Proceedings of Global 2015
September 20-24, 2015 - Paris (France)
Paper 5428

A GEOSTATISTICAL APPROACH TO CALCULATE VOLUMES OF CONTAMINATED SOIL AND GROUNDWATER AT SELLAFIELD

Yvon Desnoyers[1], Claire Faucheux[1], Nick Jefferies[2], Ian Teasdale[3], Julian Cruickshank[3]

[1]*Geovariances*
49bis avenue Franklin Roosevelt, 77215 AVON Cedex, France
Tel: +33 (0)1 6074 9090, Fax: +33 (0)1 6422 8728, Email:desnoyers@geovariances.com

[2]*Amec Foster Wheeler plc- Nuclear Services / Clean Energy Europe*
Building 150, Harwell Oxford, Didcot, Oxfordshire, OX11 0QB, United Kingdom
Tel: +44 (0)1635 280 300, Fax: +44 (0)1635 280 421

[3]*Sellafield Ltd, Seascale, Cumbria, CA20 1PG United Kingdom*

Abstract – *Sellafield is the UK facility for nuclear fuel reprocessing. It is an operational site located adjacent to the coast in west Cumbria and occupies an area of around 3 km^2. Radioactive material has entered the sub-surface environment following accidental leaks during historical operations. Sellafield Ltd manages contaminated land and groundwater arising from such leaks through a strategy of characterisation, monitoring and risk modelling prior to a final hazard reduction and remediation phase. The monitoring and characterisation programmes generate a large quantity of important environmental data gathered at public cost. Sellafield Ltd wishes to ensure that appropriate methods are being used for the analysis of these data. This paper describes the benefits of a geostatistical approach for achieving these aims and summarises the main steps in the analysis procedures. Prior to geostatistical analysis, both soil and groundwater quality datasets were validated to ensure the quality of the data. Geostatistical models were then developed and applied to determine both the best estimates of contaminated volumes of soil and groundwater and the uncertainties in those estimates. Spatial mapping of contaminated soil and groundwater has also been performed, with an emphasis on calculating the probability of contaminated materials being present at each location.*

I INTRODUCTION

Sellafield is the UK facility for nuclear fuel reprocessing. It is an operational site located adjacent to the coast in west Cumbria and occupies an area of around 3 km^2. Radioactive material has entered the sub-surface environment following accidental leaks during historical operations. Sellafield Ltd, the operator of the Sellafield site, manages contaminated land and groundwater arising from such leaks through a strategy of characterisation, monitoring and risk modelling prior to a final hazard reduction and remediation phase.

The monitoring and characterisation programmes generate a large quantity of important environmental data gathered at public cost. Sellafield Ltd estimates that more than 200,000 data items relating to soil and groundwater quality at the Sellafield site have been generated over a period of approximately 30 years. They include results from the Sellafield Contaminated Land Study (2004), the Sellafield Contaminated Land and Groundwater Management Project (SCLGMP, 2007-2010) and the routine groundwater sampling programme [1 & 2].

As part of a previous study (SCLGMP), the volume of radioactively contaminated soil at the Sellafield site was estimated by assuming that a single radionuclide fingerprint describes the contaminated soil and that the proportions of contaminated soil in the land quality dataset were equal to those in the ground beneath the Sellafield site. The total volume of soil contaminated with radioactivity from past activities at Sellafield was estimated to be approximately 14 million cubic metres. This is a very large volume, and uncertainty in this estimate (which could not be assessed using the approach adopted) is a major cause of uncertainty when assessing options for managing the radioactively contaminated land. Further, this approach suffers from a number of drawbacks, most notably the assumption that the measurements on soils from Sellafield Ltd represent an unbiased sample from the population.

Sellafield Ltd wishes to ensure that appropriate methods are being used for the analysis of these data. This paper describes the benefits of a geostatistical approach for achieving these aims and presents an understanding of spatial correlations and spatial structure of the data. The output from these calculations is an improved understanding of the volumes and total radioactivity inventories of contaminated soil and groundwater at the Sellafield site and of the uncertainties in these calculated volumes and inventories.

II DATA REVIEW

II.A. Soils datasets

In reality, the spatial distribution of soil samples from the Sellafield site is not uniform. See Figure 1. The main reason for this is limitations on borehole locations due to the presence of buildings and extensive sub-surface features such as cables, pipes and drains. As a result of the above constraints, most boreholes (particularly those in the more built-up central part of the site) are instead constructed on or adjacent to roadways and other open areas.

In addition, the expected presence of high levels of radioactive contamination in soil in some localised areas can also constrain drilling locations. Conventional drilling is generally not undertaken in the most contaminated areas because of concerns about the ability to subsequently release the drilling equipment from the Sellafield site.

The combined soil quality dataset considered for use in the geostatistical analysis contains over 14,000 measurements of each of gross alpha and gross beta activity. By comparison, there are only approximately 3,000 gamma spectrometry measurements; caesium-137 (Cs-137) is the most important radionuclide detected in soil by this technique. There are even fewer soil samples where radionuclide activities have been measured following radiochemical separation. For example, there are only a few hundred measurements for strontium-90 (Sr-90), which is the most important pure beta-emitting radionuclide in the contaminated soil. There are two possible approaches for undertaking geostatistical analysis on soils at Sellafield. Firstly, use radionuclide-specific data; second, use ‘gross alpha activity’ and ‘gross beta activity’ data.

