AIChE Wiley InterScience DISCOVER SOMETHING GREAT

Process Trend Monitoring Using Key Sensitive Index

Jyh-Cheng Jeng, An-Jhih Su, Cheng-Ching Yu, and Hsiao-Ping Huang
Dept. of Chemical Engineering, National Taiwan University, Taipei 106, Taiwan

DOI 10.1002/aic.11269
Published online August 1, 2007 in Wiley InterScience (www.interscience.wiley.com).

Process monitoring is essential to maintain product quality in semiconductor manufacturing. However, unlike chemical processes, semiconductor manufacturing processes exhibit the following characteristics: (1) much shorter (minutes) and often variable (deliberately adjusted) batch time, (2) multiple processing steps (10–20) in each batch, (3) only some particular processing steps constituting the quality-determining steps, and (4) mixed products for the same batch processing. In this work, instead of incorporating a large number of trajectory data with variable batch time and possibly "missing" data for some process variables using multivariate statistics, a process-insight-based approach, key sensitive index (KSI), is taken. From process knowledge, the key sensitive time-slot in the recipe is identified. Next, possible key sensitive process variables (KSV) are selected and validated according to the process trend correlation. A simple batch reactor example is used to illustrate the KSV selection procedure. Then, an index for these variables is sought. Two integrated circuit processing examples from real fab data are used to illustrate the KSI-based approach and results clearly indicate that process trend is captured using KSI-based approach.
 AIChE J, 53: 2339–2350, 2007
Keywords: process monitoring, batch processes, semiconductor manufacturing, time-series model

Introduction

Batch processes play an important role in the production and processing of chemicals, pharmaceutical, and semiconductor devices. Generally, a batch process is characterized by prescribed processing of raw materials for a finite duration to convert them to products. A high degree of reproducibility is necessary to obtain successful batches. Some batch processes include a single step, whereas many others are carried out in a sequence of discrete steps, which are usually referred as recipes in semiconductor manufacturing. Events taking place in each step have an impact on the final product yield and quality. For chemical process industry (CPI), upon completion of a batch, a range of quality measurements is usually made at the quality control laboratory, often hours later. For semiconductor industry, the quality measurements are usually not available at the end of a single processing unit until the processes in several processing units have been completed. Su et al.[1] provide an overview of IC process characteristics including disturbance, measurement, and control architecture. Therefore, monitoring and control of the intermediate stages (units) of operation and intermediate product quality is as important as monitoring and control at the final stage.

Online process performance monitoring and product quality prediction in real time can reduce quality variations. Multivariate statistical projection methods such as principal component analysis (PCA) and partial least squares (PLS) are capable of utilizing massive amounts of process data. These methods compress the information in this data down into a lower dimensional latent space in which process monitoring and results interpreting are much easier. These methods were first utilized in developing multivariate statistical process

Correspondence concerning this article should be addressed to C.-C. Yu at ccyu@ntu.edu.tw.

monitoring and quality prediction techniques for continuous processes.[2] Nomikos and MacGregor[3] pioneer the use of a large number of trajectory measurements collected from batch process and develop multiway projection methods for statistical process monitoring. Batch data are recorded in terms of batch runs, variables, and time. They are arranged into a two-dimensional array in the multiway projection methods and, therefore, analysis of batch processes data give rise to a variety of challenges. Because of the nature of batch processes, every batch may come to completion at a different termination time, resulting in unequal batch data length. The unequal length batch data should be equalized prior to forming a two-dimensional array. The simplest way for equalizing batch lengths is cutting batch data lengths to the length of the variable with the shortest data sequence, but this is not recommended because of the possible significant information loss generated by discarding data. Several methods for equalizing batch lengths, such as the indicator variable technique (IVT),[4,5] dynamic time warping (DTW),[6] and curve registration,[7] have been suggested in the literature. IVT is based on selecting a process variable to indicate the progress of the batch instead of the time axis. This variable should be chosen such that it progresses monotonically in time and has the same starting and ending value for each batch, but IVT does not account for the locations and the alignment of critical local features (landmarks) of each trajectory. DTW is used to synchronize two trajectories by appropriately translating, expanding, and contracting localized segments within the trajectories to achieve a minimum distance between the two trajectories. However, warping the batch trajectories to have the minimum distance with the reference batch trajectory may distort their critical features or fault patterns and, hence, reduce the monitoring ability. Curve registration is a two-step process of identifying landmarks within a trajectory (or set of trajectories) and then warping the test trajectory to the reference trajectory containing reference landmarks. Identifying multivariate landmarks is challenging because the number and location of landmarks may be different for different process variables. In addition to unequal batch data length, the other difficulty for implementing multiway projection methods is that the process variables are not necessarily recorded at regular intervals. Thus, interpolation is often needed to find measurements at regular intervals. This has to be done properly; otherwise, the auto-correlation and cross-correlation might be seriously affected.[8]

Multiway PCA (MPCA) has been successfully applied for batch process analysis and monitoring in chemical process industry.[4,9] However, implementing MPCA in the field of semiconductor manufacturing will encounter some difficulties. The main reason is that, unlike chemical processes, the IC processing has much shorter (seconds–minutes for a wafer as opposed to minutes–hours for chemicals) and often variable (often deliberately adjusted) time duration. Preprocessing and arrangement of process data to deal with the problem of unequal batch data length using aforementioned methods will impose a relatively high computational burden and may significantly distort the critical features in the batch trajectories. Moreover, not the entire batch trajectory, but some particular processing sequence constitutes the quality-determining steps, and this is especially true for crystallization processes in CPI and thermal processes in IC industries. Using the entire batch trajectory for analysis might cloud the relationship between variables in these critical steps and the final product quality and, subsequently, reduces the resolution in process trend monitoring. These characteristics make the data-based approaches difficult in semiconductor manufacturing industry.

In this work, an alternative for batch process monitoring is sought. Instead of incorporating large number of trajectory data with variable batch time and possibly "missing" data for some process variables using MPCA, a key sensitive index (KSI)-based approach is proposed. From process insight or the experience of the process operators, a certain period time within a batch where measurements have significant effect on product quality, the key sensitive time-slot (KST), is identified. Next, based on the KST, possible key sensitive process variables (KSV) are chosen. The candidate KSV is not limited to measured values themselves in KST, but can be some quantities computed from the raw measurements over a period of time. These KSVs are then evaluated by autocorrelation analysis for validation. Based on the validated KSV(s), a KSI is defined and computed. By monitoring the KSI, the process status can be realized and possible maintenance action can therefore be called for, whenever necessary. This provides dynamical capability for process trend monitoring while possibly maintaining the simplicity of single-variate analysis. Two IC processing examples are used to illustrate the KSI-based approach.

