A comparative study of empirical model with machine learning based model for GaN HEMTs

1st Harsh Kumar
Roll No. 23095039
Department of Electronics Engineering
IIT (BHU), Varanasi.
email: harsh.kumar.ece23@itbhu.ac.in

2nd Mohd Sharique
Roll No. 23095060
Department of Electronics Engineering
IIT (BHU), Varanasi.
email: mohd.sharique.ece23@itbhu.ac.in

3rd Rajat Bansal
Roll No. 23095080
Department of Electronics Engineering
IIT (BHU), Varanasi.
email: rajat.bansal.ece23@itbhu.ac.in

4th Shivaji Rao
Roll No. 23095141
Department of Electronics Engineering
IIT (BHU), Varanasi.
email: shivaji.rao.cd.ece23@itbhu.ac.in

***Abstract*—This project implements a neural network approach for accurate prediction of MOSFET drain current characteristics based on two key voltage relationships: I_D-V_D and I_D-V_G. Two distinct approaches are explored for each characteristic: 1) splitting the datasets on both drain and gate voltages, and 2) splitting based on single-voltage parameters. The neural network architecture comprises three hidden layers with 256, 512, and 256 neurons respectively. Experimental results demonstrate excellent agreement between measured and predicted I-V characteristics with Mean Absolute Errors (MAE) of 0.0020 and 0.0018 for I_D-V_D and I_D-V_G models respectively on training sets. Both models successfully capture the complex relationships in their respective domains: the I_D-V_D model accurately represents linear and saturation regions across various gate voltages, while the I_D-V_G model precisely captures subthreshold, linear, and saturation regions across different drain voltages. This comprehensive modeling approach demonstrates the versatility of neural networks for semiconductor device modeling applications.**

***Index Terms*—MOSFET modeling, neural networks, machine learning, device characteristics, I-V prediction, semiconductor modeling, I_D-V_D, I_D-V_G**

I. Introduction

Metal-Oxide-Semiconductor Field-Effect Transistors (MOSFETs) are fundamental building blocks in modern integrated circuits. Accurate modeling of their electrical characteristics is essential for circuit design and analysis. Traditional physics-based models often involve complex equations and parameters that are challenging to extract. This project explores a data-driven approach using artificial neural networks (ANNs) to predict MOSFET drain current (I_D) characteristics based on two key relationships: 1) Drain current versus drain voltage (I_D-V_D) at different gate voltages, and 2) drain current versus gate voltage (I_D-V_G) at different drain voltages.

The advantages of neural network-based approaches include:

- Ability to capture complex non-linear relationships without explicit physics equations
- Potential for higher accuracy compared to traditional compact models
- Adaptability to different device technologies through retraining

This exploratory project aims to develop and evaluate neural network models that can accurately predict MOSFET drain current characteristics across various operating conditions for both I_D-V_D and I_D-V_G relationships, providing a comprehensive modeling approach for MOSFET behavior.

II. Empirical Model Analysis

A. Parameter Count in Empirical Model

The empirical model for the GaN HEMT presented in the report uses a significant number of independent parameters. Analyzing the equations from the model formulation, we can identify the following independent parameters:

- $V_{\text{pk}},\ V_{\text{pk0}},\ \alpha_R,\ W$
- $I_{\text{pkt}},\ M,\ \lambda,\ Q_m$
- $V_{\text{gsm}},\ M_0,\ I_{\text{pk0}},\ P_{10}$
- $P_1,\ P_{20i},\ P_2,\ P_{30i}$
- $P_3,\ \alpha_S$

These 18 independent parameters are used in the formulation of the compact model for the GaN HEMT device. Despite this considerable number of fitting parameters, the compact model was unable to provide a good fit to the experimental data, primarily due to the short channel effects present in the device. The device under study has a gate length of approximately 300 nm, which introduces significant short channel effects that are challenging to capture accurately with traditional compact models.

This limitation of the empirical approach, even with numerous fitting parameters, motivated the exploration of machine learning techniques, particularly neural networks, as an alternative modeling methodology.

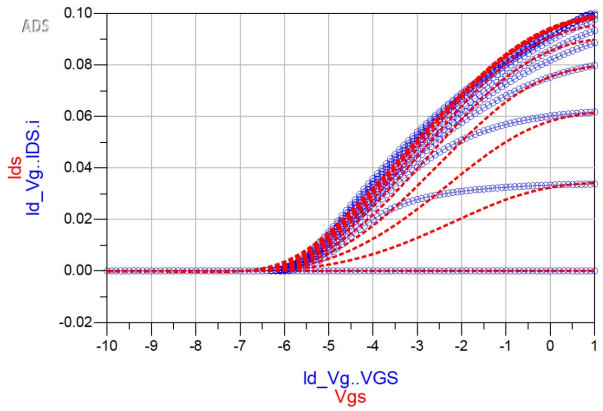

Fig. 1. Drain Current vs. Gate Voltage (I_D-V_G) characteristics for the Empirical model. Blue circles represent experimental data while red dashed lines show model predictions at different drain voltages.