If radionuclide data were to be used, geostatistical calculations would need to be undertaken on individual radionuclides. Either one radionuclide, such as Cs-137, or a group of radionuclides could be evaluated. The latter approach, which is more rigorous as it would take account of the variability of radionuclide proportions across the site, would be limited to samples where data are available for the main radionuclides in the soil; only a few hundred samples

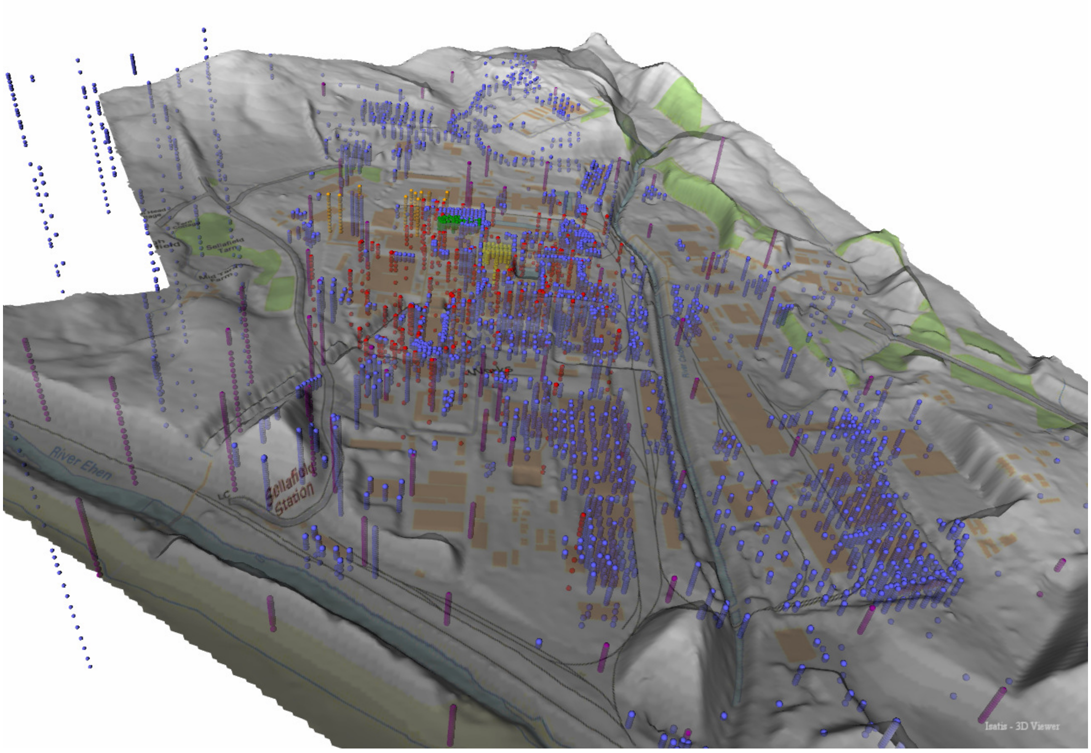


Fig. 1. Soil samples location and DTM surface.

At Sellafield, gross beta activities are determined by infinite depth counting using Geiger-Müller counting. The counting efficiency is lower (or zero) for low energy beta particles such as those emitted from hydrogen-3 (tritium), carbon-14 and plutonium-241. The sample is also generally dried and homogenised prior to counting, so volatile radionuclides are lost. Second, the measured gross activity in the soil includes contributions from a number of radionuclides (albeit weighted according to counting efficiency). Therefore, it is necessary to use an approach to calculate the total activity (i.e. the sum of all radionuclides) of the material. This requires knowledge of the radionuclide fingerprint.

Gross alpha/beta has been analysed on five times as many soil samples as Cs-137, and on more than more than 60 times as many soil samples as Sr-90 and Cs-137. Further, Cs-137 mobility in soil is much lower than Sr-90 mobility, and hence a higher proportion of Cs-137 analyses than gross beta analyses in soils are at background levels. Therefore, notwithstanding the limitations of gross alpha and gross beta measurements (see above), we decided to undertake the geostatistical analysis on the gross alpha and gross beta data.

To provide a cautious estimate for contaminated soil volumes, we assume that no remediation of soils has taken place after sampling.

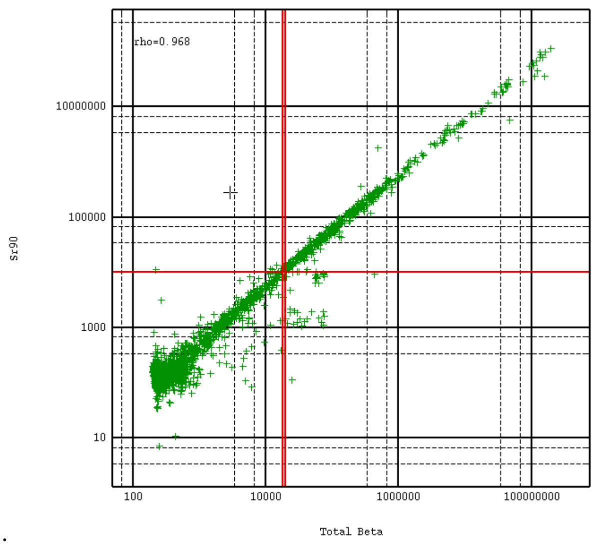
A quality control is first done thus ensuring the reliability of data: coordinates problems (checking location and altitude of boreholes along with depths of samples), erroneous dates, LOD handling, homogenization of samples lengths.

II.B.Groundwater datasets

Routine groundwater quality data from the beginning of 2004 to end of 2014 were made available to the project. The largest dataset is for the first quarter of 2009, and we used this time interval for the geostatistical calculations that we present here.

We initially considered using a Water Index approach to calculate volumes of contaminated groundwater. This would have involved summing together the activities of key radionuclides in groundwater and identifying the volume where WHO drinking water guideline value for the water was exceeded. A difficulty was that not all groundwater samples were analysed for all determinands; this meant that the Water Index could not be calculated on a consistent basis. Therefore, we instead selected gross beta activity as the parameter to use in the geostatistical calculations.

Fig. 2 shows the very good linear correlation between gross beta activity and Sr-90 activity in groundwater; the coefficient between variables is about two. This good correlation is consistent with Sr-90 and its short-lived daughter Y-90 being the dominant beta-emitting radionuclides measured in groundwater by the gross beta measurement technique used at Sellafield.



.

