# Fault Detection, Diagnosis, and Optimization of Wafer Manufacturing Processes utilizing Knowledge Creation

**Hyeon Bae, Sungshin Kim*, Kwang-Bang Woo, Gary S. May, and Duk-Kwon Lee**

**Abstract:** The purpose of this study was to develop a process management system to manage ingot fabrication and improve ingot quality. The ingot is the first manufactured material of wafers. Trace parameters were collected on-line but measurement parameters were measured by sampling inspection. The quality parameters were applied to evaluate the quality. Therefore, preprocessing was necessary to extract useful information from the quality data. First, statistical methods were used for data generation. Then, modeling was performed, using the generated data, to improve the performance of the models. The function of the models is to predict the quality corresponding to control parameters. Secondly, rule extraction was performed to find the relation between the production quality and control conditions. The extracted rules can give important information concerning how to handle the process correctly. The dynamic polynomial neural network (DPNN) and decision tree were applied for data modeling and rule extraction, respectively, from the ingot fabrication data.

**Keywords:** Data mining, data model, knowledge creation, process optimization, rule extraction, wafer fabrication.

## 1. INTRODUCTION

Wafer is an important material in semiconductor industries. In recent years, the size of wafers has been enlarged up to 300 mm, so that management is essentially required and applied. The wafer manufacturing process includes certain chemical processes; there is a time delay that causes difficulty in measurement and control. Among the chemical processes, ingot fabrication is the most important, as the quality of the ingot will definitely affect the quality of the wafer.

Over decades, many studies have been performed for fault detection and yield improvement. An adaptive resonance theory network [1] was used to develop an intelligent system that will be able to recognize defect spatial patterns to aid in the diagnosis of failure causes. A data warehouse approach to the automation of process zone-by-zone defect-limited yield analysis [2], and SOI wafer-specific behavior related to the intrinsic limitations of laser-scattering defect detection [3], was presented. The modeling of wafer fabrication was carried out, with the calculations and results of random defect-limited yield (DLY) using the deterministic yield model [4-7]. An advanced methodology using intentionally created defect arrays was implemented to enhance the understanding of defect detection tools [8].

However, in past research, the relations between qualities and control conditions have not been ascertained for the topic. In this study, data mining methods were applied to extract and gather the information from the process data. The results can improve the yield and quality of the wafer products.

It is difficult to select a proper method from various data mining methodologies. In this study, the noble data mining roadmap was proposed to assist in the selection of an appropriate methodology. Based on the roadmap, the selected methodologies were the data models to predict process quality.

After selecting the method and procedure, data acquisition from the target process is used in data mining; in addition, the collected data should be

Manuscript received February 28, 2005; revised January 4, 2006; accepted February 10, 2006. Recommended by Editor Keum-Shik Hong. This paper was supported by Research Institute of Computer, Information and Communication (PNU) and Pusan National University in the program, Post-Doc. 2006.

Hyeon Bae and Sungshin Kim are with the School of Electrical and Computer Engineering, Pusan National University, 30 Jangjeon-dong, Geumjeong-gu, Busan 609-735, Korea (e-mails: {baehyeon, sskim}@pusan.ac.kr).

Kwang-Bang Woo is with Automation Technology Research Institute, Yonsei University, 134 Shinchon-dong, Seodaemun-gu, Seoul 120-749, Korea (e-mail: kbwoo@yonsei.ac.kr).

Gary S. May is with the School of Electrical and Computer Engineering, Georgia Institute of Technology, Atlanta, GA 30332-0250, U.S.A. (e-mail: gmay@ece.gatech.edu).

Duk-Kwon Lee is with the Dept. of Computer System, MEMC Korea Co., Chungnam 330-831, Korea (e-mail: dklee@memc.co.kr).

* Corresponding author.

sufficient in number and clean enough to perform the data mining. The data on the quality of the wafer, prepared for this paper, was not sufficient because quality evaluation was performed according to a sampling inspection, not a total inspection. To solve these problems, the bootstrap method, an appropriate data preprocessing method, was proposed to generate data sufficient for a total inspection. This is a new approach in terms of the industrial applications.

The final objectives of this study were focused on detecting faults, adjusting the operational conditions for process optimization and producing wafers having no defects. To detect a fault, data mining tools to analyze input-output data using models are required.

In Section 2, the target process is described. In Section 3, important results such as the proposed road map for data mining are explained. Section 4 explains the applied data mining techniques. Section 5 shows the experimental results. Finally, Section 6 concludes the paper.

## 2. WAFER FABRICATION

### 2.1. Wafers for semiconductors

Wafers are used in manufacturing memory or non-memory semiconductor chips. Several circuit masks are mounted on one wafer by UV rays or electron beams in assembly lines. As semiconductor technology has developed, the wafer size has been enlarged to mount more circuits on the wafer. In order to enlarge the capacity of memory and non-memory chips, larger-diameter wafers and strict quality assessment are required from wafer manufacturers. To cope with these requirements, optimization of wafer fabrication is essential.

Wafer fabrication processes consist of crystal growth, wafer slicing, wafer polishing and cleaning, and epitaxial deposition. Some factors of the wafer fabrication process cause defects. Nevertheless, it is difficult to return and maintain the optimal solutions for a given process condition, because real-time analysis cannot be achieved in wafer fabrication. In this study, we develop a management system and evaluate its performance, especially as it analyzes process data related to yield and quality in wafer manufacturing and also as it controls the operating parameters based on the process status.

### 2.2. Quality inspection of wafers

Ingot is the first material manufactured in wafer fabrication. In ingot fabrication, the set-points for handling of the position or rotation of ingots and the control parameters are adjusted for quality management. These operating parameters play an important role in wafer quality and size control. Thus, proper handling of the parameters is essential for improvement of productivity and yield. The operating parameters are used as inputs in modeling and rule extraction. The quality parameters consist of five concentration values, and six defect values. Four of these were used for outputs in modeling and rule extraction for the present study.

