Global Journal of Researches in Engineering: E
Civil And Structural Engineering
Volume 14 Issue 3 Version 1.0 Year 2014
Type: Double Blind Peer Reviewed International Research Journal
Publisher: Global Journals Inc. (USA)
Online ISSN: 2249-4596 & Print ISSN: 0975-5861

Evaluation of the Parameters of SOM Model for Selecting Design Section

By Eun-Sang Im & Dae-Hyeon Kim
K-water Institute, Korea

Abstract- In order to support the basic information for designing the soft ground improvement, it is highly recommended to understand the spatial distribution of the geotechnical parameters in the landfill site. Therefore, in the present study, we applied Self-Organizing Map (SOM) to detect the characteristics of spatial distribution of the parameters which have been measured in Songsan Green City. For the purpose, the input dataset for SOM was constructed with the classification by USCS, initial Void ratio, unconfined compression strength, compression index from the stations of 41 in the site. Consequently, the methodology based on the SOM proposed in the present study can be considered that it is highly applicable to detect the spatial distribution of the Parameters and it can be used effectively for the further utilization as a data analysis tool.

Keywords: consolidation, self-organizing map (som), soft ground, design section, clustering.

GJRE-E Classification : FOR Code: 090599, 090506

EVALUATIONOFTHEPARAMETERSOFSOMMODELFORSELECTINGDESIGNSECTION

Strictly as per the compliance and regulations of :

Evaluation of the Parameters of SOM Model for Selecting Design Section

Eun-Sang Im [α] & Dae-Hyeon Kim [σ]

Abstract- In order to support the basic information for designing the soft ground improvement, it is highly recommended to understand the spatial distribution of the geotechnical parameters in the landfill site. Therefore, in the present study, we applied Self-Organizing Map (SOM) to detect the characteristics of spatial distribution of the parameters which have been measured in Songsan Green City. For the purpose, the input dataset for SOM was constructed with the classification by USCS, initial Void ratio, unconfined compression strength, compression index from the stations of 41 in the site. Consequently, the methodology based on the SOM proposed in the present study can be considered that it is highly applicable to detect the spatial distribution of the Parameters and it can be used effectively for the further utilization as a data analysis tool.

Keywords: consolidation, self-organizing map (som), soft ground, design section, clustering.

I. Introduction

Due to the scarcity of industry land resulting from recent development of national industry and from securing new construction sites, a large size of industrial sites are being constructed in many places along the shore of the Korean peninsula. As the area of such construction sites is usually so huge, the ground condition is not homogeneous and the depth of the soft ground that needs to be improved varies in different areas of the site.

Therefore, it is uncommon to apply the same soil properties to all the area of the site, but it is common to do design for several divided sections of the site that have similar soil profile and soil properties. In practice, a commonly used selection method for design zone is that we divide design zones depending on the parameters such as compression index and depth of soft ground, or depending on the purpose of use of the site. These selection methods of design zone cannot differentiate the design zone without considering all the parameters, but they take into account one individual factor, leading to a designer's subjective selection of the design zone. Such an inaccurate determination of design parameters may result in uncertainties in the estimation of settlement of soft ground.

The purpose of the study is to improve the selection of the design zone considering many factors. To accomplish the goal of the study, we employed a SOM (Self-Organizing Map) model, one of the neural network models to Songsang Green City Project. Since so many data and calculations are involved in representing engineering properties and site conditions, we applied the developed selection technique of the design zone to four groups of representative parameters.

II. SOM Model

SOM allows mapping from multi-dimensional data to two dimensional data. SOM is one of the neural network models and is used for grouping. In particular, it is very applicable to the grouping of the multi-dimensional data. Also, it can visualize the data, predetermine the structure of the grouping result, and acquire good results very quickly although there are so many data.

Kohonen (1982) of Helsinki University introduced a self structuring system. This system allows to map the input data that exist in the external space to the internal space that the system defines, without user's assistance. Kohonen called this system Self-Organizing Feature Map or (adaptive vector quantizer AVQ), and many kinds of SOM have been proposed and developed.

SOM has a similar algorithm to VQ (Vector Quantum) such as K-Means technique and is a kind of unsupervised learning algorithm. It maps independently n-dimensional data to two dimensional data and the typical structure of SOM is shown in Figure 1.

The structure of SOM has n number of input nodes when it uses n-dimensional input data. When input data are categorized as k number of regions, it has k number of output nodes. All the input nodes are connected to all the output nodes and they have their own weight. Generally input nodes transmit input data to the network, and output nodes calculate the distance using the input data and weight.

Figure 1. Typical structure of SOM

The philosophy of Kohonen's network is that only winner can output- that is, the principle of "winner take all", and that their neighbors only control their weight. To do this, the weight vector of node has certain values and should be appropriately initialized. Also,

Author α: Ph.D., Senior Researcher, Infrastructure Research Center, K-water Institute, K-water, Korea. e-mail: esim89@kwater.or.kr
Author σ: Ph.D., P.E., Assistant Professor, Department of Civil Engineering, Chosun University, Korea.

each node controls the weight in the course of three steps such as competition process, close radius (근접반경) process, adaptation learning process.

a) Competitive Process

Each neuron competes to have privilege for possessing to learn and in this process, a neuron wins that has the nearest distance between the weight vector and the input vector. The input pattern with m number of input and the weight vector can be defined as follows.

$$X = [x_1, x_2, \cdots, x_m]^T \qquad (1)$$

$$W_j = [\omega_{j1}, \omega_{j2}, \cdots, \omega_{jm}]^T, \quad j = 1, 2, \cdots, l \qquad (2)$$

Where l= the number of neuron

The winner neuron ($i(X)$) out of output neuron is determined by the following condition.

$$i(X) = \mathrm{arg}min \parallel X - W_j \parallel, \quad j = 1, 2, \cdots, l \qquad (3)$$

Thus, the selection of the winner neuron is to select the weight vector that is most similar pattern. Euclidean Distance is used to measure similarity matching.

b) Close Radius' Adjustment Process

Kohonen's system uses lateral inhibition through cooperative process with neighboring neuron, which can be similarly observed from the biological model. Learning about the input vector is allowed to both the winner and the neighboring neuron in the competition process. The radius for determining the neighboring neuron becomes small, and a very few neuron have chances to learn. Finally, the winner controls the weight.

