

# MODEL TRANSFER ACROSS ADDITIVE MANUFACTURING PROCESSES VIA MEAN EFFECT EQUIVALENCE OF LURKING VARIABLES

BY ARMAN SABBAGHI<sup>1</sup> AND QIANG HUANG<sup>2</sup>

*Purdue University and University of Southern California*

Shape deviation models constitute an important component in quality control for additive manufacturing (AM) systems. However, specified models have a limited scope of application across the vast spectrum of processes in a system that are characterized by different settings of process variables, including lurking variables. We develop a new effect equivalence framework and Bayesian method that enables deviation model transfer across processes in an AM system with limited experimental runs. Model transfer is performed via inference on the equivalent effects of lurking variables in terms of an observed factor whose effect has been modeled under a previously learned process. Studies on stereolithography illustrate the ability of our framework to broaden both the scope of deviation models and the comprehensive understanding of AM systems.

## 1. Introduction.

1.1. *The challenge of model transfer across different processes in an additive manufacturing system.* Recent advances in the industrial Internet of Things and Cyber-Physical Systems have resulted in greater connections and accessibility of distinct manufacturing processes. A particularly exciting consequence is the development of a new paradigm of additive manufacturing (AM) systems that seamlessly integrate computing, manufacturing, and services [Buckholtz, Ragai and Wang (2015), Germany Trade & Invest (2014), Wu et al. (2015)]. Each individual AM, or three-dimensional (3D) printing, process in such a system enables direct manufacturing of complex shapes from computer-aided design (CAD) models with reduced labor and costs compared to traditional manufacturing methods [Campbell et al. (2011), Gibson, Rosen and Stucker (2009)]. The impacts of such systems are not yet fully realized in practice because their constituent processes typically yield shapes whose dimensions are discrepant with those specified in the CAD models. These discrepancies, referred to as geometric shape deviations, occur due to the rapid phase changes inherent in an AM process. A fundamental issue that remains

---

Received September 2017; revised January 2018.

<sup>1</sup>Supported by NSF Grant CMMI-1744123.

<sup>2</sup>Supported by NSF Grant CMMI-1744121.

*Key words and phrases.* 3D printing, Bayesian learning, transfer learning.

to be addressed is comprehensive and efficient deviation control across processes in an AM system.

Several quality control methods have been developed to address the deviation issue for AM and are concisely described in Huang et al. [(2014), pages 061008–061009]. One particularly effective method is to specify a statistical model for a shape's deviations and devise changes to the CAD model, referred to as a compensation plan, that are predicted to reduce deviations in new printings of the shape. Tong, Lehtihet and Joshi (2003) and Tong, Joshi and Lehtihet (2008) applied this method using separate polynomial regression deviation models for each direction of a shape. In contrast, Huang et al. (2015) conceived of a distinct functional modeling framework that effectively accounts for the correlation in deviation between different directions and decouples geometric shape complexity from deviation modeling and compensation. Further advances and experiments under their framework include in-plane deviation modeling for polygons and free-form shapes [Huang et al. (2014), Luan and Huang (2017), Sabbaghi, Huang and Dasgupta (2018)], interference modeling for discretized compensation plans [Sabbaghi et al. (2014)] and out-of-plane deviation modeling for 3D shapes [Jin, Qin and Huang (2016)].

The common limitation of these existing methods is that their learned deviation models are limited in their scope of application to the particular processes for which they were specified, in the sense that they typically fail to describe, and hence control, deviations generated under other processes. Comprehensive control of an AM system based on these methods is then impractical due to the high operating costs incurred by collecting a large amount of data and constructing new models for distinct processes in the system. In addition, the number of test cases that could possibly be manufactured for a particular process is typically in the single digits because of its nature and capability of one-of-a-kind manufacturing [Sabbaghi, Huang and Dasgupta (2018)]. More importantly, modeling deviations separately for every process fails to yield deeper insights on the entire AM system. Therefore, a significant challenge for comprehensive shape deviation control in an AM system is whether the deviation model for one process can be transferred to model deviations under a different process based on a small number of experimental runs. The more general category of engineering problems is transferring a quality model established for a process  $A$  to another process  $B$ , where  $B$  has substantially fewer trials to avoid wasteful and repetitive model building (i.e., the product design set  $D_B$  for  $B$  is a small subset of the corresponding product design set  $D_A$  for  $A$ , with a potentially smaller dimension as well).

This challenge is illustrated by consideration of in-plane deviation profiles under two distinct stereolithography processes  $A$  and  $B$  in the AM system of Figure 1. Let  $\mathcal{X}$  denote the set of compensation plans and  $D_A$  and  $D_B$  the respective design sets for the two processes. The component of a statistical deviation model that is important for deviation control, and hence constitutes our primary focus, is its expectation. We denote the expected deviation models for these processes by

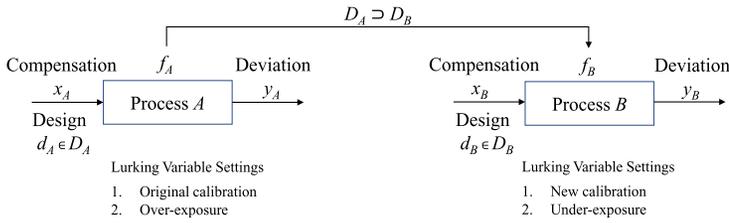


FIG. 1. Model transfer across processes in an AM system characterized by different settings of lurking variables.

$f_A : \mathcal{X} \times D_A \rightarrow \mathcal{P}$  and  $f_B : \mathcal{X} \times D_B \rightarrow \mathcal{P}$ , where  $\mathcal{P}$  is the set of in-plane deviation profiles [Huang et al. (2015), page 432]. We consider engineering processes to be distinct if their designs, parameters, or any other factors are different. Needless to say, processes involving completely different AM machines are distinct. Even if the observed factors’ settings for A and B are identical, these processes and their respective expected deviation models can be distinguished by a wide spectrum of unobserved factors related to their product designs, materials, parameters and conditions. Such lurking variables, whose settings are completely unobserved due to infeasibility of measurement or insufficient knowledge [Box (1966)], are ubiquitous in AM systems and complicate the task of model transfer. Figure 2(a) displays the in-plane deviation profiles for four cylinders of nominal radii 0.5”, 1”, 2”, and 3” manufactured under process A. Each point on a cylinder is identified by its angle under the polar coordinate system [Huang et al. (2015), page 432]. This par-

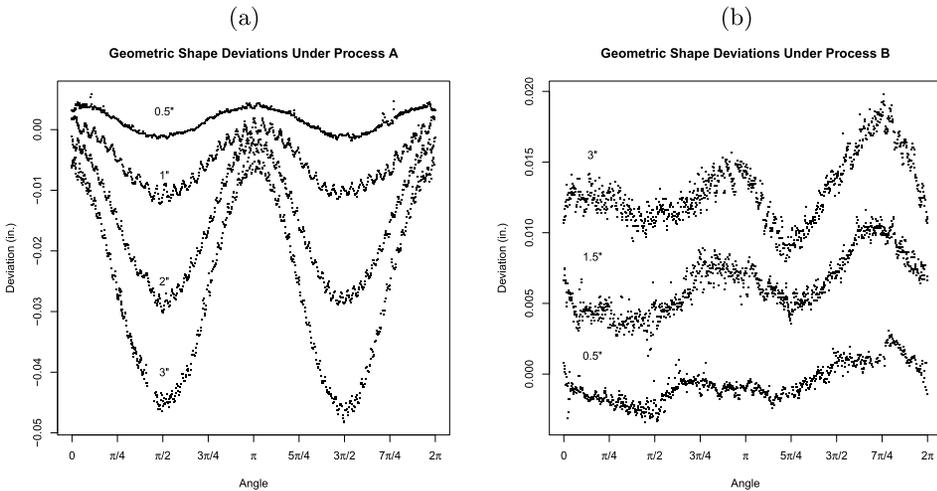


FIG. 2. (a) Deviation profiles of four cylinders of nominal radii 0.5”, 1”, 2”, and 3” manufactured under process A. (b) Deviation profiles of three cylinders of nominal radii 0.5”, 1.5”, and 3” manufactured under process B which has a new calibration setting.

ticular process was studied by [Huang et al. \(2015\)](#), and a simple specification of  $f_A$  was formulated to enable the construction of compensation plans that reduce in-plane deviation by one order of magnitude. Figure 2(b) displays the different deviation profiles for three cylinders of nominal radii 0.5", 1.5", and 3" manufactured under a new process  $B$  that operates with an unknown and distinct calibration setting. The lurking calibration factor clearly impacts deviation in a complicated manner, and  $f_A$  fails to directly capture the deviation profiles under  $B$ . Resource constraints may prevent the specification of an appropriate model  $f_B$  based only on the data from  $B$ . It thus becomes of interest to extend the previously learned model  $f_A$  to  $B$ . In this way we can leverage all of the data and knowledge across the AM system to acquire a better understanding of the change in the lurking calibration factor for process  $B$ .

