

ORIGINAL RESEARCH PAPER

On temperature-dependent small-signal modelling of GaN HEMTs using artificial neural networks and support vector regression

Anwar Jarndal¹  | Saddam Husain¹ | Mohammad Hashmi²¹Electrical Engineering Department, University of Sharjah, Sharjah, United Arab Emirates²Electrical and Computer Engineering Department, Nazarbayev University, Nur-Sultan, Kazakhstan**Correspondence**

Anwar Jarndal, Electrical Engineering Department, University of Sharjah, Sharjah, United Arab Emirates.

Email: ajarndal@sharjah.ac.ae**Funding information**

University of Sharjah

Abstract

Machine learning-based efficient temperature-dependent small-signal modelling approaches for GaN high electron mobility transistors (HEMTs) are presented by the authors here. The first method is an artificial neural network (ANN)-based and makes use of the well-known multilayer perceptron (MLP) architecture whereas the second technique is developed using support vector regression (SVR). The models are trained on a large set of measurement data obtained from a 2-mm GaN-on-silicon device operating under varying operating conditions (bias voltages and ambient temperatures) over a wide frequency range of 0.1 to 20 GHz. An excellent agreement is found between the measured and the simulated S-parameters for both models over the entire frequency range. It is identified that the training process and prediction capability of ANN is superior to SVR. However, the SVR is more robust when compared to the artificial neural network (ANN) in term of its sensitivity to local minima and uniqueness of the final solution. Subsequently, the performances of the proposed ANN- and SVR-based models are improved by incorporating particle swarm optimization (PSO) in the model development process. The PSO improves the uniqueness of the ANN model whereas it enhances the performance of the SVR by optimising its control parameters. The proposed models exhibit very good accuracy and scalability.

1 | INTRODUCTION

GaN high electron mobility transistor (HEMT) possesses excellent features such as higher electron saturation velocity, electron mobility, breakdown voltage and operating temperature [1, 2]. These features make it an optimal device for the design of advanced communication circuits such as power amplifiers (PAs) and low noise amplifiers (LNAs) [3–5]. However, the GaN devices face multiple technical challenges such as reduction in the output power, gain and power efficiency due to thermal effect and DC-RF dispersion [6, 7]. Furthermore, the lower resistivity and thermal conductivity of Si (silicon) with respect to other substrates, such as SiC and Diamond, result in relatively inferior RF characteristics for the HEMTs due to self-heating and substrate-loading effects [8]. This self-heating in addition to the ambient temperature has strong impact on the small- and large-signal characteristics of GaN HEMT devices. It is also important to note that the reliability of circuits designed without consideration of these heating effects is also questionable.

Keeping the above aspects in perspectives, a number of electrothermal modelling techniques of GaN devices subjected to large- and small-signal induced thermal effects have been reported [9–15]. In recent times, the artificial neural network (ANN)-based modelling techniques are getting popular for both large-signal [9–12] and small-signal modelling [13–15]. One of the most important features of ANN is its ability to learn complex nonlinear relationship and adapt well to the new data [16]. In addition, ANN has the ability to learn and mimic the behaviour of devices, without looking into physics of the device, and hence finds usefulness in the modelling of GaN HEMT transistor device having strong nonlinearities. The support vector regression (SVR) is also being employed for solving RF and microwave device behavioural modelling problems [17–23]. The SVR is often preferred as it looks for global optima in contrast to the conventional structural risk minimization (SRM) algorithms, which look for local minima as in the case of ANN. The ANN-based modelling for GaN-on-SiC HEMT using S-parameter measurements at a limited temperature range of 20 to 80°C

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2021 The Authors. *IET Microwaves, Antennas & Propagation* published by John Wiley & Sons Ltd on behalf of The Institution of Engineering and Technology.

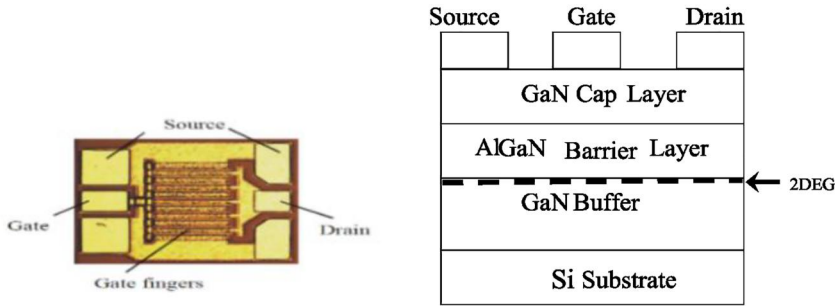


FIGURE 1 Photograph of 2-mm ($10 \times 200 \mu\text{m}$) AlGaIn/GaN on Si high electron mobility transistors and its general epitaxial structure

was reported in [13]. The model consisted of two hidden layers with 20 to 28 neurons in each layer and this complicates the model implementation in computer-aided-design (CAD) tools. The other ANN-based models either considered only the voltage and frequency dependence or used simple linear approach [14, 15] to predict the temperature-dependence, which could be acceptable at low temperatures. However, at high temperatures, some S-parameters such as S_{22} show nonlinear behaviour [11] and the reported techniques may not be scalable for such scenarios.

This paper proposes small-signal modelling techniques based on ANN and SVR for GaN-on-Si HEMT device strongly impacted by the thermal effects. The ANN-based technique utilises an enhanced procedure to implement a double hidden layer of five neurons. This drastically reduces the model complexity, simplifies its implementation and enhances simulation-speed of the CAD implemented model. This technique, developed using wide range of temperature data of 25 to 175°C, provides more insight as compared to the earlier reported ANN-based small-signal modelling techniques [13, 15]. Here, the temperature is the third input in addition to the voltage and frequency achieving excellent agreement between the measured and the modelled performance. During the model development process, the back propagation (BP)-based procedure is used to train the ANN model. One of the main limitations of BP is its initial guess dependency. The final solution may not converge to provide best fitting if the initial values of the model parameters (weights and biases) are far from the optimal ones. To overcome this limitation, particle swarm optimization (PSO) [24, 25] has been used in conjunction with the ANN here. PSO as a global technique is used to initiate the training process by exploring the search space to find the area of optimal solution. Then in the next phase the BP as a local technique could provide stronger exploitation with higher rate of convergence to find the optimal final solution. This implemented PSO-BP technique improves the efficiency and accuracy of the NN small-signal modelling and this could be considered as an additional contribution with respect to the previously reported works. Furthermore, this paper also develops for the first time a SVR-based temperature-dependent small-signal modelling technique for GaN HEMT. It has been reported that SVR-based model with properly selected parameters could provide almost similar performance as compared to the ANN-based model [26]. However, inappropriate selection of the parameters could result in an over- or under-fitting and this may deteriorate the SVR-based model

performance. This necessitates careful user intervention to set the parameters of SVR to overcome this issue of over- and under-fitting. This paper therefore makes use of PSO with the SVR model to address this problem.

The next section of this paper succinctly describes the device and the relevant measurement. Sections 3 and 4 provide the model development process using ANN and SVR respectively. Results and discussion are provided in Sections 5 and 6, whereas Section 7 concludes the paper.

2 | DEVICE PHYSICS AND CHARACTERISATION

The gate width of the considered device is 2 mm and is composed of 10 fingers of 200- μm gate-width. A photograph of the coplanar waveguide (CPW) on-wafer device and its general structure is shown in Figure 1 [27]. This device has been grown on Si substrates and fabricated by Nitronex Corporation using NRF1 process. The performance of this device has been improved by adopting source field plate technique. The device is characterised using vector network analyser (VNA) and the data are represented in terms of real and imaginary parts of S-parameters.

The device is mounted on a temperature controlled thermal chuck to carry out the S-parameters at different external temperature and its representative performance in terms of real and imaginary parts of S_{22} are shown in Figure 2. The internal temperature of the device at active bias condition is the summation of the temperature defined by the thermal chuck and the device self-heating. The measurements were carried out at the grid of bias conditions listed in Table 1. The same grid of measurements is conducted at different external temperature starting from 25 to 175°C in steps of 25°C. All measurements have been taken over the frequencies from 100 MHz to 20 GHz.

3 | ANN-BASED MODEL

3.1 | Theoretical preliminaries

ANN consists of interconnecting layers of neurons as their fundamental constituent. These neurons can learn simple and sophisticated patterns associated with any data if they are fed

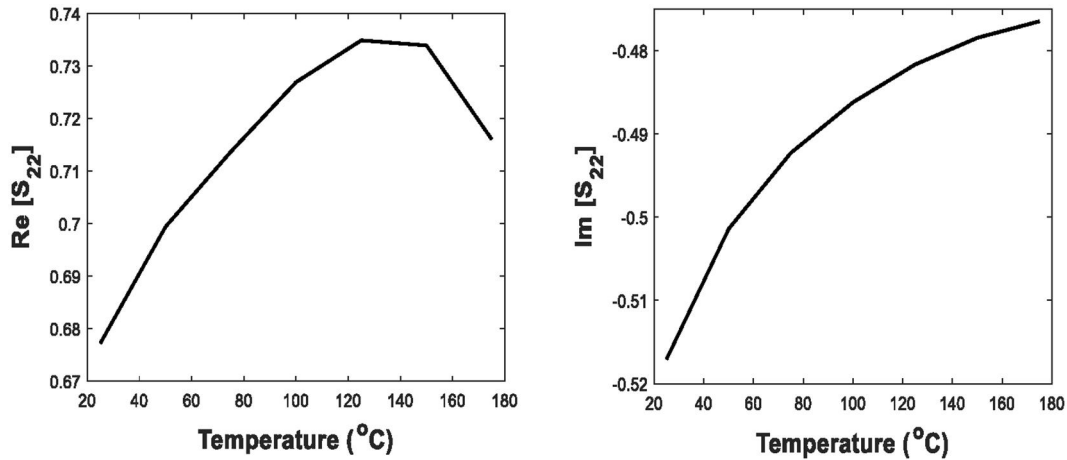


FIGURE 2 Measured S22 of $10 \times 200 \mu\text{m}$ GaN-on-Si high electron mobility transistors at different ambient temperatures

TABLE 1 Gate and drain bias voltages of the S-parameters measurements

V_{GS} (V)		V_{DS} (V)
Range	Step size	
-2 to 2	0.2	7
-2 to -0.4	0.1	28
-2 to -0.8	0.1	48

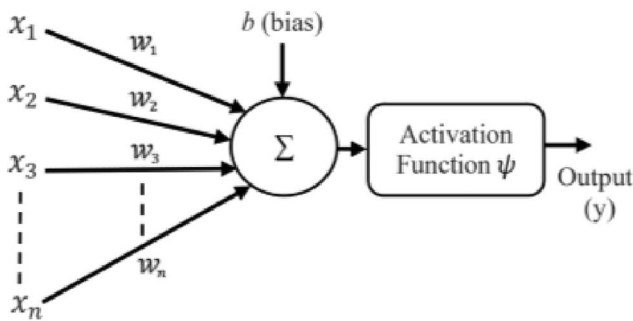


FIGURE 3 The basic model of a unit neuron

with balanced and sufficient training data. Having learnt properly the information stored, it can be easily exploited for prediction on new sets of data [28]. The ANN has been widely used for solving modelling problems and it has many advantages with respect to other techniques. In particular, its ability to exploit the nonlinearity stands out the most. In its basic embodiment, the structure of a unit neuron is shown in Figure 3. Here, x_1, x_2, \dots, x_n are inputs, w_1, w_2, \dots, w_n are synaptic weights and b is the bias term.

