Artificial Neural Networks as a Tool for Improving Microwave Transistor Empirical Noise Models

Zlatica Marinković, Olivera Pronić-Rančić and Vera Marković

Faculty of Electronic Engineering
University of Niš, Serbia
1 Introduction

Computer-aided design (CAD) methods of modern communication devices require accurate and reliable noise models of microwave transistors (metal semiconductor field effect transistors - MESFETs, high electron mobility transistors – HEMTs, heterojunction bipolar transistors – HBTs, etc.) (Dobrowolski, 1991; Ladbrooke, 1989; Golio, 1991; Pozar, 2005). During the last few decades plenty of work has been carried out in the field of signal / noise modeling of microwave transistors. The most accurate models are physically based ones, but the applications of these models require knowledge of device physics and a number of technological parameters. Therefore, the empirical models, mostly based on equivalent circuit representations, are usually used, as more convenient, for a transistor signal and noise prediction in the microwave circuit design (Fukui, 1979; Pospieszalski, 1989; Wedge & Rutledge, 1992; Pronić & Marković, 2002).

Most of the existing transistor noise models are valid only for a specific temperature and bias conditions. Namely, for each new operating point it is necessary to repeat extraction of the model parameters from the measured data, which can be very time consuming if simulations / optimizations of the circuit containing a transistor have to be done in a wide range of operating conditions.

Artificial neural networks (ANNs) have appeared as an efficient alternative to standard empirical models of microwave transistors (Wang & Zhang, 1997; Shirakawa et al., 1997; Watson et al., 1998a; Devbahaktuni et al., 1998; Giannini et al., 2002; Gunes et al., 1998; Schreurs et al., 2002; Joodaki & Kompa, 2002; Munishi et al., 2003; Caddemi et al., 2005; 2007; Taher et al., 2005; Marinković & Marković, 2004; 2005a; 2005b; Marinković et al., 2004; 2006; 2007; 2008a; 2008b; 2009a; 2009b; 2010a; 2010b; 2011a; 2011b; 2012; Kabir et al., 2010). ANNs are suitable to be used as a modeling tool since they have an ability to learn from the presented data, and therefore they are interesting for extracting models directly from the measured data. Although there are microwave transistor models based completely on “black-box” approach (Gunes et al., 1998; Watson et al., 1998a; Marković & Marinković, 2004; 2005a; Caddemi et al., 2007), in this Chapter, we put attention on application of ANNs in improvements of the empirical equivalent circuit based noise models in terms of the accuracy and the range of validity.

The Chapter is organized as follows: after Introduction, the basic background of the ANNs is briefly introduced in Section 2. Section 3 starts with the description of typical empirical noise models and continues with the discussion on different ANN approaches for improving equivalent circuit based transistor noise models. A review of the most important results related to the considered topics is given in Section 4. Finally, in Section 5, the main conclusions are reported.

2 Artificial Neural Networks

For applications in microwave transistor small-signal and noise modeling standard multilayer perceptron (MLP) artificial neural networks are mostly used. An MLP artificial neural network consists of neurons (circles in Figure 1) grouped into layers: an input layer (layer 0), an output layer (layer NL) as well as several hidden layers, (Haykin, 1994; Zhang & Gupta, 2000). Each neuron is connected to all neurons from the next layer, while there are no connections among neurons within a layer. A transfer (activation) function is assigned to a neuron and each connection between neurons is weighted.
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:

$$Y_l = F(W_lY_{l-1} + B_l)$$

(1)

where $Y_l$ and $Y_{l-1}$ are outputs of $l$-th and $(l-1)$-th layer, respectively, $W_l$ is a weight matrix between $(l-1)$-th and $l$-th layer and $B_l$ is a bias matrix between $(l-1)$-th and $l$-th layer. Function $F$ is an activation function of each neuron and, usually, it is linear for the input and output layer and sigmoid (log-sigmoid or tan-sigmoid) for hidden layers.

