Neural Network-based Time Series Modeling of Optical Emission Spectroscopy Data for Fault Prediction in Reactive Ion Etching

Sang Jeen Hong[*†]

[*†] Department of Semiconductor Engineering, Myongji University

ABSTRACT

Neural network-based time series models called time series neural networks (TSNNs) are trained by the error back-propagation algorithm and used to predict process shifts of parameters such as gas flow, RF power, and chamber pressure in reactive ion etching (RIE). The training data consists of process conditions, as well as principal components (PCs) of optical emission spectroscopy (OES) data collected *in-situ*. Data are generated during the etching of benzocyclobutene (BCB) in a SF_6/O_2 plasma. Combinations of baseline and faulty responses for each process parameter are simulated, and a moving average of TSNN predictions successfully identifies process shifts in the recipe parameters for various degrees of faults.

Key Words : Reactive Ion Etching, Optical Emission Spectroscopy, Time Series Neural Network

1. Introduction

As the size of microelectronic devices decreases and component density increases, the role of reactive ion etching (RIE) cannot be overemphasized. RIE, a low-pressure type of plasma etching, is a critical technology for current and future GSI fabrication. Due to the inherent variability of RIE, stringent process control is required in order to maximize process yield. Although a certain amount of process fluctuation can be neglected as random noise, significant degradations in product quality can occur when such variability becomes large compared to process set points. Thus, timely and accurate prediction of process shifts is necessary for successful manufacturing.

Although statistical process control (SPC) has been historically used to identify process shifts [1], it is limited to detecting shifts after the process step in question is completed and measurements have been made. Recently, neural networks have been introduced to model the behavior of real-time tool data in the RIE process [2].

Recent advancement of plasma process fault detection and diagnosis using machine learning technique achieved a great amount of improvement. Kim et. al reported the machine learning based-fault detection and classification employing plasma information in terms of plasma physics [3], Choi et. al proposed virtual metrology by interpretation of the optical emission spectroscopy data into wafer process results [4]. More recently, a comparison study of machine learning techniques for semiconductor process fault detection problems. In this paper, we propose a predictive fault detection, herein, a prognosis, instead of diagnostic concept. TSNNs that incorporate the principal components of *in-situ* optical emission spectroscopy (OES) data are proposed as a means to forecast RIE process conditions.

2. Experimental Apparatus

The RIE tool employed in this experiment is a Plasma Therm 700 series dual chamber as depicted in Figure 1. A Chromex 3020 OES system is used for spectral monitoring inside the RIE chamber during etching. The Chromex system consists of three sensor units, but only two sensor units were used in this experiment due to geometric limitations of the chamber window. The data

[†]E-mail: samhong@mju.ac.kr

collected from the most sensitive sensor, which was aimed at the middle of the sample, was used for actual data acquisition. The second sensor, aimed at the middle of the etch chamber, was used to validate the readings of the first sensor.

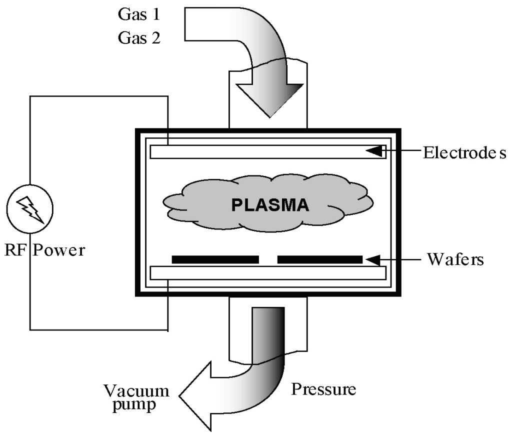


Fig. 1. Typical RIE schematic.

The material etched was benzocyclobutene (BCB), marketed by Dow Chemical under the commercial name *Cyclotene* 3022-46. BCB has been widely used as an inter-level dielectric (ILD) material. BCB offers distinctive advantages for an ILD, including a 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. Metal or organic masks have been used in most of applications of BCB patterning, but a soft mask for via patterning that greatly simplifies processing has been proposed [6].

