NEURAL NETWORKS FOR FDTD-BACKED PERMITTIVITY RECONSTRUCTION

E. Eugene Eves[1], Ethan K. Murphy[2,3] and Vadim V. Yakovlev[3]

[1] The Ferrite Company, Inc., Hudson, NH 03051
[2] Department of Mathematics, Colorado State University, Fort Collins, CO 80521
[3] Department of Mathematical Sciences, Worcester Polytechnic Institute, Worcester, MA 01609

ABSTRACT. The paper outlines different versions of a novel method for determining the dielectric properties of arbitrarily shaped materials. Complex permittivity is found using a neural network procedure to control a 3D FDTD model computation of *S*-parameters and to process their measurements. Network architectures are based on multilayer perceptron and radial basis function nets. The method is cavity-independent and handles frequency-dependent media parameters. High accuracy of permittivity reconstruction is demonstrated by numerical and experimental testing.

INTRODUCTION

Recently, microwave power engineers have taken a particular interest in complex permittivity, $\varepsilon = \varepsilon' - i\varepsilon''$. While modern electromagnetic simulators allow the engineers to extensively characterize a constructed device prior to making a physical prototype, in order to perform a trustworthy simulation, it is necessary to have reliable knowledge of the dielectric properties of the materials being modeled.

Determination of dielectric constant ε' and loss factor ε'' of practical materials is a difficult problem. Perturbation and transmission/reflection techniques and other known methods may give satisfactory results under conditions which are either difficult to follow, or simply not acceptable; samples typically require the laborious preparation to comply with strict dimensional tolerance requirements.

With further progress in numerical methods, it has become feasible to develop techniques in which the more difficult tasks are assigned to a simulator while the experimental part is reduced to an elementary measurement. This approach has been taken in the methods using FEM [1-3] and FDTD [4] modeling of the entire experimental fixtures. To further explore this trend, the present paper outlines the principal aspects of a novel efficient technology for permittivity reconstruction.

In our previous work [5], we have proposed an approach involving an experimental setup (a closed cavity with an embedded measured sample), whose *S*-parameters are computed by the FDTD method and measured by a network analyzer. Dielectric constant and loss factor are determined in the course of processing of simulated and measured data by an optimization procedure based on artificial neural networks (ANN).

In the present contribution, we generalize the capabilities of this method by considering other network architectures built on the multilayer perceptron (MLP) and the radial basis function (RBF) ANN, checking different options in network training, and expanding the class of suitable materials to the ones with frequency-dependent media parameters. Since the underlying modeling technique easily handles arbitrary sample/fixture geometry and ANN technology is capable of generalizing the processed data and adjusting to the physical
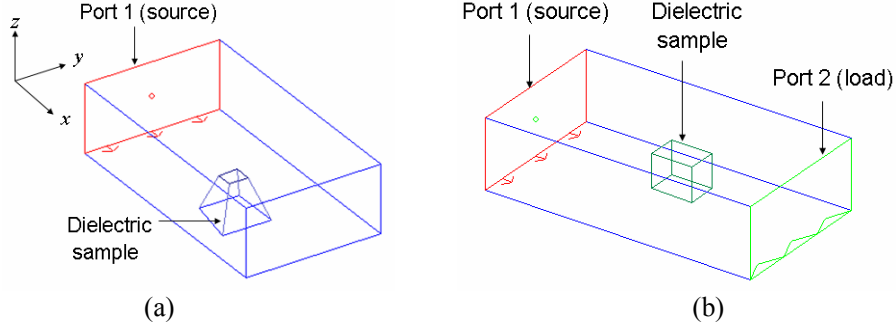
Fig. 1. One- (a) and two-port closed systems (b) with dielectric samples of arbitrary configuration.

characteristics of the cavity, our method is proposed as a flexible and efficient technique of permittivity reconstruction well suited to practical applications.

