Surrogate-Based Infill Optimization Applied to Electromagnetic Problems

I. Couckuyt[†], F. Declercq[∗], T. Dhaene[†], H. Rogier[∗], L. Knockaert[∗]

April 15, 2010

[†]Department of Information Technology (INTEC), Ghent University - IBBT, Gaston Crommenlaan 8, 9050 Ghent, Belgium

[∗]Department of Information Technology (INTEC), Ghent University, Sint Pietersnieuwstraat 41, 9000 Ghent, Belgium

Abstract

The increasing use of expensive computer simulations in engineering places a serious computational burden on associated optimization problems. Surrogate-based optimization becomes standard practice in analyzing such expensive black-box problems. This paper discusses several approaches that use surrogate models for optimization and highlights one sequential design approach in particular, namely, expected improvement. The expected improvement approach is demonstrated on two electromagnetic problems, namely, a microwave filter and a textile antenna.

Keywords: surrogate model, surrogate-based optimization, expected improvement, kriging, inverse problem

1 Introduction

For many problems in science and engineering it is impractical to perform experiments on the physical world directly. Instead, complex, physics-based simulation codes are used to run experiments on computer hardware. While allowing scientists more flexibility to study phenomena under controlled conditions, computer experiments require a substantial investment of computation time. This is especially evident for routine tasks such as optimization, sensitivity analysis and design space exploration [1]. Regardless of the rapid advances in High Performance Computing (HPC) and multi-core architectures, it is rarely feasible to explore the complete design space using high-fidelity computer simulations. As a result researchers have turned to various approximation methods that mimic the behavior of the simulation model as closely as possible while being computationally cheap(er) to evaluate.

This work concentrates on the use of data-driven approximations using compact surrogate models (otherwise known as metamodels or response surface models). Examples of surrogate models include: rational functions, Gaussian Process (GP) models, and Support Vector Machines (SVM). It is crucial to stress the distinction between local and global surrogate modeling. With the latter, an approximation model of the output behavior of the simulator is built over the entire design space. By contrast, local surrogate models, often used in trust-region optimization frameworks, approximate only a small part of the design space and are discarded after use.

Most often, surrogate models are used to solve so-called forward problems. The practitioner is interested in the performance characteristics of a complex system, given the input parameters. The surrogate models create a mapping between the design space (input parameters) and the performance space (output parameters). In contrast, the focus of the reverse (inverse) problem is on exploring the design space. Hypothetically, a surrogate model could be created that maps the output parameters to the input parameters (as opposite to forward modeling) of the complex system over the entire design space. However, many inverse problems are typically ill-posed. Considering Hadamard's definition of ill-posedness [2], the two outstanding problems hampering the creation of a full inverse surrogate model are non-uniqueness and instability. A good overview of the associated intricacies is presented by Barton in [3]. For all the above reasons, the inverse problem is often reduced to the task of finding an input parameter combination for a certain output characteristic. Still, it is possible that,

1. no such input parameter combination exists

2. more than one input parameter combination satisfies the given output characteristic

A popular solution is to convert the inverse problem to a forward optimization problem, as is done in this paper.

The construction of highly efficient surrogate models is an entire research domain in itself. In order to arrive at an acceptable model, numerous problems and design choices must be overcome (what data collection strategy to use, what model type is most applicable, how should the model parameters be tuned, which variables are relevant, how to integrate problem-specific knowledge, etc.).

This paper describes a popular optimization method for expensive black-box simulators based on kriging surrogate models, namely, expected improvement (EI) [4]. We provide a freely available implementation of the EI approach as a data collection (= sequential design) strategy in a flexible research platform for surrogate modeling, the **SU**rrogate **MO**deling (SUMO) Toolbox[1] [5]. The SUMO Toolbox is used to solve two complex problems both originating from Electromagnetics (EM). Previously, kriging surrogate models have been used for EM device optimization by creating a global accurate kriging surrogate model [6]. Afterwards, the computational cheap surrogate model is optimized instead of the

---

[1]The SUMO Toolbox can be downloaded from: `http://sumo.intec.ugent.be`. An open source license will be available soon.

expensive simulation. Although the EM device has been successful optimized, creating such one-shot kriging surrogate models does not result in the most efficient use of expensive function evaluations. Siah et al. [7] try to minimize the number of function evaluations by applying the EI approach on two EM applications.

Section 2 provides related work of Surrogate-Based Optimization (SBO), including the expected improvement (EI) function. In Section 3, this EI approach is used to design and optimize an inter-digital filter (the forward problem). In Section 4, the material properties of a textile antenna are identified using EI (the inverse problem).

2 Surrogate-Based Optimization (SBO)

2.1 Introduction

SBO techniques are concerned with accelerating the optimization of expensive simulation problems. To speedup the optimization process other existing optimization algorithms have been adapted to minimize the number of function evaluations and to utilize parallel computing. A good overview is given in [8]. These existing optimization methods can still be significantly improved by taking advantage of surrogate models. The extra information provided by the surrogate models helps avoiding local optima and efficiently guides the search to the global optimum. Various directions have been undertaken to incorporate surrogate models in the optimization process.

In the context of evolutionary optimization surrogate models are used to provide a rough approximation to guide the global search, or a local accurate surrogate model is used to speedup the local search step, or a combination of both [9, 10]. For instance, Zhou et al. [11] apply a data parallel Gaussian Process for the global approximation and a (simple) Radial Basis Function (RBF) model for the local search. Lim et al. [12] benchmark different local surrogate modeling techniques (quadratic polynomials, GP, RBF and extreme learning machine neural networks) including the use of (fixed) ensembles, in combination with evolutionary computation.

An important concept in global optimization is trust regions, introduced in surrogate modeling by [13]. Trust region-frameworks manage local surrogate models throughout the design space. A set of mathematics based [13] or pure heuristic [14] rules determines the size and location of the surrogate model. While trust region-frameworks are widely used in large scale optimization problems they have the disadvantage of sometimes overlooking the global optimum, as only a small part of the design space is approximated by the local surrogate model. On the other hand, by approximating only one local part of the design space at a time, it is possible to optimize very complex systems exhibiting non-linear behavior.

