

Automatic real-time interpolation of radiation hazards: prototype and system architecture considerations*

Paul H. Hiemstra^a, Edzer J. Pebesma^b, Chris J.W. Twenhöfel^c,
Gerard B.M. Heuvelink^d

^aUniversity of Utrecht, Department of Physical Geography,
P.O. Box 80.115, 3508 TC Utrecht; (p.hiemstra@geo.uu.nl)

^bUniversity of Münster, Institute for Geoinformatics,
Robert-Koch-Str. 26-28, D-48149 Münster;
(edzer.pebesma@uni-muenster.de)

^cNational Institute for Public Health and the Environment (RIVM),
Antonie van Leeuwenhoeklaan 9, 3721 MA Bilthoven;
(chris.twenhofel@rivm.nl)

^dEnvironmental Sciences Group, Wageningen University and Research Centre,
P.O. Box 47, 6700 AA Wageningen; (gerard.heuvelink@wur.nl)

Abstract

Detecting and monitoring the development of radioactive releases in the atmosphere is important. In many European countries monitoring networks have been established to perform this task. In the Netherlands the National Radioactivity Monitoring network (NRM) was installed. Currently, point maps are used to interpret the data from the NRM. Automatically generating maps in real-time would improve the interpretation of the data by giving the user a clear overview of the present radiological situation and provide an estimate of the radioactivity level at unmeasured locations. In this paper we present a prototype system that automatically generates real-time maps of radioactivity levels and presents results in an interoperable way through a Web Map Service. The system defines a first step towards a emergency management system and is suited primarily for data without large outliers. The automatic interpolation is done using universal kriging in combination with an automatic variogram fitting procedure. The focus is on mathematical and operational issues and on architectural considerations on how to improve the interoperability and portability of the prototype system.

* This work is licensed under the Creative Commons Attribution-Non commercial Works 3.0 License. To view a copy of this license, visit <http://creativecommons.org/licenses/by-nc-nd/3.0/> or send a letter to Creative Commons, 543 Howard Street, 5th Floor, San Francisco, California, 94105, USA.

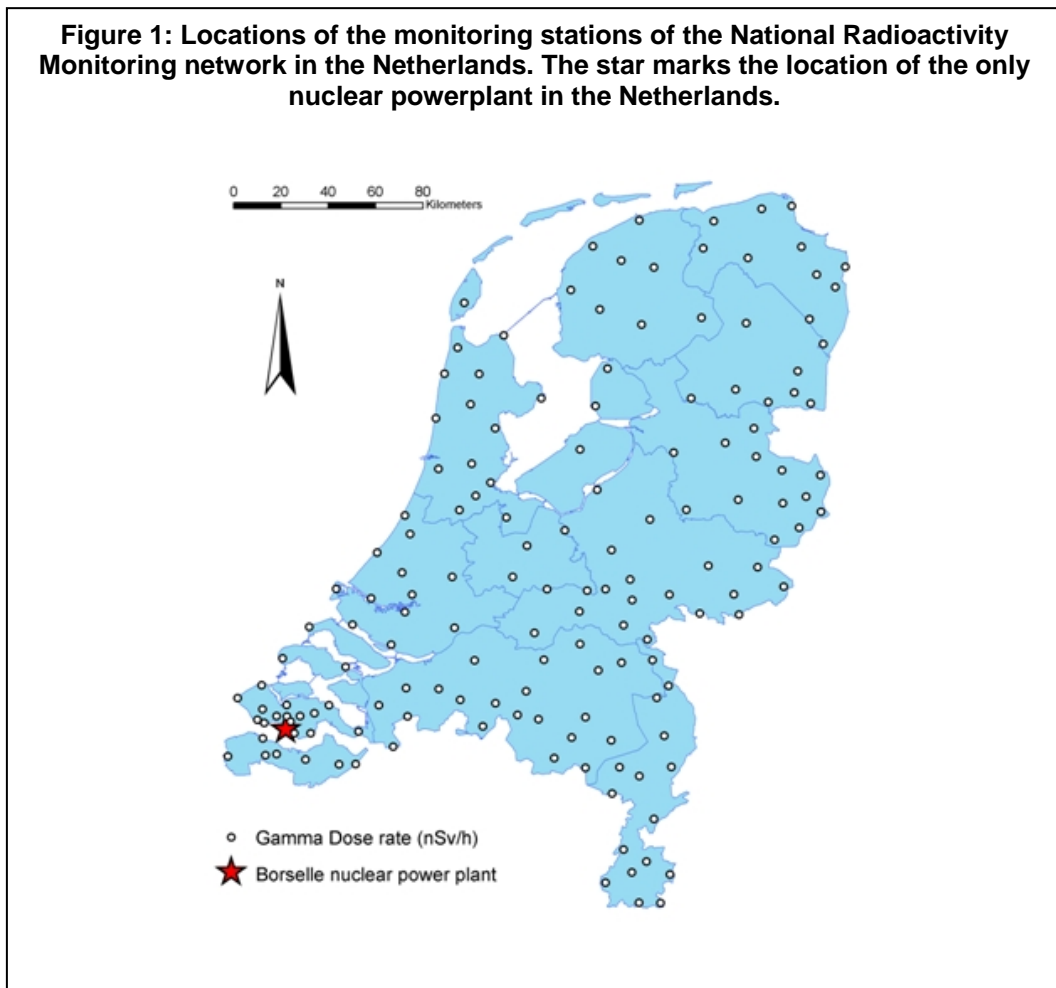
Keywords: Automatic interpolation, Kriging, Web services, Interoperability, Nuclear emergency management.

