Real-time automatic interpolation of ambient gamma dose rates from the Dutch Radioactivity Monitoring Network

Paul H. Hiemstra\textsuperscript{a,}\textsuperscript{*}, Edzer J. Pebesma\textsuperscript{b}, Chris J.W. Twenhöfel\textsuperscript{c}, Gerard B.M. Heuvelink\textsuperscript{d}

\textsuperscript{a}University of Utrecht, Department of Physical Geography, P.O. Box 80.115, 3508 TC Utrecht, The Netherlands
\textsuperscript{b}University of Münster, Institute for Geoinformatics, Weseler Straße 253, 48151 Münster, Germany
\textsuperscript{c}National Institute for Public Health and the Environment (RIVM), Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven, The Netherlands
\textsuperscript{d}Environmental Sciences Group, Wageningen University, PO Box 47, 6700 AA Wageningen, The Netherlands

Abstract

Detection of radiological accidents and monitoring the spread of the contamination is of great importance. Following the Chernobyl accident many European countries have installed monitoring networks to perform this task. Real-time availability of automatically interpolated maps showing the spread of radioactivity during and after an accident would improve the capability of decision makers to accurately respond to a radiological accident. The objective of this paper is to present a real-time automatic interpolation system suited for natural background radioactivity. Interpolating natural background radiation allows us to better understand the natural variability, thus improving our ability to detect accidents. A real-time automatic interpolation system suited for natural background radioactivity presents a first step towards a system that can deal with radiological accidents. The interpolated maps are produced using a combination of universal kriging and an automatic variogram fitting procedure. The system provides a map of 1) the kriging prediction, 2) the kriging standard error and 3) the position of approximate prediction intervals relative to a threshold. The maps are presented through a Web Map Service (WMS) to ensure interoperability with existing Geographic Information Systems (GIS).

Key words:
Automatic interpolation, Kriging, Monitoring networks, Interoperability, Web services

1. Introduction

Detection and monitoring of radiological accidents is of great importance. This has become clear after the Chernobyl accident in 1986. Possible radiological threats include releases from nuclear power plants, crashes of nuclear powered satellites, dirty bombs, releases from nuclear powered vessels and transport of radiological material. Many European countries have installed monitoring networks to detect radiological accidents. In the Netherlands the National Institute for Public Health and the Environment (RIVM) operates the National Radioactivity Monitoring Network (NRM).

In this paper we present a prototype system that provides real-time automatically interpolated maps of ambient gamma dose rate values from the NRM. Note that the monitoring stations from the NRM included in our analysis only measure total gamma dose rate and not a full gamma-ray spectrum. The prototype system is most suitable for mapping natural radioactivity data without major outliers. It represents a first step toward an automatic real-time interpolation system for emergency situations, when large outliers are to be expected. The goal of both the prototype system and the system for emergency situations is to improve the interpretation of variability in gamma dose rates measured by the NRM. Interpolated maps improve the interpretation of variability by providing a good overview and summary of the NRM data, making it easier to detect patterns and anomalies in the data. The interpolation algorithm not only provides a prediction but also an estimate of the prediction error, the \textit{kriging standard error}. Under the assumption that the kriging error is normally distributed, approximate 95% prediction intervals can be calculated to indicate where predictions exceed thresholds with 95% confidence.

Output from the system is presented using a Service Oriented Architecture (SOA) to ensure that the system is platform independent, flexible and easy to integrate into ex-
isting applications. We use a Web Map Service (WMS) to distribute images of interpolated maps. WMS is a standard defined by the Open Geospatial Consortium.

Examples of research into automatic interpolation include interpolation of meteorological data (Pardo-Iguzquiza et al., 2005), atmospheric pollution (Abraham and Comrie, 2004), seismic activity (Wald et al., 1999) and ionospheric activities (Stanislawksa et al., 2002; Turley and Gardiner-Garden, 2006). Two large statistical exercises (EUR 20667 EN, 2003; Dubois and Galmarini, 2005; EUR 21595 EN, 2005) have dealt with the mapping of radioactivity in routine and emergency situations. The second exercise (EUR 21595 EN, 2005) focused on automatic interpolation. Results from all of these studies underline the fact that real-time automatic interpolation of environmental data —especially in case of an accident— is not an easy nor straightforward task (see also Brenning and Dubois (2008)).