If a number of simulation models are available, each with varying accuracy (= fidelity), multi-fidelity methods [15], also known as variable-fidelity methods, can be used to solve more complex problems. There

are several approaches to exploit multi-fidelity models. Without loss of generality, we can assume that only two simulation models are available, a low-fidelity and a high-fidelity model. An additive or multiplicative scaling factor [16] can be introduced based on a single (or a few) data point(s). The underlying idea is that these scaling factors correct the output of the low-fidelity model to agree with the output of the high-fidelity model near the vicinity of these points (zero-order scaling). One may also use higher order scaling strategies, e.g., where the derivatives of the low-fidelity model are also modified to agree with the high-fidelity model. A more complex combination of both approaches is also possible; for instance Eldred et al. [17] propose to write the low-fidelity model as a weighted combination of additive and multiplicative scaling factors. Alternatively, space mapping methods can be utilized. Instead of approximating the output space directly, space mapping [18, 19, 20, 21] maps the input space of a low-fidelity model to the input space of the high-fidelity model, basically employing an input correction between multiple fidelity models causing the optima to align in the design space. Recently it has been proposed to apply a similar technique for output correction, denoted by output space mapping [22, 23]. Moreover, combinations of input and output space mapping are also possible. In addition, recently [24] proposed a new variant called manifold mapping which can be seen as a generalization to output space mapping. Actually, the co-kriging surrogate model [25, 26] is inherently a multi-fidelity surrogate model that essentially applies a correction to the output of the low-fidelity model. Multi-fidelity optimization methods, such as space mapping, are able to significantly improve on other methods by reducing computation time and/or generating better optimal designs. However, computational cheap low-fidelity models may not always be available to the practitioner. In that case, one may turn to pure black-box methods to optimize expensive simulation codes.

2.2 Expected improvement

Another optimization approach is to use specific tailored adaptive sampling strategies while building global surrogate models. As the focus of the sampling algorithm is on optimization, global surrogate models are not necessarily accurate over the whole design space. In engineering, adaptive sampling strategies are also known as infill criteria. An infill criterion is a function, also known as figure of merit, that measures how *interesting* a data point is in the design space. Starting from an initial approximation of the design space, identifying new data points (infill or update points) to update the approximation model is then done by optimizing the infill criterion. In global SBO it is crucial to balance between exploration[2] and exploitation[3]. A well-known infill criterion that is able to effectively solve this trade-off is Expected Improvement (EI), which has been popularized by Jones et al. [4, 27] as the Efficient Global Optimization (EGO) algorithm. EI has been suggested in the literature as early as 1978 [28]. Jones wrote an excellent discussion regarding the infill criteria approach in [29]. Subsequently, Sasena

---

[2]enhancing the general accuracy of the surrogate model
[3]enhancing the accuracy of the surrogate model solely in the region of the (current) optimum
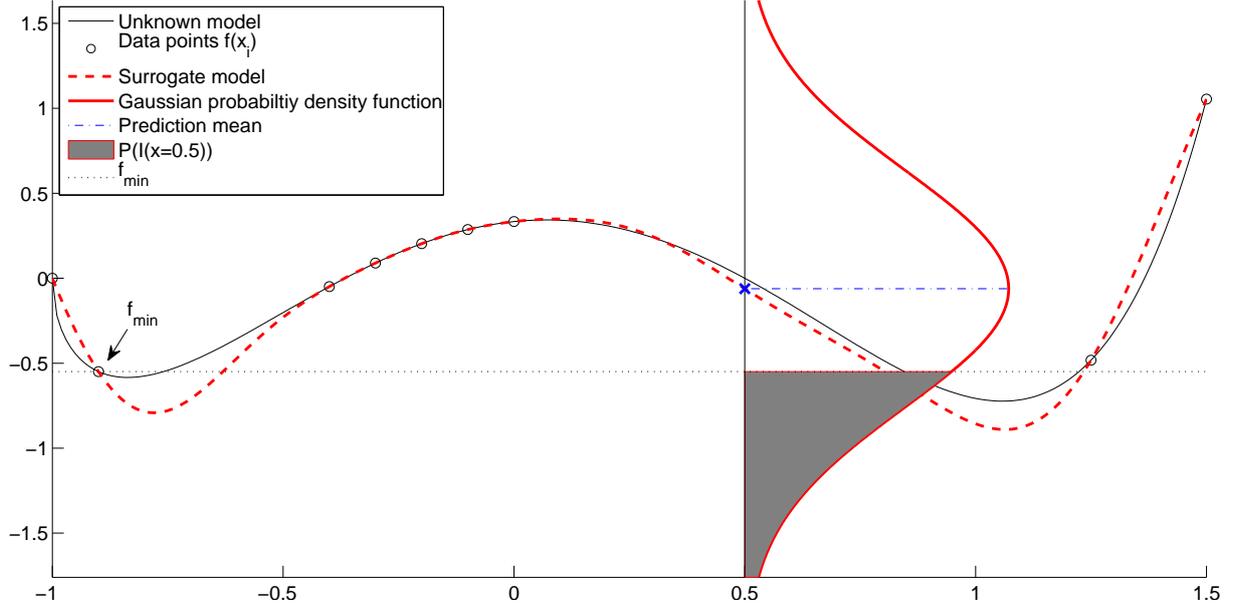
Figure 1: Graphical illustration of a Gaussian Process and expected improvement. A surrogate model (dashed line) is constructed based on some data points (circles). For each point the surrogate model predicts a Gaussian probability density function (PDF). At $x = 0.5$ an example of such a PDF is drawn. The volume of the shaded area is the probability of improvement and the first moment of this area is the expected improvement.

compared different infill criteria for optimization and investigated extensions of those infill criteria for constrained optimization problems in [30].

The EI criterion can easily be interpreted graphically (see Figure 1). At $\mathbf{x} = 0.5$, a Gaussian probability density function is drawn and expresses the uncertainty about the predicted function value of a sampled and unknown function $y = f(\mathbf{x})$. Thus, the uncertainty at any point $\mathbf{x}$ is treated as the realization of a random variable $Y(\mathbf{x})$ with mean $\hat{y} = \hat{f}(\mathbf{x})$ (= prediction) and variance $\hat{s}^2 = \hat{\sigma}^2(\mathbf{x})$ (= prediction variance). Assuming the random variable $Y(\mathbf{x})$ is normally distributed, then the shaded area under the Gaussian probability density function is the Probability of Improvement (PoI) of $Y(\mathbf{x})$ over the intermediate minimum function value f_{min}(the dotted line), denoted as $P(Y(\mathbf{x}) \leq f_{min})$, i.e.,

$$
PoI(\mathbf{x}) = P(Y(\mathbf{x}) \leq f_{min}) = \int\limits_{-\infty}^{f_{min}} \phi(Y(\mathbf{x}))\, dY
$$
$$
= \Phi\left(\frac{f_{min} - \hat{y}}{\hat{s}} \right), \tag{1}
$$

where $\phi(\cdot)$ and $\Phi(\cdot)$ are the normal probability density function and normal cumulative distribution function, respectively. The probability of improvement is already a very useful infill criterion. However, while this criterion describes the possibility of a better minimum function value, it does not quantify how large this improvement will be.

