Comparison of Latin Hypercube Sampling and Simple Random Sampling Applied to Neural Network Modeling of HfO_2 Thin Film Fabrication

Jung Hwan Lee[a], Young-Don Ko, and Ilgu Yun
Department of Electrical and Electronics Engineering, Yonsei University, 134 Sinchon-dong, Seodaemun-gu, Seoul 120-749, Korea

Kyonghee Han
Center for the Innovation of Engineering Education, College of Engineering, Yonsei University, 134 Sinchon-dong, Seodaemun-gu, Seoul 120-749, Korea

[a]E-mail : backta@yonsei.ac.kr

(Received March 2 2006, Accepted June 16 2006)

In this paper, two sampling methods which are Latin hypercube sampling (LHS) and simple random sampling were compared to improve the modeling speed of neural network model. Sampling method was used to generate initial weights and bias set. Electrical characteristic data for HfO_2 thin film was used as modeling data. 10 initial parameter sets which are initial weights and bias sets were generated using LHS and simple random sampling, respectively. Modeling was performed with generated initial parameters and measured epoch number. The other network parameters were fixed. The iterative 20 minimum epoch numbers for LHS and simple random sampling were analyzed by nonparametric method because of their non-normality.

Keywords : LHS, Neural network, Optimization, Nonparametric test

1. INTRODUCTION

Since powerful learning algorithms such as error back-propagation(BP) learning were proposed, neural network has been applied and successfully implemented into many fields such as semiconductor manufacturing, control, image processing, and so on[1]. Ko *et al.* used neural network for modeling of PL intensity in PLD-grown ZnO thin films[2]. Hong *et al.* applied neural network modeling to reactive ion etching using optical emission spectroscopy data[3].

However neural network has two major drawbacks in its learning efficiency, such as slow learning speed and convergence to local minima[4]. Many researches have investigated and presented in how to increase the neural network's modeling speed without convergence to local minima. Abbass studied improving modeling speed using multi-objective evolutionary algorithms[5]. Yu *et al.* considered efficient BP learning using learning rate and momentum factor[6].

In this paper, initial weights and bias generating method was studied to improve modeling speed. Two methods, LHS and simple random sampling, were compared. Modeling speed is defined as the time that root mean square error (RMSE) for training data and testing data are reached at desired value. Epoch number was used as modeling speed. Epoch is defined as a single pass through a fixed training set that is going to be used repeatedly. Measured epoch numbers were analyzed by nonparametric method. MATLAB was used as modeling tool.

2. TEST STRUCTURE FABRICATION

2.1 Fabrication

HfO_2 thin film was grown on a p-type Si (100) substrate and the native oxide was chemically eliminated by (50:1) H_2O:HF solution prior to growth by Metal-organic molecular beam epitaxy. Hafnium-tetra-butoxide [Hf $(O{\cdot}t\text{-}C_4H_9)_4$] was chosen as the MO precursor because it has an appropriate vapor pressure and relatively low decomposition temperature. High –purity (99.999 %) oxygen gas was used as the oxidant. Hf-t-butoxide was introduced into the main chamber using Ar as a carrier gas through a bubbling cylinder. The bubbler was maintained at a constant temperature to supply the constant vapor pressure of Hf-source. The apparatus of the system is schematically shown in Fig. 1.

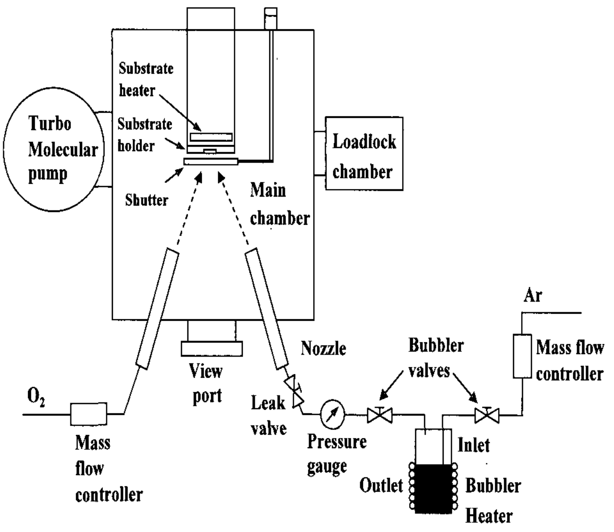


Fig. 1. The schematic of MOMBE system.