2. Natural outdoor gamma radiation and spatial variability

Natural outdoor gamma radiation originates from cosmic rays and from the decay of radionuclides present in the air, on the ground surface (e.g. radon daughter products deposited during rain storms) or in the soil (e.g. $^{40}$K). The measure of activity, the Becquerel (Bq), indicates the number of disintegrations per second of the radionuclide of a given sample. The ambient radiation levels are usually expressed in terms of an ambient dose equivalent rate at 10 mm depth, $H^*(10)$ (International Commission on Radiation Units and Measurements, ICRU (1993)). $H^*(10)$ is commonly abbreviated to ambient dose rate, measured in nSv/h.

The average ambient dose rate for the period 1990-1994 in the Netherlands is between 55 and 115 nSv/h (Smetsers and Blauuboer, 1997). There are three major sources of natural background gamma radiation in the outdoor environment in the Netherlands (Smetsers and Blauuboer, 1997):

1) Cosmic gamma radiation originates from outside the Earth’s atmosphere and accounts for about 50% of the annual dose. This source is partly screened by the Earth’s magnetic field and the atmosphere. The dose due to cosmic radiation depends on atmospheric pressure. Since the Netherlands spans just a few degrees latitude and has an almost constant atmospheric pressure, the cosmic radiation dose is nearly constant spatially.

2) Terrestrial gamma radiation consists primarily of radiation from the natural decay chain of $^{232}$Th, $^{235}$U, $^{238}$U and $^{40}$K and accounts for about 50% of the annual dose. The amount of radionuclides present in the subsoil depends on soil type (Van Dongen and Stoute, 1985). In general, clay and loess contain more radionuclides than sand and peat and thus produce more terrestrial radiation. Figure 1 shows a box-and-whisker plot of soil type versus terrestrial gamma dose rate. This figure was derived from a terrestrial gamma dose rate map made by Smetsers and Blauuboer (1996) and serves as an illustration of the effect of soil type on terrestrial radiation. Figure 2 shows a simplified soil map of the Netherlands based on a 1:250,000 soil map (Steur et al., 1985). We aggregated the 13 soil types in this map to match the classes shown in figure 2. All sand types (e.g. sands rich in clay or sands rich in peat) were put into a single class. Similarly, all built-up areas, types of peat, types of marine clay and types of fluvial clay were aggregated into each of their respective classes for a total of 5 soil types.

3) Radon ($^{222}$Rn, $^{220}$Rn and their short-lived progeny) is an important source of natural radioactivity in air. The isotopes $^{222}$Rn and $^{220}$Rn originate from the $^{238}$U and $^{232}$Th series respectively. Especially $^{222}$Rn will be transported through the atmosphere after exhalation from the soil. Locally the radiation dose in air is correlated to the amounts of radioactivity of the $^{238}$U and $^{232}$Th series in the soil and thus to soil type. However, due to variation in exhalation because of changes in the groundwater table for instance and due to atmospheric transport of $^{222}$Rn over large distances, the correlation of the $^{222}$Rn concentration in air to soil type is less significant. Furthermore, the contribution to the radiation dose in air from $^{222}$Rn and its progeny is small, but after precipitation events when significant amounts of progeny are deposited on the ground, significant temporary rises of dose rate occur that could easily double or triple the background radiation dose during an hour or more (Smetsers and Blauuboer, 1997).

3. The Dutch Radioactivity Monitoring Network

The primary objective of the NRM (Twenhöfel et al., 2005) is to act as an early warning system in the event of an accidental radioactive release. To realise this objective a network of 153 ambient gamma dose rate monitoring stations has been set up. The network provides 10-minute averaged values of the ambient gamma dose rate. Figure 3 shows the locations of the monitoring stations. The network is designed to have a high probability of detecting a radioactive release. Some areas have an increased density of monitoring stations, for example the area around the nuclear power plant at Borselle. The density is also higher close to country borders, especially near foreign nuclear power plants, to increase the probability of detecting any foreign releases. When a station reports a value exceeding 200 nSv/h (Twenhöfel et al., 2005), a warning is generated. The network automatically notifies an expert from the RIVM to validate the alarm and start emergency procedures if required. If an accident occurs, RIVM also has mobile measuring stations that can be deployed anywhere in the field.