METHOD OF PERMITTIVITY RECONSTRUCTION
Network Architectures

We consider two basic one-hidden-layer architectures associated with two types of the experimental setups – a one-port structure intended for measurement of the reflection coefficient S_{11} and a two-port system which also handles the transmission coefficient S_{21}. The configurations of the closed systems considered in our analysis are shown in Fig. 1 while the corresponding networks are presented in Figs. 2 and 3.

For network training and testing, we use information generated in the modeling phase of the method powered by the 3D FDTD method. In the first approach, the network input receives the simulated values of $|S_{11}|$ at n points of the interval around the frequency of interest f_0 while the network output is associated with ε' and ε''. In the second architecture, the input layers get (and the output layers generate) either the simulated values of $\mathrm{Re}(S_{11})$, $\mathrm{Im}(S_{11})$, $\mathrm{Re}(S_{21})$, $\mathrm{Im}(S_{21})$, or the values of ε' and ε'' for which S-parameters are computed at the modeling stage. When the network is well trained, it is supplied with the measured values of S-parameters and determines ε' and ε'' of the sample in question.

One-Port Solution

In order to describe the computation of complex permittivity with the presented networks, we introduce the vectors $\bar{\mathbf{S}} = \begin{bmatrix} \bar{S}_1 & \ldots & \bar{S}_n \end{bmatrix}^T = \begin{bmatrix} |S_{11}|(f_1) & \ldots & |S_{11}|(f_n) \end{bmatrix}^T$ and $\boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 & \varepsilon_2 \end{bmatrix}^T = \begin{bmatrix} \varepsilon' & \varepsilon'' \end{bmatrix}^T$. Then the one-port networks generate the following output:

$$\varepsilon_l = \sum_{j=0}^{N} w_{lj}^3 \sigma \left[\sum_{i=0}^{n} w_{ji}^2 \bar{S}_i \right], \, l = 1, 2, \tag{1}$$

where $\sigma(\cdot) = \tanh(\cdot)$ is the activation function used for the hidden neurons, and $w_{pq}^{2|3}$ represents the network weights of the links between the qth neuron in the 1$^{\text{st}}$ or 2$^{\text{nd}}$ layer and the pth neuron in the 2$^{\text{nd}}$ or 3$^{\text{rd}}$ layer; the activation function for the output neurons is a linear function. In the $2 \times n$-input network, $N = N_A$ and $N = N_B$ for Net A and Net B respectively.

The training data are pairs of $(\bar{\mathbf{S}}_k, \mathrm{E}_k)$, $k = 1, \ldots, P$, where E_k is the desired outputs of the network with input $\bar{\mathbf{S}}_k$ (i.e., the values of dielectric constant and the loss factor for which $\bar{\mathbf{S}}_k$ have been simulated), and P is the number of training vectors. The aim is to adjust the vector of network weights $\mathbf{w}$ in order to reduce the errors defined as

$$e_{\varepsilon_l} = \frac{1}{2} \sum_{k=1}^{P} \left| \varepsilon_l(\bar{\mathbf{S}}_k, \mathbf{w}) - \mathrm{E}_k \right|^2 \tag{2}$$
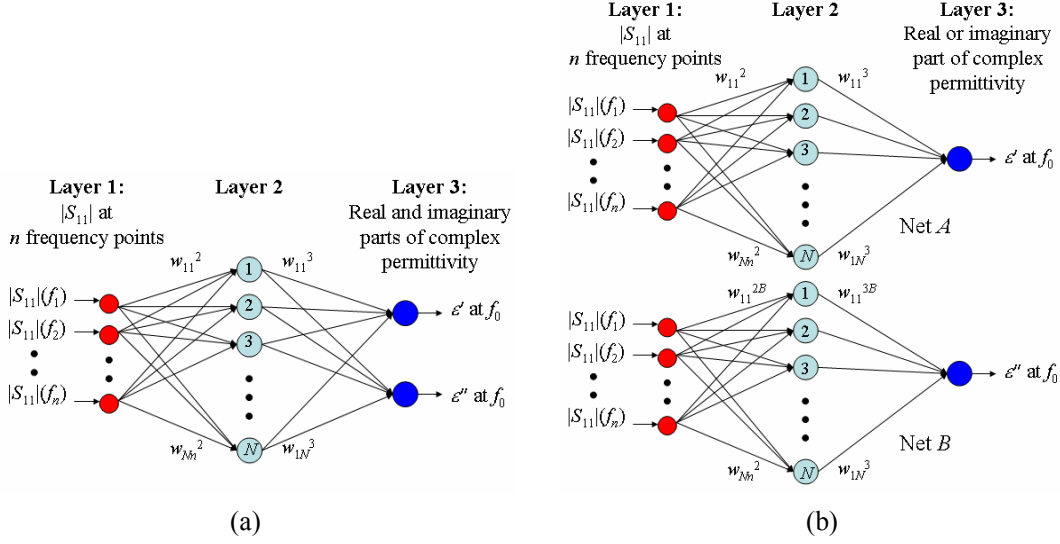
150