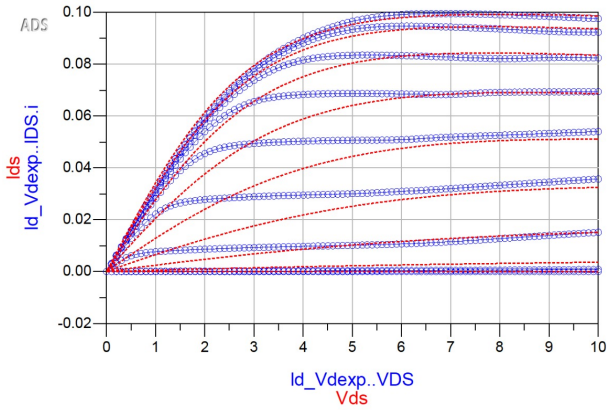

Fig. 2. Drain Current vs. Drain Voltage (I_D-V_D) characteristics for the Empirical model. Blue circles represent experimental data while red dashed lines show model predictions at different gate voltages.

III. Methodology

A. Dataset Preparation

1) I_D-V_D Dataset: The I_D-V_D dataset consists of measured I_D-V_D characteristics for a MOSFET device across multiple gate voltage conditions. Raw data was loaded from the "IdVd_D1_21_2.txt" file containing drain voltage and corresponding drain current values for nine different gate voltages. The data was preprocessed and organized into a suitable format for machine learning, with drain voltage and gate voltage as input features and drain current as the target variable.

2) I_D-V_G Dataset: The I_D-V_G dataset consists of measured I_D-V_G characteristics for the same MOSFET device across multiple drain voltage conditions. Raw data was loaded from the "IdVg_D1_21_2.txt" file containing gate voltage and corresponding drain current values for several different drain voltages. Similar to the I_D-V_D dataset, this data was preprocessed and formatted for machine learning with gate voltage and drain voltage as input features and drain current as the target variable.

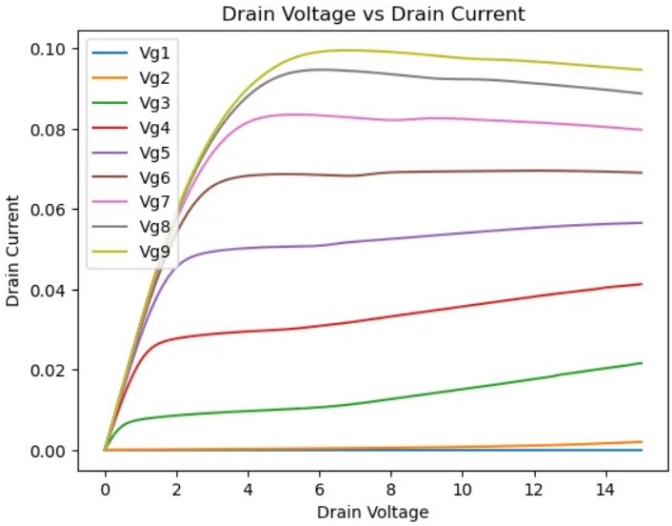

Fig. 3. Drain Current vs. Drain Voltage curves for different gate voltages in the raw I_D-V_D dataset.

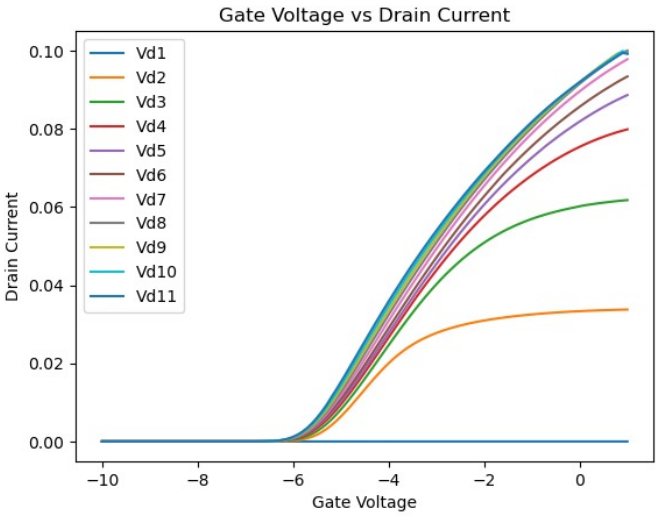

Fig. 4. Drain Current vs. Gate Voltage curves for different drain voltages in the raw I_D-V_G dataset.