Key Sensitive Time-Slot and Key Sensitive Variables

Key sensitive time-slot

In one process tool, there are many measurements, such as temperature, pressure, composition, flow rate, RF power, etc., which will be acquired into database systems. Many engineers attempt to mine useful knowledge or information from this huge amount of data. Figure 1 shows a simplified version of typical process raw data in semiconductor manufacturing. Because of the batch process characteristic, the technique of process/equipment monitoring usually treats the data as three dimensions: batches, measurements, and time. When

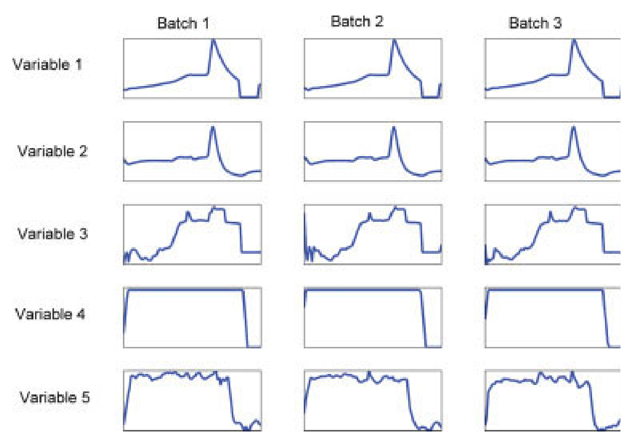


Figure 1. Sketch of general process raw data in semiconductor manufacturing.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

fetching raw data of a process tool, there are between 10 and 90 available measured variables (depending on equipment). Many variables could be neglected because of the constant behavior or relatively little influence on qualities from engineer's experience. This is the first step for dimension reduction and also the most important step to achieve successful data analysis.

Despite many successful applications of MPCA, several factors may cause difficulties. First, the step length of each batch and therefore the batch time might be different. This is quite common in tool level operation in IC processing. Second, the sampling interval may be different, and this leads to asynchronous data sequences. Moreover, each sampled discrete sequence cannot be synchronized in process time. Third, there are always problem with missing data, outliers, or other miscellanea. These difficulties necessitate substantial engineering effort on data preprocessing, to satisfy the argument format of modeling algorithms. In the worst case scenario, even great excessive engineering effort, it may be not easy to restore the true system behavior.

After selection of important measured variables, the next step is to cut each batch operation into several minor steps and filter them out. In practice, there are also many steps, typically 10–20, in a recipe (i.e., the entire batch run). Some steps are considered as pretreatment, e.g., warming-up, introducing gas flows, etc., until the chamber condition is stabilized for subsequent processing. These kinds of minor steps could be neglected. The key sensitive time-slot (KST) is defined as the steps that have significant influence on quality or tool health such as the thermal budget period in RTP or particle-generating steps in etching processes. This concept is also applicable in some chemical processes such as condensation polymerization and emulsion polymerization where the initial stage is critical for product quality.[10] For example, there are six steps in Figure 2, but only three critical steps, B, D, and F, which are KSTs. Usually, KSTs are the longest steps within a recipe. The minor steps, such as Steps A, C, and E could be treated as transition steps. Moreover, experienced engineers can identify critical steps with little difficulty.

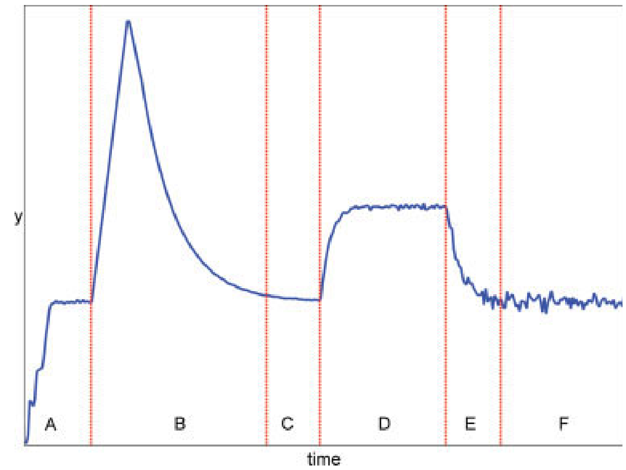


Figure 2. Schematic of key sensitive time-slot (KST).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Key sensitive variable

Candidate KSV. After the description of the KST, we introduce the concept of key sensitive variable (KSV) which is evaluated within KSTs. The KSV means certain features of process variables, and these features may have physical meanings, which will affect product specifications or will give indication on tool condition. The KSV can be "measurement data" itself or "transformation from process data" as described below.

1. *Measurement data:* In this case, the KSV is the process variable itself and it can be read directly from the raw data. The KSV can be the variable at a certain time such as the initial value or final value, or the maximum value or minimum value, etc. For example, the peak value in Step B of Figure 2 may be used as a KSV. One should note that raw data may not precisely capture true behavior in the case of low sampling frequency or fast process response.

2. *Transformation from process data:* The values computed from raw measurement, such as slope, mean, standard deviation, may have important physical implications. Again, let us use the Step F in Figure 2 to illustrate this. The mean value or standard deviation can be a sensitive measure for the process trend. Another possible example is the area of the peak or the ramp up slope in Step B (Figure 2). This type of KSV may have some physical meaning, e.g., integrated area for effective energy input. Furthermore, a prescribed model can be used to extract process features. These quantities can describe the process condition as the batch progresses. Simple models such as steady-state polynomial models or dynamical transfer function models are preferred. For example, one could use the initial value, steady-state gain, and time constant to capture the behavior of a first order response (Step D in Figure 2). This method is effective for identifying possible variation in process dynamics.

The foregoing discussion shows that the introduction of the KST and KSV actually accomplishes the dimensional reduction for subsequent analyses. In doing this, a two-dimensional data matrix (the number of observation by the number of variables) in a single batch can be reduced to a vector in which the elements are KSVs. Thus, it will be much easier to incorporate models, either steady-state or dynamic, for monitoring and/or fault detection.

Evaluation of KSV. After identifying candidate KSVs based on process knowledge, a mechanism is necessary for validation. There is no guarantee that the KSVs, which were identified based on intuition, will have strong relationship to quality specification or to the tool health as the batch runs repeatedly. For each candidate KSV, there will be a data sequence, with one point for each batch, and it can be treated as a discrete time series. A straightforward approach for validating candidate KSVs is to filter out variables that exhibit the behavior of random noise. In other words, the KSV, to a degree, should be able to capture the process trend as the batch repeats. The autocorrelation is a simple measure to discriminate a suitable KSV from a random variable, as far as process trend is concerned. The most satisfactory estimation of the kth lag autocorrelation of a time series, z_t (t = 1, 2,..., N), is[11]:

$$r_k = \frac{c_k}{c_0} \qquad (1)$$

where

$$c_k = \frac{1}{N}\sum_{t=1}^{N-k}(z_t - \bar{z})(z_{t+k} - \bar{z}) \quad k = 0, 1, 2, \ldots, K \qquad (2)$$

is the estimate of the autocovariance, and $\bar{z}$ is the sample mean value of the time series.

If the autocorrelation function of a time series for a candidate KSV is found to be significant, this KSV is then validated. Otherwise, if a weak correlation is observed, it could be treated as a white noise sequence, which offers little information on process trend. Thus, some candidate KSVs, which are not significant according to autocorrelation analysis, can be removed. If none of the KSVs is found to be significant, one should identify additional KSVs, or even possibly different KSTs. Again, it should be emphasized here that obtaining suitable KSVs depends on the understanding and experience of the engineer. However, a rigorous evaluation is critical for validation. If good KSVs are chosen, they will give better information on the process condition than a black/gray box model that receives all process measurements.

Elimination of Product Effects. Batches with mixed products are common in semiconductor manufacturing and this type of problem is very important in foundries. Therefore, it is important to consider the product effect on KSVs, so that effective process trend monitoring can be achieved. The product indices could be linewidth, pattern density, material properties, etc. If a KSV sequence contains unusual behaviors such as abrupt changes, shifts, and spikes within a trend, very likely, it may be a consequence of product effects. One can verify this effect by analyzing the correlation between KSVs and product indices. To know the selected KSV having product effects is a challenging task, but process knowledge always offers some insight. Two examples are: the removal rate of CMP varies with linewidth[12] and the geometry of a wafer surface affects the view factor for radiation related processes.