In particular, a unit neuron takes in the inputs multiplied by its corresponding weights, sums that up, and then adds a bias term to scale the weighted sum. The weighted sum could have the value in the range $(-\infty, \infty)$. To force this to a specified range, it goes through a linear or non-linear activation function. The output y can be calculated using (1), which describes the learning equation of one neuron. Each neuron uses the

same learning equation to learn the weights and biases. Now, each neuron can be defined by the same equation with additional information of layer number and its position in the same layer. The ANN is formed by interconnection of different neurons stacked on top of each other.

$$y = \psi \left(\sum_{i=1}^n w_i x_i + b \right) \quad (1)$$

3.2 | Temperature-dependent model using ANN

The implemented multilayer perceptron (MLP) ANN architecture shown in Figure 4 consists of an input layer, intermediate layers that are often known as hidden layers and output layers [28]. The model topology (number of layers and neurons) is directly related to non-linearity associated with the considered problems. The number of the hidden layers and their sizes cannot be set a priori, so they are determined during training of the model and it depends on the degree of nonlinearity. The learning equation for each ANN model can be formed by vectorization of individual neurons. It is formulated in (2) for the considered topology. The architecture used consists of four input nodes, two hidden layers with five neurons in each layer, and an output layer for each S-parameter. The model is then used to simulate the voltage, temperature and frequency dependence of the real or imaginary parts of S-parameter. The complete model consists of eight ANN models to simulate the four complex S-parameters as shown in Figure 5. Here, V_{GS}, V_{DS}, T and f are extrinsic gate voltage, extrinsic drain voltage, ambient temperature, and frequency, respectively. w_{1j}, w_{2j}, w_{3j} and w_{4j} are input weights, w_{kj} is the intermediate weights (between two hidden layers). Furthermore, w_{bj}^1, w_{kb}^2 and w_b^3 are the input-layer, hidden-layer and output-layer biases, respectively. The tanh is the non-linear activation function that is used to build the model. The

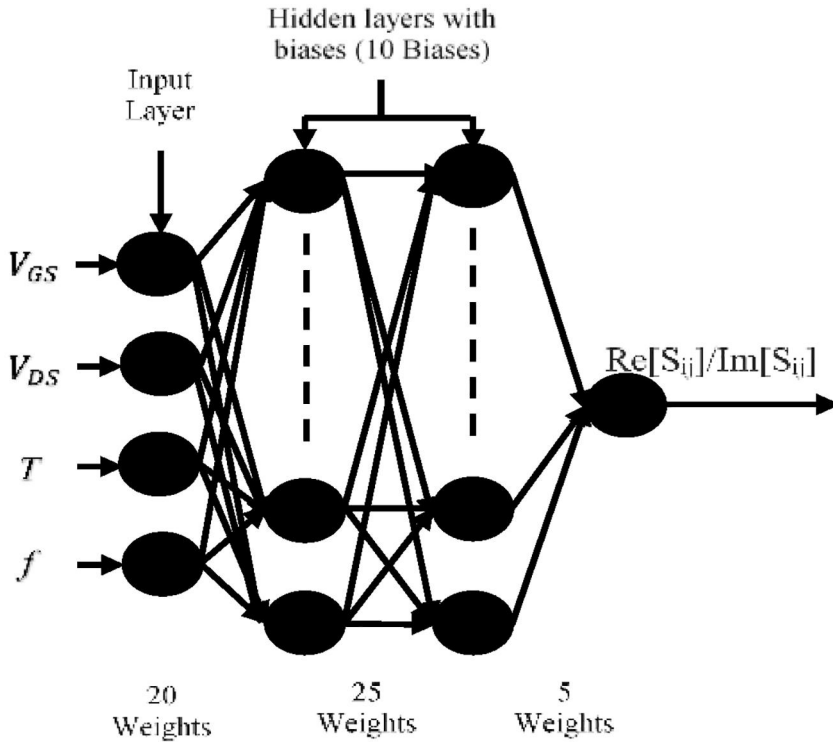


FIGURE 4 The proposed double hidden layer multilayer perceptron ANN-based model for mimicking the S-parameters

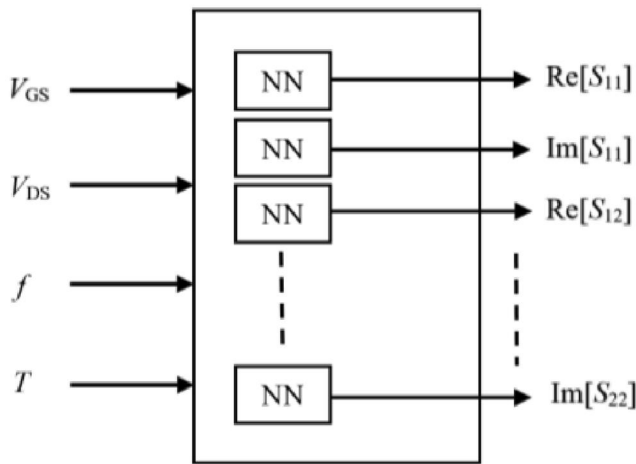


FIGURE 5 Generic depiction of ANN-based modelling for GaN high electron mobility transistors S-parameters

weights and biases of the model are later optimised using PSO algorithm to find the optimal values.

$$y = w_b^3 + \sum_{k=1}^5 w_k \tanh \left(w_{kb}^2 + \sum_{j=1}^5 w_{kj} \tanh \left(w_{1j} V_{GS} + w_{2j} V_{DS} + w_{3j} T + w_{4j} f + w_{bj}^1 \right) \right) \quad (2)$$

One of the most important parts before training the model is feature scaling also known as pre-processing of the dataset. If there is significant difference in the range of the feature vectors,

the contour of the cost function will have very skewed elliptical shapes (very tall and skinny elliptical shapes). If the training algorithm is run on this type of skinny contour, then the gradients will take very long time to reach the global optimum and sometimes they could easily get stuck into a local optimum. The model is built based on the S-parameters measurements under the listed bias conditions in Table 1 and over the frequency range of 0.1–20 GHz. The same listed measurements at ambient temperatures of 25, 75 and 150 C are used to train and test the ANN models. To check the model's generalisation capability, the model is validated using two independent datasets (which are not used for building the model). The first dataset consists of the same bias conditions and frequencies at $T = 50$ C and $T = 125$ C, while the other dataset is at $T = 175$ C. The proposed work makes use of the concept of normalisation. One of the main reasons why we used real and imaginary parts of S-parameters as opposed to the magnitude and phase of S-parameters is the distribution of data points. Except for S_{21} , all other parameters are by default within the range of $[-1, 1]$ in the case of real and imaginary parts of S-parameters.

In principle, the training algorithm initialises the random sets of weights and biases, updates them, and reassigns them again after each epoch until the algorithm converges. The cost function used for this problem is given in (3), where y_i is the measured S-parameters and \hat{y}_i is the simulated S-parameters. The Levenburg-Marquardt (LM) BP algorithm is used to train and test the model [29], whereas the Nyugen-Widrow method [30] is adopted for initialisation of weights and biases. This initialisation method significantly expedited the training and almost all the neurons are utilised in the input space. Keeping the nature of the LM-BP as a local optimisation technique in context, the ANN

model has been trained multiple times until satisfactory results are obtained. Those outcomes resulting in minimum error computed using (3), are taken into consideration.

$$MSE = \frac{1}{N} \sum_{i=1}^N (y_i - \tilde{y}_i)^2 \quad (3)$$

3.3 | ANN-based improved model

The ANN-based model suffers from the convergence issues at local minima. The BP algorithm utilises random initial weights to update and train the model, but this method sometimes produces over fitted results and this in turn does not perform well for test sets. So, in principle, the model needs to be trained multiple times until it gives satisfactory results. This needs more effort and may not be practical for strong nonlinear models [15]. Here, this major drawback of ANN model is addressed by utilising the global optimisation method of PSO. The PSO is a multiple-initial-guess based technique and therefore reduces the chance of convergence at local minimums. This attribute of PSO is very vital to improve the effectiveness of the model. The model is developed utilising the same conditions listed in Table 1, over a frequency range of 0.1–20 GHz. The flow diagram of the proposed performance enhanced ANN model is given in Figure 6. In this technique, the PSO starts by creating a set of initial particles. For each particle in the swarm, it evaluates the fitness function given in (3), set the local best and global best positions. Then it calculates and updates the velocity of the particle based on the current velocity, particle individual best position and global best. Having updated velocity, it again updates the particle position and goes forward for the next particle and repeats the same steps until termination criteria is reached. The proposed procedure is illustrated by the flow chart in Figure 6 and it can be briefly described as follows:

- First the training set is inserted with inputs and corresponding S-parameter. The optimisation algorithm starts by creating a population of particles (100 particles). The weights/biases (61) are initialised using symmetric random weight initiation function. For each variable, the lower and upper bounds are set to -1 and 1, respectively.
- The objective function is evaluated. The objective is to minimise the error between predicted values and the measured values. Utilising the fitness value, the algorithm sets the local best position and the global best position of the corresponding particle.
- Once the local best and the global best positions are known, the next step is to update the velocity using the memory values of velocity, current best position and global best position. The PSO uses (4) and (5) to update the velocity and position, respectively.

$$v_{t+1} = w * v_t + c_1 * r_1 * (p_{best} - x_t) + c_2 * r_2 * (g_{best} - x_t) \quad (4)$$

