INTERNATIONAL JOURNAL OF NUMERICAL MODELLING: ELECTRONIC NETWORKS, DEVICES AND FIELDS
Int. J. Numer. Model. 2010; **23**:425–446
Published online 15 January 2010 in Wiley Online Library (wileyonlinelibrary.com). DOI: 10.1002/jnm.743

Recent advances in space-mapping-based modeling of microwave devices

Slawomir Koziel[1,*,†] and John W. Bandler[2,3]

[1]*School of Science and Engineering, Reykjavik University, Kringlunni 1, IS-103 Reykjavik, Iceland*
[2]*Simulation Optimization Systems Research Laboratory, Department of Electrical and Computer Engineering, McMaster University, Hamilton, Ont., Canada L8S 4K1*
[3]*Bandler Corporation, Dundas, Ont., Canada L9H5E7*

SUMMARY

We review the latest developments in space-mapping-based modeling techniques with applications in microwave engineering. We discuss the two techniques that utilize a combination of standard space mapping and function approximation methodologies, in particular fuzzy systems and support vector regression (SVR). In both cases, the initial space-mapping model is enhanced by an additional term that approximates the differences between the fine model and the initial space-mapping surrogate. We compare the standard and enhanced space-mapping models, as well as the fuzzy systems and SVR directly used for modeling fine model data. A discussion of the advantages and disadvantages of the presented methods is also given. Copyright © 2010 John Wiley & Sons, Ltd.

Received 17 November 2008; Revised 27 January 2009; Accepted 24 October 2009

KEY WORDS: computer-aided design (CAD); EM modeling; space mapping; surrogate modeling; fuzzy systems; support vector regression

1. INTRODUCTION

Electromagnetic simulators offer accurate evaluation of microwave structures and devices. Unfortunately, the high computational cost of simulation typically prohibits their direct application in tasks such as statistical analysis and yield optimization, which are crucial for manufacturability-driven designs in a time-to-market development environment. Space mapping [1–10] addresses this issue by creating computationally cheap surrogate models that

*Correspondence to: Slawomir Koziel, School of Science and Engineering, Reykjavik University, Kringlunni 1, IS-103 Reykjavik, Iceland.
†E-mail: koziel@ru.is

Contract/grant sponsors: Reykjavik University Development Fund; Natural Sciences and Engineering Research Council of Canada; contract/grant numbers: T09009; RGPIN7239-06 and STPGP336760-06

in many cases are sufficiently accurate to be used instead of the simulator-based models for solving microwave design problems.

Space mapping assumes the existence of 'fine' and 'coarse' models. The 'fine' model may be a high fidelity CPU-intensive EM simulator, undesirable for direct statistical analysis and design. The cheap 'coarse' model is typically a simplified representation such as an equivalent circuit with empirical formulas. The space-mapping-based surrogate model is created by enhancing the coarse model using a limited amount of fine model data obtained by evaluating the fine model at certain base points. The enhancement is typically realized through suitable analytical formulas, which allows the surrogate model to be almost as computationally cheap as the coarse model. On the other hand, because the coarse model is supposedly physics-based, the accuracy of the space-mapping surrogate is considerably better than the accuracy of possible function approximation models [11–18] using a comparable amount of fine model data.

A number of space-mapping modeling [19–29] and neuro-space-mapping modeling [30–32] approaches have been proposed recently.

The standard SM modeling approach is based on setting up the surrogate model using a small amount of fine model data and performing extraction of the mapping parameters over the whole set of this data [20, 21]. This simple methodology gives reasonable accuracy especially for low-dimensional problems, however, introducing additional degrees of freedom to handle a larger amount of fine model data (necessary to improve the surrogate model accuracy above certain limits) is problematic [25].

Several techniques have been proposed to overcome these limitations. A space-mapping modeling approach with variable weight coefficients [24, 25] provides better accuracy than the standard method, however, at the expense of significant increase of the evaluation time, which is due to a separate parameter extraction required for each evaluation of the surrogate model. This limits potential applications of the method. In the reference [26], a combination of standard space mapping with radial basis function interpolation is described. This gives modeling accuracy comparable or better than the variable weight method [28] without compromising computational cost. Unfortunately, the problem of determining interpolation coefficients may be ill-conditioned and the method may be sensitive to some control parameters.

In this paper, we describe the two techniques that use a combination of standard space mapping with fuzzy systems [27] and support vector machines (SVM) [29]. Both approaches proved to be superior to other space-mapping-based techniques published so far [27, 29]. We compare them with the standard space mapping used here as a reference method, as well as to direct approximation of fine model data using fuzzy systems and SVM. We also provide some practical guidelines regarding applications of the methods for modeling problems of different dimensionality.

2. STANDARD SPACE-MAPPING MODELING

Let $\mathbf{R}_f: X_f \rightarrow R^m$ and $\mathbf{R}_c: X_c \rightarrow R^m$ denote the fine and coarse model response vectors, where $X_f \subseteq R^n$ and $X_c \subseteq R^n$ are design variable domains of the fine and coarse models, respectively. In particular, $\mathbf{R}_f(\mathbf{x})$ and $\mathbf{R}_c(\mathbf{x})$ may represent the magnitude of a transfer function of a microwave filter at m chosen frequencies. We denote by $X_R \subseteq X_f$ the region of interest in which we want enhanced matching between the surrogate and the fine model. Typically, X_R is an n-dimensional interval in R^n centered at reference point $\mathbf{x}^0 = [x_{0.1} \ldots x_{0.n}]^T \in R^n$,

DOI: 10.1002/jnm

i.e. $X_R = [\mathbf{x}^0 - \boldsymbol{\delta}, \mathbf{x}^0 + \boldsymbol{\delta}] = [x_{0.1} - \delta_1, x_{0.1} + \delta_1] \times \cdots \times [x_{0.n} - \delta_n, x_{0.n} + \delta_n]$, where $\boldsymbol{\delta} = [\delta_1 \ldots \delta_n]^{\mathrm{T}}$. We assume that the base set $X_B = \{\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N\} \subset X_R$ is available, such that the fine model response is known at all points $\mathbf{x}^j$, $j = 1, 2, \ldots, N$. In general, we do not assume any particular location of these base points.

We want to enhance the coarse model $\mathbf{R}_c$ and create a space-mapping-surrogate model $\mathbf{R}_s$ using auxiliary mappings with parameters determined so that $\mathbf{R}_s$ matches the fine model as well as possible at all base points. Because the coarse model is assumed to be physics-based, i.e. describes the same phenomenon as the fine model, we hope that the surrogate model will retain a good match with the fine model over the whole region of interest.

The standard space-mapping model (SM-Standard) is defined as [20]

$$\mathbf{R}_s(\mathbf{x}) = \bar{\mathbf{R}}_s(\mathbf{x}, \mathbf{p}) \tag{1}$$

where the space-mapping parameters $\mathbf{p}$ are obtained using the parameter extraction process

$$\mathbf{p} = \arg\min_r \sum_{k=1}^{N} \|\mathbf{R}_f(\mathbf{x}^k) - \bar{\mathbf{R}}_s(\mathbf{x}^k, \mathbf{r})\| \tag{2}$$

while $\bar{\mathbf{R}}_s$ is a generic space-mapping model, i.e. the coarse model composed with some suitable mappings. The model often used in practice has the form

$$\bar{\mathbf{R}}_s(\mathbf{x}, \mathbf{p}) = \bar{\mathbf{R}}_s(\mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{c}) = \mathbf{A} \cdot \mathbf{R}_c(\mathbf{B} \cdot \mathbf{x} + \mathbf{c}) + \mathbf{d} \tag{3}$$

where $\mathbf{A} = \mathrm{diag}\{a_1, \ldots, a_m\}$, $\mathbf{B}$ is an $n \times n$ matrix, $\mathbf{c}$ is an $n \times 1$ vector, and $\mathbf{d}$ is an $m \times 1$ vector.

In many cases, both fine and coarse models have parameters that are normally fixed and not used in the optimization process (so-called preassigned parameters). These parameters can be used as additional degrees of freedom in the coarse model and adjusted in order to obtain a better match with the fine model, which leads us to implicit space mapping [2]. Let us denote the coarse model exploiting the preassigned parameters $\mathbf{x}_p$ as $\mathbf{R}_c(\mathbf{x}, \mathbf{x}_p)$. The surrogate (3) enhanced by implicit space mapping could take the following form:

$$\bar{\mathbf{R}}_s(\mathbf{x}, \mathbf{p}) = \bar{\mathbf{R}}_s(\mathbf{x}, \mathbf{A}, \mathbf{B}, \mathbf{c}, \mathbf{B}_p, \mathbf{c}_p) = \mathbf{A} \cdot \mathbf{R}_c(\mathbf{B} \cdot \mathbf{x} + \mathbf{c}, \mathbf{B}_p \cdot \mathbf{x} + \mathbf{c}_p) \tag{4}$$

where $\mathbf{A}$, $\mathbf{B}$ and $\mathbf{c}$ are as in (3), while $\mathbf{B}_p$ is an $n_p \times n$ matrix, and $\mathbf{c}_p$ is an $n_p \times 1$ vector, where n_p is the number of preassigned parameters. Here, we use a generalized implicit space mapping [3] in which preassigned parameters are dependent on design variables in order to increase the flexibility of the surrogate model.

The standard space-mapping-surrogate model is very simple and fast, because once the space-mapping parameters are established, model evaluation cost is roughly the same as the evaluation cost of the coarse model, which is assumed to be much cheaper than the fine model. A limitation of this model is that linear mappings such as (3) or (4) may not be able to provide sufficient accuracy. Also, (3) or (4) may only provide a limited modification of the range of the coarse model, and this modification is basically independent of the design variables. Finally, because of the finite number of parameters that are extracted in one shot for the whole region of interest, the surrogate is, in fact, a regression model. The consequence is that the modeling error might not decrease below certain, problem dependent, non-zero limits even if the number of base points goes to infinity (cf. [25]).

