Applications of Mathematics

Pavel Krejčíř

Development of the kriging method with application

Applications of Mathematics, Vol. 47 (2002), No. 3, 217–230

Persistent URL: http://dml.cz/dmlcz/134496

47 (2002) APPLICATIONS OF MATHEMATICS No. 3, 217–230

DEVELOPMENT OF THE KRIGING METHOD WITH APPLICATION*

Pavel Krejčíř, Prague

(Received October 21, 1999)

Abstract. This paper describes a modification of the kriging method for working with the square root transformation of a spatial random process. We have developed this method for the situation where the spatial process observed is not supposed to be stationary but the assumption is that its square root is a second order stationary spatial random process. Consequently this method is developed for estimating the integral of the process observed and finally some application of the method is given to data from an environmental radioactivity survey.

Keywords: stochastic spatial process, second order stationarity, kriging, prediction

MSC 2000: 62M30, 62M20, 62P99

1. Introduction

The aim of this study is to estimate the total inventory from a set of observations of a spatial random process, which is in fact a random integral of the spatial process. Let $Z(\boldsymbol{s})$, $\boldsymbol{s} \in B \subset \mathbb{R}^2$ be a spatial random process. The inventory can be defined as

$$I(B) = \int_B Z(\boldsymbol{s}) \, \mathrm{d}\boldsymbol{s}. \tag{1}$$

The estimator of the inventory is in fact an estimator of an integral of the random function (1). This problem has been studied for example in [5], where an optimization technique is involved, or in [8] in the case that the observed point pattern forms a lattice. This assumption is not fulfilled in our situation, therefore we have to use a

* This work was supported by the Grant No. 201/98/0090 of the Grant Agency of the Czech Republic, by the project CEZ J13/98:113200008 and by the 5th frame program of the Commission of the European Community, project APPETISE, No. IST-99-11764.

different approach. We have decided in favour of optimal linear prediction, which is usually called kriging. A comparison of several methods used for spatial prediction is given in [1] and the fact that kriging is one of the most popular methods for spatial prediction encouraged us to use it.

We applied the developed method to two real data sets concerning environmental radioactivity in the United Kingdom. The data were collected by airborne gamma ray measurement. The first was sampled at the Blackpool and Ribble estuary area in north-west England and the second in the neighbourhood of Solway Firth, south-west Scotland. The data have been provided by SURRC, East Kilbride. The first area is schematically shown in Fig. 1 and the second is in Fig. 2.

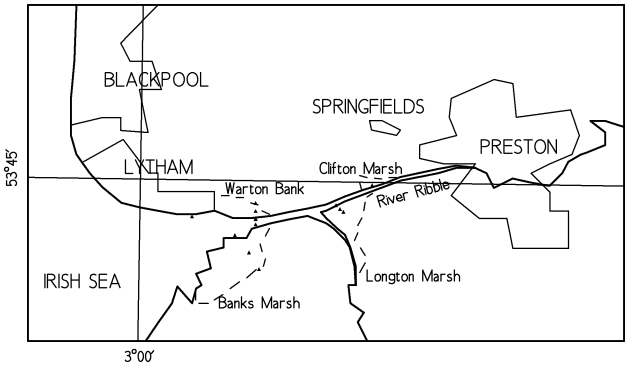

Figure 1. The map of Ribble estuary area.

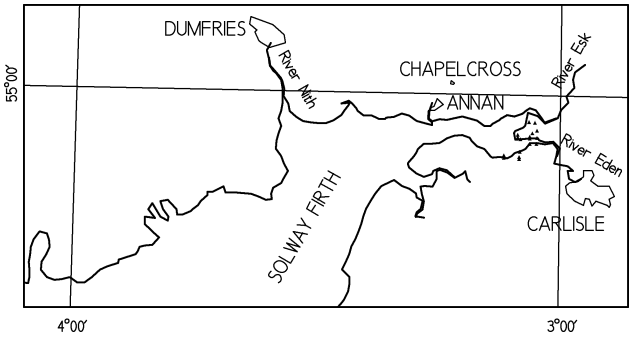

Figure 2. The map of Solway Firth area.

2. Basic tools of spatial statistics

Spatial statistics is a statistical discipline which deals with data including also an information about the 'location' of the datum. The environmental radioactivity data collected by airborne gamma ray spectrometry are suitable for spatial modeling. The data possess both the spatial location and random variation. The first major development of spatial statistics was inspired by applications in geology and mining industry and the problems and results have been presented under a common name *geostatistics*.

