



**HAL**  
open science

## Kriging for indirect measurement, with application to flow measurement

Emmanuel Vazquez, Gilles Fleury, Eric Walter

► **To cite this version:**

Emmanuel Vazquez, Gilles Fleury, Eric Walter. Kriging for indirect measurement, with application to flow measurement. IEEE Transactions on Instrumentation and Measurement, 2006, Vol.55 (N°1), pp.343-349. 10.1109/TIM.2005.861498 . hal-00257729

**HAL Id: hal-00257729**

**<https://centralesupelec.hal.science/hal-00257729>**

Submitted on 20 Feb 2008

**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 indirect measurement, with application to flow measurement

E. Vazquez, G. Fleury, E. Walter

Manuscript sent April 03, 2003

E. Vazquez and E. Walter are with the Laboratoire des Signaux et Systèmes, CNRS – Supélec – Université Paris Sud, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France, e-mail: [emmanuel.vazquez@lss.supelec.fr](mailto:emmanuel.vazquez@lss.supelec.fr), [eric.walter@lss.supelec.fr](mailto:eric.walter@lss.supelec.fr)

G. Fleury is with Supélec School, Service des Mesures, 3 rue Joliot-Curie, 91192 Gif-sur-Yvette, France, e-mail: [Gilles.Fleury@supelec.fr](mailto:Gilles.Fleury@supelec.fr)

# Kriging for indirect measurement, with application to flow measurement

## Abstract

Kriging, a technique originating from geostatistics, is employed to build black-box models to be used to predict a quantity of interest based on the values taken by some experimental factors. This attractive alternative to more popular techniques such as neural networks is first presented. It is then applied to the measurement of the flow in a water pipe from the observation of speed at given points of a cross section. A pure black-box model turns out not to be satisfactory, and two approaches are suggested for incorporating prior knowledge. The second one, which is more systematic also turns out to provide much better performance.

## Index Terms

Black-box modeling, Gray-box modeling, Flow measurement, Kriging, Software sensors.

# Kriging for indirect measurement, with application to flow measurement

## I. INTRODUCTION

In this paper, we discuss *indirect measurement* under the viewpoint of a technique primarily developed by geostatisticians under the name of *Kriging* [1], [2]. The quantity of interest is assumed not to be directly measurable, and it should be estimated by processing some experimental data. In many applications this problem turns out to be non-linear and complex. Furthermore, the knowledge about the studied process may simply not be sufficient to establish the relation between the data and the quantity of interest, or this relation may turn out to be numerically intractable.

A solution is then to analyze the dependency between the data (or inputs) and the quantities of interest (or outputs) and to design a relation, under the form of a function, between the inputs and outputs. This function will allow prediction of the outputs of the process for inputs that have not been used yet. One will talk of a *black-box model*, by opposition to a *knowledge-based model*. In the static case, where there is no time dependence, the system inputs will be called *factors*. We shall denote the set of all input factors by the vector  $\mathbf{x} \in \mathbb{R}^d$ . In this paper we assume for simplicity that if there are several scalar outputs, then they are independent and can thus be predicted separately. We shall therefore consider only one output at a time for prediction, and assume that this output is a function written as  $f(\mathbf{x})$ . Black-box modeling then translates into function approximation or interpolation.

Section II advocates Kriging as a particularly interesting method for designing black-box models in measurement and instrumentation. Section III applies it to a water flow measurement problem, which provides an opportunity for a discussion of methodological principles. Conclusions and perspectives are presented in Section IV.

## II. WHAT IS KRIGING?

The method is named after D.G. Krige, a South-African mining engineer, whose aim was to elaborate maps of ore grade from scattered samples. The method was adapted and formalized in the late 60's by the French mathematician G. Matheron. It has become one of the main tools of geostatistics [2], a branch of statistics devoted to phenomena involving spatial factors.

Surprisingly enough, Kriging remains relatively confidential among engineers, although it provides an efficient alternative to more popular black-box modeling techniques such as neural networks, radial basis functions or splines. It is a powerful yet easy to use technique of data interpolation and approximation, as illustrated on Figure 1. It not only provides predictions but also confidence intervals for these predictions. In the next section, the basics of Kriging will be presented starting from linear regression under a probabilistic viewpoint.