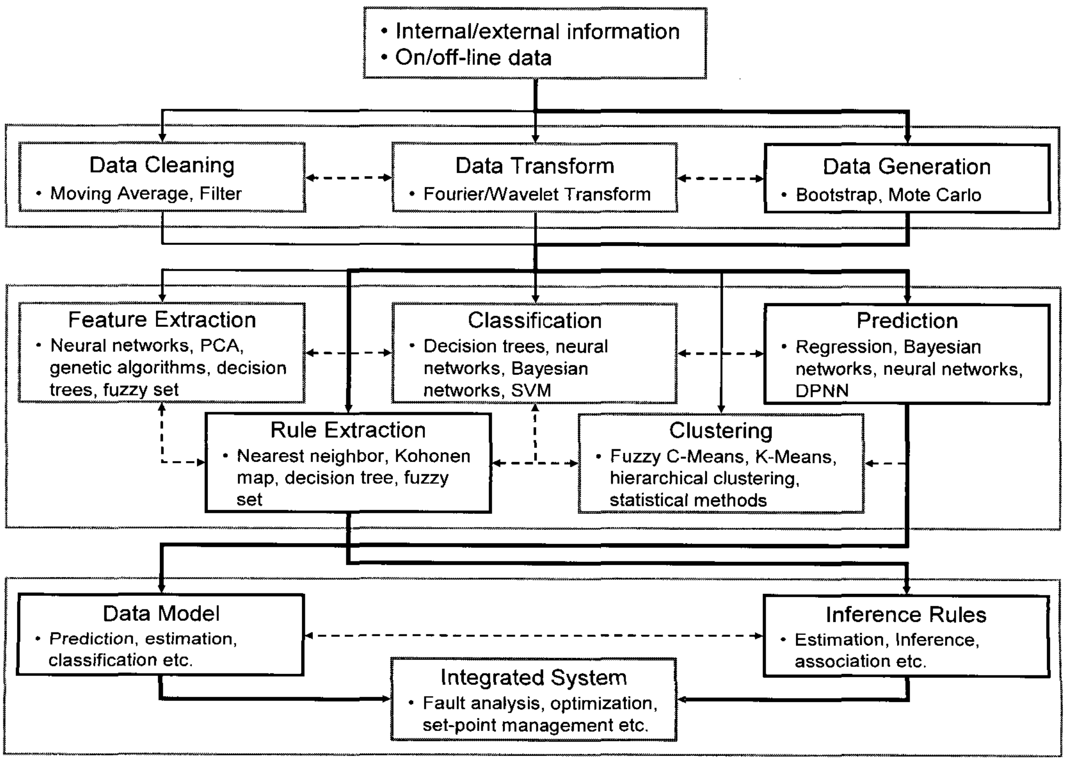



Fig. 1. Data mining roadmap proposed in this study corresponding to categories.

## 3. DESIGN OF DATA MINING ROADMAP

3.1. Data mining roadmap [9-15]

Many data mining techniques can be applied in knowledge extraction. Not all of the steps are used in data mining. Selection of the techniques depends on data features and mining targets. Therefore, selection of proper mining techniques is very important for reliable results.

In this study, we proposed a roadmap for data mining. Fig. 1 shows the proposed roadmap. A selection was made with reference to the methods and procedures for diagnosis and optimization of the ingot process by referring to the roadmap. The selected methods were data generation with the bootstrap method, prediction modeling of a dynamic polynomial neural network, and rule extraction by decision tree. Data generation was used for data preprocessing, prediction modeling was applied to predict the quality of wafers, and rule extraction was implemented to generate causal relationship rules for the control parameters.

3.2. Application of data mining

3.2.1 Data preprocessing in reducing data effects

The collected data from assembly lines may be confined to specific cases; thus, the quality data are not always uniformly distributed. Insufficient data results in unreliable model prediction. It may be difficult to extract rules that encompass the general case with limited data. Low-frequency data can be traced to error values in rule extraction. In order to solve these problems, data preprocessing is required by adding data and improving performance. In this study, the Bootstrap method, which is a type of Monte Carlo method, and multiple regression models were applied to compensate for leakage data caused by sampling inspection. This stage of data generation is part of the roadmap.

3.2.2 Data modeling in quality prediction

In prediction models, inputs can affect the performance of the models. Selection of inputs corresponding to data characteristics is necessary to improve model performance, as unnecessary inputs may strongly influence on prediction results. Therefore, the principal inputs that greatly influence model accuracy were selected. For evaluation of the function, the dynamic polynomial neural network (DPNN) was used. It has an advantage in that it requires only minimal computation and is very useful in modeling with high-dimension variables and a large amount of data. This method can also select essential inputs through the modeling stages, so that it may be able to improve accuracy of models. This stage of prediction modeling is a part of the roadmap.

3.2.3 Rule extraction in tuning parameters for quality

Following the quality prediction, the control parameters had to be adjusted in order to improve quality. The control parameter adjustment was performed by their causal relation with respect to the quality. In general, experienced operators tune the control parameters based on know-how in the physical

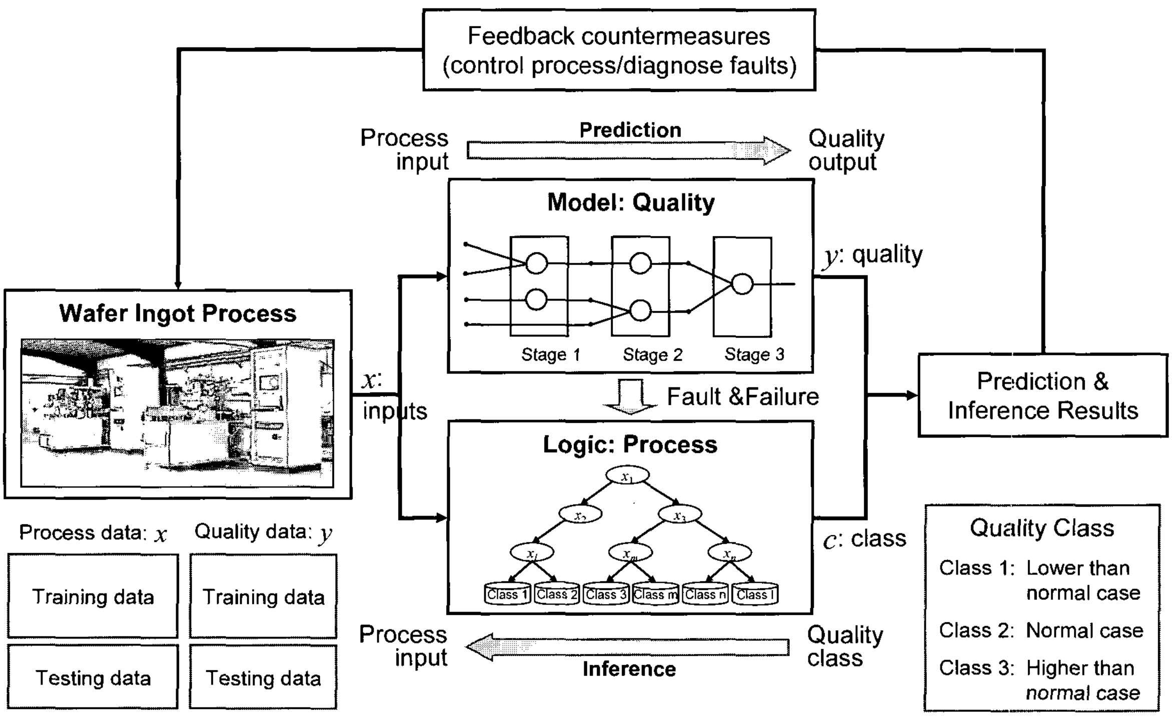



Fig. 2. Structure of the proposed process management system with model- and rule-based systems.

field. In this study, we wanted to extract rules using parameter tuning without having to acquire expert knowledge. We applied the decision tree algorithm to rule extraction. The control conditions causing the current quality could be conversely inferred based on the extracted rules. Therefore, the quality can be managed by inference rules. This stage of rule extraction was the final part of the roadmap.