As described in the literature review in Section 1.2, current statistical and machine learning methodologies cannot resolve the challenge of deviation model transfer across processes in an AM system in the presence of lurking variables. We address this challenge by incorporating effect equivalence for a lurking variable in terms of an observed factor into the model specified under a previous process. Effect equivalence refers to the common engineering phenomenon in which two factors can generate the same outcomes. [Wang, Huang and Katz \(2005\)](#) conceived of its first quantitative formulation in their study of a machining process. However, they did not consider its application for model transfer across different processes operating with distinct lurking variables settings in a system. We develop a new statistical framework and Bayesian method for such model transfer based on effect equivalence.

Under our approach, model transfer proceeds by learning a function that benchmarks a lurking variable's effect on the process mean in terms of a previously studied process and an observed factor's effect on the mean under that process. For the previous AM system, this function is denoted by  $T : \mathcal{X} \times D_B \rightarrow \mathcal{X}$ , and is specified so that  $f_B(x_B, d_B) = f_A(T(x_B, d_B), d_B)$  for all  $(x_B, d_B) \in \mathcal{X} \times D_B$ . We refer to  $T$  as the total equivalent amount of the lurking calibration in terms of compensation and interpret it as returning a compensation plan  $T(x_B, d_B) \in \mathcal{X}$  for input  $(x_B, d_B) \in \mathcal{X} \times D_B$  such that the expected deviation profile of the product with input  $(x_B, d_B)$  manufactured under process  $B$  is equivalent to the expected deviation profile of the product with shape design  $d_B$  and compensation plan  $T(x_B, d_B)$  manufactured under process  $A$ . In this manner  $T$  directly broadens the scope of the previously specified model  $f_A$  to the expectation of process  $B$  and facilitates one's understanding of  $B$ 's changed lurking variable setting. Indeed, our analysis on the total equivalent amount of calibration in terms of compensation [summarized in Figure 3(a) and described in more depth in Section 4] helps us to understand that the calibration change is essentially an attempted reproduction of the optimum compensation plan of [Huang et al. \(2015\)](#) with discrepancies for the 1.5" and 3" cylinders that explain their complicated deviations. Also, a shape manufactured under process  $B$  can now be thought of as having been equivalently manufactured

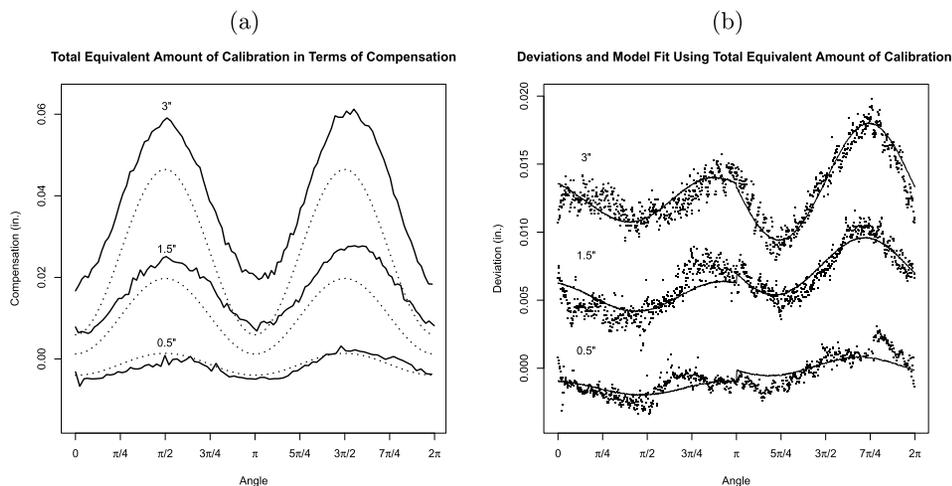


FIG. 3. (a) The inferred total equivalent amounts of calibration in terms of compensation (solid lines) for cylinders under process  $B$ , compared to the corresponding optimum compensation plans (dotted lines) from Huang *et al.* (2015). (b) Posterior predictive means (solid lines) of shape deviations under process  $B$  obtained from the transferred deviation model.

under  $A$  with a compensation plan defined by the inferred total equivalent amount. This enables us to transfer the model  $f_A$  for the mean of  $A$  to  $B$  [Figure 3(b)].

1.2. *Previous considerations of model transfer and lurking variables.* Various statistical methods exist to identify and account for lurking variables in simple models. A prime example for industrial processes is statistical process control [Shewhart (1931)]. For agricultural settings Yates and Cochran (1938) discussed how the ANOVA for a group of agronomic experiments should be conducted to account for lurking variables such as soil differences and climate. Joiner (1981) presented several instances of lurking variables and methods to detect their existence in linear models. Cook and Critchley (2000) described graphical methods for linear models that can illuminate phenomena resulting from lurking variables. However, such methods are ineffective for the nonlinear and complex deviations in AM. Furthermore, the standard practice of conducting follow-up experiments to learn about lurking variables necessarily proceeds on a case-by-case basis, and hence will be expensive, as illustrated by the cryogenic flow meters example in page 232 of Joiner (1981).

A broad class of model transfer methods belongs to inductive transfer learning [Pan and Yang (2010), page 1348]. One such method is TrAdaBoost [Dai *et al.* (2007)], which predicts the outcomes in a new setting by modifying the relative weights of instances across settings at each step of a boosting algorithm, so as to identify and utilize data from settings most similar to the new setting [Pardoe and Stone (2010)]. These methods enjoy desirable predictive and computational

performance and are simple to implement. However, they require prohibitively large samples for AM. In contrast, our framework enables model transfer with limited samples.

A topic related to model transfer is transportability, or the extrapolation of experimental findings across domains that differ both in their distributions and in their inherent causal characteristics [Bareinboim and Pearl (2016), page 7350]. Pearl and Bareinboim (2014) and Bareinboim and Pearl (2016) formalized transportability using causal diagrams and the do-calculus of Pearl (1995). A contribution of their work is the explication of conditions and transport formulae for causal inference from heterogeneous data. However, they focus on nonparametric inference and effectively assume that the relationships between factors and outcomes are sufficiently well understood from a large amount of data such that estimation error in probability distributions are not of concern. In contrast, for AM processes we must address the distinct task of transportability of parametric models across factor settings, and our formulation can handle the case of limited data. In addition, their work may be limited to linear structural equation models in practice, whereas our framework accommodates nonlinear models of interest in AM. Finally, we consider equivalence relations, not probabilistic dependencies, between observed and unobserved factors, so as to transfer models in a functional manner.

Effect equivalence was inspired by the study of a machining process in which observed and unobserved factors could yield identical outcomes. Wang, Huang and Katz (2005) considered the effects of fixture, machine tool and datum errors on quality, where the latter two are lurking variables. They described the existence of effect equivalence and specified the total equivalent amounts of the lurking variables in terms of fixture error. This approach enabled successful quality control [Wang and Huang (2006, 2007)]. However, they did not consider model transfer across different processes, characterized by distinct settings of lurking variables.

