Kriging for eddy-current testing problems

Sandor Bilicz, Emmanuel Vazquez, Szabolcs Gyimothy, Jozsef Pavo, Marc Lambert

▶ **To cite this version:**

Sandor Bilicz, Emmanuel Vazquez, Szabolcs Gyimothy, Jozsef Pavo, Marc Lambert. Kriging for eddy-current testing problems. 17th Conference on the Computation of Electromagnetic Fields (COMPUMAG'09), Nov 2009, Florianopolis, Brazil. pp.3–4. hal-00438557v1

HAL Id: hal-00438557

https://hal.science/hal-00438557v1

Submitted on 4 Dec 2009 (v1), last revised 4 Dec 2009 (v2)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Kriging for eddy-current testing problems

S. Bilicz [*‡], E. Vazquez[†], Sz. Gyimóthy[‡], J. Pávó[‡], and M. Lambert[*]
[*]Département de Recherche en Électromagnétisme,
Laboratoire des Signaux et Systèmes UMR8506 (CNRS-SUPELEC-Univ Paris-Sud),
91192 Gif-sur-Yvette cedex, France
[†]SUPELEC, 91192 Gif-sur-Yvette cedex, France
[‡]Budapest University of Technology and Economics, Egry J. u. 18, H-1521 Budapest, Hungary

***Abstract*—Accurate numerical simulation of *Eddy-Current Testing* (ECT) experiments usually requires large computational efforts. So, a natural idea is to build a cheap approximation of the expensive-to-run simulator. This paper presents an approximation method based on *functional kriging*. Kriging is widely used in other domains, but is still unused in the ECT community. Its main idea is to build a random process model of the simulator. The extension of kriging to the case of functional output data (which is the typical case in ECT) is a recent development of mathematics. The paper introduces functional kriging and illustrates its performance via numerical examples using an ECT simulator based on a surface integral method. A comparison with other classical data interpolation methods is also carried out.**

***Index Terms*—Eddy-current testing; functional kriging; simulator approximation**

I. Introduction

Eddy-current testing is applied in a wide range of industrial problems. However, the accurate numerical simulation of the electromagnetic (EM) phenomenon is still challenging. Finite element and/or integral equation methods are able to provide acceptable accuracy, at a quite high price however. Nowadays, more and more emphasis is being put on the *emulation*, or *surrogate modeling* of EM phenomena [1].

The method proposed herein enables us to get a cheap approximation for a *specific* ECT problem. The main idea is to build a database of ECT configurations, i.e., a set of pairs consisting of a vector made of the parameters of a defect and the corresponding output signal, and then to use kriging to predict the output signal at any untried vector. Once the database is built, the expensive simulator is not needed any more, but only the kriging model.

Kriging is a prediction method developed in the 60s in the domain of geostatistics. By now, several variants of the method have been designed and a comprehensive literature exists on it (see [2], [3]). Some applications in electromagnetics are presented in [4], whilst the authors have applied kriging so as to solve ECT inverse problems [5].

All approaches discussed in the references above deal with a single or a couple of scalar outputs being modeled. However, the ECT experiments usually have a functional output which needs the use of a *functional kriging model* [6] for the approximation of the ECT simulator. Functional kriging is a recent extension of the original theory, an exhaustive overview of the state-of-art is found in [7].

II. The proposed method

A. Model of the ECT setup

Let us assume that a parameterized defect model with p parameters in a given configuration is available. Typically, those are dimensions, position and electric properties of the defect. The input parameters are collected in the vector $\mathbf{x}$, living in the p dimensional input space $\mathbb{X}$: $\mathbf{x} \in \mathbb{X} \subset \mathbb{R}^p$. Let us define a norm in the input space:

$$\|\mathbf{x}\| = \sqrt{\sum_{k=1}^{p} \left(\frac{x_k - x_k^{\min}}{x_k^{\max} - x_k^{\min}} \right)^2} , \tag{1}$$

where $x_k^{\min}$ and $x_k^{\max}$ are the lower and upper bounds of the k^{th} input parameter, respectively. This scaled norm is chosen (putting equal emphasis on all parameters), since one might have no prior information about the sensitivity of the problem to the different input parameters. The output function of the experiment $y_{\mathbf{x}}(t)$ is defined in an interval $t \in T$, for each input parameter vector $\mathbf{x}$, and is related here to the impedance variation of the probe coil caused by a defect described by $\mathbf{x}$. The independent variable t is related to the position of the probe coil. A whole surface scan is covered by the interval T. Let $y_{\mathbf{x}}(t)$ be real for the sake of simplicity of the further explanations. The complex impedance of the coil can be represented by a well-chosen real function (e.g. by dividing T into two disjunct intervals, T_R and T_I, and the real and imaginary parts are treated separately on them).

