Modeling of Process Plasma Using a Radial Basis Function Network: A Case Study

Byungwhan Kim and Sungjin Park

Abstract: Plasma models are crucial to equipment design and process optimization. A radial basis function network (RBFN) in conjunction with statistical experimental design has been used to model a process plasma. A 2^4 full factorial experiment was employed to characterize a hemispherical inductively coupled plasma (HICP). In characterizing HICP, the factors that were varied in the design include source power, pressure, position of chuck holder, and Cl_2 flow rate. Using a Langmuir probe, plasma attributes were collected, which include typical electron density, electron temperature, and plasma potential as well as their spatial uniformity. Root mean-squared prediction errors of RBFN are 0.409 (10^{12}/cm^3), 0.277 (eV), and 0.669 (V), for electron density, electron temperature, and plasma potential, respectively. For spatial uniformity data, they are 2.623 (10^{12}/ cm^3), 5.074 (eV), and 3.481 (V), for electron density, electron temperature, and plasma potential, respectively. Comparisons with generalized regression neural network (GRNN) revealed an improved prediction accuracy of RBFN as well as a comparable performance between GRNN and statistical response surface model. Both RBFN and GRNN, however, experienced difficulties in generalizing training data with smaller standard deviation.

Keywords: Plasma model, radial basis function network, generalized regression neural network, response surface model

I. Introduction

Plasma processing plays a critical role in either depositing thin films or etching fine patterns during integrated circuit (IC) manufacturing. Rather than relying upon simulation technique, IC processes have been mostly optimized through extensive experimentation. This mainly stems from the fact that an accurate modeling of plasma or process is extremely difficult due to the highly nonlinear particle dynamics within the plasma. Historically, the plasma has been modeled using the first principle physics involving continuity, momentum balance, and energy balance inside a high frequency, high intensity electric or magnetic field or both [1][2]. First principle models currently available attempt to derive self-consistent solutions to complex physical equations by means of computationally intensive numerical simulation methods, which typically produce distribution profiles of electrons and ions within the plasma. Although simulations are somewhat useful for equipment design and optimization, they are subject to many simplifying assumptions and thus predictions from the models often deviate much from their actual values. Since IC process simulation is based on a variety of predicted plasma attributes (such as electron density or radicals), this inconsistency is likely to yield suspicious process attributes (such as etch rate or profile). Another approach is to derive statistical response surface model (RSM), which is widely used in IC process optimization and control. Despite their popularity, meanwhile, RSM are inherently limited in that they attempt to linearize nonlinear data, thus leading to a relatively large prediction error for certain operational-parameter space.

Recently, neural networks have been applied to model various IC manufacturing processes [3][9]. The network approach differs from the RSM in that it learns complex plasma data both adaptively and nonlinearly without any linearization required in the RSM. Of neural networks, backpropagation neural network (BPNN) [10] is the most widely employed paradigm. Applicability of other networks to plasma modeling has rarely been explored. In this study, thus, a radial basis function network (RBFN) [11] is introduced for plasma modeling. RBFN models were further compared to those from generalized regression neural network (GRNN) [12] and RSM. The plasma modeled is referred to as a hemispherical inductively coupled plasma (HICP). The HICP behaviors were characterized by a 2^4 factorial experiment [13], in which the factors that were varied include source power, pressure, position of chuck holder, and chroline (Cl_2) flow rate. Plasma attributes collected with Langmuir probe include electron density, electron temperature, and plasma potential as well as their spatial uniformities.

Manuscript received: June 7, 2000., Accepted: Nov. 22, 2000.

Byungwhan Kim, Sungjin Park: Department of Electrical Engineering, Chonnam National University

※This work is financially supported by Chonnam National University.

II. Experimental data

The HICP was generated from an etch system depicted in Fig. 1. Hemispherically shaped chamber has 10 turns of coils outside it and by feeding a 13.56MHz radio frequency power to it the plasma is then created.

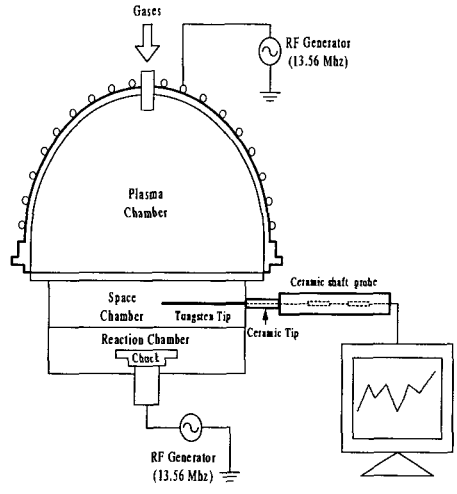

Fig. 1. A schematic diagram of inductively coupled flasma etch system.