Quantitatively, correlation analysis between KSV and the available product indices can be used to discriminate the significance of product indices. If a strong correlation exists between the KSV and a certain product index, this index will be selected.

Generalized Procedure. The generalized procedure for obtaining KSVs is summarized as follows:

P1: Select possible measurements based on experience (process insight).

P2: Define the KST in a recipe (also from process insight).

P3: Obtain possible key sensitive variables.

P4: Check autocorrelation for each KSV. If no significant KSV was found, return to Step *P2* or *P3*.

P5: Take out the product effects. If information on product type is available, take out product effect via regression.

Following this procedure, a two-dimensional raw data set (measurements vs. time for a single batch) is reduced to a vector, with one entry for each KSV. Therefore, the concept of KSV is simply a dimensional reduction based on process knowledge together with autocorrelation analysis for validation. Next, a simple batch reactor example is used to demonstrate the step-by-step procedure.

An Illustrative Example. This is a batch reactor with first-order irreversible reaction ($A \rightarrow B$) and the objective is to produce on-spec product with simple temperature control. Appendix gives the reaction kinetics and model parameters. The entire batch consists of three steps (Figure 3): (1) preheating—the temperature of the reactor is ramped up at a rate 2°C/min for 35 min, (2) reaction—the reaction begins by injecting catalyst, and the temperature is kept at 95°C via temperature control by manipulating cooling water flow rate, and (3) cooling down—the reactor cools down in a rate 1°C/min to room temperature followed by the discharge.

To illustrate bath-to-batch variations, two disturbances are introduced. One is the catalyst deactivation as the batch progresses. This is described by the following relationship:

$$a_{\text{cat},t} = 0.99a_{\text{cat},t-1} \qquad (3)$$

where a_{cat} is the activity of catalyst and t is the bath index. The other is the initial charge to the reactor (N_{A0}) which, to a degree, illustrates the product effect. Three values of N_{A0} are considered, N_{A0} = 8, 10, 12. It is clear that less initial charge or lower catalyst activity will decrease the production of B, and, consequently, less heat will be generated. This phenomenon is reflected in the usage of cooling water as shown in Figure 3. In process simulation, we assume the conversion (X) is not measurable, and available measurements are reactor temperature (T), outlet cooling water temperature (T_{co}), and cooling water flowrate (F_c) as shown in Figure 3.

We proceed with the proposed procedure, starting from Step *P1* (selection of possible measurement). (*P1*) Since the

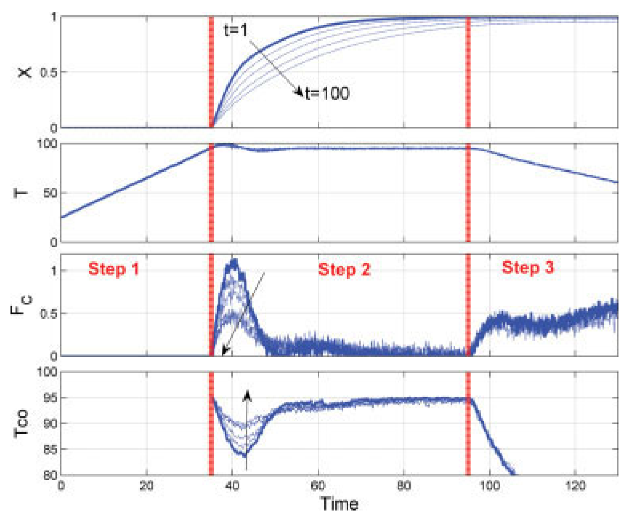


Figure 3. Profiles of conversion (X), reactor temperature (T), cooling water flowrate (F_c), and cooling water exit temperature (T_{co}) of batch reactor as batch number progressing from 1 to 100.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

reactor temperature is under closed-loop control, it is invariant from batch to batch; thus, reactor temperature is eliminated. Figure 3 shows that the cooling water flowrate (F_c) and cooling water outlet temperature (T_{co}) are dependent to each other. The more cooling water flowrate, the lower outlet temperature, and vice versa. So, one of the variable, F_c, is selected. (*P2*) The reaction stage determines the overall conversion, which is the most critical time slot as compared to the preheat and cooling periods. Thus, the reaction stage (Step 2 in Figure 3) is the obvious KST. (*P3*) For the selected process measurement, cooling water flowrate, in the KST, three possible key sensitive variables (KSV) can be devised: integrated area (overall flowrate in the reaction stage), maximum value (peak cooling water flowrate in KST), and final value (final steady-state value in KST). (*P4*) Because a trend exists in the catalyst activity, autocorrelation analysis is helpful to find the candidate KSV. Figure 4 shows that the maximum value exhibits relatively strong autocorrelation, as compared to the total amount of cooling water used and to the final value of the cooling water flowrate. So, the maximum value is the KSV in this case. Intuitively, one would think the integrated area (total amount of cooling water used) is a good candidate, but the product effect somehow distorts the correlation. (*P5*) The KSV, $Fc_{\max}$, indeed reveals process trend as batch number increases as shown in Figure 5 (top graph). However, spikes in $Fc_{\max}$ are also evident as the batches progress. Figure 5 also indicates that the spikes are correlated to the product effect, N_{A0} (actually with a correlation coefficient of 0.83 in the first 20 batches). Thus, it will be helpful to take out the product effect from the process trend. We use the following equation to modify $Fc_{\max}$ by taking the product index, N_{A0}, into account.

$$Fc'_{\max,t} = Fc_{\max,t} - k(N_{A0,t} - 10) \tag{4}$$

where t is the batch number and $k = 0.0672$ is obtained by solving the following optimization problem.

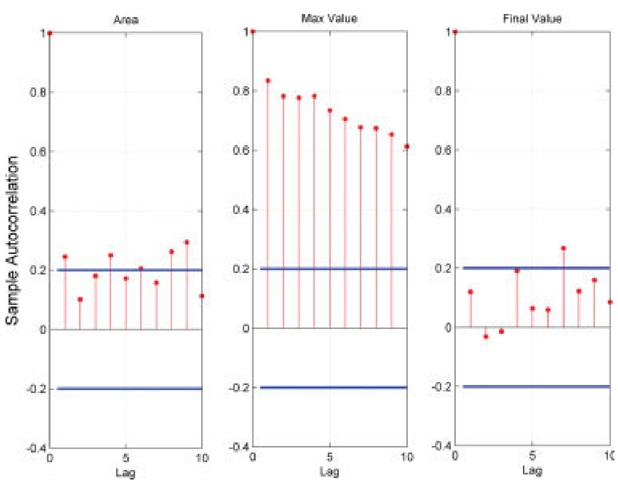


Figure 4. Autocorrelation analyses of area, maximum value, and final steady-state value of the cooling water flowrate (Step 2 of Figure 3).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

