# Appendix 5

# Spatial interpolation of settlement-average thyroid doses due to $^{131}I$ after the Chernobyl accident: 1. Feasibility study with $^{137}Cs$ deposition data in Belarus

# Spatial interpolation of settlement-average thyroid doses due to $^{131}I$ after the Chernobyl accident: 1. Feasibility study with $^{137}Cs$ deposition data in Belarus

A. ULANOVSKY[1,*,✉], R. MECKBACH[1], P. JACOB[1], S. SHINKAREV[2]

[1]GSF – National Research Center for Environment and Health,
Institute of Radiation Protection, 85764, Neuherberg, Germany
[2]State Research Center – Institute of Biophysics of the Ministry of Health,
123182, Moscow, Russia

## ABSTRACT

Measurements of $^{131}I$ activity in human thyroids were performed after the Chernobyl accident in Belarus. Settlements for such measurements were selected not randomly but depending on a level of the radioactive contamination. Then, a possibility to use geostatistical methods for spatial interpolation of settlement-average thyroidal activities from "measured" to "non-measured" settlements is under a question. To answer this question a feasibility study has been performed dealing with spatial sample determined by the geography of the thyroid measurements and $^{137}Cs$ deposition density values, which are known for both sample ("measured") and target ("non-measured") settlements. The feasibility study covers two distinct areas in Belarus: South-East (A) and East (B). To allow extrapolation from higher sample values to generally lower target ones, trends are estimated from the data using local regression techniques. Then classical kriging procedure is used to model spatially correlated residuals. Prediction is successful for the area A, while for the area B the interpolation is unsatisfactory. The results for both areas are compared and analyzed. Criteria and methods to detect potentially dangerous situations are discussed.

* On leave from: Joint Institute of Power and Nuclear Research – "Sosny", National Academy of Sciences of Belarus, 220109, Minsk, Belarus

✉ Corresponding author. Tel.: +49-89-3187-2789. Fax: +49-89-3187-3363. E-mail: ulanovsky@gsf.de

## INTRODUCTION

Radiation-induced childhood thyroid cancer is a major consequence for public health after the Chernobyl accident in 1986, therefore it appears as a subject of a number of epidemiological studies like case-control [1], cohort [2, 3], and population-based studies [4, 5, 6, 7]. Success of an epidemiological study depends on quality of dosimetric data, therefore considerable scientific efforts are concentrated now on verification and improvement of thyroid dose estimates and on assessment of their uncertainties.

The present paper describes the work done in support to a population-based risk study. While case-control and cohort type studies rely upon reconstruction of individual thyroid doses, the population-based study deals with the settlement-average thyroid doses based on historical measurements of $^{131}I$ in thyroid made soon after the accident. The so-called "measured" thyroid doses[1] are considered as a most reliable source for dose reconstruction. Consequently, the "measured" thyroid doses are preferably to be used in a subsequent risk assessment.

However, the "measured" doses are known for a limited set of settlements, only. A risk assessment study often needs in estimates of thyroid doses in the settlements where no direct measurements of thyroidal activity of $^{131}I$ had been made. A common way to assess thyroid doses in such settlements is to apply a radioecological (see e.g. [8, 9]) or 'semi-empirical' [10] models. Such approaches extensively use $^{137}Cs$ deposition density data as these data are well known from numerous spectrometric measurements made during years since the accident. Contrary, measurements of $^{131}I$ in the environment, because of its short half-life, are sparse and not sufficient for the thyroid dose assessment using radioecological approaches.

Another approach exists, which is based on statistical methods originated from geology and mining and known as geostatistical methods. These provide a way to interpolate spatial data taking into account observed spatial correlation between data points. That is, having a spatial sample of the settlement-average thyroid doses derived from direct measurements and estimating (or assuming) certain statistical

---

[1] One should understand, however, that these doses are computed from a measured activity of $^{131}I$ in the human thyroid at a given day using individual- and settlement-specific information on parameters of the intake pathway. Such individual- and settlement-specific information is often derived from questionnaires and other sources or based on expert judgment or even implied. Obviously, this means that all uncertainties related to the reconstruction of the individual pathway parameters contribute into the uncertainty of "measured" thyroid doses.

properties of this sample, one can assess average thyroid doses and their uncertainties for the settlements without “measured” thyroid doses, located in the vicinity of the sample ones.

However, a strong concern in applicability of geostatistical methods emerges as soon as one recognizes the fact that the settlements to perform thyroid measurements were selected not randomly but rather based on level of environmental radioactive contamination. In other words, one can say that the observed spatial sample is a preferential one. Working with such sample could lead to a systematic overestimation bias in predictions.

To assess predictive capabilities of the geostatistical methods in this case, it is decided to conduct a feasibility study. The feasibility study has been performed with spatial sample of “measured” settlements, however interpolation and prediction are made for the ground deposition density of $^{137}$Cs, $q_{137}$ , instead of thyroid doses due to $^{131}$I. Unlike the thyroid doses, the values of $q_{137}$ are known for both “measured” and “non-measured” settlements and a possibility exists to compare predictions with known values. The present paper describes this feasibility study and its results for the settlements in Belarus.

Nonetheless, one has to understand that spatial distributions of $^{131}$I and $^{137}$Cs need not to be identical. Generally, mechanisms of release from the damaged reactor, of atmospheric transport and deposition onto the ground are different for iodine and cesium. Because of these differences the isotopic ratios $^{131}$I/$^{137}$Cs are observed to vary spatially [10].

Spatial pattern of $^{137}$Cs deposition looks very “patchy”, with high gradients due to the fact that $^{137}$Cs was transported in the atmosphere mainly in aerosol form and any precipitations resulted in spots of high contamination. Atmospheric transport of $^{131}$I had occurred in different forms, of which aerosol form was only a fraction. Therefore, it is anticipated that spatial pattern of $^{131}$I deposition is not as strongly influenced by precipitations as the pattern of $^{137}$Cs was. One can expect smoother spatial distribution of $^{131}$I on the ground.

Moreover, the spatial distribution of $^{131}$I ground deposition is not the quantity of interest for a risk assessment study. Instead, average thyroid dose in settlement is a quantity to be interpolated. Because of the main pathway for thyroid exposure to $^{131}$I was a consumption of milk it appears plausible that thyroid dose values can be regarded as averaged over a larger spatial area than a settlement area, i.e. including pastures. Therefore, a spatial distribution of average thyroid doses seems to be even smoother and less variable on spatial scale than $^{131}$I ground deposition pattern.

## MATERIALS AND METHODS

### Description of the data

As mentioned above, the present feasibility study deals with mean values of $^{137}$Cs ground deposition density in settlements [11, 12]. These data are known [12] for practically all settlements of interest. The study deals with 460 sample settlements in Belarus. Settlements to predict in are called "target" settlements. Candidates for target settlements have been selected from those "non-measured" settlements (having less than 11 measured individuals) located within 30-km distance from any "sample" point.

Geographical coordinates of settlements are transformed to Transverse Mercator coordinates using the custom set of geographic projection parameters: ellipsoid – Krassovsky, central meridian – 28°E, zone width – 10°, scaling factor – 1.000. For convenience the coordinates are expressed in *km* relative to Chernobyl nuclear power plant. A coordinate transformation algorithm is borrowed from the program GSRUG publicly available on-line[2].

### Spatial distributions of the data points

Spatial distribution of the sample settlements is presented in Fig. 1, which shows spatial locations and indicates deciles of $q_{137}$ distribution by size and color of the points. It is seen that the sample is split in two distinct groups of spatial points. The first group, indicated as "group A", represents 308 populated places located in the vicinity of the Chernobyl power plant on the territory of Belarus. These settlements belong to Bragin, Khoiniki, Loev, El'sk, and Narovlya rayons of the Gomel oblast. The second group, indicated as "group B", is formed by 152 settlements in the Northern part of Gomel and Southern part of Mogilev oblasts. The both groups have similar spatial extensions and differ in size (the number of settlements). The group B can be characterized as poorly sampled than the group A.

---

[2] GSRUG – Geodetic Survey Routine: UTM/TM and Geographic. National Resources Canada: Geodetic Survey Division. Geomatics Canada, 615 Booth Street, Ottawa, Ontario, Canada, K1A 0E9. Available on-line:
www.geod.nrcan.gc.ca/index_e/products_e/software_e/gsrug_e.html Assessed: 9.02.2004

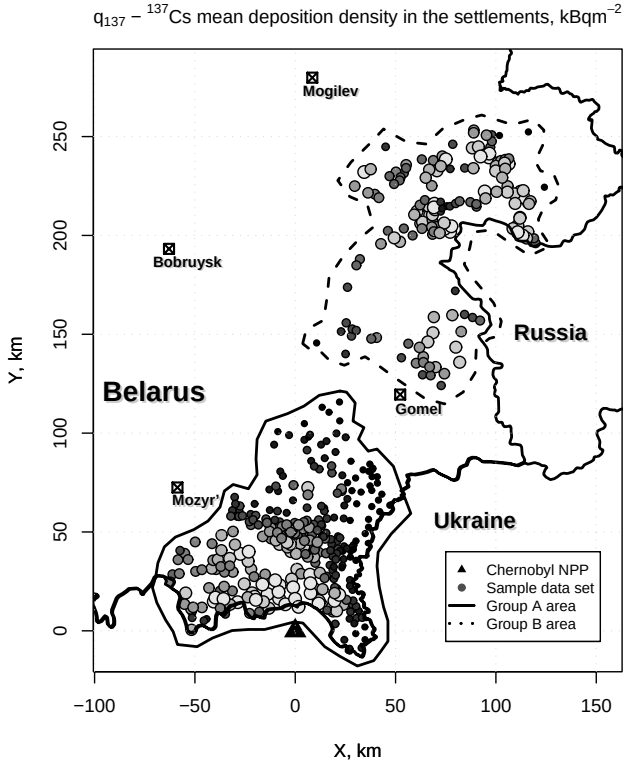


Fig. 1. Sample location map. Distribution of $^{137}$Cs ground deposition density in sample settlements in Belarus. Deciles of the distribution are marked with size and color of the sample points.

"Non-measured" settlements which fall within 30-km range from any sample point have been selected as targets for the feasibility study. Spatial distribution of the exhaustive (sample and target) data set is shown in the Fig. 2. Exhaustive data set in the Fig. 2 demonstrate that spatial distribution cannot be considered as completely random. One can see some systematic behavior between data points. It is important to note that sample points corresponding to the group A reveal systematic tendency whilst the sample points in the group B area barely demonstrate systematic seen in the exhaustive data set.

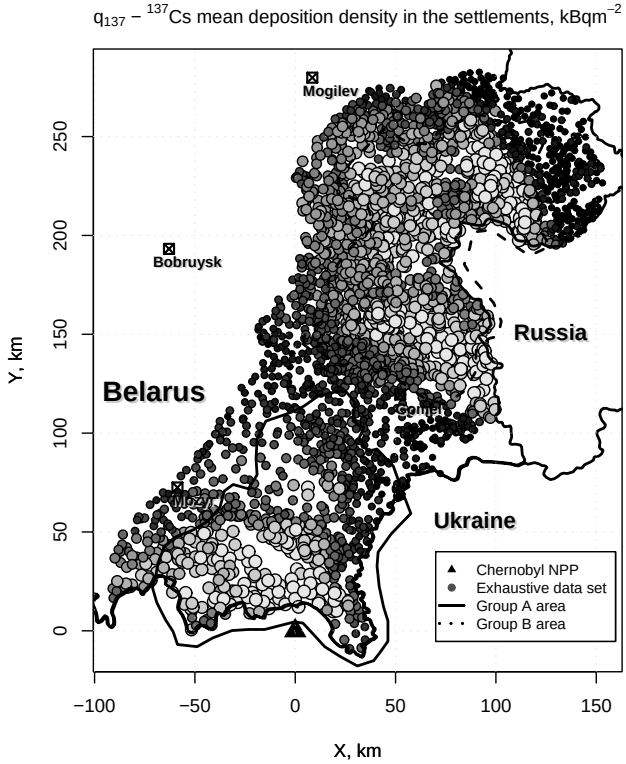


Fig. 2. Exhaustive location map. Distribution of $^{137}$Cs ground deposition density in sample and target (within 30-km range) settlements in Belarus. Deciles of the distribution are marked with size and color of the points.

