Dynamic Sampling Design for Characterizing Spatiotemporal Processes in Manufacturing

Fine-scale characterization and monitoring of spatiotemporal processes are crucial for high-performance quality control of manufacturing processes, such as ultrasonic metal welding and high-precision machining. However, it is generally expensive to acquire high-resolution spatiotemporal data in manufacturing due to the high cost of the three-dimensional (3D) measurement system or the time-consuming measurement process. In this paper, we develop a novel dynamic sampling design algorithm to cost-effectively characterize spatiotemporal processes in manufacturing. A spatiotemporal state-space model and Kalman filter are used to predictively determine the measurement locations using a criterion considering both the prediction performance and the measurement cost.

The determination of measurement locations is formulated as a binary integer programming problem, and genetic algorithm (GA) is applied for searching the optimal design. In addition, a new test statistic is proposed to monitor and update the surface progression rate. Both simulated and real-world spatiotemporal data are used to demonstrate the effectiveness of the proposed method. [DOI: 10.1115/1.4036347]

Keywords: spatiotemporal processes, dynamic sampling design, intelligent sensing, manufacturing

1 Introduction

Spatiotemporal processes exist widely in manufacturing. For example, ultrasonic metal welding is adopted for battery joining in electric vehicle manufacturing [1–4]. The surface geometry of the welding tools progresses over time [2,3,5], as illustrated by Fig. 1. In the machining of automobile engine components, machined surfaces exhibit varying spatial variation patterns as cutters wear out [6]. Fine-scale characterization of spatiotemporal processes is essential for process monitoring and control in manufacturing, because the spatiotemporal evolution could significantly affect the process dynamics or directly reflect the product quality. As an example, in ultrasonic metal welding of batteries for electric vehicle manufacturing, tool surface degradation critically affects the process dynamics, e.g., the vibration frequency [2,3], and is an important root cause for low-quality joints in production. In the example of high-precision machining processes for automotive engine components, surface quality is of great importance for assembly performance and the overall engine quality [7,8]. The spatial variations in both examples evolve over time and fall into the category of spatiotemporal processes. High-resolution characterization of such processes is crucial for both understanding the underlying process mechanism and enabling high-performance monitoring and control.

Critical challenges exist in efficient and effective characterization of spatiotemporal process in manufacturing. First, high-resolution measurement data are expensive and time-consuming to acquire. For instance, a 3D microscope is used by an electric vehicle manufacturer for tool wear monitoring of ultrasonic metal welding [2,3,5]. It takes the microscope approximately 8 h to measure a single tool surface with the dimension of 43 mm x 8 mm. The low measurement efficiency not only leads to a large amount of production downtime but also greatly delays the decision-making in tool condition monitoring. Hence, the high cost has limited the usage of high-resolution surface measurement systems for online process monitoring.

Spatial interpolation is one possible solution to reducing the cost induced by high-resolution measurement systems [7–12]. In spatial statistics, interpolation refers to the procedure of estimating the value of properties, e.g., surface height, at unsampled locations using measurement data from their vicinities. Examples of spatial interpolation techniques include kriging, cokriging, and regression kriging. The accuracy and precision of interpolation largely depend on the number of sampling/measurement sites and the distribution of measurement points. Therefore, sampling design is critically needed to achieve satisfactory interpolation results.

Sampling design has received extensive attention in the fields of natural resources and environmental science [13–20] and is recently investigated in manufacturing [21]. A sampling design problem can be generally formulated as follows. Assume the location set of interest is denoted by $\mathcal{S}$, and the measured location set is $S_m = \{s_1, \ldots, s_n\} \subset \mathcal{S}$. The goal is to adaptively select $S_m$ such that the inference at unobserved locations, denoted as $S_o = \mathcal{S} \setminus S_m$, is optimized according to some design criterion. Prediction error variance has been popularly used as a performance measure of one sampling design. Suppose the prediction error variance is denoted by $V$, then the design target can take the form of the following equation:

$$\min_{S_m \subset \mathcal{S}} f(V)$$

where $f(\cdot)$ is some function of the prediction error variance. Examples of $f(\cdot)$ include minimum, maximum, and average.

In general, sampling design approaches can be divided into two main categories according to whether temporal dependence is considered: (1) static sampling design and (2) dynamic sampling design. Static sampling design is often adopted when one is interested in characterizing a spatial process at a fixed time point (e.g., Refs. [13–16] and [22]) or the temporal dependence of spatial variations is negligible (e.g., Ref. [21]). Furthermore, based on the extent to which spatial variability is taken into consideration, static sampling design methods may be classified into random
sampling, spatial variability-based method, and sequential design approaches based on computer experiment methods. For a detailed comparison of three types of methods, refer to Ref. [6].

