2022 IEEE 22nd International Conference on Nanotechnology (NANO)
July 4-8, 2022
Palma de Mallorca, Spain

Estimating the Process Variation Effects of Stacked Gate All Around Si Nanosheet CFETs Using Artificial Neural Network Modeling Framework

Rajat Butola, Yiming Li*, Member, IEEE, Sekhar Reddy Kola, Min-Hui Chuang, and Chandni Akbar

***Abstract*— We for the first time report a novel machine learning (ML) approach to model the effects of varying process parameters on DC characteristics of stacked gate all around (GAA) Si nanosheet (NS) complementary-FETs (CFETs) using an artificial neural network (ANN) model. Process parameters that have predominant effects on device characteristics are considered and used as input features to the ANN model; and, their effects on DC characteristics are modeled. Major figures of merit (FoMs) are further extracted accurately from the transfer characteristics in much less computational time as compared to 3D device simulation. The performance of the ANN model is further evaluated using the coefficient of determination, R^2-score, which is more than 96%. It shows that the ANN model successfully learned the information from the dataset; thus, the ANN model exhibits the competency in device modeling of emerging CFETs.**

I. Introduction

For nodes 5-nm and beyond, several options have been proposed. The complementary field-effect transistor (CFET) is one of them. CFET aims at continuing Moore's law by folding the gate-all-around (GAA) nanosheet (NS) FET. It consists of a vertically stacked n-FET on top of a p-FET [1]. It is a more complex version of a GAA device and a promising alternative to continue the device scaling [2]. CFET can also provide excellent electrostatic integrity by using the GAA NS structure. Due to vertically stacked n-FETs and p-FETs, the cell area is decreased significantly [3]. However, for highly scaled devices even the small variations affect the characteristics of these devices severely, therefore, they cannot be ignored. These variations are the main obstacle for further scaling down of technology node [4].

This work was supported in part by the Ministry of Science and Technology (MOST), Taiwan, under Grant MOST 110-2221-E-A49-139, MOST 110-2218-E-492 -003-MBK, Grant MOST 109-2221-E-009-033, and Grant MOST 109-2634-F-009-030, and in part by the "Center for mm Wave Smart Radar Systems and Technologies" under the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan.

R. Butola, Y. Li, and S. R. Kola are with the Parallel and Scientific Computing Laboratory, Electrical Engineering and Computer Science International Graduate Program, Department of Electrical Engineering and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan.

Y. Li is also with the Department of Electrical Engineering and Computer Engineering, Institute of Communications Engineering, Institute of Biomedical Engineering, and Center for mmWave Smart Radar System and Technologies, National Yang Ming Chiao Tung University, Hsinchu 30010, Taiwan, (e-mail: ymli@nycu.edu.tw).

M.-H. Chuang and C. Akbar are with the Parallel and Scientific Computing Laboratory, and Institute of Communications Engineering, National Yang Ming Chiao Tung University, Hsinchu 300, Taiwan.

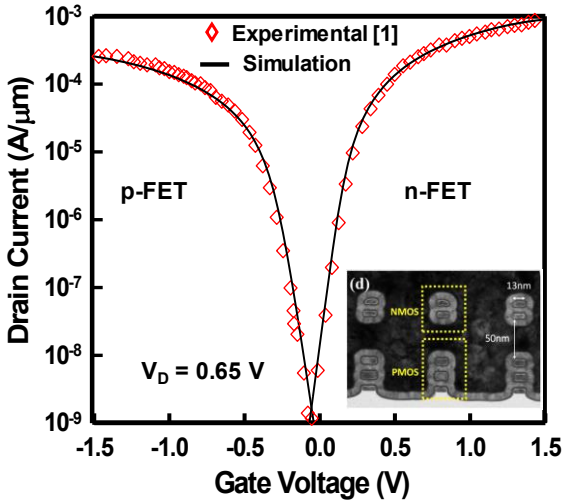

Figure 1. Simulated (line) and experimental (symbol) I_D-V_G characteristics for the accuracy calibration, for both N-FET and P-FET, where the inset is a transmission electron microscope (TEM) image of fabricated CFET.

They affect the characteristics and the parameters of the nano-devices such as threshold voltage (V_{TH}), on-current (I_{ON}), off-current (I_{OFF}), subthreshold slope (SS), etc. The important variations consist of process variation effect (PVE), work function variation (WKV), interface trap variations (ITV), random dopant variations (RDV), etc. In the recent past, a lot of work is done where the influence of these variations is studied for nano-devices [5-8].

