

# A Condition Change Detection Method for Solar Conversion Efficiency in Solar Cell Manufacturing Processes

Juan Du, *Student Member, IEEE*, Xi Zhang, *Member, IEEE* and Jianjun Shi

**Abstract**—A solar cell manufacturing process typically has multiple, serial fabrication stages. In general, there are limited quality inspection and control points along the multiple fabrication stages, resulting in inconsistent product quality at the final stage. The existing studies on solar cell manufacturing focus on offline quality monitoring after completion of all manufacturing stages. Thus, an effective in-process method of detecting changes in key stages in terms of product quality is necessary to mitigate the effect of the quality fluctuation. This work proposes an in-process method of detecting the process changes in terms of product quality by studying multichannel sensing signals collected in the epitaxy stage. This stage is an important stage of material layer formation and considerably influences the solar conversion efficiency of the final product. Interpretable geometric features are obtained by considering the process knowledge for each channel of sensing data. A decision-level data fusion method is further proposed to integrate the changes detected from each sensing signal. The proposed method is demonstrated via a real case from solar cell manufacturing processes.

**Index Terms**—multiple change point detection, feature extraction, decision-level data fusion, process monitoring, solar cell manufacturing

## I. INTRODUCTION

Solar cells have been recognized as a promising energy conversion technology. Through the photovoltaic effect [1], light energy is directly converted into electricity without any pollution or waste material. To evaluate the conversion performance of solar cells, solar conversion efficiency (SCE), which refers to the percentage of solar energy that is converted into electricity, is one of the most important quality metrics of solar cells. Owing to the complexity of solar cell fabrication, SCE varies considerably under diverse process conditions. In addition, a low SCE of fabricated solar cells may shorten the lifespan of the equipment in which it is used [2]. Thus, ensuring

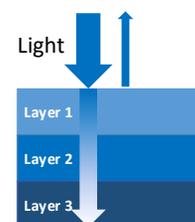


Fig. 1. Light-ray view in three growth layers.

a stable run-to-run SCE of solar cells is important to achieve high efficiency and stable usage of solar energies.

Although multiple stages in the manufacturing process may jointly influence the SCE of finished solar cells, one key upstream stage, namely epitaxy, is mainly related to the SCE of finished products where photovoltaic conversion is conducted. Fig. 1 shows the surface effect between light ray and three growth layers in the epitaxy stage. As the light passes through the growth layer, the light entering the photovoltaic semiconductor layers is absorbed and the remainder is reflected by the surface. The absorbed light energy is converted into electricity. If the process condition changes in the epitaxy stage, material growth on the substrate is affected, leading to a change in SCE [3]. Owing to the lack of available electrodes on the surface of a fabricated wafer after the epitaxy stage, SCE testing is not feasible.

In current practice, a solar cell is tested offline after the main fabrication processes are completed. Solar cells with low SCE are discarded or degraded, leading to low production efficiency and high manufacturing costs. Therefore, it is desirable to develop an automatic detection method for condition changes during the epitaxy stage to “foresee” SCE changes. By using such an in-process condition change detection method, practitioners can take timely remedial actions to reduce production cost and improve efficiency. The current monitoring strategy used in solar cell manufacturing relies on offline inspection using numerous optical detection techniques, such as Raman spectroscopy [4, 5], photoluminescence spectroscopy [6, 7], and X-ray analyses [8, 9]. These techniques are mainly conducted by randomly sampling a few wafers within a single

Manuscript received 12, February, 2018. This work was partially supported by National Natural Science Foundation of China under Grant No. 71471005, 71690232 and 71771004 (Corresponding author: *Xi Zhang*).

J. Du and X. Zhang are with Department of Industrial Engineering and Management, Peking University, Beijing, 100871, China (e-mail: dujuan@pku.edu.cn; xi.zhang@pku.edu.cn).

J. Shi is with the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta, GA 30332 USA (e-mail: jianjun.shi@isye.gatech.edu).

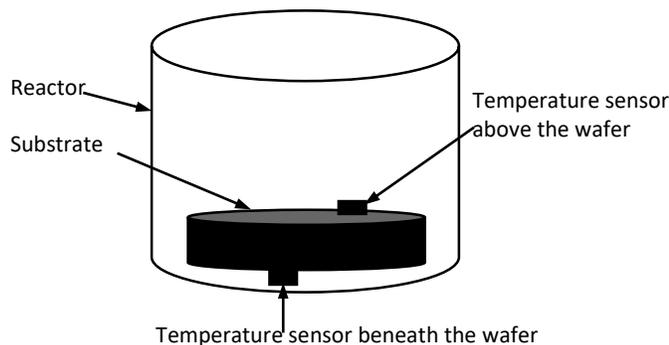


Fig. 2. Illustrations of multiple sensors installed in epitaxy stage.

batch. However, operating such devices requires substantial empirical skills and inspection time.

The rapid development of sensor technologies enables practitioners to achieve in-process condition change detection instead of traditional offline inspection. In the epitaxy stage, multiple temperature sensors are installed in the reactor chamber to collect temperature signals (shown in Fig. 2) because temperature is one of the most important factors that influences the material growth of solar cell wafers [10]. Fig. 3 shows an example of the temperature signals measured on and beneath the wafer. Three layers are designed for one type of solar cell manufacturing process whereas semiconductor materials only grow during periods between the vertical lines. The profiles of various temperature channels exhibit different signatures during material growth, which might co-characterize the process conditions, and further indicate variations of SCE in solar cell wafers. Thus, it is desirable to develop an automatic process change detection method for SCE using such sensing signals.

Significant effort has been exerted to develop effective process monitoring techniques in semiconductor fabrication processes. We divide the related literature in the following three categories.

The first category of detection techniques focuses on supervised learning methods, that is, directly classifying faults by training a model with sensor signals or the extracted features for the specific semiconductor manufacturing application. For example, Ko et al. [11] proposed a fault detection method by extracting structure features from sensor signals and using Hotelling's  $T^2$  for the recipes of similar products. Yu and Lu [12] proposed a joint local and nonlocal linear discriminant analysis method to identify wafer map defects. Lee et al. [25] proposed a deep learning model for robust wafer fault detection and classification. Rao et al. [27] proposed a process-machine interaction model that extracts features for identifying anomalies during chemical mechanical planarization (CMP). These studies developed useful wafer fault monitoring strategies by analyzing the sensing data generated from the semiconductor manufacturing process. However, the aforementioned methods can be implemented only if training data are available for each condition, which is not always possible in practice.

The second line of research identifies the change point within the sensing signals. For example, Son et al. [34] proposed a joint

prognostic model considering a change point for remaining useful life (RUL) prediction. Chen and Tsui [35] proposed a two-phase model with a change point for RUL prediction by using vibration signals of bearings. Bae et al. [36] proposed a hierarchical Bayesian change-point regression model to fit the two-phase degradation patterns with the application to plasma display panels. Du and Zhang [37] proposed a sequential piecewise linear approach to detect changes within the torque signals in threaded pipe connection processes. Du et al. [38] proposed a state-space model and use particle filter to identify a pair of change points within the torque signals in threaded pipe connection processes. All these methods identify the change points within the sensing signals to indicate the process conditions. However, this could not be analogical to the change point detection in solar cell manufacturing processes, where the change point is unnecessarily to be identified along the process signals but between adjacent samples via the signals.

