

Artificial Neural Network-Based Modeling for Estimating the Effects of Various Random Fluctuations on DC/Analog/RF Characteristics of GAA Si Nanosheet FETs

Rajat Butola^{ID}, Yiming Li^{ID}, Member, IEEE, Sekhar Reddy Kola^{ID}, Graduate Student Member, IEEE,
Chieh-Yang Chen, and Min-Hui Chuang, Student Member, IEEE

Abstract—Advanced field-effect transistors (FETs), such as gate-all-around (GAA) nanowire (NW) and nanosheet (NS) devices, have been highly scaled; therefore, they are critically affected even by a microscopic fluctuation. As the GAA NS device is considered a promising candidate beyond 5-nm technology, it is essential to analyze the effects of these fluctuations on dc and analog/radio frequency (RF) characteristics for future applications. In this article, we for the first time demonstrate that the machine learning (ML)-aided numerical device simulation approach can be used to model the effects of various fluctuations on the characteristics of GAA NS FETs (NSFETs). Among various fluctuations, we mainly focus on work function fluctuation (WKF), random dopant fluctuation (RDF), and interface trap fluctuation (ITF). The independent and combined effects of these fluctuations on the characteristics of NSFETs are studied. Except for transfer and output characteristics, analog and RF parameters, such as gate capacitance, transconductance, cutoff frequency, 3-dB frequency, and transconductance efficiency, are analyzed in detail. The main aim of this work is to show the capability and generality of ML in modeling various electrical characteristics of the explored NSFETs. The results show that the ML-based technique is fast and efficient, which accelerates the overall process and gives engineering acceptable accurate results.

Manuscript received 20 May 2022; accepted 27 July 2022. Date of publication 26 August 2022; date of current version 4 November 2022. This work was supported in part by the National Science and Technology Council, Taiwan, under Grant MOST 111-2221-E-A49-181, Grant MOST 110-2221-E-A49-139, and Grant MOST 110-2218-E-492-003-MBK; in part by the “2022 Qualcomm Taiwan Research Program (NYCU)” under Grant NAT-487835 SOW; and in part by the “Center for mm Wave Smart Radar Systems and Technologies” through the Featured Areas Research Center Program within the framework of the Higher Education Sprout Project by the Ministry of Education in Taiwan. (Corresponding author: Yiming Li.)

Rajat Butola and Sekhar Reddy Kola are with the Parallel and Scientific Computing Laboratory, Electrical Engineering and Computer Science International Graduate Program (EECS IGP), National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan.

Yiming Li is with the Parallel and Scientific Computing Laboratory, Electrical Engineering and Computer Science International Graduate Program (EECS IGP), National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan, and also with the Institute of Communications Engineering, the Institute of Biomedical Engineering, and the Department of Electrical Engineering and Computer Engineering, National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan (e-mail: yml@nycu.edu.tw).

Chieh-Yang Chen and Min-Hui Chuang are with the Parallel and Scientific Computing Laboratory and the Institute of Communications Engineering, National Yang Ming Chiao Tung University, Hsinchu 300093, Taiwan.

Color versions of one or more figures in this article are available at <https://doi.org/10.1109/TMTT.2022.3198659>.

Digital Object Identifier 10.1109/TMTT.2022.3198659

Index Terms—Characteristic fluctuation, dc/analog/radio frequency (RF), gate-all-around (GAA) nanosheet field-effect transistors (NSFETs), interface trap fluctuation (ITF), intrinsic parameter fluctuation, machine learning (ML), random dopant fluctuation (RDF), work function fluctuation (WKF).

I. INTRODUCTION

THE FinFETs were considered a revolutionary change to the way the transistor structures are built when they were commercialized at the 22-nm node [1]. In FinFETs, the channel is surrounded on three sides by the gate, which results in high scalability, better drive currents, lower leakage, faster switching time, better channel electrostatics, and all in all a better transistor of choice for semiconductor logic [2]. With bulk FinFETs, the semiconductor technology scaled from 22 to 5 nm. However, with the continuous scaling, the FinFETs reached their limit beyond 5 nm due to a lack of enough electrostatic control [3]. With every downscaling of the technology node, the goal of delivering the performance boost while maintaining the low-power reduction becomes more challenging [4], [5]. Therefore, for such extremely scaled processing nodes, the gate-all-around (GAA) transistors, nanowire (NW), and nanosheet (NS) have emerged as the successors to the FinFETs beyond 5 nm [6], [7]. The GAA structure provides the control of channel from all the sides and thus enables further device scaling possible [8]. However, the GAA NW transistor has a lower drive current due to narrower effective channel widths [9]. The GAA NS field-effect transistor (NSFET), on the other hand, provides high drive current and improved channel control as compared to FinFETs due to its high surface to volume ratio and maintain good short channel characteristics [10].

The GAA NSFET provides some key advantages, one of them is the design flexibility which allows the designer to adjust the effective width of the channel to provide more drive current and speed up the ON-OFF switching of the transistor. A sheet can be made wider to boost the current, whereas a sheet can be made narrower to limit the power consumption. This way, the GAA NSFET provides variable designing options to make it possible to match the aggressive scaling of technology nodes. These are the important features that make the GAA NSFET a suitable device for logic applications as well as for radio frequency (RF) circuits. Furthermore, the

higher cutoff frequency of the GAA NSFET makes it more appropriate for RF applications [11]. The cutoff frequency (f_t) of a transistor is proportional to the transconductance (g_m) and inversely proportional to gate capacitance (C_G). In RF applications, high gain (G) is needed for data transmission, which requires high g_m in transistors [12], [13]. The GAA NSFET provides a higher g_m as compared to FinFETs; therefore, the GAA NSFET is gaining more attention from researchers for analog/RF applications.

However, highly scaled devices are severely affected by random fluctuations, and unfortunately, the GAA NSFET is not an exception [14]. These fluctuations are the serious obstacles to further scaling down of technology nodes [15], [16], [17], [18]. They affect the major dc and analog/RF parameters of the device, such as g_m , C_G , f_t , and so on [19], [20]. Without considering the process variation effect, major sources of randomness consist of work function fluctuation (WKF), random dopant fluctuation (RDF), interface trap fluctuation (ITF), and so on. In the past, a lot of work was done, where the impact of these fluctuations was studied for the GAA NSFETs [21], [22], [23], [24], [25], [26], [27], [28], [29].

To date, however, all the reported works on the GAA NSFETs are purely based on 3-D numerical semiconductor device simulation. Simulation-based results are physically accurate; however, they are very time taking [30] and computationally inefficient [31]. The need for a technique, which can speed up this process of simulation, brings out the machine learning (ML)-based approach to the notice of device researchers [32], [33]. Some past works suggest that ML has significant value creation potential in semiconductor manufacturing, and therefore, the advancement of ML has accelerated rapidly in semiconductor manufacturing [33], [34], [35]. In the recent past, some ML work has been reported to model the variabilities in nanodevices. The ML is used successfully to predict the characteristic fluctuations that occurred in various semiconductor devices caused by different fluctuation sources, such as WKF [36], random discrete dopant [37], line-edge roughness [38], process-variation effect [39], [40], and so on. In [36], ML technique is proposed to suppress the WKF effect on the characteristics of multichannel NS devices. The deep learning techniques were applied in [41] to predict the fluctuations due to WKF on dc characteristics of the GAA NWFETs. ML is utilized to identify the WKF patterns on the metal gate to reduce the impact on dc characteristics of the GAA NSFETs [42].

All these ML-based past works have considered only dc characteristics of nanodevices, and no work talks about the high-frequency applications. Furthermore, in these works, only one fluctuation source was studied. In nanodevices, however, the effects of multiple fluctuation sources should also be evaluated simultaneously, because sometimes, considering only one fluctuation can result in unexpected variations due to other correlated fluctuation sources. Therefore, in this article, for the first time, the effects of three major fluctuation sources (WKF, ITF, and RDF) are studied individually as well as together, using the ML-based artificial neural network (ANN) model. The GAA Si NSFET is used as an example nanodevice, and its dc and analog/RF characteristics are explored in detail. The

TABLE I
LIST OF GAA Si NS MOSFET DEVICES PARAMETERS CORRESPONDING TO SUB-3-nm TECHNOLOGY NODES AND NOMINAL DEVICE CHARACTERISTICS

Parameters	Value
<i>Gate Length</i> (L_G) (nm)	12
<i>Channel Doping</i> (cm^{-3})	5×10^{17}
<i>Nanosheet Width</i> (W_{NS}) (nm)	25
<i>Nanosheet Height</i> (H_{NS}) (nm)	5
S_{ext}/D_{ext} Length (nm)	5
<i>S/D Doping</i> (cm^{-3})	1×10^{20}
<i>Sext/Dext Doping</i> (cm^{-3})	4.8×10^{18}
<i>EOT</i> (nm)	0.6
I_{SAT} (A/mm)	2.29×10^{-5}
I_{OFF} (A/mm)	4.48×10^{-12}
V_{TH} (mV)	260
<i>Gate Capacitance</i> (C_G)(F)	2.5×10^{-17}
<i>Voltage Gain</i> (dB)	13.75
<i>Cut-off Frequency</i> (f_t) (Hz)	2.27×10^{12}
<i>3-dB Frequency</i> (f_{3dB}) (Hz)	1.06×10^{11}

important parameters for high-frequency applications, such as g_m , f_t , f_{3dB} , and transconductance efficiency (g_m/I_D), are extracted using ML and compared with the simulated data by using the well-known Sentaurus¹ TCAD tool [43].

The rest of this article is organized as follows. Section II presents the structure of the GAA Si NSFET with simulation setup and data generation process including the overview of various random fluctuation sources. Section III demonstrates the ML modeling approach and the integration of ML with device simulation. In Section IV, the results from ANN are examined and compared with the simulation data. Finally, Section V concludes the work of this article.

II. DEVICE DESIGN, SIMULATION SETUP, AND DATA GENERATION

A. Nanosheet Structure and Parameters

A 12-nm GAA Si NSFET with the width (W_{NS}) and the height (H_{NS}) of 25 and 5 nm, respectively, is designed using the 3-D device simulation. The nominal device design parameters and doping concentrations of the NS device including the important figure of merits (FoMs) are summarized in Table I. To provide the best accuracy of device simulation, device characteristics are simulated after the careful calibration of parameters, as shown in Fig. 1 [8]. The gate-stack consists of a $\text{SiO}_2/\text{HfO}_2/\text{TiN}$ layer with an effective oxide thickness (EOT) of 0.6 nm. Fig. 2(a) shows a 3-D schematic plot, and Fig. 2(a1) is a 2-D cross-sectional view of the GAA Si NSFET, where the statistical generations of the three random sources, WKF, RDF, and ITF, are shown in Fig. 2. In the simulation, $V_G = 0.7$ V and $V_D = 0.05$ V are used with a frequency of 1 kHz to simulate ac characteristics. The RF characteristics are obtained by using a common-source amplifier with a sinusoidal input signal. The input has an amplitude of 0.05 V, and the biasing voltages for the gate and drain are 0.4 and

¹Registered trademark.

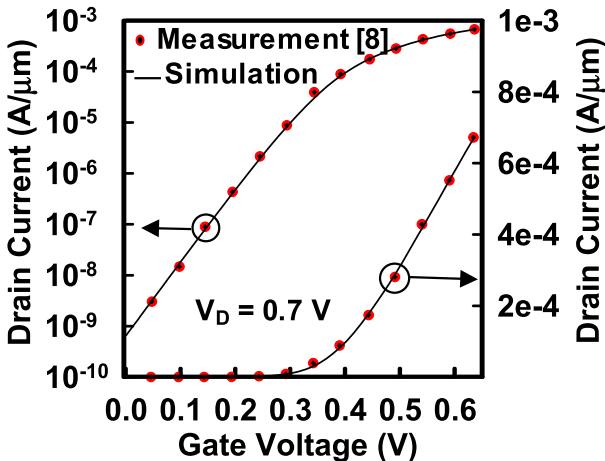


Fig. 1. Simulated (line) and measured (symbol) [8] I_D-V_G curve to calibrate the device simulation of the GAA Si NSFET.

