RESEARCH ARTICLE

Prediction of Statistical Distribution on Nanosheet FET by Geometrical Variability Using Various Machine Learning Models

JONGHYEON HA[1], (Graduate Student Member, IEEE), **SUN JIN KIM**[2], **MINJI BANG**[1],
GYEONGYEOP LEE[1], **MINKI SUH**[1], (Graduate Student Member, IEEE),
MINSEOB SHIM[3], (Member, IEEE), **CHONG-EUN KIM**[4], (Member, IEEE),
AND JUNGSIK KIM[1], (Senior Member, IEEE)

[1]Department of Electrical Engineering, Gyeongsang National University (GNU), Jinju 52828, Republic of Korea
[2]Smart Structural Safety and Prognosis Research Division, Korea Atomic Energy Research Institute, Daejeon 34057, Republic of Korea
[3]Department of Electronic Engineering, Gyeongsang National University (GNU), Jinju 52828, Republic of Korea
[4]Department of Railroad Electrical and Electronic Engineering, Korea National University of Transportation, Uiwang 16106, Republic of Korea

Corresponding author: Jungsik Kim (jungsik@gnu.ac.kr)

ABSTRACT Due to the aggressive scaling down of logic semiconductors, the difficulty of semiconductor component processes has increased. As the structure of components becomes more complex, the time and cost of processes and simulations have risen. Machine learning is now being used to analyze the electrical characteristics data of semiconductor components and apply the trained machine learning to next-generation semiconductor development. Machine learning trained on process data and simulation results can quickly and accurately predict which electrical characteristics change significantly when the component's structure changes and which parameters have a significant impact on the electrical characteristic changes. This paper presents suitable machine learning models for analyzing and predicting the electrical characteristics (on-current (I_{on}), off-current (I_{off}), threshold voltage (V_{th}), subthreshold swing (SS), and drain induced barrier lowering (DIBL)) and statistical distribution (mean and standard deviation of the electrical characteristics) resulting from geometrical variability (sheet thickness (T_{wire}), sheet diameter (D_{wire}), oxide thickness (T_{ox}), gate length (L_g), spacer length (L_{sp}), gate metal work-function (WF)) in nanosheet field-effect transistor (NSFET), which are a next-generation logic device. Machine learning models, including regulation-based models (Ridge and LASSO) and tree-based models (decision tree (DT), random forest (RF), extreme gradient boost (XGBoost), and light gradient boost machine (LGBM)), are trained on technology computer-aided design (TCAD) simulation data. The LGBM more accurately predicts the electrical characteristics and statistical distribution of the NSFET than the other models. Additionally, we analyze the effect of geometrical variability on the NSFET based on feature importance.

INDEX TERMS Nanosheet field-effect transistor, prediction, statistical analysis, technology computer-aided design simulation, machine learning.

I. INTRODUCTION

Nanosheet field-effect transistors (NSFETs) were developed to replace FinFETs, which have reached the scaling-down

limits. NSFETs, which comprise a surrounding gate, present a higher gate controllability than FinFETs. Consequently, NSFETs suppress the short-channel effect (SCE) more effectively than FinFETs [1], [2], [3]. However, as logic devices become miniaturized and technologies evolve, the associated process challenges increase. Problems arise at the front end

The associate editor coordinating the review of this manuscript and approving it for publication was Mohammad Hossein Moaiyeri.

of the line, such as local variability due to work function variation (WFV), line edge roughness (LER), and random dopant fluctuation (RDF). Furthermore, global variability issues, specifically critical dimension problems, are observed at the back end of the line [4], [5], [6], [7], [8]. These issues result in a wide distribution of electrical characteristics on wafers and lower yields. In addition, geometrical variability, such as the thickness and width of the sheet in the NSFETs, results in a wide distribution of electrical characteristics [9], [10]. To increase the yield of wafers and achieve a narrow distribution of electrical characteristics, the statistical distribution of the wafer is predicted using a simulation prior to the device fabrication process. However, as the device volume decreases, the time required for the simulation increases because the quantum-mechanical phenomena of the carriers must be considered [11], [12], [13]. Machine learning should be utilized to reduce the device processing and simulation costs and to determine the statistical distribution of the device in the wafer with high accuracy and speed. In addition, the device geometrical variability with the greatest effect on the statistical distribution of the electrical characteristics must be predicted and the importance of geometric variability must be analyzed.

The electrical characteristics of logic devices have been extensively investigated using machine learning. In another study, the effects of point defects caused by cosmic radiation on FinFETs using machine learning have been analyzed [14]. Machine learning models such as random forest (RF), extreme gradient boost (XGBoost), and light gradient boost machine (LGBM) have been used to characterize logic semiconductor devices [15], [16], [17], [18]. However, the statistical distribution of electrical characteristics caused by geometrical variability in NSFETs has not been compared using various machine learning models. Therefore, the prediction accuracy of machine learning models must be compared and a model suitable for analyzing next-generation logic devices must be identified.

In this study, we use machine learning to predict the electrical characteristics and statistical distribution caused by device geometrical variability (sheet thickness (T_{wire}), sheet diameter (D_{wire}), oxide thickness (T_{ox}), gate length (L_g), spacer length (L_{sp}), gate metal work-function (WF)). Two types of machine-learning models are used: regulation-based (LASSO, Ridge) and tree-based models (decision tree (DT), RF, XGBoost, and LGBM). The six machine-learning models, which were trained using technology computed-aided design (TCAD) simulation data, predicted the electrical characteristics (on-current (I_{on}), off-current (I_{off}), threshold voltage (V_{th}), subthreshold swing (SS), and drain-induced barrier lowering (DIBL)) and statistical distribution (mean and standard deviation of the electrical characteristics). The feature (parameter of geometrical variability) importance is extracted and analyzed using the machine learning model with the highest prediction rate to understand the effect of geometrical variability on the statistical distribution of electrical characteristics.