The connection between the input vector and the output function is formalized by the forward operator $\mathcal{F}$:

$$y_{\mathbf{x}}(t) = \mathcal{F}\{\mathbf{x}\} . \tag{2}$$

$\mathcal{F}$ is based on the underlying EM phenomena and theoretically it can be evaluated exactly. Typically, $\mathcal{F}$ is represented by a numerical simulator of the ECT experiment.

B. The forward prediction problem

Our aim is to find a cheap approximation for the computationally expensive operator $\mathcal{F}$. The proposed method involves two steps:

1) Compute some corresponding $(\mathbf{x}_i, y_{\mathbf{x}_i}(t)), i = 1, \ldots, n$ pairs ("samples") using (2) and store them in a *database* to be used at step 2. Fit a kriging interpolator to the samples.

2) Use the kriging interpolator based on the stored samples to predict the sought $y_{\mathbf{x}}(t)$ output function at any arbitrary $\mathbf{x}$ input point.

The two steps are separated, i.e., the time-consuming first step is done only once, then the fast second one (performed as many times as one needs a prediction) just relies on the stored data, yielding an off-line method.

C. Functional kriging

Functional kriging models the function $y_{\mathbf{x}}(t)$ by a Gaussian random process $\xi_{\mathbf{x}}(t)$. Let us assume that observations have been made at n input locations $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, thus the processes $\xi_{\mathbf{x}_1}(t), \xi_{\mathbf{x}_2}(t), \ldots, \xi_{\mathbf{x}_n}(t)$ are sampled. Our goal is to predict the process $\xi_{\mathbf{x}}(t)$ by using the n observed processes. Let us search for the predictor process $\hat{\xi}_{\mathbf{x}}(t)$ in the linear form

$$\hat{\xi}_{\mathbf{x}}(t) = \sum_{i=1}^{n} \lambda_i(\mathbf{x}) \xi_{\mathbf{x}_i}(t) \; . \tag{3}$$

The coefficients $\lambda_i(\mathbf{x})$ are chosen such that the prediction is optimal in a certain sense. Our requirements for the prediction are:

$$\mathbb{E}\left[\hat{\xi}_{\mathbf{x}}(t) - \xi_{\mathbf{x}}(t)\right] = 0 \; , \; \forall \; t \in T \; , \tag{4}$$

$$\int_T \mathbb{E}\left[\left(\hat{\xi}_{\mathbf{x}}(t) - \xi_{\mathbf{x}}(t)\right)^2\right] \mathrm{d}t \rightarrow \text{the smallest possible.} \tag{5}$$

Equation (4) expresses the unbiasedness of the prediction whereas criterion (5) formalizes the measure of error. This is the key difference between the classical scalar kriging methods and the functional kriging: here (5) defines the objective function (to be minimized) as an integral of the prediction error variance. Similarly with the notation used in the scalar kriging approach, the predictor (3) satisfying both conditions (4) and (5) is called the *Best Linear Unbiased Predictor (BLUP)*.

The set of Gaussian processes $\xi_{\mathbf{x}}(t)$ for each $\mathbf{x} \in \mathbb{X}$ can be considered as a set of Gaussian processes over $\mathbb{X}$ for each $t \in T$. Let us assume that the processes over $\mathbb{X}$ are stationary (for each $t \in T$), i.e., the mean $m(t)$ does not depend on $\mathbf{x}$:

$$\mathbb{E}[\xi_{\mathbf{x}}(t)] = m(t) \; , \; \forall \; \mathbf{x} \in \mathbb{X}, t \in T \; , \tag{6}$$

and the so-called semivariogram γ is shift-invariant, i.e.,

$$\frac{1}{2}\mathbb{E}\left[(\xi_{\mathbf{x}+\mathbf{h}}(t) - \xi_{\mathbf{x}}(t))^2\right] = \gamma(\mathbf{h}, t) \; , \; \forall \; \mathbf{x} \in \mathbb{X}, t \in T \tag{7}$$

holds. The difference vector $\mathbf{h}$ is called *lag* in the terms of kriging. $2\gamma(\mathbf{h}, t)$ is the variogram. The semivariogram is even in $\mathbf{h}$: $\gamma(\mathbf{h}, t) = \gamma(-\mathbf{h}, t)$.