\footnote{http://www.opengeospatial.org/standards/wms}
4. Spatial interpolation

4.1. Universal kriging

Maps of ambient gamma dose rates were calculated by making predictions on a regular 1 km × 1 km grid using universal kriging (Chil`es and Delfiner, 1999; Journel and Huijbregts, 1978). The main advantage of universal kriging is the ability to take into account trends present in the data, in our case soil type (see Hengl et al. (2004) for an example using soil type as a predictor). We chose the grid size of 1 km × 1 km because a smaller grid size was not appropriate given the density of the NRM (average distance between stations is about 12 km). We assume that measurements from the NRM network, z(x_i) with i = 1, ..., n, are a realization of a spatial random field Z(x), x ∈ D that can be described by the linear model (Christensen, 1996):

$$Z(x) = \beta_0 + \sum_{i=1}^{p} \beta_i f_i(x) + e(x),$$

where

$$E(e(x)) = 0,$$

$$\text{Cov}(e(x), e(x+h)) = C(h) \forall x \text{ and } x + h \in D$$

where $\beta_i$ is an unknown regression coefficient, $f_i(x)$ is the i-th known predictor at location $x$, $h$ is a distance vector, $C$ is the covariance function of $e$ and $D$ is the geographical domain. We assume that the residual $e$ is second order stationary, i.e. $e$ has a constant mean and the covariance between $e(x)$ and $e(x+h)$ is determined only by the distance vector $h$ (Chil`es and Delfiner, 1999).

Given the linear model in equation 1 and the observations from the NRM, the best linear unbiased prediction for grid node $x_0$ is given by:

$$\hat{Z}(x_0) = f(x_0)\hat{\beta} + \delta(z - X\hat{\beta}),$$

where $f(x_0) = (f_1(x_0), ..., f_p(x_0))'$ is the estimated regression coefficients, $z = (z(x_1), ..., z(x_n))'$ are the observations, $X$ is the $n \times (p+1)$ design matrix where the i-th row is equal to $f(x_i)$, $\delta = \Sigma_0^{-1}\Sigma$, $\Sigma$ is the variance-covariance matrix of the $e(x_i)$, $i = 1, ..., n$ and $\Sigma_0$ is a vector containing the covariance between the residual of the prediction and the residuals of the observations ($\Sigma_0 = (\text{Cov}(e(x_1), e(x_0)), ..., \text{Cov}(e(x_n), e(x_0))))'$.

The vector $\hat{\beta}$ is the best linear unbiased estimator of $\beta$ and is given by:

$$\hat{\beta} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}z(x).$$

The variance of the prediction error, or universal kriging variance, is given by:

$$\text{Var}(Z(x_0) - \hat{Z}(x_0)) = \sigma^2_n \equiv \Sigma_0'\Sigma^{-1}\Sigma_0 + (f(x_0) - \delta X)(X'\Sigma^{-1}X)^{-1} \times (f(x_0) - \delta X)'$$

where $\sigma^2_n = \text{Var}(Z(x_0))$.

We use soil type as a predictor in universal kriging. Because soil type is a categorical variable, this involves a stepwise trend. Usually the first column of $X$ is the vector $(1, ..., 1)'$, so $\beta_0$ is an intercept. In case of a stepwise trend, the remaining columns contain indicators that are 1 if the observation belongs to the associated soil type or 0 otherwise. Summing up all the columns except the intercept column results in a column equal to the intercept column, in which case $X$ is a singular matrix, and $(X'\Sigma^{-1}X)^{-1}$ does not exist. This problem is solved by dropping either the intercept or one of the categories. Here, one category is dropped. The soil map generally explains 15-20% of the variability observed in the NRM.

