Modeling Broadband Microwave Structures by Artificial Neural Networks

Zbyněk RAIDA, Zbyněk LUKEŠ, Viktor OTEVŘEL

Dept. of Radio Electronics, Brno University of Technology, Purkyňova 118, 612 00 Brno, Czech Republic

raida@feec.vutbr.cz

Abstract. *The paper describes the exploitation of feed-forward neural networks and recurrent neural networks for replacing full-wave numerical models of microwave structures in complex microwave design tools. Building a neural model, attention is turned to the modeling accuracy and to the efficiency of building a model. Dealing with the accuracy, we describe a method of increasing it by successive completing a training set.*

Neural models are mutually compared in order to highlight their advantages and disadvantages. As a reference model for comparisons, approximations based on standard cubic splines are used.

Neural models are used to replace both the time-domain numeric models and the frequency-domain ones.

Keywords

Artificial neural networks, frequency-domain finite elements, time-domain method of moments, wire antennas, microwave transmission lines.

1. Introduction

Modern communication services require wider and wider frequency bands for their operation. Since lower frequency bands are out of their capacity today, broadband services have to operate on higher microwave frequencies.

Designing broadband microwave communication systems, efficient modeling tools are required. These modeling tools have to be based on numeric solving Maxwell's equations. If a harmonic steady state is assumed, then so called frequency-domain methods are used for analysis. If a general excitation containing a large number of harmonic components is assumed, then we analyze the structure by time-domain methods [1]–[3].

Exploiting frequency-domain numerical methods, a broadband microwave structure has to be analyzed on each harmonics from the examined frequency band separately. In the case of time-domain analysis, a short pulse containing spectral components from the whole examined frequency band excites the structure of interest. Within a single analysis, information about the structure behavior over the whole frequency band is obtained [1], [2].

Numerical analysis both in the time domain and in the frequency one is rather time-consuming, which complicates their potential usage in complex design tools. Numerical models are therefore to be replaced by closed-form approximate formulae, or by neural models.

Neural networks are electronic systems, which are built according to a human brain: they contain a large number of the same non-linear building blocks (neurons), they are highly parallel, they are organized in layers, and they are able to learn [4], [5]. The structure of a typical neural network is depicted in Fig. 1.

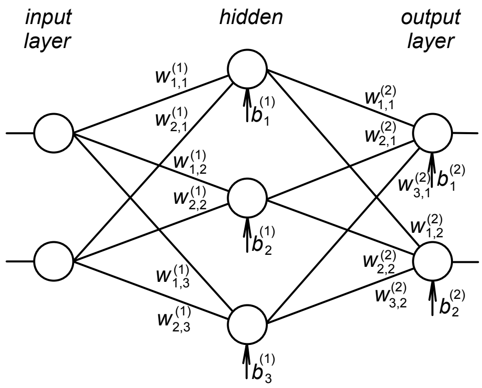

Fig. 1. An example of an artificial neural network. The symbol $w_{i,j}^{(n)}$ denotes a synaptic weight between the output of the ith neuron in the layer (n-1) and the input of jth neuron in the layer n. The symbol $b_i^{(n)}$ denotes a threshold of the ith neuron in the nth layer.

Neurons in the network (circles[1] in Fig. 1) perform a few basic mathematical operations only: they multiply incoming signals by variable coefficients (synaptic weights $w_{i,j}^{(n)}$ and thresholds $b_i^{(n)}$), they sum the products and evaluate a non-linear function for the sum of products [4]. Hence, the neuron operation is computationally efficient.

[1] In the diagram depicted in Fig. 1, the circles represent the neurons in general. In a real neural network, we can replace circles by adaptive non-linear neurons (Fig. 3), adaptive non-linear feedback neurons (Fig. 5) or other types of neurons [4].

Since the neural network is able to learn, we can train it to behave a similar way as a numerical model. Therefore, a properly trained neural network can replace a computationally inefficient numerical model in the design tools.

Let us explain the whole training procedure on an example when a neural model of a symmetric wire dipole of the length h and of the wire radius a is developed. The neural model is asked to produce input resistance $R(h,a)$ and input reactance $X(h,a)$ of the antenna on a fixed prescribed frequency f. Then, the whole training consists of the following steps:

- **Creating training patterns.** Using a numerical method, the dipole is analyzed for various lengths h_1, h_2, to h_M and for various radii a_1, a_2, ..., a_N. Input patterns are formed by the doublets $[h_m, a_n]$, where $m = 1, 2,$ to N and $n = 1, 2, \dots, N$. Output targets consist of corresponding input impedances $[R(h_m, a_n), X(h_m, a_n)]$. An input pattern and an output target give a training pattern together. All the training patterns form the training set.
- **Building neural network.** We create a neural network consisting of an estimated number of layers and of an estimated number of neurons in the layers. The number of neurons in the input layer is given by the number of parameters in the input pattern (two in our example, due to the input doublets $[h_m, a_n]$). The number of neurons in the output layer is determined by the number of parameters in the output target (again two in our example, due to the output doublets $[R(h_m, a_n), X(h_m, a_n)]$). The number of hidden layers and hidden neurons have to be estimates; the discussion is given in the section 2.1 for feed-forward neural models and in the section 2.2 for recurrent neural models. Synaptic weights are set randomly.
- **Training neural network.** During the training, input patterns $[h_m, a_n]$ are successively introduced into the inputs of the neural network, and synaptic weights are changed to reach desired output responses $[R(h_m, a_n), X(h_m, a_n)]$. The training is finished when the network reacts properly to all the input patterns from the training set.
- **Verifying neural model.** We introduce such patterns to the inputs of the neural model, which differ from the input patterns of the training set. Exploiting the numerical model, correctness of the response of the network is verified. If the response is incorrect, additional training patterns have to be prepared, and training has to be repeated over a larger training set.
- **Using neural model.** The trained neural network produces output responses with a sufficient accuracy both for training patterns and for interlaying input patterns. Therefore, the neural network can replace the numerical model.