Let us define the trace-semivariogram $\Gamma(\mathbf{h})$ as the generalization of the semivariogram:

$$\Gamma(\mathbf{h}) = \int_T \gamma(\mathbf{h}, t) \mathrm{d}t \; . \tag{8}$$

This function, not explicitly known, has the main role in the presented functional kriging approach and must be predicted from the observed data.

The coefficients λ_i in (3), to obtain the BLUP, are computed by solving the following linear system of equations:

$$\left[\begin{array}{c|c} -\mathbf{A} & \begin{matrix} 1 \\ \vdots \\ 1 \end{matrix} \\ \hline 1 \cdots 1 & 0 \end{array}\right] \left[\begin{array}{c} \lambda_1(\mathbf{x}) \\ \vdots \\ \lambda_n(\mathbf{x}) \\ \hline \mu \end{array}\right] = \left[\begin{array}{c} -\mathbf{a}(\mathbf{x}) \\ \hline 1 \end{array}\right] \; , \tag{9}$$

where the $n \times n$ matrix $\mathbf{A}$ contains the inter-sample trace-semivariogram values: $A_{ij} = \Gamma(\mathbf{x}_i - \mathbf{x}_j)$ and the $n \times 1$ column vector $\mathbf{a}(\mathbf{x})$ consists of the Γ values between the samples and the arbitrary location $\mathbf{x}$: $a_i(\mathbf{x}) = \Gamma(\mathbf{x} - \mathbf{x}_i)$. μ is a Lagrange multiplier to force the unbiasedness constraint.

Once the coefficients $\lambda_i(\mathbf{x})$ are computed, the prediction for the output function at any unsampled $\mathbf{x}$ location can be written as:

$$\hat{y}_{\mathbf{x}}(t) = \sum_{i=1}^{n} \lambda_i(\mathbf{x}) y_{\mathbf{x}_i}(t) \; . \tag{10}$$

D. Variogram estimation

Replacing the semivariogram $\gamma(\mathbf{h}, t)$ in (8) by its expression (7), changing the order of the integration, and approximating the $\mathbb{E}$-operator by the empirical mean computed from the samples leads to

$$2\hat{\Gamma}(\mathbf{h}) = \frac{1}{|N(\mathbf{h})|} \sum_{(i,j) \in N(\mathbf{h})} \int_T (y_{\mathbf{x}_i}(t) - y_{\mathbf{x}_j}(t))^2 \mathrm{d}t \; , \tag{11}$$

with $N(\mathbf{h})$ being a set of (i, j) pairs where $\mathbf{h} = \pm(\mathbf{x}_i - \mathbf{x}_j)$ holds. The number of elements in $N(\mathbf{h})$ is denoted as $|N(\mathbf{h})|$. Once the integrals of (11) are numerically computed the values of the *empirical trace-variogram* $2\hat{\Gamma}(\mathbf{h})$ are known at some discrete lags $\mathbf{h}$. However, knowing that an analytical expression of the trace-variogram is needed to compute (9) for an arbitrary $\mathbf{x}$, a parameterized variogram model is chosen and is fitted to the values of the empirical trace-variogram.

A popular parameterized variogram model is the *Matérn* function, having two parameters:

$$g(\mathbf{h}|\sigma^2, \nu) = \sigma^2 \left[1 - \frac{\left(2\sqrt{\nu}d\right)^{\nu} \mathcal{K}_{\nu}\left(2\sqrt{\nu}d\right)}{G(\nu)2^{\nu-1}}\right] \; , \tag{12}$$

with the parameter variance σ^2 and the smoothness parameter ν. $G(\cdot)$ is the gamma-function, $\mathcal{K}_{\nu}(\cdot)$ is the modified Bessel-function of the second kind, of order ν. The distance d is defined as an anisotropic measure of the length of the lag $\mathbf{h}$:

$$d = \sqrt{\sum_{k=1}^{p} \left(\frac{h_k}{\rho_k}\right)^2} \; , \tag{13}$$

where the coefficients ρ_k are the characteristic correlation distances of the k^{th} parameter. The smaller the ρ_k, the smaller the range of the trace-variogram along the k^{th} parameter, i.e., the more sensitive the output to this parameter. In this way, the trace-variogram model can be tuned to the possible anisotropic behaviour of the underlying operator $\mathcal{F}$. Consequently, one has to predict $2 + p$ hyperparameters from the samples of the empirical trace-variogram (11). This prediction task is

performed via a least squares fitting, where the resulting optimization problem is solved by the `FMINSEARCH` Matlab® routine.

