Data-Driven Model Based Design and Analysis of Antenna Structures

Selvakumar Ulaganathan[1,*], Slawomir Koziel[2], Adrian Bekasiewicz[2], Ivo Couckuyt[1], Eric Laermans[1], Tom Dhaene[1]

[1]Department of Information Technology, Ghent University, Ghent, Belgium
[2]Engineering Optimization & Modeling Center, Reykjavik University, Reykjavik, Iceland
[*]selvakumar.ulaganathan@ugent.be

Abstract: Data-driven models, or metamodels, offer an efficient way to mimic the behaviour of computation-intensive simulators. Subsequently, the usage of such computationally cheap metamodels is indispensable in the design of contemporary antenna structures where computation-intensive simulations are often performed in a large scale. Although metamodels offer sufficient flexibility and speed, they often suffer from an exponential growth of required training samples as the dimensionality of the problem increases. In order to alleviate this issue, a Gaussian process based approach, known as *Gradient Enhanced Kriging*, is proposed in this work to achieve cost-efficient modeling of antenna structures. The Gradient Enhanced Kriging approach incorporates adjoint-based sensitivity data in addition to function data obtained from electromagnetic simulations. The approach is illustrated using a dielectric resonator and an ultra-wideband antenna structures. The method demonstrates significant accuracy improvement with less number of training samples over the Ordinary Kriging approach which utilises function data only. The discussed technique has been favorably compared with Ordinary Kriging in terms of computational cost.

1. Introduction

Despite continual evolution of resources which make contemporary computing powerful, the mathematical and the computational complexities of physics-based simulation codes have grown phenomenally in recent years. For example, a full-wave electromagnetic (EM) analysis allows for accurate evaluation of antenna structures as well as for taking into account various environmental factors such as the presence of connectors, housing, installation fixtures, etc. EM simulation tools are therefore major design tools of contemporary antenna engineering [1]. However, high-fidelity EM analysis with fine mesh resolution of the structure at hand can be computationally very expensive. Consequently, carrying out design tasks that require repetitive runs of high-fidelity computer simulations, such as parametric optimization, sensitivity analysis, design space exploration, may be impractical or even prohibitive. In all such constantly evolving and technically challenging environments, the usage of metamodels (also known as surrogate models, replacement models, or response surface models) can significantly improve the efficiency of design processes while reducing the program cost (in terms of computational time and resources) considerably.

The goal of metamodeling is to mimic the behaviour of a computationally expensive simulator as accurately as possible by running a carefully designed set of high-fidelity simulations which can yield maximum information about the problem at hand with minimal computational effort [17].

Thus, a metamodeling approach is referred to as a technique that uses the sampled data to create a computationally cheap model which can predict the output of an expensive computer code at untried points in the design space. Various metamodeling techniques have been developed over the years which are applied to expedite computationally expensive tasks such as multi-disciplinary design and optimization of aircraft wing geometries, robustness- and/or reliability-based optimization of antenna structures, sensitivity analysis, etc. For example, implementation of low-cost antenna models is possible using various approximation techniques such as polynomial regression [2], radial basis function interpolation [3], Kriging [4, 5] support vector regression [6–9], fuzzy systems [10, 11], multidimensional Cauchy approximation [12], or artificial neural networks [13–16], etc. A common problem associated with most of the metamodeling approaches is the fact that they demand for a high model setup cost: in order to ensure usable accuracy a large number of training sample points is necessary, which quickly grows with the dimensionality of the design space (a problem often referred to as the curse of dimensionality) [2, 17, 18]. As a result, approximation techniques are mostly suitable for creating multiple-use library models but not so much to build surrogates for, say, one-time optimization of a specific antenna structure. An overview of the most commonly used metamodelling techniques is provided in [19], [20] and [21].

Physics-based surrogate modeling is another way of creating fast and usable-accuracy replacement models. Perhaps the most popular type of technique of this kind is space mapping (SM) [22–25], where the surrogate is constructed by means of suitable correction of an underlying low-fidelity (or so-called coarse) model. The bottleneck of space mapping in terms of antenna modeling is the lack of fast coarse models, because low-fidelity antenna representations are normally obtained through coarse-discretization EM simulations, the cost of which cannot be neglected. Another issue with SM is fixed number of extractable parameters which limits the model flexibility. This particular difficulty can be alleviated, to some extent by SM enhancement through fuzzy systems [26], radial-basis functions [23], or Kriging [27]. The problem of excessive number of training samples necessary to establish a reliable surrogate can be partially addressed by modeling methods that rely on appropriately extracted response features (e.g., shape-preserving response prediction [28], or feature-based modeling [29], however, these methods impose relatively strong assumptions on the response shapes of the structures under consideration so their applicability is limited to certain types of devices [29]. Another possibility is variable-fidelity modeling, where data of various degrees of accuracy are blended together using, e.g., co-Kriging [30, 31] or space mapping [32, 33]. Acquiring function (or response) data of various degrees of accuracy is more popular in areas where the computational cost of high-fidelity simulations is dominant such as full-wave EM analysis, computational fluid dynamics (CFD) and finite element (FE) analysis. For example, a set of EM or CFD methods with varying degrees of accuracy and complexity (circuit schematic simulations, Maxwell equations, panel equations, Euler equations, and Navier-Stokes equations) are frequently used to obtain data of varying fidelities. In addition, data of various levels of accuracy can be obtained by executing simulations on various mesh resolutions or with various convergence criteria. Various variable-fidelity metamodeling procedures have been reported in [34–37].

In this paper, we propose Gradient-Enhanced Kriging (GEK) approach that allows for low-cost modeling of antenna structures on a single PC machine. The cost reduction, in terms of the number of training sample points, is achieved by exploiting gradient information from adjoint simulations (currently, adjoint analyses are supported in a handful of commercial EM solvers such as CST or

HFSS). The cost of obtaining additional gradient data may be noticeable compared to the cost of obtaining a function response at the same sample point. For instance, various cases are recorded where k-dimensional gradients can be estimated with adjoint analyses which consume (approximately) 10% of the computational cost required for obtaining function values at a single sample point [40, 41]. We demonstrate that exploiting gradients of function values with respect to input variables in the Kriging model leads to significant reduction in the number of training data without degradation of the predictive power of the model. Our approach is illustrated using a dielectric resonator and an ultra-wideband antenna structures, and compared to conventional Kriging interpolation that utilizes the function data only. Despite the increased simulation cost of the EM model with sensitivity data, our method has been favorably compared with Ordinary Kriging in terms of CPU-time required for data acquisition.

