J Intell Manuf (2019) 30:1099–1110
https://doi.org/10.1007/s10845-017-1308-4

Nonstationary signal analysis and support vector machine based classification for vibration based characterization and monitoring of slit valves in semiconductor manufacturing

M. Musselman[1] · H. Xie[2] · D. Djurdjanovic[2]

Received: 22 July 2016 / Accepted: 7 February 2017 / Published online: 22 February 2017

Abstract Slit valves play an important role in semiconductor manufacturing, enabling creation and maintaining of a vacuum environment required for wafer processing. Due to the high volume of production in the modern semiconductor industry, slit valves could experience severe degradation over their lifetime. If maintenance is not applied in due time, degraded valves may lead to defects in finished products due to pressure loss and particle generation. In this paper, we propose methods for signal processing and feature extraction for analysis of slit valve vibration signals. These methods are then used to demonstrate the ability to reliably, accurately and efficiently distinguish between vibration patterns of each individual valve via a multi-class classification procedure. Furthermore, instantaneous time–frequency entropy of valve vibrations enabled long term monitoring of a slit valve in production, in spite of variations in valve speed and operations.

Keywords Slit valves · Semiconductor manufacturing · Vibrations based monitoring · Nonstationary signal analysis · Multi-class classification

Introduction

Slit valves play an important role in semiconductor manufacturing industry. A slit valve is a gate that separates the process chamber and the transfer chamber of a semiconductor manufacturing tool, enabling creation of a vacuum environment required for wafer processing. Due to the high volume of production, as well as high temperatures and corrosive gases encountered in the modern semiconductor industry, slit valves could experience severe degradation over their useful lifetime. Once the valve is degraded, vacuum condition inside the process chamber cannot be maintained, resulting in wafer defects. In addition, degraded slit valves may also lead to particle generation via the erosion of degraded valve seals and guide ways, causing contamination in both the chamber and finished products. Considering the very small margin of error in today's microelectronic manufacturing, there is an urgent need for establishing a monitoring and maintenance plan to set early alarms about degraded slit valves and prevent potential product flaws.

In majority of today's semiconductor fabrication plants (fabs), preventive maintenance of equipment is conducted following the reliability based maintenance (RBM) paradigm, i.e. based on the elapsed calendar time or usage and the statistical properties of the useful life distribution of the relevant population of machines (Fulton and Kim 2007; Cholettte et al. 2013). Differences between individual machines in a population based on which RBM policies are postulated cause RBM to incur losses due to unnecessary maintenance of equipment that does not really need to be maintained, or due to unexpected failures of machines whose scheduled maintenance did not occur soon enough.

An alternative paradigm to address this drawback is the condition based maintenance (CBM), in which one builds and uses a connection between the condition of the individual piece of equipment and sensor reading emitted by that machine. With such information, maintenance operations can be performed according to the actual working condition of the equipment, exactly when needed and exactly where needed.

✉ D. Djurdjanovic
dragand@me.utexas.edu

[1] Lam Research Corporation, Fremont, CA, USA

[2] Department of Mechanical Engineering, University of Texas, Austin, TX, USA

In the past few years, CBM has drawn increasing attention in the semiconductor manufacturing industry. Advanced condition diagnostics and prognostics methods have been employed for various equipment and processes, such as etching equipment (Shadmehr et al. 1992; Kim and May 1997; Hong et al. 2012), lithography processes (Facco et al. 2009; Bao and Spanos 2001; Shen et al. 2011), chemical vapor deposition (Raoux et al. 1998; Hopfe et al. 2003; Wu et al. 2003), chemical-mechanical planarization (Tang et al. 1998; Lee et al. 2006) and material handling devices (Jong and Lin 2007; Guan et al. 2011).

As for the problem of condition monitoring of valve performance, vibration-based methods found wide application for various industries. Lee et al. (2010) accomplished degradation monitoring in a 6-in check valve in a nuclear power plant by monitoring the dominant frequencies of the valve's vibration signals under different temperature conditions. Yang et al. (2005) developed a condition monitoring scheme using statistical features extracted from vibration signals together, based on which a support vector machine was used to detect the cavitation faults of a butterfly valves in pumping stations. Wang et al. (2009) were the earliest authors to employ nonstationary signal analysis methods for valve condition monitoring. They employed non-negative matrix factorization and a neural network ensemble to recognize the diesel valve train fault patterns present in the time–frequency distributions of valve vibrations. Lin et al. (2013) propose a novel health condition classification method for a reciprocating compressor using appropriate partitioning of the time–frequency plane to enhance the characterization of valve vibration signals. More recently, Pichler et al. (2016) presented an effective approach for detecting cracked or broken reciprocating compressor valves under varying load conditions by looking at vibration spectrogram difference and autocorrelation. Similarly, Flett and Bone (2016) implemented an impact strength feature from vibration signals to investigate diesel engine valve trains with spring faults and clearance faults.

Unfortunately, CBM technology has never been applied to monitoring of slit valves, in spite of their great importance and widespread use in semiconductor manufacturing. Some reasons explaining this may be the lack of appropriate sensing of relevant signals from slit valves on typical tools (valve velocities, accelerations, vibrations), as well as the high complexity and non-stationarity of those signals.

In the research presented in this paper, we sensorized 50 slit valves in a major domestic semiconductor fab with 3-dimensional accelerometers and developed advanced signal processing and feature extraction methods for analysis of their vibration signatures. These methods are then used to reliably, accurately and efficiently recognize each individual valve via its vibration signature and a novel multi-class classification procedure. Furthermore, vibration signatures from one of those valves were collected over multiple months of its operation and gradual changes in those vibration patterns were observed, indicating its degradation. This early warning information about valve degradation can be augmented with the temporal information as to where along the path of that valve changes in its vibration patterns occurred, since different parts of valve travel can be related to specific component or components of the valve system, thus pinpointing what repair/maintenance operation may be needed. In summary, the research presented in this paper is aimed at illustrating the feasibility of continuous vibration-based monitoring of slit valves on semiconductor manufacturing tools during their regular operation in industrial setting.

This novel condition monitoring application required a unique combination of signal processing, classification and condition monitoring methods, which will be described in the remainder of this paper. To that end, in "Methodology" section, methods for processing of slit valve vibration signals and multi-segment classification of valves based on their signatures are described in detail. Furthermore, condition monitoring method based on the entropy of time–frequency distributions of valve vibrations will also be introduced in "Methodology" section. "Results" section gives the results of applying these methods to identification of 50 individual slit valves operating in a major semiconductor fab, as well as vibration based monitoring of one slit valve over several months of its operation, between two overhaul maintenance operations done on that valve. Finally, conclusions and potential future work are presented in "Conclusions and future work" section.