The third category focuses on using statistical process control (SPC) methods, that is, control charts, for change detection. In the literature, conventional SPC methods [13], [14], [15] generally provide powerful change point detection tools by establishing adequate control limits with in-control historical data. On the basis of these tools, monitoring techniques for functional data have been developed from the parametric perspective and nonparametric perspective. For parametric methods, parametric statistical models are first established for the functional data and the model parameters are then used for process monitoring. For example, regression-based process monitoring techniques are widely established for linear and nonlinear profiles, and the regression parameters are used for process monitoring. Examples can be found in [32, 33]. Such regression models could capture the main trend of the temperature signal, but may fail to characterize the local patterns within the temperature signals. Additionally, various nonparametric methods have been developed in the state of the art. For example, to concurrently monitor operating deviations and process dynamic anomalies, Shang et al. [23] proposed a slow feature analysis-based monitoring technique that identifies the changes in process dynamics. Zhao et al. [24] further investigated a full-condition monitoring method for identifying the changes in nonstationary dynamic chemical processes by using cointegration and slow feature analysis. Liu et al. [26] proposed a control chart by using Dirichlet process Gaussian mixture models for anomaly detection in CMP process. Condition monitoring of CMP process according to correlated process variables was investigated in [28]. Gaussian process are applied for profile monitoring in [29, 30]. Yan et al. [31] proposed a spatio-temporal smooth sparse decomposition and a likelihood ratio test for process monitoring. All these methods need to establish the control limits first according to the normal samples and then identify the changes. Although the control limits of temperature signals can be acquired using SPC techniques in the solar cell manufacturing process, such limits cannot directly indicate whether the underlying solar cell is eligible or not in terms of SCE.

To address this limitation, single change point detection methods which do not require any information on specifications

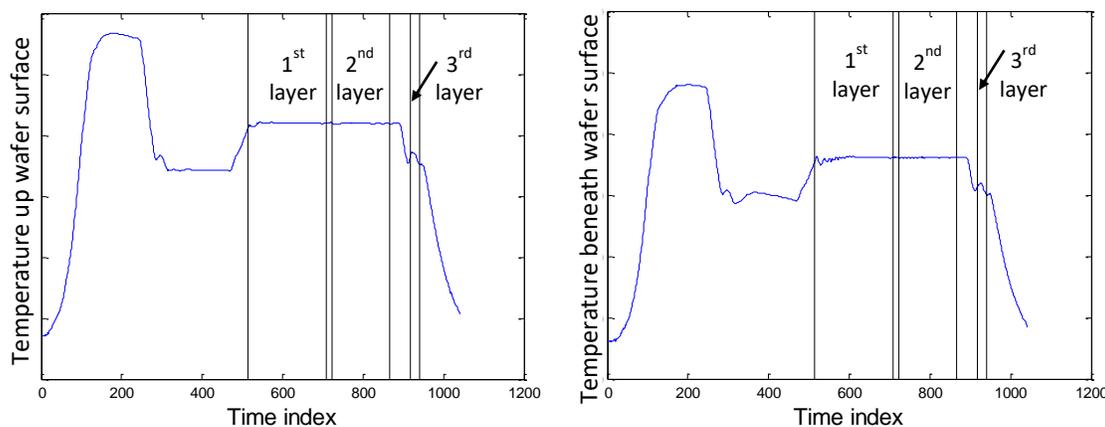


Fig. 3. Temperature signals collected in epitaxy stage.

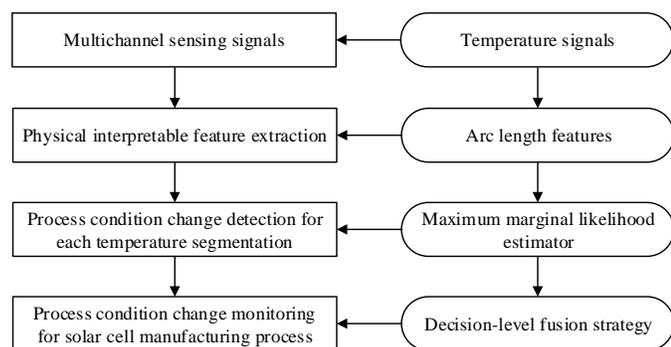


Fig. 4. The flowchart of the proposed methodology.

or control limits have been developed [16], [17]. Based on these methods, multiple condition change detection methods are developed into two categories: (i) The first approach is to use a single change point detection method repeatedly for multiple change point detection through a binary segmentation strategy. A thresholding criterion is generally required to finalize the number of multiple change points. The concept of this method is simple but the computation efficiency is low, especially when the number of condition changes is large. (ii) The second strategy is to identify multiple changes simultaneously. A typical method estimates the locations of change points by using the least squares criterion with an appropriate penalty function to avoid fake change points [18]. Alternatively, a multiple change point detection method based on the maximum of posterior estimation of change point locations has also been proposed [19].

However, existing multiple change point detection methods cannot be directly applied to the solar cell manufacturing process because the data collected from sensors are functional data varying over time, which sets a barrier to precisely estimate multiple change points. To the best of our knowledge, process condition change detection during the epitaxy stage using multichannel sensing data have not been fully investigated in the solar cell manufacturing process.

In our previous study on process monitoring in solar cell manufacturing [20], we investigated the temperature profiles and leveraged curvature of temperature signals to characterize

the process stationarity. We used these multichannel sensing data to achieve process condition change detection in the epitaxy stage. This method delivers good performance to a certain extent. However, it still has the following limitations.

- *First*, to extract curvature features, an analytical model with good smoothness to represent the temperature profiles is necessary. Additive models (e.g., polynomial) could capture the main trend of the temperature signal. Since the smoothness is highly required for the temperature signal, the established model with a moderate number of parameters might only work for part of signal. Hence, such a method uses curvature cannot fully analyze the overall process.
- *Second*, the likelihood ratio statistic used for change point detection could only be used repeatedly for multiple change point detection, leading to inappropriate local change points if the threshold of the test is not carefully chosen.
- *Third*, extracted one-after-another curvature index from the temperature signals frequently contains redundant information and may require unnecessary computations.

To address these limitations, this paper develops a general process change point detection method by introducing a model-free feature as well as a global view of temperature signals and identify multiple change points simultaneously. Our contributions lie in two aspects. *First*, we combine the engineering knowledge and the advanced statistical method to “foresee” the SCE changes. Specially, we extract the interpretable features to capture the process conditions in the epitaxy stage considering the influence of temperature stability to the quality of semiconductor material growth, which mainly causes a change of SCE [3]. *Second*, due to the multi-channel and multi-layer temperature signals of the semiconductor material growth, our contribution is to propose a decision-level fusion technique to efficiently combine the changes of features from different channels together to obtain a single index to further indicate the SCE.

