JOURNAL OF QUALITY TECHNOLOGY
https://doi.org/10.1080/00224065.2018.1541379

Dynamic quality-process model in consideration of equipment degradation

Ran Jin[a], Xinwei Deng[b], Xiaoyu Chen[a], Liang Zhu[c], and Jun Zhang[c]

[a]Grado Department of Industrial and Systems Engineering, Virginia Tech, Blacksburg, Virginia; [b]Department of Statistics, Virginia Tech, Blacksburg, Virginia; [c]Zhejiang Jing Sheng M&E Co. Ltd, Zhejiang, China

ABSTRACT

In many manufacturing processes, equipment reliability plays a crucial role for product quality assurance. It is important to consider the effect of equipment degradation for the quality-process model. In this article, we propose a dynamic quality-process model to characterize the varying effects of a process to product quality due to equipment degradation. The proposed model considers the effects of process variables on product quality as piecewise linear functions with respect to the equipment degradation. It can automatically estimate the dynamic effects via a meaningful parameter regularization, leading to accurate parameter estimation and model prediction. The merits of the proposed method are illustrated by both simulations and a real case study in a crystal growth manufacturing process.

KEYWORDS

data fusion; manufacturing automation; parameter regularization; quality and reliability interaction; varying coefficient

1. Introduction

In many manufacturing processes, automation of manufacturing processes plays an invaluable role for product quality assurance. The effectiveness of manufacturing automation in manufacturing operations relies on well-maintained equipment. When the manufacturing equipment degrades, the manufacturing operations can become unstable in controlling product quality. In current manufacturing practice, quality control or maintenance may not be jointly implemented for both product quality and equipment conditions due to limited understanding of the product quality and equipment reliability interaction. Thus, there is an emerging need to investigate the effect of equipment degradation on product quality.

This research work is motivated by a crystal growth manufacturing (a Czochralski process) with high quality standards but inefficient maintenance operations. Figure 1 shows a crystal growth furnace for silicon ingots. The furnace grows the ingot by dipping a seed crystal to the melt of polycrystalline silicon and pulling up the ingot with rotation (Dai, Wang, and Jin 2014; Zhang et al. 2014). In semiconductor manufacturing, the quality of an ingot plays a crucial role in product quality of all downstream productions. The quality of the ingot (e.g., polycrystalline defects and diameter variation) is affected by various process variables, such as furnace temperature, heating power, pulling speed, and rotation speed of the ingot. Moreover, complex chemical reactions take place inside the furnace, producing byproducts such as dioxide. The deposition of byproducts on the heaters increases their resistivity and reduces the heating efficiency. The increasing resistivity of heaters will also result in an unstable thermal field inside the furnace for ingot manufacturing, which causes quality defects of ingots. In current practice, the resistivity of the heater is used as an equipment degradation variable to guide the maintenance operations. For example, when the resistivity reaches to a certain level, the heater will be replaced by a new one. However, the knowledge on how the degradation will affect the product quality is limited. Therefore, if the process variables are optimized without considering equipment degradation effects, it could lead to many quality defects in the crystal growth manufacturing.

In this article, we focus on incorporating effects of the equipment degradation into the quality-process modeling. However, it is a challenging problem due to insufficient engineering knowledge for quantitatively characterizing the effects of equipment degradation. The importance of manufacturing equipment reliability to product quality was demonstrated in several manufacturing processes (Jin and Chen 2001). As technology advances, data of quality, process and degradation can be automatically collected. Thus, a data-driven modeling strategy can be suitable to address

CONTACT Xinwei Deng xdeng@vt.edu Department of Statistics, Virginia Tech, Blacksburg, VA 24061, USA.
Color versions of one or more figures in the article can be found online at www.tandfonline.com/ujqt

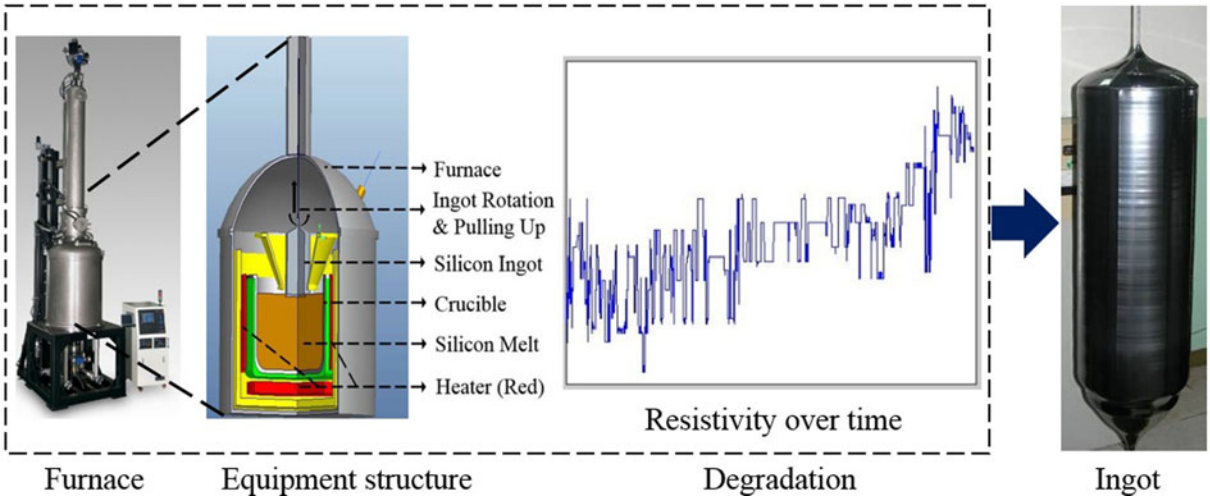

Figure 1. A crystal growth furnace for silicon ingots. Details are described in Section 4. (Redrawn from Zhang et al. (2014) with authors' permission).

the aforementioned research question on how to model quality-process relationships considering the equipment degradation effect.

In the literature, there are various methods to model the quality-process relationship in manufacturing, such as regression analysis (Agrawal, Lawless, and Mackay 1999; Fong and Lawless 1998), experimental design approaches (Jin and Ding 2004; Kang and Joseph 2009), state space models (Shi 2006; Shi and Zhou 2009), and data mining methods (Bukkapatnam et al. 2006; Jin and Shi 2012). However, these methods do not incorporate the equipment reliability information in quality-process models. On the other hand, there are various methods for modeling manufacturing equipment reliability (Moustafa, 1997; Xue and Yang, 1997). The research focused on equipment degradation modeling is also widely conducted (Gebraeel 2006), such as general path models (Lu and Meeker 1993; Meeker and Escobar 1998) and stochastic process models (Aven and Jensen 2013).

There are some recent studies on the interaction between product quality and equipment degradation (Chen and Jin 2005; Jin and Chen 2001). For example, Jin and Chen (2001) proposed a quality-reliability (QR) chain model by considering the quality-reliability effects for the manufacturing system design.

The main idea of these methods is to treat the degradation variable and its interactions with the process variables as additional predictors in the quality-process models. However, such a model formulation can be relatively restrictive due to the assumption that the degradation effects are additive for product quality. Thus, it may not be appropriate to model the dynamic effect of degradation on the quality-process relationship. For example, when the resistivity of heaters is low, the degradation (the increase of the resistivity) will not affect ingot quality significantly because the temperature can be well controlled to the target value. When the resistivity of heaters is high, the degradation will significantly affect ingot quality because the temperature cannot reach the target values due to the low heating efficiency.

To tackle the aforementioned challenges, we propose a dynamic modeling method to characterize the varying degradation effects in the quality-process relationship. Assume that the equipment degradation data are fully observed in the life cycle of the equipment. The case of partially observed degradation will be discussed in Section 5. The proposed method considers the linear quality-process relationship, while the parameters of process variables varies with the equipment degradation. That is, the degradation will change the effects (i.e., parameters in the model) of process variables to the quality response. Furthermore, such dynamic effects are modeled as piecewise constant coefficients of process variables with respect to the degradation variable, which is justified in details in Section 2. It is worth pointing out that the objective of the proposed model is not only to achieve good prediction performance but also to obtain an analytical model form with good engineering explanations. For example, the proposed model can be used to identify the change point and the values of the estimated coefficients according to the degradation variable, as illustrated in Section 5. It is known that data mining methods (Hastie, Tibshirani, and Friedman 2009) may have good prediction performance. But they may not be able to provide an appropriate model structure to understand the interactions of degradation and the quality-process relationship.