3.3. Process management system in ingot fabrication

The models designed and the rules extracted were integrated into the proposed process management system. This system will play an important quality management role in ingot manufacturing. The quality will be predicted by models and the control parameters will be modified by rules on-line. The final system is shown in Fig. 2.

## 4. APPLICATION OF DATA MINING TOOLS

The process data has two characteristics. First, trace data for control parameters are collected by real-time measurement, but measurement data for quality parameters are measured by sampling inspection after manufacturing. Therefore, input and output data cannot be one-to-one correspondent and target data are insufficient. Second, quality data are included in three cases. The problem of insufficient data results in inadequate performance of the model. The bootstrap method with data generation is then used. It is then followed by construction of the prediction model using the DPNN. The concentrated data extract unreasonable rules in knowledge extraction. A multiple regression model was used for data generation in this case. Using the bootstrap method, there is no relation between two variables.

4.1. Bootstrap method

The bootstrap method was presented by Efron and Tibshirani. In this study, the term "bootstrap" refers to a Monte Carlo simulation that treats the original sample as a pseudo-population or as an estimate of the population where no parametric assumptions are made about the underlying population that generated the random sample. Instead, we use the sample as an estimate of the population.

4.2. Dynamic Polynomial Neural Network (DPNN)

A polynomial neural network (PNN) based on the GMDH algorithm is of value to model a system from many observed data and input variables. It is widely used for modeling of dynamic systems, prediction, and artificial intelligent control because of its advantages in data handling. Fig. 3 includes the recurrent inputs with one-to-$n$ time-delayed output variables. Thus, this type of PNN is called DPNN [16-18]. The model of four-input-variables with the outputs of node 13 and node 15 are represented as

$$\begin{aligned} y_{13}&=f_1(x_1,x_2)=\omega_{01}+\omega_{11}x_1+\omega_{21}x_2+\omega_{31}x_1x_2+\omega_{41}x_1^2+\omega_{51}x_2^2, \\ y_{15}&=f_2(x_3,x_4)=\omega_{02}+\omega_{12}x_3+\omega_{22}x_4+\omega_{32}x_3x_4+\omega_{42}x_3^2+\omega_{52}x_4^2, \end{aligned} \tag{1}$$

where $y_{ij}$ is the $j$th node of the $i$th layer. The final output $z$ is represented by the polynomial equation of $y_{13}$ and $y_{15}$ as

$$z=f_3(y_{13},y_{15})=\omega_{03}+\omega_{13}y_{13}+\omega_{23}y_{15}+\omega_{33}y_{13}y_{15}+\omega_{43}y_{13}^2+\omega_{53}y_{15}^2. \tag{2}$$

The least square estimator (LSE) is applied to estimate the parameters at each node to minimize the objective function. If there is $m$ amount of data, the output equations at each node are expressed as

$$\begin{bmatrix} y_{ij}^1 \\ y_{ij}^2 \\ \vdots \\ y_{ij}^m \end{bmatrix} = \begin{bmatrix} x_{0(1)} & x_{1(1)} & x_{2(1)} & x_{1(1)}x_{2(1)} & x_{1(1)}^2 & x_{2(1)}^2 \\ x_{0(2)} & x_{1(2)} & x_{2(2)} & x_{1(2)}x_{2(2)} & x_{1(2)}^2 & x_{2(2)}^2 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ x_{0(m)} & x_{1(m)} & x_{2(m)} & x_{1(m)}x_{2(m)} & x_{1(m)}^2 & x_{2(m)}^2 \end{bmatrix} \begin{bmatrix} \omega_0 \\ \omega_1 \\ \vdots \\ \omega_5 \end{bmatrix}, \tag{3}$$

where $x_0=1$, $y_{ij}=\Phi\omega$, and $y_{ij}$ is the $j$th node of the $i$th layer.

Statistical learning networks have no loops. The network is a tree of interconnected functions that implement single input/output function. Several composition schemes for network functions and corresponding estimation algorithms are described in [19]. The parameters are estimated by

$$J(\boldsymbol{\omega}) = \| y_{ij} - \boldsymbol{\Phi\omega} \|^2 = \sum_{k=1}^{m} \left[ y_{ij}^k - \sum_{l=0}^{5} \omega_l \phi_{kl+1} \right]^2, \tag{4}$$

$$\boldsymbol{\omega} = (\boldsymbol{\Phi}^T\boldsymbol{\Phi})^{-1}\boldsymbol{\Phi}^T y_{ij}. \tag{5}$$

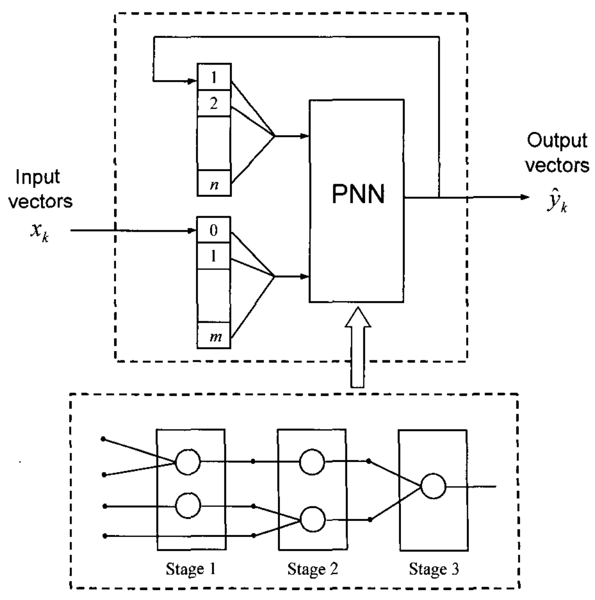



Fig. 3. Basic structure of DPNN.

