Microwave Cylindrical Cavity Applicators Modeling Using Artificial Neural Networks

ZORAN STANKOVIĆ, BRATISLAV MILOVANOVIĆ, MAJA SAREVSKA
Department of Telecommunications
University of Niš – Faculty of Electronic Engineering
Aleksandra Medvedeva 14, 18000, Niš
SERBIA AND MONTENEGRO
http://www.elfak.ni.ac.yu

Abstract: -This paper presents different approaches of microwave cavity applicators modeling using neural networks. These approaches are developed at the laboratory for Microwave Technique at the Faculty of Electronic Engineering - Niš. First, a model of loaded microwave cylindrical metal cavity based on multilayer perceptron neural network (MLP) is introduced. Then, the method for efficiency improvement (reducing the number of training samples required for neural network training) of proposed MLP models is developed. This method is based on indirect incorporation of the existing knowledge from problem domain into the neural network. This knowledge is defined by approximate-empirical cavity model. Finally, high efficient neural models of microwave cylindrical metal cavity with directly incorporated knowledge (hybrid empirical - neural model and KBNN -knowledge based neural network model) are presented.

Key-Words: Neural networks, microwave cavity, neural models, neurons, microwave applicator.

1 Introduction

The intense development of microwave technique in the last decades has lead to widespread application of microwave applicators in the science, medicine, and in industry. In most cases, these applicators have a form of the cylindrical metallic cavities with various cross-sections (circular, rectangular, elliptical, etc). Microwave cavities loaded by homogeneous dielectric layers have wide range of applications in different microwave systems. They also have a special application in the processes of dielectric material heating and drying by microwave energy. In order to manufacture an efficient cavity, it is necessary to know all the types of oscillations that may appear in it and what are the resonant frequencies [1,2].

A usual approach for theoretical analysis of the cylindrical metallic cavities is based on the application of the transverse resonance method (TRM) [3]. The resonant frequencies are determined from the transcendental characteristic equation. To calculate the resonant frequencies, an appropriate numerical technique and an efficient procedure for mode identification (especially in the case of a multilayer load) [3] are needed. Software implementation of this problem is hardware and time consuming which is main disadvantage of this approach.

The basic common disadvantage of all other numerical techniques which can be applied in cavity modeling (TLM, FDTD, etc) is that they have high demands concerning the hardware resources necessary for their software implementation [1,2,4]. The software implementation itself might be very complicated and faced with many difficulties. Also the time needed for numerical calculation when using a detailed electromagnetic (EM) model could be unacceptably long.

In order to avoid solving of a number of time-consuming complexes electromagnetic equations needed for numerical approaches an original approximate approach for cylindrical metallic cavities modeling is presented in [5]. This approach is based on a huge analytical and semi-empirical research. It avoids any kind of complicated numerical calculation in order to give fast respond. The main disadvantage of this model is that it can only be applied only when there is no need of high accuracy in modeling.

Good alternative for overcoming all these problems is modeling cavities using an artificial neural network [4,6] Neural network model in these cases can be fast as approximate model and accurate as detailed EM models.

2 Knowledge about Cavity Resonant Frequency

A number of different TM/TE$_{mnp}$ modes can be excited in a cylindrical metallic cavity loaded by homogeneous dielectric layer placed at the cavity bottom (Fig.1). Investigations conducted in reference [3] have shown that the resonant frequency f_r of excited mode in such cavity with constant dimensions depends on the relative dielectric permittivity ε_r and filling factor t_h ($t_h = t/h$, where t is thickness of dielectric layer and h is height of the cavity)

$$f_r = f(t_h, \varepsilon_r) \tag{1}$$

Using short-circuit boundary (electric wall) in a interface plane between dielectric slab and air, from the condition of resonance applied separately in air and dielectric part of the cavity, appropriate expressions for resonant frequency calculation in these regions can be easily derived

$$f_r^{(A)}(t_h) = \sqrt{\left(\ell \cdot \frac{f_0}{1-t_h}\right)^2 + f_{c0}^{\;2}} \qquad l = \begin{cases} 0,1,2,\ldots \text{ for TM}_{mnp} \\ 1,2,3,\ldots \text{ for TE}_{mnp} \end{cases} \tag{2}$$