Following this introduction and literature review, we proceed to describe in Section 2 our statistical effect equivalence framework. Section 2.1 reviews the representation and measurement of deviations for AM processes, and Section 2.2 contains the notations and assumptions for our framework. Effect equivalence with respect to the mean is formally defined in Section 2.3. Our Bayesian approach to perform model transfer via the equivalent effects of lurking variables with respect to the mean is developed in Section 2.4. We then perform model transfer for two AM processes with distinct settings of lurking overexposure and calibration factors in Sections 3 and 4 respectively. The data collected on these processes are described in Sections 3.1 and 4.1. Inferences and interpretations on the total equivalent amounts of the lurking variables in terms of compensation with respect to the mean are in Sections 3.2 and 4.2 respectively. Model transfer for these processes are performed in Sections 3.3 and 4.3. These studies demonstrate how the synergy between statistics and engineering in our framework facilitates deviation modeling across processes in an AM system more generally. We conclude in Section 5 by describing the broader impacts of our general framework.

## 2. Statistical effect equivalence framework.

2.1. *Functional in-plane deviation representation.* Throughout this article we consider in-plane deviations for 3D printed shapes with negligible heights and adopt the functional in-plane deviation representation defined by Huang et al. (2015), page 432. Our focus on in-plane deviations is due to two facts. First, in real applications a thin product or section of a product is approximated as a 2D shape with identical top and bottom surface deviations. Second, Jin, Qin and Huang (2016) demonstrated that fully 3D shapes introduce challenging complications, such as interlayer bonding effects, that detract from our fundamental question of interest for AM systems.

We measure the top boundaries of shapes manufactured by our AM processes using a single Micro-Vu Vertex system. The measurement coordinate system is aligned with the design coordinate system based on markers printed on the shapes to ensure consistency and reduce possible measurement errors or misalignment issues. The points  $i = 1, \dots, N$  on the manufactured shapes are then identified by their angles  $\theta_i$  and observed radii  $r^{\text{obs}}(\theta_i)$  under the polar coordinate system.

A CAD model is encoded as a nominal radius function  $r^{\text{nom}} : [0, 2\pi] \rightarrow \mathbb{R}$ . Deviation  $Y_i$  for point  $i$  on a product with CAD model  $r^{\text{nom}}$  is defined as

$$Y_i = r^{\text{obs}}(\theta_i) - r^{\text{nom}}(\theta_i).$$

Figures 2(a) and 2(b) illustrate this functional deviation representation for several thousands of points on each of the seven cylinders.

Studies are conducted to model deviation as a function of the continuous treatment factor compensation. This factor is formally defined for each point  $i$  as the addition or subtraction of material in the CAD model at that point. The previous notation in the case of compensation becomes

$$Y_i(x_1) = r^{\text{obs}}(\theta_i, x_1) - r^{\text{nom}}(\theta_i),$$

where  $r^{\text{obs}}(\theta_i, x_1)$  denotes the observed radius for point  $i$  under compensation  $x_1$ , and  $Y_i(x_1)$  is point  $i$ 's deviation.

2.2. *Notation and assumptions.* We denote the  $K$  factors for an AM process by  $F_k$  and their corresponding set of possible levels by  $\mathcal{X}_k$ , for  $k = 1, \dots, K$ . We consider processes whose factors possess the following two properties. First is that each is either fully observed or unobserved. Second is that the level for any factor is not an outcome of another factor. The first property ensures that each factor is well defined as either an observed factor or lurking variable, and the second eliminates factor linkages [Box (1966), pages 625–626]. Factor  $F_1$  will be reserved for compensation with  $\mathcal{X}_1 = \mathbb{R}$ .

The covariate vector for point  $i$  is denoted by  $z_i$ . An example of a covariate vector is  $z_i = (\theta_i, r^{\text{nom}}(\theta_i))^T$ . As shall be seen in our studies, such covariates serve a useful role for deviation model transfer.

We assume the Stable Unit-Treatment Value Assumption [Imbens and Rubin (2015), page 10] is satisfied. Under this assumption each point  $i$  has a well-defined deviation  $Y_i(x_1, \dots, x_K)$  for  $(x_1, \dots, x_K) \in \prod_{k=1}^K \mathcal{X}_k$  and interference does not exist.

To simplify the notations for our definition of effect equivalence, we consider  $K = 2$  in this section. We let  $p(y | z_i, x_1, x_2, \psi)$ ,  $p_1(y | z_i, x_1, \psi_1)$  and  $p_2(y | z_i, x_2, \psi_2)$  denote the probability density functions for  $Y_i(x_1, x_2)$ ,  $Y_i(x_1, c_2)$  and  $Y_i(c_1, x_2)$  respectively. The domains of  $p$ ,  $p_1$  and  $p_2$  are  $\mathbb{R}$ , and  $\psi$ ,  $\psi_1$  and  $\psi_2$  are parameter vectors with respective parameter spaces  $\Psi$ ,  $\Psi_1$  and  $\Psi_2$  (examples of which are in Section 3.1). Here,  $c_1 \in \mathcal{X}_1$  and  $c_2 \in \mathcal{X}_2$  represent fixed settings of  $F_1$  and  $F_2$  respectively. These density functions define statistical deviation models for different settings of  $F_1$  and  $F_2$  and will be used to specify likelihood functions for the parameters and total equivalent amounts in our Bayesian method. In terms of the notations in Section 1.1,  $f_A$  can be thought of as the expected deviation profile model derived from one of the statistical deviation models  $p_1$ , and  $f_B$  as the expected deviation profile model derived from the other statistical model  $p_2$ .

2.3. *Effect equivalence with respect to the mean.* Definition 2.1 formalizes effect equivalence with respect to the mean of a deviation model.

DEFINITION 2.1. Factors  $F_1$  and  $F_2$  are equivalent with respect to the mean if, for any point  $i$  and  $c_1 \in \mathcal{X}_1$ ,  $c_2 \in \mathcal{X}_2$ , functions  $T_{i,1 \rightarrow 2} : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{X}_2$  and  $T_{i,2 \rightarrow 1} : \mathcal{X}_1 \times \mathcal{X}_2 \rightarrow \mathcal{X}_1$  exist such that for all  $(x_1, x_2) \in \mathcal{X}_1 \times \mathcal{X}_2$

$$\int_{-\infty}^{\infty} yp(y | z_i, x_1, x_2, \psi) dy = \int_{-\infty}^{\infty} yp_2(y | z_i, T_{i,1 \rightarrow 2}(x_1, x_2), \psi_2) dy,$$

$$\int_{-\infty}^{\infty} yp(y | z_i, x_1, x_2, \psi) dy = \int_{-\infty}^{\infty} yp_1(y | z_i, T_{i,2 \rightarrow 1}(x_1, x_2), \psi_1) dy.$$

Function  $T_{i,k \rightarrow k'}$  is the total equivalent amount of  $F_k$  in terms of  $F_{k'}$  with respect to the mean for point  $i$ . It maps each level combination of  $F_k$  and  $F_{k'}$  to a level of  $F_{k'}$  that generates the equivalent expected outcome under a fixed setting  $c_k$  of  $F_k$ . Our consideration of fixed settings corresponds to AM systems, in which deviations for a process are generated under fixed levels of the lurking variables. Equivalence of  $F_k$  and  $F_{k'}$  is denoted by  $F_k \overset{e}{\sim} F_{k'}$ .

The total equivalent amount in Definition 2.1 enables the transfer of the expectation specified under one process to new processes. For example, if  $F_1 \overset{e}{\sim} F_2$  with respect to the mean, where  $F_1$  is observed and  $F_2$  is lurking, then the mean specification under the process in which  $F_2$  is set at  $c_2$  can be transferred to a process with a different setting of  $F_2$  by incorporating  $T_{i,2 \rightarrow 1}$  into  $p_1$ . In practice, model transfer is facilitated by means of the equivalence of a lurking variable with an observed factor that permits convenient control. Such a factor is referred to as the base factor  $F_1$ , and any  $T_{i,k \rightarrow 1}$  is shortened to  $T_{i,k}$ . Compensation is a standard

base factor for AM processes. As illustrated in our studies, domain knowledge of AM processes can yield equivalencies between their factors and compensation.