DOI: 10.1002/jnm

3. SPACE MAPPING WITH FUNCTION APPROXIMATION LAYER

The limitations of the standard space mapping can be alleviated by using a function approximation layer on top of the standard model. Let us define an enhanced space-mapping-surrogate model as

$$\hat{\mathbf{R}}_s(\mathbf{x}) = \mathbf{R}_s(\mathbf{x}) + \tilde{\mathbf{R}}_s(\mathbf{x}) \tag{5}$$

where $\mathbf{R}_s$ is a standard space-mapping-surrogate model, while $\tilde{\mathbf{R}}_s$ is a function approximation model. We can now consider $\mathbf{R}_s$ as a trend function and $\tilde{\mathbf{R}}_s$ as an output space-mapping term that models the residuals between the fine model and $\mathbf{R}_s$ at all base points. This technique has the following advantages: (i) relatively good modeling accuracy can be obtained using a limited amount of fine model data because of exploiting the concept of space mapping and the underlying physics-based coarse model, (ii) the resulting surrogate is computationally as cheap as the coarse model because the function approximation layer is typically implemented using analytical formulas, (iii) it is possible to take advantage of any amount of fine model data available, so that modeling accuracy can be as good as required provided that the base set is sufficiently 'dense'.

Several approaches exploiting (5) have been proposed so far [26–29]. It has been demonstrated that the modeling accuracy of the model (5) is better than the accuracy of the standard space-mapping surrogate, and, at the same time, better than the accuracy of the function approximation model used alone, provided that the same amount of fine model data was used to set up the model in each case. This is, of course, because of features (i) and (iii) mentioned above.

Here, we focus on two approaches: space mapping combined with fuzzy systems [27] and space mapping combined with support vector regression (SVR) [29]. In the following subsection we briefly describe both concepts, while in Section 4 we present a numerical comparison of the techniques.

3.1. Function approximation layer implemented with fuzzy systems

In [27], the function approximation layer $\tilde{\mathbf{R}}_s$ has been realized using fuzzy systems. Fuzzy systems are commonly used in machine control [33] where the expert knowledge and a set of sampled input–output (state–control) pairs recorded from successful control are translated into the set of 'IF–THEN' rules that state in what situations which actions should be taken [34]. Because of the incomplete and qualitative character of such information, it is represented using a fuzzy set theory [35], where given piece of information (element) belongs to a given (fuzzy) subset of an input space with a certain degree, according to so-called membership function [36]. The process of converting a crisp input value to a fuzzy value is called 'fuzzification'. Given the specific input state, the 'IF–THEN' rules that apply are invoked, using the membership functions and truth-values obtained from the inputs, to determine the result of the rule. This result in turn will be mapped into a membership function and truth value controlling the output variable. These results are combined to give a specific answer, using a procedure known as 'defuzzification'. The 'centroid' defuzzification method is very popular, in which the 'center of mass' of the result provides the crisp output value.

Fuzzy systems can also be used as universal function approximators [34]. In particular, given a set of numerical data pairs, it is possible to obtain a fuzzy-rule-based mapping from the input space (here, design variables) to the output space (here, surrogate model response). The

DOI: 10.1002/jnm

mapping is realized by dividing the input and output spaces into fuzzy regions, generating fuzzy rules from given desired input–output data pairs, assigning a degree to each generated rule and forming a combined fuzzy rule base, and, finally, performing defuzzification [34]. It can be shown [34] that under certain conditions, such a mapping is capable of approximating any real continuous function over the compact domain to arbitrary accuracy.

Fuzzy systems have been successfully used in the microwave area by other authors (e.g. [37–39]), but never in connection with space mapping. In the following description a fuzzy system with triangle membership functions and centroid defuzzification [34] is used. Other membership functions can also be applied resulting in similar performance (see the reference [27] for numerical comparisons).

We assume that we have data pairs $(\mathbf{x}^k, \mathbf{R}^k)$, where $\mathbf{x}^k \in X_B$ and $\mathbf{R}^k = \mathbf{R}_f(\mathbf{x}^k) - \mathbf{R}_s(\mathbf{x}_k)$, $k = 1, 2, \ldots, N$. Membership functions for the ith variable are defined as shown in Figure 1. Each interval $[x_{0.i} - \delta_{0.i}, x_{0.i} + \delta_{0.i}]$, $i = 1, 2, \ldots, n$, is divided into K subintervals (fuzzy regions). The number K corresponds to the number of base points N and is given by the formula $K = \lfloor N^{1/n} \rfloor - 1$. In particular, if X_B consists of base points uniformly distributed in the region of interest X_R then $K+1$ is exactly the number of points of this uniform grid along any of the design variable axes. In general, K is chosen in such a way that the number of n-dimensional subintervals (and, consequently, the maximum number of rules) is not larger than the number of base points. Division of $[x_{0.i} - \delta_{0.i}, x_{0.i} + \delta_{0.i}]$ into K subintervals creates $K+1$ values $x^{i.k}$, $k = 0, 1, \ldots, K$. In the case of a uniform base set, points $\mathbf{x^q} = [x^{1.q_1} \ldots x^{n.q_n}]^{\mathrm{T}}$, $\mathbf{q} \in \{0, 1, \ldots, K\}^n$ coincide with the base points. Value $x^{i.k}$ corresponds to the fuzzy region $[x^{i.k-1}, x^{i.k+1}]$ for $k = 1, \ldots, K-1$ ($[x^{i.0}, x^{i.1}]$ for $k = 0$, and $[x^{i.k-1}, x^{i.K}]$ for $k = K$). We also use the symbol $\mathbf{x}^q$ to denote the n-dimensional fuzzy region $[x^{1.q_1} \ldots x^{n.q_n}]^{\mathrm{T}}$. For any given x, the value of membership function $m_{i.k}(x)$ determines the degree of x in the fuzzy region $x^{i.k}$. In this paper we only use triangular membership functions; one vertex lies at the center of the region and has membership value unity; the other two vertices lie at the centers of the two neighboring regions, respectively, and have membership values equal to zero.

Having defined the membership functions we need to generate the fuzzy rules from given data pairs. We use if–then rules of the form IF $\mathbf{x}^k$ is in $\mathbf{x}^q$ THEN $\mathbf{y} = \mathbf{R}^k$, where $\mathbf{y}$ is the response of the rule. At the level of vector components it means

$$\text{IF } x_{k.1} \text{ is in } x^{1.q_1} \text{ AND } x_{k.2} \text{ is in } x^{2.q_2} \text{ AND} \ldots \text{AND } x_{k.n} \text{ is in } x^{n.q_n} \text{ THEN } \mathbf{y} = \mathbf{R}^k \qquad (6)$$

where $x_{k.i}$, $i = 1, \ldots, n$ are the components of vector $\mathbf{x}^k$. In general, it may happen that there are some conflicting rules, i.e. rules that have the same IF part but a different THEN part.

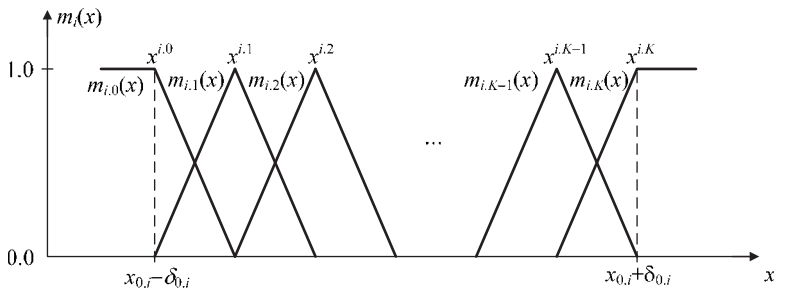

Figure 1. Division of the input interval $[x_{0.i} - \delta_{0.i}, x_{0.i} + \delta_{0.i}]$ into fuzzy regions and the corresponding membership functions.

DOI: 10.1002/jnm

We resolve such conflicts by assigning a degree to each rule and accepting only the rule from a conflict group that has a maximum degree. A degree is assigned to a rule in the following way. For the rule 'IF $\mathbf{x}_{k.1}$ is in $\mathbf{x}^{1.q_1}$ AND $\mathbf{x}_{k.2}$ is in $\mathbf{x}^{2.q_2}$ AND…AND $\mathbf{x}_{k.n}$ is in $\mathbf{x}^{n.q_n}$ THEN $\mathbf{y} = \mathbf{R}^k$', the degree of this rule, denoted by $D(\mathbf{x}^k)$ is defined as

$$D(\mathbf{x}^k, \mathbf{x}^{\mathbf{q}}) = \prod_{i=1}^{n} m_{i.q_i}(x_{k.i}) \tag{7}$$

Having resolved the conflicts we have a set of non-conflicting rules, which we denote as R_i, $i = 1, 2, \ldots, L$. We denote by $\tilde{\mathbf{R}}_s : X_R \to R^m$ the output of our fuzzy system, which is determined using a centroid defuzzification

$$\tilde{\mathbf{R}}_s(\mathbf{x}) = \frac{\sum_{i=1}^{L} D(\mathbf{x}, \mathbf{x}^i)\mathbf{y}_i}{\sum_{i=1}^{L} D(\mathbf{x}, \mathbf{x}^i)} \tag{8}$$

where $\mathbf{x}^i$ is an n–dimensional fuzzy region corresponding to the ith rule, and $\mathbf{y}_i$ is the output of the ith rule.

It should be noted that implementation of the fuzzy system described above is very simple and once the fuzzy rules are established, the evaluation of $\tilde{\mathbf{R}}_s(\mathbf{x})$ is very fast, so that the surrogate model (5) using fuzzy systems (we will refer to it as SM-Fuzzy) is virtually as cheap as the coarse model.