As mentioned above, the sample settlements had been pre-selected based on the level of radioactive contamination. This is illustrated by the data in the Table 1 below, where given are statistical characteristics of the sample and the target distributions. These data show the distributions are non-Gaussian and right-skewed. The characteristics in the table are given for all target points and for sub-sets separated according to a value of minimum distance to any sample point. It seen from the table that the target values are as lower as larger the distance. It is worth mentioning that there exists a group of target settlements, which can be attributed to both lists (A and B) of sample points. The total number of such points is 213, of which 65 points belong to target sub-group in range from 10 to 20 *km*, and 149 points – to a sub-group in range from 20 to 30 *km*.

Table 1. Statistical summaries of sample and target distributions for groups A and B. Shown are values of the $^{137}$Cs ground deposition density, kBq m$^{-2}$

| Data set | N | min | 25%-ile | median | mean | 75%-ile | max |
|---|---|---|---|---|---|---|---|
| Group A | | | | | | | |
| Sample | 308 | 9.9 | 115 | 265 | 707 | 707 | $1.6\times10^4$ |
| Target (all) | 654 | 4.1 | 43 | 75 | 126 | 122 | $5.1\times10^3$ |
| Target (0–10 km) | 266 | 7.7 | 53 | 90 | 190 | 168 | $5.1\times10^3$ |
| Target (10–20 km) | 175 | 12 | 43 | 65 | 88 | 98 | 660 |
| Target (20–30 km) | 213 | 4.1 | 30 | 72 | 77 | 105 | 348 |
| Group B | | | | | | | |
| Sample | 152 | 13 | 328 | 664 | 834 | $1.2\cdot10^3$ | $3.3\times10^3$ |
| Target (all) | 2049 | 1.5 | 45 | 159 | 269 | 303 | $2.5\times10^3$ |
| Target (0–10 km) | 958 | 1.5 | 123 | 259 | 390 | 477 | $2.5\times10^3$ |
| Target (10–20 km) | | | | | | | |

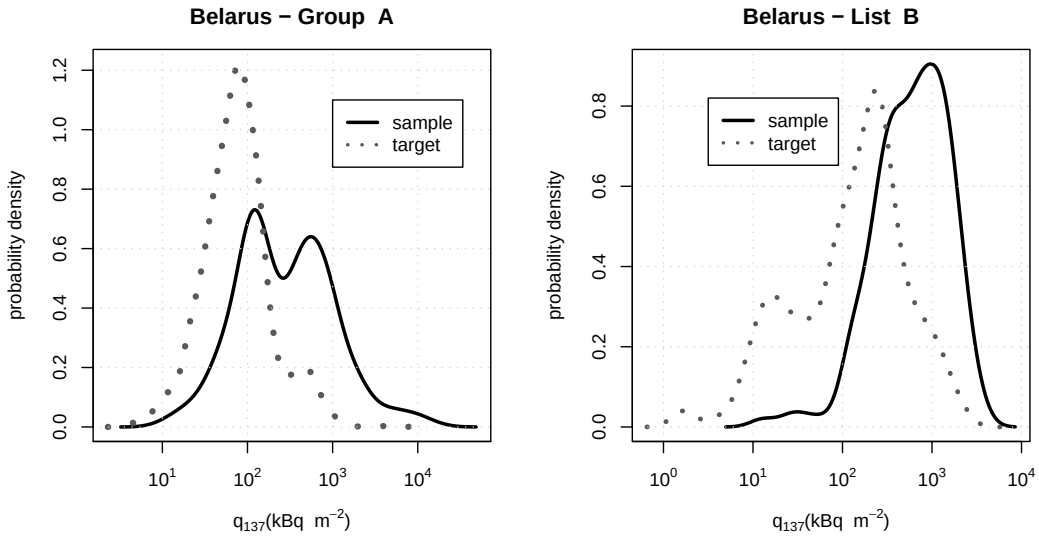


Fig. 3. Probability density distributions of the $^{137}$Cs ground deposition density for the groups A (left) and B (right) settlements in Belarus.

## Geostatistical concepts

Since 1950s, the geostatistics had been rapidly developing, mainly, due to its applications for estimation of natural reserves and mining. Nowadays, there is an extensive set of the geostatistical methods and numerous developments and generalizations are still under way. A thorough review of them is out of the scope of the paper, therefore a short summary of basic geostatistical terms and concepts relevant to the procedure adopted in the present study is described below. Of the extensive literature on this subject the books [13, 14, 15, 16, 17, 18] can be advised as comprehensive references for the current state of geostatistics.

*Correlated data*

Geostatistics deals with random spatially correlated processes defined in a spatial (1-, 2-, or 3- dimensional) domain. Consider a spatial random process, $Z$, defined in a two-dimensional domain, $\mathsf{D}$. Here, $x$ is a two-dimensional coordinate vector of a point in the spatial domain. Realizations of this process, $z(x_i)$, are observed in sampling points $x_i \in \mathsf{D}$, where $i = 1,\ldots,N$. The goal of geostatistical estimation is to predict an expectation and a variance of the random process in a target point, $x_0$, based on the observed realizations $z(x_i)$.

*Intrinsic and second-order stationarity*

Concept of stationarity of a random process plays an important role in geostatistical methods. The random spatial process is said to be second-order (or weak) stationary if

$$\mathrm{E}[Z(x)] = \mu \tag{1}$$

and its covariance depends only on a vector distance between points, $h$,

$$\mathrm{cov}[Z(x), Z(x+h)] = 2C(h). \tag{2}$$

If eq. (2) holds for scalar Euclidean distance, $h = \left\| x_i - x_j \right\|$, $i, j = 1,\ldots,N$, then the random process is said to be isotropic, also. A common name for the vector $h$ is "lag".

Widely used in geostatistics is an intrinsic hypothesis, which corresponds to a second-order stationarity of increments

$$\mathrm{E}[Z(x) - Z(x+h)] = 0 \tag{3}$$

$$\mathrm{var}[Z(x) - Z(x+h)] = 2\gamma(h) \tag{4}$$

where $\gamma(h)$ is called "semi-variogram".

*Variogram and semi-variogram*

The eq. (4) is a definition of the semi-variogram, $\gamma$. Under conditions of intrinsic stationarity (3) the variogram (4) becomes

$$2\gamma(h) = \mathrm{E}\left[(Z(x) - Z(x+h))^2\right] \tag{5}$$

and under the second-order stationarity conditions (eqs. (1) and (2)) the semi-variogram can be expressed through the covariance function

$$\gamma(h) = C(0) - C(h)\,. \tag{6}$$

An advantage of using the variogram to characterize spatial correlation is due to the fact that in the presence of a spatial trend the covariance function may not be defined while the variogram still can exist and can be used as a measure of spatial correlation between points.

In the present feasibility study variograms are calculated using *classical* estimator [19]

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} \left[z(x_i + h) - z(x_i)\right]^2\,. \tag{7}$$

Summations in eq. (7) are done over all points separated by the lag vector $h$. The lag values are grouped into arbitrary defined bins and $N(h)$ is a number of point pairs within the lag bin, therefore the eq. (7) represents a variogram in a *binned* form. Empirical variograms are fit by appropriate type of a *theoretical* variogram. There is a reach variety of functions legitimate for variogram modeling (see e.g. [14, 19, 20]). To mention just a few: *nugget* variogram for pure stochastic data

$$\gamma(h) = \begin{cases} 0, & if\ h = 0 \\ c, & \text{otherwise} \end{cases} \tag{8}$$

and *spherical* variogram

$$\gamma(h) = c \cdot \mathrm{Sph}\left(\frac{h}{a}\right) = \begin{cases} c \cdot \left[1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3\right], & if\ h \le a \\ c, & if\ h > a \end{cases} \tag{9}$$

where the parameters $a$ and $c$ are called *range* and *sill*, respectively. It is often that the empirical variogram can be represented as a sum of several theoretical models, e.g. sum of nugget and spherical variogram. Such models are called *nested*.

*Kriging*

The random spatial process is modeled as a sum of deterministic and stochastic part

$$Z(x) = m(x) + Y(x).$$

The deterministic part, $m(x)$, is commonly approximated by global polynomial surface – "trend surface"

$$m(x) = \sum_{l=0}^{L} a_l f_l(x),$$

where $f_l(\cdot)$ are power functions of independent spatial coordinates. Then, a kriging estimate of the process $Z(x)$ in a prediction point, $x_0$, is built as a weighted sample mean

$$z^*(x_0) = \sum_{i=1}^{N} \lambda_i z(x_i), \tag{10}$$

where $\lambda_i$ are *kriging weights*. The estimation must meet requirements of minimum square error

$$\mathrm{E}\left[z^*(x_0) - Z(x_0)\right]^2, \tag{11}$$

and unbiasedness

$$\mathrm{E}\left[z^*(x_0) - Z(x_0)\right] = 0. \tag{12}$$

The above requirements result in the so-called "universal kriging system", which can be written in the following form (see details in [18], p. 166)

$$\begin{aligned} &\sum_{j=1}^{N} \lambda_j \gamma(x_i, x_j) + \sum_{l=0}^{L} \mu_l f_l(x_i) = \gamma(x_i, x_0) \\ &\sum_{i=1}^{N} \lambda_i f_l(x_i) = f_l(x_0) \qquad i = 1,\ldots,N \quad l = 0,\ldots,L \end{aligned} \tag{13}$$

and the kriging variance

$$\sigma_K^2 = \sum_{i=1}^{N} \lambda_i \gamma(x_i, x_0) + \sum_{i=0}^{L} \mu_l f_l(x_0). \tag{14}$$

where $\mu_l$ are Lagrange multipliers, which appear as a result of the constrained minimization procedure.

The eq. (13) is a general form for major types of kriging, namely, *simple* kriging (corresponds to $L = 0$ and $\mu_0 = 0$), *ordinary* kriging ($L = 0$ and $\mu_0 \neq 0$), and *universal* kriging ($L \geq 1$). Recently, one can see attempts to reformulate geostatistical problems in terms of traditional statistical models: the paper [20] introduced a concept of the «model-based» geostatistics. The latter methodology is put into the core of the software used in the present study – GEOR library [22], which is a special software library for the R programming language [23].

**Representation of the data**

The sample data, $\{q_i : i = 1,\ldots,n\}$, are considered as realizations of a spatial random process, $Q(x)$, in the sample points, $x_i$, located in a two-dimensional spatial domain. That is, $x_i$ is a set of vectors. Because of apparent log-normality, transformed data, $\tilde{q} = \ln(q)$, are considered as realizations of a Gaussian spatial process $\tilde{Q} = \ln(Q)$. The sample data demonstrate both systematic behavior and random fluctuations, thus the following model of the random process is assumed

$$\tilde{Q}(x) = m(x) + Z(x) = m(x) + Y(x) + \varepsilon\,, \qquad (15)$$

where $m(x)$ represents non-stochastic spatial component of the random process $\tilde{Q}(x)$ and called hereafter "trend"; $Z(x)$ is a stochastic part of the process, which can be separated into correlated and non-correlated components, $Y(x)$ and $\varepsilon$, respectively. A variance of the non-correlated component, $\mathrm{var}(\varepsilon) = \tau^2$, is called *nugget* in the geostatistical literature and can be interpreted as a combined result of micro-scale variations and a measurement error.