EI quantifies the improvement by considering the first moment of the shaded area, i.e., every possible

improvement over f_{min} multiplied by the associated likelihood. For continuous functions EI is an integral defined as:

$$E[I(\mathbf{x})] = \int\limits_{-\infty}^{f_{min}} I(\mathbf{x}) \cdot \phi(Y(\mathbf{x}))\, dY, \tag{2}$$

where

$$I(\mathbf{x}) = max(f_{min} - Y(\mathbf{x}), 0). \tag{3}$$

Hence, EI can be rewritten in closed form as:

$$E[I(\mathbf{x})] = \begin{cases} (f_{min} - \hat{y}) \cdot \Phi\left(\frac{f_{min} - \hat{y}}{\hat{s}}\right) + \hat{s} \cdot \phi\left(\frac{f_{min} - \hat{y}}{\hat{s}}\right) & if\ \hat{s} > 0 \\ 0 & if\ \hat{s} = 0 \end{cases}. \tag{4}$$

EI (Equation 4) and PoI (Equation 1) serve as utility functions, often conceived as figures of merit, which have to be optimized over $\mathbf{x}$ to find the subsequent data point to evaluate. Note, however, that besides the prediction $\hat{y} = \hat{f}(\mathbf{x})$ of the surrogate model, a point-wise error estimation $\hat{s} = \hat{\sigma}(\mathbf{x})$ of the surrogate is also required.

Therefore, the original EGO algorithm used kriging [4] as surrogate model of choice, since kriging provides analytical formulae for prediction as well as a point-wise error estimation. A full mathematical description of kriging is beyond the scope for this paper. Kriging has been explained many times in the literature, hence, only a overview of the most influential papers is given here. A good starting point for kriging are the introductions of Matheron et al. [31] and Sacks et al. [32]. Kriging, in fact, is part of a broader class of approximation methods, namely, Gaussian Processes (GP). While traditional approximation methods only predict a single function value, GP methods predict a complete normal distribution $Y(\mathbf{x}) \sim \mathcal{N}(\hat{y}, \hat{s})$ for each point $\mathbf{x}$. The predicted distribution imparts the probability that a particular function value occurs.

For a full overview of modern GP the reader is referred to the excellent GP reference book of Rasmussen et al. [33]. Depending on the context some authors coin the term Gaussian Process (temporal) or Gaussian Random Field Metamodels (GRFM; spatial) [34], however, the underlying methods are the same.

3   Example 1: Microwave filter

3.1   Application

The first example is a microwave inter-digital filter, used for instance in cellular phones. This component can be analyzed in various ways. Circuit simulation allows for very fast evaluation with reasonable accu-

(a) CST model.

(b) Top view. The five geometric design parameters define implicitly the length of the microstrips ($off1$, $off2$ and $off3$) and the spacings between the microstrips ($S1$ and $S2$).
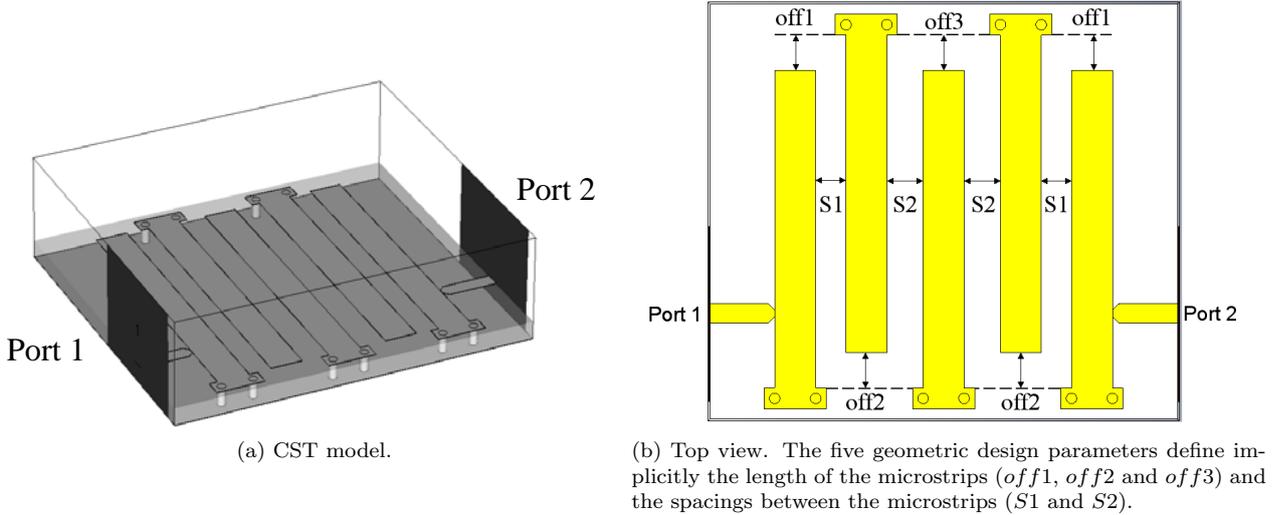
Figure 2: Microwave narrow-band filter.

racy, whereas full-wave electromagnetic (EM) simulations provide high accuracy at a high computational cost. The inter-digital filter presented in this paper has been analyzed and optimized before by Swanson [35] using a combination of analytic methods, circuit simulation and EM simulations. We use the EI criterion to optimize the filter, and use the CST MicroWave Studio® (CST MWS) as a full-wave EM simulation tool. The top view of the filter is shown in Figure 2b and consists of five quarter-wavelength parallel microstrip resonators. The scalable layout is fully parametrized by $S1 = [32, 38]$ mm, $S1 = [40, 48]$ mm, $off1 = [-3, 9]$ mm, $off2 = [-3, 3]$ mm and $off3 = [-3, 3]$ mm. The last three parameters define the offset of the quarter-wavelength microstrip resonators with respect to the horizontal dashed lines. In particular, $off1$ is the offset of the outer two strips, $off2$ of the second and fourth strip and $off3$ of the middle strip. In other words, the offsets implicitly define the length of each microstrip. While $S1$ and $S2$ (spacings) denote the gap between the outer two strips and the inner two strips, respectively. This results in a symmetric structure for the filter. In total, this adds up to five geometric design variables that must be optimized.

