Post-print of: Kozieł S., Pietrenko-Dąbrowska A.: Performance-Based Nested Surrogate Modeling of Antenna Input Characteristics. IEEE TRANSACTIONS ON ANTENNAS AND PROPAGATION. Vol. 67, iss. 5 (2019), pp. 2904 - 2912. DOI: 10.1109/TAP.2019.2896761

Performance-Based Nested Surrogate Modeling of Antenna Input Characteristics

Slawomir Koziel, *Senior Member*, *IEEE*, and Anna Pietrenko-Dabrowska

***Abstract*—Utilization of electromagnetic (EM) simulation tools is mandatory in the design of contemporary antenna structures. At the same time, conducting designs procedures that require multiple evaluations of the antenna at hand, such as parametric optimization or yield-driven design, is hindered by a high cost of accurate EM analysis. To certain extent, this issue can be addressed by utilization of fast replacement models (also referred to as surrogates). Unfortunately, due to curse of dimensionality, traditional data-driven surrogate modeling methods are limited to antenna structures described by a few parameters with relatively narrow parameter ranges. This is by no means sufficient given the complexity of modern designs. In this paper, a novel technique for surrogate modeling of antenna structures is proposed. It involves a construction of two levels of surrogates, both realized as kriging interpolation models. The first model is based on a set of reference designs optimized for selected performance figures. It is used to establish a domain for the final (second-level) surrogate. This formulation permits efficient modeling within wide ranges of antenna geometry parameters and wide ranges of performance figures (e.g., operating frequencies). At the same time, it allows uniform allocation of training data samples in a straightforward manner. Our approach is demonstrated using two microstrip antenna examples and compared to conventional kriging and radial basis function modeling. Application examples for antenna optimization are also provided along with experimental validation.**

***Index Terms*—Antenna design, surrogate modeling, approximation models, simulation-driven design, kriging interpolation.**

I. Introduction

Contemporary antenna structures need to satisfy more and more stringent specifications concerning both electrical and field performance. Furthermore, increasing demands are imposed concerning simultaneous achievement of multiple objectives, realization of additional functionalities (multi-band operation [1], band notches [2], circular polarization [3], etc.), as well as maintaining small physical dimensions [4]. Fulfilling these demands requires development of non-standard antenna geometries which are topologically complex and described by a large number of parameters. Clearly, only full-wave electromagnetic (EM) analysis is capable of providing reliable evaluation of such designs. While EM simulation is not a problem for one-time design verification, massive simulations required when carrying out tasks such as parametric optimization [5], statistical analysis [6], or robust (tolerance-aware) design [7], [8], create a serious practical challenge. The need of incorporating environmental components such as connectors, housing, etc., only adds to already high computational costs. Due to the structural complexity of modern antennas, experience-driven parameter sweeping as a way of parameter tuning is inferior to automated optimization. The latter, however, is often impractical when using conventional methods, both local [9] and global [10], [11]. Addressing these issues led to development of several techniques that offer improved computational efficiency. One of these is utilization of adjoint sensitivities within gradient-based optimization frameworks [12], [13], which is not widespread primarily because of a limited support for this technology in commercial simulation software packages. Another class of methods are surrogate-assisted algorithms where computational burden is cast into cheaper representations of the antenna at hand. The two major groups of surrogates are physics-based ones (space mapping [14], response correction techniques [15]-[18]) and data-driven models [19], [20]. Computational savings can also be obtained by exploiting a particular structure of the system response (feature-based optimization [21]).

In general, utilization of surrogate models allows for alleviating difficulties related to high cost of extensive EM simulations. In a strictly optimization context, surrogates are often constructed locally (along the optimization path). For generic purposes, global or quasi-global models are needed, which are supposed to be valid over the entire parameter space or a large portion thereof. Because of low evaluation cost and versatility, data-driven surrogates constitute the most popular class of models. Widely used techniques include polynomial regression [22], kriging [23], radial basis function interpolation [23], Gaussian process regression [24], neural networks [25], and support vector regression [26]. Unfortunately, the number of training data points necessary to construct a reliable surrogate (in many applications, RMS error below 5 percent is considered sufficient [27]) increases rapidly with the design space dimensionality but also the parameter ranges, which is a common disadvantage of approximation modeling. The latter problem is actually of higher importance because in order for the surrogate to be useful in a design work, it should cover a

The manuscript was submitted on November 28, 2018. This work was supported in part by the Icelandic Centre for Research (RANNIS) Grant 174114051, and by National Science Centre of Poland Grant 2015/17/B/ST6/01857.

S. Koziel is with Engineering Optimization and Modeling Center of Reykjavik University, Reykjavik, Iceland (e-mail: koziel@ru.is); A. Pietrenko-Dabrowska and also S. Koziel are with Faculty of Electronics, Telecommunications and Informatics, Gdansk University of Technology, 80-233 Gdansk, Poland.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2019.2896761, IEEE Transactions on Antennas and Propagation

decent range of operating condition variability (and, consequently, be valid within wide range of geometry/material parameters).

Data-driven models are normally set up in the domains defined by lower and upper bounds for the problem parameters. From the designer's perspective, vast majority of such domains are uninteresting because they contain designs that are poor with respect to whatever performance figures may be considered. For example, re-designing a multi-band antenna for various operating frequencies requires joint tuning of multiple parameters to maintain resonances at sufficient depths or operating bands at sufficient widths. Identification of the parameter space regions that contain good designs may save a lot of computational effort when setting up a surrogate. In [28], a constrained modeling technique has been proposed, in which the computational savings are achieved by restricting the model domain to a vicinity of a manifold spanned by a set of reference (base) designs. These designs could be pre-existing (e.g., from the previous design work with the same structure) or optimized specifically with respect to selected performance figures (e.g., operating frequency, bandwidth, etc.). Domain restriction permits construction of approximation models within higher-dimensional spaces without formally limiting the ranges of geometry parameters. In [29], the method [28] has been generalized for an arbitrary number of performance figures and an arbitrary allocation of the base designs. Despite the mentioned advantages, the constrained modeling techniques are rather complex in terms of the surrogate model domain geometry, applications in design optimization (implicit domain definition makes it problematic to maintain feasibility of solutions in the course of the optimization run), as well as design of experiments (uniform sampling of constrained domain is a non-trivial task [29]).

This paper proposes a novel performance-driven surrogate modeling approach. Our technique is based on a nested kriging interpolation. The first-level model maps the figures of interest (or objective) space into the geometry parameter space of the antenna in order to establish the surrogate model domain. The domain is defined as the objective space image through the first-level mapping orthogonally extended to provide it with non-zero "thickness". Subsequently, the second-level model is constructed over that domain to represent the antenna responses. Although the proposed methodology shares a similar concept of constrained modeling with the techniques of [28] and [29], it is significantly different from these in the following respects, thus constituting an important technical novelty and bringing contributions: (i) the mapping between the objective space and surrogate domain is surjective (and even homeomorphic under mild assumptions concerning the domain "thickness"), (ii) the surrogate domain has a regular geometry which is a continuous image of a unit hypercube, (iii) uniform allocation of training data samples is straightforward (as a consequence of (ii)), (iv) antenna optimization within the surrogate model domain is also straightforward as the domain is defined explicitly (as opposed to [28] and [29]). The surrogate modeling approach introduced here is demonstrated using a dual-band uniplanar dipole antenna and a ring slot antenna. In both cases, accurate models are constructed for wide ranges of operating conditions (operating frequencies for the dipole antenna and center frequency and substrate permittivity for the slot antenna) and geometry parameters using small training sets. Predictive power improvement over the benchmark methods is shown to be dramatic. Applications of the surrogates to antenna optimization as well as experimental validation of selected designs are also provided.