2.1. Spatial stochastic processes.

The basic element of spatial statistics is called a *spatial process* or a *random function* or a *random field* or a *stochastic process*. We will mostly use the first expression throughout this paper.

Definition. A *spatial process* is a set of random variables in the form

$$\{Z(\boldsymbol{s};\omega);\boldsymbol{s}\in D,\ \omega\in\Omega\}, \tag{2}$$

where D is a fixed subset of $\mathbb{R}^d$ of positive d-dimensional volume, $(\Omega,\mathscr{F},P)$ is a probability space.

We will suppose that there exists a system of functions for a given process $Z(\boldsymbol{s})$ with the following properties:

$$F_{\boldsymbol{s}_1,\dots,\boldsymbol{s}_m}(z_1,\dots,z_m)=P\big(Z(\boldsymbol{s}_1)\leqslant z_1,\dots,Z(\boldsymbol{s}_m)\leqslant z_m\big),\quad m\geqslant 1 \tag{3}$$

with

$$F_{\boldsymbol{s}_1,\dots,\boldsymbol{s}_m}(z_1,\dots,z_m)=F_{\boldsymbol{s}_{i_1},\dots,\boldsymbol{s}_{i_m}}(z_{i_1},\dots,z_{i_m}),\quad m\geqslant 1,\quad (i_1,\dots,i_m)\in\Pi_m,$$
$$F_{\boldsymbol{s}_1,\dots,\boldsymbol{s}_k}(z_1,\dots,z_k)=F_{\boldsymbol{s}_1,\dots,\boldsymbol{s}_k,\boldsymbol{s}_{k+1},\dots,\boldsymbol{s}_m}(z_1,\dots,z_k,\infty,\dots,\infty),\quad 1\leqslant k<m,$$

where Π_m denotes the set of all permutations of the set $(1,\dots,m)$. The system (3) is called a *consistent system of finite distributions* and defines the properties of the process (2). If all finite-dimensional distributions of the spatial process are Gaussian, the process is called a *Gaussian (spatial) process*.

Definition. If

$$\mu(\boldsymbol{s})=EZ(\boldsymbol{s}) \tag{4}$$

exists for all $\boldsymbol{s}\in D$, then the function $\mu(.)$ is called the *trend* of the process $Z(.)$. If

$$C(\boldsymbol{s},\boldsymbol{t})=\operatorname{cov}\{Z(\boldsymbol{s}),Z(\boldsymbol{t})\} \tag{5}$$

exists for all $\boldsymbol{s}\in D$, then the function $C(.,.)$ is called the *covariance function* of the process $Z(.)$.

The covariance function possesses the following simple property:

Lemma 1. *Let $C\colon D^2 \to \mathbb{R}$ be a covariance function of a spatial process $Z\colon D \times \Omega \to \mathbb{R}$. Then*

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j C(\boldsymbol{s}_i, \boldsymbol{s}_j) \geqslant 0 \tag{6}$$

holds for all $m \in \mathbb{N}$, $\boldsymbol{a} = (a_1, \ldots, a_m)^\top \in \mathbb{R}^m$ and $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m \in D$.

The property (6) is called the *positive semidefiniteness*.

In spatial statistics, we almost never deal with independent identically distributed (i.i.d.) random variables. Therefore it is necessary to make other assumptions about the random process observed. One common assumption about the distribution of the spatial data is usually the *stationarity*. Several types of stationarity have been defined, e.g. *strong (strict) stationarity*, *second-order stationarity* or *intrinsic stationarity*.

Definition. Let $Z(.)$ be a spatial process and let there exists a function $\gamma\colon D \to \mathbb{R}$ such that

$$\operatorname{var}\{Z(\boldsymbol{s}) - Z(\boldsymbol{t})\} = 2\gamma(\boldsymbol{s} - \boldsymbol{t}) \text{ for all } \boldsymbol{s}, \boldsymbol{t} \in D. \tag{7}$$

Then the function $2\gamma(.)$ is called the *variogram* and the function $\gamma(.)$ is called the *semivariogram*. The vector $\boldsymbol{s} - \boldsymbol{t}$ is then called the *lag*.

The variogram is a very popular mean of characterization of spatial processes. One of its properties is called the *conditional negative definiteness* (see [2]), as follows from the following lemma:

Lemma 2. *The inequality*

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j 2\gamma(\boldsymbol{s}_i - \boldsymbol{s}_j) \leqslant 0 \tag{8}$$

holds for every $m \in \mathbb{N}$, $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_m \in D$ and $\boldsymbol{a} = (a_1, \ldots, a_m)^\top \in \mathbb{R}^m$ such that $\sum_{i=1}^{m} a_i = 0$.