Static sampling design methods are applicable for a single spatial process or multiple processes with high similarities, but cannot be directly applied to the sampling design problem of spatiotemporal processes. As a matter of fact, there is relatively limited work on sampling designs for spatiotemporal processes [23]. For a spatiotemporal process, the spatial structure varies over time, and therefore the optimal design at an earlier time may not always be optimal in the future. In order to cost-effectively characterize a spatiotemporal process, the sampling design is supposed to be determined according to the time-variant spatial variation pattern. Those approaches are generally referred to as “dynamic sampling design.” Wikle and Royle [24] is one of the earliest studies on this topic. In this study, a space–time dynamic model is used to describe spatiotemporal processes, and Kalman filter is applied to obtain the prediction covariance matrix conditional on the measurements at the current time and all historical data in a recursive manner. Then, measurement locations are determined by minimizing some design criterion based on the prediction error variance. Following this study, dynamic sampling design methods have been developed for ecological monitoring [25,26] and other applications of environmental monitoring [23,24] and ecological monitoring [25,26]. However, there is a lack of studies on spatiotemporal processes in manufacturing. Compared with environmental or ecological monitoring, manufacturing processes have some unique characteristics making the design problem more complicated. Specifically,

(1) Multiple measurement systems with varied resolutions and precisions are often jointly used in manufacturing. For instance, in tool condition monitoring for ultrasonic metal welding, a microscope has various lenses with different measurement ranges and resolutions. In surface quality control for high-precision machining of automotive engine components, a coordinate measuring machine (CMM) and a profilometer are adopted to measure a path following the edges of a rectangle and several rectangular regions, respectively [7]. Therefore, the dynamic sampling design should take the measurement capability of all the gages into consideration.

(2) Existing methods have seldom considered the measurement cost during the design process. In ecological and environmental monitoring, the total number of measurement locations or monitoring devices is often determined based on budget constraints [25]. However, in manufacturing, it is desirable to achieve a balance between measurement cost and prediction precision.

(3) An algorithm for monitoring and updating the model parameters is necessary. Model parameters of spatiotemporal processes in manufacturing are generally time-variant. For example, in a process where tool wear is the major cause for the spatiotemporal variation, the tool degradation rate is often time-variant. It is critical to monitor and update the model parameters, as inaccurate parameters will likely result in nonoptimal designs and inaccurate interpolation.

To address the aforementioned challenges, this study develops a new dynamic sampling design approach to cost-effectively characterize spatiotemporal processes in manufacturing. The remainder of this paper is organized as follows. Section 2 introduces a formulation of the dynamic sampling design problem, a state-space spatiotemporal model, and recursive prediction error estimation based on Kalman filter. Then, Sec. 3 presents an innovative dynamic sampling design approach which consists of a new design criterion, a GA-based searching method, and a novel algorithm for estimating and monitoring the temporal progression rate. In Sec. 4, a simulation study and a case study using spatiotemporal data collected from an ultrasonic metal welding process are presented to demonstrate the effectiveness of our method. Finally, Sec. 5 concludes the paper.

2 Problem Formulation

This section formulates the dynamic sampling design problem for spatiotemporal processes using a state-space spatiotemporal model and Kalman filter.

2.1 Spatiotemporal Model. Various spatiotemporal models have been developed over the recent years, among which a state-space approach has attracted considerable attention [24–27]. One advantage of this model is that empirical Bayesian or spatiotemporal Kalman filter can be used to conveniently implement such models [24]. We will briefly review the model as follows.

Let \( Z_t = \left( Z(s_1, t), \ldots, Z(s_n, t) \right) \) be an \( n \times 1 \) vector for an unobservable spatiotemporal process at some fixed set of locations \( \mathcal{S} = \{ s_1, \ldots, s_n \} \). In manufacturing applications, \( s \) is normally a two-dimensional coordinate vector, namely, \( s = [x, y] \). Moreover, in manufacturing, the location set \( \mathcal{S} \) of interest is often a constant grid. The spatiotemporal model is given by the following state equation:

\[
Z_t = \Phi_t Z_{t-1} + \eta_t
\]

where \( \Phi_t \) is a first-order Markov parameter matrix, and \( \eta_t \) is a spatiotemporal process with covariance matrix \( \Sigma_{\eta} \). In spatiotemporal processes where the spatial variation is induced by tool degradation, \( \Phi_t \) represents the tool degradation dynamics and can be time-variant.

At a particular time \( t \), a subset of \( \mathcal{S} \) is measured, and the measurement data are given by \( M_t = \left( M(s_1, t), \ldots, M(s_n, t) \right) \), which is an \( m_t \times 1 \) vector. The measurement locations are \( \{ s_1, \ldots, s_{m_t} \} \).