II. SIMULATION MODELING METHODOLOGY

TCAD tool of Synopsys was used to generate an NSFET dataset [19]. The control device was calibrated based on the abovementioned dataset as a reference owing to the realistic description of the NSFET (**Fig. 1 (a)**) [1]. The I_d–V_g of the calibrated control device is indicated by the red lines in **Figs. 1 (b)** and **(c)**. The channel, substrate, and source/drain (S/D) doping concentrations of the control device were 10^{17}, 10^{17}, and 10^{20} cm^{-3}, respectively. The process parameters of the control device are indicated in bold in Table 1. To consider the physical phenomena in the logic device, physic-models were initialized. The inversion and accumulation layer mobility (IALMob) was used to reflect the carrier mobility in the inversion and S/D extension layers within the channel [20]. Carrier recombination at the recombination center in the energy band gap of silicon, which is an indirect material, was applied using the Shockley–Read–Hall and Auger models. The Hurkx tunneling model was used to consider gate-induced drain leakage [20], [21], [22]. As the channel volume decreased, discontinuous sub-bands of silicon atoms appeared in the channel. The quantum mechanical phenomenon observed when carriers were inverted and transported in the sub-band near the Si–SiO$_2$ interface was represented using modified local density approximation (MLDA) [23].
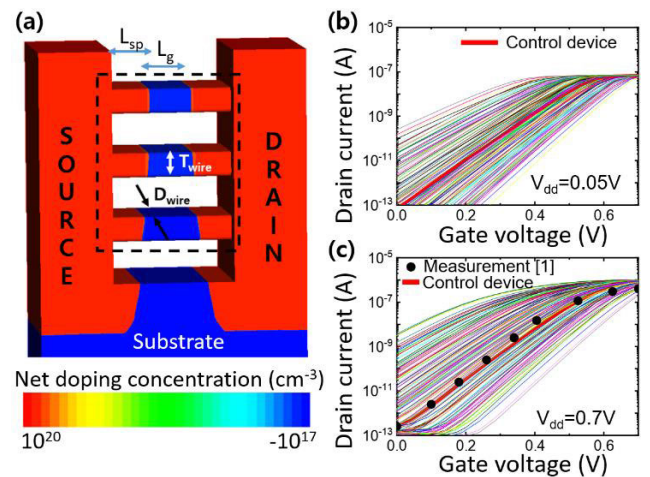

FIGURE 1. (a) Structure of nanosheet field effect transistor (NSFET). I_d–V_g graphs of 1728 data in (b) linear regime (V_{dd} = 0.05 V) and (c) saturation regime (V_{dd} = 0.7 V, Simulated I_d–V_g transfer characteristics at V_{dd} = 0.7 V compared with measurement data).

We generated 1728 NSFET datasets (the number of datasets = 4 × 3 × 3 × 3 × 4 × 4 = 1728) in each regime (linear (V_{dd} = 0.05 V) and saturation (V_{dd} = 0.7 V)) by classifying the geometrical variability relative to the calibrated control device. The range of variation of the five geometrical parameters and the gate WFs are shown in Table 1. The total of 3456 NSFET datasets were used for training and testing the six machine learning models (Five months were required to create 3456 TCAD data). The six machine learning models were organized into two categories:

TABLE 1. Geometrical parameters and work function of NSFET.

Symbol	Quantity	Range (Unit)
L_g	Physical gate length	10, 11, **12**, 13 (nm)
L_{sp}	Spacer length	**4**, 5, 6 (nm)
EOT	Equivalent oxide thickness	1.0, **1.5**, 2.0 (nm)
WF	Gate metal work function	**4.86**, 4.90, 4.94 (eV)
T_{wire}	Sheet thickness	**5**, 6, 7, 8 (nm)
D_{wire}	Sheet diameter	16, 18, **20**, 22 (nm)

regulation- and tree-based models. The regulation-based models, i.e., Ridge and LASSO, perform linear regression training using second-order and first-order formula models, respectively [24]. Decision trees (DT) refer to models that predict output values based on decision rules for various combinations of given input values [25]. Random forest (RF) is a type of ensemble learning method used for tasks such as classification and regression analysis [26]. It operates by producing classifications (for classification tasks) or average prediction values (for regression analysis) from multiple decision trees created during the training process. Extreme gradient boosting (XGBoost) is a model that significantly enhances gradient boosting. It combines estimates from a simpler and weaker set of models to accurately predict the target variable [27]. Light gradient boosting machine (LGBM) is a gradient boosting framework based on decision trees that enhances model efficiency and reduces memory usage, resulting in high prediction accuracy, alongside XGBoost [28]. The ''Results and Discussion'' section provides a detailed description of each machine-learning model. Each machine learning model was trained using optimal hyperparameters. We normalized 3456 training data using the min–max scaler. The training data were set to train the machine learning using 10 %, 30 %, 50 %, 70 %, and 90 % of the total data (training size = 10 %, 30 %, 50 %, 70 %, and 90 %). We used the mean absolute percentage error (MAPE), which is a metric that evaluates the percentage of prediction loss, to determine the error of the machine learning model in predicting the electrical characteristics and statistical distribution based on training and test data. The specifications of the hardware used for simulation and machine learning were as follows: Intel CPU (4-threads (3.60 GHz, i9-9900K)), 4 GB DRAM.

III. RESULT AND DISCUSSION

Fig. 2 shows the variation in V_{th} as the four geometric parameters and WF vary in the linear and saturation regimes. The trends in the V_{th} with T_{wire}, D_{wire}, T_{ox}, L_g, and WF were the same in both regimes. In conditions where the D_{wire} increases, the gate capacitance increases, leading to an increase of the space charge in the channel and inverted carriers. The increased gate capacitance, space charge, and number of inverted carriers reduce the device's V_{th}. Addi-

tionally, an increase in the cross-sectional area of the channel lowers its resistance. The reduced resistance enhances current conduction characteristics, increasing current. Conversely, when the D_{wire} value remains fixed, increasing T_{wire} reduces the impact of the gate bias field at the center of the channel. As the influence of the field at the channel center decreases, the space charge within the channel and inverted carriers decreases. This leads to a reduction in gate capacitance. A higher gate voltage is required to generate sufficient space charge within the channel for inverted carriers, increasing the device's V_{th}. This, in turn, weakens gate controllability and causes higher leakage current and subthreshold swing within the channel. A decrease in L_g reduces the voltage required to accumulate charge in the channel and lowers channel resistance, thereby reducing the V_{th}. With a decrease in L_g, drain bias increases, leading to V_{th} roll-off phenomena and enhanced short channel effects (SCE), ultimately impairing gate controllability. When T_{ox} decreases, the oxide capacitance increases, and gate controllability rises as the gate bias field on the channel increases. This reduces V_{th} due to the effective inversion of carriers within the channel at lower voltages. As gate WF increases, the accumulation of positive charges near the Si-SiO$_2$ interface is increased at the off-state. This decreases the voltage required to form a flat band in the channel, ultimately increasing the threshold voltage. **Fig. 3** shows the distribution of electrical characteristics in the linear and saturation regimes. The means and standard deviations of the electrical characteristics in the saturation regime are listed in Table 2.