For the proposed dynamic model, it involves the estimation of various unknown parameters, including the number of piecewise constant parameters, the corresponding change points over time, and the corresponding values of piecewise constant parameters. To address this issue, we consider the coefficient of every process variable at each time point as a parameter, leading to an over-parameterized model with more unknown parameters than the sample size. Then a penalized least-squares estimation is proposed with the penalty applied on the difference of two consecutive parameters of process variable, naturally leading to estimation of piecewise constant coefficients of the process variables in the model. Thus, we can simultaneously estimate the change points over time and values of the piecewise constant parameters. Moreover, the equipment degradation variable is also used to determine the weight of the penalization, which integrates quality-process information and degradation information in a joint framework. That is, if there is a large change in degradation at two consecutive time points, the corresponding model coefficients can change significantly. This is consistent with the perception of degradation effects in the quality-process models. Because of the dynamic model coefficients driven by the degradation variable, the proposed method can effectively characterize the dynamic manufacturing process due to the equipment degradation and leads to accurate parameter estimation and prediction of quality.

Note that the proposed model can be viewed as a varying coefficient (VC) model (Hastie and Tibshirani 1993). However, our parameter estimation procedure is different from other existing work. For example, Kolar, Song, and Xing (2009) considered a varying coefficient model with piecewise linear coefficient structure. However, their estimation of parameters is a two-stage procedure with the first step on change point detection, which may easily result in a suboptimal estimation. Huang, Wu, and Zhou (2004) developed a varying coefficient model with polynomial spline estimation for the coefficients. Such a model can work well when the underlying model coefficients are continuous over time. But it may not be applicable for the crystal growth manufacturing where the model coefficients could have sudden changes, the changes of the effects of process variables to quality variables due to the degradation effects.

The rest of the article is organized as follows. In Section 2, we describe the proposed model and estimation methods. Section 3 conducts a simulation to evaluate the performance of the proposed dynamic model. We further study the case of the crystal growth manufacturing in Section 4. Finally, we conclude this work with some discussion in Section 5.

2. Dynamic modeling

Denote the quality response as $Y(t)$ and p predictor variables as $X_1(t), \ldots, X_p(t)$. These predictors can be the process variables or their interactions. The goal is to model the quality-process relationship. We start with one equipment degradation variable, denoted as $Z(t)$, mainly responsible for the dynamic effects in the quality-process relationship. Note that the quality-process relationship can vary as the equipment degradation changes. The information of $Z(t)$ provides insights for model formulation and estimation. Therefore, we consider the following model:

$$y(t) = \beta_1(t)x_1(t) + \cdots + \beta_p(t)x_p(t) + \epsilon_t,$$
$$\beta_k(t) = f_k(Z(t)), k = 1, \ldots, p,$$

where $f_k(\cdot)$ can be a nonlinear function. Here $\beta_k(t)$ is considered to be the piecewise constant function with respect to $Z(t)$ and ϵ_t is the error term with zero mean and constant variance.

The motivation of using the piecewise constant for $\beta_k(t)$ is that the degradation effect can be reflected from the change of model coefficients on the quality response. It is a reasonable assumption to approximate the dynamic varying effects of process variables to the quality response. For a general varying effect, the piecewise constant function of $\beta_k(t)$ can well approximate the nonlinear patterns of the parameter $f_k(Z(t))$, in a similar spirit as the classification and regression tree models (Breiman et al. 1984). In manufacturing processes, the effects of process variables to quality response are usually assumed to be static, such as linear regression models (Agrawal, Lawless, and Mackay 1999) and state space models (Shi and Zhou 2009). Therefore, the piecewise constant function consents with the static assumption during a short period of time. When considering a long period of time, the dynamic effects due to degradation can be well approximated by the piecewise constant functions.

2.1. The proposed method

Suppose that the observed data contain the response $y(t), t = 1, \ldots, n$, and p predictor factor $x_1(t), \ldots, x_p(t), t = 1, \ldots, n$. We assume that $Z(t)$ is nonnegative and monotonically increasing with time t, that is, $0 \le Z(1) \le Z(2) \le \cdots Z(n-1) \le Z(n)$. In real practice, if $Z(t)$ is not monotonically increasing, we can perform transformations to $Z(t)$ or sort $Z(t)$ in an increasing order.

To model the piecewise constant function of $\beta_k(t)$, it is difficult to specify the number of piecewise constant parameters, the change points over time, and the corresponding constant parameter values in each piece. To overcome this difficulty, we consider a penalization method to automatically encourage the estimated $\beta_k(t)$ being a piecewise constant function. Specifically, we discretize the function $\beta_k(t)$ onto the observed time points and parameterize its value at each time point as a parameter. That is, we use $\boldsymbol{\beta}_{\mathrm{k}} = (\beta_{k,1}, ..., \beta_{k,n})'$ as a discretized version of $\beta_k(t)$ for $k = 1, ..., p$. The advantage of such a formation is that it enables the flexible function form of $\beta_k(t)$, while its desirable function form can be pursued by imposing proper regularization on those parameters $\beta_{k,1}, ..., \beta_{k,n}$. Clearly, it will lead to an over-parameterized model with n observations and np parameters. Now the parameter vector can be expressed as $\boldsymbol{\beta} = (\boldsymbol{\beta}_1', ..., \boldsymbol{\beta}_{\mathrm{p}}')$, and the model can be written as

$$y(t) = \sum_{k=1}^{p} x_k(t)\beta_{k,t} + \epsilon_t, t = 1, ..., n, \qquad [1]$$

where ϵ_{t} is an error term following a normal distribution with zero mean and variance as σ^2. Without any constraint on the model structure, it is not feasible to estimate the $\beta_{k,t}$'s. To encourage the piecewise constant function form of $\beta_k(t)$, we will estimate the model under the framework of penalized minimization. That is, we consider

$$\min_{\boldsymbol{\beta}} \sum_{t=1}^{n} \left[y(t) - \sum_{k=1}^{p} x_k(t)\beta_{k,t} \right]^2 + \lambda \sum_{k=1}^{p} \sum_{t=2}^{n} w_k(t)|\beta_{k,t} - \beta_{k,t-1}|, \qquad [2]$$

where $\lambda \geq 0$ is a tuning parameter. By imposing the penalty on two consecutive parameters $\beta_{k,t}$ and $\beta_{k,t-1}$, the penalization here attempts to encourage their difference to become zero. Therefore, with the appropriate choice of the tuning parameter λ, the above penalized minimization can largely reduce the number of the parameters in the model. Moreover, the weight $w_k(t)$ provides an adaptive framework of pursuing the piecewise function of $\beta_{k,t}$ to incorporate the degradation information of $Z(t)$. Regarding the choice of weight $w_k(t)$, we propose incorporating the magnitude of change in $Z(t)$ as follows. Let $\tilde{Z}(t) = Z(t) - Z(t-1)$. Then we assign

$$w_k(t) = \frac{1}{\tilde{Z}(t)}. \qquad [3]$$

It means that, for two consecutive $\beta_{k,t-1}$ and $\beta_{k,t}$, if the values of the corresponding $Z(t-1)$ and $Z(t)$ are close to each other, then the values of $\beta_{k,t-1}$ and $\beta_{k,t}$ should also be close to each other. For example, if $Z(t) = Z(t-1)$, which is $\tilde{Z}(t) = 0$, then the penalty $w_k(t)(\beta_{k,t} - \beta_{k,t-1})$ will automatically enforce $\beta_{k,t} = \beta_{k,t-1}$. This is reasonable because it is expected that the model coefficients remain the same when there is no change of the degradation variable values. This also indicates that, when the degradation levels are changing over time, the parameters will also change accordingly. One can clearly see that, when the observed $Z(t)$ is equally spaced, the proposed penalty follows the similar spirit of the fused LASSO (Tibshirani et al. 2005).