The process variations (PV) for sub 5-nm are the dominant factors that affect device-to-device performance variations. In the case of NS CFET, the prominent factors which vary the device characteristics are gate length (L_G), NS width (W_{NS}), NS height (H_{NS}), oxide thickness (T_{OX}), and spacing between n-FET and p-FET (D_{SEP}). In this work, we investigate the effect of all these parameters on the characteristics of stacked GAA Si NS CFET. We use stacked NS CFET with 2-channels therefore, we include one more important parameter, the spacing between these channels (L_{SEP}). Recently, many ML-based frameworks are proposed in semiconductor device modeling to reduce the cost and the computational time which is the bottleneck of the conventional device simulation method [9-15]. Therefore, we incorporate an ML-based ANN model, to analyze the impacts of process variation effects (PVE) on stacked NS CFET. The important FoMs: threshold voltage (V_{TH}), off-current (I_{OFF}), and on-current (I_{ON}) are also extracted from the predicted I_D-V_G characteristics as output parameters.

2022 IEEE 22nd International Conference on Nanotechnology (NANO) | 978-1-6654-5225-0/22/$31.00 ©2022 IEEE | DOI: 10.1109/NANO54668.2022.9928645

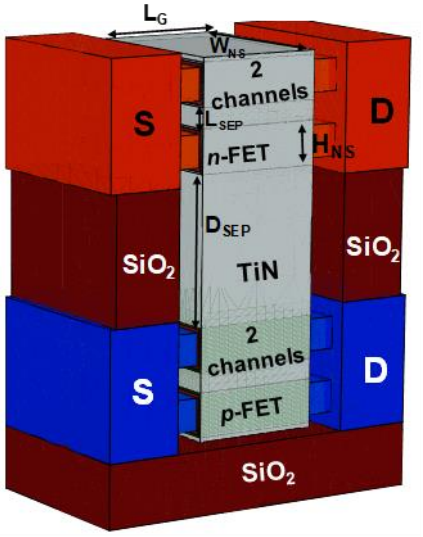

Figure 2. 3-D schematic structure of stacked GAA Si NS CFET device consisting of n-FET (on top) and p-FET (in bottom) with 2-channels each. The device parameters L_G, W_{NS}, T_{OX}, H_{NS}, L_{SEP} and D_{SEP} are specified.

TABLE I
LIST OF STACKED NS SI CFET DEVICE PARAMETERS CORRESPONDING TO SUB-5-NM TECHNOLOGY NODE

Parameters	Values
Gate Length (L_G) (nm)	20
Channel Doping (cm^{-3})	$5x10^{17}$
Nanosheet Width (W_{NS})(nm)	25
Nanosheet Height (H_{NS}) (nm)	5
Spacing between nanosheet (L_{SEP}) (nm)	10
Spacing between N- and P-FET (D_{SEP}) (nm)	20
S_{ext}/D_{ext} *Length* (nm)	5
S/D Doping (cm^{-3})	$1x10^{20}$
S_{ext}/D_{ext} *Doping* (cm^{-3})	$4.8x10^{18}$
Number of Channels	2
EOT (nm)	0.6

II. DEVICE STRUCTURE AND THE ANN APPROACH

We first calibrated the device simulation with intel measurement data for CFETs [1] to provide the best accuracy, as shown in Fig. 1. In Fig. 2, the 3-D structure of the stacked GAA Si NS CFET is represented. A similar calibration methodology, models, and doping levels are levels are applied to perform CFET PVE device simulations. For the physical model, a validated 3-D density-gradient along with a drift-diffusion model with optimal electron effective mass and quantum correction technique is adopted [16]. For the PVE simulations, six important parameters, L_G, T_{ox}, H_{NS}. W_{NS}, L_{SEP}, and D_{SEP} are varied randomly. A total of 300 stacked GAA Si NS CFET fluctuated devices are simulated by varying these parameters. The parameters of the nominal device with process parameters and doping concentrations are summarized in Table I.