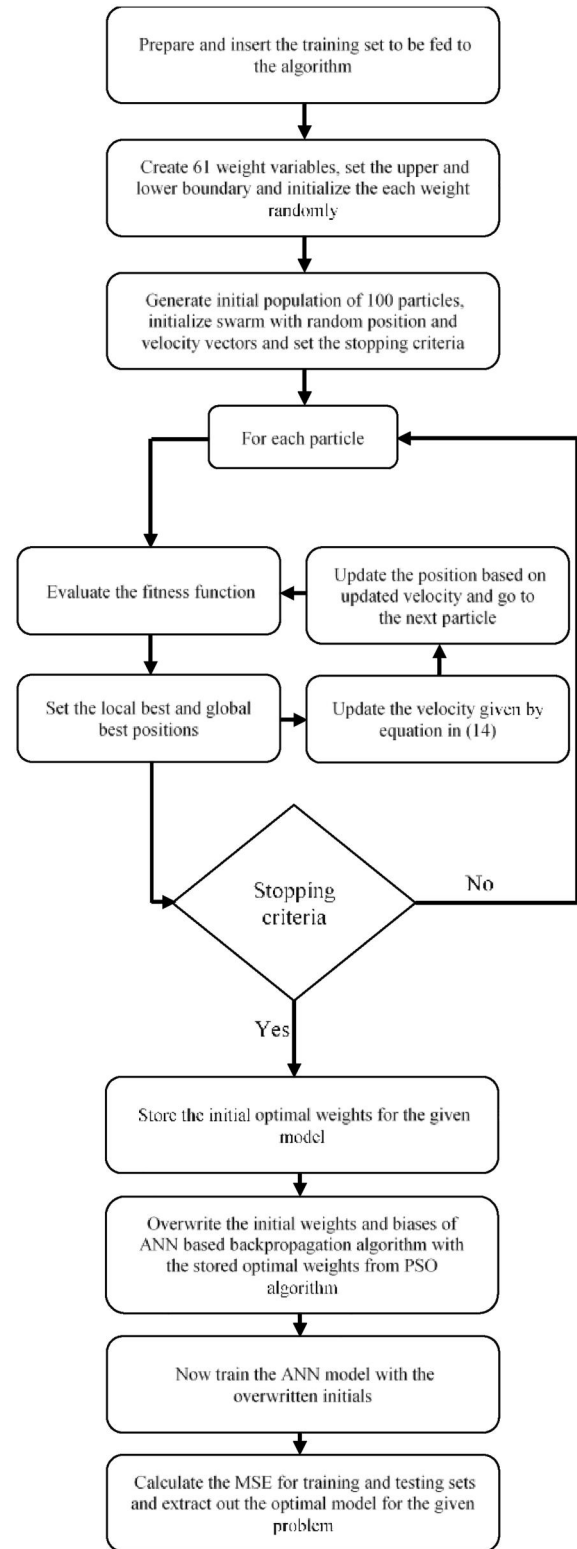


FIGURE 6 Flow chart of proposed PSO-ANN-based model

$$x_{t+1} = x_t + v_{t+1} \quad (5)$$

where r_1 and r_2 are generated randomly between 0 and 1. v_{t+1} is the new velocity and v_t is the previous velocity. Similarly, x_{t+1} is the new position and x_t is the previous position.

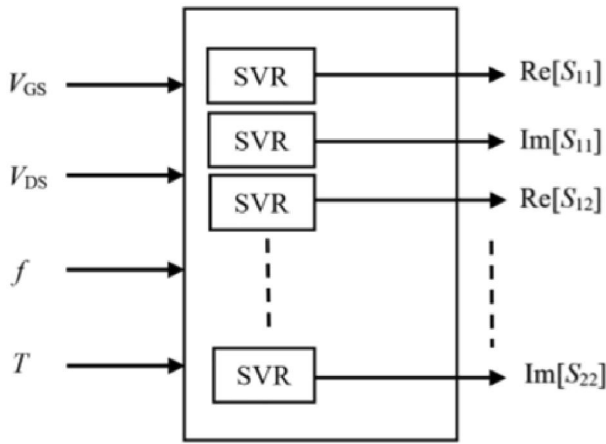


FIGURE 7 Depiction of support vector regression (SVR)-based GaN high electron mobility transistors modelling

The p_{best} and g_{best} are the local best and global best positions, respectively. The terms w , c_1 and c_2 are the inertia weight factor, self-confidence factor and swarm confidence factor, respectively. The inertia factor is calculated using (6).

$$w = w_{\text{max}} - \left(\frac{(w_{\text{max}} - w_{\text{min}})}{\text{MaxIteration}} * \text{Iteration} \right) \quad (6)$$

Here, w_{max} and w_{min} are the maximum and minimum range for the inertia factor values, which can greatly influence the updating process of new velocity based on other parameters.

- The new positions are again passed through the same fitness function and then all the steps are repeated until the termination criteria is reached. To build the ANN model with PSO initialisation, the algorithm runs for 500 iterations.
- After successful completion of all the above steps, the algorithm renders optimal set of weights/biases. These optimal values are later utilised as initial values for BP algorithm.
- The initial optimal weights overwrite the BP algorithm. Then the ANN model is trained on these weights to give the best results.

4 | SUPPORT VECTOR REGRESSION

4.1 | Theoretical preliminaries

The SVRs are mainly characterised by the use of kernels, absence of local minima, support vectors and solution determination by use of feature space [31]. The effectiveness of SVR depends mainly on the parameters such as type of kernel function, kernel parameters, support vectors decision boundary and the imposed penalty factor that is box constraint (C). The kernel trick maps the low-dimensional dataset into high-dimensional for the algorithm to learn the non-linear

behaviour of parameters such as S_{22} . In principle, the SVR algorithm tries to find the margin or regression equation, which is maximally deviated by ϵ from the true values. Suppose the training example is described as $\{x_i, y_i\}$, where i varies from 1 to n , $x_i \in \mathbb{R}^d$ and $y_i \in \mathbb{R}$, then the fitting function is given by (7). Here, $f(x)$ denotes the predicted value based on the optimal sets of weights and biases, and ω is weighing vector and b is the bias term.

$$f(x) = \omega^T \varphi(x_i) + b \quad (7)$$

The input vectors x_i can be mapped into higher-dimensional space, $\varphi(x_i)$, where $\varphi(x_i)$ serves as a kernel trick to facilitate the easier solution of the non-linear problems by mapping them into higher dimensional space and performing the simple dot products. Equation (7) is referred to as the learning equation for the SVR, which modifies according to (8). The model uses ϵ -insensitive loss function, which means the model will not care as long as error ranges between $-\epsilon$ to ϵ , which should be maximised to find the better generalisation. The loss function calculates the distance between measured y_i and the ϵ boundary using (9). The parameter C controls the weighing between the twin goals of making the $\|\omega^2\|$ small (to make the margin large) and ensuring that examples have functional margin of at least 1. To address the training examples that lie beyond the ϵ -insensitive zone, slack variables ξ_i and ξ_i^* at each point can be introduced as the soft-margin regression. Therefore, error equation can be modified to (10) [32] subject to the constraints in (11). This primal optimisation problem is solved by exploiting the Lagrange multipliers reproduced in (12) [32].

$$\text{Minimize} \left(\frac{1}{2} \|\omega^2\| + C \sum_{i=1}^n L(\epsilon) \right). \quad (8)$$

$$L(\epsilon) = \begin{cases} 0, & \text{if } |y - f(x)| < \epsilon \\ |y - f(x)| - \epsilon, & \text{otherwise.} \end{cases} \quad (9)$$

$$\text{Minimize} \left(\frac{1}{2} \|\omega^2\| + C \sum_{i=1}^n (\xi_i + \xi_i^*) \right), \quad (10)$$

$$y_i - (\omega, x_i + b) \leq \epsilon + \xi_i \quad (11a)$$

$$(\omega, x_i + b) - y_i \leq \epsilon + \xi_i^* \quad (11b)$$

$$\xi_i \geq 0 \quad \xi_i^* \geq 0. \quad (11c)$$

$$\begin{aligned} L = & \frac{1}{2} \|\omega^2\| + C \sum_{i=1}^n (\xi_i + \xi_i^*) - \sum_{m=1}^n (\eta_m \xi_i + \eta_m^* \xi_i^*) \\ & + \sum_{m=1}^n \alpha_m (f(x_m) - y_m - \epsilon - \xi_m) + \sum_{m=1}^n \alpha_m^* (y_m \\ & - f(x_m) - \epsilon - \xi_m^*) \end{aligned} \quad (12)$$

where α_m , α_m^* , η_m and η_m^* are non-negative Lagrangian multipliers. Putting the partial derivatives of (12) with

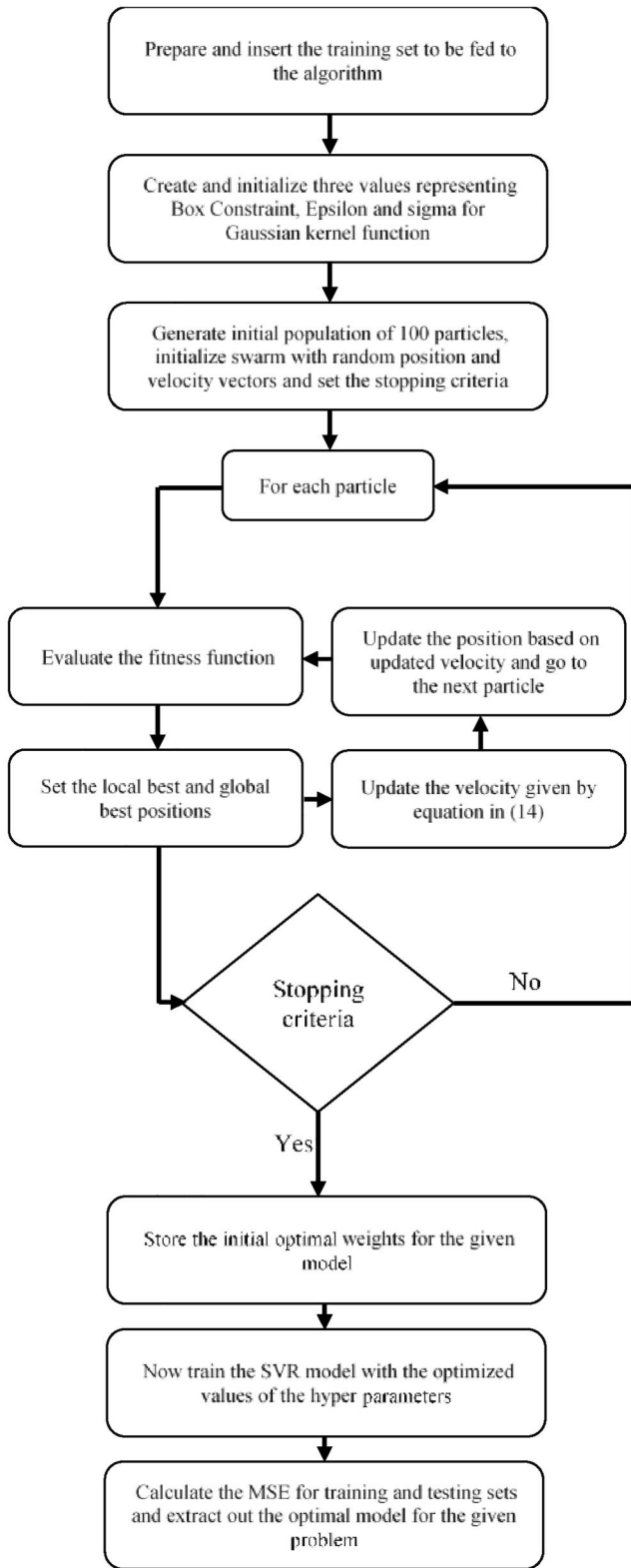


FIGURE 8 Flow chart of proposed particle swarm optimization based support vector regression model

respect to primal variables $(\omega, b, \xi_i, \xi_i^*)$ to zero for optimality and substituting those values back in (10) yields dual Formula (13). This method then needs to satisfy the