The remaining part of this paper is organised as follows. The mathematical formulations of Kriging and GEK are elaborated in Section 2. Section 3 lists the test problems and the error metrics used to assess the metamodel accuracy. Test results are presented and discussed in Section 4 which is further followed by the conclusions in Section 5.

2. Mathematical formulations of Kriging and Gradient Enhanced Kriging

Ordinary Kriging (OK) prediction $\hat{y}(\mathbf{x}^*)$ of an arbitrary function $\mathbf{f}(\mathbf{x})$ at a prediction (or untried) point $\mathbf{x}^*$ can be expressed as a summation of a constant trend function (μ) and a realization of a stationary Gaussian random process:

$$\hat{y}(\mathbf{x}^*) = \hat{\mu} + \boldsymbol{\psi}^T \boldsymbol{\Psi}^{-1}(\mathbf{y} - \mathbf{t}\hat{\mu}), \tag{1}$$

where $\boldsymbol{\psi}$ is a vector of correlations between n_s sample data points $\mathbf{X} = \{\mathbf{x}^{(1)}, ..., \mathbf{x}^{(n_s)}\}^T$ and $\mathbf{x}^*$; $\boldsymbol{\Psi}$ is a n_s x n_s symmetric matrix of correlations between the sample data points; $\mathbf{y}$ is a vector of function values and $\mathbf{t}$ is a vector of ones. The covariance structure of the sample data is captured using Matérn $\frac{3}{2}$ correlation function [42, 43] which can be expressed as,

$$\psi(d') = (1 + \sqrt{3}c)exp\left(-\sqrt{3}c\right), \tag{2}$$

where $c = \sqrt{\sum_{p=1}^{k} \theta_p d_p'^2}$, $d' = |x_p^i - x_p^j|$ and k is the dimensionality of the problem. Gradient Enhanced Kriging is considered as a multi-data extension to Kriging. GEK incorporates cheaply available gradient data in addition to function data while building surrogate models. The mathematical formulation of GEK is same as Kriging except the following variations:

- correlation matrix in GEK ($\dot{\boldsymbol{\Psi}}$) becomes a $(k+1)n_s \times (k+1)n_s$ symmetric block matrix and now contains the correlations of both response and gradient observations between the sample data points. It can be expressed as,

$$\dot{\boldsymbol{\Psi}} = \left\{ \begin{array}{cccccc} \boldsymbol{\Psi} & \dfrac{\partial \boldsymbol{\Psi}}{\partial x_1^{(i)}} & \cdots & \dfrac{\partial \boldsymbol{\Psi}}{\partial x_v^{(i)}} & \cdots & \dfrac{\partial \boldsymbol{\Psi}}{\partial x_k^{(i)}} \\[2mm] \dfrac{\partial \boldsymbol{\Psi}}{\partial x_1^{(j)}} & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(i)} \partial x_1^{(j)}} & \cdots & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(i)} \partial x_v^{(j)}} & \cdots & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(i)} \partial x_k^{(j)}} \\[2mm] \vdots & \vdots & \ddots & \vdots & \cdots & \vdots \\[2mm] \dfrac{\partial \boldsymbol{\Psi}}{\partial x_u^{(j)}} & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(j)} \partial x_u^{(i)}} & \cdots & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_u^{(i)} \partial x_v^{(j)}} & \cdots & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_u^{(i)} \partial x_k^{(j)}} \\[2mm] \vdots & \vdots & \cdots & \vdots & \ddots & \vdots \\[2mm] \dfrac{\partial \boldsymbol{\Psi}}{\partial x_k^{(j)}} & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_1^{(j)} \partial x_k^{(i)}} & \cdots & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_v^{(j)} \partial x_k^{(i)}} & \cdots & \dfrac{\partial^2 \boldsymbol{\Psi}}{\partial x_k^{(i)} \partial x_k^{(j)}} \end{array} \right\}; \tag{3}$$

- correlation vector in GEK ($\dot{\boldsymbol{\psi}}$) contains correlation of both response and gradient observations between $\mathbf{X} = \{\mathbf{x}^{(1)}, ..., \mathbf{x}^{(n_s)}\}^T$ and $\mathbf{x}^*$, and can be expressed as,

$$\dot{\boldsymbol{\psi}} = \left(\boldsymbol{\psi}^T, \left(\frac{\partial \boldsymbol{\psi}}{\partial x_1} \right)^T, ..., \left(\frac{\partial \boldsymbol{\psi}}{\partial x_k} \right)^T \right)^T; \tag{4}$$

- vector $\mathbf{y}$ in GEK contains both the response and the gradient values, and can be expressed as,

$$\dot{\mathbf{y}} = \left(\mathbf{y}^T, \left(\frac{\partial \mathbf{y}}{\partial x_1} \right)^T, ..., \left(\frac{\partial \mathbf{y}}{\partial x_k} \right)^T \right)^T; \tag{5}$$

- vector $\mathbf{t}$ in GEK contains n_s ones followed by $k \times n_s$ additional zeros, and can be expressed as,

$$\mathbf{t} = \left(1_1, ... 1_{n_s}, 0_{n_s+1}, ..., 0_{(k+1)n_s} \right)^T. \tag{6}$$

Hence, Equation 1 for GEK becomes,

$$\hat{y}(\mathbf{x}^*) = \hat{\mu} + \dot{\boldsymbol{\psi}}^T \dot{\boldsymbol{\Psi}}^{-1} (\dot{\mathbf{y}} - \mathbf{f}\hat{\mu}). \tag{7}$$

The trend function is calculated using least squares method as,

$$\hat{\mu} = (\mathbf{t}^T \dot{\boldsymbol{\Psi}}^{-1} \mathbf{t})^{-1} \mathbf{t}^T \dot{\boldsymbol{\Psi}}^{-1} \dot{\mathbf{y}}. \tag{8}$$

The maximum likelihood estimate of the hyper-parameters ($\boldsymbol{\theta}$) of the correlation function is estimated by maximizing the concentrated ln-likelihood function

$$\phi = \frac{-(k+1)n_s \; ln(\hat{\sigma}^2) - ln|\dot{\boldsymbol{\Psi}}|}{2}, \tag{9}$$

where

$$\hat{\sigma}^2 = \frac{(\dot{\mathbf{y}} - \mathbf{t}\hat{\mu})^T \dot{\boldsymbol{\Psi}}^{-1} (\dot{\mathbf{y}} - \mathbf{t}\hat{\mu})}{(k+1)n_s} \tag{10}$$

is the estimated GEK variance.