### 4.3. Decision tree

Learning by decision tree is a method to approximate discrete-valued target functions, in which a decision tree represents the learned function. Learned trees can be used for representing sets of if-then rules to improve human readability. This learning method is popular among the inductive inference methods [20]. Decision trees classify instances by sorting them down the tree from the root to some leaf node, which provides the classification of the instance.

Entropy and information gain are used at classification. In order to separate each attribute, the entropy of each attribute has to be assigned to the given data set. The entropy, a measure of the impurity in a collection of training examples, was calculated according to equation (7), while $p_i$ is the proportion of $S$ belonging to class $i$. The attributes were divided into $c$ classes with minimum entropy.

$$Entropy(S) \equiv \sum_{i=1}^{c} -p_i \log_2 p_i \,. \tag{6}$$

## 5. EXPERIMENTAL RESULTS

### 5.1. Trace and measurement data of ingots

The collected data from ingot fabrication on the factory assembly lines were provided by MEMC Korea Co. There are several hundred parameters but 14 trace parameters and 11 measurement parameters were important in terms of quality analysis of the processes. The trace parameter data was collected online but the measurement parameter data was gathered by sampling inspection.

The left column of Table 1 shows the puller trace data. Forty-eight process parameters are collected by one data set per one minute from pullers. Among the parameter data, 18 important parameter values are stored in the database and used for process analysis. The position of the trace data represents the value of the wafer position (body length) of the measurement data. The second column of Table 1 shows the puller measurement data collected off-line. The data was measured by sampling inspection after slicing the ingot into wafers, and indicates the ingot growth-related parameters. The wafer position corresponds to the position of the trace data. The trace parameter values were gathered by online measurement. The problem of insufficient data exists in modeling or at the stage of rule extraction. In this study, one set of puller data was added with data generated at the preprocessing stage at which the number of the target data can be the same as that of the input data. Fig. 4 shows the data interpolation.

Table 1. Trace and measurement parameters of wafer.

| Trace parameter | Measurement parameter |
|---|---|
| OBSERV_TIME | POSITION |
| POSITION | **OXYGEN** |
| **SD_ROT_SET** | **ORG** |
| **SD_ROTATION** | **RES** |
| **SD_LIFT_SET** | **RRG** |
| **SD_LIFT** | SPV |
| **CR_LIFT** | D_DEFECT |
| **CR_POSITION** | I_DEFECT |
| **CR_ROTATE** | OISF |
| **CZ_DIA** | SWIRL |
| **CZ_DIA_SET** | SLIP |
| **AR_GAS_FLOW** | LLPD |
| **CHAMB_PRESS** | |
| **UP_MAG_LOAD** | |
| **LO_MAG_LOAD** | |
| **HEAT_POWER** | |

| Trace data (control conditions) | | | | | | | | | | | Measure data (product qualities) | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| x1 | x2 | x3 | x4 | x5 | x6 | x7 | x8 | x9 | x10 | x11 | y1 | y2 | y3 | y4 | y5 |
| 18.00 | 1.20 | 0.11 | 56.20 | 5.00 | 204.91 | 75.55 | 24.40 | 18.03 | 17.97 | 114.88 | 15.04 | -14.11 | 10.95 | 3.30 | |
| 17.98 | 1.04 | 0.12 | 56.20 | 5.00 | 204.84 | 77.15 | 24.40 | 18.00 | 17.94 | 114.90 | 11.57 | -2.79 | 10.73 | 2.70 | |
| 18.00 | 1.08 | 0.12 | 56.30 | 5.00 | 205.25 | 78.15 | 24.60 | 17.97 | 17.91 | 114.91 | | | ~~10.71~~ | | |
| 18.01 | 1.21 | 0.14 | 56.50 | 5.00 | 206.07 | 80.10 | 24.20 | 17.95 | 17.88 | 114.82 | 11.79 | -5.85 | | | |
| 18.04 | 0.82 | 0.09 | 56.70 | 5.00 | 205.20 | 82.05 | 24.20 | 17.89 | 17.85 | 114.80 | | | 10.72 | | |
| 18.03 | 1.38 | 0.11 | 56.70 | 5.00 | 206.01 | 83.45 | 24.60 | 17.87 | 17.82 | 114.76 | • | • | 10.69 | • | |
| 18.01 | 1.39 | 0.16 | 57.00 | 5.00 | 206.34 | 85.45 | 24.40 | 17.84 | 17.79 | 114.73 | • | • | 10.71 | • | |
| 18.01 | 1.45 | 0.17 | 57.10 | 5.00 | 206.86 | 87.40 | 24.20 | 17.81 | 17.74 | 114.70 | • | • | | • | |
| 18.01 | 1.30 | 0.15 | 57.20 | 5.00 | 206.63 | 89.60 | 24.80 | 17.76 | 17.71 | 114.65 | | | 10.68 | | |
| 18.01 | 0.96 | 0.15 | 57.50 | 5.00 | 206.63 | 91.55 | 24.60 | 17.73 | 17.68 | 114.64 | | | ~~10.67~~ | | |
| 18.01 | 1.01 | 0.12 | 57.50 | 5.00 | 206.77 | 93.40 | 24.40 | 17.68 | 17.65 | 114.64 | | | 10.7 | | |
| 18.00 | 0.98 | 0.11 | 57.60 | 5.00 | 206.50 | 94.50 | 24.60 | 17.65 | 17.62 | 114.61 | • | • | 10.68 | • | |
| 17.99 | 1.16 | 0.13 | 57.70 | 5.00 | 206.45 | 96.10 | 24.40 | 17.63 | 17.59 | 114.59 | • | • | | • | |
| 18.00 | 0.84 | 0.14 | 57.90 | 5.00 | 206.50 | 97.95 | 24.60 | 17.60 | 17.56 | 114.54 | • | • | 10.69 | • | |
| 18.03 | 0.62 | 0.06 | 78.80 | 5.01 | 206.70 | 116.90 | 24.20 | 12.88 | 12.82 | 111.17 | | | 10.25 | | |
| 18.01 | 0.62 | 0.06 | 78.80 | 5.03 | 206.79 | 116.85 | 24.20 | 12.85 | 12.79 | 113.12 | 11.47 | -2.71 | 10.29 | 1.03 | |
| 18.05 | 0.62 | 0.06 | 78.90 | 5.02 | 206.83 | 116.90 | 24.20 | 12.82 | 12.79 | 113.09 | | | | | |
| 18.03 | 0.62 | 0.06 | 78.90 | 5.03 | 206.86 | 116.85 | 24.20 | 12.82 | 12.79 | 113.11 | | | 10.2 | | |
| 18.04 | 0.62 | 0.06 | 79.10 | 5.04 | 206.87 | 116.85 | 24.20 | 12.82 | 12.76 | 111.30 | | | 10.19 | | |

Insert Generated Data

Fig. 4. Data generation using the bootstrap method to fill in the missing data.