II. Performance-Based Nested Kriging Modeling

Conventional approximation-based surrogates are constructed in regular domains, typically being intervals (hypercubes), defined by lower and upper bounds for design variables. Handling such domains is straightforward. This includes allocation of training data samples: most of design of experiments techniques [30]-[32] have been developed for interval-type sets. At the same time, due to complex interactions between geometry parameters of antenna structures, design that are "good", that is, optimal with respect to any particular performance figures of choice, are concentrated within small subsets (low-dimensional manifolds) [29]. This means that vast majority of hypercube-like domains contain uninteresting designs. Consequently, constraining the modeling process to the "promising" regions may allow us to significantly reduce the training data set required to construct the model while maintaining its good predictive power.

This section outlines the proposed modeling approach. The fundamental concept is to identify the "promising" region of the parameter space and build the surrogate within it. This is explained on a generic level in Section II.*A*. Section II.*B* provides particulars concerning practical realization of the concept using a set of reference designs. A formal definition of the surrogate domain, the nested modeling framework, and the uniform sampling procedure are described in Sections II.*C* through II.*E*.

A. Objective Space and Geometry of Optimum Design Sets

We denote by X the parameter space defined in a usual manner, i.e., using the lower and upper bounds on design variables $\boldsymbol{l} \le \boldsymbol{x} \le \boldsymbol{u}$, where $\boldsymbol{x} = [x_1 \ldots x_n]^T$, $\boldsymbol{l} = [l_1 \ldots l_n]^T$, $\boldsymbol{u} = [u_1 \ldots u_n]^T$, or $X = [l_1\ u_1] \times \ldots \times [l_n\ u_n]$. Let f_k, $k = 1, \ldots, N$, denote the figures of interest to be considered in the design process (e.g., operating frequencies of a multi-band antenna). The objective space F is defined by the ranges (lower and upper bounds) for the figures f_k, $f_{k.\min} \le f_k^{(j)} \le f_{k.\max}$, $k = 1, \ldots, N$, i.e., we have $F = [f_{1.\min}\ f_{1.\max}] \times \ldots \times [f_{N.\min}\ f_{N.\max}]$. Further, let U be the scalar objective function such that

$$U_f(\boldsymbol{f}) = \arg\min_{\boldsymbol{x}} U(\boldsymbol{x}, \boldsymbol{f}) \tag{1}$$

is the optimum design of the antenna at hand for a given vector of figures of interest $\boldsymbol{f}$. For example, if $\boldsymbol{f}$ is a vector of operating frequencies of a multi-band antenna, $U(\cdot)$ may be defined as $-\min\{B_1,\ldots,B_N\}$, where B_j is the fractional operating bandwidth corresponding to the operating frequency f_j. In that case, $U_f(\boldsymbol{f})$ will be a design that maximizes the antenna fractional bandwidths while allocating them at the required frequencies $\boldsymbol{f}$. The image $U_f(F)$ of the objective space F is an N-dimensional

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2019.2896761, IEEE Transactions on Antennas and Propagation

manifold in the parameter space X as illustrated in Fig. 1. From the point of view of antenna design, only $U_f(F)$ is of interest when considering the performance figures f_1 through f_N. A practical problem is of course identification of this set and constructing the surrogate model within $U_f(F)$. This is the subject of the proposed modeling framework as explained in the remaining part of this section.

B. Reference Designs and Level I Surrogate

Identification of the manifold $U_f(F)$ is realized here using a set of reference designs $\boldsymbol{x}^{(j)} \in U_f(F), j = 1, \ldots, p$, optimized with respect to the performance figure vectors $\boldsymbol{f}^{(j)} = [f_1^{(j)} \ldots f_N^{(j)}]$. In other words, $\boldsymbol{x}^{(j)} = U_f(\boldsymbol{f}^{(j)})$, i.e., it is a solution to (1) with the objective function $U(\boldsymbol{x},\boldsymbol{f}^{(j)})$. It should be noted that the reference designs may be available beforehand (from the previous design work on the same structure) or obtained specifically for the purpose of surrogate model construction.

In order to approximate the manifold $U_f(F)$, the first-level surrogate $\boldsymbol{s}_I(\boldsymbol{f})$ is constructed, which maps the objective space F into the parameter space X. Here, it is implemented as a kriging interpolation model [23] with $\{\boldsymbol{f}^{(j)},\boldsymbol{x}^{(j)}\}$ being the training set. The conceptual illustration of $\boldsymbol{s}_I(\cdot)$ is shown in Figs. 2(a) and 2(b).

C. Surrogate Model Domain

It should be emphasized that $\boldsymbol{s}_I(F) \subset X$ is only an approximation of the manifold $U_f(F)$, obtained using limited information (here, the reference designs). In order to ensure that the entire manifold is a proper subset of the surrogate model domain to be constructed, certain "thickness" has to be provided by extending $\boldsymbol{s}_I(F)$ in all orthogonal directions to it.

We denote by $\{\boldsymbol{v}_n^{(k)}(\boldsymbol{f})\}$, $k = 1, \ldots, n - N$, an orthonormal basis of vectors normal to $\boldsymbol{s}_I(F)$ at $\boldsymbol{f}$. Moreover, we define $\boldsymbol{x}_{\max} = \max\{\boldsymbol{x}^{(k)}, k = 1, \ldots, p\}$ and $\boldsymbol{x}_{\min} = \min\{\boldsymbol{x}^{(k)}, k = 1, \ldots, p\}$, and $\boldsymbol{d}_x = \boldsymbol{x}_{\max} - \boldsymbol{x}_{\min}$ as the range of variation of antenna geometry parameters within $\boldsymbol{s}_I(F)$. Let D be a user-defined thickness parameter. Using this notation, we define extension coefficients as follows:

$$\boldsymbol{a}(\boldsymbol{f}) = [a_1(\boldsymbol{f}) \ldots a_{n-N}(\boldsymbol{f})]^{\mathrm{T}} = \frac{D}{2}\left|\boldsymbol{d}_x \boldsymbol{V}_n(\boldsymbol{f})\right|^{\mathrm{T}} = \frac{D}{2}\left[\left|\boldsymbol{d}_x \boldsymbol{v}_n^{(1)}(\boldsymbol{f})\right| \ldots \left|\boldsymbol{d}_x \boldsymbol{v}_n^{(n-N)}(\boldsymbol{f})\right|\right]^{\mathrm{T}} \tag{2}$$

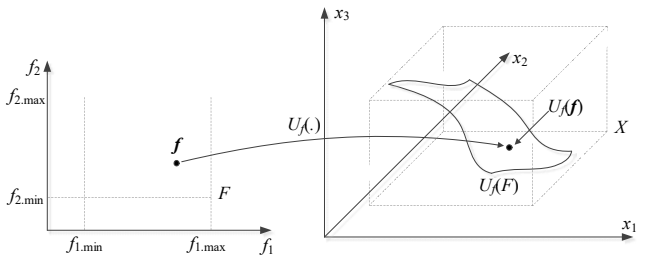

Fig. 1. The objective space F (left panel) and the design space X (right panel). The image $U_f(F)$ of F is an N-dimensional manifold in X (here, $N = 2$), which contains designs that are optimal with respect to the figures of interest f_1 through f_N. From design purposes point of view, the surrogate modeling process can be restricted to $U_f(F)$ without any loss of information concerning designs of interest with respect to the selected figures of interest.