3. Surrogate Modeling of Antenna Structures

Let $\mathbf{f}(\mathbf{x})$ be an EM-simulated antenna structure, where $\mathbf{x}$ is a vector of geometry parameters, whereas $\mathbf{f}$ is a vector-valued response, e.g., reflection coefficient versus frequency. In practice, we assume that $\mathbf{f}(\mathbf{x}) = [f(\mathbf{x}, \omega_1) f(\mathbf{x}, \omega_2)...f(\mathbf{x}, \omega_m)]^T$, i.e., the model is evaluated at a discrete set of frequencies $\omega_k, k = 1, ..., m$. The high-fidelity EM simulations of the antenna structure are computationally very expensive. The goal of the surrogate modeling process is to construct a cheaper representation of $\mathbf{f}$ that is valid in the interval $\mathbf{l} \le \mathbf{x} \le \mathbf{u}$, where $\mathbf{l}$ and $\mathbf{u}$ are the lower and the upper bounds for the geometry parameters, respectively.

3.1. Suspended Dielectric Resonator Antenna

Consider a suspended dielectric resonator antenna (DRA) shown in Figure 1 [37]. The structure is composed of a dielectric resonator ($\varepsilon_r = 10$ and $\tan\delta = 0.0001$) suspended over the ground plane on the two Teflon slabs with relative permittivity and loss tangent of 2.1 and 0.001, respectively. The antenna is fed through the ground plane slot. The substrate material is Rogers RO4003 ($\varepsilon_r = 3.38$, $\tan\delta = 0.0027$ and $h = 0.5$ mm). The DRA is covered by polycarbonate housing ($\varepsilon_r = 2.8$ and $\tan\delta = 0.01$).
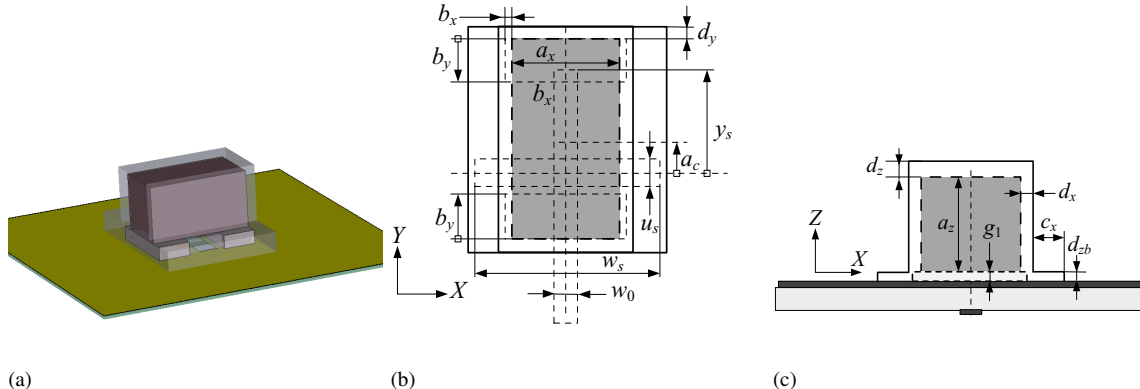
**Fig. 1.** *The suspended dielectric resonator antenna.*
(a) 3D design visualization
(b) Top view
(c) Front view

3.1.1. Case I: The vector of antenna parameters is: $\mathbf{x} = [a_c\ u_s\ w_s\ y_s\ g_1]^T$, whereas the dimensions $d_x = d_y = d_z = 1$, $a_x = 8.06$, $a_y = 14.24$, $a_z = 9.25$, $b_y = 5.87$, $d_{zb} = 2$, $b_x = 2$, and $c_x = 6.5$ remain fixed. The unit for all parameters is mm. The design variable space for the antenna parameters is defined by the following lower and upper bounds: $\mathbf{l} = [0.5\ 1.4\ 8\ 2.4\ 1.2]^T$ and $\mathbf{u} = [0.5\ 2.1\ 12\ 3.6\ 1.8]^T$. The EM model of the structure is implemented in Ansys HFSS and simulated using its frequency domain solver [38]. The antenna model consists of about 8,500 tetrahedral mesh cells and its average simulation time on a on a dual Xeon E5540 machine with 6 GB RAM is 454 s. For comparison purposes, the EM model of the antenna that does not provide sensitivity information has been also prepared and discretized using the same mesh setup. Its average evaluation time is 141 s.

3.1.2. Case II: Here, the considered DRA structure is described by a nine variable vector: $\mathbf{x} = [a_x\ a_y\ a_z\ a_c\ u_s\ w_s\ y_s\ g_1\ b_y]^T$. Parameters $d_x = d_y = d_z = 1$, $d_{zb} = 2$, $b_x = 2$, and $c_x = 6.5$ are fixed. The design bounds of the parameter space are $\mathbf{l} = [7\ 14\ 9\ -0.5\ 2.3\ 8\ 2.4\ 3.0\ 0.35]^T$ and $\mathbf{u} = [8\ 15\ 10\ 0.5\ 3.0\ 12\ 3.6\ 3.6\ 0.4]^T$. The EM models of the antenna (with and without derivative information) are the same as for Case I. The average simulation time of the model that provides sensitivity information is 459 s. For the second model, the simulation time is the same as in Case I.