III. Other interpolation methods

In this section, other approaches for the interpolation of functional data are summarized. An extensive overview of the interpolation techniques is outside the scope of the paper and we will focus on three classical schemes. Notations are kept the same as before.

- *Nearest neighbour* method (NN): the idea is to search for the nearest neighbour $\mathbf{x}_m$ among the observed locations $\mathbf{x}_1, \ldots, \mathbf{x}_n$, the predicted function being equated to $y_{\mathbf{x}_m}(t)$.
- *Multilinear* interpolation method (ML): it needs a regular Cartesian grid (not necessarily equidistant). The sampled input locations are at the nodes of this p dimensional grid, forming p dimensional cuboids. The interpolation is piecewise linear within all p coordinates of $\mathbf{x}$ in each cuboid.
- *Radial basis functions* interpolation (RBF, see e.g. [8]): The interpolation $\hat{y}_{\mathbf{x}}(t)$ is written as a linear combination of radial basis functions $\Phi(\cdot)$:

$$\hat{y}_{\mathbf{x}}(t) = \sum_{i=1}^{n} w_i(t)\, \Phi(\|\mathbf{x} - \mathbf{x}_i\|)\,. \tag{14}$$

  The functional weights $w_i(t)$ are determined (by solving a linear system of n equations) such that the prediction of $y_{\mathbf{x}}(t)$ fits the observations at the n sampled locations. The so-called thin-plate spline has been chosen as basis function:

$$\Phi(h) = h^2 \ln h \quad \text{if} \quad h > 0 \quad \text{and} \quad \Phi(0) = 0. \tag{15}$$

Both RBF and kriging are kernel-based approximation methods, thus they are very similar to each other. However, the main advantage of kriging is to provide a simple manner to estimate the parameters of the kernel (the variogram) via a stochastic framework.

IV. Illustrative ECT problem

A. The configurations

In the studied ECT configurations, a single or a pair of infinitesimally thin cracks (with zero conductivity) affects a homogeneous, non-magnetic conductive plate of conductivity $\sigma_0 = 10^6\,\mathrm{S/m}$ and depth $d = 1.25\,\mathrm{mm}$. An air-cored pancake-type coil (JSAEM benchmark probe) scans above a rectangular region of plate centered on the damaged zone. The coil is driven by a time-harmonic current with a frequency $f = 150\,\mathrm{kHz}$ and the variation of its impedance is measured on a surface scan made of 11×41 regularly spaced positions.

Three different configurations are discussed (see Fig. 1), all of them involve cracks opening to the "bottom" surface of the plate whereas the coil scans above the "top" surface (OD cracks in the terms of ECT).

#1 A 2-parameter configuration (L, A) involving two parallel cracks of equal length $L = (L_1 = L_2)$ and same depth ($D =$ 60 % of the plate thickness) separated by distance A.

Table I: Ranges of the input parameters.

Ex. #1.	$1\,\mathrm{mm} < A < 3.5\,\mathrm{mm}\,,\ 2\,\mathrm{mm} < L < 8\,\mathrm{mm}$
Ex. #2.	$1\,\mathrm{mm} < A < 3.5\,\mathrm{mm}\,,\ 2\,\mathrm{mm} < L_1, L_2 < 8\,\mathrm{mm}$
Ex. #3.	$1\,\mathrm{mm} < L < 10\,\mathrm{mm}\,,\ 5\,\% < D < 90\,\%\,,\ \lvert A\rvert, \lvert B\rvert < 1.5\,\mathrm{mm}$

#2 A 3-parameter configuration (L_1, L_2, A), an extension of the example #1, the two lengths (L_1, L_2) are allowed to vary separately.

#3 A 4-parameter configuration where a single crack with a given orientation but variable position (A, B) and sizes (L, D) (Fig. 1b).

The lower and upper bounds of the input parameters are in Table I.

The simulation of the experiments is based on the surface crack model [9] (not described here) – any other numerical model could be used.

