Run-to-Run Process Adjustment Using Categorical Observations

KAIBO WANG

Tsinghua University, Beijing 100084, P. R. China

FUGEE TSUNG

Hong Kong University of Science and Technology, Kowloon, Hong Kong

When quality characteristics cannot be measured on a continuous scale due to either inherent or outside constraints, qualitative observations can be collected alternatively. Under this situation, most conventional run-to-run (R2R) process-control algorithms that are developed based on quantitative measurements cannot be implemented. In this paper, we develop a run-to-run control scheme that uses qualitative information for process adjustments. A two-phase modeling and adjustment strategy is introduced and demonstrated by a real example from a deep reactive ion etching (DRIE) process: model building and parameter estimation is performed in Phase I, and a latent-model control law, a categorical R2R controller, is developed for process regulation in Phase II. Simulation results show that the proposed algorithm exhibits competitive control performance, less adjustment effort, and a larger stability region than the conventional exponentially weighted moving average (EWMA) controller.

Key Words: Categorical Variable; Cautious Control; Exponentially Weighted Moving Average Controller; Run-to-Run Process.

Run-to-run (R2R) control techniques have been widely adopted by the semiconductor industry for quality assurance (Sachs et al. (1995), Del Castillo and Hurwitz (1997)). The target processes usually contain uninterruptible cycles, which are termed as runs. A recipe that describes the setting of controllable factors is suggested by an R2R controller before each cycle and cannot be changed during the processing cycle. However, the recipe is allowed to vary from run to run to compensate for newly emerged deviations.

Extensive research exists reported in the literature regarding the regulation of R2R processes and several R2R controllers have been proposed. Among others, the exponentially weighted moving average (EWMA) controller (Sachs et al. (1995)) can compensate for integrated moving average (IMA) noises and weak auto-correlations of a linear process (Del Castillo and Hurwitz (1997)); the predictor corrector control (PCC) controller (Butler and Stefani (1994)) is capable of compensating for deterministic process drifts; the self-tuning controller is robust to both shifts and deterministic drifts even if the process exhibits strong autocorrelation (Del Castillo and Hurwitz (1997)). Tsung and Shi (1999) investigated more advanced proportional-integral-derivative (PID) controllers in R2R process control. All the popular controllers are designed to work on the basis of quantitative observations. A comparative analysis of the popular controllers is referred to Zhe et al. (1996), Campbell et al. (2002), and Zhang et al. (2003).

In practice, timely quantitative measurements are not always available due to various practical constraints. For example, etching is an important stage in micro-electro-mechanical systems (MEMS) fabrication, and deep reactive ion etching (DRIE) is a popular etching technique for forming desired patterns on wafers. In the DRIE process, the off-site measurement of an etched wafer needs to be carried

Dr. Wang is an Assistant Professor in the Department of Industrial Engineering.

Dr. Tsung is an Associate Professor in the Department of Industrial Engineering and Logistics Management. He is a Senior Member of ASQ. His email address is season@ust.hk.

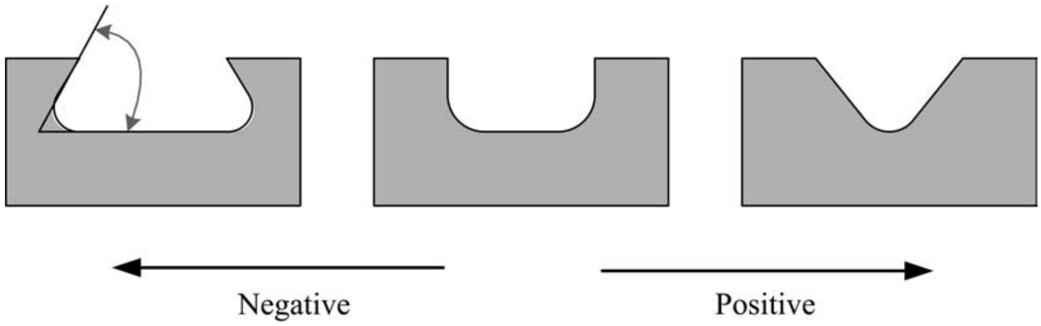

FIGURE 1. Illustrations of Positive and Negative Etching Profiles from a DRIE Process.

out with the aid of a scanning electron microscopy (SEM) in laboratories. This operation could be a major bottleneck in high-volume production lines. Alternatively, some authors have suggested utilizing enhanced algorithms to compensate the lack of timely information. Baras and Patel (1997) presented a robust R2R control approach that aims to minimize the worst-case performance. Good and Qiu (2002) introduced a multistep forecasting strategy to estimate process output and parameters in advance. Nevertheless, all these approaches have increased process variances and made the stability region smaller.

One of the compensatory techniques for metrology delay is to collect observations in a faster way by sacrificing accuracy. Qualitative observations can usually be collected much easier if no or less interactive laboratory operations are involved. Consider the DRIE process in MEMS applications, one of the most important quality characteristics of interest is the trench profile (May et al. (1991)). The desired profile is the one with smooth and vertical sidewalls, which is referred to as an anisotropic profile. However, positive or negative profiles will be seen if the machine is not finely tuned. Figure 1 outlines several deviated profiles from the DRIE process. Instead of measuring the angles of the sidewalls accurately, which usually costs longer aligning time for different shapes, the profiles can be classified as "positive," "normal," and "negative" based on the verticality of the trenches.

There are situations under which quality characteristics are measured by categorical variables in a natural way. In the DRIE example, sidewalls are not always as smooth as those shown in Figure 1. Ball-shaped, bottle-shaped, or even more complicated profiles may be generated (Ayon et al. (1999)) . In Figure 2, some irregular profiles are recorded. The sidewalls of the profiles are difficult to measure exactly in terms of degrees because of their nonlinear shapes. However, it is still possible to distinguish a positive profile from a negative one and to judge whether it is overetched or underetched. Therefore, it is advantageous to use categorical scales to convey visual impressions under these situations. Another example is given by Spanos and Chen (1997), in which the photoresist line profiles from a dry develop process with mouse bites are measured on an ordinal categorical scale as "very rough," "rough," "smooth," "very smooth," etc.

A qualitative dataset, in spite of its low accuracy, contains substantially useful information and may be used to directly affect the quality of products or the performance of processes. Conventional control methodologies have provided a framework that is re-

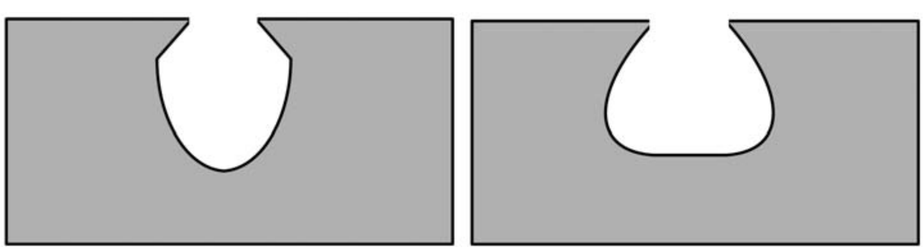

FIGURE 2. Illustrations of Irregular Etching Profiles from a DRIE Process.

stricted to quantitative measurements. However, research on R2R process control using qualitative information has not drawn much attention so far. Spanos and Chen (1997) initiated some work and demonstrated the importance and possibility to use qualitative information for R2R process monitoring and control in a real example by assuming a static model with specified noise type. If a dynamic disturbance model is presented, their method will fail to track process variations and result in poor control performance. In addition, their method involves a nonlinear optimization problem and extensive calculations are needed for model estimation and recipe generation. This feature has made it difficult for practitioners to implement and extend. Categorical variables, which by convention are called linguistic variables, are also seen in fuzzy control theory (Passino and Yurkovich (1998)). However, a fuzzy logic controller requires numerical observations as its input. Furthermore, the controller may assign each observation into more than one linguistic sets based on the membership functions, which is different from the mutually exclusive categories considered in this paper (Laviolette et al. (1995)).

