Training Data Pre-Processing for Bias-Dependent Neural Models of Microwave Transistor Scattering Parameters

Zlatica Marinković[1] and Vera Marković[2]

Abstract –**Frequency and bias dependence of scattering parameters of microwave transistors can be successfully modelled by artificial neural networks. But, sharp changes in the frequency dependence of angle of a scattering parameter may result in inappropriate modelling accuracy in the vicinity of the frequency when sharp change occurs. In this paper we are discussing pre-processing of the training data in order to make modelling accuracy better. The proposed approach is illustrated by a suitable example.**

Keywords – **Microwave transistors, scattering parameters, artificial neural networks.**

I. INTRODUCTION

Development of accurate and reliable models of microwave transistors is a very important issue in design of modern communication systems. Scattering (S-) parameters of a microwave transistor are dependent on frequency and on operating conditions. Therefore, it is important to include these dependences into a transistor model. In this paper dependence on frequency and bias condition are considered.

Artificial neural networks (ANNs), widely known as structures with huge modelling and generalization capabilities, [1], have been successfully used for modelling of S-parameters of microwave transistors, [2], [3]. One of the models, a basic one, consists of an ANN trained to predict S-parameters for the given frequency and bias conditions. Quality of the data used for the ANN training directly affects the accuracy of the model. Sharp changes in the frequency dependence of angle of a scattering parameter may result in inappropriate modelling accuracy for frequencies near the frequency when sharp change occurs. Therefore, special attention has to be paid to the training data pre-processing in order to increase modelling accuracy. In [3] a possible solution is given, training data have to be pre-processed, and the angle of each S-parameter has to be expressed in the angular range where sharp change does not exist. This kind of pre-processing requires detailed analyses of the training data, which could make modelling process more complex. Here, we are investigating a possible approach for accuracy improving based on training data pre-processing.

The paper is organized as follows: after the introductory section, a brief theoretical background on microwave transistor S-parameters and on artificial neural networks is given in Sections II and III, respectively. Section IV starts with the description of the basic ANN microwave transistor S-parameters' model, then earlier proposed improvement is discussed and the proposed way of data pre-processing aimed for accuracy improving was presented. The proposed model and modelling procedure are illustrated on a suitable modelling example. Modelling results are given in Section V and main conclusions in Section VI.

[1]Zlatica Marinković is with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: zlatica.marinkovic@elfak.ni.ac.rs

[2] Vera Marković is with the Faculty of Electronic Engineering, Aleksandra Medvedeva 14, 18000 Niš, Serbia, E-mail: vera.markovic@elfak.ni.ac.rs.

II. MICROWAVE TRANSISTOR S-PARAMETERS

Microwave transistors operating under small-signal conditions can be characterized by the scattering parameters (S-parameters) which relate the voltage waves incident on the ports to those reflected from the ports (Fig.1).

The scattering matrix, or [S] matrix, is defined in relation to these incident and reflected voltage waves as, [4]:

$$\begin{bmatrix} V_1^- \\ V_2^- \end{bmatrix} = \begin{bmatrix} S_{11} & S_{12} \\ S_{21} & S_{22} \end{bmatrix} \begin{bmatrix} V_1^+ \\ V_2^+ \end{bmatrix} \tag{1}$$

or in matrix notation

$$[V^-] = [S][V^+]. \tag{2}$$

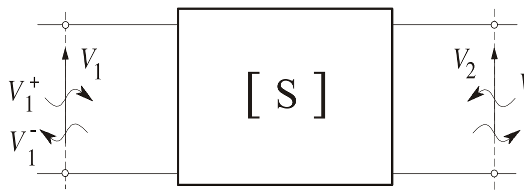

Fig.1. Incident and reflected waves in a two-port network

An element of the [S] matrix can be determined as

$$S_{ij} = \left. \frac{V_i^-}{V_j^+} \right|_{V_k^+ = 0, \text{za } k \neq j} \tag{3}$$

S_{ii} is the reflection coefficient seen looking into the port i when all other ports are terminated in load matches. S_{ij} is the transmission coefficient from port j to port i when all ports are terminated in matched loads.

The S-parameters of microwave transistors are frequency-, temperature- and bias-dependent.

III. ARTIFICIAL NEURAL NETWORKS

A standard multilayer perceptron (MLP) artificial neural network (ANN) consists of neurons grouped into layers: an input layer (layer 0), an output layer (layer N_L) as well as several hidden layers.