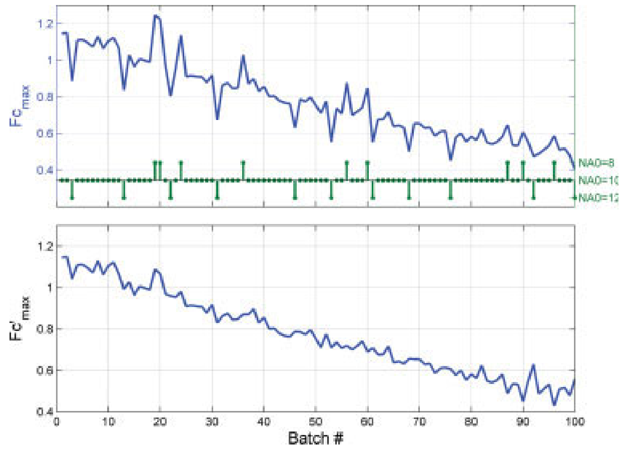


Figure 5. Original trend of *Fc*$_{\max}$ and corresponding product index (N_{A0}; top) and modified trend of *Fc*$_{\max}$ (bottom) by taking out the product effect.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

$$\operatorname*{Min}_{k} \left\| \mathbf{Fc}'_{\mathbf{max}} - \mathbf{Fc}'_{\mathbf{max,MA}} \right\| \tag{5}$$

where $\mathbf{Fc}'_{\mathbf{max}}$ is the vector with the entry of $Fc'_{max,t}$ in the training set and $\mathbf{Fc}'_{\mathbf{max,MA}}$ stands for the vector of the moving average of $Fc'_{max,t}$ computed from a sliding window with a window size (N_W) of 10.

$$Fc'_{\max,\mathrm{MA},t} = \sum_{j=1}^{N_W} Fc'_{\max,t-j+1}/N_W$$

Taking out the product effect (Eq. 2), the modified KSV, $Fc'_{max,t}$, gives better description of process trend as shown in the bottoms graph of Figure 5.

Despite being simple, the batch reactor example illustrates the step-by-step procedure to achieve the dimensional reduction using the KSV. Next, a method, based healthy index, is proposed for process monitoring.

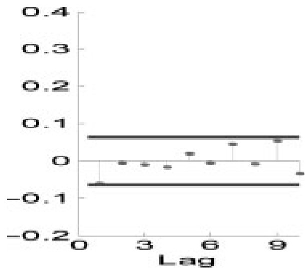
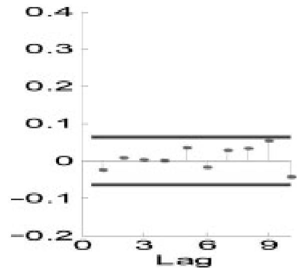
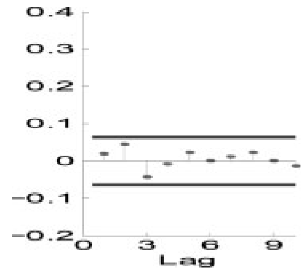
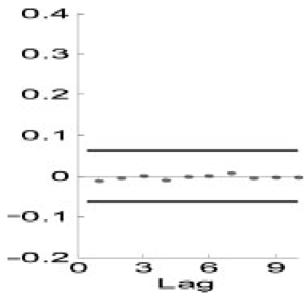
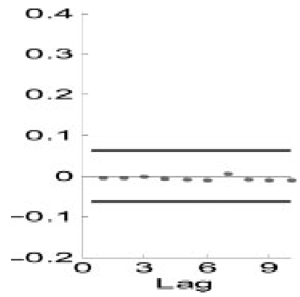
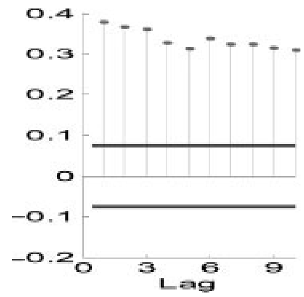
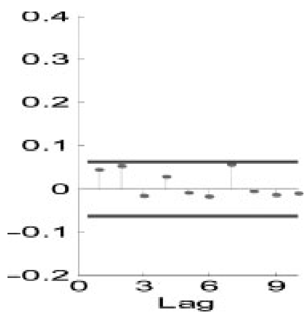
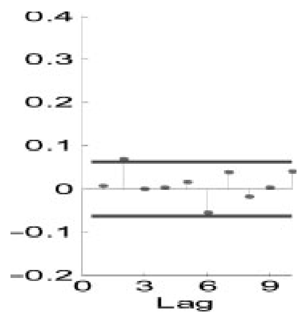
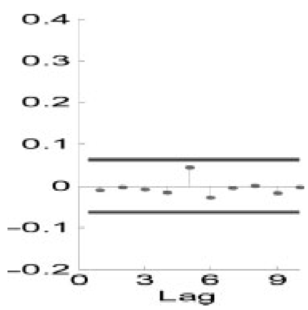
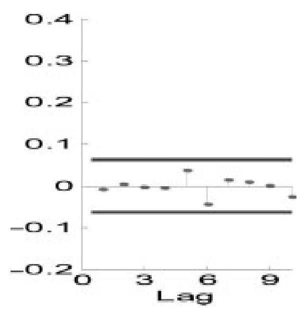
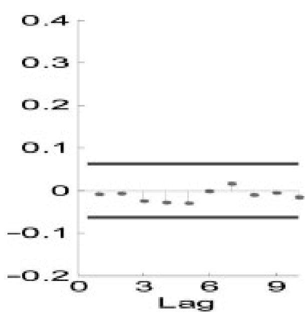
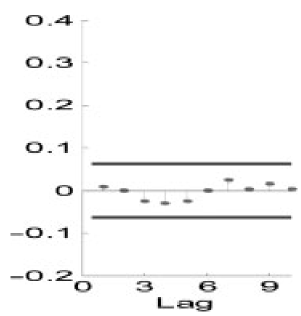
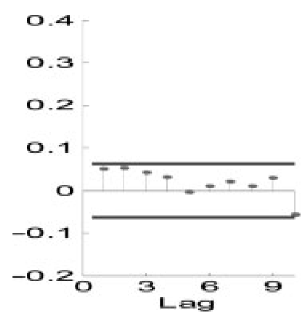
Key Sensitive Index

In the case of extreme dimension reduction, a scalar index could be constructed as a function of many KSVs in one batch. This index should reveal the current status or critical feature of a batch trajectory, and it is convenient for process/equipment monitoring. This kind of index is termed a key sensitive index (KSI) here, which is used to describe the behavior of the process. The choice of KSI depends on the analysis done on the aforementioned KSVs. Approaches for the construction of the KSI are described as follows.

A straightforward approach

If only one KSV was found and it exhibits significant autocorrelation, the simplest way is to take this KSV as the KSI. The upper and lower limits for the KSI can be established for monitoring the batch process. In addition, the batch

Table 1. Autocorrelation Analysis of Possible KSVs in Case Study 1

	Variable 1	Variable 2	Variable 3
Area			
Maximum			
Slope up			
Slope down			
Duration			

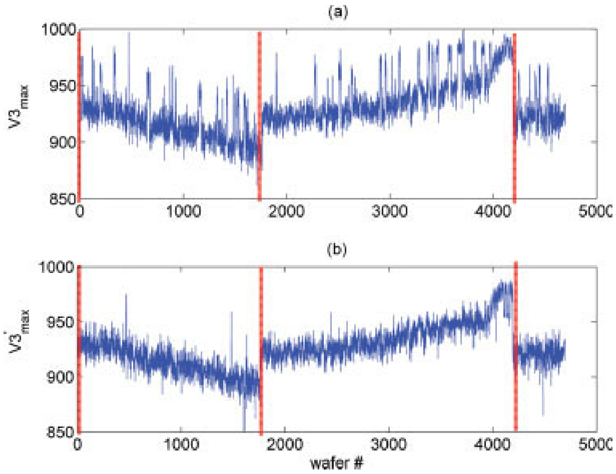


Figure 6. The trend of (a) KSV and (b) modified KSV by taking out product effects for case study 1.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

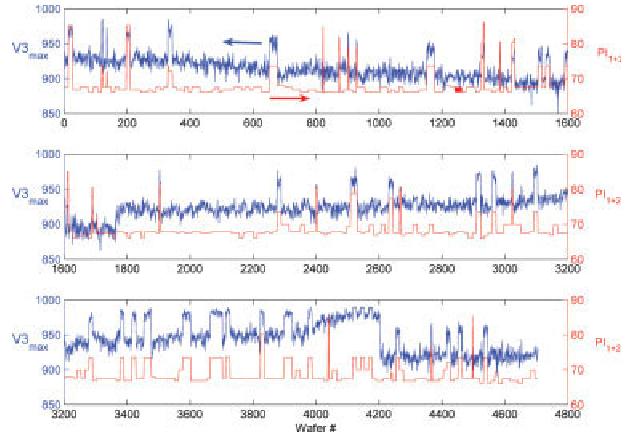


Figure 7. The influence of product index.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

process trend can be realized by inspecting the KSI of each batch as the batch progresses.