0.5 V, respectively. For numerical simulations, the experimentally calibrated 3-D quantum-mechanically corrected density gradient model together with the drift-diffusion transport model is used [44], [45]. To produce $I-V$ characteristics that closely match to measurement data, several process variability factors, such as interface trap density 10^{11} – 10^{12} cm $^{-2}$ eV $^{-1}$ and the lattice temperature from 295–300 K, are calibrated to align the measurement. The doping dependence and high-field saturation mobility models are also included to consider the scattering, mobility of carriers, and the effect of lateral and perpendicular fields. In addition to this, band-to-band tunneling from source to channel and channel to drain is considered. The thin-layer Lombardi mobility model is used for mobility degradation at the semiconductor-insulator interface. The GAA Si NSFET with three fluctuation sources that were generated statistically by following the Gaussian distribution is shown in Fig. 2 [46], [47]. The generations of fluctuation sources are given below.

B. Device Design With Work Function Fluctuation

The work function (WK) fluctuated devices are statistically generated by introducing the metal grain (MG) of different WFs due to different grain orientations and probabilities of titanium nitride (TiN) on the metal gate layer, as shown in Fig. 2(b). The size of each MG is 3.92×3 nm 2 . The Monte Carlo method is used to generate the random patterns of low/high WF for 1000 WF fluctuated devices [48]. The high WF (WKF_H) and low WF (WKF_L) have TiN<111> and TiN<200> orientations with different probabilities, respectively, as shown in Fig. 2(b1). The probability of occurrence for WKF_H and WKF_L is 60% and 40% with $WKF_H = 4.6$ eV and $WKF_L = 4.4$ eV, respectively, and the effective WF is calculated as

$$WKF_L \times 0.4 + WKF_H \times 0.6 = 4.52 \text{ eV}.$$

The total number of MG generated on the metal gate is fixed ($MGN = 80$) for all the simulated devices. The histograms in Fig. 2(b2) show that the subplanes follow the Gauss distribution, and the average grain numbers (μ) for

TiN<111> and TiN<200> orientations are 48.1 and 31.9, respectively.

C. Device Design With Random Dopant Fluctuation

In Fig. 2(c), we demonstrate the GAA Si NSFETs with different types of random dopants present at various regions, such as channel (CH), source extension (S_{EXT}), drain extension (D_{EXT}), and penetration from the S/D extensions into the channel (PE). As shown in Fig. 2(c1)–(c4), four large cuboids are formed with equivalent doping concentrations of 5×10^{17} , 1.1×10^{19} , and 3.36×10^{17} cm $^{-3}$ for the CH region, S_{EXT} region, D_{EXT} region, and PE region, respectively. The random dopants are statistically generated by the MC method in all cuboids [20]. Afterward, these large cuboids are partitioned into 1000 subcuboids with dimensions $25 \times 12 \times 5$ nm 3 for random dopants in the channel and penetration from the S/D extensions into the channel and with dimensions $25 \times 5 \times 5$ nm 3 for source extension and drain extension. These small cuboids are mapped into CH, S/D $_{EXT}$, and PE regions for each fluctuated device. The average number of discrete dopants in CH, S/D $_{EXT}$, and PE regions is 5.7, 64.6, and 5.7, respectively. To generate discrete dopants in these regions, the Gaussian distribution is used, as shown in Fig. 2(c1)–(c4).

D. Device Design With Interface Trap Fluctuation

For ITF simulations, 2700 acceptor-type random interface traps (RITs) are statistically generated in a large square plane of size 224×224 nm 2 , as shown in Fig. 2(d1), and the equivalent concentration in the entire plane is about 3×10^{13} cm $^{-2}$. The RITs are statistically generated using Gaussian distribution at the interface of SiO $_2$ and Si channel, as shown in Fig. 2(d2). The dimension of each RIT is 2×2 nm 2 with a density around 1.5×10^{13} – 7.6×10^{13} cm $^{-2}$ eV $^{-1}$, as shown in Fig. 2(d3). Thereafter, this large plane is divided into several subplanes of size 25×12 nm 2 for top and bottom areas and of size 12×5 nm 2 for side areas of the GAA Si NS MOSFETs [49], [50]. In these subplanes, the traps are varied from 0 to 14 and the average number of RITs is 9. The density of each RIT (D_{it}) on the subplane is assigned according to the relation of trap's density versus trap's energy, as represented in Fig. 2(d4). Thus, 1000 randomly generated 3-D device samples are simulated to investigate the impact of RITs on the electrical characteristics of the explored GAA Si NSFETs.

E. Device Design With Combined Fluctuation

In the end, to analyze the combined effect of all the fluctuations, i.e., WKF, RDF, and ITF together, 1000 GAA Si NS MOSFETs devices are simulated with the same distribution procedure of all sources as mentioned above. In this way, a complete dataset of 16 000 simulations is generated with 1000 samples for each fluctuation source: WKF, RDF, ITF, and combined fluctuation (CF) for the transfer (I_D-V_G), the output (I_D-V_D), C_G-V_G , and frequency response characteristics. This TCAD-generated data are intensively used for training and testing the ANN Model.

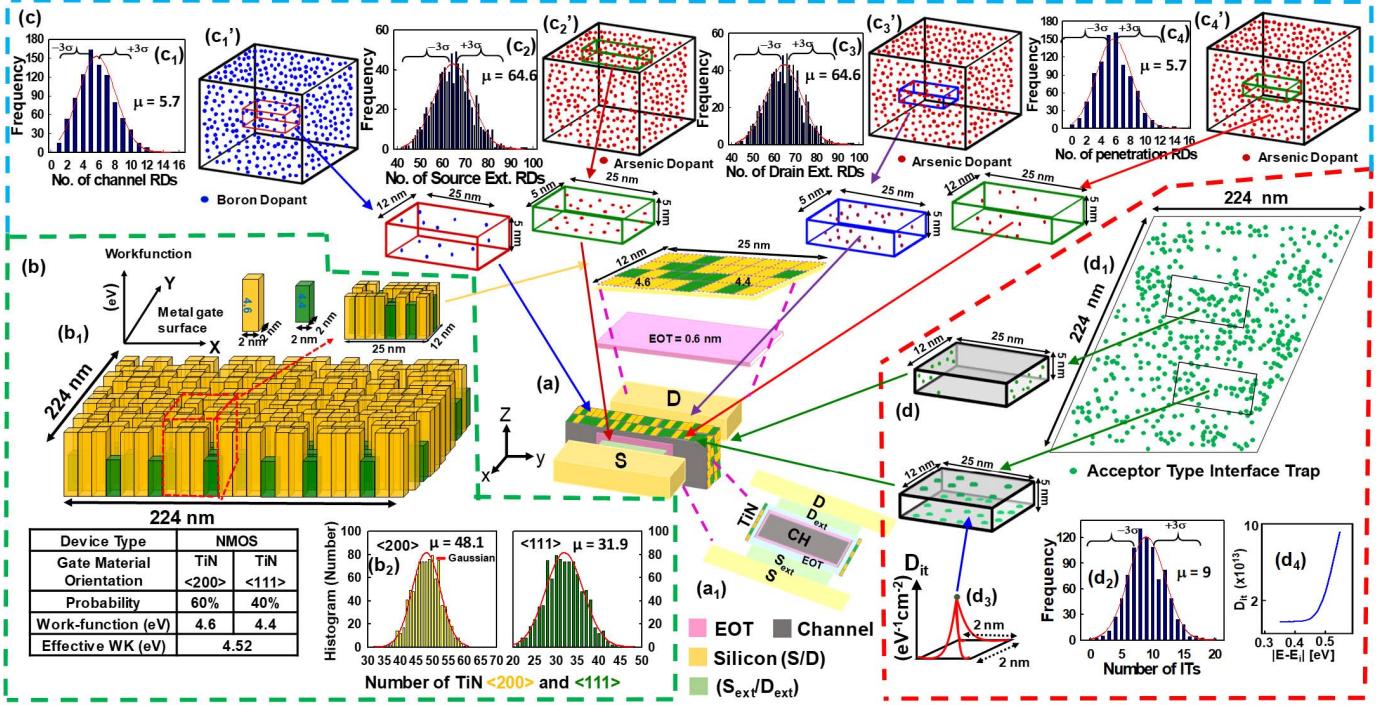


Fig. 2. (a) 3-D schematic and (a₁) 2-D cross-sectional view of the nominal GAA Si NSFET. (b) Random metal grains are generated in a large area. (b₁) Work function of every subplane is according to the material's property as shown in the inset table. (b₂) Total number of <200> (yellow) and <111> (green) orientations in the subplanes follows the Gaussian distribution with the mean value of 48.1 and 31.9, respectively. Subplanes are mapped to four sides of the metal gate. (c) 1725, 19 396, and 1718 random dopants are generated in a large cuboid with a doping concentration of 5×10^{17} , 1.1×10^{19} , and $3.36 \times 10^{17} \text{ cm}^{-3}$ for the channel (CH), source-drain extension (S/D_{EXT}), and channel penetration (PE), respectively. (c₁)–(c₄) Randomly generated dopants have Gaussian distribution, where the average dopants in each square are 5.7, 64.6, 64.6, and 5.7, respectively. (d) Statistical device simulation of ITF. (d₁) 2700 acceptor-type interface traps are generated in a 224-nm square large plane, where (d₂) each trap has a 2 nm^2 size and the density of each trap is around $1.5\text{--}7.6 \times 10^{13} \text{ cm}^{-2} \text{ eV}^{-1}$ and the plane density is $3 \times 10^{13} \text{ cm}^{-2}$. (d₃) Average of ITs is 9 in each plane. (d₄) Relation between trap density versus trap energy, where density of each trap on the plane is assigned according to distribution of trap's energy.

III. MACHINE LEARNING MODEL AND DATA PREPROCESSING

The work in this article is associated with the supervised regression problem, and after exploring many algorithms, the ANN is selected for the prediction of dc and analog/RF characteristics. The ANN is comprised of node layers, such as an input layer, one or more hidden layers, and an output layer. Each node also called an artificial neuron connects to another node and has an associated weight and bias (or threshold). When the output of any neuron is greater than the specified threshold value, then that particular neuron is activated, and the data are passed to the neuron of the next layer. This threshold is set by using the activation function. In this article, we constructed an ANN model using Python's Keras library with Tensorflow working in its backend [51] to perform the desired task and changed its hyperparameters according to the applications. Fig. 3 shows the overall ANN architecture used in this work. The models have two hidden layers, with Relu and Tanh as activation functions for different applications. The activation function suppresses the neurons whose inputs are of no relevance to the overall application of the model. Therefore, the ANN requires these functions to provide a significant performance improvement. The mathematical expressions of these activation functions are defined as follows:

$$\text{Relu} : f(x) = \begin{cases} 0, & \text{for } x < 0 \\ x, & \text{for } x \geq 0 \end{cases} \quad (1)$$

and

$$\text{Tanh} : f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}. \quad (2)$$

Therefore, in this process, first, the inputs are multiplied with their respective weights (gradient or coefficient of each variable). The weights are used to determine the importance of a given variable. After assigning the weights, the bias variable is added as follows:

$$z = \sum_{i=1}^m \omega_i x_i + b = \omega_1 x_1 + \omega_2 x_2 + \omega_3 x_3 + b \quad (3)$$

where ω_i is the weight and b is the bias. This output from the hidden layer is then passed through the activation function and then send to the next hidden layer as its input. A similar process occurred at every layer, and this process is called forward propagation

$$Y = \alpha(z) \quad (4)$$

where α is the activation function. The purpose of training an ANN is to minimize the error to make a prediction as close to the target value as possible. The training process consists of two parts: backpropagation and optimization. Backpropagation is responsible for the backward propagation of errors. It is an algorithm that computes the gradient of the loss function with respect to the weights. After each iteration, the weights are updated, such that errors should be minimum. Finally, the optimization is used to select the best weights and biases for