In addition, Gaussian function is used in the process of controlling the geometric radius, as it has symmetric and converged characteristics. The function (h_{bi}) that defines its neighboring radius and the function (d_{ji}) that represents the distance between neighboring neurons are as follows.

$$h_{bi}(t) = \exp\left(-\frac{d_{ji}^2}{2\sigma^2(t)}\right), \quad t = 0, 1, 2, \cdots \qquad (4)$$

Where d_{ji} can be defined in such a way that the distance vector r_j is expressed with an expected j neuron and a winner neuron j.

In general, an exponential damping σ can be expressed as follows.

$$\sigma(t) = \sigma_0 \exp\left(-\frac{t}{\tau_1}\right), \quad t = 0, 1, 2, \cdots \qquad (6)$$

In equation 6, σ_0 is the value of σ in the initialization of SOM algorithm, τ_1 is the time constant.

c) Process of Adaptation Learning

When the process aforementioned is completed, the adjustment of the weight such as synapse is finally done. In defining the adjusted weight vector $W_j(n+1)$ in terms of discrete time t from the unadjusted weight vector $W_j(n)$, the rule for the adjustment can be expressed using the following equation.

$$W_j(n+1) = W_j(n) + \alpha(n) \cdot h_{bi}(n) \cdot [X - W_j(n)] \qquad (7)$$

Where, α is the decreasing learning rate with increasing t and α can be expressed such that it satisfies the initial value α_0 and exponential reduction.

$$\alpha(t) = \alpha_0 \exp\left(-\frac{t}{\tau_2}\right), \quad t = 0, 1, 2, \cdots \qquad (8)$$

τ_2 is the another constant of SOM algorithm, as in equation 6, Figure 2 shows the renewal process of weight vector.

Figure 2. Renewal process of weight

The adjustment of the weight of adaptive learning in equation 7 undergoes Self-Organizing Phase and Convergence Phase. The Self-Organizing Phase is the phase that h_{bi}, defined as the adjacent radius, initiates the step including the output layer's every neuron and narrows down the radius around the winner neuron. The Convergence Phase is the fine tuning phase that very small learning rate is selected, the nearest neuron is included, and the winner neuron is ultimately selected.

III. Implementation of Som Model

a) The Site Area

The site in this study is located Songs an in Korea. It is adjacent to Lake Siwha to the north, Island Daebu and Island Yeugheung to the west, Siwha and Banwoel Industrial complexes around the site, and the West Coast Highway to the east. This site has been made from a large size reclamation project where a seawall has been constructed to prevent seawater. Due to desalination of Lake Siwha, alluvial layer is mainly found. The site consists of a sedimentary layer and a thick alluvial layer (soft soil).

Figure 3.The site area

In order to identify the applicability of SOM model, District 1 from Songs an Green City has been selected. Soil parameters used in the study are obtained from insitu field tests and laboratory tests. For reference, the site, in the master plan, was subdivided into 7 zones based on the distribution of soft soil, which was done in the designer's own judgment.

Figure 4. The site area Locations of boring and physical and mechanical properties of soil

Figure 5. Design zone considering soft soil

b) Discussion of Results of SOM Model

In this study, we used the SOM toolbox developed in University of Helsinki and Matlab to verify SOM model and its applicability. Figure 6 exhibits the flowchart of SOM and its first step is to select parameters. For the model to have a good discernment, the number of parameters being used is very important. It is not necessarily good to use many parameters and it is desirable to remove useless parameters or highly correlated parameters for avoiding, so called, the curse of dimension. As many parameters require huge data and calculation, parameters should be appropriately reduced

Figure 6. Flowchart of SOM model

In this study, we take into consideration 12 physical and mechanical properties from 41 locations: water content (Wn), specific gravity (Gs), liquid limet (LL), plastic Index (PI), percent passing #200 sieve (2μm), Unified Soil Classification System (U.S.C.S), unconfined compressive shear strength (Q_u), triaxial shear strength (C_{uu}), sensitivity(S_t), preconsolidation (P_c), modified compression index (C_{CF}), void ratio (e_O). Out of these 12 parameters, the 1st parameter group considers 5 parameters that are not affected by the ground condition, the 2nd parameter group does 3 parameters related to the initial state and mechanical properties, the 3rd parameter group does physical and mechanical properties, and the 4th parameter group does all of the above. The applicability of SOM has been studied for these 4 types of parameter group.

Table 1. Physical and consolidation properties of soils

In SOM training, as the size of map and array affects significantly the performance of SOM, commonly used hexagon array $5\sqrt{N}$ (n = number of training input data) map size has been used. This classification is done considering additional geological characteristics as it is for the design zone of the site.

(1) The 1st parameter group

Figure 7 presents training results of 6×6 SOM model for the 1st parameter group considering physical properties. Figure 7(a) shows the distribution of unified distance matrix (U-matrix) on SOM that unifies each parameter, and Figure 7(b) and 7(c) represent the optimum number of cluster and results of classification, respectively.

Figure 7. Training results of 6×6 SOM model for the 1st parameter group

The results of classification except for the clusters with small number of clusters are shown in Figure 8. In the case of separating the design zones, the constructability can be secured when designing with zones considering the location of the data. The design zones are divided into two zones considering geological conditions

Figure 8. Results of design zone of 6×6 SOM for the 1st parameter group

(2) The 2nd parameter group

For the 2nd parameter that considers the initial state and mechanical properties, 10×8 size SOM has been applied. However, it is hard to select design zones because 4 clusters are complicated.

Figure 9. Results of design zone for the 2nd parameter group

(3) The 3rd parameter group

The 3rd parameter group is the parameter group that considers physical and mechanical properties. A pattern categorization has been performed by applying SOM map with 7×8. As shown in Figure 10, it is optimum to classify two clusters. Based on the result of the categorization, two design zones, shown in Figure 11, have been categorized.

Figure 10. Results of training of 7×8 SOM model for the 3rd parameter group

Figure 11. Results of categorization of design zone for the the 3rdparameter group

(4) The 4th parameter group

The 4th parameter group considersinitial state, the second parameter group, and soil's properties under USCS. Figure 12 shows the result of SOM map with 10×10. Also, based on the result of the categorization, three design zones, shown in Figure 12, have been categorized.