3.2. Function approximation layer implemented with SVR

SVR [40] is a relatively novel technique, which is characterized by good generalization capability [41] and easy training through quadratic programming resulting in a global optimum for the model parameters [42]. SVR is a variant of the SVM methodology developed by Vapnik [43], which was originally applied to solve the classification problems. SVM exploits the structural risk minimization (SRM) principle, which has been shown to be superior [40] to traditional empirical risk minimization (ERM) principle, employed by the conventional methods used in the empirical data modeling, e.g. neural networks. SRM minimizes an upper bound on the expected risk [40], as opposed to ERM that minimizes the error on the training data, which is the difference that equips SVM with a greater ability to generalize [40].

SVR is currently gaining popularity in the electrical engineering area (e.g. [44–51]). In the reference [29], the application of SVR for enhancing the space-mapping models has been presented for the first time.

As before, let $\mathbf{R}^k = \mathbf{R}_f(\mathbf{x}^k) - \mathbf{R}_s(\mathbf{x}^k)$, $k = 1, 2, \ldots, N$, denote the differences between the fine model and the standard space-mapping model (1). We want to use SVR to approximate the residuals $\mathbf{R}^k$ at all base points. We shall also use the notation $\mathbf{R}^k = [R_1^k R_2^k \ldots R_m^k]^T$ to denote the components of vector $\mathbf{R}^k$. In the case of linear regression, we want to approximate a given set of data, in our case, the data pairs $D_j = \{(\mathbf{x}^1, R_j^1), \ldots, (\mathbf{x}^N, R_j^N)\}$, $j = 1, 2, \ldots, m$, by a linear function $f_j(\mathbf{x}) = \mathbf{w}_j^T\mathbf{x} + b_j$. The optimal regression function is given by the minimum of the functional [42]

$$\Phi_j(\mathbf{w}, \xi) = \frac{1}{2}\|\mathbf{w}_j\|^2 + C_j \sum_{i=1}^{N} (\xi_{j.i}^+ + \xi_{j.i}^-) \tag{9}$$

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 SVR is the so-called

DOI: 10.1002/jnm

ε-insensitive loss function

$$L_\varepsilon(y) = \begin{cases} 0 & \text{for } |f_j(\mathbf{x}) - y| < \varepsilon \\ |f_j(\mathbf{x}) - y| & \text{otherwise} \end{cases} \tag{10}$$

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 [39].

Here, we use non-linear regression employing the kernel approach, in which the linear function $\mathbf{w}_j^T\mathbf{x}+b_j$ is replaced by the non-linear function $\Sigma_i\gamma_{j.i}K(\mathbf{x}^k,\mathbf{x})+b_j$, where K is a kernel function. Thus, the SVR term used to enhance the standard SM is defined as [29]

$$\tilde{\mathbf{R}}_s = \begin{bmatrix} \sum_{i=1}^N \gamma_{1.i}K(\mathbf{x}^i,\mathbf{x})+b_1 \\ \vdots \\ \sum_{i=1}^N \gamma_{m.i}K(\mathbf{x}^i,\mathbf{x})+b_m \end{bmatrix} \tag{11}$$

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

$$K(\mathbf{x},\mathbf{y}) = \exp\left(-\frac{\|\mathbf{x}-\mathbf{y}\|^2}{2c^2\lambda^2}\right), \quad c>0 \tag{12}$$

where $\lambda=\lambda(\boldsymbol{\delta}, N)$—used here as an normalization factor—is a so-called characteristic distance of the base set defined as [24]

$$\lambda(\boldsymbol{\delta}, N) = \frac{2}{nN^{1/n}}\sum_{i=1}^n \delta_i \tag{13}$$

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 [14] and exponential grid search [44].

Similarly as in the case of fuzzy systems, the cost of evaluating the SVR model (11) is low and, therefore, SVR does not degrade the computational efficiency of the standard space-mapping model when both are utilized in the enhanced surrogate (5). It should be noted though, that implementation of SVR is more complicated than implementation of fuzzy systems because it involves solving a constrained non-linear optimization problem as well as the tuning of certain control parameters. On the other hand, the SVR function is smooth, which may not be the case for modeling with fuzzy systems [27]. We will refer to the model (5) using SVR enhancement as SM-SVR.

4. COMPARISON OF MODELING METHODOLOGIES

In this section we compare different aspects of the enhanced space-mapping techniques described in Section 3. In particular, we consider modeling accuracy, robustness with respect to distribution of base points, as well as applicability for modeling problems of different dimensionality. We use two examples of microwave filters. The first one is a low-dimensional problem with only two design variables; the second example is a higher-dimensional problem with five design variables. We begin, however, with a brief discussion of some design of experiments techniques that can be used to determine the distribution of the base points.

DOI: 10.1002/jnm

4.1. Design of experiments for low- and higher-dimensional modeling problems

Before the surrogate model is established, we need to decide on the amount and location of the base points $\mathbf{x}^1, \mathbf{x}^2, \ldots, \mathbf{x}^N$, at which we evaluate the fine model, and then use this information to perform the parameter extraction (2), and, subsequently, generate the data pairs $\mathbf{R}^k = \mathbf{R}_f(\mathbf{x}^k) - \mathbf{R}_s(\mathbf{x}^k)$, $k = 1, 2, \ldots, N$, used to set up the function approximation model.

Traditionally, following the classical factorial design of experiments [52], the standard space-mapping model uses the so-called star-distribution-like base set [20], where $N = 2n+1$, with $\mathbf{x}^1 = \mathbf{x}^0$, $\mathbf{x}^j = \mathbf{x}^0 + (1)^{j-1}\delta_{\lfloor (j-1)/2 \rfloor}\mathbf{e}_{\lfloor (j-1)/2 \rfloor}$ for $j = 2, \ldots, N$, where $\mathbf{x}^0$ is the reference point (cf. Section 2), and $\mathbf{e}_j = [0 \ \ldots \ 0 \ 1 \ 0 \ \ldots \ 0]^T$ is a unit vector with 1 at jth position; δ_j is the size of the region of interest along the jth axis. Figure 2 shows the star-distribution base set for $n = 2$.

As mentioned before, the star-distribution-like base set is not sufficient if we want to improve the accuracy of the surrogate model, especially when a function approximation layer is used in combination with space mapping as described in Section 3: we need more fine model data. In this paper, we do not assume any prior knowledge of the fine model behavior within the *region of interest*, X_R. Therefore, we would like the base points to be distributed more-or-less uniformly within X_R.

One of the possibilities, frequently exploited when the number of design variables n is small, is locating the base points on a uniform grid, as shown in Figure 3 for $n = 2$ and grid density $k = 5$. Obviously, this technique gives the most uniform distribution of points possible, however, the number of base points grows very fast with both n and k, and, feasible number of points that

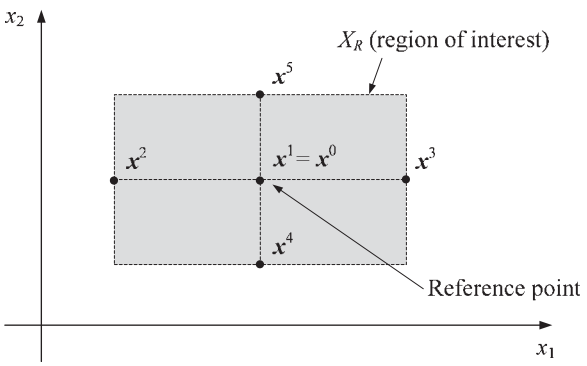

Figure 2. Region of interest and the star-distribution base set for $n = 2$.

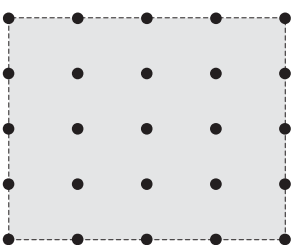

Figure 3. Uniform-grid-like base set for $n = 2$ and the grid density $k = 5$.

DOI: 10.1002/jnm

can be allocated is limited to possible values of k^n. In practice, it can be effectively applied for $n<5$.

Another technique that can be used is one of many available sampling algorithms that produce a quasi-uniform distribution of base points. In this paper, we shall use a Latin hypercube sampling [52, 53]. The basic idea of this method is to divide the range of each parameter into N intervals of equal size, which creates N^n bins in the region of interest. Next, we select N samples so that (i) each sample is randomly placed inside a bin, and (ii) for all one-dimensional projections of the N samples and bins, there will be one and only one sample in each bin [52]. Numerous improvements of this basic technique have been reported in the literature [53–56] that lead to more uniform distribution of samples. Here, we use a version described in [53]. Figure 4 shows an example of allocating 15 samples within a two-dimensional region of interest.

An attractive aspect of a Latin hypercube sampling technique is that it allows any number of base points, regardless of n. Thus, it is more suitable for modeling problems that involve a larger number of variables.

4.2. Lower-dimensional modeling problem

Consider a second-order capacitively coupled dual-behavior resonator (CCDBR) microstrip filter [57] shown in Figure 5. The design parameters are $\mathbf{x} = [L_1\ L_2\ L_3]^T$; the value of S is fixed and equal to 0.05 mm. The fine model $\mathbf{R}_f$ is simulated in FEKO [58]. The coarse model $\mathbf{R}_c$ is the circuit model implemented in Agilent ADS [59] and shown in Figure 6. The response vector consists of reflection coefficient $|S_{21}|$ in the frequency range 2–6 GHz. The reference point is $\mathbf{x}^0 = [3.448\ 4.803\ 1.036]^T$ mm, and the region size $\boldsymbol{\delta} = 0.1 \cdot \mathbf{x}^0$. The reference point corresponds, in fact, to the optimal fine model design with respect to the following specifications: $|S_{21}(\omega)| \leqslant -20$ dB for $2.0\,\text{GHz} \leqslant \omega \leqslant 3.2\,\text{GHz}$, $|S_{21}| \geqslant -3$ dB for $3.8\,\text{GHz} \leqslant \omega \leqslant 4.2\,\text{GHz}$, and $|S_{21}| \leqslant -20$ dB for $4.8\,\text{GHz} \leqslant \omega \leqslant 6.0\,\text{GHz}$.