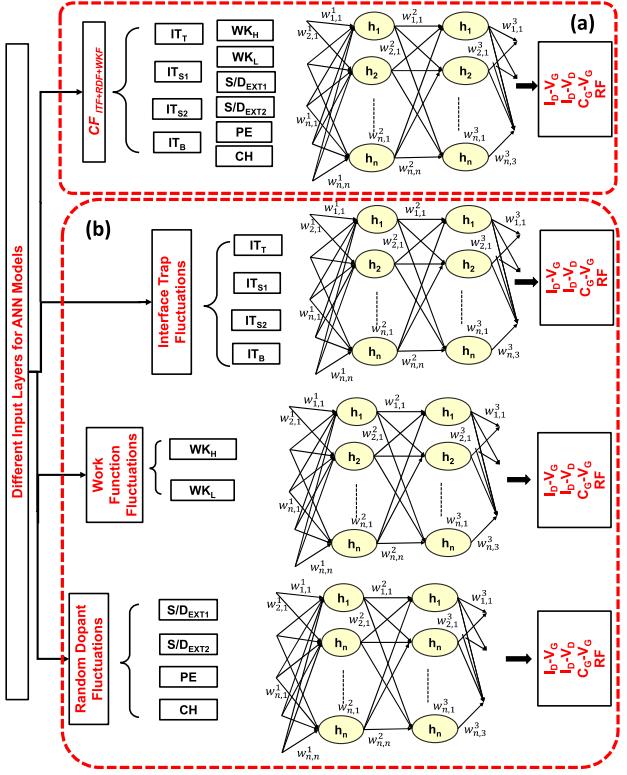


Fig. 3. Illustration of ML-based ANN model architecture for GAA Si NSFET simulated dataset. Two hidden layers, neurons, and weight and bias are shown here. (a) ANN model was trained and evaluated using CF, i.e., all sources of variations (ITF, RDF, and WKF). (b) Three ANN models were trained and evaluated for different random sources of fluctuations.

each neuron. For this work, we use the Adam algorithm as the optimizer. The weights and bias are updated as follows:

$$\omega_i = \omega_i - \left(\alpha \times \frac{\partial C}{\partial \omega_i} \right) \quad (5)$$

and

$$b = b - \left(\alpha \times \frac{\partial C}{\partial b} \right) \quad (6)$$

where C is the cost function, and $(\partial C / \partial \omega_i)$ and $(\partial C / \partial b)$ are the gradients of the cost function with respect to weights and bias, respectively. The backpropagation and gradient descent process are repeated until convergence. The loss function is used to figure out how far we are from our desired solution or final convergence. It is very important to choose a proper loss function for ANN training to achieve high accuracy. The root mean square error (RMSE) is chosen as the loss function for our problem. It is calculated for entire training dataset, and their average is called the cost function C and is defined as

$$C = \text{RMSE} = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2} \quad (7)$$

where y_i is the actual value and \hat{y}_i is the predicted value.

Furthermore, data processing is an essential integral part of ML. It ensures the quality of data and transforms the raw data into useful information that directly affects the learning ability of the ANN model. Therefore, instead of fitting the ANN on raw data, it must be transformed to meet the requirements. In this work, to model the transfer characteristics (I_D - V_G), the

raw current values (I_D) are first preprocessed before feeding them into the model as their range is too wide and values are very low from 10^{-5} to 10^{-12} . Therefore, we use $\log(I_D)$ to scale I_D since it gives better prediction [52]. The ANN model, which deals with I_D - V_G characteristics, used Tanh as an activation function because the $\log(I_D)$ transformed the I_D values to negative values. The Tanh maps the negative values to strongly negative, whereas the Relu maps all the negative values to zero. Therefore, Relu decreases the ability of the ANN model to train properly, and therefore, the prediction accuracy is affected severely. On the other hand, in I_D - V_D , C_G - V_G , and frequency response characteristics, the values have smaller range differences, and therefore, Relu is used as an activation function, and the MinMax scaler, from scikit-learn [53], is used to scale the data. The MinMax scaler is defined as

$$\text{Normalized value } x_n = \frac{(x_i - \min(x))}{(\max(x) - \min(x))}. \quad (8)$$

Fig. 4 shows the integration of two different fields: device simulation and ML. First, the data are simulated using the process as described by the flow diagram in Fig. 4(a). After generating the simulated dataset, the key parameters are extracted. The raw dataset is then preprocessed and prepared for the ANN model. The processed dataset is split into train and test components. The training dataset is used to build the ML model, as shown in Fig. 4(b), and the test dataset is used for validating the model. The model is trained using backpropagation, and the learning parameters are updated after each iteration. Once the training is done, the performance of the well-trained model is evaluated by exposing it to the test dataset and the output values are predicted. The key parameters are extracted from the predicted output curves and compared with the simulated values. Furthermore, the ML models are fully data-driven. Therefore, they learn the input–output relationship during the model training phase without requiring any device physics domain expertise [54].

IV. RESULTS AND DISCUSSION

The TCAD-simulated data are first split into train and test sets as 80% and 20%, respectively. The input of the ANN consists of the fluctuation sources governing parameters as inputs.

- 1) For WKF, counts of high and low WK [WK_H, WK_L];
- 2) For RDF, counts of random dopants in the channel, source–drain extensions, and channel penetration [CH, S/D_{EXT1}, S/D_{EXT2}, PE];
- 3) For ITF, counts of interface traps at the top, bottom, side-1, and side-2 [IT_T, IT_B, IT_{S1}, IT_{S2}]; and
- 4) For CF, the input pattern consists of all the fluctuation counts [WK_H, WK_L, CH, S/D_{EXT1}, S/D_{EXT2}, PE, IT_T, IT_B, IT_{S1}, IT_{S2}].

The ANN model is then trained and hyperparameters are tuned. To avoid overfitting, a parameter ‘‘patience’’ is set as an input argument to the early stopping function provided by Keras [51].

This approach not only restrains the overfitting but also saves computational time by stopping the training process

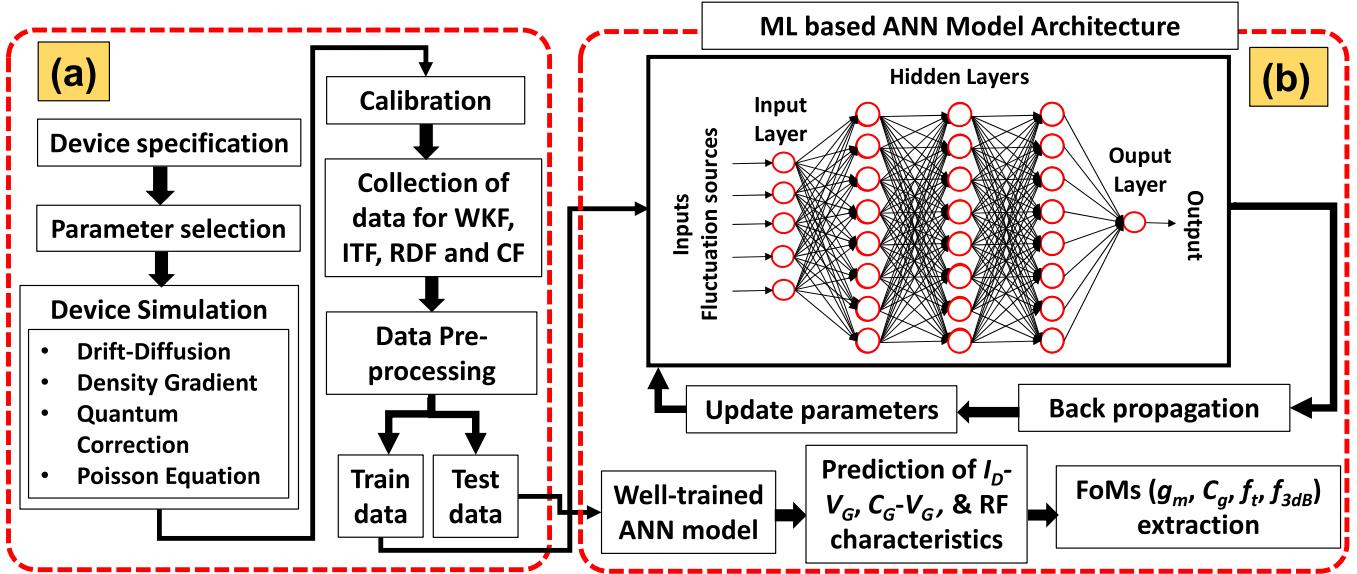


Fig. 4. Illustration of the schematic flowchart of integration of two different fields, i.e., device simulation and machine learning algorithm. (a) represents the crucial processing steps to simulate the device and generate the data and (b) depicts the application of the ML algorithm on the simulated dataset. Process including the train-test data split, parameters updating during the training phase using backpropagation, validation on testing data, predicting the characteristics, and finally extracting the key FoMs is shown in detail here.

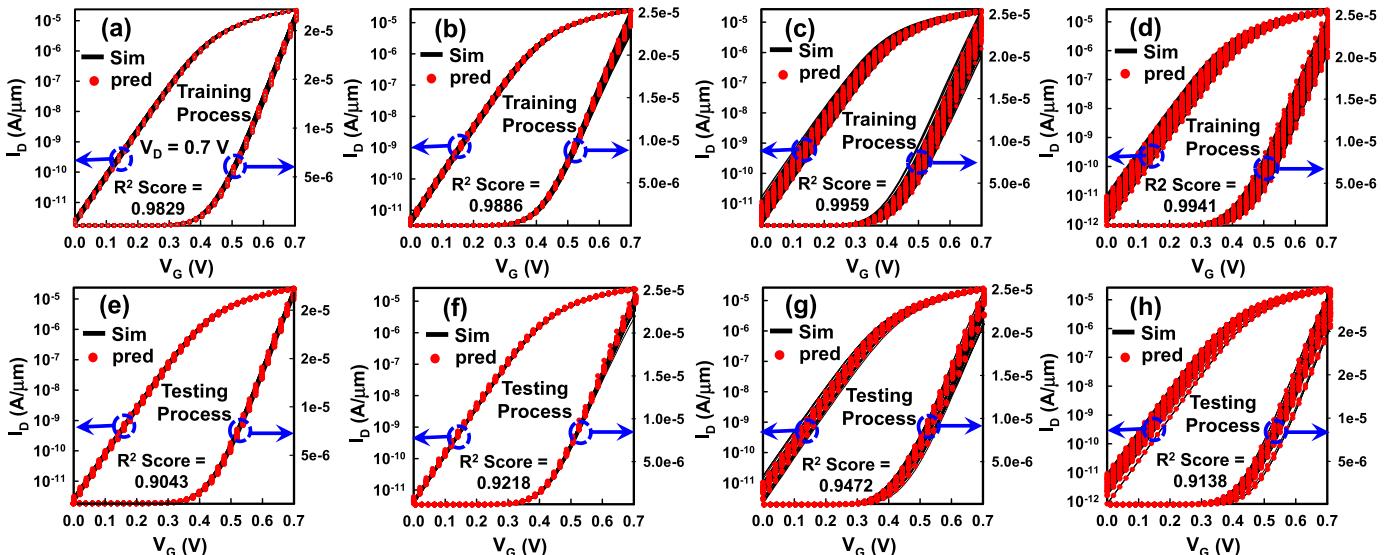


Fig. 5. Plots depict the simulated (black line) as well as predicted (red symbol) I_D - V_G curves for ITF, RDF, WKF, and CF. (a) and (e) show the effect of ITF on the electrical characteristics of GAA Si NS MOSFETs during the training and testing phase by the explored ANN model. Simulated (line) and predicted (symbol) characteristics demonstrate that the machine prediction and the device simulation are in good agreement. (b) and (f) represent the RDF effect during training and testing of the ANN model, respectively. Similarly, (c) and (g) show the WKF effect for training and testing processes, respectively. At last, (d) and (h) show the estimated results for the overall cumulative effect of all fluctuations, i.e., WKF, RDF, and ITF.

early. After the successful training, we evaluated the model using RMSE and R^2 -score. The R^2 -score, also known as the coefficient of determination, shows how well the ANN model fits on data.