In this paper, we aim to develop a novel R2R control algorithm that makes use of categorical measurements to recommend adjustments. The rest of this paper is organized as follows. The modeling issues of categorical variables are introduced in the next section. Following that, a two-phase R2R control strategy, termed a categorical controller, is developed and illustrated by a real example from DRIE. Then the corresponding stability analysis and performance comparison are conducted. The last section concludes this paper with some future research directions.

The Modeling of Categorical Variables

A categorical variable that has inherent ordered values is called an ordinal variable (Agresti (1990)). With the presence of ordinal response variables, it is essential to establish their functional relationship with influential factors. This section will focus on the modeling issues in R2R processes with categorical observations.

Latent Variables and Latent Models

When the functional relationship between a response variable and explanatory variables is difficult to establish directly, a possible solution is to introduce an intermediate variable to bridge the gap and split the function into two models. One model illustrates the way that the intermediate variable is influenced by the explanatory variables, while the other model describes the mapping relationship between the response variable and the intermediate variable. The intermediate variable is called a latent variable, as it is used for the establishment of the models. The corresponding models that comprise the latent variable are called latent models. For example, when studying the proportion employed and education level in social science, the propensity to work is a latent variable that links the response variable and the explanatory variable.

Considering an R2R process that is described by the following model (see Apley and Kim (2004), Del Castillo and Hurwitz (1997), and the references therein):

$$y_t = \alpha + \beta u_{t-1} + d_t, \tag{1}$$

where y_t, the output of the process at step t, corresponds to the control action at $(t-1)$, u_{t-1}. The vector (α, β) contains the parameters that describe the process. d_t is a disturbance series. As suggested by Apley and Kim (2004) and Montgomery et al. (2000), the IMA(1, 1) time series is appropriate for modeling disturbances in many industrial processes. Therefore, we assume d_t follows an IMA(1, 1) model, which takes the form $d_{t+1} = d_t + \varepsilon_{t+1} - \theta\varepsilon_t$, where ε_t is a white noise that follows a standard normal distribution, $\varepsilon_t \sim \mathrm{N}(0, 1)$.

Note that the linear model is rather representative in R2R process control and has been considered by many researchers. Palmer et al. (1996) studied photolithography processes using experimental methods and established linear models that take the same form to illustrate the relationship between resist thickness and spin speed. Sachs et al. (1995) used a similar linear model in their seminal work on R2R control. Ruegsegger et al. (1999) modeled the bias in an etching process with a similar linear model. May et al. (1991) obtained a linear model with a similar form to illustrate the anisotropy of an etching process. Other studies based on similar models are given in, among others, Patel and Jenkins (2000), Del Castillo and Hurwitz (1997), and Gower-Hall et al. (2002).

One practical constraint is that the complete information about the model is not always available. When there are no time constraints. For example, in the experimental phase when the purpose is to investigate a new DRIE machine, products can be carefully measured to collect accurate readings of y_t. However, during the production stage, the time-consuming metrology procedure to obtain accurate readings of y_t is not feasible. Alternatively, we may choose to complete the measurement with much shorter time by collecting qualitative data on a less accurate categorical scale.

In the following derivation, Y_t is used to denote the readings obtained on a categorical scale. Obviously, the categorical variable, Y_t, is dictated by the input, u_{t-1}. In order to update the control action to maintain the process output on target, it is necessary to identify the functional relationship between Y_t and u_{t-1}. However, as the two variables are defined on different scales, experience suggests that a linking function is difficult to establish directly. Therefore, we appeal to the modeling strategy with latent variables. We may consider y_t in Equation (1) as a latent variable to obtain the linear relationship between u_{t-1} and Y_t. We can then establish a function between y_t and Y_t. The detailed procedure is presented as follows.

At first, a qualified operator examines a resulting product after a run, possibly with the aid of equipment, and classifies it into one of the m ordered categories. The value of Y_t depends on the magnitude of y_t, which is usually too difficult or costly to be observed during R2R production. The higher the y_t, the higher the corresponding categorical value. Intuitively, there exists a set of prespecified thresholds against which the operator assigns different categories to the product. This procedure may be illustrated by the following relation:

$$Y_t = j \Leftrightarrow s_{j-1} \le y_t < s_j, \tag{2}$$

where s_i, $i = 1, \ldots, m-1$, are the cutoff parameters that are placed in the output space of y_t, (s_0, s_m), to classify each y_t into one of m categories. Thus, $Y_t = j$ indicates the output at time t falling into the jth category, with the lowest category being 1 and the highest being m. For the case with a process output unbounded (i.e., no worst and best outputs defined), the definition in Equation (1) may be modified as

$$y_t = j \Leftrightarrow \begin{cases} y_t < s_j & j = 1 \\ y_t \ge s_{j-1} & j = m \\ s_{j-1} \le y_t < s_j & \text{otherwise.} \end{cases} \tag{3}$$

By putting together Equations (1) and (2), we link the control action u_{t-1} with Y_t through the mapping of y_t. Next, we will investigate more on quantifying Y_t for control action determination.

Distance Between Categorical Observations

As y_t is usually unobservable during R2R operation and only Y_t is observed, it is important to quantify the changes of Y_t, i.e., the distance between output categories, so as to determine the corresponding control action. To quantify the distance between categories, Wu and Hamada (2000) outlined a data-based midrank score strategy for ordinal responses. According to their method, each category is assigned a score by counting the observed frequencies of the current category and all other categories with lower orders. However, the data-based rank may not reflect the real magnitude of the observations in each category. Wu and Hamada (2000) also raised concerns about this method when a middle category is the ideal one, and scores are chosen to reflect a cost that increases in both directions as the category moves away from ideal.

Here we introduce an alternative definition for the distance between categorical variables based on the conditional mean of continuous variables. Let C_j be the event that y_t falls into category j. Analogous to the definition of the Euclidean distance, we define the distance between category j and k as

$$D(j,k) = |E(y_t \mid C_j) - E(y_t \mid C_k)|. \tag{4}$$

That is, the distance is defined as the difference between the statistical mean of two continuous variables given the fact that certain categorical values have been observed. Here we assume that if C_j happens, y_t follows a normal distribution centered inside the boundaries of category j, which implies that $E(y_t \mid C_j) = (s_{j-1} + s_j)/2$. Therefore, Equation (4) can be further expressed as

$$D(j,k) = \left| \frac{s_{j-1} + s_j}{2} - \frac{s_{k-1} + s_k}{2} \right|. \tag{5}$$

If, under certain circumstances, the range of the latent variable is not specified or is unbounded, the distance in Equation (5) may become infinite when the lowest or the highest category is involved. However, given the process output, y_t, follows a normal distribution with mean T and variance σ_y^2, y_t will fall in the range $(T - 3\sigma_y, T + 3\sigma_y)$, with a probability of 99.73%. Thus, it is reasonable to denote $s'_0 = T - 3\sigma_y$, $s'_m = T + 3\sigma_y$ and use the revised boundaries in the calculation of Equation (5). The section.

finite limits are required to facilitate the control procedure to be presented in the next

The Two-Phase R2R Process Control Strategy

In practice, the latent–model-based strategy for process adjustments can be implemented in two phases. Phase I is to understand the process and to estimate process parameters, which is during the experimental and set-up stage of an R2R process. During this phase, each output will be measured on both continuous and categorical scales, and vectors in the form of (u_{t-1}, y_t, Y_t) are expected to be collected. After the Phase I modeling, Phase II is facilitated during the regular R2R operation, which is to incorporate information from the first phase and to control the process by regulating controllable factors. During this phase, each output will only be measured on the categorical scale and (u_{t-1}, Y_t) is collected on an R2R base. In the following, we will introduce the two phases in detail, and a real case from a DRIE process is adopted to illustrate their implementation in practice.