2.4. *Bayesian inference and modeling for the total equivalent amount with respect to the mean.* We now address learning of the total equivalent amount of a lurking variable  $F_2$  in terms of compensation  $F_1$  with respect to the mean, which we also refer to as mean effect equivalence. As before, we simplify the exposition by considering two processes characterized by distinct levels  $c_2$  and  $x_2$  of  $F_2$  respectively. We assume that a statistical deviation model, with corresponding probability density function  $p_1(y | z_i, x_1, \psi_1)$ , is specified for the first, and that data is collected on the second with its model yet to be specified. Let  $\psi_1 = (\mu_1, \sigma_1)^T$ , where  $\mu_1$  consists of all parameters in the mean under  $p_1$  and  $\sigma_1$  consists of the remaining parameters in  $\psi_1$ . A model for the second process is then  $p_1(y | z_i, T_{i,2}(x_1, x_2), \mu_1, \sigma_2)$ . The  $T_{i,2}$  remain to be inferred and modeled to complete the model transfer.

Let  $\mathbf{x}_1^{\text{obs}} = (x_{1,1}, \dots, x_{N,1})^T$  be the compensations for  $N$  points under the second process, and  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) = (T_{1,2}(x_{1,1}, x_2), \dots, T_{N,2}(x_{N,1}, x_2))^T$  the corresponding vector of their realized total equivalent amounts of the lurking variable in terms of compensation. We infer  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  using the Bayesian calculation of its posterior distribution. The likelihood function for  $\mu_1, \sigma_2$  and  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  follows from  $p_1$  and is denoted by  $L(\mu_1, \sigma_2, \mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) | \mathbf{z}, \mathbf{y}^{\text{obs}})$ , where  $\mathbf{y}^{\text{obs}}$  is the vector of outcomes and  $\mathbf{z}$  is the matrix of covariates for all the points. For a prior  $p(\mu_1, \sigma_2, \mathbf{T}_2(\mathbf{x}_1^{\text{obs}}))$ , the joint posterior is

$$p(\mu_1, \sigma_2, \mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) | \mathbf{z}, \mathbf{y}^{\text{obs}}) \propto L(\mu_1, \sigma_2, \mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) | \mathbf{z}, \mathbf{y}^{\text{obs}}) \times p(\mu_1, \sigma_2, \mathbf{T}_2(\mathbf{x}_1^{\text{obs}})).$$

The marginal posterior of  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  then follows by integration. It is important to incorporate data from the first process to improve the precision of our inferences. Letting  $\mathbf{D}$  denote the deviations, covariates and compensations for the first process, the posterior  $p(\mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) | \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{D})$  is similarly calculated as above. Our Bayesian approach thus enables us to leverage all of the data collected across distinct settings of a lurking variable in a straightforward manner to learn about its total equivalent amounts.

In practice, prior information on  $\mu_1$  can be elicited independently of that on  $\sigma_2$  and  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$ , because they correspond to distinct processes. Our effect equivalence framework facilitates the elicitation of appropriate priors on  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  that incorporate domain knowledge of AM processes with controlled subjectivity to enable reasonable inferences. Specifically, as compensation is readily interpretable, and experience can be acquired to understand its effect on deviation, an informative prior for  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  can be specified and justified in a straightforward manner. This is demonstrated in Section 3.2.

The posterior of  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  can be computed by the blocked Gibbs sampler that draws from the conditional posteriors  $p(\mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mu_1, \sigma_2)$ ,  $p(\mu_1 \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \sigma_2, \mathbf{T}_2(\mathbf{x}_1^{\text{obs}}))$  and  $p(\sigma_2 \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{T}_2(\mathbf{x}_1^{\text{obs}}), \mu_1)$ . If the deviations are conditionally independent given the covariates, total equivalent amounts and parameters, and if the total equivalent amounts are independent a priori, then drawing from  $p(\mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mu_1, \sigma_2)$  in the Gibbs sampler reduces to independently drawing each entry in  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  from their respective conditional posteriors. Potential computational complexities are thus reduced under this approach, with sampling of the vector of total equivalent amounts simplified into parallel sampling of one-dimensional distributions for the individual entries. This reduction will hold more generally for moderately large  $N$  and low-dimensional model parameter vectors. For very large  $N$  or high-dimensional parameter vectors associated with more complex processes and products [e.g., the 52-dimensional parameter vector for the hierarchical nonlinear regression model of an irregular polygon in [Sabbaghi, Huang and Dasgupta \(2018\)](#)], this approach may not be computationally efficient or feasible.

Such a high-dimensional inferential task can be handled in practice via the construction of a discrepancy measure [[Rubin \(1984\)](#), [Meng \(1994\)](#)] that uses Bayesian inferences on the parameters from the first process to approximate the full inference for the total equivalent amounts of the second. To illustrate, let  $p(\mu_1 \mid \mathbf{D})$  denote the posterior for  $\mu_1$  based only on data  $\mathbf{D}$  from the first process. Sampling from this distribution is less computationally complex compared to sampling from the full posterior of  $\mu_1$  because the former does not involve the unknown realized total equivalent amounts. Define  $e : \mathcal{X}_1 \rightarrow \mathcal{X}_1$  as

$$e(x_1 \mid z, \mu_1) = \int_{-\infty}^{\infty} yp_1(y \mid z, x_1, \mu_1) dy.$$

For each point  $i$  in the second process, we construct the discrepancy measure

$$(2.1) \quad T_i = \operatorname{argmin}_{t \in \mathcal{X}_1} \{y_i^{\text{obs}} - e(t \mid z_i, \tilde{\mu}_1)\}^2$$

to infer their realized total equivalent amounts, where  $\tilde{\mu}_1 \sim p(\mu_1 \mid \mathbf{D})$ . To understand this, note that the calculation of  $p(\mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{D})$  by

$$\int p(\mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{D}, \mu_1) p(\mu_1 \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{D}) d\mu_1$$

can be interpreted via sampling, with a posterior draw of  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  obtained by first drawing  $\tilde{\mu}_1 \sim p(\mu_1 \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{D})$  and then drawing from  $p(\mathbf{T}_2(\mathbf{x}_1^{\text{obs}}) \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{D}, \tilde{\mu}_1)$ . The discrepancy measure instead draws  $\tilde{\mu}_1 \sim p(\mu_1 \mid \mathbf{D})$ , which approximates  $p(\mu_1 \mid \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{D})$ , and then solves for the unknown realized total equivalent amounts as in equation (2.1) given the drawn  $\tilde{\mu}_1$ . This is less computationally complex than sampling from the conditional posterior of  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$

given  $\tilde{\mu}_1$ . Inferences obtained from the discrepancy measure will be similar to those from the full Bayesian calculation when  $p_1$  is a Normal probability density function (which is illustrated in the calibration study of Section 4.2). We recommend this discrepancy measure for practical model transfer in the case of large  $N$  or high-dimensional model parameters.

After the marginal posterior distribution of  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  is obtained, we examine it by means of exploratory data analytic and visualization methods to specify a model  $T_2(z_i, x_1; \beta)$  for the total equivalent amounts of points under the second process as a function of their covariates and compensations, with  $\beta$  denoting a parameter vector. Further elaborations and illustrations of these examinations are provided in our analyses in Sections 3.3 and 4.3. Model transfer will then be complete upon incorporating  $T_2(z_i, x_1; \beta)$  in  $p_1$ .