### A. From linear prediction to Kriging

Assume that a set of data  $(\mathbf{x}_i, f_i)$ ,  $i = 1, \dots, n$ , is available, where the  $\mathbf{x}_i$ s are vectors of  $\mathbb{R}^d$ , whose elements are the values taken by the factors, and where the  $f_i$ s are the corresponding values of the output. Let  $\hat{f}(\mathbf{x})$  be the prediction of the output for a new numerical value of the factor vector  $\mathbf{x}$ . A natural approach for performing this prediction is to establish a model of the dependency between the factors and the output. The Kriging model is a stationary random process  $F(\mathbf{x})$  defined by its second-order moments, namely its mean  $b = E[F(\mathbf{x})]$  and covariance

$$R(\mathbf{x}, \mathbf{y}) = \text{cov}[F(\mathbf{x}), F(\mathbf{y})]. \quad (1)$$

In this probabilistic framework, the  $f_i$ s are assumed to be sampled values of the random variables  $F(\mathbf{x}_i)$ ,  $i = 1, \dots, n$  (note that the observation noise is neglected in this presentation but could also be taken into account).

If  $b$  is known, the model can always be simplified to a zero-mean process by subtracting it. The *Kriging estimator* for such a zero-mean process is by definition the random variable obtained by linear combination of the random variables  $F(\mathbf{x}_i)$ , written as

$$\hat{F}(\mathbf{x}) = \sum_{i=1}^n \lambda_i(\mathbf{x}) F(\mathbf{x}_i), \quad (2)$$

that best approximates  $F(\mathbf{x})$  in quadratic mean, *i.e.*, minimizes the variance of the prediction error  $\text{Var}[\hat{F}(\mathbf{x}) - F(\mathbf{x})]$ .  $\hat{F}(\mathbf{x})$  is then the orthogonal projection of  $F(\mathbf{x})$  onto  $\text{span}\{F(\mathbf{x}_1), \dots, F(\mathbf{x}_n)\}$ , so

$$E[(\hat{F}(\mathbf{x}) - F(\mathbf{x}))F(\mathbf{x}_i)] = 0, \forall i \in \{1, \dots, n\}. \quad (3)$$

If  $R(\mathbf{x}, \mathbf{y})$  is known, this projection is obtained by the well known linear regression formula,

$$\hat{F}(\mathbf{x}) = \mathbf{r}^T(\mathbf{x}) \mathbf{R}^{-1} \mathbf{F}, \quad (4)$$

where  $\mathbf{r}^T(\mathbf{x})$  is a row vector of covariances

$$\mathbf{r}^T(\mathbf{x}) = [R(\mathbf{x}_1, \mathbf{x}), \dots, R(\mathbf{x}_n, \mathbf{x})],$$

and  $\mathbf{R}$  is the covariance matrix of the random vector

$$\mathbf{F} = [F(\mathbf{x}_1), \dots, F(\mathbf{x}_n)]^T.$$

Thus, (4) allows the computation of the  $\lambda_i(\mathbf{x})$ s of (2), and the value of the output at  $\mathbf{x}$  can be predicted from the sample values  $f_i$  of the  $F(\mathbf{x}_i)$ s as

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^n \lambda_i(\mathbf{x}) f_i. \quad (5)$$

Note that the function  $\hat{f}(\mathbf{x})$  interpolates the data in a  $d$ -dimensional space of factors, even if the  $\mathbf{x}_i$ s are not on a regular grid. Compared with splines, Kriging yields very similar results. Actually splines may be interpreted as a special case of Kriging [3], [4]. The idea that is specific to Kriging is the introduction of a covariance function that describes how the process is correlated in the space of factors: one expects that two points sufficiently close in this space should give similar output values. More will be said about this in Section II-C. As already mentioned,

it is also possible to formulate Kriging in the case of additive noise on the data [2]. The Kriging estimator then yields an approximation of the data instead of an interpolation.

When the mean  $b$  of  $F(\mathbf{x})$  is unknown, the mean of  $\hat{F}(\mathbf{x})$  is also unknown and a constraint can be added on the estimator to require that  $\hat{F}(\mathbf{x})$  be unbiased. This constraint is written as

$$\mathbb{E}[\hat{F}(\mathbf{x}) - F(\mathbf{x})] = 0, \quad (6)$$

which implies

$$\sum_{i=1}^n \lambda_i(\mathbf{x}) - 1 = 0. \quad (7)$$

Until now, it was assumed that  $b$  was a constant (known or unknown) but it is also possible to expand the Kriging model so as to take into account a dependency of  $b$  on  $\mathbf{x}$ . The next section will present *intrinsic Kriging* as a solution to deal with this dependency by introducing parametric terms.

### B. Intrinsic Kriging

The theory of intrinsic Kriging [5] is mathematically rather involved and we shall stick to its surface. However the ideas remain simple, as we shall see. Our opinion is that intrinsic Kriging actually simplifies Kriging because it enables one to deal with a larger class of admissible covariances, some of which are very simple and convenient in practice (see Section II-C). Moreover, G. Matheron has shown that intrinsic Kriging was the proper framework to deal with parametric terms [5], so intrinsic Kriging can be viewed as the semi-parametric formulation of Kriging.