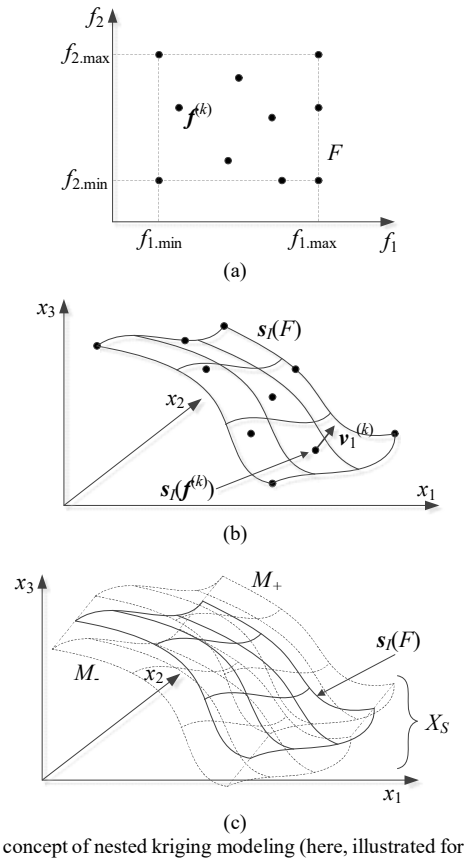

Fig. 2. The concept of nested kriging modeling (here, illustrated for two figures of interest and three-dimensional parameter space): (a) reference designs and objective space F; (b) the image $\boldsymbol{s}_I(F)$ first levels surrogate model and the normal vector $\boldsymbol{v}_1^{(k)}$ at $\boldsymbol{f}^{(k)}$; (c) manifolds M_- and M_+ as well as the surrogate model domain X_S defined as orthogonal extension of $\boldsymbol{s}_I(F)$.

Here, the matrix $\boldsymbol{V}_n$ consists of the vectors $\boldsymbol{v}_n^{(k)}(\boldsymbol{f})$ (as its columns).

The coefficients of (2) determine the boundaries of the surrogate model domain X_S, which is allocated between the following two manifolds:

$$M_- = \left\{\boldsymbol{x} \in X : \boldsymbol{x} = \boldsymbol{s}_I(\boldsymbol{f}) - \sum_{k=1}^{n-N} a_k(\boldsymbol{f})\boldsymbol{v}_n^{(k)}(\boldsymbol{f})\right\} \tag{3}$$

and

$$M_+ = \left\{\boldsymbol{x} \in X : \boldsymbol{x} = \boldsymbol{s}_I(\boldsymbol{f}) + \sum_{k=1}^{n-N} a_k(\boldsymbol{f})\boldsymbol{v}_n^{(k)}(\boldsymbol{f})\right\} \tag{4}$$

In other words, we have

$$X_S = \left\{\begin{array}{l} \boldsymbol{x} = \boldsymbol{s}_I(\boldsymbol{f}) + \sum_{k=1}^{n-N} \lambda_k a_k(\boldsymbol{f})\boldsymbol{v}_n^{(k)}(\boldsymbol{f}) : \boldsymbol{f} \in F, \\ -1 \le \lambda_k \le 1, k = 1,\ldots,n-N \end{array}\right\} \tag{5}$$

Thus, the domain X_S is defined as a set of all points of the form $\boldsymbol{x} = \boldsymbol{s}_I(\boldsymbol{f}) + \sum_{k=1,\ldots,n-N} \lambda_k a_k(\boldsymbol{f})\boldsymbol{v}_n^{(k)}(\boldsymbol{f})$ with $\boldsymbol{f} \in F$ and $-1 \le \lambda_k \le 1$, for $k = 1, \ldots, n - N$.

The concept of orthogonal extension of $\boldsymbol{s}_I(F)$ has been graphically illustrated in Figs. 2(b) and 2(c). It should be noted that the thickness parameter D determines the "amount" of the orthogonal extension of $\boldsymbol{s}_I(F)$ (into the manifolds M_- and M_+), or, in other words, the lateral size of the surrogate model domain in relation to the tangential size of $\boldsymbol{s}_I(F)$.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2019.2896761, IEEE Transactions on Antennas and Propagation

D. Level II Surrogate

The second-level surrogate is a kriging interpolation model set up in X_S, using a set of training data samples $\{\boldsymbol{x}_B^{(k)}, \boldsymbol{R}(\boldsymbol{x}_B^{(k)})\}_{k = 1, \ldots, N_B}$, where $\boldsymbol{R}$ is the EM-simulation model of the antenna at hand. The samples are allocated as described in Section II.*E*. The second-level model is our final surrogate which accounts for the antenna designs restricted to the vicinity of the manifold $U_f(F)$. Figure 3 shows the flow diagram of the proposed surrogate modeling methodology.

E. Design of Experiments for Nested Kriging

Achieving the best possible predictive power of the surrogate model requires uniform allocation of the training data samples (note that a "one-shot" design of experiments is considered here rather than sequential sampling [33]). This has been a major issue for the previous attempts to constrained antenna modeling [29]. The problem was a complex definition and geometry of the surrogate model domain, determined through triangulation of the reference designs [29].

For the proposed nested modeling methodology, uniform sampling can be realized in a convenient manner, directly using the domain definition (5) and a two-step surjective transformation of a unit hypercube onto X_S. Let $[0,1]^n$ be a unit hypercube, and $\{\boldsymbol{z}^{(k)}\}$, $k = 1, \ldots, N_B$, be a uniformly distributed data set (here, obtained using a Latin Hypercube Sampling [34]), with $\boldsymbol{z}^{(k)} = [z_1^{(k)} \ldots z_n^{(k)}]^T$.

The first step of mapping the samples onto X_S is realized as

$$\begin{aligned}\boldsymbol{y} = h_1(\boldsymbol{z}) = h_1([z_1 \ldots z_n]^T) = \\ = \left[f_{1.\min} + z_1(f_{1.\max} - f_{1.\min}) \ \ldots \ f_{N.\min} + z_N(f_{N.\max} - f_{N.\min})\right] \times \\ \times [-1 + 2z_{N+1} \ \ldots \ -1 + 2z_n]\end{aligned} \tag{6}$$

In other words, h_1 maps the hypercube onto a Cartesian product $F \times [-1,1]^{n-N}$.