Supposing the measurements are acquired with errors, a measurement equation can be written as the following equation:

\[
M_t = D_t Z_t + \varepsilon_t
\]

where \( M_t \) is a vector of \( m_t \) measurement data; \( D_t \) is an \( m_t \times n \) matrix that maps the true process, \( Z_t \), to the data at observed locations, \( M_t \); and \( \varepsilon_t \) reflects the measurement capability and is a zero-mean measurement error process with covariance matrix \( \Sigma_{\varepsilon} \). \( D_t \) determines which locations in \( Z_t \) are observed and is a sparse matrix of 0’s and 1’s [28]. Consequently, the dynamic sampling design is equivalent to the determination of \( D_t \).

In the context of manufacturing processes, we make the following assumptions regarding models (2) and (3).
Assumption 1 (Parameter Assumptions for the Spatiotemporal Model).

1. The transition matrix takes the form of
\[ \Phi_t = \phi_t I \] (4)
where \( \phi_t \) is the surface progression rate and should be monitored and updated, and \( I \) is an \( n \times n \) identity matrix. In spatiotemporal processes where tooling degradation is the dominant factor, \( \phi_t \) can be viewed as the tooling degradation rate.

2. \( \Sigma_u \) is a diagonal matrix and is fully determined by the gage repeatability and the design \( D_t \).

3. \( \Sigma_y \) is time-invariant, i.e., \( \Sigma_y = \Sigma_y \), and the parameters can be estimated and updated over time.

The rationale behind Assumption 1 is that the global change of the spatial pattern is captured by \( \phi_t \) and the local variation is quantified by the spatiotemporal term \( \eta_t \) in Eq. (2). In manufacturing, \( \phi_t \) represents the dominant effect of the physics factors, e.g., tooling degradation in ultrasonic metal welding or machining. The assumption of Eq. (4) manifests twofold benefits. First, the simple form could enable convenient monitoring and diagnosis based on engineers’ expert knowledge; thus, it is appealing to manufacturing practitioners. Second, compared with the matrix \( \Phi_t \), we only need to estimate one parameter \( \phi_t \) at time \( t \), which significantly reduces the computational complexity. Meanwhile, \( \eta_t \) takes the natural spatiotemporal variation into account. Therefore, the model still has sufficient generosity to characterize spatiotemporal processes.

2.2 Prediction Error Estimation. Given the state-space model (2) and the measurement equation (3), a state-space Kalman filter can be developed either by Bayesian arguments [29] or projection arguments [30]. Then, we can obtain recursive equations for the prediction error covariance, as shown by the following equations:

\[ A_t \equiv \text{Var}(\hat{Z}_t|M_1, ..., M_t) = B_t - B_t D_t^{-1} [D_t D_t^{-1} + \Sigma_y]^{-1} D_t B_t \]

(5)

\[ B_t \equiv \text{Var}(\hat{Z}_t|M_{t-1}, ..., M_t) = \Phi_t A_{t-1} \Phi_t^T + \Sigma_y \]

(6)

To start the recursion, \( A_0 \) needs to be specified and is typically chosen to be the unconditional variance–covariance matrix of the spatiotemporal process [34]. Moreover, by Assumption 1, \( \Sigma_y \) is constant and will be updated after each sampling; the calculation of \( B_t \) can be further simplified as the following equation:

\[ B_t = \phi_t^2 A_{t-1} + \Sigma_y \]

(7)

3 Dynamic Sampling Design

This section presents a dynamic sampling design approach. Particularly, the design problem is first formulated; a new design criterion simultaneously considering the prediction precision and the measurement cost is then presented; finally, a binary integer programming problem is formulated and GA is proposed for the solution search.

3.1 Measurement System Setting. For the measurement and monitoring of spatial variations in manufacturing, various measurement devices are available, including CMMs, profilometers, 3D microscopes, laser holographic interferometers, and touching probe sensors. Simultaneous adoption of multiple gages with different measurement resolutions is a common practice. For instance, in high-precision machining of automobile engine components, a CMM and a profilometer are used in combination to monitor the surface quality; in tool condition monitoring of ultrasonic metal welding, different lenses of a 3D microscope may be used together for surface measurement.

In this study, we assume a 3D measurement system with two resolution levels, i.e., level 1 and level 2, is applied to measure a spatiotemporal process. At each sampling time, two levels are used in combination for measurement. It is worth noting that our approach is flexible and applicable to the settings when more levels are used together. Here, we use the example of two levels as an illustration. The following assumptions are made regarding levels 1 and 2.

Assumption 2 (Measurement System Setting).

1. Level 1 is able to measure the whole surface with one-time scanning but it has a relatively poor precision, which is reflected by a large measurement error variance, or a poor repeatability. We denote the measurement error variance of level 1 as \( \sigma_1^2 \).

