Incipient Fault Detection of Reactive Ion Etching Process

Sang Jeen Hong[a] and Jae-Hyun Park
Department of Electronic Engineering & NBRC, Myongji University, San 38-2 Nam-dong, Yongin-si, Gyeonggi 449-728, Korea

Seung-Soo Han
Department of Information Engineering & NPTC, Myongji University, San 38-2 Nam-dong, Yongin-si, Gyeonggi 449-728, Korea

[a]E-mail: samhong@mju.ac.kr

(Received December 1, Accepted December 12 2005)

In order to achieve timely and accurate fault detection of plasma etching process, neural network based time series modeling has been applied to reactive ion etching (RIE) using two different in-situ plasma-monitoring sensors called optical emission spectroscopy (OES) and residual gas analyzer (RGA). Four different subsystems of RIE (such as RF power, chamber pressure, and two gas flows) were considered as potential sources of fault, and multiple degrees of faults were tested. OES and RGA data were simultaneously collected while the etching of benzocyclobutene (BCB) in a SF_6/O_2 plasma was taking place. To simulate established TSNNs as incipient fault detectors, each TSNN was trained to learn the parameters at t, $t+T$, ..., and $t+4T$. This prediction scheme could effectively compensate run-time-delay (RTD) caused by data preprocessing and computation. Satisfying results are presented in this paper, and it turned out that OES is more sensitive to RF power and RGA is to chamber pressure and gas flows. Therefore, the combination of these two sensors is recommended for better fault detection, and they show a potential to the applications of not only incipient fault detection but also incipient real-time diagnosis.

Keywords : RIE, OES, PCA, RGA, Neural network, Fault detection

1. INTRODUCTION

For the semiconductor manufacturing in this era of sub-micron technology, component density is continuously increasing as device size decreases. Due to the nature of inherent variability of sophisticated semiconductor equipment, stringent control of process as well as equipment is required in order to maximize process yield. Although a certain amount of inherent variability is regarded as random process noise, significant performance shifts may occur when this variability becomes larger compared to the acceptable noise. Such shifts are often indicative of faulty process. When operating conditions become to vary beyond an acceptable level, overall product quality can be jeopardized. Thus, timely and accurate faulty process detection is required for the success of the semiconductor manufacturing.

Statistical process control (SPC) has been used to identify out-of-control state of processes[1]. Although SPC is a useful process monitoring technique based on control chart consists of upper and lower limits, SPC techniques are able to detect undesirable process shift, they have drawback of being incapable of detecting shifts until the process step in question is completed and after the measurements are made. In addition, since the classical SPC is also not sensitive to detect a small process shifts, cumulative sum (CUSUM) control chart has been used for detecting small variation[2]. These statistical monitoring techniques are capable of off-line monitoring, but the delay from post-process measurements can result in the fabrication of semiconductor devices that do not conform to required specification. Thus, products within the delayed time period become in question until the following measurements are made. However, the proposed approach based on on-line process monitoring will provide the advantage unable to achieve from SPC techniques. Faulty process detection using in-situ sensors can offer even more benefits so as to generate malfunction alarm prior to or at the very onset of a shift.

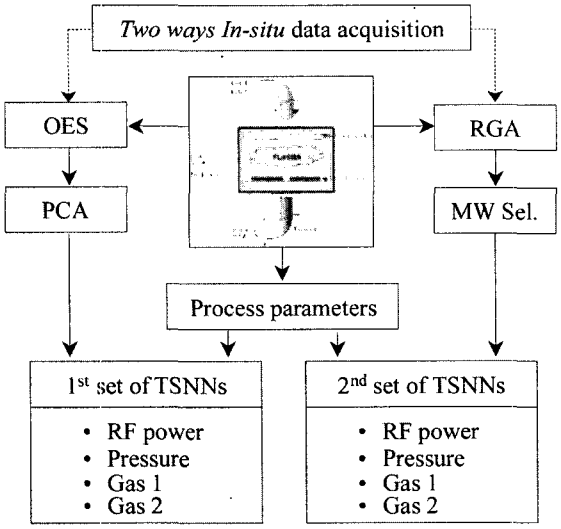

Fig. 1. Graphical overview of the proposed research using two *in*-situ sensors(MW Sel. stands for the selection of statistically significant molecular weight as data pre-processing).

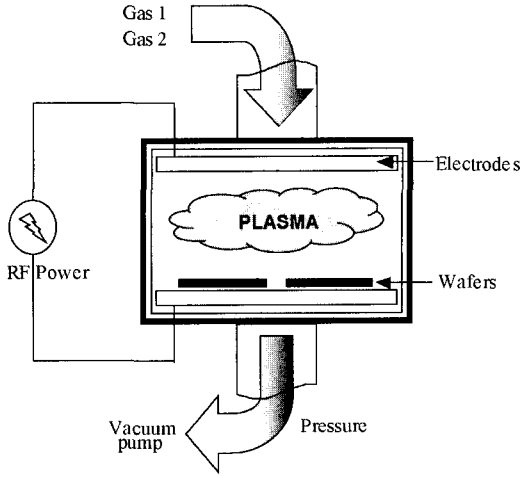

Fig. 2. A schematic of reactive ion etching.

