

Received October 6, 2020, accepted October 20, 2020, date of publication October 26, 2020, date of current version November 6, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.3033788

A Generic and Efficient Globalized Kernel Mapping-Based Small-Signal Behavioral Modeling for GaN HEMT

AHMAD KHUSRO¹, SADDAM HUSAIN², MOHAMMAD S. HASHMI², (Senior Member, IEEE), ABDUL QUAIYUM ANSARI¹, (Senior Member, IEEE), AND SULTANGALI ARZYKULOV², (Member, IEEE)

¹Department of Electrical Engineering, Jamia Millia Islamia (A Central University), Delhi 110025, India

²Department of Electrical and Computer Engineering, School of Engineering and Digital Sciences, Nazarbayev University, Nur-Sultan 010000, Kazakhstan

Corresponding author: Mohammad S. Hashmi (mohammad.hashmi@nu.edu.kz)

This work was supported in part by FDRG at Nazarbayev University, Kazakhstan, under Grant 110119FD4515, and in part by the Social Policy Grant, Nazarbayev University.

ABSTRACT The work reported in this article explores a novel Particle Swarm Optimization (PSO) tuned Support Vector Regression (SVR) based technique to develop the small-signal behavioral model for GaN High Electron Mobility Transistor (HEMT). The proposed technique investigates issues such as kernel selection and model optimization usually encountered in the application of SVR to model the GaN based HEMT devices. Here, the PSO algorithm is utilized to find the optimal hyperparameters to minimize the fitness function. To enumerate the efficiency and the generalization capability of the predictors, the performance of the model is investigated in terms of mean square error (MSE) and mean relative error (MRE). A very good agreement is found between the measured S-parameters and the proposed model for multi-biasing sets over the complete frequency range of 1 GHz-18 GHz. The proposed technique is even used to test the frequency extrapolation capability of the model. A comparative analysis indicates that the proposed PSO-SVR predictor achieves significantly improved computational efficiency and the overall prediction accuracy. To demonstrate the ready usefulness of the modeling approach, the developed model has been incorporated in CAD environment using MATLAB Cosimulation in ADS Ptolemy. Subsequently, the small-signal stability analysis is performed and gain of a power amplifier configuration designed using the proposed GaN HEMT model is determined.

INDEX TERMS GaN HEMT, hyperparameters, Kernel function, modeling, particle swarm optimization, s-parameters, support vector regression.

I. INTRODUCTION

Gallium Nitride (GaN) HEMT is getting very popular in the design of circuits and components such as RF Power Amplifier (PA) owing to its unique traits like wide energy bandgap, high breakdown field, and high electron mobility [1]–[7]. However, its use in RFPAs necessitates the development of specification-based large-signal models. It has been established that an accurate large-signal model of GaN HEMT utilizes a bottom-up modeling technique where small-signal models are extremely important [8]–[12]. A number of GaN

HEMT small-signal modeling techniques exist and many of these make use of analytical formulations, cold pinch-off concept, and de-embedding [11]–[21]. These conventional methods, although accurate, are often found to be highly cumbersome and computationally inefficient. Therefore, the alternative machine learning (ML) based small-signal modeling technique is gaining popularity as their turn-around time is fast with very good accuracy [22]–[23]. A key feature of ML is its ability to predict the outcome in real-time very quickly and this is very appealing for device modeling especially at RF and microwave frequencies where the inter-dependence of various device parameters on each other is huge [24], [25]. Furthermore, device modeling by integrating various device

The associate editor coordinating the review of this manuscript and approving it for publication was Yin Zhang¹.

geometry and learning device behavior with more features to build an automation model using numerous ML algorithms has the potential to bring a paradigm shift in the way device modelling is carried out. It is due to the fact that the ML based modelling discards the need to solve simultaneous complex equations involved in conventional methods, possesses reduced computation time, exhibits better accuracy and prediction ability, and enhanced production yield [26].

More recently, Artificial Neural Network (ANN) and Support Vector Regression (SVR) techniques have been used in the development of small-signal model for GaN HEMT [27]–[34]. However, SVR is considered better than ANN as it enables the determination of global optimum which is superior to ANN that supports local optimum [23], [25], [32]–[37]. Moreover, the SVR possesses geometrical interpretation whereas ANN is based on the tedious description of various parameters. In addition, SVR also exhibits robustness to noise, and excellent generalization capability (providing solution to underfitting and overfitting) and is therefore preferred over ANN in the device modeling regime. At this instance, it is important to note that the SVM based regression model also suffers from multiple issues in SVR [32]–[37]. Some of the issues can be attributed to selection of inappropriate kernel function and random set of hyperparameters which leads to poor prediction ability of the model. Apparently, the selection of kernel function depends on dataset and an uncertainty in dataset makes it difficult to determine the specific kernel function which is most suitable for model development. On the other hand, set of hyperparameters if not selected properly may lead to underfitting or overfitting. In addition, SVR is not appropriate while dealing with massive dataset due to large processing Gram matrices associated with kernels [25].

Keeping the above issues in perspective, therefore, a novel technique making use of Particle Swarm Optimization (PSO) and SVR is developed to model behavior of GaN HEMT under small-signal conditions. First, a kernel function library is constructed so that an optimized kernel function could be selected according to the problem requirement. Subsequently, PSO is utilized for superior optimization considering that the performance of SVR depends on its kernel function and hyperparameters. The PSO is used here due to the fact that it possesses superior features such as simple search rules, easy implementation, less parameters to adjust, no effect of initial conditions on its computational behavior and fast convergence in comparison to other techniques such as the Genetic Algorithm (GA) and the resampling technique. In addition, the PSO ensures convergence to the global optimum instead of local optimum observed in GA and other local search techniques. Furthermore, the PSO is a computational population based algorithm which evaluates the fitness function at each particle (collection of individuals) as particles move in steps throughout the region. Eventually, the algorithm evaluates the new position and velocity of the particle and iterative procedure continues as the particle moves further and algorithm re-evaluates. Overall, the main contributions of this article

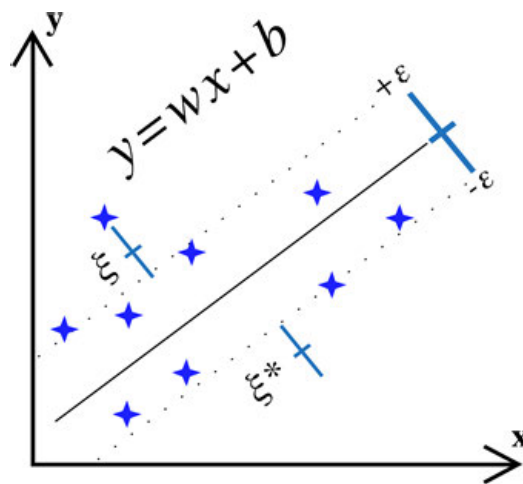


FIGURE 1. Basic Principle of Support Vector Regression (SVR).

are: (a) development of a novel PSO-SVR based small signal behavioral model for GaN HEMT device, (b) validation of the prediction ability of the model by frequency and geometric extrapolation, (c) robustness evaluation of the proposed model by mixing random noise, and (d) the integration of the proposed transistor model in CAD environment for circuit simulation and analysis as a part to demonstrate the usefulness of the model.

The next section elaborates the basic SVR method and PSO-SVR model whereas development of the framework for GaN HEMT under small signal is explained in Section III. The experimental results, model validation, and comparison with the state-of-the-art are presented in Sections IV and V respectively while Section V concludes the paper.

II. SUPPORT VECTOR REGRESSION & PARTICLE SWARM OPTIMIZATION ALGORITHM

A. SUPPORT VECTOR REGRESSION (SVR)

Support vector machine (SVM) was invented in 1963 by Vapnik and Chervonenkis (VC) and developed over three decades. Eventually, B. Boser, I. Guyon and V. Vapnik proposed an approach to create non-linear classifiers by application of kernel trick to maximize the margin of hyperplane in 1992. It is a unique class of algorithm characterised by application of kernels, sparseness of the solution, presence of global minima, and capacity controlled by altering margin. SVR, a regression technique, works on the principle of Support Vector Machine (SVM) but with few minor differences. Being a regression problem, it uses the curve to find the match between vector and position of the curve instead of decision boundary in SVM. The SVM is one of the non-parametric techniques based on computation of linear function in a high dimensional feature space. Here, the inputs are mapped through a nonlinear function called kernel. For example, if x_i is a multi-variable set of predictors of N observations with response values y_i for a given set of training sample then the idea behind SVR is to find a function $f(x)$ that maximally deviates by epsilon, ϵ , from the observed targets y_i for

complete set of observations and ensure to be as flat as possible as illustrated in Fig. 1 [38], [39]. To comprehend this aspect, let us consider a linear function $f(x)$ given as

$$f(x) = w^T \psi(x) + b, \tag{1}$$

where $\psi(x)$ is a mapping function that maps each input vector x in the input space $X = \Psi(x)$ in the high dimensional feature space with bias vector w , T is the transpose and $b \in \mathcal{R}$ is a bias value. This study focuses on ε -SVR model. This method uses a concept of ε -insensitive loss function where the training data has at most ε deviation from targets y_i [38]–[42]. Loss function calculates the distance between the observed value y and the ε boundary by treating those error equal to zero that lie within ε distance of the observed value.

To address the training samples that lie beyond the ε -insensitive zone, slack variables ξ_i and ξ_i^* at each point can be introduced as the soft margin classification. Therefore, the primal formula is restated in (2) and the applicable constraints in (3). Then the dual formula can be framed by constructing Lagrangian function from the given set of primal function and equivalent constraints by introducing positive multipliers. The terms α_m , α_m^* , η_m and η_m^* in (4) are non-negative Lagrangian multipliers [39], [40]. Then the partial derivatives of (4) with respect to primal variables (w , b , ξ_i , ξ_i^*) to zero for optimality and substitutions of those values back in (4) yields dual formula expressed in (5). It leads to expression (6) and which needs to satisfy the complementarily Karush-Kuhn Tucker (KKT) conditions expressed by equations 7(a)–(c) [39]–[41].

$$\text{Minimize } \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*), \tag{2}$$

$$y_i - ((w, x_i) + b) \leq \varepsilon + \xi_i : \xi_i \geq 0 \forall i, \tag{3}$$

$$((w, x_i) + b) - y_i \leq \varepsilon + \xi_i : \xi_i \geq 0 \forall i. \tag{3}$$

$$L = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^N (\xi_i + \xi_i^*) - \sum_{m=1}^N (\eta_m \xi_m + \eta_m^* \xi_m^*) + \sum_{m=1}^N \alpha_m (f(x_m) - y_m - \varepsilon - \xi_m) + \sum_{m=1}^N \alpha_m^* (y_m - f(x_m) - \varepsilon - \xi_m^*), \tag{4}$$

$$\text{Maximize } \sum_{i=1}^N y_i (\alpha_i^* - \alpha_i) - \varepsilon (\alpha_i^* + \alpha_i) - \frac{1}{2} \sum_{i,j=1}^N (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) K(x_i, x_j). \tag{5}$$

$$\sum_{i=1}^N y_i (\alpha_i - \alpha_i^*) = 0, \text{ where } \alpha_i, \alpha_i^* \in [0, C]. \tag{6}$$

$$\alpha_i (\varepsilon + \xi_i - y_i + f(x_i)) = 0; \xi_i (C - \alpha_i) = 0 \tag{7a}$$

$$\alpha_i^* (\varepsilon + \xi_i^* - y_i + f(x_i)) = 0; \xi_i^* (C - \alpha_i^*) = 0; \forall i \tag{7b}$$

$$\alpha_i \alpha_i^* = 0; \eta_i \eta_i^* = 0. \tag{7c}$$

TABLE 1. List of Kernel Functions.