The neural network learns relationship among sets of input-output data (training data) by adjusting network parameters (connection weights and neuron activation function thresholds). The basic algorithm used for ANN training is the backpropagation optimization algorithm (Haykin, 1994; Zhang & Gupta, 2000). As the first step in the backpropagation algorithm, input vectors are presented to the input neurons and output vectors are computed (forward phase). These output vectors are compared with desired values and errors are computed. Error derivatives are calculated and summed up for each weight and bias until whole training set has been presented to the network. The error derivatives are then used to update the weights and biases of the neurons in the model (backward phase). 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. There are several algorithms based on the backpropagation algorithm that have high order of convergence, such as Conjugate Gradient algorithm or Levenberg-Marquard algorithm (Haykin, 1994; Zhang & Gupta, 2000).

Once trained, the network provides fast response for different input vectors, even for those belonging to the same range as the training data but not included in the training set (generalization ability).

Evaluation of a trained ANN assumes evaluation of both the ANN learning ability and the ANN generalization. The ANN learning ability is evaluated by testing the ANN response for the values from
the training set, and its generalization by testing ANN response for the values differing from the ones from the training set.

Most often, the quality of the learning and generalization is estimated from the following measures, (Zhang & Gupta, 2000): average test error (ATE), worst case error (WCE) and correlation coefficient \( r \). The Pearson Product-Moment correlation coefficient, \( r \), is defined by:

\[
r = \frac{\sum (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum (x_i - \bar{x})^2 \sum (y_i - \bar{y})^2}},
\]

where \( x_i \) is the referent value, \( y_i \) is the neural network computed value, \( \bar{x} \) is the referent sample mean, and \( \bar{y} \) is the neural network sample mean. The correlation coefficient indicates how well the modelled values match the referent ones. A correlation coefficient near one indicates an excellent predictive ability, while a coefficient near zero indicates poor predictive ability.

In general, the quality of generalization depends, before all, on the size of the training set. A too small training set often may not represent properly the input space and that may lead to bad generalization although the quality of learning of the training data is high. Therefore, the training set should consist of samples properly covering all parts of input space. It is not necessary for the distribution of the samples to be uniform in the input space, moreover in some cases non-uniform distribution is recommended.

The size of the network itself (number of hidden layers and hidden neurons) makes influence on the network generalization. Namely, larger networks have greater capability of learning the training data, but too many network parameters (network parameters are connection weights and thresholds of the neuron activation functions) may lead to the overlearning. In that case the neural network output matches the training targets very well, but the generalization capability is poor. On the other hand, the networks with a small number of neurons usually are not capable to learn accurately the problem behavior even for the training set (underlearning). Since the optimal number of hidden neurons cannot be a priori set, usually ANNs with different number of hidden neurons are trained and the final one is chosen after comparison of the test statistics of the trained ANNs.

The number of epochs during the training has an important impact on the quality of learning as well. Too few epochs lead to the underlearning, while too many epochs may result in the overlearning. In order to control the generalization during the training it is advisable to use, beside the training set, a set intended for validation of the generalization. This set is used to check the errors after each epoch providing the possibility to stop the training earlier and to prevent the network overlearning.

Therefore, the final network structure has to be chosen to meet a compromise between learning and generalization. To summarize, a neural model development consists of several steps:

- problem definition (determining the number of the neurons in the input and output layer corresponding to the number of input and output parameters, respectively)
- building a training set (usually the measured data are used as the target values),
- optionally, building a validation test set for the network training purposes,
- building a test set (containing different data from the training set) to be used for the generalization estimation,
- training of the networks with the chosen number of input and output neurons and different number of hidden neurons,
- evaluation of the learning and generalization of the trained networks,
Each MLP ANNs can be described by a set of mathematical expressions that can easily be implemented within the software packages for circuit analysis and optimization, including microwave circuit simulators.

