Progress In Electromagnetics Research, PIER 92, 65–77, 2009

KNOWLEDGE-BASED SUPPORT VECTOR SYNTHESIS OF THE MICROSTRIP LINES

N. T. Tokan and F. Güneş

Department of Electronics and Communication Engineering
Faculty of Electrics and Electronics
Yıldız Technical University
Yıldız, Istanbul 34349, Turkey

Abstract—In this paper, we proposed an efficient knowledge-based Support Vector Regression Machine (SVRM) method and applied it to the synthesis of the transmission lines for the microwave integrated circuits, with the highest possible accuracy using the fewest accurate data. The technique has integrated advanced concepts of SVM and knowledge-based modeling into a powerful and systematic framework. Thus, synthesis model as fast as the coarse models and at the same time as accurate as the fine models is obtained for the RF/Microwave planar transmission lines. The proposed knowledge-based support vector method is demonstrated by a typical worked example of microstrip line. Success of the method and performance of the resulted synthesis model is presented and compared with ANN results.

1. INTRODUCTION

Nowadays, two typical nonlinear learning machines are widely employed as the fast and flexible machines in the generalization of the highly nonlinear input-output discrete mapping relations in the microwave modeling: Artificial Neural-Network (ANN) and Support Vector Machine (SVM). A detailed literature for the utilization of the ANN techniques in the CAD of a variety of microwave components and circuits can be found in [1].

On the other hand within the last decade, Vapnik's SVM theory [2] has been successfully applied in a wide range of classification and regression problems, resulting in its improved generalization performance over other classical optimization techniques. This is

Corresponding author: N. T. Tokan (nturker@yildiz.edu.tr).

mainly because, firstly, SVM solves a convex constrained quadratic optimization problem, whose error surface is free of local minima and has a unique global optimum; secondly, SVM approach is based on structural risk minimization (SRM) principle instead of empirical risk minimization (ERM) which is used in ANN approach. SRM principle implements well trade-off between the model's complexity and its generalization ability [3]. Furthermore, SVM is based on small sample statistical learning theory, whose optimum solution is based on limited samples instead of infinite sample that ensures enormous computational advantages. Typical applications of the SVRM to the microwave modeling can be found in [4–6].

Both of these nonlinear learning machines are once trained, they are capable of responding almost instantly to any input variable set, thus they are as fast as the approximate (coarse) models and can be as accurate as the detailed electromagnetic (fine) models [6–9]. Here, the key problem is to train these machines with the accurate data which may be measured or simulated data. In modeling using these nonlinear learning machines, accurate training data generation is the major constituent of the total model development time as it needs both CPU and human time. Recently, there is a new trend in the Electromagnetic (EM)-ANN area for searching the techniques using reduced number of accurate training data, thus resulting in lessened CPU and human time together with faster model development. The pioneering techniques for reducing the need for accurate traning data can be given as follows: Neural networks with knowledge such as the knowledge-based neural networks (KBNN) [7], difference method (DM) [8], prior-knowledge input (PKI) network [9] and space mapped neural networks (SMNN) [10–13]. Furhermore an efficient knowledge-based automatic model generation (KAMG) technique is proposed [14] combining automatic model generation, knowledge neural networks, and space mapping, where the two data generators —coarse and fine generators— are simultaneously employed for the first time.

In this work, for the first time, we proposed a knowledge-based SVM method and applied it to synthesis of transmission lines. Here, our motivation is to obtain a synthesis model as fast as the coarse models and at the same time as accurate as the fine models. Thus, we choose the SVM method to utilize its improved features, especially its working principle based on the small sample statistical learning theory, in lessening the need for the accurate training and validation data together with the human time. We have demonstrated our knowledge-based SVM method by a typical worked example, depending on the analysis process. In the worked example, the coarse and fine SVM synthesis models are generated using the coarse and fine generators, respectively. Since coarse data generators are approximate and fast (e.g., two and one-half dimensional EM simulators), the coarse SVM model is generated by the extensive use of the coarse data, but much less fine data are used to match accurately this coarse SVM model to the fine generator. Finally, SVR is applied to the resulted fine synthesis data again for the accurate synthesis formulation. Moreover, performance of the SVM is compared with that of the ANN.