TABLE 2. Statistical distribution of NSFET.

TCAD	I_{offsat} (nA)	I_{dsat} (μA)	V_{ths} (V)	SS (mV/dec)	DIBL (V)
Mean	0.223	0.843	0.389	100.7	0.107
Std. deviation	1.192	0.202	0.123	39.11	0.054
Min	2.09E-06	0.089	0.001	61.2	0.023
Max	14.1	1.01	0.623	350	0.306

The six machine learning models were used with 1728 datasets in linear and saturation regimes to predict the electrical characteristics. After training the training data, the test was conducted using 10 %, 30 %, 50 %, 70 %, and 90 % of the TCAD simulation data (1728 datasets). **Figs. 4** and **5** show the prediction loss, which occurs to predict the electrical characteristics. The MAPE saturated above the trainset size of 70 %. In the case with the trainset size of 90 %, undesired overfitting occurred during training because most of the data were in the plot, which reduced the prediction accuracy. Therefore, using the trainset size of 70 % was suitable.

Based on a trainset size of 70 % with a low MAPE, electrical characteristics prediction compares to the TCAD simulation data. **Figs. 6 (a)** and **(b)** show the scatter plots of the predicted electrical characteristics in the linear and
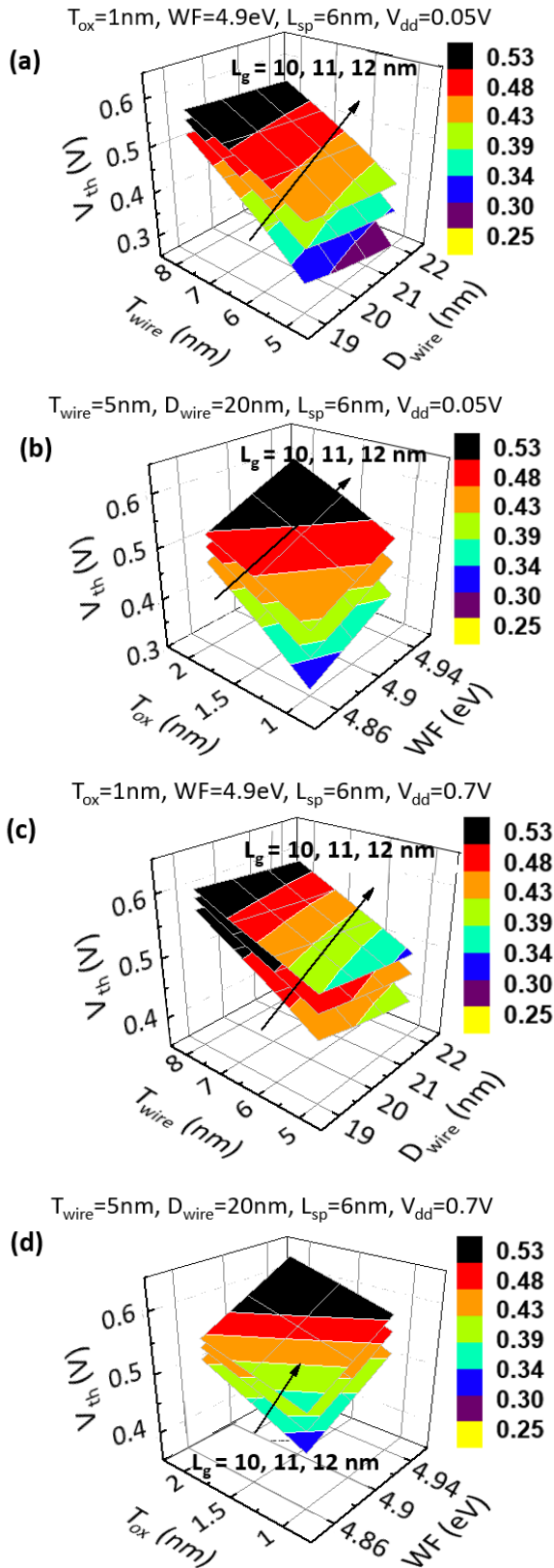
FIGURE 2. V_{th} fluctuations with varying structure parameters. (a) T_{wire}, D_{wire}, and L_g, and (b) T_{ox}, WF, and L_g fluctuation in linear regime. (c) T_{wire}, D_{wire}, and L_g, and (d) T_{ox}, WF, and L_g fluctuation in saturation regime.
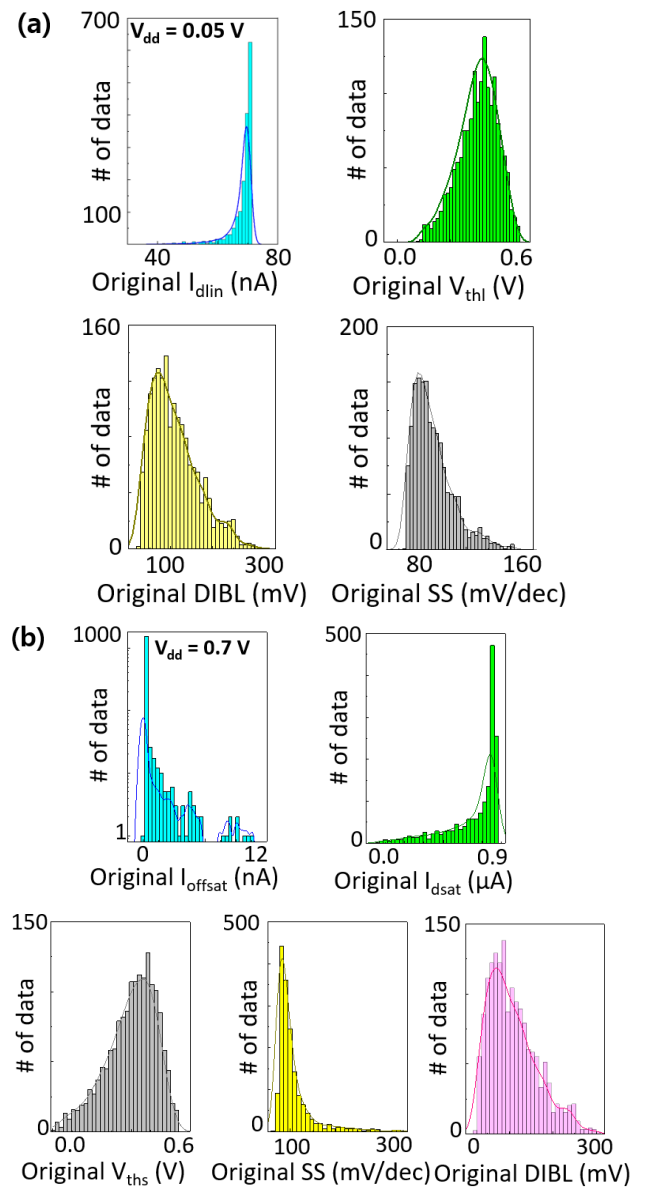