Because there is a large number of parameters in the model, an efficient algorithm is needed to address the computational efficiency for parameter estimation. For the penalty terms, we can define projection matrices for a clear expression as follows:

$$\boldsymbol{W}_k = \begin{pmatrix} -w_k(2) & w_k(2) & 0 & ... & 0 \\ 0 & -w_k(3) & w_k(3) & ... & 0 \\ \vdots & \vdots & \ddots & ... & \vdots \\ 0 & ... & 0 & -w_k(n) & w_k(n) \end{pmatrix}.$$

Then the penalty term can be written as

$$\lambda \sum_{k=1}^{p} \sum_{t=2}^{n} w_k(t)|\beta_{k,t} - \beta_{k,t-1}| = \lambda \sum_{k=1}^{p} ||\boldsymbol{W}_k \boldsymbol{\beta}_{\mathrm{k}}||_1,$$

where $||\cdot||_1$ is a vector L_1 norm. Therefore, by defining $\boldsymbol{y} = (y(1), y(2), ..., y(n))'$ and $\boldsymbol{X}_k = diag(x_k(1), ..., x_k(n))$, we can rewrite the objective function in [2] as

$$\min_{\boldsymbol{\beta}} \left(\boldsymbol{y} - \sum_{k=1}^{p} \boldsymbol{X}_k \boldsymbol{\beta}_{\mathrm{k}} \right)' \left(\boldsymbol{y} - \sum_{k=1}^{p} \boldsymbol{X}_k \boldsymbol{\beta}_{\mathrm{k}} \right) + \lambda \sum_{k=1}^{p} ||\boldsymbol{W}_k \boldsymbol{\beta}_{\mathrm{k}}||_1. \qquad [4]$$

If we further define the model matrix $\boldsymbol{X} = (\boldsymbol{X}_1, ..., \boldsymbol{X}_p)$, then the above objective function in [4] can be written as

$$\min_{\boldsymbol{\beta}} (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta})' (\boldsymbol{y} - \boldsymbol{X}\boldsymbol{\beta}) + \lambda ||\boldsymbol{W}\boldsymbol{\beta}||_1, \qquad [5]$$

where $\boldsymbol{W} = diag(\boldsymbol{W}_1, ..., \boldsymbol{W}_p)$. It is clear that the coordinate decent algorithm for a generalized LASSO penalty (Tibshirani and Taylor 2011) can be used to solve the above minimization problem in [5].

2.2. Selection of tuning parameter

Note that there is a tuning parameter λ to be specified for the proposed method. Based on the data, the common methods to select tuning parameters include

Table 1. An example of model coefficients used in the simulation with six significant predictors.

$Z(t)$	$\beta_1(t)$	$\beta_2(t)$	$\beta_3(t)$	$\beta_4(t)$	$\beta_5(t)$	$\beta_6(t)$
$[0, 1.25)$	−0.41	−2.24	−2.13	−0.96	−2.41	−1.21
$[1.25, 4.78)$	−2.88	−1.08	2.08	−3.02	−1.16	0.87
$[4.78, 6.73)$	1.15	1.04	0.72	1.78	0.79	−2.30
$[6.73, 7.47)$	2.27	1.86	−1.37	2.39	−2.39	−2.88
$[7.49, \infty)$	−2.64	−3.11	−2.79	−1.56	2.33	1.59

cross-validation, Akaike information criterion (AIC), and Bayesian information criterion (BIC; McQuarrie and Tsai 1998). When there is a large number of parameters in the model, the extended Bayesian information criterion (EBIC) can be more effective to reduce the model size (Chen and Chen 2008). In this work, we adopt the extended BIC for finding an optimal value of the tuning parameter λ. The extended BIC for the proposed model can be written as

$$EBIC(\lambda) = n\log(\hat{\sigma}^2) + q\log(n) + 2\gamma\log(\tau(q)), \quad [6]$$

where

$$\hat{\sigma}^2 = \frac{\sum_{t=1}^{n}(y(t)-\hat{y}(t))^2}{n}. \quad [7]$$

Here q is the number of nonzero estimates of parameters in the model, $\tau_q = \binom{np}{q}$, and $\gamma = 1-\log_{10}(n)/\log_{10}(p) \in [0,1]$ as recommended by Chen and Chen (2008). Specifically, we generate a grid for λ values, where $\lambda \in \Lambda = \{\lambda_1, ..., \lambda_m\}$. For each grid point λ_j in Λ, we evaluate the corresponding extended BIC score and find the optimal choice of λ that has the minimal extended BIC score among all grid points in Λ.

3. Simulation

Here we conduct simulation studies to evaluate the estimation and prediction performance of the proposed method. Suppose that the underlying quality-process model is $y(t) = \sum_{k=1}^{p} x_k(t)\beta_k(t) + \epsilon_t$, where $x_k(t), k = 1, ..., p$ are p process variables and $\beta_k(t)$ is the corresponding model coefficients. We choose $p = 20$ in this simulation study. Note that the motivation example in crystal growth manufacturing has little knowledge for complex interactions of quality and reliability. It may not be appropriate to consider the data generation based on QR chain models because the QR chain models assume that the degradation effects are additive to the quality response.

Assume that the model coefficient $\beta_k(t)$ varies with the degradation variable $Z(t)$. Specifically, we consider generating $Z(t)$ from $Z(t) = \exp(10^{-3}t)|U(t)|$, where $U(t)$ is a fractional Brownian motion with the Hurst parameter as 0.9, and "$|\cdot|$" represents the operation to take the absolute value. The fractional Brownian motion can be generated by Abry and Sellan's (1996) algorithm. To mimic the dynamic effects of the degradation variable, we generate the piecewise constant functions of $\beta_k(t)$ as follows. First, based on the values of $Z(t)$, we partition its range into five intervals, as shown in Table 1. For each $k, k = 1, ..., p$, we generate a constant value in each interval for $\beta_k(t)$. That is, we have five intervals within the range of degradation variable. For each predictor, the constant value of the parameter in each interval is randomly generated from one of the following normal distributions: $N(-3, 0.09), N(-2, 0.09), N(-1, 0.09), N(1, 0.09)$, and $N(2, 0.09)$. Note that there is not a fixed normal distribution specified for a particular interval to generate constant values of parameters. Table 1 shows an example of model coefficients used in the simulation. Because there are only six significant predictors with nonzero parameters in the model, the table shows the parameters of these six predictors. Based on the values of $Z(t)$, the model coefficients are used as constants during that region of $Z(t)$.

In this simulation, we consider two scenarios of generating data of process variables $x_k(t)$. In scenario 1 (S1), for $k = 1, ..., p$, we generate $x_k(t)$ independently over time to follow a multivariate normal $N(0, \mathbf{\Sigma})$, where $\mathbf{\Sigma}$ is a covariance matrix. In scenario 2 (S2), we generate autocorrelated $x_k(t)$ over time. That is, the first instance $x_k(1)$ is generated from $N(0, \mathbf{\Sigma})$. The rest $x_k(t)$ $(t = 2, ..., n)$ are generated from an AR(1) model as $x_{k,t+1} = 0.7x_{k,t} + w_{k,t}$, where $w_{k,t}$ follows a normal distribution $N(0, 0.04)$. Here the $\mathbf{\Sigma}$ in both scenarios are set to be $\mathbf{\Sigma} = (\rho^{|i-j|})_{p\times p}$ with $\rho^{|i-j|}$ as the element in the ith row and jth column in the matrix, and the correlation parameter $\rho \geq 0$. To conduct a comprehensive simulation study, we vary several settings, including the sample size n, the correlation parameter ρ, and the model sparsity (sp). Here the model sparsity refers to the percentage of significant predictors with nonzero parameters among p predictors in the model. We consider three cases of different sample sizes as $n = 100$, $n = 300$, and $n = 500$. The value of ρ has two levels: 0 and 0.5. The sparsity has two levels: $sp = 30$ percent and $sp = 70$ percent. The sparsity is regarding the percentage of nonzero coefficients out of p coefficients of predictors. After generating $\beta_k(t)$ and $x_k(t)$, we use the model $y(t) = \sum_{k=1}^{p} x_k(t)\beta_k(t) + \epsilon_t$ to generate $y(t)$, where ϵ_t follows an i.i.d. normal distribution $N(0, 0.01)$.