### 3. Deviation model transfer: Real scenario 1.

3.1. *Data and mean effect equivalence for overexposure.* An important case of model transfer in stereolithography involves the lurking variable of overexposure, or the unintended expansion of a shape due to the faulty spread of light beams on its boundary. This factor  $F_2$ , with its levels consisting of positive real-valued functions on  $[0, 2\pi]$ , was present in the study of Huang et al. (2015) on the cylinders in Figure 2(a). Huang et al. [(2015), pages 434, 436] hypothesized the following probability density function for the deviation of a point  $i$  on a cylinder of nominal radius  $r_i^{\text{nom}}$  under no overexposure:

$$(3.1) \quad p_1(y \mid z_i, x_1, \mu_1, \sigma_1) = (2\pi\sigma_1^2)^{-1/2} \exp \left[ -\frac{1}{2\sigma_1^2} \{y - x_1 - \alpha_0(r_i^{\text{nom}} + x_1)^{a_0} - \alpha_1(r_i^{\text{nom}} + x_1)^{a_1} \cos(2\theta_i)\}^2 \right],$$

with  $z_i = (\theta_i, r_i^{\text{nom}})^\top$ ,  $\mu_1 = (\alpha_0, \alpha_1, a_0, a_1)^\top$ ,  $\Psi_1 = \mathbb{R}^4 \times \mathbb{R}_{>0}$  and independent outcomes. In terms of the notation in Section 1.1, this setting corresponds to a previous process  $A$ , and the expected deviation model  $f_{i,A} : \mathbb{R} \times \mathbb{R}_{>0} \rightarrow \mathbb{R}$  for point  $i$  on a cylinder of nominal radius  $r_i^{\text{nom}}$  with compensation  $x_{i,1}$  manufactured under it is derived from equation (3.1) as

$$f_{i,A}(x_{i,1}, r_i^{\text{nom}}) = x_{i,1} + \alpha_0(r_i^{\text{nom}} + x_{i,1})^{a_0} + \alpha_1(r_i^{\text{nom}} + x_{i,1})^{a_1} \cos(2\theta_i).$$

Also, for  $N$  points on a cylinder of nominal radius  $r^{\text{nom}}$  with compensation plan  $\mathbf{x}_1 = (x_{1,1}, \dots, x_{N,1})^\top$ ,

$$f_A(\mathbf{x}_1, r^{\text{nom}}) = (f_{1,A}(x_{1,1}, r^{\text{nom}}), \dots, f_{N,A}(x_{N,1}, r^{\text{nom}}))^\top.$$

This model performed poorly for the data in Figure 2(a) and in validation experiments. Huang et al. [(2015), pages 436, 438] then identified overexposure as the lurking variable. They effectively provided a physical justification for  $F_1 \stackrel{e}{\sim} F_2$  that prespecified  $T_{i,2}(x_1, x_2(\cdot)) = x_1 + x_2(\theta_i)$  under overexposure  $x_2(\cdot) \in \mathcal{X}_2$ . The transferred deviation model for this process  $B$  then follows from the probability density function  $p_1(y | z_i, T_{i,2}(x_1, x_2(\cdot)), \mu_1, \sigma_2)$  with a distinct standard deviation  $\sigma_2$ . The remaining task is to learn  $T_{i,2}$ .

We proceed to infer and model the realized total equivalent amounts of overexposure in terms of compensation with respect to the mean. Following Sabbaghi et al. [(2014), pages 1401, 1411–1413] and Huang et al. [(2015), page 435], we assume independent outcomes conditional on covariates and total equivalent amounts. Approximately a thousand equally spaced points were collected from each cylinder.

3.2. *Inference for the total equivalent amount of overexposure in terms of compensation.* The likelihood function for the realized total equivalent amounts is obtained from equation (3.1). The following prior is specified:

$$\begin{aligned}
 p(\mu_1, \sigma_2^2, \mathbf{T}_2(\mathbf{x}_1^{\text{obs}})) &\propto (\sigma_2^2)^{-4} \exp\left\{-\frac{0.0016}{\sigma_2^2}\right\} \\
 &\times \prod_{i=1}^N \mathbb{I}\{0 \leq T_{i,2}(0, x_2(\cdot)) \leq 0.015\} \\
 &\times \exp\left\{-50\alpha_0^2 - 50\alpha_1^2 - \frac{(a_0 - 1)^2}{8} - \frac{(a_1 - 1)^2}{2}\right\}.
 \end{aligned}$$

This prior is based on previous literature and studies, as we describe below.

First, consider the prior for the total equivalent amounts. We observed from the work of Zhou, Chen and Waltz (2009) and Zhou and Chen (2012) that a product’s size should not impact the spread of light beams on its boundary, and that this impact in terms of the equivalent amount of compensation should be relatively small (in inches). These observations led to our bounds  $0 \leq T_{i,2}(0, x_2(\cdot)) \leq 0.015$  for all points with the upper bound not depending on the nominal radius of the cylinder on which a point resides. We consider the upper bound 0.015 a conservative prior estimate of the greatest possible total equivalent amount of overexposure in terms of compensation under our process. This value is also justified by noting that we are considering cylinders of nominal radii 0.5” to 3”, and that overexposure is at most a single-digit percentage of the 0.5” cylinder’s radius. After this range of overexposure values was elicited, we then specified the corresponding Uniform prior to reflect our uncertainty about which values within it are more likely than others. A different distribution (e.g., a truncated Normal) could also be specified given additional information on the likely overexposure values.

Now consider the prior for  $(\alpha_0, \alpha_1, a_0, a_1, \sigma_2^2)^\top$ . Our specified prior for these parameters was also informed by domain knowledge of our process. For example, the physical reasoning of Huang et al. [(2015), page 436] suggests that the priors of both  $a_0$  and  $a_1$  should be centered at 1, and we accordingly specified dispersed Normal priors for these parameters. Similarly, the priors for  $\alpha_0$  and  $\alpha_1$  are dispersed Normal distributions centered at zero, and the prior for  $\sigma_2^2$  is an Inverse- $\chi^2$  with small degrees of freedom and scale that reflect our prior conception of the level of variation for cylinder deviations.

It is important to recognize that data does not exist on a process with no overexposure in this case, and so our prior specification must necessarily be informative to prevent identifiability issues. In addition, as the deviation model in equation (3.1) is a nonlinear regression, specifying noninformative or improper prior distributions for the total equivalent amounts and model parameters and verifying that the corresponding posteriors are proper is difficult. It is simpler in practice to specify proper and informative priors that guarantee a proper posterior and are straightforward to interpret.

Our blocked Gibbs sampler to compute the posterior distribution proceeds via three steps. Let  $\mu_1^{(t)}$ ,  $\sigma_2^{(t)}$  and  $\mathbf{T}_2^{(t)}(\mathbf{x}_1^{\text{obs}})$  denote the draws in iteration  $t$ . First, a draw of  $\sigma_2^{(t)}$  from  $p(\sigma_2 | \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{T}_2^{(t-1)}(\mathbf{x}_1^{\text{obs}}), \mu_1^{(t-1)})$  follows as the square root of a scaled Inverse- $\chi^2$ . Second, a draw of  $\mu_1^{(t)}$  from  $p(\mu_1 | \mathbf{z}, \mathbf{y}^{\text{obs}}, \mathbf{T}_2^{(t-1)}(\mathbf{x}_1^{\text{obs}}), \sigma_2^{(t)})$  is obtained using a Metropolis random walk. Finally, each entry in  $\mathbf{T}_2^{(t)}(\mathbf{x}_1^{\text{obs}})$  is independently drawn from their respective posteriors conditional on  $\mu_1^{(t)}$  and  $\sigma_2^{(t)}$  using a Metropolis random walk. This Gibbs sampler was implemented with 10,000 draws obtained after a burn-in of 200,000. Convergence was verified by trace and autocorrelation plots and the Gelman and Rubin (1992) statistic of the evaluated log posterior.

3.3. *Model for the total equivalent amount of overexposure in terms of compensation.* We examine the posterior of  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  with a visualization in Figure 4 of the points' posterior means stratified according to their nominal radii. The high-frequency oscillations for the 1", 2" and 3" cylinders are an artifact of the lower resolution of the process for them. More importantly, we observe that the posterior means lie in a small range. This suggests that the total equivalent amount does not depend on  $\theta_i$  or  $r_i^{\text{nom}}$ . It also explains why the corresponding assumption in Huang et al. (2015) provided a good fit.