The paper is organized as follows: The generation of the support vector expansion of given data takes place in the next section. Third section gives general definitions about the knowledge-based support vector synthesis of RF/Microwave transmission lines. The fourth section is devoted to the worked example. The conclusions finally end the paper.

2. SUPPORT VECTOR REGRESSION MACHINES

Given a training dataset $(\mathbf{x}_i, y_i)$, $i = 1, 2, \ldots, \ell$ where $\mathbf{x}_i \in R^n$, $y_i \in R$ and ℓ is the size of training data, SVR tries to find the mapping function $f(\mathbf{x})$ between the input variable vector $\mathbf{x}$ and the desired output variable y. In formula this read as:

$$f(\mathbf{x}) = \mathbf{w}^T \cdot \boldsymbol{\phi}(\mathbf{x}) + b \tag{1a}$$

where $\mathbf{x}^T = (x_1, \ldots, x_n) \Rightarrow \boldsymbol{\phi}^T(\mathbf{x}) = (\phi_1(x), \ldots, \phi_N(x))$ and $b \in R$.

This step is equivalent to mapping the input space x into a new space, $F = \{\boldsymbol{\phi}(\mathbf{x})/\mathbf{x} \in X^n\}$. Thus, $f(\mathbf{x})$ is a nonlinear function in the x-input space and linear function in the F-feature space. Thus, we will build a nonlinear machine in two steps: First, a fixed nonlinear mapping vector $\boldsymbol{\varphi}(\mathbf{x})$ transforms the data into a feature space F, and then the linear machine built in this feature space is used to perform regression on the data. In this manner, we will refer to the quantities $\mathbf{w}$ and b in (1a) as the weight vector and bias where $\mathbf{w}$ is:

$$\mathbf{w}^T = (w_1, w_2, \ldots, w_N) \tag{1b}$$

Traditional regression method finds the regression function $f(\mathbf{x})$ by determination of $\mathbf{w}$ and b using the rule of empirical risk minimization principle:

$$R_{emp}[f] = \frac{1}{\ell}\sum_{i=1}^{\ell} L(\mathbf{x}_i, y_i, f) \tag{2}$$

where $L(\mathbf{x}_i, y_i, f)$ represents an error (loss) function. One of the familiar loss functions is ε-insensitive loss function developed by

Vapnik [4]:

$$L(f(\mathbf{x}_i) - y_i) = |y - f(\mathbf{x})|_\varepsilon = \max|0, |y - f(\mathbf{x})| - \varepsilon| \tag{3}$$

However, the actual risk minimization cannot be realized only with the empirical risk minimization [2]. A typical example is the over-fitting of ANN. By the SRM principle employed by the SVR, the generalization accuracy is optimized over the empirical error and the flatness of the regression function which is guaranteed on a small $\mathbf{w}$:

$$R_{SRM}(f, \mathbf{w}) = \frac{1}{2}\|\mathbf{w}\|^2 + CR_{emp}[f] \tag{4a}$$

where $\frac{1}{2}\|\mathbf{w}\|^2$ is the term characterizing the modeling complexity and C is a regularization parameter which determines the trade off between model complexity and empirical loss function. Substituting (2), (3) into (4a) and introducing the slack variables ξ_i, ξ_i^*, Eq. (4a) is transformed into the following soft margin primal optimization problem:

Minimize:

$$\frac{1}{2}\|w\|^2 + C\sum_{i=1}^{\ell}(\xi_i + \xi_i^*) \tag{4b}$$

subject to:

$$(< w \cdot \phi(x_i) > +b) - y_i \le \varepsilon + \xi_i \tag{5a}$$
$$y_i - (< w \cdot \phi(x_i) > +b) \le \varepsilon + \xi_i^* \tag{5b}$$
$$\xi_i \ge 0,\ \xi_i^* \ge 0,\ \varepsilon \ge 0,\ i = 1, 2, \ldots, \ell \tag{5c}$$