TABLE 2 MSE for testing set at randomly selected bias conditions and ambient temperatures (ANN model)

Bias and temperature	Re	Img	Re	Img	Re	Img	Re	Img
	S_{11}	S_{11}	S_{21}	S_{21}	S_{12}	S_{12}	S_{22}	S_{22}
$V_{GS} = 1.6 \text{ V}$	1.23	1.54	1.99	2.12	1.25	2.25	2.71	3.55
$V_{DS} = 7 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 25^\circ\text{C}$								
$V_{GS} = -1 \text{ V}$	1.81	2.51	1.23	4.58	2.25	3.55	2.46	3.59
$V_{DS} = 28 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 100^\circ\text{C}$								
$V_{GS} = -1.5 \text{ V}$	2.35	3.09	1.81	2.89	1.47	2.46	3.45	2.41
$V_{DS} = 28 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 150^\circ\text{C}$								

TABLE 3 MSE for validation using independent set at different ambient temperature for randomly selected dataset (first three: interpolation case, last three: extrapolation case) (ANN Model)

Bias and temperature	Re	Img	Re	Img	Re	Img	Re	Img
	S_{11}	S_{11}	S_{21}	S_{21}	S_{12}	S_{12}	S_{22}	S_{22}
$V_{GS} = -1.4 \text{ V}$	4.78	5.24	3.25	5.69	4.18	5.33	4.19	6.48
$V_{DS} = 48 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 50^\circ\text{C}$								
$V_{GS} = -1.8 \text{ V}$	3.16	6.31	3.14	4.27	3.29	4.07	5.55	4.59
$V_{DS} = 7 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 125^\circ\text{C}$								
$V_{GS} = 1.8 \text{ V}$	3.23	5.44	3.25	5.64	5.57	6.26	4.40	4.46
$V_{DS} = 48 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 125^\circ\text{C}$								
$V_{GS} = -1.4 \text{ V}$	6.61	5.66	5.12	6.24	4.74	6.22	9.53	5.27
$V_{DS} = 7 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 175^\circ\text{C}$								
$V_{GS} = -1.3 \text{ V}$	5.23	6.12	7.37	5.98	5.48	4.22	4.29	8.48
$V_{DS} = 28 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 175^\circ\text{C}$								
$V_{GS} = -1.8 \text{ V}$	5.67	4.82	5.26	6.29	6.57	5.99	7.40	4.42
$V_{DS} = 48 \text{ V}$	$e-4$	$e-4$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 175^\circ\text{C}$								

complementarily Karush-Kuhn Tucker (KKT) conditions in Equations (14)–(15) [32]. The optimal solution can then be found using sequential minimal optimization (SMO) algorithm. So, finally the predicted value can be expressed by (16).

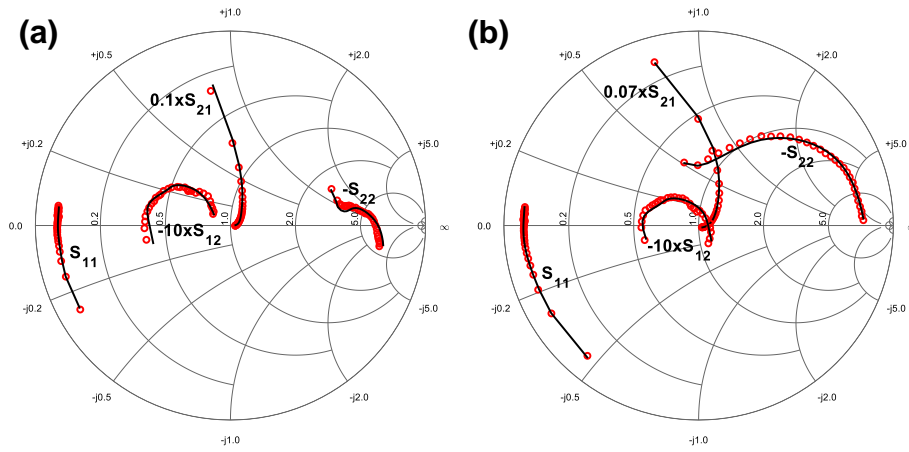


FIGURE 9 The measured and simulated S-parameters at (a) $V_{GS} = 0$ V, $V_{DS} = 7$ V and $T = 25$ C; (b) $V_{GS} = -0.8$ V; $V_{DS} = 28$ V and $T = 25$ C using ANN model

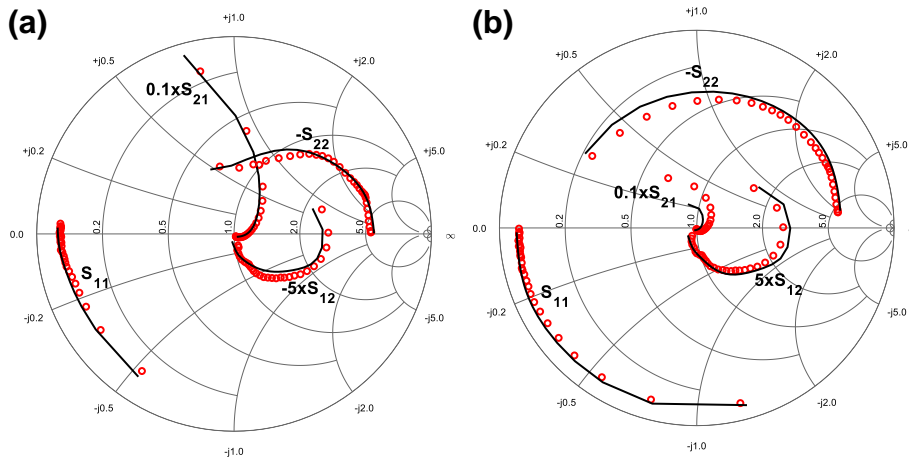


FIGURE 10 The measured and simulated S-parameters at (a) $V_{GS} = -1.4$ V, $V_{DS} = 7$ V, and $T = 175$ C; (b) $V_{GS} = -1.7$ V; $V_{DS} = 28$ V, and $T = 175$ C using ANN model

TABLE 4 MSE for testing set at randomly selected bias conditions and ambient temperatures (ANN Model with PSO initialisation)

Bias and temperature	Re S_{11}	Img S_{11}	Re S_{21}	Img S_{21}	Re S_{12}	Img S_{12}	Re S_{22}	Img S_{22}
$V_{GS} = 1.6$ V	1.37	3.21	2.25	1.22	2.63	2.52	1.71	2.25
$V_{DS} = 7$ V	e-5	e-5	e-3	e-3	e-6	e-6	e-5	e-5
$T = 25^\circ\text{C}$								
$V_{GS} = -1$ V	2.79	2.46	1.52	2.84	2.14	3.02	1.42	3.25
$V_{DS} = 28$ V	e-5	e-5	e-3	e-3	e-6	e-6	e-5	e-5
$T = 100^\circ\text{C}$								
$V_{GS} = -1.5$ V	1.15	3.88	1.73	2.52	3.18	2.41	1.55	2.46
$V_{DS} = 28$ V	e-5	e-5	e-3	e-3	e-6	e-6	e-5	e-5
$T = 150^\circ\text{C}$								

Abbreviation: PSO, particle swarm optimization.

$$\text{Maximize} \left(\sum_{i=1}^n y_i (\alpha_i^* + \alpha_i) - \varepsilon (\alpha_i^* + \alpha_i) - \frac{1}{2} \sum_{i,j=1}^n (\alpha_i^* - \alpha_i) (\alpha_j^* - \alpha_j) K(x_i, x_j) \right) \quad (13)$$

So that:

$$\sum_{i=1}^n y_i (\alpha_i - \alpha_i^*) = 0, \text{ where } \alpha_i, \alpha_i^* \in [0, C] \quad (14)$$

$$\alpha_i (\varepsilon + \xi_i - y_i + f(x_i)) = 0; \xi_i (C - \alpha_i) = 0 \quad (15a)$$

$$\alpha_i^* (\varepsilon + \xi_i^* - y_i + f(x_i)) = 0; \xi_i^* (C - \alpha_i^*) = 0; \forall i \quad (15b)$$

$$\alpha_i \alpha_i^* = 0; \eta_i \eta_i^* = 0. \tag{15c}$$

$$f(x) = \sum_{j=1}^n (\alpha_j^* - \alpha_j) K(x, x_j) + b, \tag{16}$$

where x_j is the support vector, n is the number of support vectors, the terms α_j^* and α_j in bracket are weight coefficients of support vectors, b is the bias and $K(\cdot)$ is the kernel function.

TABLE 5 MSE for validation using independent set at randomly selected different ambient temperatures (first two: interpolation case, last three: extrapolation case) (ANN Model with PSO initialisation)

Bias and temperature	Re S_{11}	Img S_{11}	Re S_{21}	Img S_{21}	Re S_{12}	Img S_{12}	Re S_{22}	Img S_{22}
$V_{GS} = -1.4$ V	4.98	4.04	5.99	4.17	4.28	6.33	3.98	6.35
$V_{DS} = 48$ V	e-5	e-5	e-3	e-3	e-6	e-6	e-5	e-5
$T = 50^\circ$ C								
$V_{GS} = -1.8$ V	5.47	5.55	4.27	4.59	4.77	4.26	4.36	4.16
$V_{DS} = 48$ V	e-5	e-5	e-3	e-3	e-6	e-6	e-5	e-5
$T = 125^\circ$ C								
$V_{GS} = -1.3$ V	6.67	5.58	5.14	5.66	5.18	5.36	7.28	6.18
$V_{DS} = 28$ V	e-5	e-5	e-3	e-3	e-6	e-6	e-5	e-5
$T = 175^\circ$ C								
$V_{GS} = -1.7$ V	5.51	5.56	5.77	4.57	4.28	7.89	4.33	5.29
$V_{DS} = 28$ V	e-5	e-5	e-3	e-3	e-6	e-6	e-5	e-5
$T = 175^\circ$ C								
$V_{GS} = -1.8$ V	4.29	7.80	5.17	6.18	5.89	5.77	5.19	7.59
$V_{DS} = 48$ V	e-5	e-5	e-3	e-3	e-6	e-6	e-5	e-5
$T = 175^\circ$ C								

Abbreviation: PSO, particle swarm optimization.

4.2 | Temperature-dependent modelling using SVR

The SVR model is built for the measured S-parameters (real and imaginary parts), as can be seen in Figure 7, under the listed bias conditions in Table 1 and over the frequency range of 0.1–20 GHz. Once again, the same listed measurements at ambient temperatures of 25, 75, 100, and 150 C are used to train and test the model. To check the model's generalisation capability, the model is validated using two independent datasets (which are not used for building the model). The first dataset consists of the same bias conditions and frequencies at $T = 50$ C and $T = 125$ C, while the other dataset is at $T = 175$ C. The training data is standardised/normalised with respect to mean and standard deviation. The model utilises Gaussian kernel function and the SMO solver in this case. The kernels convert the data into high dimension, which is typically more representative. However, the high dimension implies higher computational cost and needs more training data.

