Computers in Industry 64 (2013) 203–213

Contents lists available at SciVerse ScienceDirect

Computers in Industry

journal homepage: www.elsevier.com/locate/compind

Feature extraction, condition monitoring, and fault modeling in semiconductor manufacturing systems

Alexander Bleakie *, Dragan Djurdjanovic

The University of Texas at Austin, Department of Mechanical Engineering, 1 University Station C2200, ETC 5.124, Austin, TX, 78712, United States

ARTICLE INFO

Article history:
Received 1 December 2011
Received in revised form 25 July 2012
Accepted 16 October 2012
Available online 21 November 2012

Keywords:
Condition Monitoring
Linear Discriminant Analysis
Gaussian Mixture Model
Multivariate Gaussian

ABSTRACT

Reliable feature extraction, condition monitoring, and fault modeling are critical to understanding equipment degradation and implementing the proper maintenance decisions in manufacturing processes. Semiconductor manufacturing machines are highly sophisticated systems, consisting of multiple interacting components operating in highly variable operating conditions. This complicates performance monitoring since equipment condition must often be inferred through concurrent interpretation of multiple sensor readings originating from potentially very different subsystems of the tool. This paper presents an integrated approach to feature extraction, condition monitoring, and fault modeling applied to a set of standard built-in sensors on a modern 300-mm technology industrial Plasma Enhanced Chemical Vapor Deposition (PECVD) tool. Linear Discriminant Analysis was utilized to determine the set of dynamic features that are the most sensitive to different tool conditions brought about by chamber cleaning or various faults. Gaussian Mixture Models of the dynamic feature distributions were used to statistically quantify changes of these features as the condition of the tool changed. In addition, four highly detrimental faults were analyzed to demonstrate the fault modeling methodology. Data collected over 8 months from a PECVD tool being operated by a major microelectronics manufacturer was used to verify the methodology. Top sensitive features from various faults observed in this period were examined and physical connections to the chamber condition were interpreted through their behavior.

1. Introduction

Maintenance is one of the key issues in modern semiconductor manufacturing [1,2]. Presently, maintenance scheduling in this industry is predominantly based on historical reliability information pertaining to individual components, and short-term performance signatures for detecting sudden failures. This leads to significant waste due to interventions performed on machines that do not need it or due to missed maintenance actions that result in machine failures, bad product quality, and unscheduled downtimes [2–5].

The condition-based maintenance (CBM) paradigm has been utilized in numerous areas of engineering to address maintenance wastes that currently plague many industries. CBM establishes a connection between the equipment condition and the readings of sensors existing in that equipment [5,6]. Such information about the actual condition of the equipment can be used to make maintenance decisions that are optimally synchronized with human and material resources in the manufacturing system, and the least intrusive on the overall manufacturing operations.

Nevertheless, many challenges exist that have prevented the semiconductor manufacturing industry from pervasively implementing more proactive CBM policies. Some of those challenges are [4,5]:

1. Semiconductor manufacturing tools and equipment are highly sophisticated systems, consisting of multiple interacting components operating in highly variable operating conditions, which often hampers the existing performance monitoring methods in this area.
2. Equipment condition must be inferred through concurrent interpretation of multiple in situ tool sensors, from potentially different subsystems of the tool, measuring at different sampling rates.

The foundation of condition monitoring in any area is the identification of a set of features within the equipment sensor readings that are indicative of the degradation processes, and the utilization of those features with an appropriate condition monitoring method to deduce the equipment health [7]. In

* Corresponding author. Tel.: +1 512 471 0871.
E-mail addresses: a_bleakie@mail.utexas.edu (A. Bleakie), dragand@me.utexas.edu (D. Djurdjanovic).

0166-3615/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.compind.2012.10.002

general, the following feature extraction techniques have been widely utilized for CBM in manufacturing [8]:

- *Time domain methods*: These methods are based on extracting signal features directly from the signal samples (i.e. from the time-domain description of the signal). Those include statistical moments (minimum, maximum, expected value, variance, skewness, kurtosis and higher order moments), as well as dynamic features, such as overshoot, settling time, rise time and others [5,9].
- *Frequency domain methods*: These methods are based on extracting signal features from the frequency domain description of the signal obtained from Fourier analysis or Fast Fourier Transforms [9,10]. Statistical moments of the frequency domain distributions of signal energy and frequency peak locations and intensities are just some of the features that could be extracted in this domain.
- *Time–frequency methods*: These methods are based on extracting signal features from the domains representing the temporal evolution of frequency domain signal representations (i.e. domains representing joint distributions of signal energy over time and frequency domains). Such signal representations can be obtained using methods that include spectrograms of the signal, wavelet transforms and Cohen's class time–frequency signal transformations [10], and are perfectly suited for non-stationary signals whose frequency content varies over time.
- *Model-based methods*: These methods are based on postulating a dynamic model of the sensor readings and quantifying deviations from them [11].

Unfortunately, the standard sensors on even the most modern semiconductor manufacturing equipment are sampled at very low frequencies (1–10 Hz), compared to some other areas, such as aerospace or rotating machinery applications (sampling frequencies in these areas are often on the order of kHz or even MHz ranges) [7]. This greatly limits the feasibility of frequency and time–frequency based feature extraction in this area. Instead, great majority of semiconductor CBM applications utilize raw sensor outputs or their time-domain statistical moments (mean, variance, etc.) for monitoring purposes. This is appropriate whenever sensor readings are very closely correlated to the monitored or controlled processes. For example, the thermocouple readings from a chemical vapor deposition (CVD) chamber can be used directly as a monitored feature in statistical process control (SPC) charts [15]. In addition, the in situ particle counts can be used as an indicator of chamber contamination and are hence directly used for monitoring, if these sensors are available on the equipment. The chemical mechanical planarization (CMP) process is an example where layer thickness, copper residuals, and the wafer temperature are directly measured and used as monitoring parameters in [12]. Spanos et al. [13] depart from this trend of utilization of raw sensor readings for monitoring and present a novel feature extraction method for plasma-etch equipment sensors. Time series models [14] for each of these signals are utilized to analyze residuals of the signals after eliminating trends.

The second fundamental enabler of proactive CBM policies is the approach to condition monitoring of the equipment using the features extracted from the sensor readings. Condition monitoring techniques determine the current system or subsystem conditions based on the features extracted from the relevant sensor readings [7]. SPC and advanced process control (APC) approaches, based on statistical or pattern recognition methods, are commonly utilized in semiconductor manufacturing. Montgomery [15] gives a thorough description of SPC, along with applications from various industries. SPC methods can detect abnormalities in the process as statistically significant departures in the behavior of process signatures from their behavior during normal conditions. For example, Spanos et al. [13] raised alarms if the univariate T^2 statistic of relevant sensor features [16] falls outside the statistically determined process control limits. Bunkofske [17] used SPC for condition monitoring after reducing the number of equipment features using principal component analysis. Mai [18] used SPC in monitoring of the contamination inside a lithography tool to prevent defects on the products caused by excessive tool contamination.

APC methods for process monitoring can be seen as a special class of SPC methods that use elaborate multivariate models, such as artificial neural networks (ANNs) or multivariate regression models to capture process dynamics. These methods are increasingly being employed in recent years to deal with diagnostic challenges associated with larger wafer sizes, higher costs, and smaller critical dimensions of electronic components. Card et al. [19] utilized neural network prediction errors to monitor and control a plasma etch process. Pompier et al. [20] monitored a multi-chamber oxide deposition process by tracking individual chambers using APC methods. Velichko [21] proposed a nonlinear multiple-input, multiple-output (MIMO) model-based APC framework for semiconductor manufacturing processes, and shows its benefits over tracking raw features. Baek et al. [22] analyzed the electron collisions during plasma processes using APC methods in order to identify small, previously unknown, trends between wet cleans.

The next fundamental enabler of CBM is fault diagnosis. This level generates diagnostic records by identifying fault possibilities based on the information of current condition and known faulty conditions. This layer essentially enriches the condition monitoring information showing that "something is abnormal in the process" with the diagnostic information about what is the reason for that abnormality (what fault caused it). The predominant approach for fault diagnosis currently used in the semiconductor industry is based on the use of the SPC and APC concepts [15]. SPC charts are usually constructed in a way that each of them corresponds to one specific fault. Thus, whenever the monitored features violate the corresponding limits, the presence of a particular fault is established. These warning limits can also provide a statistical significance to give the user an assessment of how accurately the tool health is being estimated. For example, Chen et al. [23] reported using optical emission spectroscopy (OES) data to provide a real-time SPC diagnosis scheme on the plasma performance, as well as to detect faults during the etch process. Matsuda et al. [24] presented the use of APC for equipment monitoring, error detection, and CBM in semiconductor thermal process.

