



## 휴리스틱 기법을 이용한 PBGA 공정의 핵심 인자 추출

# Analysis of Key Parameters in PBGA Manufacturing Process Using Heuristic Approach and General Regression Neural Network

심현식<sup>1,#</sup>  
Hyun Sik Sim<sup>1,#</sup>

<sup>1</sup> 경기대학교 산업경영공학과 (Department of Industrial and Management Engineering, Kyonggi University)  
# Corresponding Author / E-mail: hssim@kgu.ac.kr, TEL: +82-31-249-1348  
ORCID: 0000-0001-5031-2289

KEYWORDS: Plastic ball grid array manufacturing process (인쇄회로기판 제조 공정), Key-parameter analysis (핵심 인자 분석), Heuristic approach (휴리스틱 어프로치), General regression neural network (회귀신경망), Yield analysis (수율 분석)

*In the printed circuit board (PCB) manufacturing industry, the yield is an important management factor as it significantly affects the product cost and quality. However, in real situations, it is difficult to ensure a high yield in a manufacturing process, because the products are manufactured through numerous nanoscale manufacturing operations. Thus, for improving the yield, it is necessary to analyze the key process parameters and equipment parameters that result in a low yield. In this study, critical equipment parameters that affect the yield were extracted through a mutual analysis of the equipment parameters ( $x$ ) and process parameters ( $y$ ) in the plastic ball grid array (PBGA) manufacturing process. To this end, the study uses the correlation coefficient to apply the heuristic algorithm that extracts critical parameters that keep the redundancy among the equipment parameters to a minimum and exert maximum impact on the critical process parameters. Additionally, by using the general regression neural network technique, the effects of the critical equipment parameters on the process parameters were confirmed. The test results were applied to the PBGA production line and an improvement in the yield was confirmed.*

Manuscript received: December 29, 2020 / Revised: February 1, 2021 / Accepted: February 27, 2021

### 1. Introduction

A printed circuit board (PCB) consists of electronic components, which are connected electrically through Cu circuits. PCBs are widely used in general household appliances, such as televisions, and in precision gadgets, such as smartphones. General household appliances that perform a single simple function require only a simple PCB structure. In contrast, the latest precision gadgets, such as smartphones, use more complex PCBs for wide-ranging functionalities and segmentation. These PCBs include various types of packages, i.e., leaded, chip-scale, plastic ball grid array (PBGA), and flip-chip ball grid array (FCBGA). Complex

substrate structures degrade the competitiveness of manufacturing plants by not only increasing the production cost but also reducing the yield. To maintain a high yield and quality while reducing the defect rate, it is critical to maintain stable operations and preserve the health of the equipment. To this end, it is essential to identify major equipment parameters that optimize the operational efficiency of the equipment with limited cost and resources.

The PBGA manufacturing process comprises various steps, i.e., Cu clad lamination, Cu plating, 1st patterning, lay-up and drilling, Cu plating, 2nd patterning, solder resist (SR) coating, Au plating, and routing (Refer to Fig. 1). First, the substrates, which are the raw materials, are cut out for each product type and fed into the

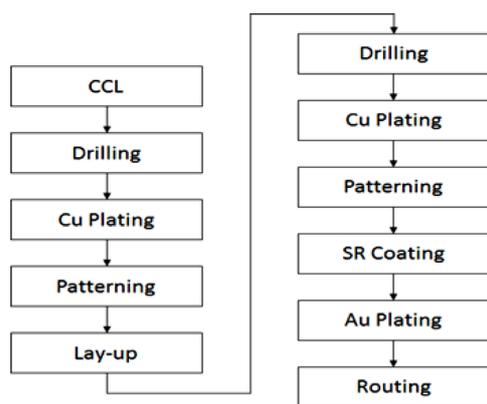


Fig. 1 Flowchart of the PBGA manufacturing process

process. In the 1st patterning step, surface plating for the conductive layer is formed on the raw material substrate; a dry film is attached; and after ultraviolet exposure, the desired pattern is formed by “developing, etching, and stripping” processes.

