JOURNAL OF QUALITY TECHNOLOGY
2018, VOL. 50, NO. 3, 279–289
https://doi.org/10.1080/00224065.2018.1474688

RESEARCH PAPER

Surrogate model–based optimal feed-forward control for dimensional-variation reduction in composite parts' assembly processes

Xiaowei Yue and Jianjun Shi

H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, Georgia

ABSTRACT
Dimension control and variation reduction are vital for composite parts' assembly processes. Due to the nonlinear properties of composites, physics-based models cannot accurately and efficiently approximate the assembly processes. In addition, conventional robust parameter design (RPD) and statistical process control (SPC) cannot actively compensate for dimensional errors or prevent defects. This article proposes a surrogate model–based optimal feed-forward control strategy for dimensional-variation reduction and defect prevention in the assembly of composite parts. The objective is accomplished by (i) developing a grouped Latin hypercube sampling approach tailored to the problem; (ii) adopting a universal Kriging model for dimensional prediction and then embedding the model into an optimal feed-forward control algorithm; and (iii) conducting a multiobjective optimization to determine the control actions. A case study reveals that the developed methodology can effectively reduce the mean and standard deviation of dimensional deviations for the assembly of composite parts.

KEYWORDS
composites assembly processes; feed-forward control; Latin hypercube sampling; surrogate model; variation reduction

1. Introduction

Composite parts have been widely used in various applications due to their superior properties of strength, stiffness, weight, thermal conductivity, temperature-dependent behavior, and fatigue life (Jones 1998). In general, multiple composite parts are assembled sequentially at multiple assembly stations to form a final functional product. Because of fabrication errors, fixture errors, and positioning errors, there always exist some dimensional variabilities among parts and subassemblies along the assembly processes. Such dimensional variation has significant impacts on the final product quality, reliability, functionality, and cost. Therefore, dimensional-variation reduction is an important yet challenging task in the process improvement and quality control of composite parts' assembly processes.

To achieve the objective of variation reduction in a multistage assembly process (MAP), three kinds of approaches are commonly used in practice. The first, robust parameter design (RPD), is used to design the product and process so that they are robust in processing disturbances to reduce variability (Wu and Hamada 2011). However, RPD cannot anticipate all potential disturbances during assembly operations; thus the variability of assembly is unavoidable. The second approach, statistical process control (SPC), has been developed to monitor the process and detect out-of-control conditions to keep the process in control (Shi 2006; Montgomery 2007). However, SPC detects process changes and then tries to find the root causes of the variability, but by that time products with larger variability have already been fabricated. The third approach, in-process quality improvement (IPQI), emphasizes process monitoring, root cause diagnosis, and active control by analyzing in-process sensing with the help of engineering domain knowledge and models (Shi 2013). IPQI aims to reduce variation and thereby increase defect prevention rather than keeping the process in-control to achieve defect detection. In this article, a surrogate model–based feed-forward control strategy is developed to reduce the in-process dimensional variation for composite parts' assembly processes.

In-process active control in multistage assembly processes with a feed-forward control strategy has been studied (Izquierdo et al. 2007; Zhong, Liu, and Shi 2010). Due to lacking autocorrelation properties of quality features in a multistage assembly process, feedback control is not effective and may cause overcontrol actions and thus even increase process variation (Zhong, Liu, and Shi 2010). In the current study, in-line dimensional measurement sensors, programmable tools, and a mathematical surrogate model are essential to realize the

CONTACT Jianjun Shi jianjun.shi@isye.gatech.edu H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, 755 Ferst Drive NW, Atlanta, GA 30332,USA.
Color versions of one or more figures in this article can be found online at www.tandfonline.com/ujqt.

Table 1. Three types of models to describe a multistage assembly process (MAP).

Model Type	Definition	Advantage	Disadvantage
Analytical physics-based models	Closed-form partial differential equations, state-space equations for a MAP	Provides process details and insights, wide applicability	Difficulty in modeling complex systems, time consuming, and error prone
Numerical-simulation models	Commercial computer software such as 3DCS Analyst or Vis VSA to simulate a MAP	High accuracy, can handle complex systems	Time consuming; must manually configure (for validation instead of optimization)
Surrogate models	Designed computer experiments offer training data; a surrogate model is used as a predictor	Simple, flexible, computationally efficient, low cost, extendable in the configuration setting	If the training points are relatively "close" to each other, the correlation matrix can become singular

feed-forward control strategy to improve dimensional quality in assembly processes.

Prior to designing an in-process active compensation control algorithm for variation reduction, modeling should be investigated to link the dimensional quality with various process parameters in multistage composite parts' assembly processes. Usually three types of models are used to analyze multistage assembly processes. The classifications, definitions, advantages, and disadvantages of those three types of models are summarized in Table 1. Analytical models developed from physical principles make full use of closed-form partial differential equations or state-space equations to encode all the input, process, and output variables in a physical system (Jin and Shi 1999; Mantripragada and Whotney 1999; Shi 2006; Izquierdo et al. 2007; Zhang and Shi 2016a, 2016b). These physics-based models provide process details and insights, which have wide applicability. However, it is time consuming, error prone, and difficult to model complex nonlinear systems with stochastic disturbances. The extended method of influence coefficients (EMIC) was proposed to predict assembly deviation for compliant composite parts and is efficient and effective in both single-station and multistation assembly processes (Zhang and Shi 2016a, 2016b). However, its linear approximation has the potential to introduce large errors for composite components with very complex structures and nonlinear properties. Numerical-simulation models can be developed by using commercial computer software such as 3DCS Analyst, Vis VSA, or ANSYS (Loose, Chen, and Zhou 2009; Zhou, Qian, and Zhou 2012; Wen et al. 2018). The software can achieve high accuracy for complex systems, yet process parameters are required before manual configuration adjustment. Thus, this type of model can be used only for variation analysis and design validation; it cannot be used for in-process active compensation control. Surrogate models are built based on training data from designed computer experiments (Loose, Chen, and Zhou 2009; Zhou, Qian, and Zhou 2012; Yue et al. 2018). Surrogate modeling has gained more attention in recent years because of its useful characteristics of flexibility, computational efficiency, and robustness to configuration changes, which are all suitable for developing in-line optimization and control algorithms.

Another driving motivation to use a surrogate model in composite parts' assembly processes is the nonlinear property of composite materials. Physics-based models like state-space equations and finite-element models are widely used to describe the multistage assembly process (MAP) for metal compliant parts (Liu and Hu 1997; Jin and Shi 1999; Mantripragada and Whotney 1999; Shi 2006; Izquierdo et al. 2007; Zhang and Shi 2016a, 2016b). In these models linear approximation can simplify the systems without a significant loss of accuracy. However, composite materials are usually anisotropic and nonlinear, which means they have different mechanical or physical properties such as stiffness, conductivity, and tensile strength in different directions. If a linear approximation of a physics-based model is adopted to analyze the nonlinear assembly processes, a large modeling error can be expected. A surrogate model is a data-driven model, and it can directly describe anisotropy and nonlinear properties of composite materials from training data sets as well as achieve higher prediction accuracy.