The trained neural network provides a special approximation in a fact: the *exact* results of the numerical analysis, which are hidden in the training patterns, are used for neural computing approximate results, which correspond to input parameters differing from input patterns. That way, a computationally modest neural network can replace a numerical analysis for parameters differing from training patterns.

In Section II, neural networks are used to replace frequency-domain finite-element model of a shielded microstrip transmission line in a layered medium in the frequency band from 10 GHz to 80 GHz.

In Section III, neural networks are exploited to approximate time-domain moment-method model of a symmetric wire dipole in the frequency band from 0 to 2 GHz.

In Section IV, we give more general conclusions and hints for building neural models of broadband structures.

2. Frequency-Domain Broadband Modeling

Frequency-domain modeling of broadband structures is going to be explained on a shielded microstrip transmission line in a layered medium (Fig. 2). The transmission line is assumed to be longitudinally homogeneous, the shielding waveguide and the microstrip are perfectly electrically conductive, and all the parts are lossles. All the parameters of the structure are fixed except of dielectric constants of layers. Dielectric constants can vary within the interval from $\varepsilon_{r1,2} = 2$ to $\varepsilon_{r1,2} = 4$.

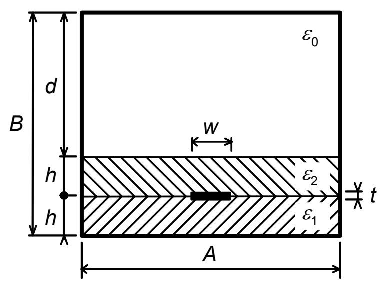

Fig. 2. A shielded microstrip transmission line in a layered medium. Height and width of the shielding waveguide $A = B = 12.7$ mm, width of the microstrip $w = 1.27$ mm, thickness of the microstrip $t \approx 0$, and height of dielectric layers $h = 0.635$ mm are assumed being fixed. Dielectric constants of a substrate and a superstrate are variable. The whole structure is lossles. Metallic parts are perfect electric conductors.

In order to get training patterns, the transmission line is analyzed by frequency-domain finite elements of the nodal-edge nature [6], [7]. The analysis results in values of the dominant-mode propagation constant on frequencies between 10 GHz and 80 GHz for various couples of dielectric constants ε_{r1} and ε_{r2}. The neural model is asked to provide propagation constants over the three-dimensional space $\varepsilon_{r1} \in \langle 2, 4 \rangle$, $\varepsilon_{r2} \in \langle 2, 4 \rangle$, $f \in \langle 10\text{ GHz}, 80\text{ GHz} \rangle$.

First, dielectric constants ε_{r1} and ε_{r2} are changed with $\Delta\varepsilon_r = 1$, and frequency is changed with $\Delta f = 10$ GHz. That way, three values of ε_{r1}, three values of ε_{r2} and eight values of f are obtained. The analysis is performed $3 \times 3 \times 8 = 72$ times to complete each input triplet [ε_{r1}, ε_{r2}, f] by a corresponding propagation constant β. On the basis of training patterns, the neural network is asked to estimate the propagation constant in the three-dimensional space [ε_{r1}, ε_{r2}, f].

2.1 Feed-Forward Neural Network

Feed-forward neural networks are characteristic by a direct signal flow from the input layer to the output one without any feedback. Therefore, this type of networks simply statically maps input patterns into output targets.

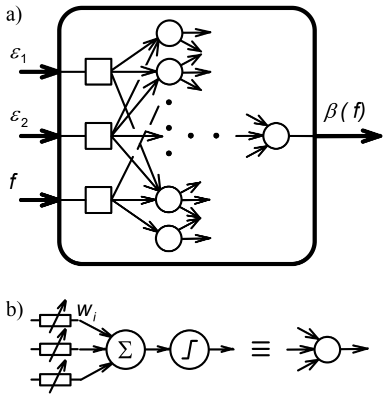

Fig. 3. a) Architecture of a feed-forward neural network for modeling a shielded microstrip transmission line in layered medium. Input layer consists of input neurons[2] (squares). Other layers contain adaptive non-linear neurons (circles). b) Structure of an adaptive non-linear neuron.

Since input patterns are triplets [ε_{r1}, ε_{r2}, f] in our case, the input layer has to consist of three neurons. Since the output target is a corresponding propagation constant β, there is a single neuron in the output layer. The structure of the neural model is depicted in Fig. 3.

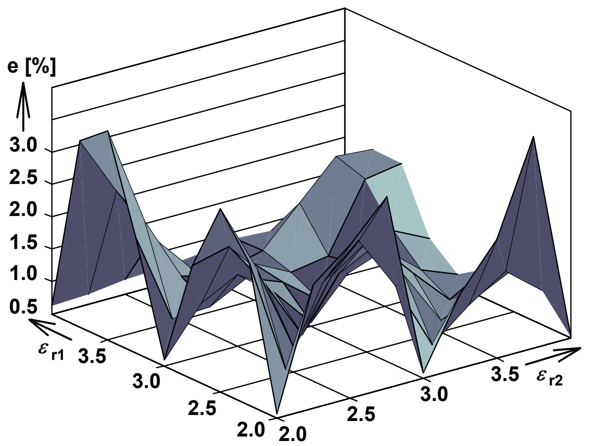

Fig. 4. Cumulative error distribution for a neural model of a shielded microstrip transmission line in a layered medium. The neural model is based on the feed-forward neural network and on the equidistant training patterns.

[2] Input neurons can be understood as dividers, which simply distribute the input signal to all the neurons in the following layer.