5.2. Quality prediction and variable selection

The process of wafer manufacturing is a chemical process, and the product quality can be measured following fabrication. When quality is predicted by current control conditions, the manufacturing process can be effectively operated. Selection of the modeling stage using DPNN is discussed, based on the roadmap.

5.2.1 Data modeling using one set of puller data (Case 1)

Figs. 5 to 8 show the test results using the trained DPNN model with unobserved data. The prediction models were designed for quality prediction corresponding to Oxygen, ORG (Oxygen Gradient), RES (Resistivity), and RRG (Resistivity Gradient). In the case of RES, the model can be designed by one puller with sufficient data. In addition, the model performance is also adequate to predict the quality of wafers with RES. However, three other parameter data are not sufficient to design a good performance model. The model was not trained well with one puller data.

5.2.2 Advanced modeling with new data (Case 2)

The preprocessing stage was required to compensate for weak points caused by insufficient data before applying the main data mining techniques. The bootstrap method is used to solve the data problem. The bootstrap method can generate reasonable data to design the data models and improve the model performance. Figs. 9 and 10 show the improved results augmented by data generation. AR gas flow, chamber pressure and heat power have strong influence on the wafer quality, so these have to be carefully handled in fabrication processes.

5.2.3 Comparison of the model performances according to data sets

Table 2 shows the comparison of results for two cases of modeling. In Case 1, the models were designed for one puller with insufficient data, so that an over-fitting problem occurred. This means that a model trained by insufficient data cannot ensure the good performance of models. In Case 2, the model

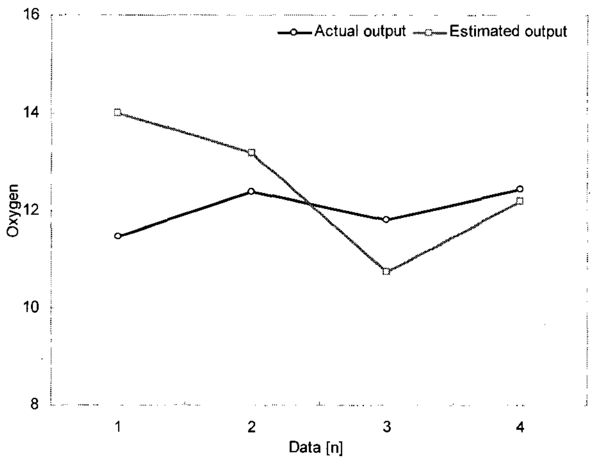



Fig. 5. Prediction results for Oxygen values using one set of puller data (Case 1).

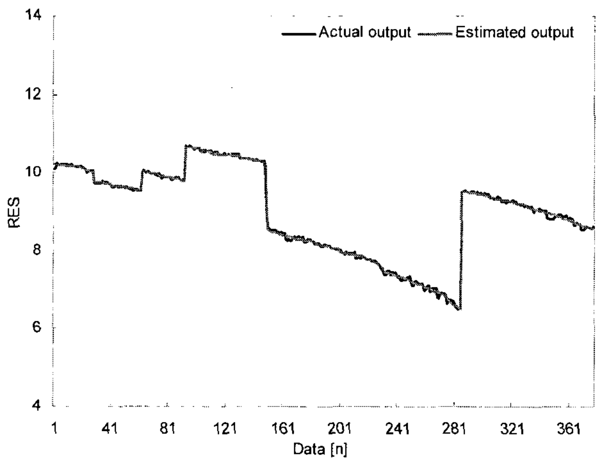



Fig. 7. Prediction results for RES values using one set of puller data (Case 1).

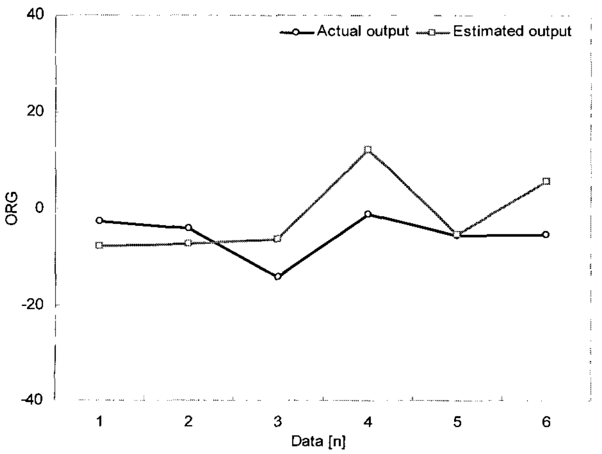



Fig. 6. Prediction results for ORG values using one set of puller data (Case 1).

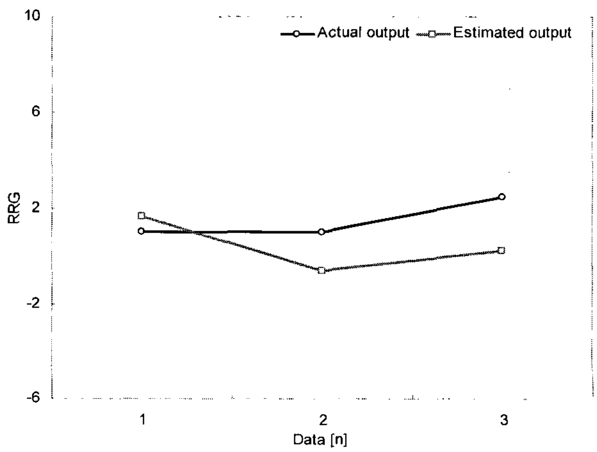



Fig. 8. Prediction results for RRG values using one set of puller data (Case 1).

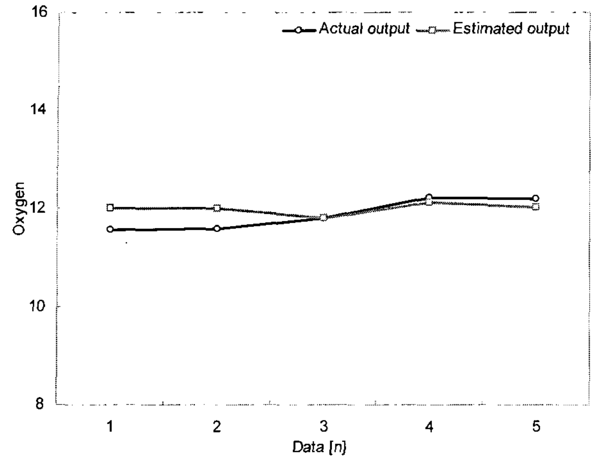



Fig. 9. Prediction results for Oxygen values using a data set generated by a bootstrap method (Case 2).