Type	Particle's Range
Linear	$K(x, x_i) = (x \cdot x_i)$
Polynomial	$K(x, x_i) = (x \cdot x_i + d)^q; d \in \mathcal{R}$
Gaussian or RBF	$K(x, x_i) = \exp(-\gamma \ x - x_i\ ^2)$

Now, the best estimated function $f(x)$ or the optimal solution of the convex optimization can be derived by solving (5) and (6) using the famous Sequential Minimal Optimization (SMO) algorithm expressed by (7). Here, $K(\cdot)$ denotes the kernel function, x_j is the support vector, n is the number of support vectors, the term in bracket represents weight coefficient of support vectors [41], [42], and b is the bias.

$$f(x) = \sum_{j=1}^n (\alpha_j^* - \alpha_j) K(x, x_j) + b, \tag{7}$$

Different kernel functions implicitly map the predictors to a high dimensional feature space and capture the non-linearity of the model or the system. The well-known widely used kernel functions with their range are included in Table 1. The parameters such as q and γ are kernel parameters. The term q , an integer quantity, determines the degree or order of a polynomial kernel whereas $\gamma = \frac{1}{2\sigma^2}$, plays the role of scaling parameter for Gaussian or Radial Basis Function (RBF) kernel.

B. PARTICLE SWARM OPTIMIZATION ALGORITHM

The PSO is an evolutionary computational algorithm based on population [15], [43]–[45]. A collection of individuals known as particles move throughout the region. The optimization process begins with randomly initialized population of solutions which are called particles. The swarm consists of n particles assigning initial velocities to each particle. It then, assesses the fitness function at each particle's location, and determines the best value and the position. Subsequently, it selects modified velocities, based on current velocity, best position of particle's individual and its neighbors. The particle's and its neighbors' locations and the particle's velocities is updated iteratively until the algorithm reaches a stopping criterion.

Let the swarm consists of n particles, with each particle position defined by a vector $x_i = (x_{i,1}, x_{i,2}, \dots, x_{i,d})$, each particle velocity defined by a vector $v_i = (v_{i,1}, v_{i,2}, \dots, v_{i,d})$ where $i = (1, 2, \dots, n)$, and d the number of dimensions of each vector. It is pertinent to mention that the PSO has two operators namely the position operator and the velocity operator. During each generation, the best position of the particle is denoted by the vector $P_{\text{best},i} = (p_{i,1}, p_{i,2}, \dots, p_{i,d})$ and global best position of the population is represented by the

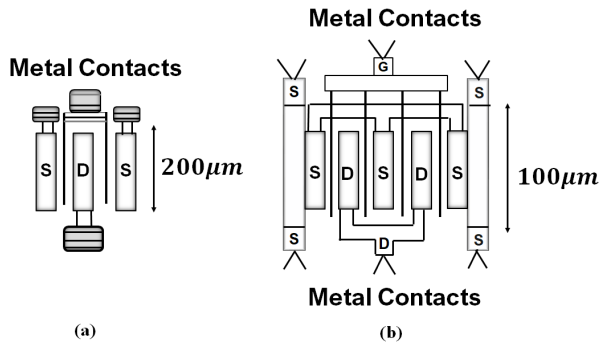


FIGURE 2. Layout of (a) $2 \times 200 \mu\text{m}$ and (b) $4 \times 100 \mu\text{m}$ GaN HEMT device.

vector $P_{\text{best},g} = (p_{g,1}, p_{g,2}, \dots, x_{g,d})$. The updated position and velocity of the particle can be respectively obtained using the following expressions

$$v_{i,d}^{\text{new}} = wv_{i,d}^{\text{old}} + c_1r_1(P_{\text{best},i,d} - x_{i,d}) + c_2r_2(P_{\text{best},g,d} - x_{i,d}), \quad (8)$$

$$x_i^{\text{new}} = x_i^{\text{old}} + v_i^{\text{new}}, \quad (9)$$

where w represents inertia weight, c_1 and c_2 are acceleration constants whereas r_1 and r_2 are the randomly generated values in the range $[0, 1]$. The inertia weight affects the current velocity whose value decreases linearly with successive generations. Higher values of inertia weight improve global search while lower values facilitate local search. The linear decreasing inertia weight (LDIW) proposed by Y. Shi is represented by (10) is used for the proposed approach.

$$w = w_{\min} - (w_{\max} - w_{\min}) \frac{\text{current iteration}}{\text{max. iteration}}. \quad (10)$$

III. DEVELOPMENT OF PSO-SVR BASED GaN HEMT BEHAVIORAL MODELLING TECHNIQUE

A. DEVICE LAYOUT AND CHARACTERIZATION

The fabricated AlGaN/GaN HEMT device, not shown here due to manufacturer's policy, of geometries $2 \times 200 \mu\text{m}$ and $4 \times 100 \mu\text{m}$ is shown in Fig. 2.

These devices, grown on SiC substrate, are fabricated by Defense Research Development Organization (DRDO, an enterprise of ministry of Defense, Govt. of India). The characterization set-up includes Microwave Network Analyzer, probe station and DC power supplies. The biasing to the device are provided at the gate and the drain terminals, with the source terminal being grounded, through DC power supplies using bias Tees internal to the network analyzer. The characterization is done under the small-signal conditions for $-7 \text{ V} \leq V_{\text{GS}} (\text{V}) \leq 0 \text{ V}$, $0 \text{ V} \leq V_{\text{DS}} (\text{V}) \leq 10 \text{ V}$ for a frequency range of 1-18 GHz. The S-parameters are recorded in terms of magnitude and phase of two-port S-parameters for the provided multi-biasing and frequency range. The RF-characteristics of any transistor can be studied in terms of unity current gain frequency (f_t) and maximum oscillation frequency (f_{max}). The f_t is extracted by converting S-parameters to current gain h-parameter h_{21} and putting it

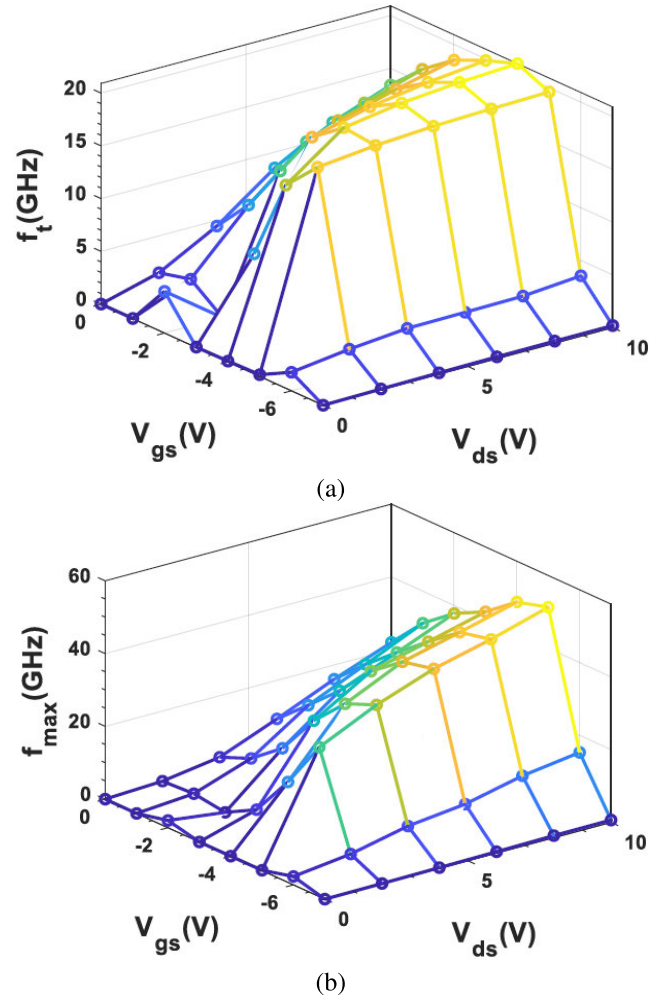


FIGURE 3. RF Characteristics (a) f_t , and (b) f_{max} of $4 \times 100 \mu\text{m}$ GaN HEMT device.

equal to unity whereas f_{max} is extracted by putting the maximum unilateral gain equal to unity. As seen from the Fig. 3, the maximum value of f_t and f_{max} obtained for class A and class AB operated $4 \times 100 \mu\text{m}$ GaN HEMT is approximately 21 GHz and 52 GHz respectively using the expression defined in [46].

The I_D - V_{DS} characteristics of GaN HEMT device characterized under DC operation is shown in Fig. 4. DC I-V characteristics is used to classify the regions of the transistor as ohmic, saturation and cold pinch-off. Under DC operation, the gate bias is stepped gradually up from pinch providing enough time for trapped electrons to response unlike to the pulsed measurement where the device is first pinched-off and electrons are injected into traps. When channel is turned on by a short pulse under high electric field, the trapped electrons are unable to response in time. Therefore, the current is reduced, knee voltage is increased and the predicted output power reduces.

B. PREPROCESSING

Modeling of small-signal behavior of GaN HEMT requires mapping of five predictor variables into high dimensional

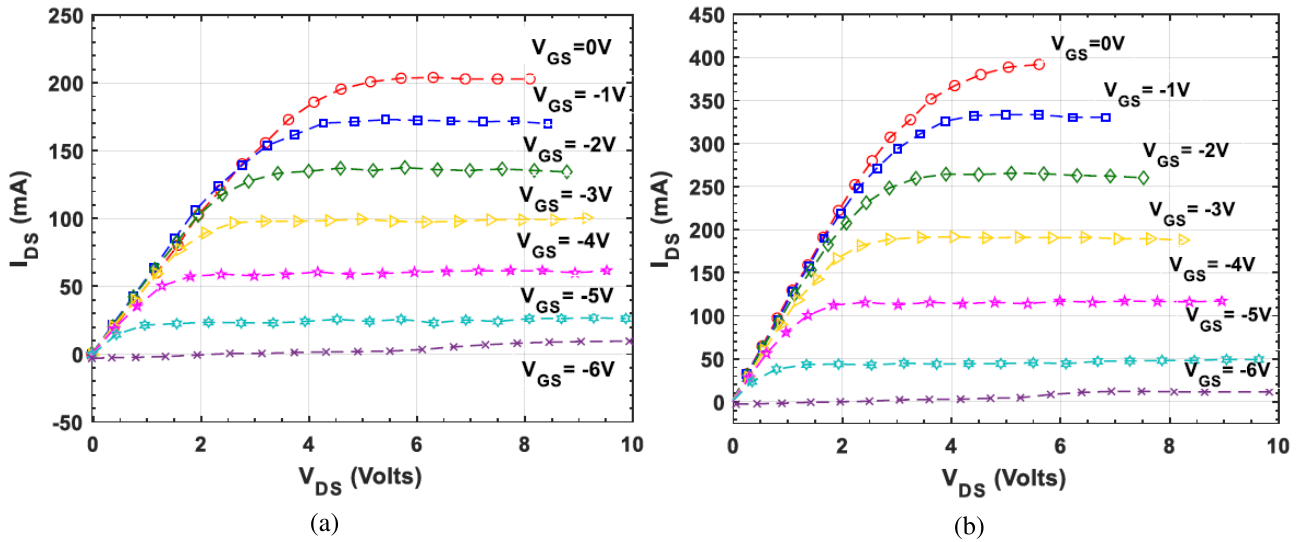


FIGURE 4. DC I-V Characteristics of GaN HEMT for (a) $2 \times 200 \mu\text{m}$ and (b) $4 \times 100 \mu\text{m}$.

feature space for the eight response variables. The five predictor variables are gate to source voltage V_{GS} (V) varying in the range $[-7\text{ V}, 0\text{ V}]$ with a step size of 1 V, drain to source voltage V_{DS} (V) varying in the range $[0\text{ V}, 10\text{ V}]$ with a step size of 2 V, frequency in the range $[1, 18]$ (GHz) with a step size of 0.1 GHz, number of fingers N ($N = 2$ and $N = 4$, respectively) and unity gate width wg (μm) ($200 \mu\text{m}$ and $100 \mu\text{m}$, respectively). On the other hand, eight desired responses are two port S-parameters, S_{11} (reflection coefficient measured at port 1), S_{22} (reflection coefficient measured at port 2), S_{12} (transmission coefficient measured at port 1) and S_{21} (transmission coefficient measured at port 2) recorded for different sets of inputs in the form of magnitude and phase.