The procedure to develop a surrogate model–based feed-forward control strategy is shown in Figure 1. Design of experiment (DOE) ensures efficient acquisition of training data. Three groups of variables describing locating fixture deviations, holding fixture deviations, and part manufacturing deviations are considered. A grouped Latin hypercube sampling (grouped LHS) is tailor-made for composite parts' assembly processes to obtain effective design configurations. Based on the grouped LHS plan and the finite-elements method (FEM), computer experiments are conducted using the designated settings of experimental configurations and provided corresponding output variables. A surrogate model is built by using the universal Kriging method, and model validation is conducted based on training and testing data sets. Finally, a multiobjective optimization is conducted to obtain the in-process feed-forward compensation control

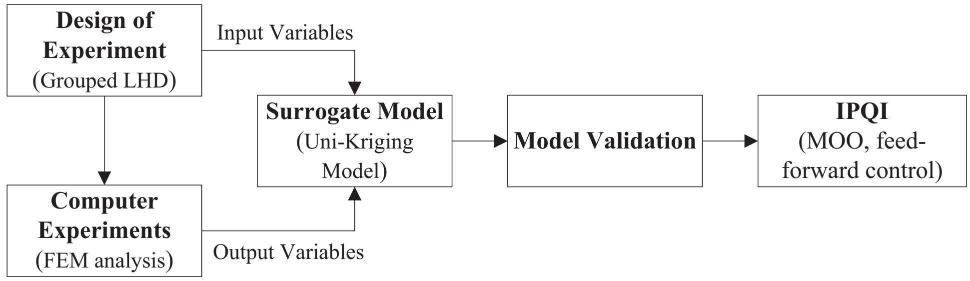


Figure 1. An overview of the proposed surrogate model–based feed-forward control.

strategies necessary to achieve dimensional-variation reduction.

The contributions of this article are threefold. (1) We propose a grouped Latin hypercube sampling (LHS) that involves two appealing properties: (i) it considers dependence among variables within each group by rank statistics and (ii) it prevents collinearity among variables between different groups. (2) We propose the framework of a surrogate model–based in-process control. Surrogate models are playing an increasingly important role in contemporary engineering because they can dramatically reduce the computational cost of the solution process. Most of the surrogate models have been applied to the prediction of response surfaces and the optimization of engineering parameters (Forrester, Sobester, and Keane 2008; Koziel and Leifsson 2013; Santner et al. 2003), and few of these articles about surrogate model–based in-process control are discussed. From the control theory perspective, models in the control systems include the state-space model, the dynamic equation model, the Volterra model, and the process-disturbance predictive model (Camacho and Alba 2013; Dorf and Bishop 2011). While few of these articles apply surrogate models to in-process control, the authors show novel applications of surrogate models; that is, to embed surrogate models into control systems. Moreover, the surrogate model–based control system achieves good performance in the composite parts' assembly process. (3) The dimensional-variation reduction in composite parts' assembly processes is very challenging due to the compliant nature and anisotropic characteristics of composite structures. Through a contribution-of-application perspective we develop a surrogate model–based feed-forward control strategy, and it succeeds in helping in-process compensation and dimensional-variation reduction of composite parts' assembly processes.

The remainder of the paper is organized as follows. In Section 2 a grouped LHS will be tailor-made for variables with properties of within-group dependence and between-group independence. Based on the DOE, a universal Kriging surrogate model is developed in Section 3 to represent the dimensional variation in a composite assembly process. In Section 4 the surrogate model is used to develop an in-process optimal feed-forward compensation control algorithm for dimensional-variation reduction. Furthermore, a case study will be implemented to generate a grouped LHS, demonstrate the prediction accuracy of the universal Kriging model, and test the performance of the optimal feed-forward control strategy. Finally, Section 5 discusses the contributions and limitations of the developed methodology.

2. Grouped Latin hypercube sampling

In this section, we describe the problem of the multistage composite parts' assembly process and the associated variable analysis. Then a tailor-made grouped Latin hypercube sampling approach is proposed to implement the design of experiment for variables in the composite assembly process. The proposed grouped LHS can not only consider the dependence among variables within each group but also prevent collinearity among variables between different groups.

2.1. Problem definition and variable analysis

Before the design of experiment, problem description and variable analysis should be conducted. In multistage assembly processes of composite parts, there are three major sources of dimensional deviations: locating fixture deviation (LFD), holding fixture deviation (HFD), and part manufacturing deviation (PMD). LFD denotes the dimensional difference between the fixture's actual position and its designed position at a locating fixture point (LFP). HFD represents the dimensional difference between the holding fixture's actual position and its designed position at a holding fixture point (HFP). PMD describes the dimensional difference between the actual part dimension and its ideal computer-aided design. Usually three locating fixture points are considered with a 3-2-1 constraint principle (Menassa and DeVries 1989).

In the design of experiment, the total input variables of composite part dimensional variations can be divided into three groups. (1) LFD group: in general, locating fixtures remain at the same assembly platform and LFDs, to some extent, are caused by platform deviations. Thus variables within an LFD group are generally dependent. (2) HFD group: engineers tend to adjust holding fixtures simultaneously to achieve a good holding performance, which results in correlations among HFDs. (3) PMD group: PMDs for the same part are also dependent because each composite part is typically fabricated at the same station with the same material and process parameter settings (Ramesh, Mannan, and Poo 2000). However, input variables in different groups are independent since their root causes and characteristics are typically different. Therefore, based on variable analysis and engineering knowledge, different groups of input variables have the following two characteristics: (i) variables within a group are dependent, but the dependent relationships among those variables within the group are unknown; and (ii) variables between different groups are independent. In the next subsection, we propose a grouped LHS for this kind of problem. One note needs to be pointed out: if the variables within a group are independent then Latin hypercube sampling considering dependence in Section 2.2 will become a regular Latin hypercube sampling (Mckay, Beckman, and Conover 1979).

2.2. Grouped Latin hypercube sampling

Latin hypercube sampling (LHS; Mckay, Beckman, and Conover 1979) has been widely used in practice to generate multivariate samples since it gives an estimator with lower variance than simple random sampling and has good projection properties for each variable. Extensions of LHS—such as OA-LHS (Tang 1993), orthogonal-maximin LHS (Joseph and Hung 2008), and sliced LHS (Qian 2012)—avoid collinearity and improve the space-filling and stratification properties. However, LHS and these extensions rely on the independence of the variables in the design of experiment. In many applications one or more variables are dependent on other variables. As an example, variables of an assembly process are dependent within each group as discussed in Section 2.1. The methods described above will not work effectively for these dependent variables due to the lack of independence.