The sample data are considered as points; however they represent values of an average contamination in an area of a settlement. That is, there is a certain class of short-range correlated variations, which are "seen" in the data as completely stochastic and uncorrelated. These are referred as micro-scale variation thus stressing the fact that such variations have correlation range smaller that size of a spatial point. In the present study spatial points are averages over settlement area.

**Trend estimation**

Selection of a method to model the trend is an important step in the present study because of the preferential sampling demonstrated above. Common in geostatistical practice is to model a trend by global polynomial spatial regression and plug-in polynomial coefficients into the system of kriging equations (universal kriging or kriging with the drift [19]. This method has shortcomings common to all polynomial interpolation techniques [24], i.e. an estimation with polynomials of high degree is extremely dangerous (and not recommended, respectively) in case of extrapolation, while low-degree polynomials may not reproduce complicated systematic variations, especially in the case of spatial or volumetric pattern. To avoid these shortcomings the technique of local polynomial regression has been used in the present study, namely LOESS method [25]. Local regression is built for

every prediction point independently, using only $N_\alpha$ closest neighbor points from $N$ sample points. That is, in a prediction point, $x_0$, trend is estimated as

$$m(x_0) = \sum_{n=0}^{L} a_n(x_0) x_0^n \ , \tag{16}$$

where coefficients $a_n(x_0)$ are defined using weighted polynomial regression, i.e. minimization in the least square sense of the following expression

$$\sum_{i=1}^{N_\alpha} W\left(\frac{x_i - x_0}{\delta}\right)\left[\tilde{Q}_i - P_L(x_i - x_0)\right]^2 , \tag{17}$$

where parameter $\delta$ is called *bandwidth*, weight function $W(u)$ is defined in the following way [25]:

$$W(u) = \begin{cases} \left(1 - \|u\|^3\right)^3 & \|u\| < 1 \\ 0 & \text{otherwise} \end{cases} , \tag{18}$$

and $P_L(x_i - x_0)$ is a polynomial of degree $L$ (usually, low-degree polynomials are used, i.e. $L \le 2$). Another parameter, $\alpha$, is called *span*; it defines the fraction of the sample points involved in the local regression procedure: $N_\alpha = \alpha N$. If the bandwidth $\delta$ is a constant then due to variable spatial density of the sample points the span varies too and, consequently, varies support for local regression. Contrary, if the span is fixed then the bandwidth, defined as $\delta = \max(x_i - x_0), i = 1, \ldots, N$, varies from one prediction point to another.

A procedure of local polynomial fit results in estimates of the deterministic spatial trend in the sample points, $m^*(x_i)$; their uncertainty is expressed by variance of the stochastic residual, $Z(x)$ (see eq. (15)).

**Estimation of the residual**

If the residual data, $z(x_i)$, are correlated then they are subjects for further geostatistical analysis. Correlations within the residual points are revealed and characterized by means of the variogram (4). If residuals are intrinsically stationary then the variogram is used to construct the kriging predictor (10). Empirical variogram is fit by an appropriate theoretical variogram and the latter is plugged into the kriging systems (13). In the present study the trend is estimated independently by local regression techniques (see eqs. (16) – (18)) and subtracted from the data and the *ordinary* kriging procedure is applied to the residual data to get kriging prediction in target points, $z_K(x_0)$, as well as kriging variance, $\sigma_K^2(x_0)$.

### Back-transformation and analysis

The predicted data in the target points need to be back-transformed. This is done using standard formulas for log-normal kriging (see e.g. [13], p. 270):

$$q(x_0) = \exp\left(m^*(x_0) + z_K(x_0) + \tfrac{1}{2}\sigma_K^2(x_0)\right) \quad (19)$$

$$\sigma^2(x_0) = q(x_0)^2\left[\exp\left(\sigma_K^2(x_0)\right) - 1\right] \quad (20)$$

The kriging estimates are compared then to the known values for the target points and an observed variance of kriging prediction is compared with estimated kriging variance (20).

## RESULTS AND DISCUSSION

Raw variograms of the sample data are shown in the Fig. 4. From the figure it is seen that the group A variogram has no sill, i.e. it is unbounded. This indicates presence of a trend in the sample data. That is, the data are not stationary and kriging is not authorized for the raw data. On the other hand, the raw variogram for group B data is bounded, i.e. the sample data do not reveal systematic behavior present in the exhaustive data set (see exhaustive location map — Fig. 2). The group B raw variogram demonstrates higher variability (fluctuations in the sill) between bins as a consequence of a poorer sampling.

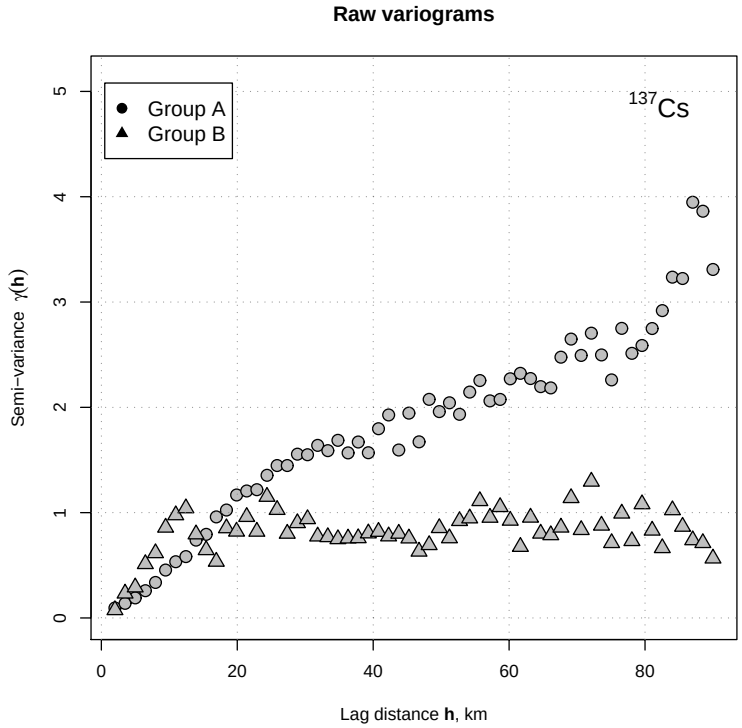


Fig. 4. Raw variograms of log-transformed sample data for the settlements in the groups A and B.

Because of preferential sampling specific to the data, it is important to assess underlying trend; however, as it can be seen from the raw variograms the group B (Fig. 4) the sample data does not demonstrate an obvious presence of a trend.

**Assessment of the trend**

Trend estimation is done by LOESS method [25]. Two most important parameters are the span, $\alpha$, and a degree of regressing polynomials, $L$,. There exist no unique quantitative recommendations on appropriate selection of these parameters ([24], p. 437).

Estimation for many target settlements falls into an extrapolation case; therefore, it appears safer to operate with low-degree polynomials. On the other hand, zero degree polynomials are simply average values and can not provide information on tendency existing in the data, which is important for extrapolation. Following this reasoning, the degree of smoothing polynomials has been chosen equal to 1.

Another difficult problem is a selection of an appropriate value for the span. For this, the following was taken into account:

- smoothness of the trend surface; the less span is used, the more variable and less smooth the trend surface;
- normality and stationarity of the residuals;
- range, sill, and nugget of residual variograms.

Contour plots of the trend surfaces for both lists A and B are shown in the Fig. 5 and 6. Both plots are made with the same values for span and degree of local regression. Residual variograms for different trend models are presented and discussed below.

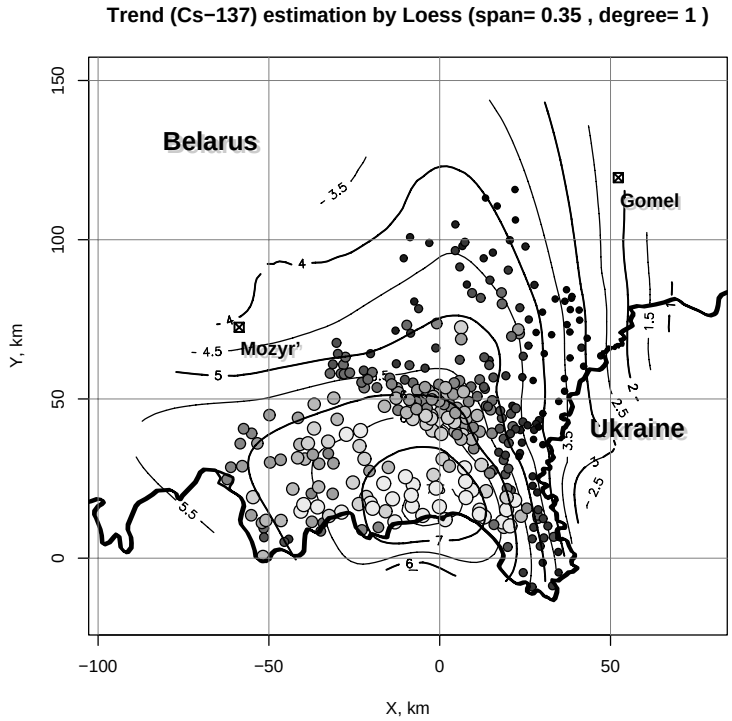


Fig. 5. Estimation of trend by local regression: group A. Points are the sample settlements. Color and size of the points represent deciles of the sample distribution.

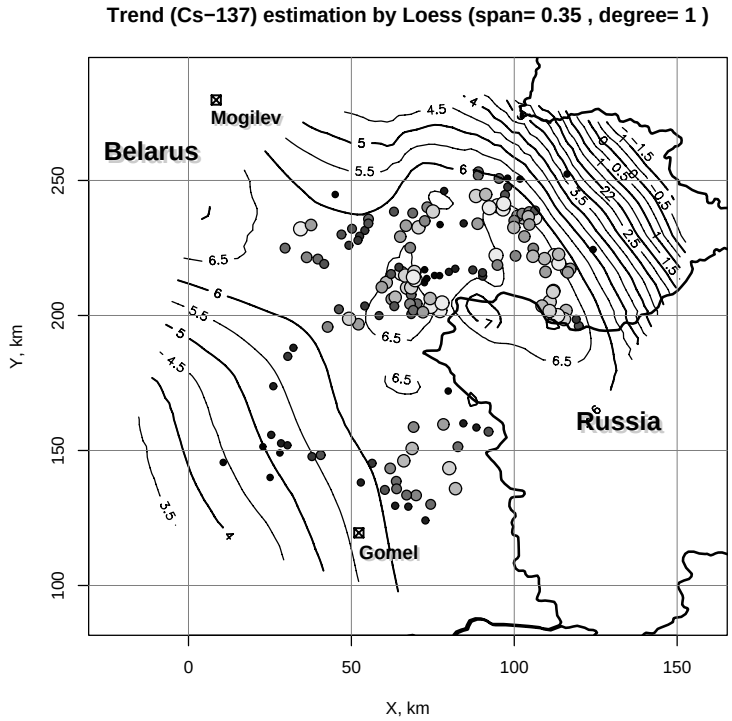


Fig. 6. Trend estimation by local regression: Group B. Points are the sample settlements. Color and size of the points represent deciles of the sample distribution.

**Modeling of the residuals**

Residuals after removal of the locally regressed trend are distributed normally and show no signs of multimodality seen in the raw data (Fig. 3). Also, the residuals are not independent; they demonstrate a spatial correlation. Shown in the following figures are residual variograms corresponding to trend models differing by the span value.