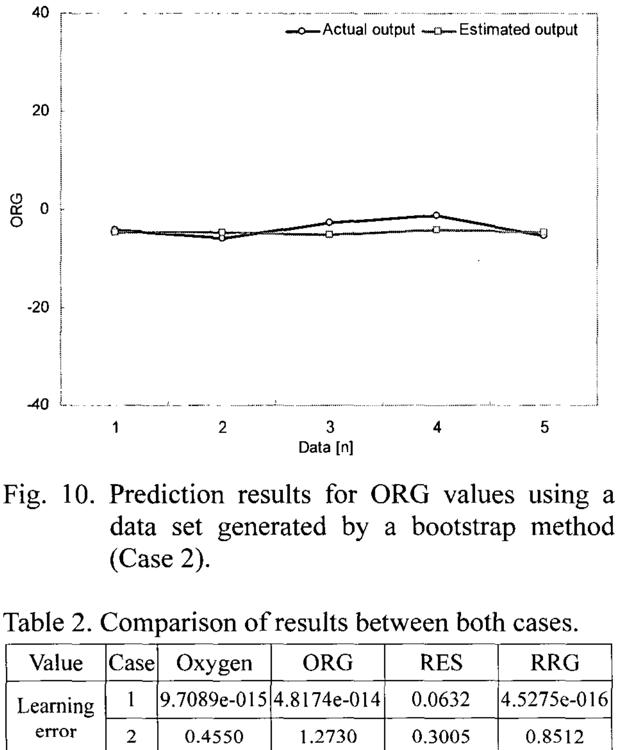



Fig. 10. Prediction results for ORG values using a data set generated by a bootstrap method (Case 2).

Table 2. Comparison of results between both cases.

| Value | Case | Oxygen | ORG | RES | RRG |
|---|---|---|---|---|---|
| Learning error | 1 | 9.7089e-015 | 4.8174e-014 | 0.0632 | 4.5275e-016 |
| | 2 | 0.4550 | 1.2730 | 0.3005 | 0.8512 |
| Prediction error | 1 | 1.4422 | 8.0759 | 0.043938 | 1.6293 |
| | 2 | 0.29528 | 1.8733 | 0.10423 | 1.2942 |
| Selected layer | 1 | 3 | 5 | 3 | 3 |
| | 2 | 4 | 4 | 4 | 3 |

trained stably by data addition using the bootstrap method showed good performance. The results indicate that statistical data generation can reduce the effect of insufficient data. It is difficult to analyze the relationship between inputs and outputs using field data as field data is often insufficient for modeling. Thus, data preprocessing is required. In this study, an adequate descriptive model was designed by data generation.

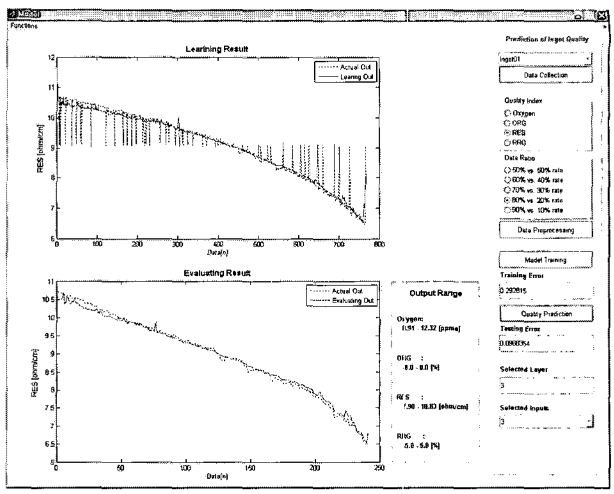



Fig. 11. Developed GUI window for data modeling.

Fig. 11 shows the GUI window for data modeling that includes the functions for data collection, preprocessing, and modeling.

5.3. Inference of control parameters based on rule extraction

The quality of wafers is evaluated according to a specific range. If the quality is out bound, the wafer will be considered defective. Therefore, the control parameters need to be determined in order to manufacture high-quality wafers.

In several fields, operators determine the control parameters by analyzing the quality and control conditions of manufactured ingots. In this case, operators usually use know-how that is defined as rules. Thus, it is suitable to extract the rules for determining operating conditions with respect to quality levels based on tree logics. The logic can combine the expert's knowledge with the rules extracted from data, and efficient systems can be constructed by knowledge and data. The rules, showing the relationship between operating conditions and qualities, are generated at the rule extraction stage of the proposed roadmap.

5.3.1 Rule generation from whole puller data

The rule extraction was not properly carried out with one puller data as they do not have several classes of qualities. A total of 10 sets of puller data are used for rule generation with respect to Oxygen, ORG, RES, and RRG, respectively, and classified as Class 1, Class 2, and Class 3. Class 1 was lower than the minimum value in the normal range, Class 2 was in the normal range, and Class 3 was higher than the maximum value in the normal range.

Figs. 12 to 13 show the rules generated by the decision tree algorithm. In Fig. 13, the rules for Oxygen, with two branches, are very simple. The tree indicates that just $x_8$ has influence on Oxygen. There

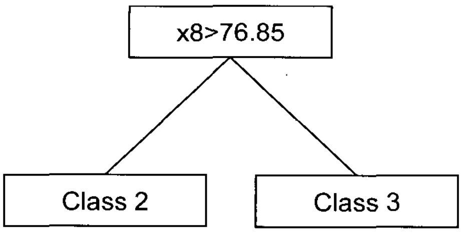



Fig. 12. Extracted rules for Oxygen with raw data sets.

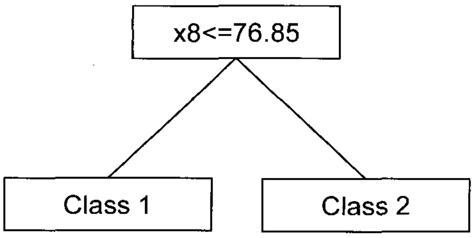



Fig. 13. Extracted rules for ORG with raw data sets.

is no rule on Class 1 as the collected data has just two range values. Fig. 13 shows that classification for ORG is similar to the one for Oxygen. Thus, one can infer that the principal control parameters of Oxygen and ORG are similar in ingot manufacturing. One difference is that the rules for ORG are divided into Class 1 and Class 2. However, if $x_8$ is higher than 76.85, the input cases of both Oxygen and ORG are classified as Class 2. The performance of classification for RES was the best among the four parameters because the RES data was sufficient to design the rules. For the RRG, the rules were not generated because all of the collected data was included in Class 2.