The goal is to design a fifth-order ($N = 5$) narrow-band filter with a flat passband response centered around 2.44 GHz and with a 10% bandwidth. In case of a lossless structure, a specific relationship between the passband ripple and return loss of the the filter allows us to minimize the ripple in the passband by minimizing the maximum of the S_{11}-parameter curve, i.e., the reflection coefficient, in the frequency range [2.32, 2.56] GHz. No specific optimization goals were set for the insertion loss and the stopband. The 'fast S-parameter' solver in CST MWS is used, and a frequency sweep takes approximately 5 to 10 minutes on a standard laptop.

This optimization problem is used to benchmark several types of kriging surrogate modeling strategies in conjunction with the EI approach. In addition, this problem serves as an example that black-box SBO methods are able to find optimal designs that compare favorably with designs obtained using domain-

7

specific knowledge [35].

3.2 Experimental setup

Version 6.1 of the SUMO toolbox is used to perform the optimization of the narrow-band filter and is configured as follows. The initial set of data points is generated by a maximin Latin Hypercube Design (LHD; implemented as in [36]) of 19 points together with 32 corner points, adding up to a total of 51 initial points. For this particular application the standard EI function (as defined in section 2) is used to select infill points. The EI function is optimized using the DIviding RECTangles (DIRECT) algorithm of Jones et al. [37] to determine the next data point to evaluate. A time budget constraint of 24 hours is applied, i.e., the overall optimization process runs for 24 hours.

The aforementioned configuration is reproduced three times with different types of the kriging surrogate model. The first two cases configure kriging as surrogate model of choice as implemented by the DACE toolbox [38]. More precisely, in one run the DACE toolbox itself performs the hyperparameter optimization, which comprises maximum likelihood estimation (MLE) using a modified Hooke & Jeeves direct search method [39] (pattern search). In another run, the hyperparameters of the kriging model are identified by Matlab's Genetic Algorithm (GA) toolbox using 10-fold cross validation to guide the search. In the last, and final configuration, a custom implementation of blind kriging [40] is employed (MLE using the DIRECT algorithm).

In addition, the cost function is also optimized using the Matlab pattern search and simulated annealing routines using the default options and initial point $\mathbf{x_0} = ([TODO])$. However, unlike the kriging configurations, a time budget of 24 hours is not applied, instead the optimization is halted after exceeding the number of samples that the best kriging configuration reached.

3.3 Results

Figure 3 shows the progress of the optimization process, i.e., the minimum cost function value versus the number of samples evaluated. The dotted line is the blind kriging configuration and performs worst in terms of the final solution. Constructing a blind kriging surrogate model is twice as expensive as standard kriging, hence, less time is available to evaluate the expensive simulation code. On the other hand, kriging (GA) is able to produce better kriging models due to a large model parameter search with a genetic algorithm guided by cross validation, and, thus, that configuration finds attracting basins more quickly. However, due to the cost of cross validation and evolutionary-based strategies it is only able to evaluate approximately 190 samples before the time budget is exceeded. Yet, at that point it is the best performing method. The standard kriging configuration finds the best solution. As it is significantly faster than the other two configurations, it is able to process more simulator runs, which proves to be more important in this application than a really accurate approximation model. After about 220 function
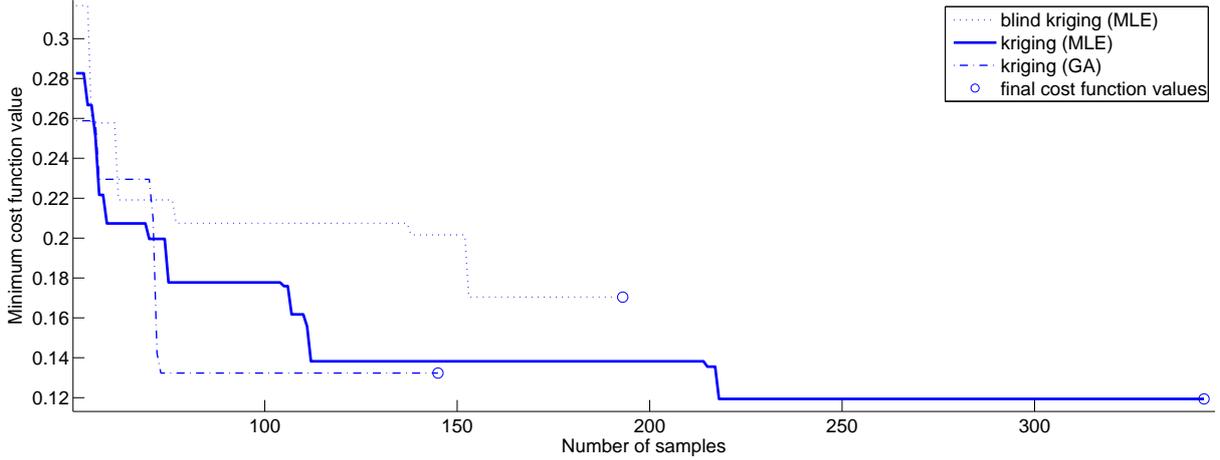
Figure 3: Evolution of the minimum cost function value versus the number of samples evaluated in 24 hours. The standard kriging (MLE) surrogate model finds the lowest cost function value. (inter-digital filter)

method	$|X|$	$\mathbf{x_{min}}$	f_{min}
blind kriging (MLE)	193	$(35.03, 41.37, 8.99, -1.14, 0.19)$	0.17038
kriging (GA)	145	$(36.00, 43.10, 5.90, -3.00, -3.00)$	0.13234
kriging (MLE)	344	$(36.68, 44.16, 6.16, -2.67, -2.47)$	0.11936
pattern search	344	TODO	TODO
simulated annealing	344	$(37.12, 43.10, 8.39, -1.87, 0.42)$	0.18074
reference [35]	$unknown$	$(37.10, 44.46, 6.30, -2.60, -2.43)$	0.12527

Table 1: Final designs of the inter-digital filter. $|X|$ is the number of samples evaluated (in 24 hours), $\mathbf{x_{min}}$ and f_{min} are the final solutions and cost function values respectively. (inter-digital filter)

evaluations it finds the best solution in the $5D$ design space and still has time to select about 100 more samples to look for an even better solution or to validate the current one. It should be noted that the usefulness of more expensive surrogate modeling strategies (such as blind kriging) may improve when the time of a single simulation run would increase to hours or even days. The final solutions found of each technique are displayed in Table 1, together with the reference optimum, as found by Swanson [35].