Methodology

Signal processing and feature extraction

Due to the non-linearity and complexity of valve motions, the vibration signals emitted by slit valves are highly non-stationary. This means that frequency contents of those signals vary significantly over time, which invokes the need for non-stationary signal analysis tools, such as Cohen's class time–frequency transform method (Cohen 1995). Cohen's general class of time–frequency distribution (TFD) for the signal $x(t)$ can be described as

$$C_x(t,\omega) = \frac{1}{4\pi^2}\int\int\int \varphi(\theta,\tau)\, x\left(u+\frac{\tau}{2}\right) x^*\left(u-\frac{\tau}{2}\right) e^{-j(\theta t+\tau w-\theta u)} d\theta d\tau du \quad (1)$$

where $x(t)$ and $x^*(t)$ denote the relevant signal and its complex conjugate respectively, while $\varphi(\theta,\tau)$ is the so-called kernel function of the TFD. The kernel determines mathematical properties of the resulting TFD, such as realness of the

resulting distribution C_x (t, ω), time and frequency support properties, upholding of the time and frequency marginals, cross-term suppression, as well as group delay and instantaneous frequency properties[1] (Jeong and Williams 1992). Fulfillment of these properties enables interpretation of the function $C_x(t, \omega)$ defined by (1) as a joint 2-dimensional distribution of signal energy in time and frequency domains. In this paper, we opted to use the binomial kernel, which can be considered to be the most advanced signal-independent reduced interference distribution (RID) kernel (Papandreou-Suppappola 2002). Its RID character and consequent ability to suppress cross terms, which are inherently present in TFDs, is highly desirable, since it is well documented that cross terms can hamper signal interpretation and classification based on TFDs riddled with cross terms (Cohen 1995). On the other hand, signal independent nature of the binomial kernel is also important because of the sheer volume of data considered in this study. Namely, signal dependent kernels, such as those introduced by Jones and Baraniuk (1995) or Coates and Fitzgerald (1999), would be computationally infeasible in the realm of multiple vibration readings from a large number of valves collected at sampling rates in the kHz ranges, which is what we dealt with in this research.

After signal processing stage, numerous features were extracted from the resulting vibration TFDs and later used as inputs for the multi-class classifier built to recognize individual valves based on their vibration signatures. These features can be partitioned into three categories: timing based features, time domain based features and time–frequency distribution based features.

Timing based features consist of times required to complete various portions of the valve movement. In this paper, for each valve opening or closing cycle, we just recorded the time interval between two fixed points along the valve motion path. Of course, a much more elaborate set of timing features describing the valve motion in finer detail could be pursued, if adequate discrete control signals are available to reliably delineate these portions of the valve motions. Such features are a good indication of machine working condition, since the time required to accomplish various portions of the designated motion for a machine with moving parts will usually vary as the condition of that machine drifts.

Time-domain based features are calculated from the time domain waveforms of the signals and include signal entropy, mean signal energy, median energy of the signal, as well as variance, skewness and kurtosis of the signal energy. Table 1 lists all time-domain based features used in this study and the formulae according to which they are calculated. These time-domain based features are intuitive and extensively used as vibration based measurements of the working condition in many previous machine monitoring works (Heng and Nor 1998; Saxena and Saad 2007).

Table 1 Time domain features

Feature	Formula
Signal energy	$SumE = \sum_{t=t_1}^{t_n} x(t)^2$
Signal entropy	$H = \sum_{t=t_1}^{t_n} -\frac{x(t)^2}{SumE} log \frac{x(t)^2}{SumE}$
Entropy of signal energy	$HE = \sum_{t=t_1}^{t_n} -\frac{x(t)^4}{\sum_{t=t_1}^{t_n} x(t)^4} \log \frac{x(t)^4}{\sum_{t=t_1}^{t_n} x(t)^4}$
Maximal energy	$max\ x(t)^2$
Time of maximal energy	$argmax_t\ x(t)^2$
Minimal energy	$min\ x(t)^2$
Time of minimal energy	$argmin_t\ x(t)^2$
Maximal amplitude	$max\ x\ (t)$
Time of maximal amplitude	$argmax_t\ x\ (t)$
Minimal amplitude	$min\ x(t)^2$
Time of minimal amplitude	$argmin_t\ x(t)^2$
Median energy	Middle value of $x(t)^2$
Mean energy	$E\left[x^2\right] = \frac{1}{n} \sum_{t=t_1}^{t_n} x\ (t)^2$
Variance of energy	$Var\left[x^2\right] = \frac{1}{n} \sum_{t=t_1}^{t_n} (x(t)^2 - E\left[x^2\right])^2$
Skewness of energy	$Skewness = \frac{\frac{1}{n} \sum_{t=t_1}^{t_n} (x(t)^2 - E\left[x^2\right])^3}{(Var\left[x^2\right])^{\frac{3}{2}}}$
Kurtosis of energy	$Kurtosis = \frac{\frac{1}{n} \sum_{t=t_1}^{t_n} (x(t)^2 - E\left[x^2\right])^4}{(Var\left[x^2\right])^2}$

Time–frequency distribution based features used in this study consisted of the so-called time–frequency distribution moments, entropy and several signal energy related features calculated from the binomial distributions of the slit valve vibrations. Following Cohen (1995), the moment terms $E\left[t^p \omega^q\right]$ can be calculated as

$$E\left[t^p \omega^q\right] = \sum_{t=t_1}^{t_n} \sum_{\omega=\omega_1}^{\omega_m} t^p \omega^q C_x\ (t, \omega) \tag{2}$$

where C_x (t, ω) is the time–frequency distribution of the signal $x(t)$. Since moments of low orders can be used to approximate the general characteristics of TFDs and successfully accomplish clarification based on those TFDs (Cohen 1995; Djurdjanovic et al. 2002). Moments up to 3 were used in this study to describe the time–frequency patterns observed in the slit valve vibrations.

Besides time–frequency moments, for each signal, entropy based on its binomial TFD $C_x(t, \omega)$ was calculated as

[1] Definitions of those properties, as well as mathematical constraints on the kernels that are necessary to achieve them are summarized in the seminal book by Cohen (1995).