Input vectors are presented to the input layer and fed through the network that then yields the output vector. The l-th layer output is:

$$\mathbf{Y}_l = F(\mathbf{W}_l \mathbf{Y}_{l-1} + \mathbf{B}_l) \quad (4)$$

where $\mathbf{Y}_l$ and $\mathbf{Y}_{l-1}$ are outputs of l-th and (l-1)-th layer, respectively, $\mathbf{W}_l$ is a weight matrix between (l-1)-th and l-th layer and $\mathbf{B}_l$ is a bias matrix between (l-1)-th and l-th layer. Function F is an activation function of each neuron and, in our case, is linear for input and output layer and sigmoid for hidden layers:

$$F(u) = 1/(1+e^{-u}) \quad (5)$$

The neural network learns relationship among sets of input-output data (training sets) that are characteristics of the component under consideration. First, input vectors are presented to the input neurons and output vectors are computed. These output vectors are then compared with desired values and errors are computed. Error derivatives are then calculated and summed up for each weight and bias until whole training set has been presented to the network. These error derivatives are then used to update the weights and biases for neurons in the model. The training process proceeds until errors are lower than the prescribed values or until the maximum number of epochs (epoch - the whole training set processing) is reached. Once trained, the network provides fast response for different input vectors, even for those not included in the training set, without additional optimizations.

IV. Bias-dependent ANN models

ANNs have been successfully applied for modeling bias- and frequency dependence of MESFET/HEMT microwave transistor S-parameters, [2], [3]. The basic ANN model of S-parameters consists of an MLP ANN trained to predict transistor S-parameters for bias voltage, bias current and frequency presented at its inputs, Fig. 2. Therefore, the network has three input neurons corresponding to

- bias voltage (dc drain-to-source voltage, V_{ds})
- bias current and (dc drain-to-source current, I_{ds}) and
- frequency f.

The output layer consists of eight neurons corresponding to magnitudes and angles of the scattering parameters.

Number of the hidden layers can be one or two. The network is trained using S- parameters' data referring to certain number of bias points in the operating frequency range. Neural networks with different number of hidden neurons are trained, tested and after their comparison, the network giving the best testing results is chosen to be the bias-dependent model of S-parameters of the modeled device.

Once the model is developed the S-parameters are obtained by calculating the ANN response to the given bias conditions and frequency. Measured values of the device S-parameters are required only for the model development.

As it is said in the introductory section and shown in details in [3], a sharp change in the frequency dependency of the angle of some S-parameter can make modeling of that parameter, and generally ANN accuracy, worse. This can be illustrated using the results given in the Section V in Figs. 6 and 8. Suppose that the dependence of the of S_{22} parameter on frequency is being modeled and that the frequency dependence of its angle has a sharp change from -180° to 180°, as it is in the case shown in Fig. 6. Furthermore, suppose that magnitude of S_{22} parameter is modeled correctly and consider only modeling of the S_{22} angle. The available measured values, represented by full dots, are used for the training process of an ANN, which will be used as a model of this frequency dependence. Since, there are not enough training data in the vicinity of the sharp change, which is often the case, the trained ANN is most likely not to achieve accurate modeling in this region. The ANN continuous response is represented with dotted line. The ANN obtained values for the frequencies corresponding to the reference values are denoted by triangles. It is obvious that satisfying modeling has not been achieved only in the region of the sharp change, moreover, it is very poor. Due to the poor modeling in this region, there is a "circle" in the modeled S_{22} plot in polar diagram, which does not exist in the actual S_{22} plot, Fig. 8.

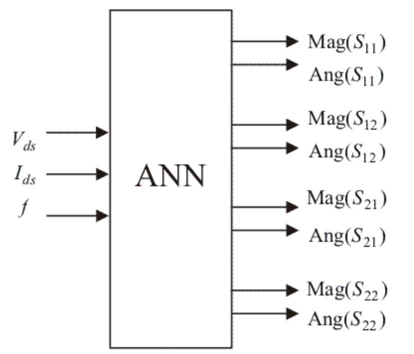

Fig.3. Bias-dependent ANN model

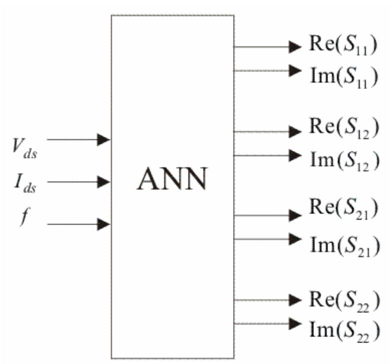

Fig.4. Proposed ANN model

In [3] a possible solution of this problem is proposed: the angle of a S-parameter that has sharp change in the frequency dependence parameter have to be expressed in the range (0°÷360°), and new ANN has to be trained by using the new data. Since the frequency dependence is smooth now, the trained ANN is able to model it with a very good accuracy. Therefore, for each S-parameter, it is recommended to express training and test values of its angle in one of (0°÷360°) or