In addition to the SPC/APC methods, artificial intelligence and classification techniques have been employed in assessing the health of semiconductor fabrication systems. For example, Salahshoor and Keshtgar [25] proposed a method which performs Independent Component Analysis followed by a Neural Network classification. This method is used to overcome false alarms and missed fault detections frequently observed when conventional monitoring techniques deal with large number of observation variables. Tu et al. [26] presented the results using principal component analysis (PCA) for fault detection and classification in 300-mm high-density plasma CVD tools.

In terms of health indicators, Blue and Chen [27] proposed the generalized moving variance as a tool health indicator, which is dependent on the changes of recipe in the semiconductor fabrication process. Djurdjanovic et al. [6] proposed a generic concept of 'confidence values' as an index to reflect how healthy the system is by evaluating the overlap between the most recently observed features and those observed during normal operation.

This paper presents integrated methods for feature extraction, condition monitoring, and fault modeling applied to multiple, simultaneously sampled sensor readings from all relevant subsystems of an industrial Plasma Enhanced Chemical Vapor Deposition (PECVD) tool commonly used in the semiconductor manufacturing industry. The data set used in our study encompasses 8 months of high volume production in a major domestic semiconductor fab. The remainder of the paper is organized as follows. In Section 2, PECVD tools are described briefly, followed by a description of the layout of the subsystems, sensors, and operation cycles of the tool. Section 3 describes the methodology utilized in this work. Sections 3.1 and 3.2 respectively list the dynamic features that were extracted from the tool sensors, and introduce the Linear Discriminant Analysis (LDA) as a viable method for determining which of these dynamic features are the most sensitive to tool condition changes. Section 3.3 introduces the overlap of probability density functions of features yielded by the tool at various stages of its degradation as a way to quantitatively monitor the tool condition. Section 3.4 presents a technique for modeling and assessing a specific condition of the tool given feature patterns corresponding to the underlying faults. Section 4 gives the results obtained through our study. The data set collected from a major domestic 300 mm fab is described in Section 4.1, while Section 4.2 presents the results of monitoring of the so-called accumulation drifts caused by short-term degradation between cleaning cycles. Section 4.3, presents the fault modeling results of four highly impacting PECVD tool faults that occurred in the tool operation period considered in this paper. Section 5 presents a discussion of the results. Finally, conclusions and future work are discussed in Section 6.

2. Description of PECVD tool

PECVD tools are used for depositing thin films onto silicon wafer substrates, which is one of the crucial steps in manufacturing of microelectronic circuits and solar cells. It is the most common method for producing conductors and dielectrics with excellent film growth properties necessary for small chip components [28]. Inside a PECVD tool chamber, reactive gases pass over silicon wafers and are absorbed onto the surfaces to form a thin layer. The gases are excited through radio frequency (RF) electrical power that creates energetic plasma used to deposit the film on the wafers. The plasma state allows the reaction to take place at lower temperatures, more suitable for large silicon wafers. Ultimately, many stacked layers of conducting and insulating films with etched patterns between them (forming thousands of microscopic electrical components) form an integrated circuit [28].

2.1. PECVD tool subsystems and sensors

A general PECVD tool is composed of a reaction chamber, radio frequency (RF) plasma generation system, gas delivery system, wafer load locks, and a robotic arm to carry wafers to and from the tool. Fig. 1 shows a diagram of the main components of a PECVD tool. The RF matching network for generating plasma is shown on the top-left of the diagram. The high frequency energy is sent through two matching capacitors (load and tune capacitors) that control the power delivered to the chamber. By varying their capacitances, the capacitors try to tune the impedance of the circuit to the load impedance of the chamber and thus deliver maximum RF power to the gases in the chamber [28]. The RF energy excites the flowing gas into the plasma state necessary for lower temperature depositions [28]. The gas delivery system is depicted on the right of Fig. 1. It consists of mass flow controllers (MFCs) for each gas used in various depositions. Gas flows over precise time intervals to ensure processing of specific thin-film recipes. A control valve (bottom of Fig. 1) controls the chamber pressure and evacuates deposition gases from the chamber. Temperature controlled top and lower chamber plates enclose the chamber and the walls are heated to minimize on-wall deposition, as well as to speed up the reaction during the automatic cleaning process.

The PECVD tool used in this study was a standard 300 mm wafer tool with numerous in-place sensors measuring the physics of the

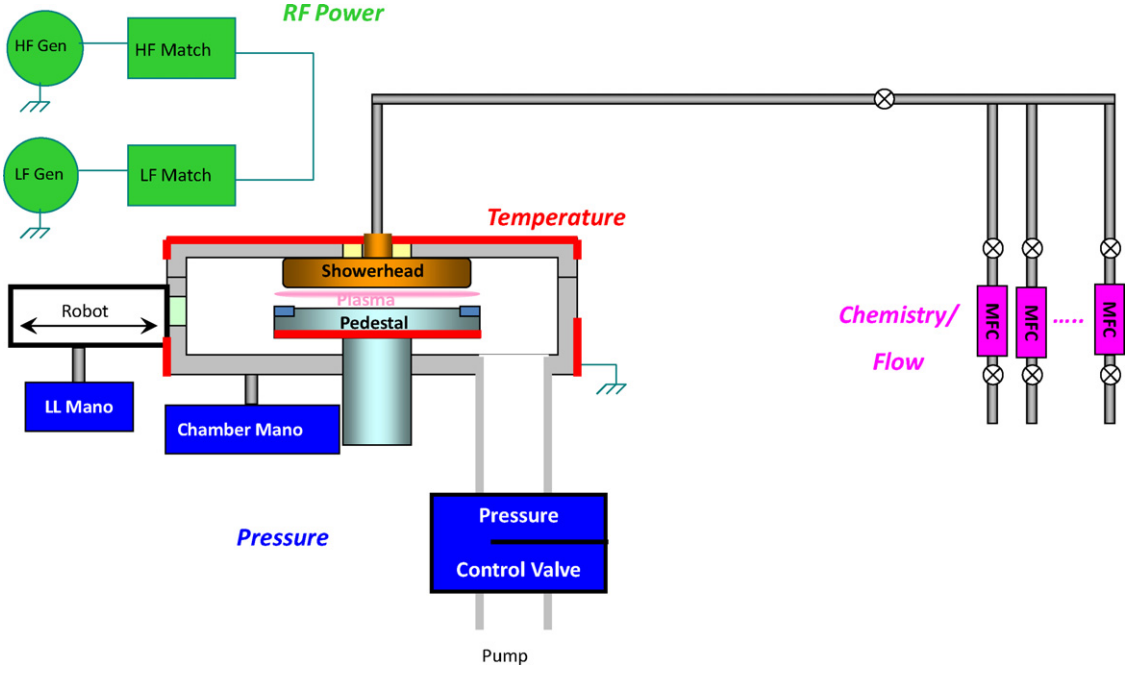

Fig. 1. Diagram representing the various components that make up the PECVD tool.

process (to ensure real-life applicability of our study, only the standard on-board sensors were analyzed and no additional sensors were considered). The signals used in this study were the RF power characteristics (forward, load, and reflected power), voltages of RF matching network capacitors, MFC flow rates, top plate temperature, chamber temperature, pedestal temperatures, chamber pressure, and the pendulum valve angle. All sensor readings were concurrently collected using a 10 Hz sampling rate, which is an order of magnitude higher than the prevalent 300 mm fab standards.

2.2. Tool operation and maintenance schedule

Various chemical compounds can be deposited using PECVD tools. Silicon nitride (Si–N) and silicon dioxide ($Si–O_2$) are some of the most common thin films deposited in these tools, even though other compounds can be used, depending on the conductivity, mechanical and reliability requirements on the film [28]. Silane (SiH_4) and tetraethyl orthosilicate (TEOS) are common reactants used to produce these films.

In addition to the deposition cycles that contribute to the process of chip-making, PECVD tools in semiconductor manufacturing facilities also perform automatic in situ cleaning programs after a predetermined total film accumulation limit (corresponding to approximately 25–100 wafers, depending on the film thickness). The usual way of performing the in situ cleans is by flowing plasma-excited Fluorine (F^-) into the chamber to eat away films deposited on the tool surfaces. These cleans are performed periodically in order to bring the tool back into a lower state of degradation. Unfortunately, this tool-cleaning procedure is imperfect since residual films can be left in parts of the chamber and at the same time, some tool surfaces can also be etched away during the process. This results in a long-term degradation of the tool, which over time leads to the production of wafers with noticeable defects, unless preventive maintenance (PM) actions are undertaken. Thus, besides the short-term accumulation drifts caused by successive wafer depositions and remedied via in situ cleans, one can also observe a long-term drift of the tool condition as numerous in situ clean cycles are executed.