Fig. 2. MLP networks for the one-port system: *n*- (a) and 2×*n*-input architectures (b).
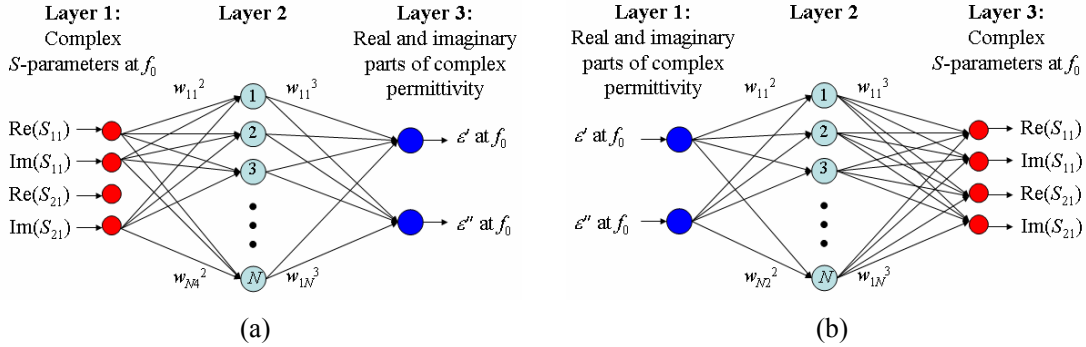

Fig. 3. MLP and RBF networks for the two-port structure: 4- (a) and 2-input architectures (b).

where $\varepsilon_l(\overline{\mathbf{S}}_k, \mathbf{w})$ is the ANN output for input $\overline{\mathbf{S}}_k$. The errors depend on the way the network is trained as well as on its configuration, i.e., on the number of hidden neurons. To minimize the errors (and improve the quality of learning), we determine this number by a standard trial-and-error process applied to the same training data set.

Two training algorithms, namely, backpropagation technique and the second-order gradient-based technique are implemented with the use of the gradient method (iterations from 1 to 200) and the Levenberg-Marquardt method (iterations beyond 200) respectively.

Since the one-port approach deals with the at frequencies different from f_0, we have a fundamental restriction on the accuracy of this version of the method applied to the materials with frequency-dependent media parameters. FDTD computation of a frequency response is performed for ε' and ε'' at f_0, and measurement of the reflection coefficient is conducted everywhere in (f_1, f_n); hence the measured values may correspond only at f_0. This provides motivation for considering alternative network architectures processing the related information only at f_0, and dealing with more parameters representing the system behavior, i.e., with the complex reflection coefficient (S_{11}) and transmission coefficient (S_{21}).