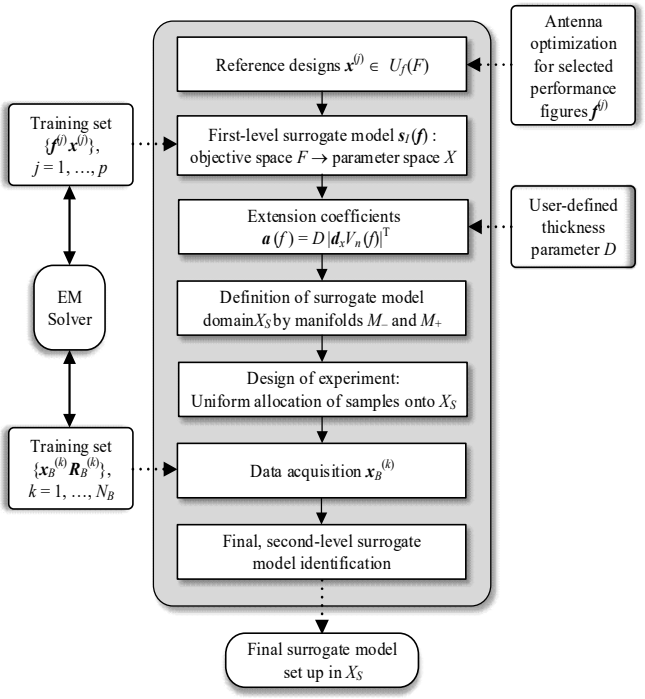

Fig. 3. Flow diagram of the proposed nested kriging modeling methodology.

The second step is mapping of $F \times [-1,1]^{n-N}$ onto X_S as

$$\begin{aligned}\boldsymbol{x} = h_2(\boldsymbol{y}) = h_2([y_1 \ldots y_n]^T) = \boldsymbol{s}_I([y_1 \ \ldots \ y_N]^T) + \\ + \sum_{k=1}^{n-N} y_{N+k} a_k([y_1 \ \ldots \ y_N]^T) \boldsymbol{v}_n^{(k)}([y_1 \ \ldots \ y_N]^T)\end{aligned} \tag{7}$$

The meaning of expansion coefficients a_k and the normal vectors $\boldsymbol{v}_n^{(k)}$ has been explained in Section II.*C*. The samples $\boldsymbol{x}_B^{(k)}$ obtained as

$$\boldsymbol{x}_B^{(k)} = H(\boldsymbol{z}^{(k)}) = h_2(h_1(\boldsymbol{z}^{(k)})) \tag{8}$$

are uniformly distributed with respect to the figures of interest f_1 through f_N. It should be noted that the same mapping can be utilized for other purposes such as parametric optimization within the domain X_S, which is convenient because regardless of the domain geometry, it is sufficient to operate within $F \times [-1,1]^{n-N}$ and only apply the mapping (8) for antenna evaluation purposes (cf. Section IV). Graphical illustration of the sampling procedure has been provided in Fig. 4.

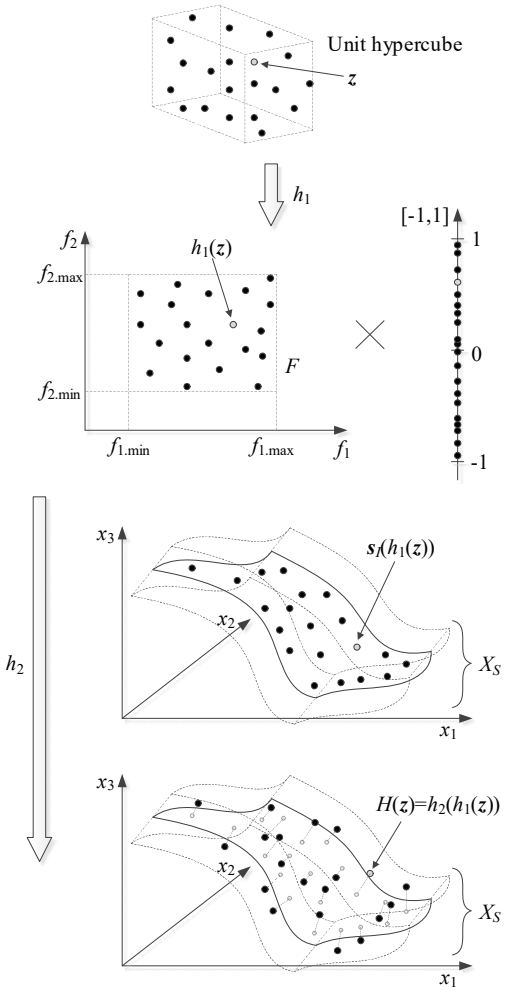

Fig. 4. Sampling procedure in the surrogate model domain X_S (two-dimensional objective space and three-dimensional parameters space is used here for illustration purposes, cf. Fig. 2). LHS-allocated samples are first mapped onto the Cartesian product of F and $[-1,1]^{n-N}$ using function h_1 (cf. (6)). The samples are then mapped onto X_S using function h_2 (cf. (7)). An additional picture (second from the bottom) also illustrates samples $\boldsymbol{s}_I(h_1(\boldsymbol{z}))$ mapped into the image $\boldsymbol{s}_I(F)$ of the objective space F (i.e., before their orthogonal relocation as in the second term of (7).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2019.2896761, IEEE Transactions on Antennas and Propagation

III. Verification Examples

In this section, operation and performance of the proposed nested modeling framework is demonstrated using two examples, a uniplanar dipole antenna and a ring slot antenna. The analysis of the modeling error versus the size of the training data set is carried out along with benchmarking using conventional kriging and radial basis function interpolation surrogates. Applications to antenna optimization are considered in Section IV.

A. Case 1: Uniplanar Dipole Antenna

Our first example is a dual-band uniplanar dipole antenna (Antenna I) shown in Fig. 5 [35]. The antenna is implemented on a Rogers RO4350 substrate ($\varepsilon_r = 3.5$, $h = 0.76$ mm). It is fed by a 50 Ohm coplanar waveguide (CPW). The variables are: $\boldsymbol{x} = [l_1\ l_2\ l_3\ w_1\ w_2\ w_3]^T$, whereas $l_0 = 30$, $w_0 = 3$, $s_0 = 0.15$ and $o = 5$ are fixed (all dimensions in mm). The EM antenna model $\boldsymbol{R}$ (~100,000 cells; 60 s simulation on a dual Xeon E5540 machine) is implemented in CST Microwave Studio and evaluated using its time-domain solver.

The goal is to construct the surrogate model of the antenna in Fig. 5 which is valid for the following ranges of operating frequencies 2.0 GHz $\le f_1 \le$ 3.0 GHz (lower band), and 4.0 GHz $\le f_2 \le$ 5.5 GHz (upper band). The allocation of ten reference designs $\boldsymbol{x}^{(j)}$, j = 1, …, 10, selected for illustration purposes, have been shown in Fig. 6. The lower and upper bounds for design variables are based on $\{\boldsymbol{x}^{(j)}\}$. We have $\boldsymbol{l} = [29\ 5.0\ 17\ 0.2\ 1.5\ 0.5]^T$, and $\boldsymbol{u} = [42\ 12\ 25\ 0.6\ 5.2\ 3.5]^T$. It should be noted that the parameter ranges are wide with the ratio between the upper and the lower bounds varying from around 1.5 to 7.0 and the average of 3.1.