The R^2 score is defined as

$$R^2 = \frac{\text{A total sum of square} - \text{Residual sum of square}}{\text{A total sum of square}}$$

$$\Rightarrow = 1 - \frac{\sum_i^n (Y_i - \hat{Y})^2}{\sum_i^n (Y_i - \bar{Y}_m)^2} \quad (9)$$

where \bar{Y} is the mean of the simulated values.

A. Effect of Fluctuations on Transconductance Efficiency (g_m/I_D)

The fluctuations present in the NSFET devices severely affect their characteristics. This results in degradation of drain current and transconductance. g_m is the crucial parameter in deciding the dc as well as analog/RF performance of a semiconductor device, and hence, it needs to be analyzed carefully. It is defined as follows:

$$g_m = \frac{\partial I_D}{\partial V_G} \quad (10)$$

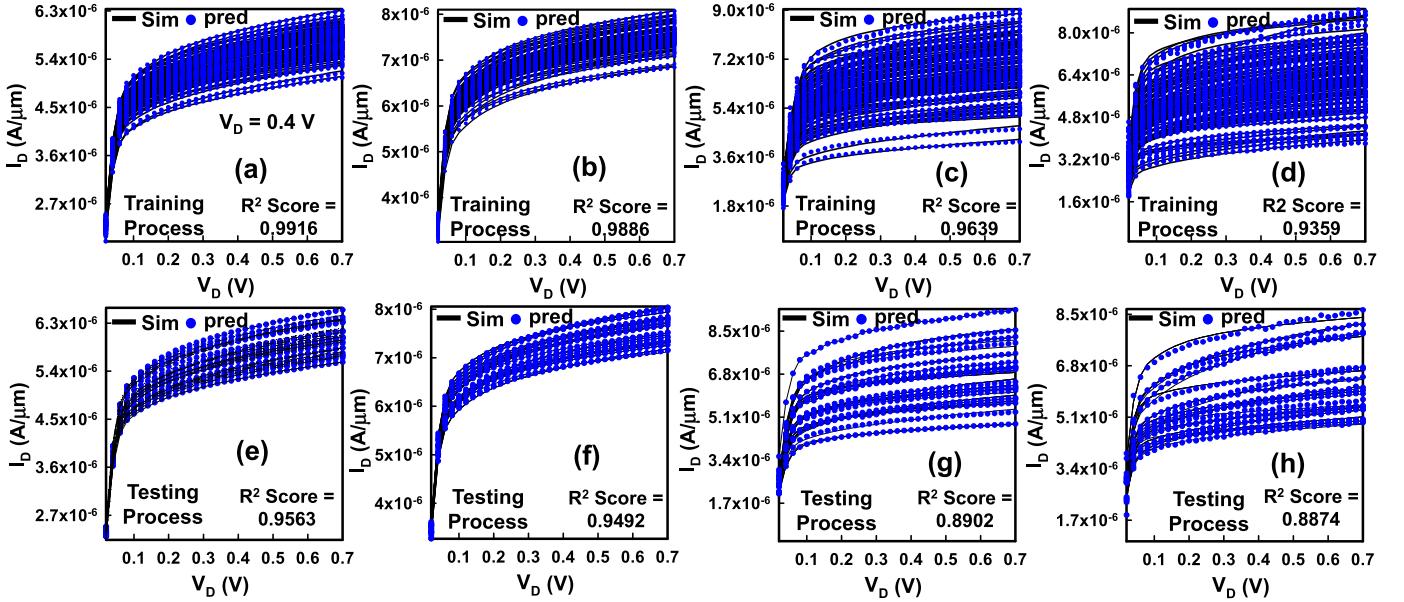


Fig. 6. Plots depict the simulated as well as predicted I_D - V_D curves induced by ITF, RDF, WKF, and CF. (a) and (e) show the effect of ITF on the electrical characteristics of GAA Si NS MOSFETs during the training and testing phases of our explored ANN model. The simulated (line) and predicted (symbol) curves demonstrate that the machine prediction and the device simulation are in good agreement. (b) and (f) represent the RDF effect during training and testing of the ANN model, respectively. (c) and (g) describe the WKF effect for training and testing process, respectively. (d) and (h) show the overall cumulative effect of all fluctuations, i.e., WKF, RDF, and ITF for training and testing of the ANN model, respectively.

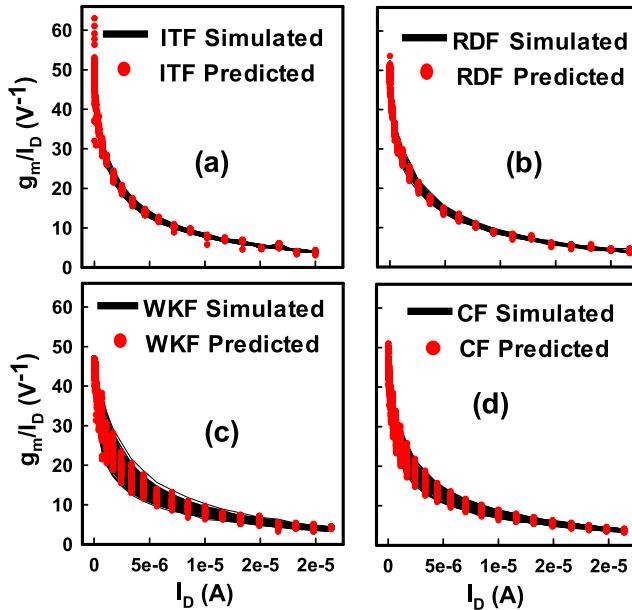


Fig. 7. g_m , extracted from the ANN predicted dc characteristics, is used to calculate the transconductance efficiency g_m/I_D that is used as an indicator to show the capability of NSFET to convert dc power into ac gain at fixed drain bias. (a)-(d) illustrate the variance of test results in the g_m/I_D under the influence of various fluctuation sources.

where ∂I_D denotes the change in drain current and ∂V_G is the change in the gate voltage. The notation g_m represents the order of derivative of I_D with respect to V_G . Therefore, the dc characteristics of the explored GAA Si NSFET are examined before discussing the RF FoMs. Notably, the variations on I_D and g_m are modeled using the ANN, and the important FoM, transconductance efficiency (g_m/I_D), is explored. The potential and the effectiveness of the ANN-based approach are also reviewed. After the successful training, the performance

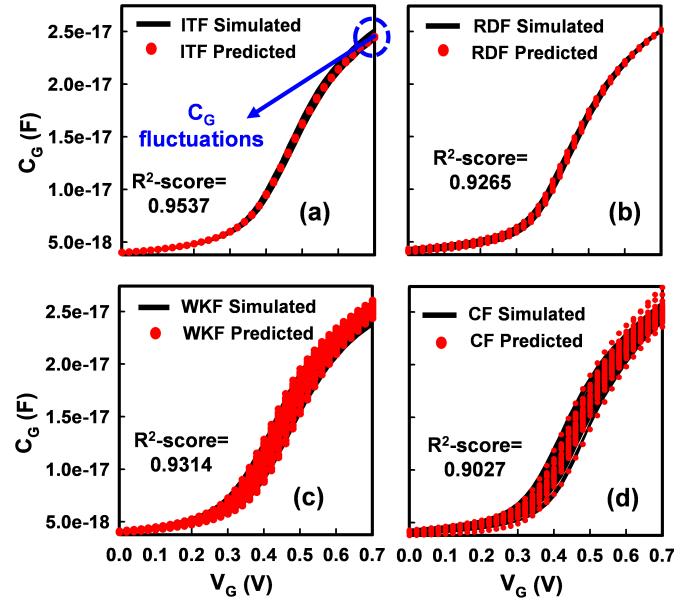


Fig. 8. Test results for gate capacitance (C_G) variations with respect to gate voltage (V_G) are analyzed using the ANN model. The comparison curves between the simulated values (line) versus predicted (line) versus predicted values (symbols) among (a) ITF, (b) RDF, (c) WKF, and (d) CF.

of the ANN is validated by using the test dataset. Fig. 5(a)–(h) represents the simulated and the predicted train-test transfer characteristics curves (I_D - V_G) of GAA Si NSFET.

Similarly, Fig. 6(a)–(h) shows the output characteristics curves (I_D - V_D) for NSFET. It can be observed from the figures that the predicted results for variability sources ITF and RDF are more precise than WKF and CF. The curves show a good agreement between the predicted and the simulated values. We can examine from these figures that the ITF has the

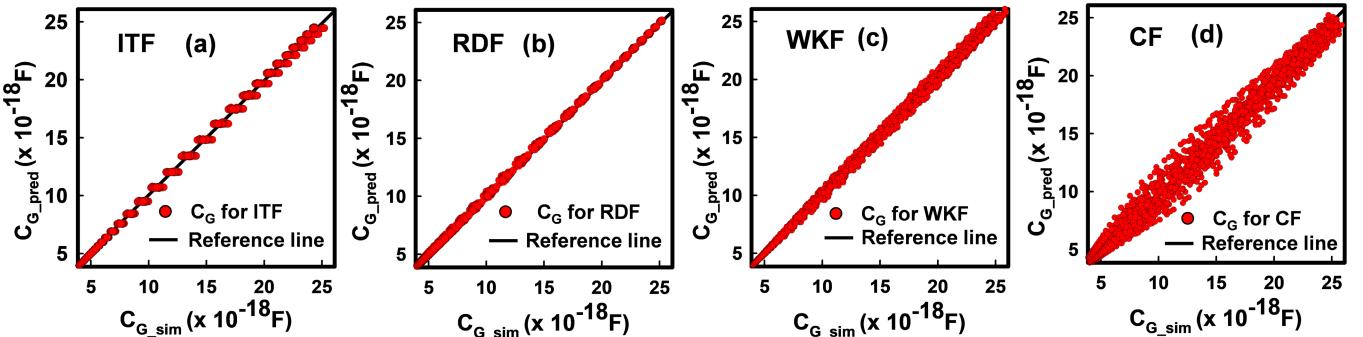


Fig. 9. Test scatter plots for the prediction of gate capacitance (C_G) for (a) ITF, (b) RDF, (c) WKF, and (d) CF. It is clearly evident from the plots that the ANN predicted values are in good agreement with the simulated values.

lowest effect, whereas the WKF has the highest effect on the transfer and output characteristics of NSFET. In the case of CF, the variation is less than the WKF owing to the fact that some effects of different fluctuations mutually cancel out each other due to the randomness of the fluctuation sources. To further confirm the prediction accuracy, the R^2 score is calculated. For the fluctuation sources, ITF, RDF, WKF, and CF, the training R^2 scores for $I_D - V_G$ are 0.9829, 0.9886, 0.9959, and 0.9941, respectively, and for $I_D - V_D$, 0.9916, 0.9886, 0.9404, and 0.9359, respectively, which ensures that the ANN models get trained efficiently. Similarly, the performance of the ANN model on test data is recorded with R^2 -score, for $I_D - V_G$ are 0.9043, 0.9208, 0.9472, and 0.9138, respectively, and for $I_D - V_D$, 0.9563, 0.9492, 0.8902, and 0.8874, respectively. From the predicted transfer characteristics, g_m is extracted to calculate the transconductance efficiency, which is the key FoM for analog/RF applications. The high transconductance values lead to high g_m/I_D that eventually provides a higher ability to convert dc power to ac gain for a fixed drain current bias (V_D), which is suitable for analog applications. However, the characteristics are severely affected when the nanodevices are suffered from various fluctuations. For a fixed V_D , the devices have variable characteristics and it also affects the g_m and g_m/I_D ratio.