We perform experiments using the following surrogate models: SM-Standard, SM-Fuzzy, SM-SVR. For comparison purposes we also consider models that directly approximate fine model data, in particular, models based on fuzzy systems (Fuzzy) and SVR. Table I shows details of the base sets used in our experiments. The base sets have growing numbers of points

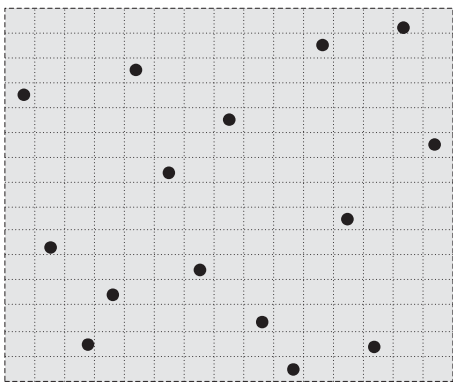

Figure 4. Latin hypercube sampling example for $n = 2$ and $N = 15$.

DOI: 10.1002/jnm

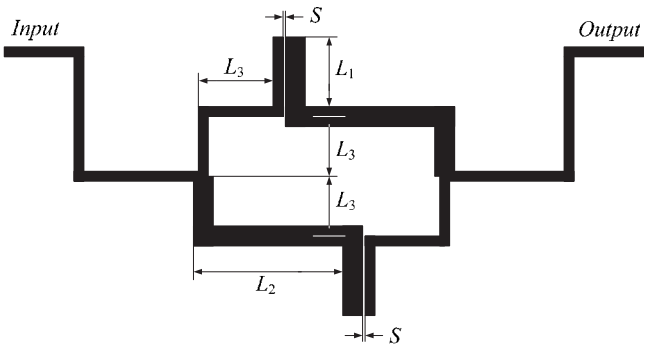

Figure 5. CCDBR filter: physical structure [57].

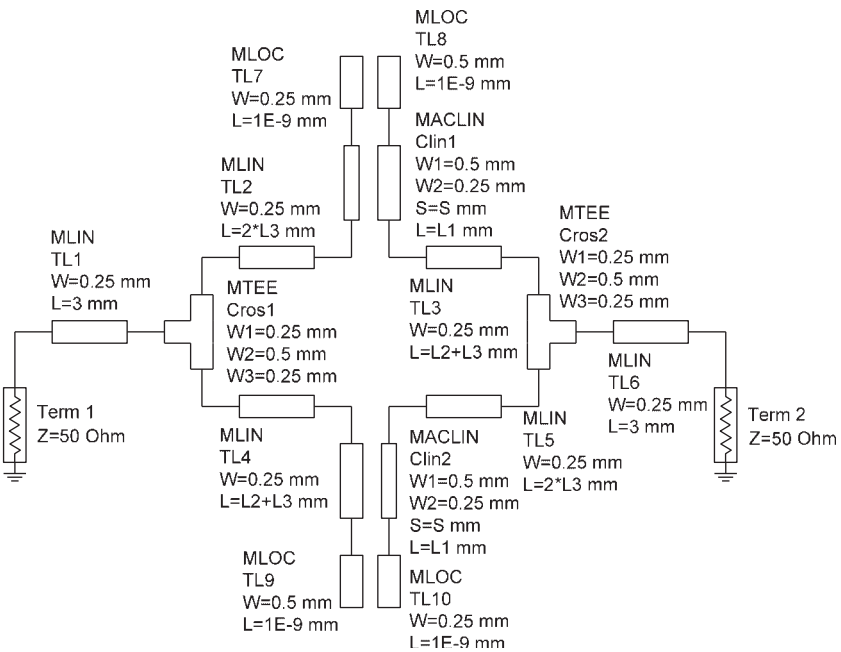

Figure 6. CCDBR filter: coarse model (Agilent ADS).

(and decreasing characteristic distances λ) in order to examine the dependence of the modeling error on the amount of fine model data used to create the model.

Accuracy was tested using 50 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.

Table II shows numerical results (error statistics) for the models with the various base sets considered. Figure 7 shows the dependence of average modeling error on the number of base points for all considered surrogate models. Figure 8 shows the error plots (the modulus of the difference between the fine model and the corresponding surrogate model response versus

DOI: 10.1002/jnm

Table I. Base set data for CCDBR filter modeling example.

Base set	Base set description	Number of base points	λ
X_{B1}	Uniform mesh of density 1	8	0.310
X_{B2}	Uniform mesh of density 2	27	0.206
X_{B3}	Uniform mesh of density 3	64	0.155
X_{B4}	Uniform mesh of density 4	125	0.124
X_{B5}	Uniform mesh of density 5	216	0.103

Table II. Modeling results for test CCDBR filter.

Model	Base set	Average error	Maximum error	Standard deviation
SM-Standard	X_{B1}	0.0981	0.1489	0.0174
SM-Fuzzy		0.0659	0.1137	0.0181
SM-SVR		0.0866	0.1386	0.0198
Fuzzy		0.4125	0.6043	0.1149
SVR		0.4107	0.6543	0.1114
SM-Standard	X_{B2}	0.0926	0.1362	0.0171
SM-Fuzzy		0.0336	0.0708	0.0118
SM-SVR		0.0356	0.0741	0.0117
Fuzzy		0.1496	0.2497	0.0534
SVR		0.1481	0.2420	0.0432
SM-Standard	X_{B3}	0.0898	0.1308	0.0165
SM-Fuzzy		0.0242	0.0636	0.0107
SM-SVR		0.0225	0.0633	0.0099
Fuzzy		0.0778	0.1466	0.0250
SVR		0.0936	0.2084	0.0432
SM-Standard	X_{B4}	0.0888	0.1304	0.0162
SM-Fuzzy		0.0211	0.0664	0.0127
SM-SVR		0.0245	0.0676	0.0148
Fuzzy		0.0506	0.0970	0.0195
SVR		0.0612	0.1725	0.0376
SM-Standard	X_{B5}	0.0887	0.1306	0.0164
SM-Fuzzy		0.0144	0.0473	0.0105
SM-SVR		0.0148	0.0486	0.0100
Fuzzy		0.0333	0.0598	0.0124
SVR		0.0485	0.1320	0.0286

Verification For 50 random test points.

frequency) for two base sets: X_{B2} and X_{B5}. Finally, Figure 9 shows actual fine and surrogate model responses at two selected test points for the models obtained with base set X_{B2}. More specifically, we show a test point representing the average modeling performance as well as a worst-case.

The results show that, as expected [25], the performance of the SM-Standard model is virtually independent of the number of base points. The performance of the SM-Fuzzy and SM-SVR models is almost the same, and better than for SM-Standard model for all base sets considered. Modeling accuracy improves with growing number of base points. We can observe the same pattern for function approximation surrogate models Fuzzy and SVR. However, both the Fuzzy and SVR models are substantially worse than SM-Fuzzy and SM-SVR and one can estimate by extrapolation that the performance of all these models might be the same

DOI: 10.1002/jnm

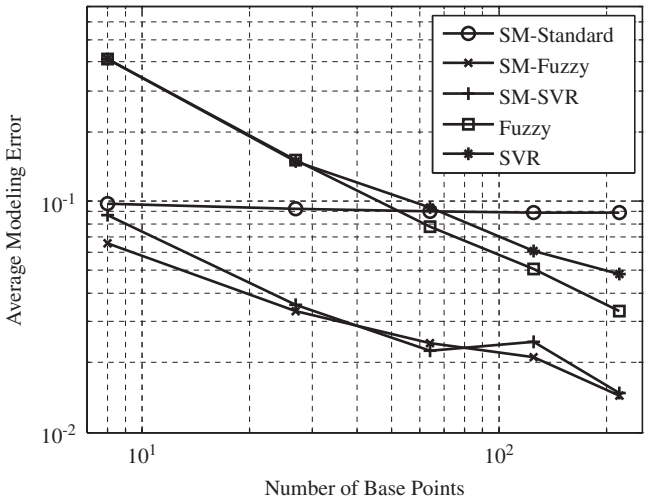

Figure 7. Average modeling error versus number of base points.

for a number of base points equal at least 10^3 or 10^4, which is rather impractical. Note also that the Fuzzy and SVR models are not better than SM-Standard until the number of base points is at least 10^2. Figure 9 shows that function approximation surrogate models are rather poor in modeling the filter response in the pass-band, which is the most important part of the frequency spectrum.

If the number of base points is severely limited, SM-Standard is probably the best choice because of its simplicity and performance comparable with more involved models using a function approximation layer.

For all other cases, where the number of base points is a few dozen or more, the best choice is either SM-Fuzzy or SM-SVR.

4.3. Higher-dimensional modeling problem

Consider the 3rd-order Chebyshev band-pass filter [60] shown in Figure 10. The design parameters are $\mathbf{x} = [L_1\ L_2\ S_1\ S_2\ W_1\ W_2]^{\mathrm{T}}$ mm. The fine model $\mathbf{R}_{\mathrm{f}}$ is simulated in Sonnet **em** [61] with a fine grid of 0.2 mm × 0.02 mm. The coarse model $\mathbf{R}_{\mathrm{c}}$ is the circuit model implemented in Agilent ADS [59] (Figure 11). The response vector consists of reflection coefficient $|S_{21}|$ in the frequency range 1–3 GHz. The reference point is $\mathbf{x}^0 = [15\ 15\ 0.4\ 0.8\ 0.4\ 0.4]^{\mathrm{T}}$ mm, and the region size $\boldsymbol{\delta} = [2\ 2\ 0.1\ 0.2\ 0.1\ 0.1]^{\mathrm{T}}$ mm.