3. Fault Definition and Data Generation

Process shifts can result from numerous causes and combinations of causes. The process faults in this paper are defined as 5% deviation from baseline values. The underlying assumption is that a single fault causes the process shifts for each parameter. Ten baseline runs and two sets of 14 different faulty runs were performed. In order to minimize run-to-run variation, all baseline runs were performed consecutively, and a second set of faulty runs were repeated for system verification purposes in Table I. OES data was collected every 10 seconds during the five-minute duration of each run.

Table I. 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 (O2)	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 (SF6)	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

4. Principal Component Analysis

Optical emission spectroscopy (OES) is often used for endpoint detection and non-invasive, *in-situ* plasma monitoring in RIE. Its primary disadvantage, however, is the large dimensionality of the spectroscopic data [7]. To alleviate this concern, principal component analysis (PCA) has been used to dramatically reduce the dimensionality of the several thousands of wavelength intensity measurements collected over time by OES systems. PCA is an established statistical method for compressing a multivariate data set.

Consider a vector $\mathbf{x}$ that consists of p random variables. Let $\mathbf{\Sigma}$ be the covariance matrix of $\mathbf{x}$. Then, for k=1, 2, …, r, the k^{th} principal component (PC) is given by $\mathbf{t}_k = \mathbf{u}_k{}^T \mathbf{x}$, where $\mathbf{u}_k$ is an eigenvector of $\mathbf{\Sigma}$ corresponding to its k^{th} largest eigenvalue, and T represents the transpose operation. Generally, if these eigenvectors are ordered from largest to smallest, then the first PCs will account for most of the variation in the original vector $\mathbf{x}$. Dimensionality reduction through PCA is achieved by transforming the original OES data to a new set of variables (i.e., the PCs), which are uncorrelated and ordered such that the first few retain most of the variation present in the original data set as shown in Figure 2.

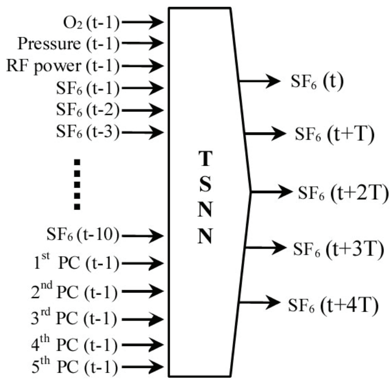


Fig. 2. I/O illustration for a time series neural network for SF_6 prediction (sampling period T=10 seconds).

In order to simulate real-time operation, in these experiments, PCA was performed on the ten most recently collected OES samples with a 10-seconds sampling period for 340 pre-selected wavelengths. Then, the first five principal components and current process parameters were simultaneously used as inputs for neural network modeling.

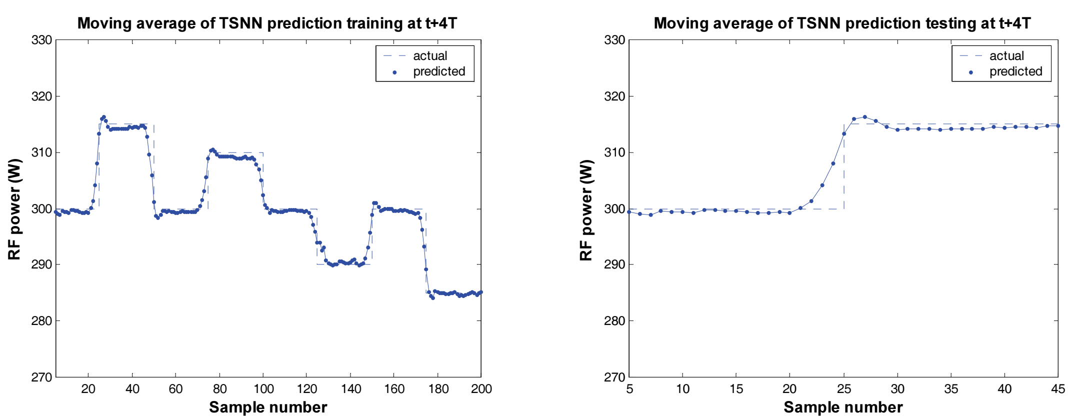
5. Neural Network Modeling

Neural networks have become useful tools in semiconductor process modeling, control and optimization. Several studies report excellent results using neural network-based techniques [8]. Time series models derived from multi-layer perceptron neural nets were employed to achieve predictive models of optical emission spectroscopy (OES) data acquired during RIE. The most relevant wavelengths were pre-selected from OES data acquired during the etching process, and the first five principal components (PCs) of this OES data, as well as time series values of the RIE process parameters were used as network inputs (Figure 3). TSNNs were trained to forecast process conditions at time *t*, as well as to predict up to 4 sampling periods in the future (*t+4T*). Separate data sets for network training and testing were established.