As a test vehicle, reactive ion etching (RIE) was employed whose role should not be overemphasized for current VLSI and future GSI technology. RIE is a plasma etching using an electrical discharge to ionize gas phase and induce ion bombardment of semiconductor wafer in order to obtain the required pattern. RIE system typically operates in low pressure. Since RIE has the capability of producing extremely minute features, it has been widely used for etching various materials. However, the lack of real-time etch control often allows unacceptable large volume of defective material to be processed in a manufacturing line before being detected resulting in significant wasted material and lower yield. Texas Instrument reported that annually estimated $135 million is lost in etch of its factories due to the lack of adequate real-time control and equipment diagnostic capabilities in RIE[3]. This shows RIE is a workhorse in semiconductor manufacturing area but plays a role as a yield limiter as well. Therefore, RIE is a good candidate for the development of a real-time fault detector.

To facilitate a better understanding of the RIE process, in-situ plasma process monitoring has risen as the key issue. Most frequently used metrology techniques are optical emission spectroscopy (OES) and residual gas analyzer (RGA). RGA is one of the well-known in-situ plasma monitoring sensors that has been used for leak detection in vacuum chamber[4]. OES systems measure the variation of the optical emission intensity of plasma as a function of the reactants and by-products inside etch chamber. This technique offers the unique capability of monitoring plasma chemistry directly in non-invasive way. Currently, OES is extensively used for etch end-point detection application[5]. It has the advantage of non-invasiveness over other monitoring methods, but it generates a huge amount of information, so the analysis of such voluminous data is a challenge. However, principal component analysis (PCA) was successfully able to reduce the dimensionality of large amount of OES data while maintaining most of the variation [4,6,8]. Therefore, OES is also a good candidate as in-situ plasma monitoring sensor for real-time application of RIE.

In order to achieve fault detection, two sets of time series neural networks (TSNNs) were developed. In each case, a set of TSNNs corresponding to the responses of interest of RIE was trained to predict the process parameters. Once networks are properly trained off-line with OES and RGA data respectively acquired from baseline and various faulty runs, the predicted parameters were filtered with the moving average (MA) technique. This will contributes to maintain consistent etching processes, increase the possibility to catch incipient faults, and eventually improve process yield in semiconductor manufacturing. Graphical overview of this research is depicted in Fig. 1. In Section 2, descriptions for experimental apparatus are explained, and detailed explanation for neural network is provided in Section 3. Processing of the two sensory data is described in Section 4. In Section 5, time series neural networks are described, and the results and conclusion are presented in the following section.

2. EXPERIMENTAL APPARATUS

2.1 Reactive ion etching (RIE)

RIE is dry-etch process using electrical discharge to ionize gas phase and induce ion bombardment of wafer

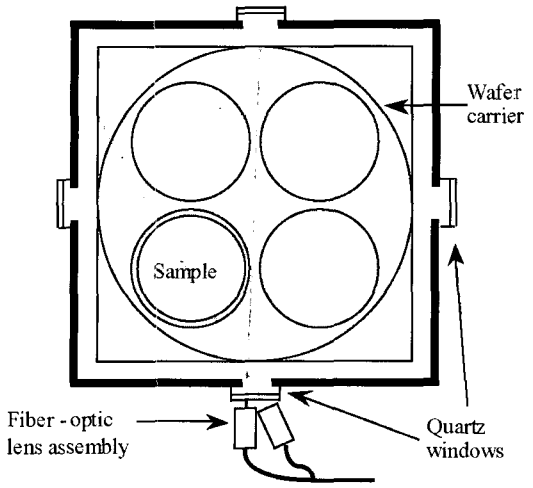

Fig. 3. OES sensor configuration for *in-situ* plasma monitoring in non-invasive way.

surface in order to obtain the required etch properties. RIE is typically operating below ~100 mTorr with 13.56 MHz RF discharge in parallel plate configuration. A schematic diagram of this etching tool is in Fig. 2. Precise etching can be achieved for smaller geometries by controlling the direction of ion bombardment. Both chemical and physical mechanisms contribute to etching processes.

The chemical reactions between reactive ion species and the material on the surface remove the surface. In addition, the physical ion bombardment of ionized gas species increases the chemical reaction by breaking chemical bonds on the surface. The chemical process has three phenomena:

1. The molecular gases, which may or may not reactive, are dissociated in the plasma to create reactive species.
2. The reactive species are absorbed and react with exposed material to form volatile products.
3. The absorbed volatile products are desorbed and removed by the vacuum system.

The physical etch process takes place due to the RF power applied to the parallel plates in etch chamber. Since electrons have higher mobility than ions, they are accelerated toward the negatively charged electrode that caused from the self DC-bias voltage between the plasma and the electrodes. The chemical and physical etch mechanisms enhance etch process and enable to achieve higher etch rate than either independent reaction. Detailed mechanisms of reactions are described in many literatures, but the literatures described the complexity of these physical and chemical reactions make RIE more complex and nonlinear multivariable process.

Table 1. RIE process fault scenarios.