(-180°÷180°) ranges where there is no sharp change in the frequency depmedence. In that way, modeling accuracy can be significantly increased.

The improvement of modeling proposed in [3] requires at the first a check if the frequency dependence of the angle of each S-parameter is smooth or is with a sharp change and then training data preprocessing if necessary. Here we propose an efficient approach to training data preprocessing.

The basic idea of the proposed approach is to model real and imaginary parts of S-parameters rather than then magnitudes and angles. In that way sharp changes in the frequency dependence can be avoided. As in the case of the basic model, the ANN used in the proposed model has three input and four output neurons, as well as several hidden neurons grouped in one or in two layers. The inputs of the model are the same as the inputs of basic model, while the outputs are real and imaginary parts of S-parameters.

For the model development it is necessary to represent S-parameters in the complex form as Real+j*Imag. The rest of the modeling procedure remains the same.

V. Modeling Example

The proposed approach was applied to Hewlett Packard pHEMT device ATF35143. The data used for the model development was taken from the device manufacturer web-site, [5]. They refer to 9 bias points in (0.5-11) GHz frequency range (23 frequency points per one bias point). The data referring to 8 bias points were used as the ANN training data and the data for the remaining bias points were used as the test data.

Firstly, a basic model trained with the original magnitude-angle represented data was developed. It is a model containing ten neurons in each of two hidden layers. Effects of sharp change of angle frequency dependence occur in the case of S_{11} and S_{22} parameter and are shown in Figs 6-9. In Figs. 6 and 7 the angles of the S-parameters were given and polar plots are given in Figs. 8 and 9. Circles represent reference data, triangles neural model prediction and dotted line continuous response of the ANN. As it is discussed in the previous section, the modeling in the vicinity of sharp change is not satisfactory due to sharp changes of the training data.