1. INTRODUCTION

Knowing when and where a radioactive release occurs and how the radioactive cloud evolves in space and time is important information in nuclear emergency management. Radioactive releases can for example originate from accidents at nuclear power stations, crashes of satellites carrying nuclear energy sources or from the use of dirty bombs. The Chernobyl accident in 1986 has acted as a catalyst in the development of monitoring networks across Europe to detect and monitor radioactive releases. In the Netherlands the National Radioactivity Monitoring network (NRM, Twenhöfel (2005)) was established. This network consists of 153 monitoring stations distributed more or less uniformly over the Netherlands, with increased density near the Borssele nuclear power plant (Figure 1). Radioactivity measurements are obtained from the NRM every 10 minutes. In case of an emergency these raw data provide input for a risk assessment of the radiological situation on which a counter-measure and recovery strategy are based. Real-time automatically generated radioactivity maps would help decision makers. Two large statistical exercises (Dubois (2003); Dubois and Galmarini (2005); Dubois (2005)) have dealt with mapping radioactivity data and showed that automatic mapping in emergency cases presents a great challenge. In this paper we present a first step towards an automatic mapping system by building a prototype system most suitable for natural radioactivity, i.e. background radiation without large local extremes. Automatic means that no human intervention is required to perform the interpolation. The maps provide a better overview of the radiological situation than the point maps that are currently used for interpretation. The interpolation is done using *kriging* (Chilès and Delfiner, 1999; Journel and Huijbregts, 1978). An advantage of kriging over deterministic methods such as inverse distance interpolation is the ability to quantify interpolation accuracy and include trends such as precipitation and soil type into the interpolation. In this study we use the fact that different types of soil contain different concentrations of radionuclides (Smetser and Blaauboer, 1997) and thus the amount of background radiation differs between soil types. Figure 2 shows the soil map that was used in this study. In general loess and clay have higher radiation levels. In this study we concentrate on (background) gamma radiation, expressed in *ambient dose rate* (ICRU, 1993) with unit nano Sievert per hour (nSv/h).

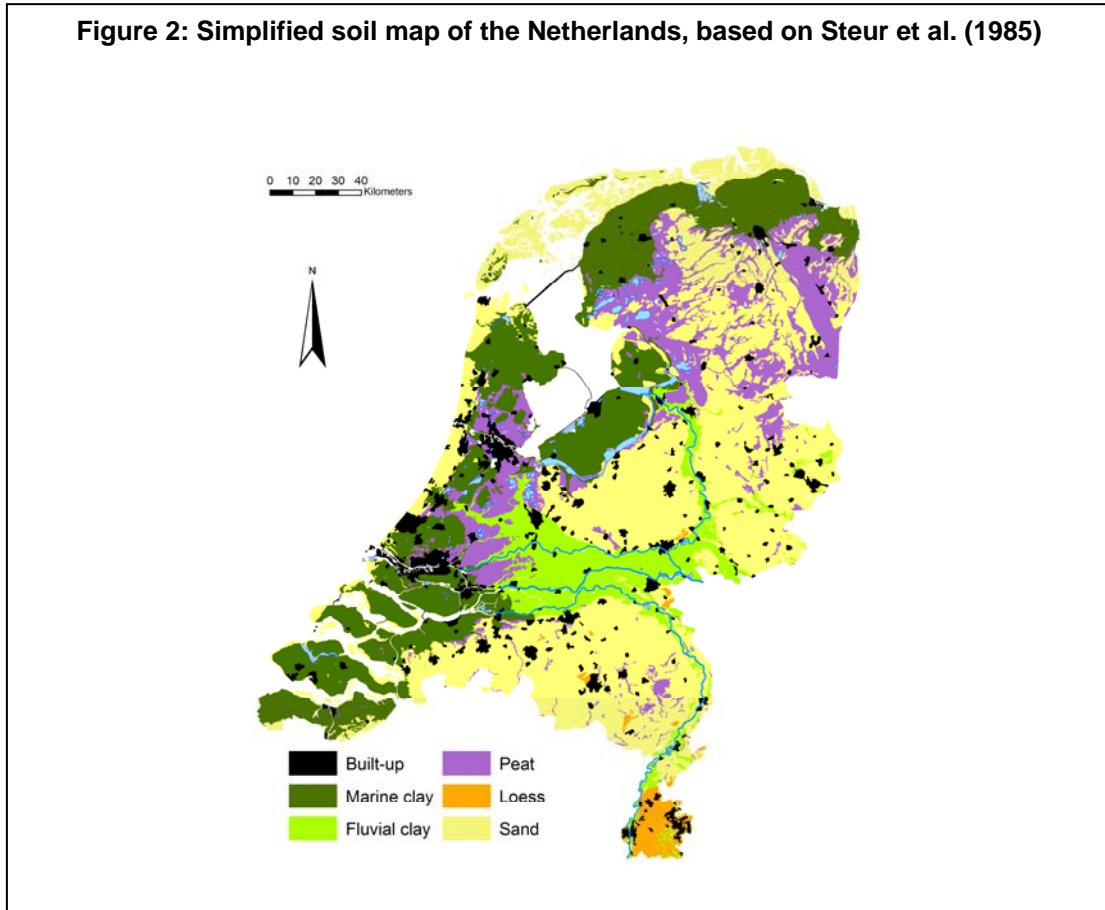
In emergency management systems many types of information are needed, such as vector maps showing municipal boundaries, locations of potential release sources or population density information. These data typically come from a range of sources (e.g. MySQL databases, files) and may use different data formats (e.g. GeoTIFF, Shapefile). A further complication is that some data are