3.2. Ultra-wideband Antenna

Our second design example is a planar ultra-wideband (UWB) monopole antenna shown in Fig. 2 [1]. The structure consists of two trapezoids and a rectangle stacked together into a radiator. It is fed through 50 Ohm microstrip line. The antenna is implemented on a 0.762 mm thick Taconic RF-35 dielectric substrate ($\varepsilon_r = 3.5$ and $\tan\delta = 0.0018$). The geometry of the structure is represented by a vector $\mathbf{x} = [a_2\ a_3\ b_1\ b_2\ b_3]^T$, whereas parameters $a_1 = 9$, $l = 8.5$, $d = 0.6$, $w = 20$, $w_0 = 1.7$, remain fixed. The unit for all dimensions is mm. The design variable space is defined by the following lower and upper bounds: $\mathbf{l} = [17\ 17\ 4\ 4\ 4]^T$ and $\mathbf{u} = [26\ 26\ 10\ 10\ 10]^T$. The antenna is implemented in Ansys HFSS. The simulation time of the EM model that provides sensitivity information is 592 s, whereas the evaluation time of the model without sensitivity data is 172s. Both consist of about 15,000 tetrahedral mesh cells.

(a)   (b)
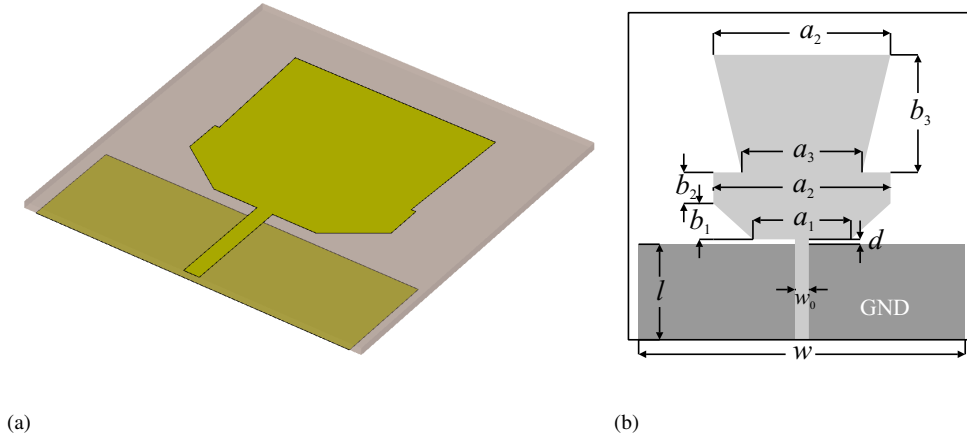
Fig. 2. The ultra-wideband Antenna.
(a) 3D visualization
(b) Geometry with highlighted parameters

3.3. Error Metrics

Two error metrics, Normalized Root Mean Square Error (NRMSE) and Relative Average Absolute Error (RAAE), are used to assess the accuracy of the surrogate models on a validation data set of n_p uniformly distributed pseudorandom points. The error metrics are expressed as,

$$NRMSE = \frac{\sqrt{\frac{\sum_{i=1}^{n_p}\left(y_t^i - \hat{y}^i\right)^2}{n_p}}}{max(y_t) - min(y_t)} \tag{11}$$

and

$$RAAE = \frac{\sum_{i=1}^{n_p} \left(y_t^i - \hat{y}^i\right)^2}{std(\boldsymbol{y}_t)n_p},\qquad(12)$$

where $\mathbf{y}_t$ is the vector of true response values, $\hat{\mathbf{y}}$ is the vector of predicted response values and std stands for standard deviation. NRMSE and RAAE show the overall surrogate modeling accuracy and their values approach zero as the overall surrogate model accuracy increases.

4. Numerical Results

Two different methods were considered to model the reflection coefficient of various types of antennas utilized in this work: OK method which incorporates the function values of the reflection coefficient only and the GEK method which incorporates the additional adjoint-gradient values of the reflection coefficient. Since the reflection coefficient results in a vector of values at any given sample point, multiple OK and GEK models were constructed to model the whole reflection characteristic $|\mathbf{S}_{11}|$ versus frequency. Figures 3 through 5 compare the actual reflection characteristic $|\mathbf{S}_{11}|$ versus frequency to its OK and GEK approximations. GEK requires only half (or even less) of the training sample points demanded by OK to accurately approximate the actual $|\mathbf{S}_{11}|$ curve (Figures 3-5 and Table 1). GEK benefits from the additional gradient data which allow the model to accurately capture the covariance structure of the training data [35, 36]. This is achieved by the fact that the GEK model is forced to interpolate both the function and the gradient values at the training samples whereas OK interpolates the function values only. Moreover, it is also important to note that GEK models exhibit an accuracy level which is never reached by OK models for the given number of training sample points (Figure 6). Additionally, the fact that OK approximations which are trained with more than 50% (or even more) additional training samples do not overlay the accurate GEK approximations confirms the same (Figures 3-5).

Table 1 shows the percentage of reduction in training samples achieved by GEK over OK while providing equally (or more) accurate approximations. It can be observed that GEK is more efficient when the number of training samples is less. This particular feature of GEK can be very useful to use GEK-based surrogate models to expedite routine activities such as design space exploration, sensitivity analysis, robustness based optimization, etc. Moreover, although the percentage of reduction in training samples gradually reduces as the training data increase, the results in Table 1 confirm the earlier perception that the GEK models achieve more accurate approximations than OK models for the given number of training samples. Similar facts can be extracted from the results of Table 3 which show the evolution of rate of accuracy improvement achieved by GEK over OK for a range of number of training samples. The reason for gradient data to become less advantageous at higher training data is the intuitive fact that the function value is more informative than a gradient value. The effect of incorporating partial set of gradients (only in three geometrical dimensions of the suspended dielectric resonator antenna (Case I) - a_c u_s w_s) in Figure 6 confirms the same conjecture. The advantage of incorporating additional gradient data in GEK is overshadowed by the incorporation of more function data in OK (for 120 and more samples in Figure 6 (d)).