3 Noise Modeling of Microwave Transistors

3.1 Empirical Noise Models of Microwave FETs

The complete characterization of microwave transistors for small-signal and low-noise applications includes knowledge about both scattering (S) and noise parameters. In that case microwave transistors can be considered as two-port linear networks. Any two-port linear noisy network can be characterized by a set of four noise parameters and describe inherent noise behavior of the component, independent of a connected circuit. The most commonly set of noise parameters consists of: the minimum noise figure, \( F_{\text{min}} \), the equivalent noise resistance \( R_n \), and magnitude (\( \text{Mag}(\Gamma_{\text{opt}}) \)) and angle (\( \text{Ang}(\Gamma_{\text{opt}}) \)) of the optimum reflection coefficient, \( \Gamma_{\text{opt}} \). Usually, a normalized equivalent noise resistance \( r_n = R_n / Z_0 \) is used instead of \( R_n \) (\( Z_0 \) is a normalizing impedance, typically equals 50Ω).

Microwave transistor noise models used in microwave CAD packages are usually based on an equivalent circuit representation. Noise effects can be included in equivalent circuits in different ways (Dobrowolski, 1991; Golio, 1991; Pospieszalski, 1989; Dobrowolski, 1996; Pronić & Marković, 2002). Namely, any noisy linear two-port can be replaced by a noiseless two-port network and two additional noise sources that are generally correlated. Noise is typically characterized by equivalent voltage and/or current sources. Therefore, the impedance and admittance matrix representations of noisy two-port networks, the chain matrix representation and a few others are often used in CAD of noisy networks. On the other hand, in the noise wave representation, a noisy two-port network is described by using a noiseless linear equivalent circuit and the waves that emanate from its ports.

The two-parameter Pospieszalski’s noise model (Pospieszalski, 1989), one of the most frequently used transistor noise models, is considered to be very suitable for implementation into the standard commercial microwave circuit simulators. The equivalent circuit of a microwave FET in the packaged form including noise sources according to Pospieszalski’s approach is shown in Figure 2. The intrinsic circuit is denoted by a dashed line and it is embedded in a network representing the device parasitics. The noise sources \( e_{gs} \) and \( i_{ds} \) represent the effects of noise generating inside the device. The equivalent gate and drain temperatures (\( T_g \) and \( T_d \)) are assigned to these noise sources, respectively. The noise parameters of a transistor intrinsic circuit are calculated from simple expressions as the functions of the noise equivalent circuit parameters (ECP): transistor intrinsic circuit elements and gate and drain noise temperatures (Pospieszalski, 1989). When the noise parameters of the intrinsic circuit are determined, the equivalent circuit parasitic elements have to be added to the intrinsic circuit with the aim of determining the noise parameters of the whole device.

For high frequency circuit applications, a wave interpretation of noise also seems very appropriate (Meys 1978; Wedge & Rutledge, 1992; Dobrowolski, 1996). The wave interpretation of noise allows the use of scattering matrices for noise computations, leading to the advantages in CAD of microwave networks. It was shown that the wave approach could be useful for applications in both noise modeling and
measurements of microwave FETs (Wedge & Rutledge, 1992). The noise wave modeling procedure of microwave FETs based on \( T \) representation of transistor intrinsic circuit was presented in (Pronić & Marković, 2002). In that case three noise wave temperatures are introduced as empirical noise model parameters. The noise parameters of transistor intrinsic circuit are calculated in terms of the noise wave temperatures. The noise parameters of the complete circuit are computed after adding the parasitics.

![Microwave FETs equivalent circuit including noise sources](image)

**Figure 2:** Microwave FETs equivalent circuit including noise sources

### 3.2 ANNs in Noise Modeling of Microwave Transistors

Most of the existing equivalent-circuit-based noise models of microwave transistors are valid only for a single operating point. Having in mind that both the \( S \)-parameters and noise parameters of microwave transistors are bias-, frequency- and temperature-dependent, recalculation of elements / parameters of the equivalent circuit for each operating temperature and/or bias point is required. It means that for each new operating point, time consuming extraction procedures (optimizations or analytical procedures in a microwave circuit simulator) should be repeated.