Inputs	Example	Values	% dev.	Run No.	
Baseline process	RF pow/Pressure/Gas1/Gas2 300/200/40/5		0	1-10	
RF Power	High RF power	315	+5	11	12
	Mid-high RF power	310	+3.3	13	14
	Mid-low RF power	290	-3.3	15	16
	Low RF power	285	-5	17	18
Chamber Pressure	High Pressure	220	+10	19	20
	Mid-high Pressure	210	+5	21	22
Gas 1 (O_2)	High Gas 1	44	+10	23	24
	Mid-high Gas 1	42	+5	25	26
	Mid-low Gas 1	38	-5	27	28
	Low Gas 1	36	-10	29	30
Gas 2 (SF_6)	High Gas 2	7	+40	31	32
	Mid-high Gas 2	6	+20	33	34
	Mid-low Gas 2	4	-20	35	36
	Low Gas 2	3	-40	37	38

2.2 Optical emission spectroscopy (OES)

OES has been proposed as a useful monitoring technique for real-time monitoring of plasma process[6-9]. This technique provides in-situ analysis without disturbing the plasma. By combining an imaging spectrograph and CCD camera with integrated fiber optics, it is possible to acquire information on the radial or axial distribution of critical plasma species which control the uniformity, selectivity and critical dimensions of etch features on wafer.

The Chromex OES 3020 system, consists of two split fixed gratings for broad spectral coverage: 300 g/mm blazed at 250 nm (covers 230~495 nm) and 300 g/mm blazed at 600 nm (covers 525~790 nm), is used for spectral monitoring inside of RIE chamber during the series of etch periods. It consists of three sensor units, but only two sensor units were used in this experiment due to the geometric limitation of the chamber window. The data collected from a more sensitive sensor that aimed at the middle of the sample to monitor more accurately is used for actual data acquisition. The second sensor that pointed to the middle of the etch chamber to achieve overall intensity monitoring is used for supporting the validity of the first sensor. OES sensor configuration in this experiment is in Fig. 3.

2.3 Residual gas analyzer (RGA)

Special mass spectrometer designed to analyze the gases in a vacuum system is called residual gas analyzer (RGA). RGA is one of the well-known in-situ plasma

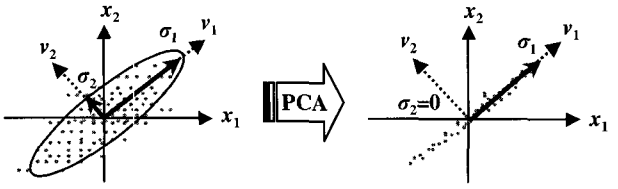

Fig. 4. An illustration of principal component analysis of two measurement variables: x_1 and x_2 indicated mean centered sample data, v_1 and v_2 are eigenvectors, and σ_1 and σ_2 are corresponding standard deviations[15].

monitoring sensors that has been used for leak detection in vacuum chamber and end point detection in plasma vapor deposition (PVD), chemical vapor deposition (CVD) and RIE process monitoring[10]. The RGA system employed in this experiment is the Partial pressure transducer (PPT), manufactured by MKS instruments, that utilizes Quadrupole mass analyzer (QMA) technology to provide an instrument specifically designed for the production environment.

Quadrupole mass spectrometers are the most common type of gas analyzer for vacuum application. They are often used in vacuum systems where it is beneficial to know the partial pressures of component gases. Partial pressures are determined by ionizing atoms and molecules, mass filtering ions, and then detecting unfiltered ions. Atoms and molecules are first ionized to provide an electronic handle by which ions may be separated and subsequently detected. After ionization, the quadrupole rods of RGA serve as sets of tunable high and low-band pass filter, allowing only the ion species of interest to reach the detector. The detector measures an amplified signal proportional to the ion current. It is possible to quickly determine the partial pressures of gases in a vacuum system, typically from 1 to 200 atomic mass units (AMU) in the range from 10-4 to 10-9 Torr. A continuous gas sample is introduced to the partial pressure sensor through a valve or a pressure reduction aperture while the vacuum pumping system maintains a sufficiently low-pressure sensor side of the aperture for optimum sensor operation[11].

2.4 Materials

The test vehicle for etching is Benzocyclobutene (BCB) under the commercial name Cyclotene 3022-46 that has been widely used in MCM-D fabrication as inter-level dielectrics (ILD). BCB offers distinctive advantages for ILD such as low dielectric constant (K=2.65), excellent planarization properties, low moisture absorption, good compatibility with copper and the ability to be cured by rapid thermal techniques. However, wet etching is not alternative in fully-cured BCB etching. Metal or organic mask has been used in

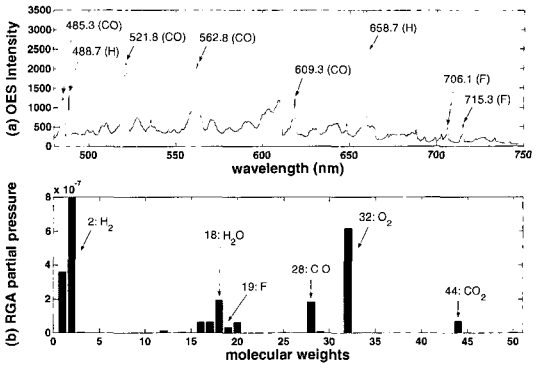

Fig. 5. Particles reside in etch chamber corresponding to BCB etching with O_2/SF_6 plasma. Statistically significant peaks are indicated with their corresponding wavelength and molecular weights.

most of applications, but soft mask for via patterning that greatly simplifies via patterning in BCB film has been proposed[12].