Figure 2 shows an example of the simulated data, including the degradation variable, model coefficients,

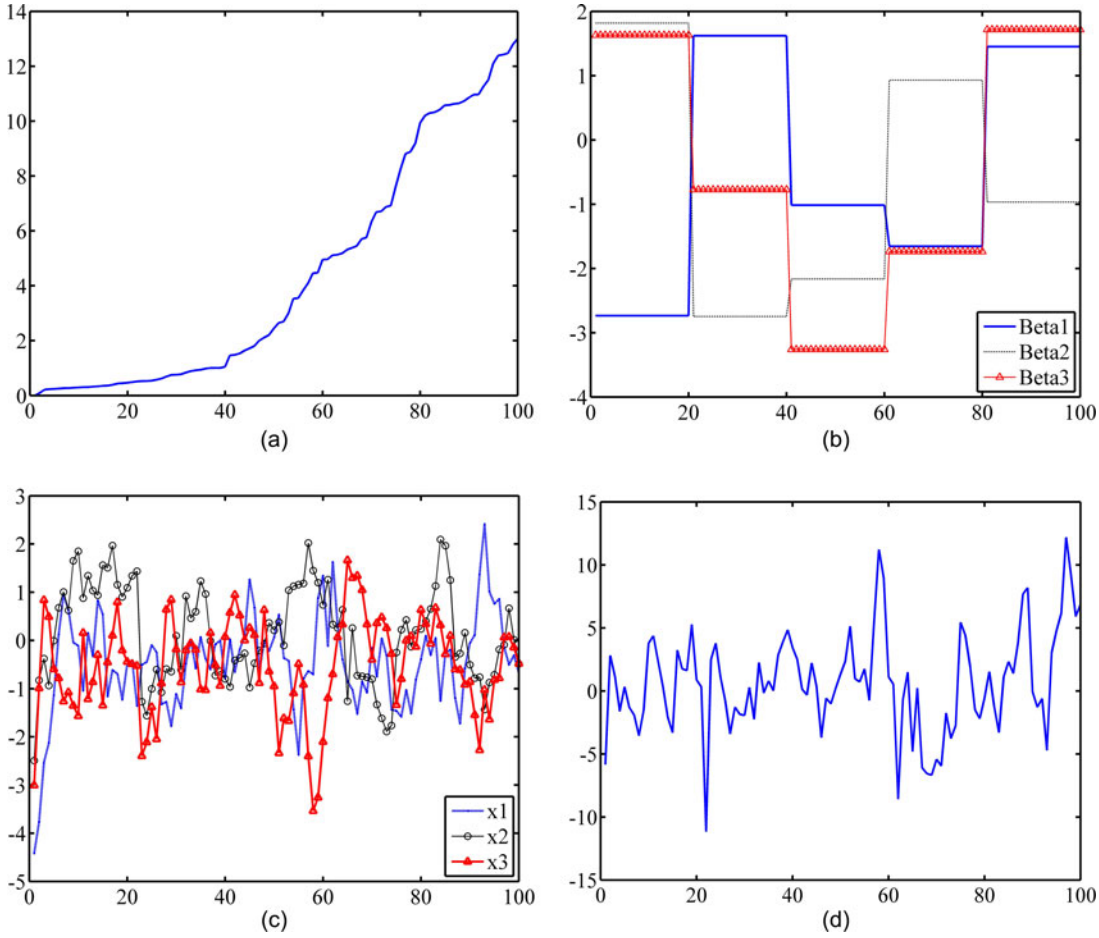

Figure 2. An example of simulated data. (a) Degradation variable over time. (b) Model coefficients as piecewise constant functions over time (for three predictors as an example). (c) Three autocorrelated significant predictors over time. (d) Quality response over time.

process variables, and quality response in S1 with $n = 100, \rho = 0$, and $sp = 30$ percent. It is clear that the model coefficients are generated as piecewise constant functions over time, which is based on the values of the degradation variables.

For each simulation setting, we generate 100 replications of the data sets. The proposed model is compared with two benchmark models in the spirit of QR chain models. The first benchmark model (BM_1) is a linear regression model to predict $y(t)$ based on all process variables, and the degradation variable $Z(t)$ as an additional predictor, that is,

$$BM_1: \quad y(t) = \sum_{k=1}^{p} x_k(t)\beta_k + Z(t)\beta_z + \epsilon,$$

where β_z is the parameter for the predictor $Z(t)$ and ϵ is the model error following a normal distribution with zero mean and a constant variance. The second benchmark model (BM_2) is a linear regression model to predict $y(t)$ based on all process variables, the degradation variable $Z(t)$, and the pairwise interactions between $Z(t)$ and $x_k(t)$, that is,

$$BM_2: \quad y(t) = \sum_{k=1}^{p} x_k(t)\beta_k + Z(t)\beta_z + \sum_{k=1}^{p} x_k(t)Z(t)\eta_k + \epsilon,$$

where η_k's are the parameters for the pairwise interactions and ϵ is the model error following a normal distribution with zero mean and a constant variance. For both BM_1 and BM_2, we adopt the LASSO penalty (Tibshirani 1996) to perform the variable selection, where the tuning parameters are chosen using BIC. The third benchmark model is the MARS method (Friedman 1991) to model $y(t)$ based on all process variables and the degradation variable $Z(t)$ as an additional predictor, that is,

$$\text{MARS}: \quad y(t) = \sum_{m=1}^{M} a_m B_m(\mathbf{x}) + \epsilon,$$

Table 2. Averages and standard errors (in parentheses) of root mean squared prediction error from 100 simulation replications (lowest average values in bold).