For validation purposes, the nested surrogate model has been set up using the training data sets of the sizes of 50, 100, 200, 400, and 800 samples. This has been done for three different values of the thickness parameter, D = 0.05, 0.1, and 0.15. The model accuracy has been tested using a split-sample method [36] with 100 independent random test points. Table I shows the average RMS errors for the nested kriging model, conventional kriging model [36], and radial basis function (RBF) [23] surrogate set up in the interval [$\boldsymbol{l}$, $\boldsymbol{u}$]. The RBF model utilizes Gaussian basis function with the scaling parameter adjusted through cross validation. Responses of the EM simulation model of the antenna as well as conventional and nested kriging surrogates (set up with 200 training samples) at the selected testing designs have been shown in Figs. 7 and 8, respectively.

The results gathered in Table I indicate that the nested kriging surrogate exhibits significantly better predictive power than the conventional kriging model. For D = 0.1, its accuracy with only 50 training samples is the same as the accuracy of the conventional kriging model at 200 samples, whereas the accuracy comparison for the same corresponding training data sets reveals dramatic advantage of the proposed approach by a factor between two and three.

It can also be observed that the dependence between the model accuracy and the thickness parameter D is weak considering dimensionality of the design space. The reason for this is that the surrogate model domain is a relatively "thin" set, i.e., the amount of the orthogonal extension is small compared to the "tangential" dimensions of X_S. Consequently, the effective dimension of the domain is close to that of the objective space, and, in practice, we have $N << n$. This is a highly desirable feature which makes the nested surrogate relatively insensitive to D.

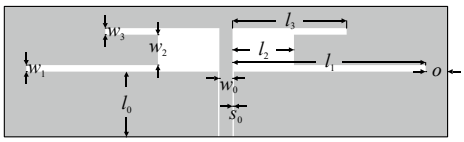

Fig. 5. Geometry of dual-band uniplanar dipole antenna (Antenna I) [35].

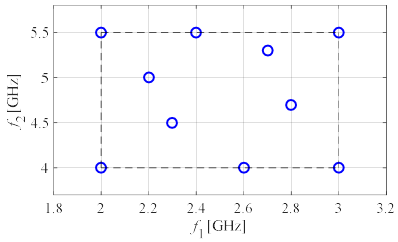

Fig. 6. Allocation of the reference designs for nested kriging modeling of Antenna I.

TABLE I Modeling Results For Antenna I

Number of training samples	Relative RMS Error				
	Conventional Models		Nested Kriging Model [this work]		
	Kriging	RBF	D = 0.05	D = 0.1	D = 0.15
50	21.7 %	24.9 %	6.9 %	9.9 %	13.0 %
100	17.3 %	19.8 %	4.5 %	6.4 %	9.6 %
200	12.6 %	14.3 %	2.8 %	4.4 %	6.5 %
400	9.3 %	10.5 %	2.6 %	3.8 %	5.5 %
800	7.2 %	8.7 %	2.4 %	3.4 %	4.7 %

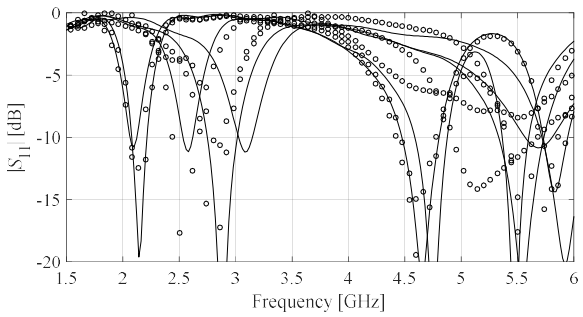

Fig. 7. Responses of Antenna I at the selected test designs for N = 200: EM model (—), conventional kriging surrogate (o).

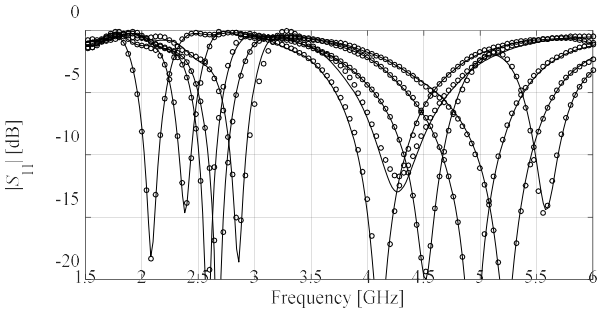

Fig. 8. Responses of Antenna I at the selected test designs for N = 200: EM model (—), proposed nested kriging surrogate with D = 0.1 (o).

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2019.2896761, IEEE Transactions on Antennas and Propagation

B. Example 2: Ring Slot Antenna

The second example is a slot antenna shown in Fig. 9 [28]. The structure comprises a microstrip line that feeds a circular ground plane slot with defected ground structure (DGS). The low-pass properties of the DGS allows for suppression of the antenna harmonic frequencies [28]. The substrate thickness is set to 0.76 mm. The geometry parameter set is: $\boldsymbol{x} = [l_f\ l_d\ w_d\ r\ s\ s_d\ o\ g]^T$; ε_r is an additional variable representing relative permittivity of the substrate. The feed line width w_f is computed for each ε_r to ensure 50 ohm input impedance. The antenna model $\boldsymbol{R}$ is implemented in CST (~300,000 cells, simulation 90 s).

For this antenna, the goal is to construct the surrogate model for the operating frequencies f within the range 2.5 GHz $\le f \le$ 6.5 GHz, and substrate permittivity ε within the range of $2.0 \le \varepsilon \le 5.0$. The allocation of ten reference designs $\boldsymbol{x}^{(j)}, j = 1, \ldots, 10$, have been shown in Fig. 10. The lower and upper bounds for design variables are based on $\{\boldsymbol{x}^{(j)}\}$. We have $\boldsymbol{l} = [22.0\ 3.5\ 0.3\ 6.5\ 3.0\ 0.5\ 3.5\ 0.2]^T$, and $\boldsymbol{u} = [27.0\ 8.0\ 2.3\ 16.0\ 7.0\ 5.5\ 6.0\ 2.3]^T$. Note that the parameter ranges are even wider than for Antenna I, with the ratio between the upper and the lower bounds varying from around 1.2 to 11.5 and the average of 5.0.

The verification setup was the same as for Antenna I. The numerical data has been gathered in Table II whereas the selected reflection characteristics of the EM model and the surrogates are shown in Figs. 11 and 12. The results are consistent with those obtained for the first example. In particular, the nested surrogate exhibits significantly better accuracy than conventional models for all sizes of the training data set. The dependence of the modeling error on the thickness parameter is more pronounced because the modeling problem is much more challenging than for Antenna I. It should be emphasized that the accuracy of conventional models make them unusable for practical purposes (error larger than 25 percent even for the largest training data set), see also Fig. 11.

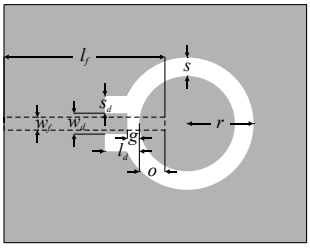

Fig. 9. Geometry of the ring slot antenna (Antenna II) with a microstrip feed (dashed line) [28].