4.3 | SVR-based improved model

One of the main reasons for the deterioration of the performance of the SVR model is the improper values of the parameters. For a given SVR model, there are six parameters which primarily affect the performance of the model. These parameters are called hyper parameters. Proper tuning of these parameters for a given problem plays a significant role in building the model. In the SVR, the central parameters are the box constraint or regularisation parameter (C), tube size of ϵ -insensitive loss function (ϵ), kernel scale, kernel function, polynomial order and presence and absence of the standardisation of the measured set. The tube size of ϵ -insensitive loss function (ϵ) defines the margin of tolerance. Higher and lower values of ϵ , controls the number of the

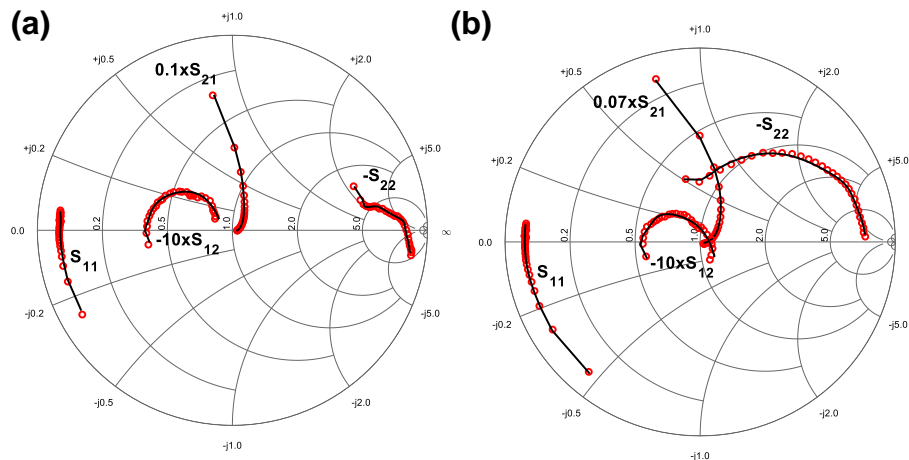


FIGURE 11 The measured and simulated S-parameters at (a) $V_{GS} = 0$ V, $V_{DS} = 7$ V, and $T = 25$ C; (b) $V_{GS} = -0.8$ V, $V_{DS} = 28$ V, and $T = 25$ C for the ANN model with PSO initialisation

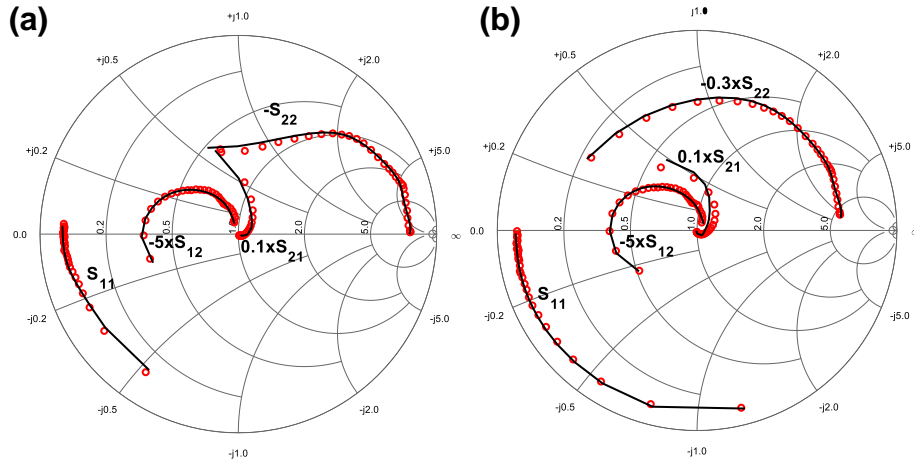


FIGURE 12 The measured and simulated S-parameters at (a) $V_{GS} = -1.4$ V, $V_{DS} = 7$ V, and $T = 175$ C; (b) $V_{GS} = -1.7$ V, $V_{DS} = 28$ V, and $T = 175$ C for the ANN model with PSO initialisation

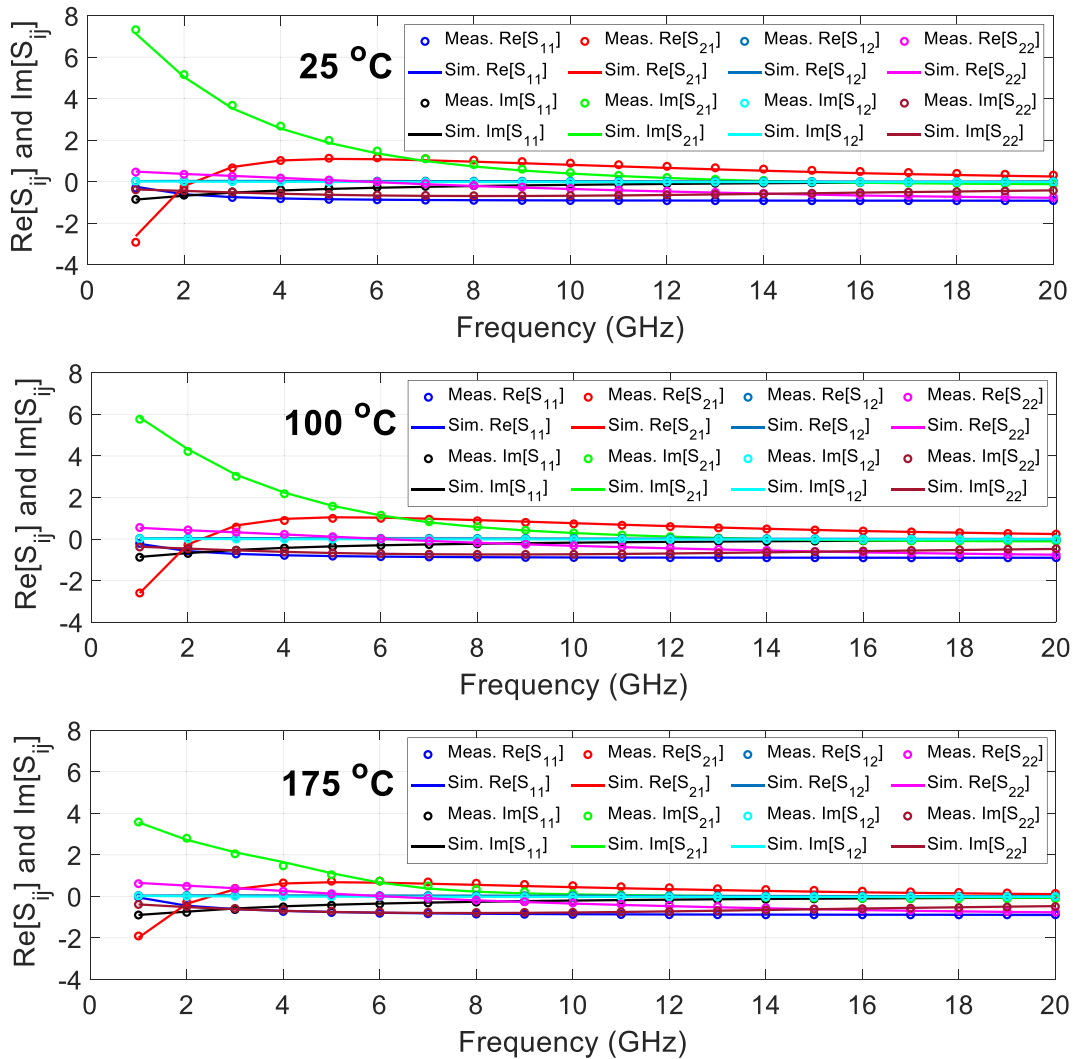


FIGURE 13 Modelled and measured S-parameters at $V_{GS} = -1.5$ V, $V_{DS} = 48$ V, and $T = 25, 100,$ and 175°C using particle swarm optimization-ANN model

TABLE 6 MSE for testing set at randomly selected bias conditions and ambient temperatures (SVR model)

Bias and temperature	Re S_{11}	Img S_{11}	Re S_{21}	Img S_{21}	Re S_{12}	Img S_{12}	Re S_{22}	Img S_{22}
$V_{GS} = 1.6$ V	2.942	4.98	1.16	1.14	4.46	4.17	2.19	2.56
$V_{DS} = 7$ V	$e-3$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 25^\circ\text{C}$								
$V_{GS} = -1$ V	3.59	5.36	1.27	1.45	3.28	4.07	2.54	3.69
$V_{DS} = 28$ V	$e-3$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 100^\circ\text{C}$								
$V_{GS} = -1.5$ V	1.55	4.84	2.14	1.45	2.57	3.96	4.49	2.66
$V_{DS} = 28$ V	$e-3$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 150^\circ\text{C}$								

Abbreviation: SVR, support vector regression.

support vectors in the decision boundary, which directly controls the optimal values of weights and biases and consequently controls the optimal hyperplane for the problem. The term C outlines the imposed penalty on the observations which go beyond the ϵ range. It controls the underfitting and overfitting of the model. The Kernel function plays a significant part in solving complex non-linear problems using linear methods. The intuition is to convert the low-dimensional dataset to high-dimensional dataset and perform dot products. Pre-processing and normalisation of the data is also a very critical step in getting optimal results. In the proposed model, Gaussian kernel function is used. The hyper parameters are initialised with random values. C and sigma (σ) are set to be within the range of $[1e-3 \ 1e3]$. The range of error (ϵ) is set between $[1e-6 \ 1e3]$.

The process is repeated many times until we get the optimal results. The flow chart of the proposed model is summarised in Figure 8. The same PSO parameters are chosen as discussed in Section 3. All steps are the same, except here only three parameters are used with defined upper and lower boundaries. These optimised particle values obtained using PSO are then utilised to get the optimal model.