Fig. 2 illustrates the scheduling for different levels of PECVD tool cleaning and maintenance. An automatic in situ clean program is performed after depositions on a predetermined number of wafers (approximately, every 25–100 wafers). The tool loops through these programs using different chemistries and physical parameters until the fixed-time maintenance schedule requires a long-term PM intervention (approximately, every 25,000–100,000 wafers). This long-term PM action usually consists of a physical wipe down of the chamber and replacements of various critical tool components.

The next section details the methods for feature extraction, condition monitoring, and fault modeling used in this study.

3. Architecture of feature extraction, condition monitoring and fault modeling

3.1. Feature extraction methodology

Extraction of signal features that are the most descriptive of machine performance is one of the key elements of CBM [7]. Useful information can be gathered from extracting dynamic features by utilizing them in prediction models for capturing the features' time-series performance. Predictions of system degradation, remaining useful life, and sudden developing failures are all examples of critical information that are found through dynamic features directly extracted from raw data. In this study, we extracted numerous features from multiple sensor readings, including dynamic features such as rise-time, overshoot, and steady state values [9], along with statistical features, such as mean value, variance and range [5,6]. These features were chosen due to their commonplace use in dynamic systems and stochastic process theory. The 40 features extracted from PECVD tool sensor readings in this study are summarized in Table 1.

Each feature is calculated from its corresponding sensor monitoring the tool, while these sensors were selected by the manufacturer to monitor the physics relevant to the processes. For each signal, one can visually look at the raw traces to determine which dynamic features can be extracted from the profile and in which domain (time or frequency). As an example, sensors with higher sampling rates will be able to utilize other domains (e.g. frequency or wavelet parameters) for extracting dynamic features. The chosen features in this study are extracted from signals corresponding to consistent processes (for example, depositions, pre-coats, in situ clean, etc.), by comparing the features of the same process throughout the tool operation. Signal features that are the most descriptive of tool condition and degradation were sought within this comprehensive feature set.

Frequently, more information can be gathered from the system if additional sensors and/or dynamic features are incorporated that capture the needed dynamics for the application, while insignificant ones can be filtered through principal component analysis or discriminant analysis as to be described in Section 3.2. Depending on the complexity of the subsystem dependencies, monitoring an exhaustive set of dynamic features can bring out relationships that would otherwise be unknown if one was monitoring the subsystems or sensors separately. By extracting dynamic features inspired from dynamic systems and stochastic processes, one can capture the dynamics of a large part of the system physics and the results will be the same if one extracted more features from the existing sensors.

In order to perform the multivariate analysis of the numerous features extracted, the features were standardized to eliminate the physical units and thus make them dimensionally homogeneous. Standardization of each feature was accomplished by subtracting

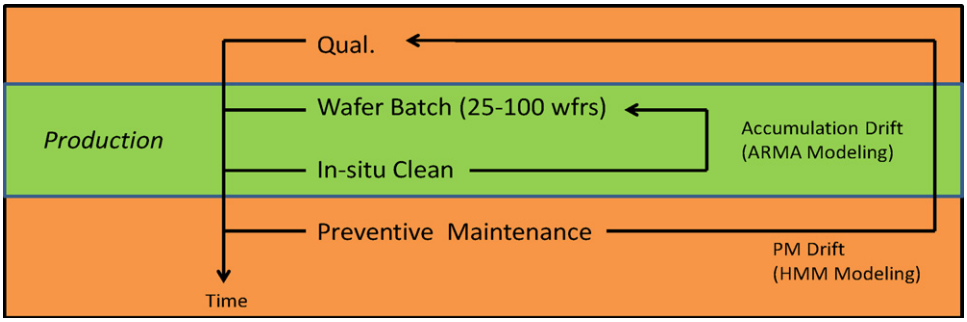

Fig. 2. Production cycle of the PECVD tool.

Table 1
Features extracted from the data analyzed in this study.

Signal	Signal features			
Top Plate Temperature	Mean	Minimum	Amplitude	
Chamber Temperature	Mean	Minimum	Amplitude	
Pedestal 1 Temperature	Mean	Minimum	Amplitude	
LF Forward Power	Steady State Error	Tune Time		
LF Load Power	Steady State Error	Tune Time		
LF Reflected Power	Steady State Error	Tune Time	Maximum	
HF Forward Power	Steady State Error	Tune Time		
HF Load Power	Steady State Error	Tune Time		
HF Reflected Power	Steady State Error	Tune Time	Maximum	
Load Capacitor Voltage	Steady State	Tune Time	Overshoot High	Overshoot Low
Tune Capacitor Voltage	Steady State	Tune Time	Overshoot High	Overshoot Low
Pendulum Valve Angle	Steady State	Maximum		
Process Chamber Pressure	Steady State Error	Rise Time	Overshoot	Minimum
Liquid Flow Rate TEOS	Steady State Error	Rise Time	Overshoot	

its mean and dividing it by its standard deviation, where the mean and standard deviation were calculated from the data set considered to be representative of normal operation.

3.2. Sensitivity analysis methodology: Linear Discriminant Analysis

The analysis of sensitivity of the feature set to various equipment conditions was accomplished via comparison of features extracted from the sensor readings emitted during times when the tool operated under normal and various faulty conditions. Changes in the dynamic features that are caused by the changes in the tool conditions are examined using statistical sensitivity analysis based on evaluating the amount of change seen in the feature data as the tool shifts from one condition to another (presence or absence of a fault).

More specifically, Linear Discriminant Analysis (LDA) [29] is used for finding the most sensitive features between two classes of data. LDA is a method used in statistics and machine learning to find a linear combination of features that separate two or more classes of data the best. The resulting combination may be used as a linear classifier, or, more commonly, for dimensionality reduction before later classification. Other classifiers such as a Quadratic Classifier [29] or a Support Vector Machine [30] may be superior in dealing with multi-class problems and non-linear boundaries between classes, but they require solving of large-scale optimization problems, while LDA requires the solutions to a simple Eigenvalue problem [29]. This is particularly important when one deals with a feature set of large dimensionality, as we do in this study. A brief description of LDA adapted from [29] is enclosed below:

Suppose the unit-less feature vector extracted at a processing cycle is denoted by

$$\mathbf{x} = [\, f_1 \quad f_2 \quad f_3 \quad f_4 \quad \dots \quad f_N \,]^T, \tag{1}$$

where f_i represent individual features from sensor readings (for us, this vector consists of 40 normalized features listed in Table 1). Let the data set be partitioned into two classes (clean/dirty tool, or pre/post maintenance) denoted ω_1 and ω_2, with n_1 and n_2 samples in each class, respectively. Then, the mean vector for the ith class ω_i, of all the feature vectors belonging to the class is

$$\boldsymbol{\mu}_\mathbf{i} = \frac{1}{n_i} \sum_{x \in \omega_i} \mathbf{x} \tag{2}$$

The so-called "within-class" scatter matrices are calculated as

$$\mathbf{S}_\mathbf{i} = \sum_{\mathbf{x} \in \omega_i} (\mathbf{x} - \boldsymbol{\mu}_\mathbf{i})(\mathbf{x} - \boldsymbol{\mu}_\mathbf{i})^T \tag{3}$$

while the overall within class scatter matrix is found as

$$\mathbf{S_w} = \mathbf{S}_1 + \mathbf{S}_2 \tag{4}$$

LDA finds a vector pointing in the direction that can be interpreted as the direction in which the features having the largest between class separation and smallest within class localization. More specifically, the principal discriminant unit vector is found by maximizing the so-called Rayleigh quotient, which leads to the solution in the form

$$\mathbf{w} = \mathbf{S}_\mathbf{w}^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_2) \tag{5}$$

The features with the most weight in the principal discriminant unit vector can be interpreted as being the most sensitive between the two classes analyzed. Furthermore, one can also use projections of feature vectors onto the principal discriminant vector **w** in order to analyze the data in one-dimensional space where the classes are most separated. The sensitive features or their projections can then tracked using the overlaps of their probability density functions characterizing different tool conditions, as described in the next section.