$$f_r^{(D)}(t_h, \varepsilon_r) = \sqrt{\left(k \cdot \frac{f_0}{\sqrt{\varepsilon_r}} \frac{1}{t_h}\right)^2 + \left(\frac{f_{c0}}{\sqrt{\varepsilon_r}}\right)^2} \qquad k = 1,2,3\ldots \tag{3}$$

where: $f_{c0} = ck_c/(2\pi)$ represents the cutoff frequency of a waveguide with the same cross-section as cavity and filled with air, while k_c is a constant that depends on mode of oscillation and waveguide cross-section shape and dimensions; $f_0 = c/(2h)$; and integers l and k are the number of half waves of standing wave for electric field in corresponding part of the cavity [5]. For the cavity of rectangular cross-section with dimension $a \times b$ for TM/TE$_{mnp}$ modes constant k_c is

$$k_c = \sqrt{\left(\frac{m\pi}{a}\right)^2 + \left(\frac{m\pi}{b}\right)^2} \tag{4}$$

while for the cavity of circular cross-section with radius r, constant k_c is

$$k_c = \frac{x_{mn}}{r} \tag{5}$$

where x_{mn} is n-th zero of the Bessel function of the first kind of order m for TM$_{mnp}$ modes and n-th zero of the derivation of the same function for TE$_{mnp}$ modes [2].

Applying open-circuit boundary (magnetic wall) in the interface plane between dielectric and air, from anti-resonant condition in air and dielectric part of the cavity, expressions for anti-resonant frequency calculation in these parts can be found as

$$f_a^{(A)}(t_h) = \sqrt{\left(\frac{2\ell - 1}{2} \cdot \frac{f_0}{1-t_h}\right)^2 + f_{c0}^{\;2}} \qquad l = 1,2,3,\ldots \tag{6}$$

$$f_a^{(D)}(t_h, \varepsilon_r) = \sqrt{\left(\frac{2k-1}{2} \cdot \frac{f_0}{\sqrt{\varepsilon_r}} \cdot \frac{1}{t_h}\right)^2 + \left(\frac{f_{c0}}{\sqrt{\varepsilon_r}}\right)^2} \quad k = 1,2,\ldots \tag{7}$$

Let note characteristic point of considered TM/TE$_{mnp}$ mode family (m=const, n=const) in t_h - f_r plane as RR$_k^l(\varepsilon_r)$ which represents a crossing point of k-th resonant curve in dielectric part of the cavity (3) and l-th resonant curve in air part of the cavity (2), and characteristic point AA$_k^l(\varepsilon_r)$ as a crossing point of k-th

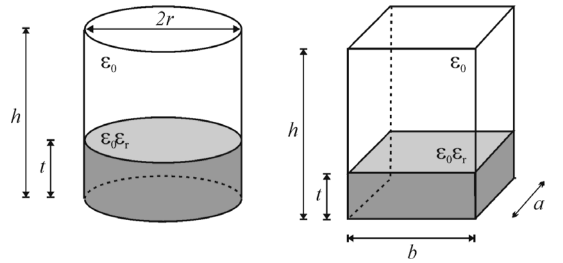

Fig. 1. Microwave cylindrical metallic cavity with (a) circular (b) rectangular cross-section loaded by dielectric layer of thickness t placed at the cavity bottom

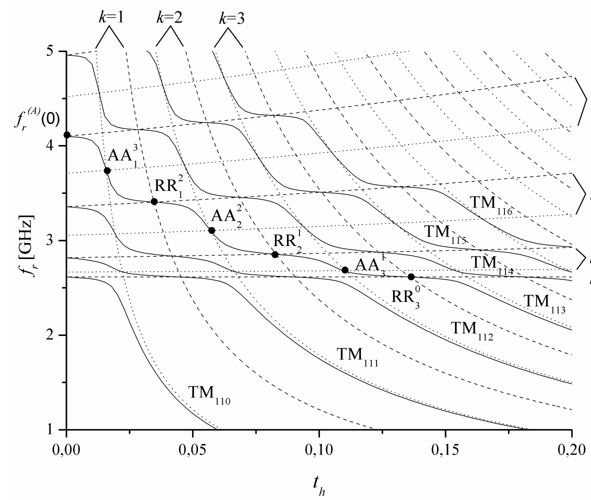

Fig. 2. Family of the resonant frequencies for TM11$_p$ mod obtained using TRM for the cylindrical metallic cavity with circular cross-section (r = 7 cm and h=14.24) loaded with water (ε_r = 80).