As it will be described later, by applying ANNs in microwave transistor modeling, models valid in a wide range of operating conditions can be developed, often providing higher modeling accuracy comparing to the standard models. Furthermore, ANNs can be used for efficient determination of equivalent circuit parameters without optimization procedures in microwave simulators.

The simplest transistor noise model based on ANNs referred in the literature is completely “black-box” model. Namely, a “black-box” model is an ANN trained to predict noise parameters for given combination of operating conditions (biases and/or ambient temperature) and frequency, as shown in Figure 3 (Gunes et al., 1998; Marković & Marinković, 2004; 2005a, Caddemi et al., 2007). The ANN has \((n+1)\) inputs, i.e. \((n+1)\) neurons in the input layer, corresponding to \(n\) parameters representing operating conditions and to the frequency. There are four neurons in the output layer corresponding to the four noise parameters. For the model development it is necessary to acquire a set of noise parameters measured at different operating conditions in the considered frequency range. Namely, each data sample from the train-
ing set consists of one combination of the \((n + 1)\) input parameters and corresponding measured noise parameters used as the target values for the network outputs. Using the trained ANN, the noise parameters for an arbitrary operating point and frequency are easily obtained by calculating the ANN response. It is important to note that for development of this model, there is no need for knowledge about the device physics and processes in the device.

![Figure 3: Black-box ANN noise model](image)

The focus in this study is on the applications of ANNs for improving the accuracy and range of the validity of the existing equivalent circuit based empirical noise models. Namely, all effects contributing to the device behavior are often not completely included into empirical transistor noise models, and therefore there are deviations of simulated values from the measured ones. One of the simplest ways to model all effects contributing to the device noise is to add an ANN to the empirical noise model aimed to correct the noise parameters obtained by the empirical model, as illustrated in Figure 4.

![Figure 4: Empirical noise model with improved accuracy (PKI approach)](image)

The inputs of this ANN are, besides the frequency, the noise parameters obtained by the empirical noise model \((F'_{\text{min}}, \text{Mag}(\Gamma'_{\text{opt}}), \text{Ang}(\Gamma'_{\text{opt}}), r'_n)\). This means that the empirical model equivalent circuit has to be extracted in a microwave circuit simulator before the further development of the neural part of the model. Therefore, the ANN has five neurons in the input layer. Four neurons in the output layer correspond to the final values of the device noise parameters \((F_{\text{min}}, \text{Mag}(\Gamma_{\text{opt}}), \text{Ang}(\Gamma_{\text{opt}}), r_n)\). The training dataset used for development of this ANN refers to a list of frequency values over the operating frequency range. Each training data sample consists of: frequency and four noise parameters obtained in a microwave simulator from the empirical noise model as well as of the four targets – the noise parameters measured at that frequency.
After the training and evaluation of the network, it is added on the equivalent circuit noise model in microwave circuit simulator by means of a block dealing with variables and expressions (VAR) whose contents are equations that describe the trained network. Inputs of the VAR block are the same as the inputs of the neural network, while the noise parameters calculated from these expressions have to be assigned to the device. Such schematics with the added neural network can be used as a new library noise model.

Since the target training data for this ANN are the measured values of noise parameters, the ANN can compensate the inaccuracies that may arise in standard empirical models because of neglecting some of the effects contributing to the device noise. The ANN itself belongs to so-called prior knowledge input (PKI) ANNs (Watson et al., 1998b; Marinković et al., 2008b). PKI ANNs are ANNs that have, besides the original input values (in this case the frequency), additional inputs representing prior knowledge about the problem being modeled, which can be represented by approximate values of the output parameters (for example, obtained by empirical or simplified models) or by certain parameters correlated with the modeled output parameters (obtained either by an appropriate external model or by measurements). In the considered case, the noise parameters obtained by the empirical noise model represent the ANN prior knowledge.