3.3. Performance monitoring methodology: Gaussian mixture models and confidence value analysis

In order to quantify the changes in the most sensitive features, one can track the probability density functions (PDFs) of the projections of feature vectors along the direction in which changes in the feature space are the most prominent (**w**-vector described in the previous section). Alternatively, we can also track a multi-dimensional PDF of a subset of features that were determined by the sensitivity analysis to be the most sensitive to a given tool condition (features associated with the highest weights in the **w**-vector).

Following [6], we will track the overlapping volume between feature PDFs characterizing normal and current system behavior. As in [6], we will refer to this overlap as the performance confidence value (CV). The CV between the two probability distributions is calculated as

$$CV = \frac{\int \mathbf{g}(\mathbf{x})\mathbf{f}(\mathbf{x})d\mathbf{x}}{\sqrt{\int \mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})d\mathbf{x}} \cdot \sqrt{\int \mathbf{f}(\mathbf{x})\mathbf{f}(\mathbf{x})d\mathbf{x}}} \tag{6}$$

where $\mathbf{g}(\mathbf{x})$ and $\mathbf{f}(\mathbf{x})$ represent the feature PDFs corresponding to the normal and current system behaviors, respectively.

The integral, $\int \mathbf{g}(\mathbf{x})\mathbf{f}(\mathbf{x})d\mathbf{x}$ can be seen as the inner product in the space of PDFs, while $\sqrt{\int \mathbf{g}(\mathbf{x})\mathbf{g}(\mathbf{x})d\mathbf{x}}$ and $\sqrt{\int \boldsymbol{f}(\boldsymbol{x})\boldsymbol{f}(\boldsymbol{x})d\boldsymbol{x}}$ denote the corresponding norms of **g(x)** and **f(x)** respectively. Thus, the concept of a CV can be interpreted as the "cosine" of the angle between any two vectors in the Euclidean vector space of PDFs [31]. In the case of a perfect match (the currently observed tool

operation is identical to that observed during normal behavior), the distributions will be directly on top of one another and the CV will be equal to one. As the system degrades, the PDF of the features describing the most recent behavior will drift away from the PDF corresponding to the normal system behavior, thus causing the CV to fall toward zero. Fig. 3 graphically illustrates the concept of performance CVs.

The feature PDFs described in this work were approximated using Gaussian Mixture Models (GMMs) [32] because of their universal approximation capabilities and their analytical tractability. This enables easy tracking of feature PDFs through on-line updating of the mean vectors and covariance matrices of the Gaussian components in the GMM description. In addition, analytical tractability of GMM approximations also enables one to analytically calculate performance CVs, which is of particular importance in large-dimensional feature spaces, such as the one dealt with in this study.

More specifically, let feature PDFs $\mathbf{f}(\mathbf{x})$ and $\mathbf{g}(\mathbf{x})$ in Eq. (6) be described with GMMs containing respectively M and N components, i.e. let

$$\mathbf{f}(\mathbf{x}) = \sum_{i=1}^{M} \alpha_i \mathbf{f}_i(\mathbf{x}) \tag{7}$$

and

$$\mathbf{g}(\mathbf{x}) = \sum_{j=1}^{N} \beta_j \mathbf{g}_j(\mathbf{x}) \tag{8}$$

where α_i and β_j are the mixing coefficients, and $\mathbf{f}_i(\mathbf{x})$ and $\mathbf{g}_j(\mathbf{x})$ are the corresponding Gaussian components in a D-dimensional feature space. These components are described as

$$\mathbf{f}_i(\mathbf{x}) = \frac{1}{(2\pi)^{D/2}\sqrt{det(\bar{\Sigma}_{\mathbf{f}i})}} e^{[-1/2(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{f}i})^T \bar{\Sigma}_{\mathbf{f}i}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{f}i})]} \tag{9}$$

$$\mathbf{g}_j(\mathbf{x}) = \frac{1}{(2\pi)^{D/2}\sqrt{det(\bar{\Sigma}_{\mathbf{g}j})}} e^{[-1/2(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{g}j})^T \bar{\Sigma}_{\mathbf{g}j}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{\mathbf{g}j})]} \tag{10}$$

where $\boldsymbol{\mu}_{\mathbf{f}i}$, $\boldsymbol{\mu}_{\mathbf{g}i}$, $\bar{\Sigma}_{\mathbf{f}i}$ and $\bar{\Sigma}_{\mathbf{g}j}$ denote the mean vectors and covariance matrices of the corresponding Gaussian PDF component. In order to estimate a GMM, the expectation maximization (EM) algorithm can be used to iteratively update the means, covariance matrices, and mixing coefficients of each Gaussian component [32]. A similar approach was used in [33] for monitoring of various automotive systems. Furthermore, in this context, the CV described by (6) can be calculated analytically as

$$CV = \frac{\sum_{i=1}^{M}\sum_{j=1}^{N}\left(\alpha_i\beta_j\right)\left[\frac{1}{(2\pi)^{D/2}\sqrt{det(\bar{\Sigma}_{\mathbf{f}i}+\bar{\Sigma}_{\mathbf{g}j})}}\right] e^{[-\frac{1}{2}(\mu_{\mathbf{f}i}-\mu_{\mathbf{g}j})^T(\bar{\Sigma}_{\mathbf{f}i}+\bar{\Sigma}_{\mathbf{g}j})^{-1}(\mu_{\mathbf{f}i}-\mu_{\mathbf{g}j})]}}{\sqrt{\sum_{i=1}^{M}\sum_{j=1}^{M}(\alpha_i\alpha_j)\left[\frac{1}{(2\pi)^{D/2}\sqrt{det(\bar{\Sigma}_{\mathbf{f}i}+\bar{\Sigma}_{\mathbf{f}j})}}\right] e^{[-\frac{1}{2}(\mu_{\mathbf{f}i}-\mu_{\mathbf{f}j})^T(\bar{\Sigma}_{\mathbf{f}i}+\bar{\Sigma}_{\mathbf{f}j})^{-1}(\mu_{\mathbf{f}i}-\mu_{\mathbf{f}j})]}}\sqrt{\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\beta_i\beta_j\right)\left[\frac{1}{(2\pi)^{D/2}\sqrt{det(\bar{\Sigma}_{\mathbf{g}i}+\bar{\Sigma}_{\mathbf{g}j})}}\right] e^{[-\frac{1}{2}(\mu_{\mathbf{g}i}-\mu_{\mathbf{g}j})^T(\bar{\Sigma}_{\mathbf{g}i}+\bar{\Sigma}_{\mathbf{g}j})^{-1}(\mu_{\mathbf{g}i}-\mu_{\mathbf{g}j})]}} \tag{11}$$

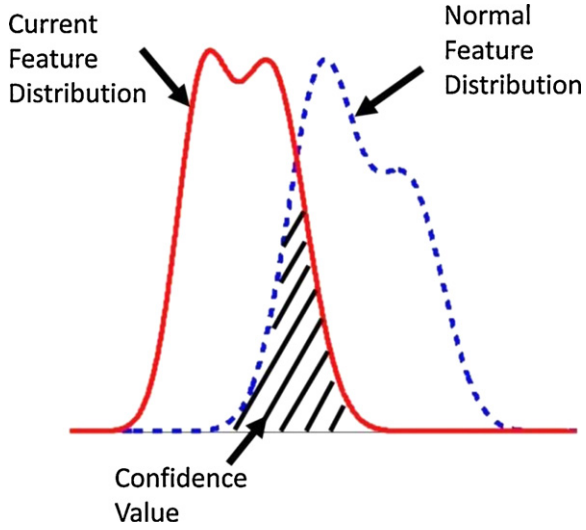

Fig. 3. The concept of confidence values (CV), characterizing the changes in the distributions of features sensitive to system degradation [6].

3.4. Fault modeling through characterization of documented failures using LDA and GMMs

Diagnosis of machine performance is based on recognizing a fault in the monitored system through comparison of the most recently observed features with feature patterns corresponding to various faults available from the past system operation [5,6]. Similarity of the current system condition to a given fault can be evaluated using the overlap between PDFs of the most recently observed feature vectors and feature vectors observed in the presence of that fault, similarly to how CVs are used to monitor the proximity of the current system behavior to the normal one [6]. Fig. 4 illustrates this concept. The feature PDFs corresponding to normal and various faulty conditions known from previous operation can be identified from the training set by, for example, estimating the means, covariance matrices, and mixing coefficients of a GMM. Through their comparison with the most recently observed feature PDFs (current system behavior), one would expect to see the system degrading (reduction in CV) and drifting into a specific faulty regime (or perhaps drifting to an area never seen during training, which is also important diagnostic information indicating that a new condition has been observed). Fig. 4 illustrates the situation in which system degradation is visible through the drifting of the monitored features toward the area characteristic of Fault 1 behavior (indicating that Fault 1 is developing in the monitored system). Consequently, the drop in CVs (indicating system degradation) is accompanied with the rise in the feature PDF overlap with Fault 1 features (indicating increasing proximity to Fault 1), while the overlap with features representative of Fault 2 remains low.