Time series approach

The KSI is the prediction error of time series models, constructed from KSVs. For a KSV with significant autocorrelation, it is appropriate to identify time series models (based on this KSV) for forecasting. When the time series model is built using nominal operating data, the difference between its predicted value and actual KSV of a batch can serve as the KSI. A large KSI (i.e., residual) means the process has deviated from its normal operating condition. Moreover, we can identify different kind of time series models (e.g., ARMA and ARIMA) and use them for prediction or for checking possible drifting, as the batch process repeats.

Multivariate statistical approach

The KSI is a statistical value (e.g., T^2, Q) determined from PCA modeling on the KSVs. Since a batch trajectory is represented by several KSVs, multivariate projection methods such as PCA can be applied to these KSVs for batch process monitoring. The time (observation) dimension of data matrix for continuous processes now becomes the batch number. Thus, the KSI could be the statistical quantities, T^2 or Q, frequently used in PCA monitoring. Such a KSI can be used not only for batch process monitoring, but also for fault diagnosis.

Table 2. Correlation Analyses of Product Indices in Case Study 1

Product Index	Correlation Coefficient
PD_1	0.388
PD_2	0.484
PD_3	0.115
$PD_1 + PD_2$	0.514
$PD_1 + PD_3$	0.27
$PD_2 + PD_3$	0.387
$PD_1 + PD_2 + PD3$	0.429

Applications

In this section, we introduce two case studies based on the method we proposed. The data are from a fab in Hsinchu, Taiwan. For confidential reason, all variables are deliberately disguised.

Case study 1—thermal process

Filtering KSVs with Product Effects. This is a front-end thermal process, which is a critical processing step.[13–15] The objective here is to track the process trend such that preventive maintenance can be executed at an appropriate time. This is important for semiconductor manufacturing, because we can achieve high throughput while maintaining acceptable tool performance as long as possible. At the present time, the tool is shut down for maintenance after a certain number of wafers have been processed or after the tool have been oper-

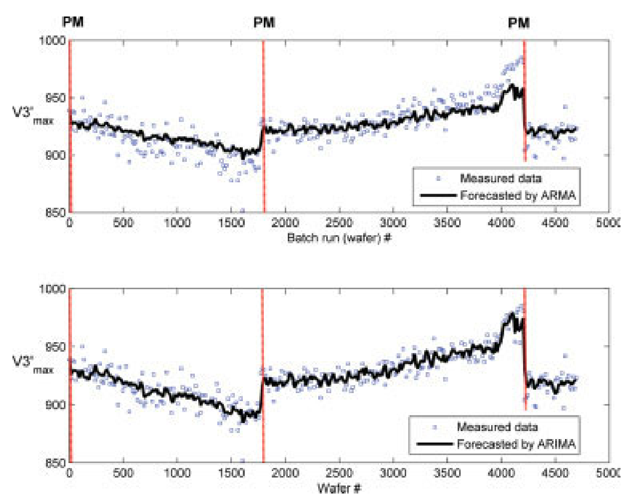


Figure 8. Comparison of ARMA and ARIMA prediction as compared to the true measurement (only one out of 15 points shown for better resolution).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

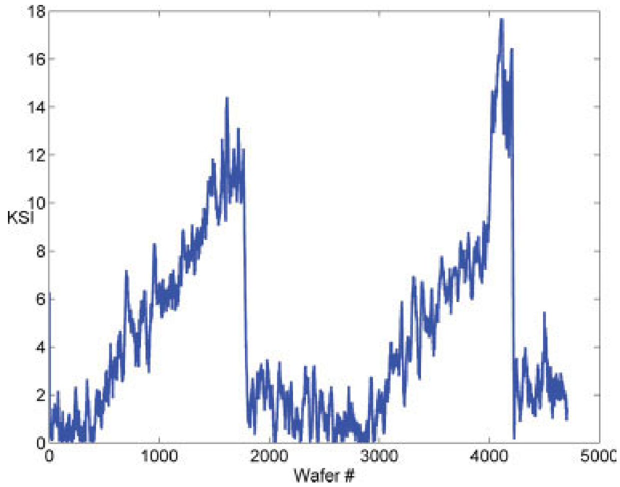


Figure 9. The resultant key sensitive index (KSI) for Case Study 1.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

ated for a certain period of time. Therefore, it is desirable to realize the tool condition and to generate an index for preventive maintenance (PM) only when necessary. The engineer has provided three likely process measurements. The autocorrelation analysis of possible KSV candidates is summarized in Table 1. Variable 3 does not have features of slopes. It is seen that only significant autocorrelation is for the maximum value of variable 3, which is termed as $V3_{\max}$ and is plotted in Figure 6a. As this tool has been proceed with almost 5000 wafers, this window size covers two complete PM periods, which are respectively located at vertical dash lines. Obviously, this KSV approximately follows the PM behavior, because there are two sudden shifts just after each PM.

The appearance of many spikes in $V3_{\max}$ (Figure 6a) may cause difficulty in control application. The most likely reason for these spikes is that there are many different products in a foundry fab. The available information describing different products has been collected as three product indices (PI_1, PI_2, and PI_3). Consequently, we check the correlation coefficients between various combinations of product indices and $V3_{\max}$ as shown in Table 2. The result indicates that the