First, write  $F(\mathbf{x})$  as the sum of two terms,

$$F(\mathbf{x}) = W(\mathbf{x}) + b(\mathbf{x}, \mathbf{b}). \quad (8)$$

where  $W(\mathbf{x})$  is a second-order random process with zero mean and a known covariance  $R(\mathbf{x}, \mathbf{y})$ , and where  $b(\mathbf{x}, \mathbf{b})$  is a parametric model, which may be as simple as the constant of the previous section and whose parameter vector  $\mathbf{b}$  is unknown. Introducing such parametric terms in the Kriging model  $F(\mathbf{x})$  may help one incorporate prior knowledge about the process. For instance, an affine dependency with respect to the factors (*i.e.*  $b(\mathbf{x}, \mathbf{b}) = \mathbf{b}_1^T \mathbf{x} + b_0$ ) can be taken into account.

*Remark:* Back-fitting, that is, estimating the parameters of the parametric model before approximating the resulting residues by Kriging is not advisable. It would be as bad an idea as estimating  $b_0$  in the parametric model before  $\mathbf{b}_1$ . Instead, it is important to estimate all parameters and residues simultaneously.

The classical approach in intrinsic Kriging is to choose a parametric model under the form of a low-degree polynomial, which may be written as

$$\sum_l b_l f_l(\mathbf{x}), \quad (9)$$

where  $\{f_l\}_{l \in \mathcal{L}}$  constitutes a finite basis, typically a family of  $d$ -variate monomials in the factors of degree less than some specified integer  $k$ . This family generates a space of functions, which will be called  $\mathcal{N}$ . (This notation is chosen because this space corresponds to the null space, the kernel, of the smoothing operator in the theory of splines.)

The intrinsic Kriging estimator  $\hat{F}(\mathbf{x})$  remains of the form of the linear combination (2). Moreover,  $\hat{F}(\mathbf{x})$  will be constrained in the same way as in (6) when  $b$  was an unknown constant. As a best linear approximation property, we want the error  $\hat{F}(\mathbf{x}) - F(\mathbf{x})$  to be orthogonal to  $\mathcal{N}$ . This translates into a constraint on the  $\lambda_i(\mathbf{x})$ s, a generalization of (7) that can be written as

$$\sum_i \lambda_i(\mathbf{x}) f_i(\mathbf{x}_i) - f_l(\mathbf{x}) = 0 \quad \forall l \in L. \quad (10)$$

We now give, without derivation, the result to be used in this paper in Section III; for a proof, see [2]. Intrinsic Kriging builds a prediction from the data as in (5). (Note that (5) does not show the parametric terms as they are “intrinsically included”.) The  $\lambda_i(\mathbf{x})$ s are obtained by solving the linear system:

$$\begin{pmatrix} \mathbf{R} & \mathbf{M}^T \\ \mathbf{M} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\lambda}(\mathbf{x}) \\ \boldsymbol{\mu}(\mathbf{x}) \end{pmatrix} = \begin{pmatrix} \mathbf{r}(\mathbf{x}) \\ \mathbf{m}(\mathbf{x}) \end{pmatrix}, \quad (11)$$

where  $\mathbf{M}$  is the  $p \times n$  matrix with entries  $f_l(\mathbf{x}_j)$ ,  $\mathbf{m}(\mathbf{x}) \in \mathbb{R}^p$  is the vector of elements  $f_l(\mathbf{x})$ , and  $\boldsymbol{\mu}$  is a vector of Lagrange coefficients.

So far, the most critical part of Kriging, namely the choice of the covariance function, has been put aside. It will be considered in the next section.

### C. Choosing and estimating the covariance

A proper parameterization of the covariance is essential for correctly modeling the process. This is known in geostatistics as *structural analysis*.

A first important step is to choose the regularity of the model, which will depend on the behavior of the covariance at the origin. For instance, the resulting interpolation will be differentiable if the covariance is differentiable at the origin. On the contrary, if the covariance is not even continuous at the origin, discontinuities will be allowed in the result, which is called in geostatistics a *nugget effect* [2].

A second important step is to decide whether the covariance will be stationary and isotropic. Although not necessary, this will facilitate numerical computation since the covariance could then be written as a function of the Euclidean distance depending on a small number of parameters:

$$R(\mathbf{x}, \mathbf{y}) = R(\|\mathbf{x} - \mathbf{y}\|_2). \quad (12)$$

Many classes of covariances can be considered. They should correspond to positive functions, in order to ensure that the variance of  $\hat{F}(\mathbf{x})$  be positive. However in intrinsic Kriging this requirement can be relaxed by using *generalized covariances*, which are only *conditionally positive* functions, that is, yield positive variances of  $\hat{F}(\mathbf{x})$  provided that (10) holds.