From Fig. 7(a)–(d), we can inspect the difference in g_m/I_D that occurred due to various fluctuations. Even for similar process parameters and bias conditions, the characteristics of semiconductor nanodevices are varied under the influence of fluctuations. The g_m is extracted from the predicted transfer characteristics, and then, ANN predicted g_m/I_D ratio is compared with simulated values. The figures show that the ANN predicted values are close to the simulated values. We can also observe from the figure that the g_m/I_D ratio is varied differently for each fluctuation source. The GAA Si NSFET exhibits superior g_m/I_D , but it is highly suppressed due to WKF fluctuation. The effect of CF is also prominent; however, the effect of RDF and ITF has comparatively very less influence on NSFET devices.

B. Effect of Fluctuations on Gate Capacitance (C_G)

The gate capacitance is an important parameter since the cutoff frequency (f_t) of the device is inversely proportional to it. The low parasitic capacitance means high f_t and it also improves the device speed. The fluctuations on the GAA

Si NSFET change the charge distribution under the gate and degrade the C_G . Fig. 8(a)–(d) shows the ANN fitting results for $C_G - V_G$ characteristics of NSFET. The behavior of parasitic capacitance is illustrated in the presence of various fluctuations. From Fig. 8(c) and (d), we can observe that C_G is increased due to the effect of WKF and CF fluctuations, which results in the reduced device speed. The large C_G is a dominant factor that significantly degrades the characteristics of the GAA Si NSFETs. Furthermore, the investigation of ITF and RDF effect, in Fig. 8(a) and (b), shows that their impact is low on C_G , which leads to fewer variations. The fitting parameter R^2 -score is calculated for each fluctuation. For ITF, RDF, WKF, and CF, the training R^2 -scores are 0.9810, 0.9824, 0.9639, and 0.9226, respectively, and test R^2 -scores are 0.9537, 0.9265, 0.9265, and 0.9226, respectively, which ensures the better fitting of the ANN model. Additionally, the scatter plots in Fig. 9(a)–(d) show the relationship between the actual simulated versus the predicted C_G characteristics for different fluctuation sources. The ANN model achieved the predicted values close to the simulated values effectively except for a few outliers.

C. Effect of Fluctuations on Cutoff and 3-dB Frequencies

To examine the influence of fluctuations on the RF characteristics of the GAA Si NSFET devices, we explored the important FoMs, the cutoff frequency, and the 3-dB frequency. The f_t of a device is the most important parameter for high-frequency applications and is directly proportional to the transconductance and inversely proportional to the gate capacitance

$$f_t = \frac{g_m}{2\pi C_G} \quad (11)$$

where f_t is the frequency, where the gain of the device reduces to unity, and $f_{3\text{dB}}$ is the frequency calculated at 3-dB point gain, where the power is half of its maximum value.

The 3 dB is calculated by

$$-3 \text{ dB} = 20\log_{10}\left(0.707 \times \frac{v_{\text{out}}}{v_{\text{in}}}\right). \quad (12)$$

Fig. 10(a)–(d) shows the simulated versus predicted testing results of the RF characteristics. The black thick line is the nominal device characteristic, the green lines show the simulated values, and the red lines show the ANN predicted

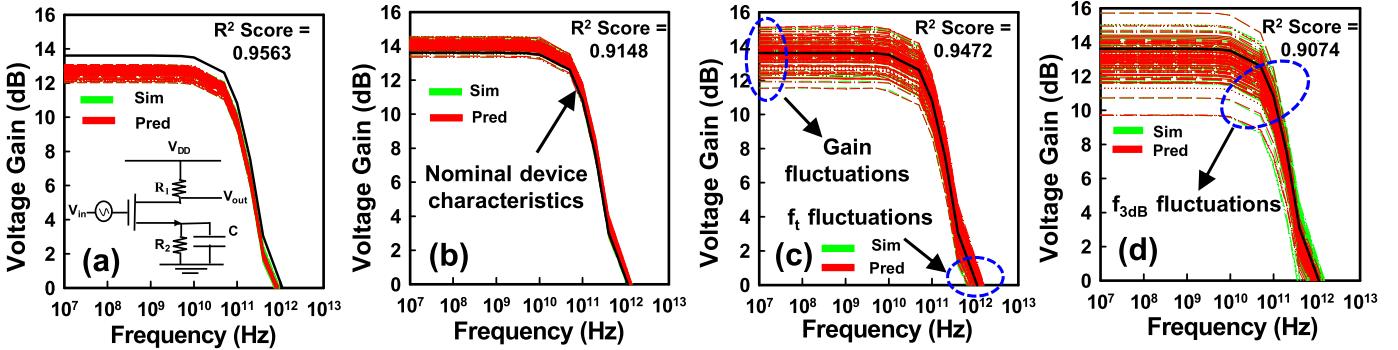


Fig. 10. Comparison of voltage gain versus frequency characteristics generated through device simulation and their respective predicted curves through the ANN models. Circuit inset in (a) is a common-source amplifier used for RF simulations, where the values of R_1 is 50 K Ω , R_2 is 10 K Ω , and C is 10^{-6} F. Both V_{in} and V_{out} are the input and output signals of the circuit. Plots (a)–(d) show the nominal device (black line), various fluctuation-based simulated (green lines), and predicted (red lines) curves. Key areas of FoMs are also pointed out in plots.

TABLE II

R²-SCORE AND RMSE VALUES OF TRAINING AND TESTING OF DIFFERENT CHARACTERISTICS FOR VARIOUS FLUCTUATIONS (ITF, RDF, WKF, AND CF)

Characteristics	Fluctuations	R ² -score (Train)	R ² -score (Test)	rmse (Train)	rmse (Test)
I_D - V_G characteristics	ITF	0.9829	0.9043	1.98e-07	2.55e-07
	RDF	0.9886	0.9208	4.44e-07	4.79e-07
	WKF	0.9959	0.9472	1.68e-07	2.66e-07
	CF	0.9941	0.9138	3.77e-07	5.29e-07
I_D - V_D characteristics	ITF	0.9916	0.9563	0.1997	0.2276
	RDF	0.9886	0.9492	0.2483	0.2907
	WKF	0.9404	0.8902	0.1264	0.1937
	CF	0.9359	0.8874	0.4159	0.4841
C_G - V_G characteristics	ITF	0.9810	0.9537	1.6023	1.8134
	RDF	0.9824	0.9265	1.3867	1.5707
	WKF	0.9639	0.9314	0.9488	1.0642
	CF	0.9226	0.9027	2.1407	2.3428
Frequency response characteristics	ITF	0.9751	0.9563	0.0595	0.0649
	RDF	0.9454	0.9148	0.0550	0.0692
	WKF	0.9778	0.9472	0.2253	0.2806
	CF	0.9607	0.9074	0.2135	0.3396

outputs. The important areas, from where the FoMs are calculated, are also highlighted in the figures. The results show that the trained ANN model predicted the variations efficiently. The fluctuations induced by WKF, ITF, RDF, and CF are captured accurately. From the figure, we also deduce that the gain has the largest variations due to CF fluctuations from 9.8 to 16 dB. The ITF has the lowest variation range from 12 to 13.1 dB. For WKF and RDF, the range is from 11.8 to 15 dB and 13.7 to 15 dB, respectively. The training R^2 -scores for ITF, RDF, WKF, and CF are 0.9751, 0.9454, 0.9778, and 0.9607, respectively, and test R^2 -scores are 0.9563, 0.9148, 0.9472, and 0.9074, respectively. The complete list of performance parameters is summarized in Table II.

It is worth mentioning at this point that although the nature of variations of all four types of fluctuation is very different, still the ANN model is able to perform well and imitate the target values. It is also important because the accurate modeling of device fluctuations is critical to circuit performance. Furthermore, the FoMs, f_t , and f_{3dB} values are also extracted from simulated and predicted RF characteristics.

In Fig. 11(a)–(l), the scatter plots for f_t , f_{3dB} , and gain are shown. It can be analyzed from the figures that the ANN model learned the trend precisely by utilizing the training knowledge. To further examine the accuracy of the ANN model, the extracted f_t and f_{3dB} values are compared with the actual simulated values by calculating the relative standard deviation (RSD). RSD or the coefficient of variation is often more convenient. It is useful when comparing the predicted results with the actual simulated data. In Fig. 12(a) and (b), the RSD between actual and predicted f_t and f_{3dB} is shown using a bar graph for various fluctuation sources. The yellow bar shows the actual simulated RSD values, and the green bars show the predicted RSD value of f_t and f_{3dB} , for ITF, RDF, WKF, and CF. The small differences between the simulated and predicted RSD represent that the ANN model predicted the data more precisely.

Until now, we analyze the prediction capability of the ANN model against the simulated technique by using the curve comparison, and we also calculate the R^2 -score that shows the fitting ability of the model. However, besides such

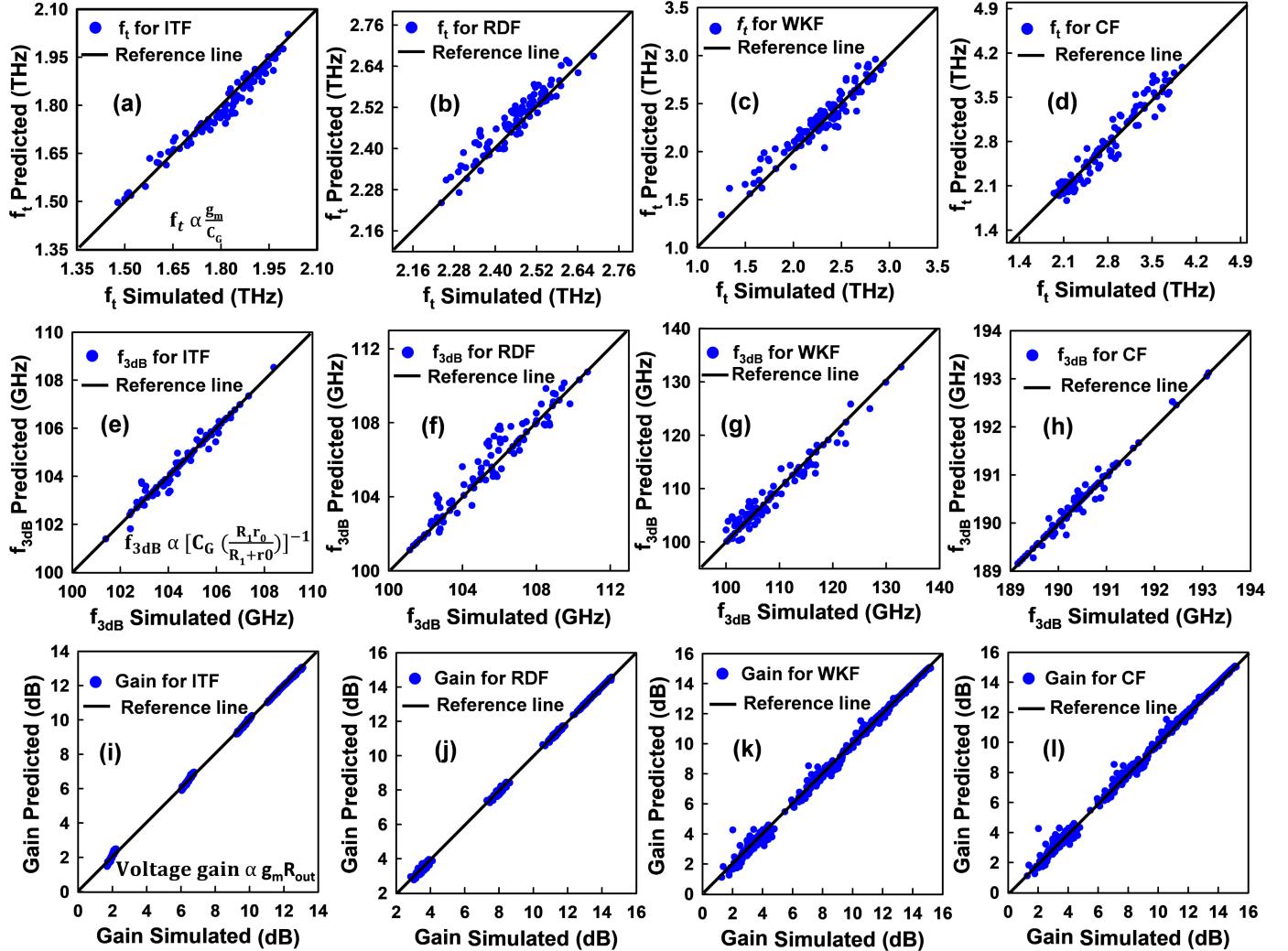


Fig. 11. Scattered plots for comparing the simulated versus predicted f_t for (a) ITF, (b) RDF, (c) WKF, and (d) CF. Similarly, (e), (f), (g), and (h) show the simulated versus predicted f_{3dB} of ITF, RDF, WKF, and CF, respectively. Lastly, (i) ITF, (j) RDF, (k) WKF, and (l) CF represent the comparison of simulated versus predicted gain values extracted from the curves of the frequency response plot in Fig. 10 for different sources of variations. Results represented that the ANN model is effectively learning the trend from the frequency response characteristics. Black line shows the reference line or ideal line and the blue symbols show the predicted values.

