Support-Vector-Regression-Based Output Space-Mapping for Microwave Device Modeling

Slawomir Koziel, *Senior Member*, *IEEE*, and John W. Bandler, *Life Fellow*, *IEEE*

School of Science and Engineering, Reykjavik University, Kringlunni 1, IS-103 Reykjavik, Iceland
Department of Electrical and Computer Eng., McMaster University, Hamilton, ON, Canada L8S 4K1

***Abstract*—An enhancement of the space mapping (SM) surrogate model through support vector regression is presented. This technique uses a standard SM model (trend function) and support vector regression to model the residuals between the fine model and the standard model. The latter is implemented as a additive output SM term. The proposed methodology offers efficient utilization of the available fine model data (not possible in the standard SM modeling) and accuracy comparable or better than the recently published modeling techniques combining SM with radial basis functions and fuzzy systems. Examples demonstrate the robustness of our approach.**

***Index Terms*—Computer-aided design (CAD), EM modeling, space mapping, surrogate modeling, support vector regression.**

I. Introduction

Accurate and computationally efficient models of microwave components and devices are crucial in many areas such as signal processing, wireless communication and biomedical engineering. Full-wave EM simulations of microwave structures offer high accuracy at the cost of CPU effort, which is undesirable from the point of view of direct statistical analysis and design. The space mapping (SM) concept [1]-[4] addresses this issue. Space mapping assumes the existence of "fine" and "coarse" models. The "fine" model may be a high fidelity CPU-intensive EM simulator. The "coarse" model can be a simplified representation such as an equivalent circuit with empirical formulas. SM modeling [5]-[9] and neuro-space-mapping modeling [4], [10], [11] exploit the speed of the coarse model and the accuracy of the fine model to develop fast, accurate enhanced models (surrogates) valid over a wide range of parameter values.

The standard SM modeling [6] sets up the surrogate model using a small amount of fine-model data with extraction of the model parameters performed over the whole set of this data. This methodology is simple and gives reasonable accuracy, which, however, may not be sufficient for some applications.

This work was supported in part by the Natural Sciences and Engineering Research Council of Canada under Grants RGPIN7239-06, STPGP336760-06, and by Bandler Corporation.

S. Koziel was with the Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada L8S 4K1. He is now with the School of Science and Engineering, Reykjavik University, Kringlunni 1, IS-103 Reykjavik, Iceland.

J.W. Bandler is with the Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, ON, Canada L8S 4K1 and also with Bandler Corporation, Dundas, ON, Canada L9H 5E7.

SM modeling with variable weight coefficients [7] provides better modeling accuracy, however, at the expense of some computational overhead related to a separate parameter extraction required for each evaluation of the surrogate model. This limits potential applications of the method.

SM modeling enhanced by radial basis function interpolation [8] and SM modeling with fuzzy systems [9] give modeling accuracy comparable with [7] without compromising computational cost. Unfortunately, the problem of determining the interpolation coefficients in [8] may be ill-conditioned and the method may be very sensitive to some control parameters. Model [9], on the other hand, may not be differentiable, which makes it difficult to optimize and hence not suitable for some applications. Also, model [9] works well if the base set is a rectangular grid; otherwise its performance may be degraded.

In this paper, we present other approach that uses standard space mapping enhanced by support vector regression (SVR) [12]. SVR is implemented as an additive output SM term that models the differences between the fine and standard SM model responses at the base points. SVR is characterized by good generalization capability [13] and easy training through quadratic programming resulting in a global optimum for the model parameters [14]. We demonstrate that the accuracy of the new surrogate model is competitive with the accuracy of previously published SM modeling approaches as well as the direct support vector regression of the fine model data.