2. The measurement locations of level 1 are fixed.

3. Level 2 has a smaller measurement error variance, denoted as \( \sigma_2^2 \), but the measurement range is small. We have

\[ \sigma_1^2 > \sigma_2^2 \] (8)

4. The measurement locations of level 2 are programmable.

5. When designing the sampling locations for level 2 measurement, a surface needs to be segmented into a number of grids, and each grid is measured by one single scan. The “grid segmentation” process will be explained later. The design problem is then equivalent to selecting a subset of the predefined grids.

6. The unit measurement costs of level 1 and level 2 are assumed to be \( c_1 \) and \( c_2 \), respectively. Level 2 measurement is more expensive to acquire, so we have

\[ c_1 < c_2 \] (9)

Figure 2 illustrates the process of grid segmentation. The points in the prediction location set \( \mathcal{S} \) are assumed to align with an equally spaced grid, the size of which is \( N_x \times N_y \), as represented by the dots in Fig. 2. The measurement range of level 2 is \( n_{d,x} \times n_{d,y} \), indicating that one single scan of level 2 can measure \( n_{d,x} \times n_{d,y} \) points. Then, the total number of grids is calculated as

\[ N_{\bar{k}} = \left[ \frac{N_x}{n_{d,x}} \right] \times \left[ \frac{N_y}{n_{d,y}} \right] \] (10)

where \( N_{\bar{k}} \) is the total number of grids, and \( \left[ \cdot \right] \) is a ceiling function. In the example illustrated by Fig. 2, \( N_x = N_y = 11, n_{d,x} = n_{d,y} = 2 \), so the total number of grids is \( N_{\bar{k}} = \left[ \frac{11}{2} \right] \times \left[ \frac{11}{2} \right] = 36 \), and the grids are indexed as 1, 2, ..., 36.

The sampling design is to select a certain number of grids from the \( N_{\bar{k}} \) grids. The design target is to determine (1) the number of measurement grids \( (n_{\bar{k}}) \) and (2) where to allocate these grids. The number of all possible designs, \( N_{\text{design}} \), is calculated by the following equation:

\[ N_{\text{design}} = \sum_{n_{\bar{k}=0}}^{N_{\bar{k}}} {N_{\bar{k}} \choose n_{\bar{k}}} = 2^{N_{\bar{k}}} \] (11)

For the example in Fig. 2, the number of all possible designs is \( 2^{36} \approx 687 \times 10^{9} \).

The design matrix, \( D_t \), can be rewritten as

\[ D_t = \begin{bmatrix} D_1 \\ D_2 \end{bmatrix} \] (12)
where $D_1$ is time-invariant, but $D_{2,t}$ is time-variable and is designed predictively.

The number of measurement points of level 1 and level 2 can be calculated as

$$m_1 = \text{trace}(D_1 D_1^T)$$

$$m_{2,t} = \text{trace}(D_{2,t} D_{2,t}^T)$$

(13)

where $m_1$ and $m_{2,t}$ are the number of measurement points of levels 1 and 2, respectively.

Then, the total measurement cost, $C_t$, is given by the equation below

$$C_t = c_1 m_1 + c_2 m_{2,t}$$

(14)

With the representation of Eq. (12), the measurement equation, i.e., Eq. (3), is rearranged, and we obtain the following equation:

$$M_t = \begin{bmatrix} M_{1,t} \\ M_{2,t} \end{bmatrix}$$

(15)

where $M_{1,t}$ and $M_{2,t}$ are the measurements of level 1 and level 2, respectively. $M_{1,t}$ and $M_{2,t}$ can be obtained using

$$M_{1,t} = D_1 Z_t$$

(16)

$$M_{2,t} = D_{2,t} Z_t$$

(17)

Accordingly, the covariance matrix of the measurement error process, $\varepsilon_t$, can be estimated with the following equation:

$$\Sigma_n = \begin{bmatrix} \sigma^2 \mathbf{I}_{m_1 \times m_1} & 0_{m_1 \times m_{2,t}} \\ 0_{m_{2,t} \times m_1} & \sigma^2 \mathbf{I}_{m_{2,t} \times m_{2,t}} \end{bmatrix}$$

(18)

where $\mathbf{I}_{m_1 \times m_1}$ and $\mathbf{I}_{m_{2,t} \times m_{2,t}}$ are identity matrices, and $0_{m_1 \times m_{2,t}}$ and $0_{m_{2,t} \times m_1}$ are zero matrices.

Note: if some locations are measured by both level 1 and level 2, we will only use the measurement from level 2 due to its superior precision. Meanwhile, the specific rows corresponding to these locations will be removed from $D_1$, and we denote the modified design matrix as $D_{1,t}$. Moreover, the formulation in the equations above will be revised accordingly.