The rest of this paper is organized as follows. Section II presents the proposed methodology for process condition change detection in the epitaxy stage. In Section III, we illustrate our proposed method using industrial data from a solar cell manufacturing process, in which our process condition change detection in the epitaxy stage is validated by

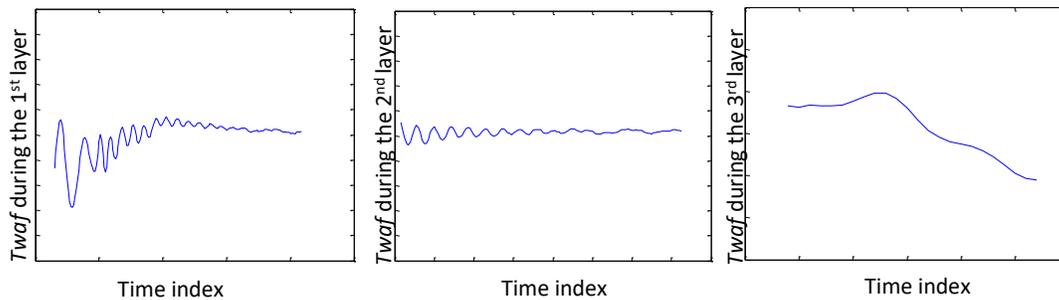


Fig. 5. One sample of segmented temperature signals during three semiconductor material growth layers in epitaxy stage.

corresponding offline tested SCE of solar cells. Finally, a summary of the developed methods is provided in Section IV.

## II. RESEARCH METHODOLOGY

### A. Overview of Proposed Methodology

Fig. 4 shows the overall framework of the proposed methodology. After collecting the multichannel sensing signals, physically meaningful features are extracted from the sensing signals during semiconductor material growth. Specifically, arc length features are extracted from temperature signals by considering material engineering knowledge. Maximum marginal likelihood estimator is then applied for multiple process condition change detection of extracted features. Finally, a decision-level data fusion strategy is proposed to integrate the multiple changes from the different segmented sensing signals to make the final determination on the true process change locations.

### B. Feature Extraction

Due to the complexity of the epitaxial film growth process, a physical model linking temperature and semiconductor material growth is not available. To extract useful information to represent the semiconductor material growth condition from temperature signals, material science as well as the associated engineering knowledge are considered. Although some small changes in temperature are acceptable during semiconductor material growth [21], the best growth condition requirement is constant temperature [22].

From this perspective, the goal of feature extraction is to propose a physical quantity to measure the deviation between temperature signals and constant values. The intuition of arc length is to measure the signatures of temperature signals (fluctuations or trends). These data thereby indicate that the arc length feature of temperature signals is an effective measure of deviation of temperature signals from a constant value. The arc length of a constant value can be proved to be always smaller than any non-constant curves for any given range  $[a, b]$ . Thus, we first propose to use the arc length features to represent temperature signals for the solar manufacturing process.

Let  $C(t)$  be a curve on range  $[a, b]$ . The arc length  $l(C)$  of  $C(t)$  is calculated as

$$l(C) = \int_a^b \left\| \frac{dC}{dt} \right\| dt. \quad (1)$$

In implementation, arc length is approximated by summation of the Euclidean distance between pairs of adjacent sampling points. The usage of arc length enables two key advantages: (1)

no models are needed to fit various segmented temperature signals. Fig. 5 shows a sample of the temperature signals at three different semiconductor material growth layers, where  $Twaf$  denotes the temperature signals beneath the wafer surface. Establishing universal models to describe the profiles of temperature signals at different layers is a challenging task. Our proposed arc length features can directly capture the deviation information of temperature signals from the best temperature environment condition without any model. (2) Arc length is also physically interpretable, which can be easily accepted by practitioners. Therefore, we pursue a physically interpretable quantity to extract effective information from temperature signals.

Without loss of generality, we assume that  $q$  growth layers exist in the epitaxy stage and  $p$  channels of temperature signals are collected during semiconductor material growth in each layer. Then, a total number of  $q \times p$  arc length features are extracted for each sample by using our proposed method. For the following process condition change detection, these  $q \times p$  arc length features are used instead of the original temperature signals.

### C. Process Condition Change Detection via Marginal Likelihood

In general, one or multiple changes may occur in a solar cell manufacturing process. Most existing multiple change point detection methods optimize a loss function with penalty terms, such as that in [18]. The penalty tuning parameter in the loss function is often selected by using cross-validation, thereby resulting in theoretical challenges, such as the effect of adaptation [19]. A Bayesian method for multiple change point detection problem is more flexible and can be widely used in many models. Du et al. [19] proposed a maximum marginal likelihood estimator to determine the number and locations of change points, which guarantees the true number and locations of change points of the signals asymptotically. Thus, we adopt this method for multiple change point detection based on the extracted features. This subsection briefly introduces this method whereas additional information and details of this method can be found in [19].

#### 1) Problem formulation

We assume that the feature  $\mathbf{f} = \{f_1, f_2, \dots, f_n\}$ , is extracted from sample indices  $\mathbf{s} = \{s_1, s_2, \dots, s_n\}$ ,  $s_1 < s_2 < \dots < s_n$ , and  $\mathbf{f} = \{f_1, f_2, \dots, f_n\}$  comes from a family of distribution functions  $g(\mathbf{f}|\theta)$ ,  $\theta \in \Theta$ . The  $(N - 1)$  change points occurred at sample indexes  $\boldsymbol{\tau} = \{\tau_1, \tau_2, \dots, \tau_{N-1}\}$ ,  $\tau_i \in [s_1, s_{n-1}]$  for  $i \in$

$\{1, \dots, N-1\}$ , dividing the signal into  $N$  segments,  $N \leq n$ . The multiple change point detection problem can be formulated as

$$\theta(s) = \theta_i, \text{ if } s \in (\tau_{i-1}, \tau_i]. \quad (2)$$

Given the  $N$  segments and associated parameters  $\theta_i$ ,  $i \in \{1, \dots, N\}$ , the features are assumed to be independently distributed as follows:

$$P(\mathbf{f}|\boldsymbol{\tau}, \boldsymbol{\theta}_{1:N}) = \prod_{i=1}^N \prod_{s_j \in (\tau_{i-1}, \tau_i]} g(f_j|\theta_i), j = 1, \dots, n. \quad (3)$$

After the problem formulation, the following estimation procedure is proposed to determine the number of feature change segments  $N$  and locations of change sample indexes  $\boldsymbol{\tau} = \{\tau_1, \tau_2, \dots, \tau_{N-1}\}$ .

### 2) Maximum marginal likelihood estimator

The number  $N$  and locations of change sample indexes  $\boldsymbol{\tau}$  are estimated using the maximum marginal likelihood method in which family distribution parameters  $\theta_i$ ,  $i \in \{1, \dots, N\}$  are integrated. Given the change sample indexes  $\boldsymbol{\tau} = \{\tau_1, \tau_2, \dots, \tau_{N-1}\}$ ,  $\theta_i$ ,  $i \in \{1, \dots, N\}$  are assumed to be independently and identically drawn from a prior distribution  $\pi(\cdot|\alpha)$ . The marginal likelihood can be formulated as

$$P(\mathbf{f}|\boldsymbol{\tau}) = \prod_{i=1}^N \int_{\theta_i} \prod_{s_j \in (\tau_{i-1}, \tau_i]} g(f_j|\theta_i) \pi(\theta_i|\alpha) d\theta_i \\ = \prod_{i=1}^N D(\mathbf{f}_{(\tau_{i-1}, \tau_i]}|\alpha), \quad (4)$$

where  $\tau_0$  and  $\tau_N$  are set as 0 and  $s_n$ , respectively.  $D(\mathbf{f}_{(\tau_{i-1}, \tau_i]}|\alpha)$  denotes the probability that no change occurs in the features during the sample indexes  $(\tau_{i-1}, \tau_i]$ . Note that the maximum of the aforementioned marginal likelihood is equivalent to the maximum posterior estimation of the change points with uniform prior  $P(\boldsymbol{\tau}) \propto 1$  due to the following formula:

$$P(\boldsymbol{\tau}|\mathbf{f}) \propto P(\mathbf{f}|\boldsymbol{\tau})P(\boldsymbol{\tau}) = P(\mathbf{f}|\boldsymbol{\tau}). \quad (5)$$

The uniform prior is also viewed as noninformative, indicating that the prior probability of every  $f_j$  being a change point is the same for  $j = 1, \dots, n$ . Such a choice is reasonable, especially when we do not know where a change occurs. Notably, in practice, the number of change points may be restricted. Thus, we suppose an upper bound  $M \leq n$  on the number of segments, i.e.,  $N \leq M$ . The extreme case is  $M = n$ , which means that every observed sample is a segment.