- - - resonant curves (monotonous increasing in air part and monotonous. decreasing in dielectric part of the cavity)

· · · anti-resonant curves (monotonous increasing in air part and monotonous. decreasing in dielectric part of the cavity)

anti-resonant curve in dielectric part of the cavity (7) and l-th anti-resonant curve in air part of the cavity (6) loaded with dielectric relative permittivity ε_r (Fig. 2). Detail analysis of cylindrical metallic cavities in reference [5] has shown that resonant frequency curves, for considered TM/TE$_{mnp}$ mode excited in such cavities, are approaching either resonant or anti-resonant curves, passing though characteristic points with the following order: AA$_1^p$, RR$_1^{(p-1)}$, AA$_2^{(p-1)}$, RR$_2^{(p-2)}$,..., AA$_p^{min(l)+1}$, RR$_p^{min(l)}$. The characteristic points are easily found from Eqs. (2), (3), (6) and (7) for known relative permittivity ε_r. The fact that these points describe the behavior of resonant frequency curves (mode tuning behavior) and that they are determined directly by resonant and anti-resonant frequency functions in air and dielectric part of the cavity, given in analytical form, represents a partial knowledge from the problem domain implemented in structure of knowledge based neural models discussed in section 5.

3 Neural Models for Loaded Microwave Cavities Based on MLP Network

Multilayer perceptron (MLP) neural network is high-parallel and high-adaptive feed-forward structure that is consisted of mutually connected neurons with nonlinear activation functions in hidden layers [4,6,7]. Researching of MLP application in microwave technique has showed that this network is able to approximate highly nonlinear functions with satisfactory accuracy and high level of generalization. Using this structure there is no need of knowledge for the explicit functional connection between the output and input parameters. Namely the beginning of the researching concerning the neural network application in microwave cavity's modeling was based on MLP neural model [8,9,10]

According to the equation (1), MLP network that models the cavity will give the resonant frequency f_r at the output, while at the input will have variable parameters of the dielectric slab: the filling factor t_h, and relative permittivity ε_r. The neural model is given by $\mathbf{y}=y(\mathbf{x},\mathbf{w})$. Where $\mathbf{w}$ is a connection weight matrix among neurons [4,6,7], $\mathbf{x}=[t_h, \varepsilon_r]^T$ is the input vector, and output vector is $\mathbf{y}=[f_r]$. The architecture of the corresponding MLP model is presented in Fig.3. The general symbol of this type of MLP neural model is MLPH-N_1-N_2-...-N_H , where H is the number of hidden layers, and N_i is the number of neurons in the i-th hidden layer. Activation functions of the hidden layers are sigmoid [6], while the output layer has linear activation function.

In [8] this MLP models the resonant frequency of the experimental cylindrical metallic cavity with circular cross-section with dimensions r=7 cm and h=14.24, for different TM/TE$_{mnp}$ modes in the range of input parameters: $0 \leq t/h \leq 0.1$ and $2 \leq \varepsilon_r \leq 82$. Training samples, testing samples, and referent curves during the simulations are generated using the transverse resonance method. Training samples are uniformly distributed in the input space. Neural network training is performed with different number of training samples in order to investigate its influence on the modeling accuracy. It is concluded that for satisfactory MLP model accuracy there is a need of more than 1000 uniformly distributed samples [8].

4. Method for MLP model efficiency improving

Because MLP is getting all the information from the training samples it is necessary a large number of input samples to be generated. This could be a difficult

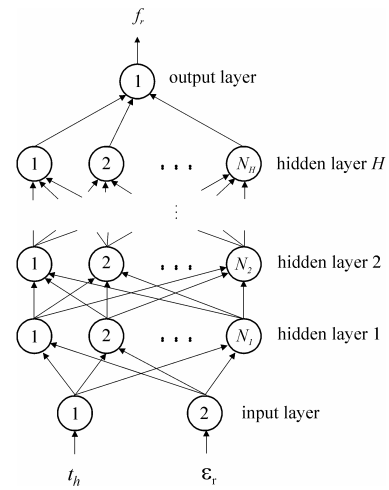

Fig. 3. MLP neural model for cylindrical metallic cavity with fixed dimensions and dielectric layer at the bottom