Phase I Modeling of an R2R Process

The modeling and parameter estimation of an R2R process in Equation (1) has been studied thoroughly by using response surface experimental designs (Myers et al. (2004)), ordinary least-square estimation (Wu and Hamada (2000)), Ultramax sequential process optimization software (Moyne et al. (1994)), and other treatments (Apley and Kim (2004)). A literature review can be found in Del Castillo and Hurwitz (1997) and the references therein.

The second model to be estimated, Equation (2), has a slightly different form from the usual linear regression model. The initial values of the cutoff parameters may come from prior knowledge, such as customer specifications. Samples are generated based on these initial values to train operators. A well-trained operator can then classify the future products into proper categories. These classification results during Phase I need to be studied to verify their consistency with the initial values and to calibrate the model.

We now take a DRIE process as an example to illustrate the implementation of this phase. DRIE is a process that involves complex chemical–mechanical reactions. Here the machine to be studied is an inductive-coupled plasma (ICP) silicon etcher from Surface Technology System Ltd. (STS) (see McAuley et al. (2001), Rauf et al. (2002), and Zhou et al. (2004) for more details about this system). A schematic diagram of the system is shown in the Appendix. The central part of the machine is a process chamber, within which wafers are loaded and processed. The system first releases etching plasma into the chamber to generate trenches subject to designed mask patterns; then in the deposition step, different gases are introduced into the chamber to generate a protective film on the sidewalls. The etching and deposition steps repeat alternately until the preset processing time is reached or the end-point detection module confirms the correct etching depth. The etching/deposition procedures are illustrated in Figure 3a to Figure 3d.

The sidewall profile generated by the process is one of the major quality characteristics of customers' concern. Among others, the etching/deposition time ratio (ED ratio) is an effective factor that dominates the profile. Therefore, in this case, the ED ratio will be adjusted from run to run to control the desired profiles.

In this research, open-loop experiments have been conducted in Phase I to study the intrinsic property of the machine. The control factor of the process (u_{t-1}) is the ED ratio. Three levels of the ED ratio are chosen in the experiment. Corresponds to each level, six replicates are conducted. The accurate

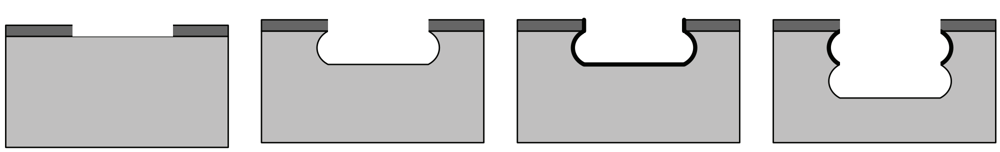

FIGURE 3. Etching and Deposition Steps of the DRIE Process. (a) Before Etching, (b) an Etching Step, (c) a Deposition Step, (d) Another Etching Step.

TABLE 1. Experimental Data from the DRIE Process

Sample number	Etching time (s)	Deposition time (s)	ED ratio (u_t)	Sidewall slope (degree) (y_t)	Category (Y_t)
1	5	7	0.71	90.44	2
2	5	7	0.71	90.48	2
3	5	7	0.71	90.45	2
4	5	7	0.71	90.82	3
5	5	7	0.71	90.70	3
6	5	7	0.71	90.73	3
7	7	7	1.00	89.75	2
8	7	7	1.00	89.57	2
9	7	7	1.00	89.57	2
10	7	7	1.00	89.76	2
11	7	7	1.00	89.80	2
12	7	7	1.00	89.51	2
13	11	7	1.57	89.29	2
14	11	7	1.57	88.85	1
15	11	7	1.57	88.97	1
16	11	7	1.57	88.99	1
17	11	7	1.57	88.79	1
18	11	7	1.57	88.93	1

slope reading of each wafer (y_t) is shown in Table 1. Note that the wafer measurements on both continuous and categorical scales are considered as a set-up cost, which only applies to Phase I, not to Phase II, and thus is affordable in this case.

A regression analysis of the data has suggested $\alpha = 91.7$ and $\beta = -1.8$ in model (1). The R^2 value is 88.5%, which shows that the data is adequately represented by the linear model. Regarding the IMA(1, 1) disturbance series, the procedure suggested by Apley and Kim (2004) has been followed. The process input was held constant, so that the output was the effect of a disturbance process plus a constant. Based on that, an IMA(1, 1) model was fitted to the output series to estimate θ. Results show that an IMA(1, 1) disturbance with $\theta = 0.6$ fits our process well.

Besides the accurate reading y_t, each wafer is also classified into one of three categories: negative (1), normal (2), positive (3) by a trained operator. Using the categorical information, Y_t, as a dependent variable and the accurate readings, y_t, as an explanatory variable, a classification tree can be fitted to the dataset using standard tree algorithms (Hastie et al. (2001)). In the DRIE process, we have used the commercial software S-Plus™ to generate the classification tree. The analysis results show that the negative and normal categories are separated by $s_1 = 89.14$, while the normal and positive categories are separated by $s_2 = 90.59$, which are quite consistent with the customer specifications. The desired target value is $T = 90$, and the output space of the slope is chosen to be $s_0 = 87$ and $s_3 = 93$, according to the knowledge and experience of process engineers. All these process parameters will be used for R2R process regulation in Phase II.

Note that as the classification is usually based on human operators, a gage repeatability and reproducibility (gage R&R) study is required to ensure the consistency of the operators' performance before moving to Phase II. The procedure of measurement system analysis (MSA) for bounded ordinary data may be found in de Mast and van Wieringen (2004) and the references therein. In our demonstrated case, the operator's performance is rather consistent and the gage R& R is satisfactory. The gage R&R performance can surely be improved by purchasing a metrology system producing an image and also automatically quantifying the attributes of that image, although this alternative could be costly.

Phase II Adjustments of an R2R Process

Phase II of the proposed R2R process control strategy is to generate a recipe for each run to meet a specific criterion, minimize process variance, and also lessen regulation efforts, according to available categorical information. The major difference of the proposed strategy from conventional R2R strategies is the assumption that no accurate/continuous output measure, but categorical information, is available from run to run. A novel control algorithm for such a situation is developed as the following. Again, the DRIE process is used for illustration in this Phase.

Let F_t be a set that comprises the cumulative information collected up to the present step, $F_t = \{Y_t, Y_{t-1}, \ldots, u_{t-1}, u_{t-2}, \ldots\}$. That is, at step t, all the responses and covariates in F_t are known. The purpose of the controller is to minimize the deviation between the real process output and the target. Therefore, we define the following quadratic loss function by conditioning on known information:

$$L = E[(y_{t+1} - T)^2 \mid F_t], \tag{6}$$

where y_{t+1} is the output that will happen at step $t+1$ and T is the target value, which is also the center of the target category. The loss function penalizes the squared one-step ahead deviation between the real output and the target value given all the past control actions and categorical responses. Thus, the objective of the new recipe, u_t, is to minimize such a function.

Here we subtract T from both sides of model (1) and multiply it by $(1-\theta B)^{-1}(1-B)$ to obtain

$$(1-\theta B)^{-1}(1-B)(y_{t+1}-T) = \beta(1-\theta B)^{-1}\Delta u_t + \varepsilon_{t+1}, \tag{7}$$

where B is a back-shift operator such that $By_{t+1} = y_t$, $(1-B)\alpha = 0$, and $(1-B)T = 0$. Knowing that $(1-\theta B)^{-1}(1-B) = 1-(1-\theta)(1-\theta B)^{-1}B$, Equation (7) can be rewritten as

$$\begin{aligned} y_{t+1} - T = {} & (1-\theta)(1-\theta B)^{-1}(y_t - T) \\ & + \beta(1-\theta B)^{-1}\Delta u_t + \varepsilon_{t+1}, \end{aligned} \tag{8}$$

where $\Delta u_t = u_t - u_{t-1}$ is the relative amount of adjustment comparing with the previous run, i.e., the adjustment of the ED ratio in the DRIE process.