The HICP behaviors were characterized with a 2^4 full factorial experiment. The factors that were varied in the design are contained in Table 1with their experimental ranges. Using the Langmuir probe, fundamental plasma attributes were collected and they are electron density, electron temperature, and plasma potential. The tungsten tip of the probe has a length of 6 mm and radius of 0.18 mm. The probe shaft with a 20 uH inductor placed inside acts as a filter to suppress rf fluctuations in the probe current (I), induced by fluctuations in the plasma potential (V_p). The probe was inserted radially into the plasma, and subsequently current-voltage (I-V) characteristics of the probe were acquired by sweeping the probe voltage (V) from −100 V and +100 V. The plasma potential was determined as a voltage at the intersection of two tangent lines fit to the regions below and above the characteristic " knee" in the curve. The electron temperature (T_e) was measured from the slope of In(I) versus V, where a tangent line was fitted to the voltage region less than, but close, to the potential. A formula from which the electron density (N_e) is estimated is given by:

$$N_e = \frac{I(V_p)}{A_p}\left(\frac{2\pi m_e}{e^3 k T_e}\right)^{1/2}, \quad (1)$$

where $I(V_p)$ is the probe current with the probe voltage set to the V_p, and A_p is the probe surface area; e and m_e indicate the electron charge and mass, respectively, and k denotes the Boltzman constant.

Table 1. Experimental factors for statistical design.

Factor	Range	Units
Source Power	700-900	Watts
Pressure	5-10	MTorr
Chuck Position	-30-90	Mm
Cl_2	60-120	Sccm

For each attribute and a given factor combination, a total of 35 measurements were performed across the wafer diameter and then the measured data were averaged. Each averaged attribute was used as the model output (or response). Apart from the 2^4 factorial experiment, eight experiments were additionally preformed to provide the test data for model evaluation. Consequently, a total of 24 experiments were required for model build up and evaluation. The spatial plasma uniformity was estimated using the metric defined as:

$$uniformity = \frac{X_{\max} - X_{\min}}{2 \times X_{avg}} \times 100(\%), \quad (2)$$

where X denotes individual plasma attribute, $X_{\max}$ and $X_{\min}$ indicate the maximum and minimum values among the measured data for each X. Remaining X_{avg} is the average of the measured data for each X.

III. Plasma modeling using rbfn

1. Basics of RBFN

As depicted in Fig. 2.A, RBFN is a two-layer fully connected network.

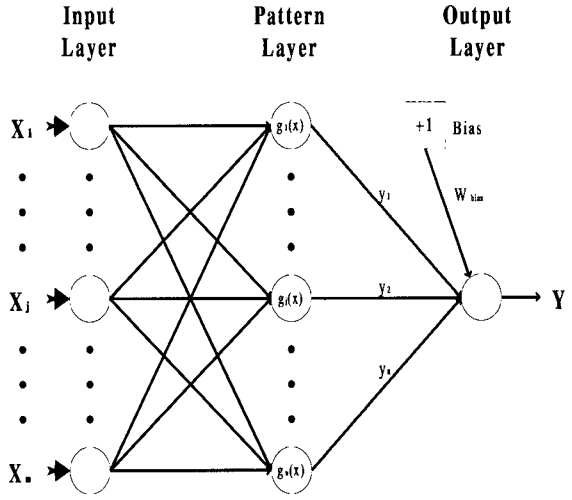

Fig. 2.A A typical architecture of radial basis function network.

Each hidden neuron is parameterized by two quantities: a center μ of a receptive field and a width parameter σ_w^2 of the field. The center corresponds to the vector defined by the weights between the node and the input nodes. The receptive field is an area in the input space that activates the hidden neurons and serves to cluster similar input vectors. In this sense, the RBFN performs a local mapping rather than a global mapping as in BPNN. As an activation function, hidden nodes employ radial basis function that is radially symmetric in input space. For the j th hidden neuron in pattern layer, the activation is defined as:

$$g_j(x) = \exp[-\frac{(x-\mu_j)^2}{\sigma_w^2}]. \quad (3)$$

The output layer is a layer of linear neurons and transforms linearly the hidden neuron outputs. Prediction from RBFN is thus computed as:

$$\hat{y}_i = w_{bias} + \sum_j y_j g_j(x). \quad (4)$$

In this work, the RBFN coded originally by MATLAB was utilized and modified to accommodate the plasma data given. Here, the width parameter was set to 0.8326/spread, where the spread was experimentally adjusted to optimize model prediction performance.