An important characteristic of spatial processes is described by the behaviour of their variograms close to $\mathbf{0}$. If $\lim_{\boldsymbol{h}\to\mathbf{0}} \gamma(\boldsymbol{h}) = 0$ then the process is called an *L_2-continuous process*. But experimental studies show that such a limit property

is not always met in practice. In order to cover such cases, Matheron introduced a new notion, namely the *nugget effect* δ_0 defined as

$$\lim_{\boldsymbol{h}\to\mathbf{0}} \gamma(\boldsymbol{h}) = \delta_0 > 0, \tag{9}$$

if the limit exists.

2.2. Statistical inference applied to spatial processes.

The most usual aim of statistical analysis of a spatial process is the *spatial prediction*, which means to estimate the value of the spatial process at a non-observed location. The term comes from the theory of time series and has been chosen in order to distinguish it from *estimating* a constant parameter. Another very similar task is to predict some functional of a spatial process, for example an integral, which is the case of our study. The most natural approach for spatial predicting is the linear prediction, which is called *kriging* in spatial statistics.

Let a random sample of a spatial process be denoted by

$$\boldsymbol{Z} = (Z_1, \ldots, Z_n)^\top = (Z(\boldsymbol{s}_1), \ldots, Z(\boldsymbol{s}_n))^\top, \tag{10}$$

where $n > 1$, $n \in \mathbb{N}$ and $\boldsymbol{s}_1, \ldots, \boldsymbol{s}_n \in D$ are the location points of the sample. Furthermore suppose that $EZ(\boldsymbol{s}) = \mu$ for all $\boldsymbol{s} \in D$, μ is unknown. Matheron suggested the following 'natural' variogram estimator:

Definition. Let $N(\boldsymbol{h}) = \{(i,j)\colon\ \boldsymbol{s}_i - \boldsymbol{s}_j = \boldsymbol{h};\ i,j = 1,\ldots,n\}$ and let $|N(\boldsymbol{h})|$ denote the number of distinct pairs in $N(\boldsymbol{h})$. Then the variable

$$2\hat{\gamma}(\boldsymbol{h}) = \frac{1}{|N(\boldsymbol{h})|} \sum_{(i,j)\in N(\boldsymbol{h})} (Z(\boldsymbol{s}_i) - Z(\boldsymbol{s}_j))^2, \quad \boldsymbol{h} \in D \tag{11}$$

is called the *classical variogram estimator*.

If the assumption of the constant trend is not fulfilled, first the trend shape must be modeled, its values estimated and subtracted from the data. Only then the classical variogram estimator can be used.

2.3. Kriging-spatial prediction.

We are approaching one of the main objectives of spatial statistics—the spatial prediction. We will consider one general type of predictor—the linear predictors, e.g. predictors in the form

$$\hat{Z}(\boldsymbol{s}) = \lambda_0 + \sum_{i=1}^{n} \lambda_i Z(\boldsymbol{s}_i). \tag{12}$$

Generally the idea of kriging is to find a predictor $\hat{Z}(\boldsymbol{s}_0)$ of the value $Z(\boldsymbol{s}_0)$ so that $\hat{Z}(\boldsymbol{s}_0)$ is *unbiased* and the mean value

$$EL(Z(\boldsymbol{s}_0), \hat{Z}(\boldsymbol{s}_0)), \tag{13}$$

where $L(.,.)$ is a loss function, is minimal. Only the case of $L(z_1, z_2) = (z_1 - z_2)^2$, i.e. the quadratic loss function, will be considered here. Let us introduce a simplified notation: $Z_0 = Z(\boldsymbol{s}_0)$, $\hat{Z}_0 = \hat{Z}(\boldsymbol{s}_0)$, $\mu_i = \mu(\boldsymbol{s}_i)$, $i = 0, \ldots, n$, $\boldsymbol{\mu} = (\mu_1, \ldots, \mu_n)^\top$ and $\boldsymbol{\lambda} = (\lambda_1, \ldots, \lambda_n)^\top$.

There are three basic types of kriging considered in literature. The *simple kriging* method assumes that the trend is known. The *ordinary kriging* method assumes that the trend is constant but unknown. The most general method is the *universal kriging*. It assumes that the trend is a linear combination of known functions of the location and unknown parameters and can be formally written as

$$Z(\boldsymbol{s}) = \sum_{j=1}^{p} f_j(\boldsymbol{s})\beta_j + \delta(\boldsymbol{s}), \quad \boldsymbol{s} \in D, \tag{14}$$

where $\delta(.)$ is a zero-mean intrinsically (or second order) stationary spatial process, $\boldsymbol{\beta} = (\beta_1, \ldots, \beta_p)^\top \in \mathbb{R}^p$ is an unknown vector of parameters and the functions $f_0(.), \ldots, f_p(.)$ are known (usually they are polynomials up to a given order).