### 3) Computation algorithm via dynamic programming

The formulation of marginal likelihood (4) indicates that  $P(\mathbf{f}|\boldsymbol{\tau})$  can be calculated as a product of non-overlapping functions  $D(\mathbf{f}_{(\tau_{i-1}, \tau_i]}|\alpha)$ . Thus, dynamic programming is an effective tool to solve this problem through the following algorithm. Let  $H(f_1, \dots, f_j|N) = \max_{\boldsymbol{\tau} \subseteq \{s_1, \dots, s_{j-1}\}} P(f_1, \dots, f_j|\boldsymbol{\tau})$ ,  $N = 1, \dots, M$ .

Algorithm 1:  
 Given a family of distribution functions  $g(f_j|\theta_i)$  and a prior distribution  $\theta_i \sim \pi(\cdot|\alpha)$   
 For  $N = 1$   
     For  $j = 1:n$   
          $H(f_1, \dots, f_j|N) = D(f_1, \dots, f_j|\alpha)$   
     End  
 End  
 For  $N = 2:M-1$   
     For  $j = N:n$

$$H(f_1, \dots, f_j|N) = \max_{N-1 \leq i \leq j-1} H(f_1, \dots, f_i|N-1) \cdot D(f_{i+1}, \dots, f_j|\alpha)$$

End  
 End  
 For  $N = M$   
     For  $j = M:n$   
          $H(f_1, \dots, f_j|N) = \max_{M-1 \leq i \leq j-1} H(f_1, \dots, f_i|M-1) \cdot D(f_{i+1}, \dots, f_j|\alpha)$   
     End  
 End

Through above procedure in Algorithm 1,  $\boldsymbol{\tau}$  can be estimated by  $\hat{\boldsymbol{\tau}} = \arg \max_{\boldsymbol{\tau}} P(\mathbf{f}|\boldsymbol{\tau})$ .

Based on the aforementioned  $q \times p$  arc length features from one sample, the marginal likelihood estimator is conducted for the process condition change detection. Assume that  $m$  samples are collected, arc length features, namely,  $l_{1k}, l_{2k}, \dots, l_{mk}$ ,  $k = 1, \dots, q \times p$  can be extracted from these  $m$  samples. For the  $k^{th}$  type of arc length feature, multiple change points can be identified using the maximum marginal likelihood estimator. Assume a number of  $M_k$  change points are detected at locations  $c_{kh}$ ,  $h = 1, \dots, M_k$  for the  $k^{th}$  type of arc length features.  $G_{kh}$  is the corresponding arc length mean shifts at change points  $c_{kh}$ ,  $h = 1, \dots, M_k$ ,  $k = 1, \dots, q \times p$ . After obtaining change point locations and corresponding mean shifts for each semiconductor material layer of each channel, the last goal is fusing these change points and formulating the final determination of significant change points in the epitaxy stage that may induce mean shifts in SCE of produced solar cells.

### D. Decision-level Fusion Strategy

Several process changes may be identified according to the sensing signals measured at different channels during different growth layer periods using the above method. However, only a few significant changes among them may induce significant mean shifts in SCE of the finished solar cells. Thus, the values of mean shifts  $G_{kh}$  characterize the level of process changes. Based on this concept, a decision-level fusion strategy is proposed to integrate all identified changes from the sensing signals measured at various channels in different growth layers. We compose the final determination of significant changes during the epitaxy stage, which may lead to mean shifts in SCE.

In the epitaxy stage, due to the gas non-uniformity in the chamber, different locations in the chamber may show different temperature conditions. The temperature on and beneath the substrate would jointly affect the quality of the grown layers [3]. Hence, we first propose to extract arc length features from multi-channel temperature signals for multiple change detection. Then, we propose a decision-level fusion strategy to generate a single index for the final determination of changes in terms of SCE. Specifically, we define  $W_{kh}$  to measure the significance of each identified change as follows:

$$W_{kh} = \sum_{k', h'} y_{kh} |G_{k' h'}|, \quad (6)$$

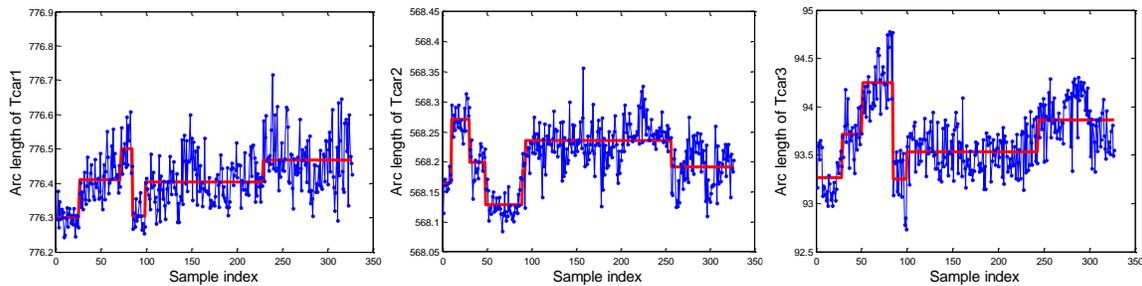


Fig. 6. Change detection results for arc length features of first growth layer (left), second growth layer (middle), and last growth layer (right) of temperature signals above the wafer surface.

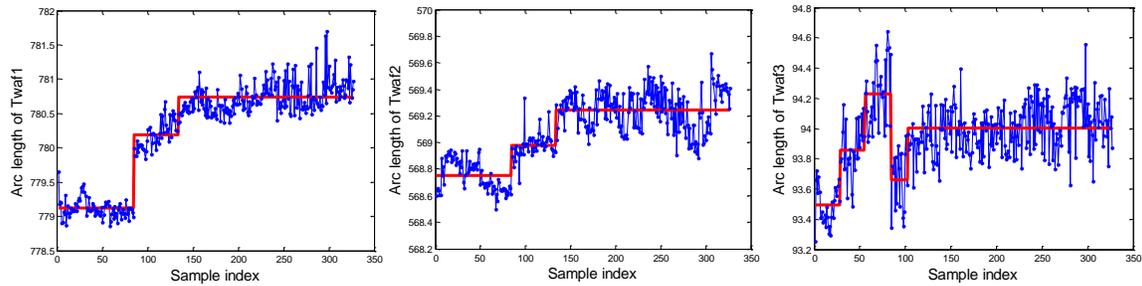


Fig. 7. Change detection results for arc length features of first growth layer (left), second growth layer (middle), and last growth layer (right) of temperature signals beneath the wafer surface.

where  $y_{kh} = \begin{cases} 1 & c_{kh} = c_{k'h'} \\ 0 & c_{kh} \neq c_{k'h'} \end{cases}$ , which is an indicator function,  $k = 1, \dots, q \times p$ ;  $k' = 1, \dots, q \times p$ ;  $h = 1, \dots, M_k$ ;  $h' = 1, \dots, M_{k'}$ . This indicates that  $W_{kh}$  is a summation of weighted absolute mean shifts at the same change sample index from different growth layers or different channels. Notably, the mean shifts  $G_{kh}$  at change points  $c_{kh}$  can be negative or positive. The negative  $G_{kh}$  value indicates that the mean of the arc length features after the change point  $c_{kh}$  is smaller than the previous mean value. That is, the temperature signals become less variable after the process change and provide a better temperature environment for semiconductor material growth, and vice versa.