The covariance function satisfies $C(h) = C(0) - \gamma(h)$, where $C(0)$ is the variance of $e$, equal to the sill of the variogram and $\gamma(h)$ is the variogram of $e$ (Cressie, 1993; Christensen, 1991). Note that we refer to the variogram model when we use the term variogram. Assuming the kriging error is normally distributed we can approximate the 95% prediction intervals for $Z$ at prediction location $x_0$ by:

$$[\hat{Z}(x_0) - 1.96\sigma(x_0), \hat{Z}(x_0) + 1.96\sigma(x_0)]$$

with $\hat{Z}(x_0)$ the kriging prediction and $\sigma(x_0)$ the kriging standard error, defined as the square root of (4).

4.2. Automatic variogram fitting

The variogram model is a function of semivariance versus distance that provides a fit to the sample variogram. For a function to be a permissible variogram model, the variance of the prediction based on this variogram model should always be positive. This is ensured if the function is conditionally positive definite (Chil`es and Delfiner, 1999). One way to fit a variogram model is by visual comparison of the variogram model and the sample variogram. In this study however, one of the goals was to automate the full process and so we need an automatic fitting procedure.

The sample variogram is calculated from the data for distance intervals (bins) $(h_k, h_k + \delta_k)$ by the semivariance

$$\hat{\gamma}(h_k) = \frac{1}{2N_k} \sum_{i=1}^{N_k} (\hat{e}(x_i) - \hat{e}(x_i + h))^2,$$

$$\forall (x_i, x_i + h) : h \in [h_k, h_k + \delta_k]$$

where $\hat{e}(h)$ is the average of all separation distances $h$ within bin $k$, $N_k$ is the number of point pairs in bin $k$, and $\hat{e}(x_i)$ and $\hat{e}(x_i + h)$ are the estimated residuals at locations $x$ and $x + h$ respectively. The variogram model is fit to the sample variogram by iterative reweighted least squares (Gauss-Newton fitting, Cressie (1993)) with weights equal to $N_k/h^2_k$. Gauss-Newton fitting requires an initial guess of the variogram parameters and after which the algorithm iteratively converges toward an optimal solution. We obtain an initial guess of the variogram model based on the sample vari-
ogram. For the initial guess of the variogram parameters the following rules seemed appropriate:

- The initial nugget (semi-variance when the separation distance approaches zero) is estimated using the minimum semi-variance in the sample variogram.
- The initial range (the distance at which the variogram becomes constant) is set at 0.35 times the diagonal of the area that is covered by the measurements, i.e. 137 km in this study.
- The initial sill (semi-variance at distances beyond the range) is the average of the maximum semi-variance and the median semi-variance in the sample variogram.

Note that the lags are distributed such that the bins are smaller at smaller distances. This leads to the following bin boundaries for the Netherlands (km): 0, 3, 6, 8, 12, 15, 21, 34, 48, 68, 89, 110, 137. Note also that we use an isotropic spherical variogram model. When we applied this procedure to the gamma dose rate data from the NRM we faced some problems with regard to the short range spatial variability. These problems will be discussed in the following section.

4.3. Short range spatial variability

The NRM monitoring stations are distributed relatively evenly across the country. As a result, few stations fall into the shortest distance class of a separation distance of less than 12 km. This causes two specific problems for variogram modeling.

First, all monitoring station pairs with a separation distance under 12 km are located near the nuclear power plant at Borselle (see figure 3). These stations behave differently from the remainder of the monitoring stations as is apparent from the sample variogram shown in figure 4 (top left). The second bin in the sample variogram had a much higher semi-variance than the other bins. We solved this problem by excluding the five stations around the power plant in order to fit the variogram model. These stations were included however, for prediction based on the fitted variogram model.

Second, fitting the nugget based on the sample variogram was difficult considering the lack of monitoring stations at small separation distances. We performed the automatic fitting procedure described in section 4.2 for the first three months of 2005. Figure 5(a) plots the fitted nugget against time. Figure 5(b) shows the first six variograms of 2005. The time interval between the variograms was ten minutes.