3. Model

The data investigated in this study are airborne gamma-ray data. To make an inventory estimator using this data we need an appropriate model for it. The nature of the random variation of the data has basically two sources. One is the Poisson-like dispersion caused by the counting process and the second is the variation of the inventory itself—it can be assumed to be a two-dimensional stationary random process. Radionuclides are not spread uniformly in the environment but tend to be concentrated in some types of soils. Therefore sharp boundaries of radioactivity concentration exist in the environment copying the boundaries of soils.

A complex model for environmental gamma ray data is given in [4]. The model assumes that the measured Poisson random variables are of the form

$$Z(\boldsymbol{s}_i) = \mu(\boldsymbol{s}_i) + X(\boldsymbol{s}_i) + e_i, \quad i = 1, \ldots, n \tag{15}$$

where $\boldsymbol{s}_i$ is the location of the i-th observation, $\mu(.)$ is a non-random trend, $X(.)$ is a zero-mean second-order stationary Gaussian process and e_i is an independent zero

mean random error. For reasons explained in greater detail in Section 5, we prefer to work with the variance stabilised process. A new random variable is introduced (see [7]) and its decomposition is proposed:

$$Y(\boldsymbol{s}_i) = \sqrt{Z(\boldsymbol{s}_i)} = \Lambda(\boldsymbol{s}_i) + U(\boldsymbol{s}_i) + e_i, \tag{16}$$

where $\Lambda(.)$ is a non-random trend, $U(.)$ is a zero-mean second-order stationary Gaussian process and e_i is an independent random error.

Let $Y(\boldsymbol{s})$ be a process and $\Lambda(\boldsymbol{s})$ a trend. We are going to make an assumption about the random process $U(\boldsymbol{s}) = Y(\boldsymbol{s}) - \Lambda(\boldsymbol{s})$, namely $EU(\boldsymbol{s}) = 0$ and $\operatorname{var}\{U(\boldsymbol{s})\} = \sigma^2$ for all $\boldsymbol{s} \in B$ independently of the location (for more detail see [6]). The process $U(s)$ is called the *residual process*. Finally, for some methods of statistical inference we will suppose that U is a *Gaussian second order stationary random spatial process*.

4. Kriging inventory prediction

Let $\boldsymbol{Z}$ be the sample as in (10). Denote $Y_i = Y(\boldsymbol{s}_i) = \sqrt{Z(\boldsymbol{s}_i)}$, $i = 1, \ldots, n$, $\boldsymbol{Y} = (Y_1, \ldots, Y_n)^\top$. The kriging inventory estimator is similar to the kriging spatial predictor. We are looking for the best linear estimator, e.g. an estimator of the form:

$$\hat{I}_K(B) = \lambda_0 + \sum_{I=1}^{n} \lambda_i Z_i = \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{Z} \tag{17}$$

where the vector $\boldsymbol{\lambda}$ is to be estimated in order to fit the criteria (18) (see [2]):

$$\begin{gathered} E\hat{I}_K(B) = EI(B) = I_0(B) \\ \operatorname{var}(I(B) - \hat{I}_K(B)) \quad \text{is 'minimal'}. \end{gathered} \tag{18}$$

The estimator of the trend and the covariance matrix is essential for successful calculation of kriging. After estimating the trend and the covariance matrix for the process Y, the trend of Z can be expressed quite easily as

$$EZ(\boldsymbol{s}) = EY^2(\boldsymbol{s}) = \operatorname{var}\{Y(\boldsymbol{s})\} + (EY(\boldsymbol{s}))^2 = \sigma^2 + \Lambda^2(\boldsymbol{s}), \quad \boldsymbol{s} \in B \tag{19}$$

where $\sigma^2 = C(\boldsymbol{s}, \boldsymbol{s})$. To calculate the covariance matrix for the vector $\boldsymbol{Z}$ is not so straightforward.

Denote by $\hat{I}_K(B)$ the kriging estimator of $I(B)$ and

$$\begin{aligned} I_0(B) = EI(B) &= E\int_B Z(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} = E\int_B Y^2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} = \int_B \{\sigma^2 + \Lambda^2(\boldsymbol{s})\}\,\mathrm{d}\boldsymbol{s} \\ &= \sigma^2\nu(B) + \int_B \Lambda^2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \end{aligned} \tag{20}$$

where $\nu(B)$ means the area of B. Let $\hat{I}_0(B)$ be an estimator of $I_0(B)$.