II. Surrogate Modeling with Space Mapping and SVR

Let $\boldsymbol{R}_f : X_f \to R^m$, $X_f \subseteq R^n$, and $\boldsymbol{R}_c : X_c \to R^m$, $X_c \subseteq R^n$, denote the fine and coarse model response vectors. For example, $\boldsymbol{R}_f(\boldsymbol{x})$ and $\boldsymbol{R}_c(\boldsymbol{x})$ may represent the magnitude of a transfer function at m chosen frequencies. We denote by $X_R \subseteq X_f$ the region of interest in which we want an enhanced matching between the surrogate and the fine model. Typically, X_R is an n-dimensional interval in R^n with center at reference point $\boldsymbol{x}^0 = [x_{0.1} \ldots x_{0.n}]^T \in R^n$ and size $\boldsymbol{\delta} = [\delta_1 \ldots \delta_n]^T$ [6].

Suppose we have the base set $X_B = \{\boldsymbol{x}^1, \boldsymbol{x}^2, \ldots, \boldsymbol{x}^N\} \subset X_R$, where N is the number of base points, such that the fine model response is known at all points $\boldsymbol{x}^j$, $j = 1, 2, \ldots, N$.

Let $\bar{\boldsymbol{R}}_s : X_R \times X_p \to R^m$ be a generic SM surrogate model where X_p is a parameter domain. For any given base set X_B the standard surrogate model $\boldsymbol{R}_{s.SM}$ is defined as

$$\boldsymbol{R}_{s.SM}(\boldsymbol{x}) = \bar{\boldsymbol{R}}_s(\boldsymbol{x}, \bar{\boldsymbol{p}}) \tag{1}$$

where

$$\bar{\boldsymbol{p}} = \arg\min_{\boldsymbol{p}\in X_p} \sum_{k=1}^{N} \| \boldsymbol{R}_f(\boldsymbol{x}^k) - \bar{\boldsymbol{R}}_s(\boldsymbol{x}^k, \boldsymbol{p}) \| \quad (2)$$

A variety of SM surrogate models is available [1]-[4]. The model often used in practice (e.g., [6]) employs both input and output SM, i.e., $\bar{\boldsymbol{R}}_s(\boldsymbol{x},\boldsymbol{p}) = \bar{\boldsymbol{R}}_s(\boldsymbol{x},\boldsymbol{A},\boldsymbol{B},\boldsymbol{c}) = \boldsymbol{A}\cdot\boldsymbol{R}_c(\boldsymbol{B}\cdot\boldsymbol{x}+\boldsymbol{c})$. It is often enhanced by a frequency SM [6]. More general SM models can be found, e.g., in [2].

Let $\boldsymbol{R}^k = \boldsymbol{R}_f(\boldsymbol{x}^k) - \boldsymbol{R}_{s.SM}(\boldsymbol{x}^k)$ for $k = 1, 2, ..., N$. We want to enhance the standard SM model by an additive term approximating the residuals $\boldsymbol{R}^k$ at all base points. We shall also use the notation $\boldsymbol{R}^k = [R_1^k\ R_2^k \ldots R_m^k]^T$ to denote components of vector $\boldsymbol{R}^k$. Approximation of $\boldsymbol{R}^k$ is implemented using so-called support vector regression [12]. This technique is a variant of the support vector machines methodology developed by Vapnik [15], which was originally applied to solve classification problems. Support vector regression is gaining popularity in the microwave engineering area (e.g., [13]). In the case of linear regression, we want to approximate a given set of data, in our case, the data pairs $D_j = \{(\boldsymbol{x}^1,R_j^1),\ldots,(\boldsymbol{x}^N,R_j^N)\}$, $j = 1, 2, \ldots, m$, by a linear function $f_j(\boldsymbol{x}) = \boldsymbol{w}_j^T\boldsymbol{x} + b_j$. The optimal regression function is given by the minimum of the functional [12]

$$\Phi_j(\boldsymbol{w},\xi) = \frac{1}{2}\|\boldsymbol{w}_j\|^2 + C_j\sum_{i=1}^{N}(\xi_{j.i}^+ + \xi_{j.i}^-) \quad (3)$$

where C_j is a user-defined value, and $\xi_{j.i}^+$ and $\xi_{j.i}^-$ are slack variables representing upper and lower constraints on the output of the system. The typical cost function used in support vector regression is the so-called ε-insensitive loss function

$$L_\varepsilon(y) = \begin{cases} 0 & \text{for } |f_j(\boldsymbol{x}) - y| < \varepsilon \\ |f_j(\boldsymbol{x}) - y| & \text{otherwise} \end{cases} \quad (4)$$

The value of C_j determines the trade-off between the flatness of f_j and the amount up to which deviations larger than ε are tolerated [12].