The Fig. 7 clearly demonstrates how the trend takes out variability from the data. One can see that decrease in value of $\alpha$ results in a shorter range of the correlation and smaller sill. Another important feature of this figure is an effect of sill stabilization as the span decreases. This means that residual variance beyond the correlation range is almost constant and does not depend on distance between sample points. A comparison of the raw variograms for the group A (Fig. 4) and the variogram for $\alpha = 1$ clearly shows effect of removal of a strong trend[3]. It is also remarkable that the variogram of residuals for $\alpha = 0.1$ shows very weak correlation. That means that the trend model takes out almost all correlated part of the variability, leaving out non-correlated (almost) "white noise".

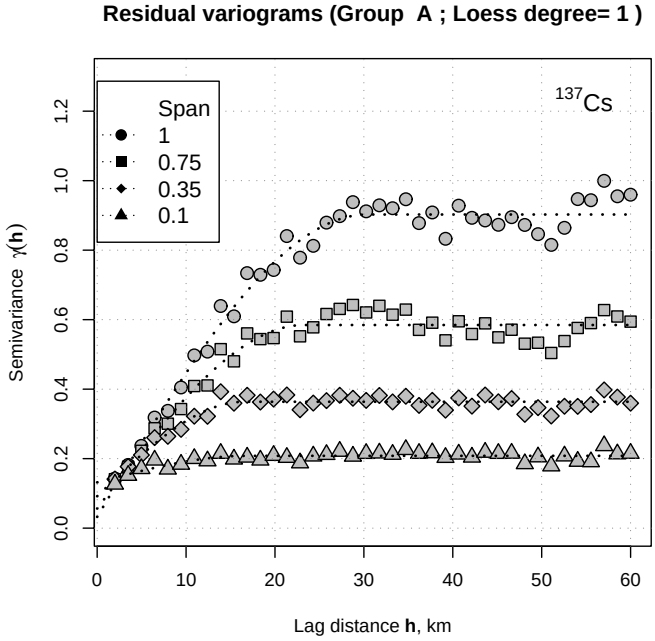


Fig. 7. Residual variograms for different trend models: group A

[3] Local regression with span $\alpha = 1$ uses all data points, however, weighting is done individually for every prediction point according to distance between the prediction and a sample point.

Residual variograms for the group B shown in the Fig. 8 demonstrate similar behavior as for the group A, i.e. the sill decreases as the span becomes smaller. However, an important difference from the previous figure is that large span trend models fail to model a trend in these data. One can see from the Fig. 8 that the variograms for the span values $\alpha \geq 0.75$ are practically identical and coincide with the raw variogram (Fig. 4). For the smaller span values ($\alpha \leq 0.55$) the figure demonstrates reduction of the sill and shortening of the correlation range. Strong fluctuations between variogram points are likely due to poor sampling statistics.

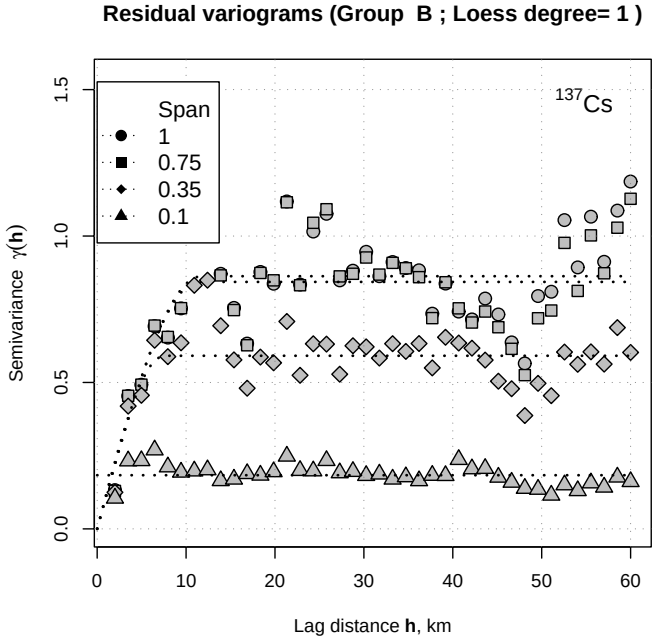


Fig. 8. Residual variograms for different trend models: group B

The residual variograms shown in the Fig. 7 and 8 are fit using spherical model for theoretical variogram (9). The residual variograms are taken isotropic as no meaningful anisotropy was found from angular variograms. Parameters of the fitted theoretical variograms are given in the Table 2.

Table 2. Parameters of the theoretical variograms for groups A and B.

| Group | Trend parameters | | Variogram parameters | | | $GSD_{total}$ | $GSD_{meas}$ |
|---|---|---|---|---|---|---|---|
| | | | Range, km | Sill $\gamma(\infty)$ | Nugget[a] $\gamma(0)$ | $\left(e^{\sqrt{\gamma(\infty)}}\right)$ | $\left(e^{\sqrt{\gamma(0)}}\right)$ |
| | $\alpha$ | $L$ | | | | | |
| A | 0.35 | 1 | 16.6 | 0.364 | 0.141 | 1.83 | 1.46 |
| B | 0.35 | 1 | 7.5 | 0.591 | 0.125 | 2.16 | 1.42 |

[a] 'Empirical' nugget, i.e. a value of the empirical variogram in the first bin: $\gamma(h_1)$

**Prediction by kriging**

Estimation of $^{137}$Cs deposition density in target points has been done by performing the following steps:

- normalization of the data by log-transform;
- estimation of the trend by local linear ($n = 1$) regression (eqs. (16) and (17)) with span $\alpha = 0.35$;
- validating spatial stationarity and isotropy of the residual and modeling of the residual variograms;
- spatial interpolation of the residual data and estimation of their uncertainty by kriging;
- back-transformation of the kriging results combined with trend estimates using eqs. (19) and (20).

The following figures (Figs. 9 and 10) illustrate how the trend model compensates a bias, which exists in sample data due to the preferential sampling. Shown in the Fig. 9 are probability distributions for group A data. Sample data (dashed line) are obviously biased compared to target data (solid line), however the predictions (dotted line) derived from the sample are distributed remarkably similar to the true target data.

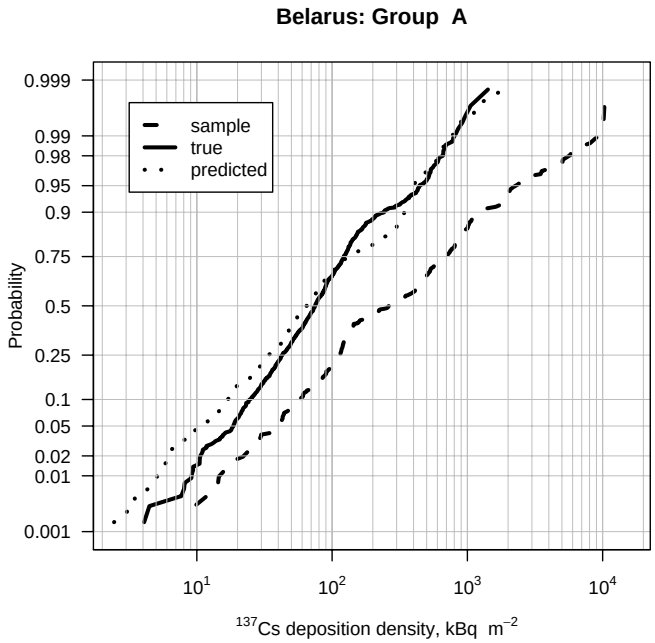


Fig. 9. Cumulative probability distributions for sample, true, and predicted values for group A settlements.

The group B results are drawn in the Fig. 10. The probability distribution of predicted values (dotted line) fails to reproduce the distribution of true values (solid

line), although it is shifted toward lower values compared to the sample one (dashed line). Then, one can say that the group B sample is poor and does not provide sufficient information for the trend assessment.

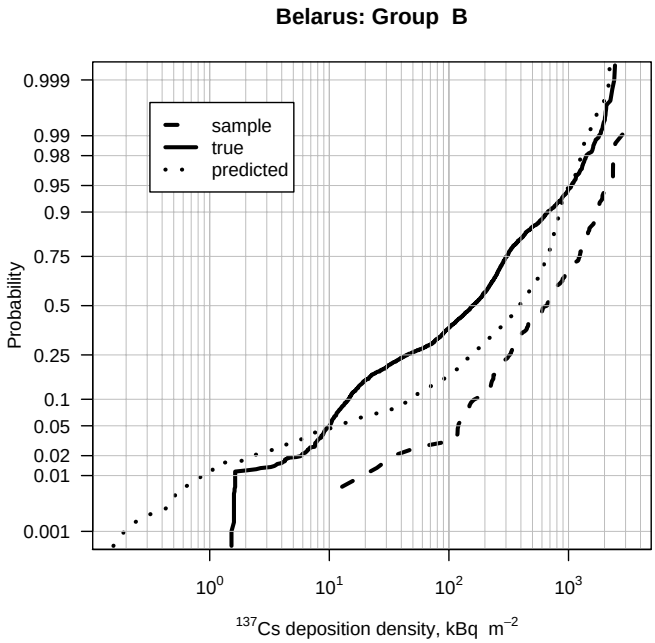


Fig. 10. Cumulative probability distributions for sample, true, and predicted values for the group B settlements.

This behavior can be compared with variograms shown in Figs. 4, 7, and 8. First, the raw variogram for the group A is unbounded; this demonstrates existence of systematic trend in the sample data. The raw variogram for the group B show correlation range approximately equal to 10 *km* and no evidence of a trend. The residual variograms (Fig. 7 and 8) support these observations. Namely, trend models of different spans for group A data effectively remove variability from the sample data. Residuals for the accepted trend model with span $\alpha = 0.55$ show correlation range of approximately 20 *km*. Contrary, residuals for the group B data are not affected by trend models unless for spans $\alpha \le 0.35$. Correlation range in these data is very short, less than 10 *km*. In other words, the sample data for the group B do not provide enough information for estimation of a trend and, consequently, trend model can be inadequate for prediction of the target data within 30-*km* range. It is likely, that the predictions made for the group B data will be better at target points located near (i.e. within the observed correlation range) the sample points. In the next sub-section the quality of predictions will be analyzed depending on the distance between target and sample points.

**Cross-validation**

Cross-validation is a general term which is applied generally to two different situations. One is a so-called "external" cross-validation when target points are completely different from the sample ones. Another method is a kind of resampling procedure, i.e. a point (or a group of points) is removed from the sample dataset and a prediction is made for the removed point based on the rest of the sample. This procedure is repeated for every sample point and then one can compare the distributions of the true sample and the predicted values. This method is known as "leaving-one-out" cross-validation. Unique feature of the present feasibility study is a possibility to compare the external and 'leaving-one-out' cross-validation techniques, because in the true values are known for the prediction (target) points.

For comparison purposes, the target points are characterized by minimum distance to any sample point, $H_{\min}$. This criteria is not sufficient to unambiguously distinguish between inter- and extrapolation cases, however it still provides some clue for this. Namely, one can say that the target points with higher $H_{\min}$ are more likely to fall in extrapolation case and *vice versa*.

The true and the predicted values are compared and analyzed by various ways. First, the predicted values are plotted against the true ones. All target points are divided into three subgroups depending on the value of $H_{\min}$: (a) 0…7 *km*; (b) 7…15 *km*; and (c) 15…30 *km*. The values 7 and 15 *km* are close to the ranges of the variograms for the groups B and A, respectively. Additionally, fourth sub-group is created from the "leaving-one-out" cross-validation data. Plots for the group A are presented in Figs. 10–13 and plots for the group B — in Figs. 14–17.