Table 2 Time–frequency domain features

Feature	Formula
TFD moments of order up to 3	$\mathrm{E}\left[\mathrm{t}^{\mathrm{p}}\omega^{\mathrm{q}}\right]$ with p+q≤3
Entropy	$H = \sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m} -\frac{C_x(t,\omega)}{\sum_{\tilde{t}=t_1}^{t_n}\sum_{\tilde{\omega}=\omega_1}^{\omega_m} C_x(\tilde{t},\tilde{\omega})} log \frac{C_x(t,\omega)}{\sum_{\tilde{t}=t_1}^{t_n}\sum_{\tilde{\omega}=\omega_1}^{\omega_m} C_x(\tilde{t},\tilde{\omega})}$
Maximal energy	$max\ C_x\ (t,\omega)$
Time of maximal energy	$argmax_{\mathrm{t}}\ C_x\ (t,\omega)$
Frequency of maximal energy	$arg\max_{\omega}\ \mathrm{C_x}\ (\mathrm{t},\omega)$
Minimal energy	$\min\ \mathrm{C_x}\ (\mathrm{t},\omega)$
Time of minimal energy	$arg\min_{\mathrm{t}}\ \mathrm{C_x}\ (\mathrm{t},\omega)$
Frequency of minimal energy	$arg\min_{\omega}\ \mathrm{C_x}\ (\mathrm{t},\omega)$
Median energy	Middle value of $\mathrm{C_x}\ (\mathrm{t},\omega)$
Mean energy	$E\left[C_x\ (t,\omega)\right] = \frac{1}{nm}\sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m} C_x\ (t,\omega)$
Variance of energy	$Var\left[C_x\ (t,\omega)\right] = \frac{1}{nm}\sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m}(C_x\ (t,\omega) - E\left[C_x\ (t,\omega)\right])^2$
Skewness of energy	$Skewness = \frac{\frac{1}{nm}\sum_{t=t_1}^{t_n}\sum_{w=w_1}^{w_n}(C_x(t,w)-E[C_x(t,w)])^3}{(Var[C_x(t,w)])^{3/2}}$
Kurtosis of energy	$Kurtosis = \frac{\frac{1}{nm}\sum_{t=t_1}^{t_n}\sum_{w=w_1}^{w_n}(C_x(t,w)-E[C_x(t,w)])^4}{(Var[C_x(t,w)])^2}$

$$H = \sum_{t=t_1}^{t_n}\sum_{\omega=\omega_1}^{\omega_m} -\frac{C_x\ (t,\omega)}{\sum_{\tilde{t}=t_1}^{t_n}\sum_{\tilde{\omega}=\omega_1}^{\omega_m} C_x\ (\tilde{t},\tilde{\omega})} log\frac{C_x\ (t,\omega)}{\sum_{\tilde{t}=t_1}^{t_n}\sum_{\tilde{\omega}=\omega_1}^{\omega_m} C_x\ (\tilde{t},\tilde{\omega})} \quad (3)$$

In addition, various signal energy related features were also extracted from their binomial TFDs, including maximal energy, as well as the time instance and frequency at which that maximal energy appeared in the TFD. A complete list of time–frequency domain features used in this paper is summarized in Table 2. Within this plethora of signal features, not all will provide useful information for valve recognition. To the contrary, the so-called curse of dimensionality plagues the performance of classifiers based on such highly dimensional feature space and a feature reduction process is needed to improve the classification process by removing the redundant and irrelevant features (Kittler 1975). The feature reduction procedure employed in this study is closely associated with the classification strategy and is therefore embedded in the description of the classification framework given below.

Identification of individual valves via classification of vibration patterns

Individual valve differentiation via their vibration signatures is a multi-class classification problem. One straightforward method dealing with this type of problems is to train a universal multi-class classifier that could take care of all of the classes simultaneously. In the recent literature, we see such approaches in Wuxing et al. (2004), as well as in Mahamad and Hiyama (2011). However, this strategy requires that the information needed to separate classes 1 and 2 also be suitable for differentiating classes 2 and 3 and all other possible class-pairs. Unfortunately, this assumption turns out to be overly constraining, especially when the number of classes involved becomes very large, as it is in the case considered in this paper (we are dealing with 50 valves and hence must realize a multi-class classification problem involving 50 classes).

In this study we adopted a classification approach introduced by Kreßel (1999) and recently utilized by Musselman and Djurdjanovic (2012) for classification of electroencephalogram (EEG) signals based on their time–frequency signatures. The method is based on repeated pairwise classifications that successively distinguish between all possible class-pairs in a multi-class classification problem. This method increases classification accuracy by enabling pairwise distinction between any given pair of classes, using a feature set specifically selected to optimize that particular classification problem. Thus, this approach uses a divide-and-conquer paradigm and a variable, customized feature set, rather than utilizing one universal feature set to tackle the entire multi-class classification.

More specifically, this method decomposes an n-class problem into $\binom{n}{2}$ one-against-one pairwise classification sub-problems, with a specific classifier being trained for each one of those sub-problems using the most discerning features for that sub-problem. When a query signal is to be classified, it is passed through all the pairwise classifiers, each providing a vote for one of the two classes involved. Eventually, all the outputs of these $\binom{n}{2}$ subclassifiers are aggregated and the query signal is assigned to the class receiving the most votes. Fig. 1 pictorially illustrates this entire process.

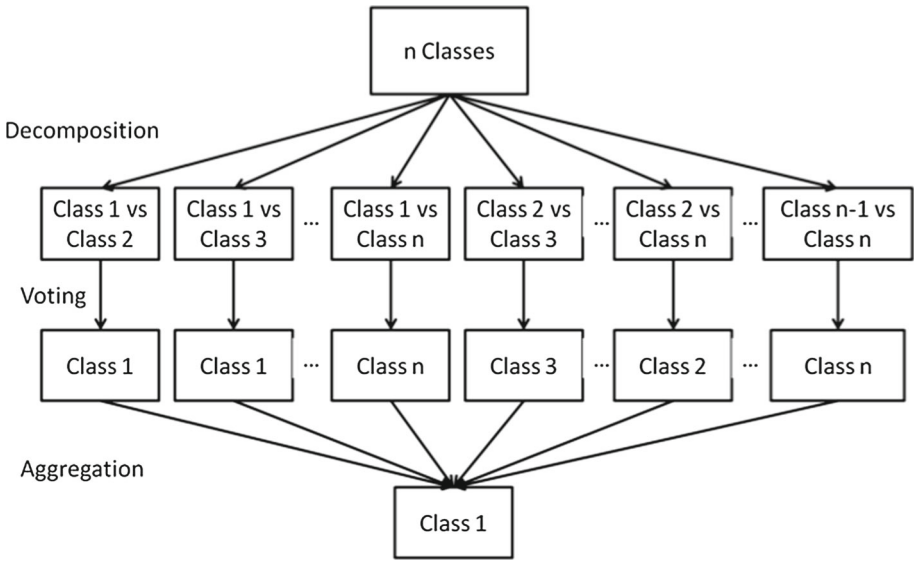

Fig. 1 Pairwise classification process

For each of the pairwise classification problems, the features relevant for that problem were selected from the exhaustive feature set described earlier in this section, using an inter/intra class distance ratio in the feature space, similar with the Index introduced by Dunn (1973). Namely, given an n-class problem and an exhaustive feature set consisting of m features, the selection process is conducted as follows:

- For each possible pair of classes $(\nu_i, \nu_j), i \in \{1, 2, \ldots, n\}, j \in \{1, 2, \ldots, n\}, i \neq j$, and each feature $l, l \in \{1, 2, \ldots, m\}$, calculate the maximum intra-class Mahalanobis distance[2] $d\left(\nu_i, \nu_j, l\right)$ between all points in classes ν_i and ν_j, as well as the minimum interclass Mahalanobis distance[3] $D\left(\nu_i, \nu_j, l\right)$ between points in classes ν_i and ν_j. Then, the inter/intra class distance ratio is obtained as

$$\mathrm{r}\left(\nu_i, \nu_j, l\right) = \frac{D\left(\nu_i, \nu_j, l\right)}{d\left(\nu_i, \nu_j, l\right)} \tag{4}$$

- For a specific pair of classes $(\nu_i, \nu_j), i \in \{1, 2, \ldots, n\}, j \in \{1, 2, \ldots, n\}, i \neq j$, features with the highest inter/intra class distance ratios are selected as the feature for this class-pair (ν_i, ν_j). In this way, each pair-wise classification sub-problem is addressed via features that have high intra-class localization and inter-class separation.