Fig. 2. Correlation between Sr-90 and gross beta, with threshold of 10,000 Bq/m^3 for Sr-90 and 18,000 Bq/m^3 and 20,000 Bq/m^3 for gross beta

II.C.Geometry of 3D domain

This section defines and justifies the boundaries of the volume of ground within which the geostatistical calculations are undertaken. For contaminated soil, the boundaries are:

- the maximum area considered is bounded by the Sellafield site perimeter (see Fig. 3). This is a reasonable assumption, as no significant radioactive contamination of soil has been observed outside the Sellafield site boundary. In addition, calculations are also undertaken on the area bounded by the Separation Area perimeter (the inner boundary on Fig. 3)
- the upper surface is ground level, which is a reasonable assumption given that contamination of the ground surface is known to occur. Further, we assume that all building foundations are at ground level (i.e. we assume that all sub-surface material is soil); this is a cautious assumption
- the lower surface is the bedrock surface. That is, we calculate the volume of contaminated ground within Made Ground and the superficial deposits that overlie bedrock. Beneath the Sellafield site, bedrock is always overlain by Made Ground and superficial strata. Beneath Separation Area, depth to bedrock is at least 15m, and is up to 55m within the area of a buried valley feature. We have reviewed the extent of radioactive contamination within bedrock; only a small area of bedrock, in the vicinity of one facility, contains levels of radionuclides that would categorise the material as radioactive waste (predominantly VLLW) if excavated. For the purposes of the waste volume calculations, we have disregarded this small

volume. Finally, choosing bedrock as the lower surface is also a pragmatic choice, as substantial excavation of bedrock is unlikely to be feasible.

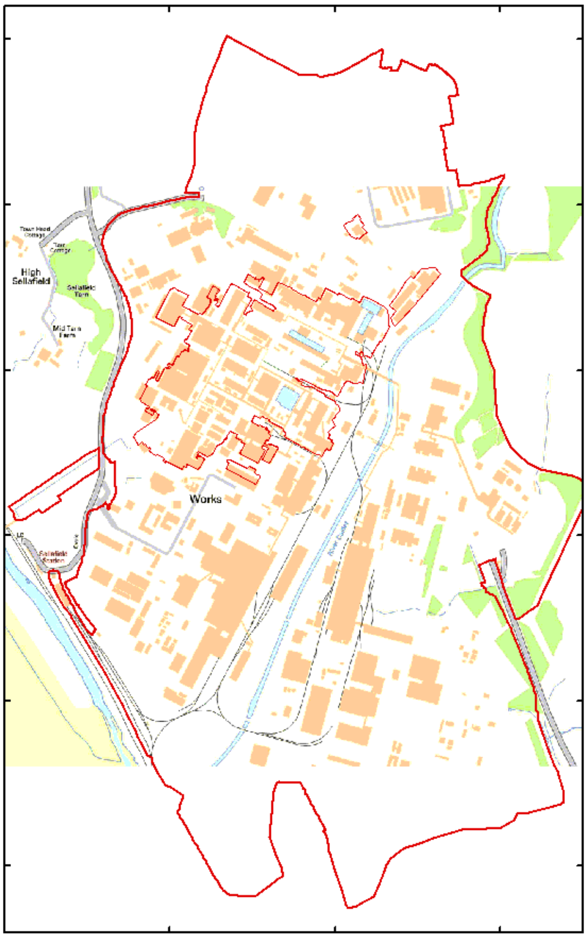


Fig. 3. Delineation of the site (bold line) and Separation Area (inside) boundaries.

Sellafield Ltd has provided a Digital Terrain Model (DTM) of the area on a 10m mesh basis. This has been used to generate the ground surface elevation. For the bedrock surface elevation the mesh is 5m and is obtained from the ConnectFlow groundwater flow model of the Sellafield area. For the purpose of the geostatistical analysis, these surfaces have been interpolated using the linear kriging technique on a 5m mesh grid. They constitute the vertical limits of the interpolated domain for volume estimates.

For the geostatistical calculations, the above volume has been divided into two sub-volumes: above the groundwater table and below the groundwater table. The rationale is twofold:

- the mechanisms of radionuclide transport are different in the two sub-volumes. Above the groundwater table, transport is predominantly downwards, driven by infiltrating water. Below the groundwater table, transport is predominantly sub-horizontal within flowing groundwater. The spatial structure of contamination in the two zones may be different, and this should be explored
- it would be substantially more difficult to excavate soils from below the groundwater table. Determining the overall volumes of contaminated soil above and below the groundwater table is useful when considering practicable remediation options that might be considered.

Interpolation of the groundwater table elevation is based on the January or October 2009 data, using data from the shallowest piezometer in the case of boreholes installed with several piezometers. Interpolation is done using linear kriging on the same area as the DTM. Inconsistencies after interpolation where water table is greater than ground level are removed and the altitude set equal to DTM.

II.D.Contaminated soil and groundwater categories

For the purposes of this study, radioactively contaminated soil at Sellafield has been divided into categories based on the UK solid radioactive waste categories. The following are relevant:

- Intermediate Level Waste. Greater than 4GBq/te (Alpha activity) or 12 GBq/te (Beta activity). Upper limits are not relevant to this study.
- Low Level Waste. Less than 4GBq/te (Alpha activity) or 12 GBq/te (Beta activity) and greater than 4 MBq/te. Other than limits on Total Alpha and Beta activity, there are no radionuclide-specific limits on LLW. Sellafield Ltd further sub-divides this category based on existing arrangements for solid LLW disposal at Sellafield:
    - >200 MBq/te (LLW Upper)
    - 40 MBq/te - 200 MBq/te (LLW Middle)
    - 4 MBq/te - 40 MBq/te (LLW Lower)
- High Volume VLLW. Less than 4 MBq/te and greater than the 'Out Of Scope value' of Environmental Permitting (Amendment) Regulations 2011. The Out of Scope value for contaminated soil at Sellafield (0.80 MBq/te) has been calculated using the radionuclide fingerprint of the contaminated soil and the sum of quotients approach described in the guidance to the Environmental Permitting Regulations.