As for the previous example, we perform experiments using the space-mapping models SM-Standard, SM-Fuzzy, SM-SVR, and the two function approximation models: Fuzzy and SVR. Table III shows details of the base sets used in our experiments.

Accuracy was tested using 50 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.

Table IV shows numerical results (error statistics) for the models with the various base sets considered. Figure 12 shows the dependence of average modeling error on the number of base points for all considered surrogate models. Figure 13 shows the error plots (the modulus of the

DOI: 10.1002/jnm

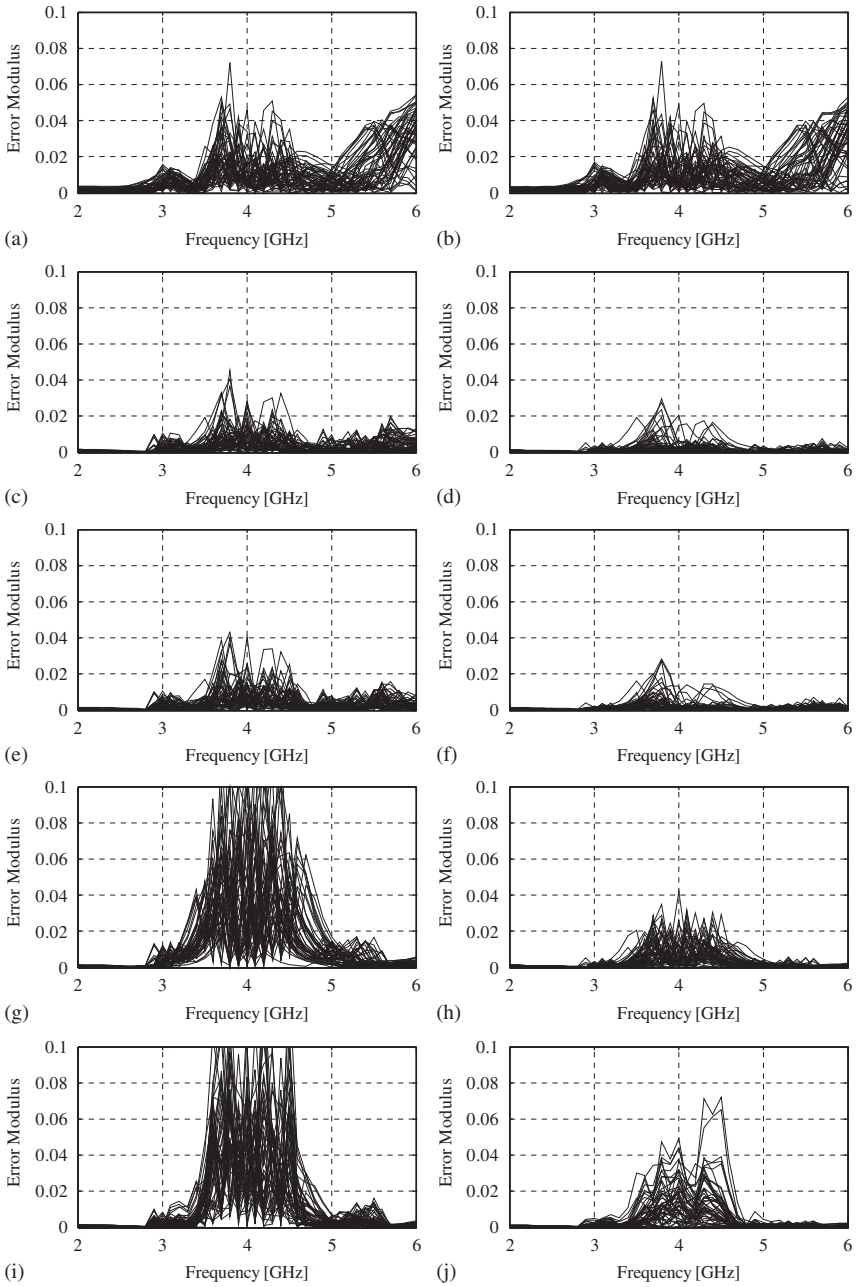

Figure 8. Error plots for the base set X_{B2} and X_{B5} for SM-Standard (a, b), SM-Fuzzy (c, d), SM-SVR (e, f), Fuzzy (g, h), and SVR (i, j); 50 random test points used in each case.

DOI: 10.1002/jnm

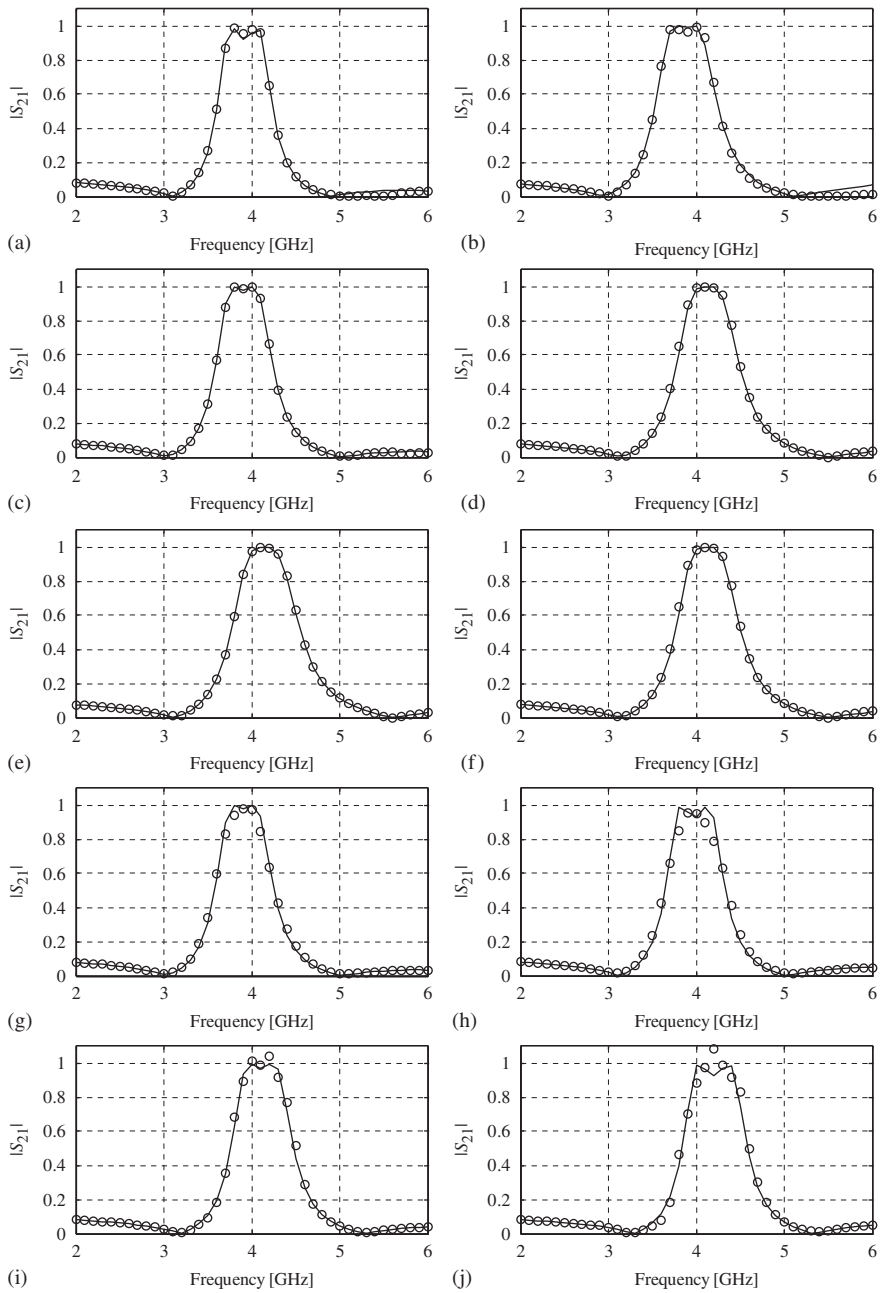

Figure 9. Fine model (solid line) and surrogate model (circles) responses at a typical and a worst-case test point for SM-Standard (a, b), SM-Fuzzy (c, d), SM-SVR (e, f), Fuzzy (g, h), and SVR (i, j), for the base set X_{B2}.

DOI: 10.1002/jnm

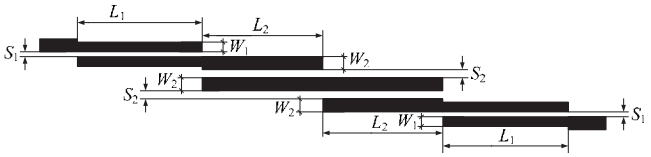

Figure 10. Third-order Chebyshev band-pass filter: physical structure [60].