2.5 Fault definition and data generation

Process shifts can take place by numerous sources and the combination of the sources. However, the process fault, in this instance, is defined as approximately more than 5 % deviation from the baseline. Underlying assumption is that a single source contributes the process shifts for each parameter. 10 baseline runs and two sets of 14 different faulty runs were performed. In order to avoid the run-to-run variation, all the baseline runs were performed consecutively, and the second set of faulty runs was repeated for system verification purpose. Table I contains the fault process scenarios and their degrees of fault. OES data was collected every 10 seconds during the five minutes of each run, and altering baseline and faulty runs formed faulty pattern library for each parameter.

3. DATA PROCESSING

3.1 Principal component analysis of OES data

Principal component analysis (PCA) is an established statistical method for compressing a multivariate data set [13]. PCA has been suggested as a method for analyzing optical emission spectra due to its ability to dramatically reduce the dimensionality of large amounts of OES data [14]. Consider a vector $\mathbf{x}$ that consists of p random variables. Let Σ be the covariance matrix of $\mathbf{x}$. Then, for $k = 1,2,..., p$ the k^{th} principal component (PC) is given by:

$$\mathbf{t}_k = \mathbf{u}_k^T \mathbf{x} \qquad (1)$$

where $\mathbf{u}_k$ is an eigenvector of Σ corresponding to its k^{th} largest eigenvalue, and T represents the transpose operation. Dimensionality reduction through PCA is achieved by transforming the OES data to a new set of coordinates (i.e., selected k eigenvectors), which are uncorrelated and ordered such that the first few retain most of the variation present in the original data set. Generally, if the eigenvalues are ordered from largest to smallest, then the first few PCs will account for most of the variation in the original vector **x**. A simplified example of PCA with two measurement variables, x_1 and x_2, is presented in Fig. 4.

In order to simulate in real-time fashion, PCA was performed on the most recently collected 10 samples with 20 seconds of sampling period with pre-selected 340 of most significantly varying wavelengths. The Eq. (2) represents the selected data matrix **S**:

$$\mathbf{S} = \begin{bmatrix} s_{11}, \cdots, s_{1w} \\ \ddots \\ s_{o1}, \cdots, s_{ow} \end{bmatrix} \tag{2}$$

where the index $o = 1,2,...,10$ represents one of 10 consecutive observations collected every 20 seconds. The pre-filtered and expanded data matrix **S** was mean centered with respect to each column using the equation:

$$m_{xy} = s_{xy} - \bar{\mathbf{s}}_y, \quad \text{for } 1 \le x \le o \text{ and } 1 \le y \le w \tag{3}$$

where m_{xy} and s_{xy} are the components of the mean centered matrix **M** and the sample matrix **S** respectively, is the mean of the column y over the total number of rows (o). The variance-covariance of matrix **C** was computed in the Matlab as:

$$\mathbf{C} = \left(\frac{1}{c-1}\right)\mathbf{M}^T\mathbf{M} \tag{4}$$

Employing singular value decomposition, the variance-covariance matrix **C** was decomposed into a linear combination of the eigenvectors and eigenvalues using:

$$\mathbf{C} = \mathbf{U}\Delta\mathbf{U}' \tag{5}$$

where $\mathbf{U} = [\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_w]$ is a $w \times w$ orthogonal (unitary) matrix containing c eigenvectors, and Δ is the diagonal matrix of eigenvalues such that $\delta_1 \ge \delta_2 \ge \cdots \ge \delta_w \ge 0$. Using singular value decomposition, the eigenvectors were arranged in descending order associated with the magnitude of the eigenvalues. The first few eigenvectors capture most of the variation in the original data.

Table 2. Detected molecules with RGA. Bold numbers are regarded as statistically significant.

AMU	Species	% Variation
1	H	45.5177
2	**H_2**	**27.8169**
12	C	0.0054
16	O	0.0516
17	OH	0.0820
18	**H_2O**	**1.0646**
19	F	0.0265
20	HF	0.0614
28	**CO**	**2.2751**
32	**O_2**	**17.6028**
44	**CO_2**	**5.4386**
69	Cyclohexane	0.0440

The cumulative percentage of total variation was employed as a criterion for choosing the number of principal components. The definition of the cumulative percentage of variation is:

$$\sigma_k = \left(\sum_{q=1}^{k}\delta_q \Big/ \sum_{q=1}^{w}\delta_q\right)\cdot 100\,(\%) \tag{6}$$

where c is the total number of eigenvalues as same as the total number of columns of **C**, and δ_k is the k^{th} diagonal element of eigenvalues matrix Δ (k denotes the number of subset of principal components). Using this approach, the mean-centered OES data set was compressed into five principal component vectors by transposing **M** onto the selected new set of coordinates, or:

$$\hat{\mathbf{T}} = \mathbf{M}\hat{\mathbf{U}} \tag{7}$$

where $\hat{\mathbf{U}}$ is a $w \times k$ orthogonal matrix. The five of the ten-element principal component vectors (k=5) were subsequently used as the input data for neural network modeling of the etch responses.