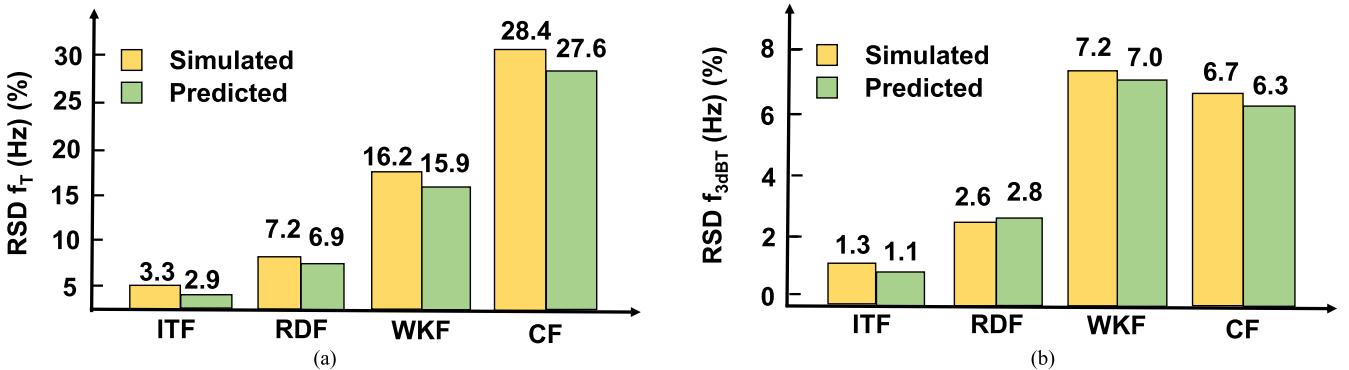


Fig. 12. Comparison between the RSD values for device simulation and the ML-based ANN model prediction of (a) f_t and (b) f_{3dB} , for the test dataset. RSD is calculated as $(\sigma/\mu) \times 100\%$.

visual comparison, it is imperative to calculate the mean value and standard deviation of essential parameters for quantitative assessment. For transfer characteristics, we extracted the key parameters, such as I_{off} , I_{on} , SS, and V_{TH} , from the simulated as well as predicted I_D-V_G curves. The mean values and standard deviations of these parameters are calculated, and the error rate between the simulated and predicted value is

examined for different fluctuation sources. The error rate in percentage is defined as follows:

$$\text{Error}(\%) = \frac{\text{Simulated value} - \text{Predicted value}}{\text{Simulated value}} \times 100. \quad (13)$$

Table III lists the mean value and standard deviation of I_D-V_G characteristics with their corresponding error rates.

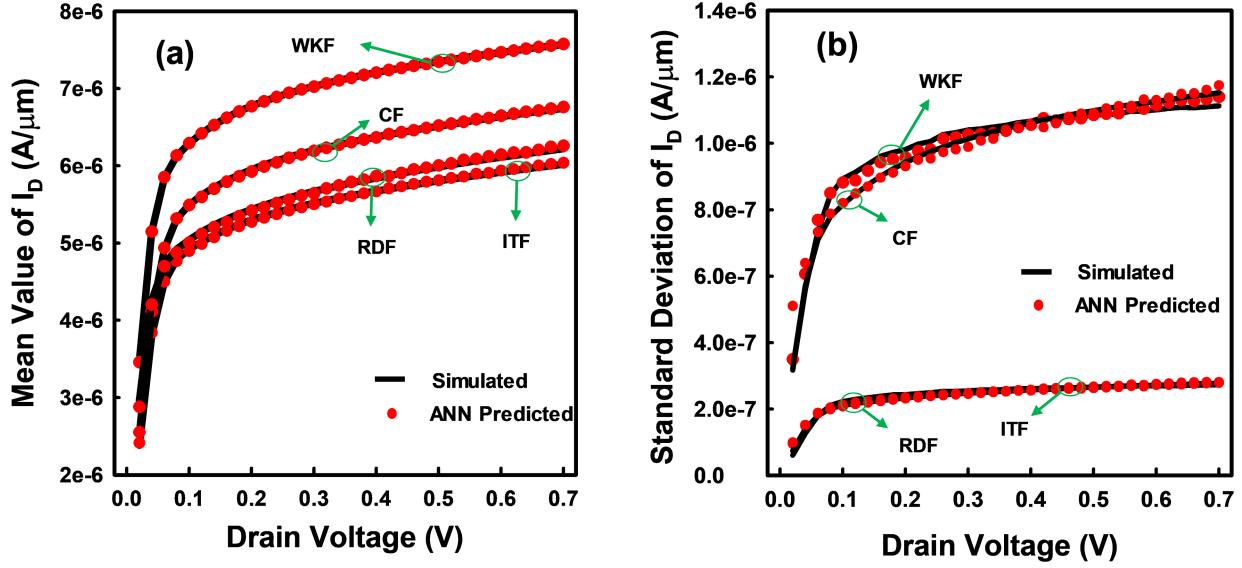


Fig. 13. Illustration of (a) mean value and (b) standard deviation of I_D - V_G curves for all sources of variations, i.e., WKF, ITF, RDF, and CF. In (a) and (b), the solid line (-) shows the data collected through device simulation of GAA Si NS MOSFET, whereas symbol (o) is the data predicted from the explored ML algorithms.

TABLE III
MEAN VALUE AND STANDARD DEVIATION OF KEY PARAMETERS EXTRACTED FROM THE I_D - V_G CHARACTERISTICS FOR VARIOUS FLUCTUATIONS (ITF, RDF, WKF, AND CF)

Parameters	Fluctuation Type	Mean Value		Error (%)	Standard Deviation		Error (%)
		Sim	ANN		Sim	ANN	
<i>Off-current</i> (I_{OFF})	ITF	2.43E-12	2.31E-12	4.94	3.16E-13	3.40E-13	7.41
	RDF	4.32E-12	4.59E-12	6.25	5.66E-13	5.30E-13	6.37
	WKF	6.13E-12	6.02E-12	1.79	2.44E-12	2.28E-12	6.63
	CF	3.99E-12	3.91E-12	2.01	2.27E-12	2.32E-12	2.23
<i>On-current</i> (I_{ON})	ITF	2.16E-05	2.15E-05	0.23	3.37E-07	3.32E-07	1.64
	RDF	2.38E-05	2.40E-05	0.83	6.75E-07	7.10E-07	5.08
	WKF	2.29E-05	2.31E-05	1.09	9.53E-07	9.94E-07	4.31
	CF	2.28E-05	2.27E-05	0.66	1.47E-06	1.50E-06	2.02
<i>Subthreshold Slope</i> (SS)	ITF	66.24619	66.29281	0.07	0.776544	0.721846	7.04
	RDF	66.83865	66.78594	0.08	0.944410	0.958536	1.50
	WKF	67.19244	67.20690	0.02	1.401944	1.409398	0.53
	CF	68.18785	67.98432	0.30	1.151313	1.094673	4.92
<i>Threshold Voltage</i> (V_{TH})	ITF	0.260533	0.258961	0.60	0.002599	0.002433	6.40
	RDF	0.274840	0.275175	0.12	0.003952	0.004158	5.21
	WKF	0.268414	0.264733	1.37	0.017858	0.017338	2.91
	CF	0.253018	0.252987	0.01	0.014323	0.013402	6.43

For the output characteristics, the mean value and standard deviation plots of I_D - V_D curves for simulated and predicted values are compared in Fig. 13 for all the fluctuated devices. Similarly, the key parameters, $f_{3\text{dB}}$ and f_t , are extracted from the curves of frequency response for simulated and predicted values. The mean value and standard deviation of these values along with the percentage error rate are listed in Table IV.

The ML-based models are highly dependent on the quality and quantity of the dataset. To know how much data are required by the ANN model to achieve maximum accuracy, we calculate the RMSE values of I_D - V_G and I_D - V_D , and frequency response curves are predicted by the ANN model for

different data samples, as shown in Fig. 14. The model training is started from 300 samples for each fluctuation source and increased by 100 samples if the accuracy is not satisfactory. We examine that up to 500 samples, the model is unable to capture the relationship between the input and output variables accurately and leads to poor predictions. The model generates a high error rate and, therefore, results in an underfit model. As the training samples are increased, the performance and accuracy of the ANN model are improved, and error starts reducing. This trend continues till 700–800 samples, where the error rate is minimum. This region is called the optimum fit, and the model achieves the highest accuracy. If we further train

TABLE IV
MEAN VALUE AND STANDARD DEVIATION OF KEY PARAMETERS EXTRACTED FROM THE FREQUENCY RESPONSE CHARACTERISTICS FOR VARIOUS FLUCTUATIONS (ITF, RDF, WKF, AND CF)

Parameters	Fluctuation Type	Mean Value		Error (%)	Standard Deviation		Error (%)
		Sim	ANN		Sim	ANN	
f_{3db}	ITF	1.05E+11	1.03E+11	1.60437	1.29E+09	1.23E+09	4.33429
	RDF	1.05E+11	1.07E+11	1.87200	2.52E+09	2.25E+09	10.38536
	WKF	1.08E+11	1.07E+11	1.01927	6.49E+09	5.85E+09	9.85118
f_t	CF	1.90E+12	1.90E+12	0.03947	7.30E+09	6.61E+09	9.42265
	ITF	1.80E+12	1.76E+12	2.42958	1.26E+11	1.20E+11	5.06331
	RDF	2.45E+12	2.49E+12	1.35093	9.33E+10	8.82E+10	5.41681
	WKF	2.28E+12	2.31E+12	1.21728	3.72E+11	3.46E+11	6.82509
	CF	1.93E+12	1.94E+12	0.33471	5.50E+11	5.31E+11	3.40687