Figure 12. Results of training of 10×10 SOM model for the 4th parameter group

Figure 13. Results of categorization of design zone for the4th parameter group

c) Evaluation of Applicability for Design Zone

In order to find applicability of selection of design zone, the mean value and standard deviation for modified compression index (C_{CF}), one of the most important parameters, have been calculated. Figure 14 shows the distribution of modified compression index for all the site, and its mean value and standard deviation were 0.350 and 0.107, respectively. The statistics for the zone according to the distribution of soft ground is presented in Table 2. As shown in Table 2, the mean value and standard deviation are in the range of 0.289 ~ 0.429, and 0.055 ~ 0.169, respectively. Note that for some zones, the number of data is too small, the reliability of the mean value is questionable. As Zone 1, 2, 4 and 6 have small data and relatively high standard deviation, it seems unreliable to do consolidation analysis with the mean values of these zones.

Figure 14. mean value and standard deviation for modified compression index (C_{CF})for all the site

Table 2. mean value and standard deviation for existing design zone

(1) The 1st parameter group

In order to find applicability of SOM model, the mean value and standard deviation for the modified compression index of the main cluster in each design zone been calculated. Based on the results, design zone 1 tends to have high standard deviation, but in general categorization of the design zone has been done reasonably.

Figure 15. mean value and standard deviation for modified compression index (C_{CF}) for each design zone

(2) The 3rd parameter group

The mean value and standard deviation for the modified compression index have been calculated. The results have been found to be improved.

Figure 16. mean value and standard deviation for modified compression index (C_{CF}) for each design zone

(3) The 4th parameter group

In order to find applicability of the 4th parameter group SOM model, the mean value and standard deviation for the modified compression index of the main cluster in each design zone has been also calculated. It is found that design zones 1-3 have small standard deviation.

Figure 17. mean value and standard deviation for modified compression index (C_{CF}) for each design zone

Based on the results, the 4th parameter group was found to be the most appropriate model because modified compression index has been used as categorization factor. However, it is more desirable to select the 4th parameter group considering the initial state, physical and mechanical properties.

Table 3. mean value and standard deviation for each parameter group

IV. Conclusion

In this study, we attempted to develop a more reasonable method of estimating consolidation settlement using SOM for Songsang Green City Project in Korea. On the basis of the information obtained from the site investigation, categorization using SOM has been made. The following conclusions can be drawn based on the results of the study.

- We were able to improve the design zone by using SOM technique. In dividing the design zone for a wide area of the site, many parameters have been considered unlike the existing method considering one parameter.
- When using SOM technique for the categorization of design zone, it is desirable to utilize initial state, physical and mechanical properties of soft ground.
- For modified compression index, compared with other parameter groups such as the 1st parameter group considering only physical properties, the 2nd parameter group considering initial state and mechanical properties, the 3rd parameter group considering mechanical and physical properties, the 4th parameter group considering initial state, physical and mechanical properties obtained more reliable results with less standard deviation.
- It is better to use the 3rd parameter group as we need to take into account physical and mechanical properties of soils as well as initial states.

References Références Referencias

1. Kim, K. B, (2003), Development of Algorithm for recognizing car plates and parking management system using SOM, Journal of Korean Marine Information Communication, Vol 7, no. 5, pp. 1052-1061.
2. Kim, Y. G, Jin, Y. H, and Park, S. C (2006), Application of SOM for analysis of rainfall-discharge characteristics, Journal of KSCE, Vol 2, no. 1, pp. 61-67.

3. Kim, J. T, Lee, K. H, Lim, Y. S and Ku, J. Y(2004), A study on Partial Blackout Pattern Recognition using SOM, Journal of Korean Electrical Engineering, Vol 53, no 10, pp. 515-522.
4. Park, W. J, Jung, S. U, Kim, K. Y, Ahn, W. K, and Shin, M. S(2008), Investigation of Content-Based Image using Data Integration, Journal of Information Management, Vol 25, No 2, pp. 49-68.
5. Jin, Y. H, Kim, Y. G, Noh, K. B, and Park, S. C. (2009), Application of SOM for locating Space Distribution based on Water Quality and Quantity, Journal of Korean Water Environment, Vol 25, No.5, pp. 735-741.
6. Korea Water Resource Cooperation (2009), Site Investigation Report, Songsan Green City Zone.
7. Gatciar, H. L. and Gonzalez, I. M. (2004). "Self-organizing map and clustering for wastewater treatment monitoring." *Engineering Applications of Artificial Intelligence,* Vol. 17, pp. 205-225.
8. Ian H. Witten and Eibe Frank (2005). "Data Mining: Prectical mechine learning tools and techniques." *Elsevier.*
9. Juha Vesanto, Johan Himberg, Esa Alhoniemi, and Juha Parahankangas (2000) "SOM toolbox for matlab 5." *Technical Report.*
10. Kohonen, T. (1982). "Self-organized formation of topologically correct feature maps." *Biological Cybernetics,* Vol. 43, pp. 59-69.
11. Luk, K.C., Ball, J.E. and Sharma, A. (2000). "A study of optimal model lag and spatial inputs to artificial neural network for rainfall forecasting." *Journal of Hydrology,* Vol. 227, pp. 56-65.