The number of hidden layers and neurons in hidden layers has to be estimated. As shown in [8], the neural network should consist of so many neurons so that the number of effectively used parameters (weights and biases) varies between 70 % and 90 % of the total number of the network parameters. The number of effectively used parameters can be estimated by the Bayesian regularization [9].

In our development, we started with two hidden layers consisting of eight neurons. The Bayesian regularization estimated 50 % of network parameters being effectively used. Consecutively, we decreased the number of neurons in hidden layers up to five in each, which corresponds to 82 % of effectively used parameters.

Quality of the neural model can be expressed by the cumulative error [8]

$$e(\varepsilon_{r1}, \varepsilon_{r2}) = 100 \cdot \sum_{n=1}^{8} \frac{\left| \beta_a(\varepsilon_{r1}, \varepsilon_{r2}, f_n) - \beta_n(\varepsilon_{r1}, \varepsilon_{r2}, f_n) \right|}{\beta_n(\varepsilon_{r1}, \varepsilon_{r2}, f_n)} \qquad (1)$$

In (1), we compute the absolute value of the percentage error of the propagation constant computed by the neural model β_n with respect to the propagation constant obtained by the finite-element analysis β_a, and we sum the percentage errors over all frequencies $f_1 = 10$ GHz, $f_2 = 20$ GHz to $f_8 = 80$ GHz.

Cumulative error of the above-described neural model of the shielded transmission line is depicted in Fig. 4. Whereas the cumulative error is negligible in the training points it reaches up to 3 % in the interlaying points.

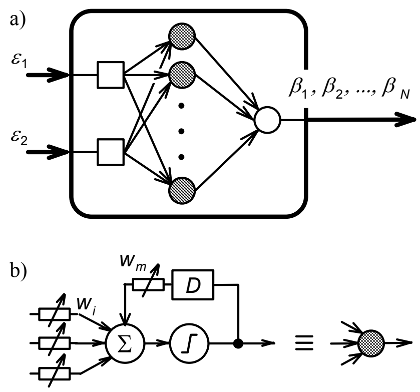

Fig. 5. a) Architecture of a recurrent neural network for modeling a shielded microstrip transmission line in a layered medium. The input layer consists of input neurons (squares). In a hidden layer, adaptive non-linear feedback neurons are used (hatched circles). The output layer consists of usual adaptive non-linear neurons (see Fig. 3b). b) Structure of an adaptive non-linear feedback neuron (D denotes a delay block).

Now, we try to repeat the described neural model development when using a recurrent neural network.

2.2 Recurrent Neural Network

Due to the feedback in the structure (see Fig. 5), the recurrent neural network is able to map an input sequence into an output sequence. In case of broadband modeling of the shielded microstrip transmission line, doublets of dielectric constants can play the role of input patterns, and a sequence of propagation constants on frequencies 10 GHz, 20 GHz, …, 80 GHz is a target sequence.

Structure of the recurrent model of the transmission line is depicted in Fig. 5. The input layer consists of two neurons (the input for frequency was removed), the output layer is identical with the output of the feed-forward neural network. Dealing with hidden layers, recurrent networks contain a single hidden layer of feedback neurons.

In case of our model, the optimum number of hidden neurons was estimated to 5. For higher number of neurons, behavior of the neural model did not change. For lower number of neurons, the neural model exhibited unacceptably high error.

Cumulative error of the recurrent neural model is depicted in Fig. 6. The maximum of the error is half-size compared to the feed-forward model.

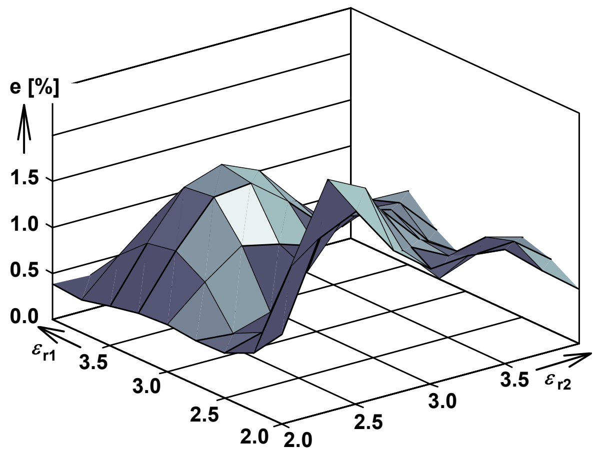

Fig. 6. Cumulative error distribution for a neural model of a shielded microstrip transmission line in a layered medium. The neural model is based on the recurrent neural network and on the equidistant training patterns.

Now, we compare results of the neural modeling and a standard approximation based on cubic splines.

2.3 Cubic Splines

If functional values in interlaying points should be evaluated using functional values in equidistantly distributed independent points, splines are usually exploited [10]. In case of cubic splines, the unknown function is approximated by piecewise-cubic function.

Applying cubic splines to computing propagation constants of shielded microstrip transmission line, we use two-dimensional approximation over the plane of dielectric constants ε_{r1} and ε_{r2}. Such an approximation is applied to each frequency from [10 GHz, 20 GHz, …, 80 GHz] separately: 8 approximations for triplets [ε_{r1}, ε_{r2}, 10 GHz], [ε_{r1}, ε_{r2}, 20 GHz], …, [ε_{r1}, ε_{r2}, 80 GHz] are performed.

For the described approximation technique, the distribution of the cumulative error is depicted in Fig. 7. Comparing the error of splines to the error of the feed-forward model (Fig. 4) and the recurrent model (Fig. 6), the cumulative error can be concluded to be the same for the feed-forward model and the splines whereas for the recurrent model, the cumulative error is half-sized.