			$sp = 30\%$		$sp = 70\%$	
n	ρ	Method	S1	S2	S1	S2
		BM_1	4.47 (0.06)	3.91 (0.07)	7.04 (0.09)	6.12 (0.10)
		BM_2	4.44 (0.15)	3.50 (0.07)	7.62 (0.23)	5.31 (0.10)
$n = 100$	$\rho = 0$	MARS	5.33 (0.10)	4.33 (0.09)	8.68 (0.16)	6.86 (0.15)
		VC	5.60 (0.16)	4.77 (0.22)	9.83 (1.00)	8.81 (0.41)
		Proposed	**3.00 (0.06)**	**3.13 (0.11)**	**6.26 (0.10)**	**5.23 (0.16)**
		BM_1	3.98 (0.06)	4.17 (0.07)	7.28 (0.10)	6.17 (0.09)
		BM_2	3.91 (0.09)	3.77 (0.10)	7.51 (0.19)	5.58 (0.11)
$n = 100$	$\rho = 0.5$	MARS	4.94 (0.09)	4.34 (0.10)	8.74 (0.17)	6.85 (0.13)
		VC	5.11 (0.15)	4.23 (0.23)	9.37 (0.31)	7.18 (0.33)
		Proposed	**3.07 (0.06)**	**3.41 (0.09)**	**5.98 (0.11)**	**5.22 (0.14)**
		BM_1	4.35 (0.03)	3.87 (0.04)	6.53 (0.06)	6.55 (0.06)
		BM_2	3.75 (0.03)	2.91 (0.04)	5.72 (0.06)	5.43 (0.06)
$n = 300$	$\rho = 0$	MARS	4.83 (0.05)	3.77 (0.04)	7.37 (0.08)	6.53 (0.06)
		VC	6.54 (1.24)	2.74 (0.13)	8.49 (0.28)	9.69 (0.70)
		Proposed	**1.32 (0.04)**	**1.52 (0.07)**	**3.37 (0.05)**	**3.39 (0.11)**
		BM_1	3.58 (0.04)	4.31 (0.05)	6.54 (0.05)	6.88 (0.07)
		BM_2	3.21 (0.03)	3.10 (0.04)	5.30 (0.06)	5.57 (0.07)
$n = 300$	$\rho = 0.5$	MARS	3.98 (0.05)	4.16 (0.05)	7.41 (0.08)	6.83 (0.06)
		VC	3.82 (0.13)	7.96 (0.69)	10.56 (0.93)	12.80 (1.13)
		Proposed	**1.56 (0.04)**	**1.52 (0.06)**	**3.05 (0.05)**	**3.46 (0.11)**
		BM_1	4.67 (0.03)	4.58 (0.03)	6.54 (0.04)	6.72 (0.04)
		BM_2	4.26 (0.03)	3.60 (0.03)	5.78 (0.04)	5.00 (0.05)
$n = 500$	$\rho = 0$	MARS	4.97 (0.03)	4.45 (0.04)	7.00 (0.05)	6.61 (0.05)
		VC	8.68 (0.88)	3.68 (0.23)	8.21 (0.61)	7.40 (0.68)
		Proposed	**1.12 (0.03)**	**1.37 (0.06)**	**2.51 (0.05)**	**2.26 (0.06)**
		BM_1	4.36 (0.03)	3.59 (0.03)	6.58 (0.04)	6.30 (0.04)
		BM_2	3.49 (0.03)	3.11 (0.03)	6.05 (0.05)	5.29 (0.04)
$n = 500$	$\rho = 0.5$	MARS	4.67 (0.04)	3.45 (0.03)	7.16 (0.05)	6.28 (0.05)
		VC	4.45 (0.65)	9.10 (0.83)	11.79 (2.00)	18.81 (4.26)
		Proposed	**1.02 (0.03)**	**1.30 (0.06)**	**2.47 (0.04)**	**2.35 (0.05)**

where M is the dimension of knots, which is automatically selected by the algorithm of MARS; a_m is the mth coefficient of the expansion whose values are jointly adjusted to give the best fit to the data, $B_m(\mathbf{x})$ is an indicator function having value one determined by the algorithm of MARS, $\mathbf{x} = (x_1(t), x_2(t), \ldots, x_p(t), Z(t))$, and ϵ is the model error. The fourth benchmark model is the VC model in terms of time t (Kolar, Song, and Xing 2009) with a similar model structure as the model in [1] but with a different loss function.

$$\text{VC: } \min_{\boldsymbol{\beta}} \sum_{t=1}^{n} \left[y(t) - \sum_{k=1}^{p} x_k(t)\beta_{k,t} \right]^2 + \lambda_1 \sum_{k=1}^{p} \sum_{t=1}^{n} |\beta_{k,t}| + \lambda_2 \sum_{k=1}^{p} \sum_{t=2}^{n} |\beta_{k,t} - \beta_{k,t-1}|,$$

where a two-step algorithm proposed is implemented to minimize the loss function.

To check the prediction performance of the methods in comparison, we conduct five-fold cross validation (Hastie, Tibshirani, and Friedman 2009) to compute the root mean squared prediction error (RMSPE). Specifically, we randomly divide the simulated data into five portions $C_1, \ldots, C_5$ based on time instances. Then we use the four portions to fit the models and evaluate the prediction error based on the unused portion. By iteratively using every portion for prediction, we can calculate RMSPE as

$$RMSPE = \sqrt{\frac{1}{n} \sum_{k=1}^{K} \sum_{t \in C_k} (y(t) - \hat{y}(t))^2}, \qquad [8]$$

where $\hat{y}(t)$ is the predicted response. Table 2 summarizes the average RMSPEs and their standard errors over 100 replications (in parentheses) for the three models. Clearly, the proposed model has the smallest RMSPEs in most scenarios. This is because the advantage of the proposed model lies in the flexibility of dynamically adjusting the parameters of process predictors according to the degradation. Note that, when the sample size n is relatively small with $n = 100$, the proposed model has comparable prediction performance as BM_2, under the scenario S2 with the observations $x_k(t)$ being autocorrelated. One possible reason is that the proposed method involves a large number of parameters for estimation. When the observations are autocorrelated over time, there are less new information for parameter estimation. As the sample size increases in the case of $n = 300$ and $n = 500$, the proposed method performs much better than the benchmark models BM_1, BM_2, MARS, and VC. The

Table 3. The average and standard errors (in parentheses) of parameter estimation accuracy of $\boldsymbol{\beta}$ from 100 simulation replications (numbers are in percentage and lowest values are in bold).

			$sp = 30\%$		$sp = 70\%$	
n	ρ	Method	$S1$	$S2$	$S1$	$S2$
		BM_1	115.21 (1.72)	133.47 (2.86)	105.93 (0.83)	112.42 (1.02)
		BM_2	157.74 (5.21)	157.71 (6.30)	128.88 (3.86)	136.51 (3.75)
$n = 100$	$\rho = 0$	MARS †	–	–	–	–
		VC	129.07 (4.80)	133.03 (3.68)	123.07 (10.34)	118.73 (1.92)
		Proposed	**108.13 (1.22)**	**116.59 (1.34)**	**104.61 (0.80)**	**103.32 (0.88)**
		BM_1	**107.53 (1.28)**	137.27 (3.06)	**105.30 (1.01)**	112.35 (1.17)
		BM_2	149.48 (5.47)	183.29 (7.75)	140.60 (5.11)	139.62 (4.93)
$n = 100$	$\rho = 0.5$	MARS	–	–	–	–
		VC	124.88 (2.91)	132.14 (9.69)	115.40 (1.51)	123.97 (3.90)
		Proposed	118.26 (1.53)	**119.90 (1.84)**	107.85 (0.87)	**108.35 (0.81)**
		BM_1	104.72 (0.82)	118.07 (1.46)	100.59 (0.37)	105.18 (0.55)
		BM_2	109.12 (0.99)	127.49 (1.53)	102.72 (0.48)	112.59 (1.02)
$n = 300$	$\rho = 0$	MARS	–	–	–	–
		VC	127.43 (8.42)	103.64 (2.87)	100.41 (1.25)	111.73 (1.79)
		Proposed	**56.09 (0.79)**	**50.69 (0.76)**	**64.56 (0.42)**	**69.07 (0.62)**
		BM_1	100.23 (0.51)	118.17 (1.53)	99.06 (0.28)	104.32 (0.52)
		BM_2	106.55 (0.85)	133.65 (1.99)	112.90 (1.68)	116.04 (1.12)
$n = 300$	$\rho = 0.5$	MARS	–	–	–	–
		VC	116.19 (2.29)	138.74 (2.47)	107.41 (1.01)	108.26 (1.41)
		Proposed	**66.07 (1.04)**	**55.57 (0.87)**	**61.91 (0.46)**	**60.64 (0.60)**
		BM_1	101.12 (0.59)	115.32 (1.29)	99.29 (0.29)	103.20 (0.38)
		BM_2	103.50 (0.63)	121.30 (1.30)	100.53 (0.31)	111.27 (0.87)
$n = 500$	$\rho = 0$	MARS	–	–	–	–
		VC	120.58 (3.03)	98.88 (2.47)	86.50 (1.24)	90.51 (3.03)
		Proposed	**40.73 (0.49)**	**37.39 (0.70)**	**58.92 (0.54)**	**53.40 (0.58)**
		BM_1	100.29 (0.41)	108.93 (1.04)	97.99 (0.29)	102.92 (0.41)
		BM_2	104.20 (0.76)	112.36 (1.23)	100.58 (0.49)	106.13 (0.57)
$n = 500$	$\rho = 0.5$	MARS	–	–	–	–
		VC	92.98 (1.49)	127.33 (2.65)	103.76 (1.42)	99.03 (1.84)
		Proposed	**48.12 (0.62)**	**50.05 (0.68)**	**60.68 (0.50)**	**60.97 (0.48)**

Note. †The parameter estimation accuracy metric is not applicable for MARS model.

explanation for relatively large standard errors of RMSPE in the VC model may be due to the use of randomized LASSO, which may result in suboptimal estimations, as we argued in Section 1.