The problem under this work is associated with the supervised regression problem and after exploring many algorithms the Artificial Neural Network (ANN) is selected for the estimation of PVE effects on stacked GAA Si NS CFET device characteristics. The ANN architecture, illustrated in Fig. 3, has one input layer (6 neurons each), two

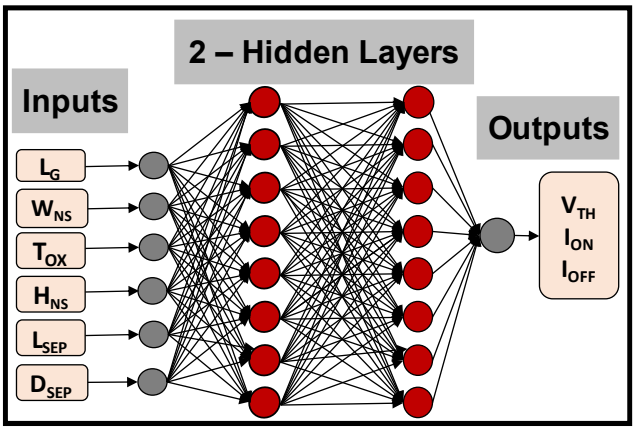

Figure 3. It illustrates the explored ANN structure. It consists of four layers, i.e., one input layer, two hidden layers and one output layer. L_G, W_{NS}, T_{OX}, H_{NS}, L_{SEP} and D_{SEP} are the input parameters whereas V_{TH}, I_{ON} and I_{OFF} are the extracted output parameters from the predicted I_D-V_G characteristics.

TABLE II
LIST OF THE TUNED HYPERPARAMETERS UTILIZED IN THE ANN MODEL DURING TRAINING AND TESTING PROCEDURE.

Hyperparameters	Values
Hidden Layers	2
Optimizer	RMSprop
Epochs	100
Activation Function	Tanh
Neurons	45
Learning Rate	0.001
(I/P, O/P) Dimension	(6,36)
Random State	42
Batch Size	5

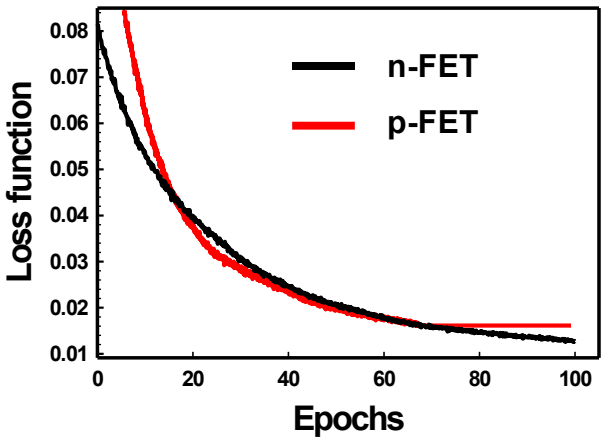

Figure 4. It represents the loss function versus number of epochs during training of the ANN model for n-FET (blue line) and p-FET (red line).

hidden layers (45 neurons each), and one output layer (36 neurons each). Before the training of the ANN model, the hyperparameters are set. The list of various hyperparameters such as activation function, type of optimizer, number of epochs, number of neurons in each hidden layer, learning rate, and batch size are summarized in Table II. After choosing the appropriate hyperparameters, the model is trained. During the training, the weights and the bias are updated to be fit to the training dataset for the specified number of epochs. The biasing conditions for simulating devices are as follows: the drain voltage V_D = 0.7 V and the

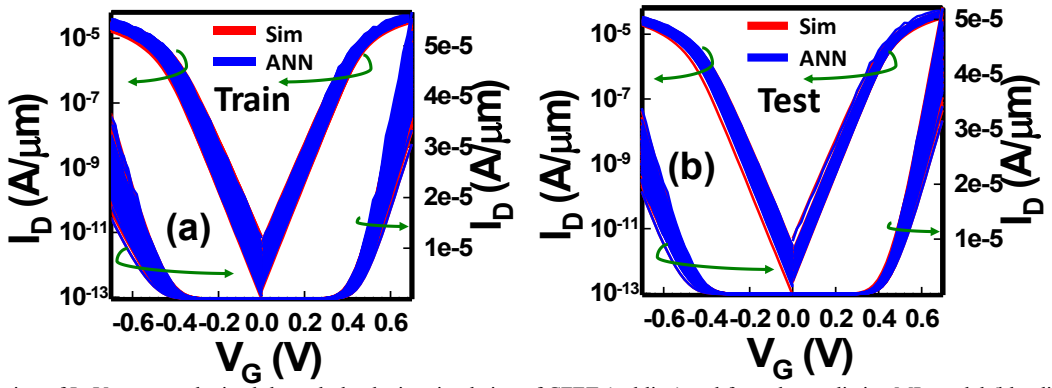

Figure 5. Illustration of I_D-V_G curves obtained through the device simulation of CFET (red line) and from the predictive ML model (blue line). (a) represents the training of the ANN model based on the CFET device using n-and p-FET. It also shows the linear as well as logarithmic I_D-V_G curves. Similarly, (b) shows the testing of the ANN model. These I_D-V_G curves for both training and testing for n-and p-FET depict the well-trained ANN model.

