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\textsubscript{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\textsuperscript{12}/cm\textsuperscript{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\textsuperscript{12}/cm\textsuperscript{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 chlorine (Cl\textsubscript{2}) flow rate. Plasma attributes collected with Langmuir probe include electron density, electron temperature, and plasma potential as well as their spatial uniformities.

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
The HICP behaviors were characterized with a $2^4$ full factorial experiment. The factors that were varied in the design are contained in Table 1 with 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 ($eT_e$) was measured from the slope of $\ln(I)$ 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 ($eN_e$) is estimated is given by:

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

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.

<table>
<thead>
<tr>
<th>Factor</th>
<th>Range</th>
<th>Units</th>
</tr>
</thead>
<tbody>
<tr>
<td>Source Power</td>
<td>700-900</td>
<td>Watts</td>
</tr>
<tr>
<td>Pressure</td>
<td>5-10</td>
<td>MTorr</td>
</tr>
<tr>
<td>Chuck Position</td>
<td>-30-90</td>
<td>Mm</td>
</tr>
<tr>
<td>Cl$_2$</td>
<td>60-120</td>
<td>Sccm</td>
</tr>
</tbody>
</table>

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 performed 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:

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

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 $X_{\text{avg}}$ 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.

- **Input Layer**
  - $x_1$ to $x_i$
  - $bias$

- **Pattern Layer**
  - $g(x)$
  - $g_j(x)$

- **Output Layer**
  - $y_1$ to $y_i$
  - $w_{bias}$

#### 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

\[ y_i = \frac{\sum y_j \exp[-D(x_i, x_j)]}{\sum \exp[-D(x_i, x_j)]}, \]  

(5)

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

\[ D(x_i, x_j) = \sum_{j=1}^{p} \left( \frac{x_i - x_j}{\sigma} \right)^2, \]  

(6)

where \( p \) indicates the total number of training patterns and the \( \sigma \) 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:

\[ y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i} \beta_i x_i^2 + \sum_{i,j} \beta_{ij} x_i x_j, \]  

(7)

where \( y \) is a plasma attribute, \( \beta_i \) and \( \beta_{ij} \) 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:

\[ \text{RMSE} = \sqrt{\frac{1}{n-1} \sum_{i=1}^{n} (y_i - \hat{y}_i)^2}, \]  

(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.
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 means. 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.

<table>
<thead>
<tr>
<th>Regression Coefficient (s)</th>
<th>Electron Density (10^{11}/cm^3)</th>
<th>Electron Temperature (eV)</th>
<th>Plasma Potential (V)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \beta_0 )</td>
<td>1.812</td>
<td>0.708</td>
<td>2.832</td>
</tr>
<tr>
<td>( \beta_1 )</td>
<td>0.812E-2</td>
<td>0.205E-2</td>
<td>0.150E-1</td>
</tr>
<tr>
<td>( \beta_2 )</td>
<td>0.231E-1</td>
<td>-0.132E-1</td>
<td>-0.295E-1</td>
</tr>
<tr>
<td>( \beta_3 )</td>
<td>-0.268</td>
<td>0.364</td>
<td>1.800</td>
</tr>
<tr>
<td>( \beta_4 )</td>
<td>-0.205E-1</td>
<td>-0.468E-2</td>
<td>-0.175E-1</td>
</tr>
<tr>
<td>( \beta_{12} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \beta_{13} )</td>
<td>0.338E-4</td>
<td>0.177E-4</td>
<td>0.480E-4</td>
</tr>
<tr>
<td>( \beta_{14} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \beta_{23} )</td>
<td>-0.212E-3</td>
<td>-0.405E-3</td>
<td>-0.229E-2</td>
</tr>
<tr>
<td>( \beta_{24} )</td>
<td>-0.166E-2</td>
<td>0.833E-5</td>
<td>0.190E-2</td>
</tr>
<tr>
<td>( \beta_{34} )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>( \beta_1^2 )</td>
<td>0.364E-4</td>
<td>0.541E-5</td>
<td>0.160E-4</td>
</tr>
<tr>
<td>( \beta_2^2 )</td>
<td>0.809E-4</td>
<td>0.118E-4</td>
<td>-0.204E-4</td>
</tr>
<tr>
<td>( \beta_3^2 )</td>
<td>-0.208E-3</td>
<td>-0.450E-3</td>
<td>-0.592E-3</td>
</tr>
<tr>
<td>( \beta_4^2 )</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Attributes</th>
<th>RBNN</th>
<th>GRNN</th>
<th>RSM</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ne(U) (10^{12}/cm^3)</td>
<td>2.623</td>
<td>2.89</td>
<td>3.17</td>
</tr>
<tr>
<td>Te (U) (eV)</td>
<td>5.074</td>
<td>5.92</td>
<td>5.92</td>
</tr>
<tr>
<td>Vp (U) (V)</td>
<td>3.481</td>
<td>3.58</td>
<td>3.57</td>
</tr>
</tbody>
</table>
Fig. 4. Prediction error comparison of electron density model.

Fig. 5. Prediction error comparison of electron temperature model.
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 Cl₂ 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

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.