In the lay-up step, a multilayer substrate is formed by adding Cu foils to the patterned substrate. The lay-up step is followed by the drilling step, wherein holes for electrical connections are drilled up and down, inside, and outside on the multilayer substrate. After the drilling, the Cu plating process is used to create the circuit connections. In the 2nd patterning step, the circuit is configured on the newly formed layer from the lay-up process after the development, etching, and stripping processes. In the SR coating process, a solder resisting layer is applied to protect the surface circuitry of the product. Finally, the Au plating process is conducted to improve the electrical conductivity and corrosion resistance of the soldered ball. Then, the production process is completed via routing.

In PCB manufacturing, quality control has been performed for many years through statistical process controls, which identify faults by measuring the substrate plating thickness or wire width after the production process is complete.<sup>1</sup> However, this measurement method has a disadvantage: it is expensive and time-consuming to inspect the entire product. Furthermore, a defect occurring in an intermediate process can have a cascading effect on the subsequent processes, reducing the productivity. To reduce the measurement process and improve the quality, sensors are installed in the production equipment to measure the (i) equipment parameters ( $x$ ), such as the pressure, temperature, and concentration, and (ii) process parameters ( $y$ ) through a sample inspection.

Montgomery and Douglas<sup>2</sup> noted that traditional univariate quality control methods, such as the Shewhart and CUSUM chart, have been widely used in recent years; however, these methods have faced frequent false alarms because of correlated control variables. Additionally, a control chart considering the skewness-

kurtosis correction was presented for the efficient management of quality characteristics following a non-normal distribution among the electrical quality characteristics of unit devices, which are typically measured in the semiconductor manufacturing processes.<sup>3</sup> It also presented optimal sampling intervals and control limits for monitoring changes in production processes,<sup>4</sup> as well as a long-tail acceptance judgment chart linking the control limits and specifications.<sup>5</sup> However, these methods were used for controlling the quality characteristics of critical factors in production processes.

There are numerous equipment parameters in the PBGA manufacturing process, and considering the equipment performance, analysis resources, and time, it is difficult to extract and analyze them all.<sup>6</sup> To enhance corporate competitiveness, it is most effective to quantitatively extract and analyze such equipment data and manage crucial equipment parameters. For effective equipment parameter management, it is essential to reduce the analysis time and resources by reducing the number of factors to the greatest extent possible while analyzing the process data.<sup>7,8</sup> Ma et al.<sup>9</sup> identified important parameters affecting the yield by applying statistical methods in a complex chemical vapor deposition process. First, a specification limit to find defective wafers is set using the Mahalanobis distance, which takes into account the correlations among the variables. The Mahalanobis distance is a standardized value based on the distance from each data point to the mean of an independent variable. Furthermore, stepwise regression was used to express important parameters in a form similar to a DNA microarray, providing an easy framework for the process analysis. However, this method has limitations: it is impossible to consider several processes and multiple equipment simultaneously; rather, only one process is analyzed at a time.

The objective of the present study was to analyze equipment and processing data of a PCB manufacturing line that produces PBGA products to identify important equipment parameters and to intensively manage selected equipment parameters for stabilizing process variables, improving the yield of PCB products.

To this end, a factor extraction heuristic analysis method based on the correlation coefficient for selecting important equipment parameters is proposed.

The selected parameters were checked for conformity using the multiple linear regression (MLR),<sup>10,11</sup> partial least squares (PLS),<sup>12</sup> and general regression neural network (GRNN) methods,<sup>13</sup> and the effects of the selected critical equipment parameters on the process variables were confirmed. The remainder of this paper is organized as follows. Section 2 describes the equipment and parameters of the PBGA process. Section 3 describes the heuristic techniques used to select the critical equipment parameters. Section 4 briefly explains the process of verifying the MLR, PLS, and GRNN results.

