[Turkish Journal of Electrical Engineering and Computer Sciences](https://journals.tubitak.gov.tr/elektrik)

Volume 13 | Number 1 Article 4

1-1-2005

Prior Knowledge Input Method In Device Modeling

SERDAR HEKİMHAN

SERDAR MENEKAY

N. SERAP ŞENGÖR

Turk J Elec Engin, VOL.13, NO.1 2005, © TÜBİTAK

Prior Knowledge Input Method In Device Modeling

Serdar HEKİMHAN, Serdar MENEKAY, N. Serap ŞENGÖR
İstanbul Technical University, Faculty of Electrical & Electronics Engineering,
80626, Maslak, İstanbul-TURKEY
e-mail: shekimhan@hotmail.com, smenekay@mynet.com, neslihan@ehb.itu.edu.tr

Abstract

The artificial neural networks are being used in modelling electronic elements and devices especially at microwave frequencies where non-linearity and dependence on frequency cannot be neglected. In this paper, instead of using artificial neural networks as a unique modelling device prior knowledge input method based on feed-forward artificial neural network structures as multi-layer perceptrons and wavelet-based neural networks is investigated. The benefits of prior knowledge input method over plain usage of artificial neural networks in modelling BJT is explored by comparing models obtained with and without prior knowledge input. The novelty of the paper is utilizing wavelet-based neural networks as feed-forward structure in prior knowledge input method. The training and test data used in simulations are obtained by HP 4155 parameter analyser.

1. Introduction

The electronic devices' response at RF/microwave frequencies can be obtained solving physical equations using Electromagnetic (EM) simulators or by measurement. As these techniques are time consuming and not much versatile for design purposes, methods as empirical models, look-up tables etc., are used during CAD based design procedures [1]. Artificial Neural Networks (ANNs) are being used as a mean of model generator for the last decade especially for microwave elements and devices. ANNs give less accurate results than EM simulators but are very fast once training phase is complete. Moreover they are much more accurate than conventional approximate methods and suitable for CAD applications. So, models obtained via ANNs eliminate drawbacks of time consumption in EM simulators and non-accuracy in conventional techniques The ANN approach is coming to be common as applications at RF/microwave frequencies are more needed. Besides its comparable fastness and accuracy, ANN has the advantage of giving black-box models in case of insufficient knowledge to obtain a model of a device.

The aim of this work is to investigate Prior Knowledge Input (PKI) method, especially with different ANN structures and observe its advantages over using plain ANN structures in modelling. In this work, as an example transistor is modelled. Transistors, especially BJT's, are widely used in discrete circuits as well as in IC design, in analogue and digital applications. As, BJT has different responses at different frequency ranges modelling BJT requires solving different equations at different frequencies. ANN models can be used to obtain its response at a broad spectrum. However in this work, as training and test sets are not sufficient, model only at DC working conditions are obtained. This is a drawback of ANN, as models depend on training set dramatically.

However in literature, various methods based on ANNs such as space mapping [2], knowledge-based [3], prior knowledge input [4,5] are proposed to overcome this inconvenience and to prevent long training time. On account of its simplicity PKI is considered in this work to obtain a model of BJT to prevent long training. In this work, the purpose is to investigate the performance of Wavelet-based Neural Network (WNN) within the context of PKI method. WNN is claimed to have better performance for non-linear regression applications [6]. The comparison of WNN with Multi-Layer Perceptron (MLP) is exploited for the considered application, as MLP is the most popular structure.

In section II, MLP [7] and WNN [6] used in obtaining models of devices and circuits especially used at RF/microwave frequencies are presented. In section III, PKI method is introduced and in section IV simulation results obtained are given.

2. Multi-layer Perceptron and Wavelet-based Neural Network

In most applications of ANN in modelling, a feed forward neural network is used to get a non-linear input-output mapping related with a device or circuit [1].

In this paper two kinds of feed forward neural networks, MLP and WNN, will be presented. Besides using these ANN structures in PKI method, for comparison MLP based modelling will also be given.