Another key point in a classification problem is the selection of the classification algorithm. In this study, the k-Nearest Neighbor (kNN) classifier is chosen to discriminate between different valves. kNN is a non-parametric classification algorithm that determines the class memberships of an unknown testing point according to the k closest training points in the feature space (Duda et al. 2000). Because of its simplicity, kNN is a highly suitable classification algorithm for a classification strategy based on numerous pairwise classification problems, such as the one encountered here.

Long-term monitoring of valve vibration patterns

During long-term operation of a slit valve in a semiconductor manufacturing fab, significant variability in the timing and speed and speed of its operation could be observed due to normal changes in operating regimes (recipes). These changes affect vibration patterns of the valve, but should not be seen as a sign of degradation and should not lead to alarms. Our examinations have shown that the instantaneous entropy of the vibration signal, evaluated from its time–frequency distribution at any given time t as

$$H_t = \sum_{\omega=\omega_1}^{\omega_m} -\frac{C_x\left(t, \omega\right)}{\sum_{\omega=\omega_1}^{\omega_m} C_x\left(t, \omega\right)} \ln\left(\frac{C_x\left(t, \omega\right)}{\sum_{\omega=\omega_1}^{\omega_m} C_x\left(t, \omega\right)}\right) \tag{5}$$

is immune to the aforementioned operating regime-induced changes in the vibration signal, while still depicting the changes that occur as the valve deteriorates over time. The reason for this is that a well-maintained valve shows very little vibration (it is quiet) and as it degrades, at specific

[2] Observe all Mahalanobis distances between training vectors in class ν_i and all Mahalanobis distances among training vectors in class ν_j. Then $d(\nu_i, \nu_j, l)$ is the *maximum of those distances*. In a way, this is a measure of intra-class localization feature l provides for classes ν_i and ν_j.

[3] Observe Mahalanobis distances from any training vector in class ν_i to any training vector in class ν_j. Then $d(\nu_i, \nu_j, l)$ is the *minimum of those distances*. In a way, this is a measure of inter-class separation feature l provides for classes ν_i and ν_j.

time-instances, "disorder" in the vibration patterns appears and grows, which is readily captured by the instantaneous entropy (5). Information as to where exactly in the travel of the valve this "disorder" appears is of high importance, since it can be used to localize the source of this degradation and eliminate it through maintenance.

Results

In this section, the signal processing, feature extraction, valve classification and vibration monitoring methods described in the previous section will be applied to vibration monitoring of slit valves in a major domestic semiconductor manufacturing facility.

All valves considered in this paper were pneumatic valves of identical design, produced by the same manufacturer (i.e. nominally, they are supposed to be identical). When the valve closes, a pneumatic cylinder drives the valve plate down guide rails. Near the bottom of the valve travel, the valve head encounters a cam, which directs the valve head motion from downward to forward. At the end of the valve motion, the valve head makes contact with a base plate of the chamber to create the seal. The motion is reversed when the valve opens and is schematically illustrated in Fig. 2.

The data acquisition system was based on the sbRIO-9636 embedded control and acquisition device from National Instruments Corporation (2015). The valve vibrations were captured using 3-dimensional (3D) accelerometers ADXL327 (Analog Devices 2015) mounted on the valve housing and their readings were sampled at 5 kHz. The results in this paper are based on the root mean square (RMS) of the 3D vibration signals provided by this sensor, as well as the original 3D signals, since specific vibration directions can be used to provide characterization and monitoring of specific portions of the valve motion (descent down/ascent up the rail, camming, valve closing/opening).

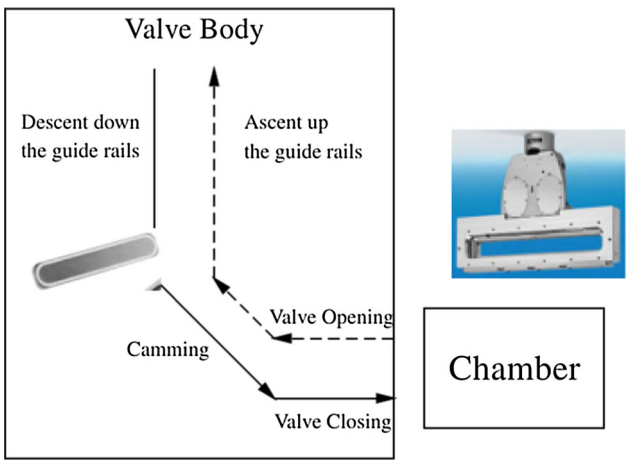

Fig. 2 Illustration of valve operation and motion during opening (*dashed line*) and closing (*solid line*)

In addition, since control signals from the valve were not available, the timing information about the valve motion was obtained using two photo-resistors placed over status lights on the valve housing. One of those lights indicated valve body passing by a fixed point near the top of the guide rail, while the other indicated the valve body passing a fixed point near the end of the valve movement, where the valve seal makes contact with the chamber wall. Signals from the photo-resistors were used as automatic markers for valve motions and for normalization of the valve travel time, which reduced variability of valve signatures caused by operating-mode (recipe) induced variations in the valve travel times. Normalization was accomplished as follows. During valve closing motion, turning on of the light near the top of valve motion denoted normalized time 0, while turning on of the light near the bottom of the valve motion signified normalized time 1. Conversely, during the valve opening motion, normalized time 0 occurred when status light near the bottom of the valve motion turned off, while normalized time 1 occurred when the status light at the top of the valve motion path turned off. One should note that valve movement pre and post these light indicators were also collected, resulting in valve vibrations for normalized times −0.2 to 1.2. Each vibration signal was divided into three stages in the following manner. During valve closing, the three segments were valve motion before the valve-up signal (normalized time −0.2 to 0), valve motion between the valve up and valve down signals (normalized time 0 to 1) and valve motion after the valve down signal (normalized time 1 to 1.2). Conversely, during valve opening, we observed valve motion before the valve-down signal turned off (normalized time −0.2 to 0), valve motion between the valve down and valve up signals turned off (normalized time 0 to 1) and valve motion after the valve up signal turned off (normalized time 1 to 1.2). Accordingly, features described in the previous section were extracted using signal portions from these three stages separately or from the whole period (normalized times −0.2 to 1.2), yielding a rich feature library consisting of 1122 features.[4]

Identification of individual valves

From 50 individual valves, vibration signals corresponding to 10 openings and 10 closings were collected, yielding the total of 1000 signals. From each of those signals, timing based, time domain and time–frequency domain characteristics

[4] From each of the 4 segments, we got 16 time-domain based features and 19 time–frequency domain based features, for each of the three directions of vibrations, as well as the vibration RMS. In addition, the feature set also included the movement times of both closing and opening motions, yielding 1122 features.