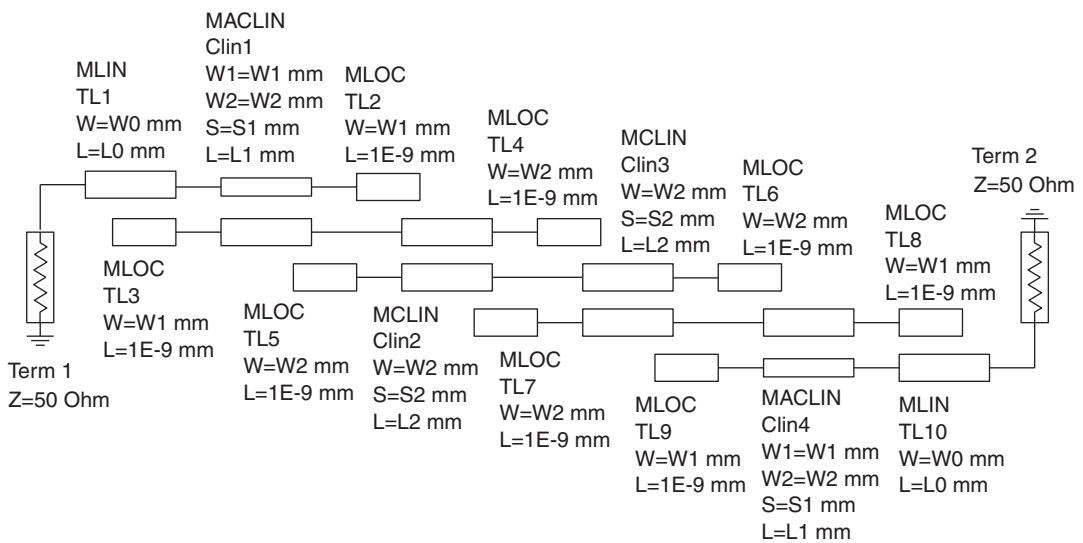

Figure 11. Third-order Chebyshev band-pass filter: coarse model (Agilent ADS).

Table III. Base set data for Chebyshev filter modeling example.

Base set	Base set description	Number of base points	λ
X_{B1}	Latin hypercube sampling	15	0.96
X_{B2}	Latin hypercube sampling	60	0.76
X_{B3}	Latin hypercube sampling	200	0.62

difference between the fine model and the corresponding surrogate model response versus frequency) for two base sets: X_{B1} and X_{B3}. Finally, Figure 14 shows actual fine and surrogate model responses at two test points (for the models obtained with the base set X_{B3}): a test point representing the average modeling performance as well as a test point representing the worst-case (largest modeling error).

As in the previous example, the performance of the SM-Standard model is virtually independent of the number of base points. The performance of SM-Fuzzy and SM-SVR models is better than for the SM-Standard model for all base sets considered except X_{B1}, and modeling accuracy improves with growing the number of base points. It should be emphasized that the good accuracy of SM-Fuzzy and SM-SVR model is a result of combining the respective function

DOI: 10.1002/jnm

Table IV. Modeling results for test Chebyshev filter.

Model	Base set	Average error	Maximum error	Standard deviation
SM-Standard	X_{B1}	0.2083	0.4718	0.074
SM-Fuzzy		0.2172	0.3957	0.071
SM-SVR		0.2021	0.4695	0.065
Fuzzy		1.2306	2.9121	0.705
SVR		0.6561	0.9697	0.146
SM-Standard	X_{B2}	0.2018	0.4681	0.073
SM-Fuzzy		0.1879	0.3370	0.065
SM-SVR		0.1603	0.3794	0.066
Fuzzy		0.8349	2.3477	0.345
SVR		0.3515	0.8901	0.127
SM-Standard	X_{B3}	0.1985	0.4190	0.072
SM-Fuzzy		0.1740	0.3281	0.065
SM-SVR		0.1086	0.2893	0.044
Fuzzy		0.8319	0.2809	0.588
SVR		0.2170	0.4473	0.075

Verification for 50 random test points.

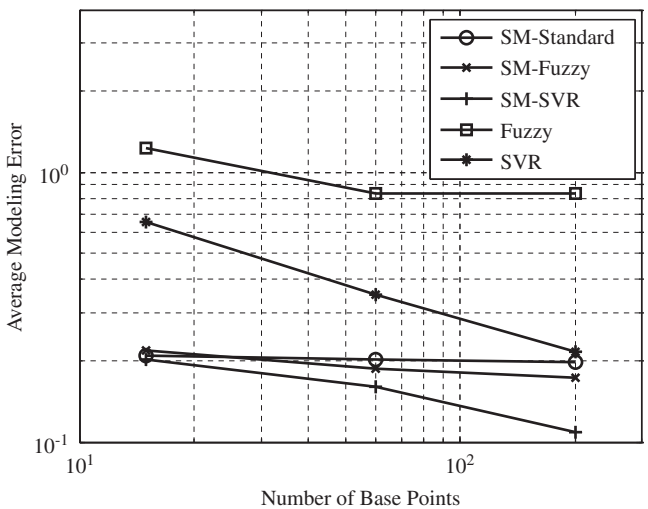

Figure 12. Average modeling error versus number of base points.

approximation techniques with space mapping: performance of the function approximators (both Fuzzy and SVR) acting as stand-alone models is rather poor. In general, it is expected that any kind of universal function approximation method will enhance performance of the standard space-mapping model (see, e.g. [26] where the combination of SM with radial basis function interpolation is described).

Note also that SM-SVR substantially outperforms SM-Fuzzy. The improvement of modeling accuracy for SM-Fuzzy is very slow with the growth of the number of base points. The reason is that for this example, we cannot afford a uniform-grid-like base set and we use a Latin hypercube sampling. While it does not change the performance of the SM-SVR model (as well

DOI: 10.1002/jnm

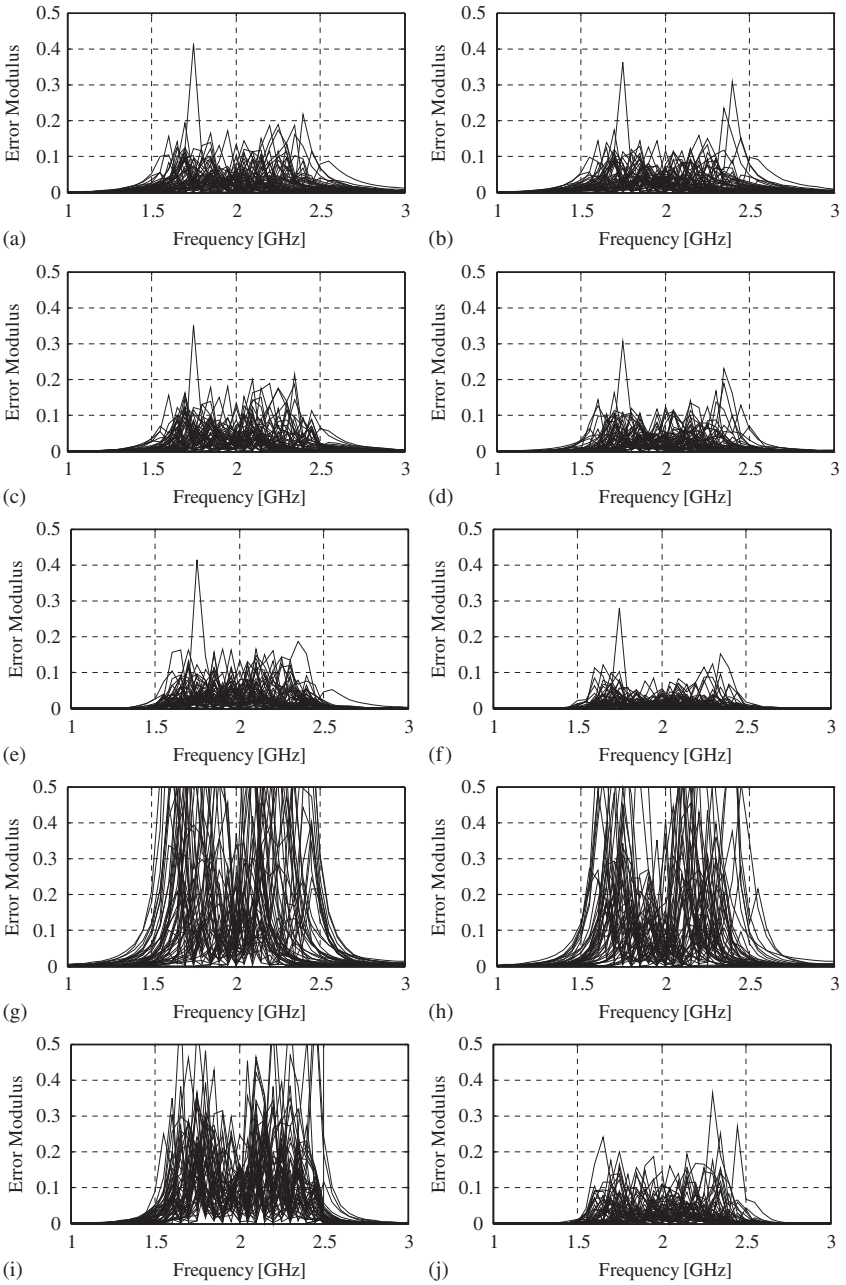

Figure 13. Error plots for the base set X_{B1} and X_{B3} for SM-Standard (a, b), SM-Fuzzy (c, d), SM-SVR (e, f), Fuzzy (g, h), and SVR (i, j); 50 random test points used in each case.