2.1. Multi-layer perceptrons

MLP as shown in Figure 1 has three layers input, hidden and output, respectively. Hidden layers, which can be more than one, are formed using neurons having continuously differentiable non-linear activation functions, while output neurons have linear activation functions since MLP is used for modelling purpose. Input layer has no function other than fan out the data. Input data move ahead through the layers and at output layer, output of the ANN is obtained. These outputs, y, are compared to desired values, y_d, and weights are changed by means of back propagation algorithm minimizing error function depending on error, e, between desired and obtained outputs. Error, instantaneous error of k-th training data and error function are given below by Equations (1), (2) and (3), respectively.

$$e = y_d - y \tag{1}$$

$$\varepsilon^{(k)} = \frac{1}{2}\sum_j e_j^{(k)^2} \tag{2}$$

$$\varepsilon_{ort} = \frac{1}{p}\sum_{k=1}^{p} \varepsilon^{(k)} \tag{3}$$

Updating weights is carried on mostly until average error of training set, is in a predetermined range. Stopping criteria given in Equation (4) is considered in this work where change in average error is preferred rather than average error itself.

$$|\Delta\varepsilon_{ort}| < \varepsilon \tag{4}$$

As back-propagation algorithm is a well-known method [7] it will not be mentioned further here.

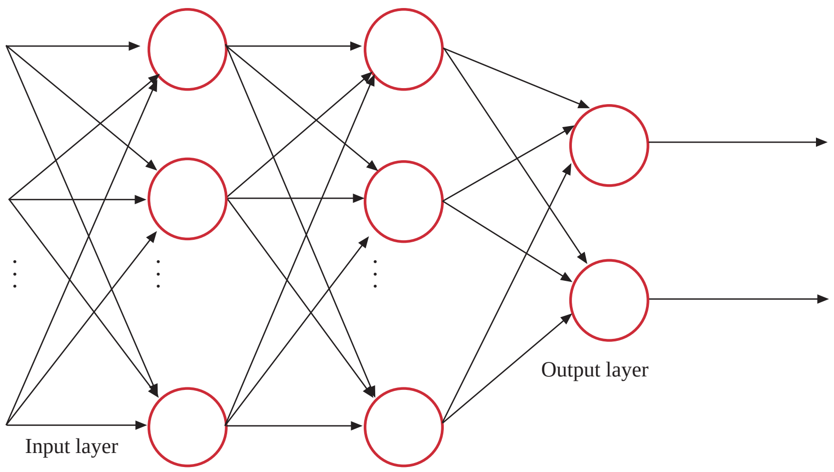

Figure 1. Multi-layer perceptron structure.

2.2. Wavelet based network

Another feed-forward structure used as universal approximator in different applications is WNN. MLPs' hidden layers can be more than one but WNN has only one hidden layer as shown in Figure 2. WNN, which has been obtained by incorporating wavelet theory into ANN is a new feed-forward structure [6]. Another difference between MLP and WNN is that activation functions of hidden layer of are wavelet functions. Generally used inverse Mexican hat function given in Equation (5), (6) is utilized in this work.