It is imperative to note that multiple transitions are observed in the measured phase values of S_{11} and S_{22} , due to few multi-bias condition usually at low V_{GS} and high V_{DS} are shown in Fig. 5, considering that the phase values of S-parameters are recorded by the VNA in the range of $[-\pi, +\pi]$. However, no such anomaly is observed in the magnitude of S-parameters as these are recorded on a linear scale. This situation necessitates pre-processing of data. It is well-known that data preprocessing is an important unit of machine learning based model as shown in Fig. 6, and it is employed for cleaning and smoothing of the data before training the model. This stage deals with the issues like presence of inconsistency, outliers and random noise present in the dataset. Scaling the data is also one of the steps of the pre-processing required as per the requirement of the regression problems.

In this article, the magnitude of two port S-parameters has been scaled logarithmically to ensure the data-set must be a uniform distribution of data within small range which therefore will expedite the search of global optimum and fast convergence. It is well observed that phase of S_{11} and S_{22}

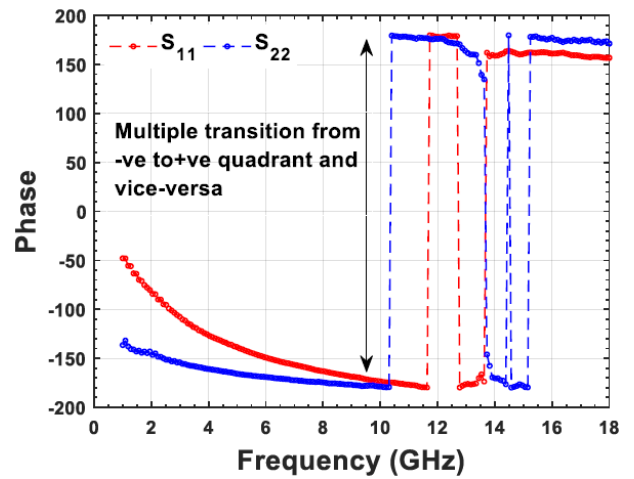


FIGURE 5. Observed Multiple Transition in Phase of Measured S_{11} and S_{22} .

is highly discontinuous curve past 11 GHz. Trigonometric adjustment for positive phase angles ($\theta_{\text{new}} = \theta - 360^\circ$) is made to restore the data consistency. Subsequently, dataset is moved to outlier detection unit. An outlier is an unwanted part that behaves differently with another portion of dataset. Outliers play a major role and have the capability to develop poor model. Margin will shrink and sub-optimal decision boundary may result in poor prediction capability of the model or issues like underfitting and overfitting may arise. This may affect regularization constant C , as the outlier affected model may generate more errors, converting soft-margin SVM problem to hard-margin SVMs.

In this work, the condition for detection of an outlier, expressed in (11), is based on statistical (stochastic) assumption that data is normally distributed and therefore idea is to learn a model fitting the provided dataset and therefore objects detected in error region of the model are termed as

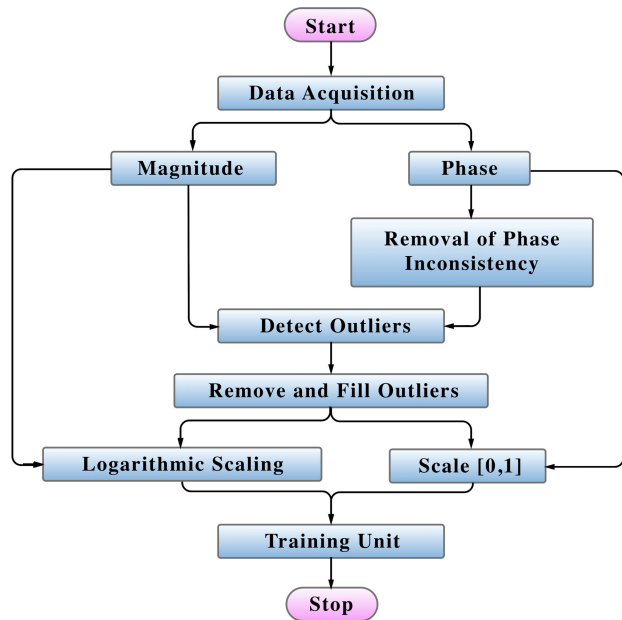


FIGURE 6. Pre-processing Unit for the proposed technique.

an outlier.

$$y_{element} > \frac{-3}{\left(\sqrt{2}erfc^{-1}\left(\frac{3}{2}\right)\right)} \text{median}(|y_{element} - \text{median}(y)|), \quad (11)$$

where $erfc^{-1}(x) = erfc^{-1}(1 - x)$ and known as inverse complementary error function. The error function is obtained by integrating the normalized Gaussian distribution. The normal distribution function $N(x)$ gives the probability that a variate assumes a value in the interval of $[0, x]$

$$erfc^{-1}(x) = 1 - erf(x) = \frac{2}{\sqrt{\pi}} \int_x^\infty e^{-t^2} dt. \quad (12)$$

Subsequently, the outliers detected are filled using linear interpolation method of adjacent non-outlier values. The phase variables are scaled in the range $[0.01, 1]$ using (13)

$$c = (c_{max} - c_{min}) \frac{b - b_{min}}{b_{max} - b_{min}} + c_{min}, \quad (13)$$

where c_{max} is the maximum value of the defined range, c_{min} is the minimum value of the defined range, b is the value to be scaled, b_{max} is the maximum value of the response variables to be scaled, b_{min} is the minimum value of the response variables to be scaled. Past pre-processing, the data can now be utilized to train the model.

C. TRAINING OF THE MODEL: SVR HYPER-PARAMETER OPTIMIZATION USING PSO

The flowchart for PSO-SVR based behavioral model of GaN HEMT is shown in Fig. 7. This section discusses the training stage of the model development. Two-third of the data is utilized for training of the model while the rest one-third is reserved for testing the prediction ability of the model. Two devices with distinct geometries having different number of

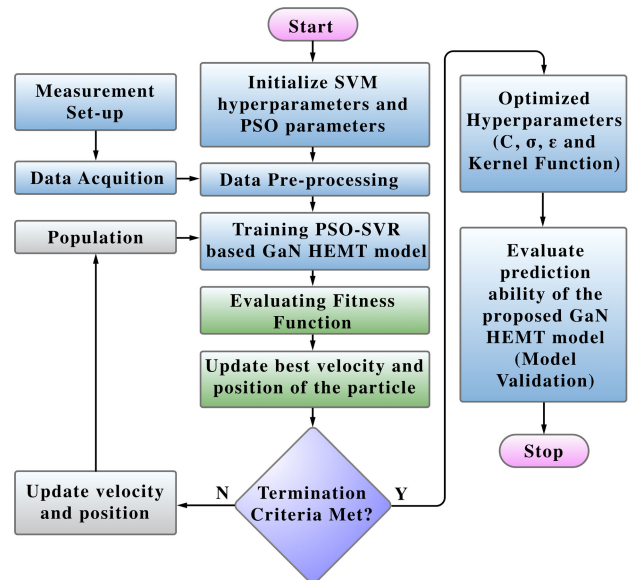


FIGURE 7. Flowchart for PSO-SVR based behavioral modeling of GaN HEMT.

fingers N and unity gate width w_g is utilized to develop the model. However, the effective gate width w_{eff} of both the devices are equal. The training of the proposed model is set up with S-parameters recorded for varying gate to source voltage V_{GS} (V) in the range $[-7V, 0V]$ with a step size of $-1V$, drain to source port voltage V_{DS} (V) in the range $[0V, 10V]$ with a step size of $2V$ and broad range sweeping frequency from $[1GHz-13GHz]$ with a step size of $0.09GHz$. The remaining $5GHz$ is reserved for testing the frequency extrapolation capability of the model.

The main idea here is to develop a small-signal behavioral model of GaN HEMT that utilizes optimal kernel function to transform the data into high-dimensional feature space that aids in linear separation of the optimal hyperplane (maximum margin). Effective kernel transformation, an architecture shown in Fig. 8, depends on couple of critical factors, also known as hyperparameters, such as penalty factor C , margin of tolerance (epsilon width) (ϵ) and kernel scaling parameter (q and γ) which eventually decides optimal hyperplane. The overall optimization problem is a combination of achieving the best fit of the model subjected to the condition that most of the samples should stay outside the margin. Therefore, regularization parameter C is a user defined input, tuned in a manner that margin neither completely overfit nor loses its large margin property. An optimization problem is modified such that it optimizes the fitness function by penalizing the number of samples inside the margin at the same time. Here, C defines the weight of samples inside the margin resulting in contribution to an overall error. Apparently, a lower value of C penalizes less sample as compared to higher values of C . The other vital parameter for defining the support vector regression is epsilon (ϵ) which is defined as a margin of tolerance where errors are not penalized. Support vectors (SVs) are instances across the margin where

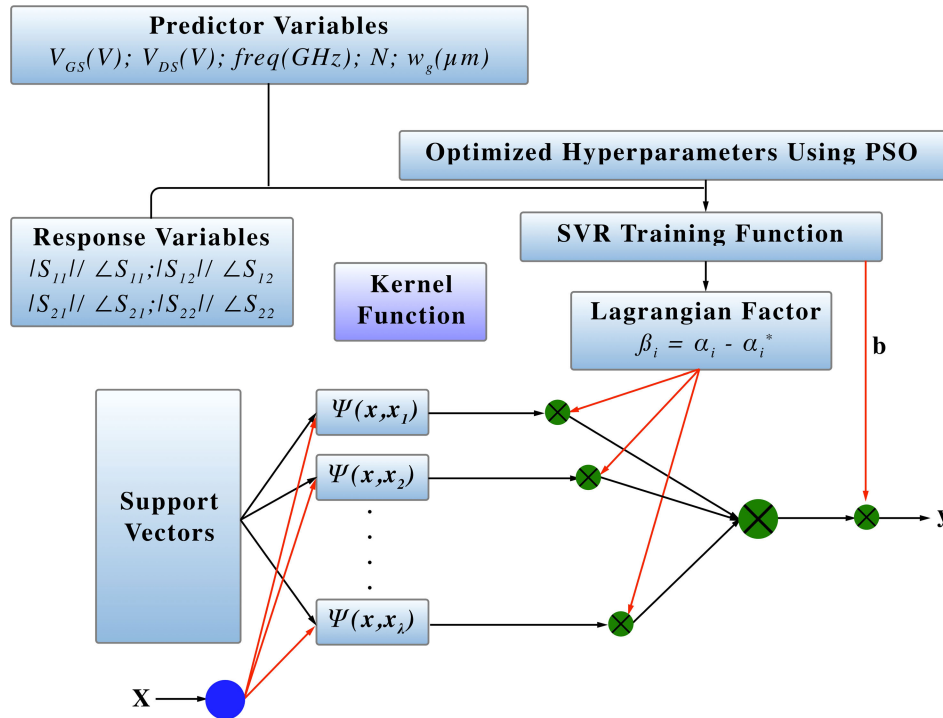


FIGURE 8. Architecture flow for PSO-SVR based GaN HEMT behavioral model.

samples are penalized for which slack variables (ξ_i and ξ_i^*) are non-zero as discussed in Section II. It is to be noted that a large value of ϵ implies more errors admitted in the solution. It implies that a greater number of SVs are required if every error is penalized. Moreover, the epsilon-insensitive loss function ensures existence of global minima and optimized reliable generalization bound [38].