The accuracy of the considered model is improved by using the above described procedure, but it is still valid for only one operating bias point. By adding new ANN inputs that correspond to the operating conditions, as shown in Figure 5, the validity of the previous noise model (shown in Figure 4) can be extended to a range of operating conditions. Unlike the ANN from the previous model, which has five input neurons, the new ANN has \( n \) neurons more, corresponding to the \( n \) parameters representing the operating conditions. The number of output neurons remains four.

The model is developed as follows: first, one operating point from a modeled device operating range is chosen. Then measured S-parameters and noise parameters at the chosen reference operating point have to be acquired for several operating frequencies. From the measured data, the elements and parameters of the empirical noise model for the chosen operating point are extracted in a microwave simulator (usually by using an optimization procedure). The next step is training of the neural network. For that purpose, as target data, measured values of the noise parameters for certain number of operating points should be supplied. Therefore, each training data sample refer to a particular combination of the operating conditions and the frequency, the noise parameters simulated from the equivalent circuit for the

![Figure 5: Noise model dependent on operating conditions (extended PKI approach)]
given frequency, and finally the targets – measured noise parameters at the given combination of the operating conditions and frequency. After the neural network training, it is added on the schematics of the noise model in microwave simulator as previously explained. As in the case of the previous model, since the target training data for this ANN are the measured values of noise parameters, the ANN models all effects contributing to the device noise.

An alternative way for including the dependence on bias conditions in the equivalent-circuit-based noise models of microwave transistors is through modeling of dependence of the equivalent circuit parameters (ECPs) on operating conditions by ANNs. For this purpose, an ANN trained to predict ECP dependence on operating conditions is assigned to the empirical noise model, as shown in Figure 6.

This approach belongs to so-called space mapping (SM) approach (Bandler et al., 1994), i.e. the ANN is used for transformation of the space of input parameters (operating conditions) to the input space of the empirical noise model (ECPs). The ANN used for space-mapping has as many inputs as the number of parameters describing the operating conditions (n), and as many outputs as the number of ECPs (N). For training of this ANN, it is necessary to extract ECPs in a standard way at certain number of combinations of the operating conditions. Training samples consist of a combination of the operating conditions and ECPs extracted in the simulator from the scattering and noise parameters measured at those operating conditions. Once trained, the ANN can be used for direct determination of ECPs for any combination of operating conditions without new extraction procedures and without need for measured data. As in the previous case, the ANN is assigned to the empirical model by adding the corresponding mathematical expressions in the model. Outputs of the expressions are assigned to the corresponding ECPs. After assigning the ANN (i.e., the corresponding mathematical expressions) to the empirical model, the model becomes dependent on operating conditions. However, as far as the model accuracy is concerned, it can be noticed that the accuracy of an SM ANN model is limited by the accuracy of the empirical noise model (Marinković et al., 2007). Namely, the accuracy of this neural model is limited by the accuracy of the chosen transistor empirical noise model and the accuracy of extraction of the ECPs used for the network training. The better ECP extraction is, the better training data are and therefore the better neural model is. Therefore, the accuracy of the neural model could not be better than the accuracy of the original transistor empirical noise model and this is its main limitation concerning the model accuracy.

The advantages of the two previously mentioned approaches (PKI and SM approaches) can be combined within a single model. Namely, this model consists of an empirical noise model and two ANNs (ANN 1 and ANN 2), Figure 7. The first network, ANN1, plays a role of mapping of the original input space to the input space of the empirical model. The second network, ANN 2, produces the final values
of the noise parameters and improves the accuracy of the model. The noise parameters obtained by the SM model are led to the ANN 2 as the prior knowledge. ANN 2 has operating conditions and frequency as inputs as well.