task because they should be generated using long time numerical method (transverse resonance method) or by complex experimental measurements. Also a large number of training samples might cause a long time necessary for MLP training. One way to avoid this problem is to decrease the number of training samples. This could be done by appropriate non-uniform distribution of the training samples on the input space, according to the behavior of the resonant frequencies in t_h-ε_r space [9]. This procedure is based on the fact that the dynamics of resonant frequency's changes is not same everywhere in the two-dimensional input space and it is determined by characteristic points defined in approximate model [5]. Where the dynamics is larger, the number of training samples is larger, and where it is smaller, a smaller number of training samples are picked. In this paper for training samples generating a new modification of the distribution presented in [9] is used. For given ε_r, the values which correspond to the characteristic points AA and RR of the modeled TM/TE$_{mnp}$ mode, to one intermediate added point between them, as well as to the boundary points (for t_h=0 and t_h=0.2) are used for input parameter t_h. Values of ε_r are generated in the following way

$$\varepsilon_{ri} = 1 + i^2, \; i = 1, 2, ..., 9 \tag{8}$$

5. Knowledge based neural models

Previously exposed technique for MLP efficiency improvement provides a satisfactory accuracy for network training while using a smaller number of training samples (in researching [9,10] the number of training samples was decreased from about 800-1200 to 300-500 samples for TM/TE$_{11p}$ $1 \leq p \leq 5$ modes). But in the basis there is still the black box-principle for modeling, where MLP is learning only from the input data. The existing empirical knowledge about the problem is indirectly presented to the MLP trough the training samples and not by its direct incorporation in the neural model. The main goal of performed investigation was steered to the realization of a model based on the knowledge. This model will be able to incorporate the knowledge about the cavity in itself in more direct way. In this way a higher accuracy might be achieved using less samples in training. For this model two approaches are used:

- Approach that uses hybrid neural-empirical model (HEN) [4,10]
- Approach that uses knowledge based neural network (KBNN) [7]

The first approach is a transition to a fully neural approach and uses the integration of the existing approximate model [5] as empirical knowledge holder and MLP network. The basic idea in this approach is that the empirical model with corresponding connection to the neural network provides higher generalization and extrapolation capabilities of the network [4]. This is achieved by presenting extra information about the problem at the input of the network. According to that, HEN model is developed for loaded microwave cavity whose architecture is presented in Fig.4. Approximate model determines the resonant frequency f_r^e in the following way: in the first step for given mode and given ε_r, according to section 2, determines AA and RR characteristic points; in the second step the resonant frequency between the characteristic points is approximated with spline function whose parameters are determined empirically according to [5]. The output from the approximate model f_f^e is brought to MLP as additional input. General symbol for this HEN model is HENH-N_I-…-N_l-…-N_H where H is the total number of hidden layers and N_i is the number of neurons in the i-th hidden layer. This model is applied for TM_{112} mode resonant frequency calculation for experimental cylindrical metallic cavity with circular cross-section (with dimensions r=7 cm and h=14.24) in the wider range of input parameters ($0 \le t/h \le 0.2$ i $2 \le \varepsilon_r \le 82$). A training set of 82 samples has been obtained by the non-uniform distribution (8). In order to obtain a model as good as possible, training of various HEN models is done, where $1 \le H \le 3$, and $1 \le N_l \le 30$, using the same training set. Levenberg Marquardt's training algorithm [4] with prescribed error value $E_c = 10^{-4}$ is chosen. For comparison the same training set was used for different MLPH-N_I-…-N_l-…-N_H models.

The both HEN and MLP model have been tested using testing data set of 60 uniformly distributed samples (not used in the training process). The testing results for eight HEN models and for eight MLP models with the lowest average test error (ATE) are shown in Table 1 and Table 2, respectively. It can be seen that the HEN models show significantly lower ATE as well as worst case error (WCE) compared to MLP model.

Two models are selected for the TM_{112} mode simulation: one from the HEN group (HEN4-12-11) and one from the MLP group (MLP2-12-12). A three-dimensional (3D) presentation of the resonant frequency dependence versus the cavity filling factor and relative permittivity obtained by these models is presented in Fig. 6.a (HEN4-12-11) and in Fig. 6.b (MLP2-12-11). The comparison of these 3D plots with the referent surface obtained using transverse resonance method (shown in Fig. 6.c), shows that the surface obtained by HEN model is more similar to the referent surface than the surface obtained by MLP model. Moreover the surface obtained by MLP model in some areas is so irregular and largely deviates from referent one. The reason for this is that the number of training samples is not sufficient for MLP training. Using 10000 points calculation for this surface, transverse resonance method needs time between 20 and 30 hours (Pentium III 450 MHz, 128 MB RAM), while the HEN model needs 2 to 3 minutes (MLP model needs only 5 seconds).