5.3.2 Advanced rule extraction using data generation

Unlike data modeling, the bootstrap method for data generation is not suitable at the stage of rule extraction. As for output data that are generated with no consideration of the relationship to input parameters, the bootstrap method is also not appropriate for the mechanism of rule extraction. In this study, multiple regression models were used to generate target data with respect to control inputs. In practice, the performance of rule generation was improved by data addition generated by the regression model, and the classification model was stable to the several cases.

Fig. 14 shows the extracted rules for Oxygen data; they are pruned rules because the original rules are too complicated to illustrate. The rules were pruned by the cross-validation method, as shown in Fig. 15. The best node was determined by the validation, and then pruning was achieved. In practical cases, pruned rules are applied for simple application and high performance.

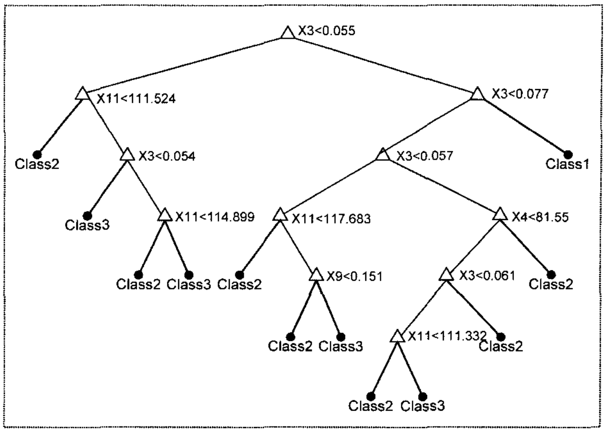



Fig. 14. Constructed tree for Oxygen with generated data sets.

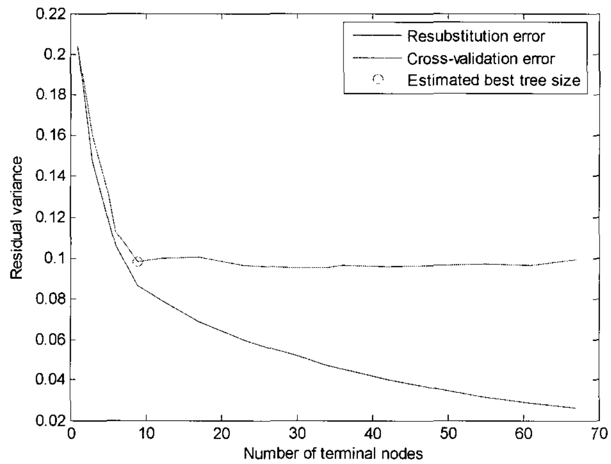



Fig. 15. Selection of the best node for Oxygen.

5.3.3 Comparison of performance of inference rules

As shown in Table 3, without data generation, only RES data could be extracted to reasonable rules. Also, the performance of inference rules is satisfied, but the proper rule extraction of the other three data sets was difficult because of unfair data. In particular, all RRG data are included in the normal range as Class 2, so the rules for Class 1 and Class 2 were not generated in this study. To solve this problem, the regression model was applied for data generation that is different from that of modeling.

By examining the results of rule extraction, the performance of rules can be improved by data generation using the regression model. The RRG data was not generated into rules caused by unfair data, but the rules were extracted by data generation. To be sure the results cannot be completely reliable, because the virtual data (generated data) was used for compensation. However, the relationship between input and output parameters in data generation was considered by using the regression model; thus, this processing is reasonable for a goal such as decision support.

Table 3. Comparison of rule generation for three variables.

| Variables / Values | Oxygen | | ORG | | RRG | |
|---|---|---|---|---|---|---|
| | Size | Error | Size | Error | Size | Error |
| 10 puller data | 2 | 4(3.8%) | 2 | 2(1.9%) | 1 | 0 (0%) |
| Generated data | 53 | 37(2.6%) | 3 | 3(0.2%) | 16 | 8(0.6%) |

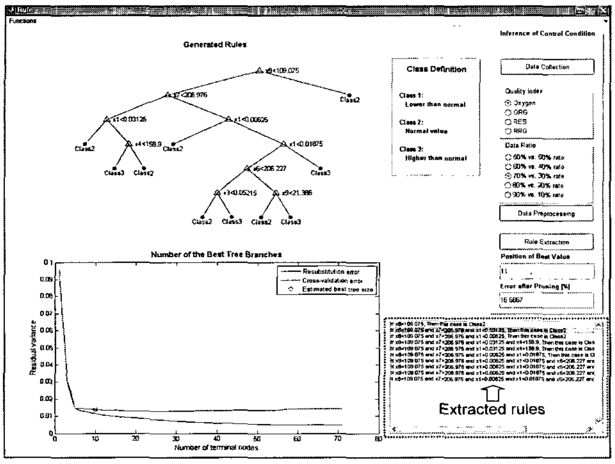



Fig. 16. Developed GUI window for rule generation.

Fig. 16 shows the GUI window for rule extraction that also includes several functions such as data collection, preprocessing, and generating rules.

## 5. CONCLUSIONS

The ingot fabrication process is one of the most important sub-processes in wafer manufacturing. In ingot fabrication, quality inspection is accomplished by product sampling testing. Afterwards, the control parameter is adjusted by an operator's action corresponding to the quality. Therefore, it is necessary to predict the quality with respect to current control parameters and to handle the parameters effectively. This function can be useful for low-defect wafer fabrication. However, it is difficult to design models or rules using collected data from the field because the data is gathered by sampling inspection. In this study, we used the bootstrap method and multiple regression models for data generation. Next, we designed models, and extracted rules using the DPNN and decision tree, respectively. Through various stages, the models and rules can be improved and their performance was reasonable.

One aim of this study was to design a roadmap for data mining, because it is difficult to determine which method is the best for a given target plant. Here, we proposed a roadmap, based on which the applied methods were selected. The models will be utilized in the future to integrate both the diagnosis and the optimization systems of the ingot fabrication process.

## REFERENCES


[1] F.-L. Chen and S.-F. Liu, "A neural-network approach to recognize defect spatial pattern in semiconductor fabrication," *IEEE Trans. on Semiconductor Manufacturing*, vol. 13, no. 3, pp. 366-373, Aug. 2000.