We shall focus on a particular class of generalized covariances that are stationary and isotropic: the so-called *polynomial covariances of order  $k$* , which may be written as

$$R(h) = \sum_{p=0}^k (-1)^{p+1} a_p |h|^{2p+1}, \quad (13)$$

with  $h = \|\mathbf{x} - \mathbf{y}\|_2$ . These covariances correspond to a kernel  $\mathcal{N}$  generated by a family  $\{f_l\}_{l=1,p}$  of  $p$  independent monomials of degree at most  $k$ . The advantage of (13) is that these covariances are linear in the coefficients  $a_p$ . These coefficients must satisfy some conditions that hold true in all dimensions  $d$  of the factor vector, provided that all the  $a_p$ s be nonnegative [5]. This formulation also makes it easy to choose the regularity of the interpolation according to the regularity of  $R(h)$  at the origin.

Once the model of covariance has been chosen, an important remaining task is to estimate its parameters, the  $a_p$ s in (13). Several methods are available, but *maximum likelihood* estimation is particularly attractive when the number of observations used for the estimation is limited. For large data sets the method of the variogram, which is most frequently used by geostatisticians, may be preferred, or approximate maximum likelihood [6]. Note that for intrinsic Kriging a modified version of maximum likelihood, known as *REstricted Maximum Likelihood estimation* (REML), should rather be considered. Details may be found in [7].

### III. APPLICATION TO WATER FLOW MEASUREMENT

Kriging will be applied to the estimation of the flow in a water pipe from the observation of speed at given points of a cross section. The estimation should achieve a good precision given constraints about the maximum number of observations allowed (see Figure 2). Our approach was to reconstruct the entire water speed profile from the observed raw data, and then to integrate this profile to obtain an estimate of the flow rate. The problem is difficult because a large set of configurations must be considered including angled, converging or diverging pipes, and small or important flows. Several complex physical phenomena then occur such as turbulence, secondary modes, friction with the inner surface of the pipe.

We have simulated by finite differences a set of 3000 possible configurations, corresponding to different geometries and flows. Figure 3 shows examples of resulting speed profiles. In this paper, we shall focus on the methodological aspects and shall not discuss every aspects of the measurement device. The choice of the location of the speed observations is an interesting problem of experiment design, which will not be discussed either. We have chosen an empirical distribution based on 16 points corresponding to a good overall performance and satisfying a physical constraint concerning the minimal distance of the measurements to the inner surface of the pipe (see Figure 2).

#### A. Basic procedure

Once the profile has been reconstructed, the flow rate is estimated by integration of the function modeling the profile, say  $\hat{f}(\mathbf{x})$ , where  $\mathbf{x} \in \mathcal{X} = \{\mathbf{x} \in \mathbb{R}^2, \|\mathbf{x}\|_2 \leq r\}$  is the position on the section. This integration was carried out numerically by a Monte Carlo method, although an exact computation would have been possible, at least in principle (for analytically complicated covariance this is not trivial).

So the main step is to achieve the reconstruction of the speed profile, which is viewed here as an interpolation problem, neglecting speed measurement noise (the speed observations are actually very precise in the device). In the framework of Kriging, the reconstruction is easy. The speed at  $\mathbf{x}$  is estimated by

$$\hat{f}(\mathbf{x}) = \sum_{i=1}^{16} \lambda_i(\mathbf{x}) f_i, \quad (14)$$

where  $\{f_i\}_{i=1}^{16}$  are the observed speeds.

The  $\lambda_i(\mathbf{x})$ s are easily computed using (11). We have chosen a generalized polynomial covariance of order 2, which involves terms up to  $|h|^5$ . The estimation of the coefficients of this covariance is important in order to obtain the proper regularity. It was carried out by the REML method. However, it may not be necessary to estimate these coefficients for each speed profile if estimating the variance of the interpolation error is not required. In fact, the Kriging interpolation itself does not depend on the covariance  $R(\mathbf{x}, \mathbf{y})$ , but rather on the correlation  $\gamma(\mathbf{x}, \mathbf{y})$ . To see this, write the covariance as the product of a variance and a correlation function

$$R(\mathbf{x}, \mathbf{y}) = \sigma_0^2 \gamma(\mathbf{x}, \mathbf{y}), \quad (15)$$

with  $\sigma_0^2 = R(0, 0)$ . One can notice from (11) that the  $\lambda_i$ s are independent of  $\sigma_0^2$ . The regularity of interpolation is connected with the correlation structure  $\gamma(\mathbf{x}, \mathbf{y})$ , which is chosen independently of the flow rate. If the variance of the prediction error is to be estimated,  $\sigma_0^2$  becomes important and should be adjusted according to the flow rate.