Eventually, the goal is to develop a model with minimization of fitness function, an optimal hyperplane (better generalization bound) which maximizes the margin if part of error is tolerated. The Support Vector Regression (SVR) is regression problem and hence it becomes very difficult to obtain optimal set of kernel functions and hyperparameters that will achieve the desired objective stated above by employing hit and trial method which consists of infinite possibilities. In the present work, therefore, Particle Swarm Optimization (PSO) algorithm is used to obtain and select optimal kernel and associated hyperparameters. Here, optimal refers to the function and hyperparameters that yield the best fitness function for the model. Owing to attributes such as excellent interpretability, easy implementation and fast convergence, the PSO is preferred over other algorithms. Flow of the PSO-tuned based SVR algorithm, depicted in Fig. 6, is described in the subsequent steps.

Step 1: The parameters of PSO such as acceleration coefficients (c_1, c_2), and randomly generated coefficient (r_1, r_2) are initialized. Later, the hyperparameters, that needs to be optimized, are generated in feature space. Each i th particle is represented by $\{C, \sigma, \epsilon\}$ and kernel functions. The initialization parameters and particle position range are defined

in Table 2, whereas p_{best} is assigned as the initial population. A wide range of particle's range is chosen to ensure the precision in selection of optimal hyperparameters.

TABLE 2. Parameter Initialization.

PSO Parameters		Particle's Range	
Parameters	Value	Parameters	Range
Population Size	$n=50$	C, ϵ and σ	[1e-5, 1e5]
Acceleration	$c_1 = 1.8$	Kernel Function	Linear, polynomial
Coefficients	$c_2 = 2.1$		Gaussian
Inertia Weight	$\omega = 1$	Polynomial order	[2,5]

Step 2: The hyperparameters, namely C, σ , and ϵ , are set as optimized variables. For this purpose, the fitness function defined in terms of mean square error (MSE) expressed by (14) needs to be evaluated. Here, N is sample size, y_{meas} represents the measured value and y_{pred} is the predicted value. Each hyperparameter (i.e., particle) is a potential solution to the problem in search space. The particles share information among each other and tend to shift towards promising search region by adjusting their search direction. The hyperparameters have their own optimal experience represented as best position of particle $p_{best,i}$ in the feature space and optimal experience derived from population is represented as best position from the population as $p_{best,g}$. Apart from the performance criteria in (14), the proposed model performance is also evaluated in terms of correlation coefficient, R , given in (16). The model fits better for R closer to 1.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_{meas} - y_{pred})^2, \quad (14)$$

$$R = \sqrt{\frac{\left(N \sum_{i=1}^N y_m y_p - \sum_{i=1}^N y_p \sum_{i=1}^N y_m\right)^2}{\left(N \sum_{i=1}^m y_p^2 - \left(\sum_{i=1}^N y_p\right)^2\right) \left(N \sum_{i=1}^N y_m^2 - \left(\sum_{i=1}^N y_p\right)^2\right)}, \quad (15)$$

where $R \in [0, 1]$.

Step 3: At every optimization step, each hyperparameter evaluates its own fitness as well as fitness of its neighbouring hyperparameters. During this process, each set of hyperparameters does accelerate towards $p_{best,i}$ and $p_{best,g}$. Subsequently, the algorithm evaluates and searches the minimum value of MSE. If the value of MSE for given set of hyperparameters i for the current iteration surpasses that of $p_{best,i}$, then $p_{bestnew,i}$ replaces $p_{best,i}$. The MSE is evaluated correspondingly for each set of hyperparameters and the best MSE value of for a set of hyperparameters is chosen among all set of hyperparameters are considered as global best. Essentially, $p_{best,g}$ replaces $p_{best,i}$, if the MSE value of $p_{best,i}$ is greater than that of $p_{best,g}$. The hyperparameters are obtained for which the minimum value of fitness function is achieved.

Step 4: The velocity and the position of a particle are updated according to (9) and (10) respectively.

Step 5: The optimization process is carried out in sequential manner as mentioned above until the maximum number of iterations is reached or the objective function attains constant value for about next 30 iterations. In this work, maximum number of iterations is set to 1000.

Subsequently, the proposed GaN HEMT model is developed using a distinct algorithm called Sequential Minimal Optimization (SMO) [25] for optimal set of kernel function and associated hyperparameters considering various additive noise functions. The SMO algorithm has a unique feature of solving quadratic programming (QP) without any numerical optimization steps and even without extra matrix storage. Often, it consists of couple of stages in which large QP problem is decomposed into subsets of small QP problems. The optimized parameters for the proposed GaN HEMT behavioral model is listed in Table 3.

D. EXPERIMENTAL DISCUSSION

In this section, series of experiments are conducted to investigate the effect of noise on the proposed behavioral model of transistor. The proposed model is exposed to different types of noises, gaussian and non-gaussian, to identify their effect and to elaborate the robustness of the proposed approach. As a fact, it is well-known that the conventional lumped element based electrical equivalent model of any device contributes noise of different types and levels at different frequencies when integrated in CAD tool for circuits and systems design. Considering noise as a random process, the study assumes that the proposed model is corrupted by an additive noise of different nature, defined as Gaussian noise and Rayleigh symmetric (non-gaussian) noise densities, in (16) and (17) respectively.

TABLE 3. Optimized Hyperparameters Using PSO.

Response/Parameters	Kernel Functions	Convergence Time (sec)	C	Hyperparameters	
				σ	ϵ
S ₁₁	Gaussian	0.063	1.419	0.9618	7.6649e-4
∠S ₁₁	Gaussian	2.237	0.5252	2.6203	0.010336
S ₁₂	Gaussian	4.1943	3237	3.1777	7.9524e-5
∠S ₁₂	Gaussian	3.1306	0.1306	1.6322	1.5751e-4
S ₂₁	Gaussian	0.2774	0.00572	2.0539	5.0584e-3
∠S ₂₁	Gaussian	0.117	95347	89151	3.3012e-4
S ₂₂	Gaussian	1.1487	1.0158	0.9975	2.4588e-3
∠S ₂₂	Gaussian	8.5366	75.87	21.242	1.0814e-5

1) GAUSSIAN NOISE FUNCTION

The proposed model with same value of obtained optimal hyperparameters in Table 3 are further trained assuming that the proposed model is corrupted by noise term N contributing to it additively, obtained from the probability density function (PDF) known as Gaussian distribution as

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad (16)$$

where μ and σ^2 denote mean and variance. respectively.

2) RAYLEIGH DISTRIBUTION FUNCTION (NON-GAUSSIAN)

Similarly, another noise function known as Rayleigh distribution function is considered to test the effect of noise on the proposed model. A continuous distribution non-Gaussian function with shape parameter σ is defined using probability density function for $x > 0$ as

$$f_x(x) = \frac{x}{\sigma^2} e^{-\frac{x^2}{2\sigma^2}}. \quad (17)$$

The motive behind conducting this experiment is to initially analyze the performance of the proposed behavioral model of GaN HEMT in various noisy environment and its effect on prediction ability, accuracy, generalization capability of the model and various training parameters such as MSE, loss function, support vectors. In this work, the proposed model is influenced by noise. As a data based model, random Gaussian samples are generated for couple of mean and variance sets of {0, 1} and {2, 5}, whereas Rayleigh samples are generated for shape parameter (b) and variance set of 1, 1 and 2, 5. Overall, 11520 training samples and 5760 testing samples are generated to perform this experiment.

The proposed model is re-trained and now six predictor variables (including noise) is mapped to feature space with the same set of optimized hyperparameters obtained using PSO in Table 3. It is worthwhile to mention that the initial conditions of all the training are stored and this ensures that the performance of the model is enumerated and assessed under the noise environment without further training.

TABLE 4. Comparative Analysis of PSO-SVR GaN HEMT Model Under Noiseless and Gaussian Noise Function.

Response\ Parameters	Training MSE			Loss Function L(ϵ)			SVs (%)		
	PSO-SVR (SVR)	$\mu = 0;$ $\sigma = 1$	$\mu = 2;$ $\sigma = 2.23$	PSO-SVR (SVR)	$\mu = 0;$ $\sigma = 1$	$\mu = 2;$ $\sigma = 2.23$	PSO-SVR (SVR)	$\mu = 0;$ $\sigma = 1$	$\mu = 2;$ $\sigma = 2.23$
$ S_{11} $	2.697e-5	5.252e-6	1.781e-6	1.1e-3	1.8e-5	3.62e-6	34.77	38.80	43.09
$\angle S_{11}$	2.175e-4	2.015e-4	1.896e-4	2.5e-3	2.3e-3	1.9e-3	18.63	21.33	25.02
$ S_{12} $	7.757e-5	1.268e-4	4.811e-5	2.7e-3	7.2e-3	5.4e-3	38.77	40.71	42.34
$\angle S_{12}$	1.082e-5	4.772e-5	4.469e-4	1.4e-3	1.9e-3	4.4e-3	26.49	27.93	29.67
$ S_{21} $	1.788e-4	2.599e-4	5.678e-4	2.5e-3	3.4e-3	6.6e-3	30.26	32.71	35.43
$\angle S_{21}$	8.628e-5	9.573e-5	1.029e-4	1.8e-4	2.2e-4	2.8e-4	51.87	52.91	55.07
$ S_{22} $	1.053e-4	7.284e-5	2.831e-5	2.6e-3	8.9e-4	8.5e-5	23.88	27.63	28.97
$\angle S_{22}$	2.231e-4	2.693e-4	2.891e-4	3e-3	3.9e-3	5.23e-3	34.11	36.83	38.17

TABLE 5. Comparative Analysis of PSO-SVR GaN HEMT Model Under Noiseless and Rayleigh Noise Function.

Response\ Parameters	Training MSE			Loss Function L(ϵ)			SVs (%)		
	PSO-SVR (SVR)	b=1; $\sigma = 1$	b=2; $\sigma = 2.23$	PSO-SVR (SVR)	b=1; $\sigma = 1$	b=2; $\sigma = 2.23$	PSO-SVR (SVR)	b=1; $\sigma = 1$	b=2; $\sigma = 2.23$
$ S_{11} $	2.697e-5	1.202e-5	5.019e-6	1.1e-3	5.03e-4	2.43e-4	34.77	40.42	44.23
$\angle S_{11}$	2.175e-4	2.061e-4	2.039e-4	2.5e-3	2.4e-3	2.3e-3	18.63	22.27	25.58
$ S_{12} $	7.757e-5	2.258e-4	6.461e-5	2.7e-3	4.3e-3	5.4e-3	38.77	42.59	44.05
$\angle S_{12}$	1.082e-5	2.866e-5	6.454e-5	1.4e-3	1.8e-3	2.0e-3	26.49	30.43	32.26
$ S_{21} $	1.788e-4	2.243e-4	3.037e-4	2.5e-3	3.1e-3	3.8e-3	30.26	32.85	36.51
$\angle S_{21}$	8.628e-5	8.628e-5	8.628e-5	1.8e-4	1.8e-4	1.8e-4	51.87	53.14	55.48
$ S_{22} $	1.053e-4	7.906e-5	5.944e-5	2.6e-3	1.3e-3	6.7e-4	23.88	26.93	30.33
$\angle S_{22}$	2.231e-4	2.838e-4	3.041e-4	3e-3	2.73e-3	1.69e-3	34.11	38.26	39.95

The trained model is then tested under noiseless condition, and also with the incorporation of Gaussian noise and Rayleigh Distribution for different sets of mean and variance. Then the training parameters such as training error, loss function, and support vectors are enumerated and compared between the noiseless model and the proposed model under noise in Table 4 and Table 5. It can be clearly inferred that the increasing variance of noise samples increases the support vectors slightly even though one can see reduction in the training error. The increase in the support vectors indicate that the regularization parameter C is penalizing more errors and this leads to change in the generalization bound of the model due to the noise and this is obvious for a pre-trained model. The overall impact of this is the reduced generalization capability of the model which may result in few degree of data overfit. Furthermore, almost every error is penalized with less errors admitted in the solution due to low epsilon and this indicates a good generalization capability of the model under noiseless condition.