FIGURE 3. Statistical distribution of electrical characteristics in (a) linear regime and (b) saturation obtained via technology computer-aided design (TCAD) simulation.

saturation regimes, respectively. The X-axis represents the TCAD prediction of the electrical characteristics under geometry variability. The Y-axis indicates the predicted electrical characteristics using machine learning. The Ridge and LASSO models demonstrated much lower prediction accuracies than the tree-based models. The Ridge and LASSO models were regulated to prevent overfitting [24]. An ideal linear regression uses data around the mean and generalizes them to exclude noise. The LASSO model trains a linear regression using a polynomial of degree one (linear function). Ridge uses a polynomial of degree two (quadratic function)
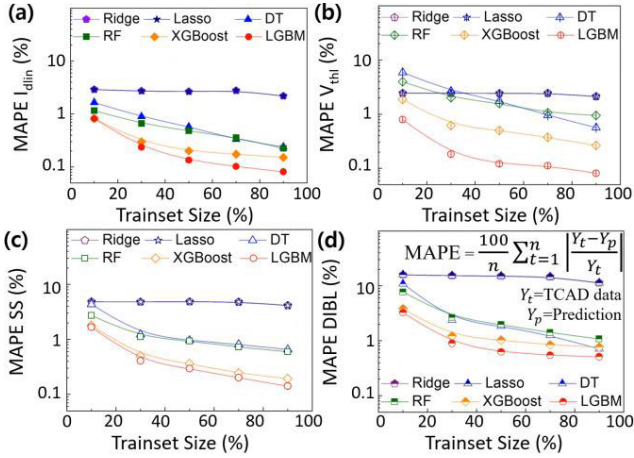
FIGURE 4. Mean absolute percent error (MAPE) based on training dataset size for each machine learning model in linear regime.
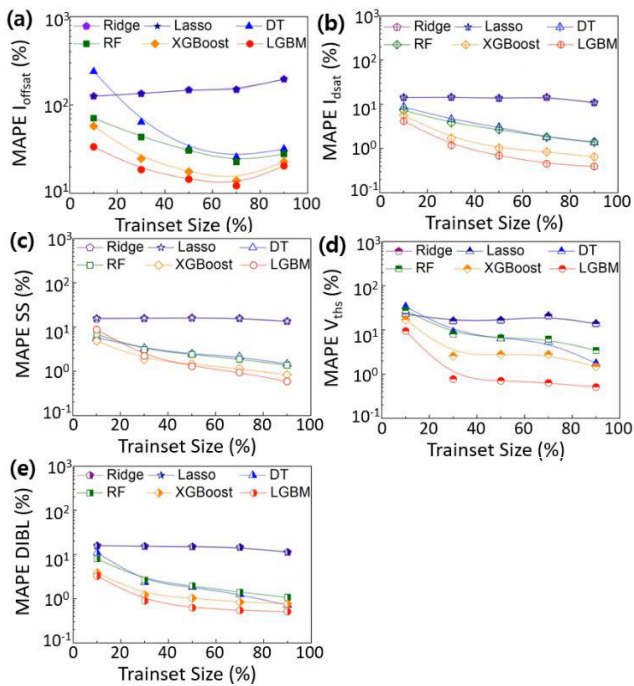

FIGURE 5. Mean absolute percentage error (MAPE) based on training dataset size for each machine learning model in saturation regime.

FIGURE 6. Comparison of scatter plots and correlations of electrical characteristics between the TCAD data and ML approach in the (a) linear and (b) saturation regimes (The prediction accuracy of each model is shown as an R^2 score in the lower right corner of the graph).

FIGURE 7. Effects of variation in geometry parameters on electrical characteristics in (a) linear and (b) saturation regime.