Combining the objective function given in (4b) with the constraints in (5), we have the corresponding Lagrangian function. By applying saddle point conditions with respect to the primal variables w_i, b, ξ_i, ξ_i^* leads to the optimal regressor (i.e., optimal set of the weighting values), given by

$$\mathbf{w} = \sum_{i=1}^{\ell}(\alpha_i^* - \alpha_i)\boldsymbol{\phi}(\mathbf{x}_i) \tag{6}$$

where α_i, α_i^* are positive Lagrangian multipliers obtained by maximization of the following dual space objective function [2]:

$$\begin{aligned} W(\alpha, \alpha^*) \ = \ & -\varepsilon\sum_{i=1}^{\ell}(\alpha_i + \alpha_i^*) + \sum_{i=1}^{\ell} y_i(\alpha_i - \alpha_i^*) \\ & -\frac{1}{2}\sum_{i,j=1}^{\ell}(\alpha_i^* - \alpha_i)(\alpha_j - \alpha_j^*) < \boldsymbol{\phi}(\mathbf{x}_i)\boldsymbol{\phi}(\mathbf{x}_j) > \end{aligned} \tag{7}$$

subject to:

$$0 \leq \alpha_i \leq C,\ 0 \leq \alpha_i^* \leq C,\ i = 1, \ldots, n,\ \sum_{i=1}^{\ell} \alpha_i = \sum_{i=1}^{\ell} \alpha_i^* \tag{8}$$

The corresponding Karush-Kuhn-Tucker complementary conditions are:

$$\alpha_i(< \mathbf{w}\boldsymbol{\phi}(\mathbf{x}_i) > +b - y_i - \varepsilon - \xi_i) = 0 \tag{9a}$$
$$\alpha_i^*(y_i - < \mathbf{w}\boldsymbol{\phi}(\mathbf{x}_i) > +b - \varepsilon - \xi_i^*) = 0 \tag{9b}$$
$$\xi_i \cdot \xi_i^* = 0 \tag{9c}$$
$$\alpha_i \cdot \alpha_i^* = 0 \tag{9d}$$
$$(\alpha_i - C) \cdot \xi_i = 0 \tag{9e}$$
$$(\alpha_i^* - C) \cdot \xi_i^* = 0,\ i = 1, 2, \ldots, \ell \tag{9f}$$

From (9a) and (9b) of the Karush-Kuhn-Tucker conditions, it follows that for only the samples satisfying $|f(x_i) - y_i| \geq \varepsilon$, the Lagrangian multipliers may be nonzero, and for the samples of $|f(x_i) - y_i| < \varepsilon$, the Lagrangian multipliers α_i, α_i^* vanish. Since the products of α_i with α_i^* and ξ_i with ξ_i^* according to (9c) and (9d) are zero, at least one term in the couples of (α_i, α_i^*); (ξ_i, ξ_i^*) is zero. Therefore, we have a sparse expansion of $\mathbf{w}$ in terms of input variable vector $\mathbf{x}$; thus we do not need all data to describe $\mathbf{w}$. The samples $(\mathbf{x}_i, y_i)$ that come with nonvanishing coefficients are called Support Vectors (SVs). The idea of representing the solution by means of a small subset of training points has also enormous computational advantages. This reduced number of nonzero parameters together with the guaranteed global minimum gains superiority to SVM over the alternative methods. A detailed mathematical background together with the literature can be found in [2].

Thus, substituting the calculated nonzero Lagrangian multiplier (α_is) into (6) and then into (1a), the mapping function $f(\mathbf{x})$ between the input variable space and the desired output variable can be expressed in terms of the SVs as follows:

$$f(\mathbf{x}) = \sum_{i=1}^{ns} \alpha_i y_i < \boldsymbol{\phi}(x_i)\boldsymbol{\phi}(x) > +b \tag{10a}$$

where b can be found making use of the primal constraints in (9a), (9b) and ns is the number of SVs. The inner product $< \boldsymbol{\phi}(x_i)\boldsymbol{\phi}(x) >$ in the feature space is called kernel function K, which can be given for all x, z, X as:

$$K(x, z) = \langle \phi(x)\phi(z) \rangle \tag{10b}$$

where ϕ is a mapping from X to an inner product feature space F. Substituting (10b) into (10a), we have the SV expansion of $f(\mathbf{x})$ in terms of the kernels as:

$$f(\mathbf{x}) = \sum_{i=1}^{ns} \alpha_i y_i K(\mathbf{x}_i, \mathbf{x}) + b \tag{10c}$$

In the next section, the continuous domain SV expansion of the microstrip transmission line geometry will be given in terms of the required characteristic impedance $\mathbf{Z}_0$ and dielectric material ε_r provided as accurate as that of their fine models.