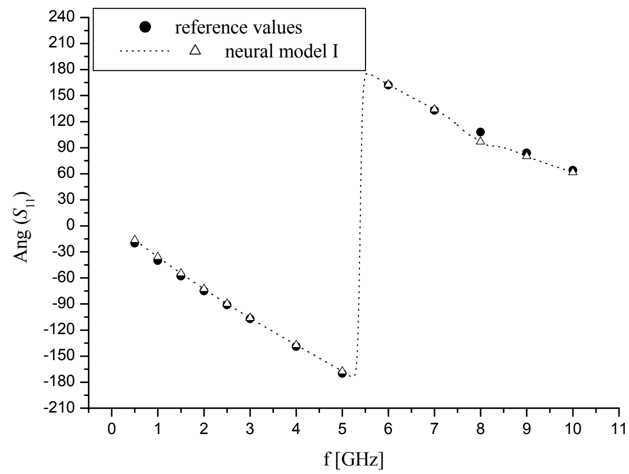

Fig. 5. S_{11} parameter angle

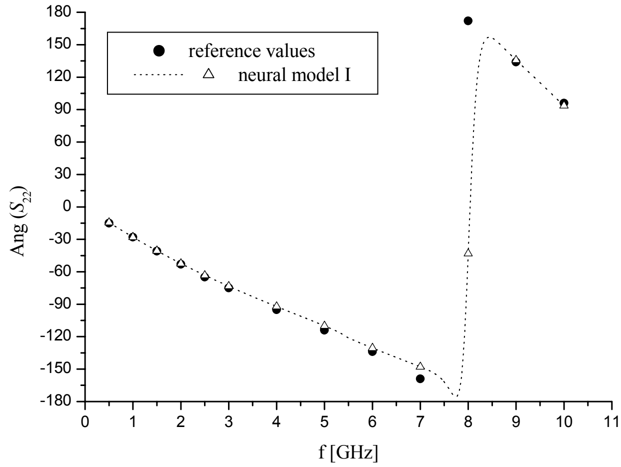

Fig. 6. S_{22} parameter angle

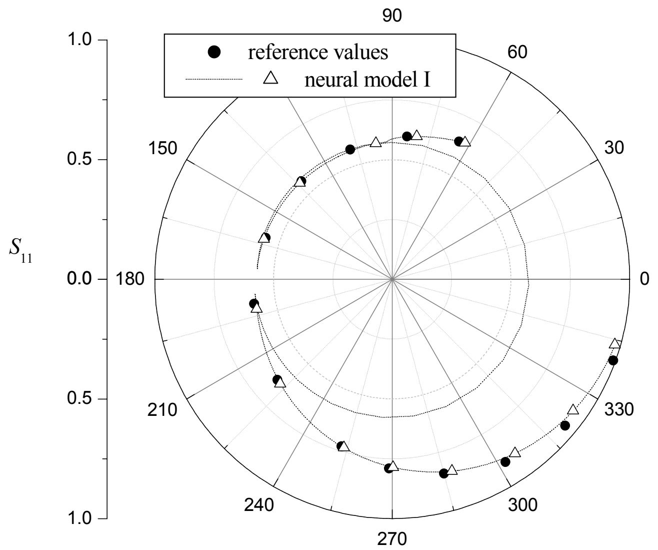

Fig. 7. S_{11} parameter – polar plot

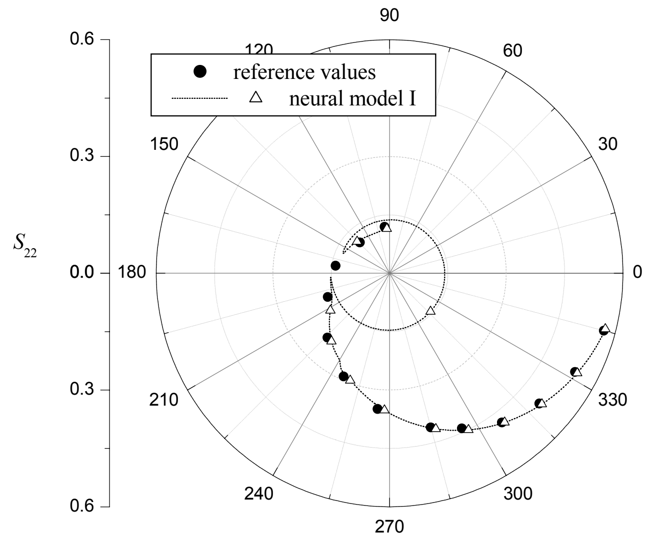

Fig. 8. S_{22} parameter – polar plot

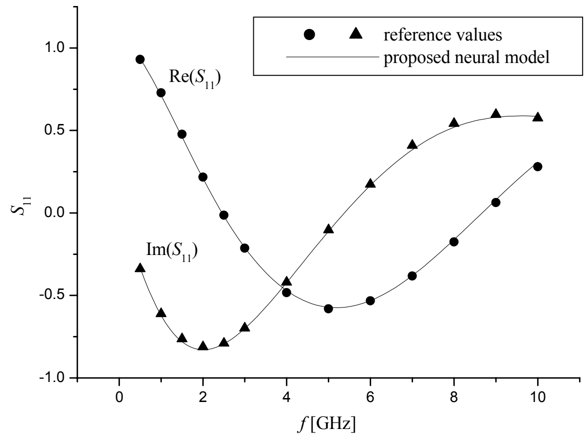

Fig. 10. S_{11} parameter – real and imaginary part

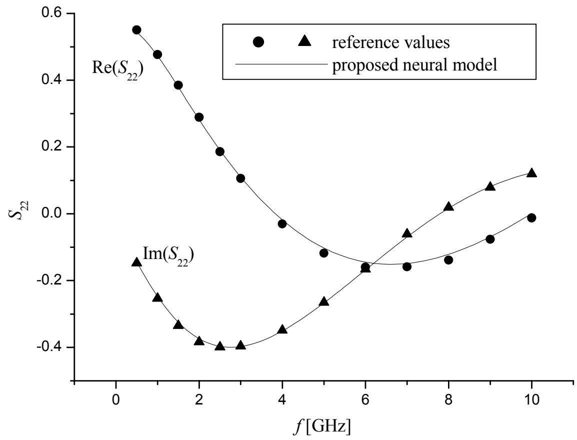

Fig. 11. S_{22} parameter – real and imaginary part

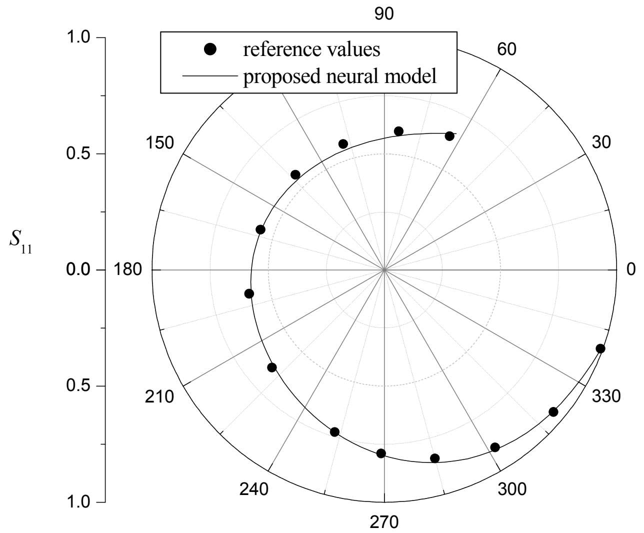

Fig.12. S_{11} parameter – polar plot

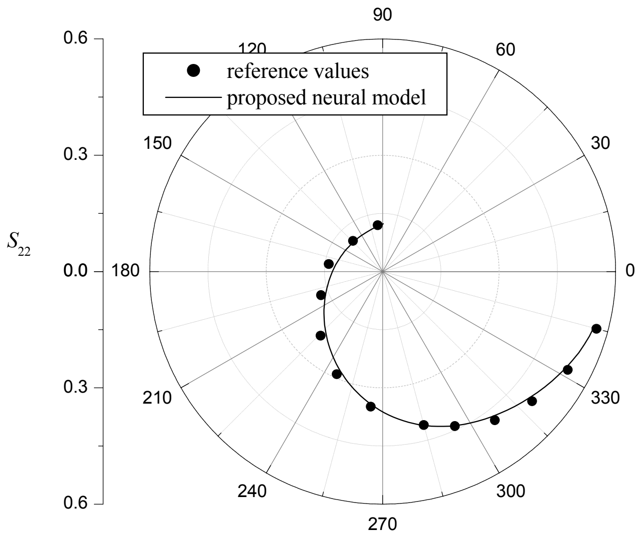

Fig. 13. S_{22} parameter – polar plot

Further, the newly proposed model was developed. S-parameters' data aimed for training purposes were preprocessed, i.e. represented in the form Real+j*Imag, and ANNs were trained. One of the trained ANNs with the best modeling results is a network with two hidden layers with ten neurons in each layer. Results for a bias point not used for the training are given in Figs. 10-13. Figs 10 and 11 refer to the real and imaginary part modeling and Figs 12 and 13 to polar plots of S_{11} and S_{22} parameters. It can be observed that the continuous ANN response (solid line) behaves smoothly and is very close to the reference values (symbols).

VI. Conclusion