We further examine the parameter estimation accuracy of the proposed model and the two benchmark models. Note that the parameters of the underlying models are piecewise constant functions over time. To evaluate the accuracy of the estimated parameters, we compare the percentage of difference between the estimated parameters and the true parameters. Specifically, we define the performance measure as

$$PM = \frac{\sum_{k=1}^{p}\sum_{t=1}^{T}|\hat{\beta}_{k,t} - \beta_{k,t}|}{\sum_{k=1}^{p}\sum_{t=1}^{n}|\beta_{k,t}|} \times 100\%,$$

where $\hat{\beta}_{k,t}$ is the value of the estimated parameters for the kth predictor at time t and $\beta_{k,t}$ is the true parameter for the kth predictor at time t in the underlying model. The quantity represents the percentage of the difference area between two curves $\hat{\beta}_{k,t}$ and $\beta_{k,t}$ in the total areas under $\beta_{k,t}$ for all predictor variables. The smaller the percentage is, the better the estimation accuracy of the parameters is. For BM_1 and BM_2, the parameters are estimated as constants over time, so they will be treated as constant lines in the comparison. For BM_2, because it involves both predictors $x_k(t)$, $Z(t)$ and interactions $x_k(t)Z(t)$, we only compare the parameters of $x_k(t)$ in the model. For the MARS model, because it is a polynomial spline-based regression and the parameter estimation accuracy metric is not applicable for it, we do not compare its parameter estimation errors in Table 3. For the VC model, which involves both predictors $x_k(t)$ and $Z(t)$, here we only compare the parameters of $x_k(t)$ in the model in Table 3. Table 3 summarizes the performance measure of the parameter estimation accuracy for three models. Clearly, the proposed dynamic model also gives the best parameter estimation accuracy in most scenarios. However, most of the accuracies are shown to be larger than 50 percent. A possible explanation is that the piecewise constant on parameter coefficients and independent assumption of the residuals of the proposed model is not fully satisfied (see Section 5 for discussion). In summary, the results from the simulation study indicate that the proposed dynamic models can outperform the

Table 4. Process variables, degradation variable, and quality response in the crystal growth process.

Types	Variable names	Physical explanations
Process variables	Power	The power of the heater
	Temperature ingot	Temperature of the top part of ingot
	Temperature heater	Temperature of the heater
	Pulling speed	The pulling speed of the ingot
	Rotation speed	The rotation speed of the ingot
Degradation variable	Resistivity	The resistivity of the heater
Quality response	Ingot diameter	The diameter of the ingot

conventional models on both prediction and parameter estimation.

4. Case study: Crystal growth manufacturing

In this section, we apply the proposed dynamic model to a real case study of the crystal growth process. Recall the crystal manufacturing described in Section 1. Because the heating efficiency is affected by the degradation of the heater, the effects of the process variables on the quality are expected to be dynamic over time. Table 4 summarizes the variables of data collected from the manufacturing process. The ingot diameter is chosen as the quality response because poor control of the diameter will cause huge waste in crystal manufacturing (Zhang et al. 2014). In total, there are over 1000 samples collected over time in the case study. Note that the flow time of crystal growth usually is much shorter than the life cycle of the equipment. Because of the continuous growth of the ingot in different batches, both the quality response (the diameter) and the degradation variable (the resistivity) are measured at every time instance. When the time scales for quality response and degradation variable are different, the collected data are sampled and aligned to make the time scale consistent. The response here is a centered response (i.e., a zero-mean response).

To model the quality-process relationship, we use the main effects of process variables as predictors in the models. The variable "resistivity" is considered as the degradation variable. For the benchmark models, we consider the degradation variable "resistivity" as an additional predictor in BM_1 and consider all the pairwise interactions of the resistivity and the process variables as additional predictors in BM_2. Figure 3 shows the estimated model coefficients obtained from the proposed dynamic modeling method. Specifically, Figure 3(a) shows the change of the resistivity sorted in an increasing order. Figures 3(b)–(f) show the estimated model coefficients for power, temperature ingot, temperature heater, pulling speed, and rotation speed. Clearly, the estimated parameters vary significantly at different time points, which is driven by the degradation of the heater. Taking the power as an example (Figure 3(b)), it generally has bigger impact on the ingot quality (ingot diameter) at the early stage when the resistivity is small than at the later stage. This is consistent with the engineering knowledge that, when the resistivity is small, the power can be easily adjusted to target values. Thus, it can affect the ingot quality significantly. When the resistivity of heaters becomes larger, the heating efficiency is limited. In this case, adjusting the power would not largely change the heating performance. Therefore, the parameter of power becomes a small constant in the later stage of the manufacturing process when the resistivity become large, as shown in Figure 3(b). Taking the pulling speed as another example, it is clear that the pulling speed has significant impact on the ingot quality when the resistivity is small, but the variable becomes less influential to the ingot quality when the resistivity is large (i.e., the parameter of pulling speed becomes close to zero). The result implies that no matter how the pulling speed is adjusted, the ingot quality will not be significantly changed when the resistivity is large. Besides, it can be easily noticed that Figures 3(b)–(f) start with small values near zeros. This can be attributed to the sorting of $Z(t)$. Specifically, because the degradation measure $Z(t)$ is sorted in increasing order along the time t, it means that the sorted degradation measures start with small values. As a result, small quality defects (i.e., small $y(t)$ with small variation) will also occur in the beginning. Therefore, the fitted coefficients will start with small values. In summary, the estimated varying coefficients from the proposed model can reflect the dynamic effects of the process variables due to the degradation of the heater. Driven by the degradation variable, abrupt changes commonly occurred in the estimated varying coefficients. The conventional varying coefficient model in the literature may not be directly applicable in this case.

To further show the effectiveness of the proposed model, we use cross validation with 100 folds to evaluate the prediction performance. For each iteration, we