Two-Port Solution

With introduction of the vector $\overline{\overline{\mathbf{S}}} = \left[\overline{\overline{\mathbf{S}}}_1 \ ... \ \overline{\overline{\mathbf{S}}}_4 \right]^T = \left[\mathrm{Re}(S_{11}) \ ... \ \mathrm{Im}(S_{21}) \right]^T$, the output of the two-port MLP and RBF networks is represented by the formulas

$$\varepsilon_l = \sum_{j=0}^{N} w_{lj}^3 \sigma\left[\sum_{i=0}^{4} w_{ji}^2 \overline{\overline{S}}_i\right], \; l = 1, 2, \text{ and} \tag{3}$$

$$\overline{\overline{S}}_l = \sum_{j=0}^{N} w_{lj}^3 \sigma\left[\sum_{i=0}^{2} w_{ji}^2 \varepsilon_i\right], \; l = 1, \ldots, 4, \tag{4}$$

associated with the 4- and 2-input networks; the hidden neuron activation functions are the hyperbolic tangent and Gaussian function $\sigma(\gamma) = e^{-\gamma^2}$ for MLP and RBF ANN respectively. A linear activation function is used for the output layer in the networks of both types.

The training data for the 4-input MLP and RBF architectures are pairs of $(\overline{\overline{S}}_k, E_k)$, and the training error is defined as

$$e_{\varepsilon_l} = \frac{1}{2}\sum_{k=1}^{P}\left|\varepsilon_l(\overline{\overline{S}}_k, \mathbf{w}) - E_k\right|^2 \tag{5}$$

where $\varepsilon_l(\overline{\overline{S}}_k, \mathbf{w})$ is the ANN output for input $\overline{\overline{S}}_k$. In the 2-input networks, the training data are pairs of $(\boldsymbol{\varepsilon}_k, \Sigma_k)$, where Σ_k is the desired outputs of the network for inputs $\boldsymbol{\varepsilon}_k$ (i.e., the values of S-parameters simulated for given $\boldsymbol{\varepsilon}_k$). Computation of error in this case is preceded by minimization of the function

$$G_k = \left|\overline{\overline{S}}_l(\boldsymbol{\varepsilon}_k, \mathbf{w}) - \Sigma_k\right|^2, \; k = 1, \ldots, P, \; l = 1, \ldots, 4, \tag{6}$$

where $\overline{\overline{S}}_l(\boldsymbol{\varepsilon}_k, \mathbf{w})$ is the ANN output for input $\boldsymbol{\varepsilon}_k$. The solution to this minimization problem is a set of approximated complex permittivity values. Therefore, the network error is determined from:

$$e_{\overline{\overline{S}}_l} = \frac{1}{2}\sum_{k=1}^{P}[\text{Min}(G_k) - E_k]^2 \tag{7}$$

For the training, the backpropagation technique and the second-order gradient-based technique are used in the two-port networks just as in the one-port ones.

NUMERICAL TESTING
One-Port Structure
All the above ANN algorithms have been implemented in a MATLAB 6 environment. For modeling, we use the full-wave 3D conformal FDTD simulator *QuickWave-3D* (*QW-3D*) [6]. Data required for network training is collected by a special procedure that repeatedly runs *QW-3D* to compute S-parameters for various values of ε' and ε'' of the sample.

The one-port scheme has been tested numerically for a section of 72×34 mm waveguide with a rectangular ($20 \times 20 \times 30$ mm) dielectric block in the corner near the shorting wall. The FDTD model representing this scenario was built with a non-uniform mesh with 7.5 and 3 mm cubic cells in air and in a dielectric sample respectively (8,463 cells total).

The networks were trained using vectors of $|S_{11}|$ frequency responses with $n = 3$, $f_1 = 2.4$ GHz, $f_2 = f_0 = 2.45$ GHz, $f_3 = 2.5$ GHz, and with 27 values of complex permittivity from the intervals $5 \le \varepsilon' \le 9$ and $0.2 \le \varepsilon'' \le 1.0$. The graphs in Fig. 4 show the typical sum-squared error produced by the n- and $2 \times n$-input networks for different number of neurons in the 2nd layer. It is seen that for more than 10 hidden neurons, the networks are characterized by errors not larger than 10^{-5}.