In this paper, we use nonlinear regression employing the kernel approach, in which the linear function $\boldsymbol{w}_j^T\boldsymbol{x} + b_j$ is replaced by the nonlinear function $\Sigma_i\gamma_{j.i}K(\boldsymbol{x}^k,\boldsymbol{x}) + b_j$, where K is a kernel function. Thus, the SVR term used to enhance the standard SM is defined as

$$\boldsymbol{R}_{s.SVR} = \begin{bmatrix} \sum_{i=1}^{N}\gamma_{1.i}K(\boldsymbol{x}^i,\boldsymbol{x}) + b_1 \\ \vdots \\ \sum_{i=1}^{N}\gamma_{m.i}K(\boldsymbol{x}^i,\boldsymbol{x}) + b_m \end{bmatrix} \quad (5)$$

with parameters $\gamma_{j.i}$ and b_j, $j = 1, \ldots, m$, $i = 1, \ldots, N$ obtained according to a general support vector regression methodology. In this paper we use Gaussian kernels of the form

$$K(\boldsymbol{x},\boldsymbol{y}) = \exp\left(-\frac{\|\boldsymbol{x}-\boldsymbol{y}\|^2}{2c^2\lambda^2}\right) \qquad c > 0 \quad (6)$$

where $\lambda = \lambda(\boldsymbol{\delta},N)$—used here as an normalization factor—is a so-called characteristic distance of the base set defined as [7]

$$\lambda(\boldsymbol{\delta},N) = \frac{2}{nN^{1/n}}\sum_{i=1}^{n}\delta_i \quad (7)$$

The scaling parameter c as well as parameters C_j and ε are adjusted to minimize the generalization error calculated using a cross-validation method and exponential grid search.

The overall surrogate model is defined as follows

$$\boldsymbol{R}_s(\boldsymbol{x}) = \boldsymbol{R}_{s.SM}(\boldsymbol{x}) + \boldsymbol{R}_{s.SVR}(\boldsymbol{x}) \quad (8)$$

Similarly as in the case of combining space mapping with radial basis functions [8] and fuzzy systems [9], the surrogate model (8) ensures good accuracy and, at the same time, computational efficiency almost the same as the underlying coarse model. Model parameters are determined as a convex optimization problem, which result in a unique global optimum, which is in contrast to [8] where the problem of obtaining model parameters may be ill-conditioned. Also, the support vector regression function is smooth, which may not be the case for modeling with fuzzy systems [9].

III. Examples

In this section, we compare the modeling accuracy for the standard SM modeling methodology [6] (SM-Standard), SM modeling with variable weight coefficients [7] (SM-VWC), SM with radial basis function interpolation [8] (SM-RBF), the SM with fuzzy systems [9] (SM-Fuzzy) and the combination of SM with SVR described in Section II (SM-SVR). In our comparison we also include direct approximation of the fine model data using SVR.

A. Test Problem Description

Problem 1: Microstrip right-angle bend [5]. The fine model, Fig. 1(a), is analyzed by Sonnet's ***em*** [16]. The coarse model is an equivalent circuit shown in Fig. 1(b). The design parameters are $\boldsymbol{x} = [W\ H\ \varepsilon_r]^T$. The response vector consists of reflection coefficient $|S_{11}|$ in the frequency range of 1 to 31 GHz. The reference point is $\boldsymbol{x}^0 = [25\ 12\ 9]^T$, and the region size is $\boldsymbol{\delta} = [6\ 4\ 1]^T$.