Table 3 Pairwise classification results

	Timing features (%)	Time domain features (%)	TFD features (%)	Fusion of timing, time domain, TFD features (%)
Opening	22.88	96.88	96.99	98.18
Closing	27.54	96.64	96.35	97.93
Opening and closing	60.58	97.15	98.73	98.74

described in "Methodology" section were extracted, based on which this 50-class classification study was conducted.[5] The training set was constructed by randomly picking half of the recordings from each valve, while the remaining half was used for testing. Such selection of training and testing sets was repeated 100 times to objectively evaluate the performance of the proposed classification method, independently of the choice of the training set.

Classification accuracies based on different feature sets and movement directions of the valves are listed in Table 3. It is obvious that the use of time-domain features and time–frequency domain features greatly outperformed the accuracy of classification based on the timing features alone (performance of the timing based features is actually quite poor). In addition, it is visible that the elaborate feature extraction and classification methods introduced in this paper enabled improved classification results via the fusion of features from various domains and valve motion stages, yielding the best performance when features from all domains and all motion stages are included. In that case, average accuracy of 98.74% was reached over the 100 tests (100 randomly selected sets of training and testing data), with perfect results obtained in 7 out of those 100 tests. Furthermore, the consistency of this method in differentiating individual slit valves was evident in the low variance of classification accuracies observed over all 100 tests (7.58×10^{-5}% of accuracies observed over the 100 random training/testing sets). From these results, one can conclude that vibration patterns of slit valves are so distinctive that they identify individual valves, similarly to how human's speech can be used to identify an individual person. Advanced time–frequency analysis and sophisticated feature extraction methods introduced in this paper were able to expose those discerning vibration patterns and enable almost perfect valve identification.

Delving deeper into how this remarkable performance was achieved, we can notice that many of the 1122 features generated from feature extraction stage were never selected or were only rarely selected for pairwise classification subproblems, while a few happened to be selected and used more frequently. For the case of jointly using opening and closing valve motions, and features from all domains (highest classification accuracy, as per in Table 3), top 10 most frequently used features are identified in the pareto-chart shown in Fig 3, and are explained in Table 4. As can be seen in Fig. 3, median energy in the time–frequency domain of the RMS of the entire vibration signal is the feature used most commonly. It was used in 800 out of 1225, or more than 68% of the pairwise classification sub-problems. Mean energy in the time–frequency domain calculated for the RMS of the vibration signal between the valve up and valve down signals ranks second in the chart and covers 100 out of 1225 pairwise subproblems, which is about 8.16% of sub-problems. Besides these two features, none of the remaining ones was used in more than 3.1% classification sub-problem, which implies that the aforementioned two features can be considered to be the most discerning features in differentiating the valves from each other.

In addition, the 10 most utilized features listed in Fig. 3 and Table 4 happen to be selected for use in around 90% of all the pairwise classification sub-problems, with 51 features covering all 1225 sub-problem pairs. It is an indication that valve identification information could be stored in 51 or even fewer vibration based metrics instead of the whole vibration signature from three channels, implying that efficient valve monitoring could be accomplished without excessive data storage requirements.

Long-term monitoring of an individual valve

Vibrations from one of the slit valves in the fab were collected over the period of several months, starting with an overhaul maintenance of that valve (consisting of lubrication of the valve body and placement of a fresh seal) and ending with its next overhaul maintenance. As mentioned in "Long-term monitoring of valve vibration patterns" section, due to significant variabilities in timing and speed due to normal changes in operations (recipes) executed on the tool, instantaneous time–frequency entropy of valve vibrations was used for its monitoring and Fig 4. shows the progression of those entropies in the valve closing motion, as the valve degraded during the monitoring period. Horizontal axis in this figure corresponds to the time, or rather the increasing number of cycles this valve executed, while the vertical axis denotes the

[5] One should note that our classification method transformed this 50-class classification into 1225 pairwise classification problems, which were solved using the kNN classification algorithm.

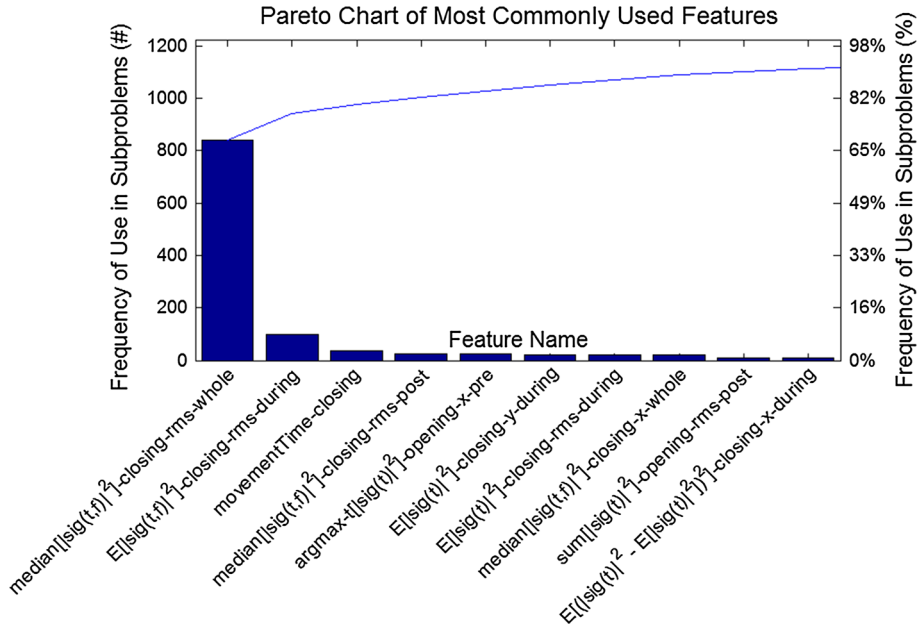

Fig. 3 Pareto Chart of most commonly used features in the classification scheme that led to 98.74% accuracy (highest accuracy achieved in this study)

position of the valve within any given motion cycle. Gradual increase in the instantaneous time–frequency entropies is evident in several sections of the valve travel and those portions are emphasized in the figure (encircled with dashed lines). This change in entropies denotes increased "disorder" in the time–frequency distributions of valve vibrations corresponding to those portions of its travel (notably, most pronounced changes were observed toward the end of the valve closing travel, indicating deterioration in the portion of valve motion corresponding to the seal making contact with the chamber door just before normalized valve travel time equal to 1, as well as valve system movements just after the seals have made contact with the chamber door, corresponding to valve travel times just after 1).

Figures 5 and 6 provide further characterization of what causes the entropy (disorder) growth apparent in Fig. 4. They respectively show instantaneous frequency and instantaneous frequency variance and are organized similarly to Fig. 4, in the sense that their horizontal axes correspond to the increasing number of cycles this valve executed, while their vertical axes denote the position of the valve within any given motion cycle.