4. Application of feature extraction, condition monitoring and fault modeling methods on a factory dataset

4.1. Description of dataset

The data set used in this study was gathered from a standard 300 mm PECVD tool operating in the facilities of a major domestic

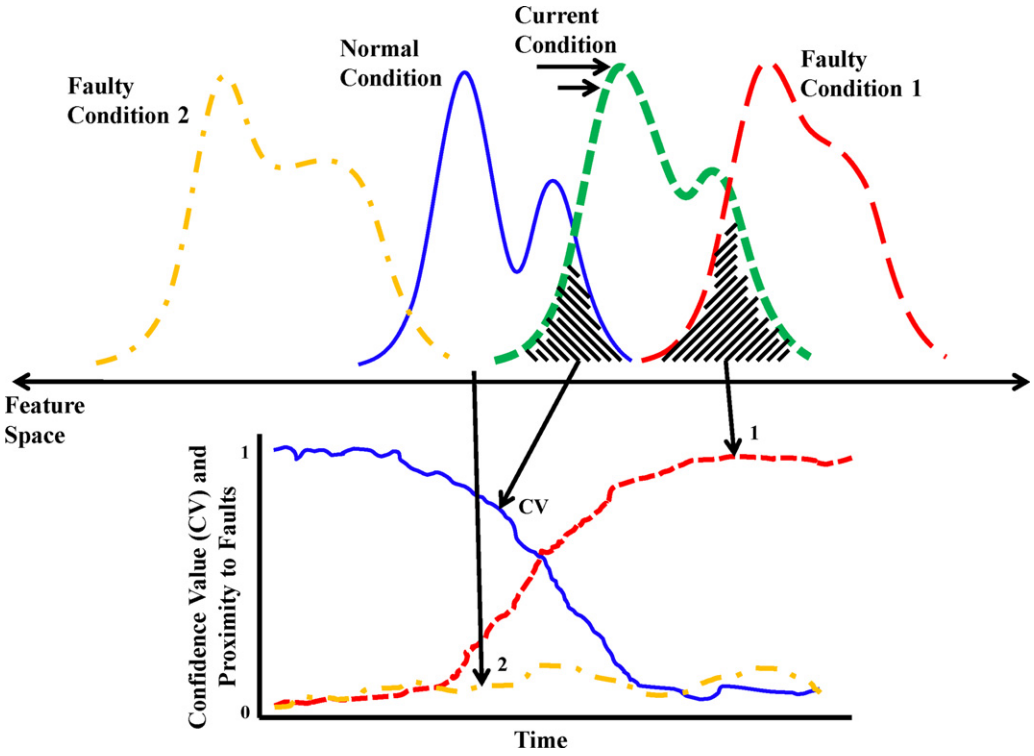

Fig. 4. Illustration of fault modeling, where the current system behavior drifts toward Fault 1. The current system degradation and proximity to all known faults can be quantified using the overlaps of PDFs.

manufacturer of integrated circuits. It was obtained from a single tool continuously depositing recipes of $Si–O_2$ films with various thicknesses (TEOS is used as the main reactant). In this study, we focused on roughly 8 months of production data and on the recipe (film thickness) that significantly dominated the operations (about 80% of operation). All the signals in Table 1 are recorded with a sampling rate of 10 Hz. The approximate size of the entire data set for the recipe considered here corresponds to about 40,000 wafers, with 1500 in situ cleans, and two scheduled preventive maintenance interventions (open chamber wipe downs, often referred to as "wet cleans"). Four events of unacceptable tool behavior occurred within the dataset. These faults corresponded to out-of-specification wafer thicknesses, bi-modal wafer thicknesses, particle contamination, and abnormal plasma formations, referred to as "Coulomb crystals" [34]. These faults had a high impact on the productivity of the tool, with particle formation and plasma formation events keeping the tool down for more than 2 weeks each. In addition to the impact on production, significant effort was spent by technicians in finding the root causes of the faults, incurring further costs on the company. More details on the faults will be given later. The methodologies described in Section 3 were used to analyze the chamber degradation between in situ cleaning events and as well as to model the behavior around the four documented faults.

4.2. Analysis of scheduled in situ clean events

The following are the results of the analyses aimed at finding the features that are the most sensitive to short-term accumulation drifts in between in situ cleans (i.e. features that probabilistically change the most between two in situ cleans).

The data utilized for the in situ clean analysis corresponds to about 80 wafer batches (40 wafers per batch), with in situ cleans performed between them. In this relatively short period, corresponding to a few weeks of production,[1] long-term feature drifts or sudden changes were not detected, and the analysis can be considered to address only the short-term accumulation drifts between in situ cleans.

Sensitivity analysis to degradation between in situ cleans was performed by applying LDA to classes formed with the feature set obtained *just prior to the in situ cleans* (the last 5 depositions just before each of the cleans in the training set) and *depositions just after the in situ cleans* (the first 5 depositions just after each of the cleans in the training set). In total, there were 400 deposition cycles composing each of the classes. Table 2 lists the top 10 sensitive features identified by the LDA. All other features had significantly smaller weighting in the principal discriminant vector and showed no visible changes as depositions progressed.

Table 2 indicates that the load capacitor features make up the top sensitive features. This is plausible because the load capacitor signal sees a different chamber impedance before and after the in situ cleaning, and hence its tuning characteristics change in order to match the impedance of the chamber and deliver the maximum RF power to the plasma. Several temperature features also appear to be sensitive to degradation between in situ cleans. This also matches the engineering intuition since progressive depositions leave increasingly thicker films on the pedestal and chamber walls,

Table 2
Results of the sensitivity analysis between pre and post in situ clean features.

Principal vector (w)	Feature name
0.60	Load Capacitor Overshoot Low
0.41	Load Capacitor Overshoot High
0.40	Load Capacitor Steady State
0.27	HF Load Power Steady State Error
0.26	Top Plate Temperature Mean
0.25	Top Plate Temperature Minimum
0.22	Pedestal 1 Temperature Minimum
0.21	Pedestal 1 Temperature Mean
0.13	LF Reflected Power Tune Time
0.06	Process Pressure Minimum

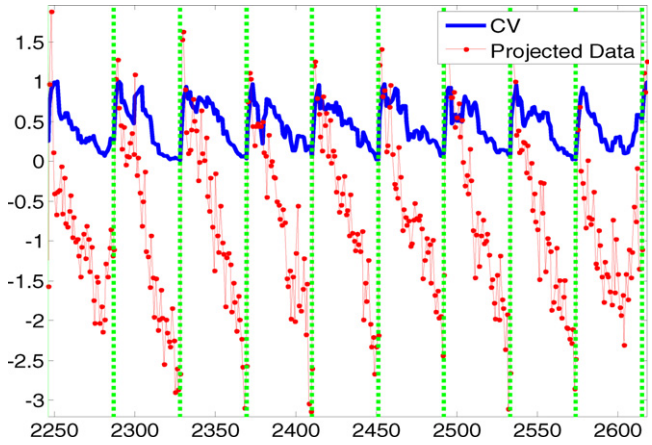

Fig. 5. Sensitive features and CV displaying statistical drift between in situ cleans (projected features).

thus gradually changing their thermal emissivity, which is mirrored in the corresponding temperature features.

The changes in these sensitive signal features are analyzed using the CV concept described in Section 3.3. A feature PDF representative of normal tool operation was constructed by fitting a GMM to the features emitted during depositions just after the in situ cleans. The CV was updated with each new deposition by re-evaluating the overlap between the GMM of features corresponding to normal tool operation (just after the in situ cleans) and the GMM of features corresponding to most recent deposition cycles.[2] Throughout the study, we used bi-modal GMMs based on the visual examination of the feature distributions, though the number of Gaussians in a GMM can also be determined more formally [29].

Fig. 5 presents the results of condition monitoring corresponding to 9 wafer batches in our dataset. The figure illustrates results of tracking of feature projections onto the principle discriminant vector yielded by the LDA, as well as the performance CVs corresponding to those projections.