As it can be seen from Sections 3.1 and 3.2 the evaluation cost of the EM antenna model that provides sensitivity information is noticeably higher compared to the simulation cost of the model that returns only the function response. On the other hand, the results collected in Table 3 indicate

7

that the GEK model with similar RAAE can be constructed using a limited number of samples compared to OK. For the considered design cases, computational savings from construction of GEK model are from 20% to over 40% with respect to OK. Consequently, utilization of the EM antenna model which provides sensitivity information is justified despite of its increased evaluation cost with respect to the model that only returns the reflection response. It should be noted that the memory consumption of HFSS solver is similar regardless the selected EM model (with or without sensitivity data). For the DRA structure (cf. Section 3.1) typical memory usage is below 280 MB, whereas evaluation of the UWB antenna (cf. Section 3.2) requires about 350 MB of RAM.
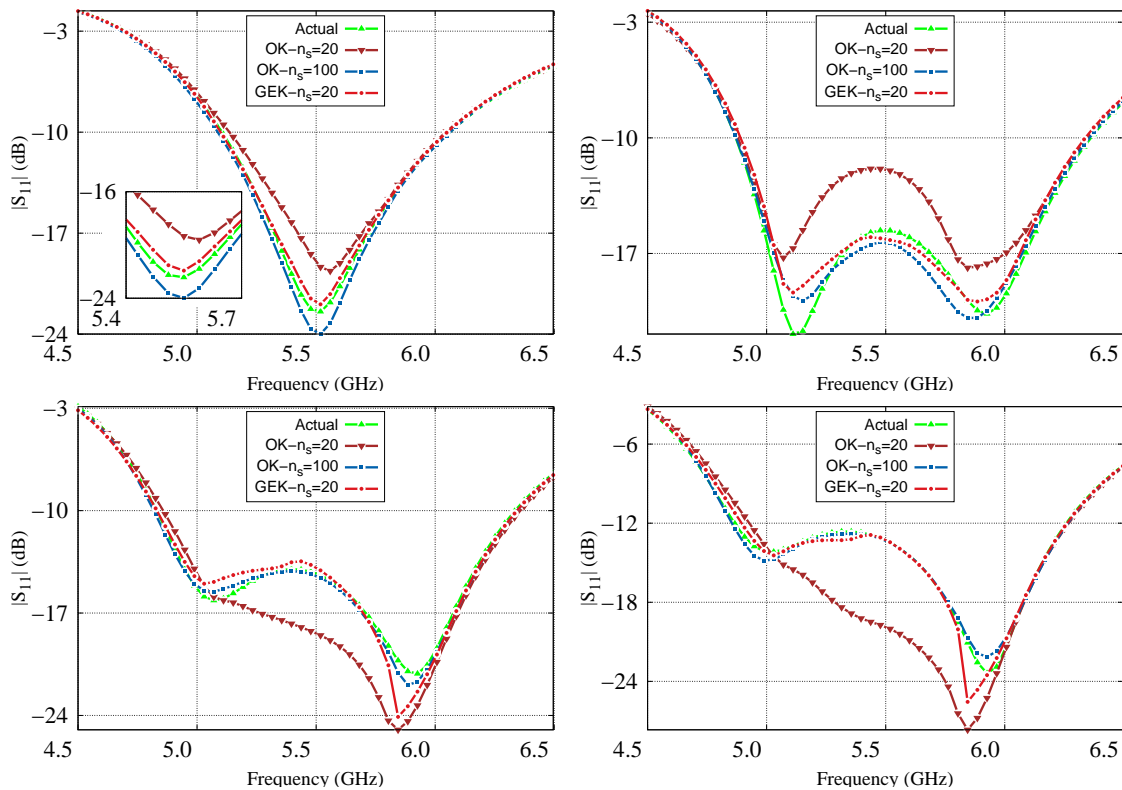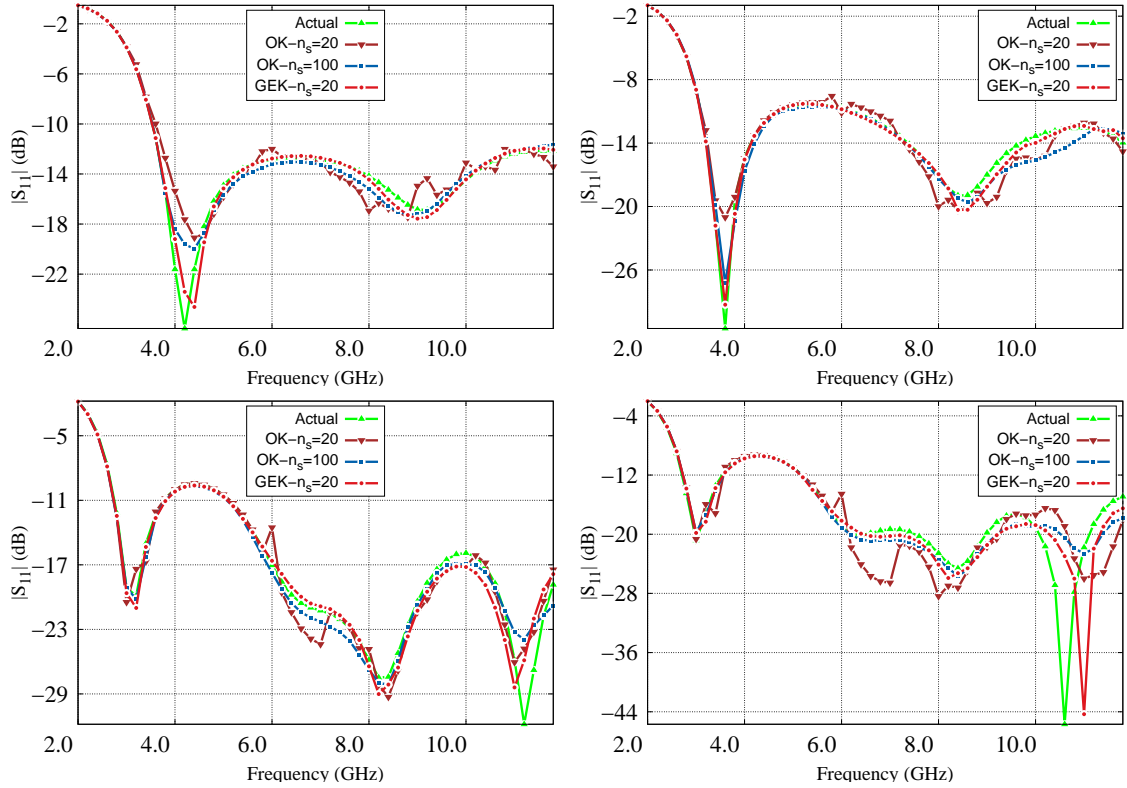