B. Feature Engineering

1) Extracting gate voltage values and drain voltages from the raw data
2) Reshaping the data into a proper format with columns for drain voltage, gate voltage, and drain current
3) Handling missing values by dropping any problematic data points
4) Scaling features and target using MinMaxScaler to normalize the data

C. Neural Network Architecture

A consistent neural network architecture was implemented for both I_D-V_D and I_D-V_G models using TensorFlow/Keras with the following structure:

- Input layer: 2 neurons (drain voltage and gate voltage)

- First hidden layer: 256 neurons with ReLU activation
- Second hidden layer: 512 neurons with ReLU activation
- Third hidden layer: 256 neurons with ReLU activation
- Output layer: 1 neuron (drain current prediction)

Both models were compiled with:

- Optimizer: Adam with a learning rate of 0.001
- Loss function: Mean Squared Error (MSE)
- Metrics: Mean Absolute Error (MAE)

D. Training Strategy

For both I_D-V_D and I_D-V_G models, the datasets were split into training (70%), validation (15%), and test (15%) sets using the scikit-learn train_test_split function. To prevent overfitting, an early stopping callback was implemented with a patience of 50 epochs, monitoring validation loss. The models were trained for up to 50 epochs with a batch size of 32.

E. Evaluation Metrics

Model performance for both I_D-V_D and I_D-V_G was evaluated using Mean Absolute Error (MAE), which measures the average absolute difference between predicted and actual drain current values.

IV. Results and Discussion

A. I_D-V_D Model Results

1) Training Performance: The I_D-V_D model converged rapidly within the first 10 epochs, as shown in the training history plots (Fig. 5).

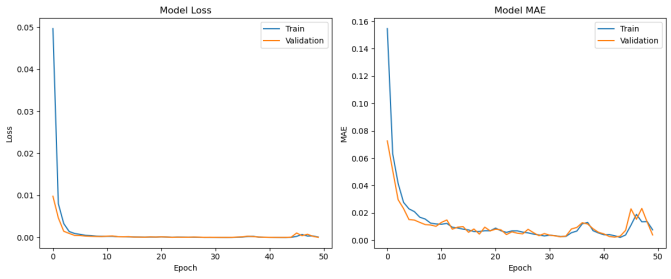

Fig. 5. I_D-V_D model training history showing loss and MAE over epochs for training and validation sets.

B. I_D-V_G Model Results

1) Training Performance: Similar to the I_D-V_D model, the I_D-V_G model showed efficient convergence during training, as depicted in Fig. 6.

C. Error Analysis

1) I_D-V_D Error Analysis: Analysis of prediction errors in the I_D-V_D model shows that it performs consistently well across the operating range, with slightly higher relative errors in regions of very low current.

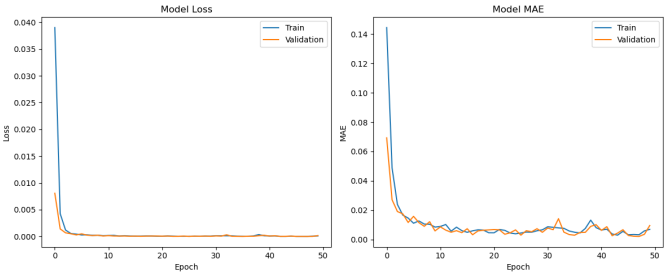

Fig. 6. I_D-V_G model training history showing loss and MAE over epochs for training and validation sets.

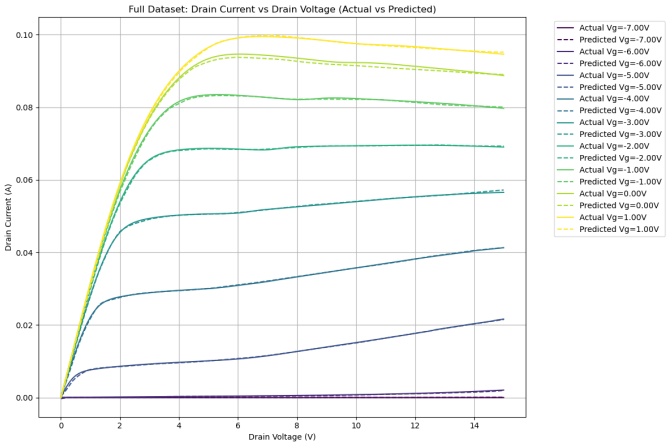

Fig. 7. Comparison of actual vs. predicted drain current characteristics for the full I_D-V_D dataset.