Figure 7: Improved empirical noise model dependent on operating conditions (combined PKI-SM approach)

The steps in the model development are the following. First, it is necessary to develop equivalent circuit empirical noise model in a microwave simulator. The next step is extraction of the empirical model ECPs for several combinations of the operating conditions. These ECPs are then used as the targets for the training of the ANN 1. After the training of the ANN 1, the mathematical expressions describing it are added to the equivalent circuit, which in that way become dependent on the operating conditions. Namely, for any new operating point, it is not necessary to extract ECPs in a standard way; they are automatically calculated from the operating conditions by the implemented expressions. For training of ANN 2 it is necessary to simulate the noise parameters by the empirical model with added expressions for ECP determination for several operating frequencies at all considered operating conditions. The training target values in this case are the noise parameters measured at the corresponding operating conditions and frequency. The mathematical expressions describing the ANN 2 are added to the equivalent circuit and the final calculated values of the noise parameters are assigned to the component.

As the ANN 2, which produces the final values of the noise parameters, is trained by the measured data, this type of the models shows higher modeling accuracy than SM models (Marinković & Marković, 2005b).

4 Modeling Examples

A review of the most illustrative modeling examples related to applications of ANNs for improving empirical transistor noise models in terms of the accuracy and the range of validity is presented here. As previously said, the optimal number of hidden neurons cannot be a priori set. Therefore, for any trained structure of ANNs, ANNs with one and two hidden layers and a different number of neurons in hidden layers were trained and the final ANN was chosen after comparison of the test statistics of the all trained ANNs. Moreover, since the initial parameters of the ANNs were set randomly, each structure was trained independently few times in order to avoid local minima during the optimization.
In addition to the numerical test criteria mentioned earlier (test errors and correlation coefficients) as well as correlations plots, the quality of generalization was checked also from the network response visually by plotting the output parameters versus the input parameters with a step smaller than the step of change of the input parameters in the training/test set. In that way, possible deviations from the desired behaviour were easily noticed, especially overlearning effects.

4.1 Example 1

Accuracy improvement of noise models of packaged MESFETs / HEMTs and dual-gate MESFETs applying the neural approach illustrated in Figure 4 is presented in (Marinković et al., 2008a). The proposed model is based on a basic transistor noise wave model as the empirical noise model (Pronić & Marković, 2002). The numerical results obtained for a packaged HEMT transistor, type CFY65A, are presented here to illustrate the validity and effectiveness of this approach.

The modeling procedure was performed as follows. First, the elements and parameters of the basic noise wave model were extracted in the microwave simulator. The obtained noise parameters are presented in Figure 8 as dashed lines, while the corresponding measured (reference) values are shown as symbols. It can be observed that the modeled values deviate from the measured ones almost in the whole frequency range. Further, the training process of the additional ANN was done, using the measured noise parameters as the target training data. ANNs with different number of hidden neurons were trained and the best modeling results were obtained using an ANN with one hidden layer which has nine hidden neurons. This ANN is added to the device noise wave model in the circuit simulator. The final values of the noise parameters are shown in Figure 8 as solid lines. Unlike the values obtained by basic noise wave model, the noise parameters values obtained by the ANN are very close to the measured ones, confirming that in this way more accurate noise modeling is achieved.

![Figure 8: Noise parameters of packaged HEMT type CFY65A](image-url)
4.2 Example 2

As mentioned earlier, the most empirical noise models of microwave FETs are valid only for one specific bias point. To include bias dependence in the empirical noise model, the ANN approach illustrated in Figure 5 was applied (Marinković et al., 2009b; 2011a). Pospieszalski’s noise model, shown in Figure 2, is considered as the empirical noise model. The considered operating conditions in this case are bias conditions (dc drain-to-source voltage, \( V_{ds} \), and dc drain-to-source current, \( I_{ds} \)). The model consists of the empirical noise model developed for one representative (reference) bias point and an ANN through which dependence on bias conditions is included in the model. The ANN has four network outputs corresponding to the noise parameters and seven inputs corresponding to bias conditions, frequency and the noise parameters obtained by the empirical noise model for that value of frequency.