Problem 2: Bandstop microstrip filter with open stubs [10] shown in Fig. 2. The fine model is simulated with Sonnet's ***em*** [16] using a high-resolution grid with a 0.2 mil × 1 mil cell size. The coarse model, Fig. 3, is the equivalent circuit model implemented in Agilent ADS [17]. The design parameters are $\boldsymbol{x} = [W_1\ W_2\ L_0\ L_1\ L_2]^T$. The response vector consists of transmission coefficient $|S_{21}|$ in the frequency range of 5 to 15 GHz. The reference point is $\boldsymbol{x}^0 = [5.6\ 10.4\ 119.2\ 118\ 112]^T$ mil and the region size is $\boldsymbol{\delta} = [0.4\ 0.4\ 2\ 2\ 2]^T$ mil.

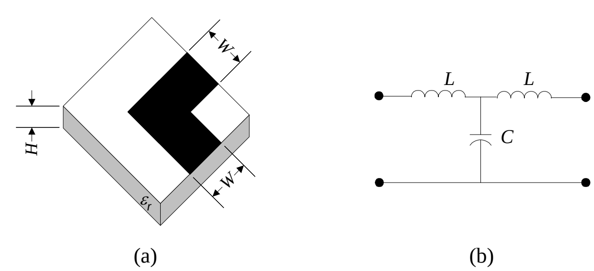


Fig.1. The microstrip right-angle bend [5]: the fine model (a) and the coarse model (b).

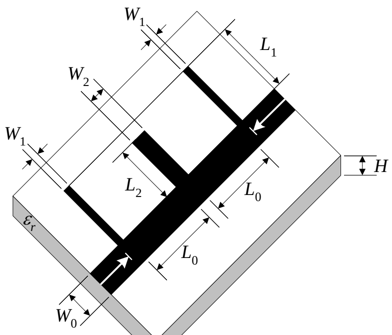


Fig.2. Bandstop microstrip filter with open stubs [10].

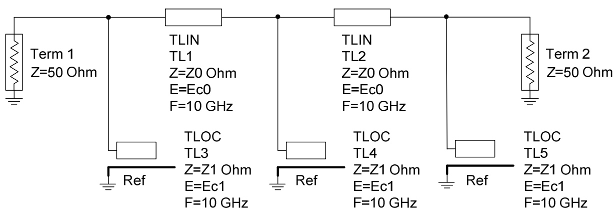


Fig.3. Bandstop microstrip filter with open stubs: coarse model (Agilent ADS).

B. Experimental Setup

For both test problems we performed a number of experiments using models: SM-Standard, SM-VWC, SM-RBF, SM-Fuzzy, SM-SVR, and direct support vector regression.

Table I shows details of the base sets used in our experiments. The base sets have growing numbers of points in order to examine the dependence of the modeling error on the amount of fine model data used to create the model. The standard SM model uses the generic model $\bar{\boldsymbol{R}}_s(\boldsymbol{x},\boldsymbol{p})=\bar{\boldsymbol{R}}_s(\boldsymbol{x},\boldsymbol{A},\boldsymbol{B},\boldsymbol{c})=\boldsymbol{A}\cdot\boldsymbol{R}_c(\boldsymbol{B}\cdot\boldsymbol{x}+\boldsymbol{c})$ enhanced by frequency SM. Accuracy was tested using 30 test points randomly distributed in the region of interest. The error measure used was the l_2 norm of the difference between the fine model response and the corresponding surrogate model response.

C. Experimental Results and Discussion

Tables II and III show numerical results (error statistics) for the considered surrogate models with the base sets X_{B1} to X_{B3}. Figs. 4 and 5 show error plots (the modulus of the difference between the fine model and the corresponding surrogate model response versus frequency) for the SM-Standard and SM-SVR with base set X_{B3}, for Problems 1 and 2 respectively. Figs. 6 and 7 show dependence of average modeling error on the characteristic distance λ for all surrogate models considered.