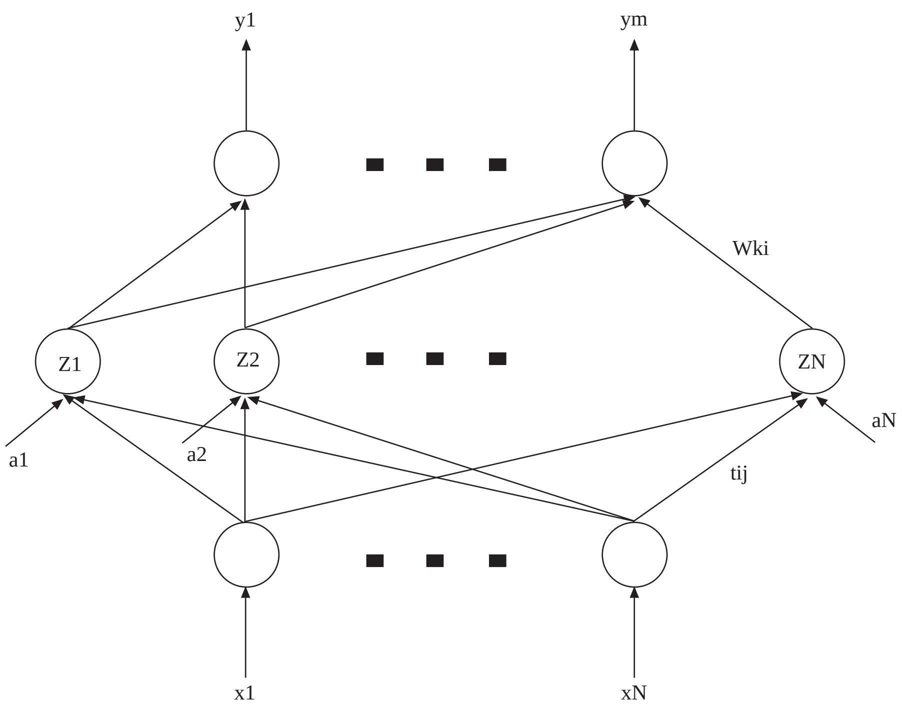

Figure 2. Wavelet-based neural network structure.

$$Z_i = \Psi\left(\frac{x - t_i}{a}\right) = \sigma(\gamma_i) = \left(\gamma_i^2 - n\right)\exp\left(-\frac{\gamma_i^2}{2}\right) \tag{5}$$

$$\gamma_i = \left\| \frac{x - t_i}{a} \right\| = \sqrt{\sum_{j=1}^{n} \left(\frac{x - t_i}{a} \right)^2} \tag{6}$$

The outputs of WNN are obtained by Equation (7), where m denote the number of hidden neurons.

$$y_i = \sum_{j=1}^{m} w_{ij} z_j \tag{7}$$

Outputs are compared to desired outputs same as in MLP, and parameters w_{ij}, t_{ij} and a_{ij} are updated to minimize error function given in Equation (2) using steepest descent method. As back-propagation algorithm depends on the same optimisation method updating of parameters by Equations (8-10) are very similar to back-propagation but this time as different type of parameters are of concern at each layer there is no need to back-propagate error. η in these equations is the learning rate.

$$W_{ji}^{(k+1)} = w_{ji}^{(k)} - \eta \frac{\partial E^{(k)}}{\partial w_{ji}^{(k)}} \tag{8}$$

$$a_i^{(k+1)} = a_i^{(k)} - \eta \frac{\partial E^{(k)}}{\partial a_i^{(k)}} \tag{9}$$

$$t_{ij}^{(k+1)} = t_{ij}^{(k)} - \eta \frac{\partial E^{(k)}}{\partial t_{ij}^{(k)}} \tag{10}$$

When Equation (4) is satisfied, training is terminated. In order to understand whether network is well trained, it has to be tested with test data.

How ANN structures are trained for modelling purpose is shown in Figure 3. Scaling unit is needed to carry the circuit/device parameters and responses into a range meaningful for ANN structure considered. Once training is over, the weights and bias terms of trained ANN is saved and ANN model is used for design purposes. It has to be pointed out that even though training phase is long, using ANN model takes very short time, as there are no complex calculations.

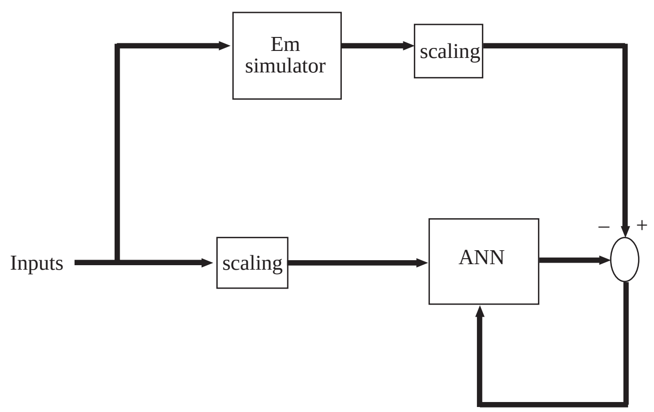

Figure 3. Neural network structure during training phase for model generating.

3. Prior Knowledge Input (PKI) Method

For good training of neural networks large amount of data are needed. But, on the other hand, training time increases with the number of data. Furthermore, obtaining large set of data is not an easy process as it depends either on measurement or time taking difficult EM computations. So, if any information exists about component, circuit etc., to be modelled, time for training can be reduced with addition of this knowledge into modelling method. PKI is a method, which not only reduces training time but also produces more accurate results than modelling based on plain ANN structures.