From the iterated conditional expectation formula $E(Y \mid X_1) = E[E(Y \mid X_1; X_2) \mid X_1]$, the loss function (6) can be written as

$$\begin{aligned} L &= E[E((y_{t+1} - T)^2 \mid y_t; F_t) \mid F_t] \\ &= E[E(((1-\theta)(1-\theta B)^{-1}(y_t - T) \\ &\qquad + \beta(1-\theta B)^{-1}\Delta u_t + \varepsilon_{t+1})^2 \mid y_t; F_t) \mid F_t] \\ &= E[((1-\theta)(1-\theta B)^{-1}(y_t - T) \\ &\qquad + \beta(1-\theta B)^{-1}\Delta u_t)^2 + \sigma_\varepsilon^2 \mid F_t]. \end{aligned} \tag{9}$$

In Equation (9), y_t is a random variable with unknown value. However, taking expectation over y_t will remove the dependence on it. We can then take the derivatives of (9) with respect to Δu_t, set it to zero, and obtain the following optimal adjustment action:

$$\Delta u_t^* = -\frac{(E(y_t \mid F_t) - T)(1-\theta)}{\beta}. \tag{10}$$

The minimized loss function is obtained by applying the optimal control action, Equation (10), to the objective function, Equation (9):

$$\begin{aligned} L^* &= e[((1-\theta)(1-\theta B)^{-1}(y_t - T) \\ &\qquad + \beta(1-\theta B)^{-1}\Delta u_t^*)^2 + \sigma_\varepsilon^2 \mid F_t] \\ &= E\left[\left((1-\theta)(1-\theta B)^{-1}(y_t - T) \right.\right. \\ &\qquad \left. + \beta(1-\theta B)^{-1}\frac{(E(y_t \mid F_t) - T)(1-\theta)}{\beta}\right)^2 \\ &\qquad \left. + \sigma_\varepsilon^2 \mid F_t \right] \\ &= E[((1-\theta)(1-\theta B)^{-1}(y_t - E(y_t \mid F_t)))^2 \\ &\qquad + \sigma_\varepsilon^2 \mid F_t] \\ &= (1-\theta)^2(1-\theta B)^{-2}E[(y_t - E(y_t \mid F_t))^2 \mid F_t] \\ &\qquad + \sigma_\varepsilon^2 \\ &= (1-\theta)^2(1-\theta B)^{-2}\mathrm{Var}(y_t \mid F_t) + \sigma_\varepsilon^2. \end{aligned} \tag{11}$$

From the definition of the categorical distance in Equation (4), one can observe that the nominator of the optimal controller, Equation (10), is dictated by the distance between the observed category and the target category. Assuming the last observation falls into category j, $Y_t = j$, the expected location of the real output will take the following form:

$$E(y_t \mid F_t) = E(y_t \mid C_j) = (s_{j-1} + s_j)/2. \tag{12}$$

Therefore, the control action, Equation (10), can be expressed as

$$\Delta u_t^* = -\frac{((s_{j-1} + s_j)/2 - T)(1-\theta)}{\beta}. \tag{13}$$

By substituting the values estimated from Phase I for the parameters in Equation (13), the categorical controller can be implemented easily as the following:

$$\Delta u_t^* = \begin{cases} 0.4289 & \text{if } Y_t = 1 \\ 0 & \text{if } Y_t = 2 \\ -0.3989 & \text{if } Y_t = 3. \end{cases} \tag{14}$$

Note that, if the observation falls into the target category, no adjustment would be necessary.

In practice, the true values of the parameters in Equation (13) are not known exactly. Therefore, it is of interest to study the impact of estimation uncertainties on the categorical controller. We now assume θ, α, and β are all random variables and denote their estimated values by $\hat{\theta}$, $\hat{\alpha}$, and $\hat{\beta}$, respectively. The controller aims to minimize the following objective function:

$$L = \underset{\theta,\alpha,\beta,\varepsilon_{t+1}}{E}[(y_{t+1} - T)^2 \mid F_t], \tag{15}$$

in which the expectation is taken over all the unknown parameters.

Although a nonlinear function between $(y_{t+1} - T)$ and θ has been obtained in Equation (8), this function is too complex to be solved analytically. Following the method suggested by Apley and Kim (2004), we adopt the first-order Taylor expansion to approximate the loss function. As in Equation (8), α is cancelled out, we take partial derivatives of (8) with respect to parameters β and θ that yield

$$\left.\frac{\partial(y_{t+1} - T)}{\partial\beta}\right|_{\theta=\hat{\theta},\beta=\hat{\beta}} = (1 - \hat{\theta}B)^{-1}\Delta u_t \tag{16}$$

and

$$\begin{aligned}
&\left.\frac{\partial(y_{t+1} - T)}{\partial\beta}\right|_{\theta=\hat{\theta},\beta=\hat{\beta}} \\
&\quad = -(1 - \hat{\theta}B)^{-1}(y_t - T) \\
&\qquad + (1 - \hat{\theta})(1 - \hat{\theta}B)^{-2}B(y_t - T) \\
&\qquad + \hat{\beta}(1 - \hat{\theta}B)^{-2}B\Delta u_t \\
&\quad = -(1 - \hat{\theta}B)^{-1} \\
&\qquad \times ((y_t - T) - (1 - \hat{\theta})(1 - \hat{\theta}B)^{-1}B(y_{t-1} - T) \\
&\qquad\quad - \hat{\beta}(1 - \hat{\theta}B)^{-1}\Delta u_{t-1}) \\
&\quad = -(1 - \hat{\theta}B)^{-1}\hat{\varepsilon}_t.
\end{aligned} \tag{17}$$

The last equality holds from Equation (8). Thus, the first-order Taylor approximation of $y_{t-1} - T$ about $\theta = \hat{\theta}$, $\beta = \hat{\beta}$ is

$$\begin{aligned}
y_{t+1} - T &\approx (1 - \hat{\theta})(1 - \hat{\theta}B)^{-1}(y_t - T) \\
&\quad + \hat{\beta}(1 - \hat{\theta}B)^{-1}\Delta u_t + \varepsilon_{t+1} \\
&\quad + (\beta - \hat{\beta}) \left.\frac{\partial(y_{t+1} - T)}{\partial\beta}\right|_{\theta=\hat{\theta},\beta=\hat{\beta}} \\
&\quad + (\theta - \hat{\theta}) \left.\frac{\partial(y_{t+1} - T)}{\partial\beta}\right|_{\theta=\hat{\theta},\beta=\hat{\beta}} \\
&= (1 - \hat{\theta})(1 - \hat{\theta}B)^{-1}(y_t - T) \\
&\quad + \hat{\beta}(1 - \hat{\theta}B)^{-1}\Delta u_t + \varepsilon_{t+1} \\
&\quad + (\beta - \hat{\beta})(1 - \hat{\theta}B)^{-1}\Delta u_t \\
&\quad - (\theta - \hat{\theta})(1 - \hat{\theta}B)^{-1}\hat{\varepsilon}_t.
\end{aligned} \tag{18}$$