Table 1 PBGA process equipment parameters and classification

Process	Eqp. Parameter	Group	Process	Eqp. Parameter	Group	Process	Eqp. Parameter	Group	Process	Eqp. Parameter	Group
Developing	d01	Speed	Plating	P01-p09	Temperature	Etching	e01	Temperature	Solder Resistor	s01	Temperature
	d02	Temperature					e02	Speed		s02	Temperature
	d03	Temperature					e03	Pressure		s03-s07	Pressure
	d04	Speed		e04	Pressure						
	d05	Speed		P10-p14	Concentration		e05	Concentration		s08-s12	Concentration
	d06	Temperature					e06	Concentration			
	d07	Speed					e07	Concentration			
e08			Concentration								

In Section 5, the selected critical equipment parameters are checked, and the verification results of MLR, PLS, and GRNN are presented. Section 6 presents the conclusions and future research challenges.

**2. BGA Process Data**

As described previously, there are numerous equipment parameters in the PBGA production process. However, it is difficult to periodically analyze and manage all the processing data. In this study, four core processes, i.e., “developing,” “plating,” “flesh etching,” and “SR coating,” were selected, reflecting the opinions of the engineers in charge of these processes in the production line, for accurate factor analysis.

The equipment parameters applicable to the developing process were denoted as d01-d07, Similarly, the Cu plating process parameters were denoted as p01-p14, the etching process parameters were denoted as e01-e08, and the SR process parameters were denoted as s01-s12 (Refer to Table 1). The equipment parameters corresponded to the following properties: the temperature, pressure, speed, and concentration. The width was an analysis process variable.

In this study, the PBGA production process analysis of the target company was conducted on a “lot” basis; the collected equipment parameters were classified, stored, combined into lots, and then used for analysis.

We analyzed the process capability index Cpk and the average for combining the lots (or Units) of distributed data. After confirming that most of the data in the lot were constant and the variance was small, individual data were combined using the average (Fig. 2).

For the combined data in the lots, the outliers were removed through Mahalanobis and Cook distance measurements.

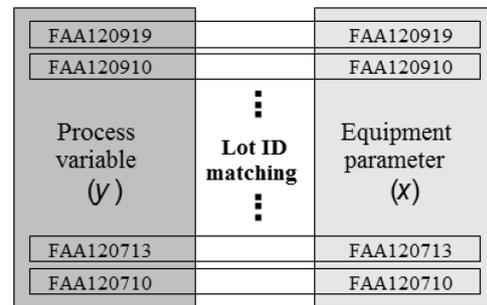


Fig. 2 Lot matching example

Furthermore, a normalization process was used to analyze the correlation of variables with different units.

**2.1 Removal of Outliers**

In any analytical technique, outliers in the data can significantly affect the analysis results. Outliers can be divided into univariate and multivariate. Multivariate outliers occur when two or more variables are combined in a sample composed of heterogeneous groups. In this study, the Mahalanobis distance and Cook distance were measured via regression analysis to detect and eliminate multivariate outliers.

The Mahalanobis distance is the standardized distance from each data point to the mean of the independent variable in the case of multiple regression.<sup>14</sup> A large value indicates that each data point is farther from the linear combination of variables. The Cook distance, which predicts the dependent variables through regression analysis, reflects the changes in the residuals if certain data points are removed.<sup>15</sup> Generally, if the Cook distance is > 1, the data point is an outlier.

From the combinations of the lots, 114 system data and processing datasets were extracted for 41 equipment parameters. The Mahalanobis distance and Cook distance were measured to remove the outliers, and 111 critical datasets were analyzed, from which three datasets were removed.

### 2.2 Normalization

The equipment parameters distributed in all the processes were largely classified into the following: the required temperature for the process; the temperature and concentration of the liquid chemical used in the manufacturing; the speed of the equipment (Such as a Conveyor); the pressure required for various bonding processes; and the rotational speed of the motor. When these parameters are analyzed together, the differences in their units may have a significant impact on the analysis results.