Surprisingly, the design found using the EI criterion and the kriging (MLE) surrogate models outperforms the reference design, i.e., with respect to the same cost function. To compare the different designs in a fair way the S_{11}-parameter curves are drawn in Figure 4. The reference design is constructed so that the ripples in the S_{11}-parameter curve are of equal height (exact equal ripple tuning). This might also be important as it guarantees consistent performance over the whole frequency range of interest. Note however that in this paper the cost function does not punish or favor equal ripple tuning. Thus, the solution found in this paper is better on the cost function but has not exact equal ripples (though very close to), while the reference design is slightly more consistent over the frequency range with regard to the ripples.

A huge advantage of SBO is the ability to easily explore the final (and intermediate) approximation models. The practitioner is able to cheaply analyze the robustness of the solution, locate other interesting
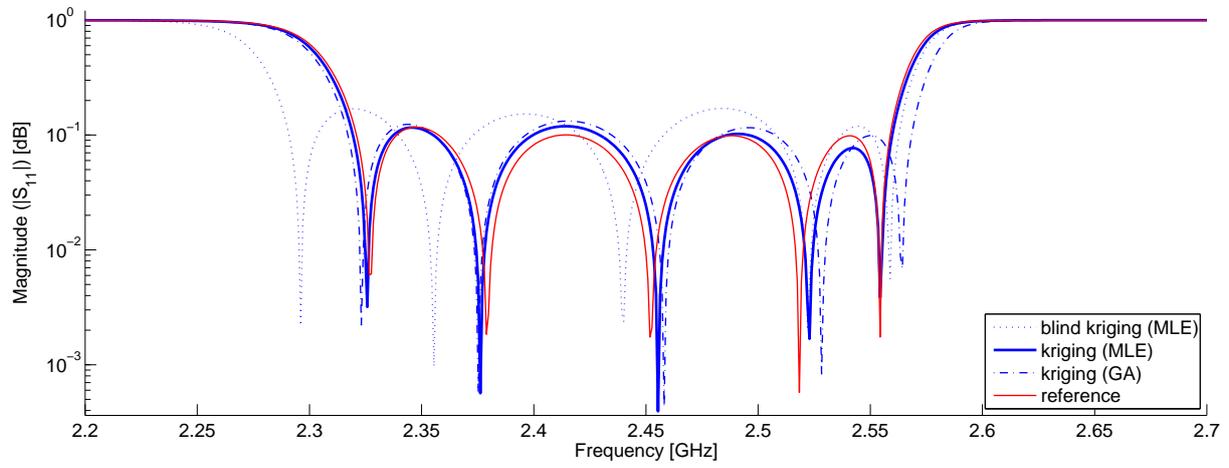
9

Figure 4: S_{11}-parameter magnitude plots of the different solutions. (inter-digital filter)
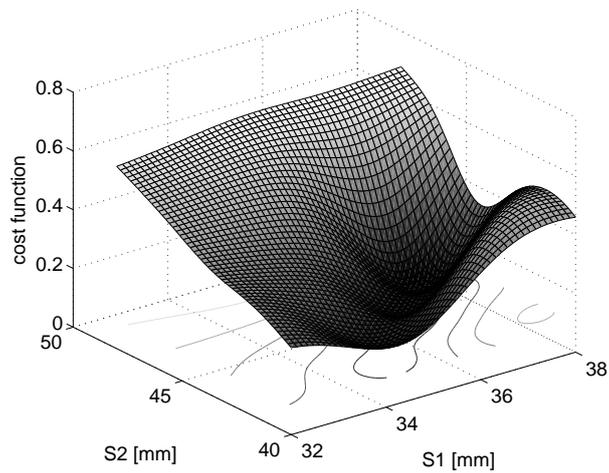

Figure 5: 2D slice plot of the kriging (MLE) surrogate model of the cost function. The offset parameters are set to the values of the final solution, i.e., $off1 = 6.16$ mm, $off2 = -2.67$ mm and $off3 = -2.47$ mm. (inter-digital filter)

regions (e.g., local optima), etc. For illustration purposes the kriging (MLE) surrogate model of the cost function is shown in Figure 5. This plot is a $2D$ slice of the $5D$ design space where the offset parameters are fixed, i.e., $off1 = 6.16$ mm, $off2 = -2.67$ mm and $off3 = -2.47$ mm.

4 Example 2: Textile antenna

4.1 Application

Material property identification is a well-known inverse problem. In particular, we address the characterization of the electrical properties of textile materials for the use in textile antennas. These antennas are constructed out of a non conductive textile substrate, a conductive ground plane and a conductive patch as shown in Figure 6. The textile substrate used here has a vegetable cellulose based origin and has a thickness of 0.805 mm. The textile substrate's electrical properties of interest are the permittivity ε_r and loss tangent $tan\delta$, and together with the patch geometry they determine the antenna performance indicators such as resonance frequency and bandwidth. The goal is to accurately characterize these two material properties of the textile antenna substrate based on the measured performance characteristics, provided by the reflection coefficient measurement of the antenna. The exact knowledge of the textile substrate's properties will then be exploited in the computer-aided design of complex wearable antenna topologies. Previously, manual fitting of the simulated and measured data has been reported in [41] for extracting the permittivity and loss tangent of the textile substrate. Based on a rough estimation (rule of thumb) of the textile substrates' electrical properties a full-wave EM simulation is performed to design a single-mode textile antenna with a sharp resonance. Therefore, the length L of the patch antenna is chosen such that a resonance is obtained in the vicinity of 2.4 GHz. Impedance matching is obtained by optimizing the width W and the coaxial feed positions x_f and y_f. The resulting patch dimensions L, W and feed positions x_f and y_f, based on the estimated permittivity and loss tangent of the textile substrate are 45.5 mm, 33 mm, 11 mm and 16.5 mm respectively. The way the real resonance peak of the textile antenna has shifted and changed in form allows to determine the actual permittivity and loss tangent of the substrate. Therefore, the textile antenna's reflection coefficient is measured and compared to simulations for multiple substrate parameters using ADS Momentum in the $[2, 3]$ GHz frequency range.