Substituting Equation (18) for $y_{t+1} - T$ in the loss function, Equation (15), yields

$$\begin{aligned}
&\underset{\theta,\alpha,\beta,\varepsilon_{t+1}}{E}[(y_{t+1} - T)^2 \mid F_t] \\
&\quad = \underset{y_t}{E}\{[(1 - \hat{\theta})(1 - \hat{\theta}B)^{-1}(y_t - T) \\
&\qquad + \hat{\beta}(1 - \hat{\theta}B)^{-1}\Delta u_t]^2 \\
&\qquad + \sigma_\varepsilon^2 + (1 - \hat{\theta}B)^{-2}\Delta u_t^2\sigma_\beta^2 \\
&\qquad - (1 - \hat{\theta}B)^{-2}\hat{\varepsilon}_t^2\sigma_\theta^2 \mid F_t\},
\end{aligned} \tag{19}$$

where σ_ε^2, σ_β^2, and σ_θ^2 are the variances of ε_{t+1}, β, and θ, respectively. Here we may assume that β and θ are independent. That means, in Phase I, the experiments are conducted with open-loop operations and without feedback. Setting the partial derivatives of Equation (19) with respect to Δu_t equal to zero and solving for (Δu_t) gives the following optimal controller:

$$\Delta u_t^* = -\frac{(1 - \hat{\theta})(E(y_t \mid F_t) - T)}{\hat{\beta}(1 + \hat{\sigma}_\beta^2/\hat{\beta})}, \tag{20}$$

which is the optimal control action when considering parameter uncertainties.

If we substitute the values estimated from Phase I for the parameters in Equatiion (20), the alternative control action is given by the following:

$$\Delta u_t^* = \begin{cases} 0.4254 & \text{if } Y_t = 1 \\ 0 & \text{if } Y_t = 2 \\ -0.3957 & Y_t = 3. \end{cases} \tag{21}$$

We call this a cautious categorical controller, as it is consistent with the idea of cautious control in the engineering process control (EPC) literature (Astrom and Wittenmark (1995) and Jin and Ding (2004)). Equation (20) shows that, with the consideration of uncertainties in parameter estimation, $\hat{\sigma}_\beta^2 > 0$, the control action is always more conservative than that obtained from the original categorical controller in Equation (10). This is confirmed by comparing Equation (21) to Equation (14). If the variance of parameters can be ignored, i.e., $\hat{\sigma}_\beta^2 = 0$, the cautious categorical controller would reduce to the ordinary categorical controller in Equation (10).

A Discussion on the Categorical Controller Design

A practical issue to be solved is the determination of the number of categories. There may be no universal guidelines for various applications. However, there are two critical criteria that should be considered in choosing the number of categories: one is that they have to be physically meaningful and the other is that they should be practically distinguishable.

In the DRIE example, all profiles that fall into the (89.14, 90.59) range are considered acceptable by the customer. Therefore, wafers with slopes in this rage are classified as one category. Further classification within this range may not be physically meaningful, as they are all good wafers. Furthermore, if the slope of a wafer falls outside this range, we classify it as either a negative or a positive wafer. In a stable etching process, most slopes are between (87, 93). For negative wafers, their slopes are within (87, 89.14); for positive wafers, their slopes are within (90.59, 93). It would be hard for us to further classify within the two degree ranges, as human operators cannot tell such a small difference. The ranges may become practically indistinguishable.

As a prerequisite, accurate output data should be collected in Phase I in order to build a useful process model for the R2R control in Phase II. If in some circumstances the quality characteristic of interest cannot be accurately measured even in Phase I, one may consider utilizing a surrogate variable to replace the original variable. The technique of using surrogate variables may be found in Kwon et al. (2001) and p. 442 of Wu and Hamada (2000).

Performance and Stability Analysis

After proposing the categorical R2R controller, in this section, we investigate its performance in terms of minimizing the mean square error (MSE) deviated from the desired target. We also investigate the controller's stability, as the stability of a controller determines its robustness to uncertainties in parameter estimation and unexpected changes of processes.

Because wafer etching is a time-consuming, costly, and unrepeatable operation, it is not practical for us to apply different controllers to the same wafer to compare their performances. Therefore, we apply Monte Carlo simulations based on the real context and process model of DRIE from the previous section to study the performance and stability of the proposed controller. The model parameters estimated from the Phase I experiments are treated as their true values in the simulation. Each simulated process is run for 200 steps, and 100 replicates for each parameter setting are used to calculate the average MSE.

Given an initial control action u_0, a categorical controller updates recipes using Equation (10). At time t, the process input becomes

$$u_t = u_0 - \sum_{i=1}^{t} \frac{(E(y_t \mid F_t) - T)(1-\theta)}{\beta} = u_0 - \frac{(1-\theta)}{\beta} \sum_{i=1}^{t} (E(y_i \mid F_i) - T), \quad (22)$$

which shows an integral form that sums up all output deviations.

Consider a conventional EWMA controller with smoothing parameter $1-\theta$, the control action can be written as (see Sachs et al. (1995)):

$$\Delta u_t = -\frac{1-\theta}{\beta}(y_t - T), \quad (23)$$

or equivalently,

$$u_t = u_0 - \frac{1-\theta}{\beta} \sum_{i=1}^{t} (y_i - T). \quad (24)$$

Comparing Equation (22) to Equation (24) suggests that the categorical controller, similar to the EWMA controller, is also an integral-type controller. It is known that the EWMA controller is popular for the R2R process with accurate/continuous outputs, and with the smoothing parameter, $1-\theta$, it is also the optimal in terms of minimizing MSE for a process with IMA disturbances (Ingolfsson and Sachs (1993)). Thus, in the following study, we will use the optimal EWMA controller as a benchmark (i.e., the best performance assuming that the accurate output information is available) to compare to the proposed categorical controller (where only categorical information is available).

Performance Analysis Under Parameter Estimation Uncertainties

In the categorical controller (13), estimated parameters are used for recipe generation. However, as estimation uncertainties always exist, it is worth knowing how the controller will perform if the estimated parameters are different from their true values. In the first simulation study, both parameters, $\hat{\beta}$ and $\hat{\theta}$, vary around their respective true values. Here

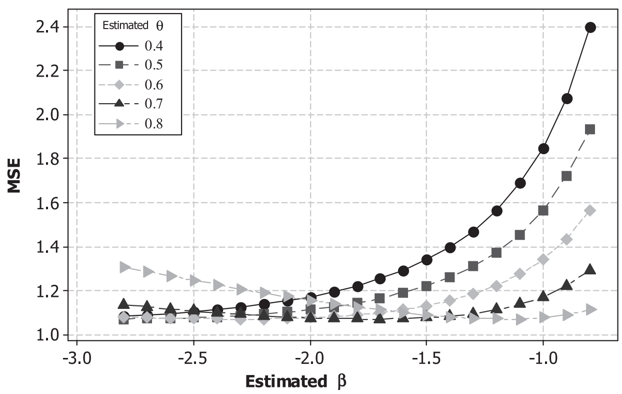

FIGURE 4. Etching and Deposition Steps of the DRIE Process. (a) Before Etching, (b) an Etching Step, (c) a Deposition Step, (d) Another Etching Step.

the MSE is used as a criterion to evaluate the performance and stability of each process and is shown in Figure 4.

As can be seen in Figure 4, both changes in $\hat{\theta}$ and $\hat{\beta}$ have significant impact on the process MSE. In addition, interactions exist between $\hat{\theta}$ and $\hat{\beta}$ as the curves are not parallel. More specifically, when $\hat{\beta}$ is close to zero, a small increase in the parameter will cause a significant increase in MSE. This may be explained by Equation (13) of the categorical controller that, when the absolute value of $\hat{\beta}$ is small, any deviation in the output will be magnified and lead to excess adjustments, which in turn results in serious output oscillations. While if $\hat{\beta}$ is overestimated, a more conservative control action will be made and the resulting controller would be rather insensitive to the parameter changes.