Two typical approaches have been proposed to solve a DOE problem for dependent variables: (i) distance-based designs for nonrectangular regions (Hayeck 2009), and (ii) LHS with dependence (Stein 1987). However, a prerequisite of distance-based designs for nonrectangular regions is that nonrectangular regions are known, which cannot be satisfied for the problem described in Section 2.1. Lekivetz and Jones (2015) proposed a fast-flexible space-filling design for nonrectangular regions. This method achieves good space filling by taking a number of random points from the design region and clustering them such that each cluster is used to form a design point according to the MaxPro space-filling criterion (Joseph, Gul, and Ba 2015). The method has been embedded into the JMP (statistical software) implementation. By introducing a categorical factor, the design can consider variables from different groups (LFD, HFD, or PMD groups). The limitation of this method is that the design will be highly dependent on the defined factor constraints. In the problem discussed in Section 2.1, the constraints are not clear for a particular group of variables. The assumption of LHS with dependence is that independent and identically distributed (i.i.d.) samples with a joint distribution of all variables can be generated. Based solely on engineering knowledge, it is difficult to obtain i.i.d. samples with a joint distribution of all the variables in practice for the problem discussed in Section 2.1.

The grouped LHS approach is tailored to solve the engineering problem illustrated in Section 2.1. An overview of the main steps in the grouped LHS is shown in Figure 2 and are detailed below.

Assume that there are a total of q random variables that are further classified into K groups according to the engineering knowledge. Variables within each group are dependent while variables between groups are independent. The number of variables in different groups is $q_1, q_2, \ldots, q_K$, and $q = \sum_{k=1}^{K} q_k$. Assume that we are able to obtain samples $\boldsymbol{Y}_1^k, \boldsymbol{Y}_2^k, \ldots, \boldsymbol{Y}_N^k$ for random variables in the kth group such that each $\boldsymbol{Y}_i^k$ has a joint distribution $\boldsymbol{F}^k$. This assumption is reasonable because a sufficient number of samples can be obtained from an assembly process during regular production. Let

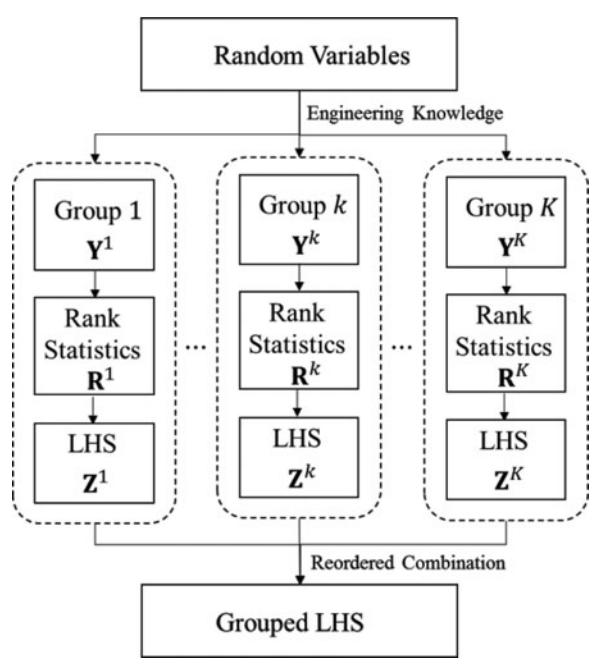


Figure 2. An overview of the grouped LHS.

$\boldsymbol{Y}^k = (\boldsymbol{Y}_1^k, \boldsymbol{Y}_2^k, \ldots, \boldsymbol{Y}_N^k)^T$ where $\boldsymbol{Y}^k$ is an $N \times q_k$ matrix. The jth column of $\boldsymbol{Y}^k$ is $Y_{1,j}^k, Y_{2,j}^k, \ldots, Y_{N,j}^k$. Reorder the jth column of $\boldsymbol{Y}^k$ as $Y_{(1),j}^k \le Y_{(2),j}^k \le \cdots \le Y_{(N),j}^k$. The index of $Y_{i,j}^k$ within $Y_{(1),j}^k, Y_{(2),j}^k, \ldots, Y_{(N),j}^k$ is the ith rank statistic, which can be expressed as

$$r_{i,j}^k \left(Y_{1,j}^k, Y_{2,j}^k, \ldots, Y_{N,j}^k\right) = \sum_{m=1}^{N} 1_{\left\{Y_{m,j}^k \le Y_{i,j}^k\right\}} \qquad [1]$$

Thus, an $N \times q_k$ matrix of rank statistics of the kth random variables group $\boldsymbol{R}^k$ can be produced. For simplification, we use $r_{i,j}^k$ instead of $r_{i,j}^k(Y_{1,j}^k, Y_{2,j}^k, \ldots, Y_{N,j}^k)$ to denote the ijth element of $\boldsymbol{R}^k$.

Based on the above formulations, an LHS for the kth random variables group $\boldsymbol{Z}^k = (\boldsymbol{Z}_1^k, \boldsymbol{Z}_2^k, \ldots, \boldsymbol{Z}_N^k)^T$ can be obtained by defining the ijth element of $\boldsymbol{Z}^k$as

$$Z_{i,j}^k = (\boldsymbol{F}^k)^{-1}\left(\frac{r_{i,j}^k + \xi_{i,j}^k - 1}{N}\right) \qquad [2]$$

where $(\boldsymbol{F}^k)^{-1}$ denotes the inverse joint distribution of random variables in the kth group and $\xi_{i,j}^k$ $(i = 1, \ldots N;\ j = 1, \ldots, q_k)$ is an $N \times q_k$ matrix including i.i.d. random variables that are uniformly distributed on the interval [0,1], independent of $\boldsymbol{R}^k$ and $\boldsymbol{Y}^k$. LHS with dependence is designed within groups. If random variables within a group are independent, the LHS with dependence will generate sampling points that have the same distribution as a regular LHS (Stein 1987).

Since we have obtained the LHS with dependence for each group of variables, we need to generate one combined LHS for all variables. The simplest way to get the grouped LHS is to produce the sample matrix $\boldsymbol{Z}$ by combining $\boldsymbol{Z}^1, \boldsymbol{Z}^2, \ldots, \boldsymbol{Z}^K$ row by row, which is given as

$$\boldsymbol{Z} = \cup_{i=1}^{K} \boldsymbol{Z}^i \qquad [3]$$

However, the matrix $\boldsymbol{Z}$ generated by Eq. [3] can be collinear, which means the effects of variables in different groups are merged and hard to distinguish.

Let $\boldsymbol{P}$ represent an $N \times k$ matrix and each column $\boldsymbol{P}_{\cdot j}(j = 1, \ldots, K)$of $\boldsymbol{P}$ is an independent random permutation of $\{1, \ldots, N\}$. The matrix $\boldsymbol{Z}$ can be reordered by the independent random permutation, shown in Eqs. [4] to [5]:

$$\boldsymbol{Z}^{(i)} = f_{\text{reorder}}\left(\boldsymbol{Z}^i, \boldsymbol{P}_{\cdot i}\right) \qquad [4]$$

$$\boldsymbol{Z} = \cup_{i=1}^{K} \boldsymbol{Z}^{(i)} \qquad [5]$$

The matrix $\boldsymbol{Z}$ generated by Eq. [5] denotes the grouped LHS.

In summary, the main steps of the grouped LHS are elaborated above and shown in Figure 2. The two properties can be realized by using a grouped LHS. First, the grouped LHS considers the dependence among variables within each group by rank statistics. Compared with space-filling designs in non-rectangular input regions, it can reflect the dependence by adaptively choosing non-rectangular regions corresponding to historical samples. Second, the grouped LHS can prevent collinearity among variables between different groups through a reorder function with random permutation.