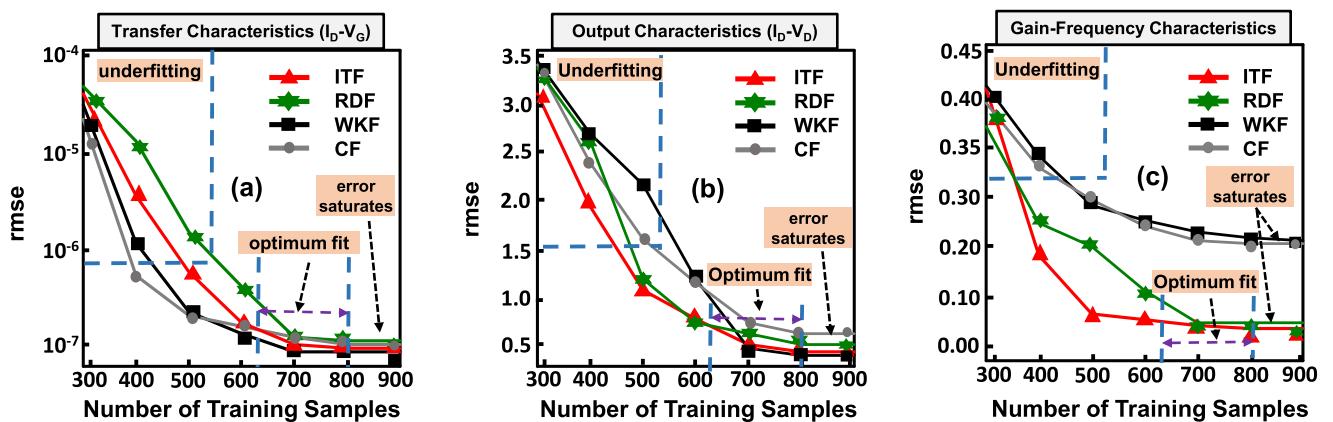


Fig. 14. Illustration of dependency of the ANN model accuracy on the dataset sizes for training. Curves show the number of variability device characteristics required for high accuracy. Plots of accuracy for (a) transfer characteristics, (b) output characteristics, and (c) frequency response characteristics are shown in terms of RMSE values. Dataset size is represented in terms of the number of samples of fluctuated devices for ITF (red-triangle straight line), RDF (green-star straight line), WKF (black-rectangle straight line), and CF (gray-circle straight line).

the model with more data, the error stays at the minimum value and saturates without any significant improvement in accuracy. Therefore, the dataset size has to be optimized to achieve the balance between ANN model accuracy and computational efficiency. The results show that the ANN models learned the trends accurately for different fluctuation sources with good predictive accuracy and perform effectively on training and test datasets.

V. CONCLUSION

In this article, we evaluated the ANN-based modeling of advanced NSFET for dc and high-frequency applications. The characteristics of NSFET are comprehensively analyzed, and the important FoMs are examined under the influence of various fluctuation sources. An exhaustive simulation process is run to generate the database of 16 000 fluctuated devices for training and testing the ANN model. The results show that when fluctuations are considered independently, the WKF produces very high variations on dc and analog/RF characteristics of nanodevices among others and affects the device performance severely. The ITF and RDF have very less effects as compared to WKF. Even when the composite effects of all the fluctuations are analyzed simultaneously, we see that the WKF still has more variations than their combined effect. This happened because the fluctuations are random in nature and

some effects are mutually canceled out each other. The ANN effectively predicts all these variations with a good R^2 -score and small RMSE.

Moreover, the prediction capability of the ANN model is assessed by predicting the important analog/RF FoMs including g_m , g_m/I_D , C_G , f_t , and f_{3dB} , and compared with the simulated values. The results show that the ANN model still performs exceptionally well despite the variable nature of the data, fluctuation sources, and the values of the parameters. The ANN model predicted that the values of these FoMs have shown a good agreement with the simulated values. In addition to this, the analysis of analog/RF applications requires the use of simulation, which is a very time-consuming process and requires a huge computational power. Therefore, the ML-based ANN model provides a computationally inexpensive alternative and shows an enormous potential to provide a deeper understanding of future nanodevices.

REFERENCES

- [1] D. Hisamoto *et al.*, “FinFET—A self-aligned double-gate MOSFET scalable to 20 nm,” *IEEE Trans. Electron Devices*, vol. 47, no. 12, pp. 2320–2325, Dec. 2000, doi: [10.1109/16.887014](https://doi.org/10.1109/16.887014).
- [2] J. Kedzierski *et al.*, “Extension and source/drain design for high-performance FinFET devices,” *IEEE Trans. Electron Devices*, vol. 50, no. 4, pp. 952–958, Apr. 2003, doi: [10.1109/TED.2003.811412](https://doi.org/10.1109/TED.2003.811412).
- [3] Y. M. Lee *et al.*, “Accurate performance evaluation for the horizontal nanosheet standard-cell design space beyond 7nm technology,” in *IEDM Tech. Dig.*, Dec. 2017, p. 29, doi: [10.1109/IEDM.2017.8268474](https://doi.org/10.1109/IEDM.2017.8268474).

- [4] I. Kwon, M. Je, K. Lee, and H. Shin, "A simple and analytical parameter-extraction method of a microwave MOSFET," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 6, pp. 1503–1509, Jun. 2002, doi: [10.1109/TMTT.2002.1006411](https://doi.org/10.1109/TMTT.2002.1006411).
- [5] C. Enz, "An MOS transistor model for RF IC design valid in all regions of operation," *IEEE Trans. Microw. Theory Techn.*, vol. 50, no. 1, pp. 342–359, Jan. 2002, doi: [10.1109/22.981286](https://doi.org/10.1109/22.981286).
- [6] S. R. Kola, Y. Li, and N. Thoti, "Characteristics of gate-all-around silicon nanowire and nanosheet MOSFETs with various spacers," in *Proc. Int. Conf. Simulation Semiconductor Processes Devices (SISPAD)*, Sep. 2020, pp. 1–4, doi: [10.23919/SISPAD49475.2020.9241603](https://doi.org/10.23919/SISPAD49475.2020.9241603).
- [7] D. Nagy, G. Espineira, G. Indalecio, A. J. García-Loureiro, K. Kalna, and N. Seoane, "Benchmarking of FinFET, nanosheet, and nanowire FET architectures for future technology nodes," *IEEE Access*, vol. 8, pp. 53196–53202, 2020, doi: [10.1109/ACCESS.2020.2980925](https://doi.org/10.1109/ACCESS.2020.2980925).
- [8] N. Loubet *et al.*, "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](https://doi.org/10.23919/VLSIT.2017.7998183).
- [9] W. Lu, P. Xie, and C. M. Lieber, "Nanowire transistor performance limits and applications," *IEEE Trans. Electron Devices*, vol. 55, no. 11, pp. 2859–2876, Nov. 2008, doi: [10.1109/TED.2008.2005158](https://doi.org/10.1109/TED.2008.2005158).
- [10] D. Jang *et al.*, "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](https://doi.org/10.1109/TED.2017.2695455).
- [11] J.-S. Yoon and R.-H. Baek, "Device design guideline of 5-nm-node FinFETs and nanosheet FETs for analog/RF applications," *IEEE Access*, vol. 8, pp. 189395–189403, 2020, doi: [10.1109/ACCESS.2020.3031870](https://doi.org/10.1109/ACCESS.2020.3031870).
- [12] S. N. Ong, K. S. Yeo, K. W. J. Chew, and L. H. K. Chan, "Substrate-induced noise model and parameter extraction for high-frequency noise modeling of sub-micron MOSFETs," *IEEE Trans. Microw. Theory Techn.*, vol. 62, no. 9, pp. 1973–1985, Sep. 2014, doi: [10.1109/TMTT.2014.2340375](https://doi.org/10.1109/TMTT.2014.2340375).
- [13] P. Kushwaha, S. Khandelwal, J. P. Duarte, C. Hu, and Y. S. Chauhan, "RF modeling of FD-SOI transistors using industry standard BSIM-IMG model," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 6, pp. 1745–1751, Jun. 2016, doi: [10.1109/TMTT.2016.2557327](https://doi.org/10.1109/TMTT.2016.2557327).
- [14] S. R. Kola, Y. Li, and N. Thoti, "Random telegraph noise in gate-all-around silicon nanowire MOSFETs induced by a single charge trap or random interface traps," *J. Comput. Electron.*, vol. 19, no. 1, pp. 253–262, Jan. 2020, doi: [10.1007/s10825-019-01438-9](https://doi.org/10.1007/s10825-019-01438-9).
- [15] Y. Sun *et al.*, "Investigation of process variation in vertically stacked gate-all-around nanowire transistor and SRAM circuit," *Semicond. Sci. Technol.*, vol. 36, no. 5, Jan. 2021, Art. no. 055009, doi: [10.1088/1361-6641/abe01b](https://doi.org/10.1088/1361-6641/abe01b).
- [16] D. Nagy, G. Indalecio, A. J. García-Loureiro, M. A. Elmessary, K. Kalna, and N. Seoane, "Metal grain granularity study on a gate-all-around nanowire FET," *IEEE Trans. Electron Devices*, vol. 64, no. 12, pp. 5263–5269, Dec. 2017, doi: [10.1109/TED.2017.2764544](https://doi.org/10.1109/TED.2017.2764544).
- [17] Y. Li and C. H. Hwang, "High-frequency characteristic fluctuations of nano-MOSFET circuit induced by random dopants," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 12, pp. 2726–2733, Dec. 2008, doi: [10.1109/TMTT.2008.2007077](https://doi.org/10.1109/TMTT.2008.2007077).
- [18] S. R. Kola *et al.*, "Effects of a dual spacer on electrical characteristics and random telegraph noise of gate-all-around silicon nanowire p-type metal-oxide-semiconductor field-effect transistors," *Jpn. J. Appl. Phys.*, vol. 59, pp. 1–5, Apr. 2020, doi: [10.7567/1347-4065/ab5b7c](https://doi.org/10.7567/1347-4065/ab5b7c).
- [19] 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 *IEDM Tech. Dig.*, Dec. 2015, p. 34, doi: [10.1109/IEDM.2015.7409827](https://doi.org/10.1109/IEDM.2015.7409827).
- [20] W.-L. Sung and Y. Li, "DC/AC/RF characteristic fluctuations induced by various random discrete dopants of gate-all-around silicon nanowire n-MOSFETs," *IEEE Trans. Electron Devices*, vol. 65, no. 6, pp. 2638–2646, Jun. 2018, doi: [10.1109/TED.2018.2822484](https://doi.org/10.1109/TED.2018.2822484).
- [21] P. H. Vardhan, 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](https://doi.org/10.1109/TED.2019.2933061).
- [22] J.-S. Yoon, S. Lee, J. Lee, J. Jeong, H. Yun, and R.-H. Baek, "Reduction of process variations for sub-5-nm node fin and nanosheet FETs using novel process scheme," *IEEE Trans. Electron Devices*, vol. 67, no. 7, pp. 2732–2737, Jul. 2020, doi: [10.1109/TED.2020.2995340](https://doi.org/10.1109/TED.2020.2995340).
- [23] C. K. Jha, P. Yogi, C. Gupta, A. Gupta, R. A. Vega, and A. Dixit, "Comparison of LER induced mismatch in NWFET and NSFET for 5-nm CMOS," *IEEE J. Electron Devices Soc.*, vol. 8, pp. 1184–1192, Sep. 2020, doi: [10.1109/JEDS.2020.3026534](https://doi.org/10.1109/JEDS.2020.3026534).
- [24] 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)*, Apr. 2019, pp. 1–2, doi: [10.1109/VLSI-TSA.2019.8804637](https://doi.org/10.1109/VLSI-TSA.2019.8804637).
- [25] Y. Sun *et al.*, "Impact of process fluctuations on RF small-signal parameter of gate-all-around nanosheet transistor beyond 3 nm node," *IEEE Trans. Electron Devices*, vol. 69, no. 1, pp. 31–38, Jan. 2022, doi: [10.1109/TED.2021.3130009](https://doi.org/10.1109/TED.2021.3130009).
- [26] C. K. Jha, C. Gupta, A. Gupta, R. A. Vega, and A. Dixit, "Impact of LER on mismatch in nanosheet transistors for 5nm-CMOS," in *Proc. 4th IEEE Electron Devices Technol. Manuf. Conf. (EDTM)*, Apr. 2020, pp. 1–4, doi: [10.1109/EDTM47692.2020.9117816](https://doi.org/10.1109/EDTM47692.2020.9117816).
- [27] N. Thoti, Y. Li, and W.-L. Sung, "Significance of work function fluctuations in SiGe/Si hetero-nanosheet tunnel-FET at sub-3 nm nodes," *IEEE Trans. Electron Devices*, vol. 69, no. 1, pp. 434–438, Jan. 2022, doi: [10.1109/TED.2021.3130497](https://doi.org/10.1109/TED.2021.3130497).
- [28] N. Seoane, J. G. Fernandez, K. Kalna, E. Comesana, and A. Garcia-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](https://doi.org/10.1109/LED.2021.3109586).
- [29] J.-S. Yoon *et al.*, "DC performance variations by grain boundary in source/drain epitaxy of sub-3-nm nanosheet field-effect transistors," *IEEE Access*, vol. 10, pp. 22032–22037, 2022, doi: [10.1109/ACCESS.2022.3154049](https://doi.org/10.1109/ACCESS.2022.3154049).
- [30] H.-C. Choi, H. Yun, J.-S. Yoon, and R.-H. Baek, "Neural approach for modeling and optimizing Si-MOSFET manufacturing," *IEEE Access*, vol. 8, pp. 159351–159370, 2020, doi: [10.1109/ACCESS.2020.3019933](https://doi.org/10.1109/ACCESS.2020.3019933).
- [31] J. Gao and A. Werthof, "Scalable small-signal and noise modeling for deep-submicrometer MOSFETs," *IEEE Trans. Microw. Theory Techn.*, vol. 57, no. 4, pp. 737–744, Apr. 2009, doi: [10.1109/TMTT.2009.2015075](https://doi.org/10.1109/TMTT.2009.2015075).
- [32] T. Wu and J. Guo, "Speed up quantum transport device simulation on ferroelectric tunnel junction with machine learning methods," *IEEE Trans. Electron Devices*, vol. 67, no. 11, pp. 5229–5235, Nov. 2020, doi: [10.1109/TED.2020.3025982](https://doi.org/10.1109/TED.2020.3025982).
- [33] C.-W. Teo, K. L. Low, V. Narang, and A. V.-Y. Thean, "TCAD-enabled machine learning defect prediction to accelerate advanced semiconductor device failure analysis," in *Proc. Int. Conf. Simulation Semiconductor Processes Devices (SISPAD)*, Sep. 2019, pp. 1–4, doi: [10.1109/SISPAD.2019.8870440](https://doi.org/10.1109/SISPAD.2019.8870440).
- [34] S.-C. Han, J. Choi, and S.-M. Hong, "Acceleration of semiconductor device simulation with approximate solutions predicted by trained neural networks," *IEEE Trans. Electron Devices*, vol. 68, no. 11, pp. 5483–5489, Nov. 2021, doi: [10.1109/TED.2021.3075192](https://doi.org/10.1109/TED.2021.3075192).
- [35] 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 Lett.*, vol. 41, no. 10, pp. 1548–1551, Oct. 2020, doi: [10.1109/LED.2020.3016119](https://doi.org/10.1109/LED.2020.3016119).
- [36] 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 Trans. Electron Devices*, vol. 68, no. 11, pp. 5490–5497, Nov. 2021, doi: [10.1109/TED.2021.3084910](https://doi.org/10.1109/TED.2021.3084910).
- [37] H. Carrillo-Núñez, 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 Lett.*, vol. 40, no. 9, pp. 1366–1369, Sep. 2019, doi: [10.1109/LED.2019.2931839](https://doi.org/10.1109/LED.2019.2931839).
- [38] J. Lim, J. Lee, and C. Shin, "Probabilistic artificial neural network for line-edge-roughness-Induced random variation in FinFET," *IEEE Access*, vol. 9, pp. 86581–86589, 2021, doi: [10.1109/ACCESS.2021.3088461](https://doi.org/10.1109/ACCESS.2021.3088461).
- [39] R. Ghoshhajra, K. Biswas, and A. Sarkar, "A review on machine learning approaches for predicting the effect of device parameters on performance of nanoscale MOSFETs," in *Proc. Devices Integr. Circuit (DevIC)*, May 2021, pp. 489–493, doi: [10.1109/DevIC50843.2021.9455840](https://doi.org/10.1109/DevIC50843.2021.9455840).
- [40] H. Y. Wong *et al.*, "TCAD-machine learning framework for device variation and operating temperature analysis with experimental demonstration," *IEEE J. Electron Devices Soc.*, vol. 8, pp. 992–1000, 2020, doi: [10.1109/JEDS.2020.3024669](https://doi.org/10.1109/JEDS.2020.3024669).
- [41] C. Akbar, Y. Li, and W.-L. Sung, "Deep learning algorithms for the work function fluctuation of random nanosized metal grains on gate-all-around silicon nanowire MOSFETs," *IEEE Access*, vol. 9, pp. 73467–73481, 2021, doi: [10.1109/ACCESS.2021.3079981](https://doi.org/10.1109/ACCESS.2021.3079981).