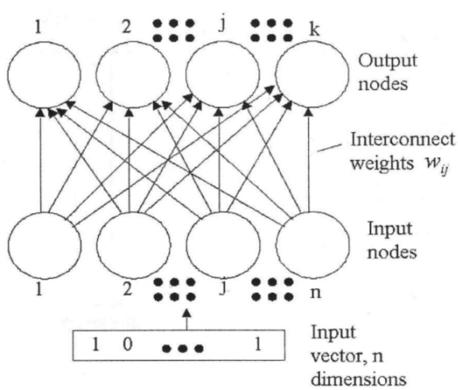


Figure 1 : Typical structure of SOM

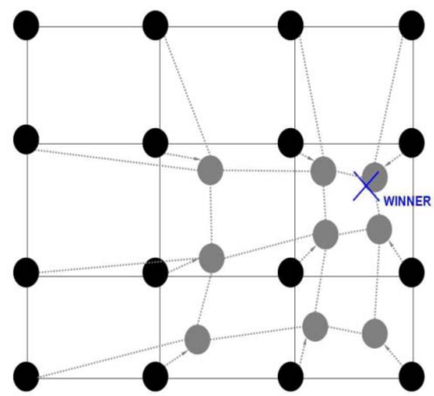


Figure 2 : Renewal process of weight

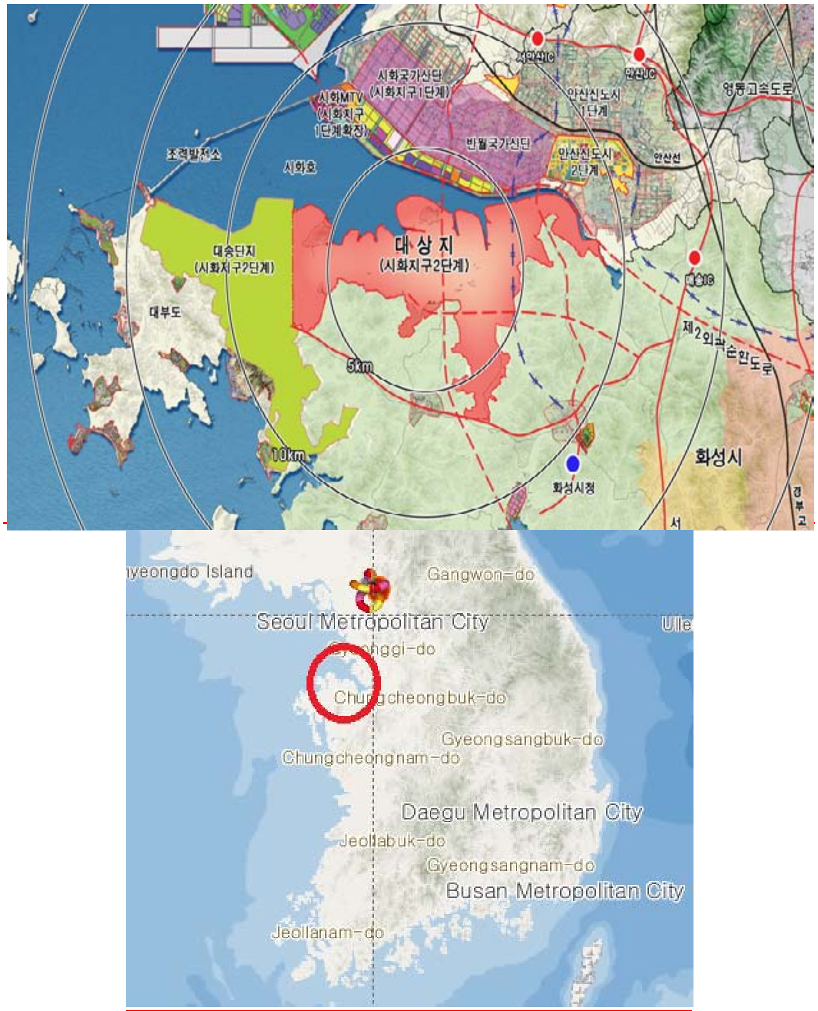


Figure 3 : The site area

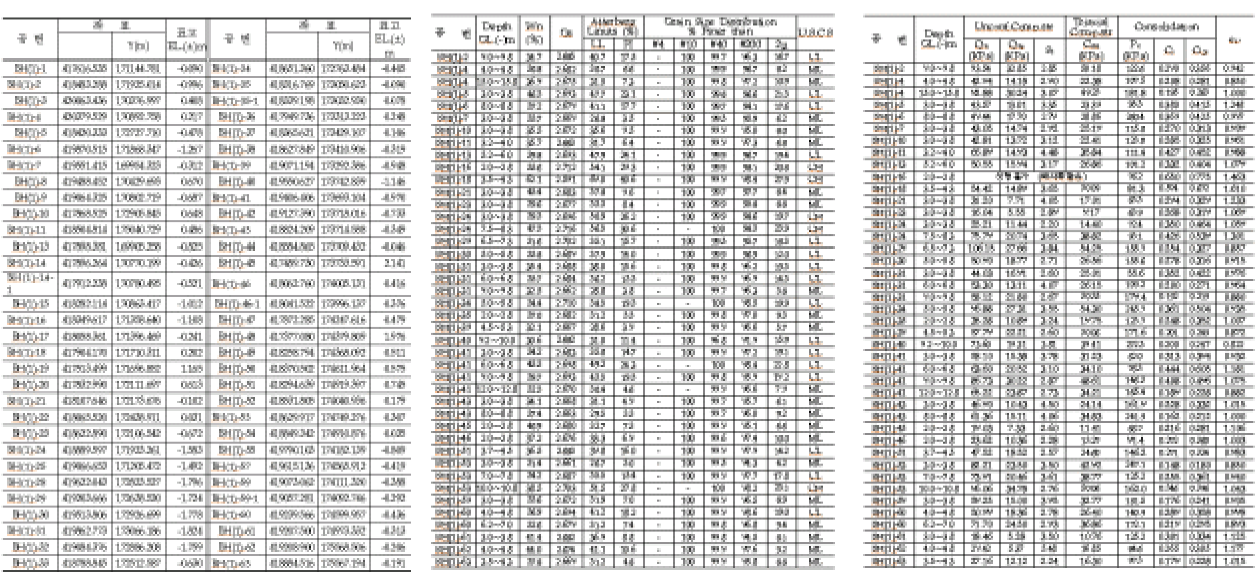


Figure 4 : The site area Locations of boring and physical and mechanical properties of soil