It is important to note that these categories are used for illustrative purposes, and should not be taken to imply that Sellafield Ltd's preferred approach is to excavate the soil and produce it as waste.

For the purposes of this report, we follow the Sellafield Ltd approach that the measured gross beta activity for a soil sample is assumed to be equal to the sum of the activities of strontium-90, yttrium-90, caesium-137 and barium-137m, plus the activity of naturally occurring beta-emitting radionuclides. As these radionuclides generally represent more than 90% of all beta-emitting radionuclides in the soil, it follows that the gross beta activity is approximately equal to the total beta activity (i.e. the calculated sum of the activities of all beta-emitting radionuclides). Similarly, we assume that the measured gross alpha activity of soil is equal to the total alpha activity (i.e. the sum of the activities of all alpha-emitting radionuclides).

Given these assumptions, the waste categories described in Section II D can be defined in terms of measured gross alpha and gross beta activities.

For groundwater, the WHO drinking water guidance value of 10,000 Bq/m^3 for Sr-90 was chosen as the threshold value, above which groundwater was considered to be contaminated. Figure 2 shows how the Sr-90 threshold value of 10,000 Bq/m^3 was used to determine a corresponding threshold for gross beta activity. Building the linear regressions based on all available data or only on data for the first quarter of 2009 gives corresponding gross beta thresholds of 18,000 Bq/m^3 and 20,000 Bq/m^3 respectively.

III VOLUME OF CONTAMINATED SOIL

Estimation of contaminated soil is based on the geostatistical analysis of gross alpha and gross beta analyses. All geostatistical calculations have been undertaken using the 'as measured' gross alpha and gross beta activities. This preserves the integrity of the data. Subtraction of natural background activity and calculation of the volumes of the waste categories described in Section II D is undertaken as a 'post-processing' activity. This ensures that it will not be necessary to rerun the geostatistical calculations in the event that some of the assumptions are changed in the future. The following natural backgrounds have been derived from the data: 600 Bq/kg gross alpha activity; 800 Bq/kg gross beta activity.

III.A. Exploratory Data Analysis

Histograms (Fig. 4) of Gross Alpha and Beta show skewed statistical distributions with many low values and few high values.

Some correlations have also been checked:

- gross alpha and gross beta show different behaviour and no linear correlation,
- gross beta and Sr-90 present a good correlation with a linear coefficient of 0.92.

Gaussian anamorphosis modelling is then performed for both gross alpha and gross beta so as to get a better definition and understanding of the spatial behaviour of the data from these very skewed distributions. It is also a necessary precursor to performing the geostatistical simulations.

III.B. Variography

The whole point of the geostatistical methodology is to take into account the spatial continuity of the phenomenon in order to predict it at unsampled locations and to quantify the prediction uncertainty. The characterization of this spatial continuity, or spatial variability, is an essential stage which is performed through the variographic analysis [3].

The experimental variogram is calculated by averaging, within classes of distance, the variability contribution of each pair of data points; this contribution is usually quantified by the half squared difference of the measured values.

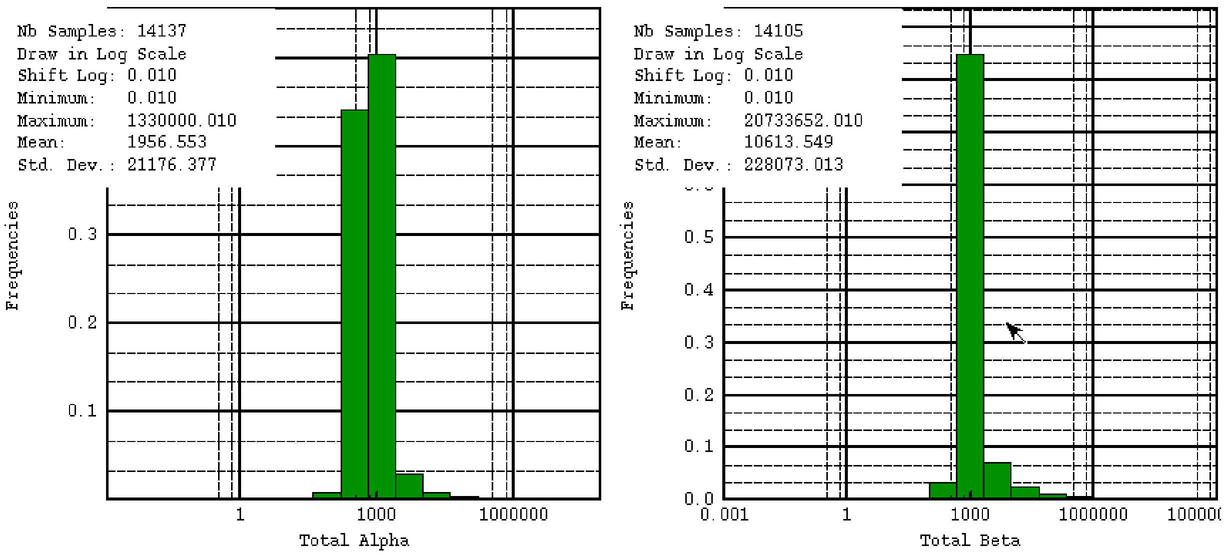


Fig. 4. Histograms of Gross Alpha and Gross Beta (log scale).

Generally, for a structured phenomenon, the spatial variability increases with distance and tends to stabilize ("sill") at a distance named "range". Data separated by a distance larger than the range are no longer spatially correlated.

In the multivariate case, the geostatistical methodology remains the same except that this is now a multivariate variographic analysis with two simple variograms for the two single variables and one cross-variogram which describes the spatial behavior of the correlation between the two variables. It can be extended to a larger number of input data but then requires a more complex inference of the spatial structure.