2. Basics of GRNN

A schematic diagram of GRNN is depicted in Fig. 2.B. The GRNN consists of four layers: input layer, pattern layer, summation layer and output layer. The number of input units in the first layer is equal to independent factors (x_i), which is four here. The first layer is fully connected to the second, pattern layer, where each unit represents a training pattern and its

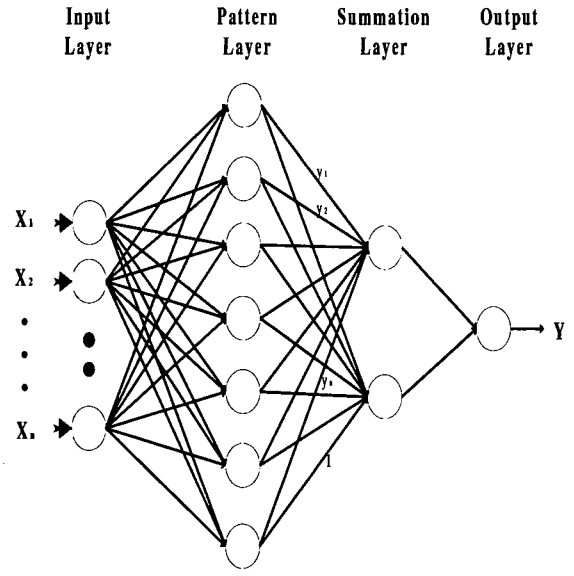

Fig. 2. B A typical architecture of generalized regression neural network.

output is a measure of the distance of the input from the stored patterns. Each pattern layer unit is connected to the two neurons in the summation layer: S-summation neuron and D-summation neuron. The S-summation neuron computes the sum of the weighted outputs of the pattern layer while the D-summation neuron calculates the unweighted outputs of the pattern neurons. The connection weight between the i th neuron in the pattern layer and the S-summation neuron is y_i, the target output value corresponding to the i th input pattern. For D-summation neuron, the connection weight is unity. The output layer merely divides the output of each S-summation neuron by that of each D-summation neuron, yielding the predicted value expressed as:

$$\hat{y}_i = \frac{\sum_{i=1}^{n} y_i \exp[-D(x,x_i)]}{\sum_{i=1}^{n} \exp[-D(x,x_i)]}, \tag{5}$$

where n indicates the number of independent input variables and the Gaussian D function is defined as:

$$D(x,x_i) = \sum_{j=1}^{p} \left(\frac{x_j - x_{ij}}{\sigma_j}\right)^2, \tag{6}$$

where p indicates the total number of training patterns and the σ is generally referred to as the smoothing parameter, whose optimal value is often determined experimentally. Besides RBFN and GRNN, RSM was also constructed for a comparison, which is typically expressed as:

$$y = \beta_o + \sum_{i=1}^{k} \beta_i x_i + \sum_{i}^{k} \beta_{ii} x_i^2 + \sum_{i} \sum_{j} \beta_{ij} x_i x_j, \tag{7}$$

where y is a plasma attribute, β_i and β_{ij} are regression coefficients, and x_i is a regressor variable representing a factor in Table I. An index k denotes the total number of the factors, four in this study.

IV. Performance evaluation

1. Model development

Plasma attributes were modeled individually and thus a total of six RBFN models were constructed. Each RBFN was trained on the sixteen experiments obtained from the 2^4 full factorial design. Prediction performances of trained models were subsequently tested on the eight experiments. To optimize model performance, the spread parameter was incrementally adjusted until the root mean-squared error (*RMSE*) is minimized. Here the *RMSE* is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{n-1}}, \tag{8}$$

where n is the size of the test set, y_i is the measured value of the plasma attribute, and $\hat{y}_i$ is the prediction from the model. In Fig. 3, prediction errors of electron density model are depicted as a function of the parameter. As indicated in Fig. 3, one best error is obtained at the spread parameter 1.05 and the corresponding *RMSE* is 0.409. For the other electron temperature and plasma potential including electron density, optimized parameters for RBNN are contained in Table 2 along with those for GRNN. Meanwhile, the regression coefficients in (7) were estimated using the SAS statistical package and those estimated ones are included in Table 3.