Once the PCA was done for each sampling period, the selected first five principal components and auto- and cross-correlation of the concurrent process parameters were simultaneously used as inputs of neural networks modeling. Emission intensity was recorded every ten seconds, and the collected data was stored in a local computer as ASCII files.

3.2 Molecular weight selection of RGA data

In this experiment, RGA was simultaneously employed to monitor the etching chamber with OES.

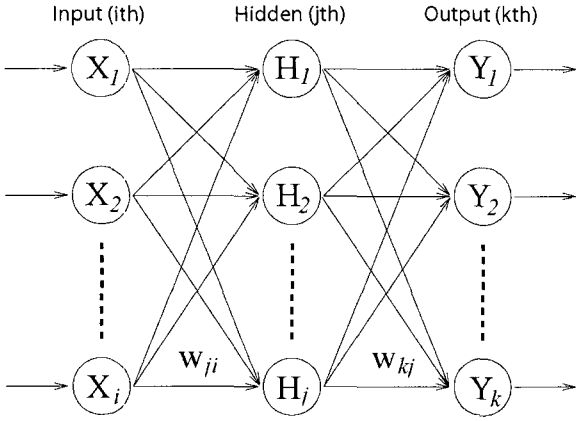

Fig. 6. A multi-layer perceptron typed neural network.

RGA provides the partial pressure of specific components gases present in the chamber, which can be interpreted as the concentration of the components. While the partial pressure is maintained 5×10^{-5} mTorr, the PPT Series RGA provides from 1 to 200 atomic mass units (AMU). Total 12 different kinds of AMUs were detected as in Fig. 5(b), but some of them were found to be statistically significant over the etching process. The selected AMUs were continuously used for neural networks modeling. Detected and selected AMUs are shown in Table 2. Hydrogen contributes the most of the variation in RGA scanning. However, it is highly suspicious that the detected hydrogen atoms were residue build up the chamber wall during the previous processes. Therefore; hydrogen atom (AMU=1) was excluded from the selection of significantly varying components.

4. NEURAL NETWORKS

Neural networks (NNs) have become useful tools in semiconductor process modeling and have demonstrated the capability of learning complex relationships between groups of related parameters. A NN is a structured interconnection of computational nodes called "neurons" that contribute to parallel computation in a manner similar to the human brain. The interconnection of neurons establishes knowledge that is acquired by the network through a learning process, and that knowledge is stored in the form of inter-neuron connection strengths known as "weights". Each neuron contains the weighted sum of its inputs filtered by a neuron activation function, providing NNs with the ability to generalize with an added degree of freedom that is not available in statistical regression techniques[15]. Due to the inherent ability to learn complex nonlinear mapping, NNs have been applied to semiconductor process modeling, control and optimization, and those studies reported excellent results using neural network-based techniques[16].

Many of different neural networks structures and learning algorithms are available. The learning algorithm used in this study is the feed forward error back-propagation (BP). A multilayer perceptron neural network structure is depicted in Fig. 6. In the BP learning algorithm, a single iteration consists of two parts: a forward and a backward computation[17]. In the forward computation, the outputs from the $(l\text{-}1)^{th}$ layer are weighted and summed, and the weighted sum is filtered through a sigmoid function. The outputs of neurons in lth layer become inputs to the neurons in the next layer k. The internal activity level $s_j^{(l)}(n)$ for neuron j in layer l is

$$s_j^{(l)}(n)=\sum_{i=0}^{p} w_{ji}^{(l)}(n)o_i^{(l-1)}(n) \tag{8}$$

Where $o_i^{(l-1)}(n)$ is the function signal of neuron i in the previous layer (l-1) at iteration n, $w_{ji}^{(l)}(n)$ is the synaptic weight of neuron j in layer l that is fed from neuron i in layer l-1, and p is the number of neurons in the l^{th} layer. For i=0, $o_i^{(l-1)}(n)=-1$ and $w_{i0}^{(l)}(n)=\theta_j^{(l)}(n)$, where $\theta_j^{(l)}(n)$ is the threshold applied to neuron j in layer l. Then, the output signal of neuron j in layer l is

$$o_j^{(l)}(n)=\begin{cases} x_j(n), & l=1 \\ \dfrac{1}{1+\exp[-s_j^{(l)}(n)]}, & 1<l<L \\ y_j(n), & l=L \end{cases} \tag{9}$$

where $x_j(n)$ is the j^{th} element of the input vector in first hidden layer (i.e., l=1), and L denotes the last layer.