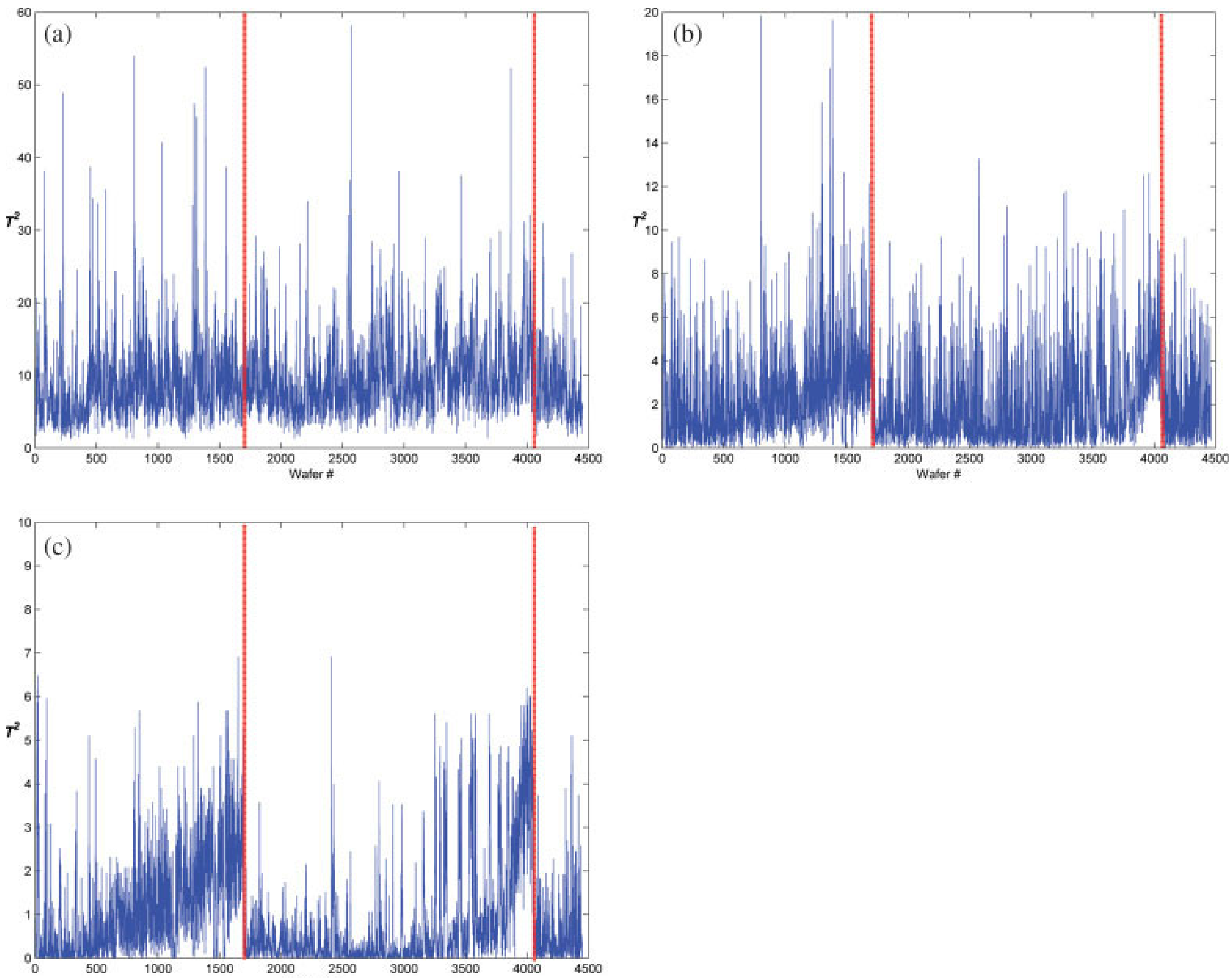


Figure 10. The T^2 plot of MPCA for Case Study 1: (a) the trajectory data of three variables in KST used; (b) the maximum values of three variables used; (c) the maximum values of the third variable used.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

Table 3. Detail Description of 16 KSVs in Case Study 2 for Fault Detection

KSV No.	Source Variable	KST	Feature Description
1, 2	Variable A	Step 11	Initial value, stabilized value
3, 4	Variable A	Step 14	Initial value, stabilized value
5, 6	Variable A	Step 18	Initial value, stabilized value
7, 8, 9	Variable B	Step11 to Step 20	Initial value, gain, time constant*
10, 11, 12	Variable C	Step11 to Step 20	Initial value, gain, time constant*
13, 14	Variable D	Steps 11 and 14	Mean value
15, 16	Variable E	Steps 11 and 14	Mean value

*Parameters of the first-order transfer function model.

combination of PI_1 and PI_2 (PI_{1+2}) has the strongest relationship with $V3_{max}$. In Figure 7, the spikes of $V3_{max}$ occur when product index has higher value. Based on this observation, we do a modification including a least square regression, in which, if PI_{1+2} is beyond a specific threshold value, $V3_{max}$ will be reduced by the value of kPI_{1+2}. The threshold value and parameter k are obtained by least square regression. The modified value, $V3'_{max}$, is shown in Figure 3b. It is seen that, this KSV more clearly indicates the PM behavior by reducing product effects.

A Time Series-Based KSI. Roughly, $V3'_{max}$ can be treated as a monitoring index with given upper and lower control limits. However, if the tool's behavior (e.g., health index) can be captured, it is more useful for the scheduling of PM. Time series analysis[11] is helpful for modeling a discrete sequence. An autoregressive moving average (ARMA) model is built for $V3'_{max}$ based on measurements from 500 wafers:

$$(1 - 1.744q^{-1} + 0.776q^{-2})V3'_{max}(t) = (1 - 1.346q^{-1} + 0.476q^{-2})e(t) \quad (6)$$

where q^{-1} is the backward shift operator and $e(t)$ is white noise. It is found that one root of the autoregressive polynomial is close to unity, which means that this time series of $V3'_{max}$ exhibits nonstationary behavior. For this reason, an autoregressive integrated moving average (ARIMA) model is then built as the following to describe this behavior.

$$(1 + 0.942q^{-1})\nabla V3'_{max}(t) = (1 + 0.452q^{-1} - 0.553q^{-2})e(t) \quad (7)$$

where $\nabla = (1 - q^{-1})$. These two time-series models are then used for forecasting the values of $V3'_{max}$ as the batch process repeats. The result is shown in Figure 8 where two abrupt changes are observed due to scheduled tool preventive maintenance. Initially, both the forecasts of ARMA and ARIMA models can follow the process trend well. However, as the batch process repeats, the forecast of ARMA model gradually deviates from the actual $V3'_{max}$, while the forecast of ARIMA model remains accurate. This phenomenon disappears after PM and then can be observed again as the batch process repeats. To capture the drifting behavior of this batch process, the KSI is thus defined as the absolute value of difference between residuals of these two time series models:

$$\mathrm{KSI} = |\mathrm{Residual}_{\mathrm{ARMA}} - \mathrm{Residual}_{\mathrm{ARIMA}}| \quad (8)$$

This KSI is shown in Figure 9. The results clearly indicate that the process trend can be realized using the proposed KSI and tool maintenance is required once this KSI is greater than a prescribed limit. The reason for that is an ARIMA model can capture the nonstationary behavior while the ARMA model fails to achieve (only for stationary processes).

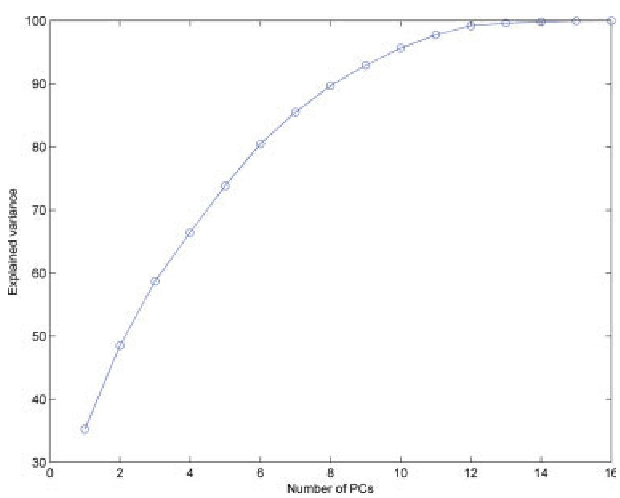


Figure 11. The explained variance with number of PCs for the fault detection of Case Study 2.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