### 3.2 Design Criterion

The design criterion can be constructed using some function of the prediction covariance matrix, $A_t$, such as average or maximum prediction error variance [24].

Then, the design can be conducted using the following two-step procedure.

**DESIGN PROCEDURE**

**Step 1:** Calculate $B_t$ based on $A_{t-1}$.

**Step 2:** Minimize the prespecified design criterion over all possible designs $\mathcal{D}$.

Average or maximum prediction variance is easy to implement in practice; however, there are two major limitations in using these criteria. First, the measurement cost cannot be incorporated in the design process. Although it may be reasonable in environmental and ecological monitoring to predetermine the total number of observations according to budget constraints, it is not desirable to use such criteria in manufacturing. Second, these criteria cannot accommodate the scenario where one has varying precision requirement on a surface. For example, in tool condition monitoring for ultrasonic metal welding, the knurl peaks are of more concern because these areas affect the weld quality more critically than other areas [3].

Here, we propose a new design criterion to simultaneously consider the measurement cost and the “weighted average prediction variance.” The weighted average prediction variance is defined as

$$V_t = \text{trace}(A_t W_t)$$

(19)

where $A_t$ is the prediction covariance matrix and $W_t$ is an $n \times n$ weight matrix, and it is given by the following equation:

$$W_t = \begin{bmatrix} w_1 & 0 & \ldots & 0 \\ 0 & w_2 & \ldots & \ldots \\ \ldots & \ldots & \ldots & \ldots \\ 0 & \ldots & \ldots & w_n \end{bmatrix}$$

(20)

where the diagonal entry of $W_t$, $W_{ii}$, is the weight assigned to the $i$th location in $Z_t$. Higher weight corresponds to stricter precision requirement. If the precision requirement is time-variant, one can instead adopt a varying weight matrix, $W_t$.

Note: Expert knowledge on manufacturing processes should be incorporated to determine the values in the weight matrix. Locations that more significantly affect the product quality are generally assigned higher values. For instance, in ultrasonic metal welding, knurl peaks are directly involved in the bonding formation mechanism [2,3]; hence, it is important to obtain high-precision estimation of such areas. Larger values can be assigned to those locations such that the design search algorithm is guided to yield better prediction precision accordingly.
The design criterion is defined by a loss function, $L_t$, as given by the following equation:

$$L_t = C_t + \lambda V_t$$

where $\lambda$ is a tuning parameter. Larger $\lambda$ places more emphasis on the prediction precision and will generally lead to better precision but higher measurement cost. In practice, $\lambda$ can be determined based on expert knowledge.

### 3.3 GA-Based Binary Integer Programming

The dimension of the design space $\mathcal{D}$ in manufacturing could be extremely high. In the example shown by Fig. 2, there are $687 \times 10^6$ possibilities in selecting measurement grids for level 2. This number will increase exponentially with the number of grids. Exchange algorithms, which have been commonly applied in existing studies on dynamic sampling design, might not be effective in dealing with such a high-dimensional design space. Therefore, a more effective programming technique is needed. In this section, we first formulate the design process as a binary integer programming problem and then use GA to search for the solution.

The design problem is essentially equivalent to determining whether to conduct measurement at a certain grid, which is a binary decision. Assume there are in total $N_g$ grids in the surface of interest. We define an $N_g \times 1$ design vector $d_t$, and each element can only take the values of 0 or 1. A “0” value indicates that there is no measurement at the corresponding grid, while “1” suggests that a level 2 measurement is performed at that grid. Note that the design vector $d_t$ and the level 2 design matrix, $D_{2,t}$, are one-to-one mapped. Additionally, given that the level 1 design matrix, $D_{1}$, is fixed, $d_t$ and $D_{t}$ are also one-to-one mapped. Hence, the dynamic sampling design problem can be formulated as a binary integer programming problem.
There are various techniques available in the literature to solve integer programming problems [31], including exact methods, such as branch-and-bound [32], branch-and-reduce [33], and heuristic methods, which may be based on tabu search [34] and GA [35,36]. This study adopts a GA-based integer programming approach due to its computational efficiency. Furthermore, if more computationally intensive design problems are encountered, e.g., larger surfaces, higher resolution, and larger design space, we provide three suggestions to improve the computational efficiency: (1) adopting sparse matrix representation, (2) parallelizing GA, and (3) using high-performance computing.

3.4 Estimation and Monitoring of the Temporal Transition Parameter. Accurate estimation of the temporal transition parameter, \( \phi_t \), is crucial in dynamic sampling design, as an inaccurate estimation will lead to incorrect evaluation of \( A_t \) and \( B_t \), and the error will accumulate over time, deteriorating the overall design and prediction performance. Therefore, an algorithm for estimating and monitoring \( \phi_t \) is critically needed.