3. Optimal feed-forward control for dimensional quality

The designed experiments from Section 2 are implemented in computer experiments to obtain the output variables. Next the input and output data is used to build a surrogate model. A universal Kriging model is developed for the composite assembly process, and the surrogate model is then embedded into the optimal feed-forward control algorithm. A multiobjective optimization will be conducted to determine the optimal control actions.

3.1. Surrogate model: A universal Kriging model for each output

According to the engineering specifications of a composite assembly process, there are a limited number of points whose dimensional variations are critical to the assembled product. Those dimensional variations are functions of part manufacturing errors, fixture errors, and tooling adjustments. Thus, a Kriging model can be built for each critical point, which represents the dimensional deviation in two aspects: mean outputs that show the basic trend of the response, and stationary Gaussian random processes that denote the uncertainty associated with the prediction (Santner, Williams, and Notz 2003).

$$\boldsymbol{Q}(\boldsymbol{x}) = \boldsymbol{g}^T(\boldsymbol{x}) \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon}(\boldsymbol{x}) \qquad [6]$$

where $\boldsymbol{x} = (\delta\boldsymbol{F}_1, \delta\boldsymbol{P}_1, \delta\boldsymbol{H}_1, \delta\boldsymbol{F}_2, \delta\boldsymbol{P}_2, \delta\boldsymbol{H}_2)$ represents the input variable vector; $\delta\boldsymbol{F}_i$ denotes the locating fixture deviation vector of part i; $\delta\boldsymbol{P}_i$ denotes the part manufacturing deviation vector of part i $(i = 1,2)$; and $\delta\boldsymbol{H}_i$ denotes the holding fixture deviation vector of part i. $\boldsymbol{\beta}$ is a vector representing unknown regression parameters. $\boldsymbol{\varepsilon}(\cdot)$ is a stationary Gaussian random process having zero mean. $\boldsymbol{Q}(\cdot)$ represents the output variable vector that corresponds to dimensional variations at the chosen measurement points after the assembly processes.

There are various parameterizations of the correlation function in a Kriging model. In this article, a Gaussian correlation function is used in the model because the squared-exponential formulation has properties of easy interpretation and being smoothly differentiable.

$$c(\phi, \boldsymbol{x}, \boldsymbol{z}) = \prod_{i=1}^{n} \exp\left[-\phi_i (x_{1i} - x_{2i})^2\right] \quad [7]$$

The correlation function represents the basic principle that when two inputs are close to each other in Euclidean distance, the correlation between the corresponding outputs will be high. This principle is commonly observed in the dimensional errors of composite parts. As a result, the uncertainty related to the predictions is small for input values that are similar to the training data set and large for input values that are far from the training data set.

$\boldsymbol{R}$ is the matrix that reveals stochastic process correlations between input samples in the training data set. The component of $\boldsymbol{R}$ is shown in Eq. [8]. $\boldsymbol{r}'$ is the vector that illustrates correlations between a new input sample $\boldsymbol{x}'$ and the input samples in the training data set. As the component in $\boldsymbol{r}'$ increases, the uncertainty related to the prediction becomes smaller. Here

$$R_{ij} = c(\phi, \boldsymbol{x}_i, \boldsymbol{x}_j),\ i, j = 1, \ldots, n_s \quad [8]$$

$$\boldsymbol{r}' = \left[c(\phi, \boldsymbol{x}_i, \boldsymbol{x}'), \ldots, c(\phi, \boldsymbol{x}_{n_s}, \boldsymbol{x}')\right]^T \quad [9]$$

Before using the model for prediction, the parameters in the Gaussian process model should be estimated. Maximum likelihood estimation (MLE) is used in parameter estimation. The optimal estimation of $\boldsymbol{\beta}$ that maximizes the likelihood function is equivalent to the generalized least square estimator:

$$\hat{\boldsymbol{\beta}} = \left(\boldsymbol{G}^T \hat{\boldsymbol{R}}^{-1} \boldsymbol{G}\right)^{-1} \boldsymbol{G}^T \hat{\boldsymbol{R}}^{-1} \boldsymbol{Q}^{n_s} \quad [10]$$

The basic strategy of predicting $\boldsymbol{Q}(\boldsymbol{x}')$ can be represented by

$$\hat{\boldsymbol{Q}}(\boldsymbol{x}') = \hat{\boldsymbol{Q}}_0 = \boldsymbol{g}^T(\boldsymbol{x}') \cdot \hat{\boldsymbol{\beta}} + \hat{\boldsymbol{r}}'^T \hat{\boldsymbol{R}}^{-1} (\boldsymbol{Q}^{n_s} - \boldsymbol{G} \cdot \hat{\boldsymbol{\beta}}) \quad [11]$$

where $\boldsymbol{G}$ denotes the matrix whose row vectors correspond to the regression function $\boldsymbol{g}^T(\boldsymbol{x}')$from the training data. $\boldsymbol{x}'$represents a new testing data vector. $\hat{\boldsymbol{R}}$ and $\hat{\boldsymbol{r}}'$ are determined from the MLE estimator of the correlation function.

3.2. Optimal feed-forward control strategy

After a surrogate model is built based on and validated by the testing data set, the surrogate model is used to develop a feed-forward control strategy. Control actions are implemented by programming tools (PTs) or shimming adjustment for dimensional error compensation.

An overview of the optimal feed-forward control strategy is presented in Figure 3. Here we use a two-part assembly problem to illustrate the control concepts and procedures. In the kth station of a composite assembly line, the deviations of locating fixtures and holding fixtures are collected from the measurements at a previous assembly station. Part manufacturing deviations of part 1 and part 2 are measured at preceding part manufacturing stations. The feed-forward controller collects these measurements and sets up virtual programming tools adjustment in the algorithm. The embedded surrogate model is used to predict the dimensional errors of the assembled parts. Since multiple responses are considered in composite parts' assembly processes, the feed-forward controller solves a multiobjective optimization problem before the determination of optimal PT adjustment. Afterward programming tools are applied to adjust the holding fixtures and parts are assembled into a new subassembly.