Table 1. Dielectric samples used in numerical testing of the networks for the two-port scheme

Sample	x-, y-, z-dimensions (mm)	position	distance (mm) from…
A	$50 \times 50 \times 20$	A	…the 2nd port: 120, …central line: 0
B	$42 \times 30 \times 50$	B	…the 2nd port: 120, …central line: 30
C	$20 \times 25 \times 62$	C	…the 2nd port: 120, …central line: 60
D	$20 \times 25 \times 20$	D	…the 2nd port: 150, …central line: 30

When tested with the training sets of 51 vectors, the networks demonstrated sufficiently accurate permittivity reconstruction. The desired and actual responses from the $2 \times n$-input MLP are presented in Fig. 5: mean square error (MSE) is of order 10^{-3}.

Two-Port Structure

The two-port scheme dealing with the S-parameters at f_0 has been numerically tested with vectors of Re(S_{11}), Im(S_{11}), Re(S_{21}), and Im(S_{21}) at $f_0 = 915$ MHz for the 497 mm section of a 248×124 mm waveguide containing a rectangular dielectric sample (Table 1).

We built the training sets for the values of relative complex permittivity in the ranges $54 \le \varepsilon' \le 74$ and $6 \le \varepsilon'' \le 30$. The 4- and 2-input MLP and the 4-input RBF ANNs were trained with the sets obtained for 48 equally spaced points in the complex (ε', ε'')-plane and additional points on the border (68 samples total). For the 2-input RBF, where the number of vectors in the training set is equal to the number of hidden neurons, the decision as to how many vectors (i.e., points from the (ε', ε'')-plane) in the database to use was made dynamically. The network was given a small database and the error was computed. The three test points with the greatest error were chosen, and for each point an average was taken between the supposed and the ANN-generated values. This average was then taken for the computation of the next sample for the database. For example, for Sample B in Position B, the optimal number of training vectors (and hidden neurons) turned out to be 57 (Fig. 6). In the 4- and 2-input MLP, N was taken 13 and 14 respectively.

Although all MLP/RBF 4-/2-input networks have demonstrated good performance, some of them were found to be more accurate. In Fig. 7, the desired and actual responses are shown for the 2-input RBF network with a corresponding mean square error 0.013 while for the 4- and 2-input MLP ANNs, MSE's are 0.029 and 0.073 respectively.

Training sets for the ranges of $36 \le \varepsilon' \le 56$ and $4 \le \varepsilon'' \le 26$ have also been created. The MLP and RBF networks were trained as described above. The 2-input nets have again shown somewhat lower errors. In Fig. 8, typical examples of the desired and actual responses from the MLP networks are presented: in both cases the MSE values are of order 10^{-3}.

The detailed error analysis has been carried out to evaluate the accuracy of the two-port systems with a ± 2 mm divergence in the sample's geometry in each dimension. Numerical experimentation has been performed for $\varepsilon = 57 - i8$ (apple, 88 % moisture contents), $\varepsilon = 68 - i14$ (cantaloupe, 92 %), $\varepsilon = 62 - i22$ (potato, 79 %), and $\varepsilon = 55 - i16$ (sweet potato, 80 %) [7]. A typical example of this computation is shown in Fig. 9.

Generalizing the results of the analysis conducted for these materials as Samples A to D at Positions A to D, we conclude that the 2-input networks can give an error in ε'' less than 5 % if the sample's geometrical deviation in the longitudinal and transverse directions does not exceed 0.5-1.0 mm. A 10 % error results from a 1.2-1.5 mm deviation. For ε', the error is less than 5 % when the deviation is less than 1.2-1.5 mm and always less than 10 % in the considered 2 mm deviation. At the same time, a notable variation of accuracy is observed when the sample's height changes – even if variation in the vertical dimension is very small. So, for accuracy, this experimental setup should be constructed to minimize accidental deviations of the sample size in the z-direction.