Therefore, for an accurate analysis, it is necessary to normalize all the data to values between 0 and 1. Eq. (1) gives the normalization process for the data. Assuming that the  $i^{th}$  individual data point of a specific equipment parameter  $x$  is  $x_i$ , we have the following:

$$x_i^{new} = \frac{x_i - x^{min}}{x^{max} - x^{min}} \quad (1)$$

### 3. Selection of Key Factors via Heuristic Approach

In this section, we introduce a heuristic correlation-based factor selection procedure using the normalized process data. The heuristic approach involves the selection of a greedy method that increases the number of multivariate factors one by one while correlating individual equipment parameters with the process parameters.<sup>7</sup> In addition to adding the equipment parameters to the factor set each time, a correlation analysis with the equipment parameters that are already selected is performed in parallel, to reduce the redundancy among the equipment parameters. In other words, the correlation coefficients (Influence) between the equipment parameters and the process parameters are calculated. The redundancy in the equipment parameters is reflected in the model by subtracting the mean (Redundancy) average of the correlation coefficient between variables in the existing factor list, as well as the equipment parameters.

The parameter with the largest influence value among the correlation coefficients with redundancy is selected as a critical factor.

This algorithm is expressed as follows.

Step 1) Initialization: Set the number of critical factors as  $k$ , insert a total of  $n$  equipment parameters into the parameter set  $F$ , and initialize  $S$  for  $m$  sets of critical factors.

Step 2) Calculation of the correlation coefficient with process variable  $y$ : Calculate the correlation coefficient  $CC(y; \hat{f}_i)$  between individuals  $\hat{f}_i \in F$  and  $y$ .

Step 3) First critical factor selection: Select the factor  $\hat{f}_i$  with the highest correlation coefficient, i.e.,  $CC(\hat{f}_i, y)$ , and insert it into the

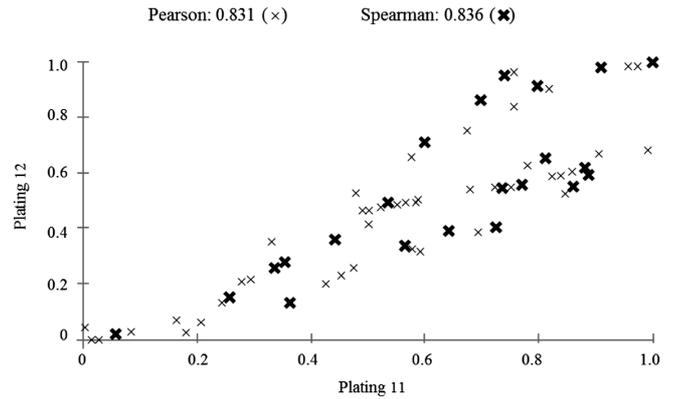


Fig. 3 Comparison of pearson and spearman correlation coefficients

corresponding process  $S$ .

$$F \leftarrow F - \{f_i\}, S \leftarrow \{\hat{f}_i\}$$

Step 4) Greedy selection method: Repeat until  $|S| = k$ . Select  $\hat{f}_i$  to maximize  $I$  in Eq. (2) and insert it into the process;  $F \leftarrow F - \{f_i\}, S \leftarrow \{\hat{f}_i\}$

$$I = CC(y; \hat{f}_i) - \frac{1}{|S|} \sum_{f_s \in S} CC(f_s; \hat{f}_i) \quad (2)$$

Step 5) Output the set  $S$ .

In the foregoing greedy selection method, the left-hand side  $I$  in Eq. (2) is the correlation coefficient with the process variable  $y$ . The right-hand side indicates the redundancy between the critical factors selected for set  $S$  in the current process and equipment parameters ( $\hat{f}_i$ ) being verified. Eq. (2) gives a pure correlation that excludes the redundancy with the equipment parameters  $\hat{f}_i$  and  $y$  in the greedy selection method. The correlation coefficients of Eq. (2) are linearly correlated with Pearson's correlation coefficient;<sup>16</sup> however, it may be preferable to use Spearman's correlation coefficient to reflect the nonlinear correlations.

In this study, Spearman's correlation coefficient was used to reflect the nonlinearity between the variables, and the difference between the two correlation coefficients was very small owing to the characteristics of the process data. Therefore, the Pearson correlation coefficient was used for all the analyses.

The analysis of the Pearson and Spearman correlation coefficient for equipment parameters p13 and p14 confirmed that there was no difference between the two, as shown in Fig. 3.