[2] H. Iwata, M. Ono, J. Konishi, S. Isogai, and T. Furutani, "In-line wafer inspection data warehouse for automated defect limited yield analysis," *Proc. of IEEE/SEMI Advanced Semiconductor Manufacturing Conference and Workshop*, pp. 124-129, 2000.

[3] C. Maleville, E. Neyret, L. Ecarnot, T. Barge, and A. J. Auberton, "Defect detection on SOI wafers using laser scattering tools," *Proc. of IEEE International SOI Conference*, pp. 19-20, 2001.

[4] A. Singh and J. Rosin, "Random defect limited yield using a deterministic model," *Proc. of IEEE/SEMI Advanced Semiconductor Manufacturing Conference*, pp. 161-165, 2001.

[5] J. H. Lee, S. J. Yu, and S. C. Park, "Design of intelligent data sampling methodology based on data mining," *IEEE Trans. on Robotics and Automation*, vol. 17, no. 5, pp. 637-649, Oct. 2001.

[6] C.-J. Huang, C.-F. Wu, and C.-C. Wang, "Image processing techniques for wafer defect cluster identification," *IEEE Design & Test of Computers*, vol. 19, no. 2, pp. 44-48, Mar.-Apr. 2002.

[7] T. Kubota, P. Talekar, T. S. Sudarshan, X. Ma, M. Parker, and Y. Ma, "An automated defect detection system for silicon carbide wafers," *IEEE Southeast Con*, pp. 42-47, 2002.

[8] A. Engbrecht, R. Jarvis, and A. Warrick, "An approach for improving yield with intentional defects," *Proc. of IEEE/SEMI Conference and Workshop Advanced Semiconductor Manufacturing*, pp. 284-288, 2002.

[9] J. Han, *Data Mining: Concepts and Techniques*, Morgan Kaufmann Publishers, San Francisco, 2001.

[10] P. Adriaans and D. Zantinge, *Data Mining*, Addison-Wesley, Harlow, England, Reading, Massachusetts, 1996.

[11] K. J. Cios, *Data Mining Methods for Knowledge Discovery*, Kluwer Academic, Boston, 1998.

[12] B. V. Dasarathy (Eds.), *Data Mining And Knowledge Discovery: Theory, Tools, and Technology II*, SPIE, Orlando, Florida, Bellingham, Wash., 2000.

[13] S. Dzeroski and Lavrac (Eds.), *Relational Data Mining*, Springer, Berlin, New York, 2001.

[14] U. M. Fayyad, G. Piatetski-Shapiro, P. Smyth, and R. Uthurusmy (Eds.), *Advances in Knowledge Discovery and Data Mining*, AAAI Press, MIT Press, Menlo Park, Calif., 1996.

[15] U. Fayyad, G. Grinstein, and A. Wierse (Eds.), *Information Visualization in Data Mining and Knowledge Discovery*, Morgan Kaufmann, San Francisco, CA, 2002.

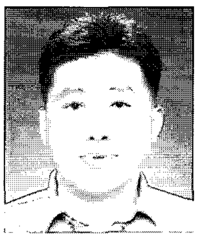
[16] S. Kim and G. J. Vachtsevanos, "An intelligent approach to integration and control of textile processes," *Information Sciences*, vol. 123, pp. 181-199, 2000.

[17] S. Kim and G. Vachtsevanos, "Polynomial fuzzy neural network for identification and control," *Proc. of North American Fuzzy Information Processing Society*, Berkeley, CA, pp. 5-9, June 19-22, 1996.

[18] G. E. Fulcher and D. E. Brown, "A polynomial network for predicting temperature distributions," *IEEE Trans. on Neural Networks*, vol. 5, no. 3, pp. 372-379, May 1994.

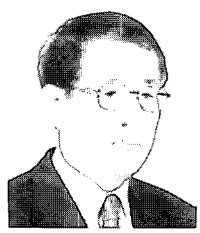
[19] F. A. Graybill, *Theory and Application of the Linear Model*, Duxbury Press, CA, 1976.

[20] T. M. Mitchell, *Machine Learning*, McGraw-Hill, 1997.

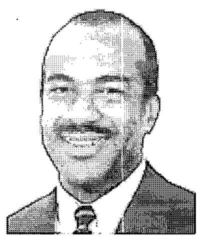


**Hyeon Bae** received the M.S. and Ph.D. degree in Electrical Engineering from Pusan National University, in 2001 and 2005, respectively. He is currently a post-doc at Information and Knowledge Control Lab, Pusan National University. His research interests include intelligent control and system, fault-tolerant control, and data mining.

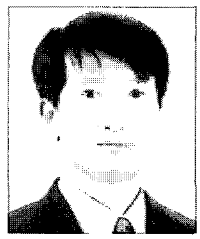
**Sungshin Kim** received the B.S. and M.S. degrees from Yonsei University, and the Ph.D. degree in the School of Electrical and Computer Engineering from Georgia Institute of Technology, Atlanta, in 1984, 1986, and 1996, respectively. He is currently an Associate Professor in the School of Electrical and Computer Engineering, Pusan National University. His research interests include intelligent control, fuzzy logic control, manufacturing systems, and data mining.



**Kwang-Bang Woo** received the B.S. and M.S. degrees from Yonsei University, and the Ph.D. degree in Electrical Engineering from Oregon State University, in 1957, 1959, and 1964, respectively. He is currently a Research Professor in Automation Technology Research Institute, Yonsei University. His research interests include intelligent control, automation science and engineering, systems biology.



**Gary S. May** received the B.E.E. degree in Electrical Engineering from the Georgia Institute of Technology in 1985 and the M.S. and Ph.D. degrees in Electrical Engineering and Computer Science from the University of California at Berkeley in 1987 and 1991, respectively. He was a National Science Foundation and is Editor-in-Chief of IEEE Transactions on Semiconductor Manufacturing. He was a National Science Foundation and an AT&T Bell Laboratories graduate fellow, and has worked as a Member of the technical staff at AT&T Bell Laboratories in Murray Hill, NJ. He is Chairperson of the National Advisory Board of the National Society of Black Engineers (NSBE). His research interests include the field of computer-aided manufacturing of integrated circuits.



**Duk-Kwon Lee** received the B.S. degree in Physics from Hankuk University of Foreign Studies, and the M.S. degree in Electrical Engineering from Kumoh National Institute of Technology, in 1986, and 1989, respectively. He is currently CIO of MEMC Korea Co. His interests include MES, System engineering, system analysis and automation.