DOI: 10.1002/jnm

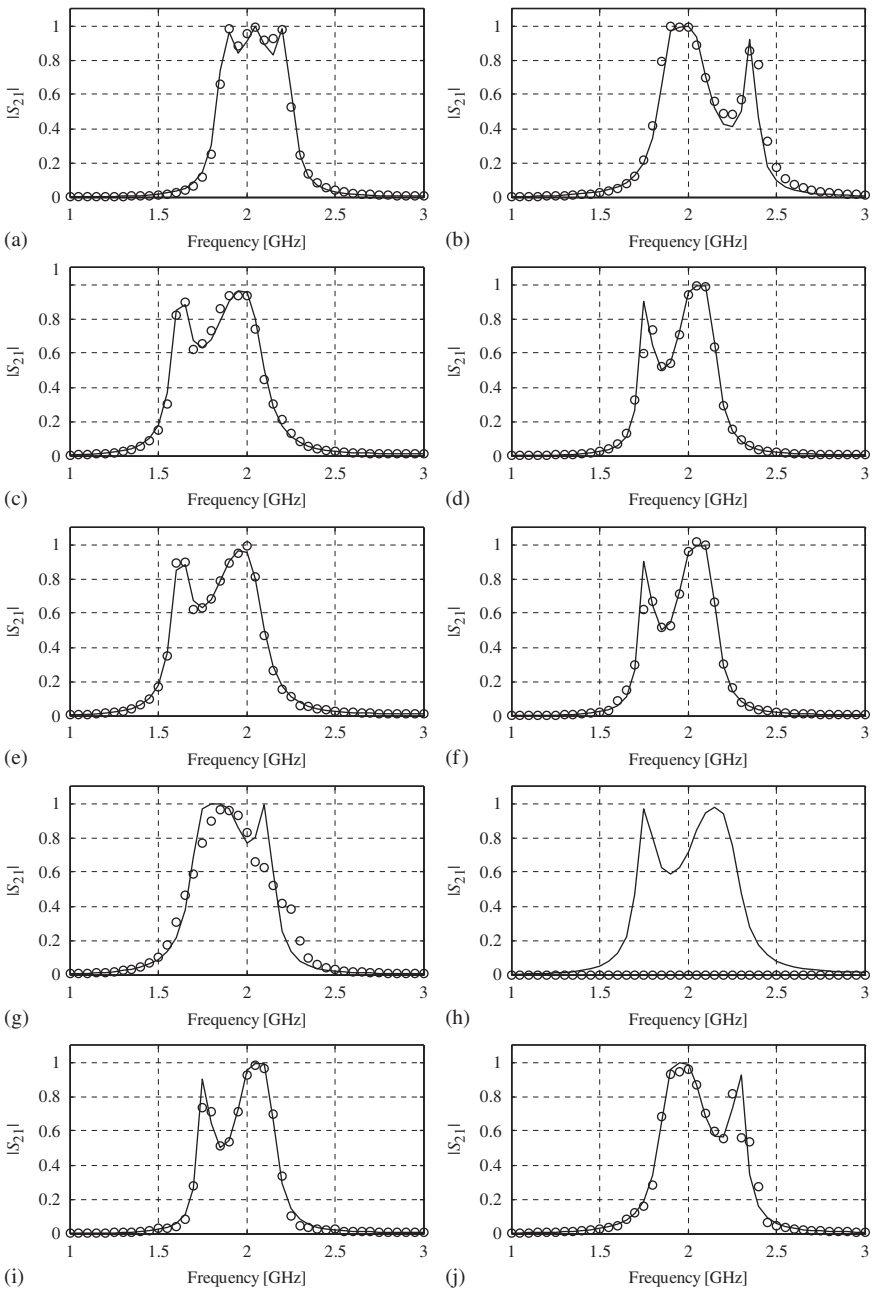

Figure 14. Fine model (solid line) and surrogate model (circles) responses at a typical and a worst-case test point for SM-Standard (a, b), SM-Fuzzy (c, d), SM-SVR (e, f), Fuzzy (g, h), and SVR (i, j), for the base set X_{B3}.

DOI: 10.1002/jnm

as the SVR model), it seriously affects the SM-Fuzzy model (as well as the Fuzzy model), because fuzzy systems require, for best performance that the base points are located at maxima of the membership functions (here, we use triangle functions). Also, given a division of the design space into fuzzy regions, it may happen that a given region has no corresponding fuzzy rule (because of the lack of a data point in the region). This is especially visible for the Fuzzy model where it may happen that it gives zero response for certain values of the design variable vector (see Figure 14(h)). We can also observe that the accuracy of function approximation models is much worse than the accuracy of the space-mapping models and we would need at least 10^3 or 10^4 base points in order to obtain comparable accuracy of both types of models.

Similarly as in the previous example, the SM-Standard model is probably the best choice if the number of base points is limited.

If more fine model data is available, we recommend the SM-SVR model but not the SM-Fuzzy model for the reasons mentioned before.

We should also mention that the SM-SVR model has another advantage with respect to the SM-Fuzzy model, which is that the accuracy of the SM-SVR model can be assessed using a cross-validation method so that no 'external' test points are necessary [14]. This cannot be done for the SM-Fuzzy model because of the fact that the fuzzy rules are strictly localized to the neighborhood of the respective base points, which makes a cross-validation of this kind of model unreliable.

5. CONCLUSION

A review of recent advances in space-mapping modeling of microwave devices is presented. A detailed comparison of space-mapping models enhanced by a function approximation layer involving fuzzy systems and SVR is given. Both techniques substantially improve the modeling accuracy when compared with standard space mapping. The observed improvement is a result of combining the respective approximation techniques with space mapping, which is confirmed by a poor performance of both fuzzy systems and SVR acting as stand-alone surrogate models.

It is shown that the performance of the space-mapping model exploiting fuzzy systems is heavily dependent on the structure of the base set used to establish a surrogate model, which is not the case for the model using SVR as a function approximation layer. Therefore, both models are equally good for lower-dimensional problems where the uniform-grid-line base set can be utilized, however, the space-mapping model enhanced by SVR is preferred for higher-dimensional problems where other design-of-experiment techniques, such as Latin hypercube sampling, must be utilized to select base points.

In any case, a standard space-mapping model is probably the best choice if the amount of fine model data available to set up the surrogate model is severely limited.

ACKNOWLEDGEMENTS

This work was supported in part by the Reykjavik University Development Fund under Grant T09009 and the Natural Sciences and Engineering Research Council of Canada under Grants RGPIN7239-06 and STPGP336760-06. The authors thank Sonnet Software, Inc., Syracuse, NY, for **em**^{TM} and Agilent Technologies, Santa Rosa, CA, for making ADS available.

DOI: 10.1002/jnm

REFERENCES

1. Bandler JW, Biernacki RM, Chen SH, Grobelny PA, Hemmers RH. Space mapping technique for electromagnetic optimization. *IEEE Transactions on Microwave Theory and Techniques* 1994; **4**(12):536–544.
2. Bandler JW, Cheng QS, Dakroury SA, Mohamed AS, Bakr MH, Madsen K, Sondergaard J. Space mapping: the state of the art. *IEEE Transactions on Microwave Theory and Techniques* 2004; **52**(1):337–361.
3. Koziel S, Bandler JW, Madsen K. A space mapping framework for engineering optimization: theory and implementation. *IEEE Transactions on Microwave Theory and Techniques* 2006; **54**(10):3721–3730.
4. Echeverria D, Hemker PW. Space mapping and defect correction. *CMAM The International Mathematical Journal Computational Methods in Applied Mathematics* 2005; **5**(20):107–136.
5. Choi H-S, Kim DH, Park IH, Hahn SY. A new design technique of magnetic systems using space mapping algorithm. *IEEE Transactions on Magnetics* 2001; **37**(5):3627–3630.
6. Ismail MA, Smith D, Panariello A, Wang Y, Yu M. EM-based design of large-scale dielectric-resonator filters and multiplexers by space mapping. *IEEE Transactions on Microwave Theory and Techniques* 2004; **52**(1):386–392.
7. Amari S, LeDrew C, Menzel W. Space-mapping optimization of planar coupled-resonator microwave filters. *IEEE Transactions on Microwave Theory and Techniques* 2006; **54**(5):2153–2159.
8. Rayas-Sánchez JE, Gutiérrez-Ayala V. EM-based Monte Carlo analysis and yield prediction of microwave circuits using linear-input neural-output space mapping. *IEEE Transactions on Microwave Theory and Techniques* 2006; **54**(12):4528–4537.
9. Zhang XJ, Fang DG. Using circuit model from layout-level synthesis as coarse model in space mapping and its application in modelling low-temperature ceramic cofired radio frequency circuits. *Microwaves, Antennas and Propagation, IET* 2007; **1**(4):881–886.
10. Crevecoeur G, Dupre L, Van de Walle R. Space mapping optimization of the magnetic circuit of electrical machines including local material degradation. *IEEE Transactions on Magnetics* 2007; **43**(6):2609–2611.
11. Burrascano P, Dionigi M, Fancelli C, Mongiardo M. A neural network model for CAD and optimization of microwave filters. *IEEE MTT-S International Microwave Symposium Digest*, Baltimore, MD, 1998; 13–16.
12. Peik SF, Mansour RR, Chow YL. Multidimensional Cauchy method and adaptive sampling for an accurate microwave circuit modeling. *IEEE Transactions on Microwave Theory and Techniques* 1998; **46**(12):2364–2371.
13. Simpson TW, Peplinski J, Koch PN, Allen JK. Metamodels for computer-based engineering design: survey and recommendations. *Engineering with Computers* 2001; **17**(2):129–150.
14. Queipo NV, Haftka RT, Shyy W, Goel T, Vaidynathan R, Tucker PK. Surrogate-based analysis and optimization. *Progress in Aerospace Sciences* 2005; **41**(1):1–28.
15. Mullur AA, Messac A. Metamodeling using extended radial basis functions: a comparative approach. *Engineering with Computers* 2006; **21**(3):203–217.
16. Simpson TW, Maurey TM, Korte JJ, Mistree F. Kriging models for global approximation in simulation-based multidisciplinary design optimization. *AIAA Journal* 2001; **39**(12):2233–2241.
17. Powell MJD. Radial basis functions for multivariate interpolation: a review. In *Algorithms for Approximation*, Mason JC, Cox MG (eds). Clarendon Press: Oxford, 1987.
18. Buhmann MD, Ablowitz MJ. *Radial Basis Functions*: *Theory and Implementations*. Cambridge University Press: Cambridge, 2003.
19. Bandler JW, Georgieva N, Ismail MA, Rayas-Sánchez JE, Zhang QJ. A generalized space mapping tableau approach to device modeling. *IEEE Transactions on Microwave Theory and Techniques* 2001; **49**(1):67–79.
20. Koziel S, Bandler JW, Mohamed AS, Madsen K. Enhanced surrogate models for statistical design exploiting space mapping technology. *IEEE MTT-S International Microwave Symposium Digest*, Long Beach, CA, June 2005; 1609–1612.
21. Bandler JW, Cheng QS, Koziel S. Simplified space mapping approach to enhancement of microwave device models. *International Journal of RF and Microwave Computer-Aided Engineering* 2006; **16**(5):518–535.
22. Rautio JC. A space mapped model of thick, tightly coupled conductors for planar electromagnetic analysis. *IEEE Microwave Magazine* 2004; **5**(3):62–72.
23. Wu KL, Zhao YJ, Wang J, Cheng MKK. An effective dynamic coarse model for optimization design of LTCC RF circuits with aggressive space mapping. *IEEE Transactions on Microwave Theory and Techniques* 2004; **52**(1):393–402.
24. Koziel S, Bandler JW. Space-mapping-based modeling utilizing parameter extraction with variable weight coefficients and a data base. *IEEE MTT-S International Microwave Symposium Digest*, San Francisco, CA, June 2006; 1763–1766.
25. Koziel S, Bandler JW, Madsen K. Theoretical justification of space-mapping-based modeling utilizing a data base and on-demand parameter extraction. *IEEE Transactions on Microwave Theory and Techniques* 2006; **54**(12):4316–4322.
26. Koziel S, Bandler JW. Microwave device modeling using space-mapping and radial basis functions. *IEEE MTT-S International Microwave Symposium Digest*, Honolulu, Hawaii, June 2007, 799–802.