PKI method is presented first by Gupta. In this method inputs of ANN are no longer only physical/process parameters related with the device to be modelled but also outputs of equivalent circuit model/empirical equations as shown in Figure 4. This is the main difference of the PKI method compared to using unique ANN structures in modelling. This method is used for modelling finline [4] empirical expressions are used as prior knowledge, and for microwave components [5] where an ANN model trained beforehand is used as prior knowledge.

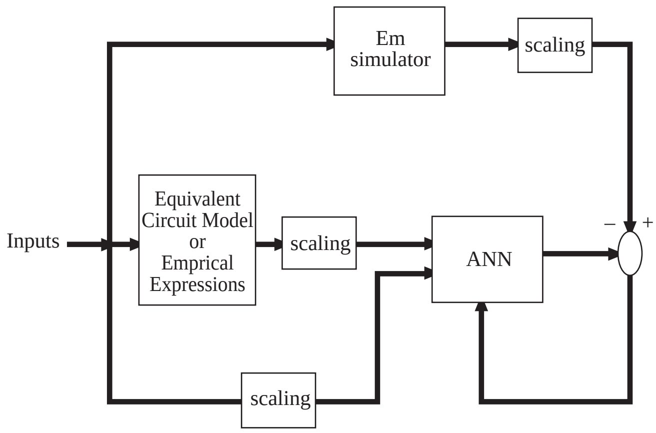

Figure 4. Structure for PKI method in model generating.

The difference between using ANN with PKI and without PKI during training can be seen from Figures (3) and (4). In case of PKI method, inputs to ANN is enlarged, but as will be seen in simulation results this has no disadvantage on training phase, indeed training phase is dramatically shortened. Once training is over, model obtained by PKI is composition of two sub-systems, equivalent circuit/empirical relations and ANN.

4. Simulation Results

It is expected that ANN is more accurate than emprical relations and faster than fine model. To observe this MLP based ANN structure explained in section 2.1 is trained and model of BJT is obtained. Also to benefit the knowledge about BJT given by empricial relations PKI method is utilized. BJT Ebers-Moll 3 model is considered as emprical relations to obtain collector and base currents and two different models are obtained by PKI method where two different ANN structures, namely MLP and WNN, are used. Thus on total three models for BJT are obtained and these are compared with experimental data.

Desired output data is obtained by HP 4155 parameter analyser. After data is obtained, they must be scaled. Scaling plays very important role for the ANN since if data is not seperately distributed in the region of interest, formed according to activation functions used covergence may not occur. So, in this study, first logarithmic scaling and then linear scaling, given by Equations (11), (12) respectively, are used. Scaling ranges are -0.9 and 0.9 for MLP and -1 and 0.5 for wavelet-based type. As mentioned above these range values depend on activation functions used so, as in MLP with and without PKI, hyperbolic tangent function is used range values are –1 and 1. The range of inverse Mexican hat function is on the interval of [-1, 0.5] the values are choosen appropriately.

$$\hat{X} = \ln(X) \tag{11}$$

$$\hat{X} = \hat{X}_{\min} + \frac{X - X_{\min}}{X_{\max} - X_{\min}} \left(\hat{X}_{\max} - \hat{X}_{\min} \right) \tag{12}$$

where,

X : value taken from data set

$\hat{X}$: scaled value

$X_{\min}$: minimum of the data set

$\hat{X}_{\min}$: minimum scaled value

$X_{\max}$: maximum of the data set

$\hat{X}_{\max}$: maximum scaled value

One of inputs of neural networks scaled is base-emitter voltage, and other two inputs scaled are collector and scaled base currents obtained by Ebers-Moll 3 model given in Equations (13,14). Due to PKI method ANN structure is in a way restricted with Equations (13,14).

$$I_c = I_s(0) \left[\exp\left(\frac{V_{BE}}{V_T} \right) - 1 \right] \tag{13}$$

$$I_b = \frac{I_s(0)}{\beta_{FM}} \left[\exp\left(\frac{V_{BE}}{V_T} \right) - 1 \right] + C_2 I_s(0) \left[\exp\left(\frac{V_{BE}}{n_{EL}.V_T} \right) - 1 \right] \tag{14}$$

where, $I_s = 69,144.10^{-14}A$, $V_T = 25,875.10^{-3}V$, $\beta_{FM} = 506$, $C_2 = 74,165$, $n_{EL} = 2.94$.