Thus, from this exploratory visualization of the posterior of  $\mathbf{T}_2(\mathbf{x}_1^{\text{obs}})$  we model the total equivalent amount as

$$(3.2) \quad T_2(z_i, x_1; \beta_0) = \beta_0 + x_1,$$

with the additive nature of this model prespecified as before. Parameter  $\beta_0$  corresponds to  $x_0$  in the model of Huang et al. (2015), page 438. The corresponding

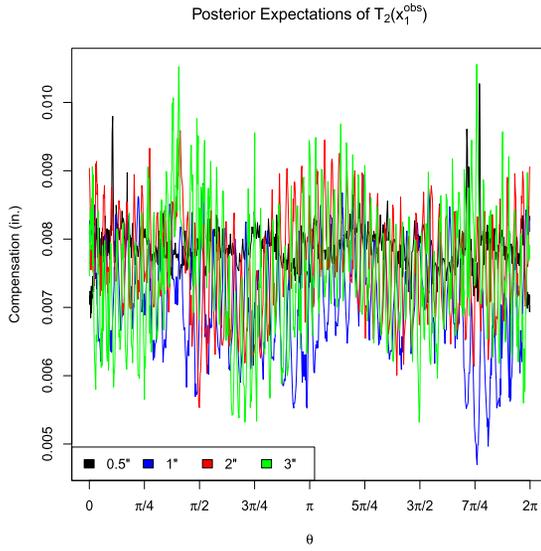


FIG. 4. Posterior expectations of the realized total equivalent amounts of overexposure in terms of compensation with respect to the mean.

transferred model is defined by the probability density function

$$\begin{aligned}
 (3.3) \quad p(y \mid z_i, x_1, \mu_2, \sigma_2) &= (2\pi\sigma_2^2)^{-1/2} \exp \left[ -\frac{1}{2\sigma_2^2} \left\{ y - \beta_0 - x_1 \right. \right. \\
 &\quad \left. \left. - \alpha_0(r_i^{\text{nom}} + \beta_0 + x_1)^{a_0} \right. \right. \\
 &\quad \left. \left. - \alpha_1(r_i^{\text{nom}} + \beta_0 + x_1)^{a_1} \cos(2\theta_i) \right\}^2 \right],
 \end{aligned}$$

with  $\mu_2 = (\alpha_0, \alpha_1, a_0, a_1, \beta_0)^T \in \mathbb{R}^4 \times \mathbb{R}_{>0}$ . Note that for points  $i = 1, \dots, N$ ,

$$\begin{aligned}
 f_{i,B}(x_{i,1}, r_i^{\text{nom}}) &= f_{i,A}(\beta_0 + x_{i,1}, r_i^{\text{nom}}) \\
 &= \beta_0 + x_{i,1} + \alpha_0(r_i^{\text{nom}} + \beta_0 + x_{i,1})^{a_0} \\
 &\quad + \alpha_1(r_i^{\text{nom}} + \beta_0 + x_{i,1})^{a_1} \cos(2\theta_i), \\
 f_B(\mathbf{x}_1, r^{\text{nom}}) &= (f_{1,B}(x_{1,1}, r^{\text{nom}}), \dots, f_{N,B}(x_{N,1}, r^{\text{nom}}))^T.
 \end{aligned}$$

The validation experiments of Huang et al. [(2015), pages 437, 439] effectively illustrate the utility of this model for deviation control under overexposure. As  $T_{i,2}(x_1, x_2(\cdot)) = x_1 + x_2(\theta_i)$  was prespecified and  $x_1$  is known, we have explicitly inferred and derived the fixed, unknown overexposure.

**4. Deviation model transfer: Real scenario 2.**

4.1. *Data and mean effect equivalence for calibration.* An important case of model transfer for general types of AM processes involves the lurking variable of calibration, that is, unknown hardware settings, which is denoted by  $F_3$ . In general  $F_3 \stackrel{e}{\sim} F_1$ , because a change to the process calibration, which yields a new process  $B$ , can be viewed as equivalent to a compensation plan under a previous calibration setting  $A$ . This effect equivalence for the processes  $A$  and  $B$  in Section 1.1 can be justified by the physics of stereolithography.

In this section, we formally infer and model the total equivalent amount of calibration in terms of compensation with respect to the mean. We incorporate the data and model for process  $A$  specified in equation (3.3), which operated with overexposure and the original calibration setting, for this inference, and transfer the model to process  $B$ . Approximately a thousand equally spaced points were collected from each of the cylinders under  $B$ .

4.2. *Inference for the total equivalent amount of calibration in terms of compensation.* Our prior distribution for  $(\mu_2, \sigma_2^2, \sigma_3^2, \mathbf{T}_3(\mathbf{x}_1^{\text{obs}}))^\top$  is

$$p(\mu_2, \sigma_2^2, \sigma_3^2, \mathbf{T}_3(\mathbf{x}_1^{\text{obs}})) \propto \sigma_2^{-2} \sigma_3^{-2} \beta_0^{-1} \prod_{i=1}^N \mathbb{I} \left\{ -\frac{r_i^{\text{nom}}}{10} \leq T_{i,3}(0, x_3) \leq \frac{r_i^{\text{nom}}}{10} \right\} \\ \times \exp \left\{ -\frac{(a_0 - 1)^2}{8} - \frac{(a_1 - 1)^2}{2} - \frac{(\log \beta_0)^2}{2} \right\}.$$

We specify noninformative priors for  $\alpha_0, \alpha_1, \sigma_2^2$  and  $\sigma_3^2$  because our incorporation of data from the previous process eliminates possible identifiability issues. Our priors for the  $T_{i,3}(0, x_3)$  now depend on the nominal radius  $r_i^{\text{nom}}$  of the cylinder on which a point resides because our prior knowledge of process calibration does not preclude this dependence. The lower and upper bounds are our prior estimates of their possible magnitudes.

We calculate the posterior of  $\mathbf{T}_3(\mathbf{x}_1^{\text{obs}})$  based on all seven cylinders by blocked Gibbs sampling and summarize it in Figure 5 with visualizations that stratify it according to the nominal radii and halves of the cylinders. We also calculate the expectations of the discrepancy measure defined in Section 2.4, where  $\mathbf{D}$  is the data on cylinders under overexposure, and

$$e(t \mid z_i, \mu_2) = \beta_0 + t + \alpha_0(r_i^{\text{nom}} + \beta_0 + t)^{a_0} + \alpha_1(r_i^{\text{nom}} + \beta_0 + t)^{a_1} \cos(2\theta_i).$$

The posterior  $p(\mu_2 \mid \mathbf{D})$  has been calculated from the previous process. The discrepancy measure  $T_i$  for point  $i$  under the new process is

$$T_i = \operatorname{argmin}_{t \in \mathcal{X}_1} \{ y_i^{\text{obs}} - \tilde{\beta}_0 - t - \tilde{\alpha}_0(r_i^{\text{nom}} + \tilde{\beta}_0 + t)^{\tilde{a}_0} \\ - \tilde{\alpha}_1(r_i^{\text{nom}} + \tilde{\beta}_0 + t)^{\tilde{a}_1} \cos(2\theta_i) \}^2.$$

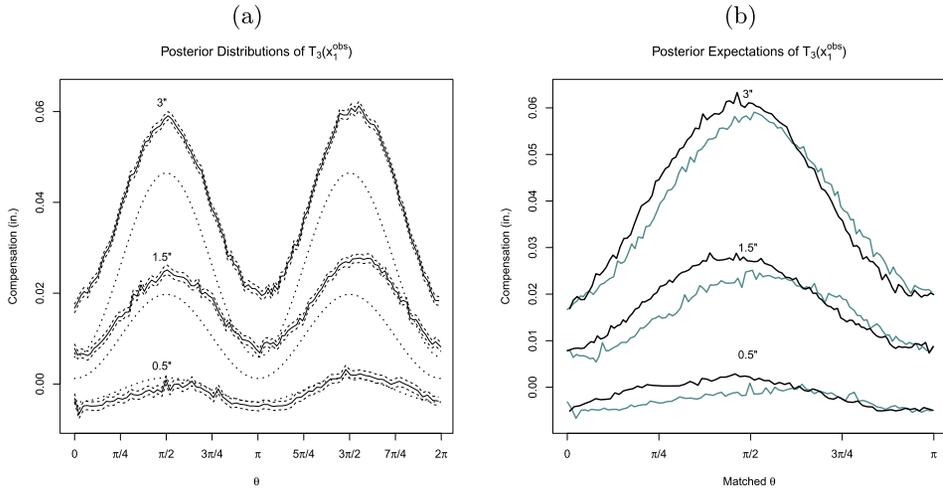


FIG. 5. (a) *Posterior expectations (solid) and 95% central credible intervals (dashed) of the total equivalent amounts of calibration in terms of compensation with respect to the mean. The compensation plans from Huang et al. (2015) are dotted lines.* (b) *Posterior expectations of the total equivalent amounts for the upper (gray) and lower (black) cylinder halves, where the angles on the lower halves are matched to those on the upper halves directly above them.*

Our visual comparison of the results obtained from the blocked Gibbs sampler and the discrepancy measure in Figure 6 demonstrates that they yield identical inferences. In either case they enable the specification of a total equivalent amount model, as described next.