Nevertheless, accuracy of neural models can be increased by adding new training patterns into the initial equidistant training set (the training set becomes non-equidistant). Such an approach is inapplicable in the case of splines.

2.4 Non-Equidistant Training Patterns

Following [8], accuracy of a feed-forward neural model is increased by finer spatial sampling in areas, which are associated with the largest values of the cumulative error. In our case (Fig. 4), we can identify five above-described areas, which surround points [ε_{r1}, ε_{r2}] = [2.5, 2.0], [2.5, 3.0], [2.5, 4.0], [3.5, 2.0], [3.5, 4.0]. Therefore, 40 new analyses are performed (for each new point, phase constant on eight frequencies is evaluated), and the so-far existing training set is completed by newly constructed patterns. The structure of the neural model stays unchanged (two hidden layers consisting of five neurons). Cumulative error of such a model is depicted in Fig. 8.

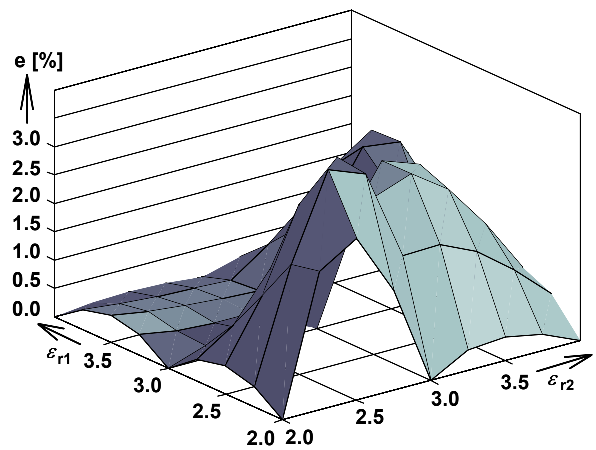

Fig. 7. Cumulative error distribution for a spline model of a shielded microstrip transmission line in a layered medium.

Comparing the cumulative error distribution of the equidistant feed-forward model and the non-equidistant one, the maximum value of the cumulative error is concluded being reduced for approximately 50 %. The result is twice better compared to cubic splines, which cannot be improved by the non-equidistant sampling technique.

In the next step, the same approach is applied to improving accuracy of the recurrent model.

Investigating Fig. 6, the maximum cumulative error is revealed in the vicinity of points [ε_{r1}, ε_{r2}] = [2.0, 2.5], [2.5, 3.5], [3.5, 2.5], [3.5, 3.0]. Therefore, 32 new analyses are performed (for each of four new points, phase constant is evaluated on eight frequencies), and the so-far existing training set is completed by those new patterns. Structure of the neural network stays unchanged again (a single hidden layer consisting of five feedback neurons), and the neural model is re-trained. The resultant cumulative error is depicted in Fig. 9.

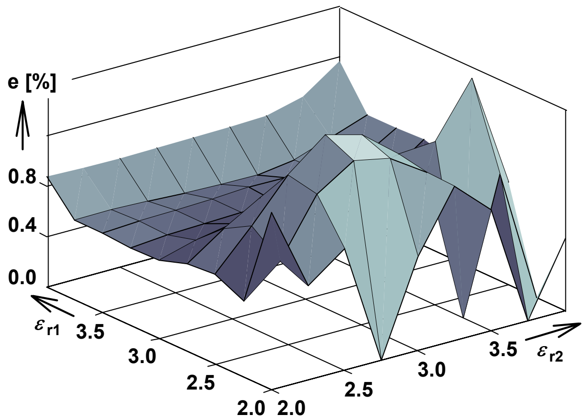

Fig. 8. Cumulative error distribution for a neural model of a shielded microstrip transmission line in a layered medium. The neural model is based on the feed-forward neural network and on the non-equidistant training patterns.

Maximum value of cumulative error reaches 0.5 %, which is twice better compared to a non-equidistant feed-forward model and 4-times better compared to the spline model.

In the next step, CPU-time demands of creating models were compared. Comparisons were performed on a PC equipped with the processor AMD Athlon 2 GHz, with the 1 024 MB RAM and the operating system Windows 2000. For computations, MATLAB 6.1 and the Neural Network Toolbox 4.0 were used. Creating feed-forward model consumed 4.47 seconds of CPU-time[3]. When creating recurrent model, 63.61 seconds of CPU-time were consumed, which is approximately 15-time longer time. On the other hand, the longer creation time of recurrent models is compensated by more stable training process compared to the feed-forward networks: training feed-forward networks is more sensitive to setting a proper number of hidden neurons. The spline model does not require any training, and therefore, creation CPU-time demands are zero.

2.5 Conclusions

The above-described models of the shielded microstrip transmission line in a layered medium can be evaluated from the following points of view:

- Model accuracy;
- CPU-time demands of creation process;
- Elaborateness of model creation;
- Approximation abilities of models.

Model accuracy. Recurrent models are more accurate than feed-forward ones (accuracy is approximately twice better). Whereas feed-forward models suffer from a possible increase of the cumulative error in interlaying points, recurrent models do not exhibit such a disadvantage.

CPU-time demands of creation process. Training recurrent models is approximately 15-times more CPU-time demanding compared to feed-forward ones. This disadvantage is compensated by higher accuracy and by the immunity to the increase of the cumulative error in interlaying points.

Elaborateness of model creation. Both the feed-forward neural model and the recurrent one were trained using very similar training sets (in the case of the recurrent model, one training pattern is saved), and therefore, both time and energy needed for their preparation are comparable. Whereas in the case of feed-forward models the care has to be taken for the approximation error increase in interlaying points, recurrent models are more robust from that point of view.

Approximation abilities of models. Approximation abilities of recurrent models are better compared to the abilities of feed-forward models: recurrent models do not suffer from approximation error increase in interlaying points.