5 | ANN-BASED MODEL VALIDATION

5.1 | Evaluation of simple model

The proposed ANN model is implemented using MATLAB. Subsequently, the mean squared error (MSE) is calculated using (3) for the training/validation and testing set (interpolation and extrapolation sets). The minimum errors using LM-BP algorithm for the two sets for randomly selected bias and temperature values are listed in Table 2. Figures 9 and 10 compare the modelled and measured S-parameters, randomly selected from training/validation and testing sets, at two extreme ambient temperatures under distinct bias conditions. An

TABLE 7 MSE for independent validation set at different ambient temperature (first three: interpolation case, last three: extrapolation case) (SVR Model)

Bias and temperature	Re S_{11}	Img S_{11}	Re S_{21}	Img S_{21}	Re S_{12}	Img S_{12}	Re S_{22}	Img S_{22}
$V_{GS} = -1.4$ V	5.32	7.12	4.19	3.47	4.32	3.59	6.91	3.66
$V_{DS} = 48$ V	$e-3$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 50^\circ\text{C}$								
$V_{GS} = -1.8$ V	4.88	6.77	4.52	3.64	5.64	6.51	5.51	5.27
$V_{DS} = 7$ V	$e-3$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 12^\circ\text{C}$								
$V_{GS} = -1.5$ V	4.19	3.95	4.97	4.26	4.19	7.58	6.29	6.16
$V_{DS} = 48$ V	$e-3$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-4$	$e-4$
$T = 125^\circ\text{C}$								
$V_{GS} = -0.6$ V	4.48	5.55	5.19	6.12	4.25	3.91	3.41	7.37
$V_{DS} = 7$ V	$e-2$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-2$	$e-3$
$T = 175^\circ\text{C}$								
$V_{GS} = -1.7$ V	7.02	7.21	6.99	8.58	5.23	5.31	4.51	5.17
$V_{DS} = 28$ V	$e-2$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-2$	$e-3$
$T = 175^\circ\text{C}$								
$V_{GS} = -1.8$ V	5.18	4.33	5.78	5.77	5.84	2.58	5.09	4.24
$V_{DS} = 48$ V	$e-2$	$e-3$	$e-2$	$e-2$	$e-5$	$e-5$	$e-2$	$e-3$
$T = 175^\circ\text{C}$								

Abbreviation: SVR, support vector regression.

excellent agreement between the measurements and simulation is obtained over all considered bias voltages and temperatures. To reiterate, the prediction capability of the model has been evaluated by independent measured data sets than the one used for model development (see Table 3). It can thus be inferred from the validation, the developed ANN model provides a very good accuracy for both interpolation and extrapolation cases.

5.2 | Evaluation of improved ANN model

The improved model makes use of the PSO for initialisation of optimal weights and biases corresponding to global minimum. Once again, the same conditions are taken to check the model's ability and robustness over entire frequency range. It can be inferred from the results in Tables 4 and 5, the MSE for randomly selected ambient temperature and bias conditions for ANN model with PSO initialisation, is significantly superior as compared to the accuracy achieved by simple ANN model. This enhanced accuracy can be attributed to the improved initial values. A visual depiction in Figures 11–12 for the extreme ambient temperatures shows an excellent consonance between the modelled and measured S-parameters. This demonstrates the effectiveness of the PSO in the ANN-based model development process. It is thus safe to convey that the use of PSO is very

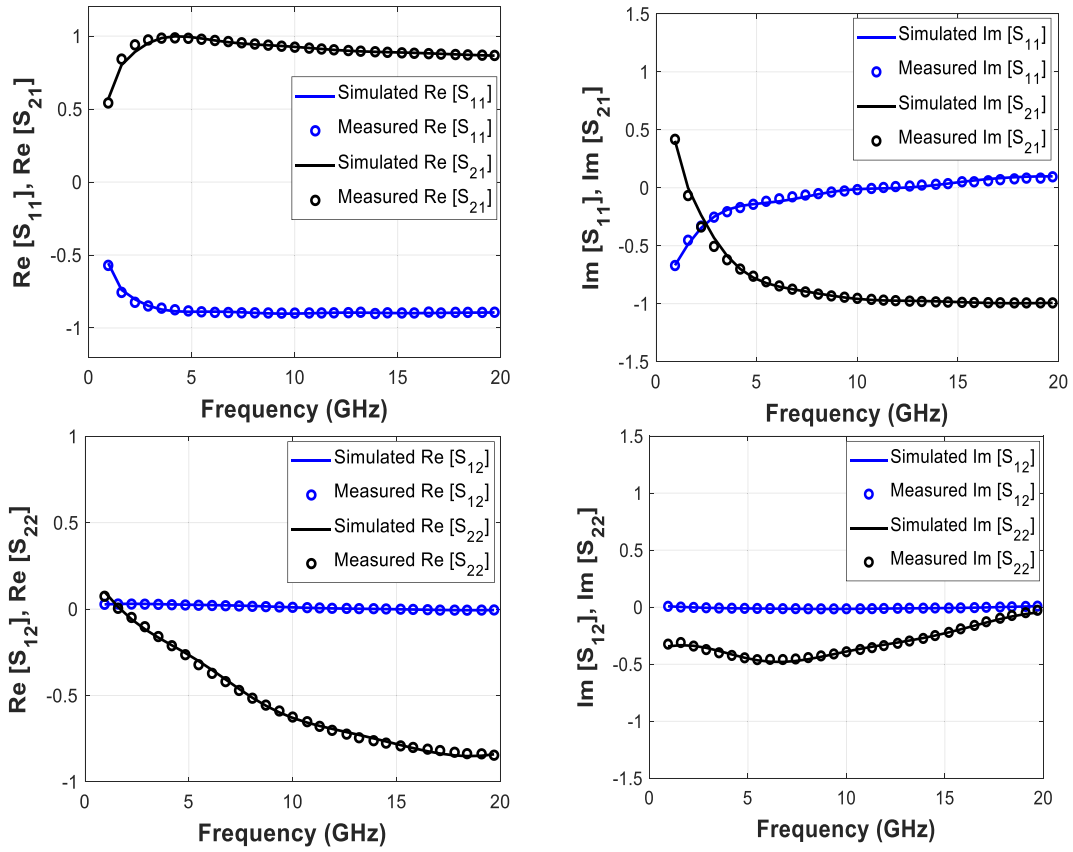


FIGURE 14 Comparison of modelled and measured S-parameters at $V_{GS} = -0.8$ V, $V_{DS} = 28$ V, and $T = 25$ C using support vector regression

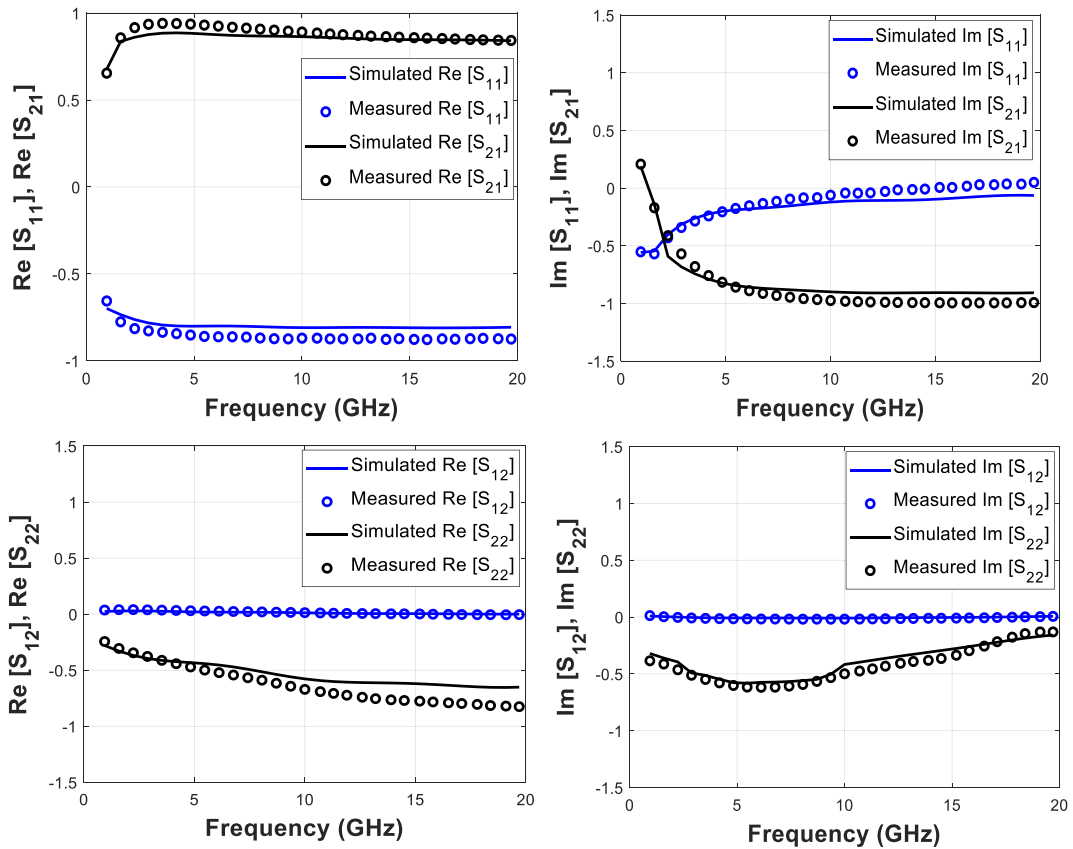


FIGURE 15 Comparison of modelled and measured S-parameters at $V_{GS} = -1$ V, $V_{DS} = 28$ V, and $T = 175$ C using support vector regression

TABLE 8 The optimal values of hyper parameters using PSO

Response/parameter	Kernel function	Box constraint (C)	Epsilon (ϵ)	Sigma (σ)
Re (S_{11})	Gaussian	2.53	2.68e-4	1.04
Img (S_{11})	Gaussian	1.99	2.49e-4	0.98
Re (S_{21})	Gaussian	3.14	5.73e-4	1
Img (S_{21})	Gaussian	3.23	6.84e-4	1.03
Re (S_{12})	Gaussian	1.57	1.92e-5	1.15
Img (S_{12})	Gaussian	1.98	1.91e-5	2.38
Re (S_{22})	Gaussian	2.91	3.24e-4	0.91
Img (S_{22})	Gaussian	3.89	2.33e-4	1.05

Abbreviation: PSO, particle swarm optimization.

TABLE 9 MSE for testing set at randomly selected bias conditions and ambient temperatures. (PSO-based SVR Model)

Bias and temperature	Re S_{11}	Img S_{11}	Re S_{21}	Img S_{21}	Re S_{12}	Img S_{12}	Re S_{22}	Img S_{22}
$V_{GS} = 1.6$ V	1.56	1.22	1.16	2.32	2.25	2.31	1.29	1.35
$V_{DS} = 7$ V	e-4	e-4	e-3	e-3	e-6	e-6	e-4	e-4
$T = 25^\circ\text{C}$								
$V_{GS} = -1$ V	2.39	1.35	1.98	1.93	2.25	2.88	1.93	1.53
$V_{DS} = 28$ V	e-4	e-4	e-3	e-3	e-6	e-6	e-4	e-4
$T = 100^\circ\text{C}$								
$V_{GS} = -1.5$ V	2.25	1.84	1.47	1.81	1.64	3.25	1.52	2.61
$V_{DS} = 28$ V	e-4	e-4	e-3	e-3	e-6	e-6	e-4	e-4
$T = 150^\circ\text{C}$								

Abbreviations: PSO, particle swarm optimization; SVR, support vector regression.