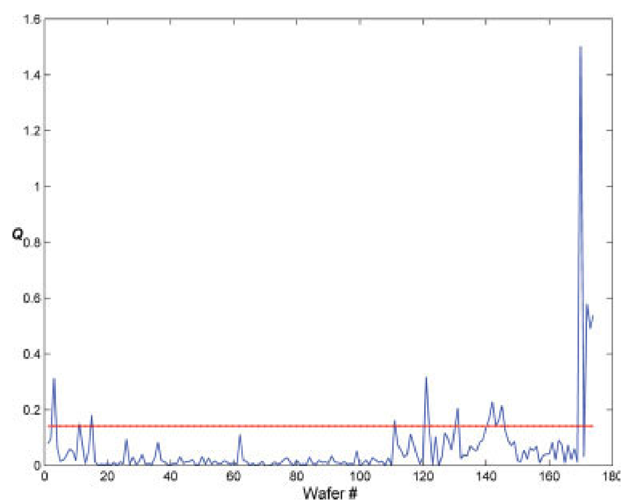


Figure 12. Q statistics for the fault detection of Case Study 2 with scrapped wafers, Nos. 170, 172, 173, and 174.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

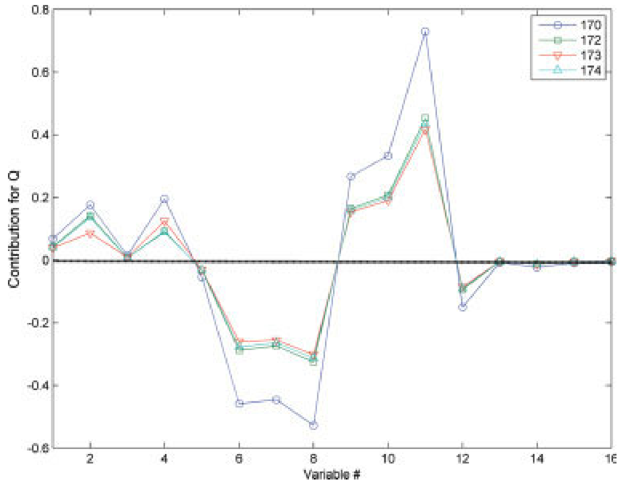


Figure 13. Contribution plot of faults for wafer nos. 170, 172, 173, and 174 in Case Study 2 with KSV 6-11 (see Table 3 for KSV detail).

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

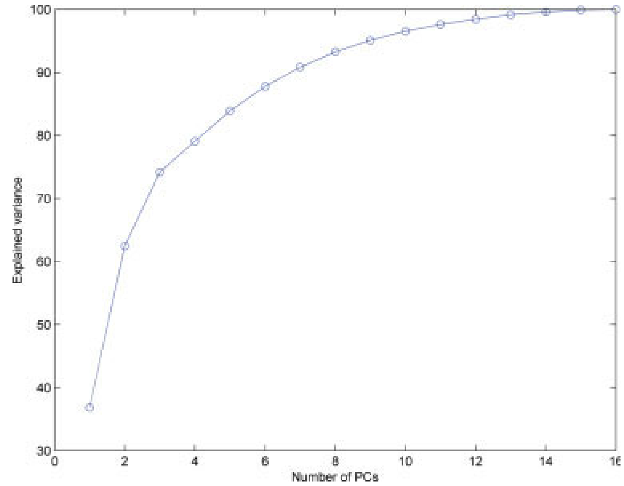


Figure 14. The explained variance with number of PCs for KSI of Case Study 2.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

A large KSI indicates that the process deviates from the stationary behavior with noticeable trend. Thus, Eq. 6 is applicable to the process with noticeable trend. Therefore, this KSI-based approach not only can be used for batch process trend monitoring, but also it is helpful for the engineers to decide when to call for tool maintenance.

Comparison with the Results of MPCA. As mentioned previously, using the entire batch data for process monitoring (e.g., MPCA) might cloud the relationship between important variables in critical steps and the final product quality. Consequently, it may reduce the resolution in process trend monitoring. Here, we repeat the process trend monitoring using MPCA by including more data points. Three data sets are used for such a comparison. The first data set consist of three variables in the KST in each batch, i.e., 29 × 3 matrix for each wafer. The second data set selects only the maximum values of these three variables, i.e., 1 × 3 matrix per wafer, and third set selects the maximum values of the third variable, i.e., 1 × 1 matrix per wafer. The plots of T^2 statistics are shown in Figure 10. It can be seen that no process trend can be observed when the entire trajectory measurements are used for analysis (data set 1). However, as data points are reduced, the process trend becomes more apparent (data set 2). As for the third case, one can see the process trend a little better, similar to that of the KSI-based approach. This example illustrates that using the entire trajectory data for all variables may not be appropriate for this particular application, and the proposed KSI-based approach, on the other hand, can capture the process trend with much less data.

Table 4. Detail Description of 16 KSVs in Case Study 2 for Trend Monitoring

KSV No.	Source Variable	KST	Feature Description
1, 2, 3	Variable A	Step 5	Initial, min, max value
4, 5, 6	Variable A	Step 11	Initial, max, min value
7	Variable A	Step13	Stabilized value
8, 9, 10	Variable A	Step15 to Step17	Initial, max, final value
11, 12	Variable B	Step11	Initial, final value
13, 14	Variable B	Step 13	Initial, final value
15, 16	Variable C	Step 11	Initial, final value

Case study 2—thin film process

The concept of KSV is applied to a thin film tool. There are different objectives from the previous case: virtual metrology, which relates tool raw data and quality index, is considered here. The objectives are: (1) to perform fault detection (for possible severe process failure) and (2) to provide process trend monitoring. The raw data of these three objectives involve three different products respectively from the same tool. So the time step in recipes may have different

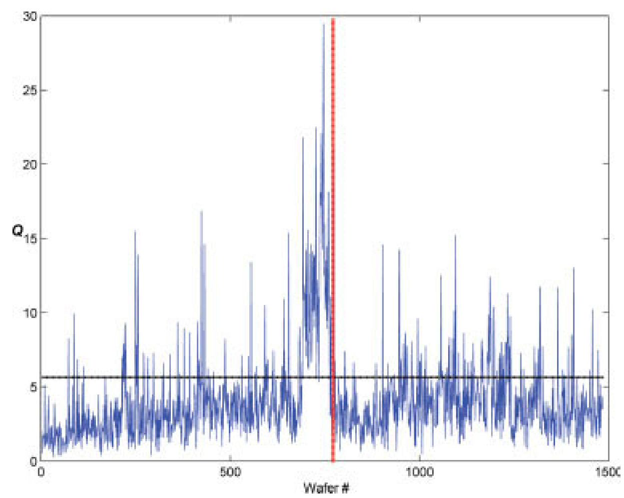


Figure 15. The resultant key sensitive index (KSI, Q statistics) for Case Study 2 with runs of 1500 wafers while the preventive maintenance (PM) indicated by the dashed line.

[Color figure can be viewed in the online issue, which is available at www.interscience.wiley.com.]

meanings. This process tool has 23 variables in raw measurement, and one batch spans over 400 sampling instances.

Fault detection is considered for a set of process data with a known fault. This process tool may produce defects, e.g., voids, which can not be detected until several further process steps have been completed. This is certainly not desirable, because many scrapped wafers may result. Thus, a KSV-based approach is proposed for fault detection. First, 41 KSV candidates are identified, followed by autocorrelation analyses. Then, 16 KSVs with significant autocorrelation are selected, which include wafer temperatures, chamber side temperatures, etc. Table 3 gives these 16 KSVs. Because the number of KSVs is relatively large, a multivariate statistical approach is taken here. A PCA model is built based on KSVs from 100 wafers at normal operation, where 12 PCs are selected with 99% variance explained (Figure 11), and consequently Q statistics is used for fault detection as shown in Figure 12. Toward the end of the data series, the values of the last four wafers spike abnormally, and these wafers have also been identified as scraped wafers. For diagnosis, the contribution plot is shown in Figure 13, and KSVs with significant contribution to Q are numbers 6–11. From Table 3, the number 6–11 KSVs are obtained from disguised variable B and variable C, and these two variables actually are the measurements of chamber temperature. According to the analysis of the process engineer, there was a leakage of a gas tube at that time, and it resulted in abrupt increase in the chamber side temperature. This coincides with our contribution analysis. Hence, the KSV-based approach can be applied to fault detection and diagnosis in a straightforward manner.