available on the local computer, but much of it is stored on remote servers. In order to make use of these data, the computer systems must be able to work together (i.e. be interoperable). To achieve interoperability we need strict definitions how the user requests data and how a remote server provides them. The output maps of our system are published in an interoperable way through a Web Map Service (Open Geospatial Consortium, 2002). A WMS serves images of the data, not the actual interpolated data themselves.



We first describe the automatic interpolation algorithm followed by a section on how the algorithm was incorporated into an operational system. The discussion will point out some mathematical and operational issues and then focus on architectural considerations on how to improve interoperability and portability to other environmental variables (e.g. ozone, water pollution).

Figure 2: Simplified soil map of the Netherlands, based on Steur et al. (1985)



2. AUTOMATIC SPATIAL INTERPOLATION

2.1 Theory of spatial prediction

The real-time automatic interpolation algorithm generates maps of ambient dose rate using data from the NRM. Interpolation refers to estimation of the radioactivity level at unknown locations based on surrounding observations. The interpolation method we use is called universal kriging (UK) (Chilès and Delfiner, 1999; Journel and Huijbregts, 1978). UK can take into account trends that are present in the data, based on predictors such as soil type or precipitation. Predictions are made on a regular, 1 km x 1 km grid. The observations from the NRM $z(x_i)$, $i = 1, \dots, n$ are considered to be taken from a realization of the random function $Z(x)$, $x \in D$ that satisfies the linear model (Christensen, 1996):

$$\begin{aligned} Z(x) &= \sum_{i=1}^p \beta_i f_i(x) + e(x), \\ E(e(x)) &= 0, \\ \text{Cov}(e(x), e(x+h)) &= C(h) \quad \forall x, x+h \in D \end{aligned} \tag{1}$$

where $f_i(x)$ is the i -th known predictor at x , β_i an unknown regression coefficient, h a distance vector, C the covariance function of e and D the geographical domain of interest. Note that we assume intrinsic stationarity of the residual e , i.e. we assume that it has a constant mean and that the covariance between $e(x)$ and $e(x+h)$ only depends on the separation distance h (Chilès and Delfiner, 1999), and not on their location in the domain of interest. The best linear unbiased prediction for an arbitrary unobserved grid node x_0 , given this model and observations $z(x_i)$, $i = 1, \dots, n$, is given by:

$$\hat{Z}(x_0) = f(x_0)' \hat{\beta} + \Sigma_0' \Sigma^{-1} (z - X \hat{\beta}) \tag{2}$$

where $z = (z(x_1), \dots, z(x_n))'$ is the vector of observations, $f(x_0) = (f_1(x_0), \dots, f_p(x_0))$, X is the $n \times p$ design matrix whose i -th row equals $f(x_i)$, Σ is the variance-covariance matrix of the $e(x_i)$, $i = 1, \dots, n$ and Σ_0 is the covariance vector between the residuals of the observations and the residual of the prediction ($\Sigma_0 = (\text{Cov}(e(x_1), e(x_0)), \dots, \text{Cov}(e(x_n), e(x_0)))'$).

The best linear unbiased estimator of β is given by

$$\hat{\beta} = (X' \Sigma^{-1} X)^{-1} X' \Sigma^{-1} z(x) \tag{3}$$

The prediction error variance (kriging variance) is

$$\begin{aligned} \text{Var}(Z(x_0) - \hat{Z}(x_0)) &= \sigma_{Z(x_0)}^2 - \Sigma_0' \Sigma^{-1} \Sigma_0 + \\ &\quad (f(x_0) - \Sigma_0' \Sigma^{-1} X) (X' \Sigma^{-1} X)^{-1} (f(x_0) - \Sigma_0' \Sigma^{-1} X)' \end{aligned} \tag{4}$$

where $\sigma_{Z(x_0)}^2 = \text{Var}(Z(x_0))$. The kriging variance provides a measure of the uncertainty of the prediction. In this study we include the soil type as a predictor. This involves a stepwise trend because soil type is a categorical variable. In the stationary case, covariances are defined as $C(0) - \gamma(h)$, where $\gamma(h)$ is the variogram of e and $C(0)$ is the variance of e , which is identical to the sill of the variogram (Cressie, 1993; Christensen, 1991). The term variogram refers to the variogram model that is fitted to the sample variogram, calculated from the data. Under the assumption that the kriging prediction error is normally distributed

around the kriging prediction, approximate 95% prediction intervals can be computed by:

$$\left[\hat{Z}(x_0) - 2\sigma(x_0), \hat{Z}(x_0) + 2\sigma(x_0) \right] \quad (5)$$

where $\sigma(x)$ equals the square root of equation 4.