As a full inverse surrogate model is infeasible, the inverse problem is converted to a forward optimization problem. Specifically, the problem is reduced to the minimization of an error function ($=$ cost function) between the simulation results $\mathbf{y}$ and the measured data $\tilde{\mathbf{y}}$ (see Figure 7). The error function is the popular Mean Squared Error (MSE) defined by,

$$MSE(\mathbf{y}, \tilde{\mathbf{y}}) = \frac{1}{n} \sum_{i=1}^{n} (y_i - \tilde{y}_i)^2, \tag{5}$$

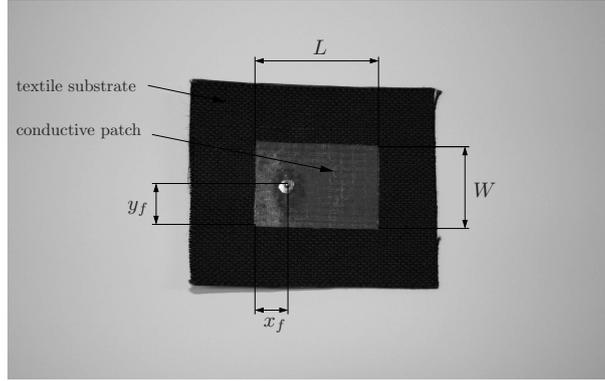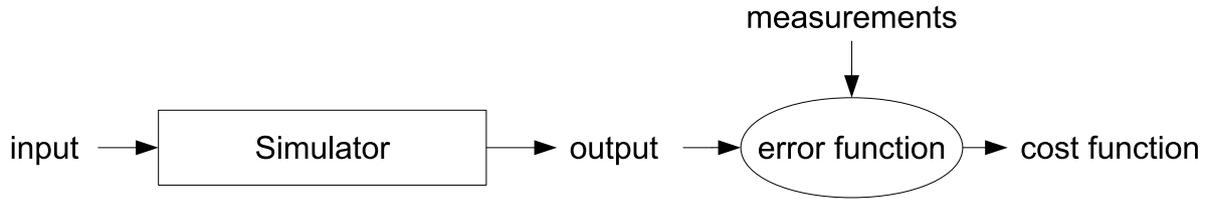Figure 6: Photograph of the textile antenna.

Figure 7: The inverse problem is solved by minimizing the error function between the simulation results $\mathbf{y}$ and the measured data $\tilde{\mathbf{y}}$.

with $n = 401$ the number of frequency points (samples). In the optimization process ϵ_r and $\tan\delta$ are bounded by $[1.1, 2.5]$ and $[0.020, 0.090]$, respectively. Hence, the correct material properties are identified by minimizing the MSE between the simulated S_{11} curve and the measured S_{11} curve.

A difficult problem often encountered with inverse problems is the presence of noise in the cost function. Obviously, the reflection coefficient measurements used in the error function contain noise. The error function (MSE) reduces the noise in the cost function. Any remaining noise is handled by the kriging surrogate model.

4.2 Experimental setup

Version 6.2 of the SUMO toolbox is utilized to solve the inverse problem of the textile antenna. The configuration is quite similar to the previous (forward) problem. An initial set of samples is generated by an optimal maximin Latin Hypercube Design (LHD; [42]) of 10 points together with four corner points, adding up to a total of 14 initial points. Subsequently, infill points are selected using the EI as a figure of merit which is optimized by the DIRECT algorithm. The optimization is halted when the number of samples exceeds 70.

The surrogate model of choice is a custom implementation of kriging. The hyperparameters are determined using Maximum Likelihood Estimation (MLE). The actual optimization is accomplished by a Sequential Quadratic Programming method (SQPLab[4] [43]), taking into account derivative information.

---

[4]SQPLab is found at `http://www-rocq.inria.fr/~gilbert/modulopt/optimization-routines/sqplab/sqplab.html`
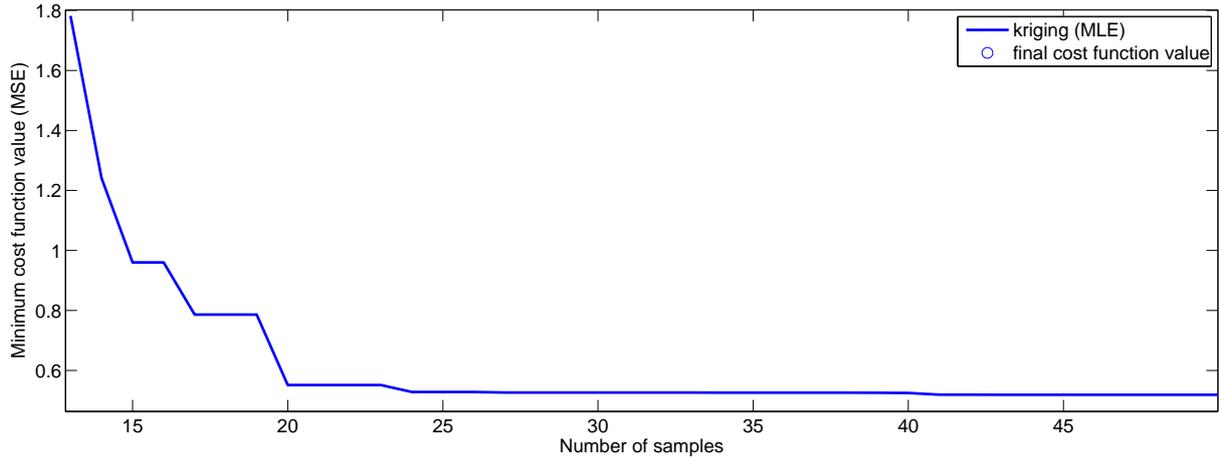
Figure 8: Evolution of the minimum cost function value versus the number of samples evaluated. The EI approach quickly locates the region of the global optimum and, subsequently, explores this region further, fine-tuning the final solution. (textile antenna)

method	$|X|$	ϵ_r	$\tan\delta$	MSE
kriging (MLE)	71	1.691	0.054	0.5185
pattern search	71	1.675	0.056	0.8963
simulated annealing	71	1.705	0.065	1.1513
reference	$unknown$	1.694	0.06	0.6974

Table 2: Final material parameters. $|X|$ is the number of samples, ϵ_r and $\tan\delta$ are the material parameters with the associated cost function value (MSE). (textile antenna)

Finally, kriging is modified to approximate the (noisy) cost function, instead of using interpolation.

To provide a comparison against traditional black-box optimization techniques the cost function is also optimized using the Matlab pattern search and simulated annealing routines using initial point $\mathbf{x_0} = (1.80, 0.05)$ and the same sample budget as the kriging configuration. The remaining options are left to their default values.

4.3   Results

An evolution plot of the minimum cost function values versus the number of samples is depicted in Figure 8. Starting from 14 samples the EI criterion quickly locates the region of the global optimum. At 20 samples the EI function starts exploring other parts of the design space (the flat parts), occasionally fine-tuning the current solution At approximately 45 samples the design space has been sufficiently explored and the final solution has been found. Still, the sampling continues until the sample budget is met, though no improvement is made.