To investigate anisotropy in the datasets, experimental variograms for gross alpha and gross beta activity have been constructed for the vertical direction and four horizontal directions (oriented at 0°, 45°, 90° and 135°). Fig. 5 shows that there is anisotropy in both gross alpha and gross beta between the vertical and horizontal directions. Further, horizontal behaviours are the same for all four directions considered. In subsequent analysis, two experimental variograms are considered: one for the vertical direction and one for the horizontal plane.

In order to check whether the spatial structure of the data is different in the unsaturated and saturated zones (i.e. above and below the groundwater table position in October 2009), experimental variograms were computed for the vertical and horizontal directions in each region. Results show some small differences between variograms. Total variability appears smaller below the groundwater table, which is probably due to the more limited number of points available for the computation of the vertical variogram in this region. Given these results, we decided that all further geostatistical calculations would work with a single variogram calculated using all data (i.e. from above and below the groundwater table).

Thresholds being used for the study relate to total activity, which is computed as a linear combination of gross alpha and gross beta activities. If these variables are modelled separately and present a spatial correlation, computing the linear combination and volumes on the result is biased as explained below.

If

$$Total\ Activity\ (x_i) = a \times Gross\ Alpha\ (x_i) + b \times Gross\ Beta(x_i),$$

Then

$$Total\ activity^*(x) = a \times Gross\ Alpha(x)^{CoK} + b \times Gross\ Beta\ (x)^{CoK}$$ with CoK for cokriging,

But

$$Total\ activity^*(x) \neq a \times Gross\ Alpha(x)^{K} + b \times Gross\ Beta\ (x)^{K}$$ with K for kriging.

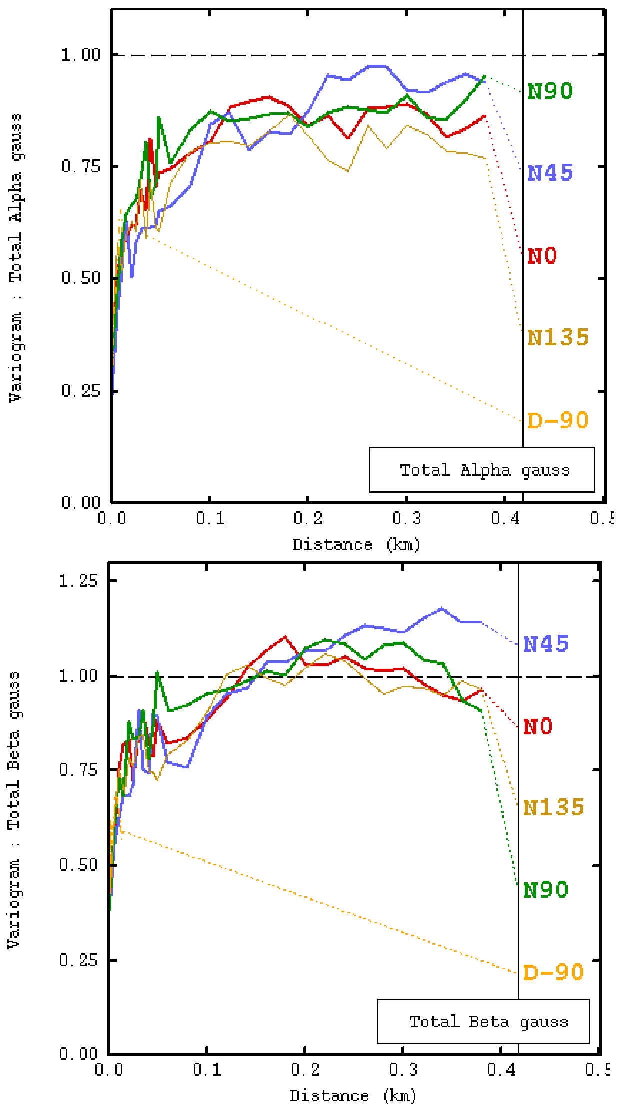


Fig. 5 Experimental variograms for Gross Alpha and Gross Beta, computed for the vertical direction (D-90) and four horizontal directions (N0 to N135).

Fig. 6 presents experimental variograms (left: horizontal, right: vertical) and the corresponding models for gross alpha (bottom), gross beta (top) and the cross-variogram (middle). The results clearly shows that the cross variogram presents a spatial correlation. Therefore it is necessary to build a multivariate variogram to take account of the spatial variability of gross alpha and gross beta in the contaminated soil at Sellafield.