EXPERIMENTAL TESTING

To show the method in full operation, we have designed the experimental setup implementing the concept of the one-port solution and thus measuring reflections from a
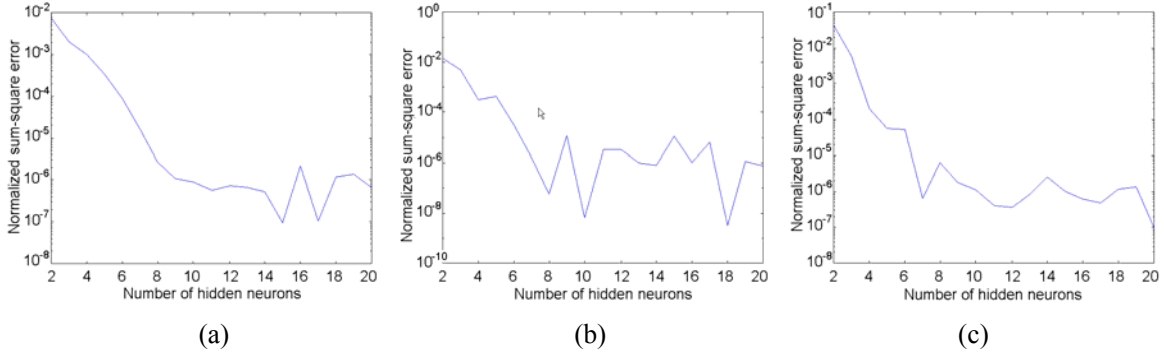
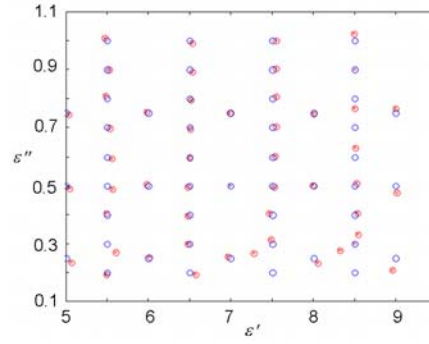Fig. 4. Training and testing error of *n*-input MLP (a) and Net *A* (b) and Net *B* (c) of 2×*n*-input MLP.

Fig. 5. Complex permittivity reconstructed with the 2×*n*-input MLP with $N_A = N_B = 10$: circles and crossed circles mark the test data and the actual responses respectively.

cavity with a dielectric sample (Fig. 10). Using a rectangular (70 × 70 × 50 mm) Teflon block with a cylindrical cutout (radius 25 mm, height 40 mm) suitable for holding liquids, we have determined complex permittivity of tap and saline water. The container filled with water was placed on the center line of the waveguide section at 40 mm from the waveguide's shorting wall in the opposite end with respect to the coaxial-waveguide transition.

We used a *QW-3D* model consisting of 71,442 cells with a non-uniform mesh (cell sizes in air, Teflon and water are 15, 5, and 2 mm respectively) for the entire cavity and dielectric inclusions. The permittivity of Teflon was taken as 2.06 – *i*0. The database of the training and testing sets was created with $n = 3, f_1 = 0.91$ GHz, $f_2 = f_0 = 0.915$ GHz, and $f_3 = 0.92$ GHz for $60 \leq \varepsilon' \leq 90$ and $1 \leq \varepsilon'' \leq 20$ and included 108 and 224 vectors respectively. For the 2×*n*-input network, the optimal structure was found as having $N_A = 15$ and $N_B = 19$. The normalized sum of squared differences between the desired and actual network responses at the training stage was less than 10^{-4} for both Net *A* and Net *B*.

The values of $|S_{11}|$ measured at $f_1 = 0.91$ GHz, $f_2 = 0.915$ GHz, and $f_3 = 0.92$ GHz for the Teflon container filled with water were given to the trained network, and it generated water's dielectric constant and the loss factor. For the sample of known temperature and salinity, ε' and ε'' have been also determined from the model whose average error is 0.3% for ε' and 1.8% for ε'' [8]. As one can see from Table 2, the results are in very close agreement. This confirms the capability of the proposed ANN-based method for accurate reconstructing of complex permittivity of materials.