3. A KNOWLEDGE-BASED SUPPORT VECTOR SYNTHESIS OF RF/MICROWAVE TRANSMISSION LINES

The proposed knowledge-based SVM synthesis method is given as block diagram in Fig. 1, which is demonstrated by the typical example of commonly used MIC line. In the worked example, SV expansion for the synthesis of the microstrip line will be obtained employing its coarse model. Then this coarse SVM model will be matched to its fine model using an EM simulator.

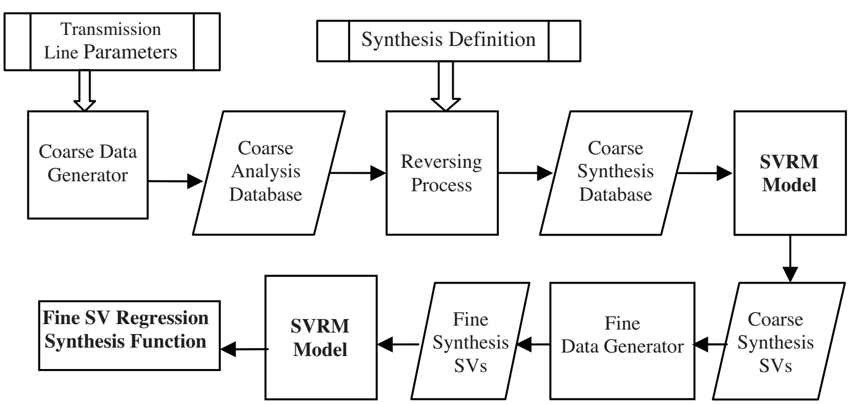


Figure 1. Knowledge-based SVRM approach applied to the synthesis of transmission lines.

It should also be noted that SVM performs a selection onto the training data to obtain much less data (SVs) to represent the whole data in the regression. These SVs can also be used in training another type of learning machine to obtain an efficient regression on the whole

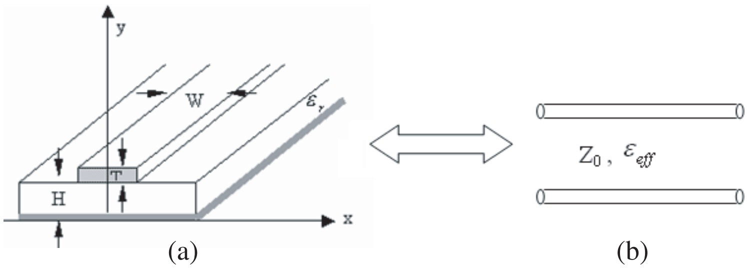


Figure 2. (a) Microstrip line (**M**); (b) its transmission line equivalence (**T**).

data. Similarly, in this work only SVs are taken into account to match the coarse SVRM model to the fine generator; thus a large amount of reduction in data and computational time is obtained.

4. WORKED EXAMPLE

4.1. Definition of the Problem

A microstrip transmission line given in Fig. 2(a) can be characterized in **M**-space defined by $\mathbf{M} = \{\mathbf{G}, \boldsymbol{\varepsilon}\} \in R^5$ where $\mathbf{G}$ and $\boldsymbol{\varepsilon}$ are the geometry and permeability vectors, respectively defined as:

$$\mathbf{G}^t = [W, H, T], \quad \boldsymbol{\varepsilon}^t = [\varepsilon_r, \varepsilon_y] \tag{11}$$

which are the width of the strip, height of the dielectric substrate, and strip thickness respectively, and ε_r, ε_y are the dielectric constants. Similarly, $\mathbf{Z}_0$ and $\varepsilon_{\mathit{eff}}$ are the characteristic impedance and effective dielectric parameters of the equivalent transmission line (Fig. 2(b)) and can be represented in a **T**-space:

$$\mathbf{T} = \{\mathbf{Z}_0, \varepsilon_{\mathit{eff}}\} \in R^2 \tag{12}$$

So analysis of the microstrip lines can be achieved by mapping A:

$$A : \mathbf{M}(R^5) \to \mathbf{T}(R^2) \tag{13}$$

Similarly, synthesis of the microstrip line can be defined by mapping S:

$$S : \mathbf{T}(R^2) \to \mathbf{M}(R^5) \tag{14}$$

4.2. Error Analysis for the Black-box Models

To evaluate the quality of the fit to target data and to make comparison between the SVRM and neural models, the following error terms are found to be convenient as in [4]:

$$\text{Accuracy}_X = 1 - \frac{\sum_{k=1}^{n} \left| X_{k_{\text{target}}} - X_{k_{\text{predicted}}} \right|}{\sum_{k=1}^{n} X_{k_{\text{target}}}} \tag{15}$$

where $X_{k_{\text{predicted}}}$ is the kth predicted value for the kth target value $X_{k_{\text{target}}}$, n is the total number of data.

4.3. Knowledge-based Support Vector Synthesis of the Microstrip Line

Approach for knowledge-based SVRM synthesis is given in Fig. 1, where in the first stage, the coarse data generator provides the coarse analysis data depending on the given transmission line geometry and dielectric permittivity parameters. In this work, we applied this approach to the microstrip lines with the empirical analysis formulas [15] used as the coarse data generator. The employed transmission line parameter range is given in Table 1. Thus 1290 data are found to be sufficient to analyze the microstrip lines within the given parameter range. Then the coarse synthesis data is generated by reversing the coarse analysis data subject to the definition of the synthesis process. In our work, this corresponds to determining width, W on a chosen dielectric substrate (H, ε_r) for the required characteristic impedance, $\mathbf{Z}_0$. These coarse synthesis data are inputted to the SVRM model to obtain the coarse synthesis SVs. Accuracy of

Table 1. Microstrip line parameter range.

Variable	Training and test data	
	Min	**Max**
H (mm)	0.1	10
ε_r	2	13
$\mathbf{Z}_0$ (Ω)	11	200
W (mm)	0.1	10

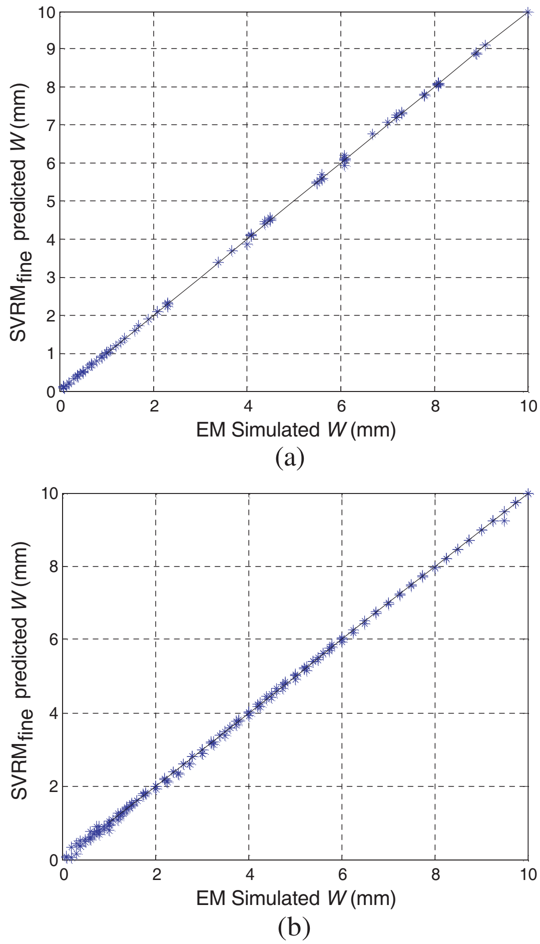


Figure 3. Scatter plots of the EM simulated and $\mathrm{SVRM_{fine}}$ predicted W for the (a) training data; (b) test data.

this $\mathrm{SVRM_{coarse}}$ (trained with coarse data) model is given depending on ε tolerance-parameter in Table 2, where $\varepsilon = 0.25$ and 151 SVs with the accuracy of % 99.10 are assumed to be sufficient to be input to the fine data generator.