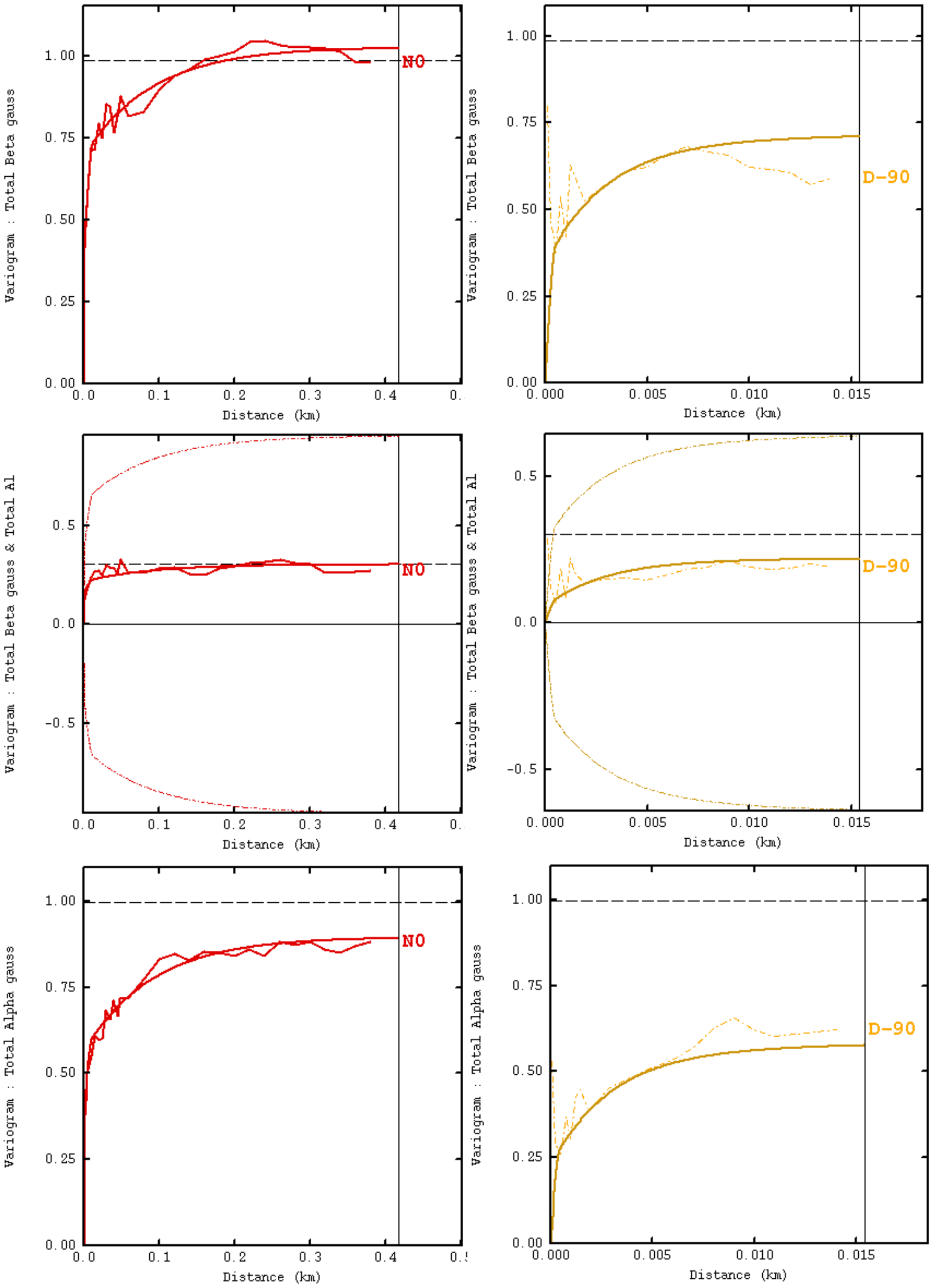


Fig. 6. Experimental variograms (left: horizontal, right: vertical) and corresponding model for Gross Alpha (bottom), Gross Beta (top) and the cross-variogram (middle).

III.C.Estimation of contaminated soil

Quantifying contaminated volumes exceeding a regulatory threshold or estimating the global pollutant mass require additional tools than just kriging [3]. Indeed, the latter is smoothing the real variability of the target parameter and could lead to significant bias if used to estimate contaminated volumes. Therefore, the use of stochastic simulations is usually recommended. We have used such simulations to estimate the volume of contaminated soil and its uncertainty, and also to determine the probability of soil at any location in the modelled volume exceeding a target threshold.

Simulations are made using the Turning Bands simulations method [4]. One hundred simulations are computed over a 5 x 5 x 0.5m grid (cells of $12.5m^3$). The volume is bounded by the ground surface (the upper surface), the bedrock surface (the lower surface) and a perimeter (the horizontal extent). The perimeter is either the Sellafield site boundary or the Separation Area boundary. See Fig. 3.

Neighbourhood is chosen so as to ensure there are enough data to perform the simulation and not too many to avoid excessive computation times. Simulations are made using a neighbourhood with a horizontal radius of 400m and a vertical radius of 10m, in accordance with the ranges seen on the variograms. The minimum number of neighbours requested to perform the estimation is 15 and maximum is set to 50. Estimations are not made in a small number of locations because of the distance to the nearest neighbours. Even in Separation Area, where the density of boreholes is high, there are a few areas at depth where the nearest neighbours are too distant (there are fewer samples in depth than close to the surface) and where the simulations cannot be performed. In the area bounded by the Sellafield site perimeter, $38{,}948{,}675m^3$ are estimated, corresponding to 96.26 % of the total volume; in the Separation Area, $8{,}659{,}500m^3$ is estimated, corresponding to 97.45 % of the total volume.

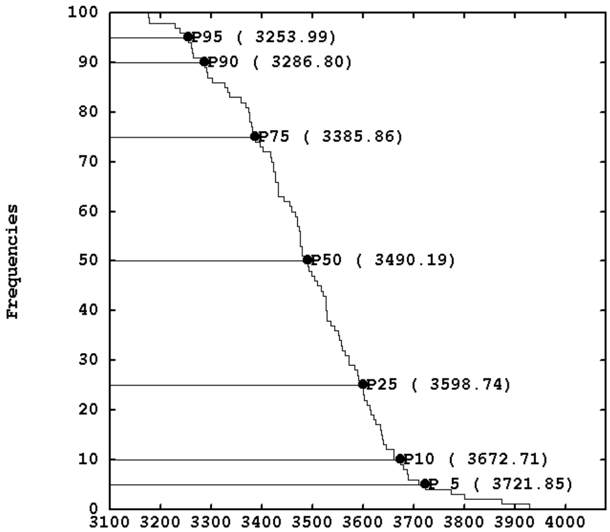


Fig. 7. Volume quantification for one of the contaminated soil categories. Units are $1{,}000m^3$

Fig. 7 shows an example of the output from geostatistical calculations. It presents the cumulative distribution function (CDFs) for one of the categories described in Section II D, and the uncertainties in the volume. The statistical distribution of the volumes is skewed for high values and extreme quantiles. Similar CDFs have been produced for each soil category and have been use to determine the volume of contaminated land at Sellafield and the associated uncertainties. An equivalent set of CDFs has also been produced to show the total radioactivity inventory for each soil category and its uncertainty.