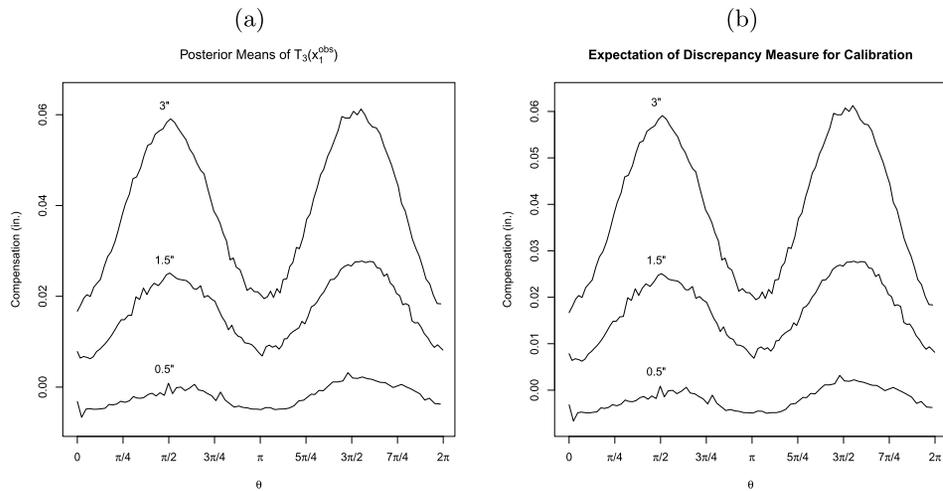


FIG. 6. (a) *Posterior means of the total equivalent amounts of calibration in terms of compensation.* (b) *Expectations of the discrepancy measure.*

4.3. *Model for the total equivalent amount of calibration in terms of compensation.* We observe in Figure 5(a) that the posterior trends of the realized total equivalent amounts are similar to the optimum compensation plans of Huang et al. (2015) but exhibit discrepancies that explain the complicated deviations. For example, these discrepancies explain the apparent underexposure of  $B$  compared to  $A$ . Figure 5(b) highlights the asymmetry in the realized total equivalent amounts between the lower and upper halves of the cylinders, which explains the asymmetrical profiles in Figure 2(b).

We incorporate our observations from the exploratory visualizations of the posterior of the total equivalent amounts against angles, nominal radii and cylinder halves to model the total equivalent amount of calibration as

$$\begin{aligned}
 T_3(z_i, x_1; \beta) = & T_2(z_i, x_1; \beta_0) \\
 & + \mathbb{I}(0 \leq \theta < \pi)[x_{0,U} + \beta_{0,U}(r_i^{\text{nom}} + x_{0,U})^{b_{0,U}} \\
 & + \beta_{1,U}(r_i^{\text{nom}} + x_{0,U})^{b_{1,U}} \cos\{2(\theta_i - \psi_U)\}] \\
 & + \{1 - \mathbb{I}(0 \leq \theta < \pi)\}[x_{0,L} + \beta_{0,L}(r_i^{\text{nom}} + x_{0,L})^{b_{0,L}} \\
 & + \beta_{1,L}(r_i^{\text{nom}} + x_{0,L})^{b_{1,L}} \cos\{2(\theta_i - \psi_L)\}],
 \end{aligned}$$

where  $T_2(z_i, x_1; \beta_0)$  is defined in equation (3.2), and  $\beta$  is a vector of length 12 containing all parameters in the above equation. Flat priors are placed on  $\beta_{0,U}, \beta_{1,U}, \beta_{0,L}, \beta_{1,L}$ , and the priors for the other parameters are

$$\begin{aligned}
 b_{0,U}, b_{0,L} & \sim N(1, 2^2), & b_{1,U}, b_{1,L} & \sim N(1, 1^2), \\
 \log\left(\frac{0.5 + x_{0,U}}{0.5 - x_{0,U}}\right), \log\left(\frac{0.5 + x_{0,L}}{0.5 - x_{0,L}}\right) & \sim N(0, 1^2), \\
 \log\left(\frac{\psi_U/\pi}{1 - \psi_U/\pi}\right), \log\left(\frac{\psi_L/\pi}{1 - \psi_L/\pi}\right) & \sim N(0, (2^{1/2})^2).
 \end{aligned}$$

We assume all parameters in  $\beta$  are mutually independent a priori.

The transferred deviation model was fit using Hamiltonian Monte Carlo [Duane et al. (1987)], which is a Markov chain Monte Carlo algorithm for sampling from a distribution based on Hamiltonian dynamics [Neal (2011)]. Figure 7 demonstrates that this transferred model provides a good fit for the new calibration setting. Thus, although the change in calibration introduced strikingly different and complicated deviation profiles, we are able to transfer the model from process  $A$  to  $B$  in a simple manner.

**5. Concluding remarks.** We developed a novel framework for deviation model transfer across distinct AM processes characterized by different settings of lurking variables in an AM system. A key component of our framework is the

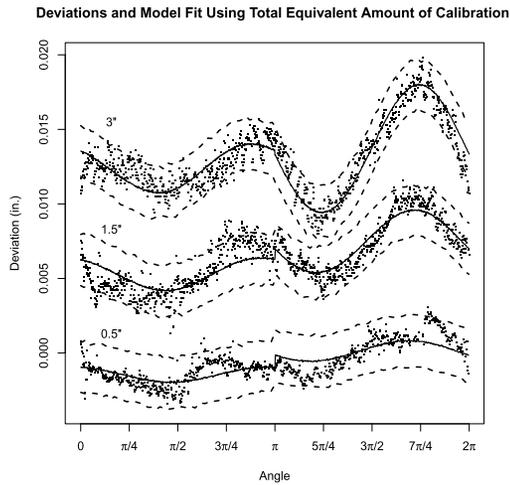


FIG. 7. Posterior predictive means (solid) and 95% central posterior predictive intervals (dashed) of shape deviations under process B obtained from the transferred deviation model.

total equivalent amount of a lurking variable in terms of a base factor with respect to the process mean. We described a Bayesian method for learning the total equivalent amount in terms of compensation for a new process. Inferences on total equivalent amounts facilitate one's understanding of changes in the lurking variables. Once the total equivalent amount is modeled, it can be directly incorporated into a previously learned process' deviation model to effectively transfer it to the new condition. Our two real studies illustrate deviation model transfer across the three stereolithography process conditions of no overexposure, constant overexposure and the combination of constant overexposure and a new calibration setting. They also demonstrate that, if considerable resources have been expended in the specification of a model for one process, its features can then be transferred in a simple manner, with little further expenditure, to a new process. This addresses the fundamental challenge of deviation modeling in an AM system.

Our general framework can have broader impacts in engineering and statistics. As our framework effectively extends engineering insights across processes, it can permit comprehensive modeling of distinct processes connected in a distributed manufacturing environment. In addition, the construction of a catalog of total equivalent amounts in terms of a base factor can enable insightful statistical conclusions to be made on a system's lurking variables. For example, if it is thought that a system is operating under a new condition, then its inferred total equivalent amount can be compared to the catalog's entries to obtain statistical assessments of the most likely change in its lurking variables. Our framework can also provide a new approach to address the statistical problem of external validity or the generalizability of empirical findings to new settings in the presence of lurking variables.