Table 1. Summary of process conditions.

Process variable	Range
Substrate temperature	450 °C ~ 550 °C
Bubbler temperature	130 °C (Fixed)
Nozzle temperature	270 °C (Fixed)
Base pressure	10^{-9} Torr
Working pressure	10^{-7} Torr
Gas flow rate (Ar)	3~5 sccm
Gas flow rate (O_2)	3~5 sccm
Growth time	30 min

High-purity Ar carrier gas passed through the bubbler containing the Hf-source. The gas line from the bubbler to the nozzle was heated to the same temperature. The mixture of Ar and metal-organic gases heated at the tip of the nozzle flows into the main chamber. The introduced Hf-source decomposed into Hf and ligand parts when it reached a substrate maintained at high temperature and the Hf ion was combined with O_2 gas supplied from another nozzle. The base pressure and working pressure were ~10^{-9} and ~10^{-7} Torr, respectively. The HfO_2 films grown by MOMBE were annealed at 700 °C for 2 min. in N_2 ambient. The process variables and conditions are listed in Table 1.

Au dots were deposited to evaluate the electrical properties of grown HfO_2 sample. The stainless shadow mask was used to make regular Au dots and the hole diameter in the mask was 0.2 mm. The determination of the electrode metal and accurate definition of electrode area has influence on the analysis of the hysteresis index of HfO_2.

Table 2. Factorial design matrix.

Run	Substrate temperature [°C]	Ar [sccm]	O_2 [sccm]
1	450	3	3
2	450	3	5
3	450	5	3
4	450	5	5
5	550	3	3
6	550	3	5
7	550	5	3
8	550	5	5
9	500	4	4

2.2 Design of experiments

Three input factors (the substrate temperature, Ar gas flow and O_2 gas flow) were explored via two level full factorial design that consists of the combinations of the high level (+) and the low level (-) with one center point to estimate the effect of process conditions. The design requires 9 runs and all experimental runs were made in random order. The experimental design matrix of input factors used in each run is summarized in Table 2. Without designed samples, two more samples were made to use test data. The hysteresis index is used as output factor. The hysteresis index is defined as the width of the hysteresis loop generated by the bi-directional voltage sweep.

3. MODELING SCHEME

Neural network based modeling of the hysteresis index was performed with 10 initial weights and bias set.

Initial parameter sets were generated using Latin Hypercube Sampling (LHS) and simple random sampling, respectively.

Modeling was performed until RMSE for training data is reached 0.03 or epoch number is 100000 for each initial weights and bias set.

Without initial weights and bias, network parameters are fixed to compare effect of LHS with simple random sampling. Fixed neural network parameters are summarized

Table 3. Summary of parameters.

Network Parameters	
Network structure	3-4-3-1
Network learning rate	0.003
Network momentum	0.04

in Table 3. Modeling was repeated 20 times and measured minimum epoch number.

3.1 Latin Hypercube sampling

The Latin Hypercube sampling (LHS) is used in this study to select randomized values for the weights and the biases, which are parameters of neural networks. The LHS method is a stratified sampling technique where the random variable distributions are divided into equal probability intervals. The LHS method generates a sample size *N* from the *n* variables. A 1/*N* probability is randomly selected from within each interval that is partitioned into *N* nonoverlapping ranges for each basic event[7].

3.2 Neural networks

Neural networks are utilized to model the nonlinear relationship between inputs and outputs in many fields. The networks consist of the three layers that are the input layer, the hidden layer and the output layer. That is comprised of simple processing units called neurons, interconnection, and weights that are assigned to the interconnection between neurons[8]. Each neuron contains the weighted sum of its inputs filtered by a nonlinear sigmoid transfer function. The schematic of general feed-forward neural networks are shown in Fig. 2.