The nugget shows strong fluctuations, leading to large differences in the interpolated maps. If we look at figure 5(b), the sample variograms do not vary enough to explain the rapid change in the nugget. To characterize the spatial behavior at smaller distances, we used a different dataset collected using mobile measuring devices. We assumed that the semi-variances extracted from this dataset were comparable to those from the NRM. This dataset included separation distances between 0 and 1500 m. Figure 6(a) shows the sample variogram calculated from this dataset. Figure 6(b) shows the mean of figure 6(a) in a typical sample variogram. We fixed the nugget in the automatic fitting procedure to the mean semi-variance of the short range dataset, 11 (nSv/h)². This assumes that the nugget is constant over time, in addition to the second order stationarity assumption. The bottom two panels in figure 4 show examples of results from the automatic fitting procedure.

4.4. Generating the map

Generating the prediction map involved calculating the best linear unbiased prediction for all gridnodes using equation 2. The kriging error map was generated in a similar manner using equation 4. Information on soil type is needed at both the monitoring locations and at the prediction locations for UK. All observations were used for prediction, i.e. we used global kriging.

5. Technical implementation

The automatic mapping system consisted of three operations: (1) getting the data from the NRM database, (2) interpolating these data as described in section 4 and (3) serving the results through a Web Map Service (WMS). The WMS is a method for distributing static maps, adopted as a standard by the Open Geospatial Consortium². A flowchart of the system is shown in figure 7. The flowchart consists of two distinct parts. One part processes the data from the NRM and generates maps. The second part of the system interacts with the user, processing requests and returning the appropriate map.

5.1. Data processing

The data processing system was implemented partly in Python (Van Rossum and Drake, 2001)³ and partly in R (R Development Core Team, 2008)⁴. The system connects to an FTP server containing the NRM data and downloads the most recent data (fig. 7 (1)). These data are interpolated using an R extension package called automap (fig. 7 (2)), developed in this study. This package implements the automatic interpolation routine described in section 4. The automap package⁵ makes use of the R packages gstat (Pebesma, 2004) and sp (Pebesma and Bivand, 2005). The automatic mapping system results in three maps: a map of the kriging prediction, a map of the kriging standard error and a map showing the approximate 95% prediction intervals, classified relative to 95 nSv/h. Intervals were classified into three categories: lower, if the entire prediction interval was below 95 nSv/h; higher, if the entire prediction

² http://www.opengeospatial.org/standards/wms
³ http://www.python.org
⁴ http://www.r-project.org
⁵ available at http://intamap.geo.uu.nl/~paul/Downloads.html
interval was above 95 nSv/h; or not distinguishable, if 95 nSv/h was included in the prediction interval. The map of classified prediction intervals indicates the areas exceeding a predefined threshold, e.g. an intervention level. The value of 95 nSv/h was chosen because it clearly demonstrates the use of this output type in non-emergency situations. Figure 8 shows examples for all three output types. Maps were stored in the GeoTIFF format \(^6\) (fig. 7 (3)). Finally, the catalogue linking the GeoTIFF files to a date and time was updated (fig. 7 (4)).

5.2. User processing

The maps produced by the data processing system were served to the user through a WMS, implemented in UMN MapServer (Vatsavai et al., 2006). A WMS supports three operations: GetCapabilities, GetFeatureInfo and GetMap. The first operation reports what maps are available on the WMS. GetFeatureInfo allows users to gather information on specific maps. The GetMap request returns the map from the WMS to the user by checking the catalogue (fig. 7 A) and retrieving the appropriate map (fig. 7 B). The WMS only advertises the maps, a client application is needed to read and display these maps. In addition to the WMS, we developed a simple CGI-based client application in Python. The client allows a user to select an interpolated map from the WMS server for one of the output types in combination with a date and time. Figure 9 shows a screen dump of the client application.

6. Discussion

The aim of this project was to develop an automatic real-time mapping system for ambient dose rates in the Netherlands. The system automatically gets the data from the NRM database, fits a variogram to the data, interpolates the data and stores the resulting maps. This process takes just a few seconds to complete. We will now shortly discuss assumptions and shortcomings.

6.1. Operational issues

Stations with anomalous short-range variability around Borselle were excluded from the calculation of the sample variogram (section 4.3). The automatic interpolation routine does not automatically detect these anomalies, they have to be identified beforehand. Ignoring anomalies may not be the best way to deal with them. If we know what causes the anomalies, it may be better to apply a correction to the data. The anomalous short range behaviour around Borselle may be caused by parking lots or slightly more radioactive rocks close to the monitoring station (personal communication with A. van Lunenburg, RIVM).