by excluding terms exceeding degree two in the prediction formula. However, prediction of the electrical characteristics using linear and quadratic functions is difficult owing to the various types of geometrical variability. Consequently, the Ridge and LASSO models became underfitting in the linear and saturation regimes, and the prediction of electrical characteristics, which is predicted by Ridge and LASSO models, deviated from $y = x$. Tree-based models offer a higher prediction accuracy than regulation-based models. The DT categorizes data through a single tree and continuously generates branches in a downward direction [25]. However, unlike other tree-based models, the DT presents an asymmetrical
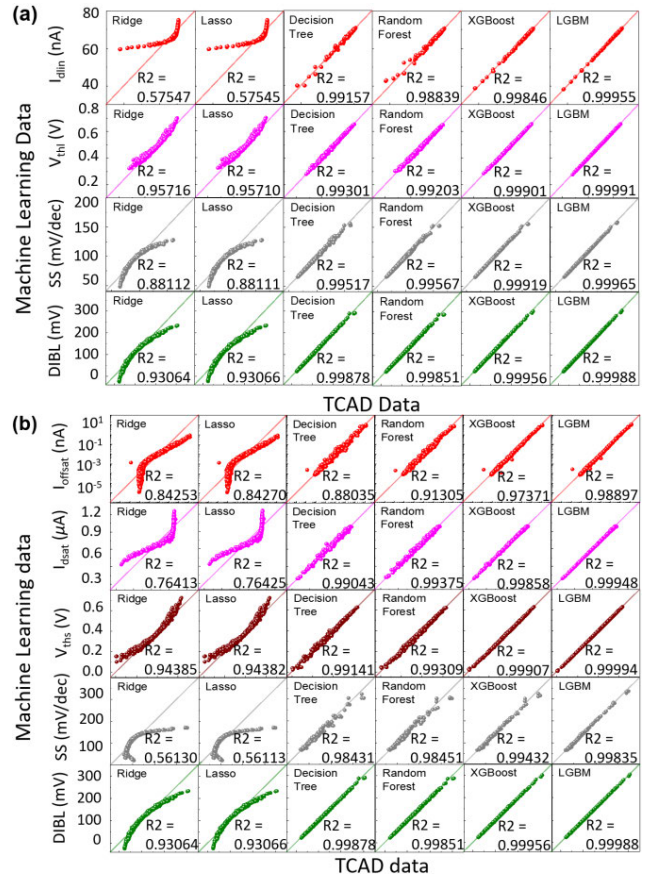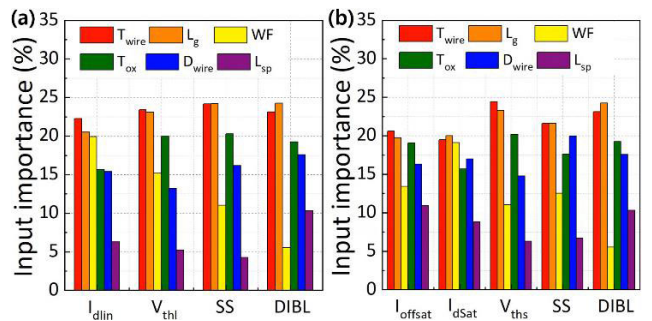
tree shape. Pruning variables such as the maximum depth of the tree and the limit on the number of nodes in the tree can be set in the hyperparameters of the DT model to prevent asymmetric tree growth. However, limiting the maximum depth of the tree and number of nodes can reduce the predictive accuracy of machine learning. Because of this tradeoff, the DT has a relatively lower prediction accuracy than other tree-based models. RF is an ensemble learning method that employs various learning algorithms. The RF forms various decision
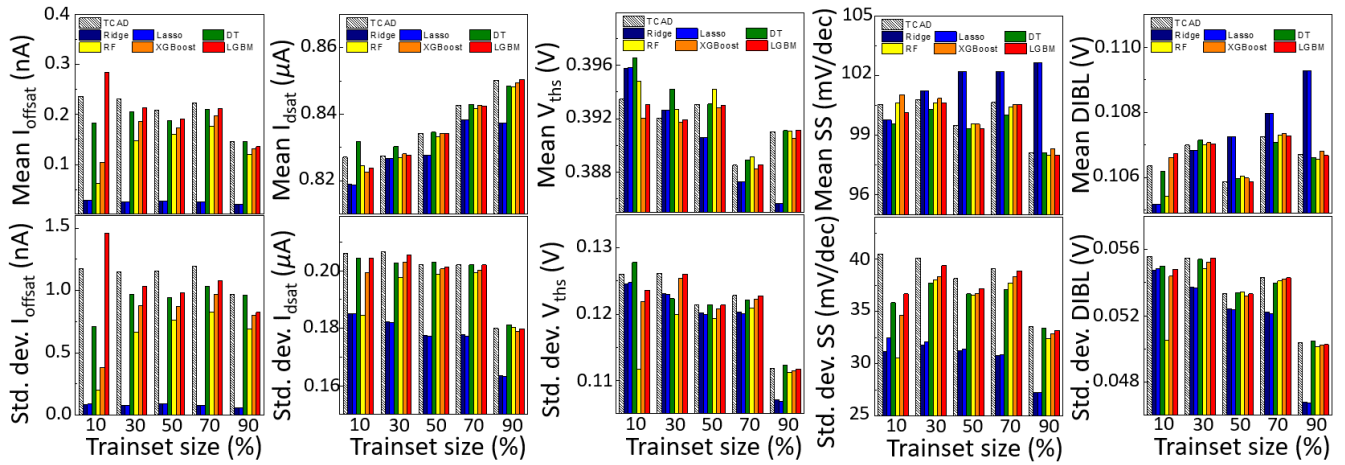
FIGURE 8. Statistical distribution (mean and standard deviation (std. dev.)) of electrical characteristics predicted using machine learning model vs. TCAD.

	Parameters	ML Models						
		TCAD	Ridge	Lasso	DT	RF	XGBoost	LGBM
Mean	I_{offsat} (nA)	0.223	0.026	0.025	0.210	0.176	0.198	0.212
	I_{dSat} (µA)	0.843	0.838	0.838	0.843	0.842	0.843	0.842
	V_{ths} (V)	0.389	0.387	0.387	0.389	0.389	0.388	0.389
	SS (mV/dec)	100.7	102.2	102.2	100.0	100.4	100.6	100.5
	DIBL (V)	0.107	0.108	0.108	0.107	0.107	0.107	0.107
Std. deviation	I_{offsat} (nA)	1.192	0.080	0.079	1.034	0.826	0.971	1.077
	I_{dSat} (µA)	0.202	0.178	0.177	0.202	0.199	0.200	0.202
	V_{ths} (V)	0.123	0.120	0.120	0.122	0.121	0.122	0.123
	SS (mV/dec)	39.11	30.80	30.85	37.11	37.72	38.29	38.89
	DIBL (V)	0.054	0.052	0.052	0.054	0.054	0.054	0.054

