. *Reliab Eng Syst Saf* 2005;87(2):223–32.
 - [25] Rebba R, Mahadevan S. Validation of models with multivariate output. *Reliab Eng Syst Saf* 2006;91(8):861–71.
 - [26] De Maesschalck R, Jouan-Rimbaud D, Massart DL. The mahalanobis distance. *Chemom Intell Lab Syst* 2000;50(1):1–18.
 - [27] Halder A, Bhattacharya R. "Model validation: A probabilistic formulation," Decision and Control and European Control Conference (CDC-ECC) In: *Proceedings of the 50th IEEE Conference on, IEEE*, pp. 1692–1697; 2011.
 - [28] Kocabas V, Dragicevic S. Assessing cellular automata model behaviour using a sensitivity analysis approach. *Comput, Environ Urban Syst* 2006;30(6):921–53.
 - [29] Xiong Y, Chen W, Tsui K-L, Apley DW. A better understanding of model updating strategies in validating engineering models. *Comput Methods Appl Mech Eng* 2009;198(15):1327–37.
 - [30] Li W, Chen W, Jiang Z, Lu Z, Liu Y. New validation metrics for models with multiple correlated responses. *Reliab Eng Syst Saf* 2014;127:1–11.
 - [31] Mullins J, Ling Y, Mahadevan S, Sun L, Strachan A. Separation of aleatory and epistemic uncertainty in probabilistic model validation. *Reliab Eng Syst Saf* 2016;147:49–59.
 - [32] Pan R, Yang T, Seo K. Planning Constant-Stress Accelerated Life Tests for Acceleration Model Selection. *Reliab, IEEE Trans on* 2015;64(4):1356–66.
 - [33] Elsayed EA, Zhang H. Design of PH-based accelerated life testing plans under multiple-stress-type. *Reliab Eng Syst Saf* 2007;92(3):286–92.
 - [34] Pascual FG, Montepiedra G. Model-robust test plans with applications in accelerated life testing. *Technometrics* 2003;45(1):47–57.
 - [35] Pascual FG. Accelerated life test plans Robust to Misspecification of the stress–life relationship. *Technometrics* 2012.
 - [36] Meeker WQ. A comparison of accelerated life test plans for Weibull and lognormal distributions and type I censoring. *Technometrics* 1984;26(2):157–71.
 - [37] Nelson W, Meeker WQ. Theory for optimum accelerated censored life tests for Weibull and extreme value distributions. *Technometrics* 1978;20(2):171–7.
 - [38] Nelson WB. Accelerated testing: statistical models, test plans, and data analysis. John Wiley & Sons; 2009.
 - [39] Nelson W. Accelerated life testing-step-stress models and data analyses. *IEEE Trans Reliab* 1980;29(2):103–8.
 - [40] Croarkin C, Tobias P, Zey C. *Engineering statistics handbook*, NIST iTL; 2002.
 - [41] Liao H, Elsayed EA. Reliability inference for field conditions from accelerated degradation testing. *Nav Res Logist (NRL)* 2006;53(6):576–87.
 - [42] Liao H, Li Z. Multiobjective design of equivalent accelerated life testing plans. *Int J Reliab, Qual Saf Eng* 2008;15(06):515–38.
 - [43] Berger JO. Statistical decision theory and Bayesian analysis. Springer Science & Business Media; 2013.
 - [44] Long Q, Scavino M, Tempone R, Wang S. Fast estimation of expected information gains for Bayesian experimental designs based on Laplace approximations. *Comput Methods Appl Mech Eng* 2013;259:24–39.
 - [45] Fard N, Li C. Optimal simple step stress accelerated life test design for reliability prediction. *J Stat Plan Inference* 2009;139(5):1799–808.
 - [46] Stroud AH, Secrest D. Gaussian quadrature formulas. Englewood Cliffs, NJ: Prentice-Hall; 1966.
 - [47] Swiler LP, Hough PD, Qian P, Xu X, Storlie C, Lee H. Surrogate models for mixed discrete-continuous variables [Constraint Programming and Decision Making]. Springer; 2014. p. 181–202.
 - [48] Rasmussen CE. Gaussian processes for machine learning. The MIT Press; 2006, [ISBN 0-262-18253-X].
 - [49] Santner TJ, Williams BJ, Notz W. The design and analysis of computer experiments. Springer; 2003.
 - [50] Lophaven SN, Nielsen HB, Sondergaard J. "DACE-A Matlab Kriging toolbox, version 2.0."; 2002.
 - [51] Deb K, Pratap A, Agarwal S, Meyarivan T. A fast and elitist multiobjective genetic algorithm: NSGA-II. *Evolut Comput* 2002;6(2):182–97.
 - [52] Lin CD, Anderson-Cook CM, Hamada MS, Moore LM, Sitter RR. Using genetic algorithms to design experiments: a review. *Qual Reliab Eng Int* 2015;31(2):155–67.
 - [53] Mahadevan S, Zhang R. Fatigue test planning using reliability and confidence simulation. *Int J Mater Prod Technol* 2001;16(4–5):317–32.
 - [54] Mahadevan S, Dey A, Tryon R, Wang Y, Rousseau C. Reliability analysis of rotorcraft composite structures. *J Aerosp Eng* 2001;14(4):140–6.