### 4. Verification of Results of Heuristic Approach

In this study, to determine the appropriate number of critical

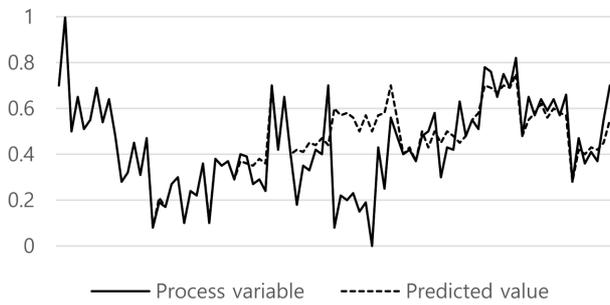


Fig. 4 Prediction results of the GRNN method for the process variables (With nine equipment parameters)

equipment parameters, the variable  $k$  was specified as the total number of equipment parameters. We performed verification using the MLR, PLS, and GRNN methods for all  $k$  values as  $k$  changed. Furthermore, a set of key equipment parameters was determined by adopting the  $k$  with the smallest root-mean square error (RMSE). MLR and PLS are suitable models when the processing data follow a multivariate normal distribution. Therefore, in this study, Q-Q plots were drawn for all the equipment parameters to check for normality. With the exception of two, all the variables followed the normal distribution. According to these results, it was assumed that the multivariate processing data followed an approximate model for the multivariate normal distribution. This section presents results obtained by using the MLR, PLS, and GRNN techniques to validate a set of critical equipment parameters that were selected via the heuristic approach.

First, the MLR analysis was performed. The purpose of the MLR model was to confirm the assumption that the derived equipment parameters did not have co-linearity. The analysis of the processing data revealed that the MLR model was not suitable for this problem, because its coefficient of determination  $R^2$  was  $< 0.5$ , with a value of 0.39. Therefore, a PLS regression analysis was performed to check for co-linearity among the equipment parameters. However, as the PLS regression yielded similar results to the MLR analysis, it was concluded that there was a nonlinear correlation among the equipment parameters.

Finally, GRNN was used as the nonlinear model. The GRNN applies the regression analysis technique to a neural-network model. With a formal structure, it has the advantage of learning a nonlinear model in a short time. Fig. 4 shows the process variables predicted with a learned GRNN using nine equipment parameters. If the input data were insufficient or the parameter setting was incorrect, a prediction error occurred, as shown in Fig. 4. In this case, the RMSE was large. To verify the analysis results, the learning and test data that were to be included in the GRNN were cross-validated. In the process of validating the results of a different set  $k$ , no random sampling was used, to eliminate the use

Table 2 Results of the heuristic parameter selection approach

k	Equipment parameter	Group
1	s06	Pressure
2	d04	Speed
3	e03	Pressure
4	p10	Concentration
5	e02	Speed
6	e04	Pressure
7	p13	Concentration
8	e01	Temperature
9	e08	Concentration
10	p12	Concentration
11	p02	Temperature
12	e07	Concentration
13	e06	Concentration
14	d03	Temperature

of a different learning and test dataset. In this verification process, 111 processing datasets were classified into five learning-test groups. Of these five groups, four were learning datasets, and one was a test group. We set  $k$  as the most suitable set of equipment parameters, with the smallest RMSE of the GRNN measured in each group.

### 5. Analysis Results

Next, we utilized the GRNN for the results obtained using the critical factor selection algorithm for the data pertaining to 41 parameters and selected the most important set of equipment parameters following the verification. When the heuristic approach was executed, the 41 parameters were sorted in the order of importance. As increased, the size of the set increased by one. The heuristic approach for this experiment was implemented using VBA 2013 (Visual Basic for Applications), and the GRNN learning and RMSE calculations were per-formed using Palisade Neural Tools 5.5 for Excel.<sup>17</sup>

Table 2 presents the result of ordering the factors using the proposed heuristic approach for equipment-parameter selection.

The “developing” and “SR coating” processes were not widely distributed but still contributed significantly to the prediction of the process variables.