DOI: 10.1002/jnm

27. Koziel S, Bandler JW. A space-mapping approach to microwave device modeling exploiting fuzzy systems. *IEEE Transactions on Microwave Theory and Techniques* 2007; **55**(12):2539–2547.
28. Koziel S, Bandler JW. Modeling of microwave devices with space mapping and radial basis functions. *International Journal of Numerical Modelling* 2008; **21**(3):187–203.
29. Koziel S, Bandler JW. Support-vector-regression-based output space-mapping for microwave device modeling. *IEEE MTT-S International Microwave Symposium Digest*, Atlanta, GA, June 2008; 615–618.
30. Zhang L, Xu JJ, Yagoub M, Ding RT, Zhang QJ. Neuro-space mapping technique for nonlinear device modeling and large signal simulation. *IEEE MTT-S International Microwave Symposium Digest*, Philadelphia, PA, June 2003, 173–176.
31. Devabhaktuni VK, Chattaraj B, Yagoub MCE, Zhang Q-J. Advanced microwave modeling framework exploiting automatic model generation, knowledge neural networks, and space mapping. *IEEE Transactions on Microwave Theory and Techniques* 2003; **51**(7):1822–1833.
32. Zhang L, Xu J, Yagoub MCE, Ding R, Zhang Q-J. Efficient analytical formulation and sensitivity analysis of neuro-space mapping for nonlinear microwave device modeling. *IEEE Transactions on Microwave Theory and Techniques* 2005; **53**(9):2752–2767.
33. Passino KM, Yurkovich S. Fuzzy Control. Addison Wesley Longman Inc.: Menlo Park, CA, USA, 1998.
34. Wang L-X, Mendel JM. Generating fuzzy rules by learning from examples. *IEEE Transctions on Systems, Man, and Cybernetics* 1992; **22**(6):1414–1427.
35. Zadeh LA. Fuzzy sets. *Information and Control* 1965; **8**(3):338–353.
36. Li YF, Lan CC. Development of fuzzy algorithms for servo systems. *IEEE Control System Magazine* 1989; **9**(3):65–72.
37. Miraftab V, Mansour RR. Computer-aided tuning of microwave filters using fuzzy logic. *IEEE Transactions on Microwave Theory and Techniques* 2002; **50**(12):2781–2788.
38. Miraftab V, Mansour RR. A robust fuzzy-logic technique for computer-aided diagnosis of microwave filters. *IEEE Transactions on Microwave Theory and Techniques* 2004; **52**(1):450–456.
39. Miraftab V, Mansour RR. EM-based microwave circuit design using fuzzy logic techniques. *IEE Proceedings Microwaves, Antennas and Propagation* 2006; **153**(6):495–501.
40. Gunn SR. Support vector machines for classification and regression. *Technical Report*, School of Electronics and Computer Science, University of Southampton, 1998.
41. Angiulli G, Cacciola M, Versaci M. Microwave devices and antennas modeling by support vector regression machines. *IEEE Transactions on Magnetics* 2007; **43**(4):1589–1592.
42. Smola AJ, Schölkopf B. A tutorial on support vector regression. *Statistics and Computing* 2004; **14**(3):199–222.
43. Vapnik VN. *The Nature of Statistical Learning Theory*. Springer: New York, 1995.
44. Ceperic V, Baric A. Modeling of analog circuits by using support vector regression machines. *Proceedings of the Eleventh International Conference on Electronics, Circuits, Systems*, Tel-Aviv, Israel, 2004, 391–394.
45. Rojo-Alvarez JL, Camps-Valls G, Martinez-Ramon M, Soria-Olivas E, Navia-Vazquez A, Figueiras-Vidal AR. Support vector machines framework for linear signal processing. *Signal Processing* 2005; **85**(12):2316–2326.
46. Yang Y, Hu SM, Chen RS. A combination of FDTD and least-squares support vector machines for analysis of microwave integrated circuits. *Microwave and Optical Technology Letters* 2005; **44**(3):296–299.
47. Meng J, Xia L. Support-vector regression model for millimeter wave transition. *International Journal of Infrared and Milimeter Waves* 2007; **28**(5):413–421.
48. Martinez-Ramon M, Christodoulou C. Support vector machines for antenna array processing and electromagnetics. In *Synthesis Lectures on Computational Electromagnetics*. Morgan & Claypool Publishers, 2006.
49. Ding M, Vemuri RI. An active learning scheme using support vector machines for analog circuit feasibility classification. *Eighteenth International Conference on VLSI Design*, Taj Bengal, India, 2005; 528–534.
50. Xia L, Xu RM, Yan B. Ltcc interconnect modeling by support vector regression. *Progress in Electromagnetics Research* 2009; **PIER 69**:67–75.
51. Xia L, Meng J, Xu R, Yan B, Guo Y. Modeling of 3-D vertical interconnect using support vector machine regression. *IEEE Microwave and Wireless Components Letters* 2006; **16**(12):639–641.
52. Giunta AA, Wojtkiewicz SF, Eldred MS. Overview of modern design of experiments methods for computational simulations. *American Institute of Aeronautics and Astronautics*, *Paper AIAA 2003-0649*, 2003.
53. Beachkofski B, Grandhi R. Improved distributed hypercube sampling. *American Institute of Aeronautics and Astronautics*, *Paper AIAA 2002-1274*, 2002.
54. Leary S, Bhaskar A, Keane A. Optimal orthogonal-array-based latin hypercubes. *Journal of Applied Statistics* 2003; **30**(5):585–598.
55. Ye KQ. Orthogonal column latin hypercubes and their application in computer experiments. *Journal of the American Statistical Association* 1998; **93**(444):1430–1439.
56. Palmer K, Tsui K-L. A minimum bias latin hypercube design. *IIE Transactions* 2001; **33**:793–808.
57. Manchec A, Quendo C, Favennec J-F, Rius E, Person C. Synthesis of capacitive-coupled dual-behavior resonator (CCDBR) filters. *IEEE Transactions on Microwave Theory and Techniques* 2006; **54**(6):2346–2355.

DOI: 10.1002/jnm

58. FEKO® *User's Manual*, Suite 5.4, EM Software & Systems-S.A. (Pty) Ltd, 32 Techno Lane, Technopark, Stellenbosch, 7600, South Africa, http://www.feko.info, 2008.
59. Agilent ADS, Version 2008, Agilent Technologies, 1400 Fountaingrove Parkway, Santa Rosa, CA 95403-1799, 2008.
60. Kuo JT, Chen SP, Jiang M. Parallel-coupled microstrip filters with over-coupled end stages for suppression of spurious responses. *IEEE Microwave and Wireless Components Letters* 2003; **13**(10):440–442.
61. **em**™ Version 11.53, Sonnet Software, Inc., 100 Elwood Davis Road, North Syracuse, NY 13212, USA, 2008.

AUTHORS' BIOGRAPHIES

Slawomir Koziel received the MSc and PhD degrees in electronic engineering from Gdansk University of Technology, Poland, in 1995 and 2000, respectively. He also received the MSc degrees in theoretical physics and in mathematics, in 2000 and 2002, respectively, as well as PhD in mathematics in 2003, from the University of Gdansk, Poland. He is currently an Associate Professor in the School of Science and Engineering, Reykjavik University, Iceland. His research interests include surrogate-based modeling and optimization, space mapping, circuit theory, analog signal processing, evolutionary computation and numerical analysis.

John W. Bandler studied at Imperial College and received the BSc(Eng.), PhD, and DSc(Eng.) degrees from the University of London, England, in 1963, 1967, and 1976, respectively. He joined McMaster University, Canada, in 1969. He is now Professor Emeritus. He was a President of Optimization Systems Associates Inc., which he founded in 1983, until November 20, 1997, the date of acquisition by Hewlett-Packard Company. He is the President of Bandler Corporation, which he founded in 1997. He is a Fellow of several societies, including the Royal Society of Canada. In 2004, he received the IEEE MTT-S Microwave Application Award.

DOI: 10.1002/jnm