B. Interpolation results

The accuracy of an interpolation method is evaluated via the interpolation error, defined by

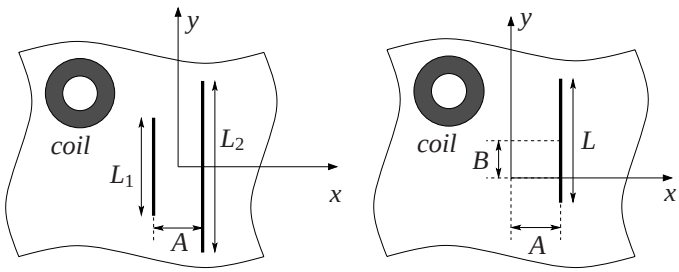
$$\epsilon_{\mathbf{x}} = \sqrt{\frac{\int_T (\hat{y}_{\mathbf{x}}(t) - y_{\mathbf{x}}(t))^2\,\mathrm{d}t}{\int_T (y_{\mathbf{x}^*}(t))^2\,\mathrm{d}t}}\,, \tag{16}$$

where $\hat{y}_{\mathbf{x}}(t)$ is the approximated, $y_{\mathbf{x}}(t)$ is the simulated output, whereas $\mathbf{x}^*$ is the so-called reference defect. In a real case $\mathbf{x}^*$ can represent, e.g., the smaller detectable defect, thus, the precision of the interpolation can be compared to the sensitivity of the related measurement. In our case, the reference defects are chosen arbitrarily according to Table II. Since the coil impedance is known at discrete coil locations, thus the discretized form of $y_{\mathbf{x}}(t)$ is used and the integrals are evaluated numerically.

Table II: Parameters of the reference defects.

Ex. #1	$L = 2\,\mathrm{mm}$, $A = 0.5\,\mathrm{mm}$
Ex. #2	$L_1 = L_2 = 2\,\mathrm{mm}$, $A = 0.5\,\mathrm{mm}$
Ex. #3	$A = B = 0\,\mathrm{mm}$, $L = 2\,\mathrm{mm}$, $D = 60\,\%$

The input samples are placed at the nodes of a regular Cartesian grid, covering the whole input space $\mathbb{X}$. The interpolation error (16) is computed on a fine regular grid (51×61, $11 \times 13 \times 13$, and $11 \times 11 \times 11 \times 11$, respecitvely



(a) Examples #1 and #2. (b) Example #3.

Figure 1: Sketch of the presented examples.

Table III: Maximal interpolation error over the input space.

Ex.	Grid	Error ($\max \epsilon_{\mathbf{x}}$)			
		NN	ML	RBF	Krig.
#1	3×3	0.833	0.229	0.734	0.248
	5×5	0.403	0.061	0.201	0.059
	7×7	0.283	0.034	0.098	0.026
#2	$3\times3\times3$	0.818	0.216	0.376	0.199
	$5\times5\times5$	0.270	0.059	0.099	0.039
	$7\times7\times7$	0.275	0.032	0.042	0.022
#3	$3\times3\times3\times3$	9.071	2.706	2.789	2.564
	$5\times5\times5\times5$	5.159	1.132	1.266	0.782
	$7\times7\times7\times7$	3.160	0.659	0.723	0.395

in the different examples) over $\mathbb{X}$. The maximum of $\epsilon_{\mathbf{x}}$ is determined on these grids and is given in Table III, by means of the four interpolation methods and different grids of input samples. One can see that the NN method is outperformed by all others. RBF and ML methods show similar performance in the sense of maximal error. However, in the case of RBF, the error is concentrated within certain regions of $\mathbb{X}$ whereas ML yields more evenly distributed error. The kriging prediction slightly outperforms the ML technique in almost all cases. The maximal error decreases more quickly using kriging interpolation when the number of samples increases. The distribution of the error over $\mathbb{X}$ is also favorable: similarly with RBF, the high-error regions are small. One can see that the similar kriging and RBF approaches yield quite different precision: the fitted variogram as a kernel for kriging yields much more precise approximations than RBF with a non-parametric basis function.

The computational cost of the NN and ML interpolators is practically negligible for a small number of samples whereas RBF and kriging methods need the inversion of an $n\times n$ matrix (or $(n+1)\times(n+1)$, respectively) once at the beginning. To make one single prediction for $y_{\mathbf{x}}(t)$, one has to evaluate n times the variogram model (12) or the basis function (15) and then compute the involved linear combination (10) or (14). In addition, the variogram has to be predicted once in the case of kriging.