FIGURE 9. Statistical distribution of electrical characteristics predicted using machine learning vs. TCAD at the train size 70 %.
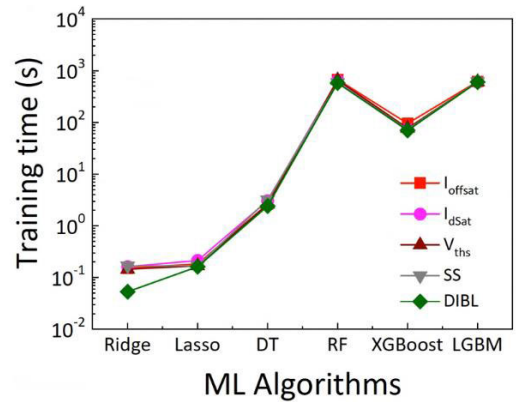
Trainset Size = 50%	Training time (s)					
	Ridge	Lasso	DT	RF	XGBoost	LGBM
I_{offsat}	0.152	0.185	2.64	675	95.5	618
I_{dSat}	0.163	0.215	3.07	646	75.2	604
V_{ths}	0.146	0.170	2.64	660	77.2	609
SS	0.168	0.171	3.18	587	72.7	614
DIBL	0.054	0.163	2.42	580	70.3	605

FIGURE 10. Training time for machine learning model at the trainset size of 50 %.

trees and passes data through each tree based on points to select the decision tree with the most significant weight [26]. In contrast to the DT, some trees in the RF are over fitted, but they prevent overfitting from affecting the results by generating many trees. Therefore, the RF offers a higher prediction accuracy than the DT. XGBoost and LGBM enhance machine learning using a gradient boost algorithm based on the ensemble learning method of multiple trees [27], [28]. Unlike the general gradient boost machine (GBM) models, XGBoost and LGBM are designed for parallel computation, enabling rapid and accurate machine learning. **Figs. 4** and **5** show that the models based on the GBM algorithm offer a lower prediction loss ratio than the other models. Unlike XGBoost, LGBM has higher prediction accuracy than XGBoost because it partitions nodes in a direction that reduces prediction loss [28].

Fig. 7 shows the importance of input parameters (T_{wire}, T_{ox}, L_g, D_{wire}, L_{sp}, and WF), which is the contribution of input parameters to fluctuating electrical characteristics. LGBM, with the highest prediction accuracy of the electrical characteristics, is used to extract input importance. **Figs. 7 (a)** and **(b)** are the importance (%) of the linear and saturation

regimes, respectively. T_{wire} and L_g affected the change in the electrical characteristics in both regimes more significantly than the other geometrical parameters. Owing to the nature of the nanosheet structure, the gate around the interface of the Si–SiO$_2$ can be controlled well. However, gate controllability decreases in the center of the sheet due to the gate electric field weakened by receding from Si–SiO$_2$ interface. The change in the gate electric field at the center of the sheet owing to in T_{wire} variation was more significant than that in D_{wire}. The importance of T_{wire} in changing the electrical characteristics was higher than D_{wire}. L_g affected the electrical characteristics to the same extent as T_{wire}. As L_g decreased, the I_{off}, SS, DIBL, and V_{th} fluctuated significantly because the SCE increased in the subthreshold region due

to V_{th} roll off. In the saturation region, the resistance in the channel changed depending on the channel length, which significantly affected I_{on}. T_{ox} affected the subthreshold region because it is directly related to gate controllability. As T_{ox} changes, causing C_{ox} variation, directly affecting the off-state current, DIBL, and V_{th}. The WF primarily affected the off- and on-currents because it contributed to the carrier accumulation and inversion in the energy band of the silicon channel. As L_{sp} varied, the S/D length varied while L_g remained the same; therefore, L_{sp} contributes primarily to V_{th} roll-off and DIBL changes due to electric field changes in the drain bias. However, because L_{sp} exerts an indirect effect compared with the other geometric parameters, it imposes a less prominent effect on the variation in the electrical characteristics.

We compared the mean and standard deviation of the electrical characteristics predicted using the machine learning model with the mean and standard deviation of the TCAD results at a trainset size of 70 %, which resulted in the lowest prediction loss ratio. In **Figs. 8** and **9**, the mean and standard deviation of the electrical characteristics are shown in a bar graph and tabulated, respectively. In **Fig. 9**, when the difference in mean value and standard deviation between machine learning and TCAD decreases, the background color of the cell equalizes the background color of the TCAD section. The mean value and standard deviation of the electrical characteristics predicted using machine learning were significantly less accurate for the Ridge and LASSO models than the other models. The extremely low accuracy of Ridge and LASSO in predicting the electrical characteristics resulted in a significant error in the mean and standard deviation of the TCAD results. The tree-based models predicted the mean values and standard deviations more accurately than the regulation-based models, primarily the ensemble tree-based model using a gradient boost machine. The LGBM with the lowest prediction loss ratio showed the lowest deviation from the mean and standard deviation of the TCAD result because it predicts the electrical characteristics more accurately than XGBoost.

Fig. 10 shows the time for the six machine learning models to learn based on the trainset size of 50 %. The average training time of each machine learning (Ridge, LASSO, DT, RF, XGBoost, and LGBM) trained through training data is as follows: Ridge, LASSO, DT, RF, XGBoost, and LGBM are 0.137, 0.181, 2.79, 629.6, 78.2, and 610 s. By fitting the second- and first-order equations, Ridge and LASSO required significantly less time than the other tree-based models. DT requires less time than the ensemble-based models because it prunes downward as a single tree, unlike RF, XGBoost, and LGBM, which are ensembles. The training time of the ensemble model is longer than those of the previous three models owing to the generation of various trees and training due to the selection of the optimal tree. XGBoost requires less training time than RF and LGBM due to the pruning nature of XGBoost [27].

IV. CONCLUSION

This study used machine learning to predict electrical characteristics and statistical distribution by varying the device structure parameters. MAPE and R2 scores were used to evaluate how well the machine learning models predicted electrical characteristics variation due to geometric variability. In addition to the general machine learning prediction evaluation, the statistical distribution of electrical characteristics predicted by machine learning was qualitatively compared with the statistical distribution of TCAD simulation results. Consequentially, in the regulation-based models (Ridge and LASSO) and tree-based models (DT, RF, XGBoost, and LGBM), LGBM showed the lowest prediction loss ratio. The electrical characteristics and statistical distribution were predicted with high accuracy. Additionally, the high accuracy of the LGBM model confirmed the extent to which the geometrical variability affects variation in electrical characteristics. The training time of the machine learning model is increased as more accurate prediction of mean and standard deviation of electrical characteristics by the machine learning. This study shows that the electrical characteristics and statistical distribution owing to geometrical variability in NSFET with complex structures can be predicted, and input parameter importance can be analyzed using machine learning.