Attempts to develop microstrip synthesis models with 3-D EM accuracy using fine data from a 3-D EM Simulator alone have proven to be computationally prohibitive, as each 3-D EM simulation needs a lot of CPU time. This led to a need for examining possibilities of using inexpensive coarse data together with fine data for efficient

Table 2. Parameter selection for the $SVRM_{coarse}$ model.

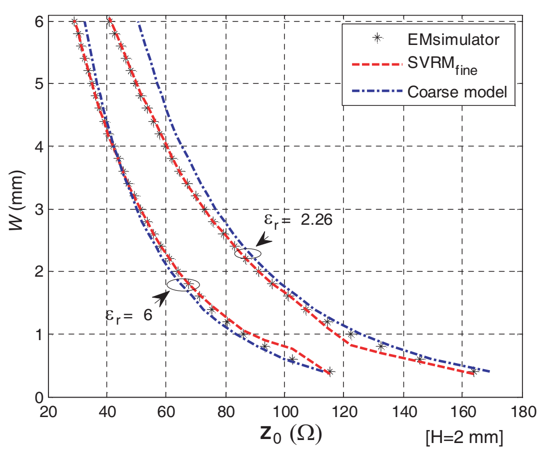
ε	Number of SVs	Accuracy (%)
0.1	361	99.49
0.15	264	99.28
0.2	195	99.29
0.25*	151	99.10
0.3	113	98.94

* ε=0.25 is used in the worked example.

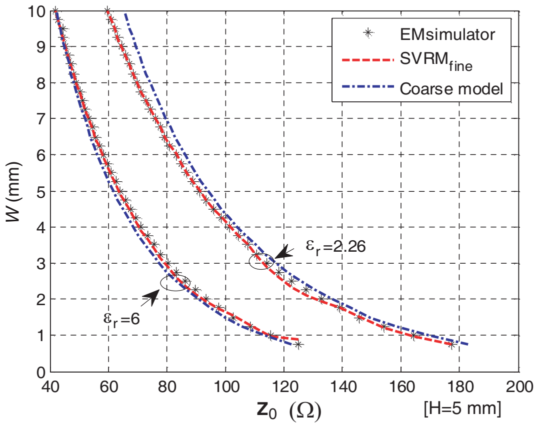
Table 3. Accuracies for the synthesis $SVRM_{fine}$ model of the microstrip line.

$\varepsilon = 0.01$	**Accuracy %**
Training	99.6859
Testing	98.9155

model development. Thus in the proposed method, fine synthesis SVs are generated by applying only the coarse synthesis SVs to the 3-D EM Simulator instead of the whole coarse synthesis data. Consequently, CPU time for data generation is significantly reduced proportionally with the reduced data, as investigated in the work [14]. Since fine data generation time is the major constituent of the total model development time, it can be concluded that model development using the proposed knowledge-based SVRM is much faster than the existing



(a)



(b)

Figure 4. SVRM$_{\text{fine}}$ model compared with EM simulator and the coarse model for the microstrip line onto the dielectric materials: PTFE\microfiberglass with $\varepsilon_r = 2.26$ and mica with $\varepsilon_r = 6$ for substrate thickness (a) $H = 2\,\text{mm}$; (b) $H = 5\,\text{mm}$.

modeling approaches.

Finally, fine SV expansion is obtained inputting these fine synthesis SVs to the SVRM model box. Thus the resulted accuracies are given for training and testing in Table 3. Furthermore the related scatter plots are presented in Fig. 3. It can be seen from the Table 3 and Fig. 3 that % 98.9155 of accuracy is achieved with the % 11.7 (151/1290) of the original data. This reduction is also valid for CPU time of data generation. Besides, data reduction also results in reduction in training time of the model. Figs. 4(a), (b) give plots of functions $W(\mathbf{Z}_0, \varepsilon_r, H)$ resulted from the models of the SVRM$_{\text{fine}}$ (trained with fine data), fine and coarse data generators.