The section further calculates and discusses the prediction ability (accuracy) and generalization capability of the model under noiseless and various noise functions with different mean and variance for complete test set. A very good MSE for the test set of the order of 10^{-4} , which is in close proximity to the MSE of training set, and correlation coefficient (R) of over 99% is obtained for test samples as shown in Table 6. These results demonstrate a good fitting and generalization over the complete range for noiseless as well as noisy environment. Apparently, the prediction ability of the proposed model for

the test set decreases by around 1.5% in the presence of noise over the complete sample and this can be considered very good outcome thus confirms that the proposed model is highly immune to noise effects.

Finally, the proposed small signal modeling approach of GaN HEMT is also compared with the conventional small-signal behavioral model approach for GaN HEMT developed using Artificial Neural Network (ANN). Similar to the work in [29]–[31], the ANN based small-signal behavioral model is developed for the same set of predictors and responses discussed in Section III-B. The desired responses are scaled using the expression in (14). In this context, a feed-forward three layer Multi-Layer Perceptron (MLP) based ANN model is trained using Levenberg-Marquardt algorithm with 15 neurons, tan-sigmoid as an activation function, and initialization of weights and bias, as discussed in [31], with the same objective function defined in (15) and (16). The objective function tends to minimize the objective function for optimal set of weights and biases. Then the ANN based network is further trained and developed with additional noise inputs described in (17) - (18) using the similar optimal set of weights and biases obtained in noiseless environment. To validate the performance of ANN based model, the trained network is subjected to novel test inputs. The prediction ability of the model is enumerated for noiseless and Rayleigh noise and the obtained results are given in Table 6. It is quite evident that the MSE of phase variables are significantly increased, of the order of 10^{-3} and 98%, for the ANN based small-signal behavioral model when compared to the proposed SVR based

TABLE 6. Prediction Abilities of Various GaN HEMT Behavioral Model.

Response/ Model	PSO-SVR (Noiseless)		PSO-SVR (Gaussian) $\mu = 0; \sigma = 1$		PSO-SVR (Non-Gaussian) $b = 1; \sigma = 1$		ANN (Noiseless)		ANN $b = 1; \sigma = 1$		
	Test MSE	R (%)	Test MSE	R (%)	Test MSE	R (%)	Train MSE	R (%)	Test MSE	Test MSE	R (%)
$ S_{11} $	2.635e-4	99.71	2.935e-4	99.88	2.891e-4	99.52	1.49e-4	99.63	2.11e-4	4.52e-4	98.99
$\angle S_{11}$	1.433e-4	99.97	1.672e-4	99.61	1.923e-4	99.17	1.24e-3	99.08	4.83e-3	6.41e-3	97.62
$ S_{12} $	5.582e-4	99.47	6.117e-4	99.42	6.143e-4	99.26	3.56e-3	98.42	5.41e-3	7.45e-3	96.01
$\angle S_{12}$	2.339e-4	99.62	2.877e-4	99.41	3.402e-4	99.29	2.68e-3	99.13	6.47e-3	8.41e-3	97.91
$ S_{21} $	2.478e-4	99.68	2.611e-4	99.34	2.774e-4	99.31	6.82e-4	98.84	8.23e-4	9.87e-4	97.92
$\angle S_{21}$	6.221e-4	99.33	6.568e-4	99.26	6.721e-4	99.18	7.56e-4	98.97	7.21e-4	9.12e-4	98.58
$ S_{22} $	1.639e-4	99.81	1.795e-4	99.67	1.935e-4	99.48	8.43e-4	98.55	1.07e-3	2.97e-3	98.22
$\angle S_{22}$	1.861e-4	99.74	2.227e-4	99.59	2.302e-4	99.47	7.91e-3	98.22	9.07e-3	9.17e-3	96.08

modeling approach under similar Rayleigh noise. The MSE and R of test set for ANN model is not close enough to the training set and is also greater than the proposed SVR model which signifies that there is degradation in the prediction and generalization capability of the ANN model.

The excellent performance of the proposed model is due to the fact that the SVR works on the principle of structural risk minimization. Here, the parameters such as C and epsilon-insensitive zone keeps regular check on the generalization capability of the model and loss function complemented with the PSO ensure global optimum. In contrast, the ANN works on the principle of empirical risk minimization. In the ANN, the primary objective is to only minimize the errors and it relies on local optimum. In addition, the ANN does not have in-built provision of keeping regular check on generalization capability. This is the reason that ANN may perform well on training set but is unable to validate and generalize the performance for test set as well. Furthermore, more number of parameters are required to tune and optimize the performance of ANN based model when compared to the proposed SVR model that relies on just three parameters.

As an experiment, the performance of the ANN based model can be improved by increasing the number of hidden layers and tuning several parameters. However, the generalization capability and the performance under non Gaussian noise still remains an issue with the conventional ANN based modeling technique. Moreover, it is pertinent to note that the performance of ANN based small-signal behavioral model can be further improved by increasing the number of hidden layers and neurons within it and selecting more appropriate activation function. In that case, the device engineer needs to consider trade-off with other factors such as modelling and computational complexity and convergence rate. At the same time it is worthwhile to mention that such an issue doesn't arise in the SVR based technique considering that global optimization in this technique is achieved by tuning of either two or three hyperparameters.

IV. MODEL VALIDATION

In the previous section, framework for the development of behavioral modeling of GaN HEMT is established and the model is trained with optimal hyperparameters and

TABLE 7. Mean Relative Error (MRE) for various bias condition of the proposed model $2 \times 200 \mu\text{m}$.

Response/ Bias	$V_{GS}=-2V;$ $V_{DS}=6V$	$V_{GS}=-4V;$ $V_{DS}=2V$	$V_{GS}=-5V;$ $V_{DS}=0V$	$V_{GS}=-7V;$ $V_{DS}=4V$
$ S_{11} $	2.59	2.03	0.97	1.66
$\angle S_{11}$	0.78	1.38	0.86	2.32
$ S_{12} $	2.64	1.91	1.07	0.47
$\angle S_{12}$	1.55	0.68	1.35	1.00
$ S_{21} $	1.61	0.75	0.48	1.88
$\angle S_{21}$	0.99	0.37	0.61	0.57
$ S_{22} $	2.09	2.02	3.01	2.98
$\angle S_{22}$	3.14	1.22	2.42	0.23

kernel functions. This section investigates the reliability and effectiveness of the trained model by subjecting the model to novel set of inputs for broad frequency range from 1-18 GHz (with the frequency extrapolation) including noise and without noise utilizing one-third of measured data reserved for testing and validation purpose. To study the accuracy of the model more precisely, the prediction ability of the proposed model is enumerated in terms of mean relative error (MRE), expressed in (18), where y_{meas} is the measured value and $y_{\text{predicted}}$ is the predicted value of the response variable of the proposed model. The relative error is plotted against the whole frequency range for four different set of multi-bias and different operating region of the transistor under noiseless condition as depicted in Fig. 9 and Fig. 10 respectively. Similarly, MRE calculated for $2 \times 200 \mu\text{m}$ GaN HEMT for different multi-bias condition over complete frequency range is given in Table 7.

$$MRE(\%) = \frac{1}{N} \sum_{i=1}^N \frac{|y_{\text{meas}} - y_{\text{predicted}}|}{y_{\text{meas}}} \times 100. \quad (18)$$

The calculated MRE shows very less error within the range of training frequency (upto 13 GHz), and a rise in spike from 2-5% at the the start of extrapolation set before the error starts decreasing and curve settles down. Altogether, with the mean relative error varying between 1.5-3.5% for extrapolation set shows a very promising accuracy of the proposed model. At this moment it can be seen that very high MRE in the plot of S_{21} in Fig. 10(a) occurs. It can be attributed to the high value of C and σ in Table 3. The large value of

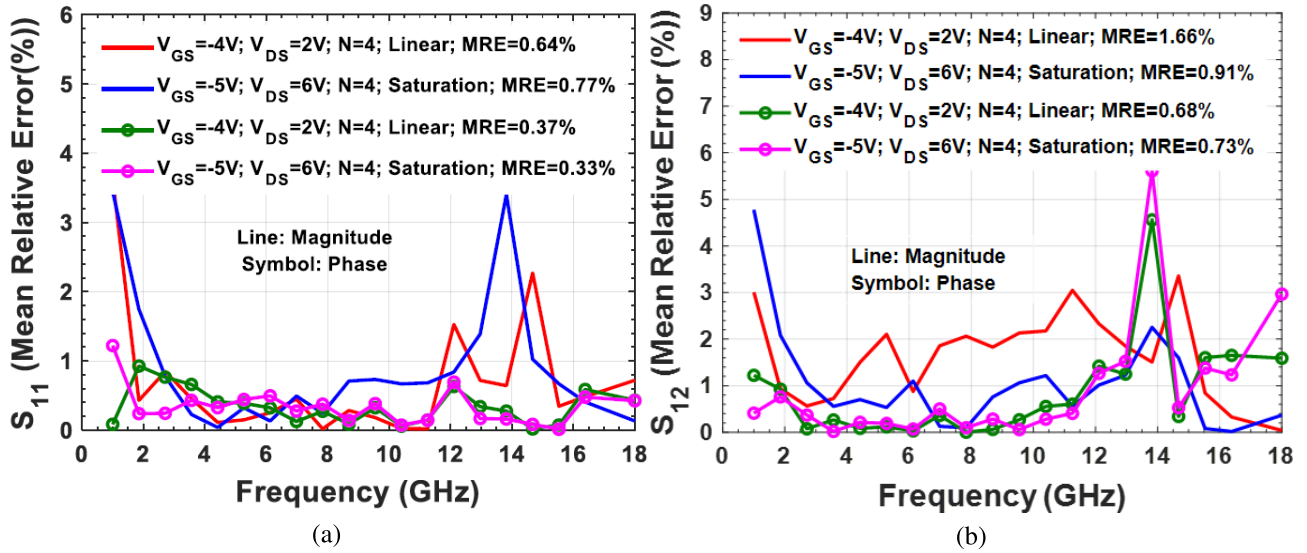


FIGURE 9. Mean Relative Error (MRE) (%) for (a) S_{11} , and (b) S_{12} .

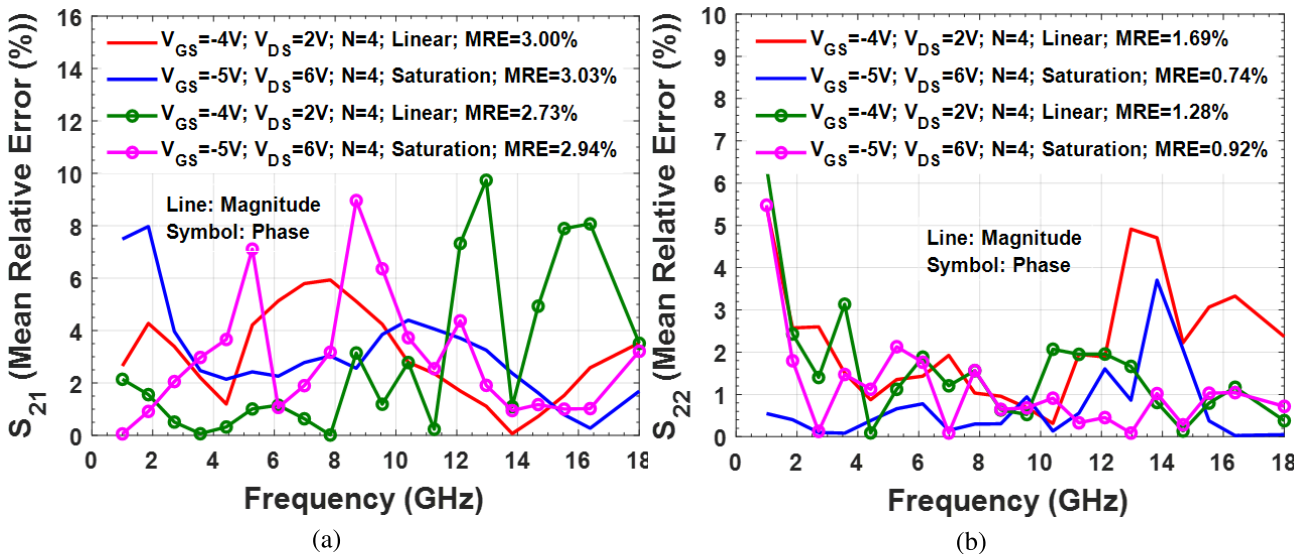


FIGURE 10. Mean Relative Error (MRE) (%) for (a) S_{21} and (b) S_{22} .