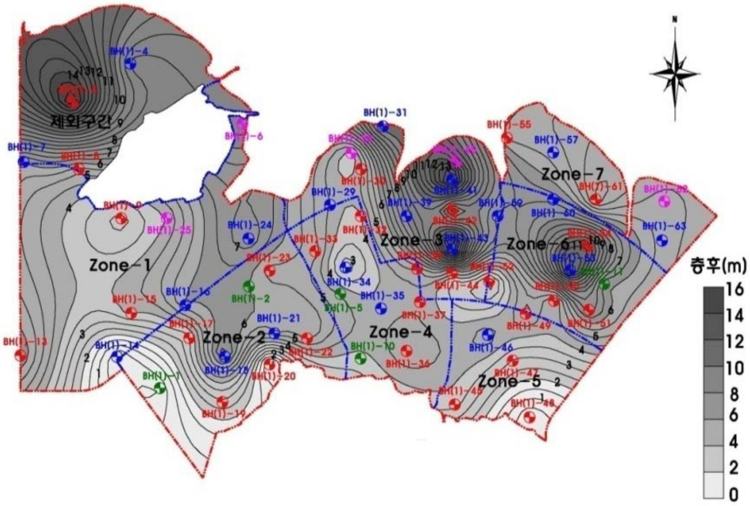


Figure 5 : Design zone considering soft soil

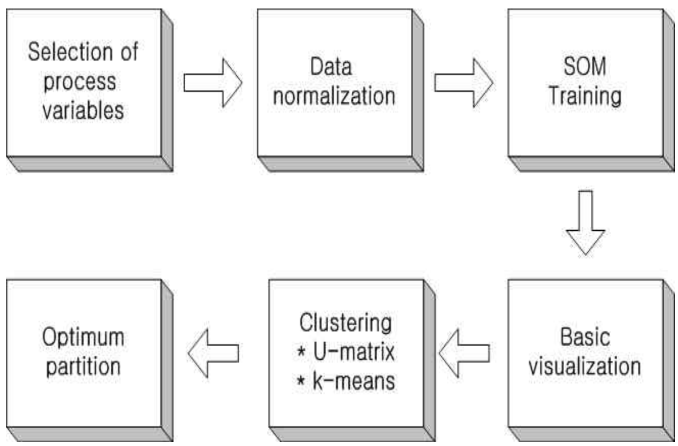


Figure 6 : Flowchart of SOM model

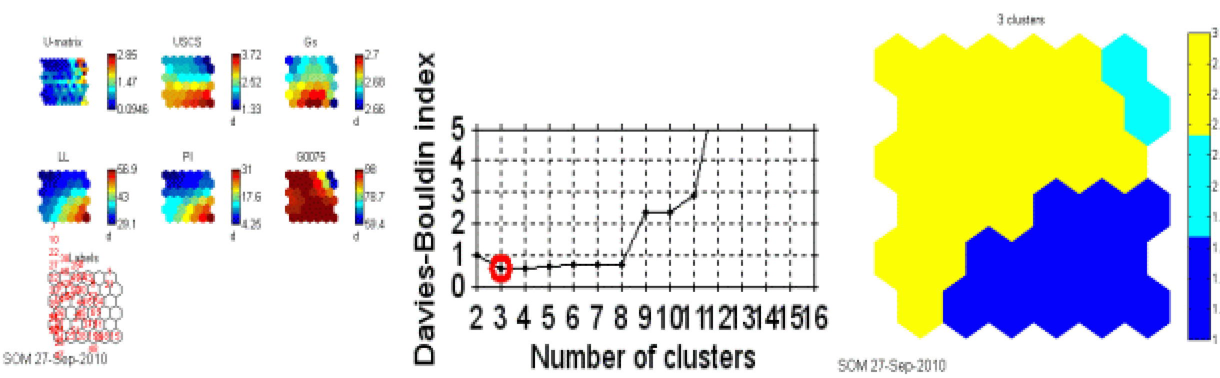


Figure 7(a) Figure 7(b) Figure 7(c)