In the dynamic sampling design literature, Hooten et al. [25] suggest using a method of moments estimator for \( \phi_t \) using all historical measurement data. Nevertheless, the estimator is developed based on the assumption that the transition parameter is constant, which may not be a valid assumption in manufacturing. On the monitoring of transition parameters, possible solutions include system identification techniques for state-space models [37] and methods for monitoring autoregressive parameters in time series models [38]. However, these methods cannot be directly applied in dynamic sampling design either because of their high-complexity or large detection delays.

In the rest of this section, we will develop an algorithm to monitor and update the temporal transition parameter, \( \phi_t \), based on hypothesis testing using the measurement data.

We first define the set of locations measured by level 1 to be \( S_1 \), which is time-invariant by assumption. At time \( t \), we denote the set of locations measured by level 2 as \( S_2 \). Then, at time \( t - 1 \), the level 1 and level 2 measurement locations are \( S_1 \) and \( S_{2,t-1} \), respectively. The locations mutually measured by level 2 at both time \( t - 1 \) and \( t \) can be obtained by

\[
S_{2,t(t-1)} = S_{2,t-1} \cap S_{2,t}
\]

where \( S_{2,t(t-1)} \) represents the set of mutually measured locations by level 2 at both time \( t - 1 \) and \( t \). Then, we denote the measurements at locations of \( S_{2,t(t-1)} \) at time \( t - 1 \) and \( t \) to be \( M_{t-1}^{2,t(t-1)} \) and \( M_t^{2,t(t-1)} \), respectively.

The mutual measurements at time \( t - 1 \) and \( t \) can be represented as stacked measurement vectors of level 1 and level 2, i.e.,

\[
\begin{pmatrix}
M_{t-1}^{1} \\
M_t^{2,t(t-1)}
\end{pmatrix} = 
\begin{pmatrix}
M_1 \\
M_{2,t-1}^{2}
\end{pmatrix}
\]

\[
\begin{pmatrix}
M_t^{1} \\
M_t^{2,t(t-1)}
\end{pmatrix} = 
\begin{pmatrix}
M_t \\
M_{2,t(t-1)}
\end{pmatrix}
\]

where \( M_1 \) and \( M_t \) are measurement vectors of level 1 and level 2 at time \( t \), respectively.
where \( M_{t-1} \) and \( M_t \) are the measurement data at mutually measured locations at time \( t-1 \) and \( t \), including both level 1 and level 2 measurements.

Statistics are defined as the average height of the measurements at these mutually locations at time \( t-1 \) and \( t \), as given by the following equation:

\[
\bar{u}_{t-1} = \frac{1}{t-t_0} \sum_{j=t_0}^{t} \frac{u_{t-1,j}}{h_{t-1,j}}
\]

A one-time temporal transition parameter, which is denoted by \( h_t \), is defined as follows:

\[
h_t = \frac{u_t}{h_{t-1}}
\]

When no shift occurs in the temporal transition parameter, \( \phi_t \), it is estimated using the following equation:

\[
\hat{\phi}_t = \frac{1}{t-t_0} \sum_{j=t_0}^{t} \frac{u_{t-1,j}}{h_{t-1,j}} = \frac{1}{t_0} \left[ (t-t_0) \hat{\phi}_{t-1} + h_t \right]
\]

where \( t_0 \) is the time right after the previous shift occurs. When no shift has occurred, \( t_0 = 1 \).

When a shift is detected, we will estimate \( \phi_t \) using the one-time temporal transition parameter, \( h_t \), namely

\[
\hat{\phi}_t = h_t
\]

The monitoring of \( \phi_t \) is conducted using the following hypothesis test.\(^2\)

**SHIFT DETECTION TEST FOR \( \phi_t \).**

Conclude that \( H_0: \phi_t = \hat{\phi}_{t-1} \) is not supported when

\[
\left| \frac{u_{t-1}(\phi_t \hat{\phi}_{t-1})}{\sigma^2_{\phi_t(\phi_t \hat{\phi}_{t-1})}} \right| > Z_{a/2}
\]

where \( Z_{a/2} \) is the standard normal critical value, and

\[
\sigma^2_{\phi_t(\phi_t \hat{\phi}_{t-1})} = \left[ m_1 + m_2 \frac{\hat{\phi}_{t-1}^2}{1 + \hat{\phi}_{t-1}^2} \right] + \frac{m_2}{(1 + \hat{\phi}_{t-1}^2)}
\]

where \( m_1 = \frac{2}{S_{\phi_t(\phi_t \hat{\phi}_{t-1})}} \) and \( m_2 = \frac{2}{S_{\phi_t(\phi_t \hat{\phi}_{t-1})}} \).

The estimation and monitoring procedure is summarized by Fig. 3.