To avoid the calculation of covariance of Z_i's, the calculation of the inventory estimator will be based on the square root process Y. Then the kriging estimator takes on the form

$$\hat{I}_K(B) = \lambda_0 + \sum_{i=1}^{n} \lambda_i \sqrt{Z(\boldsymbol{s}_i)} = \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{Y}. \tag{21}$$

4.1. Estimating of $\hat{I}_0(B)$.

The integral $\hat{I}_0(B)$ has to be calculated for the kriging estimator. It can be estimated e.g. using Voronoi cells generated by points $\boldsymbol{s}_1 \dots \boldsymbol{s}_n$. Voronoi tesselation of the plane is a set of cells, where the Voronoi cell belonging to the generators $\boldsymbol{s}_i$ is the set $\{\boldsymbol{s}, \|\boldsymbol{s} - \boldsymbol{s}_i\| < \|\boldsymbol{s} - \boldsymbol{s}_j\| \text{ for all } i \neq j\}$. An example is shown in Fig. 3. The tesselation covers the whole plane (the cells generated by the boundary points are unbounded) and then it is limited by a suitably selected window, which represents the set $\boldsymbol{B}$.

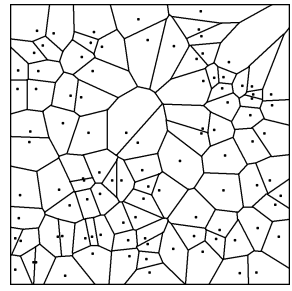

Figure 3. The Voronoi tesselations.

Let $\mathscr{T} = \{B_1, \dots, B_n\}$ be a tesselation of B and $\boldsymbol{\nu} = (\nu(B_1), \dots, \nu(B_n))^\top$ the vector of areas. Then the last integral in (20) can be estimated as

$$\int_B \Lambda^2(\boldsymbol{s})\, \mathrm{d}\boldsymbol{s} \approx \sum_{i=1}^{n} \Lambda^2(\boldsymbol{s}_i)\nu(B_i). \tag{22}$$

4.2. Calculating the kriging parameters.

The inventory has to be predicted in a way different from the ordinary kriging as a constant trend is not assumed. Since the structure of the trend is rather complicated, we have decided to use the simple kriging instead of the universal kriging and the trend was replaced by its linear estimate calculated by omitting the covariance structure of the data.

Because the mean value of the linear estimator is

$$E\hat{I}_K(B) = \lambda_0 + E\boldsymbol{\lambda}^\top \boldsymbol{Y} = \lambda_0 + \boldsymbol{\lambda}^\top \boldsymbol{\Lambda}, \tag{23}$$

where $\boldsymbol{\Lambda} = (\Lambda(\boldsymbol{s}_1), \dots, \Lambda(\boldsymbol{s}_n))^\top$, together with condition (20) it provides the condition

$$\lambda_0 = \hat{I}_0(B) - \boldsymbol{\lambda}^\top \boldsymbol{\Lambda}. \tag{24}$$

Let $\mathsf{W} = (c_{ij})$ be the covariance matrix of the vector $\boldsymbol{Y}$, $c_{ij} = \operatorname{cov}\{Y(\boldsymbol{s}), Y(\boldsymbol{t})\}$. We will calculate the variance of $I(B) - \hat{I}_K(B)$ now:

$$\begin{aligned}
\operatorname{var}(I(B) - \hat{I}_K) &= \operatorname{var}\{I(B)\} - 2\operatorname{cov}\{I(B), \hat{I}_K\} + \operatorname{var}\{\hat{I}_K\} = \\
&= \int_B \int_B \operatorname{cov}\{Z(\boldsymbol{s}), Z(\boldsymbol{t})\}\,\mathrm{d}\boldsymbol{s}\,\mathrm{d}\boldsymbol{t} - 2\operatorname{cov}\{I(B), \boldsymbol{\lambda}^\top \boldsymbol{Y}\} \\
&\quad + \operatorname{var}\{\boldsymbol{\lambda}^\top \boldsymbol{Y}\} \\
&= C_0 - 2\sum_{i=1}^{n} \lambda_i \int_B \operatorname{cov}\{Y(\boldsymbol{s}_i), Z(\boldsymbol{s})\}\,\mathrm{d}\boldsymbol{s} + \boldsymbol{\lambda}^\top \mathsf{W} \boldsymbol{\lambda} \qquad (25) \\
&= C_0 - 2\boldsymbol{\lambda}^\top \boldsymbol{c} + \boldsymbol{\lambda}^\top \mathsf{W} \boldsymbol{\lambda}, \qquad (26)
\end{aligned}$$

where the i-th component of the vector $\boldsymbol{c}$ is equal to

$$c_i = \int_B \operatorname{cov}\{Y(\boldsymbol{s}_i), Z(\boldsymbol{s})\}\,\mathrm{d}\boldsymbol{s} = \int_B \operatorname{cov}\{Y(\boldsymbol{s}_i), Y^2(\boldsymbol{s})\}\,\mathrm{d}\boldsymbol{s}. \tag{27}$$

To calculate the covariance inside the integral the following lemma will be needed.