The second approach is using the whole advantages of the first one, while eliminating the connection to the empirical model as independent part that brings limitations in the definition range, simulation speed, and software implementation. The realization of specialized neural network architecture is the basic idea in this approach. This structure will be appropriate to the cavity's model in the sense that it incorporates the existing relations and functional dependences investigated with approximate model (mode tuning behavior discussed in section 2). According to this a new Knowledge Based Neural (KBN) model is developed and it is presented in Fig.5. It has the ability to incorporate the functional dependences of resonant and anti-resonant frequencies in the air and dielectric part of the cavity using the additional specialized knowledge based neurons (KN) [7]. This incorporation is done through the transfer functions of the knowledge neurons. The transfer function of the KN neurons is a general fitness form of relations (2) and (6), or general fitness form of relations (3) and (7). According that the first type of KN neurons has the output:

$$z_i^a = \sqrt{\frac{w_{1,i}^a}{(1-t_n)^2} + w_{2,i}^a} \qquad i = 1,\dots,Z^a \tag{9}$$

while the second type of neurons has the output:

$$z_j^d = \sqrt{\frac{w_{1,j}^d}{\varepsilon_r} + \frac{w_{2,j}^d}{\varepsilon_r \cdot (t_n)^2}} \qquad j = 1,\dots,Z^d \tag{10}$$

where Z^a and Z^d are number of KN neurons of first and second type respectively. Parameters $w_{1,i}{}^a$, $w_{2,i}{}^a$, $w_{1,j}{}^d$ and $w_{2,j}{}^d$ are internal weights of KB neuron. These weights are joined to the weight matrix **w** and they also change trough the iterative process of network training. Neurons in the hidden layers have sigmoid transfer function [6]. At the input they receive the outputs from all neurons in the previous hidden layer and the outputs from all KB neurons. The output layer has one linear neuron that corresponds to the output f_r. The general symbol of this neural model is KBNH-Z-N_l-…-N_l-…-N_H where H is the number of hidden layers, Z is the total number of knowledge based neurons, and N_i is the number of neurons in the i-th hidden layer. The training and testing was performed for various KBN models, (where $1 \le H \le 3$, $1 \le Z \le 6$ and $1 \le N_l \le 30$) under the same conditions as for HEN model. Only symmetrical structures were considered where the number of KN neurons of first type is equal to the number of the neurons of second type (Z^a=Z^d=Z/2). Testing results of KBN model are shown in Table 3. It can be seen that they have ATE and WCE close to HEN models.

3D presentation of the resonant frequency dependence versus the cavity filling factor and relative permittivity obtained by KBN3-4-16-16-16 model is presented in Fig. 6.d. This surface is notably closer to the referent one compared to that for MLP model. This is similarly as for HEN model. Also the elimination of empirical model as knowledge holder and using the fully neural architecture provides higher simulation speed: KBN model for previously mentioned Pentium configuration needs less than 7 seconds for surface generation, which is much faster than HEN model. Moreover, KBN model provides easier increase of the number of input parameters compared to HEN model. The inputs of HEN model are determined with empirical model and it is more convenient for general model building.

6. Conclusions

EM modeling of microwave cavities is facing with many limitations. Good alternative is to use neural network modeling. MLP model is easy to be developed, has high simulation speed, but it needs a large number of training samples. The number of training samples could be decreased by their non-uniform distribution in the input space, which will correspond to the existing knowledge about the resonant frequencies gained from the approximate model. Further decrease and modeling efficiency improvement could be achieved using direct incorporation of the existing knowledge with the HEN and KBN model. Both models give a high accuracy even when the number of training samples is too small for satisfactory training of MLP model (smaller than 100 samples for the given experimental cavity). KBN model is advantageous compared to HEN model because it eliminates the external empirical model. The empirical model causes lower speeds and brings