### 4 Case Studies

This section shows two case studies to demonstrate the effectiveness of the proposed method. The first case study uses

\(^2\)Detailed derivation of the hypothesis test is given in the Appendix.
simulated data to compare four candidate methods, and the second case study utilizes real-world spatiotemporal data from an electrical vehicle battery manufacturing plant.

4.1 Simulation Study. A spatiotemporal process is simulated and used to compare different sampling design methods. A summary of the candidate methods is given by Table 1. For the convenience of comparison, when conducting random sampling, we predetermine the number of level 2 measurements as the same with the dynamic sampling design and randomly allocate these points on the surface. Therefore, methods 1 and 3 (same with methods 2 and 4) will always have the same measurement cost.

The experiment setting is summarized as follows:

\[ N_x = N_y = 16; \quad n_{d,x} = n_{d,y} = 2 \]
\[ \sigma_x^2 = 1, \quad \sigma_y^2 = 0.1; \quad c_1 = 1, \quad c_2 = 1.5 \]
\[ \lambda = 1.3; \quad t = 1, 2, \ldots, 10 \]

Additionally, level 1 measures 16 points, the coordinates of which are as given below

\[(1, 1), (1, 6), (1, 11), (1, 16)\]
\[(6, 1), (6, 6), (6, 11), (6, 16)\]
\[(11, 1), (11, 6), (11, 11), (11, 16)\]
\[(1, 16), (16, 6), (16, 11), (16, 16)\]
function, as given in Eq. (33). The surface degradation rate is 0.8. The temporal transition parameter is assumed to be a piecewise function values, and the prediction root mean squared errors (RMSEs) are smaller for method 4, indicating better prediction performance.

4.2 Case Study for Ultrasonic Metal Welding. In this section, a case study is presented to show the effectiveness of our method using spatiotemporal data collected from a real-world ultrasonic metal welding process. Ultrasonic metal welding has been used for joining lithium-ion batteries of electric vehicles [1,4,39–41], e.g., Chevrolet Volt manufactured by General Motors. Surface degradation of anvils is a major concern in this process [2,3,5]. In an electrical vehicle battery manufacturing plant, an anvil was measured six times in the tool life. In this study, a part of the anvil surface is utilized, the spatiotemporal progression of which is shown by Fig. 8.

The knurl peak areas significantly affect the joining quality [3]; thus, higher weights are assigned to such locations. The weight matrix is shown by Fig. 9, where Fig. 9(a) is the contour plot of the first stage, and Fig. 9(b) is the heat map of the weight matrix. The peak areas in the center two knurls are assigned weights of 3, the surrounding peaks have weights of 2, and all other locations have weights of 1. Suppose we are interested in 25 peak areas in the center two knurls are assigned weights of 3, the design space is a 25 lattice grid. Additionally, level 1 has a step size of 6 in both x and y directions. Other assumptions are summarized by the equation set below

\[
\begin{align*}
\eta_{d, t} &= 3; \quad \lambda = 1.3; \quad t = 1, 2, \ldots, 6 \\
\sigma_1^2 &= 1; \quad \sigma_2^2 = 0.1; \quad c_1 = 1; \quad c_2 = 5
\end{align*}
\]

In this case study, we compare the random sampling and dynamic sampling approaches, and both methods are equipped
with a monitoring algorithm to update $\phi_t$. Since random sampling cannot automatically determine the number of measurement points, we use the numbers from dynamic sampling for random sampling. Consequently, two methods have exactly the same measurement costs. The temporal trend of the precision index is displayed by Fig. 10. Clearly, the dynamic sampling design method has better prediction performance than the random sampling approach.

The heat maps depicting the spatial distributions of level 2 measurement frequency and average prediction variances are given by Figs. 11 and 12, respectively. In both figures, the variable magnitude is represented by the color. More specifically, dark corresponds to small values and light indicates large values. Figure 11 shows that while the random sampling allocates measurement efforts randomly, our method systematically distributes the measurements and places more samples at high-weight areas. Moreover, Fig. 12 indicates that the dynamic sampling method is able to achieve overall smaller prediction variances. Particularly, high precision is obtained at the high-weight regions.

5 Conclusion

In this paper, a generic framework for dynamic sampling design is developed to cost-effectively characterize spatiotemporal processes in manufacturing. Specifically, a new design criterion is proposed to jointly consider the prediction precision and the measurement cost, and GA is proposed for solution search. In addition, an innovative monitoring algorithm is developed to estimate, monitor, and update the surface evolution rate based on hypothesis testing. The proposed framework is flexible and applicable to generic spatiotemporal processes and different gage settings in manufacturing.