The objective of the feed-forward control is to minimize the deviations of all key output measurement points. As there are multiple measurement points of interest and the control adjustment affects all of them, a trade-off of the optimized deviation should be made in the control adjustment. Thus, a multiobjective optimization problem

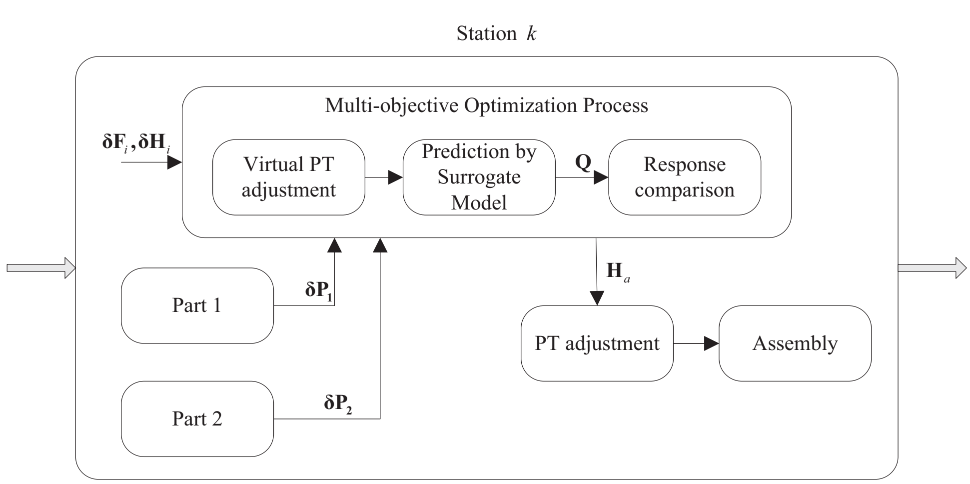


Figure 3. An overview of the feed-forward control strategy for dimension quality.

is formulated, and an engineering knowledge–based Pareto optimal weighted-sum approach is proposed to design the feed-forward control strategy.

The cost function to be minimized is

$$J = \sum_{i=1}^{N} q_i^2(\boldsymbol{x}) \cdot w_i = \boldsymbol{Q}^T(\boldsymbol{x}) \cdot \boldsymbol{W} \cdot \boldsymbol{Q}(\boldsymbol{x}) \quad s.t.\ \boldsymbol{x} \in \boldsymbol{S} \qquad [12]$$

where $q_i(\cdot)$, $i = 1, \ldots, N$ denotes the output functions in $\boldsymbol{Q}(\cdot)$. $\boldsymbol{S}$ represents the optimal Pareto Set of the multiobjective optimization problem shown in Eq. [13]. w_i, $i = 1, \ldots, N$ are the weighting coefficients given according to the engineering knowledge and $\boldsymbol{W}$ denotes the weighting coefficients matrix. The criterion to choosing weighting coefficients is to consider the relative importance of specific quality features in the products. The main idea of the optimization problem shown in Eq. [12] is to pick an optimal solution among all the Pareto solutions.

The multiobjective optimization problem to minimize the deviations of all output measurement points is formulated as

$$\min_{\boldsymbol{x}_a} |\boldsymbol{Q}(\boldsymbol{x}_a)| = \left|\boldsymbol{g}^T(\boldsymbol{x}_a) \cdot \boldsymbol{\beta} + \boldsymbol{\varepsilon}(\boldsymbol{x}_a)\right| \quad s.t.\ |\delta \boldsymbol{H}_{a1}| \le \boldsymbol{L}_{a1},\ |\delta \boldsymbol{H}_{a2}| \le \boldsymbol{L}_{a2} \qquad [13]$$

where $\boldsymbol{x}_a = (\delta\boldsymbol{F}_1, \delta\boldsymbol{P}_1, \delta\boldsymbol{H}_1, \delta\boldsymbol{H}_{a1}, \delta\boldsymbol{F}_2, \delta\boldsymbol{P}_2, \delta\boldsymbol{H}_2, \delta\boldsymbol{H}_{a2})$ denotes the adjusted input variable vector. $\delta\boldsymbol{H}_{a1}$ and $\delta\boldsymbol{H}_{a2}$ are the adjustable vectors of holding fixtures by programming tools. $\boldsymbol{L}_{a1}$ and $\boldsymbol{L}_{a2}$ are the threshold vectors for the adjustment of holding fixtures that constitute a feasible decision-variable space determined by the engineering design of the system. The target of the optimization problem is to make the multipoint deviations after assembly as small as possible under the feasible adjustments of PTs.

The NSGA-II procedure (Deb 2001; Deb et al. 2002), one of the typical evolutionary multiobjective optimization (EMO) methods, was used to find a Pareto-optimal frontier corresponding to the minimizations of the objective function with the constraints. The NSGA-II algorithm has four advantages: (i) parallel nondominated solutions are computed at the same time; (ii) it does not require any derivative information; (iii) the algorithm is relatively simple and flexible in implementation; and (iv) the NSGA-II uses an elitist principle and an explicit diversity-preserving mechanism.

The Pareto-optimal front holds a number of solutions. Engineering knowledge–based weighted sum minimization as shown in Eq. [12] is used to determine the optimal solution from the Pareto-optimal solutions. By considering the weights of different outputs the optimal solution effectively and efficiently reduces all deviations, especially in important regions. As a result, the product quality is improved and dimensional variability is reduced.

4. Case study

A case study is conducted to illustrate the design of experiment by using a grouped LHS, to test the prediction accuracy of the universal Kriging model, and to evaluate the effectiveness of the optimal feed-forward control strategy for dimensional-variation reduction of composite parts' assembly processes.

As an example, we focus on the assembly process of two composite laminated plates (Figure 4), which was also used in the case study of (Zhang and Shi 2016a). Part 1 is $1{,}000 \times 1{,}000 \times 4\ \text{mm}^3$ with four layers. The orientations of four layers are 90°/0°/90°/0° and the thickness of each layer is 1 mm. Material properties of each layer are summarized in Table 2. Part 2 is $1{,}000 \times 1{,}000 \times 6\ \text{mm}^3$ and has six layers with orientations of 90°/0°/90°/0°/90°/0°. Each layer shares the same thickness and material properties as part 1.

We consider three kinds of dimensional deviations: locating fixture deviations, holding fixture deviations, and part manufacturing deviations. As shown in Figure 4, LFP_j^i denotes the jth locating fixture point in part i; HFP_j^i denotes the jth holding fixture point in part i; and JP_j^i denotes the jth joint point in part i. Three locating fixture points are considered with the 3-2-1 constraint principle. Two holding fixture deviations are set up along the Z axis. When the part is positioned at the assembly station based on three locating fixtures, part manufacturing deviations at six well-distributed points (including two holding fixture points, three joint points, and the midpoint between

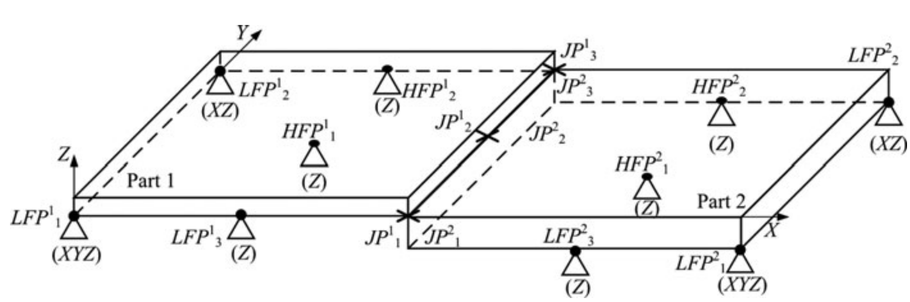


Figure 4. Diagram of assembly of part 1 and part 2.

Table 2. Material properties.

EX/MPa	EY/MPa	EZ/MPa	GXY/MPa	GYZ/MPa	GXZ/MPa	PRXY	PRYZ	PRXZ
1,250	300	300	50	20	50	0.25	0.25	0.01

Table 3. Input variable summary.