To jointly investigate the effects of $\hat{\theta}$ and $\hat{\beta}$, we further define a gain parameter,

$$g = \beta(1 - \hat{\theta})/\hat{\beta}. \tag{25}$$

Consequently, the categorical controller Equation (13) can be rewritten as

$$\Delta u_t^* = -\frac{g}{\beta}((\hat{s}_{j-1} + \hat{s}_j)/2 - T). \tag{26}$$

The purpose of adopting the gain parameter is to encompass all parameter-estimation uncertainties and to standardize the figures, as being utilized in the EPC literature, e.g., Apley and Kim (2004). In the following, the controller stability is investigated by changing the gain parameter. Based on that, the figures presented will be influenced by the relative estimation accuracy only, but not by the absolute value of the parameters. Therefore, these figures become more representative and comparable.

Figure 5a presents the MSE of both the proposed categorical controller and the EWMA controller with the smoothing parameter, $1 - \theta$, under various gain parameters. As expected, the EWMA controller, which is optimal for the IMA disturbance, achieves the minimum variance at $g = 0.4$, on which the parameters take their true values. When the gain parameter goes toward either direction, MSE increases in both controlled processes. We observe that, when the deviation of the gain parameter is small, both the categorical and EWMA controllers perform rather similarly. However, when the gain parameter deviates over a large range, the performance of the EWMA controller degrades rapidly, while that of the categorical controller is not affected as much. This suggests that the proposed categorical controller performs well and more stable over a larger range of g and thus is more robust to parameter estimation errors.

Performance and Stability Influenced by Measurement System Capability

As many categorical observation-generation procedures involve human operators, it is critical to investigate how such a measurement system influences the control performance and stability. In this study, gage reproducibility and repeatability are not explicitly separated because the operators are considered as part of the gages for output evaluation and classification.

Equation (2) is the model upon which qualitative observations are generated. The measurement system could be questionable if the cutoff parameters are not quite consistent from run to run, although it is usually the case for a human operator system. In the simulation, we assume $s_i \sim \text{Unif}(s_{iT} - \upsilon/2, s_{iT} + \upsilon/2)$. That is, each cutoff parameter varies within an interval of width υ around its true value. The larger the υ, the more unstable the classification system.

Again, the true cutoff parameters are assumed to be $s_1 = 89.14$ and $s_2 = 90.59$, and the interval between these two lines is $H = 1.45$. Using this value as a reference, we choose υ to be a certain percentage of H. Each curve in Figure 5b shows the MSE of a particular υ value. When υ increases, the curve moves upward accordingly. This is not unexpected because a larger υ means a less stable measurement system. Poor consistency in the measurement system leads to an increase in process variance. However, the

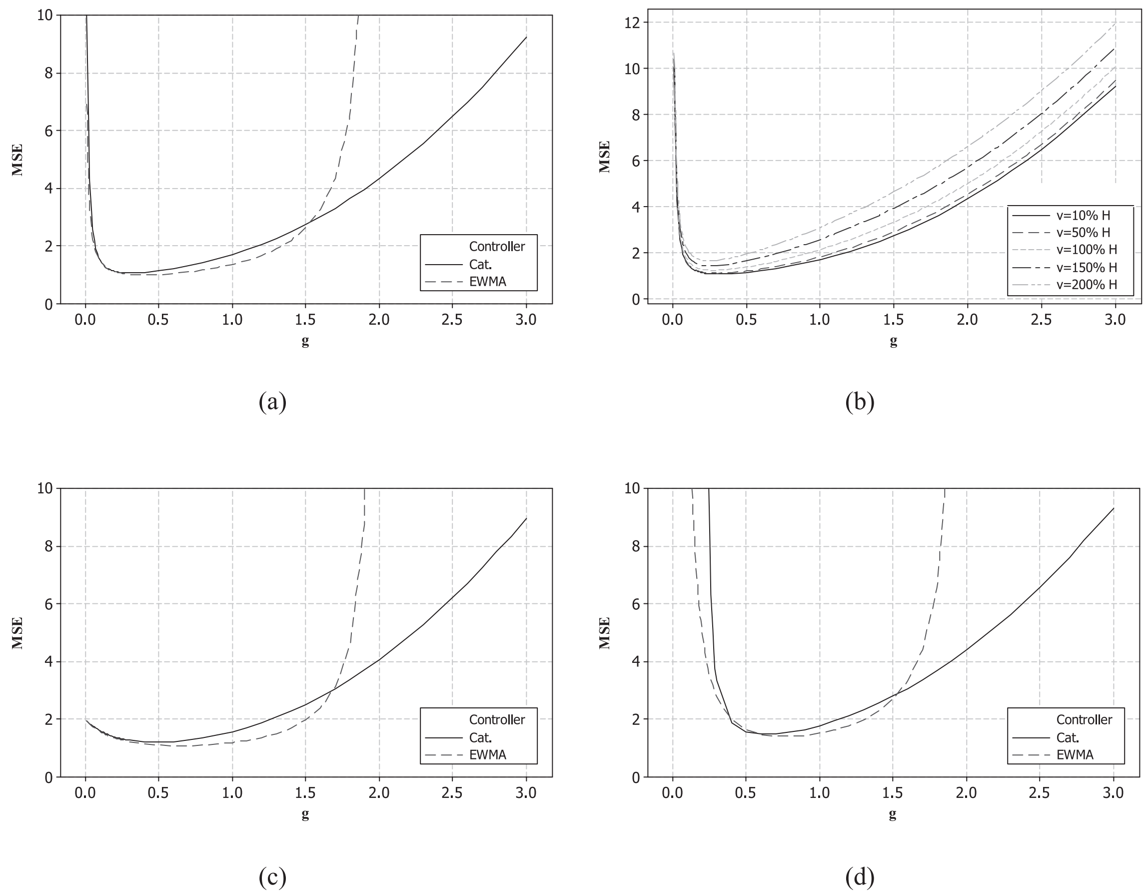

FIGURE 5. MSE Comparison of the Categorical Controller and the EWMA Controller. (a) With an IMA(1, 1) Noise Series, (b) with Measurement System Uncertainties, (c) with an ARMA(1, 1) Noise Series, (d) with Deterministic Drifts.

results show that, even if $\upsilon = 200\% H$, which means the cutoff parameter varies as wide as twice H, the increase of MSE in a categorical controlled process is still relatively small, especially for a small g.

Performance and Stability Under Model Structure Uncertainties

The optimal control action, Equation (10), is derived from model (1), in which the noise is assumed to be an IMA(1, 1) process. Although the IMA model is a very popular and useful assumption about process disturbance (see Apley and Kim (2004), Tseng et al. (2003)), the real model may not always be the same in practice. Therefore, it is worth studying the performance of the categorical controller when a different disturbance model is presented.

In the following simulation, we assume that the noise series follows a first-order autoregressive moving average (ARMA) model, which is another popular model in the literature (Apley and Kim (2004)), i.e., $d_{t+1} = \phi d_t + \varepsilon_{t+1} - \theta \varepsilon_t$. Figure 5c shows the MSE of the categorical controller and the EWMA controller when the disturbance process is an ARMA(1, 1) model with, say, $\phi = 0.8$ and $\theta = 0.2$ as an example.

As shown in Figure 5c, when the gain parameter varies from zero to a larger value, MSE decreases first and then slowly increases. A careful examination on this figure indicates that the minimal MSE of both controllers is not reached at $g = 1 - \theta$, which suggests that, when the model structure changes, the optimal setting obtained from the assumed model cannot achieve the best performance. In addition, the MSE curve of the categorical controller is much flatter than that of the EWMA controller, which suggests that the categorical controller has a larger stability region and is more robust to model-structure changes.