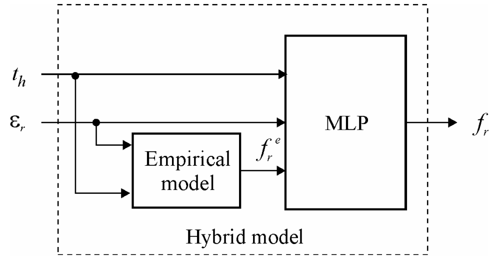

Figure 4. The architecture of HEN model

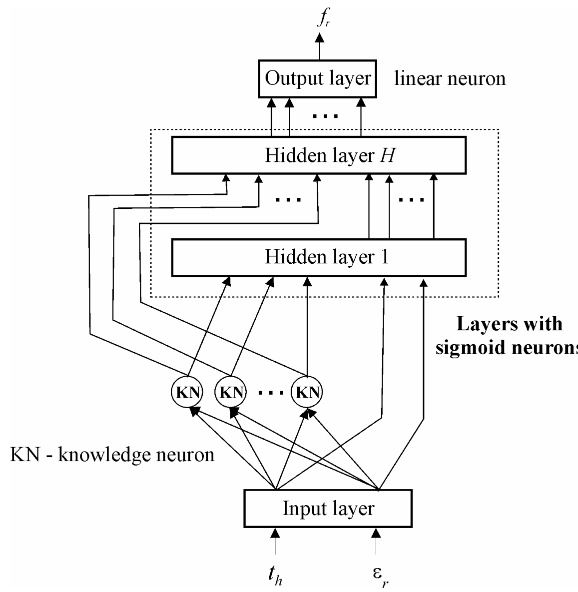

Figure 5. The architecture of KBN model

Table 1. The testing results for eight HEN models

HEN model	WCE [%]	AE [%]
HEN2-12-10	3.38	0.41
HEN2-14-10	4.05	0.49
HEN2-10-10	4.81	0.50
HEN2-12-9	4.91	0.43
HEN3-12-10-10	5.06	0.54
HEN2-11-10	5.48	0.60
HEN2-18-18	5.55	0.61
HEN2-15-9	7.27	0.69

Table 2. The testing results for eight MLP models

MLP model	WCE [%]	AE [%]
MLP2-12-12	8.96	1.09
MLP3-10-10-10	7.50	1.12
MLP2-8-8	9.74	1.11
MLP2-14-7	7.67	1.14
MLP2-11-7	10.28	1.22
MLP2-10-8	9.27	1.23
MLP2-12-8-4	11.03	1.28
MLP2-10-10	15.26	1.37

Table3. The testing results for eight KBN models

KBN model	WCE [%]	AE
KBN3-4-16-16-16	2.03	0.40
KBN2-2-16-16	2.31	0.37
KBN2-4-14-7	2.06	0.39
KBN2-2-12-12	2.43	0.49
KBN3-2-12-12-12	3.07	0.42
KBN2-4-16-16	3.27	0.51
KBN3-4-10-10-10	2.63	0.60
KBN3-2-10-10-10	3.50	0.55