Inspection of expected instantaneous frequencies in the areas in Fig. 5[6] emphasized by dashed lines implies that instantaneous frequencies near the normalized travel time corresponding to 1 (during the end stages of camming and seal contact with the chamber door) evolve toward lower values as the valve executes more and more cycles (as we go from left to right in Fig. 5). Namely, from the frequency point of view, quiet valve vibrations look like white noise and display frequencies in various parts of the spectrum, including high frequencies, leading to relatively high expected instantaneous frequencies. On the other hand, when some time–frequency structure appeared in the signals (e.g. when we have metal to metal contact or poor vibration dissipation due to a degraded seal), the instantaneous frequency was lowered compared to the "white noise-like" vibrations.

Inspection of same emphasized areas in Fig. 6 shows that instantaneous variances of frequencies evolve towards lower values just before the normalized valve travel time 1 (dark lines pointed by arrows in that figure), which corresponds to the period when seal contacts the chamber door. This is indicative of an increase in valve vibrations at a fairly localized frequency, which illustrates decreased ability of the seal to dissipate vibrations (i.e. a degraded seal). Furthermore, one can note growth of instantaneous variances of vibration frequencies just after the normalized valve travel 1, which corresponds to the movements of the valve mechanism after the seal makes contact with the chamber door. It is likely indicative of impulse-like metal-to-metal contacts (impulse-like signals have a wide spread frequency content), which are occurring because of degraded lubrication of the portion of the valve mechanism responsible for that motion.

Information about the time of appearance and character of anomalies in valve vibrations, as well as their location in the normalized valve travel (where in the normalized travel anomalies occur) yielded by Figs. 4, 5, 6 is of paramount importance for scheduling and planning of maintenance operations on those valves (when to take a valve down for maintenance *and* what to do with it, in the sense should one lubricate the valve guides, replace the seal or both).

[6] Those areas correspond to the portions of valve travel where increases in entropies were observed.

Table 4 Most commonly used features

Feature	Feature name	Direction of valve motion	Channel of the vibration signal	Portion of the signal	Usage
median $\left[\|\text{sig (t, f)}\|^2\right]$	Median energy in the time–frequency domain	Closing	RMS	Entire Signal	68.65%
$E\left[\|\text{sig (t, f)}\|^2\right]$	Mean energy in the time–frequency domain	Closing	RMS	Between	8.16%
$T_{Movement\_Closing}$	Movement time	Closing			3.02%
median $\left[\|\text{sig (t, f)}\|^2\right]$	Median energy in the time–frequency domain	Closing	RMS	Post	2.12%
$argmax_{\text{t}}\,\|sig\,(t)\|^2$	Position of maximum energy in the time domain	Opening	X-direction	Pre	2.04%
$E\left[\|sig\,(t)\|^2\right]$	Mean energy in the time domain	Closing	Y-direction	Between	1.80%
$E\left[\|sig\,(t)\|^2\right]$	Mean energy in the time domain	Closing	RMS	Between	1.71 %
median $\left[\|\text{sig (t, f)}\|^2\right]$	Median energy in the time–frequency domain	Closing	X-direction	Entire Signal	1.63 %
$Sum\,\|sig\,(t)\|^2$	Total energy in the time domain	Opening	RMS	Post	0.90 %

In the column describing portions of the signal, "Pre" denotes portion of the signals before the valve open signal (normalized times −0.2 to 0), "Between" denotes portion of the vibration signals between the valve-open and valve closed signals (normalized times 0 to 1), "Post" denotes portion of the vibration signals after the valve closed signal (normalized times 1 to 1.2) and "Entire Signal" denotes the entire vibration signal obtained during a given valve motion (normalized times −0.2 to 1.2)

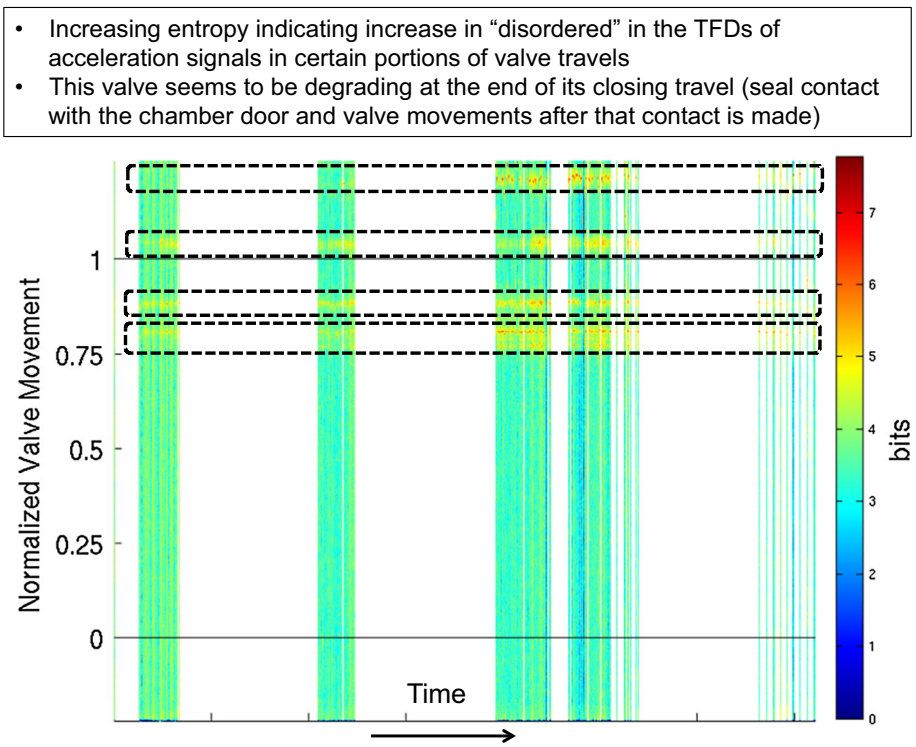

Fig. 4 Progression of instantaneous time–frequency entropies of valve vibrations, starting from an overhaul maintenance event (*left side of the figure*), until the next overhaul maintenance event (*right hand side of the figure*)

Instantaneous expected value of the normalized frequency (also known in signal processing as the instantaneous normalized frequency) evolves downwards in portions of the valve path corresponding to the end of valve travel during closing (seal contact with the chamber door and valve movement after the seals contacted the chamber door)