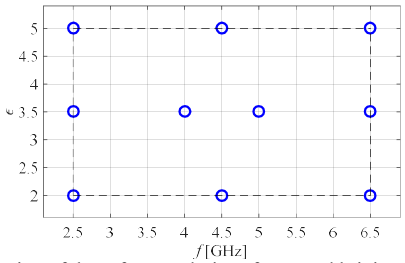

Fig. 10. Allocation of the reference designs for nested kriging modeling of Antenna II.

TABLE II MODELING RESULTS FOR ANTENNA II

Number of training samples	Relative RMS Error				
	Conventional Models		Nested Kriging Model [this work]		
	Kriging	RBF	$D = 0.05$	$D = 0.1$	$D = 0.15$
50	56.9 %	61.0 %	12.9 %	19.4 %	25.1 %
100	50.8 %	53.2 %	6.9 %	12.9 %	16.5 %
200	35.8 %	37.9 %	4.9 %	7.7 %	12.7 %
400	31.5 %	34.1 %	3.1 %	5.1 %	9.0 %
800	25.6 %	27.2 %	2.2 %	3.7 %	6.2 %

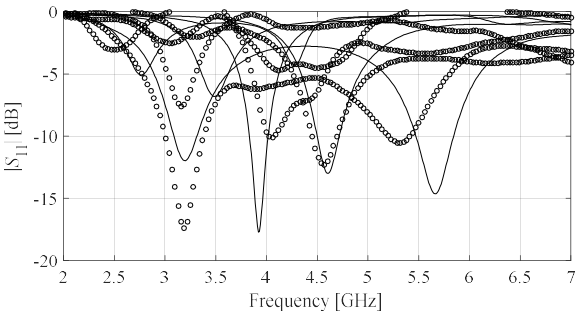

Fig. 11. Responses of Antenna II at the selected test designs for $N = 400$: EM model (—), conventional kriging surrogate (o).

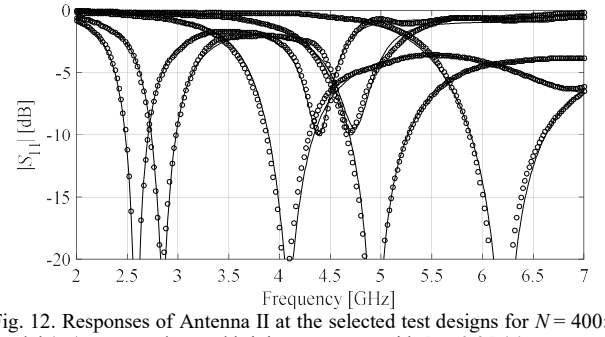

Fig. 12. Responses of Antenna II at the selected test designs for $N = 400$: EM model (—), proposed nested kriging surrogate with $D = 0.05$ (o).

IV. APPLICATION EXAMPLES. ANTENNA OPTIMIZATION

In this section, the application of the nested surrogate for antenna optimization is investigated. We explain how the definition of the model domain can be utilized to facilitate the search for the optimum design, provide numerical results, as well as discuss experimental validation.

A. Optimization Methodology

Design optimization of the antenna has to be carried out within the domain X_S of its surrogate model, which is a rather complex set as explained in Section II. However, the optimization process can be made straightforward by using the transformation (8) between the unit hypercube $[0,1]^n$ and X_S (cf. Section II.*E*). More specifically, given the original design problem (1) and the target vector $\boldsymbol{f}_t = [f_{1.t}\ \ldots\ f_{N.t}]$ the antenna is to be optimized for, one can solve an equivalent problem

$$\boldsymbol{x}^* = \arg\min_{\boldsymbol{x}\in[0,1]^n} U(H(\boldsymbol{z}), \boldsymbol{f}_t) \tag{9}$$

Note that (9) is solved within a hypercube, which is a regular set defined by box constraints, rather than directly in X_S.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2019.2896761, IEEE Transactions on Antennas and Propagation

Further, because H is surjective (cf. Section II.*E*), operating in the hypercube will cover the entire surrogate model domain.

Finally, a good initial design for (9) can be easily identified as follows

$$\boldsymbol{x}^{(0)} = \boldsymbol{s}_I(\boldsymbol{f}_t) \tag{10}$$

In other words, the initial design is the best available approximation of $U_f(\boldsymbol{f}_t)$, here, obtained by the first-level surrogate.

B. Numerical Results

Antennas I and II have been optimized for selected target vectors, using the nested surrogate obtained with the thickness parameter D = 0.1. Tables III and IV gather the numerical data, whereas Figs. 13 and 14 show the reflection characteristics at the initial and optimized designs evaluated using the surrogate and the EM model. It can be observed that reliability of the surrogate is excellent in all cases. Also, the quality of the initial design found using (10) is good, which allows for applying local optimizer instead of global routines. This is another benefit of the nested modeling approach proposed in this work.

C. Experimental Validation

Selected designs of the considered antennas have been fabricated and measured for the sake of additional validation. Both antennas have been implemented on the same substrate, RO4350. Figure 15 shows the photographs of the prototypes, whereas Fig. 16 compares the simulated and measured reflection characteristics. The agreement is very good, slight discrepancies (frequency shifts) are mostly due to not including the SMA connectors in the computational models of the antenna structures.

TABLE III OPTIMIZATION RESULTS FOR ANTENNA I

Target operating conditions		Geometry parameter values [mm]					
f_1 [GHz]	f_2 [GHz]	l_1	l_2	l_3	w_1	w_2	w_3
2.45	5.30	33.0	8.74	17.9	0.26	2.62	1.51
2.20	4.50	34.7	5.96	18.5	0.44	3.71	1.48
3.00	5.00	28.5	9.60	19.6	0.37	2.40	0.58
2.10	4.20	35.1	5.49	19.4	0.47	4.40	1.94

TABLE IV OPTIMIZATION RESULTS FOR ANTENNA II

Target operating conditions	Geometry parameter values [mm]								
f_0 [GHz]	ε	l_f	l_d	w_d	r	s	s_d	o	g
3.4	3.5	26.0	5.91	0.79	11.9	4.94	3.22	4.74	1.04
4.8	2.2	22.8	4.79	0.56	9.46	3.21	3.56	5.16	1.23
5.3	3.5	21.0	4.57	0.31	8.63	3.35	4.94	5.55	1.76
2.45	4.3	26.8	6.46	2.03	13.5	6.22	1.54	4.93	0.22

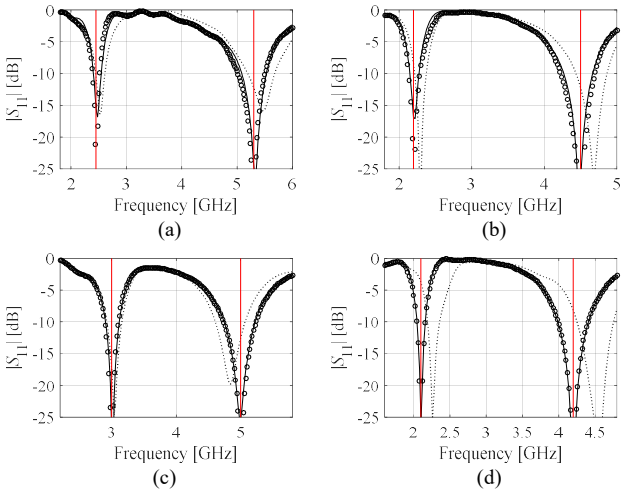

Fig. 13. Nested surrogate (o) and EM-simulated responses (—) of Antenna I at the designs obtained by optimizing the surrogate model with D = 0.1 for (a) f_1 = 2.45 GHz, f_2 = 5.3 GHz, (b) f_1 = 2.2 GHz, f_2 = 4.5 GHz, (c) f_1 = 3.0 GHz, f_2 = 5.0 GHz, and (d) f_1 = 2.1 GHz, f_2 = 4.2 GHz. Required operating frequencies are marked using vertical lines. Initial design obtained using (10) marked using a dotted line.