The performance of the TSNNs was evaluated in terms of root-mean-squared error shown in Table II. The predicted TSNN outputs were obtained using a moving average (MA) technique. The MA is a statistic used for detecting small process shifts [9]. Suppose that individual observations have been collected, and let $x_1, x_2, x_3, \cdots, x_t$ denote the observations. The moving average for a span w at time t is $M_t = \frac{x_t + x_{t-1} + x_{t-2} + \cdots + x_{t-w+1}}{w}$. At time t the oldest observation in the moving average set is dropped and the newest one added to the set. Successful prediction results were achieved in this case for w=5 (Figures 4 (a)-7 (b)).

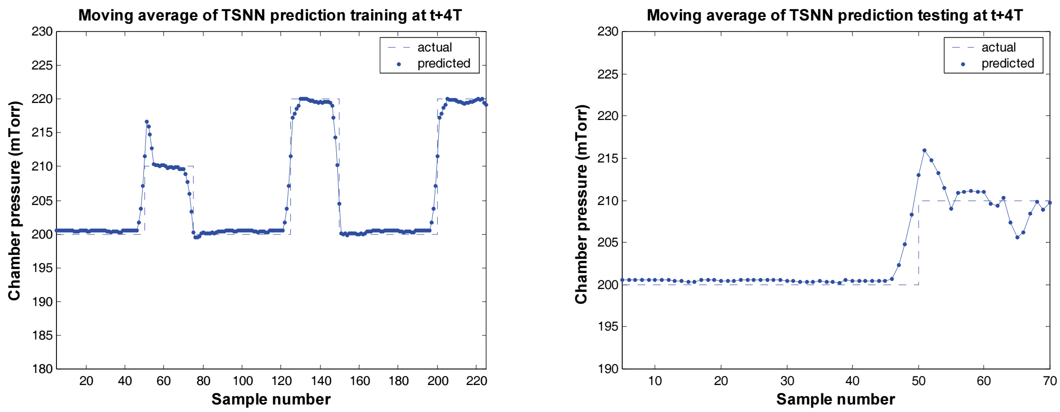
Table II. Performance Evaluation of TSNNs

	RMSE		Unit
	Training	Testing	
RF power	1.6861	1.4797	Watt
Pressure	1.4797	2.5283	mTorr
Gas 1	1.0489	1.2848	sccm
Gas 2	0.2397	0.7821	sccm



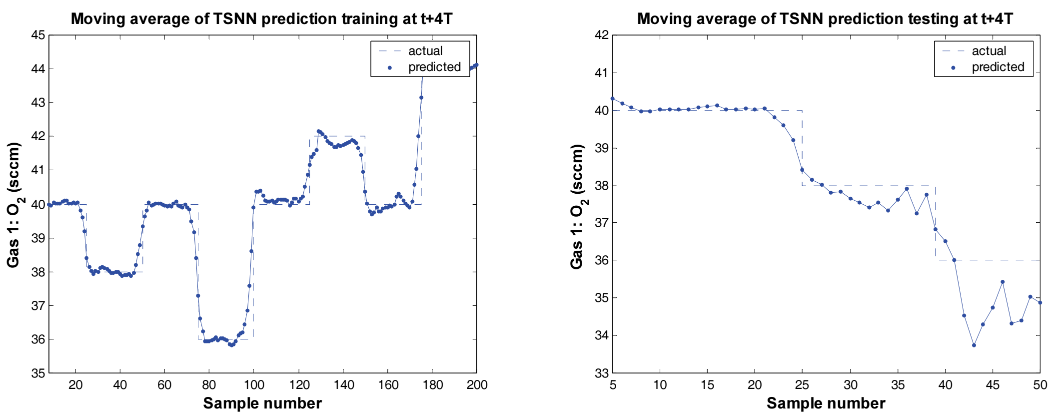
(a) (b)

Fig. 4. Moving average of RF power: (a) Training data and (b) Testing data.