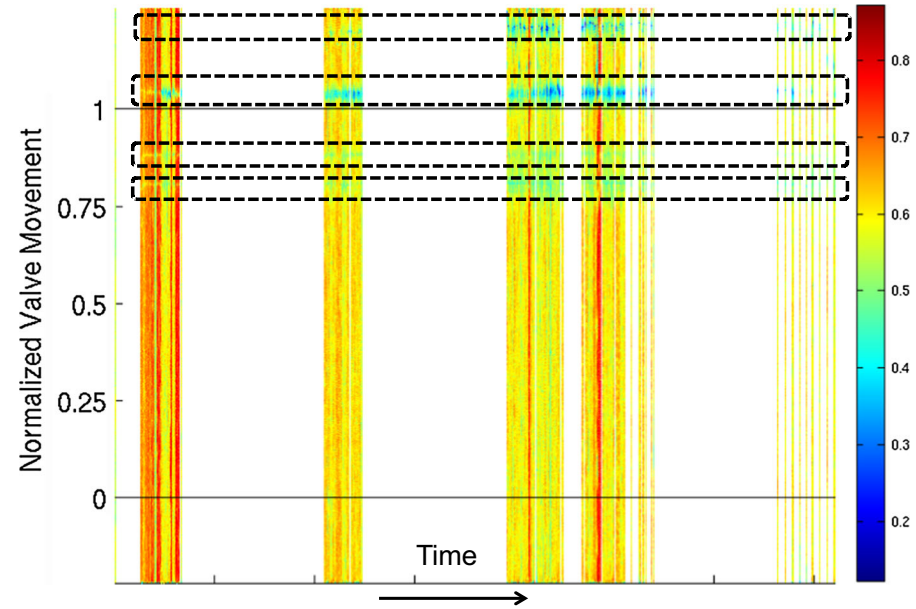

Fig. 5 Evolution of instantaneous expected frequencies evaluated from time–frequency distributions of valve vibrations

- Instantaneous variance of the normalized frequency evolves towards lower values just before the normalized valve travel time 1 (corresponding to the period when seals contact the chamber door)
- Instantaneous variance of the normalized frequency evolves upward just after the normalized valve travel time 1 (corresponding to movements of the valve mechanism after the seals made contact with the chamber door)

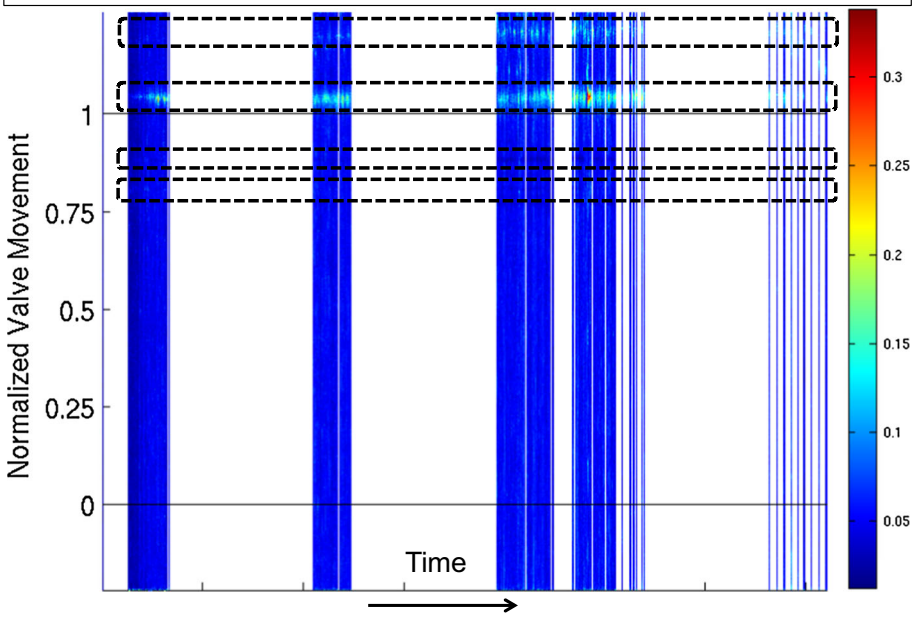

Fig. 6 Evolution of instantaneous frequency variances evaluated from time–frequency distributions of valve vibrations

Conclusions and future work

In this paper, we conducted a classification study in which numerous slit valves used in a major semiconductor manufacturing fab were individually recognized using their vibration signatures. By applying an advanced time–frequency signal processing and feature extraction method, the slit valve vibration signals were transformed into a set of descriptive metrics that were used for characterization of vibration patterns of each individual valve. A kNN based multi-class classification approach was used to recognize the source of the unclassified vibration signals, leading to almost perfect recognition of 50 individual valves in the fab. The few mis-classification occurred within the "quiet" valves, whose motion did not awaken a lot of noticeable vibrations, and thus the valves ended up being confused.

Moreover, long term monitoring of one of the slit valves was accomplished using instantaneous entropy of the time–frequency distributions of its vibrations, in spite of the operational variations in the speed and timing of its travel. Time localization of instantaneous entropies enabled localization of the valve degradation (what part of its travel was affected by it). Furthermore, a closer look at the expected values and variances of instantaneous frequencies of valve vibrations further characterizes the valve degradation in the sense that it can be used to detect metal-to-metal contact or seal degradation. Such information is highly important for subsequent maintenance and repair of the degraded valve.

In summary, novel contributions of this paper can be summarized as follows. Vibration based characterization and condition monitoring of slit valves in semiconductor manufacturing is a new application that necessitated the use of a novel combination of nonstationary signal analysis, pattern recognition and monitoring methods. In addition, implementation of the methodology in a real fab is an important contribution because collecting and processing of signals coming from fab machines rather than from valves operating in a lab carries significant challenges in terms of enabling automated triggering of signal collection and removing fab noise.

Future work consists of implementing the monitoring solution across all the valves in the relevant fab and studying their degradation patterns (what portions of the valve travel are susceptible to degradation), as well as how to more efficiently operate them (what speeds lead to slower degradation) and maintain them (optimal scheduling and planning of maintenance operations on the valves).

Acknowledgements This research is supported in part by the National Science Foundation (NSF) grant IIP 1266279. The content of this paper is solely the responsibility of the authors and does not represent the official views of the NSF.

References

Analog Devices. (2015). ADXL327 data sheet. http://www.analog.com/static/imported-files/data_sheets/ADXL327.pdf. Accessed 15 August 2015

Bao, J., & Spanos, C. J. (2001). A simulation framework for lithography process monitoring and control using scatterometry. In *AEC/APC Symposium XIII*, Abstract available via http://impact.berkeley.edu/archive/secure/archives/seminars/abstracts/Junwei100101.pdf. Accessed 15 August 2015

Cholettte, M., Celen, M., Djurdjanovic, D., & Rasberry, J. (2013). Condition monitoring and operational decision making in semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing*, *26*(4), 454–464.

Coates, M., & Fitzgerald, W. (1999). Regionally optimised time–frequency distributions using finite mixture models. *Signal Processing*, *77*(3), 247–260.

Cohen, L. (1995). *Time–frequency analysis: Theory and applications* (1st ed.). Englewood Cliffs: Prentice Hall.

Djurdjanovic, D., Ni, J., & Lee, J. (2002). Time–frequency based sensor fusion in the assessment and monitoring of machine performance degradation. In *Proceedings of the 2002 ASME international mechanical engineering congress and exposition (IMECE)* (pp. 15–22).