TABLE 10 MSE for validation independent set at different ambient temperature randomly (first two: interpolation case, last two: extrapolation case) (PSO-based SVR Model)

Bias and temperature	Re S_{11}	Img S_{11}	Re S_{21}	Img S_{21}	Re S_{12}	Img S_{12}	Re S_{22}	Img S_{22}
$V_{GS} = -1.4$ V	3.48	4.04	3.81	4.18	4.73	5.21	2.47	4.46
$V_{DS} = 48$ V	e-4	e-4	e-3	e-3	e-6	e-6	e-4	e-4
$T = 50^\circ\text{C}$								
$V_{GS} = -1.8$ V	4.36	2.38	3.99	3.89	5.51	5.79	4.98	2.66
$V_{DS} = 7$ V	e-4	e-4	e-3	e-3	e-6	e-6	e-4	e-4
$T = 125^\circ\text{C}$								
$V_{GS} = -0.6$ V	6.12	6.38	4.77	5.44	5.68	5.58	7.58	5.53
$V_{DS} = 7$ V	e-4	e-4	e-3	e-3	e-6	e-6	e-4	e-4
$T = 175^\circ\text{C}$								
$V_{GS} = -1.8$ V	6.57	5.84	6.98	7.25	7.23	7.34	8.44	7.55
$V_{DS} = 48$ V	e-4	e-4	e-3	e-3	e-6	e-6	e-4	e-4
$T = 175^\circ\text{C}$								

Abbreviations: PSO, particle swarm optimization; SVR, support vector regression.

beneficial as it significantly enhances the accuracy of the model. Figure 13 shows also the model simulation at a fixed active bias condition and different ambient temperature of 25, 100 and 175 C. This also validates the extrapolation capability of the model.

6 | SVR-BASED MODEL VALIDATION

6.1 | Evaluation of simple SVR model

The proposed SVR model is implemented using MATLAB. The predicted MSE is calculated for the training/validation and testing set (interpolation and extrapolation sets) using LM algorithm. The MSEs for these scenarios, for randomly selected temperatures and bias conditions, are given in Tables 6 and 7. Once again, modelled and measured S-parameters at different bias conditions and extreme ambient temperatures are compared in Figures 14 and 15. An excellent agreement between the measured and modelled values for both the imaginary and real parts of S-parameters exist and this proves the accuracy, effectiveness, and robustness of the developed

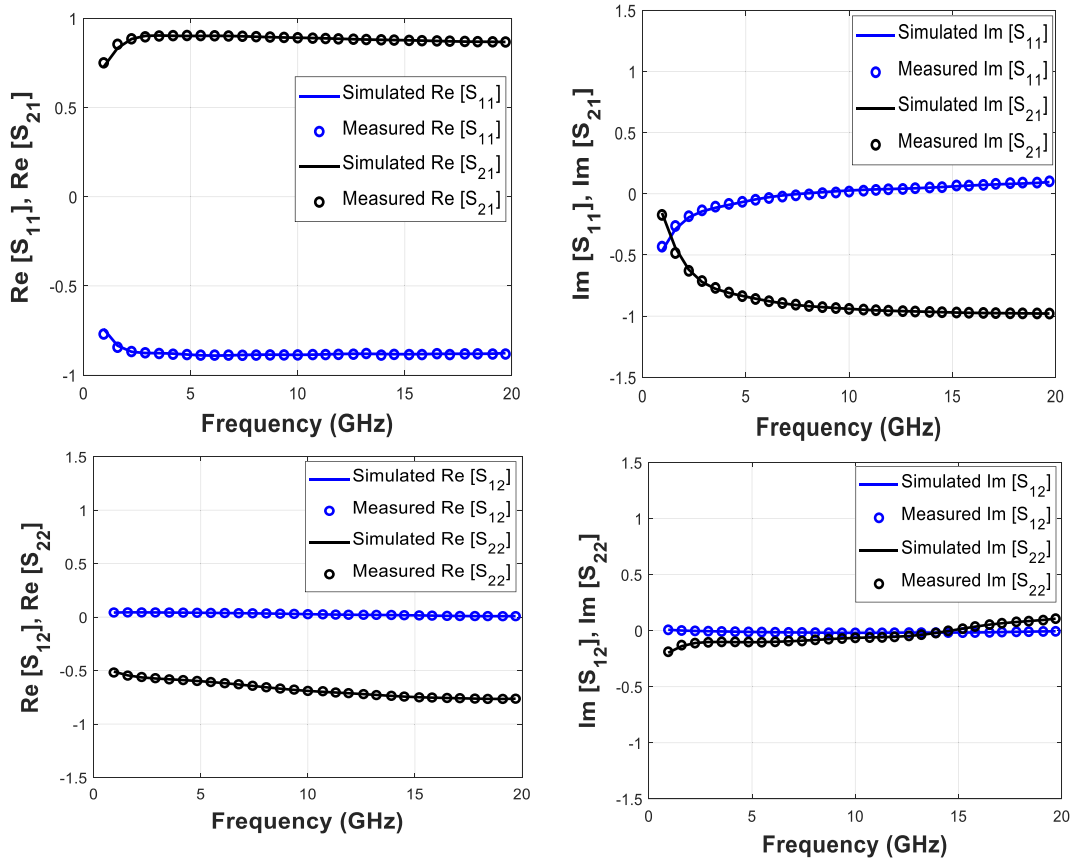


FIGURE 16 Comparison of S-parameters at $V_{GS} = 0$ V, $V_{DS} = 7$ V, and $T = 25$ C for SVR model with parameter setting using PSO

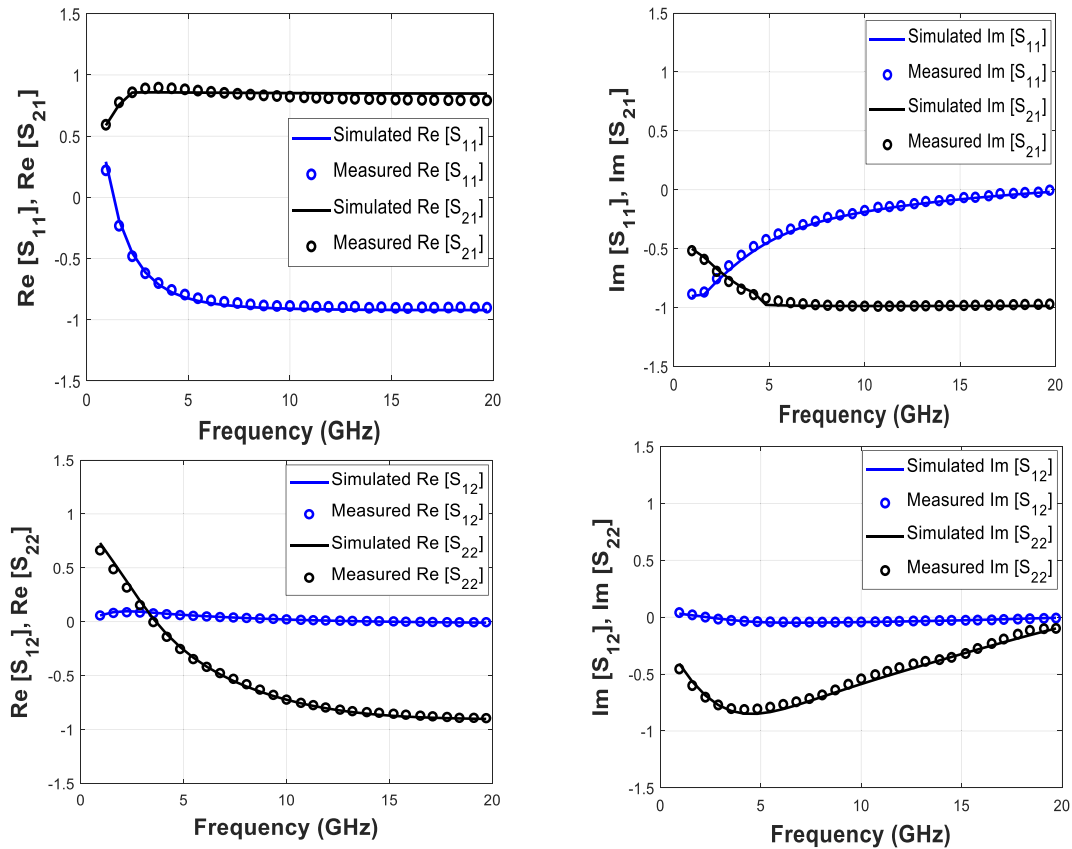


FIGURE 17 Comparison of S-parameters at $V_{GS} = -1.7$ V, $V_{DS} = 28$ V, and $T = 175$ C for SVR model with parameter setting using particle swarm optimization

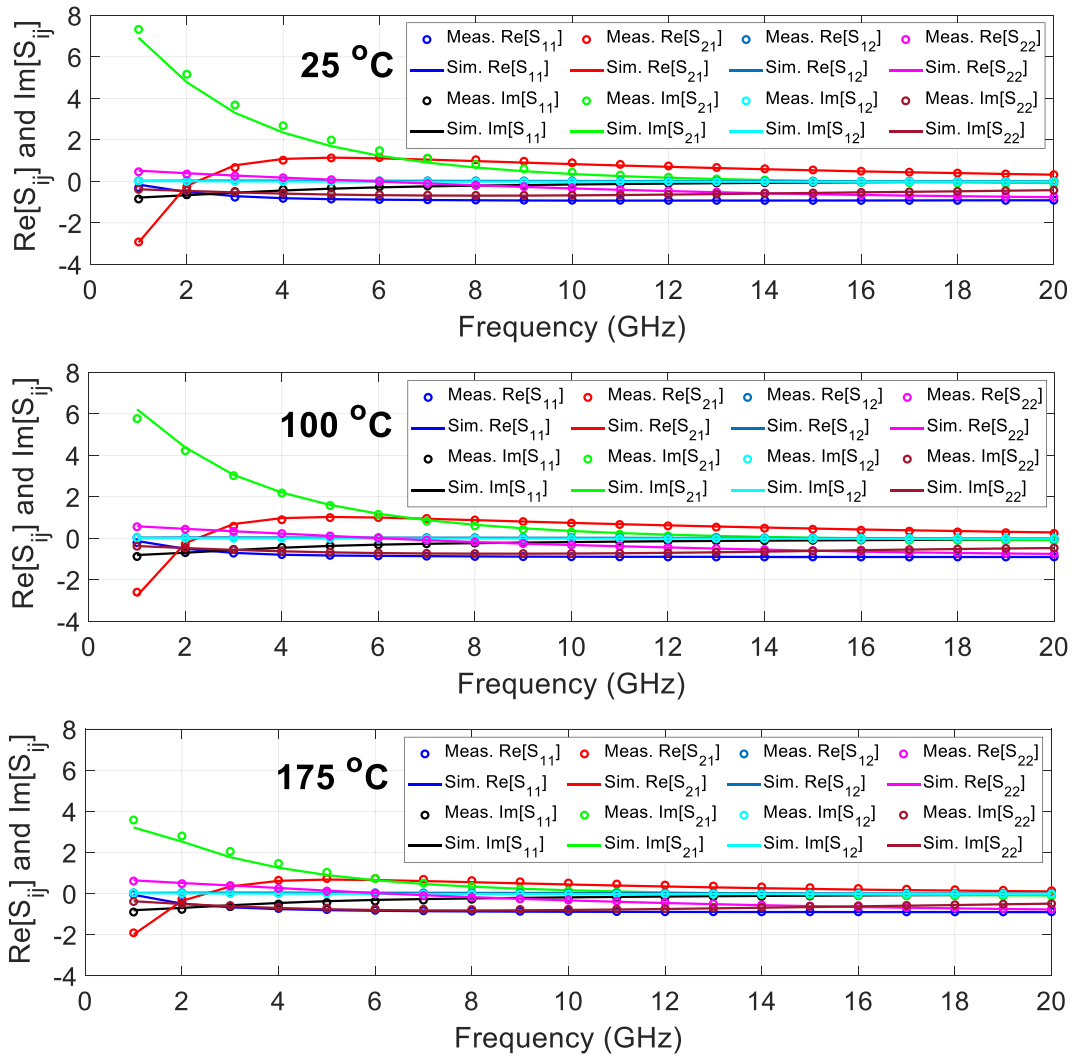


FIGURE 18 Modelled and measured S-parameters at $V_{GS} = -1.5$ V, $V_{DS} = 48$ V, and $T = 25, 100,$ and 175°C using particle swarm optimization-support vector regression model

SVR-based model for wide range of bias voltages and temperatures over the entire frequency range.