A host of interesting tasks remain for future study. One is relaxing the assumption of fixed lurking variables to address the case of dynamic lurking variables. Another is to develop automated model transfer algorithms that are applicable to a variety of systems, which can provide further impetus to the growing trend of automation in advanced manufacturing. In line with this is combining our effect equivalence framework with the adaptive Bayesian modeling approach in Sabbaghi, Huang and Dasgupta (2018) to enable deviation model building across different shapes and AM processes. When factors can affect multiple model features, an important issue is how the total equivalent amounts with respect to the different features can be learned simultaneously. One important consideration is the effect of factors on the variability of deviation in which the goal is to identify the levels that minimize it. This and similar considerations are part of our ongoing research on the theory of effect equivalence. Another issue is inference of the total equivalent amounts of multiple lurking variables across processes. For example, in uncontrolled, or observational, studies of AM systems alignment issues or device precision could constitute additional lurking variables, and their effects would be entangled with those of other lurking variables. From a statistical viewpoint it would not be possible to disentangle these effects without further information or data. Disentanglement of their effects could perhaps be accomplished by including data from a designed experiment that involved the alignments or device precision. These issues require new research on the design of experiments and observational studies for learning total equivalent amounts.

**Acknowledgments.** We are grateful to the Editor, Associate Editor and two reviewers, Tirthankar Dasgupta and Matthew Plumlee, for their constructive comments and valuable suggestions that improved this paper.

## REFERENCES

- BAREINBOIM, E. and PEARL, J. (2016). Causal inference and the data-fusion problem. In *Proceedings of the National Academy of Sciences* **113** 7345–7352.
- BOX, G. E. P. (1966). Use and abuse of regression. *Technometrics* **8** 625–629. [MR0203866](#)
- BUCKHOLTZ, B., RAGAI, I. and WANG, L. (2015). Cloud manufacturing: Current trends and future implementations. *J. Manuf. Sci. Eng.* **137** 040902.
- CAMPBELL, T., WILLIAMS, C., IVANOVA, O. and GARRETT, B. (2011). *Could 3D Printing Change the World? Technologies, Potential, and Implications of Additive Manufacturing*. Atlantic Council, Washington, DC.
- COOK, R. D. and CRITCHLEY, F. (2000). Identifying regression outliers and mixtures graphically. *J. Amer. Statist. Assoc.* **95** 781–794. [MR1803878](#)
- DAI, W., YANG, Q., XUE, G.-R. and YU, Y. (2007). Boosting for transfer learning. In *Proceedings of the 24th International Conference on Machine Learning*.
- DUANE, S., KENNEDY, A., PENDLETON, B. J. and ROWETH, D. (1987). Hybrid Monte Carlo. *Phys. Lett. B* **195** 216–222.
- GELMAN, A. and RUBIN, D. B. (1992). Inference from iterative simulation using multiple sequences. *Statist. Sci.* **7** 457–472.
- GERMANY TRADE & INVEST (2014). Industrie 4.0: Smart manufacturing for the future. Online.

- GIBSON, I., ROSEN, D. W. and STUCKER, B. (2009). *Additive Manufacturing Technologies: Rapid Prototyping to Direct Digital Manufacturing*. Springer, New York.
- HUANG, Q., NOURI, H., XU, K., CHEN, Y., SOSINA, S. and DASGUPTA, T. (2014). Statistical predictive modeling and compensation of geometric deviations of 3D printed products. *J. Manuf. Sci. Eng.* **136** 061008.
- HUANG, Q., ZHANG, J., SABBAGHI, A. and DASGUPTA, T. (2015). Optimal offline compensation of shape shrinkage for 3D printing processes. *IIE Trans. Qual. Reliab.* **47** 431–441.
- IMBENS, G. W. and RUBIN, D. B. (2015). *Causal Inference—For Statistics, Social, and Biomedical Sciences: An Introduction*. Cambridge Univ. Press, New York. [MR3309951](#)
- JIN, Y., QIN, S. and HUANG, Q. (2016). Offline predictive control of out-of-plane geometric errors for additive manufacturing. *J. Manuf. Sci. Eng.* **138** 121005.
- JOINER, B. L. (1981). Lurking variables: Some examples. *Amer. Statist.* **35** 227–233.
- LUAN, H. and HUANG, Q. (2017). Prescriptive modeling and compensation of in-plane geometric deviations for 3D printed freeform products. *IEEE Trans. Autom. Sci. Eng.* **14** 73–82.
- MENG, X.-L. (1994). Posterior predictive  $p$ -values. *Ann. Statist.* **22** 1142–1160. [MR1311969](#)
- NEAL, R. M. (2011). MCMC using Hamiltonian dynamics. In *Handbook of Markov Chain Monte Carlo* (S. Brooks, A. Gelman, G. L. Jones and X.-L. Meng, eds.). 113–162. CRC Press, Boca Raton, FL. [MR2858447](#)
- PAN, S. J. and YANG, Q. (2010). A survey on transfer learning. *IEEE Trans. Knowl. Data Eng.* **22** 1345–1359.
- PARDOE, D. and STONE, P. (2010). Boosting for regression transfer. In *Proceedings of the 27th International Conference on Machine Learning*.
- PEARL, J. (1995). Causal diagrams for empirical research. *Biometrika* **82** 669–710. [MR1380809](#)
- PEARL, J. and BAREINBOIM, E. (2014). External validity: From do-calculus to transportability across populations. *Statist. Sci.* **29** 579–595. [MR3300360](#)
- RUBIN, D. B. (1984). Bayesianly justifiable and relevant frequency calculations for the applied statistician. *Ann. Statist.* **12** 1151–1172. [MR0760681](#)
- SABBAGHI, A., HUANG, Q. and DASGUPTA, T. (2018). Bayesian model building from small samples of disparate data for capturing in-plane deviation in additive manufacturing. *Technometrics*. DOI:10.1080/00401706.2017.1391715.
- SABBAGHI, A., DASGUPTA, T., HUANG, Q. and ZHANG, J. (2014). Inference for deformation and interference in 3D printing. *Ann. Appl. Stat.* **8** 1395–1415. [MR3271337](#)
- SHEWHART, W. A. (1931). *Economic Control of Quality of Manufacturing Product*, 1st ed. Van Nostrand Reinhold, New York.
- TONG, K., JOSHI, S. and LEHTIHET, E. A. (2008). Error compensation for fused deposition modeling (FDM) machine by correcting slice files. *Rapid Prototyping J.* **14** 4–14.
- TONG, K., LEHTIHET, E. A. and JOSHI, S. (2003). Parametric error modeling and software error compensation for rapid prototyping. *Rapid Prototyping J.* **9** 301–313.
- WANG, H. and HUANG, Q. (2006). Error cancellation modeling and its application to machining process control. *IIE Trans.* **38** 355–364.
- WANG, H. and HUANG, Q. (2007). Using error equivalence concept to automatically adjust discrete manufacturing processes for dimensional variation control. *J. Manuf. Sci. Eng.* **129** 644–652.
- WANG, H., HUANG, Q. and KATZ, R. (2005). Multi-operational machining processes modeling for sequential root cause identification and measurement reduction. *J. Manuf. Sci. Eng.* **127** 512–521.
- WU, D., ROSEN, D. W., WANG, L. and SCHAEFER, D. (2015). Cloud-based design and manufacturing: A new paradigm in digital manufacturing and design innovation. *Comput.-Aided Design* **59** 1–14.
- YATES, F. and COCHRAN, W. G. (1938). The analysis of groups of experiments. *J. Agric. Sci.* **28** 556–580.
- ZHOU, C. and CHEN, Y. (2012). Additive manufacturing based on optimized mask video projection for improved accuracy and resolution. *J. Manuf. Process.* **14** 107–118.

ZHOU, C., CHEN, Y. and WALTZ, R. A. (2009). Optimized mask image projection for solid freeform fabrication. *J. Manuf. Sci. Eng.* **131** 061004.

DEPARTMENT OF STATISTICS  
PURDUE UNIVERSITY  
150 N. UNIVERSITY STREET  
WEST LAFAYETTE, INDIANA 47907  
USA  
E-MAIL: [sabbaghi@purdue.edu](mailto:sabbaghi@purdue.edu)

DANIEL J. EPSTEIN DEPARTMENT OF  
INDUSTRIAL AND SYSTEMS ENGINEERING  
UNIVERSITY OF SOUTHERN CALIFORNIA  
3715 McCLINTOCK AVENUE  
LOS ANGELES, CALIFORNIA 90089  
USA  
E-MAIL: [qiang.huang@usc.edu](mailto:qiang.huang@usc.edu)