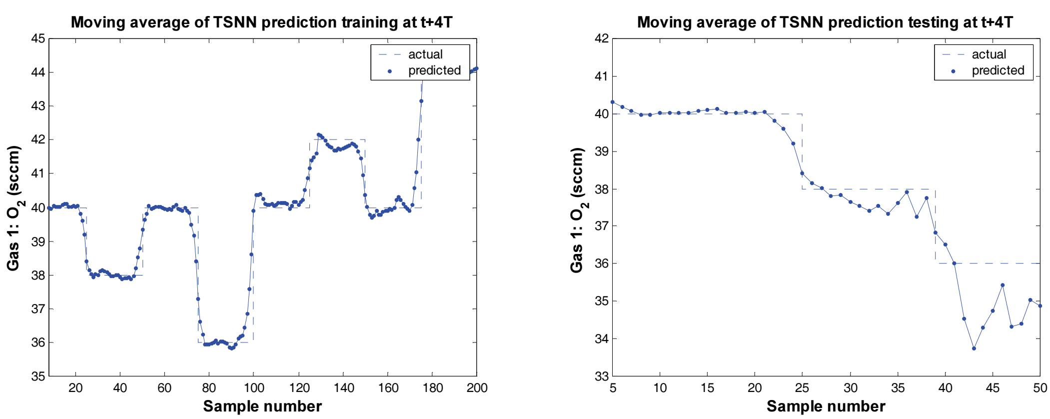
(a) (b)

Fig. 5. Moving average of chamber pressure: (a) Training data and (b) Testing data.



(a) (b)

Fig. 6. Moving average of SF6: (a) Training data and (b) Testing data.



(a) (b)

Fig. 7. Moving average of O2: (a) Training data and (b) Testing data.

6. Achievement and Conclusion

As an intermediate achievement toward real-time diagnosis and prognosis for reactive ion etching, neural network-based time series modeling of optical emission spectroscopy data has been undertaken. Principal component analysis of the OES data was employed to reduce the dimensionality of the data set. Subsequently, time series neural networks were constructed and showed the ability to predict process shifts. These ideas will be further developed for identifying and predicting the cause of RIE system malfunctions.

Acknowledgement

This work was supported by Korean Evaluation Institute of Industrial Technology (G01002665681).

References

1. D. Montgomery, *Introduction to Statistical Quality Control*, New York: John Wiley & Sons, Inc., 3rd Ed. 1997.
2. M. Baker, C. Himmel, and G. May, “Time Series Modeling of Reactive Ion Etching Using Neural Networks,” *IEEE Trans. Semi. Manufac.*, vol. 8, no. 1, pp. 62-71, Feb., 1995.
3. D. Kim, S. J. Hong, "Use of Plasma Information in Machine-Learning-Based Fault Detection and Classification for Advanced Process Control,” *IEEE Trans. Semi. Manufac.* vol. 34, no. 3, pp. 408-419, Aug., 2021.
4. J. E. Choi, H. Park, Y. H. Lee, and S. J. Hong, "Virtual Metrology for Etch Profile in Silicon Trench Etching with SF_6/O_2/Ar Plasma, *IEEE Trans. Semi. Manufac.* vol. 35, no. 35, pp. 128-137, Feb., 2022.
5. A. A. Nuhu, Q. Zeeshan, B. Safaei, and M. A. Shahzad, “Machine Learning-based Techniques for Fault Diagnosis in the Semiconductor Manufacturing Process: A Comparative Study,” *The Journal of Supercomputing,* vol. 79, pp. 2031-2081, Feb., 2023.
6. 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, pp. 210-218, 3rd Quarter, 1994.
7. 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, Feb., 1997.
8. S. Hong and G. May, “Neural Network Modeling of Reactive Ion Etching Using Principal Component Analysis of Optical Emission Spectroscopy Data,” *IEEE/SEMI Adv. Semi Manufac. Conf.*, Boston, MA, pp. 415-420, April 2002.
9. G. May, “Manufacturing ICs the Neural Way,” *IEEE Spectrum*, pp.47-51, Sept. 1994.

접수일: 2023년 11월 30일, 심사일: 2023년 12월 14일, 게재확정일: 2023년 12월 18일