Lemma 3. *Let $(X, Y)^T$ have a two-dimensional normal distribution*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \sigma^2 \begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix} \right). \tag{28}$$

Then

$$\operatorname{cov}\{X, Y^2\} = 2\varrho\mu_2\sigma^2. \tag{29}$$

P r o o f. The proof is elementary. □

Applying the previous lemma we get

Lemma 4. *Under the assumption of normality of the process Y, the kriging predictor of $I(B)$ is given by*

$$\hat{I}_K(B) = \sigma^2\nu(B) + \int_B \Lambda^2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} + \boldsymbol{c}^\top \mathsf{W}^{-1}(\boldsymbol{Y} - \boldsymbol{\Lambda}), \tag{30}$$

where the i-th component of the vector $\boldsymbol{c}$ is

$$c_i = 2\int_B C(\boldsymbol{s}_i, \boldsymbol{s})\Lambda(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}. \tag{31}$$

P r o o f. Applying Lemma 3 we obtain

$$\operatorname{cov}\{Y(\boldsymbol{s}_i), Y^2(\boldsymbol{s})\} = 2C(\boldsymbol{s}_i, \boldsymbol{s})\Lambda(\boldsymbol{s}), \tag{32}$$

which together with (27) gives

$$c_i = 2\int_B C(\boldsymbol{s}_i, \boldsymbol{s})\Lambda(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s}. \tag{33}$$

The vector $\boldsymbol{\lambda}$ is the one minimizing $C_0 - 2\boldsymbol{\lambda}^\top\boldsymbol{c} + \boldsymbol{\lambda}^\top\mathsf{W}\boldsymbol{\lambda}$, which occurs for $\boldsymbol{\lambda} = \mathsf{W}^{-1}\boldsymbol{c}$. The kriging predictor for $I(B)$ is then given by

$$\begin{aligned}\hat{I}_K &= \lambda_0 + \boldsymbol{\lambda}^\top\boldsymbol{Y} = I_0 - \boldsymbol{\lambda}^\top\boldsymbol{\Lambda} + \boldsymbol{\lambda}^\top\boldsymbol{Y}\\ &= I_0 + \boldsymbol{\lambda}^\top(\boldsymbol{Y} - \boldsymbol{\Lambda})\\ &= \sigma^2\nu(B) + \int_B \Lambda^2(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} + \boldsymbol{c}^\top\mathsf{W}^{-1}(\boldsymbol{Y} - \boldsymbol{\Lambda}).\end{aligned}$$

□

Now all the formulae necessary for calculating the kriging parameters have been expressed. Moreover, the constant C_0 has to be evaluated for estimating the kriging variance. To this end we need to know the covariances of $Z(\boldsymbol{s})$ and $Z(\boldsymbol{t})$. They are given by the following lemma.

Lemma 5. *Let $(X, Y)^T$ have a two-dimensional normal distribution*

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left(\begin{pmatrix} \mu_1 \\ \mu_2 \end{pmatrix}, \sigma^2\begin{pmatrix} 1 & \varrho \\ \varrho & 1 \end{pmatrix}\right). \tag{34}$$

Then

$$\operatorname{cov}\{X^2, Y^2\} = 4\varrho\mu_1\mu_2\sigma^2 + 2\varrho^2\sigma^4. \tag{35}$$

P r o o f. The proof is elementary. □

Lemma 6. *Under the assumption of normality of the process Y, the kriging variance is given by*

$$\sigma^2 = \int_B \int_B \{4C(\boldsymbol{s},\boldsymbol{t})\Lambda(\boldsymbol{s})\Lambda(\boldsymbol{t}) + 2C^2(\boldsymbol{s},\boldsymbol{t})\}\,\mathrm{d}\boldsymbol{s}\,\mathrm{d}\boldsymbol{t} + 2\boldsymbol{\lambda}^\top \boldsymbol{c} + \boldsymbol{\lambda}^\top \mathsf{W}\boldsymbol{\lambda}. \tag{36}$$