The neural networks in this work carried out with the error BP algorithm. The error BP neural networks consist of several layers of neurons which receive, process and transmit critical information regarding the relationships between the input parameters and corresponding responses.

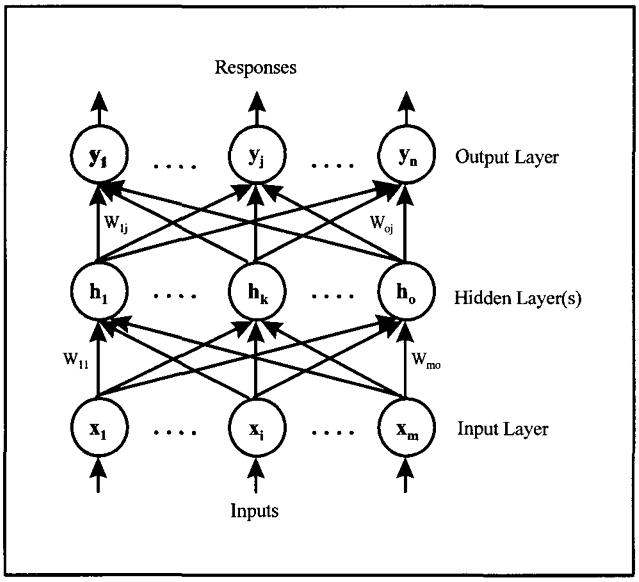


Fig. 2. Typical feed-forward neural networks.

Generally, the weight mechanism of the BP algorithm is defined by the following [9]:

$$w_{ijk}(n+1) = w_{ijk(n)} + \eta \Delta w_{ijk(n)} \tag{1}$$

where w_{ijk} is the connection strength between the j th neuron in the layer $(k-1)$ and the i th neuron in layer k, Δw_{ijk} is the calculated change in that weight which reduces the error function of the networks, and η is the learning rate.

This algorithm has been shown to be very effective in learning arbitrary nonlinear mappings between noisy sets of input and output factors.

In this study, neural network was trained on 9 experimental runs. The two trials were used for testing data in order to verify the fitness of the neural network outputs for the results of the training data. If RMSE for testing data is more than 0.03, the model was considered as over fitting model and the epoch number was not counted.

4. RESULTS

Modeling experiment was performed 20 times and each time minimum value of 10 epoch number which are result of modeling using 10 initial weights and bias was measured.

The minimum epoch number for LHS and simple random sampling are illustrated in Fig. 3 where the rounds represent minimum epoch number of LHS and the squares represent minimum epoch number of simple random sampling.

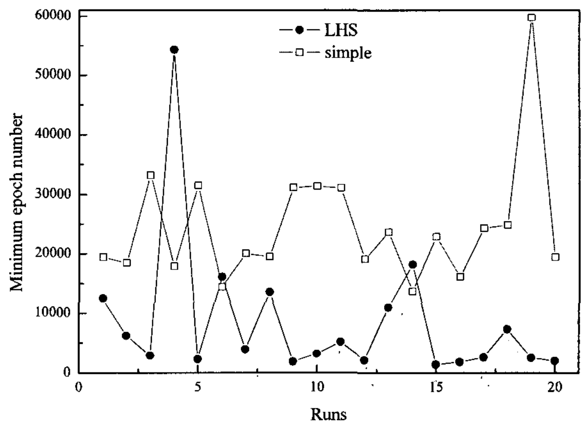


Fig. 3. Minimum epoch number.

P-values of Kolmogorov-Smirnov test which is one of normality test for LHS data and simple random sampling were 0.06 and 0.00, respectively. It indicates that the data are not sampled from a normal distribution. Because of their non-normality, minimum epoch number was analyzed by nonparametric method to compare LHS effect with simple random sampling effect in neural network.

Nonparametric method is a statistical method that allows the functional form of a fit to data to be obtained in the absence of any guidance or constraints from theory. Analyzed result is shown in Table 4.