Figures 4B and 5B show results of such an interpolation. One can easily see that the speed near the inner surface of the tube is overestimated, thus leading to a disappointing prediction of the flow (typically more than 2% of relative error). This overestimation is due to a physical constraint that prevents observation of the speed close to the surface of the pipe, thereby preventing the model to evaluate the drop in speed there correctly. A black-box approach to modeling admittedly does not give a satisfactory result, so taking into account some prior knowledge is necessary.

### B. Heuristic incorporation of prior knowledge

A rather simplistic proposal is to add fictitious observations against the inner surface of the pipe, say  $f_a$ ,  $a \in A$ . As the only available data are the 16 speed observations, a first estimation of  $f_a$ ,  $a \in A$  is carried out by Kriging using the above procedure. To take into account the friction phenomenon in the limit layer, these estimates are pulled back by multiplying them by some constant  $\rho < 1$ , that is, taking  $\hat{f}_a = \rho \hat{f}(\mathbf{x}_a)$ , with  $\hat{f}(\mathbf{x}_a)$  obtained by (14).  $\rho$  is estimated empirically from the simulation results. This cheap procedure could have been easily improved. For instance, the estimated  $f_a$  could have been refined by direct Kriging with the  $f_i$ s taken as the factors and the  $f_a$ s as the outputs. This method would have been interesting to study but is not really justified because of the quality of the results obtained with the more rigorous approach of Section III-C. Figure 4C shows adjusted points based on the estimation of Figure 4B. Next, the measured speeds are considered together with these fictitious measurements to carry out the interpolation, which leads to the significantly improved situation shown in Figure 4D. Figure 6B presents a histogram of the relative errors for all of the 3000 simulated configurations, which shows that the performance is already much improved compared with the initial solution of Figure 6A, despite the uncomfortably heuristic nature of the approach.

### C. A more systematic approach

The main drawback of the approach of Section III-B is the heuristic way in which knowledge is incorporated. As often, prior knowledge is only available under the form of a qualitative behaviour (here the knowledge that there

is friction and that the speed falls near the inner surface of the pipe).

We propose now a more elegant method to incorporate this approximate prior knowledge into the model. The semi-parametric formulation of intrinsic Kriging comes very handy for this purpose. Let  $f^*(\mathbf{x}) : \mathbf{x} \in \mathcal{X} \rightarrow \mathbb{R}$  be an assumed speed profile which represents prior knowledge, typically a standard behaviour of the speed across the section. This function is assumed to be known everywhere on  $\mathcal{X}$ . In the new procedure, the speed profile is still estimated by linear combination of the speed data as in (14). However, the previous space of factors is extended to  $\mathcal{X} \times \mathbb{R}$ , by considering the position  $\mathbf{x} \in \mathcal{X}$  and the scalar  $f^*(\mathbf{x})$  together. Now it is possible to take advantage of the framework of intrinsic Kriging to specify that:

- the (generalized) covariance does not depend on  $f^*(\mathbf{x})$ , which means that the correlation along the factor  $f^*(\mathbf{x})$  is undefined;
- most importantly, the space of orthogonal functions  $\mathcal{N}$  is taken to contain the function

$$\begin{aligned} f_1 : \mathcal{X} \times \mathbb{R} &\rightarrow \mathbb{R} \\ (\mathbf{x}, f^*(\mathbf{x})) &\mapsto f^*(\mathbf{x}), \end{aligned} \tag{16}$$

which is just a linear function of the factor  $f^*(\mathbf{x})$ .

As we have seen in Section II-B, this corresponds to choosing a parametric model that contains the term  $b_* f^*(\mathbf{x})$ . In fact, more than one assumed profile can be taken into account, by introducing as many parametric terms. There is however a limit. The more profiles there are, the more ill-conditioned the regression becomes. Since the parametric terms are not regularized, their contribution is not under control and it is easy to introduce too many parametric terms and thus over-fit the data: the increase in the number of degrees of freedom of the model may lead to a decrease in the quality of the prediction. Hence the number of parametric terms must be kept relatively low. With only two assumed profiles (as in the example of Figure 5), the application of this semi-parametric approach considerably improves the estimation of the flow, as the histograms of Figure 6 show.

For the results presented in this paper, two simulated speed profiles corresponding to different flow rates in a straight pipe were chosen as the assumed profiles. It would also have been possible to obtain these profiles analytically from fluids mechanics in a straight pipe. Such a semi-parametric approach is interesting whenever a simple model of the phenomenon is available. When a black-box model is made more accurate using the physics of the phenomenon, the term gray-box model can be employed.