σ signifies that support vectors are widely spread and this negatively influences the flatness of the curve. On the other hand, the large value of regularization parameter C penalizes the errors too much which in turn increases the number of support vectors. As a consequence, the response variable S_{21} shows a low training error but starts losing generalization properties for certain set of test inputs. Moreover, the increase in the support vectors indicate high variance and overfitting of the response variable. For all other S-parameters, the value of C and σ indicate low variance and hence there is excellent agreement between the measured and modelled values as can be seen in the plots, Figs. 9 and 10, with very low MRE.

Furthermore, an excellent agreement is also obtained between the measured data and proposed model over broad frequency range of 1-18 GHz including extrapolation for

different regions of the transistor as depicted in Fig. 11. Finally, the proposed technique is compared with the relevant state-of-the-art in the domain of small-signal modeling technique of GaN HEMT in Table 8. It is apparent that the proposed technique demonstrate superior accuracy and hence advances the state-of-the-art significantly. Overall, it can be inferred that the accuracy of GaN HEMT is mainly dependent on the selection of optimized hyperparameters.

A. STABILITY AND GAIN TEST OF POWER AMPLIFIER (Utility OF THE PROPOSED SMALL-SIGNAL Model)

Finally, the utility of the proposed small-signal model of GaN HEMT has been demonstrated by integrating the proposed model into CAD environment for circuit simulation and analysis. As already discussed, the main application of GaN based HEMT lies in the designing of RF Power

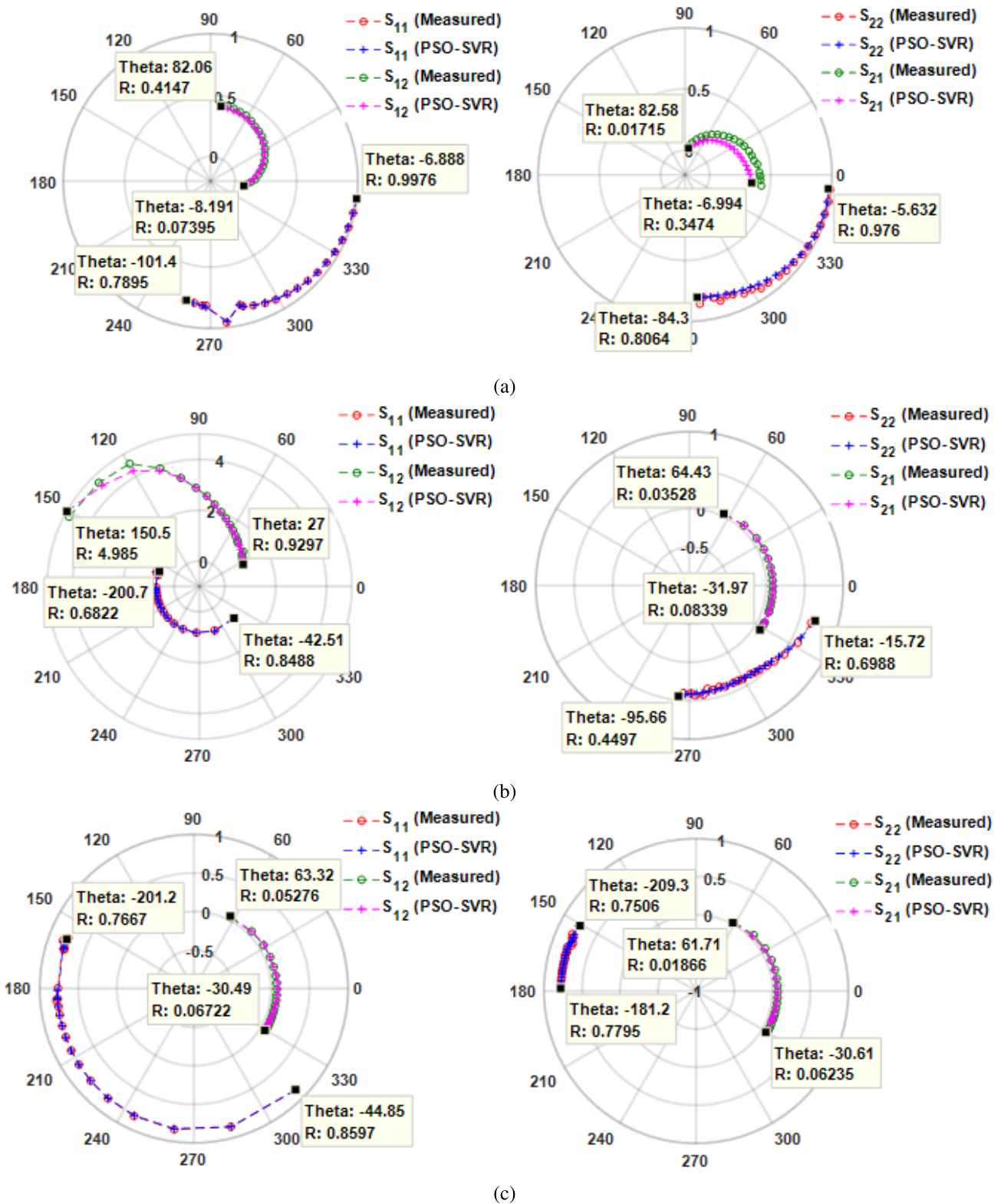


FIGURE 11. Comparison between the measured data and the proposed PSO-SVR model for 1-18 GHz under different operating regions of the transistor, (a). $V_{GS} = -7V$; $V_{DS} = 6V$, (b). $V_{GS} = -1V$; $V_{DS} = 2V$ (Ohmic), and (c). $V_{GS} = -3V$; $V_{DS} = 10V$ (Saturation) (Magnitude (R) on linear scale and phase (theta) in degrees).

Amplifiers (PAs). In this section, stability and gain test of amplifier is performed in CAD environment utilizing the

proposed model. The complete circuit for stability and gain test of amplifier is shown in Fig. 12. The proposed model

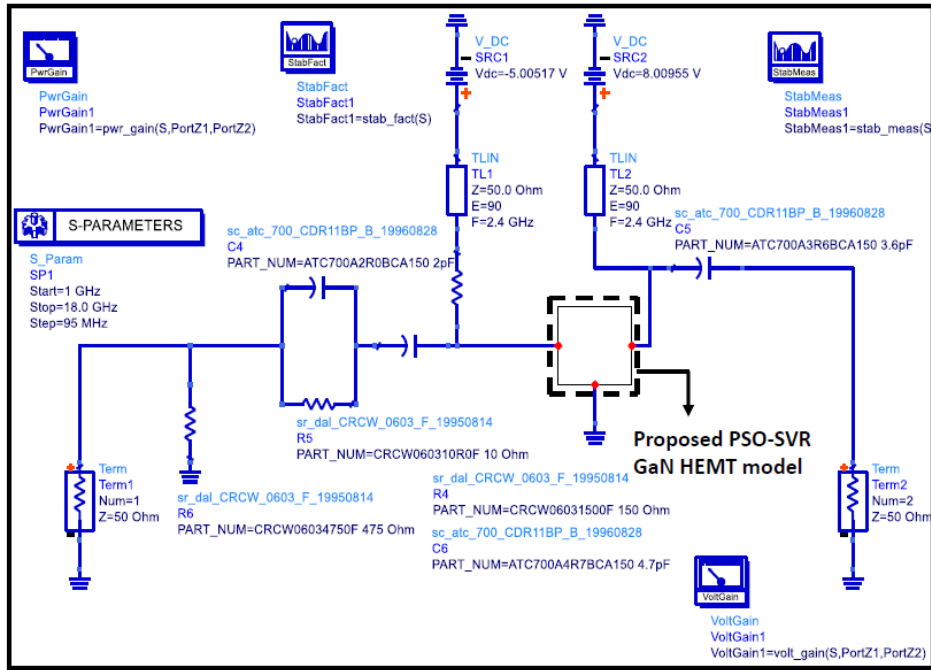


FIGURE 12. Stability and Gain Test Circuit of a Power Amplifier using proposed PSO-SVR GaN HEMT model.

TABLE 8. Comparison between Proposed Model and Small-signal Electrical Equivalent Model for GaN HEMT.

Reference/Response	Multi-biasing Condition	S ₁₁ (%)	S ₁₂ (%)	S ₂₁ (%)	S ₂₂ (%)
[16]	V _{GS} =0V; V _{DS} =8V	2.16	2.08	3.38	2.47
	V _{GS} =-3V; V _{DS} =20V	2.28	3.44	2.84	2.42
	V _{GS} =-5V; V _{DS} =6V	0.86	2.42	1.27	1.66
	V _{GS} =0V; V _{DS} =10V	1.76	3.01	1.61	0.78
[17]	V _{GS} =-5V; V _{DS} =7V	3.67	3.65	3.28	2.01
	V _{GS} =-7V; V _{DS} =0V	1.54	3.82	3.56	1.77
[18]	V _{GS} =-2.25V; V _{DS} =5V	0.32	1.35	0.96	2.56
	V _{GS} =-2V; V _{DS} =20V	1.65	2.64	4.28	1.86
[29]	All Biases	3.85	13.5	6.1	3.8
This work *	V _{GS} =-4V; V _{DS} =2V	0.64	1.66	3	1.69
	V _{GS} =-5V; V _{DS} =6V	0.77	0.91	3.03	0.74
	V _{GS} =-1V; V _{DS} =0V	0.68	2.02	2.31	0.44
	V _{GS} =-7V; V _{DS} =6V	1.19	0.37	1.05	1.92
	V _{GS} =-3V; V _{DS} =10V	0.83	1.58	0.92	0.62

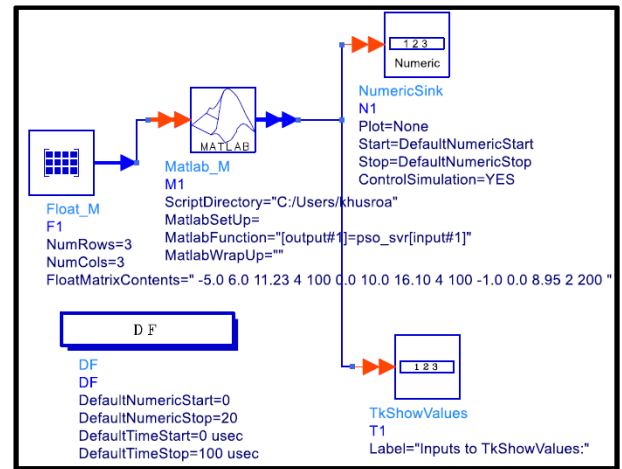


FIGURE 13. Interfacing between MATLAB and ADS environment.

has been developed in MATLAB. In this context, the section initially discusses the establishment of interface between MATLAB and CAD environment using MATLAB Cosimulation feature in Advanced Design System (ADS). MATLAB Cosimulation in ADS Ptolemy provides the MATLAB block that can actually pick the stored script or a function in MATLAB. In our work, a function was developed such that the trained model takes inputs and subsequently produces the response values. The input to the model are explicitly specified using float matrix block and passed into MATLAB function block and subsequently the responses are recorded using sink. The interfacing circuit between ADS and MATLAB is shown in Fig. 13. As the interface is established between MATLAB and ADS, the stability test for the class-F

power amplifier is performed and subsequently the amplifier gain is determined by incorporating the proposed PSO-SVR based small signal model. Determination of stability is one of the most important aspect for amplification. For a transistor amplifier to be unconditionally stable for all passive and source and load impedances, if $|\Gamma_{in}| < 1$ and $|\Gamma_{out}| < 1$ where Γ_{in} and Γ_{out} are reflection coefficient seen from the source and load side respectively and depends on source and load matching networks. In this context, one of the distinct K- Δ test is performed to test the stability of the network [47]. Where, K is known as Rollet's factor [47]. The amplifier is said to be unconditionally stable if the condition $K > 1$ and $\Delta < 1$ holds true for entire operating frequency range.