6.2 | Evaluation of improved SVR model

The parameter tuning plays an important role to improve the performance. The improved model utilises the PSO to tune the box constraint, Epsilon, and sigma values for the Gaussian kernel. The final optimised values for these parameters are listed in Table 8. Subsequently, the predicted MSE is calculated for the training/validation, and testing set (interpolation and extrapolation sets) using LM algorithm. The MSEs for some randomly selected bias and temperature conditions are given in Tables 9 and 10. The corresponding Figures 16 and 17 compare the modelled and measured S-parameters at randomly selected bias conditions and extreme ambient temperatures.

It can be observed from results that an excellent agreement between the modelled and measured values exist for the

considered bias voltages and temperatures. It is also identified that the careful selection of the hyper parameters by PSO significantly improves the SVR-based model's accuracy which can be inferred from all the achieved MSE. It can also be seen that the extrapolation capability of the SVR model is also improved by optimising the model parameters using the PSO. Figure 18 shows also the model simulation at a fixed active bias condition and different ambient temperature of 25,100 and 175 C. This apparently validates the extrapolation capability of the model.

Overall, it can be inferred that the training and interpolation capability of both ANN and SVR is almost the same but ANN has shown better extrapolation capability as opposed to SVR for this particular problem. In brief, it can be said that the ANN achieves better prediction capability for this problem. It is well known that the performance of SVR decreases with the increase in the size of the training data as it requires more memory to store the kernel and gram matrix. It is also imperative to note that the training time for SVR is more than

the training time for ANN. On the other hand, the ANN model needs to be trained many times until satisfactory results are obtained. This can be attributed to the initial guess dependency of the LM-BP and its higher chance of stacking in a local minimum. However, the introduction of PSO in the ANN model development effectively solves the initial guess dependency of LM-BP. On the contrary, the SVR does not face such a problem. Actually, the SVR-based model's performance can be greatly improved by careful tuning and optimisation of its hyper-parameters using PSO. The optimisation of hyper-parameter tuning substantially improves the overall performance and the prediction capability of the SVR-based model. Finally, it is safe to convey that the ANN model with initialisation using PSO performs the best when compared to the other models proposed in this paper owing to less training time and improved accuracy.

7 | CONCLUSION

Here, the development of temperature-dependent ANN- and SVR-based models for GaN-on-Si HEMT has been reported. The models have been trained for large set of operating conditions and later tested on two independent sets. First, the simple topology of ANN has been used efficiently to model the device and then tested using two independent sets to validate its interpolation and extrapolation capability. Subsequently, SVR has been used to develop alternative modelling technique for the same device. It makes use of SMO solver to train the model for speedy model development process. It was identified that the SVR-based technique is more robust and training need not be repeated in contrast to the ANN-based technique. It is also worth noting that the SVR takes more training time than the ANN in the case of large dataset. This could be attributed to its extra computation operations due to requirement of more memory space to store the kernel matrix. Most importantly, the extrapolation capability of SVR model is inferior. In addition, the ANN model shows very good prediction capability for the in and out of range data. The use of PSO algorithm in the model development process improves the performance of both the ANN- and SVR-based model. In summary, the ANN model with PSO-based initialisation outperforms other models in term of prediction capability, training time, and accuracy.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the support from the University of Sharjah, Sharjah, United Arab Emirates.

ORCID

Anwar Jarndal  <https://orcid.org/0000-0002-1873-2088>

REFERENCES

- Pengelly, R.S., et al.: A review of GaN on SiC high electron-mobility power transistors and MMICs. *IEEE Trans. Microw. Theor. Techn.* 60(6), 1764–1783 (2012)
- Zeng, F., et al.: A Comprehensive review of recent Progress on GaN high electron mobility transistors: devices, fabrication and reliability. *Electronics*. 7, 377–397 (2018)
- Ghannouchi, F.M., Hashmi, M.S.: Load-pull techniques and their Applications in power amplifiers design. *IEEE Bipolar/BiCMOS Circuits and Technology Meeting*, Atlanta 1-5 (2011)
- Colangeli, S., et al.: GaN-based robust low-noise amplifiers. *IEEE Trans. Electron. Dev.* 60(10), 3238–3248 (2013)
- Jarndal, A., et al.: Large-signal modelling methodology for GaN HEMTs for RF Switching-Mode power amplifiers design. *Intl. J. RF Microw. Comp. Aided Eng.* 21(1), 45–51 (2010)
- Miccoli, C., et al.: Trapping and thermal effects analysis for AlGaIn/GaN HEMTs by means of TCAD simulations. *IEEE Electron. Dev. Lett.* 34(9), 1121–1123 (2013)
- Roff, C., et al.: Analysis of DC-RF dispersion in AlGaIn/GaN HFETs using RF waveform engineering. *IEEE Trans. Electron. Dev.* 56(1), 13–19 (2009)
- Kuzmik, J., et al.: Self-heating in GaN transistors designed for high-power operation. *IEEE Trans. Electron Dev.* 61(10), 3429–3434 (2014)
- Jarndal, A., Arivazhagan, L., Nirmal, D.: On the performance of GaN-on-Silicon, Silicon-Carbide, and Diamond substrates. *Int. J. RF Microw. Comput. Aided Eng.* 30(6), e22196 (2020). <https://doi.org/10.1002/mmce.22196>
- King, J.B., Brazil T.J.: Nonlinear electrothermal GaN HEMT model Applied to high-efficiency power amplifier design. *IEEE Trans. Microw. Theor. Techn.* 61(1), 444–454 (2013)
- Jarndal, A.: On neural networks based electrothermal modelling of GaN devices. *IEEE Access*. 7, 94205–94214 (2019)
- Huang, A.D., et al.: An artificial neural network-based electrothermal model for GaN HEMTs with Dynamic Trapping effects consideration. *IEEE Trans. Microw. Theor. Techn.* 64(8), 2519–2528 (2016)
- Marinkovic, Z., et al.: A review on the artificial neural network applications for small-signal modelling of microwave FETs. *Int. J. Numer. Modell Electron. Networks Dev.* 33(3), e2668 (2019)
- Khusro, A., et al.: Small signal behavioral modeling technique of GaN high electron mobility transistor using artificial neural network: an accurate, fast, and reliable approach. *Intl. J. RF Microw. Comput. Aided Eng.* 30, e22112 (2020)
- Jarndal, A.: Neural network electrothermal modelling approach for microwave active devices. *Int. J. RF Microw. Comput. Aided Eng.* 29(9), e21764 (2019)
- Barron, I., et al.: Hessian with mini-batches for electrical demand prediction. *Appl. Sci.* 10(6) 1–14 (2020)
- Zhao, Z., et al.: Overview of microwave device modelling techniques based on machine learning. *Int. J. Adv. Comput. Technol.* 5(9), 299–306 (2013)
- Anguili, G., Cacciola, M., Versaci, M.: Microwave devices and antennas modelling by support vector regression machines. *IEEE Trans. Magn.* 43(4), 1589–1592 (2007)
- Xu, Y., et al.: Modelling of SiC MESFETs by using support vector Machine regression. *J. Electromagn. Waves Appl.* 21(11), 1489–1498 (2007)
- Khusro, A., Hashmi, M.S., Ansari, A.Q.: Exploring support vector regression for modelling of GaN HEMT. *IEEE MTT-S International Microwave and RF Conference (IMaRC)*, Kolkata (2019)
- Khusro, A., Hashmi, M.S., Ansari, A.Q.: Enabling the development of accurate intrinsic parameter extraction model for GaN HEMT using support vector regression (SVR). *IET Microw. Antenna. Propag.* 13(9), 1457–1466 (2019)
- Cai, J., et al.: Dynamic behavioural modelling of RF power amplifier based on time-Delay support vector regression. *IEEE Trans. Microw. Theor. Techn.* 67(2), 533–543 (2019)
- Xu, J., et al.: Augmented time-Delay twin support vector regression-based behavioral modeling for digital predistortion of RF power amplifier. 2019 *IEEE Access*. 7, 59832–59843 (2019)
- Hussein, A.S., Jarndal, A.H.: Reliable particle-swarm-optimization based parameter extraction method applied to GaN HEMTs. In: *Mediterranean Microwave Symposium*. Abu Dhabi pp 1-4 (2016)

25. Hussein, A.S., Jarndal, A.H.: Reliable Hybrid small-signal modelling of GaN HEMTs based on particle-swarm-optimization. *IEEE Trans. Comput.-Aided Des. Integr. Circuits Syst.* 37(9), 1816–1824 (2018)
26. Zhao, W., et al.: A Novel Hybrid method of parameters tuning in support vector regression for reliability prediction: particle swarm optimization Combined with Analytical selection. *IEEE Trans. Reliab.* 65(3), 1393–1405 (2016)
27. Haykin, S.: *Neural Networks: A Comprehensive Foundation*. Macmillan, New York. (1994)
28. Zhang, Q., Gupta, K.: *Neural Networks for RF and Microwave*. Artech House. Norwood (2000)
29. Shawash, J., Selviah, D.R.: Real-time nonlinear parameter estimation using the Levenberg-Marquardt algorithm on field programmable gate arrays. *IEEE Trans. Ind. Electron.* 60(1), 170–176 (2013)
30. Nguyen, D., Wildrow, B.: Improving the Learning Speed of 2-Layer Neural Networks by Choosing Initial Values of the Adaptive Weights. *IEEE IJCNN International Joint Conference on Neural Networks*, San Jose. (1990)
31. Scholkopf, S.: *Support Vector Learning*. OldenbourgVerlag, Munich. (1997)
32. Smola, A., Schölkopf, B.: A tutorial on support vector regression. *Statist. Comput.* 14(3), 199–222 (2003)

How to cite this article: Jarndal A, Husain S, Hashmi M. On temperature-dependent small-signal modelling of GaN HEMTs using artificial neural networks and support vector regression. *IET Microw. Antennas Propag.* 2021;15:937–953. <https://doi.org/10.1049/mia2.12112>