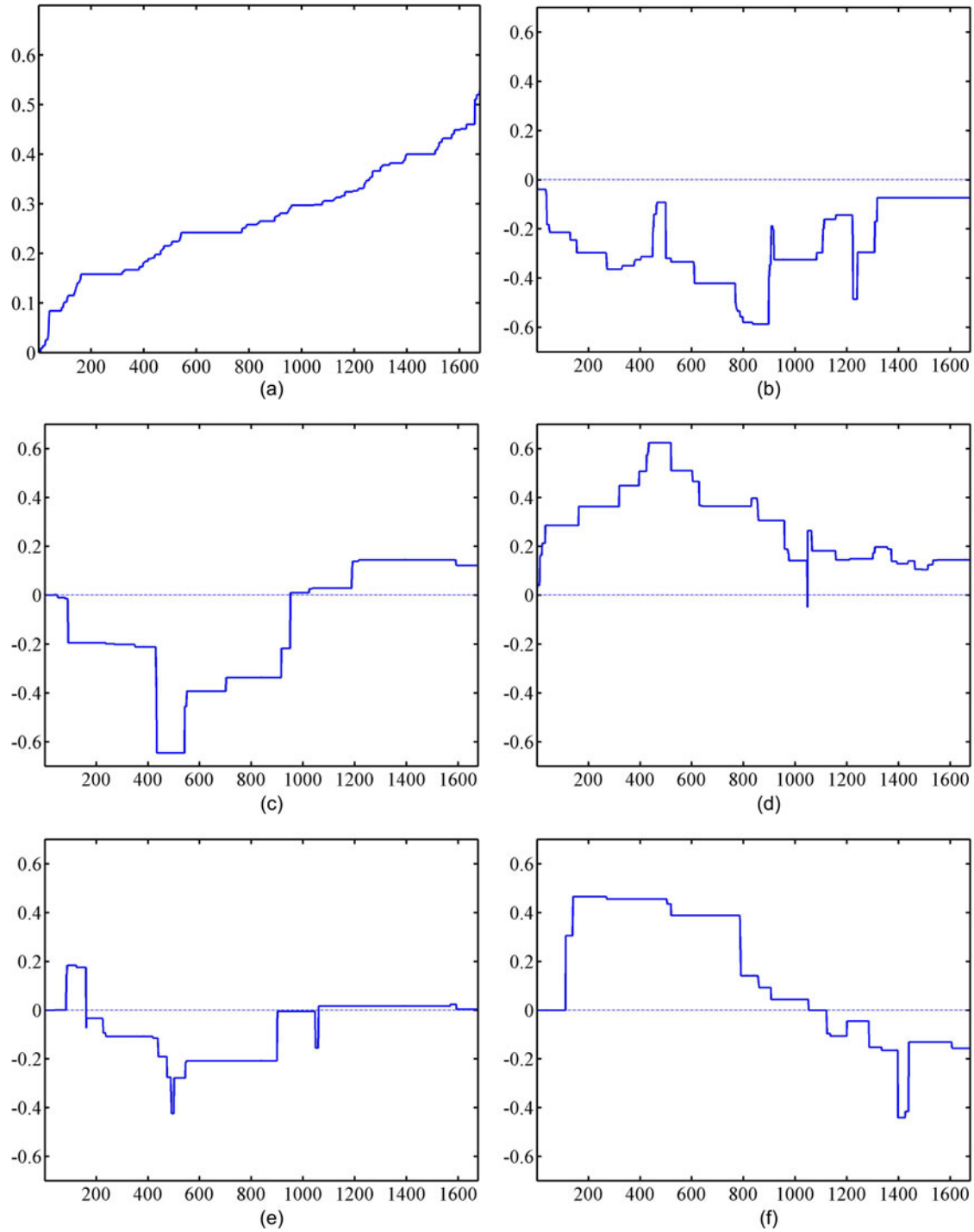

Figure 3. Performance of the proposed method in the crystal growth process. (a) Resistivity of the heater sorted in increasing order. (b) Estimated model coefficient of power. (c) Estimated model coefficient of temperature ingot. (d) Estimated model coefficient of temperature heater. (e) Estimated model coefficient of pulling speed. (f) Estimated model coefficient of rotation speed.

partition the data into the training set and the test set. Then we calculate the RMSPE in the test set using the model estimated from the training set. Figure 4 shows the box plots of the prediction errors of the proposed dynamic models and two benchmark models in 100-fold of cross validation. It should be noted that, although the box plot of prediction errors of the VC model is presented with good performance, the VC model is not applicable for prediction in this case study because it is based on the partitions of time, which are not replicable for real manufacturing processes. This VC model can be only used for process analysis. From Figure 4, it is clear that the proposed dynamic model outperforms BM_1, BM_2, and MARS for prediction,

indicating that the dynamic model can better characterize the quality-process relationship with the consideration of equipment degradation.

5. Discussion

In modern manufacturing, automation plays an important role for quality assurance. When the equipment degrades, the manufacturing process becomes less robust with low product quality. The dynamic interaction of equipment reliability and product quality is not fully understood in the literature. In this article, we propose a dynamic modeling strategy to integrate equipment degradation information for quality-process modeling, enabling effective automation of manufacturing processes to improve product quality. The effects of process variables are modeled as piecewise constant functions according to the equipment degradation, which can be useful for quality control (Qiu 2014; Woodall and Montgomery 2014). The proposed dynamic model provides an opportunity to conduct joint decision making for quality control and maintenance scheduling. Such a model can play a significant role in future quality engineering as the increasing customization of products results in fewer samples of products for quality control. Many traditional quality control methods, which require a large sample size, may not be applicable in these applications. One possible solution is to pursue high product quality through the optimized process settings and well-maintained manufacturing equipment. Therefore, the quality-process modeling needs to take account of the effects of equipment degradation.

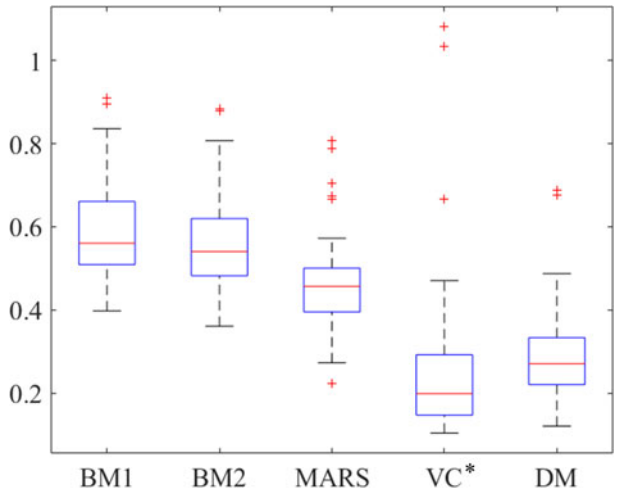

Figure 4. Box plots of prediction errors for three methods in 100-fold of cross validation for the crystal growth process. Note that, although the boxplot of VC prediction errors is presented, the VC model is not applicable for prediction.

For the proposed method, the model assumption may not be fully suitable for the real data of crystal manufacturing. Figure 5 shows two diagnostic figures for the residuals of the proposed method. It appears that the residuals may not strictly follow the normal distribution and there may be some autocorrelation in the residuals. It would be interesting to study how to incorporate the temporal information into the proposed method. For manufacturing systems with equipment degradation, we note that the degradation often changes over time. One may consider time series modeling techniques to characterize the autocorrelations of the residuals. The modeling of the autocorrelation can also characterize the trend or seasonality due to the degradation and maintenance operations.

There are also several directions for future research. First, instead of using linear models for the quality-process relationship, nonlinear models will be explored to characterize complicated manufacturing processes. Second, the dynamic effects in the quality-

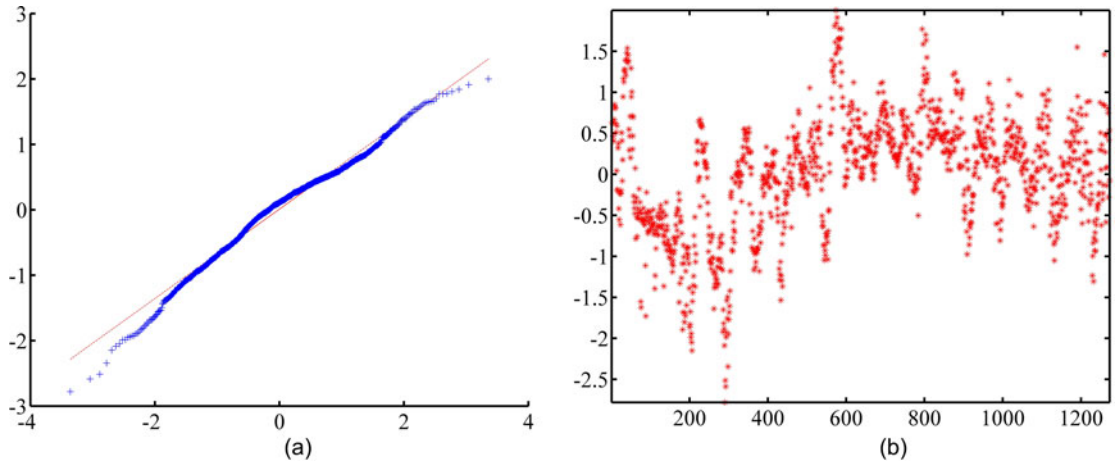

Figure 5. Diagnostics for the estimated model in the crystal growth process. (a) The QQ-plot of residuals. (b) The scatter plot of residuals versus time.

process model are not restricted to piecewise constant function of parameters of process variables. We will further explore complex model structures for dynamic effects. For example, we can incorporate piecewise linear functions to approximate more complex dynamic effects. Third, we can allow additional L_1 penalties on the model coefficients to encourage a sparse model. The estimation algorithm in this work can also be extended to maintain the computational efficiency.