CONCLUSION

Our novel technology of permittivity reconstruction which employs FDTD modeling, an ANN-based optimization technique, and elementary measurement of *S*-parameters places minimal physical requirements on fixture and sample geometry and is sufficiently accurate
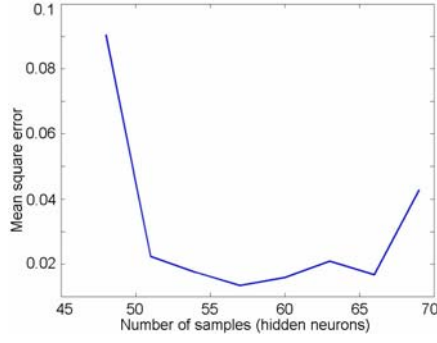
Fig. 6. MSE of the 2-input RBF with the number of training samples from 48 to 69 with step 3.
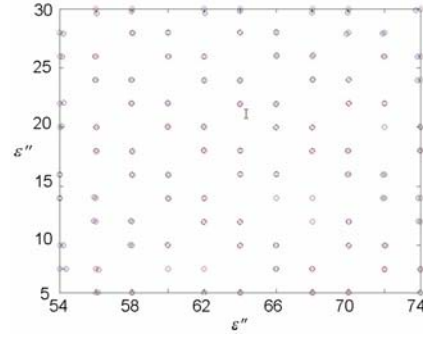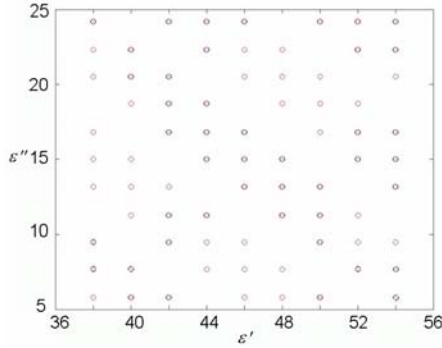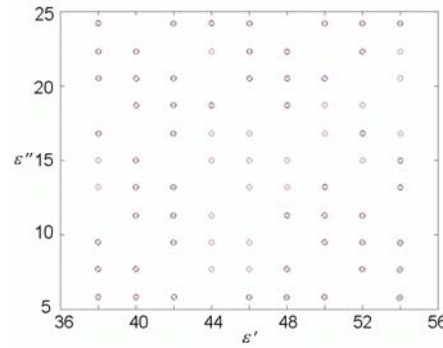
Fig. 7. Complex permittivity of Sample B in Position B: reconstructed with the 2-input RBF with $N = 57$; the test data and the actual responses.
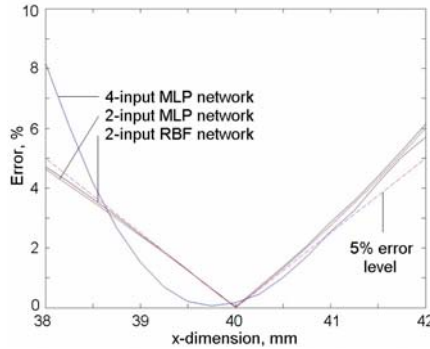

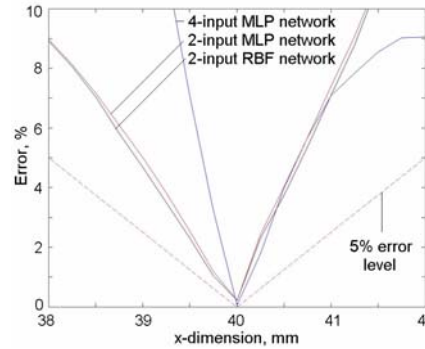(a)   (b)

Fig. 8. Complex permittivity of Sample C in Position B: reconstructed by the 4-input MLP with $N = 13$ (a) and the 2-input MLP with $N = 14$; the test data and the actual responses.