REFERENCES

[1] N. Loubet, "Stacked nanosheet gate-all-around transistor to enable scaling beyond FinFET," in *Proc. Symp. VLSI Technol.*, Jun. 2017, pp. T230–T231, doi: 10.23919/VLSIT.2017.7998183.

[2] S.-D. Kim, M. Guillorn, I. Lauer, P. Oldiges, T. Hook, and M.-H. Na, "Performance trade-offs in FinFET and gate-all-around device architectures for 7 nm-node and beyond," in *Proc. S3S*, Oct. 2015, pp. 1–3, doi: 10.1109/S3S.2015.7333521.

[3] D. Jang, D. Yakimets, G. Eneman, P. Schuddinck, M. G. Bardon, P. Raghavan, A. Spessot, D. Verkest, and A. Mocuta, "Device exploration of NanoSheet transistors for sub-7-nm technology node," *IEEE Trans. Electron Devices*, vol. 64, no. 6, pp. 2707–2713, Jun. 2017, doi: 10.1109/TED.2017.2695455.

[4] X. Wang, B. Cheng, D. Reid, A. Pender, P. Asenov, C. Millar, and A. Asenov, "FinFET centric variability-aware compact model extraction and generation technology supporting DTCO," *IEEE Trans. Electron Devices*, vol. 62, no. 10, pp. 3139–3146, Oct. 2015, doi: 10.1109/TED.2015.2463073.

[5] T. Karatsori, C. Theodorou, R. Lavieville, T. Chiarella, J. Mitard, N. Horiguchi, C. A. Dimitriadis, and G. Ghibaudo, "Statistical characterization and modeling of drain current local and global variability in 14 nm bulk FinFETs," in *Proc. Int. Conf. Microelectronic Test Struct. (ICMTS)*, Grenoble, France, Mar. 2017, pp. 1–5, doi: 10.1109/ICMTS.2017.7954263.

[6] N. Seoane, J. G. Fernandez, K. Kalna, E. Comesaña, and A. García-Loureiro, "Simulations of statistical variability in n-type FinFET, nanowire, and nanosheet FETs," *IEEE Electron Device Lett.*, vol. 42, no. 10, pp. 1416–1419, Oct. 2021, doi: 10.1109/LED.2021.3109586.

[7] X. Jiang, S. Guo, R. Wang, X. Wang, B. Cheng, A. Asenov, and R. Huang, "A device-level characterization approach to quantify the impacts of different random variation sources in FinFET technology," *IEEE Electron Device Lett.*, vol. 37, no. 8, pp. 962–965, Aug. 2016, doi: 10.1109/LED.2016.2581878.

[8] S. R. Kola, Y. Li, C.-Y. Chen, and M.-H. Chuang, "Statistical 3D device simulation of full fluctuations of gate-all-around silicon nanosheet MOSFETs at sub-3-nm technology nodes," in *Proc. Int. Symp. VLSI Technol., Syst. Appl. (VLSI-TSA)*, Hsinchu, Taiwan, Apr. 2022, pp. 1–2, doi: 10.1109/VLSI-TSA54299.2022.9771002.

[9] P. H. Vardhan, Amita, S. Ganguly, and U. Ganguly, "Threshold voltage variability in nanosheet GAA transistors," *IEEE Trans. Electron Devices*, vol. 66, no. 10, pp. 4433–4438, Oct. 2019, doi: 10.1109/TED.2019.2933061.

[10] A. Gorad and U. Ganguly, "Analytical estimation of LER-like variability in GAA nano-sheet transistors," in *Proc. Int. Symp. VLSI Technol., Syst. Appl. (VLSI-TSA)*, Hsinchu, Taiwan, Apr. 2019, pp. 1–2, doi: 10.1109/VLSI-TSA.2019.8804637.

[11] V. Moroz, J. Huang, and R. Arghavani, "Transistor design for 5nm and beyond: Slowing down electrons to speed up transistors," in *Proc. 17th Int. Symp. Quality Electron. Design (ISQED)*, Santa Clara, CA, USA, Mar. 2016, pp. 278–283, doi: 10.1109/ISQED.2016.7479214.

[12] L. Smith, M. Choi, M. Frey, V. Moroz, A. Ziegler, and M. Luisier, "FinFET to nanowire transition at 5nm design rules," in *Proc. Int. Conf. Simul. Semiconductor Processes Devices (SISPAD)*, Sep. 2015, pp. 254–257.

[13] M. Choi, V. Moroz, L. Smith, and J. Huang, "Extending drift-diffusion paradigm into the era of FinFETs and nanowires," in *Proc. Int. Conf. Simul. Semiconductor Processes Devices (SISPAD)*, Sep. 2015, pp. 242–245.

[14] J. Kim, S. J. Kim, J.-W. Han, and M. Meyyappan, "Machine learning approach for prediction of point defect effect in FinFET," *IEEE Trans. Device Mater. Rel.*, vol. 21, no. 2, pp. 252–257, Jun. 2021, doi: 10.1109/TDMR.2021.3069720.

[15] P. Bharti, H. Muthusamy, and V. Kumar, "Thermal resistance extraction of 14 nm SOI FinFET: A machine learning based approach," in *Proc. 2nd Int. Conf. Emerg. Frontiers Electr. Electron. Technol. (ICEFEET)*, Patna, India, Jun. 2022, pp. 1–5, doi: 10.1109/ICEFEET51821.2022.9847687.

[16] J. Ghosh, S. Y. Lim, and A. V. Thean, "Bridge-defect prediction in SRAM circuits using random forest, XGBoost, and LightGBM learners," in *Proc. Int. Conf. Simulation Semiconductor Processes Devices (SISPAD)*, Dallas, TX, USA, Sep. 2021, pp. 259–262, doi: 10.1109/SISPAD54002.2021.9592539.

[17] R. Butola, Y. Li, and S. R. Kola, "Machine learning approach to characteristic fluctuation of bulk FinFETs induced by random interface traps," in *Proc. 23rd Int. Symp. Quality Electron. Design (ISQED)*, Santa Clara, CA, USA, Apr. 2022, pp. 1–6, doi: 10.1109/ISQED54688.2022.9806233.