Fourth, it is likely in practice that the degradation data are partially observed. For example, at a certain time, the degradation data for future equipment condition are not observed. In this case, one can model the relationship between the degradation variable and the corresponding factors, such as process variables and ambient variables. The future degradation trend can be predicted, similar to the general path model incorporating the dynamic covariate information (Hong and Meeker 2013). Moreover, because the measurement noises typically exist for degradation signals, the effects of signal-to-noise ratio will be investigated in the future to study the robustness of the proposed model to the degradation noise. Fifth, how the product quality affects equipment degradation is also very important (Hao 2015). In the crystal growth process, higher quality requirements usually call for more demanding use of heaters, leading to quicker degradation. We will consider modeling the quality as a covariate in the degradation model (Hong and Meeker 2013). The modeling and parameter estimation on this aspect will be further investigated. Finally, the dynamic models can be used for joint quality control and maintenance scheduling. By integrating the degradation modeling (Lu and Meeker 1993) and automatic process control (Zhong, Shi, and Wu 2010), we will develop joint quality control and equipment maintenance strategies. It thus will bring significant benefit for manufacturing quality and equipment reliability improvements.

About the authors

Dr. Jin is an assistant professor in the Grado Department of Industrial and Systems Engineering. Email address: jran5@vt.edu.

Dr. Deng is an associate professor in the Department of Statistics. Email address: xdeng@vt.edu.

Mr. Chen is a Ph.D. candidate in the Grado Department of Industrial and Systems Engineering. Email address: xiaoyuch@vt.edu.

Mr. Zhu is a vice president at Zhejiang Jing Sheng M&E Co. Ltd, Zhejiang, China. Email address zhuliang@jsjd.cc

Mr. Zhang is a vice president at Zhejiang Jing Sheng M&E Co. Ltd, Zhejiang, China. Email address: zhangjun@jsjd.cc

Acknowledgment

The authors would also like to acknowledge Dr. Kaibo Wang from the Department of Industrial Engineering at Tsinghua University for his insightful comments and constructive suggestions in data collection for this work.

Funding

The authors gratefully acknowledge the funding support of NSF CMMI-1634867.

ORCID

Ran Jin http://orcid.org/0000-0003-3847-4538
Xinwei Deng http://orcid.org/0000-0002-1560-2405
Xiaoyu Chen http://orcid.org/0000-0002-1870-5290

References

Abry, P., and F. Sellan. 1996. The wavelet-based synthesis for the fractional Brownian motion proposed by F. Sellan and Y. Meyer: Remarks and fast implementation. *Applied and Computational Harmonic Analysis* 3 (4):377–83.

Agrawal, R., J. F. Lawless, and R. J. Mackay. 1999. Analysis of variation transmission in manufacturing processes-part II. *Journal of Quality Technology* 31 (2):143–54.

Aven, T., and U. Jensen. 2013. *Stochastic models in reliability*. 2nd ed. New York, NY: Springer.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. *Classification and regression trees*. Belmont, CA: Wadsworth.

Bukkapatnam, S., P. Agrawal, M. Malshe, L. Raff, M. Hagan, and R. Komanduri. 2006. Parameterization of interatomic potential functions using a genetic algorithm accelerated with a neural network. *Physical Review B* 74:224102.

Chen, J. H., and Z. H. Chen. 2008. Extended Bayesian information criteria for model selection with large model spaces. *Biometrika* 95 (3):759–71.

Chen, Y., and J. Jin. 2005. Quality-reliability chain modeling for system-reliability analysis of complex manufacturing processes. *IEEE Transactions on Reliability* 54 (3):475–88.

Dai, C., K. Wang, and R. Jin. 2014. Monitoring profile trajectories with dynamic time warping alignment. *Quality and Reliability Engineering International* 30 (6):815–27.

Fong, D., and J. F. Lawless. 1998. The analysis of process variation transmission with multivariate measurements. *Statistica Sinica* 8:151–64.

Friedman, J. H. 1991. Multivariate adaptive regression splines. *The Annals of Statistics* 19 (1):1–67.

Gebraeel, N. 2006. Sensory-updated residual life distributions for components with exponential degradation patterns. *IEEE Transactions on Automation Science and Engineering* 3 (4):382–93.

Hao, L. 2015. Residual life prediction and degradation-based control of multi-component systems. PhD diss., Georgia Institute of Technology.

Hastie, T., and R. Tibshirani. 1993. Varying-coefficient models. *Journal of the Royal Statistical Society Series B (Methodological)* 55:757–C796.

Hastie, T., R. Tibshirani, and J. Friedman. 2009. *The elements of statistical learning: Data mining, inference, and prediction*. 2nd ed. Berlin, Germany: Springer-Verlag.

Hong, Y., and W. Q. Meeker. 2013. Field-failure predictions based on failure-time data with dynamic covariate information. *Technometrics* 55 (2):135–49.

Huang, J. Z., C. O. Wu, and L. Zhou. 2004. Polynomial spline estimation and inference for varying coefficient models with longitudinal data. *Statistical Sinica* 14: 763–88.

Jin, J., and Y. Chen. 2001. Quality and reliability information integration for design evaluation of fixture system reliability. *Quality and Reliability Engineering International* 17 (5):355–72.

Jin, J., and Y. Ding. 2004. Online automatic process control using observable noise factors for discrete part manufacturing. *IIE Transactions* 36 (9):899–911.

Jin, R., and J. Shi. 2012. Reconfigured piecewise linear regression tree for multistage manufacturing process control. *IIE Transactions* 44 (4):249–61.

Kang, L., and V. R. Joseph. 2009. Bayesian optimal single arrays for robust parameter design. *Technometrics* 51 (3): 250–61.

Kolar, M., L. Song, and E. P. Xing. 2009. Sparsistent learning of varying-coefficient models with structural changes. *In Proceedings of Advances in Neural Information Processing Systems* 22:1006–14.

Lu, C. J., and W. O. Meeker. 1993. Using degradation measures to estimate a time-to-failure distribution. *Technometrics* 35 (2):161–74.

McQuarrie, A. D., and C. L. Tsai. 1998. *Regression and time series model selection*. Singapore: World Scientific.

Meeker, W. Q., and L. A. Escobar. 1998. *Statistical methods for reliability data*. New York, NY: John Wiley & Sons.

Moustafa, M. S. 1997. Reliability analysis of K-out-of-N: G systems with dependent failures and imperfect coverage. *Reliability Engineering and System Safety* 58 (1):15–7.

Qiu, P. 2014. *Introduction to statistical process control*. Boca Raton, FL: Chapman & Hall/CRC.

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

Shi, J., and S. Zhou. 2009. Quality control and improvement for multistage systems: A survey. *IIE Transactions* 41 (9): 744–53.

Tibshirani, R. 1996. Regression shrinkage and selection via the LASSO. *Journal of the Royal Statistical Society, Series B* 58:267–88.

Tibshirani, R., M. Saunders, S. Rosset, J. Zhu, and K. Knight. 2005. Sparsity and smoothness via the fused LASSO. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 67 (1):91–08.

Tibshirani, R. J., and J. Taylor. 2011. The solution path of the generalized LASSO. *The Annals of Statistics* 39 (3): 1335–71.

Woodall, W. H., and D. C. Montgomery. 2014. Some current directions in the theory and application of statistical process monitoring. *Journal of Quality Technology* 46 (1): 78–94.

Xue, J., and K. Yang. 1997. Upper and lower bounds of stress-strength interference reliability with random strength-degradation. *IEEE Transactions on Reliability* 46: 142–5.

Zhang, J., W. Li, K. Wang, and R. Jin. 2014. Process adjustment with an asymmetric quality loss function. *Journal of Manufacturing Systems* 33 (1):159–65.

Zhong, J., J. Shi, and C. F. Wu. 2010. Design of DOE-based automatic process controller with consideration of model and observation. *IEEE Transactions on Automation Science and Engineering* 7:266–73.