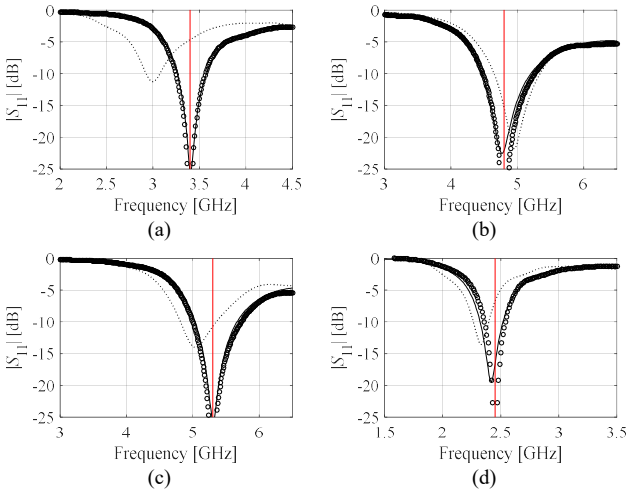

Fig. 14. Nested surrogate (o) and EM-simulated responses (—) of Antenna II at the designs obtained by optimizing the surrogate model with D = 0.1 for (a) f_1 = 2.45 GHz, f_2 = 5.3 GHz, (b) f_1 = 2.2 GHz, f_2 = 4.5 GHz, (c) f_1 = 3.0 GHz, f_2 = 5.0 GHz, and (d) f_1 = 2.0 GHz, f_2 = 4.2 GHz. Required operating frequencies are marked using vertical lines. Initial design obtained using (10) marked using a dotted line.

Fig. 15. Photographs of the fabricated antenna prototypes: (a) Antenna I, f_1 = 2.45 GHz, f_2 = 5.3 GHz, (b) Antenna II, f_0 = 3.4 GHz, ε = 3.5.

This article has been accepted for publication in a future issue of this journal, but has not been fully edited. Content may change prior to final publication. Citation information: DOI 10.1109/TAP.2019.2896761, IEEE Transactions on Antennas and Propagation

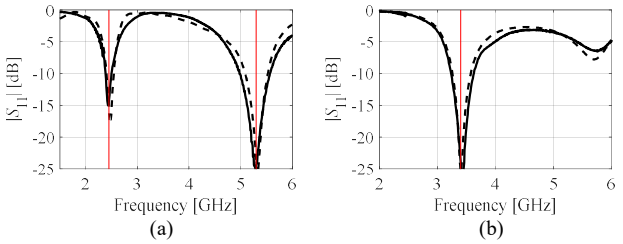

Fig. 16. Reflection responses of the antennas of Fig. 15: (a) Antenna I, f_1 = 2.45 GHz, f_2 = 5.3 GHz, (b) Antenna II, f_0 = 3.4 GHz, ε = 3.5; simulation results (- - -) and measurements (—).

V. Conclusion

The paper proposed a novel approach to surrogate modeling of antenna structures. Our methodology is based on a two-level (or nested) kriging, where the first-level model is used to establish domain of the second-level model being the actual surrogate. The first-level model approximates the manifold containing the designs that are optimum with respect to selected performance figures. This approximation is then orthogonally extended in order to capture all designs that are of interest. The proposed approach has several important advantages over conventional surrogate models as well as some recent attempts to performance-driven constrained modeling. First, by focusing the modeling process in the region containing quality designs, significant improvement of the surrogate predictive power can be achieved. Second, defining the surrogate model domain through a surjective transformation from a unity hypercube permits uniform allocation of training data points in a straightforward manner. Finally, utilization of the same mapping allows for convenient optimization of surrogate (with only box constraints on the adjustable parameters) despite nonlinear geometry of the domain itself. An additional benefit is that a good initial design corresponding to the requested performance figures (or operating condition) values can be directly extracted from the first-level surrogate. Comprehensive verification of the proposed framework provides consistent results throughout the considered benchmark antennas and training data sets, also supported by experimental validation. Furthermore, the nested surrogate has been demonstrated as a useful antenna design tool.

Acknowledgement

The authors would like to thank Dassault Systemes, France, for making CST Microwave Studio available.

References

[1] W.J. Liao, C.Y. Hsieh, B.Y. Dai, and B.R. Hsiao, "Inverted-F/slot integrated dual-band four-antenna system for WLAN access point," *IEEE Ant. Wireless Prop. Lett.*, vol. 14, pp. 847-850, 2015.

[2] I.B. Vendik, A. Rusakov, K. Kanjanasit, J. Hong, and D. Filonov, "Ultrawideband (UWB) planar antenna with single-, dual- and triple-band notched characteristic based on electric ring resonator," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, pp. 1597-1600, 2017.

[3] L. Zhang, S. Gao, Q. Luo, P.R. Young, and Q. Li, "Wideband loop antenna with electronically switchable circular polarization," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, pp. 242-245, 2017.

[4] J. Wu and K. Sarabandi, "Compact omnidirectional circularly polarized antenna," *IEEE Trans. Ant. Prop.*, vol. 65, no. 4, pp. 1550-1557, 2017.

[5] A. Lakbakhsh, M.U. Afzal, and K.P. Esselle, "Multiobjective particle swarm optimization to design a time-delay equalizer metasurface for an electromagnetic band-gap resonator antenna," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, pp. 915-915, 2017.

[6] J. Zhang, C. Zhang, F. Feng, W. Zhang, J. Ma, and Q.J. Zhang, "Polynomial chaos-based approach to yield-driven EM optimization," *IEEE Trans. Microwave Theory Tech.*, vol. 66, no. 7, pp. 3186-3199, 2018.

[7] A.S.O. Hassan, H.L. Abdel-Malek, A.S.A. Mohamed, T.M. Abuelfadl, and A.E. Elqenawy, "Statistical design centering of RF cavity linear accelerator via non-derivative trust region optimization," *IEEE Int. Conf. Numerical EM Multiphysics Modeling Opt. (NEMO)*, pp. 1-3, 2015.

[8] A. Kouassi, N. Nguyen-Trong, T. Kaufmann, S. Lallechere, P. Bonnet, and C. Fumeaux, "Reliability-aware optimization of a wideband antenna," *IEEE Trans. Ant. Prop.*, vol. 64, no. 2, pp. 450-460, 2016.