In backward computation, weights are updated in the direction that minimizes an error function defined by:

$$E_j^{(L)}(n)=\frac{1}{2}\left(\hat{y}_j-y_j^L(n)\right)^2 \tag{10}$$

where $\hat{y}_j$ is a target in j^{th} component, and $y_j^L(n)$ is the actual output value of j^{th} neuron in the last layer L. The generalized delta rule based on the gradient descent approach is applied to minimize the error function. The expressions for the weight changes (i.e., "deltas") of the output layer and other layers are:

$$\delta_j^{(L)}(n)=[\hat{y}_j-y_j^{(L)}(n)]y_j(n)[1-y_j(n)] \tag{11}$$

$$\delta_j^{(l)}(n)=o_j^{(l)}(n)[1-o_j^{(l)}(n)]\sum_k \delta_k^{(l+1)}(n)w_{kj}^{(l+1)}(n) \tag{12}$$

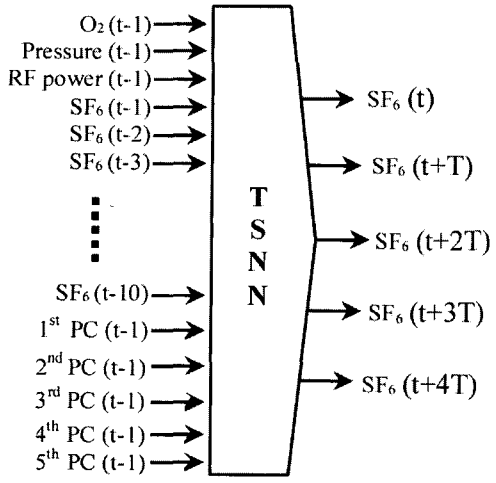

Fig. 7. I/O illustration for a time series neural network for SF_6 prediction(sampling period T=20 seconds).

The weights are initially randomized, and forward propagation is performed. Once the outputs of the last layer are calculated, weights are updated by the deltas for each node calculated from the output layer, and back-propagate to the input layer. The generalized delta rule is:

$$\Delta w_{ji}^{(l)}(n) = [w_{ji}^{(l)}(n) - w_{ji}^{(l)}(n-1)] \quad (13)$$

$$w_{ji}^{(l)}(n+1) = w_{ji}^{(l)}(n) + \eta\delta_j^{(l)}(n)o_i^{(l-1)}(n) + \alpha\Delta w_{ji}^{(l)}(n) \quad (14)$$

where n is the number of iteration, η is the learning rate and α is the momentum. The learning rate is a constant that represents the rate at which a weight will be changed along its slope to the minimum error. The momentum is a constant that includes a portion of the previous weight change to the current weights.

5. TIME SERIES NEURAL NETWORKS

Multi-layer perceptron type NNs based time series modeling was employed to achieve predictive models of reactive ion etching (RIE) using optical emission spectroscopy (OES) data and residual gas analyzer (RGA) data respectively. As it is described earlier, in OES, the most relevant wavelengths were pre-selected from OES data acquired during the etching process, and the first five principal components (PCs) of OES data and auto- and cross-correlation of the process parameters were used as inputs of the NNs. Further information for real-time fault detection is available in the previous research publication[18]. In the similar manner, pre-selected partial pressure of 5 molecules and auto- and cross-correlation of the process parameters were used as inputs of the NNs. Examples of the structure of TSNNs are depicted in Fig. 7.

Table 3. Performance evaluation of the TSNNs.

Parameter	Train	Test	Train	Test
RF power	3.73	3.92	0.78	5.49
Pressure	4.26	2.22	1.80	2.66
Gas 1 (O_2)	0.97	1.01	0.10	0.22
Gas 2 (SF_6)	0.46	0.74	0.03	0.19

Two sets of time series neural networks (TSNNs) were established to detect and predict process fault for etching parameters such as RF power, chamber pressure, and gas flows. Each set of NNs was trained using two different sensory data (OES and RGA) acquired during the etching process. To ensure the validity of established TSNNs, the performance of TSNNs at time t+4T was evaluated for all parameters in terms of root-mean-squared error (RMSE) and it appears in Table 3. The RMSES of TSNNs shows excellent results for the parameter prediction: however, the fluctuations of actual predicted parameters were observed for all parameters. In order to smooth the fluctuations, the predicted outputs of time series neural networks were then applied to moving average (MA) technique[1]. The moving average (MA) is a statistical process control (SPC) scheme that is used for detecting small process shifts. Even though the MA is not as effective as exponentially weighted moving average (EWMA) and cumulative sum (CUSUM) in detecting small shifts, but it is simpler to implement than those and more effective than Shewart chart. Suppose that individual observations have been collected, and let denote the observations. The moving average if span w at time t is defined as

$$M_t = \frac{x_t + x_{t-1} + x_{t-2} + \cdots + x_{t-w+1}}{w} \quad (15)$$

At time t the oldest observation in the moving average set is dropped and the newest one added to the set. Successful results could be achieved with the span w=5 to predict faulty process at four sampling period ahead. Figure 8-11 shows the training and testing results of the prediction results at (t+4T) for both sensory data sets.