The CVs show obvious recoveries after each clean and then gradual drops as the next clean nears. This is indicative of the features having a distribution well aligned with the template distribution immediately after the clean, and then drifting as more and more wafers are deposited. The slope and fluctuation of the CV can be controlled by the sliding window for various applications; however, the trends remain intact.

Due to the consistent behavior of the trending CV, one can use this as an indicator to perform timely in situ cleans on a PECVD tool. One possible way is by utilizing historical maintenance records, one would have to find out when improper maintenances occurred (too early, too late, or failure) and develop distributions of the CV readings at those scheduled maintenances for each category and for the proper maintenances. These distributions would show the likelihood of obtaining improper and proper maintenance using CVs and form the maintenance thresholds using basic statistics. If all the distributions overlap, then improper maintenance may be unobservable and additional data is needed to capture the dynamics. Such a condition driven maintenance trigger should reduce maintenance risks compared to the currently established purely usage-based in situ clean triggering.

[1] Due to the proprietary nature of this information, we cannot be more specific.
[2] The moving window of recent deposition cycles would not cross over the in situ clean in order not to mix cycles before and after cleans.

4.3. Modeling of documented PECVD faults

The following are the results of the analyses aimed at modeling of four specific events of unacceptable equipment behavior that were documented in our dataset. Fault 1 was a large spike in wafer thickness measurements observed on about 150 wafers (three batches), which caused a tool shutdown and significant production waste. Fault 2 was detected after noticing bi-model wafer thicknesses. These two faults are relatively common in deposition tools and are known to cause much downtime in the semiconductor manufacturing fab utilized in our work. Fault 3 was caused by particle contamination on the wafers and eventually forced a very costly loss of more than two weeks of production. Fault 4 was also a highly detrimental event caused by unwanted plasma formations below the showerheads that led to particle formations on wafers and more than three-weeks of tool shutdown. It took numerous conversations between the members our group, tool operators and tool developers, which ended with a quantum physics Ph.D. (who closely worked on the tool design) determining that the cause of the plasma formations was the phenomenon termed Coulomb crystals [34]. Since we are not the experts in quantum physics, we can only loosely describe it as a fault caused by improper evaporation of TEOS and the existence of a liquid phase in the plasma. As discussed above, Faults 3 and 4 were very costly to the fab, forcing tool shutdowns for weeks in order to investigate the root causes of the contaminated wafers and bring the wafer specifications back into control.

Our fault-modeling analysis was aimed at investigating how each of these faults is expressed in the domain of features extracted from equipment signals. The features patterns that are the most indicative of those faults can then be used to diagnose the fault by recognizing similar feature patterns in future tool operation. In order to accomplish this, LDA is performed for each of these four faults and the top *two* sensitive features (i.e. two features that probabilistically change the most in between normal tool operation and the operation around the time of the fault) are observed in the feature space to document the most significant feature shifts necessary for diagnosis of each fault.

As mentioned before, Fault 1 caused a significant spike in wafer thickness measurements, while Fault 2 caused bi-model wafer thickness measurements for several wafer batches. In both cases,

Table 3
The most sensitive features between normal tool behavior and Fault 1 (a), and Fault 2 (b).

Principal vector (w)	Feature name
(a)	
0.64	LF Load Power Steady State Error
0.44	Chamber Temperature Mean
0.36	Pedestal Temperature Mean
0.29	HF Forward Power Tune Time
0.25	HF Load Power Tune Time
0.25	Pedestal Temperature Amplitude
0.16	TEOS Flow Steady State Error
0.13	HF Load Power Steady State Error
0.11	TEOS Flow Rise Time
0.03	Load Capacitor Voltage Overshoot
(b)	
0.54	Process Pressure Rise Time
0.43	TEOS Flow Rise Time
0.41	LF Load Power Steady State Error
0.41	LF Reflected Power Tune Time
0.27	Pedestal Temperature Mean
0.23	HF Load Power Tune Time
0.20	HF Load Power Steady State Error
0.11	Process Pressure Overshoot
0.10	HF Reflected Power Maximum
0.01	Load Capacitor Voltage Overshoot

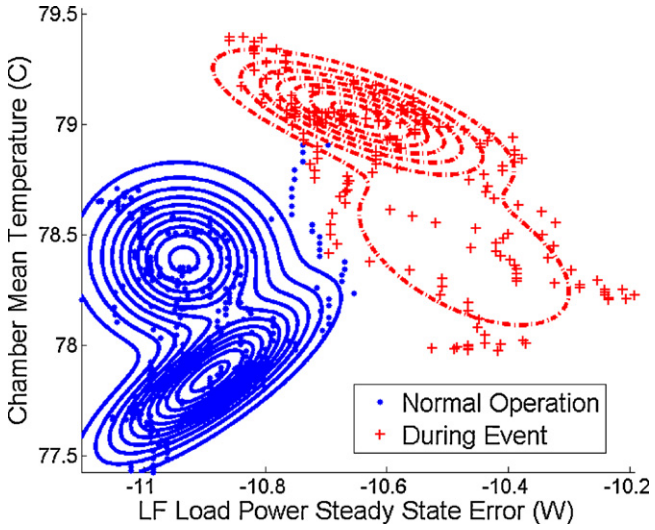

Fig. 6. Top sensitive features for Fault 1 along with normal operation, with GMM contours.

equipment operators were forced to adjust deposition times in order to continue processing within specification. We observed features of the sensor readings emitted during those faults and LDA yielded the most sensitive features listed in Table 3(a) and (b).

Fig. 6 displays the contour plot of PDFs of the two features that were the most sensitive to Fault 1 (bi-modal GMMs are fit to each distribution and their contours are shown). One can see that the Chamber Temperature slightly increases during the fault as well as that the LF Load Power increases in magnitude. These statistical results make engineering sense. Namely, the long-term chamber degradation caused by successive PECVD tool depositions led to a change in the chamber impedance, making the RF matching network adjust to it. Residual depositions on the chamber and showerhead reduced the reflected power, resulting in a higher load power output (the sum of reflected and load powers needs to be constant). Furthermore, the temperature of the chamber went up slightly due to the higher power delivery and reduced thermal emissivity of the chamber surfaces caused by residual films.

Fig. 7 displays the two features that were found by LDA to be the most sensitive to Fault 2. One can see that the Process Pressure and the TEOS flow appear to respond slower during this fault compared

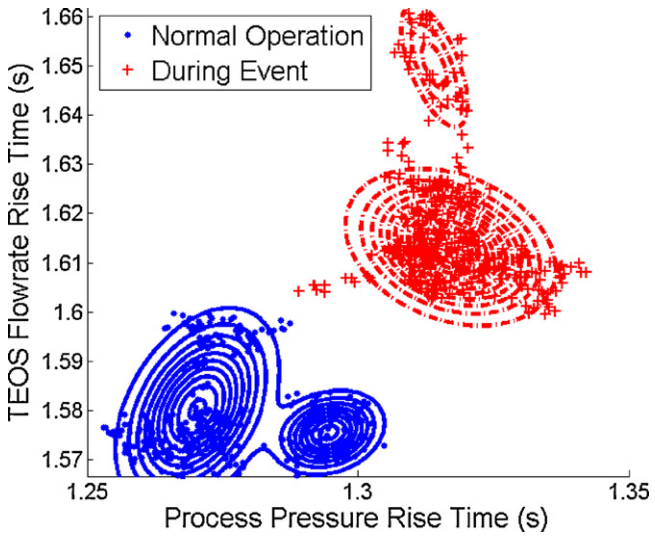

Fig. 7. Top sensitive features for Fault 2 along with normal operation, with GMM contours.

to normal operation. The slower transients could be caused by outgassing from the dirty chamber walls (residual films absorb gasses during depositions and then continuously release them into the chamber), or possible leaks in the chamber, both of which necessitate opening and comprehensive PM of the tool. Slower transients in the thermodynamics give rise to non-uniform thicknesses on the wafers due to the slower film growth at the beginning of the cycles. Furthermore, the slower dynamics in the thermal subsystem also bring about changes in the RF matching network variables during the transient stages. This causes variations in the load power during the initial stages of the depositions, further affecting the wafer thicknesses. All of these factors are contributing to the bi-modal wafer thicknesses recorded when this fault occurred.