The above-given conclusions cannot be formulated by generalizing experience with models of a single structure. They were verified by modeling a plenty of antennas, transmission lines and passive microwave circuits. In this paper, the experience was summarized and illustrated by a single example.

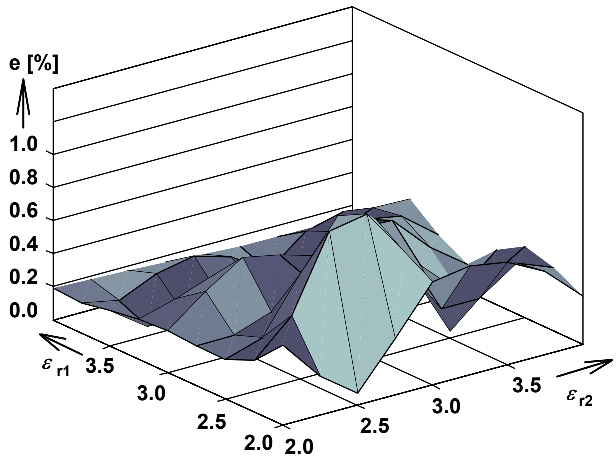

Fig. 9. Cumulative error distribution for a neural model of a shielded microstrip transmission line in a layered medium. The neural model is based on the recurrent neural network and on the non-equidistant training patterns.

In the next section, the described principles of neural modeling of broadband structures are going to be applied to developing time-domain neural models of microwave structures.

[3] The given CPU-time of creating neural models covers training only. If the training is not successful (revealed in the verification phase), new training patters have to be built and the training has to be repeated, which significantly increases CPU-time demands.

3. Time-Domain Modeling

Creating a neural model of a microwave structure in the time domain, the neural network is asked to map an input pattern into a time sample of the output quantity (in case of feed-forward models) or into a series of time samples (in case of recurrent models).

The only difference between the frequency-domain broadband modeling and time-domain modeling consists “in the horizontal axis labeling”. Whereas the horizontal axis of the frequency-domain model expresses frequency, the horizontal axis of the time-domain model denotes time.

Principles of building time-domain neural models are going to be explained for a symmetric wire dipole (Fig. 10). The radius of the antenna wire is assumed being much smaller compared to the wavelength and the length of the dipole arm. Then, a thin-wire approximation can be used (currents and charges are concentrated of the wire axis [11]). Therefore, the numerical analysis is a one-dimensional problem. During the analysis, the length of the dipole is changed in the interval $h \in$ <1 m; 3 m>, and the wire radius can vary in the interval $a \in$ <10 mm; 90 mm>.

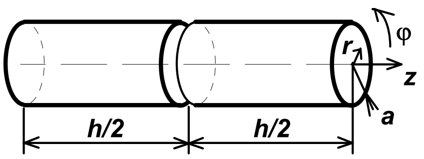

Fig. 10. Symmetric wire dipole from perfect electric conductor. The antenna is excited by plane wave of normal incidence. Electric field vector is of the same direction as antenna axis.

The antenna is analyzed by time-domain method of moments [1]. The dipole is assumed to be fabricated from the perfect electric conductor and to be placed in vacuum. The antenna is excited by a plane wave, which propagates perpendicularly to the antenna axis. The vector of electric field intensity of the excitation wave has the same orientation as the antenna wire. Time course of the excitation wave is of the character of the Gaussian pulse. The width of the pulse is 2 LM[4]. The analysis results to the time course of the current in the excitation gap of the dipole.

In the first attempt, input quantities are changed with the steps Δh = 1 m and Δa = 40 mm. The initial training set therefore consists of nine input patterns, which correspond to all the possible combinations of lengths h = [1,0 m; 2,0 millimeters; 3,0 m] and radii a = [10 mm; 50 mm; 90 mm]. The output targets I_{mn} consist of 100 time samples of a current in the excitation gap of the dipole [h_m, a_n], where m, n are 1, 2, 3. The transient current is sampled in the interval t = 40 LM with the sampling step Δt = 0.4 LM. An example of the output target is depicted in Fig. 11.

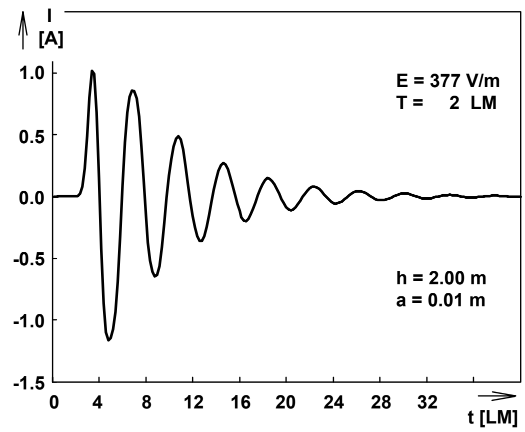

Fig. 11. Time course of the transient current in the excitation gap of the dipole.

Accuracy of time-domain neural models is quantified by the cumulative error again. Whereas a relative error was cumulated in the case of frequency-domain modeling, an absolute error has to be cumulated in the time domain because the output quantity can be positive, negative and zero. The absolute value of the absolute error is cumulated over all 100 time samples

$$e(a,h) = \frac{1}{100} \sum_{n=1}^{100} \left| I_{exc}^{app}(a,h,n) - I_{exc}^{num}(a,h,n) \right| . \quad (2)$$

Here, $I_{exc}^{app}(a, h, n)$ denotes the value of nth time sample of the current in the excitation gap of the dipole, which length is h and which radius equals to a; the current is computed by a neural model. Next, $I_{exc}^{num}(a, h, n)$ denotes the current computed by the numerical analysis.