Fixing the nugget solved the instability in fitting the nugget (figure 6). This instability results from the lack of monitoring stations at smaller separation distances. Data at smaller separation distances were collected using a different measuring device than the one used in the normal NRM network. We assumed that the semi-variance extracted from the short range dataset was comparable to the semi-variance from the NRM. This assumption was supported by the fact that the short range semivariance is in the same order of magnitude as the semivariance of the NRM data.

The system presented in this study was designed for data without major outliers. This makes the system not entirely well suited for emergency situations. If we increase the value of one station to 10,000 nSv/h we observe: (1) the kriging standard error becomes very high, (2) the approximate prediction intervals become very wide, (3) the position of the approximate prediction interval classified relative to 95 nSv/h reads not distinguishable in the whole country except close to the high value where it is classified higher. The area classified as higher gives only a rough indication of the contaminated area. This rough indication can only be given, provided that we keep the nugget fixed. If we do not fix the nugget, the automatic fitting of the variogram model is difficult (Pebesma, 2005). It is important to note that the extreme values make the assumption of second order stationarity (section 4.1) highly doubtful. Data transformation techniques such as normal score transformation or Box-Cox transformation could be used to deal with local high values.

6.2. Statistical issues

In order to provide automatically interpolated maps, it was necessary to make a number of assumptions regarding the process that generated the data. These include:

- mean gamma dose rate varies per soil type,
- stationarity of residual variation,
- the spherical variogram model used,
- an isotropic variogram,
- the fixed, time-invariant nugget value of 11 (nSv/h)\(^2\).

These assumptions were partly made to ensure that the procedure is stable and reliable when applied in an automatic setting, but potentially limit the value of the procedure for data that have different characteristics than those for which we made the assumptions. Especially in case of emergencies when one or a few extreme values occur.

In this project we only included the type of soil as a predictor. Several other predictors could be used to improve spatial predictions by reducing residual variability. In case of radon washout, total precipitation accumulated over some time frame (obtained from real-time rainfall radar images) is a candidate predictor. Another predictor useful in emergency situations could be the outcome of a process-based atmospheric dispersion model (e.g. NPK-PUFF, Van Egmond and Kesseboom (1983)). This model

\(^6\) \url{http://www.remotesensing.org/geotiff}
needs information regarding the location and the size of the release. Such information is combined with meteorological data to provide a forecast of how the contamination will spread over time. If we run the model from time of release to the time when we want to perform the interpolation, we obtain an estimate of the spatial distribution of contamination. This can be used as a predictor.

Approximate 95% prediction intervals given by equation 5 are valid if the following additional assumptions hold: kriging prediction errors are normally distributed and the variogram model is known and not subject to estimation error. Again, both of these assumptions are highly questionable in the case of a few extreme values.

Any of the assumptions made may leave room for improvement of the procedures. As a first start, and for the purpose of automated mapping of natural ambient dose rates, they seemed appropriate.

Ambient gamma dose rates vary not only in space but also in time. Measurements from other moments in time might improve interpolations. Time could become a third dimension, describing the variation in time as a combination of a trend and a zero-mean residual, similar to the description of variation in space.

6.3. Technical issues

A WMS distributes static maps, giving the user no possibility to influence the output of the system. One option to solve this problem could be to use a Web Coverage Service\(^7\), which serves the actual raster data and not just a visual depiction. With a WCS, all visualization and interpretation is done at the client side using for example desktop GIS software or a programming language. The WCS user now has many more options, for example the choice of the color scale or the threshold to which the prediction interval is classified.

Another possibility is to use a Web Processing Service\(^8\). A WPS can implement any GIS functionality, including interpolation routines. Implementing the automatic mapping system in a WPS would make it possible for the client to change parameters such as the color scale or the type of variogram model that is fit. Using a WPS allows the use of a web-based client, while allowing the user to influence the output. The INTAMAP\(^9\) project (Pebesma et al., 2006) addresses many of these issues.