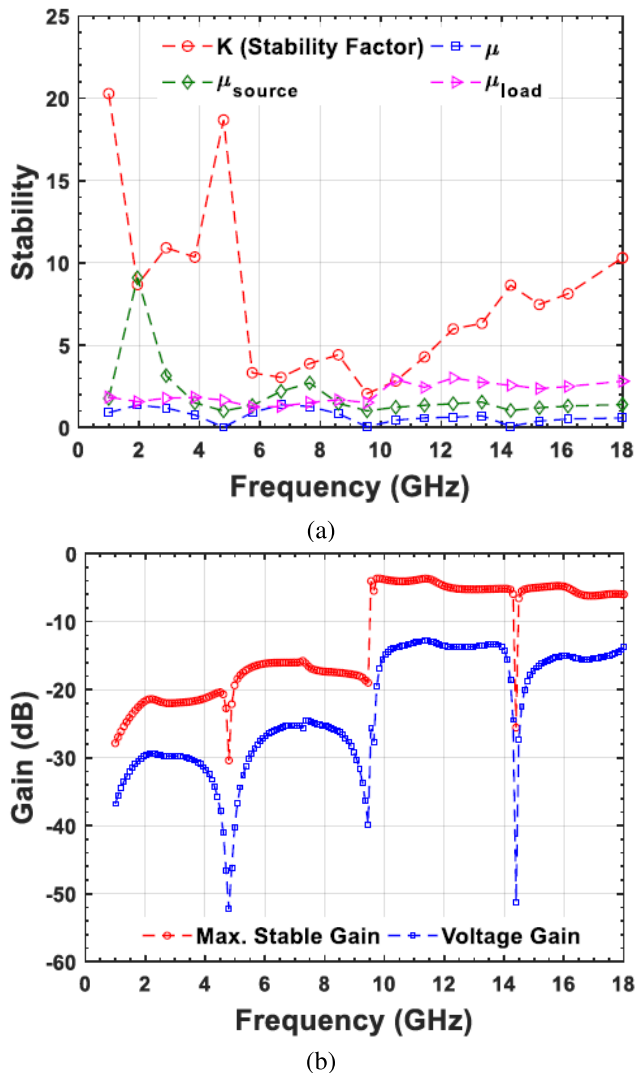


FIGURE 14. Utility of the proposed model (a) Stability Metrics, and (b) Gain of $2 \times 200 \mu\text{m}$ at $V_{GS} = 0 \text{ V}; V_{DS} = 6 \text{ V}$.

To perform the stability and gain test for amplifiers, it is vital to determine the quiescent bias point initially at which the transistor is operating. The bias points are determined from DC I-V test. The quiescent bias points and the corresponding drain current set (V_{GSQ} , V_{DSQ} , I_{DS} (mA)) obtained for $2 \times 200 \mu\text{m}$ and $4 \times 100 \mu\text{m}$ GaN HEMT are $(-5.00517, 8.00955, 24.17)$, $(-5.01596, 8.00955, 47.63)$ respectively. The simulation setup in 12 contains parallel stabilization network into bias circuitry with bias set equivalent to optimal bias values and proposed GaN HEMT model. Parallel stabilization network is a kind of high pass filter which is employed to attain the stability for lower frequency components and unattenuated high frequency component. Stability factors K and μ are studied (μ is recently proposed stability criterion well suited in simulators and it declares system to be unconditionally stable over a frequency range without any exception if $\mu > 0$). The proposed GaN HEMT is provided with a test input as $V_{GS} = 0 \text{ V}$, $V_{DS} = 6 \text{ V}$, $N=2$, $w_g=200 \mu\text{m}$ for a complete frequency range. The results related to the stability

of the amplifier is shown in Fig. 14 (a). It can be seen from the Fig.14 that $K > 1$ and $\mu > 0$ for complete frequency range which specifies that the amplifier is unconditionally stable. Apart from this, it is evident that the value of μ_{source} and μ_{load} are also greater than 1 for entire frequency range which implies that the source and load side are matched. As far as gain is concerned, the maximum available gain and small-signal voltage gain defined in [47] is obtained for the amplifier as shown in Fig. 14 (b). It can be seen from the Fig.14 (b) that the maximum stable gain or small-signal voltage gain improves to a good extent with the increase in frequency. It can be attributed to the fact that a parallel RC stabilization network is used in this simulation framework which improves attenuation at high frequencies. In practice, a combination of series RC and parallel RC stabilization network is used to achieve optimum performance.

V. CONCLUSION

In this article, a novel PSO-SVR based technique is presented to model the small-signal behavior of GaN HEMT device. The proposed modeling technique is accurate, fast, and less complex. The key contributions include the development of a kernel based modeling approach for modeling the small-signal behavior of GaN HEMT device. It picks the optimized set of kernel function to model the response variables of the model from defined kernel function library. Then the incorporation of PSO algorithm aids in the identification of the optimal sets of kernel function and associated hyperparameters for accurate modeling of GaN HEMT. The robustness of the proposed model is evaluated by testing it under various noise conditions, frequency extrapolation, and geometric interpolation. It has been shown that the proposed model is robust under various noise conditions, shows a very good frequency extrapolation ability, and excellent geometric interpolation capability. It has been also demonstrate that the proposed SVR based model exhibits far superior accuracy when compared to the conventional ANN based model. Finally, the effectiveness of the proposed model has been shown by integrating the proposed model into CAD environment and performing stability and gain test for a power amplifier.

ACKNOWLEDGMENT

The authors would also like to acknowledge the DRDO (an enterprise of Ministry of Defense) for providing access to their measurement facility and the measured data from their device.

REFERENCES

- [1] D. W. Runton, B. Trabert, J. B. Shealy, and R. Vetry, "History of GaN: high-power RF gallium nitride (GaN) from infancy to manufacturable process and beyond," *IEEE Microw. Mag.*, vol. 14, no. 3, pp. 82–93, May 2013.
- [2] A. Alt, D. Marti, and C. R. Bolognesi, "Transistor modeling: Robust small-signal equivalent circuit extraction in various HEMT technologies," *IEEE Microw. Mag.*, vol. 14, no. 4, pp. 83–101, Jun. 2013.
- [3] R. Weber, D. Schwantuschke, P. Bruckner, R. Quay, M. Mikulla, O. Ambacher, and I. Kallfass, "A 67 GHz GaN voltage-controlled oscillator MMIC with high output power," *IEEE Microw. Wireless Compon. Lett.*, vol. 23, no. 7, pp. 374–376, Jul. 2013.