This approach was applied to a pHEMT device, type ATF35143, by Agilent (HP). The measured values of \( S \)- and noise parameters were taken from the device datasheet. Data referring to the following values of bias voltage, \( V_{ds} \), and bias current, \( I_{ds} \), were available: \( V_{ds} \): 2V, 3V and 4V, \( I_{ds} \): 5mA, 10mA, 15mA, 20mA, 30mA, and 60mA.

First, the ECPs of Pospieszalski’s noise model were extracted from the available measured data for a bias point (2V, 10mA) chosen as the reference one. As the next step, a simulation of noise parameters using the developed model was performed at frequency points for which measured values of the noise parameters were available. Then, having the simulated values as the inputs and the measured noise parameters as the target values the training set was made. After the evaluation of the trained ANNs, the ANN which has one hidden layer with 10 neurons was chosen to be implemented in the final model. The chosen ANN was then added on the equivalent circuit schematics in the form of mathematical expressions, as explained earlier.

To evaluate model accuracy, Figure 9 shows the simulated noise parameters plotted against the corresponding measured data. In an ideal case of the total matching, all the points in a plot would be on the straight line \( y = x \). Therefore, the higher scattering the lower model accuracy. It is obvious that the points do not scatter much from the line \( y = x \), indicating in that way a high modeling accuracy. This is confirmed by the test statistics as well. Namely, the average test error (ATE), the worst case error (WCE) and correlation coefficient \( (r) \) were calculated according to (Zhang & Gupta, 2000) for all bias points reported in the datasheet and are given in Figure 9 as well. Small values of ATE and WCE and the correlation coefficient very close to 1 confirm that the developed model exhibits a very high accuracy.

4.3 Example 3

The approach presented in Figure 5 was also applied for including bias dependence in the noise wave transistor model. Noise modeling of GaAs FET packaged microwave transistor, type ATF21186, by Agilent (HP), was done. Measured values of \( S \) and noise parameters for biases (2V, 10mA), (2V, 15mA) and (2V, 20mA), in the frequency range \( (0.5 – 8) \) GHz, taken from the device datasheet, were used for the model development. As we considered the case where dc drain-to-source voltage is constant, the bias conditions are defined by dc drain-to-source current, \( I_{ds} \). The inputs of the neural network are: bias current and frequency, as well as the prior knowledge noise parameters for the reference bias point obtained from the noise wave model for that value of frequency. The ANN’s outputs are device noise parameters.
The bias point (2V, 15mA) is chosen to be the reference one. After extraction of the noise wave model parameters for this bias point, simulation of the noise parameters needed for building the training set was done and the training set with the measured noise parameters as the output targets was built. Then, ANNs with different number of hidden neurons were trained. After training and evaluation of the ANNs, it was found that the network with one hidden layer having 15 neurons gave the best results.

As illustration, in Figure 10 there are plots of the noise parameters simulated by the proposed PKI ANN model and compared with the measured values. Circles denote the measured values and solid lines – the simulated ones obtained by the proposed PKI ANN model. It can be seen that the predicted noise parameter values match very well the measured ones.

### 4.4 Example 4

The noise modeling approach presented in Figure 7 was applied to improve the Pospiezalski’s noise model of microwave FETs in terms of accuracy and the range of validity. The developed model includes noise dependence on bias conditions.
The proposed modeling method was applied to a GaAs FET, type ATF21186 by Agilent (HP), and the most illustrative results are presented here. We considered the case where dc drain-to-source voltage is constant, i.e., operating conditions are represented by the dc drain-to-source current, $I_{ds}$. The first step was training of the ANN 1 aimed to simulate ECP dependence on the bias current. This ANN has one input neuron (corresponding to bias current) and 26 output neurons (24 of them correspond to small-signal circuit elements and the rest two correspond to the equivalent gate and drain noise temperatures) (Marinković et al., 2009a). The number of neurons in the hidden layer was determined through the training process. The ECP values extracted from the mentioned measured $S$- and noise parameters were used as the training data. A network with three hidden neurons was chosen as the best trained one and taken as the final model of ECPs’ dependence on the bias current. Then, the model was implemented in the ADS microwave simulator.