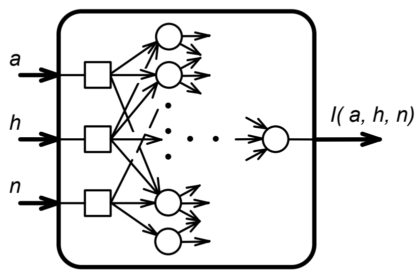

Fig. 12. Architecture of a feed-forward neural network for time-domain modeling symmetric wire dipole. Squares symbolize input neurons (a simple distribution of input signals), circles symbolize adaptive non-linear neurons (see Fig. 3).

In following sections, a way of approximating transient currents by neural networks and splines are described.

3.1 Feed-Forward Neural Network

The feed-forward neural network maps input triplets, which consist of the dipole length h, antenna wire radius a and the index of the current sample n, to the respective values of the current in the excitation gap $I(a, h, n)$.

Therefore, the feed-forward neural model consists of three input neurons and a single output neuron. Exploiting Bayessian regularization, the optimal number of hidden layers is set to three, and the optimal number of neurons in hidden layers is set to [24, 17, 24]. A relatively small num-

4 The abbreviation LM denotes light meter. Light meter equals to the time, which is consumed by electromagnetic wave in free space to travel the distance of one meter: light meter equals to $(3 \cdot 10^8)^{-1}$ s.

ber of neurons in the middle hidden layer implements the so-called bottleneck, which can prevent the neural network to over-training [4]. Over-training causes the increase of the approximation error in the interlaying points.

Cumulative error distribution[5] of the feed-forward model of the wire dipole is depicted in Fig. 13. The error is of a negligible value for nine training points. In interlaying points, the error reaches 0.15 A. This maximum error was detected for input points [*h*, *a*] = [1.6 m; 30 mm], [1.6 m; 70 mm], [2.6 m; 30 mm] and [2.6 m; 70 mm].

Now, a similar model is going to be developed when a recurrent neural network is used.

3.2 Recurrent Neural Network

Structure of the recurrent neural model of the wire dipole is depicted in Fig. 14. In the input layer, an input neuron introducing information about the index of the time sample of the excitation current *n* is missing: due to the dynamic behavior of the network, the recurrent model produces the whole sequence of 100 time samples of the current as a response to the input pattern [*a*, *h*].

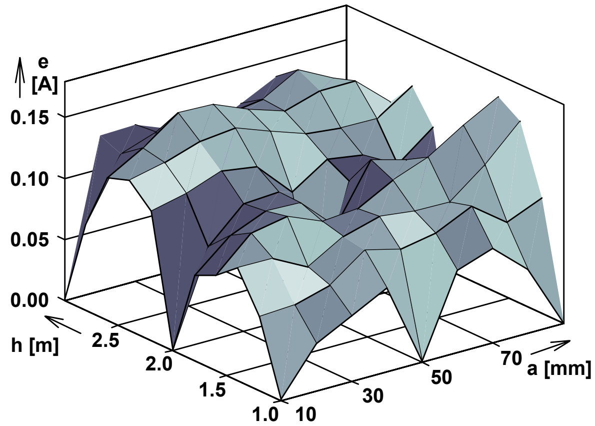

Fig. 13. Cumulative error distribution for a time-domain neural model of a wire dipole. The neural model is based on the feed-forward neural network and on the equidistant training patterns.

Hence, the recurrent network consists of two input neurons. The optimum behavior of the model was reached when the hidden layer obtained 107 recurrent neurons.

Cumulative error distribution of the time-domain recurrent model is depicted in Fig. 15. Comparing to the feed-forward model, two main differences are observed:

- Cumulative error is approximately twice smaller compared to the feed-forward model (see Figs. 13, 15);
- There are no deep minims in the error distribution in the points representing input patterns (Figs. 13, 15).

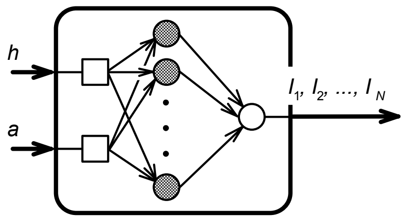

Fig. 14. Architecture of a recurrent neural network for time-domain modeling symmetric wire dipole. Squares symbolize input neurons (a simple distribution of input signals), empty circles symbolize adaptive non-linear neurons (see Fig. 3), hatched circles are recurrent neurons (see Fig. 5).

The price, which has to be paid for better properties of the recurrent model, is a longer training time. Whereas the feed-forward network was trained within 300 iterations for the training error[6] 10^{-5}, the recurrent network required 3000 iterations for the error 10^{-4}. Due to the feedbacks in the recurrent model, the training is more complicated.

The maximum error was detected for points [*h*, *a*] = [1.6 millimeters; 30 mm], [1.6 m; 70 mm], [2.6 m; 30 millimeters] a [2.6 m; 70 mm], which are identical with maxims of the feed-forward model.

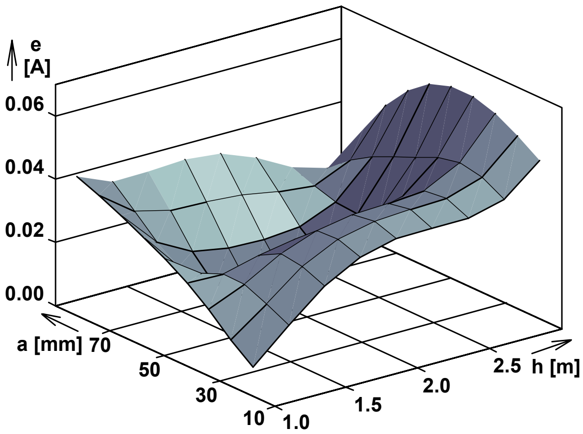

Fig. 15. Cumulative error distribution for a time-domain neural model of a wire dipole. The neural model is based on the recurrent neural network and on the equidistant training patterns.