In practice,  $W_{kh}$  can be obtained at all change points  $c_{kh}$ . We can select significant change points  $c_{kh}$  with large  $W_{kh}$  values using the scree plot or set a specified threshold by engineering specifications or simulations. If  $W_{kh}$  exceeds the threshold, then corresponding change locations induce a significant process condition change, which may lead to mean shifts in SCE. Otherwise, the wafers are assumed to be fabricated under the same condition and process condition change does not occur. Through the proposed method, all significant change points can be identified during the epitaxy stage. In addition, the samples can be segmented for further Phase II process monitoring (i.e., establishing control charts for online monitoring) and diagnosis by practitioners.

### III. INDUSTRY EXAMPLE

#### A. Data Description

In this section, we collect data of solar cell manufacturing process from a real plant. Recall that the SCE of finished solar cells is highly dependent on the epitaxy stage in the solar cell manufacturing process. Thus, in this experiment, we collected 327 two-channel temperature signals above and beneath the

wafer surface at the epitaxy stage and corresponding SCE of finished solar cells. The sampling time of these temperature signals is 4 seconds and 1,042 sampling points are collected after completion of the epitaxy stage.

In the epitaxy stage, semiconductor materials grow layer by layer. Thus, the collected sensing signals may have a transition time between different layers because the temperature condition requirement of material growth in different layers is different. For example, the temperature signals need substantial time to heat up or cool down to meet the temperature requirements of different growth layers. Thus, the first step is to segment the sensing signals according to the growth period. Then, these two-channel segmented temperature signals are used for process change detection in the epitaxy stage to validate our proposed method, which aims to “foresee” the changes in SCE of the finished solar cells. The corresponding SCE data of the finished solar cells are finally used to validate the proposed process change detection method.

#### B. Results and Analysis

In this case study, there are three types of semiconductor material growing layer by layer in the epitaxy stage. Thus, the sensing signals are divided into three segments with one segment corresponding to one layer. In this case study,  $p = 2, q = 3$ . In addition, six types of arc length features are extracted from one sample. By examining the extracted features, the features  $f_j | (\mu_i, \sigma_i^2)$  follow normal distributions with mean  $\mu_i$  and variance  $\sigma_i^2$ .

According to the guidance of prior choice in [19], we choose conjugate priors for hyperparameters  $\mu_i$  and  $\sigma_i^2$ , that is,  $\sigma_i^2 | N$  follows scaled  $\text{Inv-}\chi^2(3, 5\hat{\sigma}^2)$  and  $\mu_i | (\sigma_i^2, N)$  follows a normal distribution with mean  $\bar{f}$  and variance  $2\sigma_i^2$ , where  $\bar{f}$  and  $\hat{\sigma}^2$  are the sample mean and sample variance of the extracted features. This is a weaker prior choice and thus has an advantage to

TABLE I  
CHANGE DETECTION RESULTS OF EACH SENSING SIGNAL VIA MARGINAL LIKELIHOOD ESTIMATOR

Sensing signal	Number of changes ( $M_k$ )	Change location ( $c_{kh}$ )	Mean shift of arc length ( $G_{kh}$ )
$Tcar1$	5	25	0.110
		71	0.089
		84	-0.195
		98	0.099
		228	0.063
$Tcar2$	6	10	0.104
		30	-0.070
		48	-0.070
		89	0.057
		93	0.050
		256	-0.044
$Tcar3$	5	28	0.447
		50	0.535
		84	-1.004
		99	0.288
		243	0.331
$Twaf1$	2	84	1.065
		133	0.551
$Twaf2$	2	84	0.228
		133	0.265
$Twaf3$	4	28	0.367
		55	0.369
		84	-0.566
		102	0.341

identify the most significant mean shifts of arc length features and less sensitive to the small changes. In the epitaxy stage, the large temperature condition changes may lead to significant mean changes of SCE. Therefore, the above prior choice is designed to identify the significant mean shifts of SCE.

Given the prior distributions, the maximum marginal likelihood estimator can be calculated for each type of arc length features from 327 samples. Figs. 6 and 7 show the change detection results for arc length features from temperature signals above and beneath the wafer surface, respectively.  $Tcar1$ ,  $Tcar2$ , and  $Tcar3$  denote the temperature signals above the wafer surface during growth of the first, second, and last layers.  $Twaf1$ ,  $Twaf2$ , and  $Twaf3$  denote the temperature signals beneath the wafer surface during growth of the first, second, and last layers. The vertical line on the x-axis indicates that a change occurs. The y-axis value of the horizontal line between two changes presents the mean value of the arc length in this condition.

Table I lists the number  $M_k$ , location of condition changes  $c_{kh}$ , and corresponding mean shift  $G_{kh}$  of the arc length of each temperature signal. As shown in Table I, process change locations and corresponding mean shifts may be dissimilar for different temperature signals. After these changes and corresponding mean shifts are obtained, the decision-level fusion method is then conducted for final determination based on  $W_{kh}$ . We listed the change sample location  $c_{kh}$  that corresponds to the decreasing value of  $W_{kh}$  in Table II, thereby indicating the significance of condition changes in a decreasing order. The final determination of changes can be finalized using the scree plot of  $W_{kh}$ , as shown in Fig. 8. Three changes are finally selected for the collected samples from the scree plot based on the significant decrease of the  $W_{kh}$  value. Then,

TABLE II  
CHANGE SAMPLE INDEX  $c_{kh}$  AND CORRESPONDING  $W_{kh}$  VALUES

Order index	Change index $c_{kh}$	$W_{kh}$ value
1	84	3.059
2	133	0.816
3	28	0.814
4	50	0.535
5	55	0.369
6	102	0.341
7	243	0.331
8	99	0.288
9	25	0.110
10	10	0.105
11	98	0.099
12	71	0.089
13	48	0.071
14	30	0.071
15	228	0.063
16	89	0.057
17	93	0.050
18	256	0.044

TABLE III  
GROUP AND CORRESPONDING SAMPLE INDEXES

Group	Sample index	Mean of SCE
1	1-28	19.64
2	29-84	19.39
3	85-133	19.77
4	134-327	19.54

comparing these changes with the changes in SCE of corresponding completed solar cells is necessary to demonstrate the proposed method.

Fig. 9 shows the detected changes in SCE according to the  $W_{kh}$  values. The vertical dash lines on the x-axis indicate the change locations, whereas the horizontal lines on the y-axis indicate the mean values of SCE under varying temperature conditions. As shown in Fig. 9, the changes determined by our proposed method at the epitaxy stage can effectively characterize changes in SCE of the finished solar cells.