Duda, R. O., Hart, P. E., & Stork, D. G. (2000). *Pattern classification* (2nd ed.). London: Wiley.

Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, *3*(3), 32–57.

Facco, P., Bezzo, F., Barolo, M., Mukherjee, R., & Romagnoli, J. A. (2009). Monitoring roughness and edge shape on semiconductors through multiresolution and multivariate image analysis. *AIChE Journal*, *55*(5), 1147–1160.

Flett, J., & Bone, G. M. (2016). Fault detection and diagnosis of diesel engine valve trains. *Mechanical Systems and Signal Processing*, *72*, 316–327.

Fulton, S., & Kim, M. (2007). ISMI consensus on preventive and predictive maintenance vision Ver. 1.1. In: *International SEMATECH manufacturing initiative (ISMI)*. http://www.sematech.org/docubase/document/4819ceng.pdf. Accessed 15 August 2015

Guan, T., Kuang, Y. C., Ooi, M., Cheah, X. G., Tan, Y. S., & Demidenko, S. (2011). Data-driven condition-based maintenance of test handlers in semiconductor manufacturing. In *Proceedings of the 6th IEEE international symposium on electronic design, test and application (DELTA)* (pp. 189–194).

Heng, R., & Nor, M. (1998). Statistical analysis of sound and vibration signals for monitoring rolling element bearing condition. *Applied Acoustics*, *53*(1), 211–226.

Hong, S. J., Lim, W. Y., Cheong, T., & May, G. S. (2012). Fault detection and classification in plasma etch equipment for semiconductor manufacturing. *IEEE Transactions on Semiconductor Manufacturing*, *25*(1), 83–93.

Hopfe, V., Sheel, D., Spee, C., Tell, R., Martin, P., Beil, A., et al. (2003). In-situ monitoring for CVD processes. *Elsevier Journal on Thin Solid Films*, *442*(1), 60–65.

Kim, B., & May, G. S. (1997). Real-time diagnosis of semiconductor manufacturing equipment using a hybrid neural network expert system. *IEEE Transactions on Components, Packaging, and Manufacturing Technology—Part C*, *20*(1), 39–47.

Jeong, J., & Williams, W. J. (1992). Kernel design for reduced interference distributions. *IEEE Transactions on Signal Processing*, *40*(2), 402–412.

Jones, D. L., & Baraniuk, R. G. (1995). An adaptive optimal-kernel time-frequency representation. *IEEE Transactions on Signal Processing*, *43*(10), 2361–2371.

Jong, W. R., & Lin, T.-W. (2007). Statistical process control for e-diagnostic prediction of cluster-tool equipment. In *Proceedings of the 33rd annual conference of the IEEE industrial electronics society—IECON 2007* (pp. 2916–2921).

Kittler, J. (1975). Mathematical methods of feature selection in pattern recognition. *International Journal of Man-Machine Studies*, *7*(5), 609–637.

Kreßel, U. H.-G. (1999). Pairwise classification and support vector machines. In B. Schölkopf & C. J. C. Burges (Eds.), *Advances in kernel methods* (pp. 255–268). New York: MIT Press.

Lee, D. E., Hwang, I., Valente, C. M., Oliveira, J., & Dornfeld, D. A. (2006). Precision manufacturing process monitoring with acoustic emission. *International Journal of Machine Tools and Manufacture*, *46*(2), 176–188.

Lee, S. K., Kim, T. R., Lee, S. G., & Park, S. K. (2010). Degradation mechanism of check valves in nuclear power plants. *Annals of Nuclear Energy*, *37*(4), 621–627.

Lin, Y. H., Lee, W. S., & Wu, C. Y. (2013). A novel signal processing approach for valve health condition classification of a reciprocating compressor with seeded faults considering time-frequency partitions. *Journal of Marine Science and Technology*, *21*(5), 578–585.

Mahamad, A. K., & Hiyama, T. (2011). Fault classification based artificial intelligent methods of induction motor bearing. *International Journal Innovative Computing, Information and Control*, *7*(9), 5477–5494.

Musselman, M., & Djurdjanovic, D. (2012). Time–frequency distributions in the classification of epilepsy from EEG signals. *Expert Systems with Applications*, *39*(13), 11413–11422.

National Instruments Corporation. (2015). sbRIO-9636 OEM operating instructions and specifications. http://www.ni.com/pdf/manuals/373378d.pdf. Accessed 15 August 2015

Papandreou-Suppappola, A. (Ed.). (2002). *Applications in time–frequency signal processing* (1st ed.). Boca Raton: CRC Press.

Pichler, K., Lughofer, E., Pichler, M., Buchegger, T., Klement, E. P., & Huschenbett, M. (2016). Fault detection in reciprocating compressor valves under varying load conditions. *Mechanical Systems and Signal Processing*, *70*, 104–119.

Raoux, S., Liu, K., Guo, X., & Silvetti, D. (1998). In-situ RF diagnostic for PECVD process control. In *Proceedings of the materials research society (MRS), symposium II* (Vol. 502, pp. 53–58). Cambridge University Press.

Saxena, A., & Saad, A. (2007). Evolving an artificial neural network classifier for condition monitoring of rotating mechanical systems. *Applied Soft Computing*, *7*(1), 441–454.

Shadmehr, R., Angell, D., Chou, P. B., Oehrlein, G. S., & Jaffe, R. S. (1992). Principal component analysis of optical emission spectroscopy and mass spectrometry: Application to reactive ion etch process parameter estimation using neural networks. *Journal of the Electrochemical Society*, *139*(3), 907–914.

Shen, C. W., Cheng, M. J., & Chen, C. W. (2011). A fuzzy AHP-based fault diagnosis for semiconductor lithography process. *International Journal of Innovative Computing, Information and Control*, *7*(2), 805–815.

Tang, J., Dornfeld, D., Pangrle, S. K., & Dangca, A. (1998). In-process detection of microscratching during CMP using acoustic emission sensing technology. *Journal of Electronic Materials*, *27*(10), 1099–1103.

Wang, Q. H., Zhang, Y. Y., Cai, L., & Zhu, Y. S. (2009). Fault diagnosis for diesel valve trains based on non-negative matrix factorization and neural network ensemble. *Mechanical Systems and Signal Processing*, *23*(5), 1683–1695.

Wu, H., Chang, C., Chen, B., Lee, C., Chang, C., Ko, J., Zhou, M., & Liang, M. (2003). Fault detection and classification of plasma CVD tool. In *Proceedings of the 2003 IEEE international symposium on semiconductor manufacturing* (pp. 123–125).

Wuxing, L., Tse, P. W., Guicai, Z., & Tielin, S. (2004). Classification of gear faults using cumulants and the radial basis function network. *Mechanical Systems and Signal Processing*, *18*(2), 381–389.

Yang, B. S., Hwang, W. W., Ko, M. H., & Lee, S. J. (2005). Cavitation detection of butterfly valve using support vector machines. *Journal of Sound and Vibration*, *287*(1), 25–43.