Performance and Stability Under Process Drift

Besides the changes in the disturbance series model, it is of interest to investigate the performance of the proposed controller under different types of process faults. One common failure in R2R processes is deterministic drifts (Sachs et al. (1995), Ingolfsson and Sachs (1993)). The following model illustrates such a process,

$$y_t = \alpha + \beta u_{t-1} + \delta t + d_t, \tag{27}$$

where δ is the drifting parameter. If left uncontrolled, the process will gradually deviate away from its target.

Figure 5d shows the MSE of processes that undergo deterministic drifting disturbances. The drifting parameter δ is set to 0.4. Again, the EWMA controller is used for comparison. Simulation results show that both the categorical controller and the EWMA controller can reduce process variance if appropriate settings are utilized. In addition, the categorical controller performs closely to the EWMA controller in terms of MSE when the g value is small, while the MSE curve of the categorical controller is much flatter when the g value is large. This again indicates the robustness property of the categorical controller over a large range of g.

Adjustment Effort and Cautious Control with a Categorical Response

Most conventional controllers, including the EWMA controller, regulate a process by changing its process input recipe at the end of every run. However, in addition to the MSE reduction, practical industrial applications always favor fewer recipe changes so as to reduce operational cost (Del Castillo and Hurwitz (1997)). The proposed categorical controller, on the other hand, does not regulate the process input at every run. When the previous output falls into the target category, no regulation will be made. In the following simulation, the categorical controller and the EWMA controller are compared in terms of both their MSE and average adjustment interval (AAI). AAI is a performance measure to indicate how frequent the recipe changes, with the longer the interval being better.

Simulations based on the DRIE process with 100 replicates are conducted, and the average values are presented in Table 2. Without any adjustments, although the operational cost is reduced, the process exhibits unacceptably large variance. The EWMA controller gives the lowest MSE, although it is just slightly better than that of the categorical controller. In terms of the adjustment effort, the AAI of the categorical controller is around two times that of the EWMA controller. It means that the EWMA controller changes the recipe twice as frequently as the categorical controller and thus bears higher operational cost.

TABLE 2. A Comparison of Different Control Schemes

Controller	MSE	AAI
No adjustment	15.872	*
EWMA controller	0.995	1
Cat. controller	1.084	2.082
Cautious cat. controller	1.082	2.086

An illustration of the recipe trajectories are shown in Figure 6. As the figure indicates, the recipe of the categorical controller is flat in many small segments, which suggests that no adjustment to the recipe is made. However, the EWMA shows a zigzag recipe trajectory, where many recipe updates seem unnecessary.

A Cautious Categorical Controller

Apley and Kim (2004) conducted extensive studies on the cautious controller for the situation when the real-time continuous output information is available. Such a cautious controller has more conservative control actions but is proven to have a higher probability of closed-loop stability than the standard

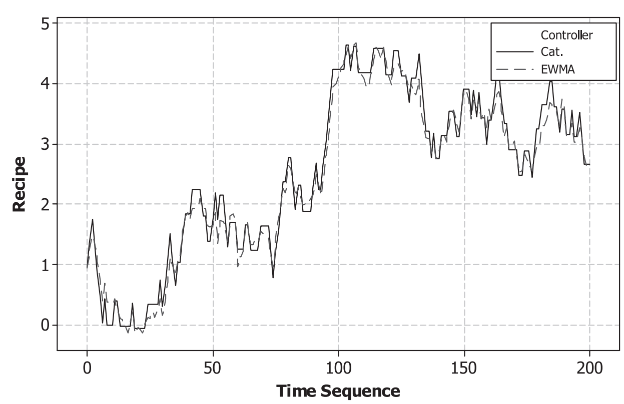

FIGURE 6. AAI Comparison of the Categorical Controller and the EWMA Controller.

minimum variance controller and can lessen the adverse impact of parameter uncertainties on closed-loop variance.

Similarly, the cautious categorical controller given by Equation (20) suggests more conservative control actions by adding the variance component to the denominator of the equation. This scenario is equivalent to overestimating the parameter β. Also, an investigation of the gain parameter in (25) reveals that, given the true value of β, the larger the $\hat{\beta}$, the smaller the gain parameter becomes for an ordinary categorical controller.

Here we compare the control action of the cautious categorical controller with that of the ordinary categorical controller and the EWMA controller in Table 2. As we can see, it seems that the cautious categorical controller performs very similarly to the ordinary categorical controller in terms of both MSE and AAI. This is because the ordinary categorical controller is already rather robust to the uncertainties of the gain parameter. Therefore, the cautious categorical controller does not bring much further improvement. However, if it is known that a high uncertainty exists in parameter estimation, a cautious categorical controller may be recommended because increasing $\hat{\beta}$ from its true value would give a slower change in MSE than decreasing $\hat{\beta}$. Thus, the cautious categorical controller could be safer in the sense that it is less likely to make the process become unstable.

Conclusion

This paper has proposed a categorical controller to regulate R2R processes when ordinal categorical observations rather than when quantitative observations are collected. The proposed method can greatly enhance the ability of operators to control an R2R process by using qualitative information in an effective way.

A two-phase design and implementation strategy has been introduced and demonstrated by a DRIE R2R process. In Phase I, with both qualitative and quantitative measurements available, off-line experiments have been designed and conducted for model building and parameter estimation. In Phase II, when only qualitative observations are available, a categorical controller or a cautious categorical controller has been developed for process regulation.

Simulation results show that the categorical controller, without using accurate output information, can significantly decrease process variance and is comparable to the optimal EWMA controller. Also, the categorical controller has been shown to be more robust to parameter estimation and model-structure uncertainties. In addition, the categorical controller requires much less frequent adjustments than the EWMA controller. This is obtained with little increase in MSE, which leads to lower operational cost.

This paper has developed a novel method that incorporates categorical observations into R2R process adjustments. This work falls into the general framework of statistical process adjustment advocated by Del Castillo (2006). Del Castillo (2006) also identified that categorical variables may sometimes appear as regressors. In practice, many applications have both qualitative and quantitative variables. Integrating delayed measurements (due to metrology delay) with real-time categorical information is also an area open for future research in R2R control.

Acknowledgments

The authors are grateful to the editor and the anonymous referees for their valuable comments, which helped improve this work greatly. The authors also thank Professor Yi-Kuen Lee in the Department of Mechanical Engineering at the Hong Kong University of Science and Technology for providing the DIRE machine for investigation and contributing his expertise in Microelectromechanical Systems during discussions. This research was supported by RGC Competitive Earmarked Research Grants HKUST6232/04E and HKUST6204/05E.

Appendix

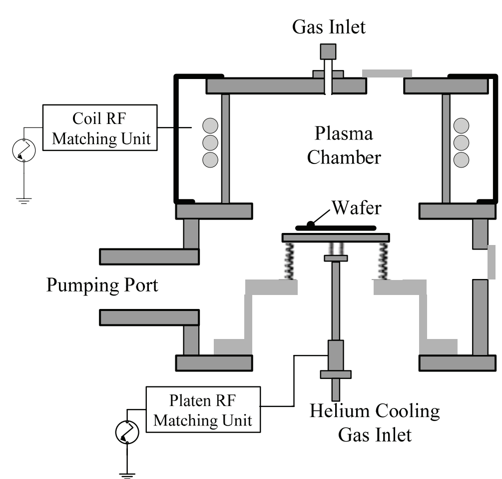

FIGURE A.1. Schematic Diagram of an STS Inductively Coupled Etch System.

References

AGRESTI, A. (1990). *Categorical Data Analysis*. Wiley, New York, NY.

APLEY, D. W. and KIM, J. (2004). "Cautious Control of Industrial Process Variability with Uncertain Input and Disturbance Model Parameters". *Technometrics* 46 (2), pp. 188–199.