2.2 Automatic fitting of the variogram model

Commonly, fitting a variogram model to a sample variogram is done interactively. An expert changes the parameters of the variogram model until a visually acceptable fit is obtained. In our interpolation system this step needs to be performed automatically. First we obtain an initial guess of the variogram model. Applying the following rules of thumb to the sample variogram to obtain the initial guess seemed appropriate:

- the minimum semivariance in the sample variogram is used as initial estimate for the nugget,
- a spherical variogram model is used,
- multiplying the diagonal of the study area times 0.35 is used as an initial estimate for the range,
- the average of the maximum semivariance and the median semivariance in the sample variogram is taken as an estimate for the sill,
- bin distances used are: 0, 3, 6, 8, 12, 15, 21, 34, 48, 68, 89, 110, 138 km.

The final sample variogram is fit using the initial guess by iteratively reweighted least squares (Gauss-Newton fitting, Cressie (1993)). The weights are dependent on the number of point pairs in a particular bin (N_k) and the bin distance (h_k), and are calculated through N_k / h_k^2 . Figure 3 shows an example of an initial guess and the associated fitted variogram model.

2.3 Automatic kriging

Applying equation 2 to all unknown grid nodes results in a map. We need information regarding all predictors on both the measurement locations and the prediction locations. For the soil type this is extracted from the soil map shown in figure 2. In this study we use global kriging, i.e. all available observations are used for prediction. The variogram model was assumed to be isotropic.

Figure 3: Example of an initial guess and the associated fitted variogram model. The number of point pairs within the bin are shown next to the points in the sample variogram.

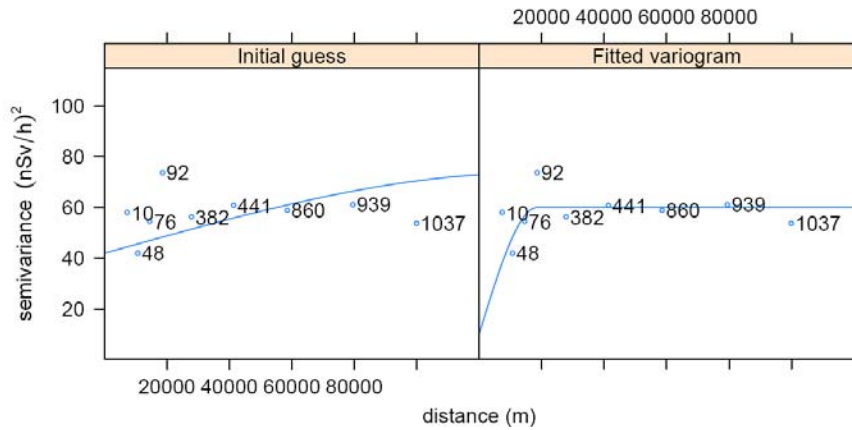
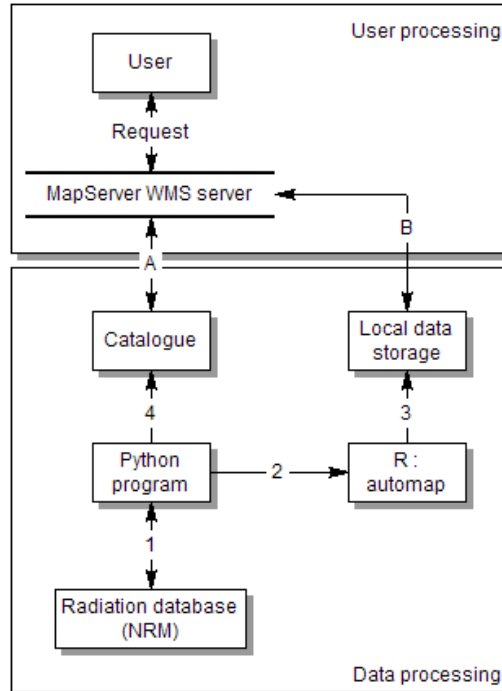


Figure 4: Flowchart describing the real-time automatic interpolation system. The data processing component completely lacks user interaction.