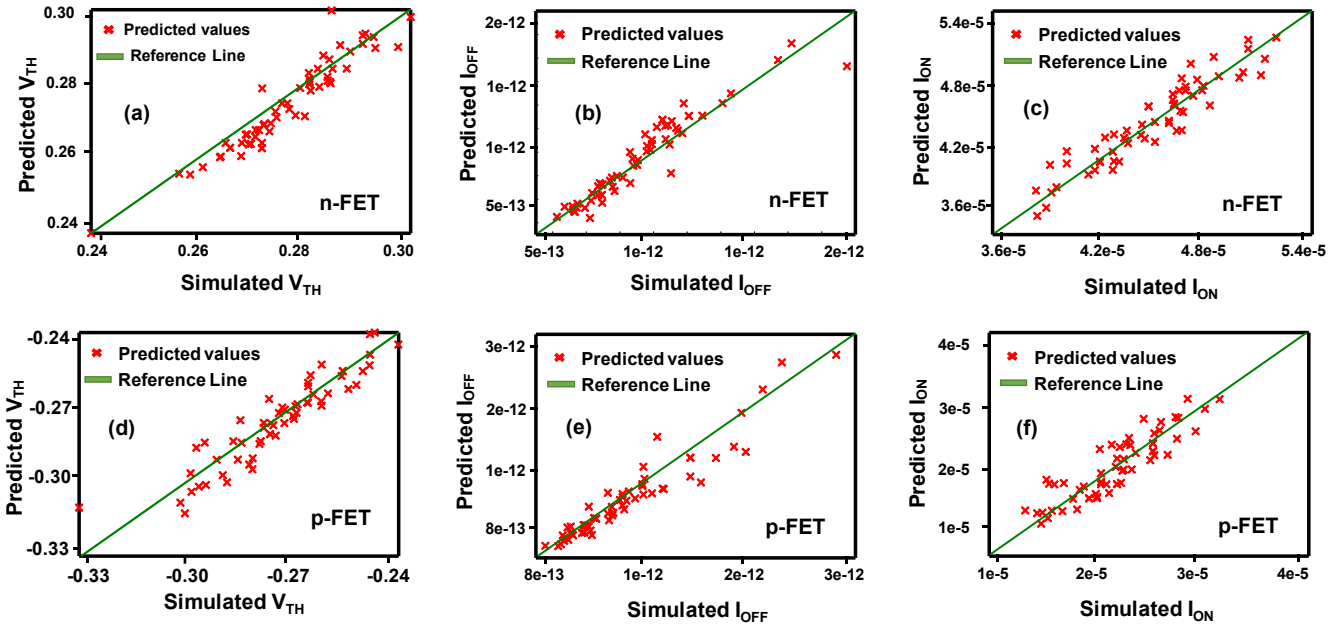

Figure 6. It shows the linear relationship between simulated and predicted values for n-FET and p-FET. (a) and (d) show the predicted and simulated V_{TH} values for n-FET and p-FET, respectively. Similarly, (b) and (e) represent the predicted and simulated I_{OFF} values for n-FET and p-FET, respectively, and (c) and (f) depict the predicted and simulated I_{ON} values for n-FET and p-FET, respectively. All extracted electrical characteristics from simulated I_D-V_G curves as well as from the predicted I_D-V_G curves are in good agreement with each other.

gate voltage V_G is swept from 0 to 0.7 V with a step size (Δ) of 0.02 V. For the ANN model, the input features consist of [L_G, T_{ox}, H_{NS}. W_{NS}, L_{SEP}, D_{SEP}] and the output structure consists of 36 [I_{D1}, I_{D2}, ….., I_{D36}] drain current values corresponding to each gate voltage V_G [0, 0.02, ….., 0.7].