#### IV. CONCLUSIONS AND PERSPECTIVES

This presentation has focused on methodological aspects and shown how Kriging could be used to design black-box models in instrumentation and measurement problems. Kriging allows the user to predict a quantity to be measured from raw data and, most importantly, allows prior knowledge to be incorporated at several steps of the estimation procedure. First, the user chooses a specific correlation structure and can take decisions about the regularity of the prediction. It is also possible to rely on automatic procedures such as maximum likelihood to estimate the parameters of this correlation structure. Second, the user can incorporate prior knowledge under the form of qualitative behaviour using the semi-parametric formulation of Kriging.

The example presented is typical of practical black-box measurement problems and is readily transposed to other situations. A number of questions were not considered in this paper. In our opinion, the most sensible improvement would be to study the influence of the number and positions of the speed measurements in the cross section. This problem of experiment design is fundamental in many disciplines. We guess that advantage could be taken in this respect of the confidence intervals provided by Kriging, and future work should focus on this aspect. However, the results obtained so far are already quite acceptable, especially since the histograms of errors presented are a very pessimistic view of the situation of interest as there are essentially singular cases (angles, shrinking, etc) in our simulations and it is unlikely that this is a condition under which one would dream to measure flow. Prior profiles could also be selected more carefully for the second method as this may contribute to further improvement of the measure.

#### REFERENCES

- [1] G. Matheron, "Principles of geostatistics," *Economic Geology*, vol. 58, pp. 1246–1266, 1963.
- [2] J.-P. Chiles and P. Delfiner, *Geostatistics : Modeling Spatial Uncertainty*, ser. Wiley Series in Probability and Statistics. New York: Wiley Interscience, 1999.
- [3] G. Matheron, "Splines and Kriging: their formal equivalence," in *Down-to-Earth Statistics: Solutions Looking for Geological Problems*, Syracuse Univ. of geology contributions ed., D. Merriam, Ed., 1981, pp. 77–95.
- [4] G. Wahba and G. Kimeldorf, "Spline functions and stochastic processes," *Sankhyā: the Indian Journal of Statistics: Series A*, vol. 32, no. 2, pp. 173–180, 1970.
- [5] G. Matheron, "The intrinsic random functions, and their applications," *Adv. Appl. Prob.*, vol. 5, pp. 439–468, 1973.
- [6] A. V. Vecchia, "Estimation and model identification for continuous spatial processes," *J. Roy. Statist. Soc. Ser. B*, vol. 50, no. 2, pp. 297–312, 1988.
- [7] M. Stein, *Interpolation of Spatial Data: Some Theory for Kriging*. New York: Springer Verlag, 1999.

## LIST OF FIGURES

1	Interpolation by Kriging (solid line) of data points materialized by circles, provided with confidence intervals (dashed lines). The data points have been sampled from the function in dotted line. . . . .	11
2	Cross section of the pipe for the flow measurement problem. Dots indicate the locations of speeds computed during simulation whereas triangles represent the observations (raw data) in the measurement device. There are only 16 observations. . . . .	12
3	Speed profiles for various configurations of pipes. For instance, (A) shows a profile in a straight pipe and (D) the profile in an angled pipe. . . . .	13
4	(A) actual profile to be estimated. (B) profile estimated from the 16 observations. (C) correction of speeds against the inner surface of the pipe. (D) final estimated profile. . . . .	14
5	(A) true speed profile, (B) reconstruction by Kriging, (C) corresponding prediction error, (D) interpolation with prior incorporated by intrinsic Kriging, (E) improved corresponding prediction error. . . .	15
6	Histograms of relative errors. (A) simple black-box modeling, showing a strong bias. Heuristic method for including prior knowledge (C) More systematic approach. . . . .	16

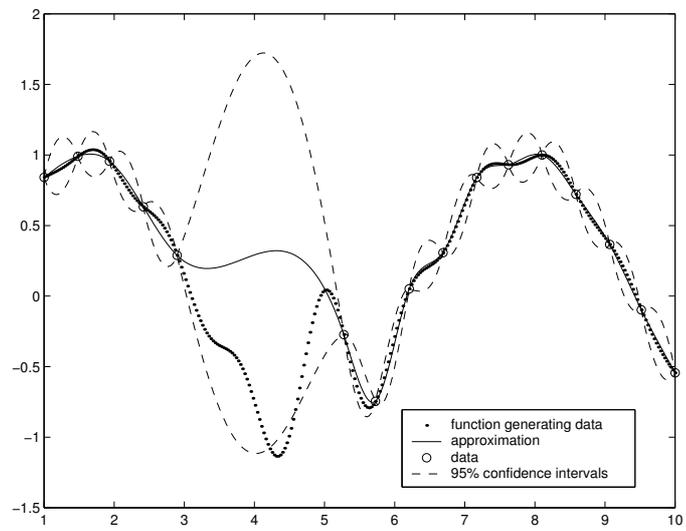


Fig. 1. Interpolation by Kriging (solid line) of data points materialized by circles, provided with confidence intervals (dashed lines). The data points have been sampled from the function in dotted line.