The results show that the new SM-SVR model provides modeling accuracy comparable or better than the best space mapping models known so far, i.e., SM-RBF and SM-Fuzzy. It should also be emphasized that the SM-SVR model does not have drawbacks of the SM-VWC, the SM-RBF and the SM-Fuzzy models, which were mentioned in the introduction. Its computational complexity is similar to the SM-RBF and the SM-Fuzzy models. Altogether, SM-SVR seems to be an attractive alternative to the existing space mapping modeling approaches.

TABLE I
BASE SET DATA FOR TEST PROBLEMS 1 AND 2

Test Problem	Base Set	Base Set Description	Number of Base Points	λ
1	X_{B1}	Uniform mesh of density 3	27	2.44
	X_{B2}	Uniform mesh of density 4	64	1.83
	X_{B3}	Uniform mesh of density 5	125	1.47
2	X_{B1}	Star distribution	11	1.68
	X_{B2}	Uniform mesh of density 2 ∪ star distribution	43	1.28
	X_{B3}	Uniform mesh of density 2 ∪ star distribution ∪ 30 edge* points	73	1.15

*30 point randomly chosen out of points in the uniform mesh of density 3 but not belonging to X_{B2}

TABLE II
MODELING RESULTS FOR TEST PROBLEM 1.
VERIFICATION FOR 30 RANDOM TEST POINTS

Model	Base set	Average Error	Maximum Error	Standard Deviation
SM-Standard [6]	X_{B1}	0.0094	0.0191	0.0037
SM-VWC [7]		0.0089	0.0541	0.0096
SM-RBF [8]		0.0013	0.0036	0.0008
SM-Fuzzy [9]		0.0051	0.0089	0.0016
SM-SVR (1)-(8)		0.0027	0.0052	0.0013
Direct SVR approximation		0.0253	0.0635	0.0142
SM-Standard [6]	X_{B2}	0.0083	0.0184	0.0038
SM-VWC [7]		0.0036	0.0093	0.0020
SM-RBF [8]		0.0011	0.0027	0.0009
SM-Fuzzy [9]		0.0015	0.0042	0.0009
SM-SVR (1)-(8)		0.00047	0.0014	0.00036
Direct SVR approximation		0.0055	0.0140	0.0037
SM-Standard [6]	X_{B3}	0.0079	0.0175	0.0038
SM-VWC [7]		0.0023	0.0066	0.0014
SM-RBF [8]		0.0009	0.0030	0.0006
SM-Fuzzy [9]		0.0008	0.0019	0.0005
SM-SVR (1)-(8)		0.00021	0.00072	0.00016
Direct SVR approximation		0.0011	0.0033	0.0008

TABLE III
MODELING RESULTS FOR TEST PROBLEM 2.
VERIFICATION FOR 30 RANDOM TEST POINTS

Model	Base set	Average Error	Maximum Error	Standard Deviation
SM-Standard [6]	X_{B1}	0.0389	0.0635	0.0084
SM-VWC [7]		0.0326	0.0483	0.0058
SM-RBF [8]		0.0055	0.0154	0.0038
SM-Fuzzy [9]		0.0222	0.0581	0.0141
SM-SVR (1)-(8)		0.0061	0.0171	0.0041
Direct SVR approximation		0.0263	0.0958	0.0216
SM-Standard [6]	X_{B2}	0.0403	0.0581	0.0073
SM-VWC [7]		0.0296	0.0440	0.0051
SM-RBF [8]		0.0024	0.0108	0.0022
SM-Fuzzy [9]		0.0176	0.0441	0.0102
SM-SVR (1)-(8)		0.0021	0.0109	0.0021
Direct SVR approximation		0.0060	0.0169	0.0044
SM-Standard [6]	X_{B3}	0.0404	0.0599	0.0074
SM-VWC [7]		0.0243	0.0343	0.0061
SM-RBF [8]		0.0022	0.0110	0.0021
SM-Fuzzy [9]		0.0139	0.0421	0.0086
SM-SVR (1)-(8)		0.0017	0.0108	0.0022
Direct SVR approximation		0.0038	0.0111	0.0023

IV. Conclusion

A new SM-based modeling methodology is presented which combines the standard space mapping with support vector regression. This method not only provides modeling accuracy competitive with recently published space mapping models enhanced by radial basis functions and fuzzy systems, but also overcomes some of the drawbacks of these techniques.