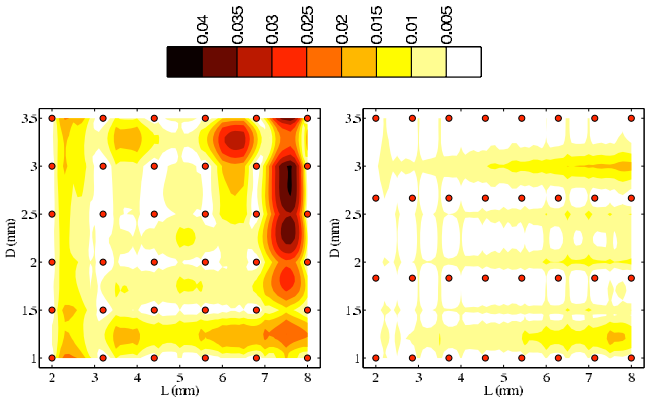
C. Adaptive gridding

One can also observe in Table III that the difference between the RBF and kriging results is the largest in the case of Ex. #1. This is due to the strong anisotropy of the problem. By using kriging, this can be predicted and taken into account easily via the anisotropic measure (13). Indeed, the parameter ρ is much higher for the distance A than for the length L. Knowing this meta-information, the precision of the approximation can be improved with a better choice of the input samples, by inserting more samples along the directions in which the problem is more sensitive. This idea is illustrated in Fig. 2 by using the so-called *adaptive grid*. The error is decreased whereas the number of samples is smaller (4×8) than by using *naive* sampling (6×6) .

V. Conclusion

A recent extension of kriging makes possible the direct prediction of functional data. This approach is applied to predict ECT output signals, yielding a cheap approximation of ECT simulations. The numerical tests show that the method has favorable interpolation properties. The underlying variogram model takes into account the anisotropy of the modeled problem, and moreover, the predicted anisotropy (as a meta-information) may improve the choice of samples to be stored.

In the authors' opinion, the method offers a new and promising way of cheap off-line approximation not only in the domain of ECT but more generally for expensive-to-simulate electromagnetic problems as well.



(a) Naive grid. (b) Adaptive grid.

Figure 2: Ex. #1, interpolation error $h_{\mathbf{x}}$ over $\mathbb{X}$ by using different grids in the kriging prediction.

Acknowledgements

The authors are grateful to Dominique Lesselier for his advices and comments on the manuscript.

This research is partially supported by the Research and Technology Innovation Fund of the Hungarian Government in the frame of the French-Hungarian Bilateral Intergovernmental S&T Cooperation (FR-1/2008) and by DIGITEO cluster's project.

References

[1] J. K. Sykulski, "New trends in optimization in electromagnetics," *Przeglad Elektrotechniczny*, vol. 83, no. 6, pp. 13–18, 2007, invited paper from ISTET'07 Conference.

[2] J. Chiles and P. Delfiner, *Geostatistics, Modeling Spatial Uncertainty*. Wiley, 1999.

[3] J. P. Kleijnen, "Kriging metamodeling in simulation: A review," *European Journal of Operational Research*, vol. 192, no. 3, pp. 707–716, February 2009. [Online]. Available: http://ideas.repec.org/a/eee/ejores/v192y2009i3p707-716.html

[4] L. Lebensztajn, C. A. R. Marretto, M. C. Costa, and J. L. Coulomb, "Kriging: a useful tool for electromagnetic device optimization," *IEEE Transactions on Magnetics*, vol. 40, no. 2, pp. 1196–1199, March 2004.

[5] S. Bilicz, E. Vazquez, M. Lambert, S. Gyimóthy, and J. Pávó, "Characterization of a 3D defect using the expected improvement algorithm," *COMPEL*, vol. 28, no. 4, pp. 851–864, 2009.

[6] R. Giraldo, P. Delicado, and J. Mateu, "Geostatistics for functional data: An ordinary kriging approach," Universitat Politècnica de Catalunya, Tech. Rep., 2007.

[7] P. Delicado, R. Giraldo, C. Comas, and J. Mateu, "Statistics for spatial functional data: some recent contributions," *Environmetrics*, 2009. [Online]. Available: http://dx.doi.org/10.1002/env.1003

[8] M. D. Buhmann, *Radial basis functions*. Cambridge Univ. Press, 2003.

[9] J. Pavo and D. Lesselier, "Calculation of eddy current testing probe signal with global approximation," *IEEE Transactions on Magnetics*, vol. 42, pp. 1419–1422, April 2006.