Faults 3 and 4 involved particle failures on the wafers and their occurrence caused enormous downtime and losses. Fault 3 was detected when successive batches of wafers started failing metrology tests due to particle contamination appearing on the films. Particles can enter the chamber from the leaks in the chamber walls, leaks in the load-locks, from residual films flaking from the chamber walls and drifting into the plasma, or more rarely (but more detrimentally) from improper plasma formations. In the end, Fault 4 was identified as one of those "hard to diagnose" cases of particle failures due to improper plasma formations referred to as Coulomb crystals, where the ions inside the plasma are trapped in an electromagnetic field in a relatively "crystalline" structure [34] (once again, weeks of downtime and expert knowledge of a quantum physics Ph.D. were needed to diagnose this problem and finally remedy it).

As before, Fault 3 and Fault 4 were analyzed using LDA to identify features that were most indicative of those faults, after which GMM approximations were used to describe the behavior of the most sensitive features before and during the fault events. Faulty classes of data were formed using features emitted during depositions just before the problems were officially detected by the tool operators. As described in Section 3.2, LDA was then used to identify the features that changed the most significantly between the normal operating mode and those faulty operating modes. The top sensitive features for Faults 3 and 4 are shown in Table 4(a) and (b), respectively. Finally, GMMs were used to describe the PDFs of the most sensitive features during normal tool

Table 4
Results of the sensitivity analysis between normal settings and with Fault 3 (a) and with Fault 4 (b).

Principal vector (w)	Feature name
(a)	
0.47	HF Load Power Steady State Error
0.44	Pedestal Temperature Amplitude
0.37	LF Load Power Steady State Error
0.32	Process Pressure Rise Time
0.28	Pedestal Temperature Mean
0.26	Tune Capacitor Voltage Overshoot
0.26	LF Forward Power Tune Time
0.22	Load Capacitor Voltage Overshoot
0.21	HF Reflected Power Maximum
0.20	Load Capacitor Tune Time
(b)	
0.51	TEOS Flow Steady State Error
0.41	LF Reflected Power Tune Time
0.38	Process Pressure Rise Time
0.37	Process Pressure Minimum
0.35	Chamber Temperature Mean
0.29	TEOS Flow Rise Time
0.17	HF Load Power Tune Time
0.15	HF Reflected Power Maximum
0.13	Load Capacitor Tune Time
0.12	LF Forward Power Tune Time

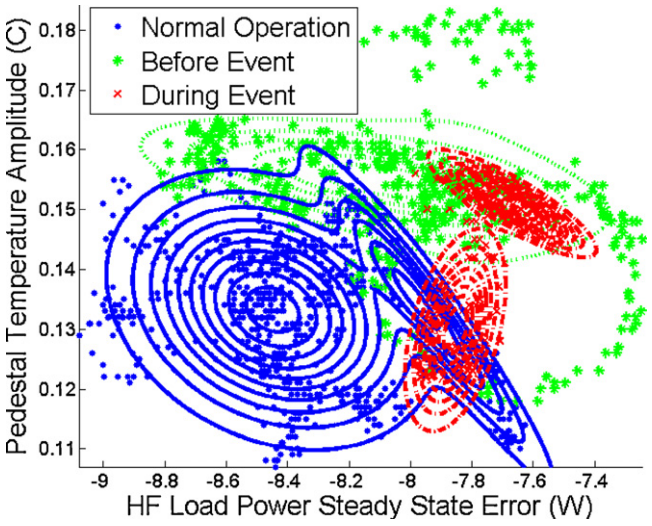

Fig. 8. Top sensitive features for Fault 3 along with normal operation, with GMM contours.

behavior, just before a fault was officially detected, and during the faults themselves.

Fig. 8 displays the two features that were found by LDA to be the most sensitive to Fault 3. The feature PDFs are shown for normal operation, just before the fault, and during the fault to see the progression of the sensitive features throughout the event. One can see that the variance of the Pedestal Temperature increases and its PDF becomes clearly bi-modal after the fault. In addition, we also see that the HF Load Power increases in magnitude throughout the fault. The temperature changes could be directly attributed to particle formation and their landing and moving on the wafer and the tool surfaces, which keeps changing their thermal characteristics and emissivity. The increase in the HF Load Power can be explained by the fact that particles floating in the plasma and landing and moving on the wafers and chamber walls keep changing the overall chamber impedance, which necessitates reactions from the RF matching network. There is some overlap between normal operation and during the event. However, one would detect large drops in the CV when moving through the event, therefore making this fault detectable by the top two sensitive features given from LDA as with the other events.

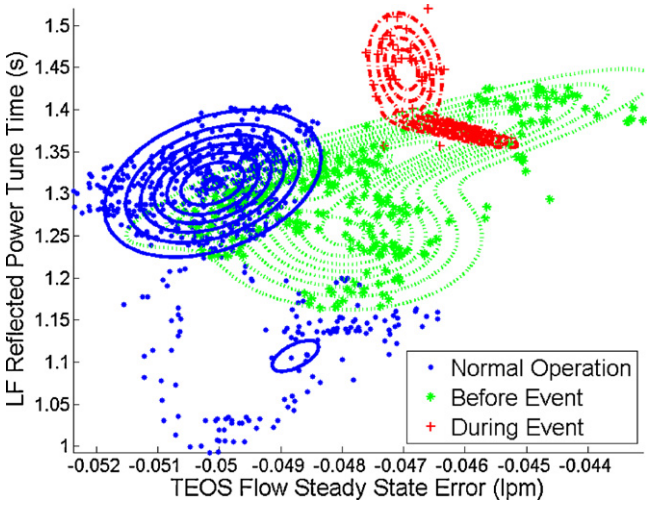

Fig. 9. Top sensitive features for Fault 4 along with normal operation, with GMM contours.

Fig. 9 displays the two features that were found by LDA to be the most sensitive to Fault 4. Once again, in order to see the progression of the sensitive features throughout the event, the feature PDFs are shown for normal operation, just before the fault, and during the fault. One can see that the steady state error in the TEOS flow increases throughout the fault, as well as that the LF reflected power tunes slower compared to normal operation. In addition, we can see that a predominantly Gaussian distribution of the most sensitive features drastically changes just before and during the faults, deteriorating into a bi-modal distribution that characterizes the faulty behavior. Once again, statistical results emanating from the LDA are supported by the engineering and physical knowledge about the faults that occurred. The particles and plasma formations are visible evidence that the RF power is generating abnormal, non-uniform plasma, which required longer tuning time of the matching network. Furthermore, improper evaporation of TEOS (which ultimately caused the Coulomb crystal formation) was the reason for the increasing trends in the steady state errors of the TEOS flow.

5. Discussion

In summary, we created statistical models of the four faults that occurred in our dataset using PDFs of the sensor features that were found by LDA to be the most sensitive to those faults. As described in Section 3.4 and illustrated in Fig. 4, the tool can be monitored by comparing the current behavior of the sensor features with PDF models of each of the four faults we observed until now. When a similar fault occurs again (which unfortunately did not happen in the dataset used in this study), a diagnosis can be performed to determine the root cause of the fault. If a previously unseen fault occurs, LDA can again be used to identify the most sensitive features, after which that new fault can be modeled using the PDF of those most sensitive features. Faults can also occur simultaneously, which can make it very difficult to decouple the faults to distinguish which components are giving problems. Due to the nonlinear nature of real dynamic systems, if mutually exclusive subsystems behave abnormally for particular faults, when the faults occur simultaneously, one cannot assume each subsystem will behave independently as if they were decoupled. Therefore, in the context of this work, simultaneous faults will be treated as unseen faults and LDA will identify a new set of sensitive features to monitor the fault (which may include features that are not inside the sensitive sets for each individual fault). When maintenance occurs for the unseen fault, one could then investigate the fault modes to see if it is a new fault or to see if one can decouple particular simultaneous fault combinations for the system at hand. Through this process, the database of known faults will grow and reactions to the current faults in operation (i.e. reactions to events for which feature patterns show high overlap with PDF model of an existing fault) can be much faster.

6. Conclusions and future work

This paper presents an integrated feature extraction, equipment monitoring, and fault modeling approach applied to in-place sensors from a commonly used industrial PECVD tool. The Linear Discriminant Analysis (LDA) and Gaussian Mixture Models (GMMs) of dynamic features extracted from numerous in-place sensors on the tool were used to recognize and track features that show the most prominent statistical changes due to tool degradation and maintenance events. The concept of performance confidence values (CVs) was used as a condition monitoring indicator expressing the overlaps of probability density functions (PDFs) of features that are the most sensitive to tool degradation

and faults. The CV combines both mean and variance changes in the feature distributions, transforming the changes in the multi-dimensional feature set into an easy-to-interpret number between zero and one. The sensitive features and CVs showed agreement with predictions made based on physical knowledge about the tool and the process. Additionally, four events of unacceptable equipment behavior were analyzed using the same methodology. Models of feature behavior in the presence of those faults were postulated as PDFs of features that were emitted by the tool immediately before the faults were officially detected. Such models are crucial for recognition of similar behavior in the future (diagnosis of those faults), which should greatly reduce the repair time for those faults. Using the feature extraction, condition monitoring, and fault modeling techniques described in this work, further research can be done to move beyond the explorations of purely diagnostic degradation models considered in this paper, toward predictive degradation models that can be used for more proactive maintenance decision making on the PECVD (and perhaps other) tools.