Appendix: Hypothesis Testing for Monitoring the Temporal Transition Parameter

$u_{t-1,(t-1)r}$ and $u_{t,(t-1)r}$ in Eq. (24) can be rewritten as below

$$u_{t-1,(t-1)r} = \frac{1}{m_1 + m_{2,(t-1)r}} \times \left( \text{sum}(M_{1,t-1}) + \text{sum}(M'_{2,(t-1)r}) \right)$$
$$u_{t,(t-1)r} = \frac{1}{m_1 + m_{2,(t-1)r}} \times \left( \text{sum}(M_{1,t}) + \text{sum}(M'_{2,(t-1)r}) \right)$$

where $\text{sum}(\cdot)$ is a summation function; $m_1$ is the number of level 1 measurement points, and it is time-invariant; and $m_{2,(t-1)r}$ is the number of measurement locations which are measured by level 2 at both time $t-1$ and $t$.

It is known that the measurement error variances of level 1 and level 2 are $\sigma_1^2$ and $\sigma_2^2$, respectively, according to Assumption 2. Therefore, we can assume the level 1 measurement points are independent normal random variables with mean as the true height and variance as the measurement error variance, $\sigma_1^2$. Similarly, level 2 measurement points are also independent normal random variables with mean as the true height and variance $\sigma_2^2$.

Both $u_{t-1,(t-1)r}$ and $u_{t,(t-1)r}$ are normal random variables. Their variances are equal and determined by the repeatability of level 1 and level 2 measurements. We denote their variances as

$$\sigma_{u_{t,(t-1)r}}^2 = \text{Var}(u_{t-1,(t-1)r}) = \text{Var}(u_{t,(t-1)r})$$

$\sigma_{u_{t,(t-1)r}}^2$ can be calculated using the following equation:

$$\sigma_{u_{t,(t-1)r}}^2 = \frac{1}{(m_1 + m_{2,(t-1)r})^2} \left( m_1 \sigma_1^2 + m_{2,(t-1)r} \sigma_2^2 \right)$$
$$= \frac{m_1 \sigma_1^2}{(m_1 + m_{2,(t-1)r})^2} + \frac{m_{2,(t-1)r} \sigma_2^2}{(m_1 + m_{2,(t-1)r})^2}$$

Define $\hat{u}_{t-1,(t-1)r}$ as the estimation of $u_{t-1,(t-1)r}$ using $u_{t-1,(t-1)r}$ and $\hat{\phi}_{t-1}$, as shown by the following equation:

$$\hat{u}_{t-1,(t-1)r} = \hat{\phi}_{t-1} u_{t-1,(t-1)r}$$

A statistic is further defined as the difference between $u_{t,(t-1)r}$ and $\hat{u}_{t-1,(t-1)r}$:

$$\Delta u_{t,(t-1)r} = u_{t,(t-1)r} - \hat{u}_{t-1,(t-1)r}$$

Under $H_0$: $\hat{\phi}_{t-1} = \phi_t$ in the shift detection test for $\phi_t$.

$$E(\Delta u_{t,(t-1)r}) = E(u_{t,(t-1)r} - \hat{\phi}_{t-1} u_{t-1,(t-1)r})$$
$$= E(h_t u_{t-1,(t-1)r} - \hat{\phi}_{t-1} u_{t-1,(t-1)r})$$
$$= (h_t - \hat{\phi}_{t-1}) E(u_{t-1,(t-1)r}) = 0$$

$$\text{Var}(\Delta u_{t,(t-1)r}) = \text{Var}(u_{t,(t-1)r} - \hat{\phi}_{t-1} u_{t-1,(t-1)r})$$
$$= \text{Var}(u_{t,(t-1)r}) + \hat{\phi}_{t-1}^2 \text{Var}(u_{t-1,(t-1)r})$$
$$= \left(1 + \hat{\phi}_{t-1}^2\right) \sigma_{u_{t,(t-1)r}}^2$$

Fig. 12 Heat map of the average prediction variance for ultrasonic metal welding: (a) random sampling and (b) dynamic sampling.
Consequently
\[ \Delta u_{t-1} \sim \mathcal{N}(0, (1 + \phi_r^2 T^{-1}) \sigma_r^2 T_{t-1}^2) \]  

(A8)

Since \( \Delta u_{t-1} \) follows a normal distribution, we can test \( H_0 \) using two-tailed hypothesis test for normal distributions and reject \( H_0 \) if the following condition holds:
\[ \frac{\Delta u_{t-1}}{\sqrt{(1 + \phi_r^2 T^{-1}) \sigma_r^2 T_{t-1}^2}} > Z_{\alpha/2} \]  

(A9)

References

Journal of Manufacturing Science and Engineering
OCTOBER 2017, Vol. 139 / 101002-11

Downloaded From: https://manufacturingscience.asmedigitalcollection.asme.org on 04/29/2019 Terms of Use: http://www.asme.org/about-asme/terms-of-use