In this paper we present results of investigation how different representations of the angles of the scattering parameters used as ANN training data affect modeling accuracy of ANN based bias dependent models of S-parameters' of microwave transistor. Earlier it was found that sharp changes in the frequency dependence might make modeling bad. Here we show that representation of S-parameters in the form Real+j*Imag, and development of the ANNs that model real and imaginary parts of the S-parameters may be a convenient way of overcoming the above mentioned problem. Beside the improvement of modeling accuracy, this modeling approach does not require previous detailed analyses of training data frequency dependence but only preprocessing of the complex training data by representing them by its real and imaginary parts instead by magnitudes and angles as usual.

References

[1] Q. J. Zhang, K. C. Gupta, *Neural Networks for RF and Microwave Design,* Artech House, 2000.
[2] F.Gunes, H.Torpi, F.Gurgen, "Multidimensional signal-noise neural network model", *Circuits, Devices and Systems, IEE Proceedings,* Vol.145, Iss.2, Apr 1998, pp. 111-117.
[3] Z. Marinković, A. Stošić, V. Marković, O. Pronić, "ANNs in Bias-Dependent Modeling of S-parameters of Microwave FETs and HBTs", *Microwave Review*, No.1, Vol. 12, June 2006, pp. 21-30
[4] D. Pozar, *Microwave Engineering*, John Wiley & Sons, Inc., New York, 1998
[5] http://www.semiconductor.agilent.com.