Part	Input Variables	Number
1	LFD: locating fixture deviations	3
1	HFD: holding fixture deviations	2
1	PMD: part manufacturing deviations	6
2	LFD: locating fixture deviations	3
2	HFD: holding fixture deviations	2
2	PMD: part manufacturing deviations	6

LFP_3^i and JP_1^i) are considered. In total 22 input variables in part 1 and part 2 are summarized in Table 3. After assembly, dimensional deviations at 17 measurement points are collected as output variables. The 17 measurement points are marked as circled numbers in Figure 4.

The total input variables are divided into six groups as shown in Table 3. Based on variable analysis and engineering knowledge, the following two characteristics for six groups of input variables can be obtained: (i) variables within the group are dependent, but the dependent relationships are unknown; and (ii) variables between groups are independent. Next we use the proposed grouped LHS approach for this case study's DOE and modeling.

The grouped LHS provides 100 sets of input variables. Variables within each group are dependent, while variables between groups are independent. Figure 5 depicts the bivariate projection of a grouped LHS of 22 input variables. On the left is the bivariate projection of variables LFD1 and LFD3, which clearly reflects the dependence within one group. There are fewer points in the upper-left and bottom-right corners because of the correlation of LFD1 and LFD3. The right graph of Figure 5 illustrates the bivariate projection of variables HFD2 and PMD5, which belong to two independent groups. All the points are relatively uniformly distributed in the bivariate-projection picture.

In the case study, FEM is conducted by using the commercial software ANSYS in the computer simulation. The results in the grouped LHS are used as the input configurations in the simulation. The nodal z-dimensions of the uniformly determined 17 points are collected as the output variables. The FEM simulation result of one training sample is shown in Figure 6.

In the case study, 100 sets of data are divided into a training data set and a testing data set along a 70/30 percent split. The surrogate-modeling approach, such as the universal Kriging model discussed in Section 3, is applied to the data analysis and modeling. The training data set is used to estimate parameters in the surrogate model; testing input variables are then introduced into the surrogate model for prediction. The surrogate model is embedded into a feed-forward control algorithm. According to the procedure in Section 3.2, the optimal feed-forward control algorithm is used to improve the dimensional quality of composite parts' assembly processes. Seventy groups of training samples are used to perform parameter estimation and obtain the surrogate model for each output variable. Then 30 groups of testing data are simulated in the assembly process with or without control. All the surrogate prediction outputs are validated in the FEM simulation. Dimensional deviations between predicted outputs and real outputs from the FEM simulation are calculated.

The box plot shown in Figure 7 compares the prediction errors with or without optimal feed-forward control. The prediction errors of different output variables are

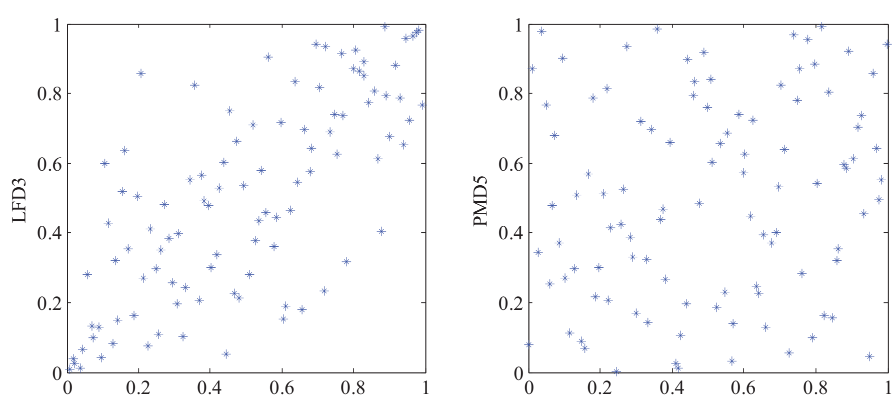


Figure 5. Left: one bivariate projection of GLHS within the LFD group (variables: LFD1 and LFD3); right: one bivariate projection of GLHS between the HFD group and the PMD group (variables: HFD2 and PMD5).

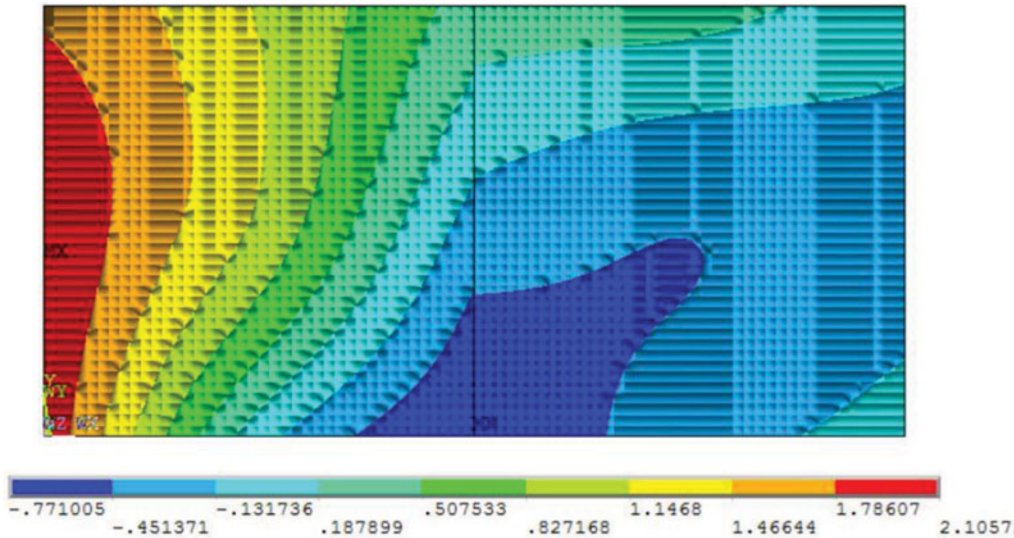


Figure 6. FEM simulation result of one training sample.

obtained by comparing the real FEM outputs and the predicted outputs obtained from using the universal Kriging model. Figure 7 shows that the universal Kriging model provides adequate prediction accuracy with a prediction error less than 0.09 mm whether or not the optimal feed-forward control is applied. This means that the surrogate model can describe the nonlinear properties of composite parts well and has good predictive capability.

The prediction and FEM results of assembly with and without optimal feed-forward control are shown in Figure 8 and Figure 9. Figure 8 shows the comparison between box plots that show dimensional deviations with or without optimal feed-forward control in the surrogate model prediction. Both the mean and standard deviation (STD) of dimensional deviations are improved with the optimal feed-forward control strategy during the assembly processes.