From the Fig. 10 one can see an example of a good agreement between predicted and true values for the group A settlements. For the target settlements in other subgroups of the group A the agreement becomes poorer as the $H_{\min}$ increases (see Figs. 11 and 12). The "leaving-one-out" cross-validation data (Fig. 13) provide another example of a good agreement while one can notice the $H_{\min}$ values match those for the first sub-group plotted in the Fig. 10. From the Figs. 10–13 one can conclude: (a) the larger distance from a prediction point to the sample ones the less adequate is a trend model and higher deviations of predictions from the true values; and (b) results of the cross-validation based on the sample points ("leaving-one-out") are in a good agreement with results of the external cross-validation for target points located within similar range of $H_{\min}$.

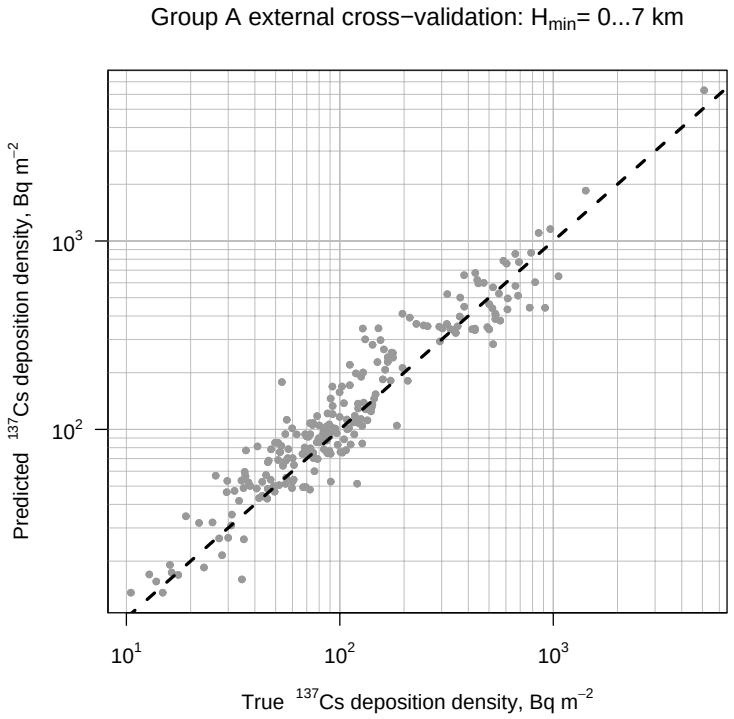


Fig. 10. External cross-validation plot for the group A target settlements ($0 \le H_{\min} < 7\ km$).

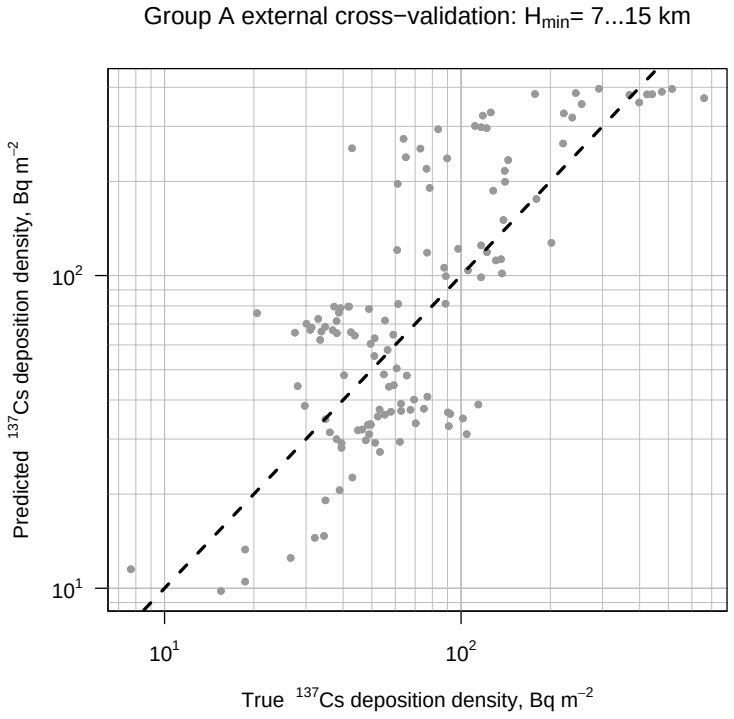


Fig. 11. External cross-validation plot for the group A target settlements ($7 \le H_{\min} < 15\ km$)

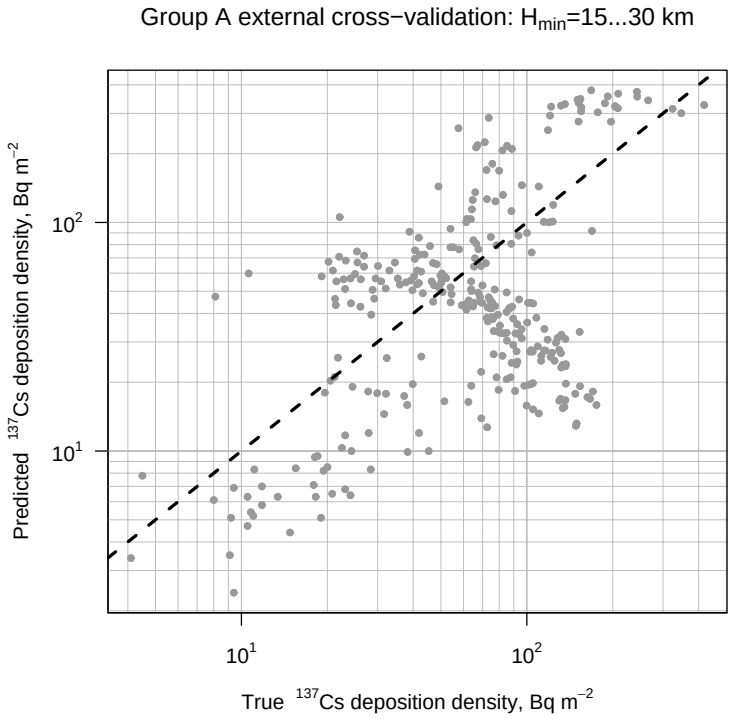


Fig. 12. External cross-validation plot for the group A target settlements ($15 \le H_{\min} < 30$ *km*)

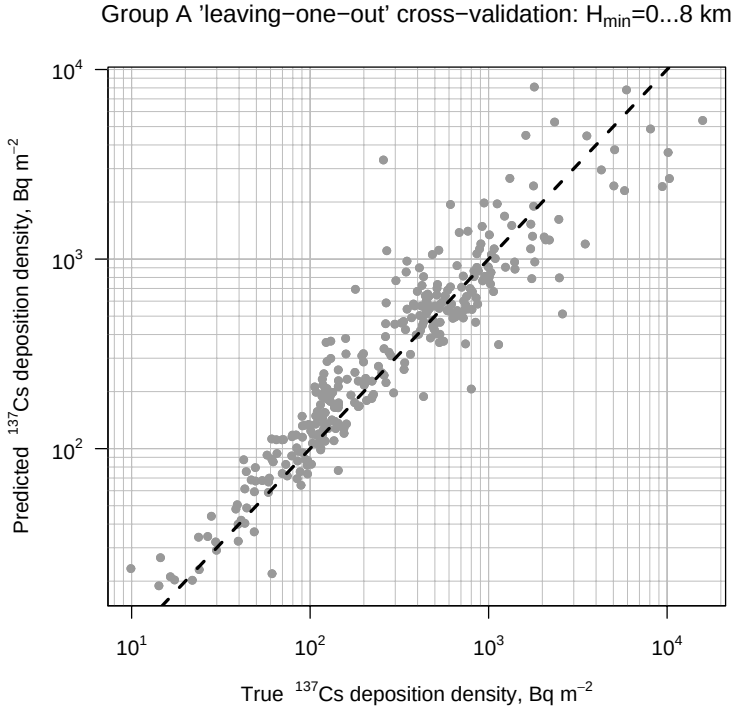


Fig. 13. "Leaving-one-out" cross-validation plot for the group A sample settlements ($0 \le H_{\min} < 8$ *km*)

The cross-validation results for the group B settlements differ considerably from those for the group A. One can see from the Figs. 14–17 the predicted values demonstrate systematic overestimation for all ranges of $H_{\min}$. The only reasonable agreement can be seen for the true values higher than $10^3$ Bq m$^{-2}$ within range $0 \le H_{\min} < 7\ km$ for external cross-validation (Fig. 14) and within range $0 \le H_{\min} < 14\ km$ for "leaving-one-out" cross-validation results (Fig. 17). For other sub-groups the overestimation is stronger as lower the true values (see Figs. 15 and 16). These observations agree with preliminary conclusions derived from the variogram analysis. Namely, the group B sample data represent mostly high contaminated places and their sampling pattern is sparse. As a result, they provide not enough information to derive the spatial trend to predict lower values.

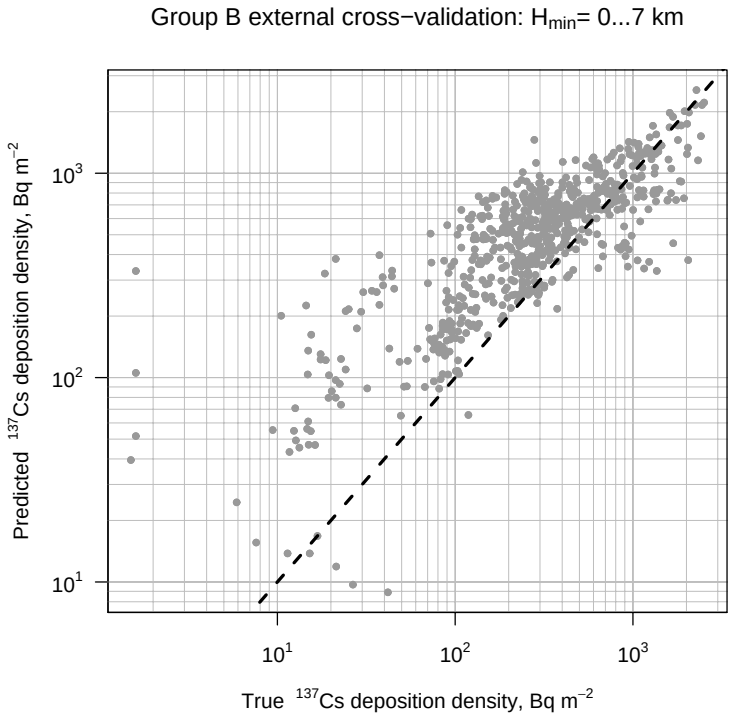


Fig. 14. External cross-validation plot for the group B target settlements ($0 \le H_{\min} < 7\ km$)

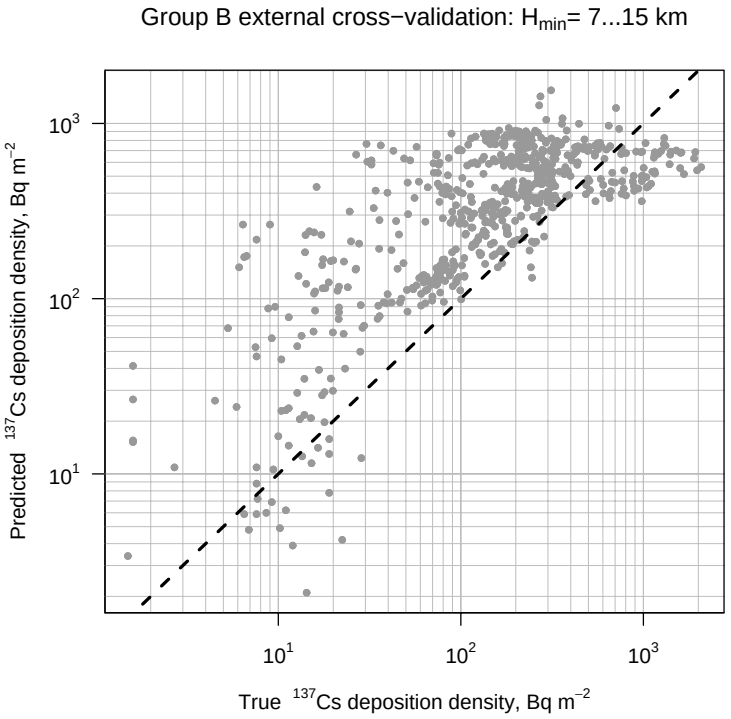


Fig. 15. External cross-validation plot for the group B target settlements ($7 \le H_{\min} < 15$ *km*)