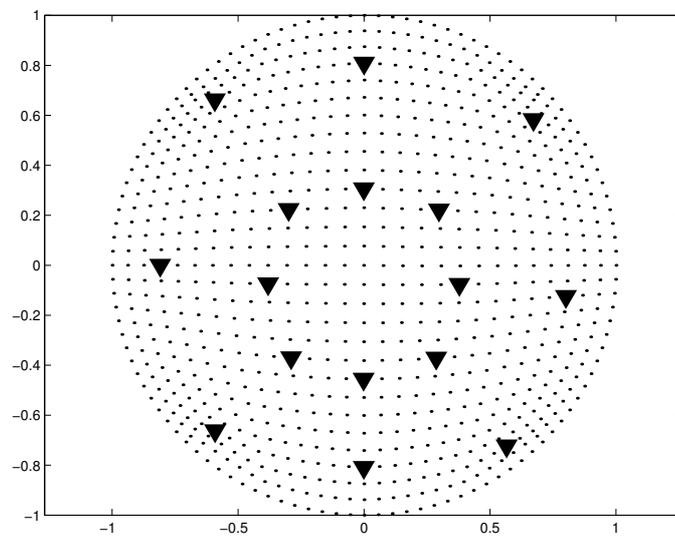


Fig. 2. Cross section of the pipe for the flow measurement problem. Dots indicate the locations of speeds computed during simulation whereas triangles represent the observations (raw data) in the measurement device. There are only 16 observations.

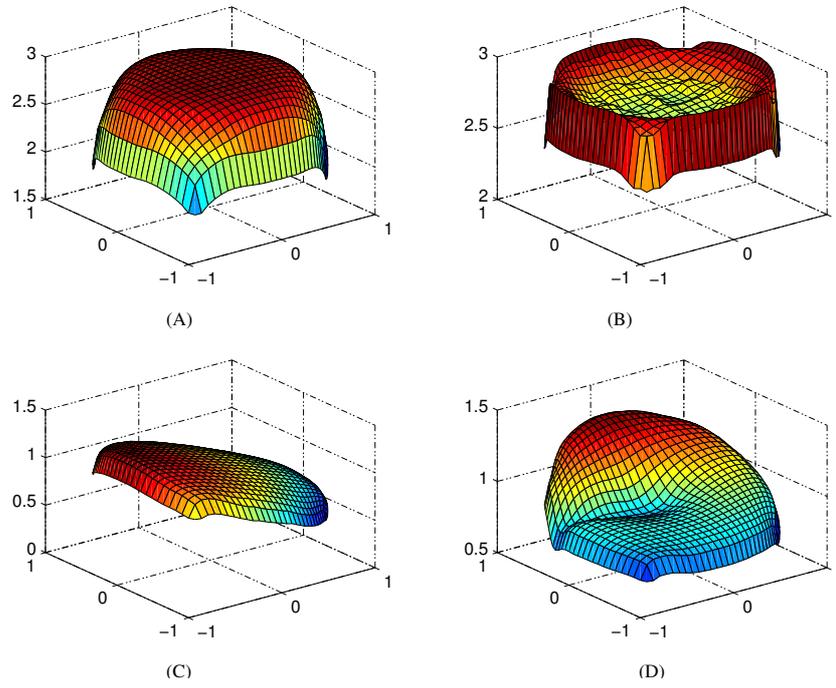


Fig. 3. Speed profiles for various configurations of pipes. For instance, (A) shows a profile in a straight pipe and (D) the profile in an angled pipe.

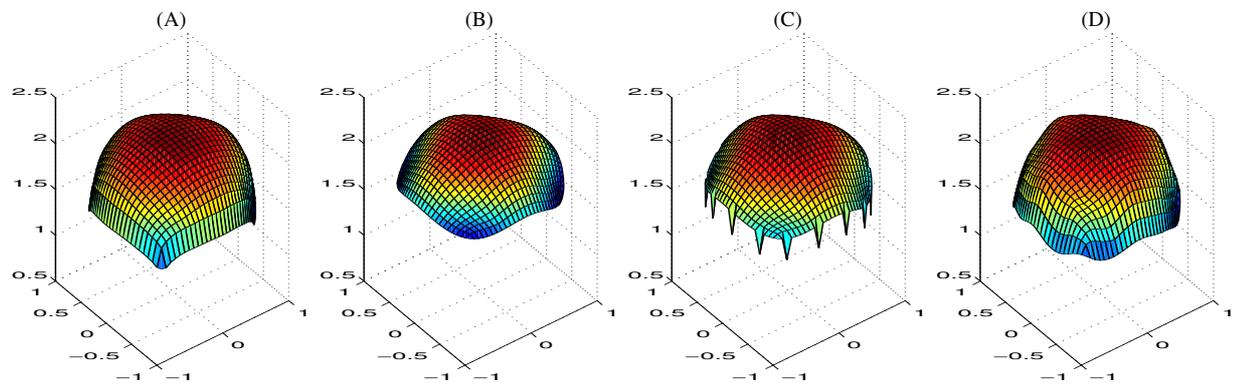


Fig. 4. (A) actual profile to be estimated. (B) profile estimated from the 16 observations. (C) correction of speeds against the inner surface of the pipe. (D) final estimated profile.