[9] J. Nocedal and S. Wright, *Numerical Optimization*, 2nd edition, Springer, New York, 2006.

[10] A. Darvish and A. Ebrahimzadeh, "Improved fruit-fly optimization algorithm and its applications in antenna array synthesis," *IEEE Trans. Ant. Prop.*, vol. 66, no. 4, pp. 1756-1766, 2018.

[11] S. K. Goudos, K. Siakavara, T. Samaras, E. E. Vafiadis, and J. N. Sahalos, "Self-adaptive differential evolution applied to real-valued antenna and microwave design problems," *IEEE Trans. Antennas Propag.*, vol. 59, no. 4, pp. 1286–1298, Apr. 2011.

[12] J. Wang, X.S. Yang, and B.Z. Wang, "Efficient gradient-based optimization of pixel antenna with large-scale connections," *IET Microwaves Ant. Prop.*, vol. 12, no. 3, pp. 385-389, 2018.

[13] S. Koziel and A. Bekasiewicz, "Rapid design optimization of antennas using variable-fidelity EM models and adjoint sensitivities," *Eng. Comp.*, vol. 33, no. 7, pp. 2007-2018, 2016.

[14] J. Zhu, J.W. Bandler, N.K. Nikolova, and S. Koziel, "Antenna optimization through space mapping," *IEEE Transactions on Antennas and Propagation*, vol. 55, no. 3, pp. 651-658, March 2007.

[15] S. Koziel and S. Ogurtsov, *Antenna design by simulation-driven optimization. Surrogate-based approach*. Springer, New York, 2014.

[16] S. Koziel and L. Leifsson, "Simulation-driven design by knowledge-based response correction techniques," Springer, 2016.

[17] S. Koziel and S.D. Unnsteinsson "Expedited design closure of antennas by means of trust-region-based adaptive response scaling," *IEEE Antennas Wireless Prop. Lett.*, vol. 17, no. 6, pp. 1099-1103, 2018.

[18] Y. Su, J. Lin, Z. Fan, and R. Chen, "Shaping optimization of double reflector antenna based on manifold mapping," *Int. Applied Computational Electromagnetic Society Symp. (ACES)*, pp. 1-2, 2017.

[19] J.A. Easum, J. Nagar, and D.H. Werner, "Multi-objective surrogate-assisted optimization applied to patch antenna design," *Int. Symp. Ant. Prop.*, pp. 339-340, San Diego, USA, 2017.

[20] D.I.L. de Villiers, I. Couckuyt, and T. Dhaene, "Multi-objective optimization of reflector antennas using kriging and probability of improvement," *Int. Symp. Ant. Prop.*, pp. 985-986, San Diego, USA, 2017.

[21] S. Koziel, "Fast simulation-driven antenna design using response-feature surrogates," *Int. J. RF & Micr. CAE*, vol. 25, no. 5, pp. 394-402, 2015.

[22] J.L. Chávez-Hurtado and J.E. Rayas-Sánchez, "Polynomial-based surrogate modeling of RF and microwave circuits in frequency domain exploiting the multinomial theorem," *IEEE Trans. Microwave Theory Tech.*, vol. 64, no. 12, pp. 4371-4381, 2016.

[23] T.W. Simpson, J.D. Pelplinski, P.N. Koch, and J.K. Allen, "Metamodels for computer-based engineering design: survey and recommendations", *Engineering with Computers*, vol. 17, pp. 129-150, 2001.

[24] J.P. Jacobs, "Characterization by Gaussian processes of finite substrate size effects on gain patterns of microstrip antennas," *IET Microwaves Ant. Prop.*, vol. 10, no. 11, pp. 1189-1195, 2016.

[25] H. Kabir, Y. Wang, M. Yu, and Q.J. Zhang, "Neural network inverse modeling and applications to microwave filter design," *IEEE Trans. Microwave Theory Tech.*, vol. 56, no. 4, , pp. 867-879, April 2008.

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

[27] I. Couckuyt, "Forward and inverse surrogate modeling of computationally expensive problems," Ph.D. Thesis, Ghent University, 2013.

[28] S. Koziel and A. Bekasiewicz, "On reduced-cost design-oriented constrained surrogate modeling of antenna structures," *IEEE Ant. Wireless Prop. Lett.*, vol. 16, pp. 1618-1621, 2017.

[29] S. Koziel and A.T. Sigurdsson, "Triangulation-based constrained surrogate modeling of antennas," *IEEE Trans. Ant. Prop.*, vol. 66, no. 8, pp. 4170-4179, 2017.

[30] S. Leary, A. Bhaskar, and A. Keane, "Optimal orthogonal-array-based latin hypercubes," *J. Applied Statistics*, vol. 30, no. 5, pp. 585-598, 2003.

[31] T.J. Santner, B.J. Williams, and W.I. Notz, "Space-filling designs for computer experiments," In *The design and analysis of computer experiments*, Springer Series in Statistics, pp. 121-161, Springer, New York, 2003.

[32] I. Steponavice, M. Shirazi-Manesh, R.J. Hyndman, K Smith-Miles, and L. Villanova, "On sampling methods for costly multi-objective black-box optimization," In *Advances in Stochastic and Deterministic Global Optimization*, pp. 273-296, Springer, New York, 2016.

[33] W. Ponweiser, T. Wagner, D. Biermann, and M. Vincze, "Multiobjective optimization on a limited budget of evaluations using model-assisted S-metric selection," *Int. Conf. Parallel Problem Solving from Nature,* pp. 784-794, 2008.

[34] B. Beachkofski and R. Grandhi, "Improved distributed hypercube sampling," *American Institute of Aeronautics and Astronautics*, paper AIAA 2002-1274, 2002.

[35] Y.-C. Chen, S.-Y. Chen, and P. Hsu, "Dual-band slot dipole antenna fed by a coplanar waveguide," *IEEE Int. Symp. Ant. Prop.*, pp. 3589-3592, 2006.

[36] N.V. Queipo, R.T. Haftka, W. Shyy, T. Goel, R. Vaidynathan, and P.K. Tucker, "Surrogate-based analysis and optimization," *Progress in Aerospace Sciences*, vol. 41, no. 1, pp. 1-28, Jan. 2005.

Slawomir Koziel received the M.Sc. and Ph.D. degrees in electronic engineering from Gdansk University of Technology, Poland, in 1995 and 2000, respectively. He also received the M.Sc. degrees in theoretical physics and in mathematics, in 2000 and 2002, respectively, as well as the PhD in mathematics in 2003, from the University of Gdansk, Poland. He is currently a Professor with the School of Science and Engineering, Reykjavik University, Iceland. His research interests include CAD and modeling of microwave and antenna structures, simulation-driven design, surrogate-based optimization, space mapping, circuit theory, analog signal processing, evolutionary computation and numerical analysis.

Anna Pietrenko-Dabrowska received the M.Sc. and Ph.D. degrees in electronic engineering from Gdansk University of Technology, Poland, in 1998 and 2007, respectively. Currently, she is an Associate Professor with Gdansk University of Technology, Poland. Her research interests include simulation-driven design, design optimization, control theory, modeling of microwave and antenna structures, numerical analysis.