The geostatistical calculations have also been used to calculate the relative proportions of contaminated soil (expressed by category and as volumes) above and below the groundwater table, and the proportion of contaminated soil that lies beneath Separation Area. From Table 1, it is evident that about half the volume of contaminated soil is above the groundwater table (position at October 2009) and that higher proportions of the more contaminated soil categories lie beneath Separation Area.

TABLE 1

Relative distributions of contaminated soil above and below the groundwater table and in Separation Area

	Above groundwater table	Below groundwater table	Proportion beneath Separation Area
LLW	51%	48%	37%
LLW Upper	42%	57%	52%
LLW Middle	49%	51%	43%
LLW South Tip	53%	48%	34%
VLLW	56%	43%	28%

For each contaminated soil category, the probability of each cell in the modelled volume belonging to that category has been computed. This analysis provides information to Sellafield Ltd on the likelihood of encountering contaminated soil in future excavations at the site, and adds to the tools that are available for planning and managing such excavation work.

Fig 8 is a plan section of the Sellafield site (refer to Fig. 3 for scale) representing the first layer of grid cells below ground surface. This type of projection is particularly useful as most excavation activities are undertaken from the ground surface, which is therefore the most useful reference point. Similar plots have been produced for horizontal and vertical sections through the Sellafield site. The probabilities of cells in this layer belonging to particular soil classes are shown: ‘VLLW’ (top left), ‘LLW lower’ (top right), ‘LLW middle’ (bottom left) and ‘LLW upper’ (bottom right). White areas are not interpolated due to the neighbourhood.

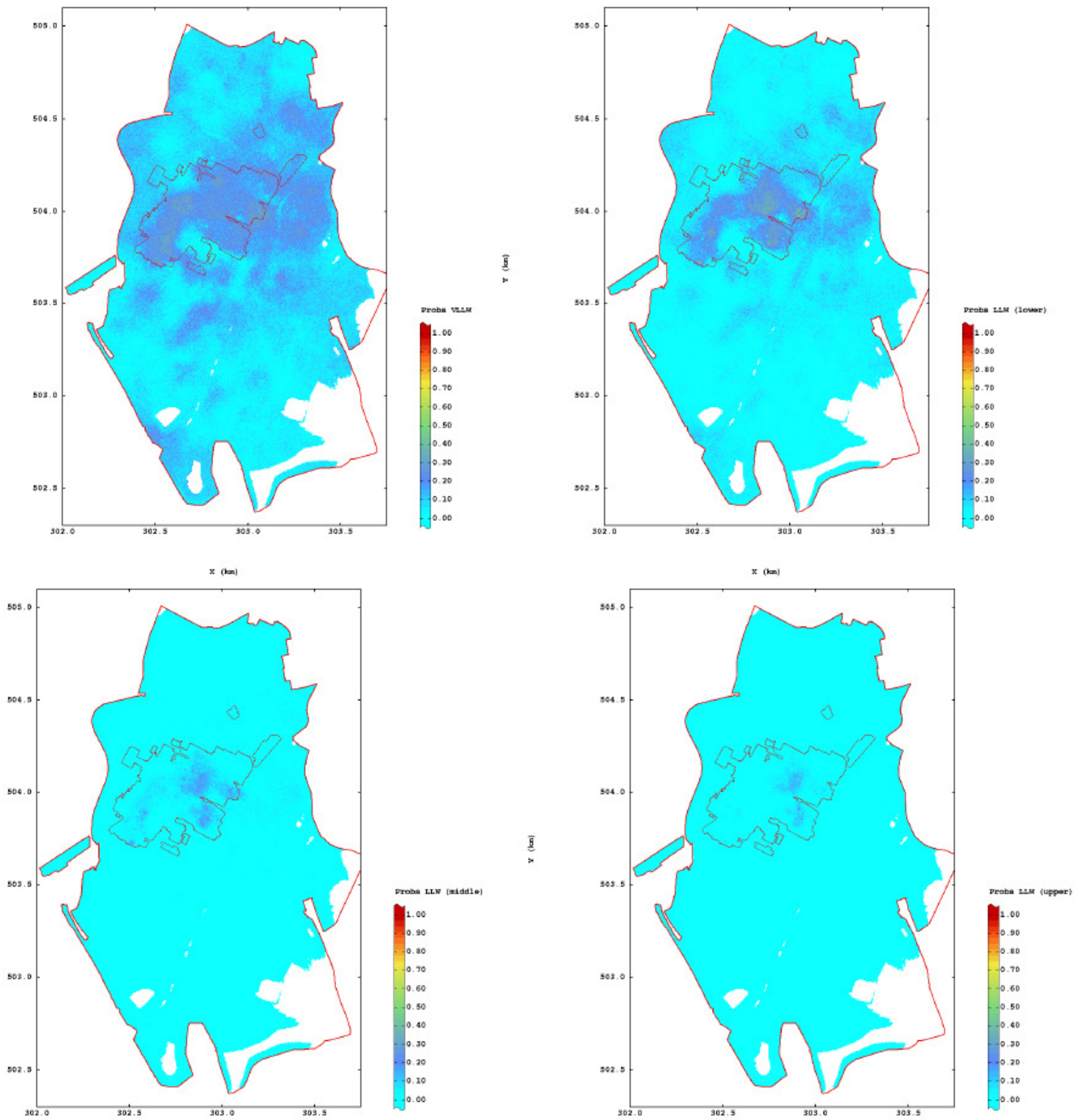


Fig. 8. Plan section of the area within the Sellafield site boundary for the first layer of grid cells below ground elevation. The probabilities of cells in this layer belonging to particular soil classes are shown: 'VLLW' (top left), 'LLW lower' (top right), 'LLW middle' (bottom left) and 'LLW upper' (bottom right). White areas are not interpolated due to the neighbourhood

IV GROUNDWATER CONTAMINATION

One hundred simulations using turning bands technique were performed for gross beta activity. For the purposes of the calculations presented here, we restricted the calculations to the part of the region defined in Section II C that was below the January groundwater table. Volumes were computed with both thresholds of 18,000 Bq/m^3 and 20,000 Bq/m^3; the calculated median volumes were within 3% of each other.