Fig. 3. *DRA (Case I): Prediction of* $\mathbf{f}(\mathbf{x})$ *at random test points. OK incorporates only the response data whereas GEK incorporates gradient data in addition to the response data.*
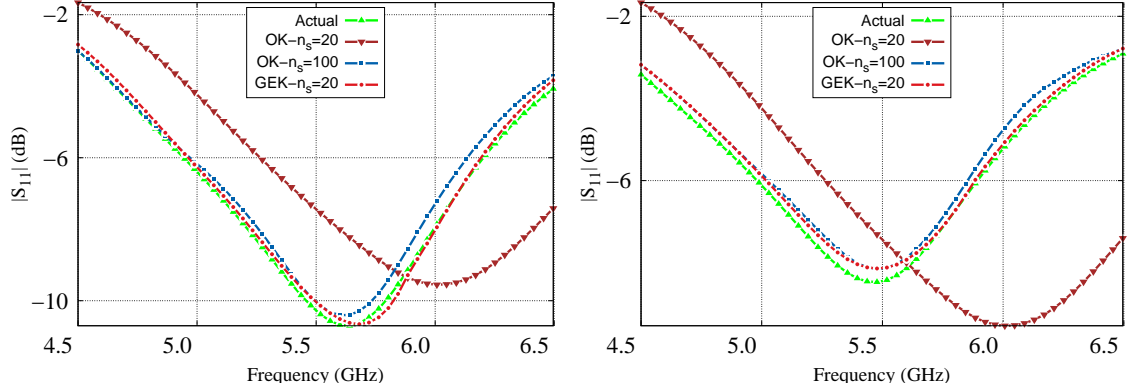
Fig. 4. *DRA (Case II): Prediction of $\mathbf{f}(\mathbf{x})$ at random test points. OK incorporates only the response data whereas GEK incorporates gradient data in addition to the response data.*
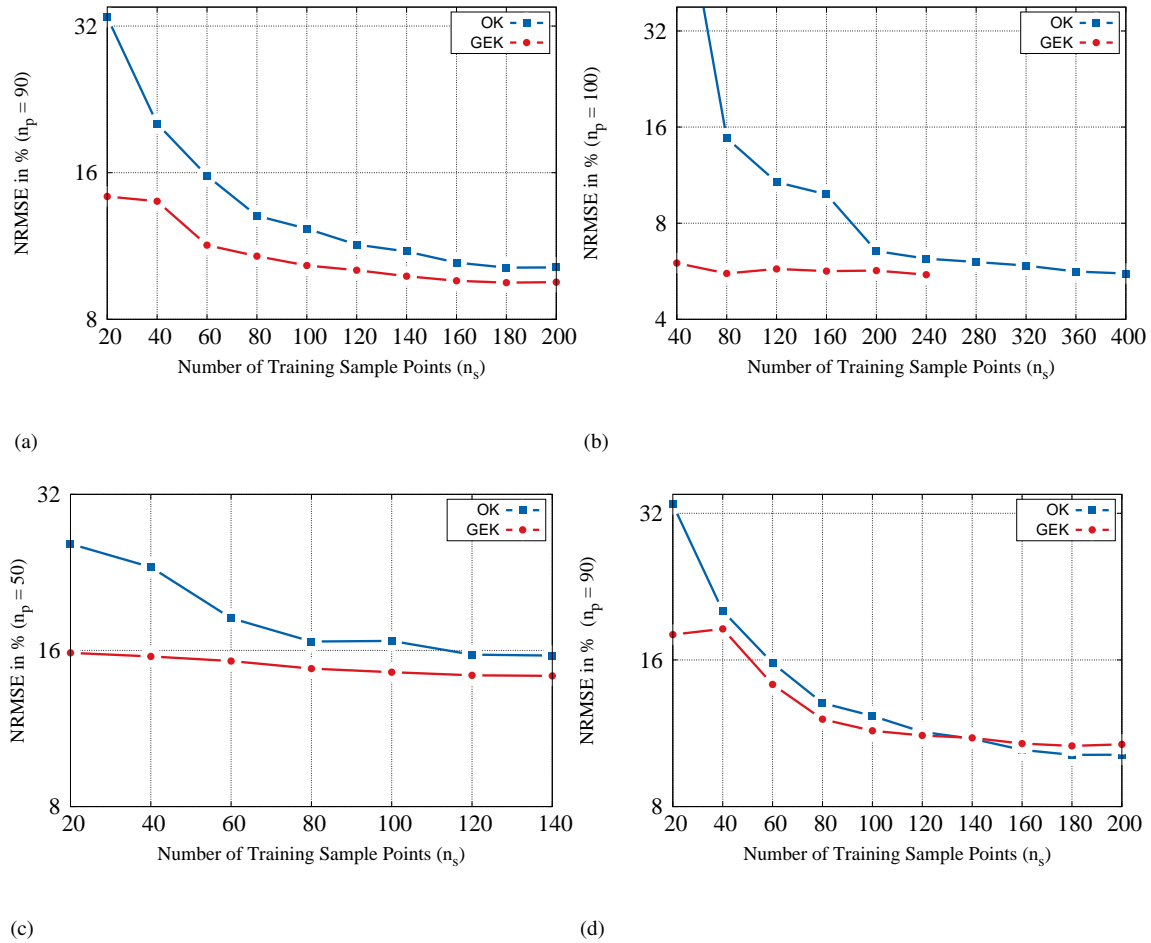

Fig. 5. *UWB: Prediction of $\mathbf{f}(\mathbf{x})$ at random test points. OK incorporates only the response data whereas GEK incorporates gradient data in addition to the response data.*

Fig. 6. *Evolution of NRMSE on a validation data set of n_p test points for a varying number of training sample points. OK incorporates only the response data wheres GEK incorporates gradient data in addition to the response data.*
(a) DRA (Case I)
(b) DRA (Case II)
(c) UWB
(d) DRA (Case I) - GEK incorporates additional gradient data in the direction of three geometric variables only

Table 1 Reduction in number of training samples with GEK based on normalized root mean square error (NRMSE) on a validation data set of n_p points.