Figure 7 : Training results of 6×6 SOM model for the 1st parameter group

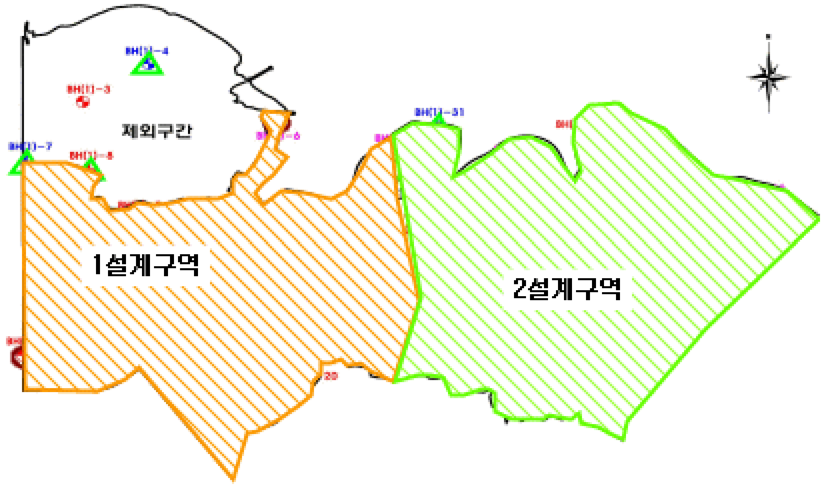


Figure 8 : Results of design zone of 6×6 SOM for the 1st parameter group

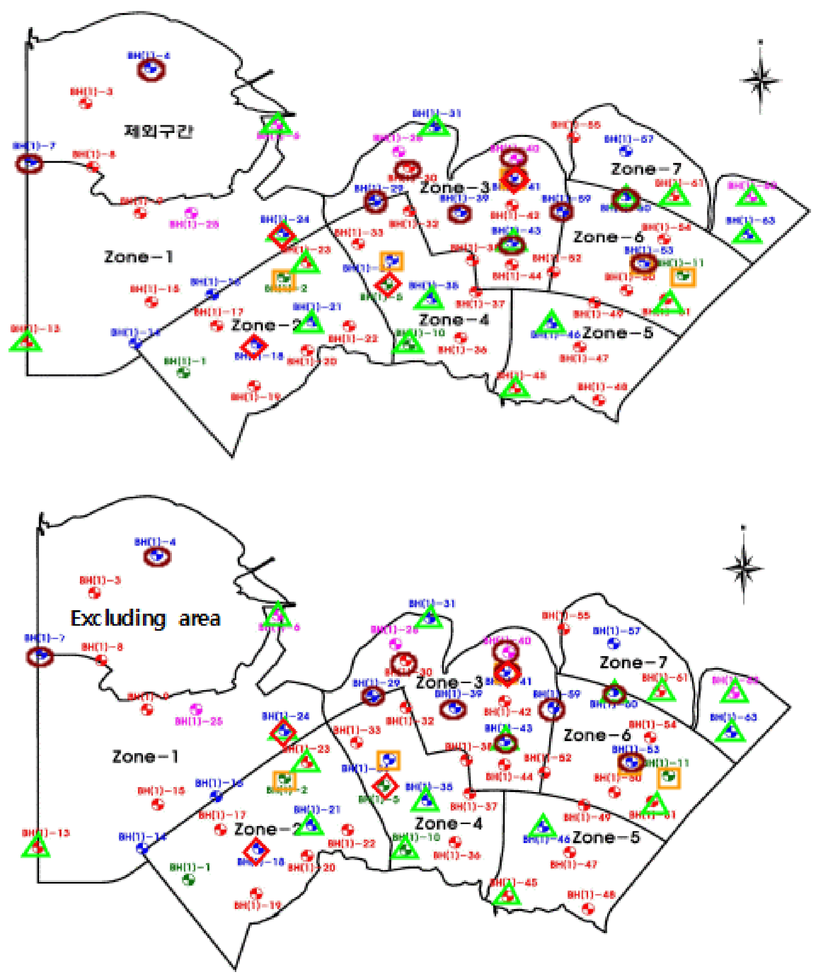


Figure 9 : Results of design zone for the 2nd parameter group

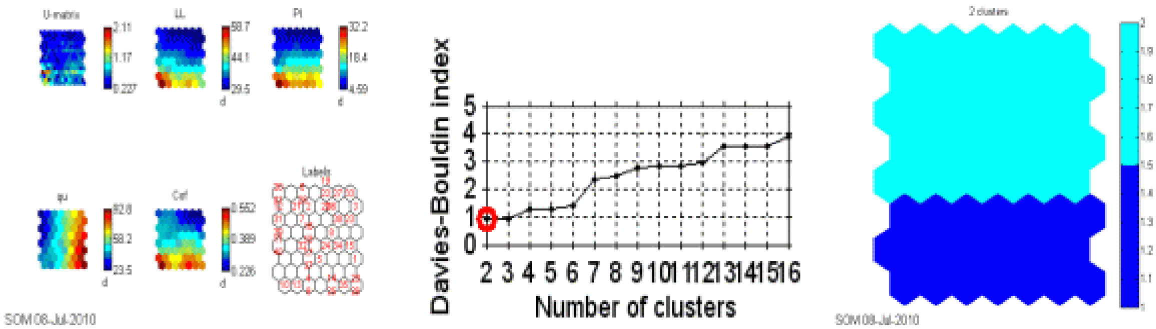


Figure 10 : Results of training of 7×8 SOM model for the 3rd parameter group

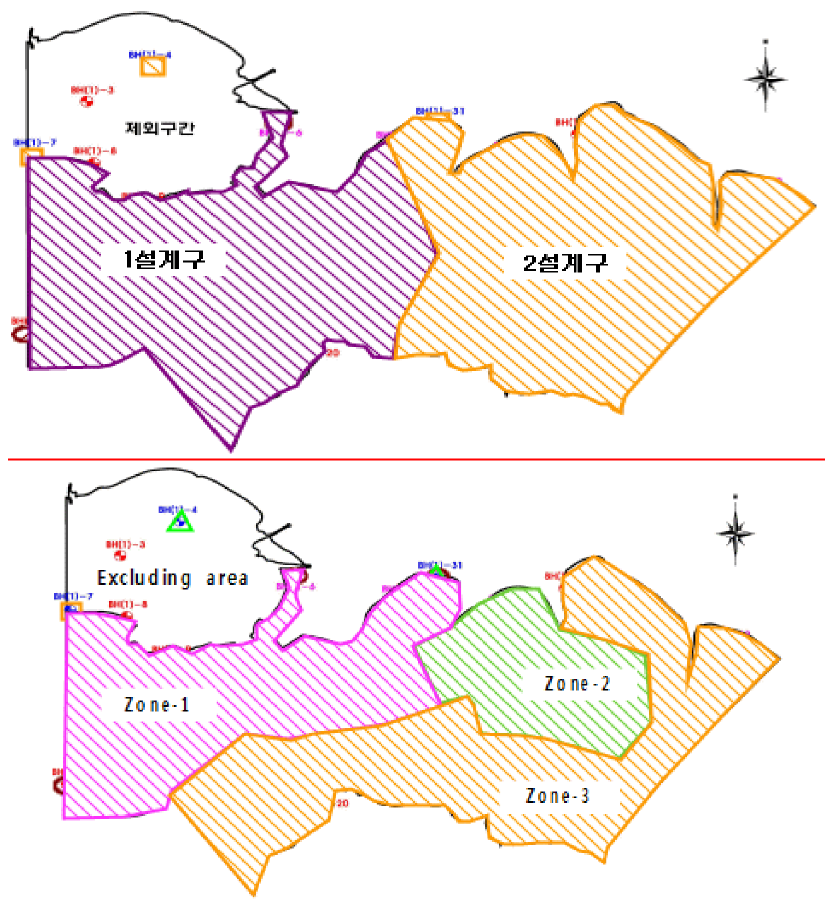


Figure 11 : Results of categorization of design zone for the the 3rd parameter group