For the second objective, process trend monitoring, we also identify 16 KSVs described in Table 4. A data set containing a PM is used for analysis. A PCA model is built based on first 100 wafers, where eight PCs are selected with 93% variance explained (Figure 14), and the Q statistic of all batches is plotted as shown in Figure 15. This KSI values increase as batch repeats and return to its normal value immediately after PM. The results again indicate that the process trend and the health condition of the tool can be realized using the proposed KSI.

Conclusions

In this work, an insight-based approach, key sensitive index (KSI), is proposed for process trend monitoring in semiconductor manufacturing. From process knowledge, a key sensitive time-slot (KST), which is a time period within a batch is identified. Next, possible key sensitive process variables (KSVs) are selected and validated based on the process trend correlation. A simple batch reactor example is used to illustrate the KSV selection procedure. Then, an index for these variables (KSI) is sought. Two integrated circuit processing examples using real fab data are used to illustrate the KSI-based approach. Applications include process trend monitoring, process quality estimation, and fault detection. The results clearly indicate that the process trend can be captured using the KSI-based approach.

Acknowledgment

This work was supported in part by the National Science Council of Taiwan.

Literature Cited

1. Su AJ, Jeng JC, Huang HP, Yu CC, Hung SY, Chao CK. Control relevant issues in semiconductor manufacturing: overview with some new results. *Control Eng Pract*. In press; doi:10.1016/j.conengprac.2006.11.003.
2. Kourti T, Lee J, MacGregor JF. Experience with industrial applications of projection methods for multivariate statistical process control. *Comput Chem Eng*. 1996;20:S745–S750.
3. Nomikos P, MacGregor JF. Monitoring of batch processes using multi-way principle component analysis, *AIChE J*. 1994;40:1361–1375.
4. Nomikos P, MacGregor JF. Multivariate SPC charts for monitoring batch processes. *Technometrics*. 1995;37:41–59.
5. Neogi D, Schlags CE. Multivariate statistical analysis of an emulsion batch process. *Ind Eng Chem Res*. 1998;37:3971–3977.
6. Kassidas A, MacGregor JF, Taylor PA. Synchronization of batch trajectories using dynamic time warping. *AIChE J*. 1998;44:846–875.
7. Williams BA, Cinar A. An On-line monitoring and curve registration system for multivariable batch processes. Presented at the AIChE Annual Meeting, Los Angeles, CA, 2000.
8. Kourti T. Multivariate dynamic data modeling for analysis and statistical process control of batch processes. *J Chemom*. 2003;17:93–109.
9. MacGregor JF, Kourti T. Statistical process control of multivariate processes. *Control Eng Pract*. 1995;3:403–414.
10. Flores-Cerrillo J, MacGregor JF. Multivariate monitoring of batch processes using batch-to-batch information. *AIChE J*. 2004;50:1219–1228.
11. Box GEP, Jenkins GM, Reinsel GC. *Time Series Analysis: Forecasting and Control*, 3rd ed. Upper Saddle River, NJ: Prentice Hall, 1994.
12. Chiu JB, Su AJ, Yu CC, Shen SH. Planarization strategy of Cu CMP: interaction between plated copper thickness and removal rate. *J Electrochem Soc*. 2004;151:G217–G222.
13. Chao CK, Hung SY, Yu CC. Thermal stress analysis for rapid thermal processor. *IEEE Trans Semicond Manuf*. 2003;13:335–341.
14. Chao CK, Hung SY, Yu CC. Effect of lamp radius on thermal stresses for rapid thermal processing system. *J Manuf Sci Eng (Trans ASME)*. 2003;125:504–511.
15. Huang CJ, Yu CC, Shen SH. Identification and Nonlinear Control for Rapid Thermal Processor. *J Chin Inst Chem Eng*. 2000;31:585–594.

Appendix: Model and Parameters for the Batch Reactor Example

Consider the irreversible order reaction for the batch reactor example (Figure A1):

$$\mathrm{A} \xrightarrow{\text{catalyst}} \mathrm{B} \tag{A1}$$

where A is the reactant and B is the product. The reaction rate can be expressed as:

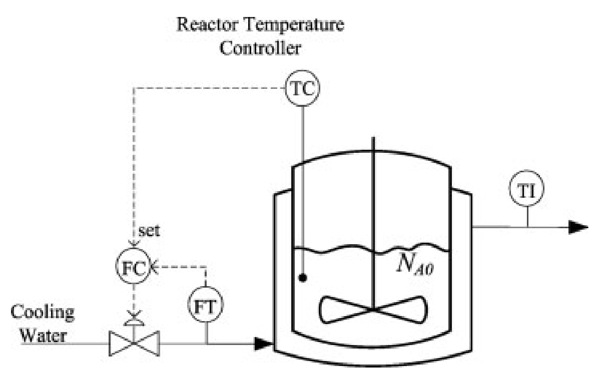


Figure A1. Batch reactor with temperature control.

$$-r_A = k_r C_A \tag{A2}$$

$$k_r = 1.33 a_{cat} \times 10^{10} \exp\left(-\frac{8500}{T}\right) \tag{A3}$$

$$\Delta H_{rxn} = -1000 + (C_{P,B} - C_{P,A}) \times (T - 273) \tag{A4}$$

where C_A is concentration of component A, k_r is the rate constant, a_{cat} describes the activity of the catalyst, ΔH_{rxn} is reaction heat (J/kmol), $C_{P,i}$ is heat capacity (J/kmol/K) of component i, and T is temperature in Kelvin. The simplified nonisothermal dynamic model is given as:

$$\frac{dX}{dt} = k_r(1 - X) \tag{A5}$$

$$\frac{dT}{dt} = \frac{(-\Delta \mathrm{H}_{rxn})(-r_A V) - UA_c(T - T_{co})}{N_{A0} C_{P,A} + (C_{P,A} - C_{P,B})X} \tag{A6}$$

$$\frac{dT_{co}}{dt} = \frac{F_c}{V_c \rho_c}(T_{ci} - T_{co}) \frac{UA_c(T - T_{co})}{V_c \rho_c C_{P,c}} \tag{A7}$$

where X is the conversion of component A, T_{co} is the outlet temperature of cooling water, T_{in} is the inlet temperature of cooling water, F_c is the cooling water flowrate, V is the holdup of the reactor, N_{A0} is initial charge of reactant A (in weight), U is heat transfer coefficient, A_c is the heat transfer area, V_c is the jacket holdup, and ρ_c is the density of cooling water.

System Feed conditions: C_{A0} = 1 (kmol/m^3), $C_{P,A}$ = 26 (J kmol^{-1} K^{-1}), $C_{P,B}$ = 33 (J kmol^{-1} K^{-1}), and N_{A0} (kmol) = V (m^3) $\times$ C_{A0} (kmol/m^3).

Cooling water conditions: T_{ci} = 298 (K), $V_c\rho_c$ = 1 (kg), $C_{P,C}$ = 15 (J kmol^{-1} K^{-1}), and UA_c = 70 (J min^{-1} K^{-1}).

Manuscript received Jan. 15, 2007, and revision received Jun. 20, 2007.