The “plating” and “etching” processes can be expected to play important roles in the overall PBGA process, given their large distribution in the list. According to the equipment parameter

Table 3 GRNN verification results (RMSE)

k	t1	t2	t3	t4	t5	Average	k	t1	t2	t3	t4	t5	Average
1	0.177	0.194	0.172	0.142	0.199	0.177	8	0.176	0.140	0.239	0.139	0.160	0.171
2	0.139	0.217	0.187	0.176	0.185	0.181	9	0.130	0.151	0.193	0.135	0.176	0.157
3	0.153	0.153	0.136	0.181	0.168	0.158	10	0.170	0.138	0.159	0.163	0.180	0.162
4	0.142	0.188	0.166	0.160	0.164	0.164	11	0.161	0.180	0.162	0.122	0.181	0.161
5	0.161	0.154	0.145	0.181	0.172	0.163	12	0.163	0.122	0.195	0.196	0.177	0.171
6	0.158	0.154	0.190	0.190	0.206	0.179	13	0.164	0.150	0.161	0.200	0.143	0.164
7	0.162	0.137	0.164	0.167	0.189	0.164	14	0.187	0.191	0.165	0.132	0.160	0.167

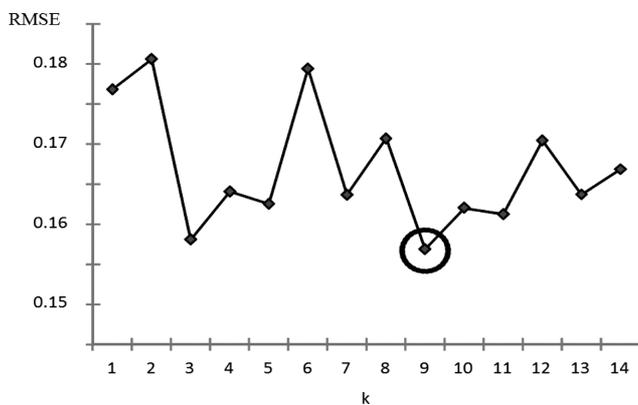


Fig. 5 RMSE measurement results

classification, the concentration of the liquid chemical was the most important parameter, regardless of the process. The importance of the other parameters was as follows: pressure, three; speed, two; temperature, three; liquid concentration, six.

Table 3 presents the results of verifying the set of equipment parameters derived from the GRNN. The datasets were placed in five categories for crossover analysis, and the RMSE of each was calculated. All the RMSEs were summarized as mean values to evaluate the set of critical factors. Fig. 5 is a graphical representation of the RMSE averages over 14 sets of  $k$ . As shown, the smallest RMSE average corresponded to  $k = 9$ , indicating that this set was the most significant of all the equipment parameter sets.

The final selected critical equipment parameters were s06, d04, e03, p10, e02, e04, p13, e01, and e08. The engineers at the site confirmed that these were important factors that affected the process variable (Line Width) for “developing,” “plating,” “etching,” and “SR coating”.

Hence, they were selected and managed intensively by engineers at the site. Additionally, the trace confirmed that the yield was improved through intensive management of these variables.

## 6. Conclusions

The objective of this study was to analyze the process data of PBGA products using new data-mining techniques with regard to various aspects, for ensuring manufacturing competitiveness by improving the yield and productivity of micro-processed products.

In this study, we gathered the equipment parameter data collected from an actual PBGA production process and analyzed how they affected the process variables measured during the processing. We proposed a heuristic technique to identify the most important parameter in predicting the process variables.

The analysis revealed that the equipment parameters that significantly affected the process variables of the PBGA process were s06, d04, e03, p10, e02, e04, p13, e01, and e08.

Among the results of this analysis, the equipment variables (Temperature and Concentration) of the plating process are managed intensively at the site, reducing the dispersion of the line width and improving the yield (Approximately 5-10%).

These parameters were selected as key factors. The equipment parameters classified as important must be managed intensively by collecting the opinions of field engineers, who are experts on the production process.

Furthermore, by delivering information about hidden critical factors that engineers may not directly perceive, we plan to conduct further factor selection studies in the future.

Finally, there is a lack of information and instances in the existing research on equipment data collection and improving the productivity in PBGA production. Therefore, it is necessary to develop various new data analysis techniques with proper consideration of the actual processing environment.

In particular, a methodology should be developed for identifying which equipment in which process has a large amount of defects in cases of dozens of processes.