2.4 NRM specific issues

A few stations around the nuclear power plant of Borssele (see figure 1) show different spatial behaviour than the other stations. This caused problems in constructing the sample variogram. Not including these stations resulted in a better sample variogram. Note that choosing the stations that need to be removed was not done automatically, but was done beforehand. The lack of observations close to each other (<15 km) prohibits the accurate fitting of the nugget. Additional short range measurements were performed using mobile measuring devices. Based on these data the nugget was fixed to 11 nSv/h.

3. REAL-TIME AUTOMATIC INTERPOLATION SYSTEM

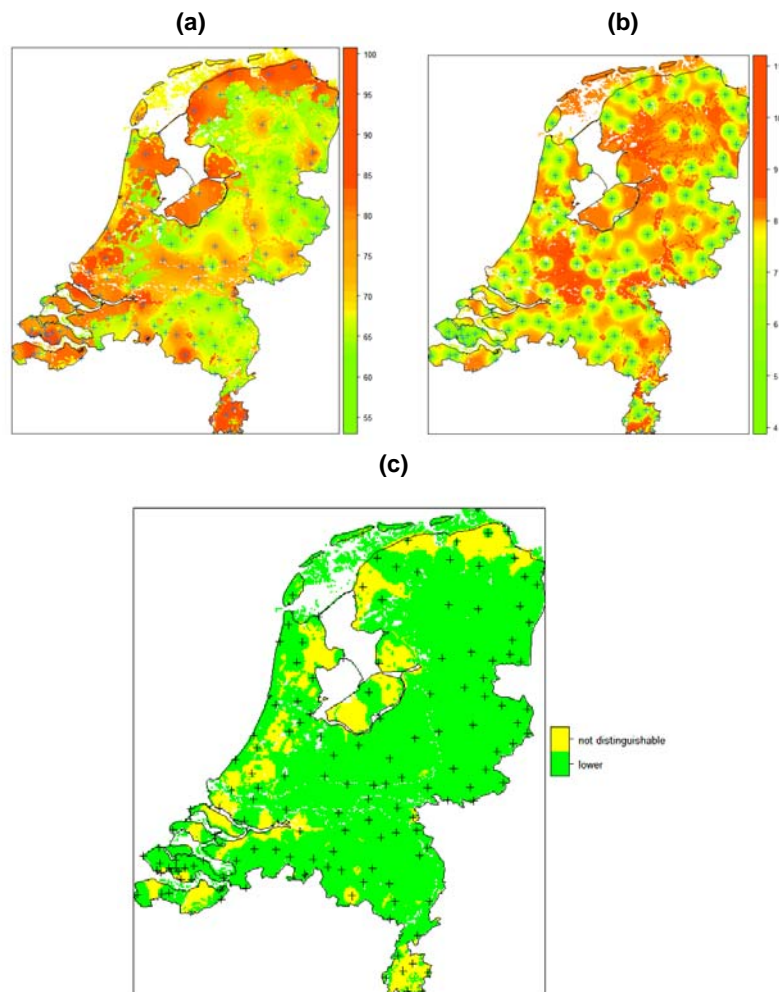
The implementation of the real-time automatic interpolation system concerns the flow of data from the monitoring network to the user who analyzes the map. Figure 4 shows a flowchart of the implemented system. The system can be divided into two parts: a part that processes the data from the National Radioactivity Monitoring network (NRM) and produces a map and a part that processes the requests made by user to view these maps.

3.1 Generating the map

We extract the NRM data, stored in text files, from an FTP server located at the RIVM using a Python (Van Rossum and Drake, 2001) program. Python then calls the statistical software package R (R Development Core Team, 2006) to perform the interpolation and produce the output maps. The interpolation in R is done using the package `automap` (Available for download at <http://intamap.geo.uu.nl/~paul/Downloads.html>). This package implements the interpolation routine described in section 2. The `automap` package uses R packages `gstat` (Pebesma, 2004) and `sp` (Pebesma and Bivand, 2005). R produces three output maps for each time step: a map with kriging predictions, a map with kriging errors ($\sigma(x_0)$) and a map where the approximate 95% prediction intervals have been classified relative to 95 nSv/h. We chose this threshold value because it clearly demonstrates the potential of this type of output. The third output type contains three classes: 1) the entire prediction interval is below the threshold, 2) the entire prediction interval is above the threshold and 3) the threshold is within the prediction interval and is not distinguishable from the predicted value. R stores the output maps in GeoTIFF files on the hard drive of the server, this is done using the R-package `rgdal` (Keitt et al., 2007). The colour scale is fixed for all GeoTIFF files. The colours were chosen to give good contrast between areas with high and low background radiation. Figure 5 shows an example for each output type. A Python script finally updates the catalogue

database that records where the maps associated with a certain time step are stored.

Figure 5: An example for each type of output from the automatic interpolation service: (a) kriging prediction (nSv/h), (b) kriging error (nSv/h) and (c) approximate 95% prediction interval classified relative to 95 nSv/h. The NRM stations are marked by crosses.