As with the soil data, a Gaussian transform of the logarithm of gross beta activity was undertaken. Variographic analysis shows a clear horizontal structure (Fig. 9). For the vertical direction, fewer points are available and the structure is derived from the variogram computed on Beta data from all available dates.

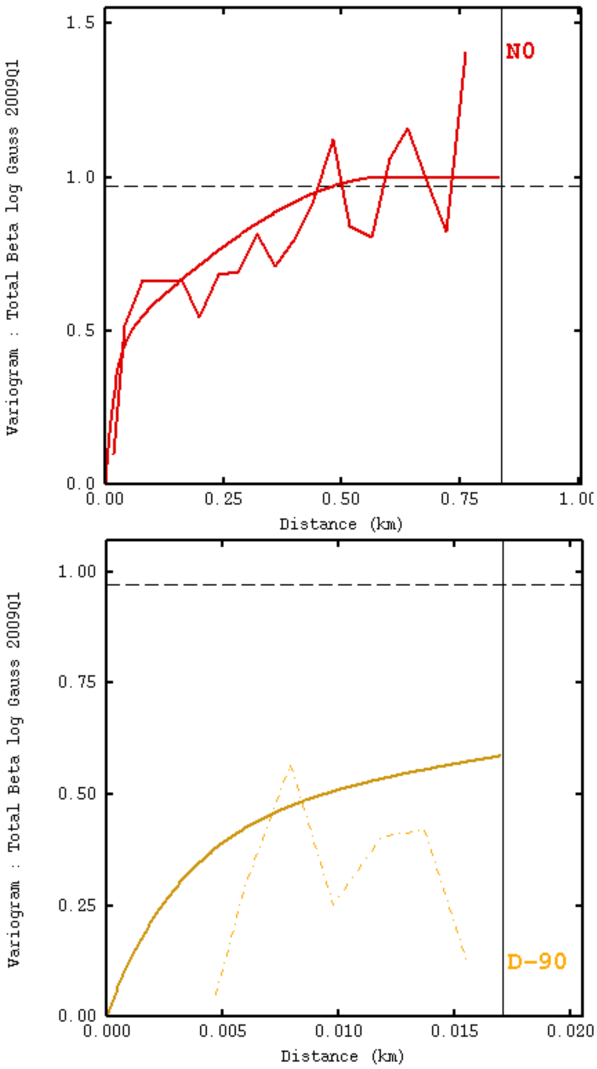


Fig. 9. Variographic analysis of the gaussian transform of log10(gross beta): Horizontal (top) and vertical (bottom) variogram.

A 3D plot of the probability of exceeding 18,000 Bq/m^3 gross beta activity is shown in Fig. 10. The 3D view shows an area within the Sellafield site boundary: compare with Fig. 3 to get the location. In Fig 10, only cells where the probability of exceeding the threshold is greater than 30% are shown.

V CONCLUSIONS

Monitoring and characterisation programmes at Sellafield generate a large quantity of important environmental data gathered at public cost. Sellafield Ltd wishes to ensure that appropriate methods are being used for the analysis of these data. Some significant progress has been made on this matter through the application of geostatistical model-based simulations, visualisation tools and spatial analysis. A geostatistical understanding of the spatial distribution of contaminated soil and groundwater has now been successfully developed through this work. Better estimates of the volume, activity, and spatial distribution of contaminants have been obtained; an improved understanding of these parameters is important in demonstrating management control over sub-surface soil and groundwater contamination.

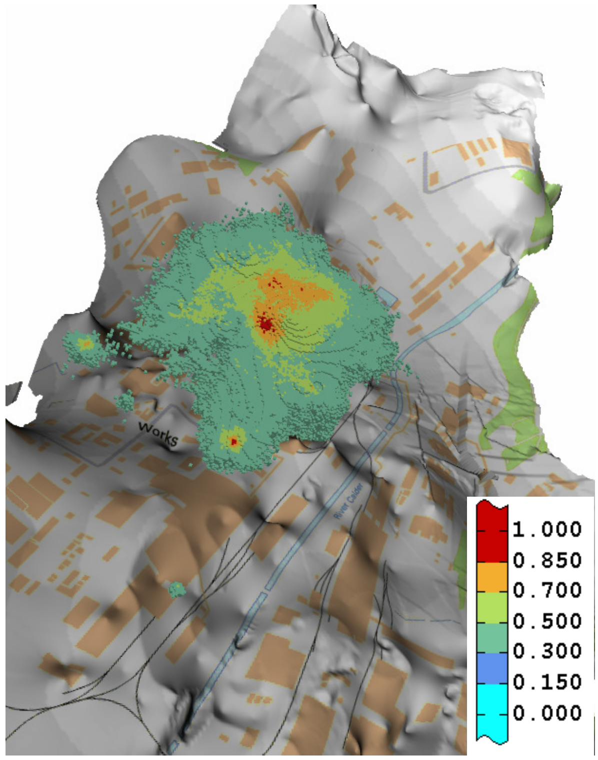


Fig. 10. Probability to exceed 18,000 Bq/m^3 in groundwater (only cells with probabilities greater than 30% are displayed). The 3D view shows an area within the Sellafield site boundary: compare with Fig. 3 to get the location

REFERENCES

1. Sellafield Ltd, Findings of the Sellafield Contaminated Land & Groundwater Management Project and the Next Steps for the Land Quality Programme, LQTRG 11 021, Report Number: MAN000405, Land Quality Technical Team (2012).
2. Sellafield Ltd, Groundwater Monitoring at Sellafield. Annual Data Review 2013, Sellafield Ltd Report LQTD000077, 22 September 2014
3. J.P. CHILES. and P. Delfiner, Geostatistics: modeling spatial uncertainty. 2nd edition, Wiley series in probability and statistics (2012).
4. Geovariances, ISATIS technical references. Geovariances (2015).