To further illustrate the relationship between changes of arc lengths and SCE mean changes, we first divide the sample indexes into four groups based on the SCE shifts, as listed in Table III. We then plot these four groups according to the arc length features of temperature signals beneath the wafer surface. As shown in Fig. 10, it can be seen that these four production groups having different arc length values and can be separated, thereby indicating that arc length features can be utilized for early stage process change detection in terms of SCE.

In cases, we are not able to predetermine the exact changes in SCE. Thus, validating the detection power becomes an issue. In this case study, we use statistical change detection methods to label the changes in SCE. We use the maximal marginal likelihood estimator to label the changes of SCE. We set the upper bound number of segments  $M = 4$  to validate the proposed method. We check that the SCE data follow normal distribution, so we also set the conjugate prior, which is the same type with the change detection of arc length. The labeled change indexes of SCE under  $M = 4$  are sample indexes 31, 86, and 143, as shown in Fig. 11. We can observe that the changes

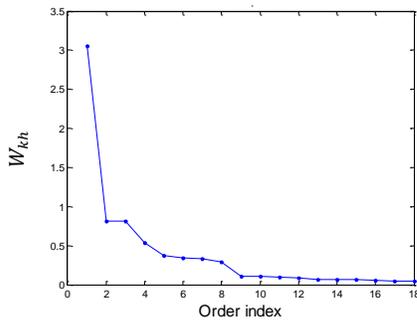


Fig. 8. Scree plot of  $W_{kh}$ .

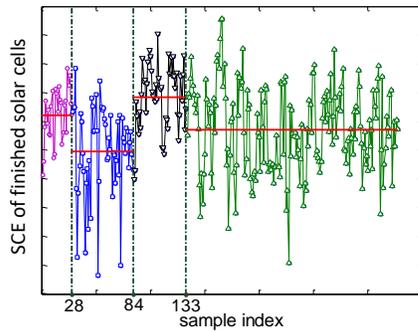


Fig. 9. Change detection results of SCE by our proposed method.

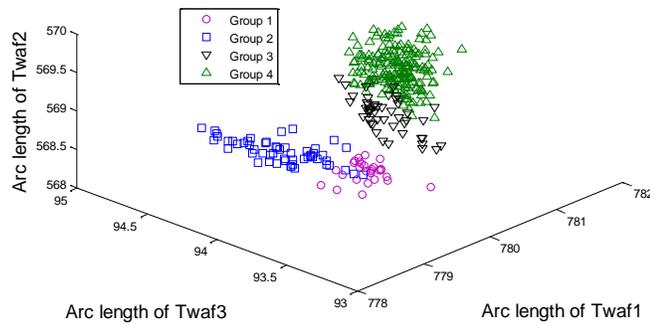


Fig. 10. Scatter plot for 4 groups of SCE productions with regard to arc length of the temperature signals beneath the wafer surface.

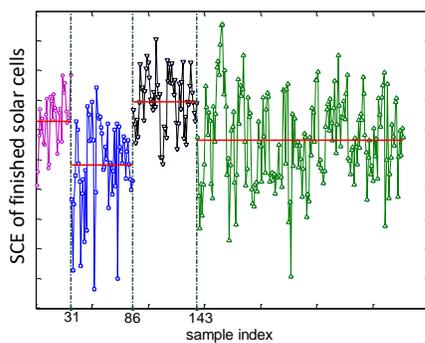


Fig. 11. Change detection of SCE via maximum likelihood estimation method under upper bound  $M = 4$ .

identified by the proposed method are quite close to the labeled significant change locations, thereby indicating the effectiveness of the proposed method.

### C. Comparison

In this section, we compare our method with two existing benchmark methods with different data fusion techniques. We first compare our proposed method with that in [20], which

TABLE IV  
COMPARISON RESULTS BETWEEN PROPOSED AND EXISTING METHODS

Method	Three major significant change sample indexes	Computation time (s)
Statistical Labels	31, 86, 143	-
Our method	28, 84, 133	0.78
Method in [20]	85, 163, 230	12.25
MFPCA [16]	132 (133), 241	246.38

also aims to achieve multiple-condition change detection in the solar cell manufacturing process. The method in [20] first extracts multiple features from different channels of sensing signals, and then uses principal component analysis (PCA) for feature-level fusion. Finally, a likelihood ratio test is used repeatedly for change detection. The three major significant changes detected by this method from the same industry data occur at sample indexes 85, 163, and 230, as listed in Table IV. We also list the computation times of our method and the method in [20] by using MATLAB R2013a on a computer with Intel Core i5-4210U @ 2.40 GHz processor, 4 GB RAM. With relation to the statistical change labels 31, 86, and 143, the method in [20] cannot identify the change sample index around 31 in the first three significant changes and is thus less effective than our method. Moreover, the computation time is much longer than our method due to the repeated calculations of the likelihood test for multiple-change point detection. Thus, we believe that our proposed method is comparably better than [20] in identifying true changes and computation time. In addition, unlike arc length features, curvature features cannot capture the linear trend of temperature signals because the curvature feature of the temperature signal in a linear trend is zero, which indicates that curvature features cannot distinguish the constant curve from the linear trend.

We subsequently compare our method with the method in [16], which is a pure data-driven method that does not consider engineering knowledge. The method in [16] proposes multidimensional functional principal component analysis (MFPCA) to achieve raw data fusion of multi-channel functional data and extracts principal component features. Such features are further used for identifying one outstanding change via a likelihood ratio test. We compare this method in two modes. In the first mode, we perform MFPCA to extract features from all temperature signals above and beneath the wafer surface. The most significant change sample index is 132. In the second mode, we use MFPCA to extract features from the segmented temperature signals according to growth layers Tcar1, Twaf1, Tcar2, Twaf2, Tcar3, and Twaf3. The most significant change sample indexes for the first, second, and last growth layers are 132, 133, and 241, respectively. In Table IV, we list the detection results of the second comparison and the computation time by using R 3.0.2 on the same computer as used for our method. Compared with the change labels in Table IV, the detected changes of temperature signals by MFPCA cannot successfully indicate the changes of SCE; our method can efficiently identify the condition changes from temperature signals in terms of SCE with the aid of engineering knowledge. In addition, our method can simultaneously identify multiple

condition changes, whereas the method in [16] can only identify single change. Hence, our method is more effective and interpretable for practitioners.

#### D. Discussion

The extracted arc length features represent the flatness of temperature, thereby indicating the temperature condition provided in the epitaxy stage for semiconductor material growth. For example, a large magnitude of arc length feature indicates the unfavorable temperature condition provided by the epitaxy stage for semiconductor material growth, and vice versa. Thus, it is reasonable for foreseeing a downward mean shift in SCE if the arc length achieves a larger magnitude at a change sample index in the epitaxy stage, and vice versa.

We proposed a quantity  $T_{kh} = \sum_{k',h'} y_{kh} G_{k'h'}$  to measure the total arc length changes in segmented temperature signals. Table V lists the value of  $T_{kh}$  at the significant change sample indexes. The following observations can be seen from this case study:

- If  $T_{kh} < 0$ , the overall arc length decreases after the change. This means there is a better temperature condition for material growth after the change, which is validated at sample index 84. We can observe that an upward mean shift in SCE exists at sample index 84 in Fig. 9.
- If  $T_{kh} > 0$ , the overall arc length increases after the change. This means there is a worse temperature condition for material growth after the change, which is validated at sample indexes 28 and 133. We can observe that downward mean shifts exist in SCE at change sample indexes 28 and 133 in Fig. 9.