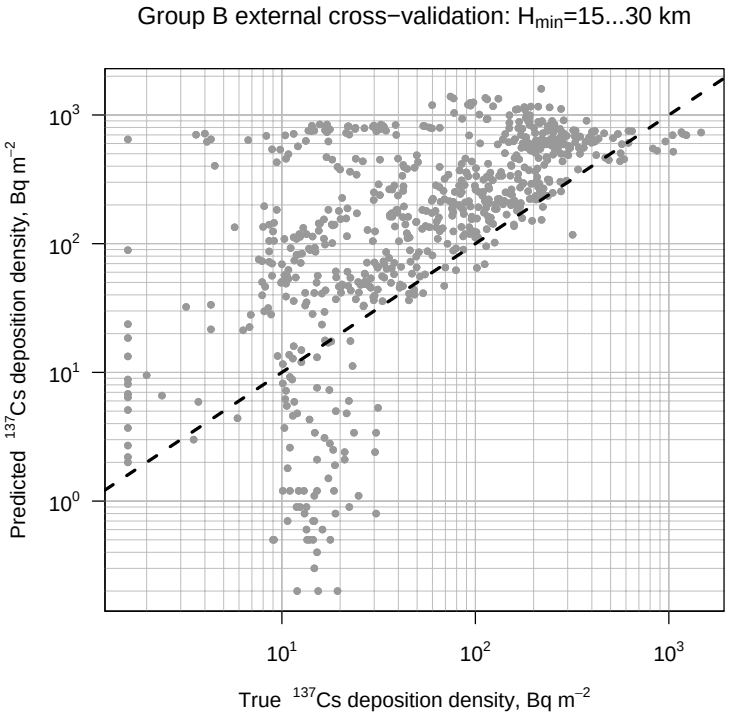


Fig. 16. External cross-validation plot for the group B target settlements ($15 \le H_{\min} < 30$ *km*)

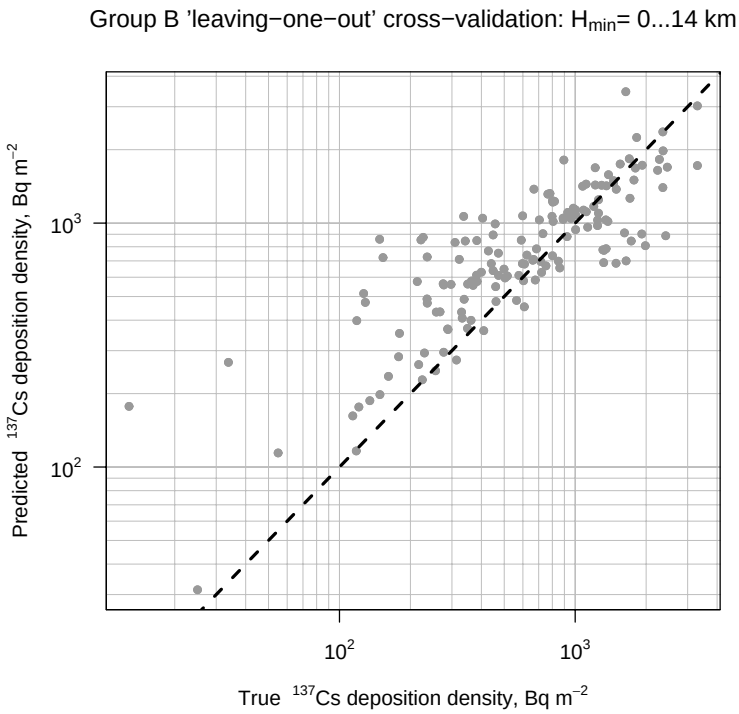


Fig. 17. "Leaving-one-out" cross-validation plot for the group B sample settlements ($0 \le H_{\min} < 14$ *km*)

Figures 10–17 provide a qualitative insight to the cross-validation results. To quantify them statistical properties of the distributions for the sub-groups with different $H_{\min}$ are given in the Table 3.

Table 3. Statistical properties of the cross-validation results. Quantiles of the distributions for the ratio "Predicted/True" as a function of $H_{\min}$ (minimum distance the target to any sample point)

| $H_{\min}$ range (km) | Quantile for probability: | | | | | | |
|---|---|---|---|---|---|---|---|
| | 0.025 | 0.05 | 0.25 | 0.5 | 0.75 | 0.95 | 0.975 |
| Group A: external cross-validation | | | | | | | |
| 0 ... 5 | 0.66 | 0.73 | 0.93 | 1.13 | 1.39 | 1.80 | 1.99 |
| 0 ... 10 | 0.54 | 0.59 | 0.89 | 1.09 | 1.40 | 1.92 | 2.15 |
| 0 ... 15 | 0.46 | 0.53 | 0.85 | 1.09 | 1.47 | 2.36 | 2.77 |
| 0 ... 20 | 0.30 | 0.38 | 0.75 | 1.07 | 1.46 | 2.45 | 3.21 |
| 0 ... 25 | 0.22 | 0.27 | 0.71 | 1.04 | 1.47 | 2.43 | 3.15 |
| 0 ... 30 | 0.16 | 0.22 | 0.60 | 1.01 | 1.50 | 2.56 | 3.20 |
| Group A: "leaving-one-out" cross-validation | | | | | | | |
| 0 ... 5 | 0.36 | 0.52 | 0.86 | 1.09 | 1.42 | 2.22 | 2.78 |
| 0 ... 8 | 0.35 | 0.53 | 0.86 | 1.11 | 1.42 | 2.24 | 2.78 |
| Group B: External cross-validation results | | | | | | | |
| 0 ... 5 | 0.54 | 0.64 | 1.07 | 1.46 | 2.18 | 4.16 | 5.3 |
| 0 ... 10 | 0.49 | 0.58 | 1.14 | 1.69 | 2.73 | 6.88 | 8.6 |
| 0 ... 15 | 0.46 | 0.54 | 1.19 | 1.77 | 2.95 | 7.58 | 12.5 |
| 0 ... 20 | 0.43 | 0.54 | 1.22 | 1.84 | 3.23 | 9.03 | 16.4 |
| 0 ... 25 | 0.37 | 0.51 | 1.25 | 1.92 | 3.32 | 10.55 | 24.1 |
| 0 ... 30 | 0.24 | 0.48 | 1.24 | 1.95 | 3.57 | 13.42 | 25.8 |
| Group B: "leaving-one-out" cross-validation | | | | | | | |
| 0 ... 5 | 0.47 | 0.57 | 0.93 | 1.17 | 1.52 | 2.68 | 3.81 |
| 0 ... 10 | 0.47 | 0.53 | 0.92 | 1.17 | 1.57 | 2.99 | 3.85 |
| 0 ... 14 | 0.47 | 0.54 | 0.93 | 1.18 | 1.59 | 3.50 | 4.21 |

The cross-validation results worth more in-depth quantitative analysis. As a measure of a deviation from the true values one can use a standardized prediction error (see e.g. [17, 18]). The standardized error

$$\delta_{std} = \frac{q_{pred} - q_{true}}{\sigma}$$

assumes a normal distribution for prediction around the true value, i.e. $\delta_{std} \propto N(0,1)$. It is shown above (see e.g. Fig. 3) that the sample data are apparently log-normal. That means that another quantity

$$\delta^*_{std} = \frac{m^* + z_K - \ln q_{true}}{\sigma_K} = \ln \frac{q_{pred}}{q_{true}} - \frac{1}{2} \ln \left( \frac{\sigma^2}{q^2_{pred}} + 1 \right) \qquad (21)$$

has the standard normal distribution. The latter quantity is plotted against $H_{\min}$ and $q_{true}$ in the Figs. 19–25. The standardized error (21) values are plotted as a point cloud in every of these figures. The solid line represents local estimate of the mean, while dashed lines are (local mean ± local standard deviation). Deviation of the local mean values from the zero line signals an existence of a bias in the predictions and serves as an indicator of a systematic error. If the local standard deviations exceed one this means the estimated kriging error does not represent the real error of the prediction.

The standardized errors are plotted for group A (Figs. 18–21) and group B (Figs. 22–25) cross-validation data. The Figs. 18, 19, 22, and 23 plot $\delta^*_{std}$ as a function of $H_{\min}$ while others (Figs. 20, 21, 24, and 25) are plots of $\delta^*_{std}$ vs. $q_{true}$.

One can see from the Figs. 18, 22, 23 that the standardized error of the cross-validations as higher as larger distance between the prediction and the sample points. The “leaving-one-out” data for the group A (Figs. 19) do not suggest such conclusions perhaps because of $H_{\min} < 8\,km$.

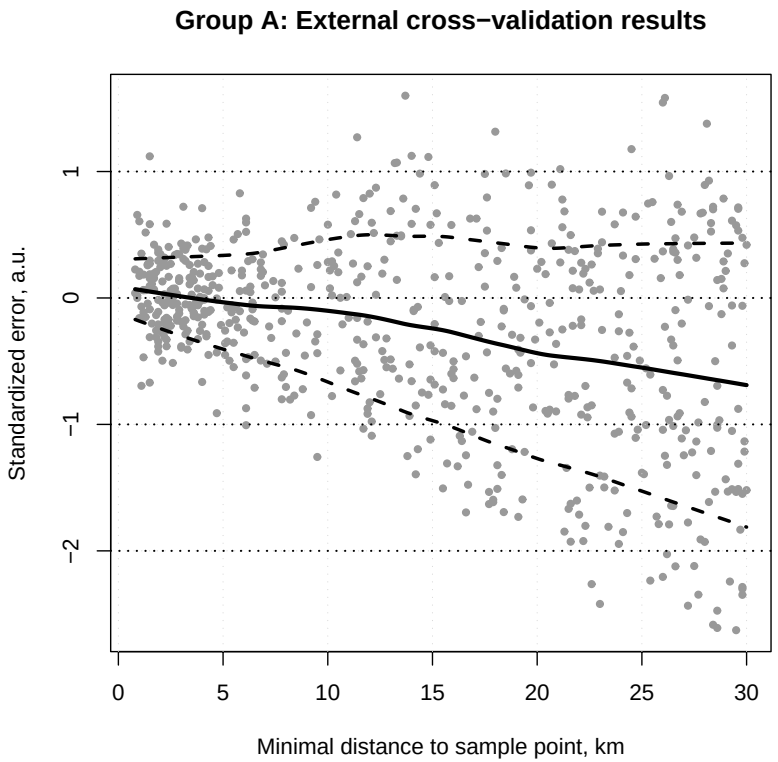
Group A: External cross−validation results
Standardized error, a.u.
1
0
−1
−2
0
5
10
15
20
25
30
Minimal distance to sample point, km

Fig. 18.