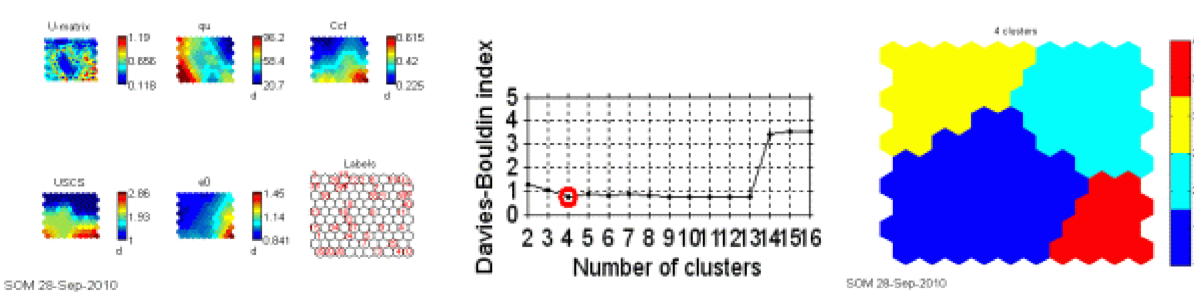


Figure 12 : Results of training of 10×10 SOM model for the 4th parameter group

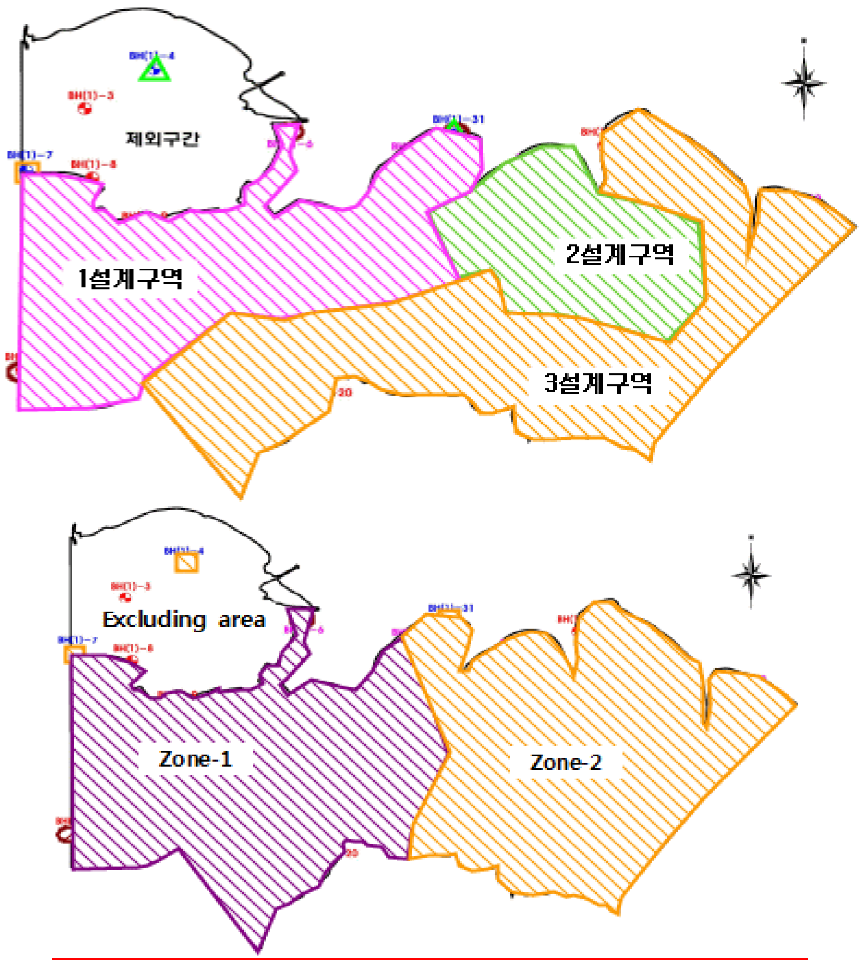


Figure 13 : Results of categorization of design zone for the 4th parameter group

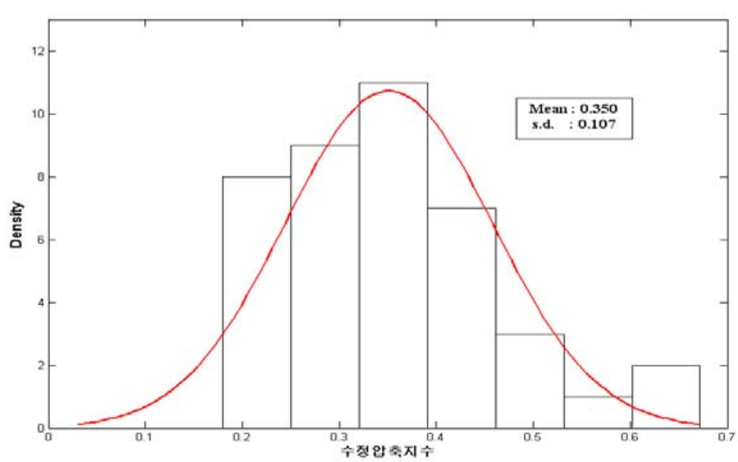
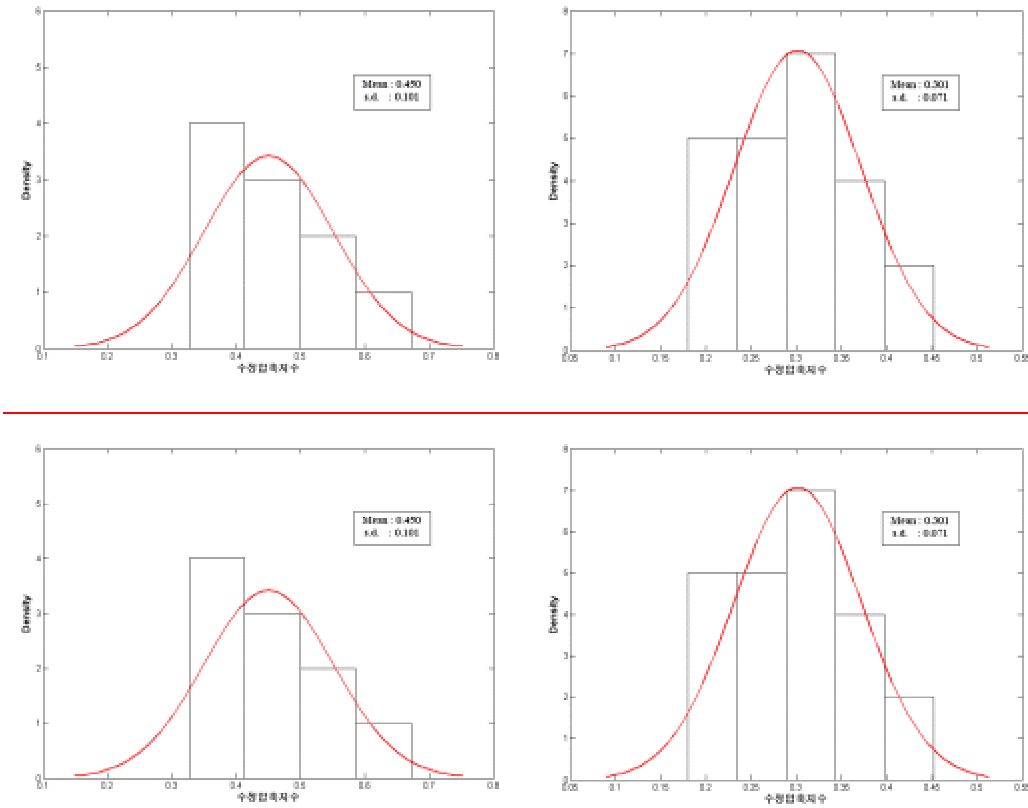
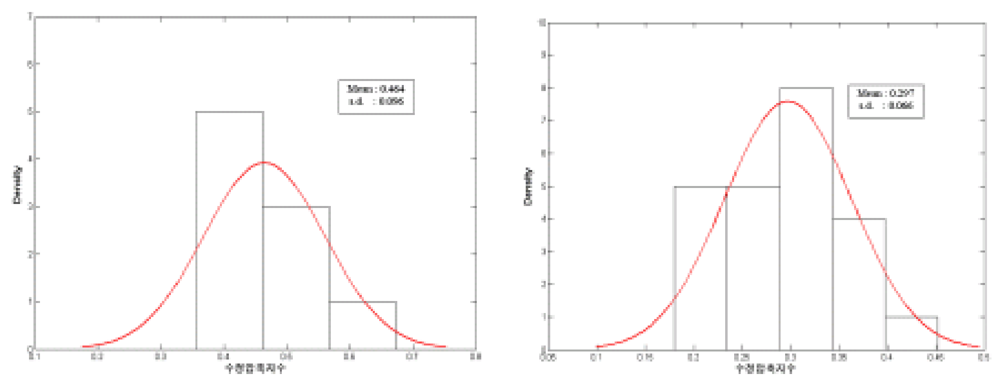