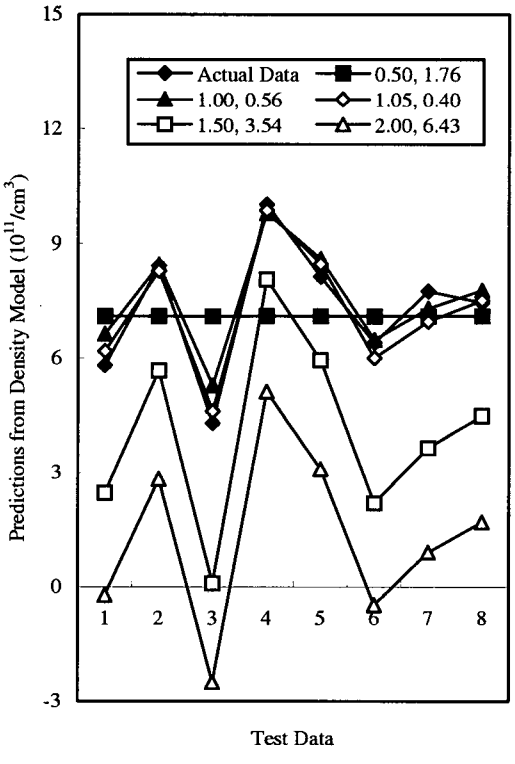

Fig. 3. Predictive performance of RBFN density model with increasing the spread parameter.

Table 2. Spread and smoothing factors optimized for RBFN and GRNN.

Plasma Attribute	RBFN	GRNN
Electron Density (10^{12}/cm^3)	1.05	0.50
Electron Temperature (eV)	1.13	0.50
Plasma Potential (V)	1.14	0.40

2. Comparison of predictive models

Figure 4 exhibits predictions obtained from density models of RBFN, GRNN, and RSM.As illustrated in Fig. 4, the RBFN provides better estimates over GRNN and RSM. This improvement is further supported by the prediction errors shown in Fig. 4. An improvement of more than 17% is demonstrated for the RBFN over GRNN and RSM. Of interest is that both predictive capabilities of GRNN and RSM are identical. This also appears in modeling electron temperature as in Fig. 5. In this case, about 6% improvement is achieved for RBFN over GRNN and RSM. Figure 6 compares models derived for plasma potential. All the models demonstrate performances commensurable. Spatial plasma data were additionally modeled and resultant prediction errors are contained in Table 4. Table 4 reveals an improvement of RBFN over GRNN and RSM. From Figs. 4-6 and Table 4, it can generally be determined that RBFN is the best predictor over the other models. Another attempt made was to correlate model prediction capability to the training data, which may facilitate choosing a model suitable for a data with an arbitrary distribution. This was accomplished by quantifying model prediction error with another R^2 metric, defined as the square of correlation coefficients while characterizing the training data via statistical mean. Resultant correlations are shown in Fig. 7. For the data with smaller standard deviation, each model exhibits relatively smaller R^2 values, implying that all the models experience difficulties in generalizing the data with smaller standard deviation. This undesirable situation seems to improve as the deviation becomes larger.

Table 3. Estimated regression coefficients of RSM.

Regression Coefficients	Electron Density (10^{11}/cm^3)	Electron Temperature (eV)	Plasma Potential (V)
β_o	1.812	0.708	2.832
β_1	0.812E-2	0.205E-2	0.150E-1
β_2	0.231E-1	-0.132E-1	-0.295E-1
β_3	-0.268	0.364	1.800
β_4	-0.205E-1	-0.468E-2	-0.175E-1
β_{12}	0	0	0
β_{13}	0.338E-4	0.177E-4	0.480E-4
β_{14}	0	0	0
β_{23}	-0.212E-3	-0.405E-3	-0.229E-2
β_{24}	-0.166E-2	0.833E-5	0.190E-2
β_{34}	0	0	0
β_1^2	0.364E-4	0.541E-5	0.160E-4
β_2^2	0.809E-4	0.118E-4	-0.204E-4
β_3^2	-0.208E-3	-0.450E-3	-0.592E-3
β_4^2	0	0	0

Table 4. RMS prediction errors of spatial uniformity (U) Data.