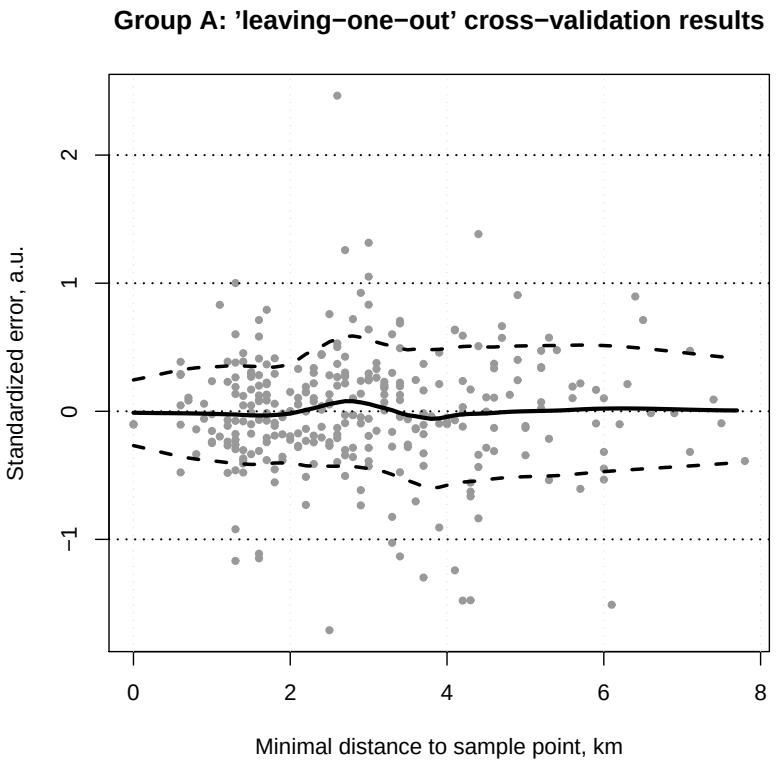
Group A: 'leaving−one−out' cross−validation results
Standardized error, a.u.
2
1
0
−1
0
2
4
6
8
Minimal distance to sample point, km

Fig. 19.

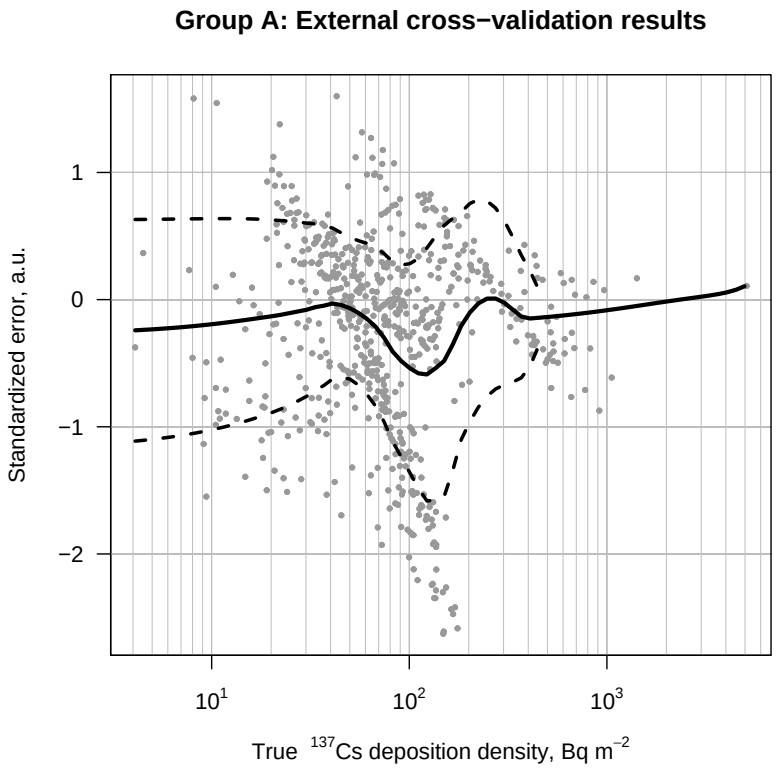
Group A: External cross−validation results
Standardized error, a.u.
1
0
−1
−2
$10^1$
$10^2$
$10^3$
True $^{137}$Cs deposition density, Bq m$^{-2}$

Fig. 20.

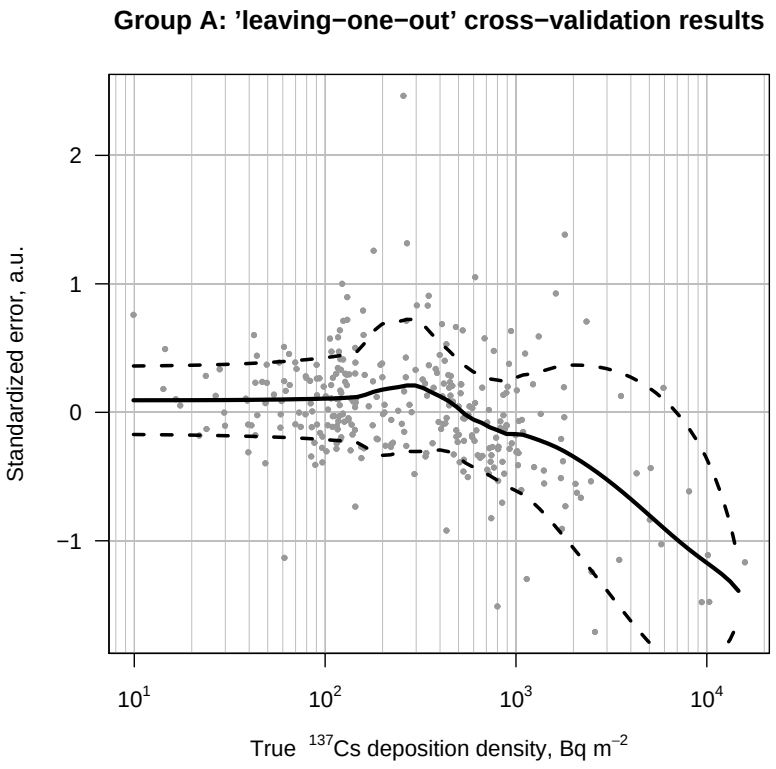
Group A: 'leaving−one−out' cross−validation results
Standardized error, a.u.
2
1
0
−1
$10^1$
$10^2$
$10^3$
$10^4$
True $^{137}$Cs deposition density, Bq m$^{-2}$

Fig. 21.

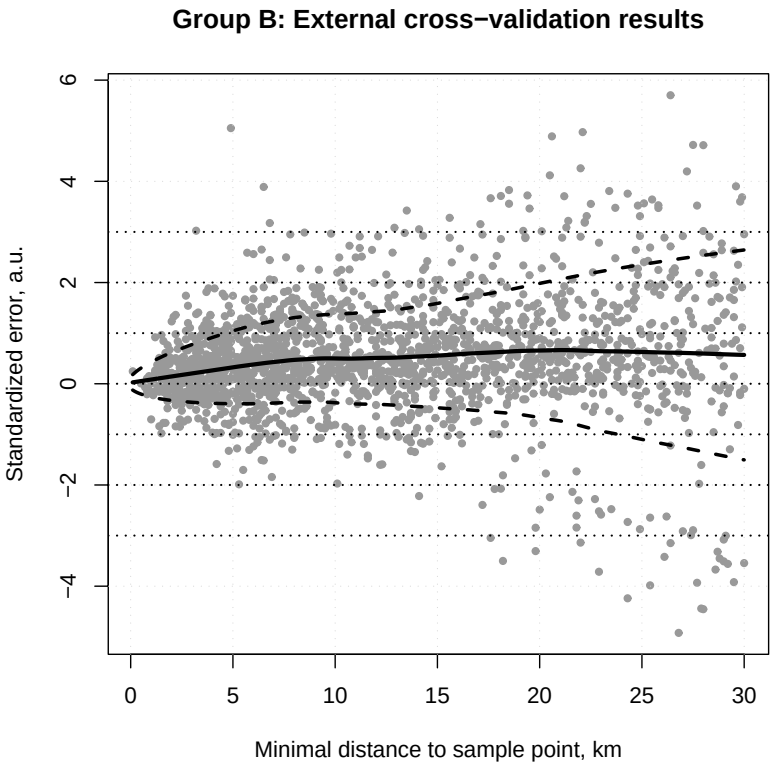
Group B: External cross−validation results
Standardized error, a.u.
6
4
2
0
−2
−4
0
5
10
15
20
25
30
Minimal distance to sample point, km

Fig. 22.

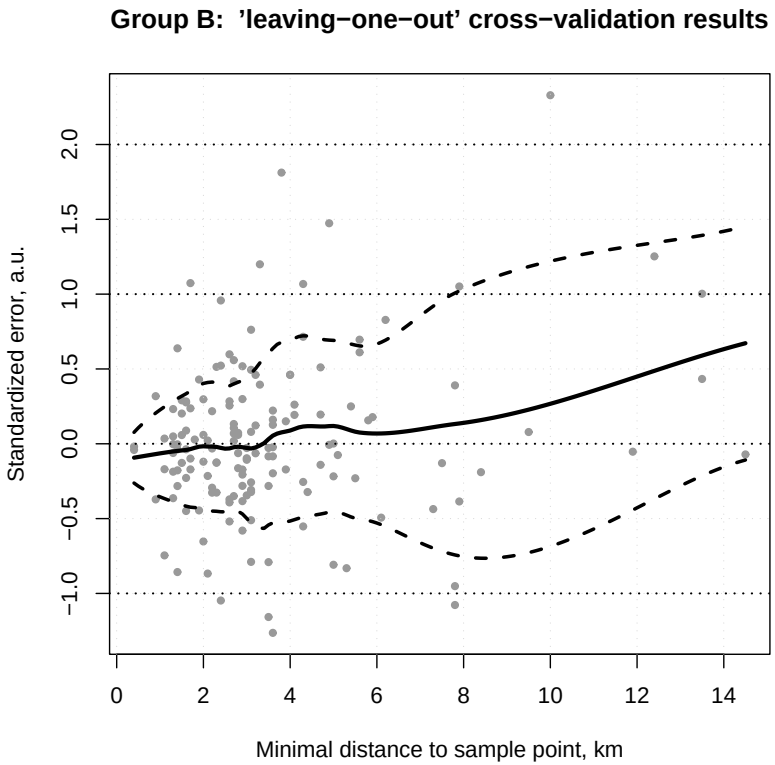
Group B: 'leaving−one−out' cross−validation results
Standardized error, a.u.
2.0
1.5
1.0
0.5
0.0
−0.5
−1.0
0
2
4
6
8
10
12
14
Minimal distance to sample point, km

Fig. 23.