Therefore, the proposed method can not only predict the mean shifts of SCE but also indicate the corresponding shift directions by using the temperature data in the epitaxy stage.

Notably, in the epitaxy stage of solar cell manufacturing processes, multiple phases have been designed during the whole process for semiconductor material growth, and the normal changes of process conditions exist between two successive phases. Practically, the normal changes between the successive material growth phases are known a priori according to the process knowledge.

Multiphase batch process monitoring methods such as [39] are substantially useful for the scenario where a large amount of historical data is available for establishment of the confidence limits in process monitoring. In the solar cell manufacturing case, historical data is generally unavailable to establish such confidence limits due to varying specification requirements of the solar cell products. Therefore, the multiphase process monitoring method cannot be directly used to solve the problem in solar cell manufacturing processes. In our method, we can directly fuse the changes of extracted features from the temperature signals to indicate SCE changes without any historical dataset.

The error accumulation before the epitaxy stage is mainly dependent on the quality of the substrates, which are fully inspected before they are released for the epitaxy process. Therefore, the errors accumulated from previous stages are not considered in this paper. The practitioners only need to

TABLE V  
CHANGE SAMPLE INDEX  $c_{kh}$  AND CORRESPONDING  $T_{kh}$  VALUES

Order index	Change index $c_{kh}$	$T_{kh}$ value
1	84	-0.472
2	133	0.816
3	28	0.813

keep the process parameters as same in the design phase in the epitaxy stage in order to ensure the consistency of product quality.

The detection performance may not be good using multi-channel signals simultaneously due to the following two reasons. *First*, few segments of temperature signals are strongly related to the layer growth because the preparation of epitaxy condition (e.g., preheating and cooling down of the chamber) would take a large amount of the time in the whole period. Hence, we only pick the temperature segments in which the material layers are growing, to extract the arc length for efficient SCE change monitoring. *Second*, there are ‘compensation effect’ of features from different layer growth periods. For example, given a sample whose arc length is larger in the first grown layer but smaller in the second grown layer than the normal samples, the overall arc length feature might be equivalent to the normal samples. In this situation, this would fail to identify the condition change of such sample if we use the arc length feature extracted from the whole piece of signals, i.e., using the multiple signals simultaneously. In this research, we adopt change detection by using multiple signals individually first and then fuse to achieve a good detection performance. As shown in the case study, our method can effectively achieve the goal of change detection in epitaxy stage.

#### IV. CONCLUSION

In this paper, we proposed an in-process change detection method in the solar cell manufacturing process by integrating the engineering domain knowledge and advanced statistical methods. Compared with other methods that use training data or labeled samples to establish a suitable classifier to achieve the process change detection, the proposed method fully considers the engineering knowledge in material science domain and proposed arc length features to characterize the temperature conditions of the epitaxy stage. The multiple process condition changes are identified by a multiple mean shift detection method based on these extracted features. A decision-level fusion method is proposed for the final determination of changes to exclude the inessentially false alarms. Results of the real industrial case study confirmed that the proposed method can simultaneously identify multiple process condition changes in the epitaxy stage of the solar cell manufacturing process and foresee the mean shifts in SCE of corresponding fabricated solar cells. Future research aims to extend the study to forecast the SCE quantitatively of fabricated solar cells and establish effective control charts for Phase II process monitoring at the epitaxy stage by using the collected multichannel sensing data.