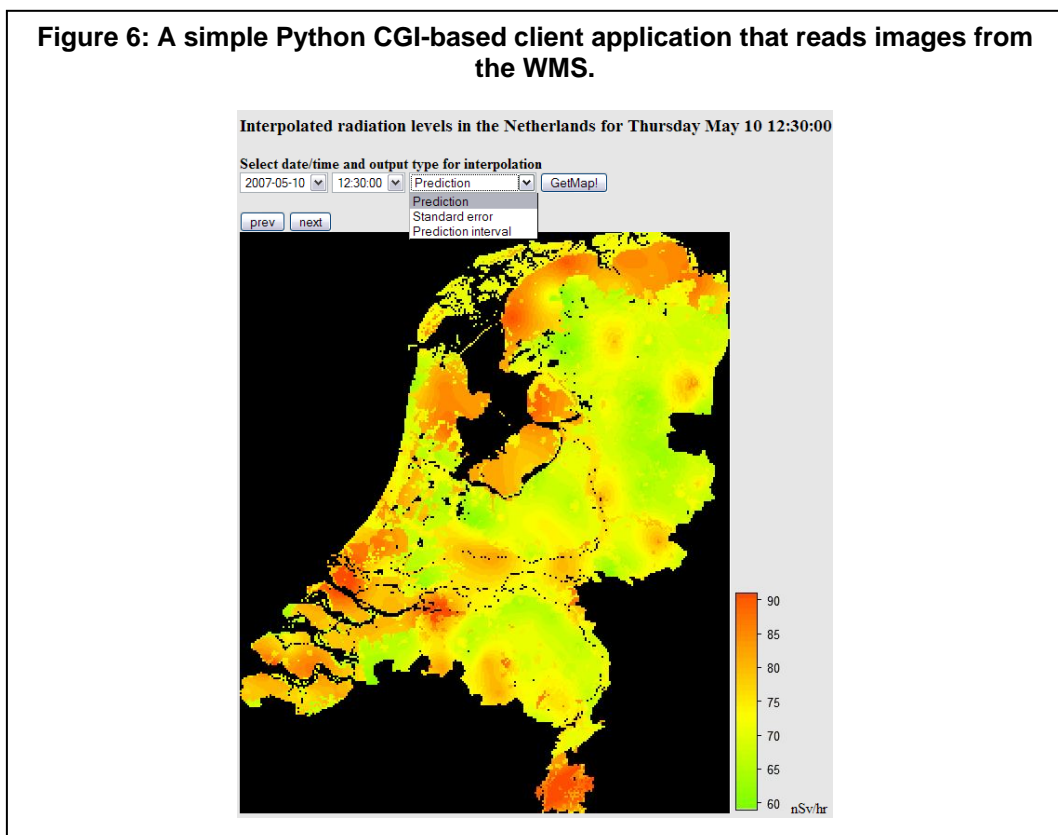


3.2 Serving the map

The maps are made available in an interoperable way through a Web Map Service (WMS, Open Geospatial Consortium, 2002). A WMS allows three types

of operations: `GetCapabilities`, `GetFeatureInfo` and `GetMap`. The first operation provides information on what the WMS has to offer, in our case three layers for each available time step, one for each type of output. A user can request more detailed information about a specific layer using the `GetFeatureInfo` operation. `GetMap` returns the map the user specifies. We chose to implement the WMS using the open source program UMN MapServer (Vatsavai et al., 2006). To read and display the maps from the WMS we developed a simple CGI-based client in Python. A screenshot of this client is shown in figure 6.

Figure 6: A simple Python CGI-based client application that reads images from the WMS.



4. DISCUSSION

4.1 Mathematical and operational considerations

An important assumption underlying kriging is intrinsic stationarity of the residual (section 2.1), $e(x)$. In case of a large, local release, a small group of stations may have values that are up to two or three orders of magnitude larger than background ambient dose rates. In that case the assumption of intrinsic

stationarity is highly unrealistic and fitting a variogram model becomes difficult (Pebesma, 2005). A possible way of dealing with these outliers is to use data transformation techniques (e.g. Box-Cox transformation (Christensen, 1996)).

Universal kriging includes trends into the prediction. In our case we only used the soil type as an additional source of information. Including other sources of additional information could improve the prediction. One example of such a source is the amount of precipitation, which can be derived from rainfall radar images. Precipitation causes wet deposition of radon progeny from the atmosphere. This results in short-lived ground surface activity, increasing the ambient dose rate (Smetsers and Blaauboer, 1997). Another candidate predictor is the height of the mixing layer, as this influences the radionuclide activity concentrations in this layer. The output of an atmospheric dispersion model (NPK-PUFF, Van Egmond and Kesseboom (1983)) could also be used as a trend in case of a radioactive release.

4.2 System architecture considerations

The interoperability of our system is limited in the way the output of the interpolation is delivered. The prototype system is highly customized towards the problem of interpolating radiation measurements:

- data from the NRM network are retrieved using text files stored on an FTP server,
- the additional data (soil map) are stored on the server's hard drive,
- the code to retrieve the data, interpolate these and save the results is customized and is not directly portable to other environmental variables,
- the results are presented through a WMS, providing the user with little flexibility with regard to the output.