Acknowledgment

This work was supported in part by the International SEMATECH Manufacturing Initiative (ISMI).

References

[1] http://www.sematech.org/meetings/archives/mfg/8194/8194.pdf.
[2] S. Fulton, M. Kim, ISMI Consensus Preventive and Predictive Maintenance Vision Ver. 1.1, International SEMATECH Manufacturing Initiative, 2007.
[3] A. Buchner, S. Anand, J. Hughes, Data mining in manufacturing environments: goals, techniques and applications, Studies in Informatics and Control 6 (4) (1997) 319–328.
[4] Y. Liu, P. Kumar, J. Zhang, D. Djurdjanovic, J. Ni, Predictive Modeling and Intelligent Maintenance Tools for High Yield Next Generation Fab, TECHCON, 2005.
[5] D. Djurdjanovic, Y. Liu, Survey of Predictive Maintenance Research and Industry Best Practice, University of Michigan, Ann Arbor, MI, 2006.
[6] D. Djurdjanovic, J. Lee, J. Ni, Watchdog agent – an infotronics based prognostics approach for product performance assessment and prediction, International Journal of Advanced Engineering Informatics, Special Issue on Intelligent Maintenance Systems 3–4 (17) (2003) 109–125.
[7] M. Lebold, K. Reichard, D. Boylan, Using DCOM in an open system architecture framework for machinery monitoring and diagnostics, in: IEEE Aerospace Conference, 2003, 1227–1235.
[8] J.C. Li, Signal processing in manufacturing monitoring, in: Condition Monitoring and Control for Intelligent Manufacturing, Springers, 2006, pp. 245–265.
[9] G. Franklin, A. Emami-Naeini, Feedback Control of Dynamic Systems, Pearson and Prentice Hall, NJ, 2006, pp. 115–121.
[10] B. Boashash, Time Frequency Signal Analysis and Processing: A Comprehensive Reference, Elsevier Ltd, 2003.
[11] R. Serway, J. Jewett, Physics for Scientists and Engineers, Brooks/Cole Publishers, 2010.
[12] J. Tang, D. Dornfeld, S.K. Pangrle, A. Dangca, In-process detection of microscratching during CMP using acoustic emission sensing technology, Journal of Electronic Materials 27 (1998) 1099–1103.
[13] C. Spanos, H. Guo, A. Miller, J. Levine-Parrill, Real time statistical process control using tool data, IEEE Transactions on Semiconductor Manufacturing 5 (1992) 308–318.
[14] S.M. Pandit, S.M. Wu, Time Series and System Analysis with Applications, John Wiley, 1983.
[15] D.C. Montgomery, Introduction to Statistical Quality Control, John Wiley, New York, 2001.
[16] R. Harris, A Primer of Multivariate Statistics, Academic Press, New York, 1975.
[17] R.J. Bunkofske, N.T. Pascoe, J.Z. Colt, M.W. Smit, Real-time process monitoring, in: Proceedings of the 1996 7th Annual IEEE/SEMI Advanced Semiconductor Manufacturing Conference, ASMC 96, November 12–14 1996, Cambridge, MA, USA, (1996), pp. 382–390.
[18] K. Mai, M. Tuckermann, SPC based in-line reticle monitoring on product wafers, in: IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop-Advancing Semiconductor Manufacturing Excellence, Munich, Germany, (2005), pp. 185–189.
[19] J.P. Card, M. Naimo, W. Ziminsky, Run-to-run process control of a plasma etch process with neural network modeling, Quality and Reliability Engineering International 14 (1998) 247–260.
[20] D. Pompier, C. Giuliani, N. Csejtei, E. Paire, J. Cholvy, Oxide deposition monitored through APC system, in: 7th AEC/APC Europe, Aix-en-Provence, 2006.
[21] S.A. Velichko, Multi-parameter model based advanced process control, advanced semiconductor manufacturing, in: IEEE Conference and Workshop, 2004, 443–447.
[22] K.H. Baek, Y. Jung, G.J. Min, C. Kang, H.K. Cho, J.T. Moon, Chamber maintenance and fault detection technique for a gate etch process via self-excited electron resonance spectroscopy, Journal of Vacuum Science; Technology B (Microelectronics and Nanometer Structures) 23 (2005) 125–129.
[23] Z. Chen, R. Li, M. Ito, J. Ding, A real time plasma monitoring and FDC method using OES, in: 7th AEC/APC Europe, Aix-en-Provence, 2006.
[24] M. Matsuda, T. Miyata, M. Shimada, T. Watanabe, Advanced process control for semiconductor thermal process, Hitachi Review 51 (2002).
[25] K. Salahshoor, A. Keshtgar, Statistical monitoring of process condition using neural network, in: 7th AEC/APC Europe, Aix-en-Provence, 2006.
[26] H.J. Tu, C.T. Chang, C.J. Chang, C.Y. Lee, M.F. Yoo, F. Ko, J. Wang, C.H. Yu, M.S. Liang, Simultaneous fault detection and classification of a 300 mm plasma CVD tool, in: 5th APC/AEC Europe, Hotel Hilton Dresden, Germany, 2004.
[27] J. Blue, A. Chen, Recipe-independent tool health indicator for predictive maintenance and tool qualification, in: 7th AEC/APC Europe, Aix-en-Provence, 2006.
[28] Geng, Hwaiyu, Semiconductor Manufacturing Handbook, McGraw-Hill Handbooks, NY, 2005.
[29] R. Duda, P. Hart, D. Stork, Pattern Classification, Wiley & Sons, NY, 2001, pp. 117–124.
[30] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge University Press, Cambridge, 2000.
[31] D. Lay, Linear Algebra and its Applications, Addison, Wesley, 2005.
[32] B. Lindsay, Mixture Models: Theory, Geometry and Applications, Institute of Mathematical Statistics, CA, 1995, pp. 1–6.
[33] J. Liu, Autonomous Anomaly Detection and Fault Diagnosis, PhD thesis, University of Michigan, 2007.
[34] J. Chu, I. Lin, Direct observation of coulomb crystals and liquids in strongly coupled RF dusty plasmas, Physical Review Letters 72 (25) (1994) 4009–4012.

Alexander Q. Bleakie was born in Houston, Texas, USA. He received a B.S. degree in Mechanical Engineering from Texas A&M University in 2008 with Summa Cum Laude Honors, and a M.S. degree in Mechanical Engineering from the University of Texas at Austin in 2010. He is currently pursuing a Ph.D. degree in Mechanical Engineering at the University of Texas at Austin with emphasis on Dynamic Systems and Control. His current research interests include dynamic modeling, control theory, fault diagnosis, and estimation.

Dragan Djurdjanovic received the B.S. degree in mechanical engineering and applied mathematics from the University of Nis, Nis, Serbia, in 1997, the M.Eng. degree in mechanical engineering from Nanyang Technological University, Singapore, in 1999, the M.S. degree in electrical engineering (systems), and the Ph.D. degree in mechanical engineering from the University of Michigan, Ann Arbor, in 2002. He is currently an Assistant Professor of mechanical engineering, and operations research and industrial engineering, with the Department of Mechanical Engineering, University of Texas, Austin. He has co-authored 37 published or accepted journal publications, 3 book chapters, and 30 conference publications. His current research interests include advanced quality control in multistage manufacturing systems, intelligent proactive maintenance techniques, and applications of advanced signal processing in biomedical engineering. Dr. Djurdjanovic was the recipient of several prizes and awards, including the 2006 Outstanding Young Manufacturing Engineer Award from the Society of Manufacturing Engineers, the 2005 Teaching Incentive Award from the Department of Mechanical Engineering, University of Michigan, Nomination for the Distinguished Ph.D. Thesis from the Department of Mechanical Engineering, University of Michigan in 2003, and the Outstanding Paper Award at the 2001 SME North American Manufacturing Research Conference.