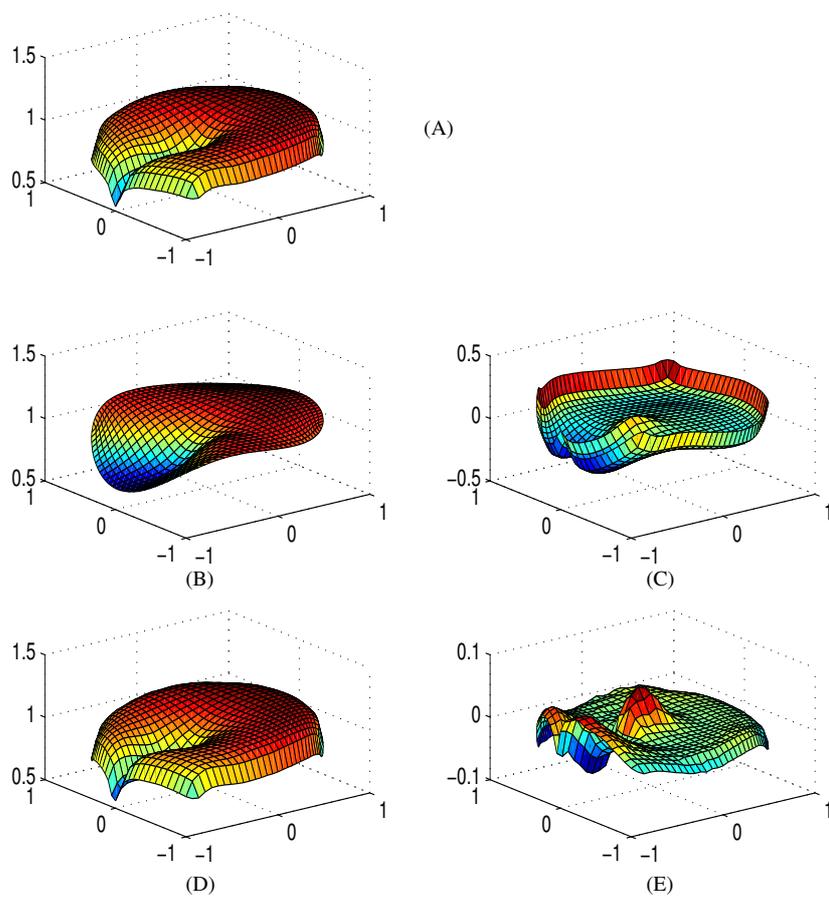


Fig. 5. (A) true speed profile, (B) reconstruction by Kriging, (C) corresponding prediction error, (D) interpolation with prior incorporated by intrinsic Kriging, (E) improved corresponding prediction error.

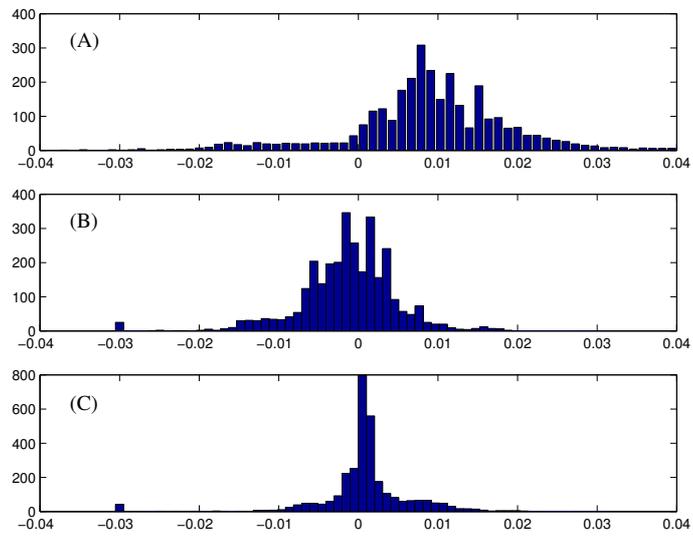


Fig. 6. Histograms of relative errors. (A) simple black-box modeling, showing a strong bias. Heuristic method for including prior knowledge (C) More systematic approach.