III. RESULTS AND DISCUSSION

The raw simulated data is collected and preprocessed for the ANN model. The MinMax scalar is used to transform the data before feeding it into the ANN model. Afterward, the processed data is randomly split into the training (80%) and test dataset (20%). The training data is used by the ANN model to learn the hidden information present in the data, whereas test data is used to validate the model performance. During training, the loss values are calculated in each epoch to optimize the weights and biases of the hidden layers until the error rate goes below a certain minimum value and is stabilized. Fig. 4 shows the loss function versus the number of epoch curves for n-FET and p-FET. We can examine from the figure that the errors for n-FET and p-FET are reduced to minimum values and become steady after approximately 90 epochs. First, the training I_D-V_G curves are plotted to check whether the model is trained perfectly or not. In Fig. 5(a), the ANN model fitting results for the training dataset in linear and logarithmic scales are presented and compared with simulated values. It can be examined from the plots that the I_D-V_G curves are in good agreement with each other and high model accuracy has been achieved for the drain current with respect to the gate voltage. However, the true performance of the fully trained model is judged when it is exposed to the unseen test dataset. From Fig. 5(b), we can evaluate that the curves are predicted accurately, and predicted results follow the simulated values closely. This confirms that the model reproduced the IV characteristics efficiently using the ANN model.

TABLE II
LIST OF THE RMSE VALUES AND R^2-SCORE OBTAINED THROUGH TRAINING AND TESTING OF ANN MODEL FOR N-FET AND P-FET.

		RMSE	R^2-Score
n-FET	*Train*	3.82e-07	0.9723
	Test	4.22e-07	0.9219
p-FET	*Train*	5.79e-06	0.9673
	Test	5.91e-06	0.9056

For quantitative assessment, root mean square error (RMSE) and R^2-score are used as descriptors. Table III lists the RMSE values and R^2-score of training and test results for both n-FET and p-FET. The training R^2-score for n-FET and p-FET is more than 96% and for testing, the R^2-score is more than 90%. Furthermore, the smaller errors in the table exhibit the better performance of the ANN model. The test plots show that the trained model predicted the I_D-V_G characteristics accurately with a small error rate. The variability in key FoMs, V_{TH}, I_{OFF}, and I_{ON}, for test curves, are also estimated using the ANN model. The comparison between the predicted and the simulated I_D-V_G curves for n-FET and p-FET are presented in Figs. 6(a)-(f). It can be examined from the results that the ANN outputs are confining towards the reference line (ideal line) for all the FoMs. The results proved that the ANN model has accurate predictions for different FoMs.

Moreover, the ANN model is a more time-efficient method as compared to the simulation method. Simulating one CFET device, in 3D device simulation software, takes 68 minutes and more than 2 weeks to simulate 300 such fluctuated devices. On the other hand, with the same computing resources, the ANN model is trained in only 210 seconds. This shows that the ANN-based modeling is much faster than conventional device simulation, and also reduces the cost as well as the huge computational power. Therefore, the ANN can be the best alternative to accelerate the nano-device modeling process.

IV. CONCLUSION

In summary, to the best of our knowledge, the ML-based ANN model is utilized for the first time to estimate the effects of process variation on electrical characteristics of stacked GAA Si NS CFET. The dataset for the ANN model is generated by 3D device simulation by varying various process parameters. The ANN architecture is constructed to model the device characteristics and to analyze the impact of the PV effects. To assure the effectiveness of the proposed ANN model, we calculate the R^2-score and RMSE values. With R^2-score of more than 96% and 90% for training and testing, respectively, show the good fitting of the ANN model. To further investigate the impact of the process variation effects on the electrical characteristics of the CFET devices, different key FoMs (e.g. V_{TH}, I_{OFF}, and I_{ON}) are extracted and compared with the ANN predicted values. The results suggest that the predicted values of FoMs are closely following the simulated values. In this study, we established that the ML-based ANN modeling is a fast and economic alternative to 3D device simulation to explore the relationship between the process variation and the transfer characteristics without compromising the accuracy. In the future, our focus is on the use of ML models in CFET based circuit simulation to minimize the huge computational power and time.

REFERENCES

[1] C.-Y. Huang et al., "3-D Self-Aligned Stacked NMOS-on-PMOS Nanoribbon Transistors for Continued Moore's Law Scaling," in IEEE International Electron Devices Meeting, pp. 20.6.1-20.6.4, Dec. 2020.