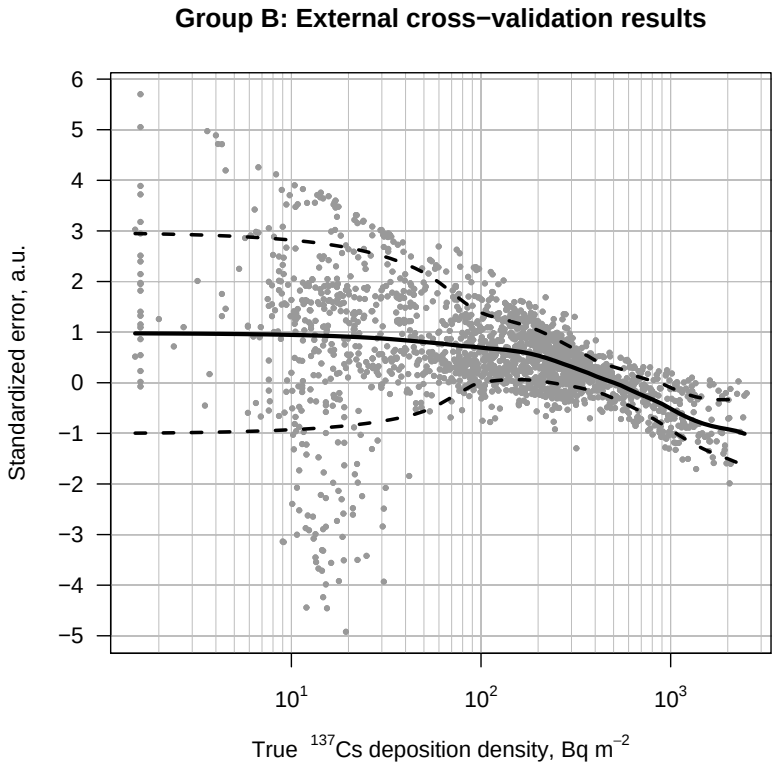
Group B: External cross−validation results
Standardized error, a.u.
6
5
4
3
2
1
0
−1
−2
−3
−4
−5
$10^1$
$10^2$
$10^3$
True $^{137}$Cs deposition density, Bq m$^{-2}$

Fig. 24.

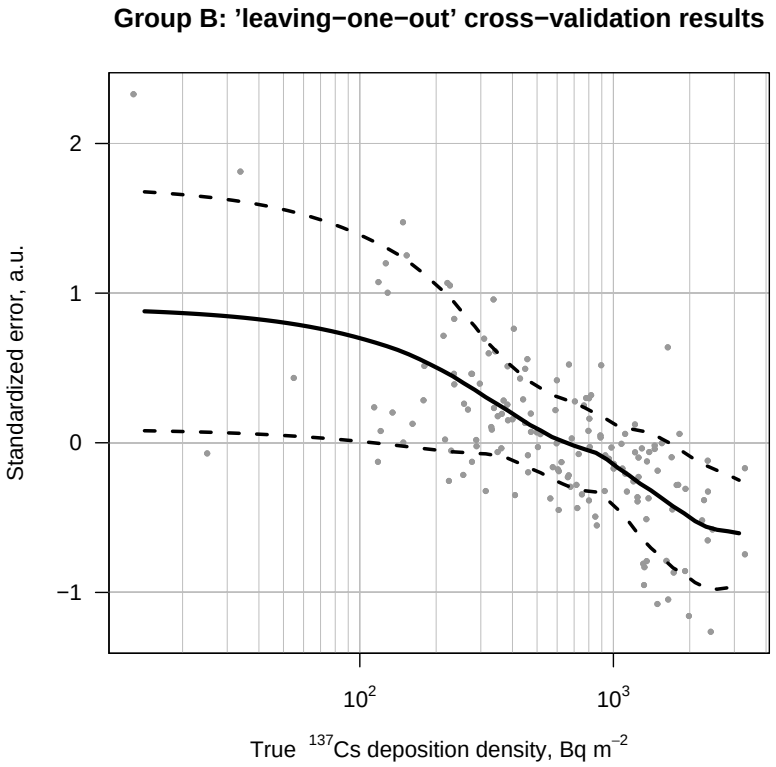
Group B: 'leaving−one−out' cross−validation results
Standardized error, a.u.
2
1
0
−1
$10^2$
$10^3$
True $^{137}$Cs deposition density, Bq m$^{-2}$

Fig. 25.

## CONCLUSIONS

The feasibility study has encompassed those settlements in Belarus where $^{131}I$ thyroidal activity had been measured among the members of public. The spatial configuration of the sample is defined by the geography of the thyroid measurements. Instead of settlement-average thyroid doses (or $^{131}I$ integrated activities) the data on $^{137}Cs$ ground deposition density are spatially interpolated to neighbour settlements, which are called "target" settlements. The data on $^{137}Cs$ ground deposition density are known for the sample and the target settlements, both. Therefore, a possibility exists to check capability of a spatial interpolation to predict values in the target settlements and to evaluate how appropriate is a spatial sample for such interpolation.

The feasibility study was also motivated by evident preferential bias in the samples. Namely, in Belarus the thyroid measurements had been taken mostly in high contaminated places. This fact had raised concerns in applicability of the spatial interpolation techniques to predict lower thyroid doses in the target settlements based on the higher values in the sample ones.

The sample settlements are separated into two groups due to their spatial location — A and B. The sampling patterns differ for these groups. Group A is sampled more densely and homogeneously while the group B sampling is sparse and clustered.

To compensate for the preferential sampling bias a special procedure has been developed. Namely, the local regression has been done on the sample data to reveal systematic spatial behaviour in the data — spatial trend. The regression residuals have been used in the ordinary kriging procedure which resulted in estimates of the residual and corresponding standard deviation for the target settlements. These, being combined with estimates of the trend in the prediction points, resulted in a set of predicted values.

Predicted data have been cross-validated using the true values for the both target and sample settlements. Results of the cross-validation are carefully evaluated and analyzed. The cross-validation has demonstrated that the spatial interpolation is possible within the range of the residual variogram. Results of the external cross-validation and "leaving-one-out" method are found to be consistent to each other. The interpolation results for the group A are found to be adequate while for the group B interpolated values are systematically higher at all distances

between the target and the sample points. A comparison of the standardized error and the true values has shown underestimation of the high values for both groups of settlements. However, both a local regression and a kriging are smoothing techniques, therefore they cannot reproduce a variability of the true values and they always underestimate high values and overestimate low ones (see e.g. [17, p.192]). For the group B predictions are adequate in the medium range of the true values from 200 to 1000 Bq $m^{-2}$ (Fig. 24).

The feasibility study has shown a capability to compensate for sample bias for the data of the group A. For the group B data the study demonstrated a scarcity of the sample data to adequately estimate the trend, thus leading to systematic overestimation in the predicted values. The established technique will be applied for interpolation of the settlement-average $^{131}I$ integrated thyroidal activity for the settlements of the group A. The feasibility study suggests not to use the group B settlements for the interpolation. However, this conclusion is validated for $^{137}Cs$ ground deposition density and could be re-evaluated for other radioecological quantities like settlement average thyroid doses.

It should be remembered that $^{137}Cs$ ground deposition density and thyroid doses to ingested $^{131}I$ do not necessarily have the same distributions and statistical properties. Fallout patterns of $^{137}Cs$ and $^{131}I$ differ considerably due to different transport forms and deposition mechanisms. The $^{137}Cs$ data are averages specific to the settlement area, while the thyroid doses are mainly resulted from a consumption of contaminated milk. Thus, the thyroid doses can be attributed to a wider area than the settlement area itself, i.e. including pasture area around the settlement. The thyroid doses are averages of individual values which have high variability because of individual diet, consumption rate, and metabolic parameters. All said above implies that the spatial distribution of the thyroid doses are anticipated to have higher variance and smoother spatial distribution with higher correlation range. That is, some conclusions made in the feasibility study for $^{137}Cs$ data could be different in case of thyroid doses to $^{131}I$. The latter ones are considered in the second part of the present study.

## ACKNOWLEDGMENT

This work has been supported by the German Federal Ministry of Environment, Nature Preservation, and Reactor Safety and the German Federal Office of Radiation Protection under the contract No. StSch 4240.

## REFERENCES


1. Astakhova L, Anspaugh L, Beebe G, et al. (1998) *Chernobyl-related thyroid cancer in children of Belarus: a case-control study*. Radiation Research **150**: 349–356.
2. Stezhko VA, Tronko MD, Beebe GW, et al. *A cohort study of thyroid cancer and other thyroid diseases following the Chornobyl accident: objectives, design, and methods*. (Radiation Research; in press).
3. Tronko MD, Bobylyova OO, Bogdanova TI, et al. (2003) *Thyroid gland and radiation (Ukrainian-American Thyroid Project)*. In: Radiation and Humankind. Proceedings of the First Nagasaki Symposium of the International Consortium for Medical Care of Hibakusha and Radiation Life Science, Nagasaki, Japan, 21-22 February 2003. International Congress Series 1258. Elsevier, pp. 91–104
4. Jacob P, Goulko G, Heidenreich H, et al. (1998) *Radioactive iodine risk to children estimated*. Nature **392**: 31–32
5. Jacob P, Kenigsberg J, Zvonova I, et al. (1999) *Chernobyl exposure during childhood and thyroid cancer incidence in Belarus and Russia*. Brit. J. Cancer **80**: 1461–1469
6. Likhtarev IA, Kayro IA, Shpak VM, Tronko ND, Bogdanova TI (1999) *Radiation-induced and background thyroid cancer of Ukrainian children (dosimetric approach)*. Int J Radiat Med **3**–**4**: 51–66
7. Jacob P, Kenigsberg J, Goulko G et al. (2000) *Thyroid cancer in Belarus after the Chernobyl accident. Comparison with external exposures*. Radiat Environ Biophys **39**: 25–31
8. Drozdovitch V, Goulko G, Minenko V, et al. (1997) *Thyroid dose reconstruction for the population of Belarus after the Chernobyl accident*. Radiat Environ Biophys **36**: 17–23
9. Kenigsberg J, Buglova E, Kruk J, Golovneva A (2002) *Thyroid cancer among children and adolescents of Belarus exposed due to the Chernobyl accident: dose and risk assessment*. In: Chernobyl: Message for the 21st Century. Elsevier, Amsterdam
10. Gavrilin Y, Khrouch V, Shinkarev S, et al. (1999) *Chernobyl accident: Reconstruction of thyroid dose for inhabitants of the Republic of Belarus*. Health Physics **76**: 105–119
11. DeCort M, Dubois G, Fridman S, et al. (1998) *Atlas on caesium contamination of Europe after the Chernobyl nuclear plant accident*. EUR 16733 EN RU. Office for Official Publication of European Communities. Brussels, Luxembourg
12. Germenchuk M (2001) *Radioactive contamination of the territories of Western Europe, Belarus, Ukraine, and Russia by $^{137}$Cs after the Chernobyl accident: analysis and prognosis*}. In: Fifteen years after the catastrophe. Ed.: A. Milyutin. Sakharov International Ecological University, Minsk [In Russian]
13. Journel A, Huijbregts C (1978) *Mining Geostatistics*. Academic Press, London
14. Isaaks E, Srivastava R (1989) *An Introduction to Applied Geostatistics*. Oxford University Press, Oxford
15. Cressie N (1993) *Statistics for spatial data, Revised Edition*. Wiley, New York
16. Wackernagel H (1995) *Multivariate Geostatistics. An Introduction with Applications*. Springer, Berlin
17. Webster R, Oliver M. (2001) *Geostatistics for environmental scientists*. Wiley, Chichester
18. Chilès JP, Delfiner P (1999) *Geostatistics: modeling spatial uncertainty*. Wiley, New York

19. Deutsch CV, Journel AG (1998) *GSLIB: Geostatistical software library and user's guide. Second Edition*. Oxford University Press, New York
20. Schlather M (1999) *Introduction to positive definite functions and to unconditional simulation of random fields*. Technical report ST 99-10. Lancaster University, Lancaster
21. Diggle P, Tawn J, Moyeed R (1998) *Model-based geostatistics*, Appl. Statist. **47**: 299–350
22. Ribeiro Jr PJ, Diggle PJ (2001) *geoR: A package for geostatistical analysis*. R-News **1** No. 2: 15–18 (ISSN 1609-3631)
23. Ihaka R, Gentleman R (1996) *R: A Language for Data Analysis and Graphics*, Journal of Computational and Graphical Statistics **5**: 299–314
24. Venables WN, Ripley BD (2000) *Modern Applied Statistics with S-Plus. Third edition*. Springer, New York
25. Cleveland W, Grosse E (1991) *Computational methods for local regression*. Statistics and Computing **1**: 47–62