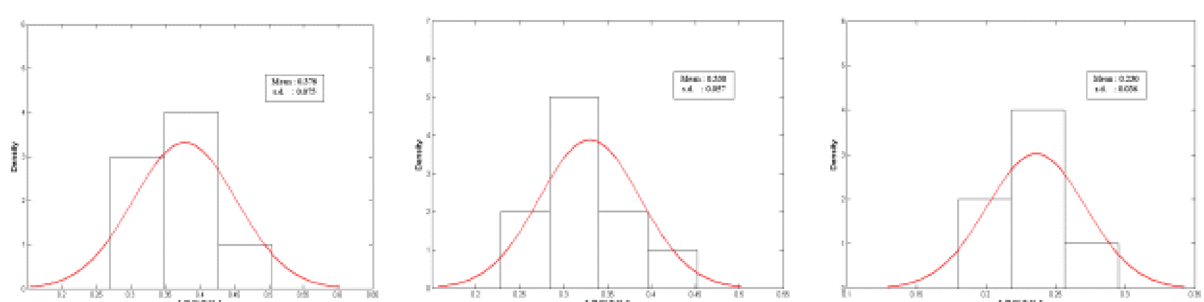
Figure 14 : Mean value and standard deviation for modified compression index (CCF) for all the site



a. 1st parameter group



b. 3rd parameter group



c. 4th parameter group

Figure 15 : Mean value and standard deviation for modified compression index (C_{CF}) for each design zone

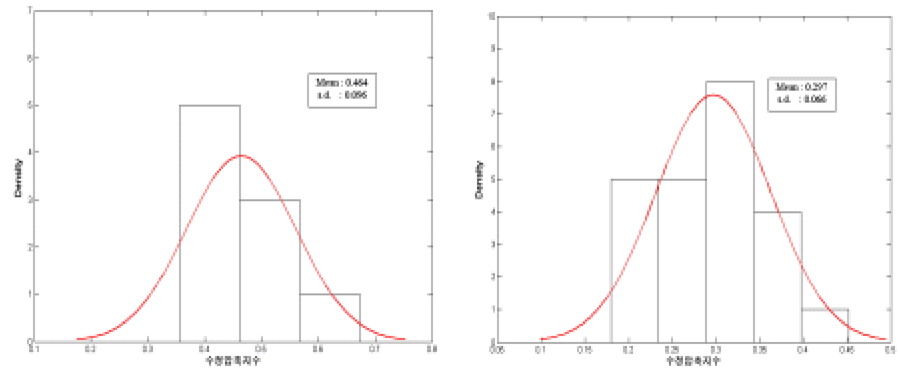


Figure 16 : Mean value and standard deviation for modified compression index (C_{CF}) for each design zone

Table 1 : Physical and consolidation properties of soils

	Parameters used	comments
1st parameter group	Specific gravity(Gs), Liquid Limit(LL), Plastic Index(PI), % passing No. 200(2μ), Soil Classification(U.S.C.S)	Considering Physical properties
2nd parameter group	Unconfined compressive strength(Qu), Modified Compressive Index(CCF), Void Ratio(eO)	Considering initial states and mechanical properties
3rd parameter group	Unconfined compressive strength(Qu), Modified Compressive Index(CCF), Liquid Limit(LL), Plastic Index(PI)	Considering physical and mechanical properties
4th parameter group	Unconfined compressive strength(Qu), Modified Compressive Index(CCF), Void Ratio(eO), Soil Classification(U.S.C.S)	Considering initial states, physical and mechanical properties

Table 2 : Mean value and standard deviation for existing design zone

	total	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7
Mean	0.350	0.429	0.419	0.326	0.399	0.321	0.338	0.289
Standard Deviation	0.107	0.083	0.167	0.117	0.080	0.056	0.086	0.055

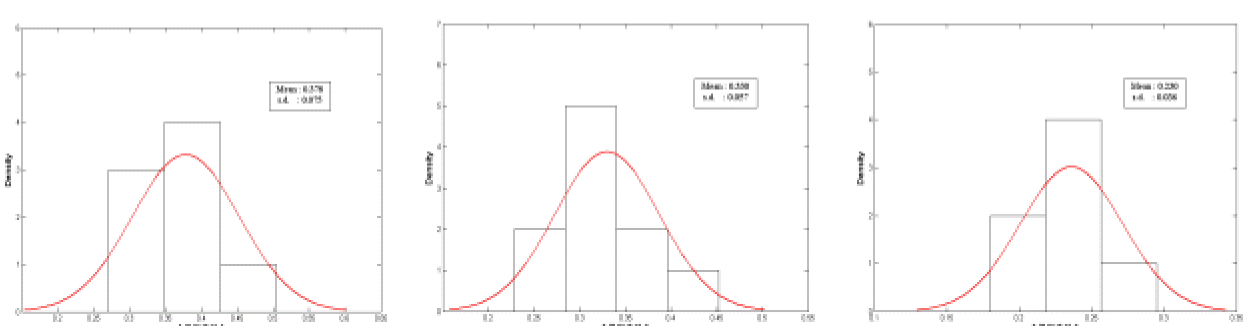


Figure 17 : Mean value and standard deviation for modified compression index (C_{CF}) for each design zone

Table 3 : Mean value and standard deviation for each parameter group

	No. of design zone	Mean value each design zone			Standard deviation for each design zone			comments
1st parameter group	2	0.450	0.301		0.101	0.071		Considering Physical properties
2nd parameter group	-	-			-			Considering initial states and mechanical properties
3rd parameter group	2	0.464	0.297		0.096	0.066		Considering physical and mechanical properties
4th parameter group	3	0.378	0.330	0.230	0.075	0.057	0.036	Considering initial states, physical and mechanical properties

This page is intentionally left blank