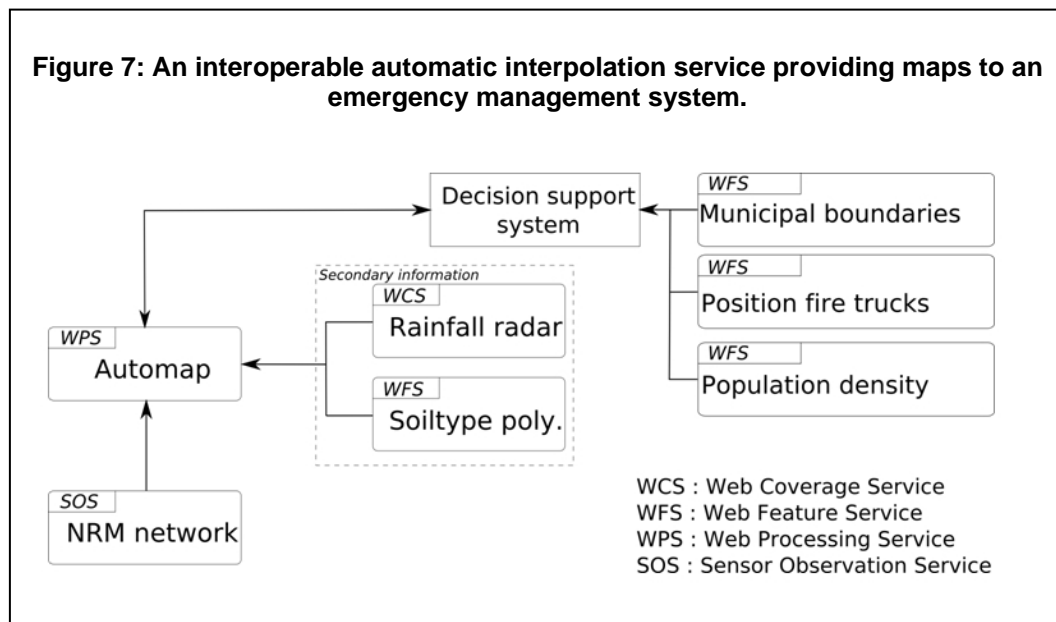
There are OGC specifications that could make the automatic interpolation system much more generic, easier to maintain and easier to port to other environmental variables. Figure 7 shows a flowchart that uses a number of OGC specifications to improve interoperability and portability. The different OGC specifications and their role in the system are now briefly discussed. The data from the NRM are implemented in a *Sensor Observation Service* (Open Geospatial Consortium, 2006c). A SOS supports three core operations:

- `GetCapabilities` returns the metadata on the SOS,
- `DescribeSensor` returns information on the sensor that performs the measurements,
- `GetObservation` gives access to the available measurements of the SOS.

A *Web Processing Service* (Open Geospatial Consortium, 2008) can be used to implement the automatic interpolation service described in section 2. A WPS can implement any kind of GIS operation, such as map algebra or buffering. It provides three operations:

- `GetCapabilities` returns the metadata on the WPS, e.g. available processes for execution,
- `DescribeProcess` returns information on a specific process such as input parameters or the resulting outputs,
- `Execute` starts the specified process. The user needs to provide all necessary input parameters and receives the output in response.

Implementing the interpolation service as a WPS makes it much easier to use for other environmental variables. A WPS also provides the user with a means of directly influencing the interpolation through the input parameters.



The additional data are usually provided by third parties, e.g. a weather service in case of rainfall radar. Additional data in raster format can be provided by a *Web Coverage Service* (Open Geospatial Consortium, 2006a). A WCS supports the same kind of operations as a WMS (see section 3.2) but returns a raster rather than an image. Additional data consisting of points, lines or polygons can be accessed through a *Web Feature Service* (Open Geospatial Consortium, 2006b). A WFS supports the following operations:

- `GetCapabilities` returns the metadata on the WFS, i.e. the available features,
- `DescribeFeatureType` returns information on the requested feature type,
- `GetFeature` returns the requested features allowing spatial and non-spatial queries.

A special WFS type is the Transaction WFS that allows users to modify features, that is to create, update or delete geographic features. In figure 7 the output of the WPS is retrieved by an emergency management system. It uses the output rasters from the WPS and combines these with other relevant data. The position of fire trucks and a municipal boundary layer map are just a few examples of the kinds of data that can be incorporated into this kind of emergency management system.

5. CONCLUSIONS

The automatic real-time interpolation system introduced in this paper presents a significant improvement in interpretation of data from the Dutch National Radioactivity Monitoring network. The system creates maps using universal kriging in combination with an automatic variogram fitting routine. Kriging delivers both a prediction and an associated prediction error estimate. Under additional assumptions the prediction and the error yield approximate 95% prediction interval maps. These intervals can be used to assess where a given threshold has been exceeded or not. The resulting maps are published through a Web Map Service, making these results interoperable with existing systems. Soil type is the only predictor currently included in the system. Adding more predictors would likely improve the prediction. Candidate predictors include rainfall radar images, height of the mixing layer and the output of atmospheric dispersion models. The interpolation routine is not suitable for data containing large outliers. The assumption of intrinsic stationarity is not very realistic in this case. A possible way to deal with outliers is to use data transformation techniques. Implementing all components of the system in OGC specified web services would make the system interoperable at a more basic level and would make the system more portable to other environmental variables.