References

[1] J.W. Bandler, Q.S. Cheng, S.A. Dakroury, A.S. Mohamed, M.H. Bakr, K. Madsen, and J. Søndergaard, "Space mapping: the state of the art," *IEEE Trans. Microwave Theory Tech.*, vol. 52, no. 1, pp. 337-361, Jan. 2004.

[2] S. Koziel, J.W. Bandler, and K. Madsen, "A space mapping framework for engineering optimization: theory and implementation," *IEEE Trans. Microwave Theory Tech*., vol. 54, no. 10, pp. 3721-3730, Oct. 2006.

[3] D. Echeverria and P.W. Hemker, "Space mapping and defect correction," *CMAM The International Mathematical Journal Computational Methods in Applied Mathematics*, vol. 5, no. 2, pp. 107-136, 2005.

[4] J.E. Rayas-Sánchez and V. Gutiérrez-Ayala, "EM-based Monte Carlo analysis and yield prediction of microwave circuits using linear-input neural-output space mapping," *IEEE Trans. Microwave Theory Tech.*, vol. 54, no. 12, pp. 4528-4537, Dec. 2006.

[5] J.W. Bandler, N. Georgieva, M.A. Ismail, J.E. Rayas-Sánchez, and Q. J. Zhang, "A generalized space mapping tableau approach to device modeling," *IEEE Trans. Microwave Theory Tech.*, vol. 49, no. 1, pp. 67-79, Jan. 2001.

[6] S. Koziel, J.W. Bandler, A.S. Mohamed, and K. Madsen, "Enhanced surrogate models for statistical design exploiting space mapping technology," *IEEE MTT-S Int. Microwave Symp. Dig.*, Long Beach, CA, June 2005, pp. 1609-1612.

[7] S. Koziel, J.W. Bandler, and K. Madsen, "Theoretical justification of space-mapping-based modeling utilizing a data base and on-demand parameter extraction," *IEEE Trans. Microwave Theory Tech*., vol. 54, no. 12, pp. 4316-4322, Dec. 2006.

[8] S. Koziel and J.W. Bandler, "Microwave device modeling using space-mapping and radial basis functions," *IEEE MTT-S Int. Microwave Symp. Dig*, Honolulu, HI, 2007, pp. 799-802.

[9] S. Koziel and J.W. Bandler, "A space-mapping approach to microwave device modeling exploiting fuzzy systems," *IEEE Trans. Microwave Theory Tech.*, vol. 55, no. 12, pp. 2539-2547, Dec. 2007.

[10] M.H. Bakr, J.W. Bandler, M.A. Ismail, J.E. Rayas-Sánchez, and Q.J. Zhang, "Neural space-mapping optimization for EM-based design," *IEEE Trans. Microwave Theory Tech.*, vol. 48, no. 12, pp. 2307-2315, Dec. 2000.

[11] L. Zhang, J.J. Xu, M. Yagoub, R.T. Ding, and Q.J. Zhang, "Neuro-space mapping technique for nonlinear device modeling and large signal simulation," *IEEE MTT-S Int. Microwave Symp. Dig.*, Philadelphia, PA, June 2003, pp. 173-176.

[12] S.R. Gunn, "Support vector machines for classification and regression," Tech. Rep., School of Electronics and Computer Science, University of Southampton, 1998.