For one side test, null hypothesis is $M_L \geq M_S$ and alternative hypothesis is $M_L < M_S$, where M_L is Median of minimum epoch number for LHS method and M_S is Median of minimum epoch number for simple random sampling method. One side P-value indicates that null hypothesis is strongly rejected and minimum epoch number for LHS is smaller than simple random sampling.

Because of property of LHS, Minimum modeling speed for LHS method is faster than simple random sampling. Unlike the simple random sampling, the LHS method can describe a full coverage of the sampling range by maximally satisfying each marginal distribution. The distributions of sampling with respect to the selecting method are illustrated in Fig. 4.

Table 4. Nonparametric test result.

	Value
U value of Mann-Whitney test	25.00
W value of Wilcoxon test	235.00
Two side P-value	0.000
One side P-value	0.000

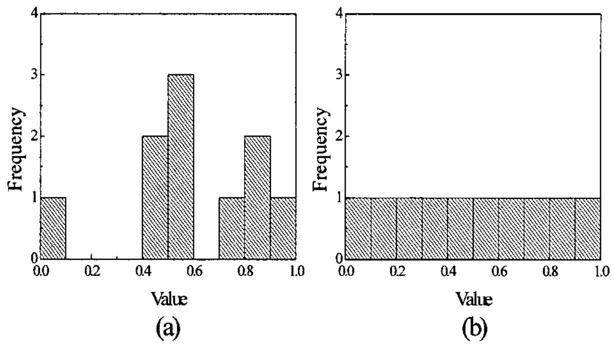


Fig. 4. Two difference distributions of the sampling values: (a) the simple random sampling and (b) LHS.

The 10 samples were generated in the range of 0 to 1. It is presented that the sampling values of the LHS method are uniformly distributed comparing to that of the random sampling. Therefore, the unbiased random values of the weights and biases for the neural networks were selected via the LHS method.

5. CONCLUSION

Comparison of LHS and simple random sampling methods with respect to the modeling speed and accuracy for neural networks with generating initial weights and bias was investigated. As an application example, the electrical characteristic for HfO_2 thin film was modeled by two suggested methods and compared the results for the modeling speed and accuracy. Minimum epoch number was used as modeling speed and the number was analyzed by nonparametric method. Using LHS method was faster than using simple random sampling since the sampling values have a uniform distribution for LHS.

ACKNOWLEDGMENTS

This work was supported by Yonsei University Research Fund of 2005.

REFERENCES

[1] Y.-J. Chen, T.-C. Huang, and R.-C. Hwang, "An effective learning of neural network by using RFBP learning algorithm", Inform. Sci., Vol. 167, p. 77, 2004.

[2] Y.-D. Ko, H. S. Kang, M.-C. Jeong, S. Y. Lee, J.-M. Myoung, and I. Yun, "Neural network based modeling of PL intensity in PLD-grown ZnO thin films", J. Mater. Process. Tech., Vol. 159, p. 159, 2005.

[3] S. J. Hong, G. S. May, and D.-C. Park, "Neural network modeling of reactive ion etching using optical emission spectroscopy data", IEEE Trans. Semi. Manufac., Vol. 16, No. 4, p. 598, 2003.

[4] Y. Lee, S.-h. Oh, and M. W. Kim, "The effect of initial weights on premature saturation in back-propagation learning", IJCNN-91-Seattle International Joint Conference on, p. 765, 1991.

[5] H. A. Abbass, "Speeding up backprobagation using multiobjective evolutionary algorithms", Neural Comput., Vol. 15, p. 2705, 2003.

[6] X.-H. Yu and G.-A. Chen, "Efficient backpropagation learning using optimal learning rate and momentum", Neural Netw., Vol. 10, No. 3, p. 517, 1997.

[7] J. F. Swidzinski and K. Chang, "Nonlinear statistical modeling and yield estimation technique for use in Monte Carlo simulations", IEEE Trans. Microw. Theory Tech., Vol. 48, p. 2316, 2000.