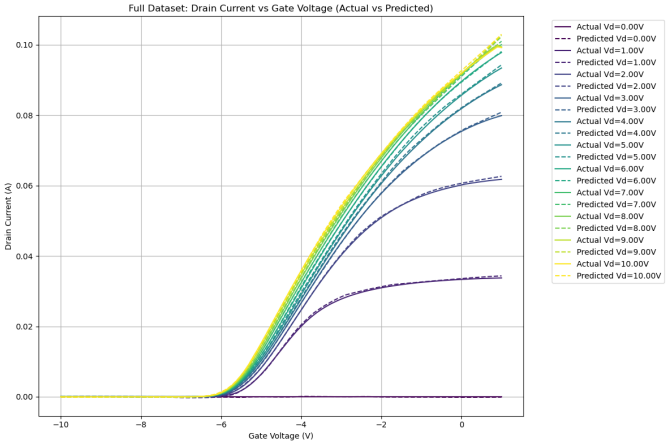

Fig. 8. Comparison of actual vs. predicted drain current characteristics for the full I_D-V_G dataset.

2) I_D-V_G Error Analysis: For the I_D-V_G model, error analysis reveals excellent performance across most of the operating range with particularly good accuracy in the subthreshold region. Minor deviations were observed in the transition region from subthreshold to strong inversion, which is expected due to the rapid change in current in this region.

D. Comparison Between I_D-V_D and I_D-V_G Models

Both models demonstrated excellent predictive capabilities for their respective voltage-current relationships. The I_D-V_D model excelled at capturing the transition between linear and saturation regions, while the I_D-V_G model performed particularly well in modeling the exponential relationship in the subthreshold region and the power-law relationship in the strong inversion region.

TABLE I
PERFORMANCE COMPARISON OF I_D-V_D AND I_D-V_G MODELS

Dataset	I_D-V_D **MAE**	I_D-V_G **MAE**
Training	0.0020	0.0018
Test	0.0019	0.0017
Validation	0.0023	0.0022

E. Advantages of Neural Network Approach Over Empirical Model

The neural network approach demonstrates superior performance compared to the empirical model, particularly for the 300 nm gate length device where short channel effects are significant. While the empirical model required 18 independent parameters and still failed to provide an accurate fit, the neural network models achieved excellent agreement with experimental data using a consistent architecture. This highlights the ability of neural networks to capture complex physical phenomena without explicit physics-based equations, making them particularly valuable for modeling devices with complex behavior that traditional compact models struggle to represent accurately.

V. CONCLUSION

This exploratory project demonstrates the effectiveness of neural networks for modeling MOSFET drain current characteristics from both the I_D-V_D and I_D-V_G perspectives. The neural network approach has successfully captured the complex non-linear relationships in both cases without relying on explicit physics-based equations. The models demonstrate excellent agreement with measured data across various operating regions, including linear, saturation, and subthreshold operations.

The complementary nature of the two models provides a comprehensive understanding of MOSFET behavior across its entire operating range. The I_D-V_D model excels at representing the transition between linear and saturation regions at fixed gate voltages, while the I_D-V_G model accurately captures the exponential subthreshold characteristics and power-law behaviors in the strong inversion region.

For the studied GaN HEMT device with a gate length of 300 nm, the machine learning approach proved significantly more effective than the empirical model which, despite using 18 independent parameters, could not adequately capture the device behavior due to short channel effects.

Future work could explore:

- Extension to other device types and technologies.
- Incorporation of temperature effects.
- Neural architecture optimization.
- Development of a unified model for both I_D-V_D and I_D-V_G characteristics.
- Integration into circuit simulation frameworks.

CERTIFICATE

This paper presents original work conducted by the authors as their **Exploratory Project** from **Jan to June 2025** at the Department of Electronics Engineering, IIT (BHU), Varanasi, under the supervision of **Prof. Jaya Jha.**

REFERENCES

[1] S. Emekar et al., "Modified Angelov model for an exploratory GaN-HEMT technology with short, few-fingered gates," Proc. SISPAD, pp. 117-120, 2017.

[2] P. Chandra and J. Jha, "Rapid LUT Modelling Technique for GaN HEMT Based MMIC Technology," Proc. IEEE Asia-Pacific Microwave Conf. (APMC), 2021, doi: 10.1109/APMC47863.2021.9661735.

[3] F. Chollet et al., "Keras," https://keras.io, 2015.

[4] M. Abadi et al., "TensorFlow," https://tensorflow.org, 2015.

[5] F. Pedregosa et al., "Scikit-learn," JMLR, 2011.

[6] J. D. Hunter, "Matplotlib," Computing in Science & Engineering, 2007.