[18] M. Hashimoto, W. Liao, and S. Hirokawa, "Soft error rate estimation with TCAD and machine learning," in *Proc. Int. Conf. Simul. Semiconductor Processes Devices (SISPAD)*, Kamakura, Japan, Sep. 2017, pp. 129–132, doi: 10.23919/SISPAD.2017.8085281.

[19] *Sentaurus Device User Guide, Version R-2020.09*, Synopsys, San Jose, CA, USA, 2020.

[20] J.-W. Han, H. Y. Wong, D.-I. Moon, N. Braga, and M. Meyyappan, "Stringer gate FinFET on bulk substrate," *IEEE Trans. Electron Devices*, vol. 63, no. 9, pp. 3432–3438, Sep. 2016, doi: 10.1109/TED.2016.2586607.

[21] J. Kim, J.-W. Han, and M. Meyyappan, "Reduction of variability in junctionless and inversion-mode FinFETs by stringer gate structure," *IEEE Trans. Electron Devices*, vol. 65, no. 2, pp. 470–475, Feb. 2018, doi: 10.1109/TED.2017.2786238.

[22] G. A. M. Hurkx, D. B. M. Klaassen, and M. P. G. Knuvers, "A new recombination model for device simulation including tunneling," *IEEE Trans. Electron Devices*, vol. 39, no. 2, pp. 331–338, Feb. 1992, doi: 10.1109/16.121690.

[23] G. Paasch and H. Übensee, "A modified local density approximation. Electron density in inversion layers," *Phys. Status Solidi (B)*, vol. 113, no. 1, pp. 165–178, Sep. 1982, doi: 10.1002/pssb.2221130116.

[24] L. E. Melkumova and S. Ya. Shatskikh, "Comparing Ridge and LASSO estimators for data analysis," *Proc. Eng.*, vol. 201, pp. 746–755, Jan. 2017, doi: 10.1016/j.proeng.2017.09.615.

[25] A. J. Myles, R. N. Feudale, Y. Liu, N. A. Woody, and S. D. Brown, "An introduction to decision tree modeling," *J. Chemometrics*, vol. 18, pp. 275–285, Jun. 2004, doi: 10.1002/cem.873.

[26] G. Biau and E. Scornet, "A random forest guided tour," *TEST*, vol. 25, pp. 197–227, Jun. 2016, doi: 10.1007/s11749-016-0481-7.

[27] T. Chen and C. Guestrin, "XGBoost: A scalable tree boosting system," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, New York, NY, USA, Aug. 2016, pp. 785–794, doi: 10.1145/2939672.2939785.

[28] G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A highly efficient gradient boosting decision tree," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 3146–3154.
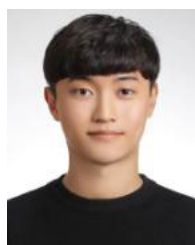
JONGHYEON HA (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from Gyeongsang National University, in 2023, where he is currently pursuing the master's degree in electrical engineering. His main research interest includes the modeling of logic FET using TCAD simulation.
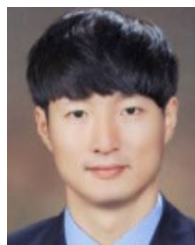
SUN JIN KIM received the Ph.D. degree from the Department of Electrical Engineering, Korea Advanced Institute of Science and Technology (KAIST), Daejeon, South Korea, in 2017. He was a Research Scientist with the Center for Nanotechnology, NASA Ames Research Center, Moffett Field, CA, USA. He is currently a Research Scientist with the Smart Structural Safety and Prognosis Research Division, Korea Atomic Energy Research Institute, Daejeon. His research interests include device design and process development for flexible thermoelectric device and radiation sensor.

MINJI BANG received the B.S. degree in electrical engineering from Gyeongsang National University, in 2022, where she is currently pursuing the master's degree in electrical engineering. Her main research interest includes the modeling of logic FET using TCAD simulation.

GYEONGYEOP LEE received the B.S. degree in electrical engineering from Gyeongsang National University, in 2021, where he is currently pursuing the master's degree in electrical engineering. His main research interest includes the modeling and reliability of silicon channel FET.

MINKI SUH (Graduate Student Member, IEEE) received the B.S. degree in electrical engineering from Gyeongsang National University, in 2022, where he is currently pursuing the master's degree in electrical engineering. His main research interest includes the modeling and reliability of silicon channel FET.

CHONG-EUN KIM (Member, IEEE) received the B.S. degree in electrical engineering from Kyungpook National University, Daegu, South Korea, in 2001, and the M.S. and Ph.D. degrees in power electronics from the Korea Advanced Institute of Science and Technology, Daejeon, South Korea, in 2003 and 2008, respectively.

MINSEOB SHIM (Member, IEEE) received the B.S. and Ph.D. degrees in electrical engineering from Korea University, Seoul, South Korea, in 2012 and 2018, respectively. From 2015 to 2016, he was a Visiting Researcher with the University of Michigan, Ann Arbor, MI, USA. In 2018, he joined the Korea Electrotechnology Research Institute (KERI), Changwon, South Korea, as a Senior Researcher. He is currently an Assistant Professor with the Department of Electronic Engineering, Gyeongsang National University, Jinju, South Korea. His research interests include integrated power management systems, gate driver integrated circuit for wide bandgap devices, low voltage analog and mixed-signal integrated circuits, and analog-to-digital converter designs.

JUNGSIK KIM (Senior Member, IEEE) received the Ph.D. degree in IT convergence engineering from the Pohang University of Science and Technology, Pohang, South Korea, in 2016. From February 2016 to March 2018, he was with SK-Hynix for the modeling of 96-stacks VNAND and Samsung Electronics for compact modeling of 1a-node DRAM, from April 2019 to February 2020. He was a Visiting Scholar with the NASA Ames Research Center, for reliability due to the radiation effect in silicon devices, from April 2018 to March 2019. He is currently an Assistant Professor with the Department of Electrical Engineering, Gyeongsang National University, Jinju, South Korea. His research interests include the modeling and reliability of nano-scale devices based on technical computer-aided simulation and measurement.

○ ○ ○