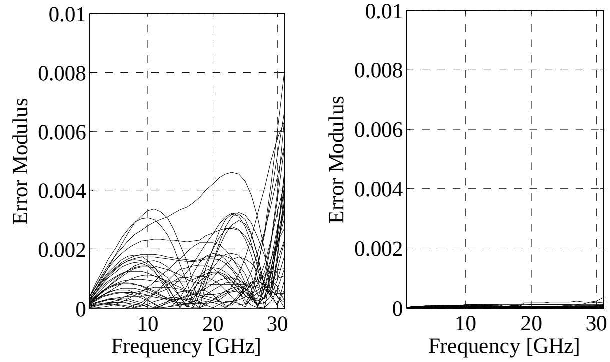
[13] G. Angiulli, M. Cacciola, and M. Versaci, "Microwave devices and antennas modeling by support vector regression machines," *IEEE Trans. Magnetics*, vol. 43, no. 4, pp. 1589-1592, Apr. 2007.

[14] A.J. Smola and B. Schölkopf, "A tutorial on support vector regression," *Statistics and Computing*, vol. 14, no. 3, pp. 199-222, Aug. 2004.

[15] V.N. Vapnik, *The Nature of Statistical Learning Theory*. New York: Springer Verlag, 1995.

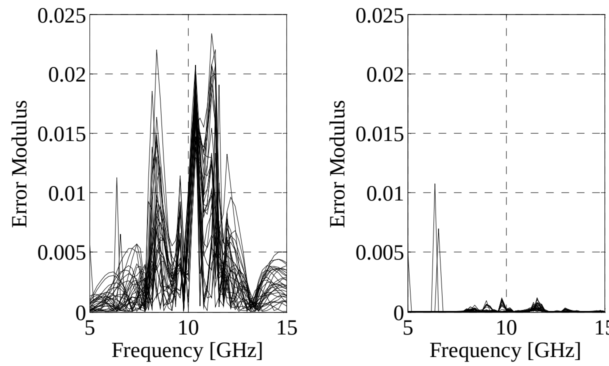
[16] ***em***™ Version 10.52, Sonnet Software, Inc., 100 Elwood Davis Road, North Syracuse, NY 13212, USA.

[17] Agilent ADS, Version 2003C, Agilent Technologies, 1400 Fountaingrove Parkway, Santa Rosa, CA 95403-1799, 2003.



(a) (b)

Fig.4. Test problem 1: error plots for the SM-Standard (a) and SM-SVR (b) surrogate models with base set X_{B3} (30 test points).



(a) (b)

Fig.5. Test problem 2: error plots for the SM-Standard (a) and SM-SVR (b) surrogate models with base set X_{B3} (30 test points).

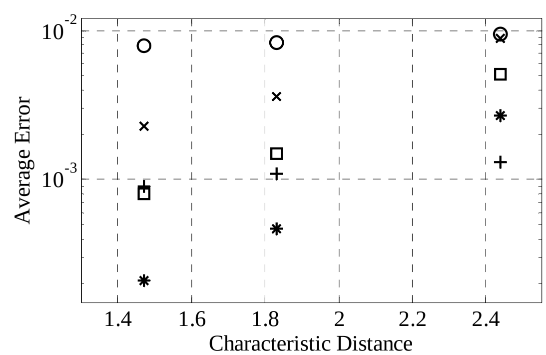


Fig.6. Test problem 1: average modeling error versus λ: SM-Standard (o), SM-VWC (×), SM-RBF (+), SM-Fuzzy (□), and SM-SVR (∗).

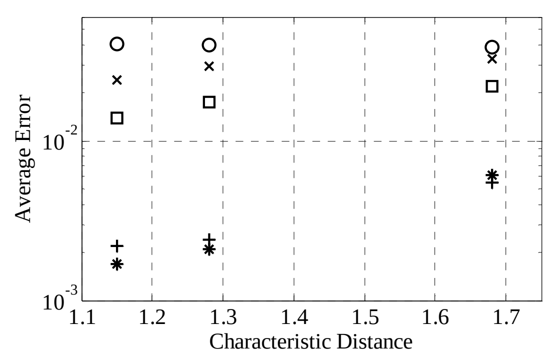


Fig.7. Test problem 2: average modeling error versus λ: SM-Standard (o), SM-VWC (×), SM-RBF (+), SM-Fuzzy (□), and SM-SVR (∗).