[8] G. S. May, "Manufacturing ICs the neural way", IEEE Spectr., Vol. 31, p. 47, 1994.

[9] C. H. Chen, "Fuzzy Logic and Neural Network Handbook", McGraw-Hill, 1996.

(2 line spacing)

Instructions for the Preparation of Article for Transactions on Electrical and Electronic Materials (bold, 15 points)

(1 line spacing)

Gil-Dong Hong[a], Gap-Soon Lee, and Nor-Ree Han
Department of Electrical and Electronic Engineering, College of Engineering, Yonsei University, 134 Shinchon-dong, Seodaemoon-ku, Seoul 120-749, Korea
(1 line spacing)
P. de Smith and Helew Pack
Department of Electrical Engineering, Old Dominion University, Norfolk, VA 23529, U.S.A.
(1 line spacing)
[a]E-mail : gdhong@ocbc.ac.kr

(1 line spacing)

(Received 2 January 2002, Accepted 15 February 2002)

(1 line spacing)

These pages provide you with an example of the layout and style which we wish you to adopt during the preparation of your paper. Make the width of abstract to be 14cm.

(1 line spacing)

Keywords : List less than 5 keywords related to this article

(2 line spacing)

1. FORMAT

(1 line spacing)

We recommend to use MS Word processor and prepare text within the dimensions shown on these pages; In A4 paper, left and right margin are 16mm respectively, 8mm middle margin, 35mm top margin, and 27mm bottom margin. When a paragraph starts, give an indent of 2 characters. In the last page of the article, make the length of left and right stage to be equal approximately.

Make use of the maximum stipulated length apart from the following two exceptions (i) do not begin a new section directly at the bottom of a page, but transfer the heading to the top of the next column; (ii) you may exceed the length of the text area by one line only in order to complete a section of text or a paragraph.

(1 line spacing)

1.1 Spacing

We normally recommend the use of 1.0 (single) line spacing. However, when typing complicated mathematical text it is important to increase the space between text lines in order to prevent sub- and super-script fonts overlapping one another and making your printed matter illegible.

(1 line spacing)

1.2 Fonts

These instructions have been produced using a 10.5 point Times Roman font. Title and subtitle are written in bold-faced characters.

(2 line spacing)

2. PRINTOUT

(1 line spacing)

Please make use of good quality plain white A4 paper size. Here we demonstrate a problem which we often experience with computer printout. Printers sometimes produce text which contains light and dark streaks, or has considerable lighting variation either between left- and right-hand margins or between text heads and bottoms. To achieve optimal reproduction quality, the contrast of text lettering must be uniform, sharp, and dark over the whole page and throughout the article.

If corrections are made to the printout, run-off completely new replacement pages. The contrast on these pages should be consistent with the rest of the paper as should text dimensions and font sizes.

(2 line spacing)

3. TABLES AND ILLUSTRATIONS

(1 line spacing)

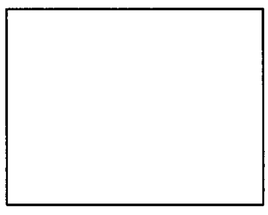
Tables and illustrations should be originals or sharp prints. They should be arranged throughout the text preferably being included on the same page as they are first discussed. They should have a self-contained caption and be positioned in center margin within the column. If they do not fit into one column they may be placed across both columns in which case place them at the top or at the bottom of a page.

(1 line spacing)

3.1 Tables

Tables should be presented in the form shown in Table1. Their layout should be consistent throughout. Horizontal lines should be placed above and below table headings, above the subheadings and at the end of the table above any notes. Vertical lines should be avoided. If a table is too long to fit onto one page, the table number and headings should be repeated on the next page before the table is continued.

(1 line spacing)



(1 line spacing)

Fig. 1. Good quality figure with clear lettering.

Table 1. Periodic table of elements.

(2line spacing)

(1 line spacing)

Fig. 2. Bad quality, distorted figure; lettering is too small.