Attributes	RBNN	GRNN	RSM
Ne(U) (10^{12}/cm^3)	2.623	2.89	3.17
Te (U) (eV)	5.074	5.92	5.92
Vp (U) (V)	3.481	3.58	3.57

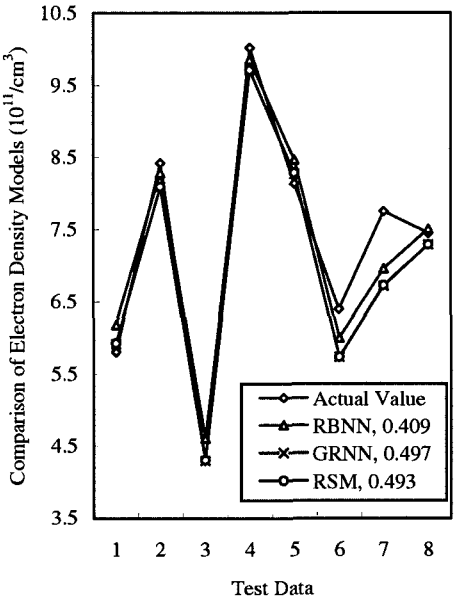

Fig. 4. Prediction error comparison of electron density model.

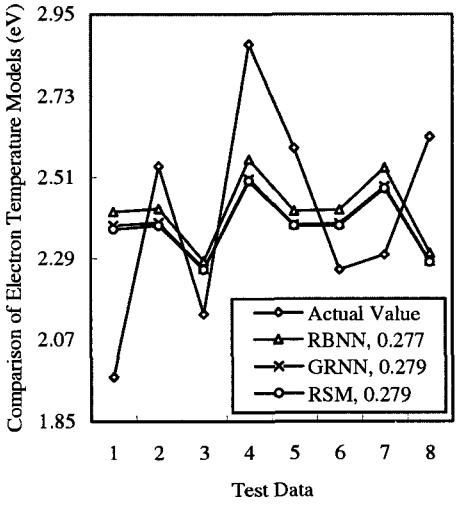

Fig. 5. Prediction error comparison of electron temperature model.

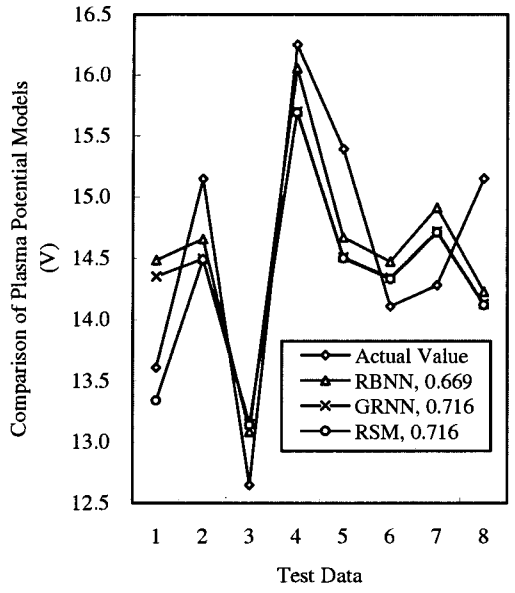

Fig. 6. Prediction error comparison of plasma potential model.

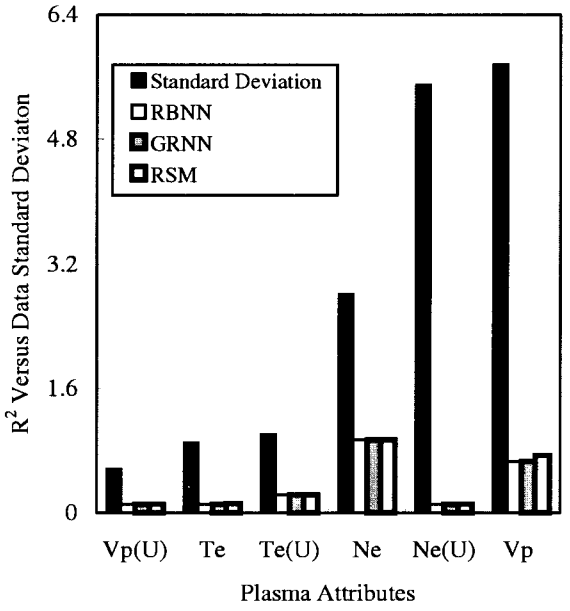

Fig. 7. Correlation between model data fit performance and statistical data distribution.