In the next section, we describe a spline model, which is analogical to the above-described neural models. For the development, cubic splines are used.

[5] In the time domain, the cumulative error has to be expressed as the absolute error due to the zero value of the current in several time instants: dividing the absolute error by the numerically computed current tends to infinity if the numerically computed current approaches zero.

[6] The training process can be understood as an optimization problem in fact: the state variables (biases and weights are set so that the squared difference between the actual output of the neural network and the required output training patterns is minimized. The difference between the actual output and the desired one is expressed by the training error.

3.3 Cubic Splines

For time-domain wire-dipole modeling, two-dimensional cubic splines are used (approximations over the plane *ah*). Since the model consists of 100 *ah* planes (for each time sample of the current in the excitation gap), the spline model can be understood as 100 parallel approximations for triplets [*a*, *h*, 1], [*a*, *h*, 2], …, [*a*, *h*, 100].

For cubic splines, cumulative error distribution is computed the same way as for neural models. The result is depicted in Fig. 16.

Fig. 16 shows the absolute value of the cumulative error being of the same level as the feed-forward neural model exhibited (error of the recurrent model is approximately twice lower). Whereas the accuracy of neural models can be improved by finer sampling in the areas of the maximum error, accuracy of a spline model cannot be improved that way.

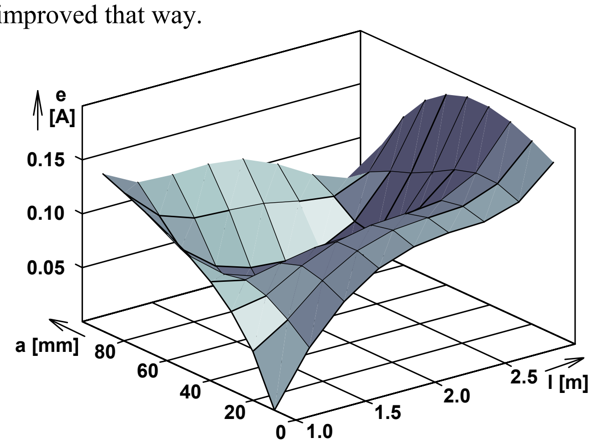

Fig. 16. Cumulative error distribution for a time-domain spline model of a wire dipole.

In the next section, the training set is completed by the additional training patterns corresponding to the doublets [*h*, *a*] = [1.6 m; 30 mm], [1.6 m; 70 mm], [2.6 m; 30 mm] and [2.6 m; 70 mm]. For those doublets of wire-antenna parameters, both the feed-forward model and the recurrent one exhibited the maximum cumulative error.

3.4 Non-Equidistant Training Patterns

In the case of the feed-forward neural model, a larger training set has to be respected by changing the number of hidden layers and their neurons: instead of three layers consisting of [24, 17, 24] neurons, only two hidden layers of [51, 51] neurons were used: the neural network of the original architecture was not able to absorb new information hidden in new training patterns.

Cumulative error distribution of the non-equidistant neural model is depicted in Fig. 17. In the distribution, four deep minims, which correspond to new training patterns, can be observed. The cumulative error in the interlaying points was twice reduced comparing to the equidistant model (Figs. 13, 17). The reached accuracy of the feed-forward non-equidistant model is approximately twice better compared to splines (Fig. 16), and approximately the same compared to the recurrent equidistant model (Fig. 15).

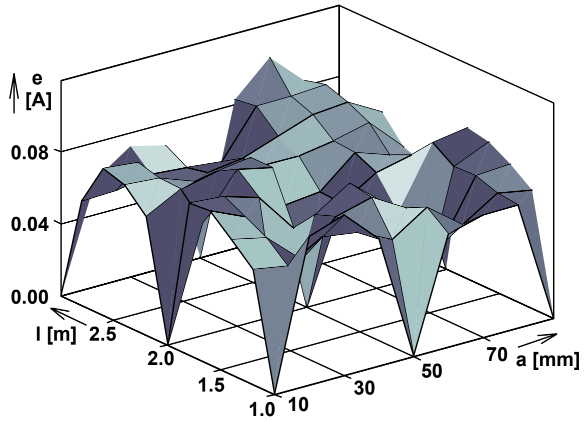

Fig. 17. Cumulative error distribution for a time-domain neural model of a wire dipole. The neural model is based on the feed-forward neural network and on non-equidistant patterns.

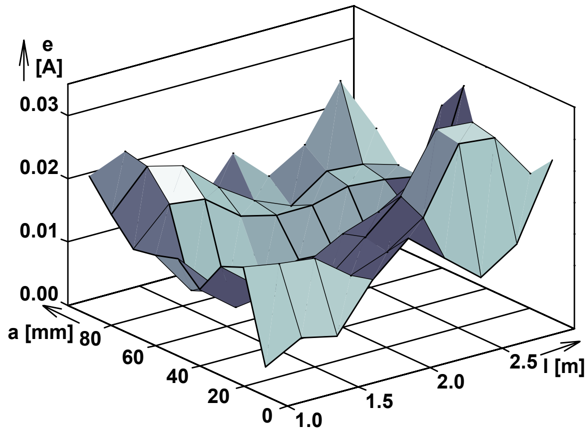

Fig. 18. Cumulative error distribution for a time-domain neural model of a wire dipole. The neural model is based on the recurrent neural network and on the non-equidistant patterns.

Dealing with the recurrent non-equidistant model, the architecture has to be modified: instead of 107 recurrent neurons, 142 recurrent neurons have to be used.

Cumulative error of the recurrent non-equidistant model is depicted in Fig. 18. Accuracy of this model is approximately twice better compared to the recurrent equidistant model (Fig. 15) and twice better compared to the feed-forward non-equidistant model (Fig. 17).