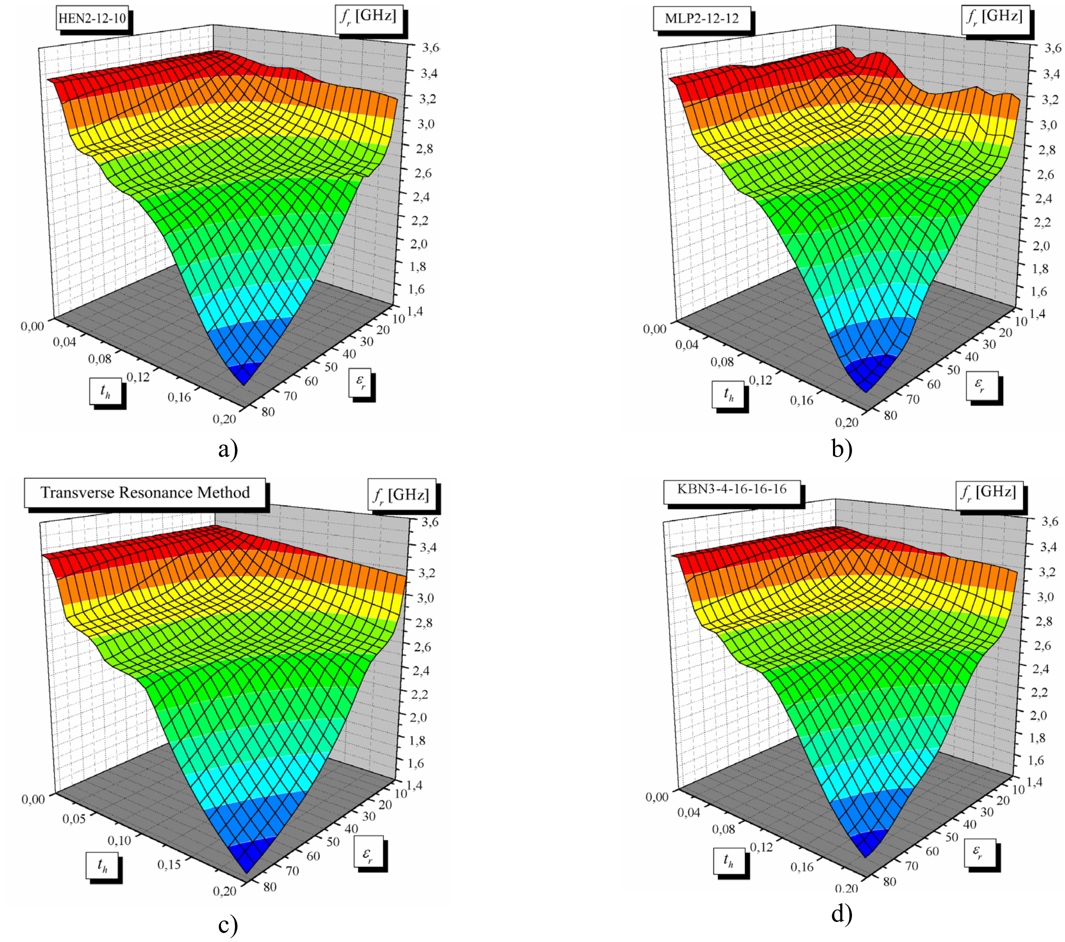

Fig. 6 Resonant frequency of TM_{112} mode vs. cavity filling factor and relative dielectric permittivity obtained by using (a) HEN, (b) MLP model, (c) TRM and (d) KBN model

different limitations, especially for increasing the number of input parameters.

References:

[1] Tse V. Chow Ting Chan, Howard C. Reader, *Understanding Microwave Heating Cavities*, Artech House, 2000.

[2] C. A. Balanis, *Advanced Engineering Electromagnetics*, John Wiley & Sons, Inc., New York, 1989.

[3] B. Milovanović, S. Ivković, D. Djordjević and N. Dončov, The Loading Effect Analysis of the Cylindrical Metallic Cavities with Various Cross-Sections, *Journal of Microwave Power and Electromagnetic Energy*, Vol. 33, No. 1, pp. 49-55, April 1998.

[4] Q. J. Zhang, K. C. Gupta, *Neural Networks for RF and Microwave Design*, Artech House, 2000.

[5] B. Milovanović, S. Ivković and A. Atanasković, Approximate Resonant Frequency Determination of the Cylindrical Metallic Cavities Loaded by Lossy Dielectric Slab, *Journal of Microwave Power and Electromagnetic Energy*, Vol. 33, No. 1, pp. 49-55, April 1999.

[6] S. Haykin, *Neural Networks,* New York, IEEE, 1994.

[7] Wang, F., and Q. J. Zhang, Knowledge Based Neural Networks for Microwave Design, *IEEE Trans. Microwave Theory and Techniques,* Vol. 45, 1997, pp. 2333-2343.

[8] B. Milovanović, Z. Stanković, S. Ivković, V. Stanković, Modelovanje opterećenih cilindričnih metalnih rezonatora pomoću neuronskih mreža, *ETRAN'99 Conference Proceedings (CD version)*, Yugoslavia 1999.

[9] B. Milovanović, Z. Stanković, S. Ivković*,* Modelling of the Cylindrical Metallic Cavity with Circular Cross-section using Neural Networks, *Proceedings of IEEE 10th Mediterranean Electrotechnical Conference - MELECON'2000,* Vol.II, Cyprus 2000, pp. 449-452.

[10] B. Milovanović, Z. Stanković, "Microwave Cylindrical Cavity Modeling Using Neural Model Incorporating Empirical Knowledge", *YUINFO 2002 Conference Proceedings (CD version)*, Kopaonik, Yugoslavia 2002.