V. Conclusion

Radial basis function network was used to model a hemispherical inductively coupled plasma (HICP). The HICP was characterized by a 2^4 full factorial experiment, the factors that were varied in the design include source power, pressure, position of chuck holder, and Cl_2 flow rate. For a comparison, other generalized regression neural network and statistical response models were further constructed. A total of six plasma attributes were modeled and compared to each other to examine their relative advantages from the standpoint of prediction accuracy. Performances of Both RBFN and GRNN were optimized by incrementally adjusting related control parameters. Comparisons revealed an improvement of RBFN over the others as well as a comparable performance between GRNN and RSM. This implies that RSM behaviors can be estimated by GRNN. A correlation study uncovered a drawback common to all the models that they are less effective in generalizing the data with relatively larger standard deviation.

References

[1] P. L. G. Ventzek, T. J. Sommerer, R. J. Hoekstra, and M. J. Kushner, "Two-dimensional hybrid model of inductively coupled plasma sources for etching," *Appl. Phys. Lett.*, 63 (5), 1994.

[2] R. A. Stewart, P. Vitello, and D. B. Graves, "Two-dimensional fluid model of high density inductively coupled plasma sources," *J. Vac. Sci. Technol.*, B 12(1), pp. 478-485, 1994.

[3] C. D. Himmel, G. S. May, "Advantages of plasma etch modeling using neural networks over statistical techniques," *IEEE Trans. Semicond. Manufact.*, vol. 6, pp. 103-111, 1993.

[4] E. A. Rietman, E. R. Lory, "Use of neural networks in modeling semiconductor manufacturing processes: An example of plasma etch modeling," *IEEE Trans. Semicond. Manufact,*," vol. 6, no. 4, pp. 343-347, 1993.

[5] B. Kim, G. S. May, "An optimal neural network process model for plasma etching," *IEEE Trans. Semicond. Manufact.*," vol. 7, no. 1, pp. 12-21, 1994.

[6] B. Kim, G. S. May, "Reactive ion etch modeling using neural networks and simulated annealing," *IEEE Trans. Comp., Pack.. Manufact. Technol., C*, vol. 19, no. 1, pp. 3-8, 1996.

[7] B. Kim, K. H. Kwon, "Modeling a magnetically RIE of aluminum alloy films using neural networks," *IEEE Trans. Semicond. Manufact.*, vol. 11, no. 4, pp. 692-95, 1998.

[8] B. Kim, K. H. Kwon, and S. H. Park, "Characterizing metal-masked silica etch process in a CHF_3/CF_4 inductively coupled plasma," *J. Vac. Sci. and Technol., A,* vol. 17, no. 5, pp. 2593-2597, 1999.

[9] B. Kim, J. H. Sun, C. J. Choi, D. D. Lee, and Y. S. Seol, "Use of neural network to model low-temperature tungsten etch characteristics in high density SF_6 plasma," *J. Vac. Sci. and Technol., A,* vol. 18, no. 2, pp. 417-422, 2000.

[10] J. Freeman, D. Skapura, *Neural Networks*, Addison Wesley, New York, 1991.

[11] S. Chen, C. F. N. Cowan, and P. M. Grant, "Orthogonal least squares learning algorithm for radial basis function networks," *IEEE Trans. Neural Networks*, vol. 2, no. 2, pp. 302-309, 1991.

[12] P. D. Wasserman, *Advanced Methods in Neural Computing*, New York: Van Nostrand Reinhold, pp. 151-61, 1993.

[13] G. Box, N. Draper. *Empirical Model Building and Response Surfaces*, New York: Wiley, 1987.

Byungwhan Kim

He was born on October 30, 1962. He received the B.S. and M.S. degrees in electrical engineering from Korea University in 1985 and 1987, and the Ph.D. degree from the Georgia Institute of Technology in 1995, respectively. From 1996 to 1998, he served as principal technical staff at Hyundai Electronics. Following a Post-Doctor at Korea University from 1998 to 1999, he is currently a full time instructor in the department of electrical engineering at Chonnam National University. His research fields include intelligent modeling, optimization, diagnosis and control of semiconductor manufacturing processes.

Sungjin Park

He was born on May 20, 1975. He received the B.S. from the University of SuNam in 1999. He is currently a candidate for the M.S. degree in the department of electrical engineering at Chonnam National University. His research interest is an intelligent process modeling and optimization.