P r o o f. It follows from Lemma 5 that

$$\operatorname{cov}\{Y^2(\boldsymbol{s}), Y^2(\boldsymbol{t})\} = 4C(\boldsymbol{s},\boldsymbol{t})\Lambda(\boldsymbol{s})\Lambda(\boldsymbol{t}) + 2C^2(\boldsymbol{s},\boldsymbol{t}), \tag{37}$$

hence (25) yields

$$C_0 = \int_B \int_B \{4C(\boldsymbol{s},\boldsymbol{t})\Lambda(\boldsymbol{s})\Lambda(\boldsymbol{t}) + 2C^2(\boldsymbol{s},\boldsymbol{t})\}\,\mathrm{d}\boldsymbol{s}\,\mathrm{d}\boldsymbol{t}, \tag{38}$$

which together with (26) gives (36). □

4.3. Calculating the remaining integrals.

The same approach as for calculating the integral in equation (20) has been used for calculating the constant C_0 and the vector $\boldsymbol{c}$ which are necessary for estimating the kriging parameters and the kriging variance. The components of the vector $\boldsymbol{c}$ were estimated as

$$c_i = \int_B C(\boldsymbol{s}_i,\boldsymbol{s})\Lambda(\boldsymbol{s})\,\mathrm{d}\boldsymbol{s} \approx \sum_{j=1}^{n} C(\boldsymbol{s}_i,\boldsymbol{s}_j)\Lambda(\boldsymbol{s}_j)\nu(B_j) \tag{39}$$

and the integral C_0 as

$$\begin{aligned} C_0 &= \int_B \int_B \{4C(\boldsymbol{s},\boldsymbol{t})\Lambda(\boldsymbol{s})\Lambda(\boldsymbol{t}) + 2C^2(\boldsymbol{s},\boldsymbol{t})\}\,\mathrm{d}\boldsymbol{s}\,\mathrm{d}\boldsymbol{t} \\ &\approx \sum_{i=1}^{n}\sum_{j=1}^{n} \{4C(\boldsymbol{s}_i,\boldsymbol{s}_j)\Lambda(\boldsymbol{s}_i)\Lambda(\boldsymbol{s}_j) + 2C^2(\boldsymbol{s}_i,\boldsymbol{s}_j)\}\nu(B_i)\nu(B_j), \end{aligned}$$

but we have to be careful about the values of $C(\boldsymbol{s}_i,\boldsymbol{s}_j)$ for $i = j$. Whenever the non-zero nugget effect occurs, the covariance function has a discontinuity at this point. However, the measure of the set where discontinuity occurs is zero (in terms of integration measure) and, therefore, the value of C must be calculated without the nugget.

5. Application

The explored data were collected by a gamma-ray spectrometer attached to a helicopter. The helicopter flew over the observed area along nearly parallel lines with spacing about 500 m. The spectra have been collected usually each 5 seconds. After some data processing, the value of the observed radionuclide signal $Z(\boldsymbol{s}_i)$ has been estimated at each point of observation $\boldsymbol{s}_i \in B$, where B means the observed area.

The probabilistic distribution of the data themselves is determined by the process of collection. It is well known that the distribution of counts in the spectra measurement has almost ideal Poisson distribution. After a complicated procedure leading to the estimation of one particular radionuclide, we still can consider it as a (continuous) non-negative distribution the variance of which is proportional to the mean value. Although a non-Gaussian approach to kriging has been developed [3], we have preferred to find a transformation of the data which would make the data more Gaussian like.

Several kinds of data transformation were examined and the best result was obtained for a simple square root transformation. The reason is that the square root function stabilizes the variance of Poisson random variables as explained theoretically in [7].

The data can naturally come from several compartments which may mutually differ very much in the mean value of the radioactivity. The trend of the data is important for the kriging process and since it was neither constant nor continuous, we had to apply some method to determine the various compartments. To describe the method is beyond the scope of this article, we can just say that it was based on the assumption of a mixture of Gaussian distributions with a constant variance. Thus the transformed data may consist of several parts with different, asymptotically Gaussian distributed data with unknown means and variances, but the variance is the same for all the data.

The two areas of data collection have been mentioned in the introduction. The Ribble estuary is situated in north-west England, in the northern direction from Liverpool. The reason of interest in this area is the Springfields plant for nuclear fuel manufacturing, which is about 3 km north from the Ribble and is considered the main source of radioactivity contamination in the estuary. The airborne gamma ray data contain 2 800 entries. The area and the data are shown in Fig. 4.