Example	Model	Number of Training Samples	NRMSE (%)	Percentage of Reduction in Training Samples
	OK	200	10.2	—
DRA	GEK	60	11.3	70%
(Case I)	GEK	80	10.7	60%
($n_p = 90$)	GEK	100	10.3	50%
	OK	400	5.56	—
DRA	GEK	40	6.00	90%
(Case II)	GEK	80	5.56	80%
($n_p = 100$)	GEK	120	5.75	70%
	OK	140	15.7	—
UWB	GEK	60	15.2	57%
($n_p = 50$)	GEK	80	14.8	43%
	GEK	100	14.5	26%

Table 2 Efficiency of GEK over OK based on relative average absolute error (RAAE) on a validation data set of n_p points.

Example	Number of Training Samples	OK		GEK	
		RAAE (%)	% of Accuracy Improvement of OK	RAAE (%)	% of Accuracy Improvement of GEK over OK
	20	13.2	—	5.9	55%
	40	8.4	37%	5.8	56%
DRA	60	6.8	48%	4.8	64%
(Case I)	80	5.9	55%	4.7	65%
($n_p = 90$)	100	5.6	57%	4.5	66%
	40	53.8	—	2.7	95%
	80	6.6	88%	2.5	95.3%
DRA	120	5.0	90%	2.6	95.2%
(Case II)	160	4.6	91%	2.5	95.3%
($n_p = 100$)	200	2.8	95%	2.4	95.2%
	20	13.2	—	7.5	43%
	40	11.3	14%	7.2	45%
UWB	60	8.5	35%	6.5	49%
($n_p = 50$)	80	8.4	36%	6.6	50%
	100	8.0	40%	6.4	51%

Table 3 Comparison of GEK and OK in terms of model preparation cost.

Example	Model	Number of Samples	RAAE (%)	Cost$^{\#}$ [h]	Cost reduction
DRA	OK	80	5.9	3.13	–
(Case I)	GEK	20	5.9	2.52	20%
DRA	OK	200	2.8	7.83	–
(Case II)	GEK	40	2.7	5.10	35%
UWB	OK	120	7.6	5.73	–
	GEK	20	7.5	3.28	43%

$^{\#}$ The cost of model identification and validation is excluded

5. Conclusions

In this work, a cost-efficient surrogate modeling of antenna structures on a single PC machine using Gradient-Enhanced Kriging has been demonstrated. This approach is applied to various antenna structures. Utilization of additional sensitivity data significantly reduces the number of training samples required to set up a fast and reliable surrogate model which can be used in routine activities such as design space exploration. Additionally, Gradient-Enhanced Kriging results in highly accurate approximations at any given amount of sample data than the conventional approach that only relies on the function data. The additional sensitivity data acquisition cost can be significantly reduced by taking advantage of adjoint sensitivities that allow for obtaining the gradients at extra computational cost. The Gradient-Enhanced Kriging method has been favorably compared with Ordinary Kriging in terms of the CPU-time required for data acquisition. Our further work will focus on application of the method for antenna arrays.

6. Acknowledgements

7. References

[1] Koziel S., Bekasiewicz, A.: 'Fast EM-driven size reduction of antenna structures by means of adjoint sensitivities and trust regions', IEEE Ant. Wireless Prop. Lett., 2015, **14**, pp. 1681–1684

[2] Forrester, A.I.J., Keane, A.J.: 'Recent advances in surrogate-based optimization', Prog. Aerospace Sci., 2009, **1-3**, (45), pp. 50–79

[3] Buhmann, M.D., Ablowitz, M.J.: 'Radial Basis Functions: Theory and Implementations', Cambridge University, 2003

[4] Couckuyt, I., Dhaene, T., Demeester, P.: 'ooDACE toolbox: A flexible object-oriented kriging implementation', J. Machine Learning Research, 2014, (15), pp. 31833186

[5] Gorissen, D., Couckuyt, I., Demeester, P., Dhaene, T., Crombecq, K.: 'A surrogate modeling and adaptive sampling toolbox for computer based design', J. Machine Learning Research, 2010, (11), pp. 2051–2055.

[6] Smola, A.J., Scholkopf, B.: 'A tutorial on support vector regression', Statistics and Computing, 2004, **3**,(14), pp. 199–222

[7] Meng, J., Xia, L.: 'Support-vector regression model for millimeter wave transition', Int. J. Infrared and Milimeter Waves, 2007, **5**, (28), pp. 413–421

[8] Xia, L., Meng, J. Xu, R., Yan, B., Guo, Y.: 'Modeling of 3-D vertical interconnect using support vector machine regression', IEEE Microwave Wireless Comp. Lett., 2006, **12**, (16), pp. 639–641

[9] Martinez-Ramon, M., Christodoulou, C.: 'Support vector machines for antenna array processing and electromagnetics', Synthesis Lectures on Computational Electromagnetics, 2006, **1**, (1)

[10] Miraftab, V., Mansour, R.R.: 'EM-based microwave circuit design using fuzzy logic techniques', IEE Microwaves, Ant. Prop., 2006, **6**, (153), pp. 495–501

[11] Zhai, J., Zhou, J., Zhang, L., Hong, W.: 'Behavioral modeling of power amplifiers with dynamic fuzzy neural networks', IEEE Microwave Wireless Comp. Lett., 2010, **9**, (20), pp. 528–530

[12] Shaker, G.S.A., Bakr, M.H., Sangary, N., Safavi-Naeini, S.: 'Accelerated antenna design methodology exploiting parameterized Cauchy models', 2009, PIER-99, pp. 279–309

[13] Rayas-Sanchez, J.E.: 'EM-based optimization of microwave circuits using artificial neural networks: the state-of-the-art', IEEE Trans. Microwave Theory Tech., 2004, **1**, (52), pp. 420–435

[14] Kabir, H., Wang, Y., Yu, M., Zhang, Q.J.: 'Neural network inverse modeling and applications to microwave filter design', IEEE Trans. Microwave Theory Tech., 2008, **4**, (56), pp. 867–879

[15] Cao, Y., Chen, X., Wang, G.: 'Dynamic behavioral modeling of nonlinear microwave devices using real-time recurrent neural network', IEEE Trans. Electron Devices, 2009, **5**, (56), pp. 1020–1026