(1 line spacing)

3.2 Line drawing

Line drawings should be drawn in India ink on tracing paper with the aid of a stencil or be glossy of the same, if they have not been prepared on your computer facility. They should be attached to your manuscript page, correctly aligned. All illustrations should be clearly displayed by leaving at least a single line of spacing above and below them. When placing a figure at the top of a page, the top of the figure should be at the same level as the bottom of the first text line of the other column.

All notations and lettering should be no less than 2 mm high. The use of heavy black, bold lettering should be avoided as this will look unpleasantly dark when printed.

(1 line spacing)

3.3 Black and white photographs

Photographs must always be sharp originals (*not screened versions*) and rich in contrast. They should be pasted on your page in the same way as line drawings.

(1 line spacing)

3.4 Color photographs

Sharp originals should be submitted close to the size expected in publication. Charges for the processing and printing of color will be passed on to the author(s) of the paper. As the costs involved are per page, care should be taken in the selection of size and shape so that two or more illustrations may be fitted together on one page.

(2 line spacing)

4. EQUATION

(1 line spacing)

Equations are placed in center and should be preceded and followed by one line of white.

(1 line spacing)

$$H_{\alpha\beta}(\omega) = E_{\alpha}(0)\,\delta_{\alpha\beta} + <\alpha:W_{\pi}:\beta> \qquad (1)$$

(1 line spacing)

If they are numbered, make sure that they are numbered consecutively. Place the numbers in parentheses. Flush with the right-hand margin of the column and level with the last line of the equation.

(2 line spacing)

ACKNOWLEDGMENTS

(1 line spacing)

This work was supported by the Ministry of Science and Technology though the Nano-Structure Technology Project.

(2 line spacing)

REFERENCES

(1 line spacing)

[1] D. A. Neamen, "Semiconductor Physics and Devices", Irwin, p. 10, 1997.

[2] T. W. Choi, "Electrical and mechanical properties of ceramics", Bulletin of KIEEME, Vol. 15, No. 1, p. 10, 2001.

[3] T.-W. Choi and S.-C. Yoo, "Electrical and mechanical properties of ceramics", J. of KIEEME(in Korean), Vol. 15, No. 1, p. 10, 2001.

[4] T. W. Choi, C. S. Lee, and S. C. Yoo, "Electrical and mechanical properties of ceramics", Trans. EEM, Vol. 15, No. 1, p. 10, 2001.

[5] Tae Wuk Choi and Sang Chul Yoo, "Electrical and mechanical properties of ceramics", J. Mater. Sci., Vol. 15, No. 1, p. 10, 2001.

[6] T. W. Choi, "Electrical properties of ceramics", Korea Report, No. KR-R250, p. 10, 2001.

[7] T. W. Choi and S. C. Yoo, "Electrical ceramics", SID'95 digest paper, p. 10, 1995.

[8] T. W. Choi and S. C. Yoo, "Electrical ceramics", Proc. 2002 Summer Conf. KIEEME, p. 10, 2002.

[9] T. W. Choi, "Electrical properties of ceramics", US Patent, 1,234,567, 2001.

Contributions

Manuscripts for publication should be sent in triplicate (along with the electronic form) to the Korean Institute of Electrical and Electronic Material Engineers. Manuscripts from countries outside Korea may be submitted through the appropriate international editors depending on the subject area, or directly to the Korean Institute of Electrical and Electronic Material Engineers.

A Submitted Manuscript is accepted with the understanding that the manuscript is not currently under consideration by another journal and that it has not been copyrighted or accepted for publication elsewhere.

General Style. Manuscripts must be prepared in accordance with the general requirements that are summarized in Instruction for Authors for Publication published in the back of the first issue of every volume.

Alteration in Proof. A limited number of alterations in proof are allowed, but the cost of making extensive corrections and changes after an article has been set in type will be charged to the author. Proofs and all correspondence concerning papers in the process of publication should be returned to the same address as that to which the initial manuscript was sent.

Subscriptions, renewals, and changes of address should be addressed to the Office of the Korean Institute of Electrical and Electronic Material Engineers. For a change of address, please send both the old and new addresses.