The second area of interest is about 2–3 times bigger than Ribble estuary. It lies on the coast of the Irish Sea called Solway Firth, which is a bay surrounded by a corner of north-west Cumbria and south-west Scotland. One point of interest and also a potential pollutant in this area is Chapelcross nuclear power station. The

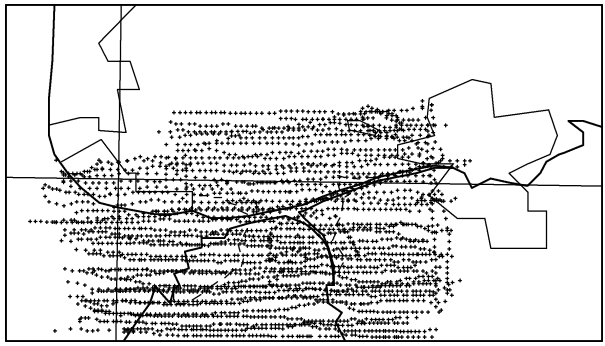

Figure 4. The data sampling points in Ribble estuary area (see Fig. 1).

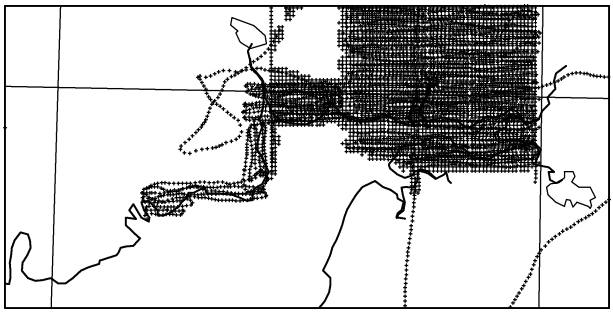

Figure 5. The data sampling points in Solway Firth area (see Fig. 2).

airborne gamma ray data contain 5 847 entries. The area and the data are shown in Fig. 5.

We have checked and compared the results with other methods for radioactivity inventory estimation. One method was the simple average multiplied by the surveyed area and the second was the regridding method. The regridding method is based on recalculating the data using an appropriate weighting function to a regular grid. Both the parameters of the grid and the weighting function were established intuitively. The results of the inventory estimators are shown in Tab. 1.

Site	Area (km^2)	Kriging (TBq)	Regridding Inv. (TBq)	Simple Mean Inv. (TBq)
Ribble	275.5	11 301	10 544	13 686
Solway	1 031.9	6 523	6 786	7 552

Table 1. The results of data processing.

We have made a set of simulated data in order to check the quality of the estimators. The results of all the three methods are comparable but the kriging provides an information on the confidence intervals. The 95% confidence intervals (10 366.456 TBq, 12 234.622 TBq) were for Ribble and (5 826.823 TBq, 7 219.915 TBq) for the Solway data.

6. Conclusion

The mathematical model in this study was very complex which inhibited us from using the standard form of kriging. Therefore a new method has been developed and tested on the real data sets. The method provides better results than two other methods used for comparison, however it is quite time consuming especially when the sample size is large (over 1 000 samples). Moreover, it provides an estimator of the confidence interval for the prediction, which the other methods do not. Further detailed information is available in [6].

References

[1] *M. Asli, D. Marcotte*: Comparison of approaches to spatial estimation in a bivariate context. Mathematical Geology *27* (1995), 641–658.

[2] *N. A. C. Cressie*: Statistic for Spatial Data. Wiley, New York, 1993.

[3] *P. J. Diggle, R. A. Moyeed and J. A. Tawn*: Non-gaussian geostatistics. Research note, Lancaster University, Technical Report MA95 (1996).

[4] *P. J. Diggle, J. A. Tawn and R. A. Moyeed*: Model-based geostatistics. J. Roy. Statist. Soc. Ser. C *47* (1998), 199–350.

[5] *U. C. Herzfeld, D. F. Merriam*: Optimization techniques for integrating spatial data. Mathematical Geology *27* (1995), 559–588.

[6] *P. Krejčíř*: The Theory and Applications of Spatial Statistics and Stochastic Geometry. PhD thesis, Charles University, Prague (2000).

[7] *R. J. Serfling*: Approximation Theorems of Mathematical Statistics. Wiley, New York, 1980.

[8] *M. L. Stein*: Predicting integrals of random fields using observation an a lattice. Ann. Statist. *23* (1995), 1975–1990.

Author's address: *P. Krejčíř*, Institute of Computer Science, Academy of Sciences of the Czech Republic, Pod vodárenskou věží 2, CZ-182 07 Praha 8, Czech Republic, e-mail: `pkrejcir@prague.ingr.com`.