The simulations are done using codes written in MATLAB® as m-files. Neural network toolbox of MATLAB® is not used since a composite system special for modeling purpose is formed. This same system can be used in obtaining models for different circuits/devices once training/test sets and for PKI method empricial equations/equivalent circuit are given.

In order to have meaningful comparison all ANN structures have same number of neurons in their hidden layer,i.e.,10 neurons. It has been tried with different number of neurons and observed that results are getting worse once the number of hidden layer neurons is over 18. This is due to over parameterization. Inputs of ANN in PKI method are more than without PKI method, they are 3 and 1, respectively. Outputs are 2, corresponding to base and collector currents. Learning rate η is choosen to be 0.5 as not much difference has been observed in results with trials in the range of [0.3, 0.7].

The comparison of base currents obtained from PKI method with MLP and WNN to measurement results are shown in Figures (5) and (6), respectively. In these figures, results obtained for scaled values are given. In Figure (7) and (8), descaled values of base current drawn on logarithmic scale is given for PKI method with MLP and WNN, respectively.

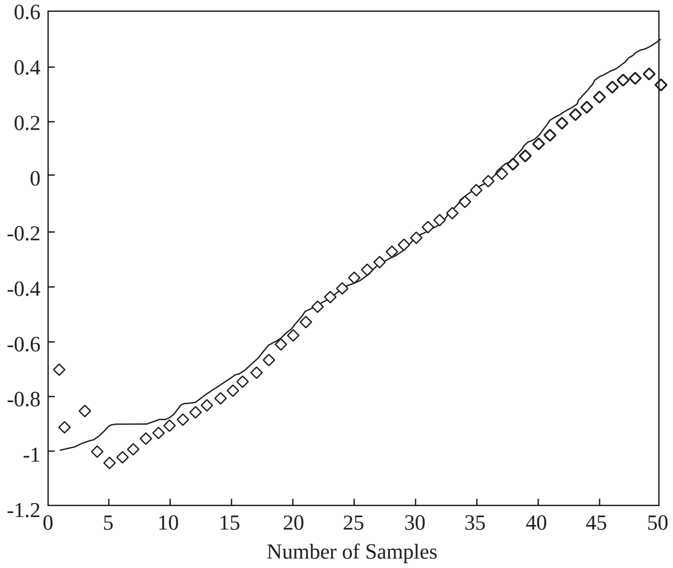

Figure 5. Scaled I_b values, PKI with MLP (desired –, obtained ◊).

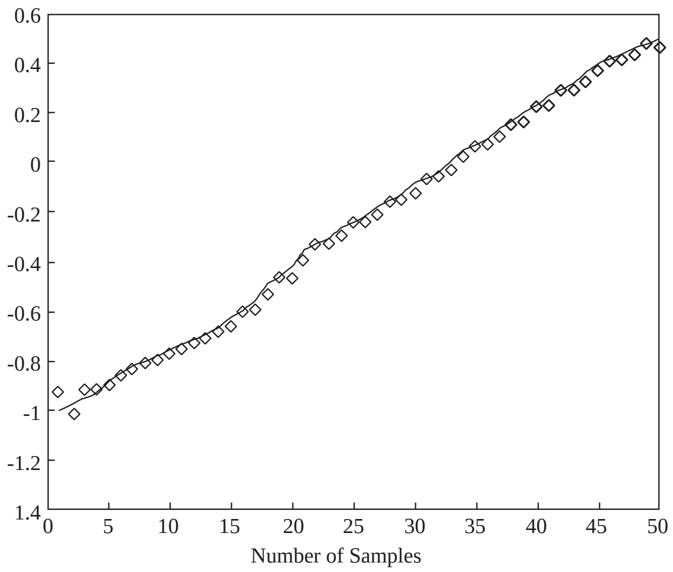

Figure 6. Scaled I_b values, PKI with WNN (desired –, obtained ◊).