6.4. Portability

The automatic interpolation system developed in this study was designed specifically for the Dutch radioactivity monitoring network. The system can also be applied to other data that are collected through a monitoring network, assuming that the residuals can be described by a second order stationary random field. One could think of air pollution data, ozone data or water quality data. However, the assumptions and solutions in section 4 can not be ported directly.

7. Conclusions

The prototype automatic mapping system developed in this project provides real-time interpolated maps of gamma dose rates based on the Dutch National Radioactivity Monitoring Network and the Dutch soil map. This is a significant improvement over the current situation where NRM data are interpreted based on point measurements only. The system gives a quick overview of the present radiological situation and provides predictions of radiation levels at unmeasured locations. In addition to a prediction, the system provides an associated prediction standard error. Combining these two outputs allows estimation of approximate 95% prediction intervals. This is a powerful way to map areas where predictions exceed thresholds. The system works efficiently and automatically. Resulting maps are presented in an interoperable way through a Web Map Service (WMS). This ensures that the system is platform independent, flexible and easy to integrate into existing applications. It is possible to port this system to other environmental variables such as ozone or air quality.

Kriging allows the use of predictors to account for trends in the data. In this study soil type was included as a predictor. Additional predictors such as precipitation, outputs of atmospheric dispersion models or height of the mixture layer could further improve the prediction by reducing residual variation.

The system presented in this paper works well when measuring natural radioactivity in the absence of extreme events, such as nuclear or radiological accidents or enhanced measurements caused by the washout of radon daughter products, e.g. 214\(^\text{Po}\), after a (heavy) rainfall. If an accident occurs, the system can give a general indication of the size of the contaminated area, provided that we keep the nugget fixed. The kriging standard error however becomes very large. The extreme values cause the assumption of second order stationarity to be highly doubtful. One possible way to deal with local high values is to use data transformation techniques.

Acknowledgements

We gratefully acknowledge financial support from the innovation programme Space for Geo-Information (RGI), projects RGI-182 and RGI-302. This work is partially funded by the European Commission, under the Sixth Framework Programme, by the INTAMAP project Contract N. 033811 with the DG INFSO, action Line IST-2005-2.5.12 ICT for Environmental Risk Management. The views expressed herein are those of the authors and
are not necessarily those of the RGI or the European Commission. We would to thank the two reviewers for their helpful comments and suggestions and Stephanie Melles for help on improving the English of this paper.

References


Fig. 1. Box-and-whisker plot (Tukey, 1977) of terrestrial radiation against soil type obtained from a terrestrial gamma dose rate map by Smetsers and Blaauboer (1996). Dots are outliers in the data.
Fig. 2. Simplified soil map of the Netherlands based on 1:250,000 soil map (Steur et al., 1985).
Fig. 3. Location of gamma dose rate monitoring stations of the Dutch National Radioactivity Monitoring network. Diamond marks the Borselle nuclear power plant.
Fig. 4. Effect of anomalous measurements on sample variogram and an example of output from the automatic fitting procedure. Unfiltered: all data; filtered: highly variable short-distance observations removed; initial guess: starting value for variogram fit; fitted variogram: weighted least squares fit, with pre-fixed nugget (see section 4.3). Numbers indicate total number of point pairs associated with each bin in the sample variogram.
Fig. 5. Instability of fitted nugget due to lack of monitoring stations at small separation distances. (a) Time series of fitted nugget for the first three months of 2005. (b) First six 10-minute variograms of 2005. Time progresses from top left to bottom right.
Fig. 6. (a) Short range variogram based on mobile measurements. Average semi-variance was 11.4 (nSv/h)$^2$ (dashed line). (b) Average short range semi-variance of (a) (cross) plotted in a sample variogram for the complete NRM dataset.
Fig. 7. Flowchart of the automatic mapping system. Numbers indicate the order in which actions take place in the data processing chain. Letters indicate the order in which actions of the WMS server take place.
Fig. 8. Examples of three types of output from the automatic mapping system. (a) Kriging prediction (nSv/h), (b) Kriging standard error (nSv/h) and (c) approximate 95\% prediction intervals classified relative to 95 nSv/h. Circles represent locations of measuring stations.
Fig. 9. Client application running inside a web browser.