- [4] R. Krishnamoorthy, N. Kumar, A. Grebennikov, and H. Ramiah, "A high-efficiency ultra-broadband mixed-mode GaN HEMT power amplifier," *IEEE Trans. Circuits Syst. II, Exp. Briefs*, vol. 65, no. 12, pp. 1929–1933, Dec. 2018.
- [5] R. J. Trew, "Wide bandgap semiconductor transistors for microwave power amplifiers," *IEEE Microw. Mag.*, vol. 1, no. 1, pp. 46–54, Mar. 2000.
- [6] J. Kang, C. McGuire, A. Margomenos, A. Kurdoghlian, D. Brown, and H. P. Moyer, "Ultra-wideband, high-dynamic range, low loss GaN HEMT mixer," *Electron. Lett.*, vol. 50, no. 4, pp. 295–297, Feb. 2014.
- [7] A. Jarndal, A. Z. Markos, and G. Kompas, "Improved modeling of GaN HEMTs on Si substrate for design of RF power amplifiers," *IEEE Trans. Microw. Theory Techn.*, vol. 59, no. 3, pp. 644–651, Mar. 2011.
- [8] G. P. Gibbino, A. Santarelli, and F. Filicori, "A GaN HEMT global large-signal model including charge trapping for multibias operation," *IEEE Trans. Microw. Theory Techn.*, vol. 66, no. 11, pp. 4684–4697, Nov. 2018.
- [9] S. Aamir Ahsan, S. Ghosh, S. Khandelwal, and Y. S. Chauhan, "Physics-based multi-bias RF large-signal GaN HEMT modeling and parameter extraction flow," *IEEE J. Electron Devices Soc.*, vol. 5, no. 5, pp. 310–319, Sep. 2017.
- [10] P. M. Cabral, J. C. Pedro, and N. B. Carvalho, "Nonlinear device model of microwave power GaN HEMTs for high power-amplifier design," *IEEE Trans. Microw. Theory Techn.*, vol. 52, no. 11, pp. 2585–2592, Nov. 2004.
- [11] A. Jarndal, R. Essaadali, and A. B. Kouki, "A reliable model parameter extraction method applied to AlGaIn/GaN HEMTs," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 35, no. 2, pp. 211–219, Feb. 2016.
- [12] M. A. Alim, M. M. Ali, A. A. Rezazadeh, and C. Gaquiere, "Temperature dependence of the Taylor series coefficients and intermodulation distortion characteristics of GaN HEMT," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 39, no. 3, pp. 552–559, Mar. 2020.
- [13] F. Y. Huang, X. S. Tang, Z. N. Wei, L. H. Zhang, and N. Jiang, "An improved small-signal equivalent circuit for GaN high-electron mobility transistors," *IEEE Electron Device Lett.*, vol. 37, no. 11, pp. 1399–1402, Nov. 2016.
- [14] S. A. Ahsan, A.-U.-H. Pampori, S. Ghosh, S. Khandelwal, and Y. S. Chauhan, "A new small-signal parameter extraction technique for large gate-periphery GaN HEMTs," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 10, pp. 918–920, Oct. 2017.
- [15] A. S. Hussein and A. H. Jarndal, "Reliable hybrid small-signal modeling of GaN HEMTs based on particle-swarm-optimization," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 37, no. 9, pp. 1816–1824, Sep. 2018.
- [16] G. Crupi, D. Xiao, D. M. M.-P. Schreurs, E. Limiti, A. Caddemi, W. De Raedt, and M. Germain, "Accurate multibias equivalent-circuit extraction for GaN HEMTs," *IEEE Trans. Microw. Theory Techn.*, vol. 54, no. 10, pp. 3616–3622, Oct. 2006.
- [17] R. G. Brady, C. H. Oxley, and T. J. Brazil, "An improved small-signal parameter-extraction algorithm for GaN HEMT devices," *IEEE Trans. Microw. Theory Techn.*, vol. 56, no. 7, pp. 1535–1544, Jul. 2008.
- [18] A. Majumder, S. Chatterjee, S. Chatterjee, S. S. Chaudhari, and D. R. Poddar, "Optimization of small-signal model of GaN HEMT by using evolutionary algorithms," *IEEE Microw. Wireless Compon. Lett.*, vol. 27, no. 4, pp. 362–364, Apr. 2017.
- [19] A. Khuro, M. S. Hashmi, A. Q. Ansari, A. Mishra, and M. Tarique, "An accurate and simplified small signal parameter extraction method for GaN HEMT," *Int. J. Circuit Theory Appl.*, vol. 47, no. 6, pp. 941–953, Jun. 2019.
- [20] A. Jarndal and G. Kompas, "An accurate small-signal model for AlGaIn-GaNHEMT suitable for scalable large-signal model construction," *IEEE Microw. Wireless Compon. Lett.*, vol. 16, no. 6, pp. 333–335, Jun. 2006.
- [21] G. Dambrine, A. Cappy, F. Heliodore, and E. Playez, "A new method for determining the FET small-signal equivalent circuit," *IEEE Trans. Microw. Theory Techn.*, vol. MTT-36, no. 7, pp. 1151–1159, Jul. 1988.
- [22] A. Khuro, M. S. Hashmi, A. Q. Ansari, and M. Auyenur, "A new and reliable decision tree based small-signal behavioral modeling of GaN HEMT," in *Proc. IEEE 62nd Int. Midwest Symp. Circuits Syst. (MWSCAS)*, Dallas, TX, USA, Aug. 2019, pp. 303–306.
- [23] J. Cai, J. B. King, J. Su, C. Yu, S. Chen, L. Sun, H. Wang, and J. Liu, "Bayesian inference-based behavioral modeling technique for GaN HEMTs," *IEEE Trans. Microw. Theory Techn.*, vol. 67, no. 6, pp. 2291–2301, Jun. 2019.
- [24] M. Hayati and B. Akhlaghi, "An extraction technique for small signal intrinsic parameters of HEMTs based on artificial neural networks," *AEU-Int. J. Electron. Commun.*, vol. 67, no. 2, pp. 123–129, Feb. 2013.
- [25] A. Khuro, M. S. Hashmi, and A. Q. Ansari, "Enabling the development of accurate intrinsic parameter extraction model for GaN HEMT using support vector regression (SVR)," *IET Microw., Antennas Propag.*, vol. 13, no. 9, pp. 1457–1466, Jul. 2019.
- [26] Z. Zhao, Y. Qu, Y. Zhang, and X. Shen, "Overview of microwave device modeling techniques based on machine learning," *Int. J. Advancements Comput. Technol.*, vol. 5, no. 9, pp. 299–306, May 2013.
- [27] A. Jarndal, "On neural networks based electrothermal modeling of GaN devices," *IEEE Access*, vol. 7, pp. 94205–94214, Jul. 2019.
- [28] A. Khuro, S. Husain, M. S. Hashmi, M. Auyenur, and A. Q. Ansari, "A reliable and fast ANN based behavioral modeling approach for GaN HEMT," in *Proc. 16th Int. Conf. Synth., Modeling, Anal. Simulation Methods Appl. Circuit Design (SMACD)*, Lausanne, Switzerland, Jul. 2019, pp. 277–280.
- [29] A.-D. Huang, Z. Zhong, W. Wu, and Y.-X. Guo, "An artificial neural network-based electrothermal model for GaN HEMT's with dynamic trapping effects consideration," *IEEE Trans. Microw. Theory Techn.*, vol. 64, no. 8, pp. 2519–2528, Aug. 2016.
- [30] S. Y. Lee, B. A. Cetiner, H. Torpi, S. J. Cai, J. Li, K. Alt, Y. L. Chen, C. P. Wen, K. L. Wang, and T. Itoh, "An X-band GaN HEMT power amplifier design using an artificial neural network modeling technique," *IEEE Trans. Electron Devices*, vol. 48, no. 3, pp. 495–501, Mar. 2001.
- [31] A. Khuro, S. Husain, M. S. Hashmi, and A. Q. Ansari, "Small signal behavioral modeling technique of GaN high electron mobility transistor using artificial neural network: An accurate, fast, and reliable approach," *Int. J. RF Microw. Comput.-Aided Eng.*, vol. 30, no. 4, pp. 1–15, Apr. 2020, doi: 10.1002/mmce.22112.
- [32] J. Cai, J. King, C. Yu, J. Liu, and L. Sun, "Support vector regression-based behavioral modeling technique for RF power transistors," *IEEE Microw. Wireless Compon. Lett.*, vol. 28, no. 5, pp. 428–430, May 2018.
- [33] G. Angiulli, M. Cacciola, and M. Versaci, "Microwave devices and antennas modelling by support vector regression machines," *IEEE Trans. Magn.*, vol. 43, no. 4, pp. 1589–1592, Apr. 2007.
- [34] Y. Xu, Y. Guo, R. Xu, and Y. Wu, "Modeling of SiC MESFETs by using support vector machine regression," *J. Electromagn. Waves Appl.*, vol. 21, no. 11, pp. 1489–1498, Jan. 2007.
- [35] S. Koziel and J. W. Bandler, "Support-vector-regression-based output space-mapping for microwave device modeling," in *IEEE MTT-S Int. Microw. Symp. Dig.*, Atlanta, GA, USA, Jun. 2008, pp. 615–618.
- [36] A. Khuro, M. S. Hashmi, and A. Q. Ansari, "Exploring support vector regression for modeling of GaN HEMT," in *IEEE MTT-S Int. Microw. RF Conf. (IMARC)*, Kolkata, India, Oct. 2019, pp. 1–3.
- [37] C. Xu, B. Gordan, M. Koopialipoor, D. J. Armaghani, M. M. Tahir, and X. Zhang, "Improving performance of retaining walls under dynamic conditions developing an optimized ANN based on ant colony optimization technique," *IEEE Access*, vol. 7, pp. 94692–94700, Jul. 2019.
- [38] V. N. Vapnik, *The Nature of Statistical Learning Theory*. New York, NY, USA, Springer-Verlag, 1997.
- [39] B. Scholkopf, *Support Vector Learning*. Munich, Germany: Oldenbourg-Verlag, 1997.
- [40] Z.-L. Qian, D.-C. Juan, P. Bogdan, C.-Y. Tsui, D. Marculescu, and R. Marculescu, "A support vector regression (SVR)-based latency model for network-on-chip (NoC) architectures," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 35, no. 3, pp. 471–484, Mar. 2016.
- [41] F. Ye, F. Firouzi, Y. Yang, K. Chakrabarty, and M. B. Tahoori, "On-chip droop-induced circuit delay prediction based on support-vector machines," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 35, no. 4, pp. 665–678, Apr. 2016.
- [42] M. Pradhan, B. B. Bhattacharya, K. Chakrabarty, and B. B. Bhattacharya, "Predicting X-sensitivity of circuit-inputs on test-coverage: A machine-learning approach," *IEEE Trans. Comput.-Aided Design Integr. Circuits Syst.*, vol. 38, no. 12, pp. 2343–2356, Dec. 2019.
- [43] C. Li, X. Wang, Z. Cheng, and Y. Bai, "Forecasting bus passenger flows by using a clustering-based support vector regression approach," *IEEE Access*, vol. 8, pp. 19717–19725, Jan. 2020.
- [44] A. Dai, X. Zhou, H. Dang, M. Sun, and Z. Wu, "Intelligent modeling method for a combined radiation-convection grain dryer: A support vector regression algorithm based on an improved particle swarm optimization algorithm," *IEEE Access*, vol. 6, pp. 14285–14297, Feb. 2018.
- [45] Y. Zhang, H. Sun, and Y. Guo, "Wind power prediction based on PSO-SVR and grey combination model," *IEEE Access*, vol. 7, pp. 136254–136267, Sep. 2019.

- [46] A. Khusro, M. S. Hashmi, and A. Q. Ansari, "Empirical device scaling and RF performance perspective: A small signal model for GaN high electron mobility transistor," in *Proc. Int. Conf. Comput. Netw. Commun. (CoCoNet)*, Astana, Kazakhstan, Aug. 2018, pp. 45–49.
- [47] M. D. Pozar, *Microwave Engineering*. Hoboken, NJ, USA: Wiley, 2012.



AHMAD KHUSRO received the B.Tech. degree in electronics and communication engineering from Uttar Pradesh Technical University, Lucknow, India, in 2011, and the M.Tech. degree from Jamia Millia Islamia (A Central University), New Delhi, India, in 2015. He is currently pursuing the Ph.D. degree in electrical engineering with Jamia Millia Islamia (A Central University). His research interests include characterization and modeling of RF/Microwave devices with particular focus on modeling of GaN HEMT for broadband power amplifiers. He worked as a Research Intern with the Integrated Microwaves Packaging Antennas and Circuits Technology (IMPACT) Laboratory, King Abdullah University of Science and Technology (KAUST), Jeddah, Saudi Arabia. He also acts as a Reviewer to several international journals.



SADDAM HUSAIN received the B.Tech. degree from the Galgotias College of Engineering and Technology, Noida, India, and the M.Tech. degree from Jamia Millia Islamia (A Central University), New Delhi, India, in 2016 and 2019, respectively. He is currently pursuing the Ph.D. degree from Nazarbayev University, Kazakhstan. He held an Internship position with the Circuit Design Research Laboratory, IIIT Delhi, India, and worked on Modeling of GaN HEMT using Machine Learning techniques. His research interests include machine learning-based modeling and optimization, HEMTs, modelling and characterization of semiconductor, and RF/Microwave devices.



MOHAMMAD S. HASHMI (Senior Member, IEEE) received the B.Tech. degree from Aligarh Muslim University, India, the M.S. degree from the Darmstadt University of Technology, Germany, and the Ph.D. degree from Cardiff University, Cardiff, U.K., all in electrical engineering and information technology. He had held research and engineering positions with the University of Calgary, Canada, Cardiff University, U.K., Thales Electronics GmbH, Germany, and the Philips Technology Center, Germany. He is currently an Associate Professor with Nazarbayev University, Kazakhstan, and also holds a faculty position at IIIT Delhi, India. His current research interests include advanced RF circuits, broadband linear and efficient power amplifiers for mobile and satellite applications, high- and low-frequency instrumentation, and wireless power transfer techniques. His research activities have led to one book, three U.S. patents (two pending), and over 185 journal and conference publications. He is an Associate Editor of *IEEE Microwave Magazine*.



ABDUL QUAUYUM ANSARI (Senior Member, IEEE) received the B.Sc.Engg. degree from Aligarh Muslim University, Aligarh, India, the M.Tech. degree from IIT Delhi, Delhi, India, and the Ph.D. degree from Jamia Millia Islamia, New Delhi, India. He held academic positions at University Polytechnic Jamia Millia Islamia, the College of Computer Science, King Khalid University Kingdom, Saudi Arabia, the Department of Computer Science, Jamia Hamdard (Hamdard University), Delhi, and the Department of Electrical Engineering, Jamia Millia Islamia. He has produced excellent results with proven records through two patent applications, 58 peer-reviewed international journal articles, one book each written and edited, 12 book chapters, and 90 peer-reviewed international conference papers. His research interests include HEMTs and microwave devices, soft computing, computer networks and data communications, embedded systems, and NoC. He is a fellow and a Chartered Engineer of the Institution of Engineering India (IEI) and IETE.



SULTANGALI ARZYKULOV (Member, IEEE) received the B.Sc. degree (Hons.) in radio engineering, electronics, and telecommunications from Kazakh National Research Technical University after K. I. Satpayev, Almaty, Kazakhstan, in June 2010, the M.Sc. degree in communication engineering from The University of Manchester, Manchester, U.K., in 2013, and the Ph.D. degree in science, engineering, and technology from Nazarbayev University, Nur-Sultan, Kazakhstan, in 2019. He is currently a Postdoctoral Scholar with Nazarbayev University, Kazakhstan. His research interests include wireless communication systems, with particular focus on cooperative communications, cognitive radio, energy harvesting, interference mitigation, and NOMA. He acts as a Reviewer for several journals/conferences and served as a Technical Program Committee Member of numerous IEEE Communication Society flagship conferences.

...