The comparison of PKI method with MLP, WNN and desired values obtained for collector current is given in Figure 9. The results exploited in Figures (5-9) are all test results. They are obtained using input values that are not in the training set. Thus, these figures also show the generalization capacity of ANN models.

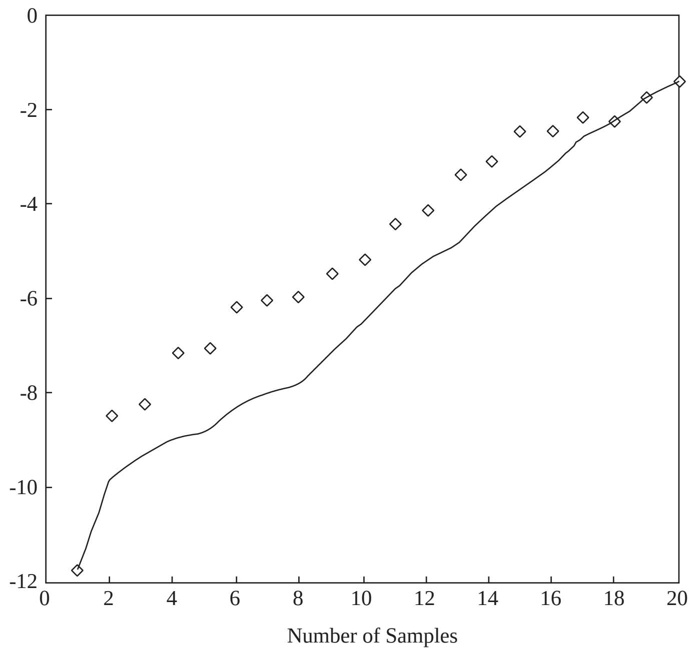

Figure 7. De-scaled I_b values, PKI with MLP (desired –, obtained ◊).

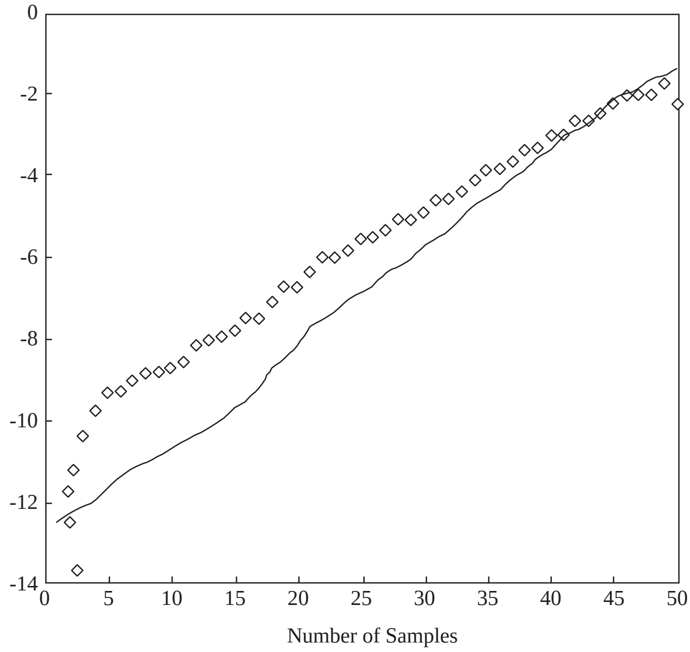

Figure 8. De-scaled I_b values, PKI with WNN (desired –, obtained ◊).

The comparison of training and test phases are given in Tables 1 and 2 for scaled values, respectively. From Table 1, it can be observed that there is a dramatic change in number of iterations during training on the behalf of PKI method. It also has to be pointed out that one iteration takes less time in PKI, so training time is shortened. In order to understand how scaling can effect the results, the descaled values for training are given in Table 3.