Figure 9 illustrates the FEM result comparison of dimensional deviations with or without optimal feed-forward control in the surrogate model prediction. In

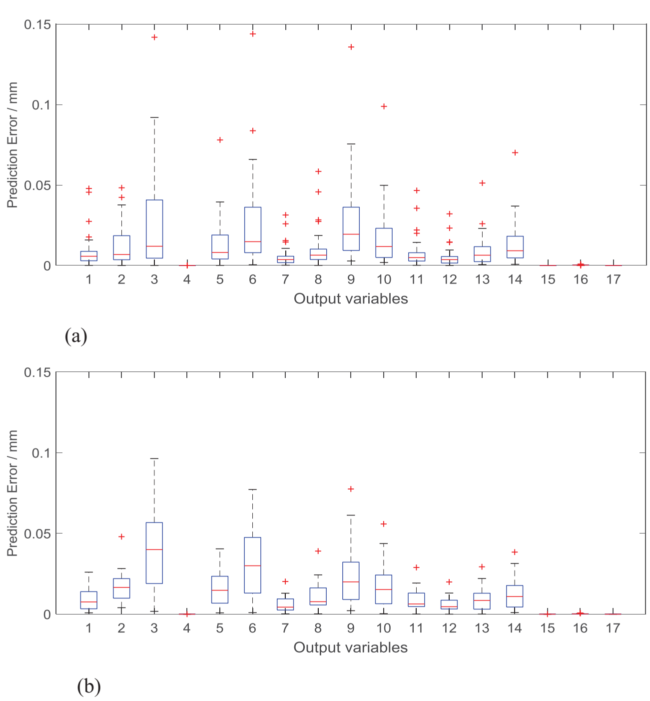


Figure 7. Prediction error of the universal Kriging model with or without optimal feed-forward control: (a) prediction errors without optimal feed-forward control; (b) prediction errors with optimal feed-forward control.

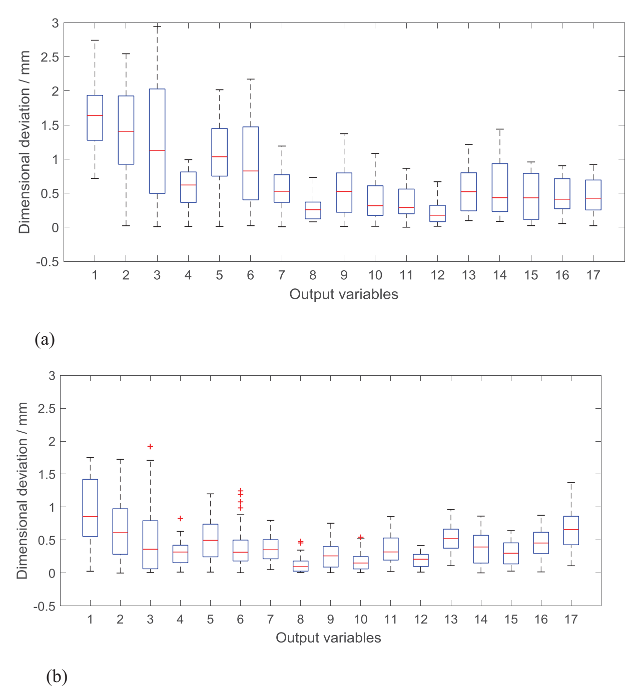
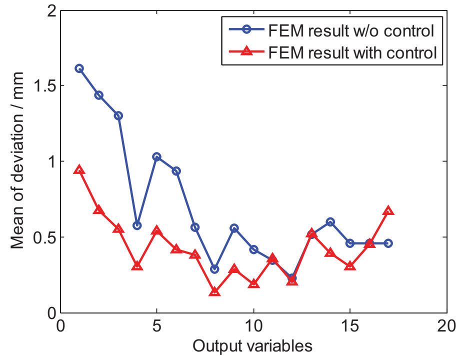
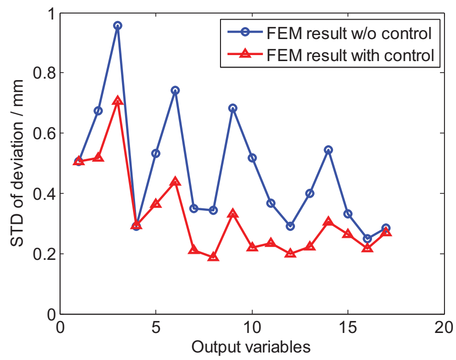


Figure 8. Prediction results of composite parts' assembly processes with or without optimal feed-forward control: (a) box plot of dimensional deviations without optimal feed-forward control; (b) box plot of dimensional deviations with optimal feed-forward control.



(a)



(b)

Figure 9. FEM results of composites' assembly processes with or without optimal feed-forward control: (a) mean of dimensional deviations; (b) STD of dimensional deviations.

Figure 9(a) the line with circle markers reflects the mean deviations of 17 output variables of the results of the FEM simulation without optimal feed-forward control. The results of the FEM simulation with optimal feed-forward control are shown in the line with triangle markers. In Figure 9(b) the dashed line shows the FEM result of the mean of assembly deviations without optimal feed-forward control. The dashed-dotted line shows corresponding FEM results of the STD of assembly deviations with optimal feed-forward control. Figure 9 shows that the optimal feed-forward control strategy can reduce both the mean and STD of the dimensional deviations of composite parts' assembly processes. Assume that $\mu_{c,i}$ and $\sigma_{c,i}$ are the mean and the STD of assembly deviations with optimal feed-forward control in the ith output variables, and $\mu_{oc,i}$ and $\sigma_{oc,i}$ are the mean and the STD of assembly deviations without optimal feed-forward control. The improvement ratio obtained through using the feed-forward control can be evaluated by $(\sum_i \mu_{c,i} - \sum_i \mu_{oc,i})/\sum_i \mu_{oc,i}$, and $(\sum_i \sigma_{c,i} - \sum_i \sigma_{oc,i})/\sum_i \sigma_{oc,i}$. Through evaluation indices the mean of dimensional deviations can be reduced by 37.94 percent and the STD of dimensional deviations 31.84 percent.

5. Conclusions

Due to the nonlinear property of composite materials, analytical physics-based models cannot accurately approximate the assembly process. Numerical-simulation models like FEM can describe the composite parts' assembly process effectively; however, their dependence on manual operations limits their scope of application to validation rather than in-process control. In this article we developed a surrogate model–based feed-forward control strategy for in-process compensation and dimensional-variation reduction of composite parts' assembly processes.

First, we proposed a tailored DOE approach—grouped Latin hypercube sampling—to obtain samples from designed experiments. The proposed grouped LHS has two appealing properties: (i) it considers dependence among variables within each group by rank statistics, and (ii) it can prevent collinearity among variables between different groups using a reordering function with random permutation. Second, we implemented FEM by the computer experiments according to the DOE of input variables, then we developed the surrogate model to link the inputs and outputs. Third, we embedded the surrogate model into the optimal feed-forward control algorithm and solved multiobjective optimization problems to determine the optimal control actions. The case study showed that the optimal feed-forward control strategy can reduce the mean of dimensional deviations by 37.94 percent and the STD of dimensional deviations by 31.84 percent.

The main contribution of the article was the framework of a surrogate model–based in-process control, which can be extended to other applications. Even though the surrogate model–based in-process control is successful in the variation reduction of the composite parts' assembly process, it may not be good for some specific systems with very high uncertainties. This is because the strategy is dependent on the accuracy of the surrogate model. If the training points are relatively "close" to each other, or the noisy training data set is highly stochastic, the surrogate model may not be good enough to be used in the control algorithm. This is a potential limitation of the proposed strategy. Further research is needed to conduct uncertainty quantification of the surrogate model, to analyze the criteria of controllability, and to determine the stability of the surrogate model–based control strategy.