[2] J. Ryckaert et al., "The Complementary FET (CFET) for CMOS scaling beyond N3," in IEEE Symposium on VLSI Tech., pp. 141-142, June 2018.

[3] S.-W Chang et al., "First Demonstration of CMOS Inverter and 6T-SRAM Based on GAA CFETs Structure for 3D-IC Applications," in IEEE International Electron Devices Meeting, pp. 11.7.1-11.7.4, Dec. 2019.

[4] S. R. Kola, Y. Li and N. Thoti, "Effects of Spacer and Single-Charge Trap on Voltage Transfer Characteristics of Gate-All-Around Silicon Nanowire CMOS Devices and Circuits," in IEEE International Conference on Nanotechnology, pp. 217-220, Sept. 2020.

[5] Y. Li, H. T. Chang, C. N. Lai, P. J. Chao, and C. Y. Chen, "Process variation effect, metal-gate work-function fluctuation and random dopant fluctuation of 10-nm gate-all-around silicon nanowire MOSFET devices," in IEEE International Electron Devices Meeting, pp. 34.4.1-34.4.4, Dec. 2015.

[6] X. Li et al., "Impact of Process Fluctuations on Reconfigurable Silicon Nanowire Transistor," IEEE Transactions on Electron Devices, vol. 68, no. 2, pp. 885-891, Feb. 2021.

[7] J. Yoon, S. Lee, J. Lee, J. Jeong, H. Yun, and R. Baek, "Reduction of Process Variations for Sub-5-nm Node Fin and Nanosheet FETs Using Novel Process Scheme," IEEE Transactions on Electron Devices, vol. 67, no. 7, pp. 2732-2737, July 2020.

[8] R. Butola, Y. Li, and S. R. Kola, "Deep Learning Approach to Modeling and Exploring Random Sources of Gate-All-Around Silicon Nanosheet MOSFETs," International Symposium on VLSI Technology, Systems and Applications, pp. 1-2, May 2022.

[9] C. Akbar, Y. Li, and W. L. Sung, "Deep Learning Approach to Inverse Grain Pattern of Nanosized Metal Gate for Multichannel Gate-All-Around Silicon Nanosheet MOSFETs," IEEE Transactions on Semiconductor Manufacturing, vol. 34, no. 4, pp. 513-520, Nov. 2021.

[10] C.-K. Cheng et al., "Design and System Technology Co-Optimization Sensitivity Prediction for VLSI Technology Development using Machine Learning," in ACM/IEEE International Workshop on System Level Interconnect Prediction, pp. 8-15, Jan. 2022.

[11] K. Ko, J. K. Lee, M. Kang, J. Jeon, and H. Shin, "Prediction of Process Variation Effect for Ultrascaled GAA Vertical FET Devices Using a Machine Learning Approach," IEEE Transactions on Electron Devices, vol. 66, no. 10, pp. 4474-4477, Oct. 2019.

[12] G. M. Seo, C. -K. Baek and B. D. Kong, "Inverse Design of Graphene FET by Deep Neural Network," in IEEE International Conference on Nanotechnology, pp. 134-137, Aug. 2021.

[13] H. Carrillo-Nuñez, N. Dimitrova, A. Asenov, and V. Georgiev, "Machine Learning Approach for Predicting the Effect of Statistical Variability in Si Junctionless Nanowire Transistors," IEEE Electron Device Letters, vol. 40, no. 9, pp. 1366-1369, Sept. 2019.

[14] C. Akbar, Y. Li, and W. L. Sung, "Machine Learning Aided Device Simulation of Work Function Fluctuation for Multichannel Gate-All-Around Silicon Nanosheet MOSFETs," IEEE Transactions on Electron Devices, vol. 68, no. 11, pp. 5490-5497, Nov. 2021.

[15] M.-H. Oh, M. -W. Kwon, K. Park, and B. -G. Park, "Sensitivity Analysis Based on Neural Network for Optimizing Device Characteristics," IEEE Electron Device Letters, vol. 41, no. 10, pp. 1548-1551, Oct. 2020.

[16] S. R. Kola, Y. Li, and N. Thoti, "Characteristics of Gate-All-Around Silicon Nanowire and Nanosheet MOSFETs with Various Spacers," in Simulation of Semiconductor Processes and Devices, pp. 1-4, Nov. 2020.