Additionally, analysis of various methods can be performed

only when sufficient data on process variables and equipment variables that affect the yield are collected by product (Lot) in the production process.

## ACKNOWLEDGEMENT

This work was supported by the GRRC program of Gyeonggi province [(GRRC KGU 2020-BO1), Research on Intelligent Industrial Data Analytics].

## REFERENCES

- Spanos, C. J., "Statistical Process Control in Semiconductor Manufacturing," Proceedings of the IEEE, Vol. 80, No. 6, pp. 819-830, 1992.
- Montgomery, D. C., "Introduction to Statistical Quality Control," John Wiley & Sons, pp. 288-506, 2009.
- Cho, B. H. and Kim, S. I., "Control Chart with Skewness-Kurtosis Correction for Non-Normal Distributions in Semiconductor Manufacturing Process," Journal of the Korea Management Engineers Society, Vol. 13, No. 1, pp. 57-71, 2008.
- Bae, H. S., "Performance Ability Comparison of Control Chart for Monitoring a Production Process Shift," Journal of the Korea Management Engineers Society, Vol. 18, No. 3, pp. 177-197, 2013.
- Choi, S. W. and Yoo, J. S., "Development of Long-Tailed Control Charts for Skew-Right Distribution," Journal of the Korea Management Engineers Society, Vol. 16, No. 2, pp. 197-205, 2011.
- Guyon, I. and Elisseeff, A., "An Introduction to Variable and Feature Selection," Journal of Machine Learning Research, Vol. 3, pp. 1157-1182, 2003.
- Estévez, P. A., Tesmer, M., Perez, C. A., and Zurada, J. M., "Normalized Mutual Information Feature Selection," IEEE Transactions on Neural Networks, Vol. 20, No. 2, pp. 189-201, 2009.
- Goodlin, B. E., Boning, D. S., Sawin, H. H., and Wise, B. M., "Simultaneous Fault Detection and Classification for Semiconductor Manufacturing Tools," Journal of the Electrochemical Society, Vol. 150, No. 12, pp. 778-784, 2003.
- Ma, M. D., Wong, D. S. H., Jang, S. S., and Tseng, S. T., "Fault Detection based on Statistical Multivariate Analysis and Microarray Visualization," IEEE Transactions on Industrial Informatics, Vol. 6, No. 1, pp. 18-24, 2009.
- Montgomery, D. C., Peck, E. A., and Vining, G. G., "Introduction to Linear Regression Analysis," John Wiley & Sons, 5th Ed., 2012.
- Shmueli, G., Patel, N. R., and Bruce, P. C., "Data Mining for Business Intelligence," John Wiley & Sons, 2009.
- Wold, S., Sjöström, M., and Eriksson, L., "PLS-Regression: A Basic Tool of Chemometrics," Chemometrics and Intelligent Laboratory Systems, Vol. 58, No. 2, pp. 109-130, 2001.
- Specht, D. F., "A General Regression Neural Network," IEEE Transactions on Neural Networks, Vol. 2, No. 6, pp. 568-576, 1991.
- De Maesschalck, R., Jouan-Rimbaud, D., and Massart, D. L., "The Mahalanobis Distance," Chemometrics and Intelligent Laboratory Systems, Vol. 50, No. 1, pp. 1-18, 2000.
- Cook, R. D., "Detection of Influential Observation in Linear Regression," Technometrics, Vol. 19, No. 1, pp. 15-18, 1977.
- Walpole, R. E., Myers, R. H., Myers, S. L., and Ye, K., "Probability & Statistics for Engineering and Scientists," Pearson Education, 9th Ed., 2016.
- Palisade Corporation. Neural Tools, "Neural Networks Add-In for Microsoft Excel. Version 5.7," <http://www.palisade.com> (Accessed 26 MARCH 2021)



### Hyun Sik Sim

He received Ph.D. degree in Information & Industrial Engineering from Yonsei University in Seoul, Korea. He is now professor in the Department of Industrial & Management Engineering at Kyonggi University, Korea. His research interest is in smart manufacturing systems.  
E-mail: hssim@kgu.ac.kr