ASTROM, K. J. and WITTENMARK, B. (1995). *Adaptive Control*, 2nd edition. Addison-Wesley, New York, NY.

AYON, A. A.; BRAFF, R.; LIN, C. C.; SAWIN, H. H.; and SCHMIDT, M. A. (1999). "Characterization of a Time Multiplexed Inductively Coupled Plasma Etcher". *Journal of the Electrochemical Society* 146 (1), pp. 339–349.

BARAS, J. S. and PATEL, N. S. (1997). "A Framework for Robust Run by Run Control with Lot Delayed Measurements". *IEEE Transactions on Semiconductor Manufacturing* 10 (1), pp. 75–83.

BUTLER, S. W. and STEFANI, J. A. (1994). "Supervisory Run-to-Run Control of Polysilicon Gate Etch Using In-Situ Ellipsometry". *IEEE Transactions on Semiconductor Manufacturing* 7 (2), pp. 193–201.

CAMPBELL, W. J.; FIRTH, S. K.; TOPRAC, A. J.; and EDGAR, T. F. (2002). "A Comparison of Run-to-Run Control Algorithms". *Proceedings of the American Control Conference* 3, pp. 2150–2155.

DE MAST, J. and van Wieringen, W. (2004). "Measurement System Analysis for Bounded Ordinal Data". *Quality and Reliability Engineering International* 20 (5), pp. 383–395.

DEL CASTILLO, E. (2006). "Statistical Process Adjustment: A Brief Retrospective, Current Status, and Some Opportunities for Further Work". *Statistica Neerlandica* 60(3), pp. 309–326.

DEL CASTILLO, E. and HURWITZ, A. M. (1997). "Run-to-Run Process Control: Literature Review and Extensions". *Journal of Quality Technology* 29 (2), pp. 184–196.

GOOD, R. and QIU, S. J. (2002). "Stability Analysis of Double EWMA Run-to-Run Control with Metrology Delay". *Proceedings of the American Control Conference*, pp. 2156–2161.

GOWER-HALL, A. E.; BONING, D. S.; ROSENTHAL, P.; and WALDHAUER, A. (2002). "Model-Based Uniformity Control for Epitaxial Silicon Deposition". *IEEE Transactions on Semiconductor Manufacturing* 15(3), pp. 295–309.

HASTIE, T.; TIBSHIRANI, R.; and FRIEDMAN, J. (2001). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*. Springer, New York, NY.

INGOLFSSON, A. and SACHS, E. (1993). "Stability and Sensitivity of an EWMA Controller". *Journal of Quality Technology* 25 (4), pp. 271–287.

JIN, J. H. and DING, Y. (2004). "Online Automatic Process Control Using Observable Noise Factors for Discrete-Part Manufacturing". *IIE Transactions* 36 (9), pp. 899–911.

KWON, H. M.; HONG, S. H.; LEE, M. K.; and KIM, S. B. (2001). "A Process Monitoring Procedure Based on a Surrogate Variable for Dichotomous Performance Variable". *IIE Transactions* 33(12), pp. 1129–1133.

LAVIOLETTE, M.; SEAMAN, J. W.; BARRETT, J. D.; and WOODALL, W. H. (1995). "A Probabilistic and Statistical View of Fuzzy Methods". *Technometrics* 37(3), pp. 249–261.

MAY, G. S.; HUANG, J.; and SPANOS, C. J. (1991). "Statistical Experimental Design in Plasma Etch Modeling". *IEEE Transactions on Semiconductor Manufacturing* 4 (2), pp. 83–98.

MCAULEY, S. A.; ASHRAF, H.; ATABO, L.; CHAMBERS, A.; HALL, S.; HOPKINS, J.; and NICHOLLS, G. (2001). "Silicon Micromachining Using a High-Density Plasma Source". *Journal of Physics D—Applied Physics* 34(18), pp. 2769–2774.

MONTGOMERY, D. C.; KEATS, J. B.; YATSKIEVITCH, M.; and MESSINA, W. S. (2000). "Integrating Statistical Process Monitoring with Feedforward Control". *Quality and Reliability Engineering International* 16 (6), pp. 515–525.

MOYNE, J.; ETEMAD, H.; and ELTA, M. (1994). "A Run-to-Run Control Framework for VLSI Manufacturing". *Proc. Microelectron. Process., Sensors, Controls, SPIE* 2091, pp. 379–390.

MYERS, R. H.; MONTGOMERY, D. C.; VINING, G. G.; BORROR, C. M.; and KOWALSKI, S. M. (2004). "Response Surface Methodology: A Retrospective and Literature Survey". *Journal of Quality Technology* 36 (1), pp. 53–77.

PASSINO, K. M. and YURKOVICH, S. (1998). *Fuzzy Control*. Addison-Wesley, Menlo Park, CA.

PATEL, N. S. and JENKINS, S. T. (2000). "Adaptive Optimization of Run-to-Run Controllers: The EWMA Example". *IEEE Transactions on Semiconductor Manufacturing* 13(1), pp. 97–107.

PALMER, E.; REN, W.; SPANOS, C. J.; and POOLLA, K. (1996). "Control of Photoresist Properties: A Kalman Filter Based Approach". *IEEE Transactions on Semiconductor Manufacturing* 9(2), pp. 208–214.

RAUF, S.; DAUKSHER, W. J.; CLEMENS, S. B.; and SMITH, K. H. (2002). "Model for a Multiple-Step Deep Si Etch Process". *Journal of Vacuum Science and Technology A* 20 (4), pp. 1177–1190.

RUEGSEGGER, S.; WAGNER, A.; FREUDENBERG, J. S.; and GRIMARD, D. S. (1999). "Feedforward Control for Reduced Run-to-Run Variation in Microelectronics Manufacturing". *IEEE Transactions on Semiconductor Manufacturing* 12(4), pp. 493–502.

SACHS, E.; HU, A.; and INGOLFSSON, A. (1995). "Run by Run Process Control—Combining SPC and Feedback-Control". *IEEE Transactions on Semiconductor Manufacturing* 8 (1), pp. 26–43.

SPANOS, C. J. and CHEN, R. L. (1997). "Using Qualitative Observations for Process Tuning and Control". *IEEE Transactions on Semiconductor Manufacturing* 10 (2), pp. 307–316.

TSENG, S. T.; YEH, A. B.; TSUNG, F.; and CHAN, Y. Y. (2003). "A Study of Variable EWMA Controller". *IEEE Transactions on Semiconductor Manufacturing* 16 (4), pp. 633–643.

TSUNG, F. and SHI, J. J. (1999). "Integrated Design of Run-to-Run PID Controller and SPC Monitoring for Process Disturbance Rejection". *IIE Transactions* 31(6), pp. 517–527.

WU, C. F. J. and HAMADA, M. (2000). *Experiments: Planning, Analysis and Parameter Design Optimization*. Wiley, New York, NY.

ZHANG, C.; DENG, H.; and BARAS, J. S. (2003). "Performance Evaluation of Run-to-Run Control Methods in Semiconductor Processes". *Proceedings of 2003 IEEE Conference on Control Applications* 1, pp. 841–848.

ZHE, N.; MOYNE, J. R.; SMITH, T.; BONING, D.; DEL CASTILLO, E.; JINN-YI, Y.; and HURWITZ, A. (1996). "A Comparative Analysis of Run-to-Run Control Algorithms in the Semiconductor Manufacturing Industry". *Proceedings of IEEE/Advanced Semiconductor Manufacturing Conference and Workshop*, pp. 375–381.

ZHOU, R.; ZHANG, H.; HAO, Y.; and WANG, Y. (2004). "Simulation of the Bosch Process with a String-Cell Hybrid Method. *Journal of Micromechanics and Microengineering* 14, pp. 851–858.