We can therefore conclude that statements presented in the paragraph 2.5 stay valid.

4. Conclusions

In the paper, exploitation of artificial neural networks to the automatic creation of models of broadband electromagnetic structures, both in the time domain and in the frequency domain, was described. We demonstrated there is no principal difference between models in both the domains.

Both the feed-forward neural networks and the recurrent ones were used for building models. Neural models were compared with cubic-spline approximations as the reference solution. The possibility of improving accuracy of neural models by generating additional training patterns was highlighted as the main advantage of the neural approach.

Recurrent neural models were shown to be more robust than feed-forward ones. Whereas feed-forward models are sensitive to the number of hidden neurons, and stand-alone realizations can exhibit dramatic differences in the reached approximation error, sensitivity of recurrent models to the number of hidden neurons is smaller, and stand-alone realizations are of very similar behavior. The price, which has to be paid for better properties of recurrent networks, are higher CPU-time demands of training.

The resultant feed-forward models differ from the recurrent ones. Whereas feed-forward models tend to over-training (high approximation error in interlaying points), recurrent models are more immune to this phenomenon.

The possibility of automating model development is the main advantage of neural networks. During the model development, approximation abilities of models are tested in interlaying points, and in the case of higher approximation error, the respective areas are sampled more finely.

Utilization of the finished neural model is very quick: output responses are formed using a relatively low number of additions, multiplications and evaluations of a non-linear function (in real arithmetic). If neural models are accurate enough, they can be used in complex design tools as an accurate and efficient model of various microwave components.

Acknowledgements

Research described in the paper was financially supported by the Czech Grant Agency under grants no. 102/01/0571 and 102/01/0573. Further financial support was obtained from the Czech Ministry of Education under programs MSM 262200011 and MSM 262200022.

References

[1] RAO, S. *Time domain electromagnetics*. San Diego: Academic Press, 1999.

[2] RAIDA, Z. et al. *Analýza mikrovlnných struktur v časové oblasti* (*Time-domain analysis of microwave structures*). Brno: VUTIUM Publishing, 2003.

[3] ČERNOHORSKÝ, D., RAIDA, Z., ŠKVOR, Z., NOVÁČEK, Z. *Analýza a optimalizace mikrovlnných struktur* (*Analysis and optimization of microwave structures*). Brno: VUTIUM Publishing, 1999.

[4] HAYKIN, S. *Neural networks: A comprehensive foundation*. Englewood Cliffs: Macmillan Publishing Company, 1994.

[5] CICHOCKI, A., UNBEHAUEN, R. *Neural networks for optimization and signal processing*. Chichester: J. Wiley & Sons, 1994.

[6] LEE, J. F., SUN, D. K., CENDES, Z. J. Full-wave analysis of dielectric waveguides using tangential vector finite-elements. *IEEE Transactions on Microwave Theory and Techniques*. 1991, vol. 39, no. 8, p. 669 – 678.

[7] LEE, J. F. Finite element analysis of lossy dielectric waveguides. *IEEE Transactions on Microwave Theory and Techniques*. 1994, vol. 42, no. 6, p. 1025 – 1031.

[8] RAIDA, Z. Modeling EM structures in Neural network toolbox of Matlab. *IEEE Antennas and Propagation Magazine*. 2002, vol. 44, no. 6, p. 46 – 67.

[9] DEMUTH, H., BEALE, M. *Neural network toolbox for use with Matlab: User's guide*. Version 4. Natick: The MathWorks Inc., 2000.

[10] DE BOOR, C. *A Practical guide to splines*. Berlin: Springer-Verlag, 1978.

[11] HARRINGTON, R. F. *Field computation by moment methods*. 2nd edition. Piscataway: IEEE Press, 1993.

[12] RAIDA, Z. Broadband design of planar transmission lines: feed-forward neural approach versus recurrent one. In *Proceedings of the International Conference on Electromagnetics in Advanced Applications ICEAA 2003*. Torino (Italy), 2003, p. 155 – 158.

[13] LUKEŠ, Z., RAIDA, Z. Design of microwave antennas: neural network approach to time domain modeling of V-Dipole. In *Proceedings of the International Conference on Electromagnetics in Advanced Applications ICEAA 03*. Torino (Italy), 2003, p. 7 – 10.

[14] RAIDA, Z. Wideband neural modeling of wire antennas: feed-forward neural networks versus recurrent ones. In *Proceedings of the Progress in Electromagnetics Research Symposium PIERS 2003*. Honolulu (Hawaii), 2003, p. 717 – 717.

About Authors...

Zbyněk RAIDA is with the Faculty of Electrical Engineering and Communication (FEEC), Brno University of Technology (BUT). He is interested in exploitation of numerical methods to the analysis of microwave systems, and to the application of genetic algorithms and artificial neural networks to the design of microwave structures.

Zbyněk Lukeš was born in 1979 in Frýdek-Místek (Czech Republic). He received the Ing. (M.Sc.) degree in electrical engineering from the Faculty of Electrical Engineering and Communication, Brno University of Technology (FEEC BUT) in 2002. At the present time, he is Ph.D. student at the Department of Radio Electronics, FEEC BUT. His research is oriented to time-domain analysis of antennas.

Viktor Otevřel was born in 1975. He received the M.Sc. degree in electronics and telecommunication from the Faculty of Electrical Engineering and Communication, Brno University of Technology in 2000. Now, he is a Ph.D student at the Department of Radio Electronics. His major research interests are aimed at the questions concerning the problems of artificial intelligence (in particular, evolutionary computation) and its usage in technical practice. Numerical modeling of electromagnetic structures in time domain using FDTD method is of minor interest to him.