- [42] 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 Trans. Semicond. Manuf.*, vol. 34, no. 4, pp. 513–520, Nov. 2021, doi: [10.1109/TSM.2021.3116250](https://doi.org/10.1109/TSM.2021.3116250).
- [43] *Sentaurus Device User Guide, Version O-2018.06*, Synopsys, Mountain View, CA, USA, 2018.
- [44] M. G. Ancona, "Density-gradient theory: A macroscopic approach to quantum confinement and tunneling in semiconductor devices," *J. Comput. Electron.*, vol. 10, nos. 1–2, pp. 65–97, Jun. 2011, doi: [10.1007/s10825-011-0356-9](https://doi.org/10.1007/s10825-011-0356-9).
- [45] A. R. Brown *et al.*, "Use of density gradient quantum corrections in the simulation of statistical variability in MOSFETs," *J. Comput. Electron.*, vol. 9, no. 3, pp. 187–196, Dec. 2010, doi: [10.1007/s10825-010-0314-y](https://doi.org/10.1007/s10825-010-0314-y).
- [46] Y. Li, S.-M. Yu, J.-R. Hwang, and F.-L. Yang, "Discrete dopant fluctuations in 20-nm/15-nm-gate planar CMOS," *IEEE Trans. Electron Devices*, vol. 55, no. 6, pp. 1449–1455, Jun. 2008, doi: [10.1109/TED.2008.921991](https://doi.org/10.1109/TED.2008.921991).
- [47] Y. Li, C. H. Hwang, and T. Y. Li, "Random-dopant-induced variability in nano-CMOS devices and digital circuits," *IEEE Trans. Electron Devices*, vol. 56, no. 8, pp. 1588–1597, Aug. 2009, doi: [10.1109/TED.2009.2022692](https://doi.org/10.1109/TED.2009.2022692).
- [48] W.-L. Sung, Y.-S. Yang, and Y. Li, "Work-function fluctuation of gate-all-around silicon nanowire n-MOSFETs: A unified comparison between cuboid and Voronoi methods," *IEEE J. Electron Devices Soc.*, vol. 9, pp. 151–159, 2021, doi: [10.1109/JEDS.2020.3046608](https://doi.org/10.1109/JEDS.2020.3046608).
- [49] Y. Li, "Random nanosized metal grains and interface-trap fluctuations in emerging CMOS technologies," *Comprehensive Nanosci. Nanotechnol.*, vol. 5, pp. 123–134, Jun. 2008, doi: [10.1016/B978-0-12-803581-8.000633-0](https://doi.org/10.1016/B978-0-12-803581-8.000633-0).
- [50] Y. Li, H.-W. Cheng, Y.-Y. Chiu, C.-Y. Yiu, and H.-W. Su, "A unified 3D device simulation of random dopant, interface trap and work function fluctuations on high- κ /metal gate device," in *IEDM Tech. Dig.*, Dec. 2011, pp. 5.5.1–5.5.4, doi: [10.1109/IEDM.2011.6131495](https://doi.org/10.1109/IEDM.2011.6131495).
- [51] A. Gulli *et al.*, "Neural networks foundations," in *Deep Learning With Keras*, 1st ed. Birmingham, U.K.: Packt, 2017, pp. 9–43. [Online]. Available: <https://keras.io/>
- [52] Y. S. Bankapalli and H. Y. Wong, "TCAD augmented machine learning for semiconductor device failure troubleshooting and reverse engineering," in *Proc. Int. Conf. Simul. Semiconductor Processes Devices (SISPAD)*, Sep. 2019, pp. 1–4, doi: [10.1109/SISPAD.2019.8870467](https://doi.org/10.1109/SISPAD.2019.8870467).
- [53] F. Pedregosa, S. Varoquaux, A. Gramfort, V. Michel, and B. Thirion, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, Dec. 2011.
- [54] K. Mehta and H.-Y. Wong, "Prediction of FinFET current-voltage and capacitance-voltage curves using machine learning with autoencoder," *IEEE Electron Device Lett.*, vol. 42, no. 2, pp. 136–139, Feb. 2021, doi: [10.1109/LED.2020.3045064](https://doi.org/10.1109/LED.2020.3045064).



Rajat Butola received the M.S. degree in electronics and communication engineering from Jaypee University, Noida, Uttar Pradesh, India, in 2012. He is currently pursuing the Ph.D. degree at the Parallel and Scientific Computing Laboratory, Electrical Engineering and Computer Science International Graduate Program (EECS IGP), National Yang Ming Chiao Tung University, Hsinchu, Taiwan.

His research focuses on the applications of machine learning and deep learning algorithms in modeling of emerging semiconductor devices.



Yiming Li (Member, IEEE) has been a Visiting Professor with Stanford University, Stanford, CA, USA; the Grenoble Institute of Technology (Grenoble INP), Grenoble, France; and Tohoku University, Sendai, Japan, since 2011. He is currently a Full Professor of electrical and computer engineering with the National Yang Ming Chiao Tung University, Hsinchu, Taiwan. He is the author or coauthor of more than 400 technical articles, including the book *Physics of Semiconductor Devices* (Wiley, 2021). His current research interests include computational electronics, device physics, semiconductor nanostructures, modeling and parameter extraction, biomedical and energy harvesting devices, and optimization techniques.

Prof. Li has been serving on Technical Program Committees of IEDM since 2011 and numerous IEEE international conferences.



Sekhar Reddy Kola (Graduate Student Member, IEEE) received the B.S. degree in mathematics, electronics, and computer science and the M.S. degree in electronics science from Sri Venkateswara University, Tirupati, Andhra Pradesh, India, in 2010 and 2012, respectively. He is currently pursuing the Ph.D. degree at the Parallel and Scientific Computing Laboratory, Electrical Engineering and Computer Science International Graduate Program (EECS IGP), National Yang Ming Chiao Tung University, Hsinchu, Taiwan. His research focuses on device simulation of FinFETs, GAA nanowire, and nanosheet FETs.

Mr. Kola received the Best Paper Award from IEDMS 2019.



Chieh-Yang Chen received the B.S. degree from the Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 2011, and the M.S. degree from the Institute of Communications Engineering, National Chiao Tung University, in 2013. He is currently pursuing the Ph.D. degree at the Parallel and Scientific Computing Laboratory, National Yang Ming Chiao Tung University, Hsinchu.

His research interests include the first-principle simulation of novel 2-D materials and TCAD device simulation of HKMG MOSFETs, FinFETs, and nanowire devices.



Min-Hui Chuang (Student Member, IEEE) received the B.S. degree from the Department of Electrical and Computer Engineering, National Chiao Tung University, Hsinchu, Taiwan, in 2019, and the M.S. degree from the Institute of Communications Engineering, National Chiao Tung University, in 2020. She is currently pursuing the Ph.D. degree at the Parallel and Scientific Computing Laboratory, National Yang Ming Chiao Tung University, Hsinchu.

Her research interests include numerical semiconductor device simulation, compact model parameter extraction, and machine learning of device modeling.

Chuang's master thesis received the first prize of Lam Research Award in 2021.