6. CONCLUSION

The ultimate goal of this research is to establish neural network models for real-time malfunction diagnosis of

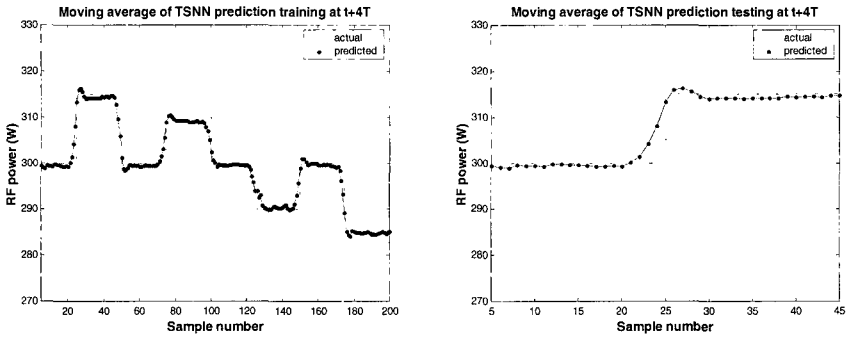

Fig. 8. Moving average of TSNN Prediction for RF power with OES data at t+4T: Left is training and right is testing with data never introduced to the network training.

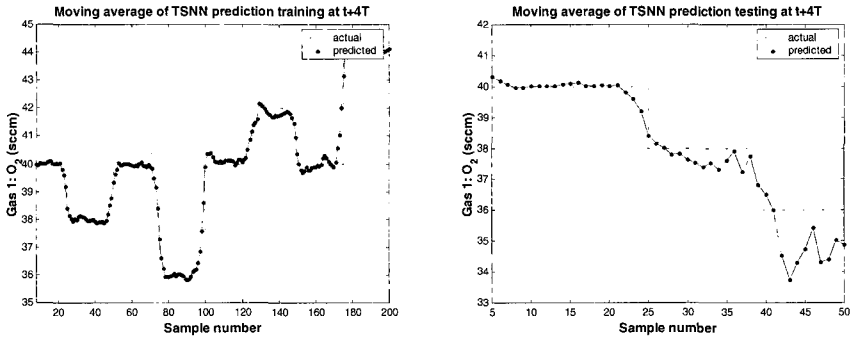

Fig. 9. Moving average of TSNN Prediction for Gas 1 with OES data at t+4T: Left is training and right is testing with data never introduced to the network training.

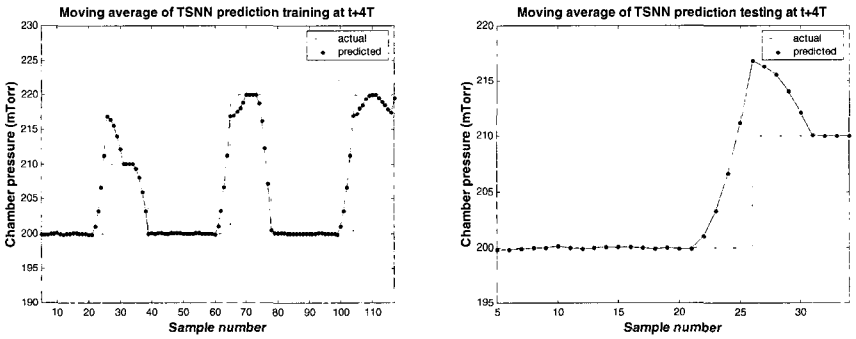

Fig. 10. Moving average of TSNN Prediction for Chamber pressure with RGA data at t+4T: Left is training and right is testing with data never introduced to the network training.

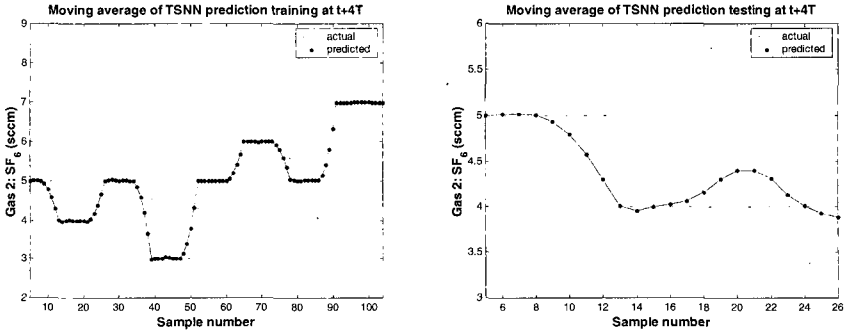

Fig. 11. Moving average of TSNN Prediction for Gas 2 with RGA data at t+4T: Left is training and right is testing with data never introduced to the network training.

reactive ion etching, and the proposed real-time fault detection/prediction has meaningful contribution to the proposed ultimate goal. Malfunction diagnosis involves two aspects. One is identifying the cause of a fault by observing the effects at monitoring points in a system.

Thus, correcting them quickly to prevent subsequent occurrences of other malfunctions is the other. Correct diagnosis is a crucial for process and equipment because subsequent malfunction might be fatal to them.

Redundant (fault-tolerance) system can alleviate this concern, but it increases the cost of equipment as well as manufacturing. In addition to diagnosing existing problems, it is highly desirable to predict malfunctions in advance of their actual occurrence. Prognosis for incipient or future malfunctions also can dramatically improve process quality by allowing end-user to correct process shifts and reduce machine down time by synchronize maintenance schedule with others. With the advent of highly proficient in-situ sensors capable of monitoring process conditions, diagnosis and prognosis in real-time basis became feasible.