The final optimal parameter combinations of each technique are presented in Table 2 along with the solution obtained through manual fitting and experimentation. The S_{11} curves of the optimal simulation run and the measurements are plotted in Figure 9. The solution found in this paper is significantly better than the reference optimum with respect to the cost function.
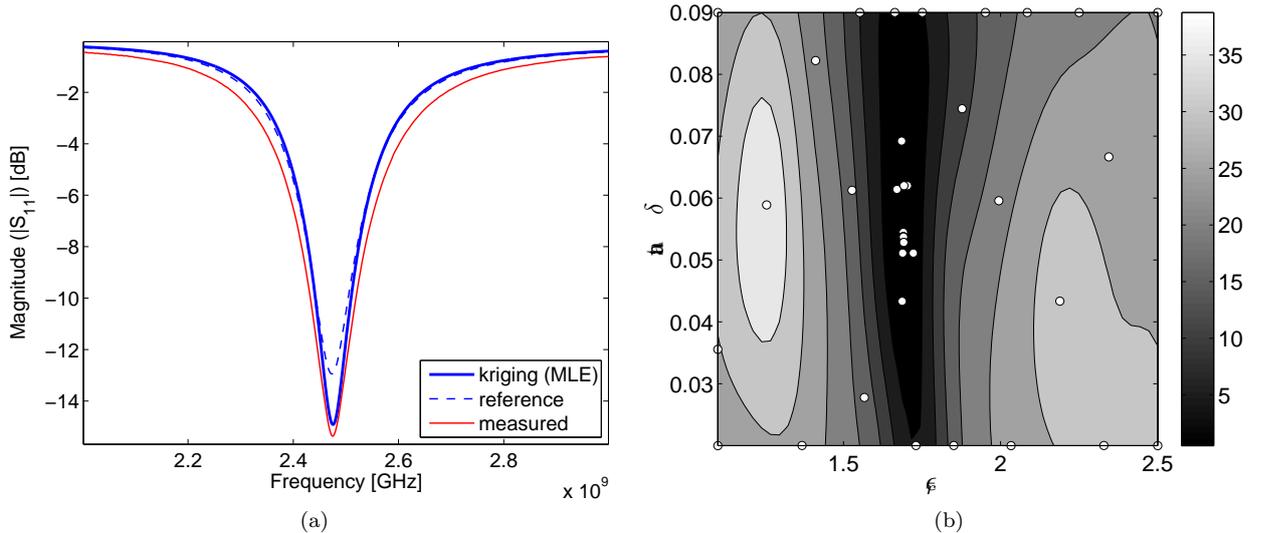
13

Figure 9: (a) S_{11}-parameter magnitude plots of the identified material parameter combinations. The design found by kriging (MLE) is significantly more accurate than the reference optimum, being able capture the measured resonance peak nicely. (b) Contour plot of the final kriging (MLE) surrogate model of the cost function based on 71 data points. (textile antenna)

Finally, the final kriging surrogate model of the cost function is displayed in Figure 9. As can be seen in the contour plot, the EI function explored the edges of the design space quite thoroughly, increasing the accuracy of the kriging model. Afterwards, more attention is paid to the valley, sampling densely near the region of the global optimum (the cluster of points).

5   Conclusion and future work

This work provides an overview of several Surrogate-Based Optimization (SBO) methods. A SBO approach based on the Expected Improvement (EI) criterion is implemented in a freely available toolbox, namely, the SUMO toolbox. The SUMO toolbox is used to benchmark different types of the kriging surrogate model on a forward EM optimization problem. Subsequently, a novel inverse EM problem is solved by minimizing the error between simulated data and measured data.

The optimization results of the first application show standard kriging (MLE) to outperform other types of kriging. In fact, it is demonstrated that the obtained design compares well against a reference design obtained by a domain expert.

The material property identification (inverse) problem is solved by optimizing an error function. The kriging surrogate model, adapted for regression, is able to approximate the noisy cost function very accurately, resulting in the identification of the material parameters with a minimal number of expensive function evaluations. The final material parameter combination largely agrees with the measured data. In particular, the identification of the resonance peak of the S_{11} curve is highly accurately. Any remaining difference between the measured data and the best solution is due to missing parameters in the simulation,

e.g., the finite conductivity of the conductive parts is not taken into account.

Future work includes full inverse electrical characterization of textile materials by including the finite conductivity and extensions to Multi-Objective Surrogate-Based Optimization (MOSBO) [44] methods.

Acknowledgments

References

[1] G. Wang and S. Shan. Review of metamodeling techniques in support of engineering design optimization. *Journal of Mechanical Design*, 129(4):370–380, 2007.

[2] J. Hadamard. Sur les problèmes aux dérivées partielles et leur signification physique. Technical Report 49–52, Princeton University Bulletin, 1902.

[3] R.R. Barton. Issues in development of simultaneous forward-inverse metamodels. In *Proceedings of the Winter Simulation Conference*, pages 209–217, 2005.

[4] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4):455–492, Nov. 1998.

[5] D. Gorissen, L. De Tommasi, K. Crombecq, and T. Dhaene. Sequential modeling of a low noise amplifier with neural networks and active learning. *Neural Computing and Applications*, 18(5):485–494, Jun. 2009.

[6] L. Lebensztajn, C.A.R. Marretto, M.C. Costa, and J-L Coulomb. Kriging: a useful tool for electromagnetic device optimization. *Magnetics, IEEE Transactions on*, 40(2):1196–1199, 2004.

[7] E.S. Siah, T. Ozdemir, J.L. Volakis, P. Papalambros, and R. Wiese. Fast parameter optimization of large-scale electromagnetic objects using DIRECT with Kriging metamodeling. *IEEE Transactions on Microwave Theory and Techniques*, 52(1):276–285, 2004.

[8] S. Andradóttir. A review of simulation optimization techniques. In *Proceedings of the Winter Simulation Conference*, pages 151–158. IEEE Computer Society Press, 1998.

[9] Y-S. Ong, P. B. Nair, and A. J. Keane. Evolutionary optimization of computationally expensive problems via surrogate modeling. *American Institute of Aeronautics and Astronautics Journal*, 41(4):687–696, 2003.

[10] Z. Z. Zhou, Y. S. Ong, M. H. Lim, and B. S. Lee. Memetic algorithm using multi-surrogates for computationally expensive optimization problems. *Soft Computing*, 11(4):957–971, 2007.

[11] Z. Z. Zhou, Y. S. Ong, P. B. Nair, A. J. Keane, and K. Y. Lum. Combining global and local surrogate models to accelerate evolutionary optimization. *IEEE Transactions On Systems, Man and Cybernetics - Part C*, 37(1):66–76, 2007.