To improve the accuracy of this bias dependent model (SM neural model, or “basic neural model” as called further) an additional neural network (ANN 2 in Figure 7) was trained to produce more accurate

**Figure 10:** Noise parameters versus frequency
values of the noise parameters. That network has 6 input neurons corresponding to the bias current, frequency and four noise parameters simulated by the SM neural model. In the output layer there are four neurons corresponding to the improved noise parameters. The neural network ANN2 is trained using a set of the measured noise parameters for a certain number of bias currents and frequencies and corresponding values of the noise parameters obtained by the basic neural model. Among the trained ANNs with different number of hidden neurons, the best-obtained ANN has two hidden layers, each containing 10 neurons. After the ANN had been implemented in the simulator, the noise parameters are simulated and compared with the available measured data.

To illustrate modeling accuracy, in Figure 12 there are plots of the minimum noise figure and the magnitude of the optimum reflection coefficient simulated by the improved ANN model (solid line) and compared with the measured values (symbols) for the bias currents from 10 mA to 20 mA in the frequency range (0.5-8) GHz. The corresponding values simulated by the basic ANN model (dashed line) are given in plots as well. It is obvious that the values obtained by the improved neural model are much closer to the corresponding measured values then the ones obtained by the basic neural model, proving the improvements of the modeling accuracy. It is important that the noise parameters for all values of the bias current (step 1 mA) were obtained directly from the model without knowledge about the measured data.

Additionally, scatter plots of the simulated minimum noise figure and normalized equivalent noise resistance values versus measured data are shown in Figure 12. The values obtained by the basic neural model are denoted by white circles, and the ones corresponding to the improved neural model are denoted by black circles. It can be observed that data points obtained by the improved neural model are much closer to a straight line $y = x$, indicating the accuracy improvement.

**Figure 11:** Minimum noise figure and magnitude of optimum reflection coefficient
5 Conclusion

Low-noise active devices and components are used in modern communication systems to ensure a low level of overall system noise. Therefore, for computer aided design of low-noise microwave active circuits it is necessary to have reliable and accurate noise models for all active devices, especially for microwave transistors. The majority of microwave transistor noise models are empirical models based on the device equivalent circuit representation. However, most of the models do not represent device noise in the whole operating frequency range with the same accuracy, which depends on a noise model itself as well as on the accuracy of the extraction of the elements of the equivalent circuit and parameters of the noise models.

Accuracy of the noise modeling can be improved by assigning an ANN to the empirical model with the aim to correct the noise parameters obtained by the empirical model. Having in mind that measured values of the noise parameters are used as the network targets, all effects contributing to the transistor noise are included in the model and possible deviations of the parameters caused by insufficiently accurate optimizations used for determination of the ECPs are corrected. As the noise parameters depend on the operating conditions, it is necessary to repeat extraction of the model ECP for each new combination of the operating conditions. To avoid repeated extractions, an ANN is trained to predict dependence of the model ECPs on operating conditions. By using the trained ANN, ECPs are determined over the considered operating range without further measurements and optimizations. After assigning the ANN to the empirical model, the model becomes dependent on operating conditions. However, the model accuracy is still limited by the accuracy of the chosen empirical noise model and the accuracy of extraction of the ECPs used for the network training. To improve the accuracy further, this approach can be combined with the previous approach, i.e., two ANNs are assigned to the model, one for determination of the ECPs
and the other for correction of the noise parameters. In that way the modeling procedure becomes more efficient and, at the same time, the model becomes more accurate.

According to the presented research examples, it can be concluded that ANNs can be efficiently used as a tool for improving the empirical models in terms of the accuracy and the range of model validity. Additionally, by combining the ANN approach with the standard empirical noise models, the noise modeling procedure becomes more efficient.

**References**