5. CONCLUSIONS

We have proposed a robust knowledge-based SVM method and applied it to generation of synthesis model for microwave applications. The aim of this work is to obtain a synthesis model as fast as the coarse models and at the same time as accurate as the fine models. The technique has integrated advanced concepts of SVM and knowledge-based modeling into a powerful and systematic framework. Motivated by the working principle based on the small sample statistical learning theory of SVM, the knowledge-based SVM method simultaneously exploits coarse and fine data generators for efficient model development. The advantages of the knowledge-based SVM method are demonstrated through practical example of microstrip line to be used in conventional monolithic or hybrid MICs. For a specified accuracy, the proposed method uses the fewest fine data compared to conventional modeling methodologies. Reduced fine training data result in significantly reduced CPU time for data generation and also faster model development process.

ACKNOWLEDGMENT

The author would like to thank The Scientific and Technological Research Council of Turkey for financially supporting this research under Contract No. (106E171).

REFERENCES

1. Zhang, Q. J. and K. C. Gupta, *Neural Networks for RF and Microwave Design*, Artech House, Norwood, MA, 2000.
2. Vapnik, V., *The Nature of Statistical Learning Theory*, Springer-Verlag, New York, 1995.
3. Bermani, E., A. Boni, A. Kerhet, and A. Massa, "Kernels evaluation of SVM based estimatiors for inverse scattering problems," *Progress In Electromagnetics Research*, PIER 53, 167–188, 2005.
4. Güneş, F., N. T. Tokan, and F. Gürgen, "Signal-noise support vector model of a microwave transistor," *Int. J. RF and Microwave CAE*, Vol. 17, 404–415, 2007.
5. Güneş, F., N. T. Tokan, and F. Gürgen, "Support vector design of the microstrip lines," *Int. J. RF and Microwave CAE*, Vol. 18, 326–336, 2008.
6. Tokan, N. T. and F. Güneş, "Support vector characterisation of resonance frequencies of microstrip antennas based on measurements," *Progress In Electromagnetics Research B*, Vol. 5, 49–51, 2008.
7. Wang, F. and Q. J. Zhang, "Knowledge-based neural models for microwave design," *IEEE Trans. Microwave Theory Tech.*, Vol. 45, 2333–2343, Dec. 1997.
8. Watson, P. M., K. C. Gupta, and R. L. Mahajan, "Applications of knowledge-based artificial neural network modeling to microwave components," *Int. J. RF Microwave Computer-aided Eng.*, Vol. 9, 254–260, 1999.

9. Jargon, J. A., K. C. Gupta, and D. C. DeGroot, "Applications of artificial neural networks to RF and microwave measurements," *Int. J. RF Microwave Computer-aided Eng.*, Vol. 12, 3–24, 2002.
10. Bandler, J. W., M. A. Ismail, J. E. Rayas-Sanchez, and Q. J. Zhang, "Neuromodeling of microwave circuits exploiting space-mapping technology," *IEEE Trans. Microwave Theory Tech.*, Vol. 47, 2417–2427, Dec. 1999.
11. Bakr, M. H., J. W. Bandler, M. A. Ismail, J. E. Rayas-Sanchez, and Q. J. Zhang, "Neural space-mapping optimization for EM-based design," *IEEE Trans. Microwave Theory Tech.*, Vol. 48, 2307–2315, Dec. 2000.
12. Devabhaktuni, V. K., M. C. E. Yagoub, and Q. J. Zhang, "A robust algorithm for automatic development of neural network models for microwave applications," *IEEE Trans. Microwave Theory Tech.*, Vol. 49, 2282–2291, Dec. 2001.
13. Bandler, J. W., J. E. Rayas-Sanchez, and Q. J. Zhang, "Yield-driven electromagnetic optimization via space mapping-based neuromodels," *Int. J. RF Microwave Computer-aided Eng.*, Vol. 12, 79–89, 2002.
14. Devabhaktuni, V. K., B. Chattaraj, M. C. E. Yagoub, and Q.-J. Zhang, "Advanced microwave modeling framework exploiting automatic model generation, knowledge neural networks, and space mapping," *IEEE Trans. Microwave Theory Tech.*, Vol. 51, No. 7, 1822–1833, Jul. 2003.
15. Edwards, T. C., "Foundations for microstrip circuit design," Wiley-Interscience, New York, 1981.