## REFERENCES

- [1] W. Shockley and H. J. Queisser, "Detailed balance limit of efficiency of p-n junction solar cells," *J. Appl. Phys.*, vol. 32, no. 3, pp. 510-519, 1961.
- [2] K. Masuko *et al.*, "Achievement of more than 25% conversion efficiency with crystalline silicon heterojunction solar cell," *IEEE J. Photovolt.*, vol. 4, no. 6, pp. 1433-1435, 2014.
- [3] A. J. McEvoy, L. Castaner, and T. Markvart, *Solar cells: materials, manufacture and operation*. Academic Press, 2012.
- [4] J. Alvarez-Garcia, E. Rudigier, N. Rega, B. Barcones, R. Scheer, A. Pérez-Rodríguez, A. Romano-Rodríguez and J. R. Morante, "Growth process monitoring and crystalline quality assessment of CuInS(Se)<sub>2</sub> based solar cells by Raman spectroscopy," *Thin Solid Films*, vol. 431, pp.122-125, 2003.
- [5] V. Izquierdo-Roca *et al.*, "Process monitoring of chalcopyrite photovoltaic technologies by Raman spectroscopy: an application to low cost electrodeposition based processes," *New J. Chem.*, vol. 35, no. 2, pp. 453-460, 2011.
- [6] S. Ostapenko, I. Tarasov, J. P. Kalejs, C. Haessler, and E. U. Reisner, "Defect monitoring using scanning photoluminescence spectroscopy in multicrystalline silicon wafers," *Semicond. Sci. and Technol.*, vol. 15, no. 8, pp. 840-848, 2000.
- [7] T. Trupke, R. A. Bardos, M. D. Abbott, F. W. Chen, J. E. Cotter, and A. Lorenz, "Fast photoluminescence imaging of silicon wafers," in *Proc. 32nd IEEE Photovoltaic Spec. Conf., 4th World Conf. Photovoltaic Energy Convers.*, 2006, pp. 928-931.
- [8] L. Zhang, X. Xu, Z. Yang, X. Sun, H. Xu, H. Liu, and H. Shen, "An efficient method for monitoring the shunts in silicon solar cells during fabrication processes with infrared imaging," *J. Semicond.*, vol. 30, no. 7, p. 076001, 2009.
- [9] A. A. Istratov, H. Hieslmair, O. F. Vyvenko, E. R. Weber, and R. Schindler, "Defect recognition and impurity detection techniques in crystalline silicon for solar cells," *Sol. Energy Mater. Sol. Cells*, vol. 72, no. 1, pp. 441-451, 2002.
- [10] Y. Nishi, and R. Doering, *Handbook of Semiconductor Manufacturing Technology*. CRC Press, 2000.
- [11] J. M. Ko, C. O. Kim, S. J. Lee, and J. P. Hong, "Structural feature-based fault-detection approach for the recipes of similar products," *IEEE Trans. Semicond. Manuf.*, vol. 23, no. 2, pp. 273-283, 2010.
- [12] J. Yu and X. Lu, "Wafer map defect detection and recognition using joint local and nonlocal linear discriminant analysis," *IEEE Trans. Semicond. Manuf.*, vol. 29, no. 1, pp. 33-43, 2016.
- [13] J. H. Sullivan, and W. H. Woodall, "A control chart for preliminary analysis of individual observations," *J. Qual. Technol.*, vol. 28, no. 3, pp. 265-278, 1996.
- [14] K. J. Worsley, "On the likelihood ratio test for a shift in location of normal populations," *J. Am. Stat. Assoc.*, vol. 74, no. 366a, pp. 365-367, 1979.
- [15] J. H. Sullivan and W. H. Woodall, "Change-point detection of mean vector or covariance matrix shifts using multivariate individual observations," *IIE Trans.*, vol. 32, no. 6, pp. 537-549, 2000.
- [16] K. Paynabar, C. Zou, and P. Qiu, "A change-point approach for phase-I analysis in multivariate profile monitoring and diagnosis," *Technometrics*, vol. 58, no. 2, pp. 191-204, 2016.
- [17] K. Paynabar, and J. Jin, "Characterization of non-linear profiles variations using mixed-effect models and wavelets," *IIE Trans.*, vol. 43, no. 4, pp. 275-290, 2011.
- [18] Z. Harchaoui, and C. Lévy-Leduc, "Multiple change-point estimation with a total variation penalty," *J. Am. Stat. Assoc.*, vol. 105, no. 492, pp. 1480-1493, 2010.
- [19] C. Du, C. L. M. Kao, and S. C. Kou, "Stepwise signal extraction via marginal likelihood," *J. Am. Stat. Assoc.*, vol. 111, no. 513, pp. 314-330, 2016.
- [20] J. Du, X. Zhang, and Q. Hu, "An automatic condition detection approach for quality assurance in solar cell manufacturing processes," *IEEE Robot. Autom. Lett.*, vol. 2, no. 3, pp. 1825-1831, 2017.
- [21] L. C. Su, I. H. Ho, and G. B. Stringfellow, "Control of ordering in Ga<sub>0.5</sub>In<sub>0.5</sub>P using growth temperature," *J. Appl. Phys.*, vol. 76, no. 6, pp. 3520-3525, 1994.
- [22] F. Smith William and J. Hashemi, *Foundations of Materials Science and Engineering*. New York, NY, USA: McGraw-Hill, 2011.
- [23] C. Shang, F. Yang, X. Gao, X. Huang, J. A. Suykens, and D. Huang, "Concurrent monitoring of operating condition deviations and process dynamics anomalies with slow feature analysis," *AIChE J.*, vol. 61, no. 11, pp. 3666-3682, 2015.
- [24] C. Zhao and B. Huang, "A full - condition monitoring method for nonstationary dynamic chemical processes with cointegration and slow feature analysis," *AIChE J.*, vol. 64, no. 5, pp. 1662-1681, 2018.
- [25] H. Lee, Y. Kim, and C. O. Kim, "A deep learning model for robust wafer fault monitoring with sensor measurement noise," *IEEE Trans. Semicond. Manuf.*, vol. 30, no. 1, pp. 23-31, 2017.
- [26] J. P. Liu, O. F. Beyca, P. K. Rao, Z. J. Kong, and S. T. Bukkapatnam, "Dirichlet process Gaussian mixture models for real-time monitoring and their application to chemical mechanical planarization," *IEEE Trans. Auto. Sci. Eng.*, vol. 14, no. 1, pp. 208-221, 2017.
- [27] P. K. Rao *et al.*, "Process-machine interaction (PMI) modeling and monitoring of chemical mechanical planarization (CMP) process using wireless vibration sensors," *IEEE Trans. Semicond. Manuf.*, vol. 27, no. 1, pp. 1-15, 2014.
- [28] H. Wang, X. Zhang, A. Kumar, and Q. Huang, "Nonlinear dynamics modeling of correlated functional process variables for condition monitoring in chemical-mechanical planarization," *IEEE Trans. Semicond. Manuf.*, vol. 22, no. 1, pp. 188-195, 2009.
- [29] Y. Zhang, Z. He, C. Zhang, and W.H. Woodall, "Control charts for monitoring linear profiles with within-profile correlation using Gaussian process models," *Qual. Reliab. Engng. Int.*, vol. 30, no. 4, pp. 487-501, 2014.
- [30] P. C. Cicorella, B. M. Colosimo, and M. Pacella, "Statistical process monitoring of complex shapes via Gaussian process modeling," *Milan, Italy*, pp. 1-6, 2013.
- [31] H. Yan, K. Paynabar, and J. Shi, "Real-time monitoring of high-dimensional functional data streams via spatio-temporal smooth sparse decomposition," *Technometrics*, vol. 60, no. 2, pp. 181-197, 2018.
- [32] K. Kim, M. A. Mahmoud, and W. H. Woodall, "On the monitoring of linear profiles," *J. Qual. Technol.*, vol. 35, no. 3, pp. 317-328, 2003.
- [33] J. D. Williams, W. H. Woodall, and J. B. "Birch statistical monitoring of nonlinear product and process quality profiles," *Qual. Reliab. Eng. Int.*, vol. 23, no. 8, pp. 925-941, 2007.
- [34] J. Son, Y. Zhang, C. Sankavaram, and S. Zhou, "RUL prediction for individual units based on condition monitoring signals with a change point," *IEEE Trans. Reliab.*, vol. 64, no. 1, pp. 182-196, 2015.
- [35] N. Chen, and K. L. Tsui, "Condition monitoring and remaining useful life prediction using degradation signals: Revisited," *IIE Trans.*, vol. 45, no. 9, pp. 939-952, 2013.
- [36] S. J. Bae, T. Yuan, S. Ning, and W. Kuo, "A Bayesian approach to modeling two-phase degradation using change-point regression," *Reliab. Eng. Syst. Safe.*, vol. 134, pp. 66-74, 2015.
- [37] J. Du, and X. Zhang, "A critical change point detection method in threaded steel pipe connection processes using two stage sequential piecewise linear approach," in *ASME Paper No. MSEC2016-8757*, 2016.
- [38] J. Du, X. Zhang, and J. Shi, "Pairwise Critical Point Detection Using Torque Signals in Threaded Pipe Connection Processes," *J. Manuf. Sci. Eng.*, vol. 139, no. 9, p. 091002, 2017.
- [39] C. Zhao, and Y. Sun, "Step-wise sequential phase partition (SSPP) algorithm based statistical modeling and online process monitoring," *Chemometr. Intell. Lab.*, vol. 125, pp. 109-120, 2013.



**Juan Du** received the B.S. degree in Composite Material Science and Engineering from Harbin Institute of Technology, Harbin, China, in 2014.

Currently, she is a Ph.D. student at the Department of Industrial Engineering and Management, Peking University, Beijing. She is currently a visiting scholar at the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta. Her research interests include data analytics for process monitoring, control and diagnosis in complex manufacturing systems.

She is a member of ASQ, INFORMS and IISE.



**Xi Zhang** (M13) received the B.S. degree in Mechanical Engineering and Automation from Shanghai Jiaotong University, Shanghai, China, in 2006, and the Ph.D. degree in Industrial Engineering from the University of South Florida, Tampa, 2010.

Currently, he is an assistant professor at the Department of Industrial Engineering and Management, Peking University, Beijing. His research interests are focused on physical-statistical modeling and analysis for monitoring and diagnosis in complex dynamic systems.

Dr. Zhang is a member of INFORMS, IISE and ASQ.



**Jianjun Shi** received the B.S. and M.S. degrees in electrical engineering from the Beijing Institute of Technology, Beijing, China, in 1984 and 1987, respectively, and the Ph.D. degree in mechanical engineering from the University of Michigan, Ann Arbor, in 1992.

Currently, he is the Carolyn J. Stewart Chair Professor in the H. Milton Stewart School of Industrial and Systems Engineering, Georgia Institute of Technology, Atlanta. His research interests include the fusion of advanced statistical and domain knowledge to develop methodologies for modeling, monitoring, diagnosis, and control for complex manufacturing systems.

Dr. Shi is a Fellow of IISE, ASME, and INFORMS, an elected member of International Statistics Institute, an Academician of the International Academy for Quality (IAQ), and a member of ASQ and ASA.