From the Fig. 8-11, we can conclude that predicting chamber pressure and gas flows can be accomplished more accurately when the RGA rather than the OES is used (Fig. 10, 11). However, since the output of the OES is more sensitive to RF power (Fig. 8) but less sensitive to pressure and gas flows than the RGA, simultaneous usage of these two sensors is highly recommended rather than individual usage. Detection/prediction has meaningful contribution to the proposed ultimate goal. Malfunction diagnosis involves two aspects. One is identifying the cause of a fault by observing the effects at monitoring points in a system. Thus, correcting them quickly to prevent subsequent occurrences of other malfunctions is the other. Correct diagnosis is a crucial for process and equipment because subsequent malfunction might be fatal to them. Redundant (fault-tolerance) system can alleviate this concern, but it increases the cost of equipment as well as manufacturing. In addition to diagnosing existing problems, it is highly desirable to predict malfunctions in advance of their actual occurrence. Prognosis for incipient or future malfunctions also can dramatically improve process quality by allowing end-user to correct process shifts and reduce machine down time by synchronize maintenance schedule with others. With the advent of highly proficient in-situ sensors capable of monitoring process conditions, diagnosis and prognosis in real-time basis became feasible.

REFERENCES

[1] D. Montgomery, Introduction to statistical quality control, New York, NY: Wiley, 1991.

[2] B. Kim and G. May, "Real-time diagnosis of semiconductor manufacturing equipment using a hybrid neural network expert system", IEEE Trans. Semi. Manufac., Vol. 20, No. 1, p. 39, 1997.

[3] C. Almgren, "The role of RF measurements in plasma etching", Semiconductor International, Vol. 19, No. 8, p. 99, 1997.

[4] D. H. Holkeboer, T. L. Karandy, F. C. Currier, L. C Frees, and R. E. Ellefson, "Miniature quadrupole residual gas analyzer for process monitoring at milliTorr pressure", J. Vac. Soc. Technol. A, Vol. 16, No. 3, p. 1157, 1998.

[5] J. Roland, P. Marcoux, G. Ray, and G. Rnakin, "Endpoint detection in plasma etching", J. Vac. Sci. Technol. A, Vol. 3, p. 631, 1985.

[6] R. Shadmehr, D. Angell, P. B. Chou, G. S. Oehrlein and R. Jaffe, "Principal component analysis of optical emission spectroscopy and mass spectrometry: Application to reactive ion etch process parameter estimation using neural networks", J. Electrochem. Soc., Vol. 139, No. 3, p. 907, 1992.

[7] S. Hong, G. May, and D. Park, "Neural network modeling of reactive ion etch using optical emission spectroscopy data", IEEE Trans. Semi. Manufac., Vol. 16, No. 4, p. 1, 2003.

[8] J. G. Shabushing and P R. Demiko, "Application of optical emission spectroscopy to semiconductor device fabrication", American Laboratory, p. 60, 1984.

[9] S. Hong, D. Park, and G. May, "Modeling optical emission spectroscopy data generated during reactive ion etching using an autoencoder neural network", Smart Engineering System Design: Neural Networks, Fussy Logic, Evolutionary Programming, Data Mining and Complex Systems, Vol. 12, p. 945, (C.H. Dagli, Ed.), New York: ASME Press, 2002.

[10] V. Comello, "RGAs Provides Real Time Process Control", Semi. International, p. 70, 1990.

[11] PPT for Wondows User Guide, The UTI Division of MKS Instruments. Inc., 1996.

[12] B. Roger, M. Berry, I. Turlik, P. Garrow, and D. Castillo, "Soft mask for via patterning in benzocyclobutene", Int. J. Micro. and Elect. Packaging, Vol. 17, No. 3, p. 210, 3rd Quarter, 1994.

[13] I. T. Jolleffe, Principal component analysis, New York, NY: Springer-Verlag, 1986.

[14] D. White, D. Boning, S. Butler, and G. Barna, "Spatial characterization of wafer state using principal component analysis of optical emissions spectra in plasma etch", IEEE Trans. Semi. Manufac., Vol. 10, No. 1, p. 52, 1997.

[15] C. Himmel and G. May, "Advantages of plasma etch modeling using neural networks over statistical

techniques", IEEE Trans. Semi. Manufac., Vol. 6, No. 2, p. 103, 1993.

[16] G. May, "Computational Intelligence in Microelectronics Manufacturing", Chapter 13, The Handbook of Computational Intelligence in Design and Manufacturing (J. Wang and A. Kusiak, Eds.), Boca Raton, FL: CRC Press, 2001.

[17] S. Haykin, Neural Networks, New York, NY: Macmillan College Publishing Company, 1994.

[18] S. Hong and G. May, "Neural network based time series modeling of optical emission spectroscopy data for fault detection in reactive ion etching", Proc. of SPIE Conf. on Adv. Microelec. Manufac., Santa Clara, CA, Vol. 5041, p. 1, 2003.