About the authors

Mr. Yue is a PhD candidate in the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology. He is a member of the American Society for Quality. His email address is xwy@gatech.edu.

Dr. Shi is the Carolyn J. Stewart Chair and Professor in the H. Milton Stewart School of Industrial and Systems Engineering at the Georgia Institute of Technology. He is a senior member of the American Society for Quality. His email address is jianjun.shi@isye.gatech.edu.

ORCID

Xiaowei Yue http://orcid.org/0000-0001-6019-0940
Jianjun Shi http://orcid.org/0000-0002-3774-9176

References

Camacho, E. F., and C. B. Alba. 2013. *Model predictive control*. New York, NY: Springer Science and Business Media.

Deb, K. 2001. *Multi-objective optimization using evolutionary algorithms*. Hoboken, NJ: John Wiley and Sons.

Deb, K., A. Pratap, S. Agarwal, and T. Meyarivan. 2002. A fast and elitist multiobjective genetic algorithm: NSGA-II. *IEEE Transactions on Evolutionary Computation* 6:182–97. doi:10.1109/4235.996017.

Dorf, R. C., and R. H. Bishop. 2011. *Modern control systems*. Upper Saddle River, NJ: Pearson.

Forrester, A., A. Sobester, and A. Keane. 2008. *Engineering design via surrogate modelling: A practical guide*. Hoboken, NJ: John Wiley and Sons.

Hayeck, G. T. 2009. The kinematics of the upper extremity and subsequent effects on joint loading and surgical treatment. Doctoral dissertation, Cornell University.

Izquierdo, L. E., J. Shi, S. J. Hu, and C. W. Wampler. 2007. Feedforward control of multistage assembly processes using programmable tooling. *Transactions of NAMRI/SME* 35:295–302.

Jin, J., and J. Shi. 1999. State space modeling of sheet metal assembly for dimensional control. *ASME Transactions, Journal of Manufacturing Science and Engineering* 121:756–62. doi:10.1115/1.2833137.

Jones, R. M. 1998. *Mechanics of composite materials*. 2nd ed. New York, NY: CRC press, Taylor and Francis Group.

Joseph, V. R., E. Gul, and S. Ba. 2015. Maximum projection designs for computer experiments. *Biometrika* 102 (2):371–80. doi:10.1093/biomet/asv002.

Joseph, V. R., and Y. Hung. 2008. Orthogonal-maximin Latin hypercube designs. *Statistica Sinica* 18:171.

Koziel, S., and L. Leifsson. 2013. *Surrogate-based modeling and optimization*. New York, NY: Applications in Engineering. Springer Science and Business Media.

Lekivetz, R., and B. Jones. 2015. Fast flexible space-filling designs for nonrectangular regions. *Quality and Reliability Engineering International* 31 (5):829–37. doi:10.1002/qre.1640.

Liu, S. C., and S. J. Hu. 1997. Variation simulation for deformable sheet metal assemblies using finite element methods. *ASME Transactions, Journal of Manufacturing Science and Engineering* 119:368–74. doi:10.1115/1.2831115.

Loose, J. P., N. Chen, and S. Zhou. 2009. Surrogate modeling of dimensional variation propagation in multistage assembly processes. *IIE Transactions* 41:893–904. doi:10.1080/07408170902906027.

Mantripragada, R., and D. E. Whitney. 1999. Modeling and controlling variation propagation in mechanical assemblies using state transition models. *IEEE Transactions on Robotics and Automation* 15:124–40. doi:10.1109/70.744608.

McKay, M. D., R. J. Beckman, and W. J. Conover. 1979. A comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 42:55–61. doi:10.1080/00401706.2000.10485979.

Menassa, R. J., and W. R. DeVries. 1989. Locating point synthesis in fixture design. *CIRP Annals-Manufacturing Technology* 38:165–69. doi:10.1016/S0007-8506(07)62676-6.

Montgomery, D. C. 2007. *Introduction to statistical quality control*. Hoboken, NJ: John Wiley and Sons.

Qian, P. Z. 2012. Sliced Latin hypercube designs. *Journal of the American Statistical Association* 107:393–99. doi:10.1080/01621459.2011.644132.

Ramesh, R., M. Mannan, and A. Poo. 2000. Error compensation in machine tools—A review: Part I: Geometric, cutting-force induced and fixture-dependent errors. *International Journal of Machine Tools and Manufacture* 40:1235–56. doi:10.1016/S0890-6955(00)00009-2.

Santner, T. J., B. J. Williams, and W. I. Notz. 2003. *The design and analysis of computer experiments*. New York, NY: Springer Science and Business Media.

Shi, J. 2006. *Stream of variation modeling and analysis for multistage manufacturing processes*. Boca Raton, FL: CRC Press, Taylor and Francis Group.

Shi, J. 2013. Data fusion for in-process quality improvement. Proceedings of the 5th Manufacturing Engineering Society International Conference, Zaragoza, Spain, June. (invited Keynote talk).

Stein, M. 1987. Large sample properties of simulations using Latin hypercube sampling. *Technometrics* 29:143–51. doi:10.1080/00401706.1987.10488205.

Tang, B. 1993. Orthogonal array-based Latin hypercubes. *Journal of the American Statistical Association* 88:1392–97. doi:10.1080/01621459.1993.10476423.

Wen, Y., X. Yue, J. Hunt, and J. Shi. 2018. Feasibility analysis of composite fuselage shape control via finite element analysis. *Journal of Manufacturing Systems* 46:272–81. doi:10.1016/j.jmsy.2018.01.008.

Wu, C. J., and M. S. Hamada. 2011. *Experiments: Planning, analysis, and optimization*. 2nd ed. Hoboken, NJ: John Wiley and Sons.

Yue, X., Y. Wen, J. H. Hunt, and J. Shi. 2018. Surrogate model based control considering uncertainties for composite fuselage assembly. *ASME Transactions, Journal of Manufacturing Science and Engineering* 140 (4):041017-041017–13. doi:10.1115/1.4038510.

Zhang, T., and J. Shi. 2016a. Stream of variation modeling and analysis for compliant composite part assembly - Part I: Single station processes. *ASME Transactions, Journal of Manufacturing Science and Engineering* 138:121003-121003–15.

Zhang, T., and J. Shi. 2016b. Stream of variation modeling and analysis for compliant composite part assembly - Part II: Multistation processes. *ASME Transactions, Journal of Manufacturing Science and Engineering* 138:121004-121004–15.

Zhong, J., J. Liu, and J. Shi. 2010. Predictive control considering model uncertainty for variation reduction in multistage assembly processes. *IEEE Transactions on Automation Science and Engineering* 7:724–35. doi:10.1109/TASE.2009.2038714.

Zhou, Q., P. Z. Qian, and S. Zhou. 2012. Surrogate modeling of multistage assembly processes using integrated emulation. *Journal of Mechanical Design* 134:011002. doi:10.1115/1.4005440.