(a)   (b)

Fig. 9. Percent error in getting right ε' (a) and ε'' (b) as a function of deviation of training data for the sample dimension in the x-direction: potato as Sample B in Position B.

for practical use. Further developments of the method may include its adjustment to non-homogeneous dielectrics and a refinement to allow sample preparation to less strict dimensional tolerances.

The practical advantages of the method are obvious. It does not depend on the associated closed system and thus can be used with any available cavity and any suitable FDTD
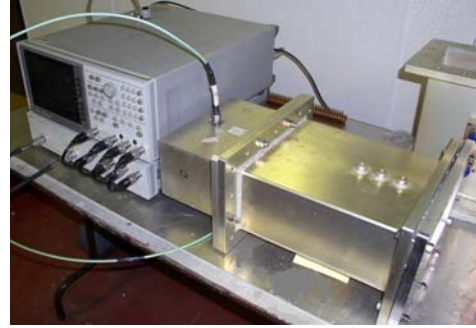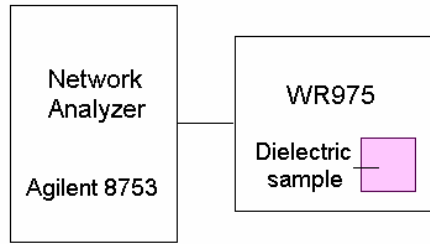
(a)　　　　　　　　　　　　　　　(b)

Fig. 10. Diagram (a) and photo (b) of the experimental setup for the one-port solution.

Table 2. Complex permittivity of fresh water with salinity 0.033 ‰ at temperature 18.6°C determined by the one-port method and the $2 \times n$-input MLP ANN

ε'			ε''		
Method	Model [8]	Accuracy, %	Method	Model [8]	Accuracy, %
80.6	80.5	0.12	4.25	4.30	1.2

simulator, not necessarily *QW-3D*. While a relatively large computational effort may be required for creation of a database for network training and testing, the process of training requires nearly negligible time. Whenever we work at a fixed frequency with materials that can take some pre-defined form, the database is created only once. One can do that prior to actual experimental testing, and each new material can be processed thereafter practically in real time – provided that ε' and ε'' of this material are within the ranges specified in the database and that the computer model is based upon the measured experimental fixture.

REFERENCES

[1] Deshpande, M.D., Reddy, C.J. (1995). Application of FEM to estimate complex permittivity of dielectric material at microwave frequency using waveguide measurements, *NASA Contractor Report CR-198203*, 23 p.
[2] Coccioli, R., Pelosi, G., Selleri, S. (1999). Characterization of dielectric materials with the finite-element method. *IEEE Trans. Microwave Theory Tech*., MTT-47, 1106-1112.
[3] Thakur, K.P., Holmes, W.S. (2001). An inverse technique to evaluate permittivity of material in a cavity. *IEEE Trans. Microwave Theory Tech*., MTT-49, 1129-1132.
[4] Wäppling-Raaholt, B., Risman, P.O. (2003). Permittivity determination of inhomogeneous foods by measurement and automated retro-modeling with a degenerate mode cavity. *Proc. 9th Conf. Microwave and HF Heating, Loughborough, U.K., Sept. 2003*, 181-184.
[5] Eves, E.E., Kopyt, P., Yakovlev, V.V. (2004). Determination of complex permittivity with neural networks and FDTD modeling, *Microwave Opt. Tech. Letters*, 40, 183-188.
[6] *QuickWave-3D*™, QWED, ul. Zwyciezcow 34/2, 03-938 Warsaw, Poland, http://www.qwed.com.pl/.
[7] Nelson, S., Datta, A.K. (2001). Dielectric properties of food materials, In: Datta, A.K., Anantheswaran, R.C., Eds., *Handbook of Microwave Technology for Food Applications*, Marcel Dekker, Inc., 69-114.
[8] Eves, E.E., Yakovlev, V.V. (2002). Analysis of operational regimes of a high power water load. *J. Microwave Power & Electromagnet. Energy*, 38, 127-144.