[12] D. Lim, Y-S. Ong, Y. Jin, and B. Sendhoff. A study on metamodeling techniques, ensembles, and multi-surrogates in evolutionary computation. In *Proceedings of the 9th Annual Conference on Genetic and Evolutionary Computation (GECCO 07)*, pages 1288–1295, New York, NY, USA, 2007. ACM.

[13] N. Alexandrov, J. E. Dennis, R. M. Lewis, and V. Torczon. A trust region framework for managing the use of approximation models in optimization. *Structural Optimization*, 15(NASA CR-201745 ICASE Report No. 97-50):16–23, October 1998.

[14] F. V. Keulen and V. V. Toropov. The multipoint approximation method in a parallel computing environment. In *Gesellschaft für angewandte Mathematik und Mechanik (GAMM)*, 1999.

[15] A. I.J. Forrester and A. J. Keane. Recent advances in surrogate-based optimization. *Progress in Aerospace Sciences*, 45:50–79, 2009.

[16] R.T. Haftka. Combining global and local approximations. *AIAA Journal*, 29:1523–1525, 1991.

[17] M.S. Eldred, A.A. Giunta, and S.S. Collis. Second-order corrections for surrogate-based optimization with model hierarchies. In *Proceedings of the 11th AIAA/ISSMO Multidsciplinary Analysis & Optimization Conference*, 2004.

[18] J. W. Bandler, R. M. Biernacki, Shao Hua Chen, P. A. Grobelny, and R. H. Hemmers. Space mapping technique for electromagnetic optimization. *IEEE Transactions on Microwave Theory and Techniques*, 42(12):2536–2544, Aug. 1994.

[19] M. Bakr, J. Bandler, K. Madsen, and J. Sondergaard. Review of the space-mapping approach to engineering optimization and modeling. *Optimization and Engineering*, 1(3):241–276, 2000.

[20] J.W. Bandler, Q.S. Cheng, S.A. Dakroury, A.S. Mohamed, M.H. Bakr, K. Madsen, and J. Sondergaard. Space mapping: the state of the art. *IEEE Transactions on Microwave Theory and Techniques*, 52(1):337–361, Jan. 2004.

[21] S. Koziel and J.W. Bandler. Space-mapping optimization with adaptive surrogate model. *Microwave Theory and Techniques, IEEE Transactions on*, 55(3):541–547, March 2007.

[22] S. Koziel and J. W. Bandler. Interpolated coarse models for microwave design optimization with space mapping. *IEEE Transactions on Microwave Theory and Techiques*, 55:1739–1746, 2007.

[23] J. W. Bandler, D. M. Hailu, K. Madsen, and F. A. Pedersen. Space-mapping interpolating surrogate algorithm for highly optimized em-based design of microwave devices. *IEEE Transactions Microwave Theory and Techniques*, 52:2593–2600, 2004.

[24] P.W. Hemker and D. Echeverría. A trust-region strategy for manifold-mapping optimization. *Computational Physics*, 224(1):464–475, 2007.

[25] Noel A. C. Cressie. *Statistics for spatial data.* John Wiley & Sons, 1993.

[26] M. C. Kennedy and A. O'Hagan. Predicting the output from complex computer code when fast approximations are available. *Biometrika*, 87:1–13, 2000.

[27] M. Schonlau. *Computer Experiments and Global Optimization.* PhD thesis, University of Waterloo, 1997.

[28] J. Mockus, V. Tiesis, and A. Zilinskas. The application of Bayesian methods for seeking the extremum. *Towards Global Optimization*, 2:117–129, 1978.

[29] D. R. Jones. A taxonomy of global optimization methods based on response surfaces. *Global Optimization*, 21:345–383, 2001.

[30] M.J. Sasena. *Flexibility and Efficiency Enhancements For Constrainted Global Design Optimization with Kriging Approximations.* PhD thesis, University of Michigan, 2002.

[31] G. Matheron. Principles of geostatistics. *Economic Geology*, 58:1246–1266, 1963.

[32] J. Sacks, W. J. Welch, T.J. Mitchell, and H. P. Wynn. Design and analysis of computer experiments. *Statistical science*, 4(4):409–435, 1989.

[33] C. E. Rasmussen and C. K. I. Williams. *Gaussian Processes for Machine Learning.* MIT Press, 2006.

[34] M. T. M. Emmerich, K. Giannakoglou, and B. Naujoks. Single- and multiobjective evolutionary optimization assisted by gaussian random field metamodels. *IEEE Trans. Evolutionary Computation*, 10(4):421–439, 2006.

[35] D. G. Swanson. Narrow-band microwave filter design. *IEEE Microwave magazine*, 8:105–114, 2007.

[36] J. V. Roshan and H. Ying. Orthogonal-Maximin Latin hypercube designs. *Statistica Sinica*, 18:171–186, 2008.

[37] D.R. Jones, C.D. Perttunen, and B.E. Stuckman. Lipschitzian optimization without the Lipschitz constant. *Optimization Theory and Applications*, 79(1):157–181, 1993.

[38] S. N. Lophaven, H. B. Nielsen, and J. Søndergaard. Aspects of the matlab toolbox DACE. Technical report, Informatics and Mathematical Modelling, Technical University of Denmark, DTU, Richard Petersens Plads, Building 321, DK-2800 Kgs. Lyngby, 2002.

[39] R. Hooke and T. A. Jeeves. "direct search" solution of numerical and statistical problems. *J. of the Association of Computing Machinery (ACM)*, 8(2):212–229, 1961.

[40] V.R. Joseph, Y. Hung, and A. Sudjianto. Blind kriging: A new method for developing metamodels. *ASME Journal of Mechanical Design*, 130(3), 2008.

[41] F. Declercq, H. Rogier, and C. Hertleer. Permittivity and loss tangent characterization for garment antennas based on a new matrix-pencil two-line method. *IEEE Transactions on Antennas and Propagation*, 56(8):2548–2554, 2008.

[42] E.R. Dam, B.G.M. van Husslage, D. den Hertog, and J.B.M. Melissen. Maximin Latin hypercube designs in two dimensions. *Operations Research*, 55(1):158–169, 2007.

[43] J.F. Bonnans, J.C. Gilbert, C. Lemaréchal, and C.A. Sagastizábal. *Numerical Optimization: Theoretical and Practical Aspects*. Springer, 2006.

[44] J. Knowles and H. Nakayama. Meta-modeling in multiobjective optimization. In *Multiobjective Optimization: Interactive and Evolutionary Approaches*, pages 245–284. Springer-Verlag, Berlin, Heidelberg, 2008.