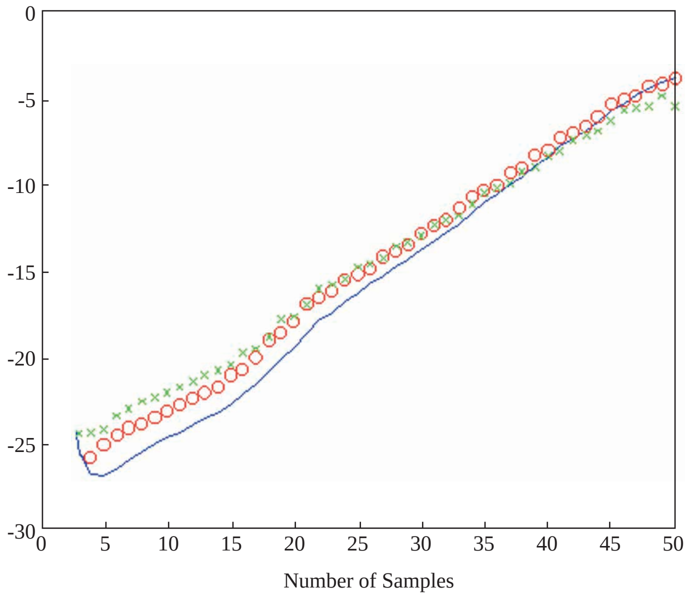

Figure 9. Scaled I_c values ($\times$-WNN, o-MLP, —Experiment results).

Table 1. Scaled training set values.

	Average error	Max. error	iteration
MLP without PKI	$1.404*10^{-2}$	0.021 (% 2.3)	27
MLP with PKI	$0.8901*10^{-3}$	$2.68*10^{-2}$ (% 1.74)	8
WNN with PKI	$3.1023*10^{-4}$	$1.76*10^{-3}$ (% 1.54)	8

Table 2. Scaled test set values.

	Average error	Max. error
MLP without PKI	$2.9*10^{-2}$	0.01874 (% 3.12)
MLP with PKI	$1.104*10^{-3}$	$2.1061*10^{-2}$ (% 2.093)
WNN with PKI	$7.805*10^{-4}$	$4.525*10^{-3}$ (% 1.97)

Table 3. De-scaled training set values.

	Average error	Max. error
MLP without PKI	$3.753*10^{-3}$	0.0485 (% 2.1)
MLP with PKI	$4.5410*10^{-4}$	0.0151 (% 1.223)
WNN with PKI	0.0024	0.0417 (% 1.087)

5. Conclusion

In this work, PKI method is investigated as a means of constructing models. Two different feed-forward structures, namely, MLP and WNN are implemented in MATLAB® environment. Using WNN in PKI method is a novel contribution as MLP is commonly used feed-forward structure in most applications. It can be observed from tables that not only shortened duration of training but also generalisation of PKI method is superior over ANN without PKI.

Even though modelling BJT is not a big challenge, the proposed advantages of PKI method in literature are recognised. By using WNN in PKI method, it is observed that there is no big difference between MLP and WNN structures. Moreover, for this application it is observed that MLP can be preferred over wavelet-based network in PKI, because even though average error and maximum error of wavelet-based network are lower than that of MLP type, slope of MLP type is very close to desired output.

References

[1] Q.J. Zhang, K.C. Gupta, Neural Networks for RF and Microwave Design , Arctech House, 2000.

[2] Bakr Muhammed H., Bandler John W., Madsen Kaj, Sondergaard Jacob, "Review of The Space Mapping Approach to Engineering Optimisation and Modelling", http://www.sos.mcmaster.ca

[3] F. Wang, Q.J. Zhang, Knowledge-Based Neural Models for Microwave Design, IEEE Trans. on Microwave Theory and Techniques, Vol. 45, No.12, December 1997.

[4] L. Chao, X. Jun, X. Liangjin, Knowledge-Based Artificial Neural Network Models for Finline, International Journal of Infrared and Millimeter Waves, Vol. 22, No. 2, 2001

[5] P.M. Watson, K.C. Gupta, R.L. Mahajan, Applications of Knowledge-Based Artificial Neural Network Modeling to Microwave Components, Int. J. RF and Microwave CAE, 9:254-260, 1999.

[6] Q. Zhang, Using Wavelet Network in Nonparametric Estimation, IEEE Trans. On Neural Networks, Vol.8, No. 2, march 1997.

[7] S. Haykin, Neural Networks: A Comprehensive Foundation, Prentice Hall, 1999.