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 like to thank the four anonymous reviewers whose comments improved this paper.

REFERENCES

- Chilès, J., Delfiner, P., (1999). *Geostatistics: Modeling Spatial Uncertainty*. John Wiley & Sons, New York.
- Christensen, R., (1991). *Linear Models for Multivariate, Time Series and Spatial Data*. Springer, New York.
- Christensen, R., (1996). *Plane Answers to Complex Questions: The Theory of Linear Models*, 2nd Edition. Springer, New York.
- Cressie, N. A., (1993). *Statistics for spatial data*. Wiley, NY.
- Dubois, G. (Ed.) (2003) *Mapping radioactivity in the environment*. Spatial Interpolation Comparison 1997 Office for Official Publications of the European Commission, Luxembourg
- Dubois, G. (Ed.) (2005) *Automatic mapping algorithms for routine and emergency monitoring data*. Report on the Spatial Interpolation Comparison (SIC2004) exercise Office for Official Publications of the European Commission, Luxembourg
- Dubois, G. and Galmarini, S. (2005) *Introduction to the Spatial Interpolation Comparison (SIC) 2004 Exercise and Presentation of the Datasets*. Applied GIS 1 (2)
- ICRU, (1993). *Quantities and units in radiation protection dosimetry*. ICRU report 51. Tech. rep., Bethesda MD.
- Journel, A., Huijbregts, C., (1978). *Mining Geostatistics*. Academic Press., London.
- Keitt, T. H., Bivand, R., Pebesma, E., Rowlingson, B., (2007). rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.5-15`<http://www.gdal.org>, <http://rgdal.sourceforge.net/>, <http://sourceforge.net/projects/rgdal/>
- Open Geospatial Consortium, (2002). *OpenGIS Web Map Service (WMS) Implementation Specification, version 1.1.1*. Available at <http://www.opengeospatial.org/standards/wms>.
- Open Geospatial Consortium, (2006a). *OpenGIS Web Coverage Service (WCS) Implementation Specification, version 1.1.0*. Available at <http://www.opengeospatial.org/standards/wcs>.

- Open Geospatial Consortium, (2006b). *OpenGIS Web Feature Service (WFS) Implementation Specification, version 1.1.* Available at <http://www.opengeospatial.org/standards/wfs>.
- Open Geospatial Consortium, (2006c). *Sensor Observation Service (SOS): Request for Public Comments.* Available at <http://www.opengeospatial.org/standards/requests/32>.
- Open Geospatial Consortium, (2008). *Web Processing Service (WPS) . Implementation Specification, version 1.0.* . Available at <http://www.opengeospatial.org/standards/wps>.
- Pebesma, E. J., (2004). *Multivariable geostatistics in S: the gstat package.* Computers and Geosciences 30 (7), 683–691.
- Pebesma, E. J., (2005). *Mapping radioactivity from monitoring data: Automating the classical geostatistical approach.* Applied GIS 1 (2).
- Pebesma, E. J., Bivand, R. S., (2005). *Classes and methods for spatial data in R.* R News 5, 9–13.
- R Development Core Team, (2006). *R: A Language and Environment for Statistical Computing.* R Foundation for Statistical Computing, Vienna, Austria, ISBN 3-900051-07-0. <http://www.R-project.org>
- Smetsers, R.C.G.M. and Blaauboer, R. O., (1997). *Source-dependent probability densities explaining frequency distributions of ambient dose rate in the Netherlands.* Radiation Protection Dosimetry, 69, 33-42.
- Steur, G., De Vries, F., Van Wallenburg, C., (1985). *Bodemkaart van Nederland 1:250.000 : beknopte beschrijving van de kaarteenheden.* Stiboka, pp 52.
- Twenhöfel, C. J. W., de Hoog van Beynen, C., van Lunenburg, A. P. P. A., Slagt, G. J. E., Tax, R. B., van Westerlaak, P. J. M., Aldenkamp, F. J., (2005). *Operation of the Dutch 3rd generation National Radioactivity Monitoring network.* In: Dubois, G. (Ed.), (2005), pp. 19–31.
- Van Egmond, N., Kesseboom, H., (1983). *Mesoscale air pollution dispersion models. II. lagrangian puff model and comparison with eulerian grid model.* Atmospheric Environment 17 (2), 267–274.
- Van Rossum, G., Drake, F., (2001). *Python Reference Manual.* PythonLabs, Virginia, USA, available at <http://www.python.org>.
- Vatsavai, R., Shekhar, S., Burk, T., Lime, S., (2006). *UMN-mapserver: A high-performance, interoperable, and open source web mapping and geo-spatial analysis system.* In: 4th International Conference on Geographic, Information Science, GIScience 2006. Vol. 4197 LNCS. Munster, pp. 400–417.