[16] Kabir, H., Wang, Y., Yu, M., Zhang, Q.J.: 'High-dimensional neural-network technique and applications to microwave filter modeling', IEEE Trans. Microwave Theory Tech., 2010, **1**, (58), pp. 145–156

[17] Koziel, S., Bekasiewicz, A., Zieniutycz, W.: 'Expedited EM-Driven Multiobjective Antenna Design in Highly-Dimensional Parameter Spaces', IEEE Ant. Wireless Propag. Lett., 2014, (13), pp. 631–634

[18] Ulaganathan, S., Couckuyt, I., Dhaene, T., Degroote, J., Laermans, E.: 'High Dimensional Kriging Metamodelling Utilising Gradient Information', Journal of Applied Mathematical Modelling, 2015, DOI: 10.1016/j.apm.2015.12.033

[19] Jin, R., Chen, W., Simpson, T.W.: 'Comparative studies of metamodeling techniques under multiple modeling criteria', 'Structural and Multidisciplinary Optimization', 2000, **23**, pp. 1–13

[20] Wang, G.G., Shan, S.: 'Review of metamodeling techniques in support of engineering design optimization', Journal of Mechanical Design, 2006, **129**, (4), pp. 370–380

[21] Simpson, T., Poplinski, J., Koch, P.N., Allen, J.: 'Metamodels for computer-based engineering design: survey and recommendations'. 2001, Eng. Comput., **17**, (2), pp. 129–150

[22] Rayas-Sanchez, E., Gutierrez-Ayala, V.: 'EM-based Monte Carlo analysis and yield prediction of microwave circuits using linear-input neural-output space mapping', IEEE Trans. Microwave Theory Tech., 2006, **12**, (54), pp. 4528–4537

[23] Koziel, S., Bandler, J.W.: 'Recent advances in space-mapping-based modeling of microwave devices', Int. J. Numerical Modelling, 2010, **6**, (23), pp. 425–446

[24] Zhang, L., Zhang, Q.J., Wood, J.: 'Statistical neuro-space mapping technique for large-signal modeling of nonlinear devices', IEEE Trans. Microwave Theory Tech., 2011, **1**, (56), pp. 2453–2467

[25] Bandler, J.W., Cheng, Q.S., Koziel, S.: 'Simplified space mapping approach to enhancement of microwave device models', Int. J. RF and Microwave Computer-Aided Eng., 2006, **5**, (16), pp. 518–535

[26] Koziel, S., Bandler. J.W.:, 'A space-mapping approach to microwave device modeling exploiting fuzzy systems', IEEE Trans. Microwave Theory Tech., 2007, **12**, (55), pp. 2539–2547

[27] Koziel, S., Bandler, J.W.: 'Accurate modeling of microwave devices using space mapping and Kriging', Int. Review Prog. Applied Comp. Electromagnetics, 2010, pp. 902–907

[28] Koziel, S., Cheng, Q.S., Bandler, J.W.: 'Fast EM modeling exploiting shape-preserving response prediction and space mapping', to appear, IEEE Trans. Microwave Theory Tech., 2014

[29] Koziel, S., Bandler, J.W.:, 'Rapid yield estimation and optimization of microwave structures exploiting feature-based statistical analysis', IEEE Trans. Microwave Theory Tech., 2015, **1**, (63), pp. 107–114

[30] Koziel, S., Ogurtsov, S., Couckuyt, I., Dhaene, T.: 'Variable-Fidelity Electromagnetic Simulations and Co-Kriging for Accurate Modeling of Antennas', IEEE Trans. Ant. Prop., 2013, **3**, (61), pp. 1301–1308

[31] Koziel, S., Bekasiewicz, A., Couckuyt, I., Dhaene, T.: 'Efficient multi-objective simulation-driven antenna design using co-kriging', IEEE Trans. Antennas Prop., 2014, **1**, (62), pp. 5900–5905

[32] Koziel, S., Bekasiewicz. A.: 'Fast multi-objective optimization of narrow-band antennas using RSA models and design space reduction', IEEE Antennas and Wireless Prop. Lett., 2015, (14), pp. 450–453

[33] Koziel, S., Bekasiewicz, A.,: 'Strategies for computationally feasible multi-objective simulation-driven design of compact RF/microwave components', Eng. Comp., 2015.

[34] Jacobs, J.P., Koziel, S.: 'Two-stage framework for efficient Gaussian process modeling of antenna input characteristics', IEEE Trans. Antennas Prop., 2014, **2**, (62), pp. 706–713

[35] Ulaganathan, S., Couckuyt, I., Ferranti, F., Laermans, E., Dhaene, T.: 'Performance study of multi-fidelity gradient enhanced kriging', Structural and Multidisciplinary Optimization, 2014, DOI: 10.1007/s00158- 014-1192-x

[36] Ulaganathan, S., Couckuyt, I., Dhaene, T., Degroote, J., Laermans, E.: 'Performance study of gradient enhanced kriging', Engineering with computers, 2014, DOI: 10.1007/s00366-015-0397-y

[37] Koziel, S., Bekasiewicz, A.: 'Variable-fidelity optimization of antennas using adjoint sensitivities', Loughborough Ant. Prop. Conf., 2014, pp. 412–415

[38] Ansys HFSS: ver. '14.0 (2012), ANSYS, Inc.', Southpointe 275 Technology Drive, Canonsburg, PA 15317

[39] CST Microwave Studio: 'ver. 2013, CST AG', Bad Nauheimer Str. 19, D- 64289 Darmstadt, Germany, http://www.cst.com, 2013

[40] Laurenceau, J., Meaux, M., Montagnac, M., Sagaut, P.: 'Comparison of gradient-based and gradient-enhanced response-surface-based optimizers', Am. Inst. Aeronaut. Astronaut. J., 2010, **48**, (5), pp.981–994

[41] Chung, H.S., Alonso, J.J.: 'Using gradients to construct Cokriging approximation models for high-dimensional design optimization problems'. In: Problems, 40th AIAA aerospace sciences meeting and exhibit, AIAA, 2002, Reno, NV, pp 2002–0317

[42] Rasmussen, C.E., Williams, C.K.I.: 'Gaussian processes for machine learning', The MIT Press, 2006, Cambridge, MA, USA

[43] Stein, M.L.: 'Interpolation of Spatial Data: Some Theory for Kriging', Springer, 1999, New York